

Introducción a la probabilidad

Luis Rincón
Departamento de Matemáticas
Facultad de Ciencias UNAM
Circuito Exterior de CU
04510 México DF

Agosto 2014

Prólogo

El presente trabajo contiene material suficiente para un primer curso a nivel universitario sobre la teoría de la probabilidad. Está dirigido a estudiantes de las carreras de actuaría, matemáticas y otras carreras científicas similares cuyos programas de estudio contemplan un curso en donde se muestren los resultados, usos y aplicaciones de la probabilidad.

Se exponen temas tradicionales de la probabilidad básica, se estudian las variables aleatorias más conocidas y sus distribuciones de probabilidad, así como algunas técnicas y resultados clásicos de la probabilidad. Se ha buscado que en el texto aparezcan numerosas gráficas y diagramas con el objetivo de hacer las explicaciones más claras. Para una lectura provechosa de este material, se requiere tener cierta familiaridad con algunos conceptos del álgebra y del cálculo diferencial e integral.

Agradezco sinceramente a los revisores anónimos, quienes emitieron dictámenes constructivos acerca de este trabajo. Agradezco también el apoyo financiero del programa PAPIME a través del proyecto PE101113, DGAPA, UNAM, con el cual fue posible la impresión de este texto y la elaboración del curso en videos, disponible en la página web

<http://www.matematicas.unam.mx/lars/0625>

Luis Rincón
Agosto 2014
Ciudad Universitaria UNAM
lars@ciencias.unam.mx

Contenido

1. Probabilidad elemental	1
1.1. Experimentos aleatorios	3
1.2. Espacio muestral	5
1.3. Operaciones con conjuntos	8
1.4. Probabilidad clásica	20
1.5. Probabilidad geométrica	23
1.6. Probabilidad frecuentista	31
1.7. Probabilidad subjetiva	34
1.8. Probabilidad axiomática	35
1.9. Sigmas álgebras	46
1.10. Sigma álgebra de Borel	50
1.11. Espacios de probabilidad	52
1.12. Análisis combinatorio	55
1.13. Probabilidad condicional	71
1.14. Teorema de probabilidad total	76
1.15. Teorema de Bayes	83
1.16. Independencia de eventos	89
1.17. Continuidad de la probabilidad	99
2. Variables aleatorias	107
2.1. Variables aleatorias	107
2.2. Función de probabilidad	118
2.3. Función de distribución	129
2.4. Teorema de cambio de variable	145
2.5. Independencia de variables aleatorias	149
2.6. Esperanza	153

2.7. Varianza	169
2.8. Momentos	178
2.9. Cuantiles	182
2.10. Moda	186
2.11. Función generadora de probabilidad	187
2.12. Función generadora de momentos	195
3. Distribuciones de probabilidad	203
3.1. Distribución uniforme discreta	204
3.2. Distribución Bernoulli	208
3.3. Distribución binomial	213
3.4. Distribución geométrica	221
3.5. Distribución binomial negativa	227
3.6. Distribución hipergeométrica	234
3.7. Distribución Poisson	238
3.8. Distribución uniforme continua	248
3.9. Distribución exponencial	253
3.10. Distribución gamma	258
3.11. Distribución beta	263
3.12. Distribución Weibull	266
3.13. Distribución normal	269
3.14. Distribución ji-cuadrada	278
3.15. Distribución t	283
3.16. Distribución F	287
4. Vectores aleatorios	291
4.1. Vectores aleatorios	291
4.2. Función de probabilidad conjunta	293
4.3. Función de distribución conjunta	308
4.4. Función de probabilidad marginal	315
4.5. Función de distribución marginal	319
4.6. Independencia de variables aleatorias	321
4.7. Distribución condicional	327
4.8. Esperanza condicional	333
4.9. Esperanza, varianza y covarianza	336
4.10. Coeficiente de correlación	341

5. Teoremas límite	345
5.1. Desigualdad de Chebyshev	345
5.2. Convergencia de variables aleatorias	349
5.3. La ley de los grandes números	353
5.4. El teorema central del límite	361
A. Fórmulas varias	375
B. Introducción a R	385
C. Solución a los ejercicios	393
Bibliografía	499
Índice analítico	501

Capítulo 1

Probabilidad elemental

En esta primera parte estudiaremos algunos de los conceptos más elementales de la teoría de la probabilidad. Esta teoría matemática tuvo como uno de sus primeros puntos de partida el intentar resolver un problema particular, concerniente a una apuesta de juego de dados entre dos personas. El problema al que nos referimos involucraba una gran cantidad de dinero y puede plantearse de la siguiente forma.

Dos jugadores escogen, cada uno de ellos, un número del 1 al 6, distinto uno del otro, y apuestan 32 doblones de oro a que el número escogido por uno de ellos aparece en tres ocasiones antes que el número del contrario al lanzar sucesivamente un dado. Suponga que el número de uno de los jugadores ha aparecido dos veces y el número del otro, una sola vez. Bajo estas circunstancias, ¿cómo debe dividirse el total de la apuesta si el juego se suspende?

Uno de los apostadores, Antonio de Gombaud, popularmente conocido como el caballero De Méré, deseando conocer la respuesta al problema, plantea la situación a Blaise Pascal (1623-1662). Pascal, a su vez, consulta con Pierre de Fermat (1601-1665) e inician, estos últimos, un intercambio de cartas a propósito del problema. Esto sucede en el año de 1654. Con ello se inician algunos esfuerzos por dar solución a éste y otros problemas similares que se plantean. Con el paso del tiempo se sientan las bases y las experiencias necesarias para la búsqueda de una teoría matemática que sintetice los

conceptos y los métodos de solución de los muchos problemas particulares resueltos a lo largo de varios años. En el segundo congreso internacional de matemáticas, celebrado en la ciudad de París en el año 1900, el matemático alemán David Hilbert (1862-1943) planteó 23 problemas matemáticos de importancia de aquella época. Uno de estos problemas era el de encontrar axiomas o postulados a partir de los cuales se pudiera construir una teoría matemática de la probabilidad. Aproximadamente treinta años después, en 1933, el matemático ruso A. N. Kolmogorov (1903-1987) propone ciertos axiomas que, a la postre, resultaron adecuados para la construcción de una teoría de la probabilidad. Esta teoría prevalece hoy en día y ha adquirido el calificativo de teoría clásica.



Blaise Pascal
(Francia 1623–1662)



Pierre de Fermat
(Francia 1601–1665)

Actualmente la teoría de la probabilidad se ha desarrollado y extendido enormemente gracias a muchos pensadores que han contribuido a su crecimiento, y es, sin duda, una parte importante y bien establecida de las matemáticas. La teoría de la probabilidad ha resultado muy útil para modelar matemáticamente fenómenos de muy diversas disciplinas del conocimiento humano en donde es necesario incorporar la incertidumbre o el azar como un elemento esencial del modelo. Así, la probabilidad puede definirse como aquella parte de las matemáticas que se encarga del estudio de los fenómenos aleatorios.

1.1. Experimentos aleatorios

Existen dos tipos de fenómenos o experimentos en la naturaleza: los deterministas y los aleatorios. Un experimento determinista es aquel que produce el mismo resultado cuando se le repite bajo las mismas condiciones, por ejemplo, medir el volumen de un gas cuando la presión y la temperatura son constantes produce, teóricamente, siempre el mismo resultado, o medir el ángulo de un rayo de luz reflejado en un espejo resulta siempre en el mismo resultado cuando el ángulo de incidencia es el mismo y el resto de las condiciones son constantes. Muchas otras leyes de la física son ejemplos de situaciones en donde, bajo idénticas condiciones iniciales, el resultado del experimento es siempre el mismo. En contraparte, un experimento aleatorio es aquel que, cuando se le repite bajo las mismas condiciones, el resultado que se observa no siempre es el mismo y tampoco es predecible. El lanzar una moneda al aire y observar la cara de la moneda que mira hacia arriba al caer, o registrar el número ganador en un juego de lotería, son ejemplos cotidianos de experimentos aleatorios.

Nuestro interés consiste en estudiar modelos matemáticos, conceptos y resultados, que nos ayuden a tener un mejor entendimiento y control de los muy diversos fenómenos aleatorios que afectan la vida del hombre.

Para ser más precisos, pediremos que los experimentos aleatorios que consideremos cumplan teóricamente las características siguientes y, con ello, restringimos sensiblemente el campo de aplicación.

- a) El experimento debe poder ser repetible bajo las mismas condiciones iniciales.
- b) El resultado de cualquier ensayo del experimento es variable y depende del azar o de algún mecanismo aleatorio.

Es necesario mencionar, sin embargo, que en algunas ocasiones no es evidente poder clasificar un experimento dado en aleatorio o determinista, esto dependerá del observador, de lo que él o ella conozca del experimento y de lo que esta persona desea observar. Así, el experimento mismo no está separado completamente del observador, pues la concepción, entendimiento y

realización del experimento aleatorio dependen, en alguna medida, del observador mismo. En la siguiente sección de ejercicios se muestran algunos casos particulares. Por otro lado, debe observarse que, convenientemente, hemos dejado sin definir el término azar. Este es un concepto difícil de capturar formalmente en una definición. Al usar este término, únicamente haremos referencia a la vaga noción intuitiva que cada uno de nosotros posee acerca del azar según nuestra propia experiencia cotidiana.

Ejercicios

1. Clasifique los siguientes experimentos en deterministas o aleatorios. Si es necesario añada hipótesis o condiciones adicionales para justificar su respuesta.
 - a) Registrar el número de accidentes que ocurren en una determinada calle de una ciudad.
 - b) Observar la temperatura a la que hierve el agua a una altitud dada.
 - c) Registrar el consumo de electricidad de una casa-habitación en un día determinado.
 - d) Registrar la hora a la que desaparece el sol en el horizonte en un día dado, visto desde una posición geográfica determinada.
 - e) Observar el precio que tendrá el petróleo dentro de un año.
 - f) Registrar la altura máxima que alcanza un proyectil lanzado verticalmente.
 - g) Observar el número de años que vivirá un bebé que nace en este momento.
 - h) Observar el ángulo de reflexión de un haz de luz incidente en un espejo.
 - i) Registrar la precipitación pluvial anual en una zona geográfica determinada.
 - j) Observar el tiempo que tarda un objeto en caer al suelo cuando se le deja caer desde una altura dada.

- k*) Registrar al ganador de una elección en un proceso de votación libre y secreto.
 - l*) Observar la posición de una molécula de oxígeno en una habitación, después de dejarla en libre movimiento durante un minuto.
2. **¿Qué es el azar?** Intente escribir una definición formal que se apegue lo más posible a su propio entendimiento de este concepto.

1.2. Espacio muestral

Hemos mencionado que la teoría de la probabilidad es la parte de las matemáticas que se encarga del estudio de los fenómenos o experimentos aleatorios. En principio no sabemos cuál será el resultado de un experimento aleatorio, así que por lo menos conviene agrupar en un conjunto a todos los resultados posibles. Esto lleva a la siguiente definición.

Definición 1.1 El espacio muestral, también llamado espacio muestra, de un experimento aleatorio es el conjunto de todos los posibles resultados del experimento y se le denota, generalmente, por la letra griega Ω (omega mayúscula). A un resultado particular del experimento se le denota por la letra ω (omega minúscula).

Más adelante mostraremos que el espacio muestral no es necesariamente único y su determinación depende de lo que desea observar o estudiar la persona que realiza el experimento aleatorio. En algunos textos se usa también la letra S para denotar al espacio muestral. Esta letra proviene del término *sampling space* de la lengua inglesa, equivalente a espacio muestral o espacio muestra. Por otro lado, y de manera preliminar, llamaremos evento o suceso a cualquier subconjunto del espacio muestral. A los eventos los denotaremos por las primeras letras del alfabeto en mayúsculas: A, B, C, \dots o bien por alguna otra letra en mayúscula que nos ayude a identificar de mejor manera al evento. A través de algunos ejemplos ilustraremos a continuación los conceptos de espacio muestral, evento y ocurrencia de un evento.

Ejemplo 1.1 Si un experimento aleatorio consiste en lanzar un dado y observar el número que aparece en la cara superior, entonces claramente el espacio muestral es el conjunto $\Omega = \{1, 2, 3, 4, 5, 6\}$. Como ejemplo de un evento para este experimento podemos definir el conjunto $A = \{2, 4, 6\}$, que corresponde al suceso de obtener como resultado un número par. Si al lanzar el dado una vez se obtiene el número “4”, decimos entonces que se observó la ocurrencia del evento A , y si se obtiene, por ejemplo, el resultado “1”, decimos que no se observó la ocurrencia del evento A . •

Ejemplo 1.2 Considere el experimento aleatorio de participar en un juego de lotería. Suponga que hay un millón de números en esta lotería y un jugador participa con un boleto. ¿Cuál es un posible espacio muestral para este experimento si únicamente uno de los posibles números es el ganador? Naturalmente, al jugador le interesa conocer su suerte en este juego y puede proponer como espacio muestral el conjunto $\Omega = \{\text{“ganar”}, \text{“perder”}\}$. Sin embargo puede también tomarse como espacio muestral el conjunto que contiene a todos los números participantes, es decir, $\Omega = \{1, 2, \dots, 10^6\}$. Este ejemplo sencillo muestra que el espacio muestral de un experimento aleatorio no es único y depende del interés del observador. •

Ejemplo 1.3 Suponga que un experimento aleatorio consiste en observar el tiempo en el que una máquina en operación sufre su primera descompostura. Si se consideran mediciones continuas del tiempo, entonces puede adoptarse como espacio muestral el intervalo $[0, \infty)$. El subconjunto $A = [1, 2]$ corresponde al evento en el que la primera descompostura se observe entre la primera y la segunda unidad de tiempo. Si se consideran mediciones discretas del tiempo, ¿cuál podría ser un posible espacio muestral para este experimento? •

Se dice que un evento es simple cuando consta de un solo elemento del espacio muestral, en cambio, se llama compuesto cuando consta de más de un elemento del espacio muestral. Puesto que los conceptos de espacio muestral y evento involucran forzosamente el manejo de conjuntos, recordaremos, en la siguiente sección, algunas operaciones entre estos objetos y ciertas propiedades que nos serán de suma utilidad. Nuestro objetivo es calcular la

probabilidad de ocurrencia de los diversos eventos en un experimento aleatorio.

Ejercicios

3. Determine un espacio muestral para el experimento aleatorio consistente en:
 - a) Observar la posición de un partícula en un instante dado, la cual se mueve sin restricciones en un espacio tridimensional.
 - b) Registrar el número de personas que requieren hospitalización en el siguiente accidente automovilístico atendido por los servicios de emergencia en una localidad dada.
 - c) Lanzar un dado hasta que se obtiene un “6”.
 - d) Registrar la fecha de cumpleaños de n personas escogidas al azar.
 - e) Observar la forma en la que r personas que abordan un elevador en la planta baja de un edificio descienden en los pisos $1, 2, \dots, n$.
 - f) Registrar la duración de una llamada telefónica escogida al azar.
 - g) Observar el número de años que le restan de vida a una persona escogida al azar dentro del conjunto de asegurados de una compañía aseguradora.
4. Proponga un espacio muestral para el experimento aleatorio de lanzar tres monedas a un mismo tiempo, suponiendo que las monedas:
 - a) son distinguibles, es decir, pueden por ejemplo ser de colores distintos.
 - b) no son distinguibles, es decir, físicamente son idénticos.
5. Considere el experimento aleatorio de lanzar dos dados distinguibles. Escriba explícitamente los resultados asociados a los siguientes eventos y determine su cardinalidad.

- a) A = “La suma de los dos resultados es 7.”
 b) B = “Uno de los dos dados cae en número impar y el otro en número par.”
 c) C = “El resultado de un dado difiere del otro en, a lo sumo, una unidad.”
 d) D = “El resultado de un dado difiere del otro en por lo menos cuatro unidades.”
 e) $E = A \cap B$.
 f) $F = B^c$.
 g) $G = C \cup D$.

1.3. Operaciones con conjuntos

Nos interesa poder identificar a todos los posibles eventos en un experimento aleatorio, pues deseamos calcular la probabilidad de ocurrencia de cada uno de ellos. Recordemos que pueden obtenerse nuevos conjuntos a partir de una colección inicial de eventos y algunas operaciones sobre ellos, como las que definiremos más adelante. Consideraremos que estos nuevos conjuntos resultantes son también eventos y deseamos poder calcular su probabilidad. Es por esto que nos será útil revisar brevemente algunas operaciones usuales entre conjuntos. Estableceremos primero varios conceptos elementales y la notación a utilizar.

Supondremos que el espacio muestral Ω de un experimento aleatorio es una especie de conjunto universal y, como hemos mencionado antes, cualquier elemento de Ω lo denotaremos por la letra ω . Al conjunto vacío lo denotaremos, como es usual, por el símbolo \emptyset . Otros símbolos usuales son los de pertenencia (\in) o no pertenencia (\notin) de un elemento en un conjunto, y los de contención (\subset, \subseteq) o no contención ($\not\subset$) de un conjunto en otro. Se dice que A es un subconjunto propio de B si $A \subsetneq B$, es decir, si A está contenido en B pero no es todo B . La igualdad de dos conjuntos A y B significa que se cumplen las dos contenciones: $A \subset B$ y $B \subset A$. Por último, si A es un conjunto, denotamos la cardinalidad o número de elementos de ese conjunto por el símbolo $\#A$. Ahora procederemos a definir algunas operaciones entre conjuntos.

Sean A y B dos subconjuntos cualesquiera de Ω . Recordamos a continuación las operaciones básicas de unión, intersección, diferencia y complemento.

$$A \cup B = \{\omega \in \Omega : \omega \in A \text{ o } \omega \in B\},$$

$$A \cap B = \{\omega \in \Omega : \omega \in A \text{ y } \omega \in B\},$$

$$A - B = \{\omega \in \Omega : \omega \in A \text{ y } \omega \notin B\},$$

$$A^c = \{\omega \in \Omega : \omega \notin A\}.$$

Cuando los conjuntos se expresan en palabras, la operación unión, $A \cup B$, se lee “A o B” y la intersección, $A \cap B$, se lee “A y B”. En la Figura 1.1 se muestran, en diagramas de Venn¹, estas dos operaciones.

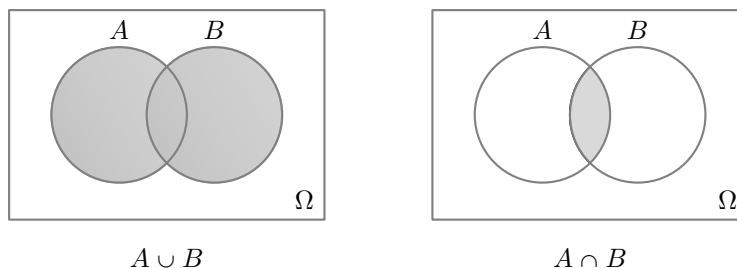


Figura 1.1

La diferencia entre dos conjuntos A y B se denota por $A - B$ y corresponde a aquel conjunto de elementos de A que no pertenecen a B , es decir, $A - B$ se define como $A \cap B^c$. En general, el conjunto $A - B$ es distinto de $B - A$, de hecho estos conjuntos son siempre ajenos, ¿puede usted comprobar tal afirmación? ¿en qué caso ambos conjuntos coinciden? Por otro lado, el complemento de un conjunto A se denota por A^c y se define como la colección de aquellos elementos de Ω que no pertenecen a A . Mediante un diagrama de Venn, ilustramos gráficamente las operaciones de diferencia y complemento en la Figura 1.2.

¹John Venn (1834-1923), filósofo y lógico inglés.

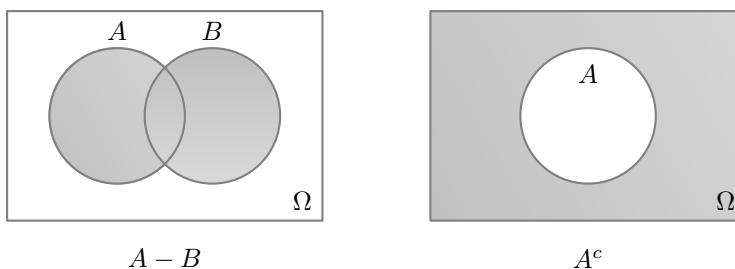


Figura 1.2

Ejemplo 1.4 Sea A el conjunto de aquellas personas que tienen hijos y B la colección de aquellas personas que están casadas. Entonces el conjunto $A \cap B$ consta de aquellas personas que están casadas y tienen hijos, mientras que el conjunto $A \cap B^c$ está constituido por aquellas personas que tienen hijos pero no están casadas. ¿Quién es $A^c \cap B$? Observe que cada persona es un elemento de alguno de los siguientes conjuntos: $A \cap B$, $A \cap B^c$, $A^c \cap B$ ó $A^c \cap B^c$. ¿A cuál de ellos pertenece usted? •

Es fácil verificar que el conjunto vacío y el conjunto total satisfacen las siguientes propiedades elementales: para cualquier evento A ,

$$\begin{array}{lll} A \cup \emptyset = A. & A \cap \emptyset = \emptyset. & A \cup \Omega = \Omega. \\ A \cap \Omega = A. & A \cup A^c = \Omega. & A \cap A^c = \emptyset. \end{array}$$

Además, las operaciones unión e intersección son asociativas, esto es, satisfacen las siguientes igualdades:

$$\begin{aligned} A \cup (B \cup C) &= (A \cup B) \cup C, \\ A \cap (B \cap C) &= (A \cap B) \cap C, \end{aligned}$$

y también son distributivas, es decir,

$$\begin{aligned} A \cap (B \cup C) &= (A \cap B) \cup (A \cap C), \\ A \cup (B \cap C) &= (A \cup B) \cap (A \cup C). \end{aligned}$$

Recordemos también la operación diferencia simétrica entre dos conjuntos A y B , denotada por $A\Delta B$ y definida como sigue

$$A\Delta B = (A \cup B) - (B \cap A).$$

En la Figura 1.3 ilustramos gráficamente el conjunto resultante de efectuar la diferencia simétrica entre los conjuntos A y B . Visualmente se puede comprobar que la diferencia simétrica también puede escribirse como $(A - B) \cup (B - A)$. ¿Cómo podría expresarse en palabras al conjunto $A\Delta B$?

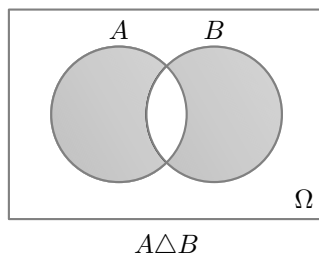


Figura 1.3

Recordemos además las muy útiles leyes de De Morgan²:

$$\begin{aligned}(A \cup B)^c &= A^c \cap B^c, \\ (A \cap B)^c &= A^c \cup B^c.\end{aligned}$$

La validez de estas dos igualdades puede extenderse a colecciones finitas e incluso arbitrarias de conjuntos. ¿Puede usted escribir estas identidades para n conjuntos?

Conjuntos ajenos

Cuando dos conjuntos no tienen ningún elemento en común se dice que son ajenos, es decir, los conjuntos A y B son ajenos o disjuntos si se cumple la igualdad

$$A \cap B = \emptyset.$$

²Augustus De Morgan (1806-1871), matemático y lógico británico.

Por ejemplo, si $\Omega = \{1, 2, 3, 4, 5, 6\}$, entonces los conjuntos $A = \{1, 2\}$ y $B = \{3, 4\}$ son ajenos pues no hay ningún elemento común entre ellos. El ejemplo general más importante de conjuntos o eventos ajenos es la pareja dada por A y A^c , para cualquier conjunto A . La propiedad de ser ajenos puede extenderse al caso cuando se tienen varios conjuntos. Decimos que n conjuntos A_1, \dots, A_n son ajenos si $A_1 \cap \dots \cap A_n = \emptyset$, y se dice que son ajenos dos a dos (o mutuamente ajenos) si $A_i \cap A_j = \emptyset$ para cualesquiera valores de los índices $i, j = 1, 2, \dots, n$, con i distinto de j . La propiedad de ser ajenos dos a dos para una colección de eventos implica que los conjuntos son ajenos, sin embargo, el hecho de que todos ellos sean ajenos no implica que sean ajenos dos a dos. Es decir, la propiedad de ser ajenos dos a dos es más fuerte que la propiedad de ser simplemente ajenos, y es la que usualmente supondremos en la mayoría de los casos. Ilustraremos la situación con el siguiente ejemplo.

Ejemplo 1.5 Los conjuntos $A = \{1, 2\}$, $B = \{2, 3\}$ y $C = \{3, 4\}$ son ajenos pues $A \cap B \cap C = \emptyset$, pero no son ajenos dos a dos pues, por ejemplo, el conjunto $A \cap B$ no es vacío. Así, los conjuntos A , B y C son ajenos en el sentido de que la intersección de todos ellos es vacía, pero no son ajenos dos a dos. •

Las operaciones entre conjuntos, que mencionaremos a continuación, no son elementales y producen nuevos conjuntos que se encuentran en un nivel distinto al de los conjuntos originales.

Conjunto potencia

El conjunto potencia de Ω , denotado por 2^Ω , es aquel conjunto constituido por todos los subconjuntos posibles de Ω . En términos estrictos, esta nueva colección deja de ser un conjunto y se le llama clase de subconjuntos de Ω , aunque seguiremos usando el primer término en nuestro tratamiento elemental de conjuntos. Por ejemplo, si $\Omega = \{a, b, c\}$, entonces el conjunto 2^Ω consta de 8 elementos, a saber,

$$2^\Omega = \left\{ \emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \Omega \right\}.$$

Observe que los elementos del conjunto potencia son, en sí mismos, conjuntos, y que en esta colección están contenidos todos los eventos que podrían

ser de interés en un experimento aleatorio. No es difícil cerciorarse que cuando $\#\Omega < \infty$,

$$\#(2^\Omega) = 2^{\#\Omega},$$

es decir, el número de elementos en el conjunto 2^Ω es exactamente 2 elevado a la potencia dada por la cardinalidad de Ω . De este hecho proviene la notación usada para el conjunto potencia. Observe que la expresión 2^Ω no tiene el significado matemático del número 2 elevado a la potencia Ω , pues ello no tiene sentido. Debe considerarse, por lo tanto, como un símbolo para denotar al conjunto potencia y que ayuda a recordar el número de elementos en dicha clase. Para el ejemplo anterior se comprueba que la cardinalidad de 2^Ω es efectivamente $2^{\#\Omega} = 2^3 = 8$.

Producto cartesiano

El producto cartesiano de dos conjuntos A y B , denotado por $A \times B$, se define como la colección de todas las parejas ordenadas (a, b) , en donde a es cualquier elemento de A y b es cualquier elemento de B . En símbolos,

$$A \times B = \{(a, b) : a \in A \text{ y } b \in B\}.$$

En el caso cuando A y B son subconjuntos de números reales, el producto cartesiano $A \times B$ puede representarse gráficamente como se muestra en el Figura 1.4

Ejemplo 1.6 Si $A = \{a_1, a_2\}$ y $B = \{b_1, b_2, b_3\}$, entonces

$$A \times B = \{(a_1, b_1), (a_1, b_2), (a_1, b_3), (a_2, b_1), (a_2, b_2), (a_2, b_3)\}.$$

Este conjunto puede representarse gráficamente como antes se mostró en la Figura 1.4 o bien mediante un diagrama de árbol como el que se ilustra en la Figura 1.5 . •

Ejemplo 1.7 Si un hombre tiene 6 camisas y 7 pantalones, ¿de cuántas maneras diferentes puede vestirse con estas prendas?

Respuesta. El hombre puede vestirse de $6 \times 7 = 42$ formas distintas. •

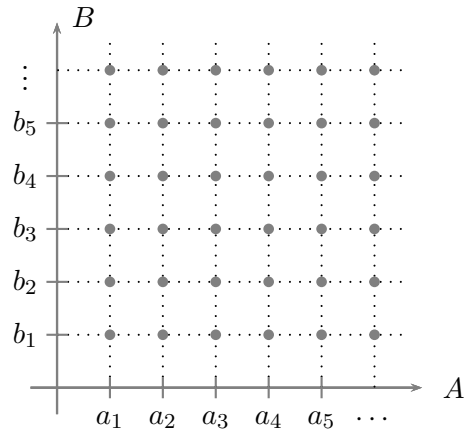


Figura 1.4

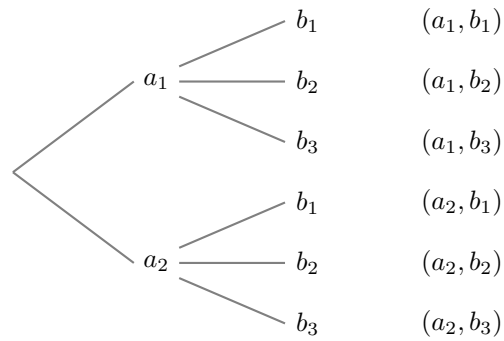


Figura 1.5

En general, los productos cartesianos $A \times B$ y $B \times A$ son distintos pues las parejas (a, b) son distintas de (b, a) , sin embargo ambos conjuntos tienen la misma cardinalidad, esto es, ambos tienen el mismo número de elementos. Si la cardinalidad de A es el número n y la cardinalidad de B es m , entonces

la cardinalidad del conjunto $A \times B$ es el producto $n \cdot m$. Este resultado es llamado principio de multiplicación y se aplica con mucha frecuencia en los procesos de conteo. Lo revisaremos nuevamente en la sección sobre análisis combinatorio.

Un poco más generalmente, si A_1, A_2, \dots, A_k son conjuntos tales que $\#A_i = n_i \geq 1$ para $i = 1, \dots, k$, entonces el producto cartesiano $A_1 \times A_2 \times \dots \times A_k$, que consta de todos los vectores de la forma (a_1, a_2, \dots, a_k) con $a_i \in A_i$, tiene un total de $n_1 \cdot n_2 \cdot \dots \cdot n_k$ elementos, es decir,

$$\#(A_1 \times \dots \times A_k) = n_1 \cdot n_2 \cdot \dots \cdot n_k.$$

Ejemplo 1.8 Si una mujer tiene 3 sombreros, 6 blusas, 8 faldas y 10 pares de zapatos, ¿de cuántas formas diferentes puede vestirse usando una prenda de cada tipo?

Respuesta. La mujer puede vestirse de $3 \times 6 \times 8 \times 10 = 1440$ maneras distintas. •

Ejemplo 1.9 Al producto cartesiano $\mathbb{R} \times \mathbb{R}$, definido como el conjunto de todas las parejas de números reales (x, y) , se le denota usualmente por \mathbb{R}^2 . Análogamente se definen los conjuntos $\mathbb{R}^3, \mathbb{R}^4, \dots, \mathbb{R}^n$. •

Concluimos aquí nuestra rápida y breve revisión de la teoría elemental de conjuntos. Recordemos que estamos interesados en calcular probabilidades de los diferentes eventos, es decir, de subconjuntos del espacio muestral que se obtienen de los diversos experimentos aleatorios. En las siguientes secciones estudiaremos algunas formas de definir matemáticamente la probabilidad de un evento.

Ejercicios

- Use las propiedades básicas de las operaciones entre conjuntos para demostrar rigurosamente las siguientes igualdades. En cada caso dibuje un diagrama de Venn para ilustrar la identidad.

- a) $A = (A \cap B) \cup (A \cap B^c)$.
- b) $A^c - B^c = B - A$.
- c) $A \cap B^c = A - (A \cap B)$.
- d) $A \cup B = A \cup (B \cap A^c)$.
- e) $(A - B) - C = A - (B - C)$.
- f) $A - (B \cap C) = (A - B) \cup (A - C)$.

7. Demuestre o proporcione un contraejemplo para las siguientes proposiciones. En cada caso dibuje un diagrama de Venn para ilustrar la situación.

- a) $A \cap B \subseteq A \subseteq A \cup B$.
- b) Si $A \cap B = \emptyset$ entonces $A \subseteq B^c$.
- c) Si $A \subseteq B$ entonces $B^c \subseteq A^c$.
- d) Si $A \cap B = \emptyset$ entonces $A \cup B^c = B^c$.
- e) Si $A \subseteq B$ entonces $A \cup (B - A) = B$.
- f) $(A^c \cap B) \cup (A \cap B^c) = (A \cap B)^c$.

8. **Diferencia simétrica.** La diferencia simétrica entre los conjuntos A y B se puede también definir como sigue

$$A \Delta B := (A - B) \cup (B - A).$$

Demuestre o proporcione un contraejemplo para las siguientes identidades.

- a) $A \Delta B = (A \cup B) - (A \cap B)$.
- b) $A \Delta B = B \Delta A$.
- c) $A \Delta \emptyset = A$.
- d) $A \Delta \Omega = A^c$.
- e) $A \Delta A = \emptyset$.
- f) $A \Delta A^c = \Omega$.
- g) $A \Delta B = A^c \Delta B^c$.
- h) $(A \Delta B)^c = A^c \Delta B^c$.

- i) $(A \Delta B) \Delta C = A \Delta (B \Delta C)$.
- j) $A \cap (B \Delta C) = (A \cap B) \Delta (A \cap C)$.
- k) $A \cup (B \Delta C) = (A \cup B) \Delta (A \cup C)$.

9. Sean A , B y C tres eventos de un experimento aleatorio. Expresa las siguientes oraciones en términos de estos conjuntos.

- a) Ocurre A o B pero no C .
- b) Ninguno de estos tres eventos ocurre.
- c) Sólo uno de ellos ocurre.
- d) Exactamente dos de ellos ocurren.
- e) Por lo menos uno de ellos ocurre.
- f) A lo sumo dos de ellos ocurren.

10. En una población humana en donde el número de mujeres duplica el número de hombres, el 42 % de los hombres son mayores de 50 años y el 38 % de las mujeres son mayores de 50 años. ¿Qué porcentaje total de la población es mayor a 50 años?

11. **Función indicadora.** La función indicadora de un evento cualquiera A se denota por $1_A : \Omega \rightarrow \mathbb{R}$ y toma el valor uno dentro del evento A y cero fuera de él, es decir,

$$1_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A, \\ 0 & \text{si } \omega \notin A. \end{cases}$$

Demuestre que:

- a) $1_\Omega(\omega) = 1$.
- b) $1_\emptyset(\omega) = 0$.
- c) $1_{A \cup B}(\omega) = 1_A(\omega) + 1_B(\omega) - 1_{A \cap B}(\omega)$.
- d) $1_{\bigcap_{i=1}^n A_i}(\omega) = \prod_{i=1}^n 1_{A_i}(\omega)$.
- e) Si A_1, \dots, A_n son eventos ajenos dos a dos, entonces

$$1_{\bigcup_{i=1}^n A_i}(\omega) = \sum_{i=1}^n 1_{A_i}(\omega).$$

12. **Señales.** Se transmiten cuatro señales consecutivas en un canal de comunicación. Debido al ruido que se presenta en el canal, cada señal se recibe bien o con distorsión. Defina el evento D_i como aquél que indica que la i -ésima señal está distorsionada. Expresé los siguientes sucesos en términos de los eventos D_i .

- a) Sólo una señal está distorsionada.
- b) Sólo dos señales están distorsionadas.
- c) Sólo hay dos señales distorsionadas y son consecutivas.
- d) No hay dos señales consecutivas distorsionadas.
- e) Por lo menos hay dos señales consecutivas distorsionadas.

13. Considere el experimento aleatorio de escoger al azar dos números x y y del intervalo unitario $(0, 1)$. El espacio muestral Ω para este experimento es entonces el producto cartesiano $(0, 1) \times (0, 1)$. Represente en un plano cartesiano este espacio muestral e identifique los siguientes eventos:

- a) $A = \{(x, y) \in \Omega : x > 1/2, y < 1/2\}$.
- b) $B = \{(x, y) \in \Omega : x < 2y \text{ ó } y < 1/2\}$.
- c) $C = \{(x, y) \in \Omega : x^2 + y^2 < 1\} - \{(x, y) \in \Omega : y < x\}$.
- d) $D = \{(x, y) \in \Omega : |x - y| < 1/4 \text{ ó } |1 - x - y| < 1/4\}$.

14. Suponga que un experimento aleatorio consiste en escoger una persona al azar dentro de una población dada. Defina los eventos:

- H = “La persona escogida es hombre.”
- E = “La persona escogida cuenta con un empleo.”
- C = “La persona escogida es casada.”

Este es un ejemplo del uso de letras adecuadas para una mejor identificación de los eventos. Expresé en palabras el tipo de personas, según las características anteriores, determinadas por los siguientes eventos:

- | | |
|---------------------|-------------------------|
| a) $H \cap E$. | e) $H \cap E \cap C$. |
| b) $H^c \cap E^c$. | f) $(H \cap C) - E$. |
| c) $H \cup E$. | g) $(H - E) \cap C^c$. |
| d) $H - E$. | h) $C^c - E^c$. |

15. Un número entero es seleccionado al azar. Defina los eventos:

- A = “El número escogido es par.”
 B = “El número escogido termina en 5.”
 C = “El número escogido termina en 0.”

Describa con palabras los siguientes eventos:

- | | |
|-----------------|-----------------|
| a) $A \cap C$. | c) $A \cap B$. |
| b) $B \cup C$. | d) $A - B$. |

16. **Circuitos.** Considere el diagrama de la Figura 1.6, el cual representa un circuito eléctrico. Los componentes A , B_1 , B_2 , B_3 y C pueden o no pueden funcionar. Denotemos por la misma letra el evento de que el correspondiente componente funcione y por su complemento el hecho de que no funcione. Sea F el evento de que el circuito completo funcione. Escriba F y F^c en términos de los eventos A , B_1 , B_2 , B_3 y C .

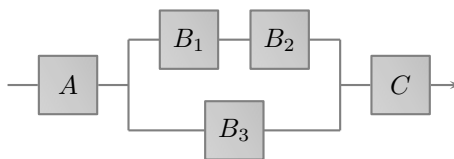


Figura 1.6

17. Sean A y B dos subconjuntos de \mathbb{R}^2 , definidos como sigue:

$$\begin{aligned}
 A &= \{(x, y) : x^2 + y^2 \leq 1\}, \\
 B &= \{(x, y) : |x| + |y| \leq 1\}.
 \end{aligned}$$

Muestre gráficamente los conjuntos:

- | | | |
|-----------------|-----------------|---------------------|
| a) A . | d) $A \cap B$. | g) $A^c \cap B^c$. |
| b) B . | e) $A - B$. | h) $A^c \cup B^c$. |
| c) $A \cup B$. | f) $B - A$. | i) $A \Delta B$. |

1.4. Probabilidad clásica

La probabilidad de un evento A es un número real en el intervalo $[0, 1]$ que se denota por $P(A)$ y representa una medida de la frecuencia con la que se observa la ocurrencia de este evento cuando se efectúa el experimento aleatorio en cuestión. Existen definiciones específicas de la probabilidad, algunas de las cuales estudiaremos en las siguientes secciones. Empezaremos estudiando la así llamada probabilidad clásica. Históricamente, esta forma de calcular probabilidades es una de las primeras en utilizarse; se aplicó con bastante éxito en problemas de juegos de azar y ayudó a sentar las bases para construir la teoría matemática. Su definición es elemental y su aplicación está restringida a situaciones cuando se satisfacen ciertas condiciones en el experimento aleatorio.

Definición 1.2 Sea A un subconjunto de un espacio muestral Ω de cardinalidad finita. Se define la probabilidad clásica del evento A como el cociente

$$P(A) = \frac{\#A}{\#\Omega},$$

en donde el símbolo $\#A$ denota la cardinalidad o número de elementos del conjunto A .

Claramente esta definición es sólo válida para espacios muestrales finitos, pues forzosamente necesitamos suponer que el número de elementos en Ω es finito. Además, el espacio Ω debe ser equiprobable, pues para calcular la probabilidad de un evento A , únicamente necesitamos contar el número de

elementos de A y dividir entre el número de elementos del conjunto total Ω , sin importar exactamente cuáles elementos particulares sean. Por lo tanto, esta definición de probabilidad puede aplicarse cuando:

- a) el espacio muestral es finito.
- b) todos los elementos del espacio muestral tienen el mismo “peso”.

Ejemplo 1.10 Considere el experimento aleatorio de lanzar un dado equilibrado. El espacio muestral es el conjunto $\Omega = \{1, 2, 3, 4, 5, 6\}$. Si deseamos calcular la probabilidad (clásica) del evento A , correspondiente a obtener un número par, es decir, la probabilidad de $A = \{2, 4, 6\}$, entonces

$$P(A) = \frac{\#\{2, 4, 6\}}{\#\{1, 2, 3, 4, 5, 6\}} = \frac{3}{6} = \frac{1}{2}.$$

•

Es inmediato verificar que esta forma de calcular probabilidades satisface, entre otras, las propiedades que se mencionan a continuación, las cuales aparecerán más adelante en la conceptualización axiomática de la probabilidad.

- a) $P(\Omega) = 1$.
- b) $P(A) \geq 0$ para cualquier evento A .
- c) $P(A \cup B) = P(A) + P(B)$ cuando A y B son ajenos.

A esta forma de definir la probabilidad también se le conoce con el nombre de probabilidad de Laplace, en honor del astrónomo y matemático francés Pierre-Simon Laplace, quien estableció de una manera sistemática y rigurosa los principios y propiedades de esta forma de calcular probabilidades. Más adelante retomaremos esta definición de probabilidad cuando revisemos algunas técnicas de conteo.

Ejercicios

18. Sean A y B dos eventos cualesquiera de un experimento aleatorio con espacio muestral finito y equiprobable. Demuestre que la definición de probabilidad clásica satisface las siguientes propiedades.
- a) $P(\emptyset) = 0$.
 - b) $P(\Omega) = 1$.
 - c) $P(A) \geq 0$ para cualquier evento A .
 - d) $P(A^c) = 1 - P(A)$.
 - e) Si $A \subseteq B$ entonces $P(A) \leq P(B)$.
 - f) $P(A \cup B) = P(A) + P(B)$ cuando A y B son ajenos.
 - g) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
19. El juego de una feria consiste en pedirle a un jugador que arroje al azar 4 monedas equilibradas, una a la vez. Suponga que las monedas son de una unidad monetaria y están marcadas con “cara” y “cruz”. Si algún lanzamiento cae “cara”, la moneda es recogida por el jugador y se le entrega una moneda adicional de la misma denominación como premio. Por otro lado, el jugador pierde cualquier moneda que caiga “cruz”. Determine el número posible de monedas que el jugador puede tener al final del juego y las probabilidades de cada uno de estos resultados.
20. Un experimento aleatorio consiste en lanzar, a un mismo tiempo, dos dados equilibrados e indistinguibles, es decir, idéntico uno del otro. Determine si a este experimento aleatorio se le puede asignar un espacio muestral finito y equiprobable.
21. **Puntos.** Suponga que un experimento aleatorio tiene como espacio muestral el conjunto de pares de números (x, y) tales que tanto x como y toman valores en el conjunto $\{1, \dots, n\}$, y que se considera que cualquiera de estos puntos en el plano cartesiano ocurre con idéntica probabilidad. Calcule la probabilidad de que, al efectuar una vez el experimento aleatorio, se obtenga un punto (x, y) :
- a) en la diagonal, es decir, $x = y$.

- b) en la orilla, es decir $x = 1$, o $x = n$, o $y = 1$, o $y = n$.
 - c) tal que $x \leq y$.
 - d) tal que $|x - y| \leq 1$.
22. Una moneda equilibrada y marcada con “cara” y “cruz” se lanza 4 veces consecutivas. Calcule la probabilidad de que:
- a) las dos caras caigan el mismo número de veces.
 - b) el número de veces que cae “cara” sea estrictamente mayor al número de veces que cae “cruz”.
23. Un experimento aleatorio consiste en lanzar una moneda equilibrada tres veces consecutivas. Escriba explícitamente a todos los elementos de un espacio muestral para este experimento y encuentre las probabilidades de cada uno de estos resultados.

1.5. Probabilidad geométrica

Esta es una extensión de la definición de probabilidad clásica, en donde ahora la probabilidad de un evento A se calcula ya no a través de su cardinalidad, sino mediante la determinación de su área, volumen o alguna característica geométrica equivalente, según el problema que se trate. Para el caso de áreas la definición es la siguiente.

Definición 1.3 Si un experimento aleatorio tiene como espacio muestral $\Omega \subset \mathbb{R}^2$ cuya área está bien definida y es finita, entonces se define la probabilidad geométrica de un evento $A \subseteq \Omega$ como

$$P(A) = \frac{\text{Área de } A}{\text{Área de } \Omega},$$

cuando el concepto de área del subconjunto A está bien definido.

Para poder aplicar la fórmula anterior, es necesario suponer que el espacio muestral es equiprobable en el sentido de que la probabilidad de observar la ocurrencia de un evento A depende únicamente de su área y no del conjunto mismo. Esta definición puede enunciarse también para el caso cuando Ω es un subconjunto de \mathbb{R} , y en tal caso no se habla de área, sino de longitud, o bien cuando $\Omega \subset \mathbb{R}^3$, y en ese caso se habla de volumen, etcétera. Ilustraremos la situación mediante algunos ejemplos.

Ejemplo 1.11 (El problema del juego de una feria) El juego de una feria consiste en lanzar monedas de radio r sobre un tablero cuadrado como el que se muestra en la Figura 1.7, en donde el lado de cada cuadrado mide a unidades. Un jugador se hace acreedor a un premio si la moneda lanzada no toca ninguna de las líneas. ¿De qué tamaño deben ser a y r para que la probabilidad de ganar en este juego sea menor a $1/4$?

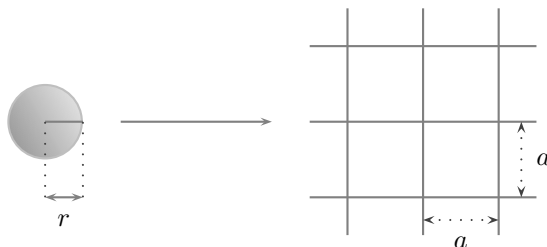


Figura 1.7

Solución. Primero debemos observar que es suficiente considerar lo que sucede únicamente en el cuadrado donde cae el centro de la moneda. No es difícil darse cuenta que la moneda no toca ninguna línea si su centro cae dentro del cuadrado interior que se muestra en la Figura 1.8.

Por lo tanto, si A denota el evento de ganar con un lanzamiento en este juego, entonces la probabilidad de A es el cociente entre el área favorable y el área total, es decir,

$$P(A) = \frac{(a - 2r)^2}{a^2} = \left(1 - \frac{2r}{a}\right)^2.$$

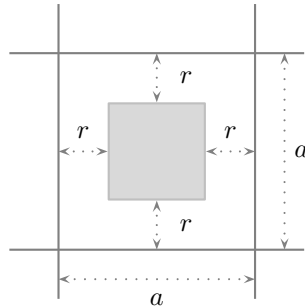


Figura 1.8

Si deseamos que esta probabilidad sea menor a $1/4$, entonces de aquí puede uno encontrar que a y r deben cumplir la relación $a < 4r$. Cuando $a = 4r$, la probabilidad de ganar es exactamente $1/4$. •

Ejemplo 1.12 (El problema de los dos amigos) Dos amigos deciden encontrarse en cierto lugar pero olvidan la hora exacta de la cita, únicamente recuerdan que la hora era entre las 12:00 y las 13:00 horas. Cada uno de ellos decide llegar al azar en ese lapso y esperar sólomente 10 minutos. ¿Cuál es la probabilidad de que los amigos se encuentren?

Solución. Sean x y y el tiempo medido en minutos en los que llegan los dos amigos. El espacio muestral del experimento consta de las parejas (x, y) tal que cada entrada de este vector es un instante entre las 12:00 y las 13:00 horas. Gráficamente este espacio muestral se puede representar como el cuadrado que se muestra en la Figura 1.9. Los amigos se encuentran si x y y son tales que $|x - y| \leq 10$, y esta región corresponde a la franja que se muestra en la Figura 1.9 y cuya área en minutos es $(60)^2 - (50)^2$. Si A denota el evento de interés, entonces tenemos que

$$P(A) = \frac{(60)^2 - (50)^2}{(60)^2} = \frac{36 - 25}{36} = \frac{11}{36}.$$

•

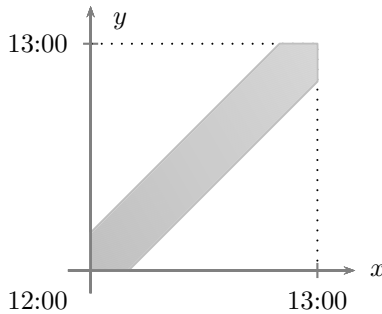


Figura 1.9

En la sección de ejercicios, el lector encontrará algunos otros problemas de aplicación de la probabilidad geométrica. Para encontrar la solución a estos problemas, y como sugerencia, se recomienda al lector, primeramente, determinar con claridad el experimento aleatorio en cuestión y establecer con precisión un espacio muestral adecuado a la pregunta que se desea contestar. Observemos finalmente que la probabilidad geométrica también cumple, entre otras, con las siguientes propiedades que hemos mencionado antes para la probabilidad clásica.

- a) $P(\Omega) = 1$.
- b) $P(A) \geq 0$ para cualquier evento A .
- c) $P(A \cup B) = P(A) + P(B)$ cuando A y B son ajenos.

Ejercicios

24. Suponga que se tiene un experimento aleatorio con espacio muestral $\Omega \subseteq \mathbb{R}^2$ cuya área está bien definida y es finita. Sean A y B dos eventos de este experimento cuya área está bien definida. Demuestre que la definición de la probabilidad geométrica satisface las siguientes propiedades:

- a) $P(\emptyset) = 0$.
 - b) $P(\Omega) = 1$.
 - c) $P(A) \geq 0$ para cualquier evento A .
 - d) $P(A^c) = 1 - P(A)$.
 - e) Si $A \subseteq B$ entonces $P(A) \leq P(B)$.
 - f) $P(A \cup B) = P(A) + P(B)$ cuando A y B son ajenos.
 - g) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
25. Se escogen dos números x y y al azar dentro del intervalo unitario $[0, 1]$. ¿Cuál es la probabilidad de que la suma de estos números sea mayor a uno y que, al mismo tiempo, la suma de sus cuadrados sea menor a uno?
26. Se escogen dos números al azar dentro del intervalo unitario $[0, 1]$. ¿Cuál es la probabilidad de que el producto de estos números sea menor a $1/2$?
27. Se escogen dos números x y y al azar dentro del intervalo unitario $[0, 1]$. ¿Cuál es la probabilidad de que la distancia
- a) de x a y sea menor a $1/2$?
 - b) de x a y sea mayor a $1/4$?
 - c) de x a cero sea, a lo sumo, la distancia de y a uno?
28. Se escoge un número a al azar dentro del intervalo $(-1, 1)$. ¿Cuál es la probabilidad de que la ecuación cuadrática $ax^2 + x + 1 = 0$ tenga dos raíces reales?
29. Se escogen dos números b y c al azar dentro del intervalo unitario $[0, 1]$. ¿Cuál es la probabilidad de que la ecuación cuadrática $x^2 + bx + c = 0$ tenga raíces complejas?
30. **El problema de la aguja de Buffón**³. Considere un conjunto infinito de líneas horizontales paralelas sobre una superficie plana como

³Georges-Louis Leclerc, Comte de Buffon (1707-1788), matemático, cosmólogo y naturalista francés.

se muestra en la Figura 1.10. La distancia entre una línea y otra es L . Se deja caer una aguja de longitud ℓ sobre la superficie. Suponga $\ell \leq L$. ¿Cuál es la probabilidad de que la aguja cruce alguna línea?

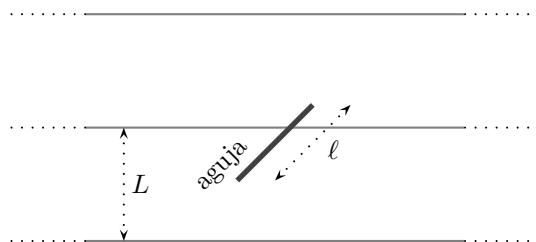


Figura 1.10

Como una simplificación, suponga el caso $\ell = L$ y sea A el evento que ocurre cuando la aguja toca alguna de las líneas. Si se efectúa este experimento n veces y n_A denota el número de ocurrencias del evento A , entonces el cociente n_A/n es cercano a $P(A)$. Así, tenemos la aproximación

$$\frac{2}{\pi} \approx \frac{n_A}{n},$$

de donde puede obtenerse una aproximación para π a partir del experimento simple de lanzar agujas en una superficie de líneas paralelas.

$$\pi \approx \frac{2n}{n_A}.$$

Suponga ahora $\ell \geq L$. Demuestre que la probabilidad de que la aguja cruce alguna línea es

$$p = 1 - \frac{2}{\pi} \arcsin\left(\frac{L}{\ell}\right) + \frac{2\ell}{\pi L} \left(1 - \cos\left(\arcsin\left(\frac{L}{\ell}\right)\right)\right).$$

Usando la identidad $\cos(\arcsin x) = \sqrt{1 - x^2}$, $-1 \leq x \leq 1$, la respuesta anterior se puede escribir también como sigue

$$\frac{2\ell}{\pi L} - \frac{2}{\pi L} (\sqrt{\ell^2 - L^2} + L \arcsin(\frac{L}{\ell})) + 1.$$

31. Se escogen al azar, y de manera independiente, tres números, a , b y c dentro del intervalo unitario $[0, 1]$. ¿Cuál es la probabilidad de que la suma de estos números sea menor a uno?
32. **El problema de la varilla de metal.** Una varilla de metal de longitud ℓ se rompe en dos puntos distintos escogidos al azar. ¿Cuál es la probabilidad de que los tres segmentos así obtenidos formen un triángulo? Véase la Figura 1.11.

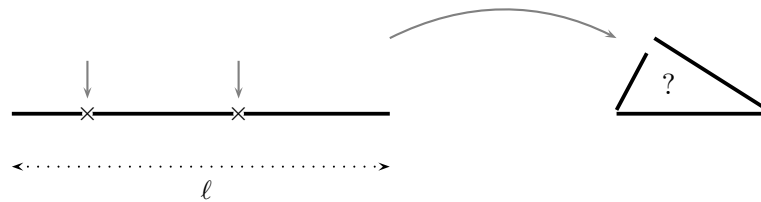


Figura 1.11

33. Un pasajero llega en autobús a la estación de trenes. La hora de llegada del autobús es aleatoria entre las 9:00 y 10:00 hrs. Por otro lado, el tren que debe tomar el pasajero parte de la estación también al azar entre las 9:00 y 10:00 hrs. El pasajero podrá subirse al tren si el autobús llega por lo menos cinco minutos antes de que el tren parta. ¿Cuál es la probabilidad de que el pasajero aborde el tren?
34. Considere nuevamente el problema de la feria del Ejemplo 1.11 en la página 24. Suponga ahora que el jugador gana si la moneda toca, a lo sumo, una línea. ¿Cuál es la probabilidad de ganar?
35. Dos personas tienen la misma probabilidad de llegar al lugar de su cita en cualquier instante dentro del intervalo de tiempo $[0, T]$ y llegan de manera independiente una de otra. Encuentre la probabilidad de que el tiempo que una persona tenga que esperar a la otra sea, a lo sumo, $t > 0$.
36. Suponga que se escoge un punto al azar dentro de un segmento de recta de longitud L de tal forma que la probabilidad de que el punto

caiga en un subsegmento es la misma de que caiga en cualquier otro subsegmento de la misma longitud. Calcule la probabilidad de que la distancia del punto al centro del segmento sea menor a ℓ .

37. Suponga que se escogen al azar, y de manera independiente, dos puntos dentro de un segmento de recta de longitud L de tal forma que la probabilidad de que cualquiera de los puntos caiga en un subsegmento es la misma de que caiga en cualquier otro subsegmento de la misma longitud. Calcule la probabilidad de que, por lo menos, uno de los puntos caiga en la primera mitad del intervalo.
38. Se escogen dos números al azar, de manera independiente uno del otro, dentro del intervalo $[0, L]$. Encuentre la probabilidad de que el promedio aritmético de estos dos números se encuentre dentro del subintervalo $[a, b] \subseteq [0, L]$.
39. Se escogen al azar, y de manera independiente, tres números reales dentro del intervalo $[0, L]$. Calcule la probabilidad de que el promedio aritmético de estos números sea menor a $L/3$.
40. **Triángulos 1.** Se escogen dos números x y y al azar de manera independiente uno del otro, dentro del intervalo $[0, \ell]$. Calcule la probabilidad de que las longitudes x , y y ℓ formen un triángulo.
41. **Triángulos 2.** Se escogen tres números x , y y z al azar, de manera independiente uno del otro, dentro del intervalo $[0, \ell]$. Calcule la probabilidad de que las longitudes x , y y z formen un triángulo.
42. **Rectángulos.** Suponga que x y y se escogen al azar dentro del intervalo unitario $[0, 1]$ y constituyen los lados de un rectángulo. Calcule la probabilidad de que el rectángulo
- a) tenga área mayor a $1/2$.
 - b) tenga perímetro menor a 1.
43. Se escoge un punto (x, y, z) al azar dentro del cubo $[-1, 1] \times [-1, 1] \times [-1, 1]$, de manera uniforme, es decir, con la misma probabilidad para dos regiones con el mismo volumen. Calcule la probabilidad del evento

$$A = \{(x, y, z) \in \Omega : |x| + |y| + |z| \leq 1\}.$$

1.6. Probabilidad frecuentista

Suponga que se realizan n repeticiones de un cierto experimento aleatorio y que se registra el número de veces que ocurre un determinado evento A . Esta información puede ser usada de la siguiente forma para definir la probabilidad de A .

Definición 1.4 Sea n_A el número de ocurrencias de un evento A en n realizaciones de un experimento aleatorio. La probabilidad frecuentista del evento A se define como el límite

$$P(A) = \lim_{n \rightarrow \infty} \frac{n_A}{n}.$$

En este caso, debemos hacer notar que no es humanamente posible llevar a cabo una infinidad de veces el experimento aleatorio y tampoco podemos garantizar, por ahora, la existencia de tal límite. Por lo tanto, mediante la definición anterior, no es posible encontrar de manera exacta la probabilidad de un evento cualquiera, aunque permite tener una aproximación empírica del valor de $P(A)$, es decir,

$$P(A) \approx \frac{n_A}{n}.$$

Las limitaciones mencionadas hacen que esta definición de probabilidad no sea enteramente formal, pero, sin duda, tiene la ventaja de que su forma de cálculo hace evidente la interpretación de la probabilidad como una medida de la frecuencia con la que ocurre un evento. Veamos un ejemplo concreto. Consideremos nuevamente el experimento aleatorio de lanzar un dado equilibrado y registrar la ocurrencia del evento A , definido como el conjunto $\{2, 4, 6\}$. Se ha llevado a cabo este experimento y después de lanzar el dado 20 veces, se obtuvieron los siguientes resultados.

Núm.	Resultado	n_A/n	Núm.	Resultado	n_A/n
1	3	0/1	11	2	7/11
2	6	1/2	12	5	7/12
3	2	2/3	13	1	7/13
4	1	2/4	14	6	8/14
5	4	3/5	15	3	8/15
6	6	4/6	16	1	8/16
7	3	4/7	17	5	8/17
8	4	5/8	18	5	8/18
9	2	6/9	19	2	9/19
10	5	6/10	20	6	10/20

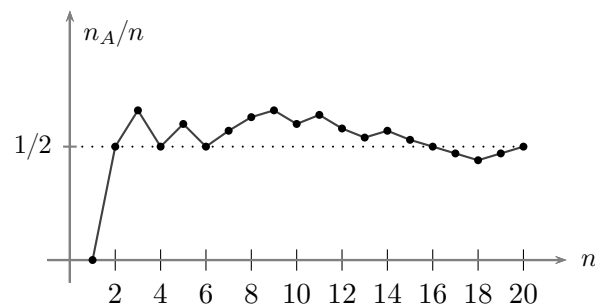


Figura 1.12

En la gráfica de la Figura 1.12 se muestra el singular comportamiento del cociente n_A/n a lo largo del tiempo, al principio se pueden presentar algunas oscilaciones, pero eventualmente el cociente parece estabilizarse en un cierto valor. Realizando un mayor número de observaciones del experimento, no es difícil verificar que el cociente n_A/n se estabiliza en $1/2$ cuando el dado está equilibrado y el número de ensayos n es grande. Se invita al lector intrigado a efectuar un experimento similar y corroborar esta interesante regularidad estadística con éste o cualquier otro experimento aleatorio de su interés. Más adelante formalizaremos este resultado mediante la así llamada ley de los grandes números que garantiza, en particular, que n_A/n efectivamente converge a la probabilidad del evento A . Esto se verifica en el Ejemplo 5.1, en la página 356.

Simulación 1.1 Arroje cien veces una moneda y registre los resultados en una lista. Calcule y grafique los cocientes n_A/n cuando el evento A corresponde a obtener alguna de las caras de la moneda. ¿Converge el cociente n_A/n a $1/2$? Para agilizar el experimento puede usted dividir los cien lanzamientos en varios grupos de personas y después juntar los resultados. Alternativamente, investigue la forma de simular este experimento en una computadora y calcule n_A/n para distintos valores de n . Véase el enunciado del Ejercicio 45. •

Para concluir esta pequeña sección, mencionaremos que es fácil comprobar que la probabilidad frecuentista también cumple las siguientes propiedades, que ya hemos mencionado antes tanto para la probabilidad clásica como para la probabilidad geométrica.

- a) $P(\Omega) = 1$.
- b) $P(A) \geq 0$ para cualquier evento A .
- c) $P(A \cup B) = P(A) + P(B)$ cuando A y B son ajenos.

Ejercicios

44. Sean A y B dos eventos de un experimento aleatorio. Demuestre que la definición de la probabilidad frecuentista satisface las siguientes propiedades:
- a) $P(\emptyset) = 0$.
 - b) $P(\Omega) = 1$.
 - c) $P(A) \geq 0$ para cualquier evento A .
 - d) $P(A^c) = 1 - P(A)$.
 - e) Si $A \subseteq B$ entonces $P(A) \leq P(B)$.
 - f) $P(A \cup B) = P(A) + P(B)$ cuando A y B son ajenos.
 - g) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

45. **Moneda.** El lanzamiento de una moneda puede simularse en R usando el comando `sample`, como se muestra en el recuadro de abajo. En el ejemplo mostrado aparece la instrucción

```
sample(0:1, 20, replace=TRUE)
```

que produce 20 selecciones al azar, una selección a la vez, de números dentro del conjunto o vector $\{0, 1\}$, en donde, naturalmente, se permite seleccionar un mismo número varias veces. Los valores 0 y 1 pueden asociarse, a conveniencia, a las dos caras de la moneda, y las probabilidades de selección son las mismas para cada uno de los números del conjunto indicado, en este caso es de $1/2$ para cada número. Realice 100 simulaciones del lanzamiento de una moneda equilibrada y compruebe, experimentalmente, que el número de veces que aparece una de las caras entre el total de lanzamientos se aproxima a $1/2$ conforme el número de lanzamientos crece.

```
# Simulación de 20 lanzamientos de una moneda equilibrada
> sample(0:1, 20, replace=TRUE)
[1] 0 1 0 1 0 0 1 1 1 0 1 0 0 0 0 1 1 1 0 1
```

46. **Dado.** El lanzamiento de un dado puede simularse en R usando el comando que aparece abajo. Realice 100 simulaciones del lanzamiento de un dado equilibrado y compruebe experimentalmente que el número de veces que aparece una de las caras entre el total de lanzamientos se aproxima a $1/6$ conforme el número de lanzamientos crece.

```
# Simulación de 20 lanzamientos de un dado equilibrado
> sample(1:6, 20, replace=TRUE)
[1] 1 5 3 4 4 6 3 4 3 3 2 4 4 5 3 1 6 5 6 6
```

1.7. Probabilidad subjetiva

En este caso, la probabilidad de un evento depende del observador, es decir, depende de lo que el observador conoce del fenómeno en estudio. Puede

parecer un tanto informal y poco seria esta definición de la probabilidad de un evento, sin embargo, en muchas situaciones es necesario recurrir a un experto para tener por lo menos una idea vaga de cómo se comporta el fenómeno de nuestro interés y saber si la probabilidad de un evento es alta o baja. Por ejemplo, ¿cuál es la probabilidad de que un cierto equipo de fútbol gane en su próximo partido? Ciertas circunstancias internas del equipo, las condiciones del equipo rival o cualquier otra condición externa, son elementos que sólo algunas personas conocen y que podrían darnos una idea más exacta de esta probabilidad. Esta forma subjetiva de asignar probabilidades a los distintos eventos debe, sin embargo, ser consistente con un conjunto de condiciones que estudiaremos a continuación.

1.8. Probabilidad axiomática

En la definición axiomática de la probabilidad no se establece la forma explícita de calcular las probabilidades, sino únicamente se proponen las reglas que el cálculo de probabilidades debe satisfacer. Los siguientes tres postulados o axiomas fueron establecidos en 1933 por el matemático ruso Andrey Nikolaevich Kolmogorov.

Axiomas de la probabilidad

1. $P(A) \geq 0$.
2. $P(\Omega) = 1$.
3. $P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$ cuando A_1, A_2, \dots son ajenos dos a dos.

Recordemos que un axioma o postulado es una proposición que se acepta como válida y sobre la cual se funda una teoría, en este caso la teoría de la probabilidad. En particular, estos postulados han sido tomados directamente del análisis cuidadoso y reflexivo de las definiciones de probabilidad mencionadas anteriormente. Y no es difícil verificar que las definiciones anteriores de probabilidad (clásica, geométrica y frecuentista) satisfacen estos

tres axiomas. Verificaremos a continuación el caso de la probabilidad frecuentista.

Ejemplo 1.13 Consideremos nuevamente la definición de probabilidad frecuentista, en donde debe realizarse una sucesión de n ensayos de un experimento aleatorio y calcular el cociente n_A/n para un evento A cualquiera. Se observa claramente que el cociente n_A/n es no negativo, esto es el primer axioma. Si A es el evento total Ω , entonces $n_\Omega = n$ y por lo tanto $n_\Omega/n = 1$, esto es el segundo axioma. Finalmente, si A y B son dos eventos ajenos, se tiene que $n_{A \cup B} = n_A + n_B$ y por lo tanto

$$\frac{n_{A \cup B}}{n} = \frac{n_A}{n} + \frac{n_B}{n},$$

lo cual puede extenderse para una colección numerable de eventos ajenos dos a dos y de esta manera obtener el tercer axioma, suponiendo la existencia de los límites correspondientes. De manera semejante puede verificarse el cumplimiento de estos axiomas para las definiciones clásica y geométrica de la probabilidad que hemos estudiado en las secciones anteriores. •

Tenemos a continuación la definición axiomática de la probabilidad.

Definición 1.5 A cualquier función P definida sobre una colección de eventos que satisfaga los tres axiomas de Kolmogorov, se le llama medida de probabilidad o, simplemente, probabilidad.

En esta definición no hemos sido muy precisos en establecer el dominio de la función P , sólo hemos dicho que tal dominio es una colección de eventos. En la siguiente sección especificaremos con más detalle a esta colección, por ahora nos interesa estudiar las propiedades generales que la función P pueda tener.

Propiedades de la probabilidad

Tomando como base los axiomas de Kolmogorov y usando la teoría elemental de conjuntos, demostraremos que toda medida de probabilidad cumple con una serie de propiedades generales e interesantes.

Proposición 1.1 $P(\emptyset) = 0$.

Demostración. Tomando $A_1 = A_2 = \dots = \emptyset$ en el tercer axioma, tenemos que

$$P(\emptyset) = P(\emptyset) + P(\emptyset) + \dots .$$

La única solución de esta ecuación es $P(\emptyset) = 0$. •

Proposición 1.2 Sea A_1, A_2, \dots, A_n una colección finita de eventos ajenos dos a dos. Entonces

$$P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k).$$

Demostración. Definiendo $A_{n+1} = A_{n+2} = \dots = \emptyset$, se verifica que esta sucesión infinita sigue siendo ajena dos a dos y por lo tanto, usando el tercer axioma, tenemos que

$$\begin{aligned} P\left(\bigcup_{k=1}^n A_k\right) &= P\left(\bigcup_{k=1}^{\infty} A_k\right) \\ &= \sum_{k=1}^{\infty} P(A_k) \\ &= \sum_{k=1}^n P(A_k). \end{aligned}$$

De esta manera, el tercer axioma incluye tanto el caso de sucesiones infinitas de eventos ajenos dos a dos como el caso de sucesiones finitas. En particular, •

cuando sólo tenemos dos eventos A y B con $A \cap B = \emptyset$, se cumple la identidad

$$P(A \cup B) = P(A) + P(B).$$

Proposición 1.3 Para cualquier evento A , $P(A^c) = 1 - P(A)$.

Demostración. De la teoría elemental de conjuntos tenemos que $\Omega = A \cup A^c$, en donde A y A^c son eventos ajenos. Aplicando el tercer axioma tenemos que

$$\begin{aligned} 1 &= P(\Omega) \\ &= P(A \cup A^c) \\ &= P(A) + P(A^c). \end{aligned}$$

•

La proposición recién demostrada establece que los eventos A y A^c tienen probabilidad complementaria, es decir, la suma de las probabilidades de estos eventos es siempre uno. Esta sencilla propiedad es bastante útil pues en ocasiones es más fácil calcular la probabilidad del complemento de un evento que del evento mismo. Más adelante tendremos múltiples ocasiones para aplicar este resultado.

Las siguientes dos proposiciones suponen la situación $A \subseteq B$, la cual se muestra gráficamente en la Figura 1.13. En particular, el siguiente resultado establece que la probabilidad es una función monótona no decreciente.

Proposición 1.4 Si $A \subseteq B$ entonces $P(A) \leq P(B)$.

Demostración. Primeramente escribimos $B = A \cup (B - A)$. Como A y $B - A$ son eventos ajenos, por el tercer axioma, $P(B) = P(A) + P(B - A)$. Usando el primer axioma concluimos que $P(B) - P(A) = P(B - A) \geq 0$. De aquí obtenemos $P(B) - P(A) \geq 0$. •

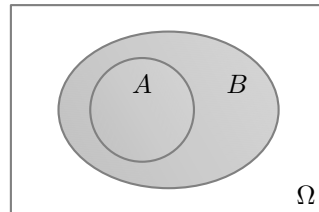


Figura 1.13

Proposición 1.5 Si $A \subseteq B$ entonces $P(B - A) = P(B) - P(A)$.

Demostración. Como $B = A \cup (B - A)$, siendo esta unión ajena, por el tercer axioma tenemos que $P(B) = P(A) + P(B - A)$. •

Por ejemplo, suponga que A y B son eventos tales que $A \subseteq B$, $P(A^c) = 0.9$ y $P(B^c) = 0.6$. Deseamos calcular $P(B - A)$. En esta situación es válida la fórmula $P(B - A) = P(B) - P(A)$, en donde $P(A) = 0.1$ y $P(B) = 0.4$. Por lo tanto, $P(B - A) = 0.4 - 0.1 = 0.3$. Observe que en este ejemplo sencillo no se especifica el experimento aleatorio en cuestión ni tampoco se definen explícitamente a los eventos A y B . El tratamiento es completamente analítico y los resultados son válidos para cualesquiera eventos A y B con las características señaladas.

Proposición 1.6 Para cualquier evento A , $0 \leq P(A) \leq 1$.

Demostración. Como $A \subseteq \Omega$, se tiene que $P(A) \leq P(\Omega) = 1$. La primera desigualdad, $0 \leq P(A)$, es simplemente el primer axioma. •

En palabras, la proposición anterior establece que la medida de probabilidad es una función que toma valores únicamente en el intervalo $[0, 1]$, y ello ha

sido consecuencia de los axiomas establecidos. El siguiente resultado proporciona una fórmula general para la probabilidad de la unión de cualesquiera dos eventos, no necesariamente ajenos.

Proposición 1.7 Para cualesquiera eventos A y B ,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Demostración. Primeramente observamos que para cualesquiera eventos A y B se cumple la igualdad $A - B = A - (A \cap B)$, en donde $A \cap B \subseteq A$, de modo que $P(A - (A \cap B)) = P(A) - P(A \cap B)$. Entonces escribimos a $A \cup B$ como la unión disjunta de los siguientes tres eventos:

$$\begin{aligned} A \cup B &= (A - B) \cup (A \cap B) \cup (B - A) \\ &= (A - (A \cap B)) \cup (A \cap B) \cup (B - (A \cap B)). \end{aligned}$$

Ahora aplicamos la probabilidad. Por el tercer axioma,

$$\begin{aligned} P(A \cup B) &= P(A - (A \cap B)) + P(A \cap B) + P(B - (A \cap B)) \\ &= P(A) - P(A \cap B) + P(A \cap B) + P(B) - P(A \cap B) \\ &= P(A) + P(B) - P(A \cap B). \end{aligned}$$

•

En la Figura 1.14 (a), el lector puede comprobar la validez de la fórmula recién demostrada identificando las tres regiones ajenas de las que consta el evento $A \cup B$. El término $P(A)$ abarca las primeras dos regiones de izquierda a derecha, $P(B)$ abarca la segunda y tercera región. Observe entonces que la región central ha sido contada dos veces de modo que el término $-P(A \cap B)$ da cuenta de ello. De esta forma, las tres regiones son contadas una sola vez y el resultado es la probabilidad del evento $A \cup B$.

Ejemplo 1.14 Sean A y B eventos ajenos tales que $P(B) = 0.3$ y $P(A \cap B^c) = 0.2$. Encuentre $P(A \cup B)$.

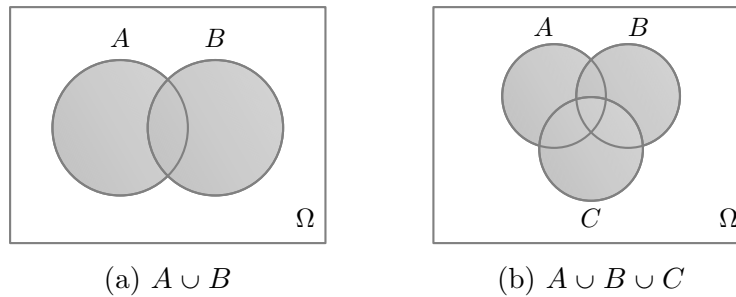


Figura 1.14

Solución. Usaremos la fórmula $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, en donde conocemos a $P(B)$; $P(A \cap B)$ es cero pues, por hipótesis, los eventos son ajenos, y $P(A) = P(A \cap B^c) = 0.2$. ¿Por qué? Por lo tanto, $P(A \cup B) = 0.2 + 0.3 = 0.5$. •

Como hemos señalado, la fórmula anterior para la probabilidad de la unión de dos eventos es válida para cualesquiera que sean estos eventos, sin embargo, cuando los eventos son ajenos, es decir, cuando $A \cap B = \emptyset$, entonces la fórmula demostrada se reduce al tercer axioma de la probabilidad en su versión para dos eventos ajenos, es decir,

$$P(A \cup B) = P(A) + P(B).$$

El siguiente resultado es una extensión natural de estas fórmulas e involucra tres eventos arbitrarios. La fórmula que a continuación se demuestra puede también verificarse usando el diagrama de Venn que aparece en la Fig 1.14 (b). Para ello se pueden seguir uno a uno los términos del lado derecho de la fórmula y comprobar que cada región es contada una sola vez, de modo que el resultado final es la probabilidad del evento $A \cup B \cup C$. La así llamada fórmula de inclusión-exclusión, que aparece en el Ejercicio 55, en la página 45, es una generalización de este resultado.

Proposición 1.8 Para cualesquiera eventos A , B y C ,

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) \\ &\quad - P(A \cap B) - P(A \cap C) - P(B \cap C) \\ &\quad + P(A \cap B \cap C). \end{aligned}$$

Demostración. Agrupando adecuadamente y usando la fórmula para dos eventos,

$$\begin{aligned} P(A \cup B \cup C) &= P[(A \cup B) \cup C] \\ &= P(A \cup B) + P(C) - P((A \cup B) \cap C) \\ &= P(A) + P(B) - P(A \cap B) + P(C) \\ &\quad - P((A \cap C) \cup (B \cap C)) \\ &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) \\ &\quad - P(B \cap C) + P(A \cap B \cap C). \end{aligned}$$

•

En particular, cuando A , B y C son mutuamente ajenos, la fórmula demostrada se reduce a la aplicación del tercer axioma, es decir,

$$P(A \cup B \cup C) = P(A) + P(B) + P(C).$$

Las propiedades anteriores son parte del estudio teórico y general de la probabilidad. En general, supondremos que la forma explícita de calcular estos números es conocida, o que se puede suponer un cierto modelo para llevar a cabo estos cálculos, dependiendo del experimento aleatorio en cuestión. Por ejemplo, cuando el espacio muestral es finito y cada resultado puede suponerse igualmente probable, entonces usaremos la definición clásica de probabilidad. En otras situaciones asignaremos probabilidades de acuerdo a ciertos modelos que especificaremos más adelante. Se espera que, a partir de la lectura cuidadosa de las propiedades enunciadas y demostradas, el lector haya podido desarrollar cierta habilidad e intuición para escribir la demostración de alguna otra propiedad general de la probabilidad. Algunas

de estas propiedades adicionales pueden encontrarse en la sección de ejercicios. Se debe también señalar que las demostraciones no son únicas y que es muy probable que se puedan producir demostraciones diferentes a las que aquí se han presentado.

Ejercicios

47. **Otras propiedades de la medida de probabilidad.** Demuestre las siguientes afirmaciones:

a) $P(A \cap B \cap C) \leq P(A \cap B) \leq P(A)$.

b) $P(A) \leq P(A \cup B) \leq P(A \cup B \cup C)$.

c) $P(A_1 \cup A_2) \leq P(A_1) + P(A_2)$.

d) $P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i)$.

e) $P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i)$.

f) $P(A - B) = P(A) - P(A \cap B)$.

g) $P(A \Delta B) = P(A) + P(B) - 2P(A \cap B)$.

h) Si $A = B$ entonces $P(A \Delta B) = 0$.

i) $P(A_1 \cap A_2) \geq 1 - P(A_1^c) - P(A_2^c)$.

j) $P(A_1 \cap A_2 \cap \cdots \cap A_n) \geq 1 - \sum_{i=1}^n P(A_i^c)$.

k) Si $A_1 \cap A_2 \subseteq A$ entonces $P(A) \geq P(A_1) + P(A_2) - 1$.

l) Si $\bigcap_{i=1}^n A_i \subseteq A$ entonces $P(A) \geq \sum_{i=1}^n P(A_i) - (n - 1)$.

m) Si $A_1 \cap A_2 \subseteq A$ entonces $P(A^c) \leq P(A_1^c) + P(A_2^c)$.

n) Si $\bigcap_{i=1}^n A_i \subseteq A$ entonces $P(A^c) \leq \sum_{i=1}^n P(A_i^c)$.

\tilde{n}) Si A_1, \dots, A_n tienen probabilidad cero entonces $P\left(\bigcup_{i=1}^n A_i\right) = 0$.

- o) Si A_1, \dots, A_n tienen probabilidad uno entonces $P\left(\bigcap_{i=1}^n A_i\right) = 1$.
- p) $\max\{P(A_1), P(A_2)\} \leq P(A_1 \cup A_2) \leq \min\{P(A_1) + P(A_2), 1\}$.
- q) $\max\{P(A_1), \dots, P(A_n)\} \leq P\left(\bigcup_{i=1}^n A_i\right) \leq \min\left\{\sum_{i=1}^n P(A_i), 1\right\}$.

48. Demuestre o proporcione un contraejemplo:

- a) Si $P(A) = 0$ entonces $A = \emptyset$.
- b) Si $P(A) = 1$ entonces $A = \Omega$.
- c) Si $P(A) = 0$ entonces $P(A \cap B) = 0$.
- d) Si $P(A) = P(B)$ entonces $A = B$.
- e) Si $P(A \Delta B) = 0$ entonces $A = B$.
- f) Si $P(A) = P(B)$ entonces $P(A \Delta B) = 0$.
- g) Si $P(A \Delta B) = 0$ entonces $P(A) = P(B)$.
- h) Si $P(A) \leq P(B)$ entonces $A \subseteq B$.
- i) Si $P(A) = P(B) = p$ entonces $P(A \cap B) \leq p^2$.
- j) Si $A \cap B = \emptyset$ entonces $P(A) \leq P(B^c)$.
- k) Si $P(A) = P(B) = 0$ entonces $P(A \cup B) = 0$.
- l) Si $P(A) = P(B) = 1$ entonces $P(A \cap B) = 1$.
- m) Si $A \cap B \subseteq C$ entonces $P(C^c) \leq P(A^c) + P(B^c)$.
- n) Si $P(A) > 1/2$ y $P(B) > 1/2$ entonces $P(A \cap B) > 0$.

49. Suponga que un cierto experimento aleatorio tiene como espacio muestral el conjunto finito $\Omega = \{1, \dots, n\}$. Determine si las siguientes funciones son medidas de probabilidad: para cualquier $A \subseteq \Omega$ se define

- a) $P(A) = \sum_{k \in A} \frac{2k}{n(n+1)}$.
- b) $P(A) = \sum_{k \in A} \frac{2^k}{2^{n+1} - 2}$, $n > 1$.
- c) $P(A) = \prod_{k \in A} \frac{k(n+1)}{k+1}$.

50. Sean A y B eventos tales que $P(A \cap B^c) = 0.2$ y $P(B^c) = 0.5$. Encuentre $P(A \cup B)$.

51. Sean A y B eventos ajenos tales que $P(A) = 0.3$ y $P(B) = 0.4$. Encuentre:

- a) $P(A \cup B)$. c) $P(A^c \cup B^c)$. e) $P(A^c \cap B)$.
 b) $P(A \cap B)$. d) $P(A^c \cap B^c)$. f) $P(A \cap B^c)$.

52. Determine si es posible una asignación de probabilidades para los eventos A y B tal que $P(A) = 1/2$, $P(B) = 1/4$ y $P(A \cap B) = 1/3$.

53. Sean A y B eventos tales que $P(A) = p$, $P(B) = q$ y $P(A \cap B) = r$. Encuentre:

- a) $P(A \cap B^c)$. c) $P(A^c \cap B^c)$.
 b) $P(A^c \cap B)$. d) $P(A \Delta B)$.

54. Sea A , B y C tres eventos tales que $P(A) = P(B) = 1/3$, $P(C) = 1/4$, $P(A \cap B) = 1/6$ y $P(B \cap C) = 0$. Encuentre $P(A \cap B \cap C)$.

55. **Fórmula de inclusión-exclusión** Use el método de inducción para demostrar que si A_1, A_2, \dots, A_n son eventos arbitrarios, entonces

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{i \neq j} P(A_i \cap A_j) + \sum_{\substack{i,j,k \\ \text{distintos}}} P(A_i \cap A_j \cap A_k) - \dots + (-1)^{n-1} P(A_1 \cap \dots \cap A_n).$$

56. Demuestre que la probabilidad de que exactamente uno de los eventos A o B ocurra es

$$P(A) + P(B) - 2P(A \cap B).$$

57. Sean A , B y C tres eventos. Demuestre que la probabilidad de que exactamente

- a) uno de estos eventos ocurra es
 $P(A) + P(B) + P(C) - 2P(A \cap B) - 2P(A \cap C) - 2P(B \cap C) + 3P(A \cap B \cap C)$.
- b) dos de estos eventos ocurran es
 $P(A \cap B) + P(A \cap C) + P(B \cap C) - 3P(A \cap B \cap C)$.
- c) tres de estos eventos ocurran es $P(A \cap B \cap C)$.

¿Cuál es el evento que se obtiene al unir estas tres condiciones?

58. Sean P y Q dos medidas de probabilidad definidas sobre la misma colección de eventos. Demuestre que, para cada $\alpha \in [0, 1]$, la función $\alpha P + (1 - \alpha)Q$ es también una medida de probabilidad.
59. La función conjuntista P asigna un número (probabilidad) a cada subconjunto A de \mathbb{R} , de la siguiente forma:

$$P(A) = \int_A f(x) dx,$$

en donde

$$f(x) = \begin{cases} 2x & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso,} \end{cases}$$

y para aquellos conjuntos A en donde la integral esté bien definida. Calcule la probabilidad de los siguientes conjuntos:

- a) $A_1 = \{x \in \mathbb{R} : 1/2 < x < 3/4\}$.
- b) $A_2 = \{x \in \mathbb{R} : x = 1/4\}$.
- c) $A_3 = \{x \in \mathbb{R} : 0 < x < 10\}$.

1.9. Sigmas álgebras

En esta breve sección vamos a definir un conjunto que agrupa a todos los eventos de un mismo experimento aleatorio y para los cuales se puede definir o calcular sus probabilidades. En este texto elemental no haremos mayor énfasis en este tipo de colecciones de eventos, de modo que el lector no debe

preocuparse si encuentra el material de esta sección, y de la siguiente, de un nivel de abstracción ligeramente más elevado que el que hasta ahora se ha supuesto en las secciones anteriores. Sin embargo, debe advertirse también que los conceptos que se presentan en estas dos secciones son fundamentales para un tratamiento completo de la teoría matemática de la probabilidad.

Definición 1.6 Una colección \mathcal{F} de subconjuntos de un espacio muestral Ω es una álgebra si cumple las tres condiciones siguientes:

1. $\Omega \in \mathcal{F}$.
2. Si $A \in \mathcal{F}$ entonces $A^c \in \mathcal{F}$.
3. Si $A_1, A_2, \dots, A_n \in \mathcal{F}$ entonces $\bigcup_{k=1}^n A_k \in \mathcal{F}$.

Veamos con más detenimiento las condiciones que aparecen en la definición anterior. La primera condición establece que el espacio muestral en su totalidad debe pertenecer a la colección \mathcal{F} . Ciertamente el espacio muestral es un evento que siempre ocurre y como hemos visto su probabilidad está bien definida y es uno. La segunda condición asegura que si algún subconjunto A es de interés y por lo tanto se le considera un evento, entonces el complemento de tal conjunto también debe ser un evento. Nuevamente, por lo estudiado antes, la probabilidad del evento A^c está siempre dada por $1 - P(A)$.

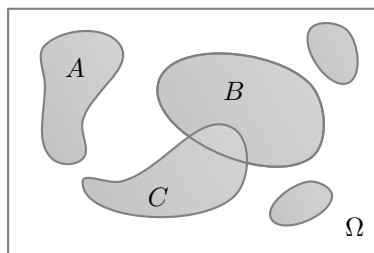


Figura 1.15

Finalmente, el tercer requisito en la definición establece que si se tiene una sucesión finita de eventos, entonces la unión de todos ellos también debe ser un evento, el cual corresponde a la ocurrencia de por lo menos uno de los eventos de la sucesión. Cuando esta tercera condición es también válida para sucesiones infinitas de eventos, a la correspondiente colección de subconjuntos de Ω se le llama σ -álgebra, en donde el prefijo σ se refiere a la operación infinita involucrada. La definición es entonces muy parecida a la anterior.

Definición 1.7 Una colección \mathcal{F} de subconjuntos de un espacio muestral Ω es una σ -álgebra si cumple las tres condiciones siguientes:

1. $\Omega \in \mathcal{F}$.
2. Si $A \in \mathcal{F}$ entonces $A^c \in \mathcal{F}$.
3. Si $A_1, A_2, \dots \in \mathcal{F}$ entonces $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

A los elementos de \mathcal{F} se les llama eventos.

Por lo tanto, a partir de ahora no llamaremos evento a cualquier subconjunto del espacio muestral, sino únicamente a aquellos elementos que pertenezcan a una σ -álgebra asociada al espacio muestral. En el siguiente ejemplo se ilustra el hecho de que pueden existir varias σ -álgebras asociadas a un mismo espacio muestral, aunque en nuestro caso sólo trabajaremos con una sola σ -álgebra de eventos a la vez y, con frecuencia, no la especificaremos con detalle. En la Figura 1.16 se muestra la relación existente entre álgebra y σ -álgebra de un mismo espacio muestral: toda σ -álgebra es una álgebra. Al respecto véanse los Ejercicios 62 y 63 en la página 50. En la siguiente sección estudiaremos un ejemplo de σ -álgebra de cierta importancia: la σ -álgebra de Borel de subconjuntos de \mathbb{R} .

Ejemplo 1.15 Sea Ω el espacio muestral de un experimento aleatorio. No es difícil comprobar que las siguientes colecciones de subconjuntos de Ω son σ -álgebras:

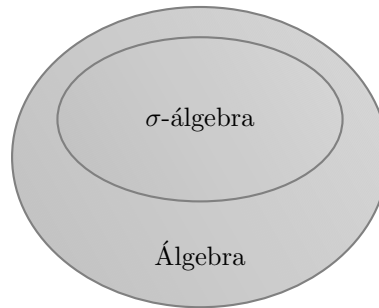


Figura 1.16

- a) $\mathcal{F} = \{\emptyset, \Omega\}$.
- b) $\mathcal{F} = \{A, A^c, \emptyset, \Omega\}$, en donde $A \subseteq \Omega$.
- c) $\mathcal{F} = 2^\Omega$.

•

De esta forma, en una σ -álgebra \mathcal{F} se agrupa a todos los subconjuntos de Ω para los que estamos interesados en calcular su probabilidad, y tal colección constituye el dominio sobre el cual se define una medida de probabilidad. Así, a cada experimento aleatorio particular se le puede asociar una pareja (Ω, \mathcal{F}) compuesta por el espacio muestral y una σ -álgebra de eventos.

Ejercicios

60. Sean A y B dos eventos. Demuestre que los siguientes conjuntos también son eventos, es decir, pertenecen a la misma σ -álgebra:

- | | | |
|------------------|-------------------|-----------------------|
| a) \emptyset . | c) $A - B$. | e) $A \Delta B$. |
| b) $A \cap B$. | d) $A \cup B^c$. | f) $A - (A \cap B)$. |

61. Sean A y B dos subconjuntos arbitrarios del espacio muestral Ω de un experimento aleatorio. ¿Qué conjuntos es necesario añadir a la colección $\{A, B\}$ para convertirla en la mínima σ -álgebra de subconjuntos de Ω que contiene a A y a B ?
62. Demuestre que toda σ -álgebra es álgebra. Mediante un contraejemplo demuestre que el recíproco es falso, es decir, no toda álgebra es una σ -álgebra. Véase la Figura 1.16. Compare este resultado general con el que se presenta en el siguiente ejercicio.
63. Sea Ω un espacio muestral finito. Demuestre que toda álgebra de subconjuntos de Ω es también una σ -álgebra.
64. Sean \mathcal{F}_1 y \mathcal{F}_2 dos σ -álgebras de subconjuntos de Ω .
- Escriba la definición de $\mathcal{F}_1 \cap \mathcal{F}_2$.
 - Demuestre que $\mathcal{F}_1 \cap \mathcal{F}_2$ es una σ -álgebra.
 - Mediante un contraejemplo demuestre que $\mathcal{F}_1 \cup \mathcal{F}_2$ no necesariamente es una σ -álgebra.
65. En un juego de tiro al blanco se pueden obtener 0 puntos, 1 punto, 2 puntos o 3 puntos. Defina el evento A_n como aquel en el que se obtienen n puntos al efectuar un tiro al blanco. Claramente los eventos A_0, A_1, A_2 y A_3 son ajenos dos a dos y constituyen una partición del espacio muestral de este experimento aleatorio. Encuentre una σ -álgebra que contenga a estos cuatro eventos simples.

1.10. Sigma álgebra de Borel

En esta sección estudiaremos brevemente un ejemplo de σ -álgebra que nos será de utilidad. Tomaremos como espacio muestral el conjunto de números reales \mathbb{R} y consideraremos la colección de intervalos de la forma $(-\infty, x]$ para cualquier número real x , es decir, sea

$$\mathcal{C} = \{(-\infty, x] : x \in \mathbb{R}\}.$$

Esta colección contiene un número infinito no numerable de elementos, pero es claro que no constituye una σ -álgebra de subconjuntos de \mathbb{R} pues, por ejemplo, no es cerrada bajo la operación de tomar complementos. La σ -álgebra de Borel de \mathbb{R} se puede construir a partir de esta colección considerando la σ -álgebra más pequeña de subconjuntos de \mathbb{R} que contiene a la colección \mathcal{C} .

Definición 1.8 La σ -álgebra de Borel de \mathbb{R} se denota por $\mathcal{B}(\mathbb{R})$ y se define como la mínima σ -álgebra de subconjuntos de \mathbb{R} que contiene a todos los intervalos de la forma $(-\infty, x]$. Esto se escribe de la forma siguiente

$$\mathcal{B}(\mathbb{R}) := \sigma\{(-\infty, x] : x \in \mathbb{R}\}.$$

A los elementos de $\mathcal{B}(\mathbb{R})$ se les llama conjuntos de Borel, conjuntos Borel medibles o simplemente borelianos de \mathbb{R} .

El símbolo σ en la expresión anterior significa que se está tomando la mínima σ -álgebra generada por la colección $\{(-\infty, x] : x \in \mathbb{R}\}$, y el adjetivo mínimo significa que si \mathcal{F} es una σ -álgebra que contiene a la colección \mathcal{C} , entonces $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{F}$, es decir, $\mathcal{B}(\mathbb{R})$ es la más pequeña. Existen otras formas equivalentes de definir a $\mathcal{B}(\mathbb{R})$, pero la que hemos presentado es suficiente para los propósitos de este texto. Es interesante mencionar también que el concepto de σ -álgebra de Borel puede extenderse a \mathbb{R}^n y aun a espacios más generales.

Para entender mejor a la σ -álgebra de Borel de \mathbb{R} , mostraremos a continuación algunos otros elementos que pertenecen a esta colección de subconjuntos de \mathbb{R} . Expresaremos a estos elementos como el resultado de operaciones de elementos que sabemos que pertenecen a $\mathcal{B}(\mathbb{R})$ y en donde las operaciones son aquellas bajo las cuales toda σ -álgebra es cerrada, según la definición. En la siguiente tabla, x y y son cualesquiera números reales tales que $x \leq y$.

Subconjunto	Se expresa como
(x, ∞)	$(-\infty, x]^c$
$(-\infty, x)$	$\bigcup_{n=1}^{\infty} (-\infty, x - 1/n]$
$[x, \infty)$	$(-\infty, x)^c$
$[x, y]$	$(-\infty, y] - (-\infty, x)$
$[x, y)$	$(-\infty, y) - (-\infty, x)$
$(x, y]$	$(-\infty, y] - (-\infty, x]$
(x, y)	$(-\infty, y) - (-\infty, x]$
$\{x\}$	$(-\infty, x] - (-\infty, x)$
\mathbb{N}	$\bigcup_{n=1}^{\infty} \{n\}$
\mathbb{Z}	$\bigcup_{n=-\infty}^{\infty} \{n\}$
\mathbb{Q}	$\bigcup_{n,m=-\infty}^{\infty} \{n/m\}$
\mathbb{I} (Irracionales)	\mathbb{Q}^c

Así, todos los conjuntos en la lista anterior son elementos de $\mathcal{B}(\mathbb{R})$, añadidos con los que se puedan producir a partir de ellos. Uno podría entonces pensar que todo subconjunto de \mathbb{R} es un boreliano, pero ello no es cierto ya que se pueden construir subconjuntos de \mathbb{R} que no son borelianos. Una vez que estudiemos en el siguiente capítulo el concepto de variable aleatoria, estaremos considerando siempre implícitamente conjuntos de Borel de \mathbb{R} como eventos.

1.11. Espacios de probabilidad

Con los elementos antes estudiados podemos ahora definir formalmente la estructura matemática diseñada para modelar un experimento aleatorio.

Definición 1.9 Un espacio de probabilidad es una terna (Ω, \mathcal{F}, P) en donde Ω es un conjunto arbitrario, \mathcal{F} es una σ -álgebra de subconjuntos de Ω y P es una medida de probabilidad definida sobre \mathcal{F} .

El conjunto arbitrario Ω representa usualmente el espacio muestral de un experimento aleatorio e inicialmente tal conjunto no tiene ninguna estructura matemática asociada, pues sus elementos pueden ser de muy diversa naturaleza. Este conjunto puede estar constituido por números (mediciones), personas, objetos, categorías, etc. Hemos señalado que no existe necesariamente un único espacio muestral para un experimento aleatorio, pero, en la mayoría de los casos, su especificación queda entendida de manera implícita.

Hemos mencionado también, en la sección anterior, que la σ -álgebra \mathcal{F} tiene el objetivo de agrupar en una sola colección a todos los subconjuntos de Ω , llamados eventos, para los cuales uno está interesado en definir o calcular su probabilidad. Así, la σ -álgebra es una clase de subconjuntos de Ω y existen varias σ -álgebras que uno puede asociar a un mismo espacio muestral.

Finalmente, la medida de probabilidad P es una función definida sobre la σ -álgebra. Tal función no se especifica de manera totalmente explícita, pero se le pide que cumpla los axiomas de Kolmogorov. Así, no existe siempre una única medida de probabilidad asignada, sino que ésta es general. Esta función indica la probabilidad de ocurrencia de cada uno de los eventos contenidos en la σ -álgebra \mathcal{F} , sin especificar la forma concreta en que estas probabilidades son calculadas.

Supondremos entonces que para cada experimento aleatorio existe un espacio de probabilidad asociado (Ω, \mathcal{F}, P) , el cual no es necesariamente único. Esta terna es un modelo matemático cuyo objetivo es capturar los elementos esenciales para el estudio científico del experimento aleatorio. Para la mayoría de los experimentos aleatorios que revisaremos, no se especifica con total detalle cada uno de los elementos del espacio de probabilidad, sino que éstos son entendidos de manera implícita, con lo cual se corre el riesgo de provocar alguna confusión o aparente paradoja en determinados casos y ello

no es raro que ocurra. Por lo tanto, si uno desea ser muy preciso en la definición del modelo a utilizar, uno debe especificar completamente los tres elementos del espacio de probabilidad.

Tomaremos también la perspectiva de no pensar demasiado en los experimentos aleatorios particulares que puedan estar detrás de un modelo. Es decir, nos ocuparemos del estudio y consecuencias del modelo matemático sin la preocupación de pensar en que tal modelo puede representar un experimento aleatorio concreto. La experiencia ha demostrado que es muy provechoso este punto de vista, pues los resultados o consideraciones abstractas pueden luego ser aplicadas con cierta facilidad a situaciones particulares. A lo largo de este texto tendremos múltiples oportunidades de ilustrar esta situación. Por ejemplo, en el tercer capítulo definiremos de manera genérica varias distribuciones o modelos de probabilidad que pueden ser aplicados en muy distintos contextos.

Ejercicios

66. Suponga que un experimento aleatorio tiene como espacio muestral el conjunto $\Omega = \{1, 2, \dots\}$ y se define a la σ -álgebra de eventos, para este experimento, como el conjunto potencia $\mathcal{F} = 2^\Omega$. Demuestre que la función $P : \mathcal{F} \rightarrow [0, 1]$, definida como aparece abajo, es una medida de probabilidad, es decir, cumple los tres axiomas de Kolmogorov.

$$P(\{k\}) = \frac{1}{k(k+1)} \quad k = 1, 2, \dots$$

Así, (Ω, \mathcal{F}, P) es un espacio de probabilidad para este experimento aleatorio. Calcule además las siguientes probabilidades:

- | | |
|-----------------------------|------------------------------|
| a) $P(\{1, \dots, n\})$. | c) $P(\{1, 3, 5, \dots\})$. |
| b) $P(\{n, n+1, \dots\})$. | d) $P(\{2, 4, 6, \dots\})$. |

67. Para cada intervalo $A = (a, b) \subseteq \mathbb{R}$ se define la función

$$P(A) = \int_a^b f(x) dx,$$

en donde

$$f(x) = \begin{cases} 2x & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

La definición de la función P puede extenderse de manera única a todos los conjuntos de Borel de \mathbb{R} . Compruebe que P cumple los tres axiomas de Kolmogorov y por lo tanto es una medida de probabilidad. El espacio muestral es el conjunto de número reales y los eventos son los distintos subconjuntos de \mathbb{R} que se obtienen a partir de los intervalos (a, b) . Calcule además la probabilidad de los siguientes eventos:

- | | |
|-----------------|----------------------|
| a) $(0, 1/2)$. | c) $(1/3, 2/3)$. |
| b) $(-1, 1)$. | d) $(1/2, \infty)$. |

68. Determine de manera completa un posible espacio de probabilidad (Ω, \mathcal{F}, P) para cada uno de los siguientes experimentos aleatorios.

- Observar el resultado del lanzamiento de un dado equilibrado.
- Observar el marcador final de un partido de fútbol.
- Observar el número de integrantes de una familia escogida al azar.
- Escoger un número al azar dentro del intervalo unitario $(0, 1)$.
- Observar la posición en la que cae un dardo lanzado sobre un círculo de radio unitario.

1.12. Análisis combinatorio

Consideraremos ahora el caso cuando el experimento aleatorio es tal que su espacio muestral es un conjunto finito y cada elemento de este conjunto tiene la misma probabilidad de ocurrir, es decir, cuando el espacio Ω es equiprobable. En estos casos hemos definido la probabilidad clásica de un evento A de la siguiente forma

$$P(A) = \frac{\#A}{\#\Omega}.$$

Para poder aplicar esta definición, necesitamos saber contar cuántos elementos tiene un evento A cualquiera. Cuando podemos poner en una lista todos y cada uno de los elementos de dicho conjunto, entonces es fácil conocer la cardinalidad de A : simplemente contamos todos los elementos uno por uno. Sin embargo, es común enfrentar situaciones en donde no es factible escribir en una lista cada elemento de A , por ejemplo, ¿cuántos números telefónicos existen que contengan por lo menos un cinco? Es poco factible que alguien intente escribir uno a uno todos estos números telefónicos y encuentre de esta manera la cantidad buscada. En las siguientes secciones estudiaremos algunas técnicas de conteo que nos ayudarán a calcular la cardinalidad de un evento A en ciertos casos particulares. El principio de multiplicación que se presenta a continuación es la base de muchos de los cálculos en las técnicas de conteo.

Proposición 1.9 (Principio de multiplicación) Si un procedimiento A_1 puede efectuarse de n formas distintas y un segundo procedimiento A_2 puede realizarse de m formas diferentes, entonces el total de formas en que puede efectuarse el primer procedimiento, seguido del segundo, es el producto $n \cdot m$, es decir,

$$\#(A_1 \times A_2) = \#A_1 \cdot \#A_2.$$

Para ilustrar este resultado considere el siguiente ejemplo: suponga que un cierto experimento aleatorio consiste en lanzar un dado y después seleccionar al azar una letra del alfabeto. ¿Cuál es la cardinalidad del correspondiente espacio muestral? El experimento de lanzar un dado tiene 6 resultados posibles y consideremos que tenemos un alfabeto de 26 letras. El correspondiente espacio muestral tiene entonces cardinalidad $6 \times 26 = 156$.

El principio de multiplicación es válido no solamente para dos procedimientos, sino que también vale para cualquier sucesión finita de procedimientos. Por ejemplo, si A_1, A_2, \dots, A_k denotan k procedimientos sucesivos, entonces este principio se puede enunciar, en símbolos, de la forma siguiente:

$$\#(A_1 \times \dots \times A_k) = \#A_1 \cdots \#A_k.$$

Ejemplo 1.16 Un hombre tiene 4 pantalones distintos, 6 camisas y dos pares de zapatos. ¿De cuántas formas distintas puede el hombre vestirse con estas prendas?

Solución. El hombre se puede vestir de manera distinta durante $4 \times 6 \times 2 = 48$ días sin repetir una combinación. •

Vamos a considerar a continuación diferentes esquemas y contextos en donde es posible encontrar una fórmula matemática para ciertos problemas de conteo. En todos ellos aplicaremos el principio de multiplicación. El esquema general es el de extraer al azar k objetos de una urna con n objetos distintos. Especificaremos más adelante varias formas en las que tal extracción puede llevarse a cabo. En cualquier caso, a la colección de objetos seleccionados le llamaremos muestra. El esquema general se ilustra en la Figura 1.17.

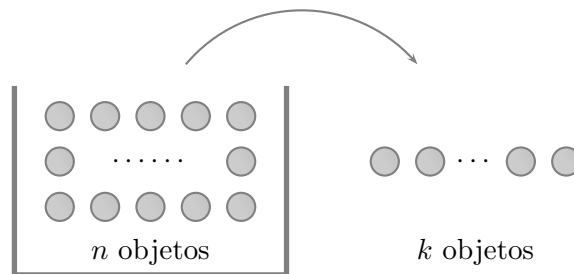


Figura 1.17

Ordenaciones con repetición: muestras con orden y con reemplazo

Suponga que tenemos una urna con n objetos distintos. Deseamos realizar k extracciones al azar de un objeto a la vez. Al efectuar una extracción, registramos el objeto escogido y lo regresamos a la urna, de esta forma el mismo objeto puede ser extraído varias veces. El total de arreglos que se pueden obtener de esta urna al hacer k extracciones es el número n^k , pues en cada extracción tenemos n objetos posibles para escoger y efectuamos k

extracciones. Esta fórmula es consecuencia del principio de multiplicación enunciado antes. A este número se le llama ordenaciones con repetición. Se dice que la muestra es con orden pues es importante el orden en el que se van obteniendo los objetos, y es con reemplazo pues cada objeto seleccionado se reincorpora a la urna.

Ejemplo 1.17 Suponga que tenemos un conjunto de 60 caracteres diferentes. Este conjunto contiene algunas letras minúsculas del alfabeto, sus correspondientes letras mayúsculas y los diez dígitos. ¿Cuántos *passwords* o palabras clave de longitud 4 se pueden construir usando este conjunto de 60 caracteres? Este es un ejemplo de una ordenación de 60 caracteres en donde se permiten las repeticiones. Como cada caracter de los 60 disponibles puede ser escogido para ser colocado en cada una de las cuatro posiciones de la palabra clave, entonces se pueden construir $60 \times 60 \times 60 \times 60 = 60^4 = 12,960,000$ *passwords* distintos de longitud 4. •

Ordenaciones sin repetición: muestras con orden y sin reemplazo

Suponga que se tiene la misma situación que antes: una urna con n objetos y de los cuales se deben extraer, uno a uno, k objetos. Suponga esta vez que la selección es sin reemplazo, es decir, una vez seleccionado un objeto, éste ya no se reincorpora a la urna. El total de arreglos distintos que se pueden obtener de este modo es el número

$$n(n-1)(n-2) \cdots (n-k+1).$$

Primeramente debemos observar que hay k factores en la expresión anterior. El primer factor es n y ello es debido a que tenemos cualesquiera de los n objetos para ser colocado en primera posición, para la segunda posición tenemos ahora $n-1$ objetos, para la tercera $n-2$ objetos, y así sucesivamente. Este razonamiento termina al escoger el k -ésimo objeto, para el cual tenemos únicamente $n-k+1$ posibilidades. Nuevamente, por el principio de multiplicación, la respuesta es el producto indicado. La expresión encontrada puede escribirse como sigue:

$$\frac{n!}{(n-k)!}$$

y se le llama permutaciones de n en k . En el caso particular cuando la selección es exhaustiva, es decir, cuando $k = n$, o bien cuando todos los objetos son extraídos uno por uno, entonces se tienen todas las permutaciones o distintos órdenes en que se pueden colocar n objetos, es decir, $n!$

Ejemplo 1.18 ¿De cuántas formas distintas pueden asignarse los premios primero, segundo y tercero en una rifa de 10 boletos numerados del 1 al 10?

Solución. Claramente se trata de una ordenación sin repetición de 10 objetos en donde se deben extraer 3 de ellos. La respuesta es entonces que existen $10 \times 9 \times 8 = 720$ asignaciones distintas para los tres primeros lugares en la rifa. •

Permutaciones: muestras exhaustivas con orden y sin reemplazo

La pregunta básica acerca del total de formas en que podemos poner en orden lineal (uno detrás de otro y por lo tanto no hay repetición) n objetos distintos tiene como respuesta el factorial de n , denotado por $n!$ y definido como sigue

$$n! = n(n-1)(n-2) \cdots 3 \cdot 2 \cdot 1.$$

A este número también se le conoce como las permutaciones de n objetos. Adicionalmente, y por conveniencia, se define $0! = 1$. Observe que las permutaciones de n objetos es un caso particular de la situación mencionada en la sección anterior sobre ordenaciones sin repetición cuando la selección es exhaustiva, es decir, cuando se extraen uno a uno todos los objetos de la urna.

Ejemplo 1.19 Si deseamos conocer el total de órdenes distintos en que podemos colocar una enciclopedia de 5 volúmenes en un librero, la respuesta es claramente $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$. El razonamiento es el siguiente: cualquiera de los cinco libros puede ser colocado al principio, quedan cuatro libros por colocar en la segunda posición, restan entonces tres posibilidades para la tercera posición, etc. Por el principio de multiplicación, la respuesta es el producto de estos números. •

Combinaciones: muestras sin orden y sin reemplazo

Supongamos nuevamente que tenemos un conjunto de n objetos distinguibles y nos interesa obtener una muestra de tamaño k . Supongamos ahora que las muestras deben ser sin orden y sin reemplazo. Es decir, en la muestra no debe haber elementos repetidos, pues no hay reemplazo, y además la muestra debe verse como un conjunto pues no debe haber orden entre sus elementos. ¿Cuántas diferentes muestras podemos obtener de estas características? Para responder a esta pregunta, seguiremos el siguiente razonamiento: cuando el orden importa, hemos encontrado antes la fórmula

$$\frac{n!}{(n-k)!}.$$

Ahora que no nos interesa el orden, observamos que cada uno de los arreglos de la fórmula anterior, está siendo contado $k!$ veces, el número de veces en que los mismos k elementos pueden ser permutados unos con otros, siendo que el conjunto de elementos es el mismo. Para obtener arreglos en donde el orden no importa, debemos entonces dividir por $k!$. La fórmula a la que hemos llegado se llama combinaciones de n en k y la denotaremos como sigue

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

A este número también se le conoce con el nombre de coeficiente binomial de n en k , pues aparece en el famoso teorema del binomio: para cualesquiera números reales a y b , y para cualquier número entero $n \geq 0$,

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k.$$

Para los casos $n = 2$ y $n = 3$, el teorema del binomio se reduce a las siguientes fórmulas, que muy seguramente el lector conoce:

$$\begin{aligned} 1. (a+b)^2 &= a^2 + 2ab + b^2 \\ &= \binom{2}{0} a^2 b^0 + \binom{2}{1} a^1 b^1 + \binom{2}{2} a^0 b^2. \end{aligned}$$

$$\begin{aligned}
 2. (a + b)^3 &= a^3 + 3a^2b + 3ab^2 + b^3 \\
 &= \binom{3}{0}a^3b^0 + \binom{3}{1}a^2b^1 + \binom{3}{2}a^1b^2 + \binom{3}{3}a^0b^3.
 \end{aligned}$$

Ejemplo 1.20 ¿Cuántos equipos distintos de tres personas pueden escogerse de un grupo de 5 personas?

Solución. Observe que el orden de las tres personas escogidas no es importante, de modo que la solución es

$$\binom{5}{3} = \frac{5!}{3!(5-3)!} = 10.$$

•

El coeficiente binomial es también una forma de generar las entradas del así llamado triángulo de Pascal, que puede observarse en la Figura 1.18. El n -ésimo renglón del triángulo de Pascal, iniciando desde cero, contiene los coeficientes del desarrollo de $(a + b)^n$. Existe una forma sencilla de construir este triángulo observando que cada uno de estos números, exceptuando los extremos, es la suma de los dos números inmediatos del renglón anterior. A este respecto véase, por ejemplo, el Ejercicio 84 en la página 67.

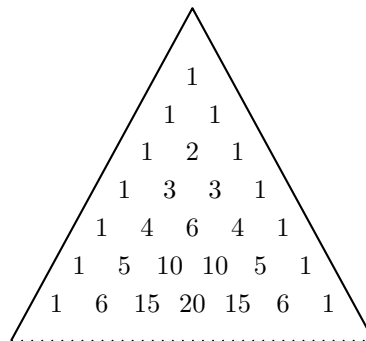


Figura 1.18

Coficiente multinomial

Ahora consideremos que tenemos n objetos no necesariamente distintos unos de otros, por ejemplo, supongamos que tenemos k_1 objetos de un primer tipo, k_2 objetos de un segundo tipo, y así sucesivamente, hasta k_m objetos del tipo m , en donde $k_1 + k_2 + \cdots + k_m = n$. Estos n objetos pueden todos ordenarse uno detrás de otro de tantas formas distintas como indica el así llamado coeficiente multinomial

$$\binom{n}{k_1 k_2 \cdots k_m} = \frac{n!}{k_1! k_2! \cdots k_m!}.$$

Un razonamiento para obtener esta fórmula es el siguiente: si consideramos que los n objetos son todos diferentes, entonces claramente las distintas formas en que pueden colocarse todos estos objetos, uno detrás de otro, es $n!$ Pero para cada uno de estos arreglos, los k_1 objetos del primer tipo, supuestos inicialmente distintos cuando en realidad no lo son, pueden permutarse entre sí de $k_1!$ formas diferentes, siendo que el arreglo total es el mismo. De aquí que debemos dividir por $k_1!$ Lo mismo sucede con los elementos del segundo tipo, y así sucesivamente hasta los elementos del tipo m . En el siguiente ejemplo se muestra una situación sencilla en donde se aplica la fórmula anterior.

Ejemplo 1.21 ¿Cuántas palabras distintas se pueden formar permutando las letras de la palabra “mamá”? (Considere que el acento no es relevante.)

Solución. Existen $\binom{4}{2,2} = 6$ palabras distintas y éstas son:

mama	amma	mmaa
maam	amam	aamm.

•

El teorema del binomio se puede extender a la siguiente fórmula, en donde aparece el coeficiente multinomial.

$$(a_1 + a_2 + \cdots + a_m)^n = \sum \binom{n}{k_1 \cdots k_m} a_1^{k_1} a_2^{k_2} \cdots a_m^{k_m}, \quad (1.1)$$

en donde la suma se efectúa sobre todos los posibles valores enteros no negativos de k_1, k_2, \dots, k_m , tales que $k_1 + k_2 + \cdots + k_m = n$. A este resultado

se le conoce como el teorema del multinomio. Por ejemplo, compruebe el lector que la fórmula (1.1) produce la siguiente expresión

$$\begin{aligned}(a + b + c)^2 &= a^2 + b^2 + c^2 + 2ab + 2ac + 2bc \\ &= \binom{2}{200} a^2 b^0 c^0 + \binom{2}{020} a^0 b^2 c^0 + \binom{2}{002} a^0 b^0 c^2 \\ &\quad + \binom{2}{110} a^1 b^1 c^0 + \binom{2}{101} a^1 b^0 c^1 + \binom{2}{011} a^0 b^1 c^1.\end{aligned}$$

¿Puede usted desarrollar $(a + b + c)^3$? Es interesante observar que cuando hay únicamente dos tipos de objetos, el coeficiente multinomial se reduce al coeficiente binomial y la notación también se reduce a la antes usada, es decir,

$$\binom{n}{k \ (n-k)} = \binom{n}{k}.$$

Muestras sin orden y con reemplazo

Finalmente, consideremos el caso de hacer k extracciones de una urna de n objetos con las condiciones de que cada objeto extraído es regresado a la urna (y entonces puede ser elegido nuevamente), y en donde el orden de la muestra no es relevante. Para encontrar una fórmula para el total de muestras que pueden obtenerse con estas características, usaremos una modelación distinta pero equivalente.

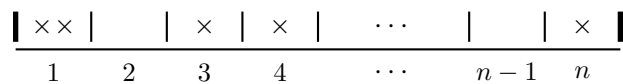


Figura 1.19

Consideremos el arreglo de n casillas de la Figura 1.19 junto con la siguiente interpretación: la primera casilla tiene dos cruces, y eso indica que la bola uno fue seleccionada dos veces, la segunda casilla está vacía, y ello significa que la bola dos no fue seleccionada, etc. El número de cruces en la casilla i indica entonces el número de veces que la bola i fue seleccionada. En total

debe haber k cruces pues es el total de extracciones. Deseamos entonces conocer el número de posibles arreglos que pueden obtenerse con estas características, y debe ser claro, después de algunos momentos de reflexión, que éste es el número de muestras de tamaño k , con reemplazo y sin orden, que se pueden obtener de un conjunto de n elementos distinguibles. Consideremos que las dos paredes en los extremos de este arreglo son fijas, estas paredes se encuentran ligeramente remarcadas como puede apreciarse en la Figura 1.19. Consideremos además que las posiciones intermedias, cruz o línea vertical, pueden ubicarse de cualquier forma u orden dentro de las paredes anteriores. En total hay $n + k - 1$ objetos movibles, y cambiar de posición estos objetos produce las distintas configuraciones posibles que nos interesan. El número total de estos arreglos es entonces

$$\binom{n + k - 1}{k}$$

que equivale a colocar dentro de las $n + k - 1$ posiciones las k cruces, dejando en los lugares restantes las paredes movibles.

Resumen de fórmulas

En el contexto de muestras de tamaño k , tomadas de un conjunto de cardinalidad n , y a manera de resumen parcial, tenemos la tabla de fórmulas en la Figura 1.20.

Muestras	con reemplazo	sin reemplazo
con orden	n^k	$\frac{n!}{(n - k)!}$
sin orden	$\binom{n + k - 1}{k}$	$\binom{n}{k}$

Figura 1.20

Debemos hacer énfasis, sin embargo, en que los problemas de conteo pueden ser difíciles de resolver y que para resolver un problema en particular no

debemos clasificarlo forzosamente y de manera mecánica en alguno de los esquemas desarrollados. Muy posiblemente el problema en cuestión requerirá de un razonamiento especial que involucre alguna combinación de las fórmulas encontradas. En algunos casos, uno puede encontrar dos o más “soluciones” distintas y aparentemente correctas de un problema. A veces las múltiples soluciones se deben a que el problema no está bien especificado y por lo tanto pueden surgir ambigüedades en su interpretación. Véase el libro de Székely [21] para conocer una amplia gama de paradojas que surgen en la probabilidad y la estadística. Mencionaremos también que la programación de computadoras puede ser una herramienta útil para resolver problemas de conteo o simplemente para verificar alguna respuesta encontrada.

Otro aspecto que amerita señalarse es que usaremos las técnicas de conteo principalmente para aplicarlas en problemas donde se use la definición de probabilidad clásica $P(A) = \#A/\#\Omega$. Y aunque estas técnicas de conteo son bastante útiles y conformaron históricamente los métodos para resolver problemas de juegos de azar, constituyen ahora sólo una parte mínima y particular de la actual teoría matemática de la probabilidad.

Ejercicios

69. Un dado equilibrado se lanza 6 veces consecutivas. ¿Cuál es la probabilidad de que
 - a) aparezcan las seis caras del dado en orden creciente o decreciente?
 - b) aparezcan las seis caras del dado en cualquier orden?
 - c) sólo aparezcan números pares?
 - d) aparezcan números pares e impares alternados?
70. ¿Cuántos enteros positivos de a lo sumo cinco dígitos son divisibles por 2? ¿Y de ellos, cuántos hay que empiecen con el dígito 1?
71. **El problema de los cumpleaños.** Calcule la probabilidad de que en un grupo de n personas al menos dos de ellas tengan la misma fecha de cumpleaños.

72. Sea $n \geq 1$ un entero. Considere el intervalo $[0, L]$, dividido en n partes de idéntica longitud. Se escogen n puntos al azar, uno por uno, en el intervalo $[0, L]$, de manera independiente uno de otro. Calcule la probabilidad de que exactamente un punto caiga en cada subintervalo.
73. Se lanza un dado equilibrado tres veces. Calcule la probabilidad de obtener tres números en orden ascendente, no necesariamente consecutivos.
74. Sean A , B y C tres eventos de un experimento aleatorio tales que cualquiera de sus intersecciones es no vacía. Determine el número máximo de formas distintas en las que el evento $A \cup B \cup C$ puede expresarse como la unión de tres eventos disjuntos. Suponga que la descomposición admite el conjunto vacío como alguno de sus componentes y que es relevante el orden de los componentes.
75. Suponga que se desea dibujar en un diagrama de Venn el evento $A_1 \cup A_2 \cup \dots \cup A_n$. Determine el número máximo de regiones simples disjuntas de las que consta este evento. Corrobore su respuesta en los casos $n = 2$ y $n = 3$.
76. **Corredores.** ¿De cuántas maneras diferentes pueden clasificarse los tres primeros lugares de una carrera de n corredores?
77. **Mesa circular.** ¿De cuántas maneras diferentes pueden sentarse n personas en una mesa circular?
78. Suponga que los conductores de 8 automóviles estacionan sus coches completamente al azar en un estacionamiento de 12 lugares y que la configuración del estacionamiento es lineal. Determine la probabilidad de que los lugares no ocupados sean adyacentes.
79. **Rumores.** En un pueblo de $n + 1$ habitantes, uno de ellos le rumorea algo a una segunda persona, ésta, a su vez, se lo cuenta a una tercera persona (que puede ser la primera persona), y así sucesivamente. Determine la probabilidad de que el rumor se transmita r veces sin que regrese a la primera persona.
80. Calcule la probabilidad de que la suma de los resultados de lanzar dos dados equilibrados sea 8, suponiendo que:

- a) los dados son distinguibles.
- b) los dados son indistinguibles.

81. **Funciones.** Sean A y B dos conjuntos finitos con cardinalidades n y m , respectivamente, como se muestra en la Figura 1.21. Determine el número total de funciones f de A en B tal que

- a) no tienen restricción alguna.
- b) son inyectivas (uno a uno), suponiendo $n \leq m$.
- c) son suprayectivas (sobre), suponiendo $m \leq n$.

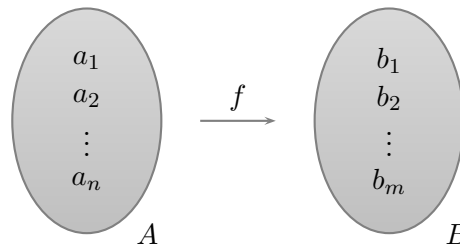


Figura 1.21

82. Un panadero elabora 100 panes en un día, en donde 10 de ellos pesan menos de lo que deberían. Un inspector pesa 5 panes tomados al azar. Calcule la probabilidad de que el inspector encuentre en su muestra exactamente un pan de peso incorrecto.
83. Sean k, n, m números naturales tales que $k \leq n + m$. Demuestre que

$$\sum_{i,j} \binom{n}{i} \binom{m}{j} = \binom{n+m}{k},$$

en donde la suma se efectúa sobre valores enteros de i y j tales que $0 \leq i \leq n, 0 \leq j \leq m$ e $i + j = k$.

84. Sean k y n números naturales tales que $k < n$. Demuestre que

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

A partir de esta fórmula se construye el triángulo de Pascal. Más generalmente, demuestre que

$$\binom{n}{k_1 k_2 \cdots k_m} = \binom{n-1}{k_1-1 k_2 \cdots k_m} + \binom{n-1}{k_1 k_2-1 \cdots k_m} + \cdots + \binom{n-1}{k_1 k_2 \cdots k_m-1},$$

en donde

$$\binom{n}{k_1 k_2 \cdots k_m} = \frac{n!}{k_1! k_2! \cdots k_m!}.$$

85. Sean n y m dos números enteros positivos y sean x_1, \dots, x_m números reales cualesquiera. Demuestre que

$$(x_1 + \cdots + x_m)^n = \sum \binom{n}{k_1 k_2 \cdots k_m} x_1^{k_1} \cdots x_m^{k_m}, \quad (1.2)$$

en donde la suma se realiza sobre todos los posibles números enteros k_1, \dots, k_m tales que $0 \leq k_i \leq n$, para $i = 1, \dots, m$, y

$$k_1 + \cdots + k_m = n.$$

86. ¿Cuántos sumandos aparecen en la fórmula (1.2)? Es decir, ¿cuántos distintos vectores existen de la forma (k_1, \dots, k_m) , en donde cada entrada es un número entero mayor o igual a cero y la suma de todos ellos es n ?
87. Sea $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una función de n variables que es infinitamente diferenciable, es decir, tiene derivadas de todos los órdenes. ¿Cuántas derivadas parciales hay de orden m ?
88. Sea $n \geq 2$ un número natural fijo. Determine la cantidad de parejas (x, y) que existen de números naturales x y y tales que

$$1 \leq x \leq y \leq n.$$

89. Sean k y n dos números naturales tales que $1 \leq k \leq n$. Determine la cantidad de parejas de números naturales (x, y) que existen de tal forma que

$$1 \leq x, y \leq n \quad \text{y} \quad |x - y| \geq k.$$

90. Sean k y n dos números naturales tales que $1 \leq k \leq n$. Determine la cantidad de vectores que existen de la forma (x_1, \dots, x_k) de tal manera que cada entrada de este vector es un número entero x_i que satisface $1 \leq x_i \leq n$ y además cumplen la condición
- $1 \leq x_1 < \dots < x_k \leq n$.
 - $1 \leq x_1 \leq \dots \leq x_k \leq n$.
91. Sean $k, n \geq 1$ dos números enteros fijos. ¿Cuántas soluciones enteras no negativas tiene la ecuación
- $x_1 + x_2 + \dots + x_k = n$?
 - $x_1 + x_2 + \dots + x_k \leq n$?
 - $x_1 + x_2 + \dots + x_k \geq n$?
92. **Las cajas de cerillos de Banach**⁴. Una persona tiene dos cajas de cerillos, cada una de las cuales contiene n cerillos. La persona coloca una caja en cada uno de sus bolsillos izquierdo y derecho, y cada vez que requiere un cerillo elije una de sus bolsillos al azar y toma de la caja correspondiente uno de los cerillos. ¿Cuál es la probabilidad de que en el momento en el que la persona se da cuenta de que la caja del bolsillo izquierdo está vacía, en la caja del bolsillo derecho haya exactamente r cerillos?
93. Se lanzan tres dados equilibrados. Encuentre la probabilidad de que exactamente en dos de ellos aparezca la cara “1”, suponiendo que:
- los dados son distinguibles.
 - los dados son indistinguibles.
94. Se lanzan dos dados idénticos y equilibrados a un mismo tiempo. Calcule la probabilidad de que la suma de las dos caras sea igual a $x = 2, 3, \dots, 12$. Compruebe que la suma de todas estas probabilidades es 1.
95. **Zapatos**. Una mujer tiene n pares de zapatos en desorden y en un viaje intempestivo escoge al azar $2r$ zapatos ($2r \leq 2n$). Calcule la probabilidad de que en el conjunto escogido:

⁴Stephen Banach (1892-1945), matemático polaco.

- a) no haya ningún par completo.
 - b) haya exactamente un par completo.
 - c) haya r pares completos.
96. **Llaves.** Una persona tiene n llaves, de las cuales únicamente una ajusta a la cerradura pero no sabe cuál de ellas es la correcta. Procede a tomar las llaves al azar, una por una, hasta encontrar la correcta. Calcule la probabilidad de encontrar la llave correcta en el n -ésimo intento suponiendo que
- a) retira las llaves que no funcionaron.
 - b) no retira las llaves que no funcionaron.
97. **Diagonales en polígonos.** ¿Cuántas diagonales se pueden trazar en un polígono convexo de n lados ($n \geq 3$)? O bien, al terminar una reunión de n personas, éstas se despiden con un saludo, ¿cuántos saludos habrá?
98. ¿Cuál es el número máximo de regiones en las que
- a) n líneas rectas pueden dividir un plano?
 - b) n círculos pueden dividir un plano?
 - c) n planos pueden dividir el espacio?
 - d) n esferas pueden dividir el espacio?
99. Sean p y q dos números primos distintos y sean n y m dos números naturales. ¿Cuántos divisores diferentes tiene el número $p^n q^m$?
100. ¿De cuántas formas diferentes se pueden ordenar los elementos del conjunto $\{1, 2, \dots, 2n + 1\}$ de tal forma que
- a) cada número impar ocupe una posición impar?
 - b) cada número impar ocupe una posición impar y de forma ascendente?
 - c) los números pares queden de forma ascendente?
 - d) los números impares queden de forma descendente?

101. La fórmula de inclusión-exclusión para la probabilidad de la unión arbitraria de n eventos aparece en el Ejercicio 55, en la página 45. En general, ¿cuántos sumandos aparecen en el lado derecho de esta fórmula?

1.13. Probabilidad condicional

La probabilidad condicional es un concepto elemental pero muy importante, que se utiliza con mucha frecuencia en el cálculo de probabilidades. En los resultados que veremos en esta sección mostraremos las situaciones en las que se aplica la probabilidad condicional para reducir ciertas probabilidades a expresiones más sencillas.

Definición 1.10 Sean A y B dos eventos y supongamos que B tiene probabilidad estrictamente positiva. La probabilidad condicional del evento A , dado el evento B , se denota por el símbolo $P(A|B)$ y se define como el cociente

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \quad (1.3)$$

El término $P(A|B)$ se lee “probabilidad de A dado B ” y es claro, a partir de la definición, que es necesaria la condición $P(B) > 0$ para que el cociente esté bien definido. Cuando $P(B) = 0$ no existe una definición establecida para $P(A|B)$. En ocasiones se usa la expresión $P_B(A)$ para denotar a esta probabilidad. En la expresión (1.3), el evento B representa un evento que ha ocurrido, y la probabilidad condicional $P(A|B)$ es la probabilidad de A modificada con la información adicional de que B ha ocurrido.

Así, uno puede imaginar que el espacio muestral del experimento aleatorio se ha reducido al evento B de tal forma que todo lo que se encuentre fuera de este evento tiene probabilidad condicional cero. La afirmación anterior es evidente a partir de observar que si A y B son ajenos, entonces el numerador de la probabilidad condicional (1.3) es cero.

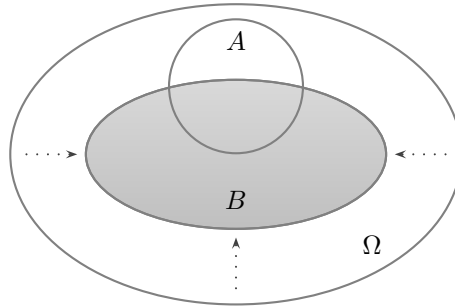


Figura 1.22

Ejemplo 1.22 Considere el experimento de lanzar un dado equilibrado y defina los eventos

$$\begin{aligned} A &= \{2\} = \text{“Se obtiene el número 2”}, \\ B &= \{2, 4, 6\} = \text{“Se obtiene un número par”}. \end{aligned}$$

Es claro que $P(A) = 1/6$, sin embargo, sabiendo que B ha ocurrido, es decir, sabiendo que el resultado es un número par, la probabilidad del evento A es ahora

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\{2\})}{P(\{2, 4, 6\})} = \frac{1/6}{3/6} = \frac{1}{3}.$$

Es decir, la información adicional de la ocurrencia del evento B ha hecho que la probabilidad de A se incremente de $1/6$ a $1/3$. •

Es interesante comprobar que la probabilidad condicional $P(A|B)$, vista como una función del evento A , cumple los tres axiomas de Kolmogorov, es decir, satisface:

- a) $P(\Omega|B) = 1$.
- b) $P(A|B) \geq 0$.
- c) $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B)$, cuando $A_1 \cap A_2 = \emptyset$.

En consecuencia, la función $A \mapsto P(A|B)$ es una medida de probabilidad y por lo tanto cumple todos los resultados conocidos para cualquier medida de probabilidad, por ejemplo:

1. $P(A | B) = 1 - P(A^c | B)$.
2. $P(A_1 \cup A_2 | B) = P(A_1 | B) + P(A_2 | B) - P(A_1 \cap A_2 | B)$.

En la sección de ejercicios se encuentran algunas propiedades generales de la probabilidad condicional y en las siguientes secciones se verán dos ejemplos importantes de aplicación de esta nueva probabilidad: el teorema de probabilidad total y el teorema de Bayes.

Ejercicios

102. A partir de la definición de probabilidad condicional, demuestre directamente las siguientes afirmaciones:

- a) $P(A | B) = 1 - P(A^c | B)$.
- b) Si $A_1 \subseteq A_2$ entonces $P(A_1 | B) \leq P(A_2 | B)$.
- c) $P(A_1 | B) = P(A_1 \cap A_2 | B) + P(A_1 \cap A_2^c | B)$.
- d) $P(A_1 \cup A_2 | B) = P(A_1 | B) + P(A_2 | B) - P(A_1 \cap A_2 | B)$.
- e) $P(\bigcup_{k=1}^{\infty} A_k | B) \leq \sum_{k=1}^{\infty} P(A_k | B)$.
- f) Si A_1, A_2, \dots son eventos ajenos dos a dos, entonces

$$P\left(\bigcup_{k=1}^{\infty} A_k | B\right) = \sum_{k=1}^{\infty} P(A_k | B).$$

103. Sean A y B dos eventos tales que $P(A) = 1/4$, $P(B | A) = 1/2$ y $P(A | B) = 1/2$. Determine y justifique si las siguientes afirmaciones son verdaderas o falsas.

- | | |
|--------------------------|------------------------------------|
| a) A y B son ajenos. | d) $P(A^c B^c) = 5/6$. |
| b) $A = B$. | e) $P(B^c A^c) = 5/6$. |
| c) $P(B) = 1/4$. | f) $P(A B) + P(A B^c) = 2/3$. |

104. Demuestre o proporcione un contraejemplo para las siguientes afirmaciones generales:

- a) $P(A|B) = P(B|A)$.
- b) $P(A|B) + P(A|B^c) = 1$.
- c) $P(A|B) \geq P(A)$.
- d) Si $P(A|B) \geq P(A)$ entonces $P(B|A) \leq P(B)$.
- e) Si $P(A) > P(B)$ entonces $P(A|C) > P(B|C)$.
- f) Si $P(A) > 0$ y $P(B) > 0$ entonces $P(A|B) > 0$.
- g) $P(A) = P(B) \Leftrightarrow P(A|C) = P(B|C)$.
- h) $A \subseteq B^c \Leftrightarrow P(A|B) = 0$.
- i) $P(A) \leq P(B) \Leftrightarrow P(A|C) \leq P(B|C)$.
- j) $A \subseteq B \Leftrightarrow P(A|C) \leq P(B|C)$.
- k) $P(A|B) = P(A|B^c) \Leftrightarrow P(B|A) = P(B|A^c)$.
- l) Si $B \cap C = \emptyset$ entonces $P(A|B \cup C) = P(A|B) + P(A|C)$.

105. Para cada inciso proporcione un ejemplo en el que se cumpla la afirmación indicada. Estos ejemplos no demuestran la validez general de estas afirmaciones.

- a) $P(A|B) = 0$ pero $P(A) > 0$.
- b) $P(A|B^c) = P(A^c|B)$.
- c) $P(A) < P(A|B)$.
- d) $P(A) = P(A|B)$.
- e) $P(A) > P(A|B)$.

106. Un grupo de personas está compuesto de 60 % hombres y 40 % de mujeres. De los hombres, el 30 % fuma y de las mujeres, el 20 % fuma. Si una persona de este grupo se escoge al azar, encuentre la probabilidad de que

- a) sea hombre y fume.
- b) sea hombre y no fume.
- c) sea mujer y fume.
- d) sea mujer y no fume.
- e) sea hombre dado que se sabe que fuma.

f) sea mujer dado que se sabe que no fuma.

107. Un dado equilibrado se lanza dos veces consecutivas. Dado que en el primer lanzamiento se obtuvo un 3, ¿cuál es la probabilidad de que la suma de los dos resultados sea mayor a 6?

108. Sean A y B eventos independientes, ambos con probabilidad estrictamente positiva. Demuestre que para cualquier evento C ,

$$P(C | A) = P(B)P(C | A \cap B) + P(B^c)P(C | A \cap B^c).$$

109. Sean B_1, \dots, B_n eventos mutuamente ajenos, cada uno con probabilidad estrictamente positiva. Defina el evento B como $\bigcup_{i=1}^n B_i$ y sea A un evento tal que $P(A | B_i) = p$ para $i = 1, \dots, n$. Demuestre que

$$P(A | B) = p.$$

110. **Regla del producto.** Sean A_1, \dots, A_n eventos tales que

$$P(A_1 \cap \dots \cap A_{n-1}) > 0.$$

Demuestre que

$$\begin{aligned} P(A_1 \cap \dots \cap A_n) &= P(A_1) P(A_2 | A_1) P(A_3 | A_1 \cap A_2) \\ &\quad \dots P(A_n | A_1 \cap \dots \cap A_{n-1}). \end{aligned}$$

111. **La urna de Polya.** En una urna se tienen r bolas rojas y b bolas blancas. Un ensayo consiste en tomar una bola al azar y regresarla a la urna junto con k bolas del mismo color. Se repite este ensayo varias veces y se define el evento R_n como aquel en el que se obtiene una bola roja en la n -ésima extracción. Demuestre por inducción sobre n que

$$P(R_n) = \frac{r}{r+b}, \quad n = 1, 2, \dots$$

112. **El problema de la ruina del jugador.** Dos jugadores, A y B , lanzan sucesivamente una moneda. En cada lanzamiento, si la moneda cae cara, el jugador B le entrega una unidad monetaria al jugador A , en caso contrario, si la moneda cae cruz, el jugador A le paga

una unidad monetaria al jugador B . El juego continúa hasta que uno de ellos se arruina. Suponga que los lanzamientos de la moneda son independientes, que en cada uno de ellos la probabilidad de obtener cara es p y que el jugador A inicia con n unidades monetarias y B inicia con $N - n$ unidades monetarias. Defina el evento E_n como aquel en el que el jugador A gana eventualmente todo el dinero cuando comienza con n unidades monetarias y sea $q = 1 - p$. Demuestre que la probabilidad $P(E_n)$, denotada por p_n , satisface la siguiente ecuación en diferencias con las condiciones de frontera especificadas.

$$\begin{aligned} (p_{n+1} - p_n) &= \frac{q}{p} (p_n - p_{n-1}), \quad n = 1, 2, \dots, N - 1, \\ p_0 &= 0, \\ p_N &= 1. \end{aligned}$$

Verifique además que la solución a este sistema de ecuaciones es

$$p_n = \begin{cases} n/N & \text{si } p = 1/2, \\ \frac{1 - (q/p)^n}{1 - (q/p)^N} & \text{si } p \neq 1/2. \end{cases}$$

1.14. Teorema de probabilidad total

Sea Ω el espacio muestral de un experimento aleatorio. Decimos que la colección de eventos $\{B_1, B_2, \dots, B_n\}$ es una partición finita de Ω si se cumplen las siguientes condiciones:

- a) $B_i \neq \emptyset$, $i = 1, 2, \dots, n$.
- b) $B_i \cap B_j = \emptyset$ para $i \neq j$.
- c) $\bigcup_{i=1}^n B_i = \Omega$.

Así, se requiere que cada uno de los elementos de una partición sea distinto del conjunto vacío, que sean ajenos dos a dos y que la unión de todos ellos

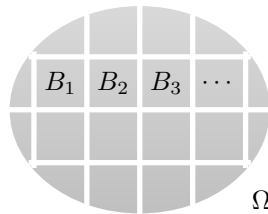


Figura 1.23

constituya la totalidad del espacio muestral. De manera gráfica podemos representar una partición finita como se muestra en la Figura 1.23.

El siguiente resultado es bastante útil y tiene una amplia aplicación en la probabilidad.

Teorema 1.1 (Teorema de probabilidad total)

Sea B_1, \dots, B_n una partición de Ω tal que $P(B_i) \neq 0$, $i = 1, \dots, n$. Para cualquier evento A ,

$$P(A) = \sum_{i=1}^n P(A | B_i)P(B_i).$$

Demostración. Cualquier evento A admite la descomposición disjunta

$$A = A \cap \Omega = A \cap \left(\bigcup_{i=1}^n B_i \right) = \bigcup_{i=1}^n (A \cap B_i).$$

De donde se obtiene

$$P(A) = \sum_{i=1}^n P(A \cap B_i) = \sum_{i=1}^n P(A | B_i)P(B_i).$$

•

Cuando la partición del espacio muestral consta de únicamente los elementos

B y B^c , la fórmula del teorema de probabilidad total se reduce a la expresión

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c).$$

En el Ejercicio 157 extenderemos ligeramente el teorema de probabilidad total al caso cuando la partición del espacio muestral consta de un número infinito numerable de elementos. La expresión es análoga,

$$P(A) = \sum_{i=1}^{\infty} P(A|B_i)P(B_i).$$

A continuación se verán algunos ejemplos de aplicación de la fórmula de probabilidad total. En la sección de ejercicios se ilustran algunas situaciones en donde puede aplicarse esta fórmula. Más adelante encontraremos problemas para los cuales no es evidente la forma de encontrar la probabilidad de un cierto evento, pero condicionando adecuadamente, como aparece en el enunciado del teorema de probabilidad total, en ocasiones se puede encontrar de manera más fácil la probabilidad buscada.

Ejemplo 1.23 Suponga que tenemos dos cajas: una con 3 bolas blancas y 7 bolas de color gris, la otra con 6 blancas y 6 grises. Esta situación se ilustra en la Figura 1.24. Si se elige una caja al azar y después se saca una bola al azar, ¿cuál es la probabilidad de que sea blanca?

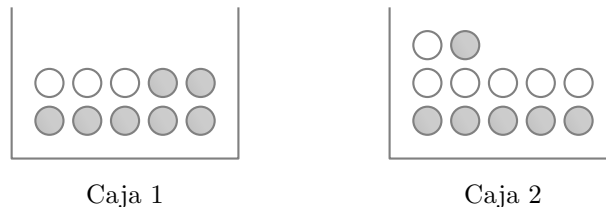


Figura 1.24

Solución. El experimento aleatorio consiste en escoger una caja al azar, con idéntica probabilidad cada una de ellas, y después escoger una bola de la caja escogida. Es claro que el espacio muestral puede escribirse como sigue:

$$\Omega = \{(C_1, B), (C_1, G), (C_2, B), (C_2, G)\},$$

en donde C_1 y C_2 denotan los eventos en donde las cajas uno y dos fueron escogidas, respectivamente, y B y G denotan los eventos en donde una bola blanca o gris fueron escogidas, respectivamente. Nos piden calcular la probabilidad de B . Observe que es fácil calcular la probabilidad de este evento cuando se conoce la caja que fue escogida. Esto sugiere condicionar sobre el resultado de escoger alguna de las dos cajas y aplicar el teorema de probabilidad total, es decir,

$$\begin{aligned} P(B) &= P(B | C_1)P(C_1) + P(B | C_2)P(C_2) \\ &= (3/10)(1/2) + (6/12)(1/2) \\ &= 2/5. \end{aligned}$$

Observe además que la partición del espacio muestral consta de dos elementos: $\{(C_1, B), (C_1, G)\}$ y $\{(C_2, B), (C_2, G)\}$. ¿Puede usted comprobar que $P(G) = 3/5$? Uno puede también preguntarse por situaciones aparentemente extrañas como la siguiente: si se obtuvo una bola blanca, ¿cuál es la probabilidad de que haya sido obtenida de la primera caja? Es posible calcular esta probabilidad a través del teorema de Bayes, el cual se tratará en la siguiente sección. •

Ejemplo 1.24 Suponga que en una población humana de igual número de hombres y mujeres, el 4% de hombres son daltónicos y el 1% de las mujeres son daltónicas. Una persona es elegida al azar, ¿cuál es la probabilidad de que sea daltónica?

Solución. Definamos primero los eventos de interés. Sea M el evento “La persona escogida es mujer”, H el evento “La persona escogida es hombre” y D el evento “La persona escogida es daltónica”. Deseamos calcular $P(D)$. Por el teorema de probabilidad total,

$$\begin{aligned} P(D) &= P(D | M)P(M) + P(D | H)P(H) \\ &= (1/100)(1/2) + (4/100)(1/2) \\ &= 1/40. \end{aligned}$$

•

Ejercicios

113. En un grupo hay m mujeres y n hombres. Suponga que $m, n \geq 2$. Se escogen a dos personas al azar, de manera secuencial y sin reemplazo. Encuentre la probabilidad de que:
- la segunda persona sea mujer.
 - la primera persona sea mujer dado que la segunda fue mujer.
 - la segunda persona sea hombre.
 - la primera persona sea hombre dado que la segunda fue mujer.
 - ambas personas sean del mismo sexo.
 - ambas personas sean de sexo distinto.
114. En un grupo hay m mujeres y n hombres. Se seleccionan al azar a k personas, una por una y sin reemplazo. Suponga que $k \leq m, n$. Encuentre la probabilidad de que la última persona escogida sea mujer.
115. La urna A contiene 2 canicas blancas y 4 rojas. La urna B contiene 1 canica blanca y 1 roja. Se toma una canica al azar, sin verla, de la urna A y se coloca en la urna B . Después se toma una canica al azar de la urna B . Calcule la probabilidad de que la canica seleccionada de la urna B sea roja.
116. Se tiene un arreglo lineal de tres cajas como se muestra en la Figura 1.25, en donde en cada caja hay 1 canica blanca y 1 azul. Se toma una canica al azar de la primera caja y, sin verla, se coloca en la segunda caja. Después se toma una canica al azar de la segunda caja y, sin verla, se coloca en la tercera caja. Finalmente se toma una canica al azar de la tercera caja. Calcule la probabilidad de que la canica escogida sea azul.
117. Se cuenta con cuatro monedas marcadas con “cara” y “cruz” tal que para la i -ésima moneda $P(\text{“cara”}) = 0.2i$, $i = 1, 2, 3, 4$. Si se escoge una moneda al azar y se lanza al aire, encuentre la probabilidad de que ésta caiga “cruz”.
118. Una persona lanza un dado equilibrado una vez, obteniendo el resultado n . Después lanza nuevamente el dado tantas veces como indicó el

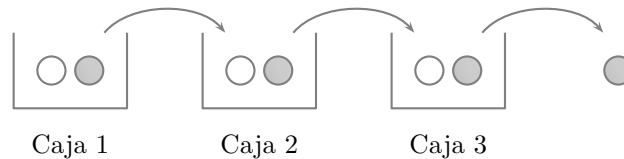


Figura 1.25

resultado del primer lanzamiento sumando los resultados de estos últimos lanzamientos y obteniendo un total de s . Calcule la probabilidad de que los números n y s coincidan.

119. **Canal binario ruidoso.** Los símbolos 0 y 1 se envían a través de un canal binario ruidoso como el que se muestra en la Figura 1.26. Debido al ruido, un 0 se distorsiona en un 1 con probabilidad 0.2 y un 1 se distorsiona en un 0 con probabilidad 0.1. Suponga que el símbolo de entrada 0 aparece el 45 % de las veces y el símbolo 1 el 55 %. Encuentre la probabilidad de que en un uso cualquiera del canal

- se reciba un 0.
- se reciba un 1.
- no haya error en la transmisión.
- se presente algún error en la transmisión.

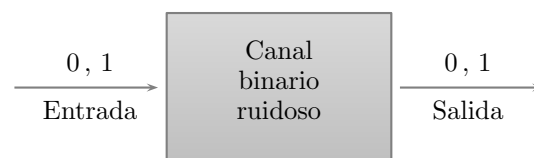


Figura 1.26

120. Una urna contiene 4 bolas blancas y 6 azules. Se lanza un dado equilibrado y se toma una muestra de la urna de tantas bolas como indicó el

dato. Suponga que la muestra es sin orden y sin reemplazo. Encuentre la probabilidad de que todas las bolas escogidas sean blancas.

121. **Examen de opción múltiple.** Un estudiante contesta un examen de opción múltiple, en el cual cada pregunta tiene cuatro opciones como respuesta pero sólo una es correcta. Cuando el estudiante conoce la respuesta correcta, la selecciona, en caso contrario, selecciona una de las opciones al azar. Suponga que el estudiante conoce la respuesta correcta al 60 % de las preguntas.
- a) Calcule la probabilidad de que el estudiante tenga correcta una de las preguntas escogida al azar.
 - b) Si el estudiante obtuvo la respuesta correcta a una pregunta escogida al azar, ¿cuál es la probabilidad de que haya sabido verdaderamente la respuesta?
 - c) Si el examen consta de 10 preguntas y es necesario tener por lo menos 6 respuestas correctas para acreditar, ¿cuál es la probabilidad de que el estudiante pase el examen?
122. Una urna contiene 3 bolas blancas y 4 negras. Se extraen dos bolas al azar, una después de la otra y sin reemplazo. Calcule la probabilidad de que
- a) la segunda bola sea negra dado que la primera fue negra.
 - b) la segunda bola sea del mismo color que la primera.
 - c) la segunda bola sea blanca.
 - d) la primera bola sea blanca dado que la segunda fue blanca.
123. Se escogen al azar dos letras del nombre CAROLINA y se retiran de su posición. Después se vuelven a colocar al azar en los dos espacios vacíos. Calcule la probabilidad de que el nombre no sea modificado.
124. **El problema de los tres prisioneros.** A tres prisioneros, a quienes llamaremos A , B y C , les informa su celador que se ha escogido al azar a uno de ellos para ser ejecutado, dejando a los otros dos en libertad. El prisionero A sabe que tiene probabilidad $1/3$ de ser ejecutado y le pide al celador que le diga en secreto cuál de sus dos compañeros

saldrá libre argumentando que por lo menos uno de ellos saldrá en libertad y que saber esta información no cambia su probabilidad de ser ejecutado. El celador, por el contrario, piensa que si el prisionero *A* sabe cuál de sus dos compañeros saldrá en libertad, la probabilidad de ser ejecutado aumenta a $1/2$. ¿Quién tiene la razón? Justifique su respuesta.

125. **El problema de las tres puertas (Monty Hall⁵)**. Se le presentan a un concursante tres puertas cerradas, detrás de una de las cuales hay un premio. El concursante debe adivinar la puerta que contiene el premio para ganarlo. Una vez que el concursante elige una puerta, y antes de abrirla, el presentador del concurso abre alguna de las puertas restantes, de la cual sabe que no contiene ningún premio. Entonces le pregunta al concursante si desea cambiar su decisión. ¿Qué debe hacer el concursante? Justifique su respuesta.

1.15. Teorema de Bayes

El resultado interesante que estudiaremos a continuación involucra nuevamente probabilidades condicionales. Fue publicado por primera vez en 1763, dos años después de la muerte de su creador: el matemático y teólogo inglés Thomas Bayes.

⁵Monte Halparin (1921-), mejor conocido como Monty Hall, narrador deportivo, productor, actor, cantante y presentador del concurso de televisión *Let's Make a Deal*.



Thomas Bayes
(Inglaterra 1702–1761)

Teorema 1.2 (Teorema de Bayes) Sea B_1, \dots, B_n una partición de Ω tal que $P(B_i) \neq 0$, $i = 1, \dots, n$. Sea A un evento tal que $P(A) \neq 0$. Entonces para cada $j = 1, 2, \dots, n$,

$$P(B_j | A) = \frac{P(A | B_j)P(B_j)}{\sum_{i=1}^n P(A | B_i)P(B_i)}.$$

Demostración. Por definición de probabilidad condicional, y después usando el teorema de probabilidad total, tenemos que para cada $j = 1, \dots, n$,

$$P(B_j | A) = \frac{P(A \cap B_j)}{P(A)} = \frac{P(A | B_j)P(B_j)}{\sum_{i=1}^n P(A | B_i)P(B_i)}.$$

•

Cuando la partición de Ω consta de los elementos B y B^c , el teorema de Bayes para el evento B adquiere la forma

$$P(B | A) = \frac{P(A | B)P(B)}{P(A | B)P(B) + P(A | B^c)P(B^c)}.$$

Y cuando la partición consta de un número infinito numerable de elementos, el teorema de Bayes tiene la siguiente extensión ligera, la cual se pide demostrar en el Ejercicio 158, para $j = 1, 2, \dots$

$$P(B_j | A) = \frac{P(A | B_j)P(B_j)}{\sum_{i=1}^{\infty} P(A | B_i)P(B_i)}.$$

Veremos ahora algunos ejemplos de aplicación del teorema de Bayes.

Ejemplo 1.25 En una fábrica hay dos máquinas. La máquina 1 realiza el 60 % de la producción total y la máquina 2 el 40 %. De su producción total, la máquina 1 produce 3 % de material defectuoso, la 2 el 5 %. El asunto es que se ha encontrado un material defectuoso, ¿cuál es la probabilidad de que este material defectuoso provenga de la máquina 2?

Solución. Conviene definir los siguientes eventos:

$$\begin{aligned} D &= \text{“El material escogido es defectuoso”}, \\ M_1 &= \text{“La máquina 1 produjo el material escogido”}, \\ M_2 &= \text{“La máquina 2 produjo el material escogido”}. \end{aligned}$$

La pregunta planteada se traduce en encontrar $P(M_2 | D)$ y observamos que la información que tenemos es $P(D | M_2)$. Por el teorema de Bayes,

$$\begin{aligned} P(M_2 | D) &= \frac{P(D | M_2)P(M_2)}{P(D | M_1)P(M_1) + P(D | M_2)P(M_2)} \\ &= \frac{(5/100)(40/100)}{(3/100)(60/100) + (5/100)(40/100)} \\ &= 10/19. \end{aligned}$$

Como un ejercicio al lector, se deja comprobar que $P(M_1 | D) = 9/19$. •

Ejemplo 1.26 En un laboratorio se descubrió una prueba para detectar cierta enfermedad, y sobre la eficacia de dicha prueba se conoce lo siguiente:

si se denota por E el evento de que un paciente tenga la enfermedad y por N el evento de que la prueba resulte negativa, entonces se sabe que

$$\begin{aligned} P(N^c | E) &= 0.95, \\ P(N | E^c) &= 0.96, \\ P(E) &= 0.01. \end{aligned}$$

Observe que esta información corresponde al caso cuando se conoce la situación médica del paciente, es decir, si está enfermo o no lo está. Con únicamente estos datos, uno podría pensar que la prueba es muy buena, sin embargo calcularemos las probabilidades $P(E | N)$ y $P(E | N^c)$ para saber la efectividad de la prueba cuando una persona recibe sus resultados. Usando el teorema de Bayes tenemos que

$$\begin{aligned} P(E | N) &= \frac{P(N | E)P(E)}{P(N | E)P(E) + P(N | E^c)P(E^c)} \\ &= \frac{0.05 \times 0.01}{0.05 \times 0.01 + 0.96 \times 0.99} \\ &= 0.000526. \end{aligned}$$

El evento al que se refiere la probabilidad anterior es llamado un falso negativo, es decir, es la situación cuando la prueba ha dado un resultado negativo pero ello es falso pues el paciente realmente tiene la enfermedad. Es bueno que esta probabilidad sea pequeña pues indica que cuando la prueba es negativa, con cierta confianza se puede asegurar que el paciente no tiene la enfermedad. Calcularemos ahora la probabilidad de un evento verdadero positivo, es decir, la probabilidad de que el paciente tenga la enfermedad cuando la prueba ha dado positivo.

$$\begin{aligned} P(E | N^c) &= \frac{P(N^c | E)P(E)}{P(N^c | E)P(E) + P(N^c | E^c)P(E^c)} \\ &= \frac{0.95 \times 0.01}{0.95 \times 0.01 + 0.04 \times 0.99} \\ &= 0.193. \end{aligned}$$

Esta última probabilidad es demasiado pequeña y por lo tanto la prueba es muy poco confiable en tales casos. Como un ejercicio simple se deja al lector el cálculo e interpretación de las probabilidades $P(E^c | N^c)$ (falso positivo) y $P(E^c | N)$ (verdadero negativo). •

Ejercicios

126. **La paradoja de la caja de Bertrand**⁶. Se tienen tres cajas y cada una de ellas tiene dos monedas. La caja C_1 contiene dos monedas de oro. La caja C_2 contiene una moneda de oro y una de plata. La caja C_3 contiene dos monedas de plata. Véase la Figura 1.27. Se selecciona una caja al azar y de allí se escoge una moneda. Si resulta que la moneda escogida es de oro, ¿cuál es la probabilidad de que provenga de la caja con dos monedas de oro? Responda los siguientes incisos.

- Argumente con palabras que la respuesta es $1/2$.
- Demuestre que la respuesta es $2/3$.

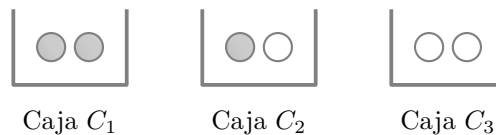


Figura 1.27

127. Una persona toma al azar uno de los números 1, 2 o 3, con idéntica probabilidad cada uno de ellos, y luego tira un dado equilibrado tantas veces como indica el número escogido. Finalmente suma los resultados de las tiradas del dado. Calcule la probabilidad de que

- se obtenga un total de 5.

⁶Joseph Louis François Bertrand (1822–1900), matemático francés.

b) se haya escogido el número 2 dado que la suma de las tiradas del dado es 8.

128. Suponga que se cuenta con dos urnas con la siguiente configuración: la urna I contiene 1 bola negra y 1 blanca, la urna II contiene 2 bolas negras y 2 blancas. Esta situación se ilustra en la Figura 1.28. Se escogen al azar dos bolas de la urna II, una a la vez, y se transfieren a la urna I. Se escoge después una bola al azar de la urna I. Calcule la probabilidad de que

- a) la bola escogida sea blanca.
 b) al menos una bola blanca haya sido transferida dado que la bola escogida es blanca.



Figura 1.28

129. **Canal binario ruidoso.** Considere nuevamente la situación en donde los símbolos binarios 0 y 1 se envían a través de un canal ruidoso como el que se muestra en la Figura 1.26 de la página 81. Debido al ruido, un 0 se distorsiona en un 1 con probabilidad 0.2 y un 1 se distorsiona en un 0 con probabilidad 0.1. Suponga que el símbolo de entrada 0 aparece el 45% de las veces y el símbolo 1 el 55%. Encuentre la probabilidad de que

- a) se haya enviado un 0 dado que se recibió un 0.
 b) se haya enviado un 1 dado que se recibió un 1.

130. Una caja contiene 3 bolas blancas y 4 bolas azules como se muestra en la Figura 1.29. Suponga que se extraen dos bolas al azar, una después de otra y sin reemplazo. Calcule la probabilidad de que

- a) la segunda bola sea azul dado que la primera es azul.
- b) la segunda bola sea del mismo color que la primera.
- c) las dos bolas sean de distinto color.
- d) la segunda bola sea azul.
- e) la primera bola sea azul dado que la segunda es azul.

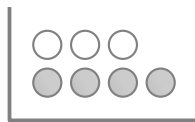


Figura 1.29

131. La urna I contiene 2 canicas blancas y 4 rojas. La urna II contiene 1 canica blanca y 1 roja. Se toma una canica al azar, sin verla, de la urna I y se coloca en la urna II. Después se toma una canica al azar de la urna II. Calcule la probabilidad de que la canica seleccionada de la urna I haya sido roja dado que la que se obtuvo de la urna II es roja.

1.16. Independencia de eventos

El concepto de independencia representa la situación cuando la ocurrencia de un evento no afecta la probabilidad de ocurrencia de otro evento. Es un concepto importante que se deriva de observaciones de situaciones reales y su utilización reduce considerablemente el cálculo de probabilidades conjuntas.

Definición 1.11 Se dice que los eventos A y B son independientes si se cumple la igualdad

$$P(A \cap B) = P(A)P(B). \quad (1.4)$$

Bajo la hipótesis adicional de que $P(B) > 0$, la condición de independencia (1.4) puede escribirse como $P(A|B) = P(A)$ y esto significa que la ocurrencia del evento B no afecta a la probabilidad de A . Análogamente, cuando $P(A) > 0$, la condición (1.4) se puede escribir como $P(B|A) = P(B)$, es decir, la ocurrencia del evento A no cambia a la probabilidad de B . Veamos algunos ejemplos.

Ejemplo 1.27 Considere un experimento aleatorio con espacio muestral equiprobable $\Omega = \{1, 2, 3, 4\}$.

- a) Los eventos $A = \{1, 2\}$ y $B = \{1, 3\}$ son independientes pues tanto $P(A \cap B)$ como $P(A)P(B)$ coinciden en el valor $1/4$.
- b) Los eventos $A = \{1, 2, 3\}$ y $B = \{1, 3\}$ no son independientes pues $P(A \cap B) = 1/2$, mientras que $P(A)P(B) = (3/4)(1/2) = 3/8$.

•

Ejemplo 1.28 (Dos eventos que no son independientes). Recordemos del Ejercicio 111 en la página 75, que un ensayo en una urna de Polya consiste en escoger una bola al azar de una urna con una configuración inicial de r bolas rojas y b bolas blancas y que la bola escogida se regresa a la urna junto con c bolas del mismo color. Considere los eventos R_1 y R_2 de la urna del Polya, es decir, obtener una bola roja en la primera y en la segunda extracción, respectivamente. Es claro que estos eventos no son independientes pues

$$P(R_2 | R_1)P(R_1) \neq P(R_2).$$

•

Inicialmente, uno podría asociar la idea de independencia de dos eventos con el hecho de que éstos son ajenos, pero ello es erróneo en general. A continuación ilustraremos esta situación.

Ejemplo 1.29 (Independencia $\not\Rightarrow$ Ajenos). Considere un evento $A \neq \emptyset$ junto con el espacio muestral Ω . Es claro que A y Ω son independientes pues $P(A \cap \Omega) = P(A)P(\Omega)$. Sin embargo, $A \cap \Omega = A \neq \emptyset$. Por lo tanto, el hecho de que dos eventos sean independientes no implica necesariamente que sean ajenos.

•

Como complemento al ejemplo anterior, mostraremos ahora el hecho de que dos eventos pueden ser ajenos sin ser independientes. Véase también el Ejercicio 137 en la página 94 para otro ejemplo de este tipo de situaciones.

Ejemplo 1.30 (Ajenos $\not\Rightarrow$ Independencia). Considere el experimento aleatorio de lanzar un dado equilibrado y defina los eventos A como obtener un número par y B como obtener un número impar. Es claro que los eventos A y B son ajenos, sin embargo no son independientes pues

$$0 = P(A \cap B) \neq P(A)P(B) = (1/2)(1/2).$$

Por lo tanto, dos eventos pueden ser ajenos y ello no implica necesariamente que sean independientes. •

La definición de independencia de dos eventos puede extenderse al caso de una colección finita de eventos de la siguiente forma.

Definición 1.12 Decimos que n eventos A_1, \dots, A_n son (mutuamente) independientes si se satisfacen todas y cada una de las condiciones siguientes:

$$P(A_i \cap A_j) = P(A_i)P(A_j), \quad i, j \text{ distintos.} \quad (1.5)$$

$$P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k), \quad i, j, k \text{ distintos.} \quad (1.6)$$

$$\vdots$$

$$P(A_1 \cap \dots \cap A_n) = P(A_1) \dots P(A_n). \quad (1.7)$$

Observe que hay en total $\binom{n}{2}$ identidades diferentes de la forma (1.5). Si se cumplen todas las identidades de esta forma, decimos que la colección de eventos es independiente dos a dos, es decir, independientes tomados por pares. Análogamente hay $\binom{n}{3}$ identidades diferentes de la forma (1.6). Si se cumplen todas las identidades de esta forma, decimos que la colección de eventos es independiente tres a tres, es decir, independientes tomados por tercias. Y así sucesivamente hasta la identidad de la forma (1.7), de la cual sólo hay una expresión.

En general, para verificar que n eventos son independientes, es necesario comprobar todas y cada una de las igualdades arriba enunciadas. Es decir, cualquiera de estas igualdades no implica, en general, la validez de alguna otra. En los ejemplos que aparecen abajo mostramos que la independencia dos a dos no implica la independencia tres a tres, ni viceversa. Por otro lado, no es difícil darse cuenta que el total de igualdades que es necesario verificar para que n eventos sean independientes es $2^n - n - 1$.

Ejemplo 1.31

(Independencia dos a dos $\not\Rightarrow$ Independencia tres a tres).

Considere el siguiente espacio muestral equiprobable junto con los eventos:

$$\begin{aligned}\Omega &= \{1, 2, 3, 4\}, & A &= \{1, 2\}, \\ & & B &= \{2, 3\}, \\ & & C &= \{2, 4\}.\end{aligned}$$

Los eventos A, B y C no son independientes pues, aunque se cumplen las igualdades $P(A \cap B) = P(A)P(B)$, $P(A \cap C) = P(A)P(C)$, y $P(B \cap C) = P(B)P(C)$, sucede que $P(A \cap B \cap C) \neq P(A)P(B)P(C)$. Esto muestra que, en general, la independencia dos a dos no implica la independencia tres a tres. •

Ejemplo 1.32

(Independencia tres a tres $\not\Rightarrow$ Independencia dos a dos).

Considere ahora el siguiente espacio muestral equiprobable junto con los eventos:

$$\begin{aligned}\Omega &= \{1, 2, 3, 4, 5, 6, 7, 8\}, & A &= \{1, 2, 3, 4\}, \\ & & B &= \{1, 5, 6, 7\}, \\ & & C &= \{1, 2, 3, 5\}.\end{aligned}$$

Estos eventos cumplen la condición $P(A \cap B \cap C) = P(A)P(B)P(C)$, pero $P(A \cap B) \neq P(A)P(B)$. Esto muestra que, en general, la independencia tres a tres no implica la independencia dos a dos. •

Finalmente, mencionaremos que el concepto de independencia puede extenderse al caso de colecciones infinitas de eventos de la siguiente forma.

Definición 1.13 Se dice que un colección infinita de eventos es independiente si cualquier subcolección finita de ella lo es.

Ejercicios

132. Sean A y B eventos tales que $P(A) = 4/10$ y $P(A \cup B) = 7/10$. Encuentre la probabilidad de B suponiendo que:
- A y B son independientes.
 - A y B son ajenos.
 - $P(A|B) = 1/2$.
133. Se lanza un dado equilibrado dos veces. Determine si los siguientes pares de eventos son independientes.
- A = “La suma de los dos resultados es 6.”
 B = “El primer resultado es 4.”
 - A = “La suma de los dos resultados es 7.”
 B = “El segundo resultado es 4.”
134. Demuestre o proporcione un contraejemplo.
- A y \emptyset son independientes.
 - A y Ω son independientes.
 - Si $P(A|B) = P(A|B^c)$ entonces A y B son independientes.
 - Si A y B son independientes entonces $P(A|B) = P(A|B^c)$.
 - Si A tiene probabilidad cero o uno, entonces A es independiente de cualquier otro evento.
 - Si A es independiente consigo mismo entonces $P(A) = 0$.

- g) Si A es independiente consigo mismo entonces $P(A) = 1$.
 h) Si $P(A) = P(B) = P(A|B) = 1/2$ entonces A y B son independientes.

135. Demuestre que las siguientes cuatro afirmaciones son equivalentes:

- a) A y B son independientes.
 b) A y B^c son independientes.
 c) A^c y B son independientes.
 d) A^c y B^c son independientes.

136. Demuestre que los eventos A y B son independientes si y sólo si

$$P(A|B) = P(A|B^c).$$

137. **Independencia y ser ajenos.** Sea $\Omega = \{a, b, c, d\}$ un espacio muestral equiprobable. Defina los eventos $A = \{a, b\}$, $B = \{a, c\}$ y $C = \{a\}$. Compruebe que los siguientes pares de eventos satisfacen las propiedades indicadas. Esto demuestra que, en términos generales, la propiedad de ser ajenos y la independencia no están relacionadas.

Eventos	Ajenos	Independientes
A, C	×	×
A, B	×	✓
A, A^c	✓	×
A, \emptyset	✓	✓

138. Sean A y B dos eventos independientes. Encuentre una expresión en términos de $P(A)$ y $P(B)$ para las siguientes probabilidades:

- a) $P(A \cup B)$.
 b) $P(A \cup B^c)$.
 c) $P(A - B)$.
 d) $P(A \Delta B)$.
 e) $P(A - (A \cap B))$.
 f) $P(A^c \cup B^c)$.

139. Sean A y B dos eventos independientes tales que $P(A) = a$ y $P(B) = b$. Calcule la probabilidad de que

- a) no ocurra ninguno de estos dos eventos.
 - b) ocurra exactamente uno de estos dos eventos.
 - c) ocurra por lo menos uno de estos dos eventos.
 - d) ocurran los dos eventos.
 - e) ocurra a lo sumo uno de estos dos eventos.
140. Sean A y B dos eventos independientes. Demuestre o proporcione un contraejemplo para las siguientes afirmaciones:
- a) A y $A \cup B$ son independientes.
 - b) A y $A \cap B$ son independientes.
 - c) A y $B - A$ son independientes.
 - d) $A - B$ y $B - A$ son independientes.
141. Demuestre que las siguientes cuatro afirmaciones son equivalentes:
- a) A, B y C son independientes.
 - b) A^c, B y C son independientes.
 - c) A^c, B^c y C son independientes.
 - d) A^c, B^c y C^c son independientes.
142. Sean A, B y C tres eventos tales que A es independiente de B y A también es independiente de C . Demuestre o proporcione un contraejemplo para las siguientes afirmaciones:
- a) B y C son independientes.
 - b) A, B y C son independientes.
 - c) A y $(B \cap C)$ son independientes.
 - d) A y $(B \cup C)$ son independientes.
143. Demuestre que si A es un evento independiente, por separado, de B , de C y del evento
- a) $(B \cup C)$, entonces A y $(B \cap C)$ son independientes.
 - b) $(B \cap C)$, entonces A y $(B \cup C)$ son independientes.

144. **Radares.** En una cierta zona geográfica estratégica se han colocado tres radares para detectar el vuelo de aviones de baja altura. Cada radar funciona de manera independiente y es capaz de detectar un avión con probabilidad 0.85. Si un avión atraviesa la zona en estudio, calcule la probabilidad de que
- no sea detectado por ninguno de los radares.
 - sea detectado por lo menos por uno de los radares.
 - sea detectado por lo menos por dos de los radares.
145. Sean A , B y C tres eventos independientes. Demuestre directamente que los siguientes pares de eventos son independientes.
- A y $B \cap C$.
 - A y $B \cup C$.
 - $(A \cap B)$ y C .
 - $(A \cup B)$ y C .
 - B y $(A - C)$.
 - B y $(A^c \cup C^c)$.
146. Sean A , B y C tres eventos independientes. Encuentre una expresión en términos de $P(A)$, $P(B)$ y $P(C)$, o sus complementos, para las siguientes probabilidades:
- $P(A \cup B \cup C)$.
 - $P(A \cup B \cup C^c)$.
 - $P(A^c \cap B^c \cap C^c)$.
 - $P(A - B - C)$.
 - $P(A \cap (B \cup C))$.
 - $P(A \cap B \cap C^c)$.
 - $P(A \cup (B \cap C))$.
 - $P(A \Delta B)$.
 - $P(A \Delta B \Delta C)$.
147. Sean A , B y C tres eventos independientes tales que $P(A) = p$, $P(B) = q$ y $P(C) = r$. Calcule la probabilidad de que
- no ocurra ninguno de estos tres eventos.
 - ocurra exactamente uno de estos tres eventos.
 - ocurra por lo menos uno de estos tres eventos.
 - ocurran exactamente dos de estos tres eventos.
 - ocurran por lo menos dos de estos tres eventos.
 - ocurra a lo sumo uno de estos tres eventos.
 - ocurran a lo sumo dos de estos tres eventos.

h) ocurran los tres eventos.

148. Use el teorema del binomio para demostrar que el total de igualdades por verificar para comprobar que n eventos son independientes es

$$2^n - 1 - n.$$

149. Sean B_1 y B_2 dos eventos, cada uno de ellos independiente del evento A . Proporcione un contraejemplo para las siguientes afirmaciones:

- a)* A y $B_1 \cup B_2$ son independientes.
- b)* A y $B_1 \cap B_2$ son independientes.
- c)* A y $B_1 - B_2$ son independientes.
- d)* A y $B_1 \triangle B_2$ son independientes.

150. Demuestre que si A_1, \dots, A_n son eventos independientes entonces

$$P(A_1 \cup \dots \cup A_n) = 1 - P(A_1^c) \cdots P(A_n^c).$$

151. Sean A_1, \dots, A_n eventos mutuamente independientes, cada uno de ellos con probabilidad p , y sea m un entero tal que $0 \leq m \leq n$. Calcule la probabilidad de que

- a)* al menos uno de estos eventos ocurra.
- b)* exactamente m de estos eventos ocurran.
- c)* al menos m de estos eventos ocurran.
- d)* a lo sumo m de estos eventos ocurran.

152. **Independencia condicional.** Se dice que los eventos A y B son independientes condicionalmente al evento C si

$$P(A \cap B | C) = P(A | C)P(B | C),$$

suponiendo de antemano que $P(C) > 0$. Proporcione contraejemplos adecuados para demostrar que

- a)* la independencia de dos eventos no implica necesariamente su independencia condicional dado un tercer evento.

- b) la independencia condicional de dos eventos, dado un tercero, no implica necesariamente la independencia de los dos primeros.

En general, pueden darse ejemplos de las situaciones que se presentan en la siguiente tabla.

A, B indep.	A, B indep. $ C$
✓	✓
✓	×
×	✓
×	×

153. Un cierto componente de una máquina falla el 5% de las veces que se enciende. Para obtener una mayor confiabilidad se colocan n componentes de las mismas características en un arreglo en paralelo, como se muestra en la Figura 1.30, de tal forma que ahora el conjunto de componentes falla cuando todos fallan. Suponga que el comportamiento de cada componente es independiente uno del otro. Determine el mínimo valor de n a fin de garantizar el funcionamiento de la máquina por lo menos el 99% de las veces.

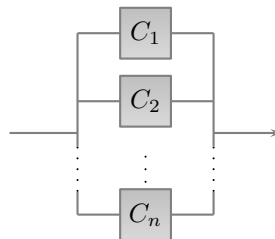


Figura 1.30

154. Dos jugadores de basketbol alternan turnos para efectuar tiros libres hasta que uno de ellos enceste. En cada intento, la probabilidad de encestar es p para el primer jugador y q para el segundo jugador, siendo los resultados de los tiros independientes unos de otros. Calcule la probabilidad de encestar primero de cada uno de los jugadores.

1.17. Continuidad de la probabilidad

Para concluir este capítulo revisaremos una propiedad que cumple toda medida de probabilidad, la cual es ligeramente más avanzada que las que hasta ahora hemos estudiado. Aunque la demostración de este resultado no es complicada, por simplicidad en esta exposición la omitiremos, para concentrarnos en su interpretación y aplicación en las siguientes secciones.

Proposición 1.10 Sea A_1, A_2, \dots una sucesión infinita creciente de eventos, es decir, $A_1 \subseteq A_2 \subseteq \dots$. Entonces

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n).$$

Para sucesiones de eventos como de las que se habla en la proposición anterior, es decir, crecientes, es natural definir su límite como la unión de todos los eventos. Tenemos entonces la siguiente definición.

Definición 1.14 Sea A_1, A_2, \dots una sucesión infinita de eventos.

- a) Si $A_1 \subseteq A_2 \subseteq \dots$ se define $\lim_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} A_n$.
- b) Si $A_1 \supseteq A_2 \supseteq \dots$ se define $\lim_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} A_n$.

Hemos incluido, en la definición anterior, también el caso cuando la sucesión de eventos es decreciente. En tal situación, el límite de la sucesión se define como la intersección de todos los eventos. Observe que la unión infinita en el

caso creciente no necesariamente es la totalidad del espacio muestral, pero es un subconjunto de él. En el caso decreciente, la intersección infinita no necesariamente es el conjunto vacío.

Por lo tanto, la proposición anterior establece que la probabilidad del límite de la sucesión coincide con el límite de las probabilidades. Este intercambio de límite y probabilidad es la definición de continuidad, en este caso para medidas de probabilidad. Este resultado puede extenderse al caso de sucesiones monótonas decrecientes.

Proposición 1.11 Sea A_1, A_2, \dots una sucesión infinita decreciente de eventos, es decir, $A_1 \supseteq A_2 \supseteq \dots$. Entonces

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n).$$

Haremos uso de estos resultados en el siguiente capítulo.

Ejercicios

155. Demuestre que las Proposiciones 1.10 y 1.11 son equivalentes, es decir, a partir de uno de estos resultados se puede obtener el otro.
156. Considere las siguientes sucesiones de subconjuntos de \mathbb{R} . En caso de que exista, encuentre el límite de cada una de estas sucesiones de eventos. Para cada número natural n se define:
- | | |
|---------------------------------|---------------------------------|
| a) $A_n = (-\infty, a + 1/n)$. | e) $A_n = (-n, n)$. |
| b) $A_n = (-\infty, a + 1/n]$. | f) $A_n = (a + 1/n, \infty)$. |
| c) $A_n = (-1/n, 1/n)$. | g) $A_n = [a - 1/n, \infty)$. |
| d) $A_n = [0, 1/n]$. | h) $A_n = (a - 1/n, b + 1/n)$. |
157. **Teorema de probabilidad total extendido.** Sea B_1, B_2, \dots una partición infinita de Ω tal que $P(B_i) \neq 0$, $i = 1, 2, \dots$. Demuestre que

para cualquier evento A ,

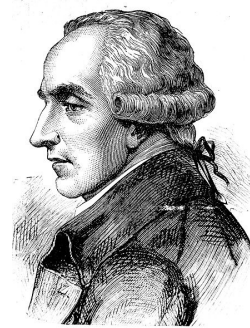
$$P(A) = \sum_{i=1}^{\infty} P(A | B_i)P(B_i).$$

158. **Teorema de Bayes extendido.** Sea B_1, B_2, \dots una partición infinita de Ω tal que $P(B_i) \neq 0, i = 1, 2, \dots$. Sea A un evento tal que $P(A) \neq 0$. Entonces para cada $j = 1, 2, \dots$

$$P(B_j | A) = \frac{P(A | B_j)P(B_j)}{\sum_{i=1}^{\infty} P(A | B_i)P(B_i)}.$$

Pierre-Simon Laplace

P.-S. Laplace (Francia, 1749–1827) nació en el seno de una familia medianamente acomodada, dedicada a la agricultura y al comercio. Su padre tuvo la intención de que Laplace siguiera una carrera eclesiástica, de modo que de pequeño asistió a una escuela local a cargo de una orden religiosa católica y a la edad de 16 años fue enviado a la Universidad de Caen para estudiar teología. En esta universidad tuvo profesores de matemáticas que despertaron vivamente su interés en esta disciplina y quienes reconocieron el potencial matemático del joven estudiante.



P.-S. Laplace

Laplace dejó la Universidad de Caen sin graduarse, pero consiguió una carta de presentación de sus profesores y viaja a París cuando tenía 19 años para presentarse ante d'Alembert. Después de una fría y escéptica recepción por parte de d'Alembert, Laplace logró demostrarle su capacidad de trabajo y su sorprendente habilidad natural para las matemáticas. Por medio de una carta de recomendación de d'Alembert, Laplace consiguió un puesto como profesor de matemáticas en la *École Militaire* de París. Una vez que Laplace aseguró un ingreso económico estable, pudo entonces establecerse en París y dedicarse con mayor empeño a sus trabajos de investigación. En efecto, de 1771 a 1787 logró producir gran parte de su monumental trabajo original en astronomía matemática. Contribuyó notablemente además al desarrollo de varias áreas de la ciencia, en particular la mecánica, la probabilidad, la estadística, la física matemática, el análisis matemático y las ecuaciones diferenciales y en diferencias. La diversidad de temas de sus trabajos científicos muestra su amplia capacidad intelectual y su dominio de muchas áreas de la física y la matemática de su tiempo. En 1773, a la edad de 24 años y después de varios intentos, logró ser nombrado miembro adjunto de la prestigiosa *Académie des Sciences* de Francia, donde, a lo largo de los años, ocupó diversos nombramientos y desde donde hizo sentir la influencia de sus opiniones y pensamientos. Es en esos años cuando Laplace consolidó su fama como astrónomo y matemático. En su obra monumental titulada *Mécanique Céleste*, la cual consta de cinco volúmenes y fue

publicada a lo largo de varios años (1799-1825), Laplace resume y extiende los resultados de sus predecesores. En este trabajo traslada los estudios geométricos de la mecánica clásica al de una mecánica basada en el cálculo diferencial e integral, abriendo con ello nuevas perspectivas de desarrollo. En 1785, siendo Laplace examinador de estudiantes de la Escuela de Artilleros, examinó y acreditó a un joven estudiante de 16 años llamado Napoleón Bonaparte (1769-1821). En 1799, 14 años después, Laplace fue nombrado Ministro del Interior de Francia por parte, justamente, de Napoleón Bonaparte. Sin embargo, Laplace ocupó tal cargo por únicamente seis semanas, y con bastante malos resultados, según un reporte del mismo Napoleón. En 1812 Laplace publicó un tratado titulado *Theorie Analytique des Probabilités*, en el cual establece varios resultados fundamentales de la probabilidad y la estadística. En particular, estudia el concepto de probabilidad clásica como aparece definida en este texto y analiza varias de las propiedades que hemos estudiado. La primera edición de este trabajo fue dedicada a Napoleón Bonaparte, sin embargo tal referencia fue omitida en ediciones subsecuentes, a partir de la abdicación de éste. Este pequeño hecho muestra los cambios de postura en cuestiones políticas que Laplace tuvo y que sus amigos y colegas le recriminaron. Se cree que Laplace buscaba alejarse de la política y de los asuntos de la vida ajenos a su quehacer científico, pero ello no era fácil dado el liderazgo que, como hombre de ciencia, naturalmente ocupaba y a los tiempos turbulentos que le tocó vivir en Francia. A Laplace se le considera como un hombre con una extraordinaria capacidad para las matemáticas. Sin duda, fue uno de los pensadores más brillantes dentro del grupo de sus contemporáneos y su influencia y contribuciones al desarrollo de la ciencia fueron notables. En honor a ello, su nombre aparece grabado en la torre Eiffel. Murió en París el 5 de marzo de 1827, a la edad de 77 años.

Fuente: Archivo MacTutor, Universidad de St. Andrews [24].

Andrey Nikolaevich Kolmogorov

A. N. Kolmogorov (Rusia, 1903–1987) quedó huérfano en el momento de nacer pues en el parto mismo murió su madre. Ella no estaba casada con el padre y éste no se ocupó del recién nacido. De su padre se conoce muy poco, se cree que fue muerto en 1919 durante la guerra civil rusa. Así, el pequeño Kolmogorov fue criado por dos de sus tías en la casa de su abuelo paterno, de quien adquiere el apellido Kolmogorov. En 1910 se mudó a Moscú con una de sus tías y realizó allí sus primeros estudios. Desde muy pequeño mostró inclinación por la ciencia y la literatura. Siendo adolescente diseñó varias máquinas de “movimiento perpetuo”, para las cuales no era evidente descubrir el mecanismo de funcionamiento. En 1920 ingresó a la Universidad Estatal de Moscú y al Instituto Tecnológico de Química, en donde estudió diversas disciplinas además de matemáticas, entre ellas metalurgia e historia de Rusia. En 1929, a la edad de 19 años empezó a tener fama mundial al publicar un trabajo matemático en donde construye una serie de Fourier que es divergente casi donde quiera. Alrededor de estas fechas es cuando decidió dedicarse completamente a las matemáticas. En 1925 se graduó en la Universidad Estatal de Moscú y en 1929 obtuvo el doctorado, habiendo publicado, para ese entonces, 18 trabajos científicos. Se incorporó a la misma universidad en 1931 como profesor. En 1933 publicó su famoso libro *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Fundamentos de la Teoría de la Probabilidad), en donde estableció los fundamentos axiomáticos de la teoría de la probabilidad y los cuales hemos estudiado en este texto. Así, Kolmogorov es uno de los fundadores de la teoría moderna de la probabilidad, quien también contribuyó, con brillantez, en muchas otras áreas de las matemáticas y la física como la topología, la lógica, los fenómenos de turbulencia, la mecánica clásica, la teoría de la información y la complejidad computacional. Es interesante mencionar que Kolmogorov tuvo particular interés en un proyecto consistente en proveer educación especial a niños sobresalientes, a quienes dedicó tiempo para crearles las condiciones materiales y de estudio para que estos tuvieran una



A. N. Kolmogorov

educación amplia e integral. Kolmogorov recibió numerosos reconocimientos por la profundidad e importancia de sus trabajos científicos, tales reconocimientos provinieron no únicamente de Rusia, sino también de otros países y de varias universidades y academias científicas internacionales.

Fuente: Archivo MacTutor, Universidad de St. Andrews [24].

Capítulo 2

Variables aleatorias

En este capítulo definiremos a una variable aleatoria como una función del espacio muestral en el conjunto de números reales. Esto nos permitirá considerar que el resultado del experimento aleatorio es el número real tomado por la variable aleatoria. En consecuencia, nuestro interés en el estudio de los experimentos aleatorios se trasladará al estudio de las distintas variables aleatorias y sus características particulares.

2.1. Variables aleatorias

Consideremos que tenemos un experimento aleatorio cualquiera, junto con un espacio de probabilidad asociado (Ω, \mathcal{F}, P) .

Definición 2.1 Una variable aleatoria es una transformación X del espacio de resultados Ω al conjunto de números reales, esto es,

$$X : \Omega \rightarrow \mathbb{R},$$

tal que para cualquier número real x ,

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}. \quad (2.1)$$

En ocasiones, el término variable aleatoria se escribe de manera breve como v.a. y su plural con una s al final. La condición (2.1) será justificada más adelante. Supongamos entonces que se efectúa el experimento aleatorio una vez y se obtiene un resultado ω en Ω . Al transformar este resultado con la variable aleatoria X se obtiene un número real $X(\omega) = x$. Así, una variable aleatoria es una función determinista y no es variable ni aleatoria, sin embargo tales términos se justifican al considerar que los posibles resultados del experimento aleatorio son los diferentes números reales x que la función X puede tomar. De manera informal, uno puede pensar también que una variable aleatoria es una pregunta o medición que se hace sobre cada uno de los resultados del experimento aleatorio y cuya respuesta es un número real, así cada resultado ω tiene asociado un único número x . De manera gráfica se ilustra el concepto de variable aleatoria en la Figura 2.1.

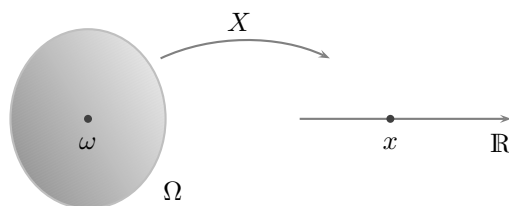


Figura 2.1

En lo sucesivo emplearemos la siguiente notación. Si A es un conjunto de Borel de \mathbb{R} , entonces la expresión $(X \in A)$, incluyendo el paréntesis, denota el conjunto $\{\omega \in \Omega : X(\omega) \in A\}$, es decir,

$$(X \in A) = \{\omega \in \Omega : X(\omega) \in A\}.$$

En palabras, la expresión $(X \in A)$ denota aquel conjunto de elementos ω del espacio muestral Ω tales que bajo la aplicación de la función X toman un valor dentro del conjunto A . A este conjunto se le llama la imagen inversa de A y se le denota por $X^{-1}A$, lo cual no debe confundirse con la función inversa de X , pues ésta puede no existir. Véase el Ejercicio 160, en donde

se pide demostrar algunas propiedades sencillas de la imagen inversa. Por ejemplo, consideremos que el conjunto A es el intervalo (a, b) , entonces el evento $(X \in (a, b))$ también puede escribirse como $(a < X < b)$ y es una abreviación del evento

$$\{\omega \in \Omega : a < X(\omega) < b\}.$$

Como otro ejemplo, considere que el conjunto A es el intervalo infinito $(-\infty, x]$ para algún valor real x fijo. Entonces el evento $(X \in (-\infty, x])$ también puede escribirse de manera breve como $(X \leq x)$ y significa

$$\{\omega \in \Omega : -\infty < X(\omega) \leq x\},$$

que es justamente el conjunto al que se hace referencia en la expresión (2.1) de la definición anterior. A esta propiedad se le conoce como la condición de medibilidad de la función X respecto de la σ -álgebra \mathcal{F} del espacio de probabilidad y la σ -álgebra de Borel de \mathbb{R} . No haremos mayor énfasis en la verificación de esta condición para cada variable aleatoria que se defina, pero dicha propiedad es importante pues permite trasladar la medida de probabilidad del espacio de probabilidad a la σ -álgebra de Borel de \mathbb{R} del siguiente modo.

Medida de probabilidad inducida

Para cualquier intervalo de la forma $(-\infty, x]$ se puede obtener su imagen inversa bajo X , es decir, $X^{-1}(-\infty, x] = \{\omega \in \Omega : X(\omega) \leq x\}$. Como este conjunto pertenece a \mathcal{F} por la condición (2.1), se puede aplicar la medida de probabilidad P pues ésta tiene como dominio \mathcal{F} . Así, mediante la función X puede trasladarse la medida de probabilidad P a intervalos de la forma $(-\infty, x]$ y puede demostrarse que ello es suficiente para extenderla a la totalidad de la σ -álgebra de Borel de \mathbb{R} .

A esta nueva medida de probabilidad se le denota por $P_X(\cdot)$ y se le llama la medida de probabilidad inducida por la variable aleatoria X . Por simplicidad omitiremos el subíndice del término P_X , de modo que adoptará la misma notación que la medida de probabilidad del espacio de probabilidad original (Ω, \mathcal{F}, P) . De esta forma, tenemos un nuevo espacio de probabilidad $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P)$, véase la Figura 2.2, el cual tomaremos como elemento base de ahora en adelante sin hacer mayor énfasis en ello.

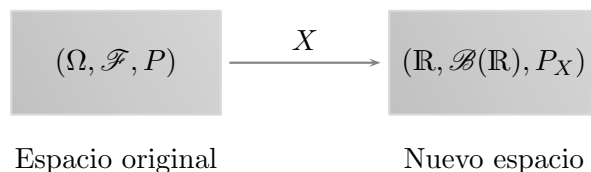


Figura 2.2

Nuestro interés es el estudio de los distintos eventos de la forma $(X \in A)$ y sus probabilidades, en donde X es una variable aleatoria y A es un conjunto de Borel de \mathbb{R} , por ejemplo, un intervalo de la forma (a, b) .

Seguiremos también la notación usual de utilizar la letra mayúscula X para denotar una variable aleatoria cualquiera, es decir, X es una función de Ω en \mathbb{R} mientras que la letra minúscula x denota un número real y representa un posible valor de la variable aleatoria. En general, las variables aleatorias se denotan usando las últimas letras del alfabeto en mayúsculas: U, V, W, X, Y, Z , y para un valor cualquiera de ellas se usa la misma letra en minúscula: u, v, w, x, y, z .

Ejemplo 2.1 Suponga que un experimento aleatorio consiste en lanzar al aire una moneda equilibrada y observar la cara superior una vez que la moneda cae. Denotemos por “Cara” y “Cruz” los dos lados de la moneda. Entonces, claramente, el espacio muestral es el conjunto $\Omega = \{\text{“Cara”}, \text{“Cruz”}\}$. Defina la variable aleatoria $X : \Omega \rightarrow \mathbb{R}$ de la forma siguiente.

$$\begin{aligned} X(\text{“Cruz”}) &= 1, \\ X(\text{“Cara”}) &= 0. \end{aligned}$$

De este modo podemos suponer que el experimento aleatorio tiene dos valores numéricos: 0 y 1. Observe que estos números son arbitrarios pues cualquier otro par de números puede ser escogido como los valores de la variable aleatoria. Véase la Figura 2.3. Se muestran a continuación algunos ejemplos de eventos de esta variable aleatoria y sus correspondientes probabilidades.

a) $P(X \in [1, 2)) = P(\{\text{“Cruz”}\}) = 1/2.$

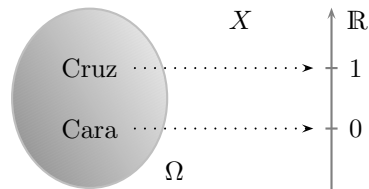


Figura 2.3

- b) $P(X \in [0, 1)) = P(\{\text{"Cara"}\}) = 1/2$.
- c) $P(X \in [2, 4]) = P(\emptyset) = 0$.
- d) $P(X = 1) = P(\{\text{"Cruz"}\}) = 1/2$.
- e) $P(X \leq -1) = P(\emptyset) = 0$.
- f) $P(X \geq 0) = P(\Omega) = 1$.

•

Ejemplo 2.2 Sea c una constante. Para cualquier experimento aleatorio con espacio muestral Ω se puede definir la función constante $X(\omega) = c$. Así, cualquier resultado del experimento aleatorio produce, a través de la función X , el número c . Véase la Figura 2.4. Decimos entonces que X es la variable aleatoria constante c y se puede verificar que para cualquier conjunto de Borel A de \mathbb{R} ,

$$P(X \in A) = \begin{cases} 1 & \text{si } c \in A, \\ 0 & \text{si } c \notin A. \end{cases}$$

•

Ejemplo 2.3 Consideremos el experimento aleatorio consistente en lanzar un dardo en un tablero circular de radio uno. El espacio muestral o conjunto

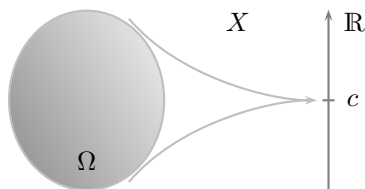


Figura 2.4

de posibles resultados de este experimento se puede escribir como $\Omega = \{(x, y) : x^2 + y^2 \leq 1\}$ y su representación gráfica se muestra en la Figura 2.5. Los siguientes son ejemplos de variables aleatorias, es decir, funciones de Ω en \mathbb{R} , asociadas a este experimento aleatorio. Para cualquier $(x, y) \in \Omega$ se define

- a) $X(x, y) = x$.
Esta función es la proyección del punto (x, y) sobre el eje horizontal. El conjunto de valores que la variable X puede tomar es el intervalo $[-1, 1]$.
- b) $Y(x, y) = y$.
Esta función es la proyección del punto (x, y) sobre el eje vertical. El conjunto de valores que la variable Y puede tomar es el intervalo $[-1, 1]$.
- c) $Z(x, y) = \sqrt{x^2 + y^2}$.
Esta función es la distancia del punto (x, y) al centro del círculo. El conjunto de valores que la variable Z puede tomar es el intervalo $[0, 1]$.
- d) $V(x, y) = |x| + |y|$.
Esta función es la así llamada “distancia del taxista” del punto (x, y) al origen. El conjunto de valores que la variable V puede tomar es el intervalo $[0, \sqrt{2}]$.
- e) $W(x, y) = xy$.
Esta función es el producto de las coordenadas del punto (x, y) . El

conjunto de valores que la variable W puede tomar es el intervalo $[-1/2, 1/2]$.

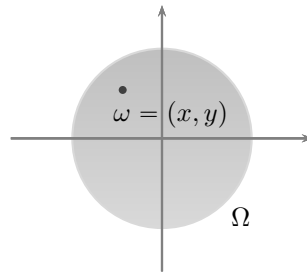


Figura 2.5

Otros ejemplos de variables aleatorias aparecen en la sección de ejercicios y muchas otras variables se definirán a lo largo del texto. Nuestro objetivo es el estudio de las distintas variables aleatorias pues éstas codifican en números los resultados de los experimentos aleatorios.

VARIABLES ALEATORIAS DISCRETAS Y CONTINUAS

Considerando el conjunto de valores que una variable aleatoria puede tomar, vamos a clasificar a las variables aleatorias en dos tipos: discretas o continuas. Decimos que una variable aleatoria es discreta cuando el conjunto de valores que ésta toma es un conjunto discreto, es decir, un conjunto finito o numerable. Por ejemplo, el conjunto $\{0, 1, 2, \dots, n\}$ es un conjunto discreto porque es finito, lo mismo \mathbb{N} pues, aunque éste es un conjunto infinito, es numerable y por lo tanto discreto. Por otra parte, decimos preliminarmente que una variable aleatoria es continua cuando toma todos los valores dentro de un intervalo $(a, b) \subseteq \mathbb{R}$. Cuando revisemos el concepto de función de distribución daremos una definición más precisa acerca de v.a. continua. Esta clasificación de variables aleatorias no es completa pues existen variables que no son de ninguno de los dos tipos mencionados. Véase el Ejercicio 200 en la página 144. Sin embargo, por simplicidad, nos concentraremos en estudiar primordialmente variables aleatorias de estos dos tipos.

Ejemplo 2.4 La variable X , definida en el Ejemplo 2.1, en la página 110, acerca del lanzamiento de una moneda, es una variable aleatoria discreta. En ese ejemplo, el espacio muestral mismo es discreto y por lo tanto las variables aleatorias que pueden allí definirse tienen que ser forzosamente discretas. En el Ejemplo 2.3, en la página 111, acerca del lanzamiento de un dardo en un tablero circular de radio uno, el espacio muestral de la Figura 2.5 es infinito no numerable, las variables X, Y, Z, V y W definidas allí son todas variables aleatorias continuas. Si se dibujan círculos concéntricos alrededor del origen y si se asignan premios asociados a cada una de las regiones resultantes, puede obtenerse un ejemplo de una variable aleatoria discreta sobre este espacio muestral. •

Ejemplo 2.5 Un experimento aleatorio consiste en escoger a una persona ω al azar dentro de una población humana dada. La variable aleatoria X evaluada en ω corresponde a conocer una de las características que aparecen en la lista de abajo acerca de la persona escogida. Debido a limitaciones en nuestras mediciones de estas características, en cada caso puede considerarse que la variable X es discreta.

- | | |
|---------------------|--------------|
| a) Edad en años. | c) Peso. |
| b) Número de hijos. | d) Estatura. |
-

Estaremos interesados también en considerar funciones de variables aleatorias, por ejemplo, si X es una v.a. entonces $Y = 5X + 2$ y $Y = X^2$ son funciones de X . En este texto supondremos que estos objetos son también variables aleatorias. Lo mismo sucederá al considerar funciones de varias variables aleatorias definidas sobre el mismo espacio de probabilidad, por ejemplo, si X y Y son dos v.a.s entonces también lo serán $X + Y$, $5X - 2Y - 4$, etc. En general, no nos preocuparemos en verificar que estas funciones cumplen la condición de medibilidad (2.1) que aparece en la definición de variable aleatoria.

En las siguientes secciones vamos a explicar la forma de asociar a cada variable aleatoria dos funciones que nos proveen de información acerca de

las características de la variable aleatoria. Estas funciones, llamadas función de densidad y función de distribución, nos permiten representar a un mismo tiempo tanto los valores que puede tomar la variable aleatoria como las probabilidades de los distintos eventos. Definiremos primero la función de probabilidad para una variable aleatoria discreta, después la función de densidad para una v.a. continua, y finalmente definiremos la función de distribución para ambos tipos de variables aleatorias.

Ejercicios

159. Suponga que un experimento aleatorio consiste en escoger un número al azar dentro del intervalo $(0, 1)$. Para cada resultado ω del experimento se expresa a este número en su expansión decimal

$$\omega = 0.a_1a_2a_3 \cdots$$

en donde $a_i \in \{0, 1, \dots, 9\}$, $i = 1, 2, \dots$. Para cada una de las siguientes variables aleatorias determine si ésta es discreta o continua y en cualquier caso establezca el conjunto de valores que puede tomar.

- a) $X(\omega) = 1 - \omega$.
- b) $Y(\omega) = a_1$.
- c) $Z(\omega) = 0.0a_1a_2a_3 \cdots$

160. **Imagen inversa.** Sean Ω_1 y Ω_2 dos conjuntos y sea $X : \Omega_1 \rightarrow \Omega_2$ una función. La imagen inversa de cualquier subconjunto $A \subseteq \Omega_2$ bajo la función X es un subconjunto de Ω_1 denotado por $X^{-1}A$ y definido como sigue

$$X^{-1}A = \{\omega \in \Omega_1 : X(\omega) \in A\}.$$

Véase la Figura 2.6 y observe que X es una función puntual mientras que X^{-1} es una función conjuntista, pues lleva subconjuntos de Ω_2 en subconjuntos de Ω_1 . Demuestre que la imagen inversa cumple las siguientes propiedades. Para A, A_1, A_2 subconjuntos de Ω_2 ,

- a) $X^{-1}A^c = (X^{-1}A)^c$.

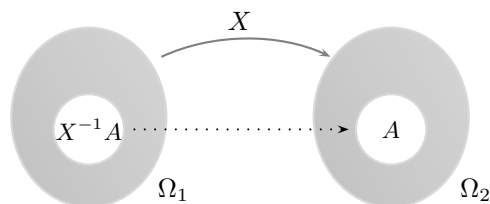


Figura 2.6

- b) Si $A_1 \subseteq A_2$ entonces $X^{-1}A_1 \subseteq X^{-1}A_2$.
- c) $X^{-1}(A_1 \cup A_2) = (X^{-1}A_1) \cup (X^{-1}A_2)$.
- d) $X^{-1}(A_1 \cap A_2) = (X^{-1}A_1) \cap (X^{-1}A_2)$.
- e) $X^{-1}(A_1 \Delta A_2) = (X^{-1}A_1) \Delta (X^{-1}A_2)$.

161. **Variable aleatoria constante.** Demuestre que la función $X : \Omega \rightarrow \mathbb{R}$, que es idénticamente constante c , es una variable aleatoria, es decir, cumple la propiedad de medibilidad (2.1) que aparece en la página 107.
162. Sea $X : \Omega \rightarrow \mathbb{R}$ una función y sean $x \leq y$ dos números reales. Demuestre que

$$(X \leq x) \subseteq (X \leq y).$$

163. **Función indicadora.** Sea (Ω, \mathcal{F}, P) un espacio de probabilidad y sea A un evento. Defina la variable aleatoria X como aquella que toma el valor 1 si A ocurre y toma el valor 0 si A no ocurre. Es decir,

$$X(\omega) = \begin{cases} 1 & \text{si } \omega \in A, \\ 0 & \text{si } \omega \notin A. \end{cases}$$

Demuestre que X es, efectivamente, una variable aleatoria. A esta variable se le llama función indicadora del evento A y se le denota también por $1_A(\omega)$.

164. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad tal que $\mathcal{F} = \{\emptyset, \Omega\}$. ¿Cuáles son todas las posibles variables aleatorias definidas sobre este espacio de probabilidad?

165. Considere el experimento aleatorio consistente en lanzar dos veces consecutivas un dado equilibrado. Sea x el resultado del primer lanzamiento y sea y el resultado del segundo lanzamiento. Identifique los siguientes eventos usando la representación gráfica del espacio muestral para este experimento que se muestra en la Figura 2.7 y calcule la probabilidad de cada uno de ellos.

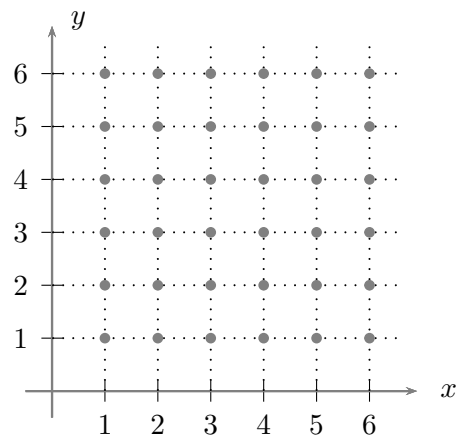


Figura 2.7

- a) $A = \{(x, y) : x \leq y\}$.
 b) $B = \{(x, y) : x + y = 6\}$.
 c) $C = \{(x, y) : x - y = 3\}$.
 d) $D = \{(x, y) : x + y \leq 4\}$.
 e) $E = \{(x, y) : x - y \geq 2\}$.
 f) $F = \{(x, y) : \text{máx}\{x, y\} = 3\}$.
 g) $G = \{(x, y) : \text{mín}\{x, y\} = 3\}$.
 h) $H = \{(x, y) : |x - y| \leq 1\}$.
 i) $I = \{(x, y) : |x - y| > 2\}$.
 j) $J = \{(x, y) : |x - 3| + |y - 3| \leq 2\}$.

2.2. Función de probabilidad

Consideremos primero el caso discreto. Debe observarse que los conceptos que se exponen a continuación no hacen referencia a ningún experimento aleatorio en particular. Sea X una variable aleatoria discreta que toma los valores x_0, x_1, \dots con probabilidades

$$\begin{aligned} p_0 &= P(X = x_0) \\ p_1 &= P(X = x_1) \\ p_2 &= P(X = x_2) \\ &\vdots \quad \vdots \end{aligned}$$

Esta lista de valores numéricos y sus probabilidades puede ser finita o infinita, pero numerable. La función de probabilidad de X se define como aquella función que toma estas probabilidades como valores.

Definición 2.2 Sea X una variable aleatoria discreta con valores x_0, x_1, \dots . La función de probabilidad de X , denotada por $f(x) : \mathbb{R} \rightarrow \mathbb{R}$, se define como sigue

$$f(x) = \begin{cases} P(X = x) & \text{si } x = x_0, x_1, \dots \\ 0 & \text{en otro caso.} \end{cases} \quad (2.2)$$

En otras palabras, la función de probabilidad es simplemente aquella función que indica la probabilidad en los distintos valores que toma la variable aleatoria. Puede escribirse también mediante una tabla de la siguiente forma

x	x_0	x_1	x_2	\dots
$f(x)$	p_0	p_1	p_2	\dots

Dado lo anterior, la probabilidad de un evento para variables aleatorias discretas se reduce al cálculo de una suma, esto es, si A es cualquier subconjunto de \mathbb{R} , entonces

$$P(X \in A) = \sum_{x \in A} f(x),$$

en donde la suma se efectúa sobre aquellos valores x dentro del conjunto A tales que $f(x)$ es estrictamente positiva. De esta manera, la función de probabilidad $f(x)$ muestra la forma en la que la probabilidad se distribuye sobre el conjunto de puntos x_0, x_1, \dots

Por otro lado, recordemos que es importante poder distinguir entre X y x pues, conceptualmente, son cosas distintas. La primera es una función y la segunda es un número real. Es importante también observar que hemos denotado a una función de probabilidad con la letra f minúscula. A veces escribiremos $f_X(x)$ y el subíndice nos ayudará a especificar que tal función es la función de probabilidad de la variable X . Esta notación será particularmente útil cuando consideremos varias variables aleatorias a la vez, pero por ahora omitiremos esa especificación. A toda función de la forma (2.2) la llamaremos función de probabilidad. Observe que se cumplen las siguientes dos propiedades.

a) $f(x) \geq 0$ para toda $x \in \mathbb{R}$.

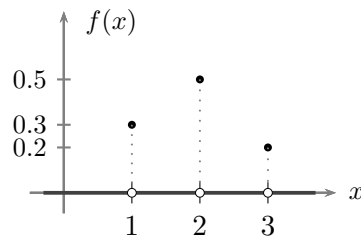
b) $\sum_x f(x) = 1$.

Recíprocamente, a toda función $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ que sea cero, excepto en ciertos puntos x_0, x_1, \dots en donde la función toma valores positivos, se le llamará función de probabilidad cuando se cumplan las dos propiedades anteriores y sin que haya de por medio una variable aleatoria. De esta manera, el estudio del modelo de probabilidad se traslada al estudio de funciones reales. Veamos un par de ejemplos.

Ejemplo 2.6 Considere la variable aleatoria discreta X que toma los valores 1, 2 y 3, con probabilidades 0.3, 0.5 y 0.2, respectivamente. Observe que no se especifica ni el experimento aleatorio ni el espacio muestral, únicamente los valores de la variable aleatoria y las probabilidades asociadas. Este

esquema genérico puede ser usado para representar cualquier experimento aleatorio con tres posibles resultados y con las probabilidades indicadas. La función de probabilidad de X es

$$f(x) = \begin{cases} 0.3 & \text{si } x = 1, \\ 0.5 & \text{si } x = 2, \\ 0.2 & \text{si } x = 3, \\ 0 & \text{en otro caso.} \end{cases}$$



(a)

x	1	2	3
$f(x)$	0.3	0.5	0.2

(b)

Figura 2.8

Esta función se muestra gráficamente en la Figura 2.8 (a) y alternativamente podemos también expresarla mediante la tabla de la Figura 2.8 (b). En esta última representación se entiende de manera implícita que $f(x)$ es cero para cualquier valor de x distinto de 1, 2 y 3. En particular, no debe ser difícil para el lector comprobar que las siguientes probabilidades son correctas.

$$\begin{aligned} P(X \geq 2) &= 0.7, \\ P(|X| = 1) &= 0.3, \\ P(X < 1) &= 0. \end{aligned}$$

•

Ejemplo 2.7 Encontraremos el valor de la constante c que hace que la siguiente función sea de probabilidad.

$$f(x) = \begin{cases} cx & \text{si } x = 0, 1, 2, 3, \\ 0 & \text{en otro caso.} \end{cases}$$

Los posibles valores de la variable aleatoria discreta (no especificada) son 0, 1, 2 y 3, con probabilidades $0, c, 2c$ y $3c$, respectivamente. Como la suma de estas probabilidades debe ser uno, obtenemos la ecuación $c + 2c + 3c = 1$. De aquí obtenemos $c = 1/6$. Este es el valor de c que hace que $f(x)$ sea no negativa y sume uno, es decir, una función de probabilidad. •

Ahora consideremos el caso continuo. A partir de este momento empezaremos a utilizar los conceptos de derivada e integral de una función.

Definición 2.3 Sea X una variable aleatoria continua. Decimos que la función integrable y no negativa $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ es la función de densidad de X si para cualquier intervalo $[a, b]$ de \mathbb{R} se cumple la igualdad

$$P(X \in [a, b]) = \int_a^b f(x) dx. \quad (2.3)$$

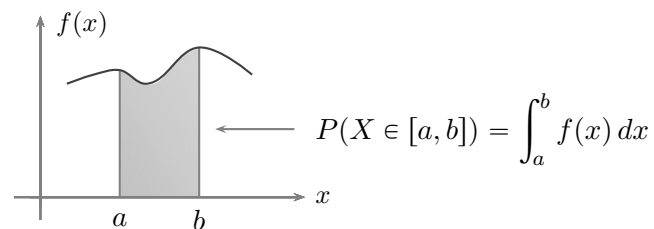


Figura 2.9

Es decir, la probabilidad de que la variable tome un valor dentro del intervalo $[a, b]$ se puede calcular o expresar como el área bajo la función $f(x)$ en

dicho intervalo. De esta forma, el cálculo de una probabilidad se reduce al cálculo de una integral. Véase la Figura 2.9, en donde se muestra esta forma de calcular probabilidades. No es difícil comprobar que toda función de densidad $f(x)$ de una variable aleatoria continua cumple las siguientes propiedades análogas al caso discreto.

a) $f(x) \geq 0$ para toda $x \in \mathbb{R}$.

b) $\int_{-\infty}^{\infty} f(x) dx = 1$.

Recíprocamente, toda función $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ que satisfaga estas dos propiedades, sin necesidad de tener una variable aleatoria de por medio, se llamará función de densidad. Hacemos la observación, nuevamente, de que la función de densidad muestra la forma en la que la probabilidad se distribuye sobre el conjunto de números reales. Veamos algunos ejemplos.

Ejemplo 2.8 La función dada por

$$f(x) = \begin{cases} 1/2 & \text{si } 1 < x < 3, \\ 0 & \text{en otro caso,} \end{cases}$$

es la función de densidad de una variable aleatoria continua que toma valores en el intervalo $(1, 3)$ y cuya gráfica aparece en la Figura 2.10. Es inmediato verificar que se trata de una función no negativa cuya integral vale uno. Observe que en el lenguaje que hemos utilizado se acepta, de manera implícita, la hipótesis de que existe una variable aleatoria con esta función de densidad.

•

Ejemplo 2.9 Encontraremos el valor de la constante c que hace que la siguiente función sea de densidad.

$$f(x) = \begin{cases} c|x| & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

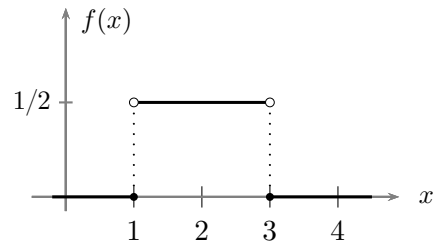


Figura 2.10

Una variable aleatoria continua con esta función de densidad toma valores en el intervalo $(-1, 1)$, pues en dicho intervalo la función de densidad es estrictamente positiva. Como esta función debe integrar uno, tenemos que

$$1 = \int_{-1}^1 c|x| dx = 2c \int_0^1 x dx = c.$$

Por lo tanto, cuando tomamos $c = 1$, esta función resulta ser una función de densidad, pues ahora cumple con ser no negativa e integrar uno. •

En el caso discreto hemos llamado función de probabilidad a $f(x)$, y en el caso continuo la hemos llamado función de densidad. Intentaremos mantener esta separación, aunque en realidad ambos nombres pueden emplearse en cualquier caso.

Concluimos esta sección observando que cualquier función de densidad $f(x)$ puede ser modificada en varios de sus puntos, e incluso puede tomar valores negativos en dichos puntos y, a pesar de ello, seguir cumpliendo la condición (2.3) que aparece en la definición de función de densidad para una v.a. continua. En tales casos las posibles modificaciones de $f(x)$ no cambian el valor de la integral y por lo tanto tampoco la probabilidad asociada. En este sentido es que debe entenderse su unicidad.

Ejercicios

166. Compruebe que las siguientes funciones son de probabilidad.

$$a) f(x) = \begin{cases} \frac{x-1}{2^x} & \text{si } x = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x) = \begin{cases} \frac{(1-p)p^{x-1}}{1-p^n} & \text{si } x = 1, \dots, n, \quad (0 < p < 1 \text{ constante}) \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f(x) = \begin{cases} \frac{1/2}{\sqrt{x}} & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$d) f(x) = \begin{cases} 3(1-|x|)^2/2 & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

167. En cada caso encuentre el valor de la constante θ que hace a la función $f(x)$ una función de probabilidad. Suponga que n es un entero positivo fijo.

$$a) f(x) = \begin{cases} \theta x & \text{si } x = 1, 2, \dots, n, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x) = \begin{cases} \theta x^2 & \text{si } x = 1, 2, \dots, n, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f(x) = \begin{cases} \theta 2^x & \text{si } x = 1, 2, \dots, n, \\ 0 & \text{en otro caso.} \end{cases}$$

$$d) f(x) = \begin{cases} \theta/2^x & \text{si } x = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

$$e) f(x) = \begin{cases} \theta/3^x & \text{si } x = 1, 3, 5, \dots \\ \theta/4^x & \text{si } x = 2, 4, 6, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

$$f) f(x) = \begin{cases} 3\theta^{2x} & \text{si } x = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

$$g) f(x) = \begin{cases} \theta x^2 & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$h) f(x) = \begin{cases} \theta x^2 + x & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$i) f(x) = \begin{cases} e^{-|x|} & \text{si } -\theta < x < \theta, \\ 0 & \text{en otro caso.} \end{cases}$$

$$j) f(x) = \begin{cases} \theta(1 - x^2) & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$k) f(x) = \begin{cases} x + 1/2 & \text{si } 0 < x < \theta, \\ 0 & \text{en otro caso.} \end{cases}$$

$$l) f(x) = \begin{cases} (1 - \theta)e^{-x} + e^{-2x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

$$m) f(x) = \begin{cases} \theta x^n(1 - x) & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$n) f(x) = \begin{cases} \theta x(1 - x)^n & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

168. Determine si la función $f(x)$ especificada abajo, puede ser una función de densidad. En caso afirmativo, encuentre el valor de la constante c .

$$a) f(x) = \begin{cases} c(2x - x^3) & \text{si } 0 < x < 2, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x) = \begin{cases} c(2x - x^2) & \text{si } 0 < x < 2, \\ 0 & \text{en otro caso.} \end{cases}$$

169. Sea X una variable aleatoria discreta con función de probabilidad como se indica en la siguiente tabla.

x	-2	-1	0	1	2
$f(x)$	1/8	1/8	1/2	1/8	1/8

Grafique la función $f(x)$ y calcule las siguientes probabilidades:

- | | |
|-------------------------|-----------------------|
| a) $P(X \leq 1)$. | d) $P(X^2 \geq 1)$. |
| b) $P(X \leq 1)$. | e) $P(X - 1 < 2)$. |
| c) $P(-1 < X \leq 2)$. | f) $P(X - X^2 < 0)$. |

170. Sea X una variable aleatoria con función de densidad

$$f(x) = \begin{cases} cx^2 & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Encuentre el valor de la constante c y calcule:

- | | |
|----------------------|--------------------------------------|
| a) $P(X \leq 1/2)$. | c) $P(1/4 < X < 3/4)$. |
| b) $P(X \geq 1/4)$. | d) a tal que $P(X \leq a) = 1/2$. |

171. Sea X una variable aleatoria continua con función de densidad

$$f(x) = \begin{cases} 3x^2/2 & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Grafique la función $f(x)$ y calcule las siguientes probabilidades:

- | | |
|----------------------|---------------------------|
| a) $P(X \leq 1/3)$. | e) $P(-1/4 < X \leq 2)$. |
| b) $P(X \geq 0)$. | f) $P(2X + 1 < 2)$. |
| c) $P(X > 1)$. | g) $P(X^2 < 1/9)$. |
| d) $P(X < 1/2)$. | h) $P(X^2 - X < 2)$. |

172. **Función de probabilidad simétrica.** Sea $f(x)$ una función de probabilidad simétrica respecto de a , es decir, $f(a+x) = f(a-x)$ para cualquier número real x . Demuestre que la siguiente función es de densidad.

$$g(x) = \begin{cases} 2f(a+x) & \text{si } x \geq 0, \\ 0 & \text{en otro caso.} \end{cases}$$

173. Sean $f(x)$ y $g(x)$ dos funciones de probabilidad. Demuestre o proporcione un contraejemplo, en cada caso, para la afirmación que establece que la siguiente función es de probabilidad.

- a) $f(x) + g(x)$.
- b) $f(x) \cdot g(x)$.
- c) $\theta f(x) + (1-\theta)g(x)$ con $0 \leq \theta \leq 1$ constante.
- d) $f(x+c)$ con c constante.
- e) $f(c-x)$ con c constante.
- f) $f(x^3)$.
- g) $f(cx)$ con $c \neq 0$ constante.
- h) $f(e^x)$.

174. Sea X una variable aleatoria discreta con valores $0, 1, 2, \dots$ y con función de probabilidad $f_X(x)$. Encuentre la función de probabilidad de la siguiente variable aleatoria, en términos de $f_X(x)$.

- a) $Y = 2X$.
- b) $Y = X + n, \quad n \in \mathbb{N}$.
- c) $Y = \cos(\pi X/2)$.
- d) $Y = X$ (módulo 2).

175. Un dado equilibrado se lanza dos veces consecutivas. Sea X la diferencia entre el resultado del primer y el segundo lanzamiento. Encuentre la función de probabilidad de X .

176. Sea X una variable aleatoria continua con función de densidad

$$f(x) = \begin{cases} \theta x & \text{si } 0 < x \leq 1/2, \\ \theta(1-x) & \text{si } 1/2 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

- a) Encuentre el valor de la constante θ que hace a $f(x)$ una función de densidad.
- b) Encuentre los valores de a y b tales que $P(a < X < b) = 7/16$ y $b - a$ es mínimo.
177. En una caja se encuentran mezcladas 19 baterías en buen estado y 6 baterías defectuosas. Se extraen las baterías al azar, una por una, hasta encontrar una batería en buen estado. Sea X la variable aleatoria que indica el número de extracciones efectuadas hasta obtener el evento de interés. Encuentre la función de probabilidad de X .
178. Los jugadores A y B llevan a cabo una sucesión de apuestas justas en donde en cada apuesta se gana o se pierde una moneda. Suponga que inicialmente el jugador A tiene dos monedas y el jugador B sólo una moneda. El juego termina hasta que uno de los jugadores gana las tres monedas.
- a) Calcule la probabilidad de ganar de cada uno de los jugadores.
- b) Defina la variable aleatoria X como el número de apuestas efectuadas hasta el final del juego. Calcule la función de probabilidad de X .
179. Se lanza un dado equilibrado hasta que aparece un "6". Encuentre la función de probabilidad del número de lanzamientos necesarios hasta obtener tal resultado.
180. Un experimento consiste en lanzar un dado equilibrado hasta obtener un mismo resultado dos veces, no necesariamente de manera consecutiva. Encuentre la función de probabilidad del número de lanzamientos en este experimento.
181. Una moneda equilibrada se lanza repetidamente hasta obtener un mismo resultado por tercera ocasión, no necesariamente de manera consecutiva. Encuentre la función de probabilidad de la variable aleatoria que registra el número de lanzamientos necesarios hasta obtener el resultado mencionado.
182. Se colocan al azar, una por una, 10 bolas en 4 cajas de tal forma que cada bola tiene la misma probabilidad de quedar en cualquiera de las

4 cajas. Encuentre la función de probabilidad del número de bolas que caen en la primera caja.

183. **Función de probabilidad condicional.** Sea X una variable aleatoria con función de probabilidad $f(x)$ y sea $A \in \mathcal{B}(\mathbb{R})$ un conjunto de Borel tal que $p := P(X \in A) > 0$. La función de probabilidad condicional de X dado el evento $(X \in A)$ se denota y define como sigue:

$$\begin{aligned} f(x|A) &:= \frac{1}{p} f(x) \cdot 1_A(x) \\ &= \begin{cases} f(x)/p & \text{si } x \in A, \\ 0 & \text{en otro caso.} \end{cases} \end{aligned}$$

Demuestre que $f(x|A)$ es, efectivamente, una función de probabilidad y calcule esta función en cada uno de las siguientes situaciones. En cada caso grafique tanto $f(x)$ como $f(x|A)$ y compare ambas funciones.

- a) $A = \{2, 3, 4, 5\}$ y $f(x) = \begin{cases} 1/6 & \text{si } x = 1, \dots, 6, \\ 0 & \text{en otro caso.} \end{cases}$
- b) $A = (0, \infty)$ y $f(x) = \begin{cases} 1/2 & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$
- c) $A = (1, \infty)$ y $f(x) = \begin{cases} e^{-x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$
- d) $A = (0, 1)$ y $f(x) = \begin{cases} 1/2 - |x|/4 & \text{si } -2 < x < 2, \\ 0 & \text{en otro caso.} \end{cases}$

2.3. Función de distribución

Otra función importante que puede asociarse a una variable aleatoria es la siguiente.

Definición 2.4 Sea X una variable aleatoria cualquiera. La función de distribución de X , denotada por $F(x) : \mathbb{R} \rightarrow \mathbb{R}$, se define como la probabilidad

$$F(x) = P(X \leq x). \quad (2.4)$$

Esto es, la función de distribución evaluada en un número x cualquiera es la probabilidad de que la variable aleatoria tome un valor menor o igual a x , o en otras palabras, que tome un valor en el intervalo $(-\infty, x]$. Para especificar que se trata de la función de distribución de la variable aleatoria X se usa la notación $F_X(x)$, pero por simplicidad omitiremos el subíndice cuando no haya necesidad de tal especificación. Así, siendo $F(x)$ una probabilidad, sus valores están siempre entre cero y uno. En el caso discreto, suponiendo que $f(x)$ es la función de probabilidad de X , la función de distribución (2.4) se calcula como sigue

$$F(x) = \sum_{u \leq x} f(u), \quad (2.5)$$

y corresponde a sumar todos los valores positivos de la función de probabilidad evaluada en aquellos números menores o iguales a x . En el caso continuo, si $f(x)$ es la función de densidad de X , por (2.4), se tiene que

$$F(x) = \int_{-\infty}^x f(u) du. \quad (2.6)$$

La función de distribución resulta ser muy importante desde el punto de vista matemático, pues siempre puede definirse dicha función para cualquier variable aleatoria y a través de ella quedan representadas todas las propiedades de la variable aleatoria. Por razones evidentes se le conoce también con el nombre de función de acumulación de probabilidad o función de probabilidad acumulada. De manera análoga a la función de densidad, la función de distribución permite llevar el estudio de un experimento aleatorio y su modelo de probabilidad al estudio de las funciones reales. Mostraremos a continuación algunos ejemplos del cálculo de la función de distribución y después estudiaremos sus propiedades.

Ejemplo 2.10 (Caso discreto) Considere una variable aleatoria discreta X con función de probabilidad

$$f(x) = \begin{cases} 1/3 & \text{si } x = 1, 2, 3, \\ 0 & \text{en otro caso.} \end{cases}$$

La gráfica de esta función se muestra en la Figura 2.11 (a). Considerando los distintos valores para x , puede encontrarse la función de distribución $F(x)$ de la siguiente forma,

$$F(x) = P(X \leq x) = \sum_{u \leq x} f(u) = \begin{cases} 0 & \text{si } x < 1, \\ 1/3 & \text{si } 1 \leq x < 2, \\ 2/3 & \text{si } 2 \leq x < 3, \\ 1 & \text{si } x \geq 3, \end{cases}$$

y cuya gráfica aparece en la Figura 2.11 (b). Para cada valor de x , $F(x)$ se calcula como la suma de los valores positivos de $f(u)$ para valores de u a la izquierda de x .

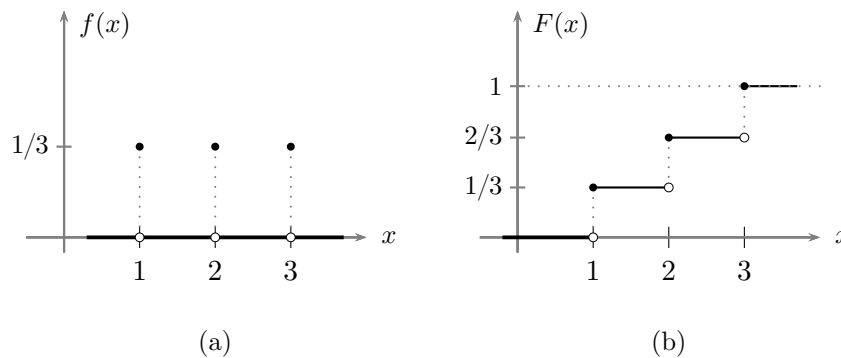


Figura 2.11

En el ejemplo anterior se ha mostrado el comportamiento típico de la función de distribución de una v.a. discreta, es decir, es una función no decreciente, constante por pedazos, y si la función tiene una discontinuidad en x , entonces esta discontinuidad es un salto hacia arriba y el tamaño del salto es

la probabilidad de que la variable aleatoria tome ese valor. Más adelante formalizaremos estas observaciones.

Ejemplo 2.11 (Caso continuo) Considere ahora la variable aleatoria continua X con función de densidad

$$f(x) = \begin{cases} |x| & \text{si } -1 \leq x \leq 1, \\ 0 & \text{en otro caso.} \end{cases}$$

La gráfica de esta función se muestra en la Figura 2.12 (a).

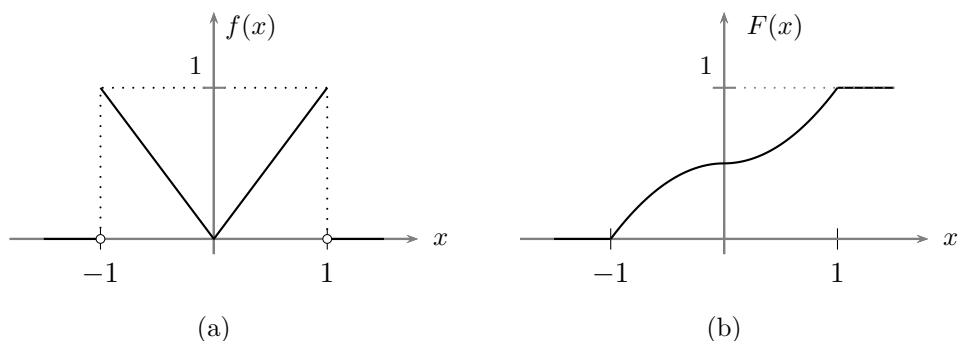


Figura 2.12

Integrando esta función desde menos infinito hasta x , para distintos valores de x , se encuentra que la función de distribución es

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(u) du = \begin{cases} 0 & \text{si } x < -1, \\ (1 - x^2)/2 & \text{si } -1 \leq x < 0, \\ (1 + x^2)/2 & \text{si } 0 \leq x < 1, \\ 1 & \text{si } x \geq 1, \end{cases}$$

y cuya gráfica se muestra en la Figura 2.12 (b). •

Ahora podemos dar una definición más precisa acerca de cuándo una v.a. es continua.

Definición 2.5 Se dice que una variable aleatoria X es continua si su función de distribución $F(x)$ es una función continua.

En los ejemplos anteriores se ha mostrado la forma de obtener $F(x)$ a partir de $f(x)$, tanto en el caso discreto como en el continuo, usando las fórmulas (2.5) y (2.6). Ahora explicaremos el proceso contrario, es decir, explicaremos la manera de obtener $f(x)$ a partir de $F(x)$. En el caso continuo tenemos que para toda x en \mathbb{R} ,

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(u) du,$$

de modo que, por el teorema fundamental del cálculo, y cuando $F(x)$ es diferenciable,

$$f(x) = F'(x). \quad (2.7)$$

De este modo, podemos encontrar $f(x)$ a partir de $F(x)$. En el caso discreto, la función de probabilidad se obtiene de la función de distribución del siguiente modo, véase el enunciado del Ejercicio 188.

$$f(x) = F(x) - F(x-), \quad (2.8)$$

en donde $F(x-)$ denota el límite por la izquierda de la función F en el punto x . Así, $f(x)$ es, efectivamente, el tamaño del salto de la función de distribución en el punto x . Análogamente, la expresión $F(x+)$ significa el límite por la derecha de la función F en el punto x . En símbolos,

$$\begin{aligned} F(x-) &= \lim_{h \searrow 0} F(x-h), \\ F(x+) &= \lim_{h \searrow 0} F(x+h). \end{aligned}$$

Haciendo referencia a la ecuación (2.8), cuando la función de distribución es continua, los valores $F(x)$ y $F(x-)$ coinciden y la función de probabilidad toma el valor cero, es decir,

$$P(X = x) = 0.$$

Cuando $F(x)$ presenta una discontinuidad, la función de probabilidad toma como valor la magnitud de dicha discontinuidad. En los siguientes ejemplos se muestra la aplicación de las fórmulas (2.7) y (2.8).

Ejemplo 2.12 Considere la función de distribución

$$F(x) = \begin{cases} 0 & \text{si } x < -1, \\ (1 + x^3)/2 & \text{si } -1 \leq x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

Se deja al lector la graficación cuidadosa de esta función. Observe que se trata de una función continua y diferenciable, excepto en $x = -1, 1$. Derivando entonces en cada una de las tres regiones de definición, se encuentra que

$$f(x) = \begin{cases} 3x^2/2 & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Esta es la función de densidad asociada a la función de distribución dada. Observe que no se especifica ninguna variable aleatoria o experimento aleatorio. •

Ejemplo 2.13 Considere la función de distribución

$$F(x) = \begin{cases} 0 & \text{si } x < -1, \\ 1/3 & \text{si } -1 \leq x < 0, \\ 2/3 & \text{si } 0 \leq x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

Al graficar esta función uno puede darse cuenta que se trata de una función constante por pedazos. Los puntos donde esta función tiene incrementos, y los tamaños de estos incrementos, determinan la correspondiente función de probabilidad, la cual está dada por

$$f(x) = \begin{cases} 1/3 & \text{si } x = -1, 0, 1, \\ 0 & \text{en otro caso.} \end{cases}$$

•

Demostremos a continuación algunas propiedades generales válidas para toda función de distribución. Haremos uso de la propiedad de continuidad de las medidas de probabilidad estudiada en la sección 1.17 del capítulo anterior.

Proposición 2.1 Toda función de distribución $F(x)$ satisface las siguientes propiedades:

- a) $\lim_{x \rightarrow \infty} F(x) = 1$.
- b) $\lim_{x \rightarrow -\infty} F(x) = 0$.
- c) Si $x_1 \leq x_2$, entonces $F(x_1) \leq F(x_2)$.
- d) $F(x) = F(x+)$.

Demostración.

- a) Sea $x_1 \leq x_2 \leq \dots$ cualquier sucesión monótona no decreciente de números reales divergente a infinito. Para cada natural n defina el evento $A_n = (X \leq x_n)$, cuya probabilidad es $P(A_n) = F(x_n)$. Entonces $A_1 \subseteq A_2 \subseteq \dots$ y puede comprobarse que el límite de esta sucesión monótona de eventos es Ω . Por la propiedad de continuidad de la probabilidad para sucesiones monótonas,

$$1 = P(\Omega) = \lim_{n \rightarrow \infty} P(A_n) = \lim_{n \rightarrow \infty} F(x_n).$$

- b) Considere ahora cualquier sucesión monótona no creciente de números reales $\dots \leq x_2 \leq x_1$ divergente a menos infinito. Defina nuevamente los eventos $A_n = (X \leq x_n)$ y observe que $P(A_n) = F(x_n)$. Entonces $A_1 \supseteq A_2 \supseteq \dots$ y puede comprobarse que el límite de esta sucesión monótona de eventos es el conjunto vacío. Por la propiedad de continuidad de la probabilidad para sucesiones monótonas,

$$0 = P(\emptyset) = \lim_{n \rightarrow \infty} P(A_n) = \lim_{n \rightarrow \infty} F(x_n).$$

- c) Si $x_1 \leq x_2$ entonces $(X \leq x_1) \subseteq (X \leq x_2)$. Por lo tanto, $P(X \leq x_1) \subseteq P(X \leq x_2)$, es decir, $F(x_1) \leq F(x_2)$.
- d) Sea $0 \leq \dots \leq x_2 \leq x_1$ cualquier sucesión monótona no creciente de números reales no negativos convergente a cero. Defina los eventos $A_n = (x < X \leq x + x_n)$ y observe que $P(A_n) = F(x + x_n) - F(x)$. Entonces $A_1 \supseteq A_2 \supseteq \dots$ y puede comprobarse que el límite de esta sucesión monótona de eventos es el conjunto vacío. Por la propiedad de continuidad de la probabilidad para sucesiones monótonas,

$$0 = P(\emptyset) = \lim_{n \rightarrow \infty} P(A_n) = \lim_{n \rightarrow \infty} F(x + x_n) - F(x).$$

Recíprocamente, definiremos a continuación una función de distribución como aquella que cumpla las cuatro propiedades anteriores.

Definición 2.6 A toda función $F(x) : \mathbb{R} \rightarrow \mathbb{R}$ que cumpla las cuatro propiedades de la Proposición 2.1 se le llama función de distribución, sin tener necesariamente una variable aleatoria que la defina.

La propiedad (c), recién demostrada, significa que $F(x)$ es una función monótona no decreciente, mientras que la propiedad (d) establece que $F(x)$ es una función continua por la derecha. Se puede demostrar que si $a \leq b$, entonces se cumplen las siguientes identidades:

Probabilidades de algunos eventos en términos de $F(x)$

$$\begin{aligned} P(X < a) &= F(a-). \\ P(a < X \leq b) &= F(b) - F(a). \\ P(a \leq X \leq b) &= F(b) - F(a-). \\ P(a < X < b) &= F(b-) - F(a). \\ P(a \leq X < b) &= F(b-) - F(a-). \end{aligned}$$

185. Considere el experimento aleatorio de lanzar un dado equilibrado. Defina una variable aleatoria de su preferencia sobre el correspondiente espacio muestral, que no sea constante ni la identidad. Encuentre la función de probabilidad y la función de distribución de esta variable aleatoria. Grafique ambas funciones.
186. **Variable aleatoria constante.** Sea X la variable aleatoria constante c . Encuentre y grafique con precisión tanto la función de probabilidad como la función de distribución de esta variable aleatoria.
187. Sea X una variable aleatoria con función de distribución $F(x)$. Demuestre que si $a \leq b$, entonces:

- a) $P(X < a) = F(a-)$.
- b) $P(a < X \leq b) = F(b) - F(a)$.
- c) $P(a \leq X \leq b) = F(b) - F(a-)$.
- d) $P(a < X < b) = F(b-) - F(a)$.
- e) $P(a \leq X < b) = F(b-) - F(a-)$.

188. Sea X una variable aleatoria con función de distribución $F(x)$. Demuestre que para cualquier número real x ,

$$P(X = x) = F(x) - F(x-).$$

Esta cantidad es cero cuando X es continua. En el caso cuando X es discreta, esta cantidad es estrictamente positiva cuando X puede tomar el valor x y representa el salto de la función de distribución en dicho punto.

189. Grafique y compruebe que las siguientes funciones son de distribución:

$$a) F(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1/2 & \text{si } 0 \leq x < 1/2, \\ x & \text{si } 1/2 \leq x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

$$b) F(x) = \begin{cases} 0 & \text{si } x < 0, \\ x^2 & \text{si } 0 \leq x < 1/2, \\ 1 - 3(1-x)/2 & \text{si } 1/2 \leq x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

190. Suponga que la variable aleatoria discreta X tiene la siguiente función de distribución

$$F(x) = \begin{cases} 0 & \text{si } x < -1, \\ 1/4 & \text{si } -1 \leq x < 1, \\ 1/2 & \text{si } 1 \leq x < 3, \\ 3/4 & \text{si } 3 \leq x < 5, \\ 1 & \text{si } x \geq 5. \end{cases}$$

Grafique $F(x)$, obtenga y grafique la correspondiente función de probabilidad $f(x)$, calcule además las siguientes probabilidades:

- | | |
|--------------------|------------------------|
| a) $P(X \leq 3)$. | d) $P(X \geq 1)$. |
| b) $P(X = 3)$. | e) $P(-1/2 < X < 4)$. |
| c) $P(X < 3)$. | f) $P(X = 5)$. |

191. Muestre que las siguientes funciones son de probabilidad y encuentre la correspondiente función de distribución. Grafique ambas funciones.

$$a) f(x) = \begin{cases} (1/2)^x & \text{si } x = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x) = \begin{cases} 2x & \text{si } 0 \leq x \leq 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f(x) = \begin{cases} 2(1-x) & \text{si } 0 \leq x \leq 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$d) f(x) = \begin{cases} 4e^{-4x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

192. Compruebe que cada una de las funciones que aparecen en la Figura 2.13 de la página 141 es una función de probabilidad. En cada caso encuentre la correspondiente función de distribución.
193. Grafique cada una de las siguientes funciones y compruebe que son funciones de distribución. Determine, en cada caso, si se trata de la función de distribución de una variable aleatoria discreta o continua. Encuentre además la correspondiente función de probabilidad o de densidad.

$$a) F(x) = \begin{cases} 0 & \text{si } x < 0, \\ x & \text{si } 0 \leq x \leq 1, \\ 1 & \text{si } x > 1. \end{cases}$$

$$b) F(x) = \begin{cases} 1 - e^{-x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) F(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - \cos x & \text{si } 0 \leq x \leq \pi/2, \\ 1 & \text{si } x > \pi/2. \end{cases}$$

$$d) F(x) = \begin{cases} 0 & \text{si } x < -a, \\ \frac{x+a}{2a} & \text{si } -a \leq x \leq a, \\ 1 & \text{si } x > a. \end{cases}$$

$$e) F(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1/5 & \text{si } 0 \leq x < 1, \\ 3/5 & \text{si } 1 \leq x < 2, \\ 1 & \text{si } x \geq 2. \end{cases}$$

$$f) F(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ x^2 & \text{si } 0 \leq x < 1/2, \\ 1 - 3(1-x)/2 & \text{si } 1/2 \leq x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

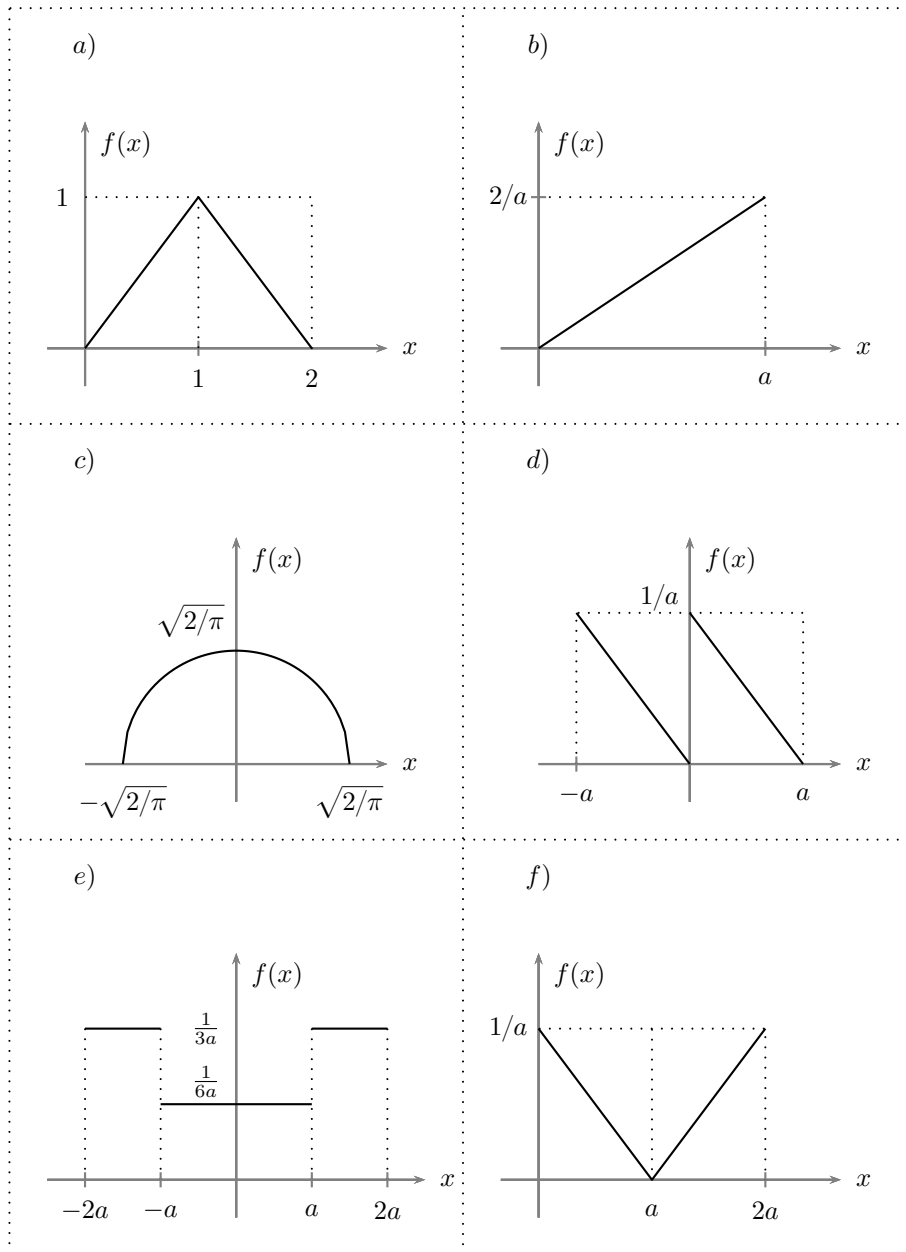


Figura 2.13

$$g) F(x) = \begin{cases} 0 & \text{si } x < -2, \\ (x+2)/2 & \text{si } -2 \leq x < -1, \\ 1/2 & \text{si } -1 \leq x < 1, \\ x/2 & \text{si } 1 \leq x < 2, \\ 1 & \text{si } x \geq 2. \end{cases}$$

194. Sea X una variable aleatoria continua con función de densidad

$$f(x) = \begin{cases} 1/x^2 & \text{si } x \geq 1, \\ 0 & \text{en otro caso.} \end{cases}$$

- Grafique $f(x)$ y compruebe que es, efectivamente, una función de densidad.
- Encuentre y grafique la función de distribución de X .
- Encuentre y grafique la función de distribución de la variable $Y = e^{-X}$.
- A partir del inciso anterior, encuentre la función de densidad de la variable Y .

195. Sea X una variable aleatoria continua con función de densidad como aparece abajo. Grafique $f(x)$ y compruebe que efectivamente es una función de densidad. Además encuentre y grafique la función de distribución correspondiente.

$$f(x) = \begin{cases} 3x^2/2 & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Calcule:

- $P(|X| < 1/2)$.
- $P(X < 0)$.
- $P(|X| < 1/n)$, $n \in \mathbb{N}$.
- $P(|X| > 1/2)$.

196. Sea X una variable aleatoria continua con función de densidad como aparece abajo. Grafique $f(x)$ y compruebe que efectivamente es

una función de densidad. Además encuentre y grafique la función de distribución correspondiente.

$$f(x) = \begin{cases} 1/4 & \text{si } |x| \leq 1, \\ 1/(4x^2) & \text{si } |x| > 1. \end{cases}$$

Calcule:

- a) $P(|X| < 3/2)$. c) $P(1/2 < X < 3/2)$.
 b) $P(X > 0)$. d) $P(|X| \leq 1)$.

197. Sea X una variable aleatoria continua con función de densidad como aparece abajo. Grafique $f(x)$ y compruebe que efectivamente es una función de densidad. Además encuentre y grafique la función de distribución correspondiente.

$$f(x) = \begin{cases} 3(1 - |x|)^2/2 & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Calcule:

- a) $P(|X| < 1/2)$. c) $P(1/2 < X < 3/2)$.
 b) $P(X < 0)$. d) $P(|X| > 1/2)$.

198. Sea X una variable aleatoria con función de distribución

$$F(x) = \begin{cases} 0 & \text{si } x < 0, \\ x/4 & \text{si } 0 \leq x < 1, \\ 1/2 + (x - 1)/4 & \text{si } 1 \leq x < 2, \\ 4/5 & \text{si } 2 \leq x < 3, \\ 1 & \text{si } x \geq 3. \end{cases}$$

- a) Encuentre $P(X = x)$ para $x = 0, 1/2, 1, 2, 3, 4$.
 b) Calcule $P(1/2 < X < 5/2)$.

199. **Distribución mixta.** Sean $F(x)$ y $G(x)$ dos funciones de distribución. Demuestre que para cualquier constante $\lambda \in [0, 1]$, la función

$$x \mapsto \lambda F(x) + (1 - \lambda) G(x)$$

es una función de distribución. Si $F(x)$ y $G(x)$ son ambas discretas entonces la función resultante es también discreta. Si $F(x)$ y $G(x)$ son ambas continuas entonces la función resultante es continua. Si alguna de $F(x)$ y $G(x)$ es discreta y la otra es continua, la función resultante no es ni discreta ni continua, se dice que es una función de distribución mixta.

200. **Variables aleatorias mixtas.** Sea X una variable aleatoria continua con función de distribución

$$F(x) = \begin{cases} 1 - e^{-x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

Sea $c > 0$ una constante. Encuentre y grafique la función de distribución de las siguientes variables aleatorias:

- a) $U = \text{mín} \{X, c\}$.
 b) $V = \text{máx} \{X, c\}$.

Estas variables no son discretas ni continuas, son ejemplos de variables aleatorias mixtas. Para estas distribuciones puede comprobarse que no existe la función de densidad, es decir, no existe una función $f(u)$, en el sentido usual, tal que para todo número real x ,

$$F(x) = \int_{-\infty}^x f(u) du.$$

201. Sea X una variable aleatoria continua con función de densidad $f_X(x)$ y sea $c > 0$ una constante. Demuestre que la función de densidad de cX está dada por

$$f_{cX}(x) = \frac{1}{c} f_X(x/c).$$

202. Sea X una variable aleatoria continua con función de densidad $f_X(x)$. Demuestre que la función de densidad de X^2 está dada por

$$f_{X^2}(x) = \begin{cases} \frac{1}{2\sqrt{x}} [f_X(\sqrt{x}) + f_X(-\sqrt{x})] & \text{si } x > 0, \\ 0 & \text{si } x \leq 0. \end{cases}$$

203. Demuestre o proporcione un contraejemplo.

- Si $X \leq Y$ entonces $F_X(x) \leq F_Y(x)$ para todo $x \in \mathbb{R}$.
- Si $F_X(x) \leq F_Y(x)$ para todo $x \in \mathbb{R}$ entonces $X \leq Y$.
- Si $X = Y$ entonces $F_X(x) = F_Y(x)$ para todo $x \in \mathbb{R}$.
- Si $F_X(x) = F_Y(x)$ para todo $x \in \mathbb{R}$ entonces $X = Y$.

204. Proporcione un contraejemplo para la siguiente afirmación: si X y Y son dos variables aleatorias con la misma distribución, entonces

$$XY = X^2.$$

205. Un dado equilibrado se lanza dos veces consecutivas. Defina la variable aleatoria X como la suma de los dos resultados. Encuentre y grafique la función de probabilidad y la función de distribución de X .
206. Una moneda equilibrada y marcada con “Cara” y “Cruz” se lanza repetidas veces hasta obtener el resultado “Cruz”. Defina la variable aleatoria X como el número de lanzamientos necesarios hasta obtener el resultado de interés. Encuentre la función de distribución de X .

2.4. Teorema de cambio de variable

Sea X una variable aleatoria con distribución conocida. Suponga que se modifica X a través de una función φ de tal forma que la composición $\varphi(X)$ es una nueva variable aleatoria, por ejemplo,

$$\varphi(X) = aX + b, \quad a, b \text{ constantes,}$$

$$\varphi(X) = X^2,$$

$$\varphi(X) = e^X.$$

El problema natural que surge es el de encontrar la distribución de esta nueva variable aleatoria. En esta sección estudiaremos un resultado que permite dar una respuesta a este problema bajo ciertas condiciones.

Proposición 2.2 Sea X una variable aleatoria continua con función de densidad $f_X(x)$. Sea $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ una función continua, estrictamente creciente o decreciente y con inversa φ^{-1} diferenciable. Entonces la función de densidad de $Y = \varphi(X)$ está dada por

$$f_Y(y) = \begin{cases} f_X(\varphi^{-1}(y)) \left| \frac{d}{dy} \varphi^{-1}(y) \right| & \text{si } y \in \text{Rango}(\varphi(X)), \\ 0 & \text{en otro caso.} \end{cases} \quad (2.9)$$

Demostración. Supongamos que φ es estrictamente creciente. Calcularemos primero la función de distribución de Y en términos de la función de distribución de X . Para cualquier valor y dentro del rango de la función $\varphi(X)$,

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(\varphi(X) \leq y) \\ &= P(X \leq \varphi^{-1}(y)) \\ &= F_X(\varphi^{-1}(y)). \end{aligned}$$

Derivando respecto de y , por la regla de la cadena, tenemos que

$$\begin{aligned} f_Y(y) &= f_X(\varphi^{-1}(y)) \frac{d}{dy} \varphi^{-1}(y) \\ &= f_X(\varphi^{-1}(y)) \left| \frac{d}{dy} \varphi^{-1}(y) \right|. \end{aligned}$$

La última identidad se obtiene al observar que, como φ es estrictamente creciente, su inversa φ^{-1} también lo es y su derivada es positiva. En el caso cuando φ es estrictamente decreciente, se puede demostrar de manera similar al análisis anterior que

$$f_Y(y) = 1 - F_X(\varphi^{-1}(y)).$$

Derivando nuevamente respecto de y ,

$$\begin{aligned} f_Y(y) &= -f_X(\varphi^{-1}(y)) \frac{d}{dy} \varphi^{-1}(y) \\ &= f_X(\varphi^{-1}(y)) \left[-\frac{d}{dy} \varphi^{-1}(y) \right] \\ &= f_X(\varphi^{-1}(y)) \left| \frac{d}{dy} \varphi^{-1}(y) \right|. \end{aligned}$$

En este caso, la inversa φ^{-1} es decreciente y su derivada es negativa. •

Por simplicidad en el enunciado del resultado anterior, hemos pedido que la función φ sea estrictamente monótona y definida en el conjunto de números reales, sin embargo únicamente se necesita que esté definida en el rango de la función X y que presente el comportamiento monótono en dicho subconjunto. Ilustraremos esta situación en los siguientes ejemplos.

Ejemplo 2.15 Sea X una variable aleatoria continua con función de densidad

$$f_X(x) = \begin{cases} 1 & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Así, la variable X toma valores únicamente en el intervalo $(0, 1)$. Consideremos la función $\varphi(x) = x^2$, la cual es estrictamente creciente en $(0, 1)$, cuya inversa en dicho intervalo es $\varphi^{-1}(y) = \sqrt{y}$ y tiene derivada

$$\frac{d}{dy} \varphi^{-1}(y) = \frac{1}{2\sqrt{y}}.$$

Por lo tanto, la variable $Y = \varphi(X)$ toma valores en $(0, 1)$ y por la fórmula (2.9) tiene función de densidad

$$f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}} & \text{si } 0 < y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

No es difícil verificar que ésta es, efectivamente, una función de densidad. •

Ejemplo 2.16 Sea X una variable aleatoria continua con función de densidad

$$f_X(x) = \begin{cases} e^{-x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

La variable X toma valores en el intervalo $(0, \infty)$. Sea la función $\varphi(x) = 1/x$, la cual es estrictamente decreciente en $(0, \infty)$. Su inversa en dicho intervalo es $\varphi^{-1}(y) = 1/y$ y tiene derivada

$$\frac{d}{dy}\varphi^{-1}(y) = -\frac{1}{y^2}.$$

Por lo tanto, la variable $Y = \varphi(X)$ toma valores en $(0, \infty)$, y por la fórmula (2.9), tiene función de densidad

$$f_Y(y) = \begin{cases} e^{-1/y} \frac{1}{y^2} & \text{si } y > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

Haciendo el cambio de variable $u = 1/y$ en la integral puede verificarse con facilidad que $f_Y(y)$ es, efectivamente, una función de densidad. •

Ejercicios

207. Encuentre una fórmula para la función de densidad de la variable Y en términos de la función de densidad de la variable X , suponiendo que esta última es continua con función de densidad conocida $f_X(x)$.

- a) $Y = aX + b$ con $a \neq 0$, b constantes.
- b) $Y = -X$.
- c) $Y = X^3$.
- d) $Y = e^X$.

208. Sea X una variable aleatoria continua con función de densidad

$$f_X(x) = \begin{cases} x^2 + x + \theta & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Encuentre el valor de la constante θ que hace a $f_X(x)$ una función de densidad. Encuentre además la función de densidad de la variable $Y = \varphi(X)$ cuando

a) $\varphi(x) = ax + b$ con $a > 0$, b constantes.

b) $\varphi(x) = \sqrt{x}$.

c) $\varphi(x) = 2x - 1$.

209. Sea X una variable aleatoria continua con función de densidad

$$f_X(x) = \begin{cases} \theta(1 - |2x - 1|) & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Encuentre el valor de la constante θ que hace a $f_X(x)$ una función de densidad. Encuentre además la función de densidad de la variable $Y = \varphi(X)$ cuando

a) $\varphi(x) = (x - 1)^2$.

b) $\varphi(x) = 1/x$.

c) $\varphi(x) = \ln x$.

2.5. Independencia de variables aleatorias

El concepto de independencia de eventos que hemos estudiado antes puede extenderse al caso de variables aleatorias de una forma natural como la que se presenta en la definición que aparece a continuación. En esta sección estudiaremos preliminarmente este concepto, el cual se revisará con mucho más detalle en la sección 4.6. Consideraremos que tenemos dos variables aleatorias, X y Y , definidas sobre un mismo espacio de probabilidad.

Definición 2.7 Se dice que las variables aleatorias X y Y son independientes si los eventos $(X \leq x)$ y $(Y \leq y)$ son independientes para cualesquiera valores reales de x y y , es decir, si se cumple la igualdad

$$P[(X \leq x) \cap (Y \leq y)] = P(X \leq x)P(Y \leq y). \quad (2.10)$$

El lado izquierdo de la identidad anterior también puede escribirse como $P(X \leq x, Y \leq y)$ o bien $F_{X,Y}(x, y)$, y se le llama la función de distribución conjunta de X y Y evaluada en el punto (x, y) . Así, observe que la identidad (2.10) adquiere la expresión

$$F_{X,Y}(x, y) = F_X(x) F_Y(y), \quad x, y \in \mathbb{R}. \quad (2.11)$$

De esta forma, para poder determinar si dos variables aleatorias son independientes, es necesario conocer tanto las probabilidades conjuntas $P(X \leq x, Y \leq y)$ como las probabilidades individuales $P(X \leq x)$ y $P(Y \leq y)$, y verificar la identidad (2.11) para cada par de números reales x y y . Por lo tanto, basta que exista una pareja (x, y) para la que no se cumpla la igualdad (2.11) para poder concluir que X y Y no son independientes. En el capítulo de vectores aleatorios explicaremos la forma de obtener las distribuciones individuales a partir de la distribución conjunta de dos variables aleatorias. Nuestro objetivo por ahora es mencionar que no es difícil demostrar, y en el Ejercicio 210 se pide hacerlo, que cuando X y Y son discretas, la condición (2.11) es equivalente a la siguiente expresión simple.

Independencia de variables aleatorias: caso discreto

$$P(X = x, Y = y) = P(X = x) P(Y = y), \quad x, y \in \mathbb{R},$$

en donde, por el teorema de probabilidad total,

$$P(X = x) = \sum_y P(X = x, Y = y),$$

$$P(Y = y) = \sum_x P(X = x, Y = y).$$

En el caso cuando X y Y son continuas y la función $F_{X,Y}(x, y)$ puede expresarse de la siguiente forma

$$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(u, v) dv du,$$

a la función $f_{X,Y}(x, y)$ se le llama la función de densidad conjunta de X y Y . En este caso puede demostrarse que la condición de independencia (2.11) es equivalente a la siguiente expresión.

Independencia de variables aleatorias: caso continuo

$$f_{X,Y}(x,y) = f_X(x) f_Y(y), \quad x, y \in \mathbb{R},$$

en donde,

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy,$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx.$$

La definición de independencia de dos v.a.s puede extenderse de manera análoga al caso de tres o más variables aleatorias. Esto significa que la función de distribución conjunta de todas ellas, o equivalentemente, la función de probabilidad conjunta, es el producto de las funciones individuales. Una discusión más completa de este tema puede encontrarse en el capítulo sobre vectores aleatorios.

Ejemplo 2.17 (Producto cartesiano de espacios muestrales) El siguiente procedimiento es una versión simple de un mecanismo general para construir variables aleatorias independientes. Consideremos que tenemos un primer experimento aleatorio con espacio muestral $\Omega_1 = \{a, b\}$ y con probabilidades p y $1 - p$ para cada uno de estos resultados, respectivamente, y que, por otro lado, también tenemos un segundo experimento con espacio muestral $\Omega_2 = \{c, d\}$ con probabilidades q y $1 - q$ para estos dos resultados. La intención es crear el experimento aleatorio consistente en llevar a cabo el primer experimento seguido del segundo, y para este experimento compuesto definir dos variables aleatorias independientes. Así, el espacio muestral para el experimento compuesto será el espacio producto

$$\Omega = \Omega_1 \times \Omega_2 = \{(a, c), (a, d), (b, c), (b, d)\}.$$

Definamos la probabilidad de cada uno de estos cuatro resultados como el producto de las probabilidades individuales, es decir, se asocian las probabilidades: $pq, p(1 - q), (1 - p)q, (1 - p)(1 - q)$, respectivamente conforme al orden en el que aparecen en el espacio muestral. Esta asignación de probabilidades es el elemento clave para verificar la propiedad de independencia.

Definamos ahora las variables aleatorias X y Y de la siguiente forma:

$$\begin{aligned} X(a, -) &= 0, & Y(-, c) &= 0, \\ X(b, -) &= 1, & Y(-, d) &= 1. \end{aligned}$$

Por lo tanto, sin importar el resultado del segundo experimento, la variable X asigna los valores cero y uno a los dos resultados del primer experimento. Simétricamente, sin importar el resultado del primer experimento, la variable Y asigna los valores cero y uno a los dos resultados del segundo experimento. Puede verificarse que X y Y son independientes pues, por construcción, para cualesquiera valores de x y y se cumple la identidad

$$P(X = x, Y = y) = P(X = x)P(Y = y).$$

•

Para algunos ejercicios que se plantean en el siguiente capítulo se necesita además tener una noción de independencia de tres o más variables aleatorias. Para ello y de manera preliminar la independencia se traduce en igualar las probabilidades conjuntas de eventos de las variables aleatorias con el producto de las probabilidades de los eventos individuales, tal y como lo hemos mencionado en esta sección para el caso de dos variables aleatorias. Reiteramos que esto se formalizará más adelante cuando estudiemos el tema de vectores aleatorios.

Ejercicios

210. **Independencia: condición equivalente caso discreto.** Sean X y Y dos variables aleatorias discretas. Sin pérdida de generalidad suponga que ambas toman valores en \mathbb{N} .

a) Usando la identidad

$$P(X \leq x, Y \leq y) = \sum_{u \leq x} \sum_{v \leq y} P(X = u, Y = v),$$

demuestre que

$$\begin{aligned} P(X = x, Y = y) &= P(X \leq x, Y \leq y) - P(X \leq x - 1, Y \leq y) \\ &\quad - P(X \leq x, Y \leq y - 1) \\ &\quad + P(X \leq x - 1, Y \leq y - 1). \end{aligned}$$

b) Usando el inciso anterior demuestre que la condición de independencia

$$P(X \leq x, Y \leq y) = P(X \leq x) P(Y \leq y), \quad x, y \in \mathbb{R}.$$

es equivalente a la condición

$$P(X = x, Y = y) = P(X = x) P(Y = y), \quad x, y \in \mathbb{R}.$$

211. Sean X_1, \dots, X_m variables aleatorias discretas e independientes, todas ellas con idéntica función de probabilidad dada por

$$f(x) = \begin{cases} 1/n & \text{si } x = 1, \dots, n, \\ 0 & \text{en otro caso.} \end{cases}$$

Más adelante nos referiremos a esta distribución como la distribución uniforme. Encuentre la función de probabilidad de la variable aleatoria:

- a) $U = \text{máx} \{X_1, \dots, X_m\}$.
 b) $V = \text{mín} \{X_1, \dots, X_m\}$.

2.6. Esperanza

En las siguientes secciones estudiaremos ciertas cantidades numéricas que pueden ser calculadas para cada variable aleatoria. Estos números revelan algunas características de la variable aleatoria o de su distribución. La primera de estas características numéricas es la esperanza, se denota por $E(X)$ y se calcula como se define a continuación.

Definición 2.8 Sea X una v.a. discreta con función de probabilidad $f(x)$. La esperanza de X se define como el número

$$E(X) = \sum_x x f(x), \quad (2.12)$$

suponiendo que esta suma es absolutamente convergente, es decir, cuando la suma de los valores absolutos es convergente. Por otro lado, si X es continua con función de densidad $f(x)$, entonces la esperanza es

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx, \quad (2.13)$$

suponiendo que esta integral es absolutamente convergente, es decir, cuando la integral de los valores absolutos es convergente.

El número de sumandos en la expresión (2.12) puede ser finito o infinito, dependiendo del conjunto de valores que toma la variable aleatoria. Además, si la suma o integral anteriores no cumplen la condición de convergencia absoluta, entonces se dice que la esperanza no existe o bien que no tiene esperanza finita. En los Ejercicios 220, 221 y 222, en la página 165, pueden encontrarse ejemplos en donde se presenta esta situación.

La esperanza de una variable aleatoria es entonces un número que indica el promedio ponderado de los diferentes valores que la variable puede tomar. A la esperanza se le conoce también con los nombre de media, valor esperado o valor promedio. En general, se usa también la letra griega μ (mu) para denotarla; es uno de los conceptos más importantes en probabilidad y tiene un amplio uso en las aplicaciones y otras ramas de la ciencia. A partir de este momento resultará muy útil conocer algunas fórmulas para llevar a cabo sumas y poseer un manejo adecuado de las técnicas de integración. Mediante algunos ejemplos ilustraremos a continuación, la forma de calcular esperanzas de variables aleatorias.

Ejemplo 2.18 (Caso discreto) Sea X una variable aleatoria discreta con función de probabilidad dada por la siguiente tabla.

x	-1	0	1	2
$f(x)$	1/8	4/8	1/8	2/8

La esperanza de X es el número

$$\begin{aligned}
 E(X) &= \sum_x x f(x) \\
 &= (-1)(1/8) + (0)(4/8) + (1)(1/8) + (2)(2/8) \\
 &= 1/2.
 \end{aligned}$$

Observe que la suma se efectúa sobre todos los valores de x indicados en la tabla, es decir, $-1, 0, 1$ y 2 . También es instructivo observar que la esperanza no es necesariamente uno de los valores tomados por la variable aleatoria. En este ejemplo, el valor $1/2$ nunca es tomado por la variable aleatoria, pero es su valor promedio. •

Ejemplo 2.19 (Caso continuo) Considere la variable aleatoria continua X con función de densidad

$$f(x) = \begin{cases} 2x & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

La esperanza de X es

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 2x^2 dx = 2/3.$$

Observe que la integral sólo es relevante en el intervalo $(0, 1)$, pues fuera de este intervalo la función de densidad se anula. •

De este modo la esperanza de una variable aleatoria es un promedio ponderado de los valores que toma la variable aleatoria: cada valor x se pondera con la función de probabilidad $f(x)$ y se suman todas estas cantidades.

Esperanza de una función de una variable aleatoria

En algunos casos es necesario calcular la esperanza de una función de una variable aleatoria, por ejemplo, si X es una variable aleatoria discreta, entonces es claro que $Y = X^2$ es una función de X , y consideraremos que Y es también una variable aleatoria. Si quisiéramos encontrar la esperanza de Y según la expresión (2.12) de la definición de esperanza tendríamos que calcular

$$E(Y) = \sum_y y f_Y(y).$$

Para lo cual se necesita encontrar primero la función de densidad de Y , y ello, en general, no es fácil. El siguiente resultado es muy útil y establece una forma de calcular esta esperanza conociendo únicamente la función de densidad de X . A este resultado se le refiere a veces como el teorema del estadístico inconsciente y lo usaremos con bastante regularidad, ya sea en esta versión discreta o en la versión continua que presentaremos más adelante.

Proposición 2.3 Sea X una variable aleatoria discreta y sea $g : \mathbb{R} \rightarrow \mathbb{R}$ una función tal que $g(X)$ es una variable aleatoria con esperanza finita. Entonces

$$E[g(X)] = \sum_x g(x) f_X(x). \quad (2.14)$$

Demostración. Sea $Y = g(X)$ y sea y uno de sus posibles valores. Así, existe por lo menos un valor x tal que $y = g(x)$. Agrupemos todos estos valores x en el conjunto $g^{-1}y = \{x : g(x) = y\}$, como se muestra en la Figura 2.14. De este modo tenemos que

$$P(Y = y) = P(X \in g^{-1}y) = \sum_{x \in g^{-1}y} f_X(x).$$

Es claro que la unión de los conjuntos disjuntos $g^{-1}y$, considerando todos los valores y , es el conjunto de valores x que la variable aleatoria X puede

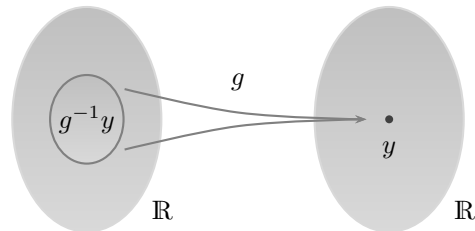


Figura 2.14

tomar. Por lo tanto, la esperanza de Y es

$$\begin{aligned}
 E(Y) &= \sum_y yP(Y = y) \\
 &= \sum_y y \sum_{x \in g^{-1}y} f_X(x) \\
 &= \sum_y \sum_{x \in g^{-1}y} g(x)f_X(x) \\
 &= \sum_x g(x)f_X(x).
 \end{aligned}$$

•

Como hemos señalado antes, debe observarse que en la fórmula (2.14) no aparece la función de probabilidad de Y , sino la de X . Allí radica la utilidad de esta fórmula, pues para calcular $E(Y)$ no es necesario conocer $f_Y(y)$. Veamos un ejemplo de aplicación.

Ejemplo 2.20 Sea X una variable aleatoria discreta con función de probabilidad dada por la siguiente tabla.

x	-2	-1	0	1	2
$f_X(x)$	2/8	1/8	2/8	1/8	2/8

Nos interesa calcular la esperanza de la variable aleatoria $Y = X^2$.

Solución 1. Un primer método consiste en calcular primero la función de probabilidad de la variable Y . No es complicado verificar que esta función es

y	0	1	4
$f_Y(y)$	2/8	2/8	4/8

Entonces, usando la definición elemental de esperanza para variables aleatorias discretas tenemos que

$$E(Y) = (0)(2/8) + (1)(2/8) + (4)(4/8) = 9/4.$$

Solución 2. Ahora calcularemos la misma esperanza, pero usando la fórmula (2.14). El resultado es el mismo, pero la ventaja es que no es necesario calcular la función de probabilidad de Y .

$$E(Y) = (-2)^2(2/8) + (-1)^2(1/8) + (0)^2(2/8) + (1)^2(1/8) + (2)^2(2/8) = 9/4.$$

•

El siguiente resultado corresponde a una versión continua de la proposición anterior. Su demostración es ligeramente más avanzada que la presentada en el caso discreto y su lectura puede omitirse sin mayores consecuencias. A partir de ahora, con frecuencia, calcularemos la esperanza de una función de una v.a. de la manera como indican estos resultados como si se tratara de una definición. En efecto, en algunos cursos y textos elementales de probabilidad se adopta tal perspectiva.

Proposición 2.4 Sea X una variable aleatoria continua y sea $g : \mathbb{R} \rightarrow \mathbb{R}$ una función tal que $g(X)$ es una variable aleatoria con esperanza finita. Entonces

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx. \quad (2.15)$$

Demostración. Supongamos primero que la función $g(x)$ es no negativa. Entonces, por la fórmula (2.18) que aparece en el Ejercicio 217, en la

página 165, la esperanza de $g(X)$ puede calcularse como sigue

$$\begin{aligned}
 E(g(X)) &= \int_0^{\infty} P(g(X) > y) dy \\
 &= \int_0^{\infty} P(X \in g^{-1}(y, \infty)) dy \\
 &= \int_0^{\infty} \int_{g^{-1}(y, \infty)} f_X(x) dx dy \\
 &= \int_{g^{-1}(0, \infty)} \int_0^{g(x)} f_X(x) dy dx \\
 &= \int_{\mathbb{R}} g(x) f_X(x) dx.
 \end{aligned}$$

De esta forma hemos demostrado el resultado para funciones no negativas. Veremos que esto es suficiente para obtener el resultado general. Para cualquier función $g(x)$, se puede definir su parte positiva y su parte negativa de la siguiente forma

$$g^+(x) = \begin{cases} g(x) & \text{si } g(x) \geq 0, \\ 0 & \text{si } g(x) < 0. \end{cases} \quad g^-(x) = \begin{cases} 0 & \text{si } g(x) \geq 0, \\ g(x) & \text{si } g(x) < 0. \end{cases}$$

Por lo tanto, $g(x)$ admite la descomposición

$$g(x) = g^+(x) - (-g^-(x)),$$

en donde, tanto $g^+(x)$ como $-g^-(x)$ son funciones no negativas. En consecuencia,

$$\begin{aligned}
 E(g(X)) &= E(g^+(X) - (-g^-(X))) \\
 &= E(g^+(X)) - E(-g^-(X)) \\
 &= \int_{\mathbb{R}} g^+(x) f_X(x) dx - \int_{\mathbb{R}} -g^-(x) f_X(x) dx \\
 &= \int_{\mathbb{R}} (g^+(x) + g^-(x)) f_X(x) dx \\
 &= \int_{\mathbb{R}} g(x) f_X(x) dx.
 \end{aligned}$$

•

Ejemplo 2.21 Calcularemos $E(X^2)$, en donde X es la variable aleatoria continua con función de densidad

$$f(x) = \begin{cases} 1 & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Solución 1. Si se desea aplicar la identidad (2.13) de la definición elemental de esperanza, se tiene que encontrar primero la distribución de la v.a. X^2 . Puede verificarse que

$$F_{X^2}(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ \sqrt{x} & \text{si } 0 < x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

De donde, derivando,

$$f_{X^2}(x) = \begin{cases} \frac{1}{2}x^{-1/2} & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Finalmente, aplicando la definición de esperanza (2.13),

$$E(X^2) = \int_0^1 x f_{X^2}(x) dx = \int_0^1 \frac{1}{2} x^{1/2} dx = 1/3.$$

Solución 2. Ahora resolveremos el mismo problema de una manera más rápida, aplicando el resultado de la proposición anterior. Por la fórmula (2.15), tenemos que

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_0^1 x^2 dx = 1/3.$$

•

Los resultados anteriores pueden extenderse al caso de dos o mas variables aleatorias como se muestra en la siguiente proposición.

Proposición 2.5 Sean X y Y dos variables aleatorias continuas definidas sobre un mismo espacio de probabilidad y con función de densidad conjunta $f(x, y)$. Sea $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ una función tal que $g(X, Y)$ es una variable aleatoria con esperanza finita. Entonces

$$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dy dx. \quad (2.16)$$

La demostración de este resultado es similar al caso presentado de una sola variable aleatoria y por lo tanto la omitiremos. Cuando las variables aleatorias son ambas discretas, el resultado es análogo, y en lugar de integrales aparecen sumas. Como ejemplos de aplicación de la fórmula (2.16) en el caso discreto, tenemos las siguientes expresiones:

a) Considerando la función $g(x, y) = x + y$, tenemos que

$$E(X + Y) = \sum_x \sum_y (x + y) f(x, y).$$

b) Considerando la función $g(x, y) = xy$, tenemos que

$$E(XY) = \sum_x \sum_y xy f(x, y).$$

Estudiaremos a continuación algunas propiedades generales de la esperanza.

Proposición 2.6 (Propiedades de la esperanza) Sean X y Y dos variables aleatorias con esperanza finita y sea c una constante. Entonces

1. $E(c) = c$.
2. $E(cX) = cE(X)$.
3. Si $X \geq 0$ entonces $E(X) \geq 0$.
4. $E(X + Y) = E(X) + E(Y)$.

Demostración. La primera propiedad es evidente, pues si X es la variable aleatoria constante c , entonces por definición, $E(X) = cP(X = c) = c1 = c$. El segundo inciso se sigue directamente de la definición de esperanza pues, tanto en el caso discreto como en el caso continuo, la constante c puede siempre colocarse fuera de la suma o integral. El tercer inciso también es inmediato pues en la integral o suma correspondiente sólo aparecerán términos no negativos. Para la última propiedad, consideraremos el caso en el que ambas variables son discretas, y por simplicidad usaremos la expresión $p(x, y)$ para denotar $P(X = x, Y = y)$.

$$\begin{aligned}
 E(X + Y) &= \sum_{x,y} (x + y) p(x, y) \\
 &= \sum_{x,y} x p(x, y) + \sum_{x,y} y p(x, y) \\
 &= \sum_x x \sum_y p(x, y) + \sum_y y \sum_x p(x, y) \\
 &= \sum_x x p(x) + \sum_y y p(y) \\
 &= E(X) + E(Y).
 \end{aligned}$$

•

Observe que la segunda y cuarta propiedad establecen que la esperanza es lineal, es decir, separa sumas y multiplicaciones por constantes. Otras propiedades de la esperanza se encuentran en la sección de ejercicios. Veamos ahora una aplicación del concepto de independencia en el cálculo de la esperanza.

Proposición 2.7 Sean X y Y independientes y ambas con esperanza finita. Entonces

$$E(XY) = E(X)E(Y).$$

Demostración. Por simplicidad consideraremos únicamente el caso en el que ambas variables aleatorias son discretas. En el caso continuo, el proce-

dimiento es análogo.

$$\begin{aligned}
 E(XY) &= \sum_{x,y} xy P(X = x, Y = y) \\
 &= \sum_x \sum_y xy P(X = x)P(Y = y) \\
 &= \left(\sum_x x P(X = x) \right) \left(\sum_y y P(Y = y) \right) \\
 &= E(X) E(Y).
 \end{aligned}$$

•

En el Ejercicio 231 se muestra una extensión de este resultado que se usará más adelante. Se debe señalar además que el recíproco de la proposición anterior no es válido en general, es decir, la condición $E(XY) = E(X)E(Y)$ no es suficiente para concluir que X y Y son independientes. Véase el Ejercicio 229 en la página 169 para un ejemplo de esta situación.

Ejercicios

212. Calcule la esperanza de la variable aleatoria discreta X con función de probabilidad:

$$\begin{aligned}
 a) \quad f(x) &= \begin{cases} 1/n & \text{si } x = 1, 2, \dots, n, \\ 0 & \text{en otro caso.} \end{cases} \\
 b) \quad f(x) &= \begin{cases} 1/2^x & \text{si } x = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases} \\
 c) \quad f(x) &= \begin{cases} \frac{2x}{n(n+1)} & \text{si } x = 1, 2, \dots, n, \\ 0 & \text{en otro caso.} \end{cases}
 \end{aligned}$$

213. Calcule la esperanza de la variable aleatoria continua X con función de densidad:

$$a) \quad f(x) = \begin{cases} |x| & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x) = \begin{cases} x^2 + 4x/3 & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f(x) = \begin{cases} e^{-x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

$$d) f(x) = \begin{cases} (x/4) e^{-x/2} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

$$e) f(x) = \begin{cases} 3(1 - x^2)/4 & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

214. **Monotonía.** Sean X y Y dos variables aleatorias con esperanza finita. Demuestre que si $X \leq Y$ entonces

$$E(X) \leq E(Y).$$

215. **Fórmula alternativa, caso discreto.** Sea X una variable aleatoria discreta con función de distribución $F(x)$, con esperanza finita y posibles valores dentro del conjunto $\{0, 1, 2, \dots\}$. Demuestre que

$$E(X) = \sum_{x=0}^{\infty} (1 - F(x)). \quad (2.17)$$

216. Use la fórmula (2.17) del ejercicio anterior para encontrar la esperanza de X cuando ésta tiene función de distribución:

$$a) F(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1/5 & \text{si } 0 \leq x < 1, \\ 3/5 & \text{si } 1 \leq x < 2, \\ 4/5 & \text{si } 2 \leq x < 3, \\ 1 & \text{si } x \geq 3. \end{cases}$$

$$b) F(x) = \begin{cases} 0 & \text{si } x < 1, \\ 1 - (1/2)^k & \text{si } k \leq x < k + 1; \quad k = 1, 2, \dots \end{cases}$$

217. **Fórmula alternativa, caso continuo.** Sea X una variable aleatoria continua con función de distribución $F(x)$, con esperanza finita y con valores en el intervalo $[0, \infty)$. Demuestre que

$$E(X) = \int_0^{\infty} (1 - F(x)) dx. \quad (2.18)$$

218. Use la fórmula (2.18) del ejercicio anterior para encontrar $E(X)$ cuando X tiene función de distribución:

$$a) F(x) = \begin{cases} 0 & \text{si } x < 0, \\ x/2 & \text{si } 0 \leq x < 2, \\ 1 & \text{si } x \geq 2. \end{cases}$$

$$b) F(x) = \begin{cases} 1 - e^{-x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

219. Demuestre o proporcione un contraejemplo.

- a) $E(E(X)) = E^2(X)$.
- b) $E(X^2 - Y^2) = E(X + Y)E(X - Y)$.
- c) $E(1/X) = 1/E(X)$.
- d) $E(X/E(X)) = 1$.
- e) $E(X - E(X)) = E(E(X) - X) = 0$.
- f) Si $E(X) = 0$ entonces $X = 0$.
- g) Si $E(X^2) = 0$ entonces $X = 0$.
- h) Si $E(X) = E(Y)$ entonces $X = Y$.
- i) $E(X) \leq E(X^2)$.
- j) $E(X + X) = 2E(X)$.
- k) $E(X^2) = E^2(X)$.

220. **Sin esperanza, caso discreto.** Sea X una variable aleatoria discreta con función de probabilidad $f(x)$ como aparece abajo. Demuestre que

$f(x)$ es efectivamente una función de probabilidad y compruebe que X no tiene esperanza finita.

$$f(x) = \begin{cases} \frac{1}{x(x+1)} & \text{si } x = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

221. **Sin esperanza, caso continuo.** Sea X una variable aleatoria continua con función de densidad $f(x)$ como en cualquiera de los casos que aparecen abajo. Demuestre que $f(x)$ es, efectivamente, una función de densidad y compruebe que X no tiene esperanza finita.

$$a) f(x) = \begin{cases} 1/x^2 & \text{si } x > 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \mathbb{R}.$$

222. **La paradoja de San Petersburgo.** Un jugador lanza repetidas veces una moneda equilibrada hasta obtener una de las caras previamente escogida. El jugador obtiene un premio de 2^n unidades monetarias si logra su objetivo en el n -ésimo lanzamiento. Calcule el valor promedio del premio en este juego.
223. **El problema del ladrón de Bagdad.** El ladrón de Bagdad ha sido colocado en una prisión en donde hay tres puertas. Una de las puertas conduce a un túnel que requiere de un día de travesía y que lo regresa a la misma prisión. Otra de las puertas lo conduce a otro túnel, aún más largo, que lo regresa nuevamente a la prisión pero después de tres días de recorrido. Finalmente, la tercera puerta lo conduce a la libertad inmediatamente. Suponga que el ladrón escoge al azar cada una estas puertas, una por una, hasta quedar libre. Encuentre el número promedio de días que le toma al ladrón quedar en libertad.
224. Un jugador lanza dos veces consecutivas un dado equilibrado y obtiene como premio tantas unidades monetarias como indica el resultado mayor de los dos lanzamientos. Encuentre el valor promedio del premio.

225. **Funciones convexas.** Se dice que una función $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ es convexa si para cualesquiera dos puntos x y y , y para cualquier $\lambda \in [0, 1]$, se cumple la desigualdad

$$\varphi(\lambda x + (1 - \lambda)y) \leq \lambda \varphi(x) + (1 - \lambda) \varphi(y). \quad (2.19)$$

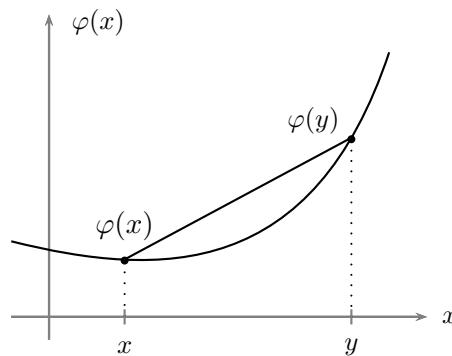


Figura 2.15

Observe que el conjunto de puntos $\{\lambda x + (1 - \lambda)y : \lambda \in [0, 1]\}$ corresponde al segmento de recta que une a los puntos x y y . Por lo tanto, esta desigualdad establece que la imagen de este segmento de recta queda por debajo de la recta que une los valores $\varphi(x)$ y $\varphi(y)$. Esta situación se muestra gráficamente en la Figura 2.15. La definición de convexidad puede ser extendida al caso cuando la función no necesariamente esté definida para la totalidad de los números reales, por ejemplo, puede adecuarse sin dificultad al caso cuando el dominio de definición sea el intervalo $(0, \infty)$. Demuestre que:

a) La condición (2.19) es equivalente a

$$\varphi(\lambda_1 x_1 + \cdots + \lambda_n x_n) \leq \lambda_1 \varphi(x_1) + \cdots + \lambda_n \varphi(x_n),$$

para cualesquiera números reales x_1, \dots, x_n y valores no negativos $\lambda_1, \dots, \lambda_n$ tales que $\lambda_1 + \cdots + \lambda_n = 1$, en donde n es cualquier número natural mayor o igual a 2.

- b) Si X es una variable aleatoria estrictamente positiva que toma un número finito de valores, entonces

$$\frac{1}{E(X)} \leq E\left(\frac{1}{X}\right).$$

226. Sea X una variable aleatoria con función de densidad $f(x)$ como aparece abajo. Encuentre el valor de la constante $a > 0$ tal que $E(X) = 0$.

$$f(x) = \begin{cases} e^{-(x+a)} & \text{si } x > -a, \\ 0 & \text{en otro caso.} \end{cases}$$

227. **Función signo.** La función signo se define sobre el conjunto de números reales de la forma siguiente:

$$\text{signo}(x) = \begin{cases} +1 & \text{si } x > 0, \\ 0 & \text{si } x = 0, \\ -1 & \text{si } x < 0. \end{cases}$$

Calcule la esperanza de la variable aleatoria $Y = \text{signo}(X)$ cuando X tiene función de probabilidad:

a)

x	-2	-1	0	1	2
$f(x)$	1/8	2/8	1/8	2/8	2/8

b) $f(x) = \begin{cases} (x+1)/2 & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$

c) $f(x) = \begin{cases} e^{-x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$

d) $f(x) = \frac{1}{2} e^{-|x|}$.

228. Sea X una variable aleatoria con función de probabilidad $f(x)$ como aparece abajo. Calcule la esperanza de la variable $Y = e^X$ encontrando primero la función de densidad de Y y aplicando la definición elemental de esperanza. Como segundo método use el teorema del estadístico inconsciente. Ambos cálculos deben coincidir.

$$a) f(x) = \begin{cases} (1-p)p^{x-1} & \text{si } x = 1, 2, \dots \quad (0 < p < e^{-1} \text{ constante}) \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0, \quad (\lambda > 1 \text{ constante}) \\ 0 & \text{en otro caso.} \end{cases}$$

229. Sea X una variable aleatoria discreta con función de probabilidad dada por la tabla que aparece abajo. Defina, por otro lado a la variable $Y = X^2$. Se cumple entonces la identidad $E(XY) = E(X)E(Y)$ y sin embargo X y Y no son independientes. Demuestre ambas afirmaciones.

x	-1	0	1
$f(x)$	1/3	1/3	1/3

230. Sea X una variable aleatoria no negativa con función de distribución $F(x)$, función de densidad $f(x)$ y con esperanza finita $\mu > 0$. Demuestre que las siguientes funciones son de densidad:

$$a) g(x) = 2(1 - F(x))f(x).$$

$$b) g(x) = (1 - F(x))/\mu.$$

231. Sean X y Y dos variables aleatorias independientes y sean $g(x)$ y $h(y)$ dos funciones de \mathbb{R} en \mathbb{R} tales que el producto $g(X)h(Y)$ es una variable aleatoria con esperanza finita. Demuestre que

$$E[g(X)h(Y)] = E[g(X)] \cdot E[h(Y)].$$

2.7. Varianza

Otra característica numérica importante asociada a las variables aleatorias se llama varianza, se denota por $\text{Var}(X)$ y se calcula de la forma siguiente.

Definición 2.9 Sea X una v.a. discreta con función de probabilidad $f(x)$. La varianza de X se define como el número

$$\text{Var}(X) = \sum_x (x - \mu)^2 f(x),$$

cuando esta suma es convergente y en donde μ es la esperanza de X . Para una variable aleatoria continua X con función de densidad $f(x)$ se define

$$\text{Var}(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx,$$

cuando esta integral es convergente.

Así, observe que se necesita conocer la esperanza de X para calcular su varianza. Es interesante observar también que la varianza se puede escribir en una sola expresión como sigue:

$$\text{Var}(X) = E(X - \mu)^2.$$

Esto corresponde a la esperanza de la función cuadrática $g(x) = (x - \mu)^2$ aplicada a una variable aleatoria X con esperanza μ . La varianza es una medida del grado de dispersión de los diferentes valores tomados por la variable. Se le denota regularmente por la letra σ^2 (sigma cuadrada). A la raíz cuadrada positiva de la varianza, esto es σ , se le llama desviación estándar. Como en el caso de la esperanza, la anterior suma o integral puede no ser convergente y en ese caso se dice que la variable aleatoria no tiene varianza finita. Veamos algunos ejemplos.

Ejemplo 2.22 (Caso discreto) Calcularemos la varianza de la variable aleatoria discreta X con función de probabilidad dada por la siguiente tabla.

x	-1	0	1	2
$f(x)$	1/8	4/8	1/8	2/8

Recordemos primeramente que, por cálculos previos, $\mu = 1/2$. Aplicando la

definición de varianza tenemos que

$$\begin{aligned}\text{Var}(X) &= \sum_x (x - \mu)^2 f(x) \\ &= (-1 - 1/2)^2(1/8) + (0 - 1/2)^2(4/8) \\ &\quad + (1 - 1/2)^2(1/8) + (2 - 1/2)^2(2/8) \\ &= 1.\end{aligned}$$

•

Ejemplo 2.23 (Caso continuo) Calcularemos la varianza de la variable aleatoria continua X con función de densidad

$$f(x) = \begin{cases} 2x & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

En un cálculo previo habíamos encontrado que la esperanza de esta variable aleatoria es $\mu = 2/3$. Por lo tanto,

$$\text{Var}(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx = \int_0^1 (x - 2/3)^2 2x dx = 1/18.$$

•

A continuación demostraremos algunas propiedades generales de la varianza, y siendo ésta una esperanza, se hará uso de lo que se estudió antes sobre la esperanza de una variable aleatoria.

Proposición 2.8 (Propiedades de la varianza) Sean X y Y dos variables aleatorias con varianza finita y sea c una constante. Entonces

1. $\text{Var}(X) \geq 0$.
2. $\text{Var}(c) = 0$.
3. $\text{Var}(cX) = c^2 \text{Var}(X)$.
4. $\text{Var}(X + c) = \text{Var}(X)$.
5. $\text{Var}(X) = E(X^2) - E^2(X)$.
6. En general, $\text{Var}(X + Y) \neq \text{Var}(X) + \text{Var}(Y)$.

Demostración. El inciso (1) es evidente a partir de la definición de varianza pues en ella aparece una suma o integral de términos no negativos. Para el inciso (2), la constante c es una variable aleatoria con un único valor, de modo que $E(c) = c$ y entonces $\text{Var}(c) = E(c - c)^2 = 0$. Para el inciso (3) tenemos que

$$\begin{aligned} \text{Var}(cX) &= E(cX - E(cX))^2 \\ &= E(cX - cE(X))^2 \\ &= c^2 E(X - E(X))^2 \\ &= c^2 \text{Var}(X). \end{aligned}$$

El inciso (4) se sigue del siguiente análisis,

$$\text{Var}(X + c) = E[(X + c) - E(X + c)]^2 = E(X - E(X))^2 = \text{Var}(X).$$

Para demostrar la propiedad (5) se desarrolla el cuadrado en la definición de varianza y se usa la propiedad de linealidad de la esperanza,

$$\begin{aligned} \text{Var}(X) &= E(X - E(X))^2 \\ &= E(X^2 - 2XE(X) + E^2(X)) \\ &= E(X^2) - 2E(X)E(X) + E^2(X) \\ &= E(X^2) - E^2(X). \end{aligned}$$

Finalmente, para demostrar la propiedad (6), es suficiente dar un ejemplo. Puede tomarse el caso $Y = X$, en general y por lo demostrado antes, no se cumple que $\text{Var}(2X) = 2\text{Var}(X)$. •

De estas propiedades generales se obtiene, en particular, que la varianza es siempre una cantidad no negativa y que no cumple la propiedad de linealidad, pues en general no separa sumas y cuando aparecen constantes como factores, las constantes se separan de la varianza elevándolas al cuadrado. Otras propiedades generales se encuentran en el Ejercicio 238, en la página 176. Veamos ahora una fórmula para el cálculo de la varianza de la suma de dos variables aleatorias bajo la hipótesis de independencia. Esta hipótesis adicional hará que aparezca una igualdad en la propiedad (6) recién demostrada.

Proposición 2.9 Sean X y Y dos variables aleatorias independientes y con varianza finita. Entonces

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

Demostración. Usaremos la propiedad de linealidad de la esperanza y el hecho de que $E(XY) = E(X)E(Y)$, pues X y Y son independientes por hipótesis.

$$\begin{aligned} \text{Var}(X + Y) &= E(X + Y)^2 - E^2(X + Y) \\ &= E(X^2 + 2XY + Y^2) - (E(X) + E(Y))^2 \\ &= E(X^2) + 2E(XY) + E(Y^2) \\ &\quad - E^2(X) - 2E(X)E(Y) - E^2(Y) \\ &= (E(X^2) - E^2(X)) + (E(Y^2) - E^2(Y)) \\ &= \text{Var}(X) + \text{Var}(Y). \end{aligned}$$

El recíproco del resultado anterior es, en general, falso, es decir, la condición $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ no es suficiente para concluir que X y Y son independientes. Un ejemplo de esta situación se muestra en el Ejercicio 476,

en la página 324, el cual requiere del concepto de distribución conjunta de variables aleatorias que estudiaremos con más detalle en el capítulo sobre vectores aleatorios.

Ejercicios

232. Calcule la media y la varianza de la variable aleatoria discreta X con función de probabilidad:

$$a) f(x) = \begin{cases} 1/n & \text{si } x = 1, 2, \dots, n, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x) = \begin{cases} 1/2^x & \text{si } x = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

233. Calcule la media y la varianza de la variable aleatoria continua X con función de densidad:

$$a) f(x) = \begin{cases} 1/3 & \text{si } 0 < |x| < 1, \\ 1/6 & \text{si } 1 < |x| < 2, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x) = \begin{cases} e^{-x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f(x) = \begin{cases} |x| & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

234. Calcule la media y la varianza de la variable aleatoria X con distribución:

$$a) f(x) = \frac{1}{2} e^{-|x|}.$$

$$b) f(x) = \begin{cases} 1 - |x| & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f(x) = \begin{cases} 6x(1-x) & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$d) F(x) = \begin{cases} 0 & \text{si } x < 0, \\ x & \text{si } 0 \leq x \leq 1, \\ 1 & \text{si } x > 1. \end{cases}$$

$$e) F(x) = \begin{cases} 0 & \text{si } x < -a, \\ \frac{x-a}{2a} & \text{si } -a \leq x \leq a, \\ 1 & \text{si } x > a. \end{cases}$$

$$f) F(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - \cos x & \text{si } 0 \leq x \leq \pi/2, \\ 1 & \text{si } x > \pi/2. \end{cases}$$

235. Encuentre la distribución de la variable aleatoria X que cumple las siguientes dos condiciones:

$$\begin{aligned} P(X = 1) &= 1 - P(X = -1), \\ E(X) &= \text{Var}(X). \end{aligned}$$

236. Encuentre el valor del parámetro θ de tal forma que la varianza de la siguiente distribución sea uno.

$$f(x) = \begin{cases} |x|/\theta^2 & \text{si } -\theta < x < \theta, \\ 0 & \text{en otro caso.} \end{cases}$$

237. Sea X una variable aleatoria con función de densidad como aparece a continuación, en donde a y b son dos constantes.

$$f(x) = \begin{cases} ax^2 + bx & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Determine el valor de las constantes a y b de tal forma que

a) la esperanza sea mínima.

b) la varianza sea mínima.

238. Demuestre las siguientes propiedades de la varianza.

a) $\text{Var}(X) \leq E(X^2)$.

b) $\text{Var}(a - X) = \text{Var}(X)$, a constante.

c) $\text{Var}(aX + b) = a^2 \text{Var}(X)$, a, b constantes.

d) $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2E[(X - E(X))(Y - E(Y))]$.

239. Sea X una variable aleatoria con media y varianza finita. Defina la función $g(x) = E[(X - x)^2]$. Demuestre que:

a) $g(x) = \text{Var}(X) + (x - E(X))^2$.

b) $g(x)$ tiene un mínimo en $x = E(X)$ y que ese valor mínimo es $\text{Var}(X)$.

240. Sea X una variable aleatoria arbitraria con posibles valores en el intervalo $[a, b]$.

a) Demuestre que $a \leq E(X) \leq b$.

b) Demuestre que $0 \leq \text{Var}(X) \leq (b - a)^2/4$.

c) Encuentre X tal que $\text{Var}(X)$ es máxima.

241. Demuestre o proporcione un contraejemplo.

a) $\text{Var}(\text{Var}(X)) = 0$.

b) $\text{Var}(E(X)) = E(X)$.

c) $E(\text{Var}(X)) = \text{Var}(X)$.

d) $\text{Var}(X - Y) = \text{Var}(X) - \text{Var}(Y)$.

e) Si $E(X)$ existe entonces $\text{Var}(X)$ existe.

f) Si $\text{Var}(X)$ existe entonces $E(X)$ existe.

g) Si $\text{Var}(X) = 0$ entonces $X = 0$.

h) Si $\text{Var}(X) = \text{Var}(Y)$ entonces $X = Y$.

i) $\text{Var}(X + Y) \leq \text{Var}(X) + \text{Var}(Y)$.

242. Sean X_1, \dots, X_n v.a.s independientes e idénticamente distribuidas con media μ y varianza σ^2 . La media muestral se define como la v.a.

$$\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k.$$

Demuestre que:

- a) $E(\bar{X}) = \mu$.
 b) $\text{Var}(\bar{X}) = \sigma^2/n$.
243. Sean X y Y dos variables aleatorias con varianza finita. Demuestre que

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) \iff E(XY) = E(X)E(Y).$$

244. Sea X una v.a. con función de distribución $F(x)$ como aparece abajo, en donde $0 \leq a \leq 1$, $\lambda_1 > 0$ y $\lambda_2 > 0$ son constantes. Encuentre la media y la varianza de X .

$$F(x) = \begin{cases} a(1 - e^{-\lambda_1 x}) + (1 - a)(1 - e^{-\lambda_2 x}) & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

245. Sean a y ℓ dos constantes con $\ell > 0$. Encuentre la esperanza y la varianza de una variable aleatoria con función de densidad dada por

$$f(x) = \begin{cases} \frac{1}{2\ell} & \text{si } |x - a| < \ell, \\ 0 & \text{si } |x - a| \geq \ell. \end{cases}$$

246. Sea Z una variable aleatoria con media 0 y varianza 1. Defina las variables $X = Z - 1$ y $Y = Z + 1$. Demuestre que $E(XY) = 0$.

247. Sea X una variable aleatoria discreta tal que $\text{Var}(X) = 0$. Demuestre que X es constante.

Nota. Compare este enunciado con el resultado más general que aparece en el Ejercicio 509, en la página 348.

248. **Distribución logarítmica.** Se dice que la variable aleatoria discreta X tiene distribución logarítmica de parámetro p , con $0 < p < 1$, si su función de probabilidad es la siguiente:

$$f(x) = \begin{cases} -\frac{1}{\log(1-p)} \frac{1}{x} p^x & \text{si } x = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

Demuestre que:

a) $f(x)$ es, efectivamente, una función de probabilidad.

b) $E(X) = -\frac{1}{\log(1-p)} \frac{p}{1-p}.$

c) $E(X^2) = -\frac{1}{\log(1-p)} \frac{p}{(1-p)^2}.$

d) $\text{Var}(X) = -\frac{1}{(1-p)^2 \log^2(1-p)} p [p + \log(1-p)].$

2.8. Momentos

En esta sección definiremos otras características numéricas de variables aleatorias llamadas momentos. Para cada número natural n se define el n -ésimo momento de una variable aleatoria X como el número $E(X^n)$, suponiendo que tal esperanza existe. Así, tenemos la siguiente definición.

Definición 2.10 Los momentos de una variable aleatoria X o de su distribución son la colección de números:

$$E(X), E(X^2), E(X^3), \dots$$

correspondientes al primer momento, segundo momento, etc., cuando tales cantidades existen.

Para variables aleatorias discretas, el n -ésimo momento se calcula como

sigue:

$$E(X^n) = \sum_x x^n f(x),$$

mientras que para variables aleatorias continuas, la fórmula es

$$E(X^n) = \int_{-\infty}^{\infty} x^n f(x) dx.$$

Suponiendo su existencia, cada uno de estos números representa una característica de la variable aleatoria o de su distribución. Por ejemplo, el primer momento es el valor promedio o la esperanza de la variable aleatoria. Recordando la fórmula

$$\text{Var}(X) = E(X^2) - E^2(X),$$

podemos decir que la varianza es el segundo momento menos el primer momento al cuadrado, de este modo el segundo momento está relacionado con la dispersión de los valores que toma la variable aleatoria. El tercer momento está relacionado con la simetría de la correspondiente distribución de probabilidad. En general, no se conoce una interpretación para cada uno de los momentos de una variable aleatoria, en el mismo sentido que no se conoce una interpretación para cada una de las derivadas de una función infinitamente diferenciable.

Por otro lado, también debemos señalar que los momentos pueden no existir y que en caso de que existan, en general no es de esperarse que se pueda encontrar una expresión compacta para ellos. En la sección 2.12 definiremos la función generadora de momentos, la cual nos permitirá calcular los momentos de una variable aleatoria de una forma alternativa al cálculo de la suma o integral de la definición.

Dada la unicidad de la suma o integral correspondiente, el n -ésimo momento de una variable aleatoria, si existe, es único. Así, cada distribución de probabilidad genera una única colección de momentos, suponiendo su existencia. En el así llamado problema de los momentos se plantea encontrar condiciones bajo las cuales una sucesión de números constituyen los momentos de una distribución de probabilidad. En estos casos, la distribución

de probabilidad puede representarse en términos de la sucesión numérica.

Ejemplo 2.24 Considere una variable aleatoria continua X con función de densidad como aparece abajo. No es difícil comprobar que el primer momento es $E(X) = 0$ y el segundo momento es $E(X^2) = 1/6$.

$$f(x) = \begin{cases} x + 1 & \text{si } -1 < x < 0, \\ 1 - x & \text{si } 0 \leq x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

•

Para concluir esta sección señalaremos que se pueden definir también los siguientes momentos para una variable aleatoria:

Expresión	Momento
$E(X - \mu)^n$	n -ésimo momento central
$E X ^n$	n -ésimo momento absoluto
$E X - \mu ^n$	n -ésimo momento absoluto central
$E(X - c)^n$	n -ésimo momento generalizado (c constante)

Ejercicios

249. Encuentre el n -ésimo momento de una variable aleatoria X con función de densidad

$$a) f(x) = \begin{cases} |x| & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x) = \begin{cases} 6x(1-x) & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f(x) = \begin{cases} x/2 & \text{si } 0 < x < 2, \\ 0 & \text{en otro caso.} \end{cases}$$

250. Sea X una variable aleatoria continua con función de densidad como aparece abajo. Encuentre el n -ésimo momento central de X

$$f(x) = \begin{cases} x & \text{si } 0 < x < 1, \\ 2 - x & \text{si } 1 \leq x < 2, \\ 0 & \text{en otro caso.} \end{cases}$$

251. Sea X una variable aleatoria con distribución uniforme en el intervalo (a, b) y sea n cualquier número natural. Encuentre:

- a) $E(X^n)$.
b) $E(X - \mu)^n$.

252. **Función de probabilidad simétrica.** Una función de probabilidad $f(x)$ es simétrica respecto de a si $f(a + x) = f(a - x)$ para cualquier número real x . Sea X una variable aleatoria con función de probabilidad simétrica respecto de a . Encuentre:

- a) $E(X)$.
b) $E(X - a)^n$ para n impar.

253. Sea X una variable aleatoria con segundo momento finito. Demuestre que

$$0 \leq E^2(X) \leq E(X^2).$$

254. Sea X una variable aleatoria con tercer momento finito. Demuestre o proporcione un contraejemplo para las siguientes afirmaciones.

- a) $E(X) \leq E(X^2)$.
b) $E(X^2) \leq E(X^3)$.

255. Demuestre que toda variable aleatoria acotada tiene todos sus momentos finitos.

256. **Distribución Rayleigh**¹. Se dice que una variable aleatoria continua X tiene distribución Rayleigh de parámetro $\sigma > 0$ si tiene la siguiente

¹John William Strutt, 3rd Baron Rayleigh (1842–1919), físico inglés.

función de densidad

$$f(x) = \begin{cases} \frac{x}{\sigma^2} e^{-x^2/2\sigma^2} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

Lo anterior se escribe como $X \sim \text{Rayleigh}(\sigma)$, en donde a σ se le conoce como parámetro de escala. Esta distribución es un caso particular de la distribución Weibull(α, λ) cuando $\alpha = 2$ y $\lambda = 1/\sqrt{2}\sigma^2$, la cual se estudiará más adelante. Para la distribución Rayleigh, arriba indicada, demuestre que:

- a) $E(X) = \sigma \sqrt{\pi/2}$.
- b) $E(X^2) = 2\sigma^2$.
- c) $\text{Var}(X) = \sigma^2 (2 - \pi/2)$.
- d) $E(X^n) = \begin{cases} \sigma^n 2^{n/2} (n/2)! & \text{si } n \text{ es par,} \\ \sigma^n \sqrt{\pi} \frac{n!}{2^{n/2} ((n-1)/2)!} & \text{si } n \text{ es impar.} \end{cases}$

2.9. Cuantiles

Los cuantiles son otras características numéricas de las distribuciones de probabilidad y se definen de la siguiente forma: sabemos que toda función de distribución $F(x)$ crece de manera continua o a través de saltos, si p es una probabilidad estrictamente positiva, entonces al valor más pequeño x tal que $F(x)$ alcanza el nivel p , o un nivel superior, se le llama cuantil p de la distribución y se le denota por c_p . Tenemos así la siguiente definición.

Definición 2.11 Sea $p \in (0, 1]$. El cuantil p de una v.a. o de su función de distribución $F(x)$ es el número c_p más pequeño, si existe, tal que

$$F(c_p) \geq p. \quad (2.20)$$

En otras palabras, el número c_p es la cantidad más pequeña tal que la función de distribución acumula por lo menos una probabilidad p hasta ese valor.

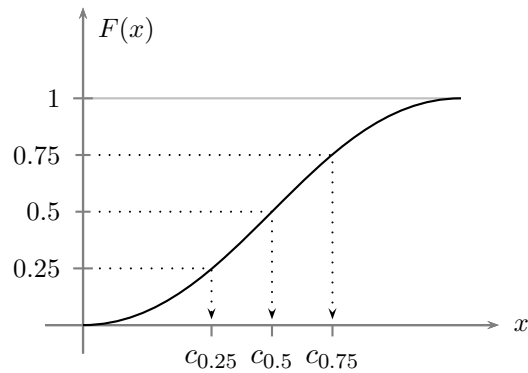


Figura 2.16

Observe que hemos omitido el caso $p = 0$ en la definición anterior pues, no existe un valor x más pequeño tal que $F(x) \geq 0$. Cuando la función de distribución es continua, la desigualdad (2.20) se reduce a la identidad

$$F(c_p) = p.$$

En la Figura 2.16 se ilustra de manera gráfica el concepto de cuantil en el caso continuo cuando p toma los valores 0.25, 0.5 y 0.75. Se acostumbra utilizar las siguientes expresiones: “ c_p es el cuantil de orden p ”, o “ c_p es el cuantil al $100p\%$ ” de la distribución. Así, por ejemplo,

$c_{0.1}$	es el cuantil al 10 %
$c_{0.2}$	es el cuantil al 20 %
...

Los cuantiles son usados con frecuencia en algunos procedimientos de la estadística. En particular, a las cantidades $c_{0.25}$, $c_{0.5}$, $c_{0.75}$ y c_1 se les llama cuartiles de la distribución. Más particularmente, al cuantil del 50 % se le llama mediana de la distribución. Debe observarse también que el número c_1 puede ser infinito, ello sucede cuando $F(x) < 1$ para cualquier valor finito de x .

Ejemplo 2.25 (Caso discreto) Sea X una v.a. con función de distribución $F(x)$ como se muestra en la Figura 2.17. Los siguientes son algunos ejemplos del cálculo de cuantiles.

$$\begin{aligned} c_{0.10} &= 1, & c_{0.50} &= 3, & c_{0.80} &= 4, \\ c_{0.20} &= 1, & c_{0.60} &= 3, & c_{0.85} &= 5, \\ c_{0.25} &= 2, & c_{0.75} &= 4, & c_{1.00} &= 5. \end{aligned}$$

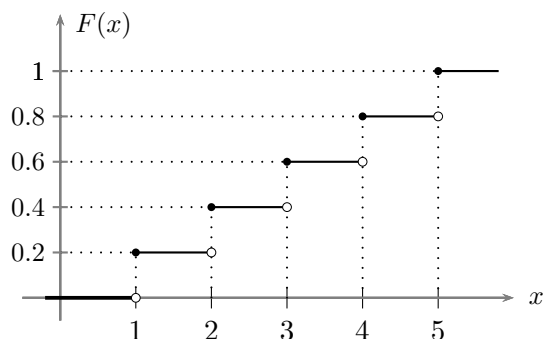


Figura 2.17

Puede definirse la función cuantil $p \mapsto c_p$ sobre el intervalo $(0, 1]$, es decir, a cada valor p en $(0, 1]$ se le asigna el cuantil c_p , en donde posiblemente el valor c_1 sea infinito. En particular, cuando la función de distribución $F(x)$ es continua y estrictamente creciente, entonces su inversa existe y por lo tanto la función cuantil es

$$c_p = F^{-1}(p).$$

En este caso la gráfica de la función $p \mapsto c_p$ puede obtenerse a partir de la gráfica de $F(x)$ al reflejar y rotar la figura de tal forma que el eje x ahora sea el eje vertical y el eje y sea el eje horizontal. Cuando no existe la función inversa de $F(x)$, puede definirse una función inversa generalizada. El aspecto importante es que al conocer un número mayor de cuantiles de una variable aleatoria, mayor información tenemos acerca de esta. Si conocemos

completamente la función cuantil, conocemos la distribución de la variable aleatoria.

Ejercicios

257. Calcule los cuantiles al 70 %, 80 % y 90 % para una variable aleatoria con función de probabilidad:

$$a) \begin{array}{c} \hline x \quad 0 \quad 1 \quad 2 \\ \hline f(x) \quad 1/2 \quad 1/4 \quad 1/4 \\ \hline \end{array}$$

$$b) \begin{array}{c} \hline x \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \\ \hline f(x) \quad 1/10 \quad 5/10 \quad 1/10 \quad 2/10 \quad 1/10 \\ \hline \end{array}$$

$$c) \begin{array}{c} \hline x \quad 1 \quad 2 \quad \dots \quad n \quad \dots \\ \hline f(x) \quad 1/2 \quad 1/4 \quad \dots \quad 1/2^n \quad \dots \\ \hline \end{array}$$

258. Calcule todos los cuantiles, si existen, de la función de distribución:

$$a) F(x) = \begin{cases} 0 & \text{si } x < -1, \\ 1/2 & \text{si } -1 \leq x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

$$b) F(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 & \text{si } x \geq 0. \end{cases}$$

$$c) F(x) = \begin{cases} 1 - e^{-x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

$$d) F(x) = \begin{cases} 0 & \text{si } x < 0, \\ x & \text{si } 0 \leq x \leq 1, \\ 1 & \text{si } x > 1. \end{cases}$$

$$e) F(x) = \begin{cases} 0 & \text{si } x \leq -a, \\ \frac{x+a}{2a} & \text{si } -a < x < a, \\ 1 & \text{si } x \geq a. \end{cases}$$

$$f) F(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - \cos x & \text{si } 0 \leq x \leq \pi/2, \\ 1 & \text{si } x > \pi/2. \end{cases}$$

259. Encuentre la distribución de la variable aleatoria discreta X con únicamente dos posibles valores y tal que satisface $c_{0,3} = 2$ y $c_1 = 4$.
260. Demuestre que si $p_1 \leq p_2$ son dos probabilidades estrictamente positivas, entonces $c_{p_1} \leq c_{p_2}$.

2.10. Moda

La moda representa otra característica numérica de las variables aleatorias o de sus distribuciones. Consideremos que X es una variable aleatoria con función de probabilidad $f(x)$. La moda de X o de su distribución es aquel número x^* , si existe, tal que la función $f(x^*)$ toma su valor máximo. Esta definición se aplica tanto en el caso continuo como en el caso discreto, pero en este último caso, la moda siempre existe y representa el punto de máxima probabilidad de la distribución.

Hemos definido a la moda como si tal punto x^* fuera único, sin embargo, en algunos casos, esto no es así. Se dice que la variable aleatoria o su distribución es unimodal en el caso cuando el punto x^* es único, en cambio, si existen dos o mas puntos x^* en donde $f(x^*)$ es máxima, entonces se dice que la distribución es multimodal. En cualquier caso, puede pensarse que la moda es en realidad un conjunto de puntos. Finalizamos esta breve sección señalando que no existe un símbolo estándar en la literatura para denotar a la moda de una distribución. Hemos usado el símbolo poco informativo x^* , pero también puede usarse la notación $\text{Moda}(X)$ o x_m .

Ejercicios

261. Calcule la moda (o modas) de las siguientes distribuciones de probabilidad:

$$a) f(x) = \begin{cases} 1-p & \text{si } x = 0, \\ p & \text{si } x = 1, \\ 0 & \text{en otro caso.} \end{cases} \quad (0 \leq p \leq 1)$$

$$b) f(x) = \begin{cases} \binom{5}{x} (1/2)^5 & \text{si } x = 0, 1, 2, 3, 4, 5, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f(x) = \begin{cases} x e^{-x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

$$d) F(x) = \begin{cases} 0 & \text{si } x < 0, \\ x & \text{si } 0 \leq x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

262. Sea X una variable aleatoria discreta con función de probabilidad $f(x)$. Demuestre que existe por lo menos un valor x^* tal que $f(x^*)$ es máxima. Esto no es necesariamente cierto en el caso continuo, por ejemplo, la distribución exponencial no tiene moda cuando se considera que la variable toma valores en el intervalo abierto $(0, \infty)$.

2.11. Función generadora de probabilidad

Sea X una variable aleatoria discreta con posibles valores dentro del conjunto $\{0, 1, 2, \dots\}$. Para este tipo de variables aleatorias vamos a asociar otra función equivalente a la función de probabilidad y a la función de distribución.

Definición 2.12 Sea X una variable aleatoria discreta con posibles valores dentro del conjunto $\{0, 1, 2, \dots\}$. A la función $G(t)$, definida como aparece abajo, se le llama la función generadora de probabilidad de X ,

$$G(t) = E(t^X) = \sum_{x=0}^{\infty} t^x P(X = x). \quad (2.21)$$

Observe que dicha función está definida por lo menos para valores reales de t dentro del intervalo $[-1, 1]$, pues en tal caso la suma que aparece en (2.21) es convergente. En forma breve, a esta función se le escribe como f.g.p. La letra G proviene del término “generadora” y, para indicar que la variable aleatoria X es la asociada, se le escribe también como $G_X(t)$. Así, la función $G(t)$ se define como una serie de potencias en t con coeficientes dados por los valores de la función de probabilidad. Estos coeficientes pueden reconstruirse nuevamente a partir de la expresión de la función $G(t)$ derivando y evaluando en cero, es decir, no es complicado demostrar el siguiente resultado.

Proposición 2.10 Sea X una variable aleatoria discreta con valores en el conjunto $\{0, 1, \dots\}$ y con función generadora de probabilidad $G(t)$. Para $x = 0, 1, \dots$

$$P(X = x) = \frac{1}{x!} G^{(x)}(0). \quad (2.22)$$

En donde $G^{(x)}(t)$ denota la derivada de orden x de $G(t)$. Esto justifica el nombre dado a esta función, pues a partir de ella se pueden generar las probabilidades de que la variable aleatoria tome sus distintos valores. La fórmula (2.22) establece también que toda la información de la distribución de probabilidad de la variable aleatoria se encuentra encapsulada en el comportamiento de $G(t)$ en $t = 0$. De esta manera, la f.g.p. proporciona una representación alterna y equivalente de la función de probabilidad. Véase el Ejercicio 265 en la página 194 para una breve lista de algunas otras propiedades de la f.g.p.

Veremos a continuación un ejemplo de la forma de encontrar la función generadora de probabilidad para una distribución concreta. Debe advertirse, sin embargo, que en ocasiones la función $G(t)$ no tiene una expresión corta o compacta como en el ejemplo y que, en general, no estaremos interesados en graficar la f.g.p. sino en utilizar sus propiedades analíticas.

Ejemplo 2.26 Considere la variable aleatoria discreta X con función de probabilidad

$$f(x) = \begin{cases} 1/2^x & \text{si } x = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

En este caso, la variable no toma valores enteros a partir de cero, sino a partir de uno, pero ello no es impedimento para poder calcular su f.g.p. Entonces

$$\begin{aligned} G(t) &= \sum_{x=1}^{\infty} t^x (1/2)^x \\ &= \sum_{x=1}^{\infty} (t/2)^x \\ &= \frac{t}{2-t} \quad \text{si } |t| < 2. \end{aligned}$$

Puede comprobarse que para valores naturales de x se cumple que

$$G^{(x)}(t) = 2x! (2-t)^{-x-1},$$

y se verifica la reconstrucción de la función de probabilidad

$$P(X = x) = 1/2^x = \frac{1}{x!} G^{(x)}(0).$$

•

El siguiente resultado nos provee de una fórmula para encontrar los momentos de una variable aleatoria a partir de su f.g.p., suponiendo la existencia de estos momentos. Para comprender mejor el enunciado, debemos recordar que la expresión $G(1-)$ se define como el límite de la función $G(t)$ cuando t se aproxima al valor 1 por la izquierda.

Proposición 2.11 Sea X una variable aleatoria discreta con función generadora de probabilidad $G(t)$. Si el k -ésimo momento de X existe entonces

$$G^{(k)}(1-) = E(X(X-1)\cdots(X-k+1)).$$

Demostración. Derivando k veces la serie de potencias (2.21) se tiene que

$$G^{(k)}(t) = \sum_{x=k}^{\infty} x(x-1)\cdots(x-k+1)t^{x-k}P(X=x).$$

Ahora se toma el límite cuando $t \nearrow 1$. El lema de Abel permite el intercambio de este límite con la suma infinita, obteniéndose así el resultado anunciado, es decir,

$$\begin{aligned} G^{(k)}(1-) &= \sum_{x=k}^{\infty} x(x-1)\cdots(x-k+1)P(X=x) \\ &= E(X(X-1)\cdots(X-k+1)). \end{aligned}$$

•

Más explícitamente, el resultado anterior dice, por ejemplo, que

$$\begin{aligned} G^{(1)}(1-) &= E(X), \\ G^{(2)}(1-) &= E(X(X-1)) = E(X^2) - E(X), \\ G^{(3)}(1-) &= E(X(X-1)(X-2)) = E(X^3) - 3E(X^2) + 2E(X), \\ &\vdots \end{aligned}$$

en donde puede apreciarse concretamente la dependencia entre el n -ésimo momento de la variable aleatoria y la n -ésima derivada de la f.g.p.

Ejemplo 2.27 En el Ejemplo 2.26 que aparece en la página 189 habíamos considerado una variable aleatoria discreta X con función de probabilidad

$$f(x) = \begin{cases} 1/2^x & \text{si } x = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

Se había encontrado que su f.g.p. es

$$G(t) = \frac{t}{2-t} \quad \text{para } |t| < 2,$$

y que entonces $G^{(x)}(t) = 2x!(2-t)^{-x-1}$. Podemos ahora encontrar los primeros momentos de esta variable aleatoria usando la f.g.p. y las relaciones:

$$\begin{aligned} E(X) &= G^{(1)}(1-) = 2, \\ E(X^2) - E(X) &= G^{(2)}(1-) = 4, \\ E(X^3) - 3E(X^2) + 2E(X) &= G^{(3)}(1-) = 12. \end{aligned}$$

De donde se obtiene $E(X) = 2$, $E(X^2) = 6$ y $E(X^3) = 26$. •

El siguiente resultado es bastante útil en las aplicaciones de la f.g.p. y establece que la f.g.p. de la suma de dos variables independientes es el producto de las f.g.p.

Proposición 2.12 Sean X y Y variables aleatorias discretas e independientes con funciones generadoras de probabilidad $G_X(t)$ y $G_Y(t)$. Entonces

$$G_{X+Y}(t) = G_X(t) G_Y(t).$$

Demostración. Usando la hipótesis de independencia tenemos que

$$\begin{aligned} G_{X+Y}(t) &= E(t^{X+Y}) \\ &= E(t^X t^Y) \\ &= E(t^X) E(t^Y) \\ &= G_X(t) G_Y(t). \end{aligned}$$

Es claro, a partir de la definición, que dos variables aleatorias con la misma distribución de probabilidad tienen asociada la misma f.g.p. Demostraremos a continuación que la relación es uno a uno, es decir, si se tienen dos variables aleatorias con la misma f.g.p., entonces éstas tienen la misma distribución

de probabilidad. Este resultado es muy importante pues establece que la f.g.p. caracteriza de manera única a la distribución de probabilidad, como se había sugerido antes.

Proposición 2.13 (Caracterización). Sean X y Y dos variables aleatorias discretas con el mismo conjunto de valores $\{0, 1, 2, \dots\}$ y con funciones generadoras de probabilidad $G_X(t)$ y $G_Y(t)$, tales que $G_X(t) = G_Y(t)$ para $t \in (-s, s)$ con $s > 0$. Entonces X y Y tienen la misma distribución de probabilidad.

Demostración. Supongamos que $G_X(t) = G_Y(t)$ para $t \in (-s, s)$ con $s > 0$. Esto significa que las dos series de potencias son iguales en dicho intervalo. Substrayendo una serie de otra se obtiene una serie idénticamente cero, es decir, para cada $t \in (-s, s)$,

$$\sum_{x=0}^{\infty} [P(X = x) - P(Y = x)] t^x = 0.$$

Esto sólo es posible cuando los coeficientes de la serie son todos cero, es decir, para cada $x = 0, 1, \dots$ se tiene que

$$P(X = x) = P(Y = x).$$

•

Ejemplo 2.28 Se dice que la v.a. X tiene distribución Poisson de parámetro $\lambda > 0$ si su función de probabilidad está dada por

$$f(x) = \begin{cases} e^{-\lambda} \frac{\lambda^x}{x!} & \text{si } x = 0, 1, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

Esto es, para cualquier valor real y fijo del parámetro $\lambda > 0$, esta función es una función de probabilidad. En el siguiente capítulo estudiaremos con más detalle esta distribución. Por ahora nuestro objetivo es calcular su f.g.p.

$$\begin{aligned}
G(t) &= E(t^X) \\
&= \sum_{x=0}^{\infty} t^x e^{-\lambda} \frac{\lambda^x}{x!} \\
&= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(t\lambda)^x}{x!} \\
&= e^{-\lambda} e^{t\lambda} \\
&= \exp\{-\lambda(1-t)\}.
\end{aligned}$$

Así, debido a lo demostrado antes relativo a la correspondencia uno a uno entre las distribuciones de probabilidad y las f.g.p., sabemos que esta función es la f.g.p. de la distribución Poisson y que cualquier v.a. discreta con f.g.p. de esta forma tiene distribución Poisson. Usaremos este hecho para demostrar con facilidad que la suma de dos variables aleatorias independientes con distribución Poisson tiene nuevamente distribución Poisson. Sean X y Y dos variables aleatorias independientes con distribución Poisson de parámetros λ_1 y λ_2 , respectivamente. Entonces, por independencia,

$$\begin{aligned}
G_{X+Y}(t) &= G_X(t) G_Y(t) \\
&= \exp\{-\lambda_1(1-t)\} \exp\{-\lambda_2(1-t)\} \\
&= \exp\{-(\lambda_1 + \lambda_2)(1-t)\}.
\end{aligned}$$

Observe que esta última expresión tiene la forma de la f.g.p. de la distribución Poisson, sólo que en lugar del parámetro λ aparece la expresión $\lambda_1 + \lambda_2$. Esto indica que $X + Y$ tiene distribución Poisson de parámetro $\lambda_1 + \lambda_2$. •

En los siguientes capítulos haremos uso de la f.g.p. y sus propiedades para caracterizar a algunas distribuciones de probabilidad específicas.

Ejercicios

263. Sea X una variable aleatoria discreta con valores en el conjunto $\{0, 1, \dots\}$ y con función generadora de probabilidad $G(t)$. Demuestre que para

$$x = 0, 1, \dots$$

$$P(X = x) = \frac{1}{x!} G^{(x)}(0).$$

264. Sea X una variable aleatoria discreta con función de probabilidad como aparece abajo. Encuentre la función generadora de probabilidad de X y, a partir de ella, encuentre la media y la varianza de X .

$$a) f(x) = \begin{cases} 2/3^x & \text{si } x = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x) = \begin{cases} 3(1/2)^{2x} & \text{si } x = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

265. Sea X una v.a. discreta con valores en el conjunto $\{0, 1, \dots\}$ y con f.g.p. $G(t)$. Demuestre las siguientes propiedades de la función generadora de probabilidad. Recuerde que $G(1-)$ se define como el límite de la función $G(t)$ cuando t se aproxima al valor 1 por la izquierda. Véase además el lema de Abel que aparece en el apéndice.

$$a) G(1-) = 1.$$

b) Si X tiene esperanza finita entonces

$$E(X) = G^{(1)}(1-).$$

c) Si X tiene varianza finita entonces

$$\text{Var}(X) = G^{(2)}(1-) + G^{(1)}(1-) - [G^{(1)}(1-)]^2.$$

266. Sea X una variable aleatoria con función de probabilidad como aparece abajo. Demuestre que la f.g.p. de X es $G(t) = (1 - t^n)/(n(1 - t))$.

$$f(x) = \begin{cases} 1/n & \text{si } x = 0, 1, \dots, n-1, \\ 0 & \text{en otro caso.} \end{cases}$$

2.12. Función generadora de momentos

Otra función bastante útil que puede calcularse para algunas variables aleatorias, ahora incluyendo por igual el caso discreto y continuo, y que está relacionada con los momentos de la variable aleatoria, es la que se define a continuación.

Definición 2.13 La función generadora de momentos de una variable aleatoria discreta o continua X es la función $M(t)$ definida como sigue

$$M(t) = E(e^{tX}),$$

para valores reales de t en donde esta esperanza existe.

En forma breve se le escribe como f.g.m. La letra M corresponde al término “momentos” y en breve justificaremos su relación con los momentos de la variable aleatoria. Cuando sea necesario especificar la variable aleatoria en cuestión se le escribe también como $M_X(t)$. Así, en el caso discreto esta función se calcula como

$$M(t) = \sum_x e^{tx} P(X = x), \quad (2.23)$$

y en el caso continuo

$$M(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx, \quad (2.24)$$

para valores reales de t en donde esta suma o integral sea finita. Los resultados que veremos en esta sección tienen la finalidad de mostrar algunas propiedades que posee la f.g.m. y que la hacen una herramienta muy atractiva para resolver algunos problemas de probabilidad. Lamentablemente, no todas las variables aleatorias tienen asociada una función generadora de momentos. Veamos un ejemplo del cálculo de esta función.

Ejemplo 2.29 Considere la variable continua X con función de densidad

$$f(x) = \frac{1}{2} e^{-|x|}, \quad x \in \mathbb{R}.$$

La función generadora de momentos de esta variable aleatoria se calcula de la siguiente forma:

$$\begin{aligned}
 M(t) &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{2} e^{-|x|} dx \\
 &= \frac{1}{2} \int_0^{\infty} e^{tx} e^{-x} dx + \frac{1}{2} \int_{-\infty}^0 e^{tx} e^x dx \\
 &= \frac{1}{2} \int_0^{\infty} e^{(t-1)x} dx + \frac{1}{2} \int_{-\infty}^0 e^{(t+1)x} dx \\
 &= \frac{1}{2} \frac{1}{(t-1)} e^{(t-1)x} \Big|_0^{\infty} + \frac{1}{2} \frac{1}{(t+1)} e^{(t+1)x} \Big|_{-\infty}^0 \\
 &= \frac{1}{2} \frac{1}{(t-1)} (-1) + \frac{1}{2} \frac{1}{(t+1)} (1) \quad \text{si } -1 < t < 1, \\
 &= \frac{1}{1-t^2} \quad \text{si } -1 < t < 1.
 \end{aligned}$$

Esta es entonces la función generadora de momentos de una variable aleatoria con la distribución indicada. Observe que su dominio de definición no es la totalidad de números reales, sino únicamente el intervalo $(-1, 1)$. •

En el siguiente capítulo veremos otros ejemplos de funciones generadoras de momentos para distribuciones particulares de interés. Reiteramos que la suma (2.23) o integral (2.24) pueden no ser convergentes para ningún valor de t distinto de cero, y en tales casos decimos que la variable aleatoria no tiene función generadora de momentos finita. Véase el Ejercicio 270 para un ejemplo de esta situación.

A partir del siguiente primer resultado, varias de las propiedades de la f.g.m. pueden encontrarse.

Proposición 2.14 Sea X una variable aleatoria con función generadora de momentos $M(t)$, definida para valores de t en el intervalo $(-s, s)$ para algún $s > 0$. Entonces todos los momentos de X existen y $M(t)$ adquiere la forma de la siguiente serie de potencias

$$M(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} E(X^n). \quad (2.25)$$

Demostración. La serie de Taylor de la función exponencial desarrollada alrededor del cero y evaluada en tX para $t \in (-s, s)$ es

$$e^{tX} = \sum_{n=0}^{\infty} \frac{(tX)^n}{n!}.$$

Tomando esperanza y suponiendo válido el intercambio de esperanza y suma infinita, tenemos que

$$\begin{aligned} M(t) &= E(e^{tX}) \\ &= E\left(\sum_{n=0}^{\infty} \frac{(tX)^n}{n!}\right) \\ &= \sum_{n=0}^{\infty} \frac{t^n}{n!} E(X^n). \end{aligned}$$

La finitud de esta suma para cada valor de t en el intervalo $(-s, s)$ implica que $E(X^n) < \infty$ para todo n natural. •

Por lo tanto, $M(t)$ tiene derivadas continuas de cualquier orden en el intervalo $(-s, s)$, y en consecuencia, se tiene el siguiente resultado que justifica el nombre para esta función.

Proposición 2.15 Sea X una v.a. con función generadora de momentos $M(t)$ finita en un intervalo $(-s, s)$ con $s > 0$. Entonces para cada $n = 0, 1, \dots$

$$\lim_{t \rightarrow 0} M^{(n)}(t) = E(X^n).$$

Demostración. Derivando término a término la expansión (2.25), es inmediato comprobar que para cualquier entero $n \geq 0$,

$$M^{(n)}(t) = E(X^n) + t E(X^{n+1}) + \frac{t^2}{2!} E(X^{n+2}) + \dots$$

Por lo tanto,

$$\lim_{t \rightarrow 0} M^{(n)}(t) = E(X^n).$$

•

Es decir, los momentos de X se encuentran derivando la f.g.m. y tomando el límite cuando $t \rightarrow 0$. De este hecho surge el nombre de esta función. Demostraremos ahora que la f.g.m. de la suma de dos variables aleatorias independientes es el producto de las funciones generadoras.

Proposición 2.16 Sean X y Y variables aleatorias independientes con funciones generadoras de momentos $M_X(t)$ y $M_Y(t)$. Entonces

$$M_{X+Y}(t) = M_X(t) M_Y(t).$$

Demostración. Usando la hipótesis de independencia tenemos que

$$\begin{aligned} M_{X+Y}(t) &= E(e^{t(X+Y)}) \\ &= E(e^{tX} e^{tY}) \\ &= E(e^{tX}) E(e^{tY}) \\ &= M_X(t) M_Y(t). \end{aligned}$$

•

La f.g.m. también tiene la propiedad de caracterizar a la distribución de probabilidad de manera única. Este es el contenido del siguiente resultado, cuya demostración no es sencilla y la omitiremos.

Proposición 2.17 (Caracterización). Sean X y Y dos variables aleatorias con f.g.m. $M_X(t)$ y $M_Y(t)$, las cuales coinciden en un intervalo $(-s, s)$, con $s > 0$. Entonces X y Y tienen la misma distribución de probabilidad.

Más aún, las funciones generadoras de momentos cumplen con la siguiente propiedad importante, que usaremos en el último capítulo para demostrar algunos teoremas límite.

Teorema 2.1 (Continuidad de la f.g.m.). Sea X_1, X_2, \dots una sucesión de variables aleatorias tal que X_n tiene f.g.m. $M_{X_n}(t)$. Sea X otra variable aleatoria con f.g.m. $M_X(t)$. Si sucede que para cada $t \in (-s, s)$, con $s > 0$,

$$\lim_{n \rightarrow \infty} M_{X_n}(t) = M_X(t),$$

entonces en cada punto x en donde $F_X(x)$ es continua se cumple que

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x). \quad (2.26)$$

A la propiedad establecida en (2.26) se le llamará convergencia en distribución de la sucesión de variables X_n a la variable X , y esto lo escribiremos como $X_n \xrightarrow{d} X$. En el último capítulo de este texto se encuentra una sección introductoria al tema de convergencia de variables aleatorias. Las demostraciones de estos dos últimos resultados más avanzados pueden encontrarse en el texto de Gut [9].

Ejercicios

267. Sea X una variable aleatoria discreta con función de probabilidad como aparece abajo. Encuentre la función generadora de momentos de X y, a partir de ella, calcule la media y la varianza de X .

$$a) f(x) = \begin{cases} 1/2^x & \text{si } x = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x) = \begin{cases} 2/3^x & \text{si } x = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

268. Sea X una variable aleatoria continua con función de densidad como aparece abajo. Encuentre la función generadora de momentos de X y, a partir de ella, calcule la media y la varianza de X .

$$a) f(x) = \begin{cases} 2x & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x) = \begin{cases} e^{-x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f(x) = \frac{1}{2} e^{-|x|}.$$

$$d) f(x) = \begin{cases} 1 - |x| & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$e) f(x) = \begin{cases} 6x(1-x) & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$f) f(x) = \begin{cases} 1/4 & \text{si } 2 \leq x \leq 4, \\ 1/2 & \text{si } 4 < x \leq 5, \\ 0 & \text{en otro caso.} \end{cases}$$

269. Sea X una variable aleatoria con f.g.m. $M_X(t)$ y sean a y b dos constantes. Demuestre que

$$M_{aX+b}(t) = e^{bt} M_X(at).$$

270. **No existencia de la f.g.m.** Se dice que la variable aleatoria X tiene una distribución t con $n = 1$ grados de libertad si su función de densidad es como aparece abajo. Demuestre que para esta distribución no existe su f.g.m.

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad x \in \mathbb{R}.$$

Capítulo 3

Distribuciones de probabilidad

Estudiaremos ahora algunas distribuciones de probabilidad particulares. Las distribuciones que mencionaremos tienen un nombre particular adquirido ya sea debido a la situación en la que surgen, o bien debido al nombre de su descubridor o a la persona que inicialmente la utilizó en alguna aplicación importante. Empezaremos con las distribuciones de tipo discreto y continuaremos con las de tipo continuo. Principalmente en este último caso omitiremos especificar el experimento aleatorio y el espacio de probabilidad en donde pueden definirse estas variables aleatorias y sus distribuciones.

Nota importante. El lector debe siempre recordar que no existe homogeneidad en la literatura acerca de la forma de escribir los parámetros de una distribución de probabilidad dada. Por lo tanto, se debe tener siempre cuidado al comparar fórmulas y resultados de una fuente bibliográfica a otra y también verificar la forma en la que los parámetros de una distribución particular son usados en sus implementaciones en los distintos lenguajes de programación y paquetes computacionales. Como se señaló antes, el objetivo es concentrarnos en el modelo matemático y sus propiedades.

3.1. Distribución uniforme discreta

Decimos que una variable aleatoria X tiene una distribución uniforme discreta sobre el conjunto de n números $\{x_1, \dots, x_n\}$ si la probabilidad de que X tome cualquiera de estos valores es constante $1/n$. Esta distribución surge en espacios de probabilidad equiprobables, esto es, en situaciones en donde tenemos n resultados diferentes y todos ellos tienen la misma probabilidad de ocurrir. Los juegos de lotería son un ejemplo donde puede aplicarse esta distribución de probabilidad. Se escribe $X \sim \text{unif}\{x_1, \dots, x_n\}$, en donde el símbolo “ \sim ” se lee “se distribuye como” o “tiene una distribución”. La función de probabilidad de esta variable aleatoria es

$$f(x) = \begin{cases} 1/n & \text{si } x = x_1, \dots, x_n, \\ 0 & \text{en otro caso.} \end{cases}$$

Es inmediato comprobar que la esperanza y la varianza para esta distribución se calculan del siguiente modo:

- a) $E(X) = \frac{1}{n} \sum_{i=1}^n x_i = \mu$.
 b) $\text{Var}(X) = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$.

Algunas otras propiedades de esta distribución se encuentran en la sección de ejercicios. Veamos ahora algunos ejemplos.

Ejemplo 3.1 La gráfica de la función de probabilidad de la distribución uniforme en el conjunto $\{1, 2, 3, 4, 5\}$ aparece en la Figura 3.1, junto con la correspondiente función de distribución. Cada salto en la función de distribución es de tamaño $1/5$. La expresión completa de $F(x)$ es la siguiente:

$$F(x) = \begin{cases} 0 & \text{si } x < 1, \\ 1/5 & \text{si } 1 \leq x < 2, \\ 2/5 & \text{si } 2 \leq x < 3, \\ 3/5 & \text{si } 3 \leq x < 4, \\ 4/5 & \text{si } 4 \leq x < 5, \\ 1 & \text{si } x \geq 5. \end{cases}$$

•

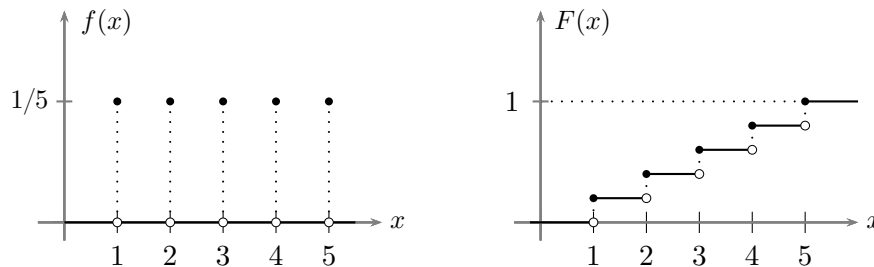


Figura 3.1

Ejemplo 3.2 Al generar un número aleatorio en una computadora dentro del intervalo unitario $[0, 1]$ y debido a que la precisión de la computadora es necesariamente finita, se obtienen siempre valores dentro de un conjunto finito de elementos. Por ejemplo, si la precisión de la computadora es de dos decimales, entonces sólo se pueden generar los números : 0.00, 0.01, 0.02, ..., 0.99, 1.00. La precisión de una computadora actual es claramente mucho mayor a la considerada, pero siempre es finita y algún grado de imprecisión prevalece, es decir, en términos prácticos se tiene una distribución uniforme discreta al generar un valor al azar en computadora dentro del intervalo $[0, 1]$. •

Simulación 3.1 En el paquete computacional R se pueden generar valores al azar de la distribución uniforme discreta usando el siguiente comando.

```
# 15 valores al azar de la distribución unif{1,...,10}
> sample(1:10,15,replace=TRUE)
[1] 7 3 4 1 1 3 2 1 6 5 3 8 2 9 1
```

Simulación 3.2 En el lenguaje de programación Python se puede crear una lista de elementos, y mediante la función predefinida `choice()` se escoge

un elemento al azar de la lista con distribución uniforme. El código aparece en el siguiente recuadro y procedimientos similares existen o pueden definirse en otros lenguajes de programación.

```
>>> import random
>>> conjunto=[1,2,3,4,5]
>>> random.choice(conjunto)
3
```

Ejercicios

271. Sea X una variable aleatoria con distribución uniforme en el conjunto $\{1, \dots, n\}$ y sean x, x_1, x_2 números dentro de este conjunto en donde $x_1 < x_2$. Encuentre las siguientes probabilidades.

- | | |
|-------------------------------|----------------------------|
| a) $P(X \leq x)$. | d) $P(x_1 < X \leq x_2)$. |
| b) $P(X \geq x)$. | e) $P(x_1 \leq X < x_2)$. |
| c) $P(x_1 \leq X \leq x_2)$. | f) $P(x_1 < X < x_2)$. |

272. Sea X una v.a. con distribución uniforme en el conjunto $\{-1, 0, 1\}$. Demuestre que las variables aleatorias X^3 y $-X$ tienen la misma distribución que X . Esto muestra claramente que puede haber varias variables aleatorias con la misma distribución y no ser iguales.

273. Sea X una v.a. con distribución uniforme en el conjunto $\{1, \dots, n\}$. Demuestre que:

- $E(X) = (n + 1)/2$.
- $E(X^2) = (n + 1)(2n + 1)/6$.
- $\text{Var}(X) = (n^2 - 1)/12$.

274. **Momentos.** Sea X una v.a con distribución unif $\{0, 1\}$. Demuestre que el n -ésimo momento de X es

$$E(X^n) = 1/2.$$

275. **Cuartiles** Sea n un número natural. Encuentre los cuatro cuartiles de la distribución $\text{unif}\{1, \dots, 4n\}$.

276. **f.g.p.** Sea X una v.a. con distribución $\text{unif}\{1, \dots, n\}$. Demuestre que la f.g.p. de X está dada por la expresión que aparece abajo. A través de esta función encuentre nuevamente la esperanza y la varianza de esta distribución como aparecen en el Ejercicio 273.

$$G(t) = \frac{t(1 - t^n)}{n(1 - t)}.$$

277. **f.g.m.** Sea X una v.a. con distribución $\text{unif}\{1, \dots, n\}$. Demuestre que la f.g.m. de X está dada por la expresión que aparece abajo. A través de esta función encuentre nuevamente la esperanza y la varianza de esta distribución como aparecen en el Ejercicio 273.

$$M(t) = \frac{e^t(1 - e^{nt})}{n(1 - e^t)}.$$

278. **Simulación.** Este es un mecanismo para generar valores al azar de una v.a. con distribución $\text{unif}\{x_1, \dots, x_n\}$ a partir de valores de una v.a. con distribución $\text{unif}(0, 1)$, la cual aparece definida más adelante. Sea u un valor al azar con distribución $\text{unif}(0, 1)$. Demuestre que la variable aleatoria X , definida a continuación, tiene distribución $\text{unif}\{x_1, \dots, x_n\}$.

$$X = \begin{cases} x_1 & \text{si } 0 < u \leq 1/n, \\ x_2 & \text{si } 1/n < u \leq 2/n, \\ \dots & \dots\dots\dots \\ x_{n-1} & \text{si } (n-2)/n < u \leq (n-1)/n, \\ x_n & \text{si } (n-1)/n < u < 1. \end{cases}$$

279. Se escogen al azar y de manera independiente dos números a y b , dentro del conjunto $\{1, \dots, 10\}$. Calcule la probabilidad de que

a) a y b coincidan.

b) a sea menor a b .

- c) a sea mayor a $b + 1$.
 - d) a y b difieran en por lo menos 2 unidades.
280. Un juego de ruleta consiste de 36 números. Un jugador apuesta, repetidas veces, a que el resultado será un número entre el 1 y el 12 inclusive, es decir, no apuesta por un número particular sino por el conjunto de estos números. Calcule la probabilidad de que el jugador
- a) pierda sus primeras cinco apuestas.
 - b) gane por primera vez en su cuarta apuesta.
 - c) gane k de n apuestas ($1 \leq k \leq n$).
281. Sea X una v.a. con distribución uniforme en el conjunto $\{1, 2, 3, 4, 5\}$. ¿Cuál es la probabilidad de que el área del rectángulo de lados X y $6 - X$ sea mayor o igual a 8?
282. Sean m y n dos números naturales tales que $m \leq n$ y sea X una variable aleatoria con distribución $\text{unif}\{1, \dots, n\}$. Encuentre la función de probabilidad de la variable aleatoria:
- a) $U = \text{mín}\{X, m\}$.
 - b) $V = \text{máx}\{X, m\}$.

3.2. Distribución Bernoulli

Un ensayo Bernoulli¹ se define como aquel experimento aleatorio con únicamente dos posibles resultados, llamados genéricamente: éxito y fracaso. Supondremos que las probabilidades de estos resultados son p y $1 - p$, respectivamente. Si se define la variable aleatoria X como aquella función que lleva el resultado éxito al número 1 y el resultado fracaso al número 0, entonces decimos que X tiene una distribución Bernoulli con parámetro $p \in (0, 1)$ y escribimos $X \sim \text{Ber}(p)$. La función de probabilidad se puede escribir de

¹Jacob (James o Jacques) Bernoulli (1654-1705), matemático suizo.

la siguiente forma.

$$f(x) = \begin{cases} 1 - p & \text{si } x = 0, \\ p & \text{si } x = 1, \\ 0 & \text{en otro caso.} \end{cases}$$

O bien de manera compacta,

$$f(x) = \begin{cases} p^x(1-p)^{1-x} & \text{si } x = 0, 1, \\ 0 & \text{en otro caso.} \end{cases}$$

La gráfica de esta función de probabilidad para $p = 0.7$ aparece en la Figura 3.2 junto con la correspondiente función de distribución, la cual tiene la siguiente forma:

$$F(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - p & \text{si } 0 \leq x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

La función de probabilidad $f(x)$ puede obtenerse en el paquete **R** usando el comando `dbinom(x,n,p)`, como se muestra en el recuadro de abajo, en donde **x** es el valor en donde se desea evaluar la función, **n** se substituye por el valor 1 y **p** es el parámetro de la distribución. El nombre asignado a este comando y sus argumentos serán justificados una vez que estudiemos la distribución binomial, pues resulta que la distribución Bernoulli es un caso particular de la distribución binomial.

```
# dbinom(x,1,p) evalúa f(x) de la distribución Ber(p)
> dbinom(0,1,0.7)
[1] 0.3
```

Para la función de distribución $F(x)$ se usa el siguiente comando.

```
# pbinom(x,1,p) evalúa F(x) de la distribución Ber(p)
> pbinom(0.2,1,0.7)
[1] 0.3
```

Para la distribución Bernoulli de parámetro p es inmediato verificar que:

a) $E(X) = p$.

b) $\text{Var}(X) = p(1 - p)$.

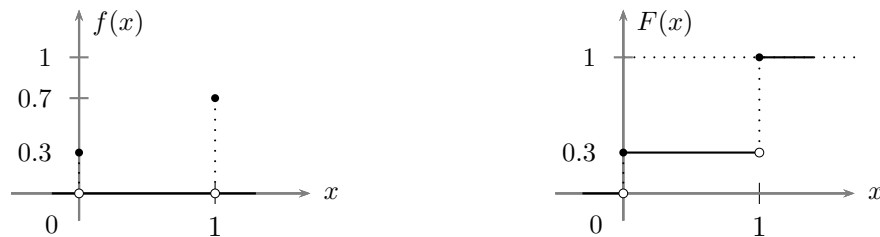


Figura 3.2

En la realización de todo experimento aleatorio siempre es posible preguntarse por la ocurrencia o no ocurrencia de un evento cualquiera. Este es el esquema general donde surge esta distribución de probabilidad. La distribución Bernoulli es sencilla, pero de muy amplia aplicación, como veremos más adelante.

Ejemplo 3.3 Sea Ω el espacio muestral de un experimento aleatorio y sea A un evento con probabilidad $p > 0$. Sea X la variable aleatoria dada por

$$X(\omega) = \begin{cases} 1 & \text{si } \omega \in A, \\ 0 & \text{si } \omega \notin A. \end{cases}$$

Entonces X tiene distribución $\text{Ber}(p)$. A esta variable aleatoria X se le llama la función indicadora del evento A y se le denota también por $1_A(\omega)$. Así, al efectuar un ensayo del experimento aleatorio, la función indicadora señala la ocurrencia del evento A tomando el valor 1, e indica que no ha ocurrido dicho evento tomando el valor 0. •

Ejemplo 3.4 Considere el experimento aleatorio de lanzar una moneda al aire. Suponga que ω_0 y ω_1 son los dos resultados posibles, con probabilidades $1 - p$ y p , respectivamente. Sea X la variable aleatoria dada por

$$\begin{aligned} X(\omega_0) &= 0, \\ \text{y } X(\omega_1) &= 1. \end{aligned}$$

Entonces X tiene distribución $\text{Ber}(p)$. ¿Puede usted encontrar la distribución de la variable $Y = 1 - X$? •

Simulación 3.3 En el paquete **R** se pueden generar k valores al azar de la distribución $\text{Ber}(p)$ usando el comando que aparece en el siguiente recuadro. Como un ejercicio de simulación, modifique los valores de los parámetros y obtenga tantos valores al azar de esta distribución como desee. ¿Cuál es la frecuencia relativa con la que aparecen los dos valores en su simulación?

```
# rbinom(k,1,p) genera k valores al azar de la distribución Ber(p)
> rbinom(25,1,0.7)
[1] 1 0 1 1 1 1 1 1 0 1 1 0 0 1 1 1 0 0 0 1 1 1 0 0 1
```

Ejercicios

283. Sea X una v.a. con distribución $\text{Ber}(p)$ y sea n un número natural. Encuentre la distribución de:

- X^n .
- $(1 - X)^n$.
- $|X - 1|^n$.

284. Sea X una v.a. con distribución $\text{Ber}(p)$ y sean a y b dos constantes con $a \neq 0$. Defina la variable aleatoria $Y = aX + b$. Encuentre

- la función de probabilidad de Y .

- b) $E(Y)$.
- c) $\text{Var}(X)$.
- d) $E(Y^n)$ para $n = 1, 2, \dots$

285. **Momentos.** Sea X una v.a. con distribución $\text{Ber}(p)$. Demuestre que el n -ésimo momento de X es constante,

$$E(X^n) = p.$$

286. Sea X una v.a. con distribución $\text{Ber}(p)$. Encuentre el valor del parámetro p que maximiza la varianza de X .

287. **Cuartiles.** Encuentre todos los cuartiles de la distribución $\text{Ber}(p)$ con $p = 1/2$.

288. **f.g.p.** Sea X una v.a. con distribución $\text{Ber}(p)$. Demuestre que la f.g.p. de X está dada por la expresión que aparece abajo. A través de esta función encuentre nuevamente la esperanza y la varianza de esta distribución.

$$G(t) = 1 - p + pt.$$

289. **f.g.m.** Sea X una v.a. con distribución $\text{Ber}(p)$. Demuestre que la f.g.m. de X está dada por la expresión que aparece abajo. A través de esta función encuentre nuevamente la esperanza y la varianza de esta distribución.

$$M(t) = 1 - p + pe^t.$$

290. **Simulación.** Este es un mecanismo para generar valores al azar de una v.a. con distribución $\text{Ber}(p)$ a partir de valores de una v.a. con distribución $\text{unif}(0, 1)$. Sea u un valor al azar con distribución $\text{unif}(0, 1)$. Demuestre que la variable aleatoria X , definida a continuación, tiene distribución $\text{Ber}(p)$.

$$X = \begin{cases} 0 & \text{si } 0 < u \leq 1 - p, \\ 1 & \text{si } 1 - p < u < 1. \end{cases}$$

291. Sean X y Y dos v.a.s independientes con idéntica distribución $\text{Ber}(p)$. Encuentre la distribución de:

- | | |
|--------------|------------------|
| a) $X + Y$. | d) $X(1 - Y)$. |
| b) $X - Y$. | e) $X(1 - X)$. |
| c) XY . | f) $X + Y - 1$. |

292. Sean X_1, \dots, X_n v.a.i.i.d. con distribución $\text{Ber}(p)$. Encuentre la distribución de la variable aleatoria producto

$$X_1 \cdots X_n.$$

3.3. Distribución binomial

Supongamos que efectuamos una serie de n ensayos independientes Bernoulli en donde la probabilidad de éxito en cada ensayo es p . Si denotamos por E el resultado éxito y por F el resultado fracaso, entonces el espacio muestral de este experimento consiste de todas las posibles sucesiones de longitud n de caracteres E y F . Así, el espacio muestral consiste de 2^n elementos. Si ahora definimos la variable aleatoria X como aquella función que indica el número de éxitos en cada una de estas sucesiones, esto es,

$$\begin{aligned} X(EE \cdots EE) &= n, \\ X(FE \cdots EE) &= n - 1, \\ &\vdots \\ X(FF \cdots FF) &= 0, \end{aligned}$$

entonces tenemos que X puede tomar los valores $0, 1, 2, \dots, n$, con las probabilidades dadas por la función de probabilidad:

$$f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{si } x = 0, 1, \dots, n, \\ 0 & \text{en otro caso.} \end{cases}$$

Decimos entonces que X tiene una distribución binomial con parámetros n y p , y escribimos $X \sim \text{bin}(n, p)$. Esta expresión para la función de probabilidad puede obtenerse de la siguiente forma: la probabilidad de obtener x

éxitos y $n - x$ fracasos en n ensayos Bernoulli es, preliminarmente,

$$\underbrace{p \cdots p}_x \underbrace{(1-p) \cdots (1-p)}_{n-x} = p^x (1-p)^{n-x},$$

pero hemos colocado los x éxitos en los primeros ensayos, cuando ello no ocurrirá necesariamente así. Las diferentes formas en que los x éxitos pueden distribuirse en los n ensayos está dada por el coeficiente binomial $\binom{n}{x}$. Por ejemplo, hay $\binom{n}{n} = 1$ manera de obtener n éxitos en n ensayos, hay $\binom{n}{n-1} = n$ formas diferentes de obtener $n - 1$ éxitos en n ensayos, etcétera. Al hacer la multiplicación de este coeficiente binomial con el término $p^x (1-p)^{n-x}$ se obtiene la expresión de la función de probabilidad para esta distribución.

La evaluación de la función $f(x)$ puede representar un reto desde el punto de vista numérico, pues se ven involucradas varias multiplicaciones, particularmente cuando n es grande. En el paquete **R**, esta función de probabilidad se obtiene usando el siguiente comando.

```
# dbinom(x,n,p) evalúa f(x) de la distribución bin(n,p)
> dbinom(8,10,0.3)
[1] 0.001446701
```

Ejemplo 3.5 Cuando el número de ensayos es $n = 10$ y la probabilidad de éxito es $p = 0.3$, se puede calcular, por ejemplo,

$$P(X = 2) = \binom{10}{2} (0.3)^2 (0.7)^{10-2} = 0.2334,$$

y de manera análoga, el resto de las probabilidades. La gráfica de esta función de probabilidad, con los parámetros n y p indicados, se muestra en la Figura 3.3.

•

La función de distribución $F(x)$ se escribe, simplemente, como la suma de los valores $f(u)$ para valores de u menores o iguales a x , pero esta fórmula no tiene una expresión reducida y por ello no la escribiremos. Los valores

x	$f(x)$
0	0.0282475
1	0.1210608
2	0.2334744
3	0.2668279
4	0.2001209
5	0.1029193
6	0.0367569
7	0.0090016
8	0.0014467
9	0.0001377
10	0.0000059

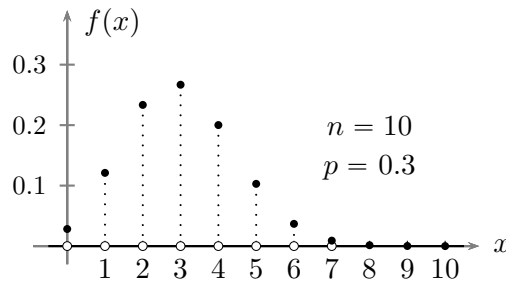


Figura 3.3

de esta función se pueden obtener en R de la siguiente forma:

```
# pbinom(x,n,p) evalúa  $F(x)$  de la distribución  $\text{bin}(n, p)$ 
> pbinom(4,10,0.3)
[1] 0.8497317
```

Por otro lado, después de algunos cálculos puede demostrarse que para una variable X con distribución $\text{bin}(n, p)$,

- a) $E(X) = np$.
- b) $\text{Var}(X) = np(1 - p)$.

Es instructivo observar que cuando el parámetro n en la distribución $\text{bin}(n, p)$ toma el valor 1 se obtiene la distribución $\text{Ber}(p)$. El siguiente resultado es muy útil y se puede demostrar por separado o bien considerarse como una consecuencia de la forma en la que se ha definido la distribución binomial.

Proposición 3.1 Sean X_1, \dots, X_n variables aleatorias independientes, cada una con distribución $\text{Ber}(p)$. Entonces

$$X_1 + \dots + X_n \sim \text{bin}(n, p). \quad (3.1)$$

Recíprocamente, toda v.a. con distribución $\text{bin}(n, p)$ puede ser expresada como una suma de la forma anterior.

Así, cada sumando toma el valor 1 o 0, dependiendo si el ensayo correspondiente fue éxito o fracaso, y la suma indica el número total de éxitos en los n ensayos. Aplicando las propiedades de la esperanza y la varianza en esta suma de variables aleatorias se pueden encontrar de forma más directa la esperanza y la varianza de la distribución binomial.

Simulación 3.4 La expresión (3.1) sugiere un mecanismo para generar valores al azar de la distribución $\text{bin}(n, p)$. Si se generan de manera independiente n valores al azar de la distribución $\text{Ber}(p)$ y se suman estos valores, se obtiene un valor al azar de la distribución $\text{bin}(n, p)$. •

Simulación 3.5 En el paquete **R** se pueden obtener valores al azar de la distribución $\text{bin}(n, p)$ usando el comando que se muestra en el siguiente recuadro. Asigne valores a los parámetros correspondientes y genere tantos valores al azar como desee. Calcule el promedio aritmético de los valores obtenidos y compare con el valor de np . ¿Son parecidos?

```
# rbinom(k,n,p) genera k valores al azar de la distribución
# bin(n,p)
> rbinom(25,10,0.3)
[1] 3 4 6 3 0 1 1 2 4 4 4 4 5 1 2 4 2 1 2 4 5 7 2 4 3
```

•

Ejemplo 3.6 Un examen tiene diez preguntas y cada una tiene tres opciones como respuesta, siendo solamente una de ellas la correcta. Si un estudiante contesta cada pregunta al azar, ¿cuál es la probabilidad de que apruebe el examen?

Solución. Si X denota el número de preguntas contestadas correctamente, entonces X tiene distribución $\text{bin}(n, p)$ con $n = 10$ y $p = 1/3$. Suponiendo que la calificación mínima aprobatoria es 6, entonces la respuesta es

$$P(X \geq 6) = \sum_{x=6}^{10} \binom{10}{x} (1/3)^x (2/3)^{10-x} = 0.07656353.$$

Esta probabilidad es sorprendentemente pequeña y por lo tanto la estrategia seguida por el estudiante para contestar el examen no parece ser la mejor. •

Ejercicios

293. Demuestre que la función de probabilidad de la distribución $\text{bin}(n, p)$ es, efectivamente, una función de probabilidad.
294. Sea X una variable aleatoria con distribución $\text{bin}(n, p)$. Encuentre los valores de los parámetros n y p cuando:
- $E(X) = 6$ y $\text{Var}(X) = 3$.
 - $E(X) = 12$ y $E(X^2) = 150$.
295. **Moda.** Sea $f(x)$ la función de probabilidad de la distribución $\text{bin}(n, p)$, en donde $0 < p < 1$. Demuestre que:

- a) Fórmula iterativa: para x entero tal que $0 \leq x \leq n$,

$$f(x+1) = \frac{p}{(1-p)} \frac{(n-x)}{(x+1)} f(x).$$

- b) $f(x)$ es creciente de x a $x+1$, es decir, $f(x) \leq f(x+1)$ para valores enteros no negativos de x en el intervalo $[0, (n+1)p - 1]$. Cuando este intervalo es vacío no se presenta este comportamiento creciente y la función de probabilidad es siempre decreciente.

- c) $f(x)$ es decreciente de x a $x + 1$, es decir, $f(x) \geq f(x + 1)$ para valores enteros no negativos de x en el intervalo $[(n+1)p-1, n-1]$. Nuevamente, existe la posibilidad de que no haya ningún valor de x para el que se cumpla este comportamiento decreciente, por ejemplo, cuando el intervalo mencionado es vacío, y por lo tanto la función es siempre creciente.
- d) si $(n + 1)p - 1$ no es un entero, entonces $f(x)$ tiene un único máximo en x^* definido como el entero más pequeño mayor o igual a $(n + 1)p - 1$, es decir,

$$x^* = \lceil (n + 1)p - 1 \rceil.$$

En este caso, x^* es la moda de la distribución y es única. Este valor se puede escribir también en la forma

$$x^* = \lfloor (n + 1)p \rfloor.$$

- e) si $(n+1)p-1$ es alguno de los enteros $0, 1, \dots, n-1$, entonces $f(x)$ alcanza su valor máximo en los puntos x^* y $x^* + 1 = (n + 1)p$, y la distribución es por lo tanto bimodal.

296. Usando directamente la definición de esperanza, demuestre que si X tiene distribución $\text{bin}(n, p)$, entonces:

- a) $E(X) = np$.
 b) $E(X^2) = np(1 - p + np)$.
 c) $\text{Var}(X) = np(1 - p)$.

297. Sea X una v.a. con distribución $\text{bin}(n, p)$. Suponiendo n constante, encuentre el valor de p que maximiza la varianza de X .

298. **f.g.p.** Sea X una v.a. con distribución $\text{bin}(n, p)$. Demuestre que la f.g.p. de X está dada por la expresión que aparece abajo. A través de esta función encuentre nuevamente la esperanza y la varianza de esta distribución.

$$G(t) = (1 - p + pt)^n.$$

299. **f.g.m.** Sea X una v.a. con distribución $\text{bin}(n, p)$. Demuestre que la f.g.m. de X está dada por la expresión que aparece abajo. A través de esta función encuentre nuevamente la esperanza y la varianza de esta distribución.

$$M(t) = (1 - p + pe^t)^n.$$

300. Demuestre la primera parte de la Proposición 3.1 de la página 216 usando

- a) la f.g.m.
- b) la f.g.p.

301. **Suma.** Sean X y Y dos variables aleatorias independientes con distribución $\text{bin}(n, p)$ y $\text{bin}(m, p)$, respectivamente. Demuestre que la variable $X + Y$ tiene distribución $\text{bin}(n + m, p)$ siguiendo los siguientes tres métodos:

- a) Calculando directamente $P(X + Y = k)$ para $k = 0, 1, \dots, n + m$.
- b) Usando la f.g.p.
- c) Usando la f.g.m.

302. Demuestre que si X tiene distribución $\text{bin}(n, p)$, entonces

$$n - X \sim \text{bin}(n, 1 - p).$$

303. Considere que se tiene un experimento aleatorio cualquiera y que A es un evento con probabilidad estrictamente positiva. Suponga que se realizan n ensayos independientes del experimento aleatorio y que X_n denota el número de veces que se observa la ocurrencia del evento A en estos n ensayos. Demuestre que para cualquier entero fijo $k \geq 1$,

$$\lim_{n \rightarrow \infty} P(X_n > k) = 1.$$

304. **Regularidades estadísticas.** Escriba un programa de cómputo que efectúe lo siguiente:

- a) Asigne un valor natural al parámetro n y una probabilidad al parámetro p con $0 < p < 1$.

- b) Genere 200 valores independientes al azar, x_1, \dots, x_{200} , de la distribución $\text{bin}(n, p)$ y calcule los promedios parciales

$$s_m = \frac{1}{m} \sum_{k=1}^m x_k, \quad \text{para } m = 1, 2, \dots, 200.$$

- c) Grafique la función $m \mapsto s_m$ y una los puntos con líneas rectas. Grafique también la línea horizontal $y = np$.

¿Qué puede decir del comportamiento de s_m ? Esta regularidad se presenta siempre para cualquier distribución con esperanza finita y se llama ley de los grandes números. Estudiaremos este resultado en la última parte de este texto.

305. Un productor de semillas conoce por experiencia que el 10% de un gran lote de semillas no germina. El productor vende sus semillas en paquetes de 20 semillas garantizando que por lo menos 18 de ellas germinarán. Calcule el porcentaje de paquetes que no cumplirán la garantía.
306. Se conoce que en una cierta población el 15% de las personas tiene un tipo específico de accidente en un año dado cualquiera. Encuentre la probabilidad de que una compañía aseguradora tenga que indemnizar a más de 5 personas de los 10 asegurados que componen su cartera para este tipo de accidentes en un año.
307. Se realizan 10 lanzamientos sucesivos e independientes de un dado equilibrado y se nos informa que han aparecido por lo menos k unos, $1 \leq k \leq 9$. Calcule la probabilidad de que
- exactamente k unos se hayan obtenido.
 - por lo menos $k + 1$ unos se hayan obtenido.
 - a lo sumo $k + 1$ unos se hayan obtenido.
 - todos hayan sido unos.
308. **El tablero de Galton.** Considere el tablero vertical que se muestra en la Figura 3.4, en donde se han colocado 5 filas de clavos en forma triangular. Una bola que se deja caer desde la parte superior choca

contra el primer clavo y baja al clavo inferior izquierdo con probabilidad $1/2$ o baja al clavo inferior derecho con probabilidad $1/2$, y así sucesivamente hasta caer en alguna de las 6 urnas que se encuentran en la parte inferior.

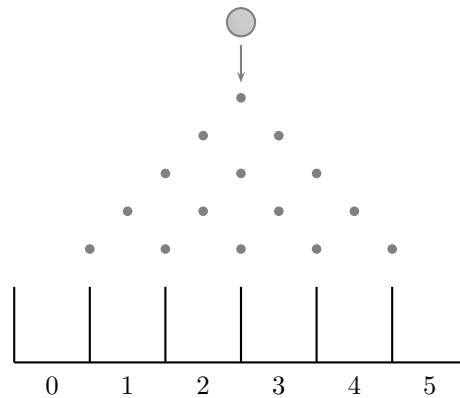


Figura 3.4

- a) Determine el número total de trayectorias distintas que la bola puede tomar. ¿Son igualmente probables estas trayectorias?
- b) Determine el número de trayectorias que llevan a cada una de las urnas.
- c) Calcule la probabilidad de que la bola caiga en cada una de las urnas. ¿Qué distribución de probabilidad es esta?
- d) Resuelva los tres incisos anteriores cuando se tienen n filas de clavos y por lo tanto $n + 1$ urnas, y la probabilidad de que la bola caiga a la izquierda es $1 - p$ y de que caiga a la derecha es p .

3.4. Distribución geométrica

Supongamos ahora que tenemos una sucesión infinita de ensayos independientes Bernoulli, en cada uno de los cuales la probabilidad de éxito es p .

Para cada una de estas sucesiones definimos la variable aleatoria X como el número de fracasos antes de obtener el primer éxito. Por ejemplo,

$$X(FEFEFF\cdots) = 1,$$

$$X(EFFEEE\cdots) = 0,$$

$$X(FFFEFE\cdots) = 3.$$

Observamos que X puede tomar los valores $0, 1, 2, \dots$. No es difícil darse cuenta que la probabilidad de que X tome el valor entero $x \geq 0$ es $p(1-p)^x$. Decimos entonces que X tiene una distribución geométrica con parámetro p y escribimos $X \sim \text{geo}(p)$ cuando su función de probabilidad es

$$f(x) = \begin{cases} p(1-p)^x & \text{si } x = 0, 1, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

La gráfica de esta función, cuando $p = 0.4$, se muestra en la Figura 3.5 y en el paquete **R** se pueden encontrar los valores de $f(x)$ de la siguiente forma.

```
# dgeom(x,p) evalúa f(x) de la distribución geo(p)
> dgeom(5,0.4)
[1] 0.031104
```

x	$f(x)$
0	0.4
1	0.24
2	0.144
3	0.0864
4	0.05184
5	0.031104
6	0.0186624
7	0.01119744
8	0.006718464
9	0.004031078
...	...

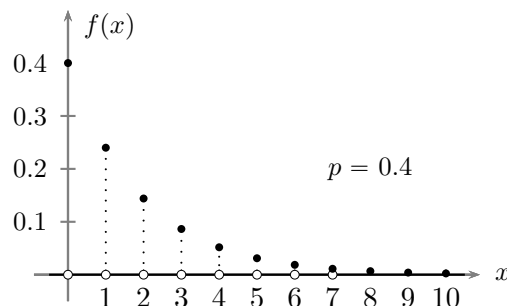


Figura 3.5

El nombre de esta distribución proviene del hecho de que al escribir la suma de todas estas probabilidades se obtiene una suma de tipo geométrico. Y dado que se conoce una fórmula para este tipo de sumas, al hacer las sumas parciales de la función de probabilidad se encuentra que la correspondiente función de distribución es

$$F(x) = \sum_{u \leq x} f(u) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - (1 - p)^{k+1} & \text{si } k \leq x < k + 1; \quad k = 0, 1, \dots \end{cases}$$

Los valores de $F(x)$ pueden encontrarse en **R** usando el siguiente comando.

```
# pgeom(x,p) evalúa F(x) de la distribución geo(p)
> pgeom(5,0.4)
[1] 0.953344
```

Para esta distribución es posible además demostrar que:

- a) $E(X) = \frac{1-p}{p}$.
- b) $\text{Var}(X) = \frac{1-p}{p^2}$.

Simulación 3.6 En el paquete **R** se pueden generar valores al azar de la distribución geométrica de la forma en que se muestra en el siguiente recuadro. Asigne un valor al parámetro p y genere valores al azar de esta distribución. Compare el promedio aritmético de los valores obtenidos con el valor $(1-p)/p$. ¿Son parecidos?

```
# rgeom(k,p) genera k valores al azar de la distribución geo(p)
> rgeom(25,0.4)
[1] 0 1 1 0 7 5 0 4 1 0 0 0 0 0 0 0 2 1 1 0 1 0 0 1 0
```

Ejemplo 3.7 Suponga la situación en donde se lleva a cabo una inspección sucesiva de artículos hasta encontrar uno defectuoso. Este proceso puede ser parte del control de calidad de una fábrica por ejemplo. El número aleatorio de artículos que se inspeccionan hasta encontrar uno defectuoso, sin contar este último, puede modelarse mediante una variable aleatoria con distribución geométrica. En este ejemplo los términos *artículo* y *defectuoso* son genéricos y pueden ser substituidos por términos adecuados al contexto de la aplicación. •

Ejemplo 3.8 Una persona participa cada semana con un boleto en un juego de lotería, en donde la probabilidad de ganar el primer premio es $p = 10^{-6} = 1/1,000,000$. ¿Cuántos años en promedio debe esta persona participar en el juego hasta obtener el primer premio?

Solución. Supongamos que X denota el número de veces que la persona participa en el juego antes de obtener el primer premio. Entonces X puede tomar los valores $0, 1, \dots$ y tiene una distribución geométrica de parámetro $p = 10^{-6}$. La variable aleatoria $X + 1$ representa, en cambio, el número de participaciones incluyendo el momento de ganar. Su esperanza es

$$E(X + 1) = \frac{1 - p}{p} + 1 = \frac{1}{p} = 10^6 = 1,000,000.$$

Lo que es equivalente a 19,230 años, aproximadamente. Esta cantidad es el número de años promedio que la persona debe jugar, semana tras semana, para obtener el primer premio. Observe que el valor directo $E(X) = (1-p)/p$ nos proporciona una aproximación de la cantidad buscada. •

Ejercicios

309. Sea $f(x)$ la función de probabilidad de la distribución $\text{geo}(p)$. Demuestre que:

- $f(x)$ es, efectivamente, una función de probabilidad.
- $f(x)$ es decreciente y por lo tanto tiene un máximo en $x^* = 0$.

310. Usando la definición de esperanza y varianza demuestre que si X es una v.a. con distribución $\text{geo}(p)$ entonces:

$$a) E(X) = (1 - p)/p.$$

$$b) \text{Var}(X) = (1 - p)/p^2.$$

311. Use la fórmula (2.17) del Ejercicio 215, en la página 164, para demostrar que si X tiene distribución $\text{geo}(p)$ entonces

$$E(X) = \frac{1 - p}{p}.$$

312. **Simulación.** Sea X_0, X_1, \dots una sucesión de v.a.s independientes con distribución $\text{Ber}(p)$. Defina

$$X = \text{mín} \{n \geq 0 : X_n = 1\}.$$

Demuestre que X tiene distribución $\text{geo}(p)$. Esto permite encontrar valores al azar de la distribución geométrica a partir de una sucesión de valores al azar de la distribución Bernoulli.

313. **f.g.p.** Demuestre que la f.g.p. de una v.a. X con distribución $\text{geo}(p)$ está dada por la expresión que aparece abajo. A través de esta función encuentre nuevamente las expresiones para la esperanza y varianza de esta variable aleatoria.

$$G(t) = \frac{p}{1 - (1 - p)t} \quad \text{para } |t| < 1/(1 - p).$$

314. **f.g.m.** Demuestre que la f.g.m. de una v.a. X con distribución $\text{geo}(p)$ está dada por la expresión que aparece abajo. A través de esta función encuentre nuevamente las expresiones para la esperanza y varianza de esta variable aleatoria.

$$M(t) = \frac{p}{1 - (1 - p)e^t} \quad \text{para } |t| < -\ln(1 - p).$$

315. **Propiedad de pérdida de memoria.** Sea X una v.a. con distribución $\text{geo}(p)$. Demuestre que para cualesquiera enteros $n, m \geq 0$,

$$P(X \geq n + m \mid X \geq m) = P(X \geq n).$$

316. Sean X y Y dos v.a.s independientes, ambas con distribución $\text{geo}(p)$. Demuestre que

$$P(X + Y = k) = \binom{k+1}{k} (1-p)^k p^2, \quad k = 0, 1, \dots$$

Compruebe que la expresión anterior corresponde, efectivamente, a una función de probabilidad. Esta es la distribución binomial negativa de parámetros (r, p) con $r = 2$, la cual veremos en la siguiente sección.

317. Dos personas lanzan alternativamente una moneda equilibrada. Se escoge previamente una de las caras de la moneda y el primero que obtenga esa cara es el ganador. Encuentre la probabilidad de ganar de cada uno de los jugadores.
318. Una moneda equilibrada y marcada con “cara” y “cruz” se lanza repetidas veces hasta que aparecen 10 “caras”. Sea X la variable que registra el número total de lanzamientos. Calcule la función de probabilidad de X .
319. Sea X una v.a. con distribución $\text{geo}(p)$ y sea n un número natural fijo. Encuentre la función de probabilidad de la variable

$$Y = \text{mín} \{X, n\} = \begin{cases} X & \text{si } X < n, \\ n & \text{si } X \geq n. \end{cases}$$

320. **Variante de la distribución geométrica.** En ocasiones es necesario considerar el número de ensayos (no el de fracasos) antes del primer éxito en una sucesión de ensayos independientes Bernoulli. En este caso la variable aleatoria en cuestión es

$$Y = 1 + X,$$

en donde X tiene una distribución $\text{geo}(p)$, es decir, la distribución se desplaza una unidad hacia la derecha. Demuestre que:

$$a) f_Y(y) = \begin{cases} (1-p)^{y-1} p & \text{si } y = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) F_Y(y) = \begin{cases} 0 & \text{si } y < 1, \\ 1 - (1-p)^y & \text{si } k \leq y < k+1; \quad k = 1, 2, \dots \end{cases}$$

$$c) E(Y) = 1/p.$$

$$d) \text{Var}(Y) = (1-p)/p^2.$$

321. Se tiene una gran cantidad de productos y se sabe que el porcentaje de defectuosos es $(100\theta)\%$, en donde $\theta \in (0, 1)$ es un número fijo. En un procedimiento de muestreo se escogen los productos al azar, uno por uno, hasta encontrar uno defectuoso. Sea X el número de elementos que se tienen que escoger hasta encontrar uno defectuoso. Entonces X tiene función de probabilidad aproximada

$$f(x) = \begin{cases} (1-\theta)^{x-1} \theta & \text{si } x = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

- a) Demuestre que $f(x)$ es una función de probabilidad.
 b) Encuentre la probabilidad de que sean necesarias más de 10 extracciones para poder obtener un producto defectuoso.

3.5. Distribución binomial negativa

Consideremos nuevamente la situación de observar los resultados de una sucesión infinita de ensayos independientes Bernoulli, en cada uno de los cuales la probabilidad de éxito es p . Sea $r \geq 1$ un número entero. Definimos ahora a la variable aleatoria X como el número de fracasos antes de obtener el r -ésimo éxito. Decimos entonces que X tiene una distribución binomial negativa con parámetros r y p , y escribimos $X \sim \text{bin neg}(r, p)$. Es claro que la variable X puede tomar los valores $0, 1, 2, \dots$ con las probabilidades dadas por la función de probabilidad

$$f(x) = P(X = x) = \begin{cases} \binom{r+x-1}{x} p^r (1-p)^x & \text{si } x = 0, 1, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

En esta fórmula aparece el término p^r pues nos interesa observar r éxitos. Por otro lado, podemos tener un número variable x de fracasos, de ahí el

término $(1-p)^x$. Finalmente, el factor $\binom{r+x-1}{x}$ indica los diferentes arreglos en los que los x fracasos y los $r-1$ éxitos se encuentran distribuidos en $r+x-1$ ensayos. Observe que el r -ésimo éxito debe aparecer en el ensayo $r+x$. En el Ejercicio 322 se deja al lector resolver el problema no trivial de verificar que la función $f(x)$, arriba indicada, es efectivamente una función de probabilidad. La gráfica de esta función aparece en la Figura 3.6 cuando los valores de los parámetros son $r = 5$ y $p = 0.5$. Los valores de $f(x)$ se pueden obtener en el paquete R de la siguiente forma.

```
# dnbinom(x,r,p) evalúa f(x) de la distribución bin neg(r,p)
> dnbinom(3,5,0.5)
[1] 0.1367188
```

x	$f(x)$
0	0.031250
1	0.078125
2	0.117187
3	0.136718
4	0.136718
5	0.123046
6	0.102539
7	0.080566
8	0.060424
9	0.043640
10	0.030548
...	...

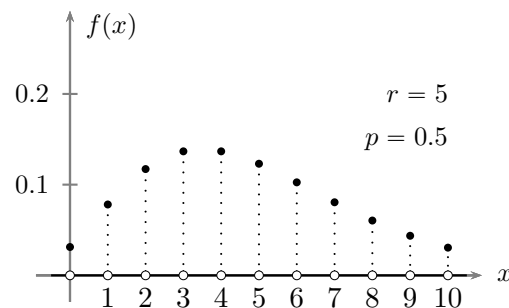


Figura 3.6

No existe una expresión compacta para la función de distribución $F(x)$ de la distribución binomial negativa y por lo tanto no intentaremos escribirla. Sus valores se pueden encontrar en R usando el siguiente comando.

```
# pnbinom(x,r,p) evalúa F(x) de la distribución bin neg(r,p)
> pnbinom(7,5,0.5)
[1] 0.8061523
```

Es claro que la distribución binomial negativa es una generalización de la distribución geométrica. Esta última se obtiene tomando $r = 1$. Se puede además demostrar que

- a) $E(X) = r(1-p)/p$.
- b) $\text{Var}(X) = r(1-p)/p^2$.

Por otro lado, el coeficiente binomial puede extenderse para cualquier número real a y cualquier entero natural x de la siguiente manera.

$$\binom{a}{x} = \frac{a(a-1)\cdots(a-x+1)}{x!}. \quad (3.2)$$

Puede entonces demostrarse la siguiente identidad, de donde adquiere su nombre la distribución binomial negativa.

$$\binom{r+x-1}{x} = (-1)^x \binom{-r}{x}. \quad (3.3)$$

El siguiente resultado establece una forma de construir la distribución binomial negativa como una suma de variables aleatorias independientes con distribución geométrica. Su demostración no es complicada. Como un ejercicio se pide desarrollar una prueba para la primera implicación. La segunda implicación es una formalización de la definición misma de una v.a. con distribución binomial negativa. Aplicando directamente esperanza y varianza a la suma indicada pueden encontrarse estas cantidades para esta distribución.

Proposición 3.2 Sea $r \geq 1$ un número entero y sean X_1, \dots, X_r variables aleatorias independientes, cada una con distribución $\text{geo}(p)$. Entonces

$$X_1 + \cdots + X_r \sim \text{bin neg}(r, p). \quad (3.4)$$

Recíprocamente, toda v.a. con distribución $\text{bin neg}(r, p)$ puede ser expresada como una suma de la forma anterior.

Simulación 3.7 En el paquete R se pueden generar valores al azar de la distribución binomial negativa, como se muestra en el siguiente recuadro. Como un ejercicio de simulación asigne valores a los parámetros r y p , y genere valores al azar de esta distribución.

```
# rnbinom(k,r,p) genera k valores al azar de la distribución
# bin neg(r,p)
> rnbinom(25,5,0.5)
[1] 1 7 7 3 1 4 2 1 3 10 4 5 3 1 11 6 7 3 5 3 9 9 6 1 7
```

Ejemplo 3.9 Se lanza repetidas veces una moneda honesta cuyos dos resultados son cara y cruz. ¿Cuál es la probabilidad de obtener la tercera cruz en el quinto lanzamiento?

Solución. Sea X el número de caras (fracasos) antes de obtener la tercera cruz. Entonces $X \sim \text{bin neg}(r, p)$ con $r = 3$, $p = 1/2$. La pregunta planteada se traduce en encontrar la probabilidad $P(X = 2)$. Se tiene entonces que

$$P(X = 2) = \binom{4}{2} (1/2)^5 = 6/32 = 0.1875.$$

Ejercicios

322. Demuestre que la función de probabilidad de la distribución binomial negativa efectivamente es una función de probabilidad.
323. Demuestre que el coeficiente binomial que aparece en la definición de la distribución binomial negativa se puede expresar de la siguiente forma.

$$\binom{r+x-1}{x} = (-1)^x \binom{-r}{x}.$$

324. **Moda.** Sea $f(x)$ la función de probabilidad de la distribución bin-neg(r, p), en donde $0 < p < 1$. Demuestre que

a) Fórmula iterativa: para $x \geq 0$ entero,

$$f(x+1) = (1-p) \frac{x+r}{x+1} f(x)$$

b) $f(x)$ es creciente de x a $x+1$, es decir, $f(x) \leq f(x+1)$ para valores enteros de x en el intervalo $[0, (r-1)(1-p)/p - 1]$. Es posible que no haya ningún valor de x para el que se presente este comportamiento creciente, por ejemplo cuando el intervalo mencionado es vacío. Este es el caso, por ejemplo, cuando $r = 1$, que corresponde a la distribución geométrica. La moda es única y es $x^* = 0$.

c) $f(x)$ es decreciente de x a $x+1$, es decir, $f(x) \geq f(x+1)$ para valores enteros de x dentro del intervalo $[(r-1)(1-p)/p - 1, \infty)$.

d) si $(r-1)(1-p)/p - 1$ no es un entero, entonces $f(x)$ tiene un único máximo en x^* , definido como el entero más pequeño mayor o igual a $(r-1)(1-p)/p - 1$, es decir,

$$x^* = \lceil (r-1)(1-p)/p - 1 \rceil.$$

En este caso, x^* es la moda de la distribución y es única. Este valor se puede escribir también en la forma

$$x^* = \lfloor (r-1)(1-p)/p \rfloor.$$

e) si $(r-1)(1-p)/p - 1$ es un entero mayor o igual a cero, entonces $f(x)$ alcanza su valor máximo en los puntos x^* y $x^* + 1 = (r-1)(1-p)/p$, y la distribución es por lo tanto bimodal.

325. **Simulación.** Sea X_1, X_2, \dots una sucesión de v.a.s independientes con distribución $\text{Ber}(p)$ y sea $r \geq 1$ un entero. Defina

$$X = \min \left\{ n \geq r : \sum_{k=1}^n X_k = r \right\} - r.$$

Demuestre que X tiene distribución bin neg(r, p). Esto permite encontrar valores al azar de la distribución binomial negativa a partir de valores al azar de la distribución Bernoulli.

326. **f.g.p.** Sea X una v.a. con distribución bin neg(r, p). Demuestre que la f.g.p. de X está dada por la expresión que aparece abajo. A través de esta función, encuentre nuevamente las expresiones para la esperanza y la varianza de esta distribución.

$$G(t) = \left(\frac{p}{1 - (1-p)t} \right)^r \quad \text{para } |t| < 1/(1-p).$$

327. **f.g.m.** Sea X una v.a. con distribución bin. neg(r, p). Demuestre que la f.g.m. de X está dada por la expresión que aparece abajo. A través de esta función, encuentre nuevamente las expresiones para la esperanza y la varianza de esta distribución.

$$M(t) = \left(\frac{p}{1 - (1-p)e^t} \right)^r \quad \text{para } |t| < -\ln(1-p).$$

328. Demuestre la primera parte de la Proposición 3.2 de la página 229 usando

- a) la f.g.m.
- b) la f.g.p.

329. **Suma.** Sean X y Y dos variables aleatorias independientes con distribución bin neg(r, p) y bin neg(s, p), respectivamente. Demuestre que la variable $X + Y$ tiene distribución bin neg($r + s, p$) siguiendo los siguientes tres métodos.

- a) Calculando directamente $P(X + Y = k)$ para $k = 0, 1, \dots$
- b) Usando la f.g.p.
- c) Usando la f.g.m.

330. **Muestreo.** Se desea encontrar a 20 personas que reúnan ciertas características para aplicarles un cuestionario. Si únicamente el 1% de la población cumple las características requeridas, y suponiendo que se consulta al azar a las personas para determinar si son adecuadas para contestar el cuestionario, determine el número promedio de personas que se necesita consultar para encontrar a las 20 personas solicitadas.

331. Un experimento aleatorio consiste en lanzar un dado equilibrado hasta obtener 6 veces el número 6. Encuentre la probabilidad de que el experimento requiera n lanzamientos.
332. **Variante de la distribución binomial negativa.** En ocasiones interesa considerar el número de ensayos (no el de fracasos) hasta obtener el r -ésimo éxito en una sucesión de ensayos independientes Bernoulli. En este caso, la variable aleatoria en cuestión es

$$Y = r + X,$$

en donde X tiene una distribución bin neg(r, p), es decir, la distribución se desplaza r unidades hacia la derecha. Demuestre que

$$a) f_Y(y) = \begin{cases} \binom{y-1}{y-r} (1-p)^{y-r} p^r & \text{si } y = r, r+1, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) E(Y) = r/p.$$

$$c) \text{Var}(Y) = r(1-p)/p^2.$$

Nota. No existe una expresión sencilla para $F_Y(y)$.

333. Un experimento aleatorio consiste en lanzar una moneda repetidas veces. Nos interesa observar el evento de obtener k resultados de la misma cara, por primera vez, en el n -ésimo lanzamiento, no necesariamente de manera consecutiva. Suponga $n = k, k+1, \dots, 2k-1$.
- Calcule la probabilidad del evento mencionado cuando la moneda es equilibrada.
 - Compruebe que la suma de las probabilidades del inciso anterior es 1 para $n = k, k+1, \dots, 2k-1$.
 - Calcule la probabilidad del evento mencionado cuando la moneda no es equilibrada.
 - Compruebe nuevamente que la suma de las probabilidades del inciso anterior es 1.
334. Sea X una v.a. con distribución bin neg(r, p) con $p = 1/2$. Demuestre que

$$F_X(r-1) = 1/2.$$

3.6. Distribución hipergeométrica

Esta distribución de probabilidad surge en el contexto de la toma de una muestra de un conjunto de objetos de dos tipos. Supongamos que tenemos N objetos dentro de una caja, de los cuales K son de un primer tipo y $N - K$ son de un segundo tipo. Véase la Figura 3.7. Los objetos del primer tipo pueden corresponder a artículos en buen estado y los del segundo tipo a artículos en mal estado, o bien a personas con una cierta característica y a aquellas que no poseen dicha característica.

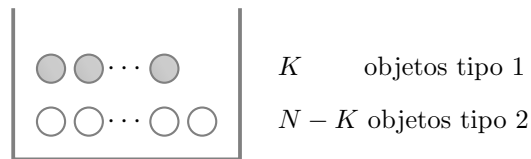


Figura 3.7

Supongamos que de esta caja tomamos al azar una muestra de tamaño n de tal forma que la selección es sin reemplazo y el orden de los objetos seleccionados no es relevante. Así, el espacio muestral de este experimento consiste de todos los posibles subconjuntos de tamaño n que se pueden obtener de esta colección de N objetos y su cardinalidad es entonces $\binom{N}{n}$. Si para cada subconjunto seleccionado se define la variable aleatoria X como el número de objetos seleccionados que son del primer tipo, entonces es claro que X puede tomar los valores $0, 1, 2, \dots, n$. Observe que X toma el valor n si y sólo si todos los objetos escogidos son del tipo 1, mientras que toma el valor 0 cuando todos los objetos escogidos son del tipo 2. Para que tales casos puedan ocurrir, supondremos que el tamaño n de muestra es suficientemente pequeño de tal forma que

$$n \leq \min \{K, N - K\}. \quad (3.5)$$

La probabilidad de que X tome un valor x está dada por la siguiente expresión.

$$f(x) = \begin{cases} \frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}} & \text{si } x = 0, 1, \dots, n, \\ 0 & \text{en otro caso.} \end{cases}$$

Decimos entonces que X tiene una distribución hipergeométrica con parámetros N , K y n , y escribimos $X \sim \text{hipergeo}(N, K, n)$. Para entender la fórmula de la función de probabilidad de esta distribución observe que el término $\binom{K}{x}$ establece las diferentes formas en que x objetos pueden escogerse de la colección de K objetos del tipo 1, mientras que el término $\binom{N-K}{n-x}$ corresponde a las diferentes formas de escoger $n-x$ objetos de los $N-K$ objetos del tipo 2. Se usa entonces el principio multiplicativo para obtener el número total de muestras diferentes, en donde x objetos son del primer tipo y $n-x$ objetos son del segundo tipo. No es un ejercicio fácil verificar que esta función de probabilidad efectivamente lo es, pero puede realizarse siguiendo la sugerencia que aparece en la solución del Ejercicio 335. La gráfica de esta función de probabilidad para $N = 20$, $K = 7$ y $n = 5$ aparece en la Figura 3.8.

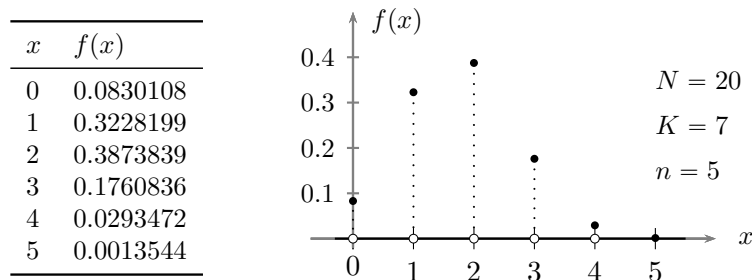


Figura 3.8

En el paquete **R** pueden obtenerse los valores de $f(x)$ como se muestra en el recuadro siguiente. Observe con cuidado la diferencia en el orden y la forma de expresar los parámetros de esta distribución en **R**: después del argumento x se especifica el número de objetos K de tipo 1, después el

número de objetos $N - K$ de tipo 2 y finalmente se especifica el tamaño de la muestra n .

```
# dhyper(x,K,N-K,n) evalúa  $f(x)$  de la distribución
# hipergeo( $N, K, n$ )
> dhyper(3,7,13,5)
[1] 0.1760836
```

Por otro lado, no presentaremos una fórmula para la función de distribución $F(x)$ pues no tiene una expresión compacta sencilla, sin embargo sus valores pueden encontrarse usando R mediante el siguiente comando.

```
# phyper(x,K,N-K,n) evalúa  $F(x)$  de la distribución
# hipergeo( $N, K, n$ )
> phyper(3,7,13,5)
[1] 0.9692982
```

Aplicando directamente la definición de esperanza, no es complicado comprobar que si X tiene distribución hipergeo(N, K, n) entonces

- a) $E(X) = n \frac{K}{N}$.
- b) $\text{Var}(X) = n \frac{K}{N} \frac{N - K}{N} \frac{N - n}{N - 1}$.

Simulación 3.8 Mediante el siguiente comando en el paquete R pueden obtenerse valores al azar de la distribución hipergeométrica. Asigne usted valores a los parámetros N , K y n , como en el ejemplo, y genere tantos valores de esta distribución como desee modificando de valor de k .

```
# rhyper(k,K,N-K,n) genera  $k$  valores al azar de la distribución
# hipergeo( $N, K, n$ )
> rhyper(25,7,13,5)
[1] 3 2 3 2 3 1 3 2 1 1 1 3 1 3 3 0 4 3 4 3 1 1 3 2 1
```

Ejercicios

335. Demuestre que la función de probabilidad de la distribución hipergeométrica efectivamente es una función de probabilidad.

336. Sea X una v.a. con distribución hipergeo(N, K, n). Demuestre que

$$\begin{aligned} a) \quad E(X) &= n \frac{K}{N}. \\ b) \quad \text{Var}(X) &= n \frac{K}{N} \frac{N-K}{N} \frac{N-n}{N-1}. \end{aligned}$$

337. Compruebe que la distribución hipergeo(N, K, n) se reduce a la distribución Ber(p) con $p = K/N$ cuando $n = 1$.

338. Sea $f(x)$ la función de probabilidad de la distribución hipergeo(N, K, n). Demuestre que

a) Fórmula iterativa: para $x = 0, 1, \dots, n$,

$$f(x+1) = \frac{(K-x)(n-x)}{(x+1)(N-K-n+x-1)} f(x).$$

b) $f(x)$ es creciente de x a $x+1$, es decir, $f(x) \leq f(x+1)$ para valores enteros de x en el intervalo $[0, (Kn - N + K + n - 1)/(N + 2)]$. Es posible que no haya ningún valor de x para el que se presente este comportamiento creciente, por ejemplo cuando el intervalo mencionado es vacío.

c) $f(x)$ es decreciente de x a $x+1$, es decir, $f(x) \geq f(x+1)$ para valores enteros de x dentro del intervalo $[(Kn - N + K + n - 1)/(N + 2), n - 1]$. Si este intervalo es vacío, este tipo de comportamiento podría no presentarse.

d) si $(Kn - N + K + n - 1)/(N + 2)$ no es un entero, entonces $f(x)$ tiene un único máximo en x^* , definido como el entero más pequeño mayor o igual a $(Kn - N + K + n - 1)/(N + 2)$, es decir,

$$x^* = \lceil (Kn - N + K + n - 1)/(N + 2) \rceil.$$

- e) si $(Kn - N + K + n - 1)/(N + 2)$ es uno de los enteros $0, 1, \dots, n - 1$, entonces $f(x)$ alcanza su valor máximo en los puntos x^* y $x^* + 1 = (n + 1)(K + 1)/(N + 2)$, y la distribución es por lo tanto bimodal.
339. Sea X una v.a. con distribución hipergeo(N, K, n). Demuestre que la función de probabilidad de X converge a la función de probabilidad bin(n, p) cuando $N \rightarrow \infty$ y $K \rightarrow \infty$ de tal forma que $K/N \rightarrow p$.
340. Suponga que en conjunto se tienen N_1 objetos de un primer tipo, N_2 objetos de un segundo tipo y N_3 objetos de un tercer tipo. Suponga que se extrae al azar un subconjunto de tamaño n de tal forma que $1 \leq n \leq \min\{N_1, N_2 + N_3\}$. Sea X el número de objetos del primer tipo contenidos en la muestra. Encuentre la distribución de X .
341. Se pone a la venta un lote de 100 artículos de los cuales 10 son defectuosos. Un comprador extrae una muestra al azar de 5 artículos y decide que si encuentra 2 o más defectuosos, entonces no compra el lote. Calcule la probabilidad de que la compra se efectúe.

3.7. Distribución Poisson

Supongamos que deseamos observar el número de ocurrencias de un cierto evento dentro de un intervalo de tiempo dado, por ejemplo, el número de clientes que llegan a un cajero automático durante la noche, o tal vez deseamos registrar el número de accidentes que ocurren en cierta avenida durante todo un día, o el número de reclamaciones que llegan a una compañía aseguradora. Para modelar este tipo de situaciones podemos definir la variable aleatoria X como el número de ocurrencia de este evento en el intervalo de tiempo dado. Es claro entonces que X puede tomar los valores $0, 1, 2, \dots$, y en principio no ponemos una cota superior para el número de observaciones del evento. Adicionalmente supongamos que conocemos la tasa media de ocurrencia del evento de interés, que denotamos por la letra λ (lambda). El parámetro λ es positivo y se interpreta como el número promedio de ocurrencias del evento por unidad de tiempo o espacio. La probabilidad de que la variable aleatoria X tome un valor entero $x \geq 0$ se definirá a continuación. Decimos que X tiene una distribución Poisson² con parámetro $\lambda > 0$,

²Siméon Denis Poisson (1781–1840), matemático, geómetra y físico francés.

y escribimos $X \sim \text{Poisson}(\lambda)$ cuando su función de probabilidad es

$$f(x) = \begin{cases} e^{-\lambda} \frac{\lambda^x}{x!} & \text{si } x = 0, 1, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

x	$f(x)$
0	0.1353353
1	0.2706706
2	0.2706706
3	0.1804470
4	0.0902235
5	0.0360894
6	0.0120298
7	0.0034370
...	...

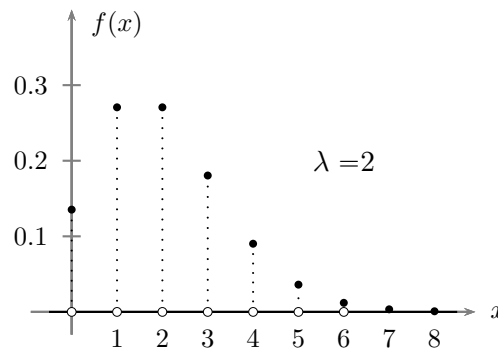


Figura 3.9

Puede demostrarse que la función $f(x)$, arriba definida, es efectivamente una función de probabilidad para cada valor de $\lambda > 0$ fijo, y para ello conviene recordar la serie de Taylor de la función exponencial alrededor del cero, es decir,

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

La forma de la función de probabilidad $f(x)$ se muestra en la Figura 3.9 cuando $\lambda = 2$. En el paquete R pueden obtenerse los valores de $f(x)$ usando el siguiente comando.

```
# dpois(x,lambda) evalúa f(x) de la distribución Poisson(lambda)
> dpois(3,2)
[1] 0.1804470
```

La función de distribución $F(x)$, como suma de los valores de $f(x)$, no tiene una expresión reducida y no la escribiremos, sin embargo sus valores pueden encontrarse con facilidad en **R** mediante el siguiente comando.

```
# ppois(x,λ) evalúa F(x) de la distribución Poisson(λ)
> ppois(3,2)
[1] 0.8571235
```

Después de algunos cálculos sencillos puede comprobarse que para esta distribución,

- a) $E(X) = \lambda$.
- b) $\text{Var}(X) = \lambda$.

Simulación 3.9 Mediante el comando en **R** que aparece en el siguiente recuadro pueden generarse valores al azar de la distribución Poisson. Especificando un valor para λ genere tantos valores al azar como desee de una v.a. con distribución $\text{Poisson}(\lambda)$ y compare el promedio aritmético de estos valores con el valor de λ . ¿Son parecidos?

```
# rpois(k,λ) genera k valores al azar de la distribución
# Poisson(λ)
> rpois(25,2)
[1] 0 3 3 3 1 1 3 0 0 2 0 3 1 1 0 3 5 4 1 0 1 2 1 2 1
```

Ejemplo 3.10 En promedio se reciben 2 peticiones de acceso a una página web durante un minuto cualquiera. Utilice el modelo Poisson para calcular la probabilidad de que en un minuto dado

- a) nadie solicite acceso a la página.
- b) se reciban mas de dos peticiones.

Solución. Sea X el número de peticiones por minuto y supongamos que X tiene distribución Poisson(λ) con $\lambda = 2$. Para el primer inciso se tiene que

$$P(X = 0) = e^{-2} \frac{2^0}{0!} = 0.135.$$

Para el segundo inciso,

$$\begin{aligned} P(X > 2) = 1 - P(X \leq 2) &= 1 - (P(X = 0) + P(X = 1) + P(X = 2)) \\ &= 1 - e^{-2} (2^0/0! + 2^1/1! + 2^2/2!) \\ &= 0.323. \end{aligned}$$

•

Puede además demostrarse que cuando $X \sim \text{bin}(n, p)$, y haciendo tender n a infinito y p a cero de tal forma que el producto np se mantenga constante igual a λ , la variable aleatoria X adquiere la distribución Poisson de parámetro λ . En el Ejercicio 346 que aparece en la página 245 puede verse el enunciado preciso de este resultado, el cual puede considerarse como una forma límite de obtener la distribución Poisson a través de la distribución binomial. Además, este mismo resultado sugiere que cuando n es grande, la distribución binomial puede ser aproximada mediante la distribución Poisson de parámetro $\lambda = np$. Esto es particularmente útil pues las probabilidades de la distribución binomial involucran el cálculo de factoriales y ello puede ser computacionalmente difícil para números grandes. El siguiente ejemplo ilustrará esta situación.

Ejemplo 3.11 En promedio, uno de cada 100 focos producidos por una máquina es defectuoso. Use la distribución Poisson para estimar la probabilidad de encontrar 5 focos defectuosos en un lote de 1000 focos.

Solución. Sea X el número de focos defectuosos en el lote de 1000 focos. Entonces X tiene distribución $\text{bin}(n, p)$ con $n = 1000$ y $p = 1/100$. Nos piden calcular

$$P(X = 5) = \binom{1000}{5} (1/100)^5 (99/100)^{995} = 0.0374531116.$$

Usando la aproximación Poisson con $\lambda = np = 1000/100 = 10$,

$$P(X = 5) \approx e^{-10} \frac{10^5}{5!} = 0.0379841747.$$

•

Hemos interpretado a una variable aleatoria con distribución Poisson de parámetro λ como aquella variable que registra el número de ocurrencias de un cierto evento dentro de un intervalo de tiempo dado. Supongamos que tal intervalo es de longitud unitaria y consideremos que se trata del intervalo $[0, 1]$. Considere ahora que nos interesa observar el número de ocurrencias del evento en un intervalo de longitud diferente, por ejemplo $[0, t]$, con $t > 0$. Tal conteo de ocurrencias también sigue una distribución Poisson pero esta vez de parámetro λt . Por ejemplo, si $t = 2$, entonces el número de ocurrencias del evento en el intervalo $[0, 2]$ tiene distribución Poisson de parámetro 2λ . El siguiente ejemplo ilustra este resultado.

Ejemplo 3.12 El número de aviones que llega a un aeropuerto internacional se considera como una cantidad aleatoria y se modela mediante una variable aleatoria con distribución Poisson con una frecuencia promedio de 3 aviones cada 10 minutos. Es decir, la unidad de medición del tiempo es de diez minutos. Se puede escribir

$$X_{[0,10]} \sim \text{Poisson}(3),$$

en donde se ha indicado el intervalo al que se hace referencia como subíndice de la variable aleatoria. Tenemos entonces los siguientes ejemplos.

- a) La probabilidad de que no llegue ningún avión en un periodo de 20 minutos (dos unidades de tiempo) es $P(X_{[0,20]} = 0)$, en donde $X_{[0,20]}$ tiene distribución Poisson de parámetro $(2)(3) = 6$, es decir,

$$P(X_{[0,20]} = 0) = e^{-6} \frac{6^0}{0!}.$$

- b) La probabilidad de que llegue sólo un avión en el minuto siguiente es $P(X_{[0,1/10]} = 1)$, en donde $X_{[0,1/10]}$ tiene distribución Poisson de parámetro $(1/10)(3) = 3/10$, es decir,

$$P(X_{[0,1/10]} = 1) = e^{-3/10} \frac{(3/10)^1}{1!}.$$

- c) La probabilidad de que lleguen dos o más aviones en un periodo de 15 minutos es $P(X_{[0,15/10]} \geq 2)$, en donde $X_{[0,15/10]}$ tiene distribución Poisson de parámetro $(15/10)(3) = 45/10$, es decir,

$$P(X_{[0,15/10]} \geq 2) = \sum_{k=2}^{\infty} e^{-45/10} \frac{(45/10)^k}{k!}.$$

•

La distribución Poisson tiene algunas propiedades que resultan muy útiles en su aplicación. La siguiente es una de ellas.

Proposición 3.3 Sean X y Y dos variables aleatorias independientes con distribución $\text{Poisson}(\lambda_1)$ y $\text{Poisson}(\lambda_2)$, respectivamente. Entonces

$$X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2).$$

Demostración. Para cualquier entero $u \geq 0$,

$$\begin{aligned} P(X + Y = u) &= \sum_{x=0}^u P(X = x, Y = u - x) \\ &= \sum_{x=0}^u P(X = x) P(Y = u - x) \\ &= \sum_{x=0}^u e^{-\lambda_1} \frac{\lambda_1^x}{x!} e^{-\lambda_2} \frac{\lambda_2^{u-x}}{(u-x)!} \\ &= e^{-(\lambda_1 + \lambda_2)} \frac{1}{u!} \sum_{x=0}^u \binom{u}{x} \lambda_1^x \lambda_2^{u-x} \\ &= e^{-(\lambda_1 + \lambda_2)} \frac{(\lambda_1 + \lambda_2)^u}{u!}. \end{aligned}$$

•

El resultado anterior ya había sido demostrado en el Ejemplo 2.28 de la página 192, usando la propiedad de caracterización única de la f.g.p. La demostración presentada ahora es directa y no hace uso de la f.g.p.

Ejercicios

342. Demuestre que la función de probabilidad de la distribución Poisson(λ) es, efectivamente, una función de probabilidad.

343. Sea X una v.a. con distribución Poisson(λ). Demuestre que:

a) $E(X) = \lambda$.

b) $\text{Var}(X) = \lambda$.

344. Sean $n, m \geq 0$ dos números enteros distintos fijos y sea X una v.a. con distribución Poisson que satisface $P(X = n) = P(X = m)$. Encuentre el parámetro de esta distribución Poisson.

345. Sea $f(x)$ la función de probabilidad de la distribución Poisson(λ). Demuestre que:

a) Fórmula iterativa: para $x \geq 0$ entero,

$$f(x+1) = \frac{\lambda}{x+1} f(x).$$

b) $f(x)$ es creciente de x a $x+1$, es decir, $f(x) \leq f(x+1)$ para valores enteros de x en el intervalo $[0, \lambda - 1]$. Es posible que no haya ningún valor de x para el que se presente este comportamiento creciente, por ejemplo, cuando el intervalo mencionado es vacío.

c) $f(x)$ es decreciente de x a $x+1$, es decir, $f(x) \geq f(x+1)$ para valores enteros de x en el intervalo $[\lambda - 1, \infty)$.

d) si $\lambda - 1$ no es un entero, entonces $f(x)$ tiene un único máximo en x^* , definido como el entero más pequeño mayor o igual a $\lambda - 1$, es decir,

$$x^* = \lceil \lambda - 1 \rceil.$$

En este caso, x^* es la moda de la distribución y es única. Este valor se puede escribir también en la forma

$$x^* = \lfloor \lambda \rfloor.$$

e) si $\lambda - 1$ es un entero, entonces $f(x)$ alcanza su valor máximo en los puntos $x^* = \lambda - 1$ y $x^* + 1 = \lambda$, y la distribución es por lo tanto bimodal.

346. **Aproximación a la distribución binomial.** Sea X una v.a. con distribución $\text{bin}(n, p)$ tal que $p = \lambda/n$ para n suficientemente grande, con $\lambda > 0$ constante. Demuestre que para cada $k = 0, 1, \dots$

$$\lim_{n \rightarrow \infty} P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Esto quiere decir que la función de probabilidad binomial converge a la función de probabilidad Poisson, o en términos de aproximaciones, lo anterior sugiere que las probabilidades de la distribución binomial pueden aproximarse mediante las probabilidades Poisson cuando n es grande y p es pequeño de la forma λ/n . En la gráfica 3.10 se comparan las funciones de probabilidad de las distribuciones binomial(n, p) y Poisson(λ) con $\lambda = 2$, $p = \lambda/n$ para $n = 4, 6, 8, 10$. Se observa que conforme el parámetro n crece, las dos funciones son cada vez más parecidas. La gráfica de barras corresponde a la distribución binomial y la gráfica de puntos corresponde a la distribución Poisson.

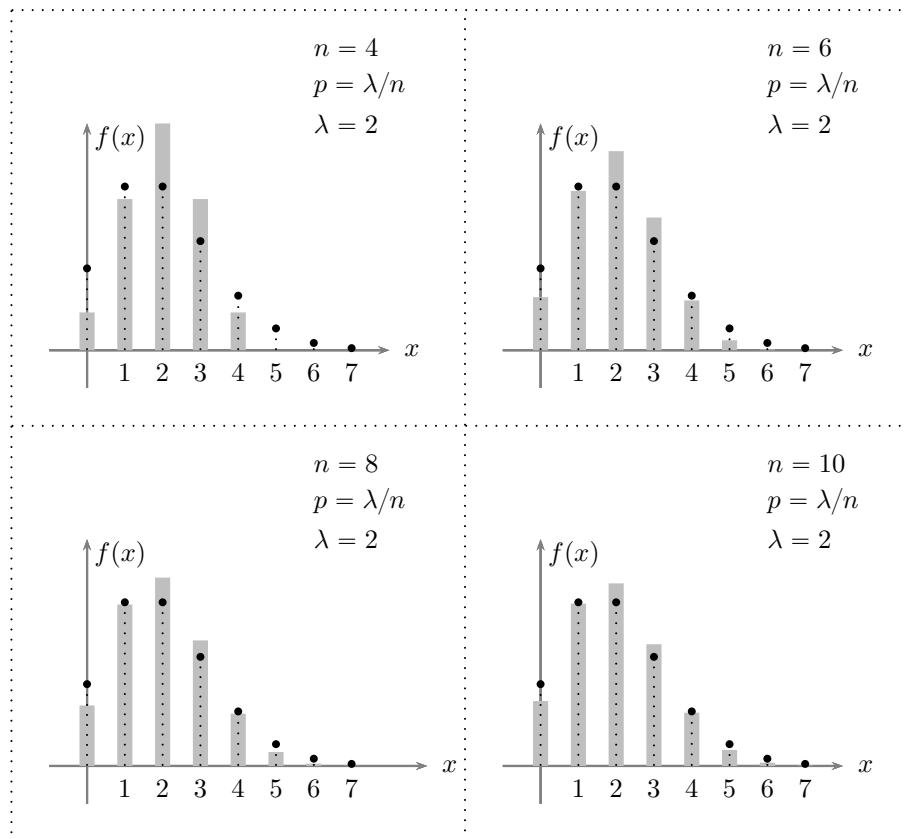
347. **f.g.p.** Sea X una v.a. con distribución Poisson(λ). Demuestre que la función generadora de probabilidad de X es la función $G(t)$ que aparece abajo. Usando esta función y sus propiedades encuentre nuevamente la esperanza y varianza de esta distribución.

$$G(t) = e^{\lambda(t-1)}.$$

348. **Momentos.** Sea X una v.a. con distribución Poisson(λ). Demuestre que para cualquier entero $n \geq 1$,

$$E(X^n) = \lambda \sum_{k=0}^{n-1} \binom{n-1}{k} E(X^k).$$

349. **f.g.m.** Sea X una v.a. con distribución Poisson(λ). Demuestre que la función generadora de momentos de X es la función $M(t)$ que aparece

Figura 3.10: Binomial *vs.* Poisson.

abajo. Usando esta función y sus propiedades encuentre nuevamente la esperanza y varianza de esta distribución.

$$M(t) = e^{\lambda(e^t - 1)}.$$

350. El siguiente resultado fue demostrado en la Proposición 3.3, en la página 243, usando un método directo. Ahora se propone un método alternativo. Sean X y Y dos variables aleatorias independientes con distribución Poisson(λ_1) y Poisson(λ_2), respectivamente. Use la función generadora de momentos para demostrar nuevamente que

$$X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2).$$

351. Sea X_1, X_2, \dots una sucesión de v.a.s independientes con distribución Ber(p) e independientes de N con distribución Poisson(λ). Demuestre que la variable X , definida a continuación, tiene distribución Poisson(λp). Cuando $N = 0$, la suma es vacía y se define como cero.

$$X = \sum_{i=1}^N X_i.$$

352. Sea X una v.a. con distribución Poisson(λ) tal que $\lambda \in (0, 1)$. Demuestre que

$$E(X!) = \frac{1}{1 - \lambda} e^{-\lambda}.$$

353. En un libro muy voluminoso, el número de errores por página se modela mediante una v.a. con distribución Poisson con media uno. Encuentre la probabilidad de que una página seleccionada al azar contenga

- a) ningún error.
- b) exactamente dos errores.
- c) al menos tres errores.

354. El número de semillas en una variedad de naranjas sigue una distribución Poisson de media 3. Calcule la probabilidad de que una naranja seleccionada al azar contenga

- a) ninguna semilla.
- b) al menos dos semillas.
- c) a lo sumo tres semillas.

355. Suponga que el número de accidentes al día que ocurren en una parte de una carretera es una variable aleatoria Poisson de parámetro $\lambda = 3$.

- a) Calcule la probabilidad de que ocurran 2 o más accidentes en un día cualquiera.
- b) Conteste el inciso anterior bajo la suposición de que ha ocurrido al menos un accidente.

§

Con esto concluimos la revisión de algunas distribuciones de probabilidad de tipo discreto. Ahora estudiaremos algunas distribuciones de tipo continuo. No construiremos estas distribuciones a partir de experimentos aleatorios particulares, como en varios de los casos de tipo discreto, sino que las definiremos sin mayor justificación.

Recordamos aquí nuevamente al lector que los parámetros de las distintas distribuciones de probabilidad pueden no ser usados de la misma forma en las diversas fuentes bibliográficas y paquetes computacionales, y hay que tener cuidado al respecto.

3.8. Distribución uniforme continua

Decimos que una variable aleatoria X tiene una distribución uniforme continua en el intervalo (a, b) , y escribimos $X \sim \text{unif}(a, b)$, cuando su función de densidad es

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } a < x < b, \\ 0 & \text{en otro caso.} \end{cases}$$

Esta distribución tiene como parámetros los números a y b . La gráfica general de esta función se muestra en la Figura 3.11 (a), y es evidente que se trata de una función de densidad pues es no negativa e integra uno. Esta es una función muy sencilla y sus valores pueden calcularse en el paquete R de

la siguiente forma.

```
# dunif(x,a,b) evalúa  $f(x)$  de la distribución  $\text{unif}(a,b)$ 
> dunif(2,-1,3)
[1] 0.25
```

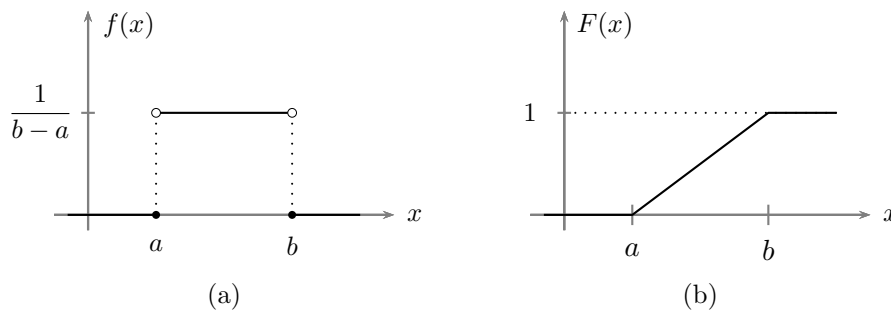


Figura 3.11

Integrando esta función de densidad desde menos infinito hasta un punto x cualquiera, puede encontrarse la función de distribución, la cual tiene la siguiente expresión y cuya gráfica se muestra en la Figura 3.11 (b).

$$F(x) = \begin{cases} 0 & \text{si } x \leq a, \\ \frac{x-a}{b-a} & \text{si } a < x < b, \\ 1 & \text{si } x \geq b. \end{cases}$$

Los valores de esta función pueden ser calculados en el paquete R usando el siguiente comando.

```
# punif(x,a,b) evalúa  $F(x)$  de la distribución  $\text{unif}(a,b)$ 
> punif(2,-1,3)
[1] 0.75
```

Por otro lado, es fácil verificar para esta distribución que

a) $E(X) = (a + b)/2$.

b) $\text{Var}(X) = (b - a)^2/12$.

Observe que la esperanza corresponde al punto medio del intervalo (a, b) . Además la varianza o dispersión crece cuando a y b se alejan uno del otro, y por el contrario, cuando estos parámetros están muy cercanos, la varianza es pequeña. Esta distribución es una de las más sencillas y se usa, naturalmente, cuando no se conoce mayor información de la variable aleatoria de interés, excepto que toma valores continuos dentro de cierto intervalo.

Ejemplo 3.13 En el experimento aleatorio teórico de generar un número al azar X en un intervalo (a, b) , se considera regularmente que X tiene distribución uniforme en dicho intervalo. Por ejemplo, algunos problemas estudiados antes sobre probabilidad geométrica hacen uso de esta distribución sin especificarlo completamente. •

Simulación 3.10 Pueden generarse valores al azar en el paquete R de la distribución uniforme continua usando el comando que aparece en el siguiente recuadro. Asigne valores de su preferencia a los parámetros a y b , y genere valores al azar de esta distribución.

```
# runif(k,a,b) genera k valores al azar de la distribución
unif(a,b)
> runif(5,-1,3)
[1] 1.5889966 -0.1420308 2.5391841 0.4416415 0.7294366
```

Ejercicios

356. Sea X una v.a. con distribución $\text{unif}(0, 4)$ y denote por μ a la media de esta distribución. Encuentre el valor de $c > 0$ tal que

- a) $P(X \leq c) = 1/8$. c) $P(|X - \mu| < c) = 1/2$
 b) $P(X > c + \mu) = 1/4$. d) $P(|X - 4| < c) = 3/4$.

357. Sea X una v.a con distribución unif(a, b). Demuestre que

- a) $E(X) = \frac{a+b}{2}$.
 b) $E(X^2) = (a^2 + ab + b^2)/3$.
 c) $\text{Var}(X) = (b-a)^2/12$.

358. Sea X una v.a. con distribución unif($-1, 1$). Demuestre que el n -ésimo momento de X está dado por

$$E(X^n) = \begin{cases} 0 & \text{si } n \text{ es impar,} \\ 1/(n+1) & \text{si } n \text{ es par.} \end{cases}$$

359. Sea X una v.a. con distribución unif(a, b). Demuestre que el n -ésimo momento de X está dado por

$$E(X^n) = \frac{b^{n+1} - a^{n+1}}{(n+1)(b-a)}.$$

360. **Cuantil.** Sea $p \in (0, 1)$. Encuentre una expresión para el cuantil al $100p\%$ de la distribución unif(a, b).

361. **f.g.m.** Sea X una v.a. con distribución unif(a, b). Demuestre que la función generadora de momentos de X es la función $M(t)$ que aparece abajo. Usando esta función y sus propiedades encuentre nuevamente la esperanza y la varianza de esta distribución.

$$M(t) = \frac{e^{bt} - e^{at}}{t(b-a)}.$$

362. **Simulación.** Sea U una variable aleatoria con distribución unif($0, 1$). Sean $a < b$ dos constantes. Demuestre que la variable $X = a + (b-a)U$ tiene distribución unif(a, b). Este es un mecanismo para obtener valores al azar de la distribución unif(a, b) a partir de valores al azar de la distribución unif($0, 1$). Esta última distribución aparece implementada en varios sistemas de cómputo.

363. Sea $\epsilon > 0$. Se escoge un número al azar X dentro del intervalo $(0, 1)$. Encuentre la probabilidad de que

a) $|X - 1/2| > \epsilon$.

b) $(2X - 1/2)^2 \leq \epsilon$.

364. Sea X una v.a. con distribución uniforme en el intervalo $(0, 1)$ y sea α un número real cualquiera. Encuentre y grafique la función de densidad de la variable

$$Y = X^\alpha.$$

365. Se escoge un número al azar X dentro del intervalo $(-1, 1)$. Encuentre y grafique la función de densidad y de distribución de la variable

a) $Y = X + 1$.

c) $Y = X^2$.

b) $Y = |X|$.

d) $Y = X^3$.

366. Se escoge un número al azar X dentro del intervalo $(-1, 1)$. Calcule la probabilidad de que

a) $X^2 < 1/4$.

b) $|X + 1| \leq |X - 1|$.

367. Sea X una variable aleatoria arbitraria cuyos posibles valores están contenidos en el intervalo (a, b) . Demuestre que

a) $a \leq E(X) \leq b$.

b) $0 \leq \text{Var}(X) \leq (b - a)^2/4$.

368. Sean α y $\beta > 0$ dos números reales fijos. Sea X una v.a. con distribución $\text{unif}(a, b)$. Encuentre el valor de los parámetros a y b de tal forma que se satisfagan las siguientes dos condiciones.

$$E(X) = \alpha,$$

$$\text{y } \text{Var}(X) = \beta.$$

369. Sea X con distribución $\text{unif}(a, b)$. Encuentre una variable Y en términos de X de tal forma que Y tenga distribución $\text{unif}(c, d)$. Encuentre además

- a) $F_Y(y)$ en términos de $F_X(x)$.
 b) $f_Y(y)$ en términos de $f_X(x)$.

370. **Una forma de aproximar π .** Escriba un programa de cómputo para generar valores al azar x, y con distribución uniforme dentro del intervalo $(0, 1)$, independiente un valor del otro. Si ocurre que $y \leq \sqrt{1 - x^2}$, entonces el punto (x, y) pertenece a la región sombreada de la Figura 3.12 y se considera un éxito, en caso contrario se considera un fracaso. Así, al repetir varias veces este procedimiento, el cociente del número de éxitos n_E entre el número de ensayos n será una aproximación del área de la región sombreada. Siendo esta región una cuarta parte del círculo unitario, su área es $\pi/4$. Grafique la función $n \mapsto 4n_E/n$ para $n = 1, 2, \dots, 100$ y comente su comportamiento conforme n crece.

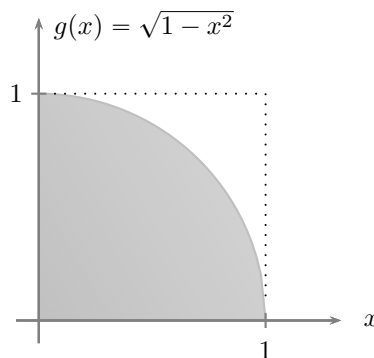


Figura 3.12

3.9. Distribución exponencial

Decimos que una variable aleatoria continua X tiene distribución exponencial con parámetro $\lambda > 0$, y escribimos $X \sim \exp(\lambda)$, cuando su función de densidad es

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

La gráfica de esta función, cuando el parámetro λ toma el valor particular 3, se muestra en la Figura 3.13 (a). La correspondiente función de distribución aparece a su derecha. Es muy sencillo verificar que la función $f(x)$, arriba definida, es efectivamente una función de densidad para cualquier valor del parámetro $\lambda > 0$. Se trata pues de una variable aleatoria continua con conjunto de valores el intervalo $(0, \infty)$. Esta distribución se usa para modelar tiempos de espera para la ocurrencia de un cierto evento. Los valores de $f(x)$ pueden calcularse en el paquete R de la siguiente forma.

```
# dexp(x,λ) evalúa f(x) de la distribución exp(λ)
> dexp(0.5,3)
[1] 0.6693905
```

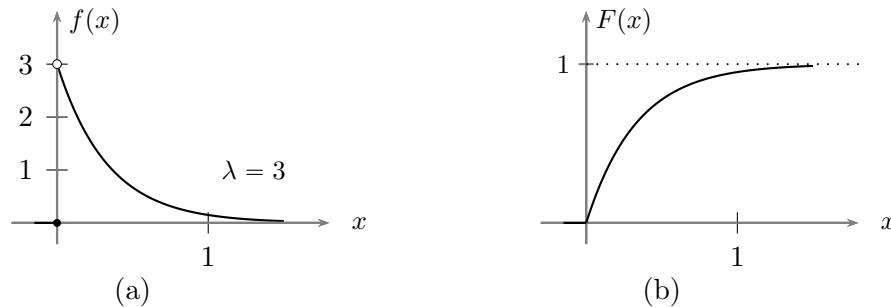


Figura 3.13

Integrando la función de densidad desde menos infinito hasta un valor arbitrario x , se encuentra que la función de distribución tiene la expresión que aparece abajo y cuya gráfica se muestra en la Figura 3.13 (b).

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

En el paquete R, la función $F(x)$ se calcula usando el siguiente comando.

```
# pexp(x,λ) evalúa F(x) de la distribución exp(λ)
> pexp(0.5,3)
[1] 0.7768698
```

Aplicando el método de integración por partes puede comprobarse que

- a) $E(X) = 1/\lambda$.
- b) $\text{Var}(X) = 1/\lambda^2$.

Simulación 3.11 Pueden generarse valores al azar en \mathbb{R} de la distribución exponencial usando el comando que aparece en el siguiente recuadro. Como un ejercicio de simulación asigne un valor de su preferencia al parámetro λ y genere valores al azar de esta distribución.

```
# rexp(k,λ) genera k valores al azar de la distribución exp(λ)
> rexp(5,3)
[1] 0.53847926 0.19371105 0.32025823 0.07144621 0.20201383
```

Ejemplo 3.14 Suponga que el tiempo en minutos que un usuario cualquiera permanece revisando su correo electrónico sigue una distribución exponencial de parámetro $\lambda = 1/5$. Calcule la probabilidad de que un usuario cualquiera permanezca conectado al servidor de correo

- a) menos de un minuto.
- b) más de una hora.

Solución. Sea X el tiempo de conexión al servidor de correo. Para el primer inciso tenemos que

$$P(X < 1) = \int_0^1 1/5 e^{-x/5} dx = 0.181.$$

Para el segundo inciso,

$$P(X > 60) = \int_{60}^{\infty} 1/5 e^{-x/5} dx = 0.0000061.$$

Ejercicios

371. Demuestre que la función de densidad exponencial efectivamente es una función de densidad. A partir de ella encuentre la correspondiente función de distribución.

372. Sea X una v.a. con distribución $\exp(\lambda)$ con $\lambda = 2$. Encuentre

- | | |
|------------------|--------------------------------------|
| a) $P(X < 1)$ | c) $P(X < 1 \mid X < 2)$. |
| b) $P(X \geq 2)$ | d) $P(1 \leq X \leq 2 \mid X > 0)$. |

373. Sea X una v.a. con distribución $\exp(\lambda)$. Use la definición de esperanza y el método de integración por partes para demostrar que

- | | |
|-----------------------------|------------------------------------|
| a) $E(X) = 1/\lambda$. | c) $E(X^3) = 6/\lambda^3$. |
| b) $E(X^2) = 2/\lambda^2$. | d) $\text{Var}(X) = 1/\lambda^2$. |

374. Sea X una v.a. con distribución $\exp(\lambda)$. Use la fórmula (2.18) del Ejercicio 217, en la página 165, para encontrar la esperanza de las variables no negativas X , X^2 y X^3 y demostrar nuevamente que

- | | |
|-----------------------------|------------------------------------|
| a) $E(X) = 1/\lambda$. | c) $E(X^3) = 6/\lambda^3$. |
| b) $E(X^2) = 2/\lambda^2$. | d) $\text{Var}(X) = 1/\lambda^2$. |

375. **Momentos.** Sea X una v.a. con distribución $\exp(\lambda)$. Demuestre que el n -ésimo momento de X es

$$E(X^n) = \frac{n!}{\lambda^n}.$$

376. Sea X una v.a. con distribución $\exp(\lambda)$ y sea $c > 0$ una constante. Demuestre que

$$cX \sim \exp(\lambda/c).$$

377. **f.g.m.** Sea X una v.a. con distribución $\exp(\lambda)$. Demuestre que la función generadora de momentos de X es la función $M(t)$ que aparece

abajo. Usando esta función y sus propiedades, encuentre nuevamente la esperanza y la varianza de esta distribución.

$$M(t) = \frac{\lambda}{\lambda - t} \quad \text{para } t < \lambda.$$

378. **Cuantil.** Sea $p \in (0, 1)$. Encuentre el cuantil al $100p\%$ de la distribución $\exp(\lambda)$. En particular, muestre que la mediana es $(\ln 2)/\lambda$.

379. **Propiedad de pérdida de memoria.** Sea X una v.a. con distribución exponencial de parámetro λ . Demuestre que, para cualesquiera valores $x, y \geq 0$,

$$P(X > x + y | X > y) = P(X > x).$$

380. **Discretización.** Sea X una v.a. con distribución $\exp(\lambda)$. Demuestre que la variable aleatoria discreta, definida a continuación, tiene distribución $\text{geo}(p)$ con $p = 1 - e^{-\lambda}$.

$$Y = \begin{cases} 0 & \text{si } 0 < X \leq 1, \\ 1 & \text{si } 1 < X \leq 2, \\ 2 & \text{si } 2 < X \leq 3, \\ \dots & \dots \end{cases}$$

381. **Simulación: método de la función inversa.** Sea U una variable aleatoria con distribución $\text{unif}(0, 1)$ y sea $\lambda > 0$ una constante. Demuestre que la variable aleatoria X , definida a continuación, tiene distribución $\exp(\lambda)$. Este resultado permite obtener valores al azar de la distribución exponencial a partir de valores de la distribución uniforme continua.

$$X = -\frac{1}{\lambda} \ln(1 - U).$$

382. **Coche en venta.** Un señor está vendiendo su coche y decide aceptar la primera oferta que exceda \$50,000. Si las ofertas son variables aleatorias independientes con distribución exponencial de media \$45,000, encuentre

- a) la distribución de probabilidad del número de ofertas recibidas hasta vender el coche.
- b) la probabilidad de que el precio de venta rebase \$55,000.
- c) el precio promedio de venta del coche.

383. El tiempo medido en horas para reparar una máquina es una variable aleatoria exponencial de parámetro $\lambda = 1/2$. ¿Cuál es la probabilidad de que un tiempo de reparación

- a) exceda 2 horas?
- b) tome a lo sumo 4 horas?
- c) tome a lo sumo 4 horas dado que no se ha logrado la reparación en las primeras 2 horas?

3.10. Distribución gamma

La variable aleatoria continua X tiene una distribución gamma con parámetros $\alpha > 0$ y $\lambda > 0$, y escribimos $X \sim \text{gamma}(\alpha, \lambda)$, si su función de densidad es

$$f(x) = \begin{cases} \frac{(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} \lambda e^{-\lambda x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

La gráfica de esta función de densidad, para varios valores de los parámetros, se muestra en la Figura 3.14. En la expresión anterior aparece el término $\Gamma(\alpha)$, el cual se conoce como la función gamma, y es de este hecho que la distribución adquiere su nombre. La función gamma se define por medio de la siguiente integral.

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt,$$

para cualquier número real α tal que esta integral sea convergente. Para evaluar la función gamma es necesario substituir el valor de α en el integrando y efectuar la integral infinita. En general, no necesitaremos evaluar esta integral para cualquier valor de α , sólo para algunos pocos valores, principalmente enteros, y nos ayudaremos de las siguientes propiedades que

no son difíciles de verificar.

- a) $\Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$.
- b) $\Gamma(\alpha + 1) = \alpha!$ si α es un entero positivo.
- c) $\Gamma(2) = \Gamma(1) = 1$.
- d) $\Gamma(1/2) = \sqrt{\pi}$.

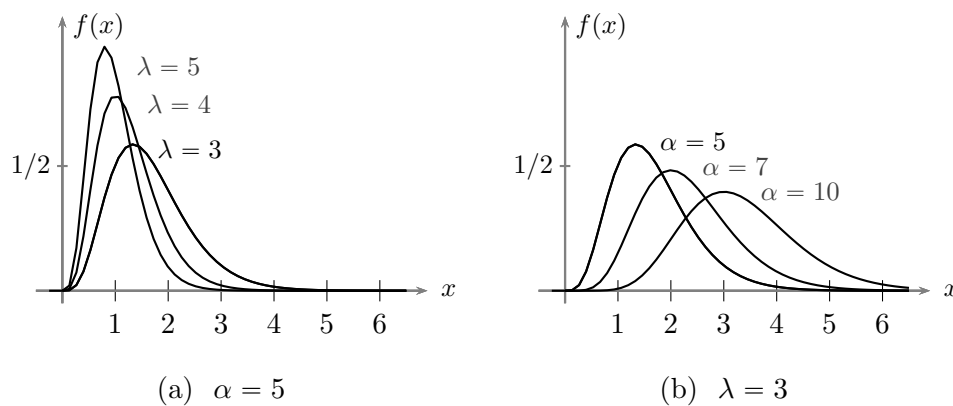


Figura 3.14

Así, a la función gamma se le puede considerar como una generalización del factorial pues coincide con éste cuando el argumento es un número entero positivo. En el paquete R pueden obtenerse los valores de la función $\Gamma(x)$ mediante el comando `gamma(x)`. Los valores de la función de densidad $f(x)$ se obtienen de la siguiente forma.

```
# dgamma(x,shape=alpha,rate=lambda) evalúa f(x) de la distribución
# gamma(alpha, lambda)
> dgamma(2.5,shape=7,rate=3)
[1] 0.4101547
```

Observemos que la distribución exponencial es un caso particular de la distribución gamma pues si en ésta se toma el parámetro α igual a 1, se obtiene

la distribución exponencial de parámetro λ .

Por otro lado, la función de distribución gamma $F(x)$ no tiene, en general, una expresión compacta, pero pueden calcularse con facilidad sus valores en R mediante el comando que aparece en el siguiente recuadro. Véase el Ejercicio 393 para conocer una fórmula para $F(x)$ en un caso particular de sus parámetros.

```
# pgamma(x,shape= $\alpha$ ,rate= $\lambda$ ) evalúa  $F(x)$  de la distribución
# gamma( $\alpha$ , $\lambda$ )
> pgamma(2.5,shape=7,rate=3)
[1] 0.6218453
```

Resolviendo un par de integrales se puede demostrar que

- a) $E(X) = \alpha/\lambda$.
- b) $\text{Var}(X) = \alpha/\lambda^2$.

Cuando el parámetro α es un número natural n , la distribución gamma(n, λ) adquiere también el nombre de distribución Erlang(n, λ), y esta distribución puede obtenerse del siguiente resultado, que es una aplicación inmediata de la f.g.m. y sus propiedades.

Proposición 3.4 Sean X_1, \dots, X_n variables aleatorias independientes, cada una de ellas con distribución $\exp(\lambda)$. Entonces

$$X_1 + \dots + X_n \sim \text{gamma}(n, \lambda).$$

Así, una variable aleatoria con distribución gamma(n, λ), cuando n es un entero positivo, puede interpretarse como el tiempo acumulado de n tiempos de espera exponenciales independientes, uno seguido del otro. Este resultado también indica un mecanismo para generar un valor al azar de la distribución gamma(n, λ) a partir de n valores al azar de la distribución $\exp(\lambda)$.

Simulación 3.12 Pueden generarse valores al azar de la distribución gamma en el paquete **R** usando el comando que aparece en el siguiente recuadro. Asigne un valor de su preferencia a los parámetros α y λ y genere valores al azar de esta distribución.

```
# rgamma(k,shape= $\alpha$ ,rate= $\lambda$ ) genera  $k$  valores al azar de la
# distribución gamma( $\alpha$ , $\lambda$ )
> rgamma(5,shape=7,rate=3)
[1] 3.170814 1.433144 2.103220 1.662244 3.025049
```

Ejercicios

384. Use la definición de la función gamma para demostrar que la función de densidad gamma efectivamente lo es.
385. Sea X una v.a. con distribución gamma(α, λ) con $\alpha = 2$ y $\lambda = 3$. Encuentre
- | | |
|------------------|-----------------------------------|
| a) $P(X < 1)$ | c) $P(X < 1 X < 2)$. |
| b) $P(X \geq 2)$ | d) $P(1 \leq X \leq 2 X > 0)$. |
386. Sea X una v.a. con distribución gamma(α, λ). Use la definición de esperanza para demostrar que
- | | |
|--|--|
| a) $E(X) = \alpha/\lambda$. | c) $E(X^3) = \alpha(\alpha+1)(\alpha+2)/\lambda^3$. |
| b) $E(X^2) = \alpha(\alpha+1)/\lambda^2$. | d) $\text{Var}(X) = \alpha/\lambda^2$. |
387. **Moda.** Sea X una v.a. con distribución gamma(α, λ). Demuestre que si $\alpha > 1$ entonces X tiene una única moda dada por

$$x^* = \frac{\alpha - 1}{\lambda}.$$

388. **Momentos.** Sea X una v.a. con distribución $\text{gamma}(\alpha, \lambda)$. Demuestre que el n -ésimo momento de X es

$$E(X^n) = \frac{\alpha(\alpha + 1) \cdots (\alpha + n - 1)}{\lambda^n}.$$

389. Sea X una v.a. con distribución $\text{gamma}(\alpha, \lambda)$ y sea $c > 0$ una constante. Demuestre que

$$cX \sim \text{gamma}(\alpha, \lambda/c).$$

390. **f.g.m.** Sea X una v.a. con distribución $\text{gamma}(\alpha, \lambda)$. Demuestre que la función generadora de momentos de X es la función $M(t)$ que aparece abajo. Usando esta función y sus propiedades encuentre nuevamente la esperanza y la varianza de esta distribución.

$$M(t) = \left(\frac{\lambda}{\lambda - t} \right)^\alpha \quad \text{para } t < \lambda.$$

391. **Suma.** Utilice la f.g.m. para demostrar que si X y Y son independientes con distribución $\text{gamma}(\alpha_1, \lambda)$ y $\text{gamma}(\alpha_2, \lambda)$, respectivamente, entonces

$$X + Y \sim \text{gamma}(\alpha_1 + \alpha_2, \lambda).$$

392. Utilice la f.g.m. para demostrar la Proposición 3.4 de la página 260.

393. **Función de distribución: caso particular.** Sea X una v.a. con distribución $\text{gamma}(n, \lambda)$ en donde n es un entero positivo. Denote por $F_n(x)$ a la función de distribución de esta variable aleatoria y defina $F_0(x)$ como la función de distribución de la v.a. constante cero. Demuestre que para $x > 0$,

$$a) F_n(x) = F_{n-1}(x) - \frac{(\lambda x)^{n-1}}{(n-1)!} e^{-\lambda x}.$$

$$b) F_n(x) = 1 - \sum_{k=0}^{n-1} \frac{(\lambda x)^k}{k!} e^{-\lambda x} = \sum_{k=n}^{\infty} \frac{(\lambda x)^k}{k!} e^{-\lambda x}.$$

394. Demuestre las siguientes propiedades de la función gamma. Para el último inciso podría ayudar consultar la distribución normal que estudiaremos más adelante.

- a) $\Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$.
- b) $\Gamma(\alpha + 1) = \alpha!$ si α es un entero positivo.
- c) $\Gamma(2) = \Gamma(1) = 1$.
- d) $\Gamma(1/2) = \sqrt{\pi}$.

3.11. Distribución beta

Decimos que la variable aleatoria continua X tiene una distribución beta con parámetros $a > 0$ y $b > 0$, y escribimos $X \sim \text{beta}(a, b)$, cuando su función de densidad es

$$f(x) = \begin{cases} \frac{1}{B(a, b)} x^{a-1} (1-x)^{b-1} & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

El término $B(a, b)$ se conoce como la función beta, y de allí adquiere su nombre esta distribución. La función beta se define como sigue.

$$B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx,$$

para números reales $a > 0$ y $b > 0$. Esta función está relacionada con la función gamma, antes mencionada, a través de la identidad

$$B(a, b) = \frac{\Gamma(a) \Gamma(b)}{\Gamma(a + b)}.$$

Véase la sección de ejercicios para una lista de propiedades de esta función. La gráfica de la función de densidad beta se muestra en la Figura 3.15 para algunos valores de sus parámetros. En el paquete R pueden calcularse los valores de $f(x)$ de la siguiente forma.

```
# dbeta(x,a,b) evalúa f(x) de la distribución beta(a,b)
> dbeta(0.3,1,2)
[1] 1.4
```

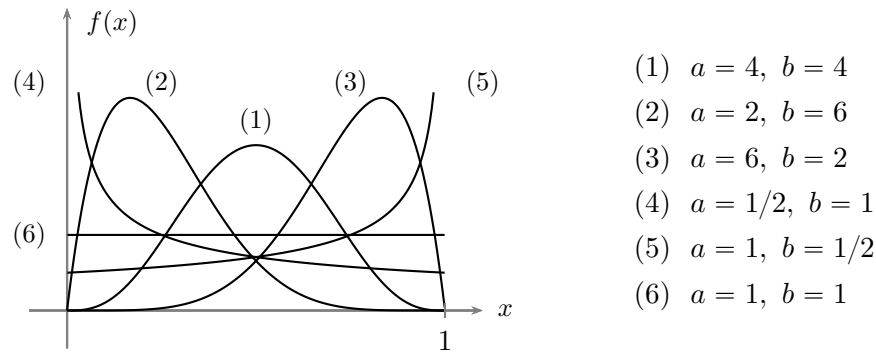


Figura 3.15

La correspondiente función de distribución no tiene una forma reducida y se escribe simplemente como sigue.

$$F(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ \frac{1}{B(a, b)} \int_0^x u^{a-1}(1-u)^{b-1} du & \text{si } 0 < x < 1, \\ 1 & \text{si } x \geq 1, \end{cases}$$

y sus valores pueden calcularse en R usando el siguiente comando.

```
# pbeta(x,a,b) evalúa F(x) de la distribución beta(a,b)
> pbeta(0.3,1,2)
[1] 0.51
```

Para la distribución $\text{beta}(a, b)$, usando una identidad que aparece en el Ejercicio 395, se puede demostrar, sin mucha dificultad, que

a) $E(X) = \frac{a}{a+b}$.

b) $\text{Var}(X) = \frac{ab}{(a+b+1)(a+b)^2}$.

La función generadora de momentos no tiene una forma compacta para esta distribución. Por otro lado, en R se pueden generar valores al azar de la distribución beta de manera análoga a las otras distribuciones, esto es,

```
# rbeta(k,a,b) genera k valores al azar de la distribución
# beta(a,b)
> rbeta(5,1,2)
[1] 0.18713260 0.07264413 0.08796477 0.15438134 0.29011107
```

La distribución beta puede obtenerse a partir de la distribución gamma como indica el siguiente resultado cuya demostración omitiremos.

Proposición 3.5 Sean X y Y dos variables aleatorias independientes con distribución $\text{gamma}(a, \lambda)$ y $\text{gamma}(b, \lambda)$, respectivamente. Entonces

$$\frac{X}{X+Y} \sim \text{beta}(a, b).$$

Finalmente observamos que la distribución $\text{beta}(a, b)$ se reduce a la distribución $\text{unif}(0, 1)$ cuando $a = b = 1$.

Ejercicios

395. Demuestre las siguientes propiedades de la función beta. La identidad (e) será particularmente útil.

a) $B(a, b) = B(b, a)$.

b) $B(a, 1) = 1/a$.

c) $B(1, b) = 1/b$.

d) $B(a+1, b) = \frac{a}{b} B(a, b+1)$.

e) $B(a+1, b) = \frac{a}{a+b} B(a, b)$.

f) $B(a, b+1) = \frac{b}{a+b} B(a, b)$.

g) $B(1/2, 1/2) = \pi$.

396. Demuestre que si $X \sim \text{beta}(a, b)$ entonces

$$1 - X \sim \text{beta}(b, a).$$

397. Sea X una v.a con distribución beta(a, b). Demuestre que

$$a) E(X) = \frac{a}{a+b}.$$

$$b) E(X^2) = \frac{a(a+1)}{(a+b)(a+b+1)}.$$

$$c) E(X^3) = \frac{a(a+1)(a+2)}{(a+b)(a+b+1)(a+b+2)}.$$

$$d) \text{Var}(X) = \frac{ab}{(a+b+1)(a+b)^2}.$$

398. Sea X una v.a. con distribución beta(a, b). Demuestre que el n -ésimo momento de X es

$$E(X^n) = \frac{B(a+n, b)}{B(a, b)} = \frac{a(a+1) \cdots (a+n-1)}{(a+b)(a+b+1) \cdots (a+b+n-1)}.$$

399. Encuentre una expresión reducida para la función de distribución $F(x)$ de una variable aleatoria con distribución beta(a, b) para

$$a) a > 0, b = 1.$$

$$b) a = 1, b > 0.$$

400. **Moda.** Sea X una v.a. con distribución beta(a, b). Demuestre que si $a > 1$ y $b > 1$, entonces X tiene una única moda dada por

$$x^* = \frac{a-1}{a+b-2}.$$

3.12. Distribución Weibull

La variable aleatoria continua X tiene una distribución Weibull³ con parámetros $\alpha > 0$ y $\lambda > 0$ si su función de densidad está dada por la siguiente expresión.

$$f(x) = \begin{cases} \lambda\alpha(\lambda x)^{\alpha-1} e^{-(\lambda x)^\alpha} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

³Ernst Hjalmar Waloddi Weibull (1887–1979), matemático e ingeniero sueco.

A la constante α se le llama parámetro de forma y a λ se le llama parámetro de escala. Se escribe $X \sim \text{Weibull}(\alpha, \lambda)$. La gráfica de la función de densidad para varios valores de sus parámetros se encuentra en la Figura 3.16 y su evaluación en el paquete R se obtiene usando el siguiente comando.

```
# dweibull(x,α,λ) evalúa f(x) de la distribución Weibull(α,λ)
> dweibull(2,8,2)
[1] 1.471518
```

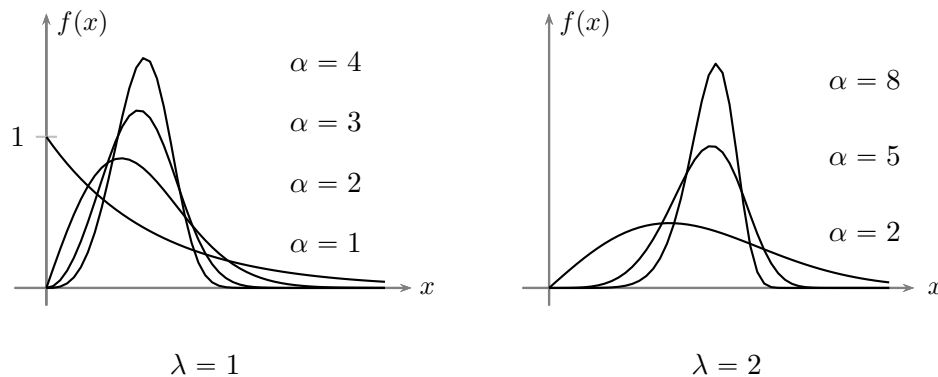


Figura 3.16

Llevando a cabo un cambio de variable (véase el Ejercicio 402) puede demostrarse que la correspondiente función de distribución adquiere la siguiente forma simple.

$$F(x) = \begin{cases} 1 - e^{-(\lambda x)^\alpha} & \text{si } x > 0, \\ 0 & \text{en otro caso,} \end{cases}$$

cuyos valores pueden calcularse en R mediante el siguiente comando.

```
# pweibull(x,α,λ) evalúa F(x) de la distribución Weibull(α,λ)
> pweibull(2,8,2)
[1] 0.6321206
```

Aplicando la definición de la función gamma, y después de algunos cálculos,

puede encontrarse que la esperanza y varianza de una variable aleatoria X con distribución Weibull(α, λ) son

$$\text{a) } E(X) = \frac{1}{\lambda} \Gamma(1 + 1/\alpha).$$

$$\text{b) } \text{Var}(X) = \frac{1}{\lambda^2} (\Gamma(1 + 2/\alpha) - \Gamma^2(1 + 1/\alpha)).$$

La distribución Weibull se ha utilizado en estudios de confiabilidad y durabilidad de componentes electrónicos y mecánicos. El valor de una variable aleatoria con esta distribución puede interpretarse como el tiempo de vida útil que tiene uno de estos componentes. Cuando el parámetro α toma el valor uno, la distribución Weibull se reduce a la distribución exponencial de parámetro λ . Por otro lado, cuando $\alpha = 2$ y $\lambda = 1/\sqrt{2}\sigma^2$ se obtiene la distribución Rayleigh(σ), la cual se menciona explícitamente en el Ejercicio 256, en la página 181. En R se pueden generar valores al azar de la distribución Weibull usando el siguiente comando.

```
# rweibull(k,α,λ) genera k valores al azar de la distribución
# Weibull(α,λ)
> rweibull(5,8,2)
[1] 1.817331 1.768006 1.993703 1.915803 2.026141
```

Ejercicios

401. Demuestre que la función de densidad Weibull efectivamente lo es.
402. Sea X con distribución Weibull(α, λ). La correspondiente función de distribución es, para $x > 0$,

$$F(x) = \int_0^x \lambda \alpha (\lambda y)^{\alpha-1} e^{-(\lambda y)^\alpha} dy.$$

Demuestre que

$$F(x) = \begin{cases} 1 - e^{-(\lambda x)^\alpha} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

403. Sea X una v.a. con distribución Weibull(α, λ). Demuestre que

$$a) E(X) = \frac{1}{\lambda} \Gamma(1 + 1/\alpha).$$

$$b) E(X^2) = \frac{1}{\lambda^2} \Gamma(1 + 2/\alpha).$$

$$c) E(X^3) = \frac{1}{\lambda^3} \Gamma(1 + 3/\alpha).$$

$$d) \text{Var}(X) = \frac{1}{\lambda^2} (\Gamma(1 + 2/\alpha) - \Gamma^2(1 + 1/\alpha)).$$

404. **Moda.** Sea X una v.a. con distribución Weibull(α, λ). Demuestre que si $\alpha > 1$ entonces X tiene una única moda dada por

$$x^* = \frac{1}{\lambda} \left(\frac{\alpha - 1}{\alpha} \right)^{1/\alpha}.$$

405. **Momentos.** Sea X una v.a. con distribución Weibull(α, λ). Demuestre que el n -ésimo momento de X es

$$E(X^n) = \frac{1}{\lambda^n} \Gamma(1 + n/\alpha).$$

406. Invertiendo la función de distribución que aparece en el Ejercicio 402, encuentre el cuantil al 100 p % para una distribución Weibull(α, λ).

407. **Simulación.** Sea U una variable aleatoria con distribución unif(0, 1) y sean $\alpha > 0$ y $\lambda > 0$ dos constantes. Demuestre la afirmación que aparece abajo. Este resultado permite obtener valores al azar de la distribución Weibull a partir de valores al azar de la distribución uniforme.

$$\frac{1}{\lambda} (-\ln(1 - U))^{1/\alpha} \sim \text{Weibull}(\alpha, \lambda).$$

3.13. Distribución normal

Esta es posiblemente la distribución de probabilidad de mayor importancia. Decimos que la variable aleatoria continua X tiene una distribución normal si su función de densidad está dada por la siguiente expresión

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}, \quad x \in \mathbb{R},$$

en donde $\mu \in \mathbb{R}$ y $\sigma > 0$ son dos parámetros. A esta distribución se le conoce también con el nombre de distribución gaussiana⁴. Escribimos entonces $X \sim N(\mu, \sigma^2)$. La gráfica de esta función de densidad tiene forma de campana, como se puede apreciar en la Figura 3.17, en donde se muestra además el significado geométrico de los dos parámetros. En ocasiones, a la gráfica de esta función se le refiere como la campana gaussiana.

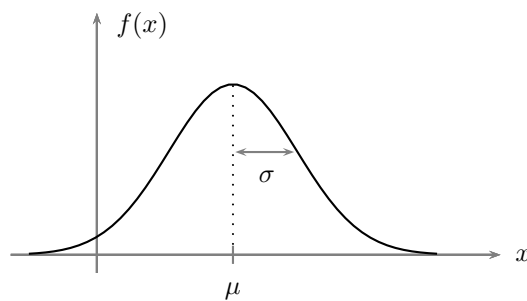


Figura 3.17

En el paquete **R**, la evaluación de la función de densidad puede obtenerse usando el siguiente comando, aunque debe observarse que se usa la desviación estándar σ en su parametrización y no σ^2 .

```
# dnorm(x, mu, sigma) evalúa f(x) de la distribución N(mu, sigma^2)
# Observe que R usa el parámetro sigma y no sigma^2
> dnorm(1.5, 3, 2)
[1] 0.1505687
```

La correspondiente función de distribución es

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/2\sigma^2} dy,$$

pero resulta que esta integral es imposible de resolver y no puede encontrarse una expresión cerrada. Más adelante explicaremos la forma en que se han encontrado valores aproximados para esta integral. En **R** es muy sencillo obtener tales evaluaciones aproximadas utilizando el siguiente comando.

⁴Carl Friedrich Gauss (1777-1855), matemático alemán.

```
# pnorm(x,μ,σ) evalúa F(x) de la distribución N(μ,σ²)
# Observe que R usa el parámetro σ y no σ²
> pnorm(1.5,3,2)
[1] 0.2266274
```

Por otro lado, usando integración por partes, es posible demostrar que para una variable aleatoria X con distribución $N(\mu, \sigma^2)$,

- a) $E(X) = \mu$.
- b) $\text{Var}(X) = \sigma^2$.

Esto significa que la campana está centrada en el valor del parámetro μ , el cual puede ser negativo, positivo o cero, y que la campana se abre o se cierra de acuerdo a la magnitud del parámetro σ^2 . El siguiente caso particular de la distribución normal es muy importante.

Distribución normal estándar

Decimos que la variable aleatoria X tiene una distribución normal estándar si tiene una distribución normal con parámetros $\mu = 0$ y $\sigma^2 = 1$. En este caso, la función de densidad se reduce a la expresión

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathbb{R}.$$

El resultado importante aquí es que siempre es posible transformar una variable aleatoria normal no estándar en una estándar mediante la siguiente operación, cuya demostración se pide hacer en el Ejercicio 416.

Proposición 3.6 Sea X con distribución $N(\mu, \sigma^2)$. Entonces

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1). \quad (3.6)$$

Al procedimiento anterior se le conoce con el nombre de estandarización y, bajo tal transformación, se dice que la variable X ha sido estandarizada. Este

resultado parece muy modesto pero tiene una gran importancia operacional pues establece que el cálculo de las probabilidades de una variable aleatoria normal cualquiera se reduce al cálculo de las probabilidades para la normal estándar. Explicaremos ahora con más detalle esta situación. Suponga que X es una variable aleatoria con distribución $N(\mu, \sigma^2)$ y que deseamos calcular, por ejemplo, la probabilidad de que X tome un valor en el intervalo (a, b) , es decir, $P(a < X < b)$. Como en el enunciado anterior de la proposición anterior, Z denotará una v.a. con distribución normal estándar. Tenemos entonces que

$$\begin{aligned} P(a < X < b) &= P(a - \mu < X - \mu < b - \mu) \\ &= P\left(\frac{a - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{b - \mu}{\sigma}\right) \\ &= P\left(\frac{a - \mu}{\sigma} < Z < \frac{b - \mu}{\sigma}\right). \end{aligned}$$

Cada una de las igualdades anteriores es consecuencia de la igualdad de los eventos correspondientes. De esta forma, una probabilidad que involucra a la variable X se ha reducido a una probabilidad que involucra a Z . De modo que únicamente necesitamos conocer las probabilidades de los eventos de Z para calcular las probabilidades de los eventos de la variable X , que tiene parámetros arbitrarios. En términos de integrales, el cálculo anterior es equivalente al siguiente análisis, en donde se lleva a cabo el cambio de variable $y = (x - \mu)/\sigma$ en la integral,

$$\begin{aligned} P(a < X < b) &= \int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} dx \\ &= \int_{(a-\mu)/\sigma}^{(b-\mu)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \\ &= P\left(\frac{a - \mu}{\sigma} < Z < \frac{b - \mu}{\sigma}\right). \end{aligned}$$

A partir de ahora, y a menos de que se diga lo contrario, usaremos la letra Z para denotar a una variable aleatoria con distribución normal estándar.

Función de distribución $N(0, 1)$

Es común denotar a la función de distribución de una variable aleatoria normal estándar como $\Phi(x)$, es decir,

$$\Phi(x) = P(Z \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du,$$

cuyo significado geométrico se muestra en la Figura 3.18 (a). Como hemos mencionado antes, no es posible resolver esta integral y para evaluarla se usan métodos numéricos para aproximar $\Phi(x)$ para distintos valores de x . Aunque en el paquete **R** pueden encontrarse estos valores con el comando `pnorm(x, 0, 1)`, en la parte final del texto aparece una tabla con estos valores aproximados. Cada renglón de esta tabla corresponde a un valor de x hasta el primer dígito decimal, las distintas columnas corresponden al segundo dígito decimal. El valor que aparece en la tabla es $\Phi(x)$. Por ejemplo, el renglón marcado con 1.4 y la columna marcada con 0.05 corresponden al valor $x = 1.45$, tenemos entonces que $\Phi(1.45) = 0.9265$. Abajo aparecen algunos ejemplos que ilustran el uso de esta tabla. Observe además que para $x \geq 3.5$, la probabilidad $\Phi(x)$ es muy cercana a uno, es decir, para esos valores de x la campana prácticamente ha decaído a cero en el lado derecho. Esto quiere decir que, con probabilidad cercana a uno, los valores que toma una variable aleatoria normal estándar están comprendidos entre -3.5 y $+3.5$.

Por otro lado, a partir del hecho de que si X tiene distribución normal estándar, entonces la variable $-X$ también tiene distribución normal estándar, puede demostrarse que

$$\Phi(-x) = 1 - \Phi(x).$$

Un argumento geométrico también puede utilizarse para darse cuenta de la validez de esta igualdad. En particular, este resultado ayuda a calcular valores de $\Phi(x)$ para x negativos en tablas de la distribución normal como la presentada al final del texto, en donde sólo aparecen valores positivos para x .

Ejemplo 3.15 Use la tabla de la distribución normal estándar para comprobar que

1. $\Phi(1.65) = 0.9505$.
2. $\Phi(-1.65) = 0.0495$.

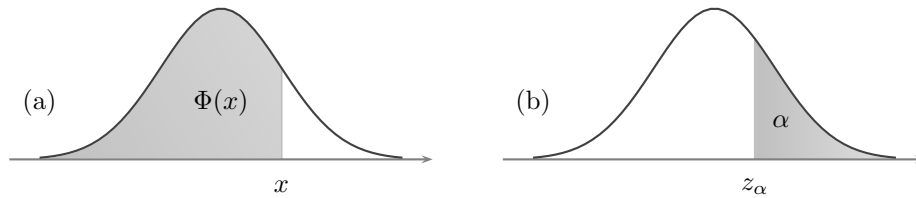


Figura 3.18

3. $\Phi(-1) = 0.1587$.

•

Ejemplo 3.16 Use la tabla de la distribución normal estándar para encontrar un valor aproximado de x tal que

1. $\Phi(x) = 0.3$.
2. $\Phi(x) = 0.75$.

Respuestas: $\begin{cases} 1. & x = -0.53. \\ 2. & x = 0.68. \end{cases}$

•

Ejemplo 3.17 Sea X una v.a. con distribución $N(5, 10)$. Use el proceso de estandarización y la tabla de la distribución normal estándar para comprobar que

1. $P(X \leq 7) = 0.7357$.
2. $P(0 < X < 5) = 0.2357$.
3. $P(X > 10) = 0.0571$.

•

A continuación definiremos el número z_α , el cual es usado con regularidad en las aplicaciones de la distribución normal.

Notación z_α . Para cada valor de α en el intervalo $(0, 1)$, el número z_α denotará el cuantil al $100(1 - \alpha)\%$ de la distribución normal estándar, es decir,

$$\Phi(z_\alpha) = 1 - \alpha.$$

El significado geométrico del número z_α se muestra en la Figura 3.18 (b), en la página 274.

Ejemplo 3.18 Usando la tabla de la distribución normal estándar puede comprobarse que, de manera aproximada,

a) $z_{0.1} = 1.285$.

b) $z_{0.2} = 0.845$.

•

Finalmente, señalaremos que en R se pueden generar valores al azar de la distribución normal haciendo uso del siguiente comando.

```
# rnorm(k, μ, σ) genera k valores al azar de la distribución
# N(μ, σ²). Observe que R usa el parámetro σ y no σ²
> rnorm(5, 3, 2)
[1] 3.0408942 0.5529831 2.3426471 2.0050003 0.4448412
```

Ejercicios

408. Demuestre que la función de densidad normal con parámetros μ y σ^2

a) efectivamente es una función de densidad.

- b) tiene un máximo absoluto en $x = \mu$. Esta es la moda de la distribución.
- c) tiene puntos de inflexión en $x = \mu \pm \sigma$.
409. Sea X una v.a. con distribución $N(\mu, \sigma^2)$. Determine, de manera aproximada, el valor de la constante $c > 0$ tal que se satisfaga la identidad que aparece abajo. Interprete este resultado.

$$P(\mu - c\sigma < X < \mu + c\sigma) = 0.99.$$

410. Demuestre que para cualquier $x > 0$,

$$\frac{x}{1+x^2} e^{-x^2/2} \leq \int_x^\infty e^{-u^2/2} du \leq \frac{1}{x} e^{-x^2/2}.$$

411. Sea X con distribución $N(\mu, \sigma^2)$. Demuestre que

- a) $E(X) = \mu$.
- b) $E(X^2) = \mu^2 + \sigma^2$.
- c) $\text{Var}(X) = \sigma^2$.

412. Encuentre la moda y mediana de la distribución $N(\mu, \sigma^2)$.

413. **f.g.m.** Sea X una v.a. con distribución $N(\mu, \sigma^2)$. Demuestre que la función generadora de momentos de X es la función que aparece abajo. Usando esta función y sus propiedades encuentre nuevamente la media y la varianza de esta distribución.

$$M(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}.$$

414. **Momentos (normal centrada).** Sea X una v.a. con distribución $N(0, \sigma^2)$. Demuestre que

$$E(X^n) = \begin{cases} \frac{n!}{(n/2)!} \left(\frac{\sigma^2}{2}\right)^{n/2} & \text{si } n \text{ es par,} \\ 0 & \text{si } n \text{ es impar.} \end{cases}$$

415. Sean X_1 y X_2 independientes con distribución $N(\mu_1, \sigma_1^2)$ y $N(\mu_2, \sigma_2^2)$. Use la función generadora de momentos para demostrar que

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

416. **Estandarización.** Calculando primero la función de distribución y después derivando para encontrar la función de densidad, o bien usando la f.g.m., demuestre los siguientes dos resultados.

- a) Si $X \sim N(\mu, \sigma^2)$ entonces $Z = (X - \mu)/\sigma \sim N(0, 1)$.
 b) Si $Z \sim N(0, 1)$ entonces $X = \mu + \sigma Z \sim N(\mu, \sigma^2)$.

417. Sean $a < b$ dos constantes positivas y sea Z con distribución normal estándar. Demuestre que

$$P(a < Z^2 < b) = 2(\Phi(\sqrt{b}) - \Phi(\sqrt{a})).$$

418. Sea X una v.a. con distribución $N(5, 10)$. Obtenga las siguientes probabilidades en términos de la función $\Phi(x)$.

- a) $P(X \leq 7)$.
 b) $P(X > 4)$.
 c) $P(|X - 2| \leq 3)$.
 d) $P(|X - 6| > 1)$.

419. Sea X una v.a. con distribución $N(10, 36)$. Calcule

- a) $P(X > 5)$.
 b) $P(4 < X < 16)$.
 c) $P(X \leq 8)$.
 d) $P(X \geq 16)$.
 e) $P(|X - 4| \leq 6)$.
 f) $P(|X - 6| > 3)$.

420. Suponga que el tiempo de vida útil X , medido en horas, de un componente electrónico se puede modelar de manera aproximada mediante una variable aleatoria con distribución normal con parámetros $\mu = 20,000$ hrs. y $\sigma = 500$ hrs.

- a) ¿Cuál es la probabilidad de que el componente dure más de 21,000 horas?

Se trata de una variable aleatoria continua con posibles valores en el intervalo $(0, \infty)$. Esta distribución tiene sólo un parámetro, denotado aquí por la letra n , al cual se le llama grados de libertad. A pesar de su aparente expresión complicada, no es difícil comprobar que $f(x)$ es, efectivamente, una función de densidad, y para ello se utiliza la definición de la función gamma. La gráfica de esta función de densidad, para varios valores de su parámetro, aparece en la Figura 3.19. Sus valores se pueden calcular en el paquete R de la forma siguiente.

```
# dchisq(x,n) evalúa  $f(x)$  de la distribución  $\chi^2(n)$ 
> dchisq(2,3)
[1] 0.2075537
```

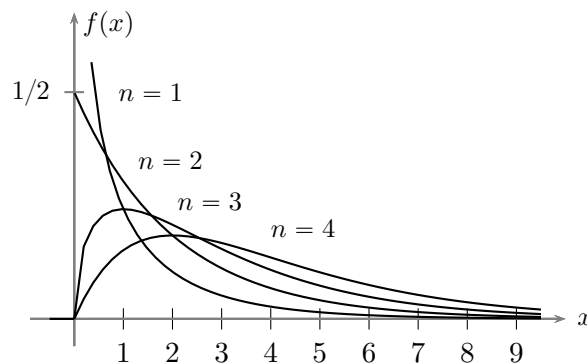


Figura 3.19

Escribiremos simplemente $X \sim \chi^2(n)$, en donde la letra griega χ se pronuncia “ji”. Por lo tanto, la expresión “ $\chi^2(n)$ ” se lee “ji cuadrada con n grados de libertad”. Es interesante observar que la distribución $\chi^2(n)$ se obtiene de la distribución gamma(α, λ) cuando $\alpha = n/2$ y $\lambda = 1/2$. La expresión para la función de distribución no tiene una forma reducida: para $x > 0$,

$$F(x) = \int_0^x \frac{1}{\Gamma(n/2)} \left(\frac{1}{2}\right)^{n/2} u^{n/2-1} e^{-u/2} du,$$

y sus valores pueden calcularse en R de la forma que aparece en el siguiente

recuadro. Alternativamente en la parte final de este texto se encuentra una tabla con algunas de estas probabilidades.

```
# pchisq(x,n) evalúa  $F(x)$  de la distribución  $\chi^2(n)$ 
> pchisq(2,3)
[1] 0.4275933
```

A través de la construcción de una distribución χ^2 con un parámetro adecuado en el integrando correspondiente, puede demostrarse, sin mucha dificultad, que

a) $E(X) = n$.

b) $\text{Var}(X) = 2n$.

La distribución ji-cuadrada puede obtenerse como indican los varios resultados que a continuación se enuncian, cuyas demostraciones se pide desarrollar en la sección de ejercicios.

Proposición 3.7 Si $X \sim N(0, 1)$, entonces

$$X^2 \sim \chi^2(1).$$

Es decir, el cuadrado de una variable aleatoria con distribución normal estándar tiene distribución ji-cuadrada con un grado de libertad. Vea el Ejercicio 425. Por otro lado, el siguiente resultado establece que la suma de dos variables aleatorias independientes con distribución ji-cuadrada tiene distribución nuevamente ji-cuadrada con grados de libertad la suma de los grados de libertad de los sumandos.

Proposición 3.8 Sean X y Y dos variables aleatorias independientes con distribución $\chi^2(n)$ y $\chi^2(m)$, respectivamente, entonces

$$X + Y \sim \chi^2(n + m).$$

En el Ejercicio 432 se pide usar la f.g.m. para demostrar este resultado, el cual puede extenderse al caso cuando se tienen varias variables aleatorias independientes con distribución χ^2 . En particular, si X_1, \dots, X_n son variables independientes con distribución normal estándar, entonces la suma de los cuadrados $X_1^2 + \dots + X_n^2$ tiene distribución $\chi^2(n)$. De este modo, si conocemos una forma de simular n valores al azar de la distribución normal estándar, la suma de los cuadrados de los números obtenidos será una observación de la distribución ji-cuadrada con n grados de libertad. Por último, mencionaremos el siguiente resultado, el cual es utilizado en algunos procedimientos estadísticos.

Proposición 3.9 Sean X_1, \dots, X_n variables aleatorias independientes, cada una de ellas con distribución $N(\mu, \sigma^2)$. Entonces

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1),$$

en donde $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ y $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$.

Para concluir esta sección mencionaremos que a través del siguiente comando en R se pueden obtener valores al azar de la distribución χ^2 .

```
# rchisq(k,n) genera k valores al azar de la distribución
#  $\chi^2(n)$ 
> rchisq(5,3)
[1] 2.5946656 6.9019593 0.7172345 4.5362704 0.7995995
```

Ejercicios

424. Compruebe que la función de densidad de la distribución $\chi^2(n)$ efectivamente lo es.
425. Sea X una variable aleatoria continua. Demuestre que

a) $F_{X^2}(x) = F_X(\sqrt{x}) - F_X(-\sqrt{x})$ para $x > 0$.

b) si $X \sim N(0, 1)$, entonces $X^2 \sim \chi^2(1)$.

426. Sea X una v.a. con distribución $\chi^2(n)$ y sea $c > 0$ una constante. Defina los parámetros $\alpha = n/2$ y $\lambda = 1/(2c)$. Demuestre que

$$cX \sim \text{gamma}(\alpha, \lambda).$$

427. Compruebe que la distribución $\text{gamma}(\alpha, \lambda)$, en donde $\alpha = n/2$ con $n \in \mathbb{N}$ y $\lambda = 1/2$, se reduce a la distribución $\chi^2(n)$.

428. Sea X una v.a. con distribución $\chi^2(n)$. Demuestre que

a) $E(X) = n$.

c) $E(X^3) = n(n+2)(n+4)$.

b) $E(X^2) = n(n+2)$.

d) $\text{Var}(X) = 2n$.

429. **Moda.** Sea X una v.a. con distribución $\chi^2(n)$. Demuestre que si $n > 2$ entonces X tiene una única moda dada por

$$x^* = n - 2.$$

430. **Momentos.** Sea X una v.a. con distribución $\chi^2(n)$. Demuestre que el m -ésimo momento de X es

$$E(X^m) = 2^m \frac{\Gamma(n/2 + m)}{\Gamma(n/2)} = n(n+2) \cdots (n+2m-2).$$

431. **f.g.m.** Sea X una v.a. con distribución $\chi^2(n)$. Demuestre que la función generadora de momentos de X es la función que aparece abajo. Use esta función y sus propiedades para encontrar nuevamente la esperanza y la varianza de esta distribución.

$$M(t) = \left(\frac{1}{1-2t} \right)^{n/2} \quad \text{para } t < 1/2.$$

432. **Suma.** Use la f.g.m. para demostrar que si X y Y son variables aleatorias independientes con distribución $\chi^2(n)$ y $\chi^2(m)$, respectivamente, entonces

$$X + Y \sim \chi^2(n + m).$$

433. Sea U una v.a. con distribución $\text{unif}(0, 1)$. Demuestre que

$$-2 \ln(U) \sim \chi^2(2),$$

en donde la distribución $\chi^2(2)$ coincide con $\exp(1/2)$.

3.15. Distribución t

Decimos que la variable aleatoria continua X tiene una distribución t de Student⁵ con $n > 0$ grados de libertad si su función de densidad está dada por la siguiente expresión

$$f(x) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi} \Gamma(n/2)} (1 + x^2/n)^{-(n+1)/2}, \quad x \in \mathbb{R}.$$

En tal caso se escribe $X \sim t(n)$, en donde n es un número real positivo, aunque tomaremos principalmente el caso cuando n es entero positivo. La gráfica de esta función de densidad aparece en la Figura 3.20 y sus valores pueden calcularse en el paquete R con ayuda del siguiente comando.

```
# dt(x,n) evalúa f(x) de la distribución t(n)
> dt(1,3)
[1] 0.2067483
```

En la Figura 3.20 puede apreciarse el parecido de la función de densidad $t(n)$ con la función de densidad normal estándar. En esta misma figura está graficada también la función de densidad normal estándar, pero ésta se empalma completamente con la función de densidad $t(n)$ con $n = 100$. En el límite cuando $n \rightarrow \infty$, ambas densidades coinciden, véase el Ejercicio 440. La función de distribución no tiene una expresión simple y la dejaremos indicada como la integral correspondiente, es decir,

$$F(x) = \int_{-\infty}^x \frac{\Gamma((n+1)/2)}{\sqrt{n\pi} \Gamma(n/2)} (1 + u^2/n)^{-(n+1)/2} du,$$

cuyos valores pueden encontrarse en una tabla al final del texto o bien en R mediante el siguiente comando.

⁵Sudónimo de William Sealy Gosset (1876-1937), estadístico inglés.

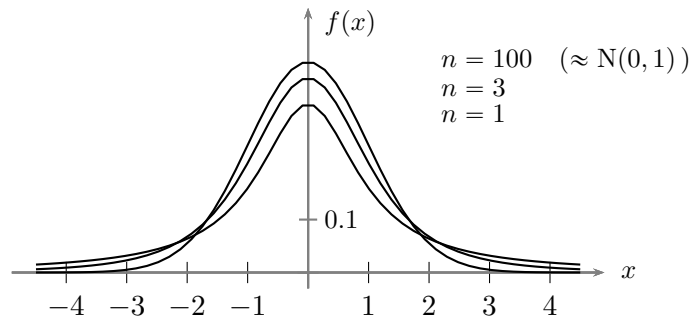


Figura 3.20

```
# pt(x,n) evalúa F(x) de la distribución t(n)
> pt(1,3)
[1] 0.8044989
```

Llevando a cabo la integral correspondiente se puede demostrar que

a) $E(X) = 0$ si $n > 1$.

b) $\text{Var}(X) = \frac{n}{n-2}$ si $n > 2$.

La distribución t es un ejemplo de distribución para la cual no existe la función generadora de momentos. Esta distribución se puede encontrar cuando se estudian ciertas operaciones entre otras variables aleatorias. Por simplicidad en la exposición, omitiremos la demostración de los siguientes resultados.

Proposición 3.10 Si $X \sim N(0, 1)$ y $Y \sim \chi^2(n)$ son dos v.a.s independientes, entonces

$$\frac{X}{\sqrt{Y/n}} \sim t(n).$$

En el estudio y aplicación de la estadística matemática se necesitan realizar operaciones como la indicada en la proposición anterior. Por otro lado, este resultado sugiere un mecanismo para generar simulaciones de los valores que toma una variable aleatoria con distribución $t(n)$. Para ello se pueden generar n observaciones de la distribución normal estándar, y conformar una observación de la distribución $\chi^2(n)$ como fue explicado antes. Se necesita una observación adicional de la distribución normal estándar, que será el valor de X según la fórmula de la proposición anterior, se hace el cociente indicado y el resultado será un valor de la distribución $t(n)$. La distribución t aparece también en el siguiente contexto.

Proposición 3.11 Sean X_1, \dots, X_n variables aleatorias independientes, cada una de ellas con distribución $N(\mu, \sigma^2)$. Entonces

$$\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1),$$

en donde $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ y $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$.

Por último, mencionaremos que se pueden generar valores al azar de la distribución t en el paquete **R** usando el siguiente comando.

```
# rt(k,n) genera k valores al azar de la distribución t(n)
> rt(5,3)
[1] 0.06769745 -0.33693291 -0.36182444 1.68520735 -0.02326697
```

Ejercicios

434. Demuestre que la función de densidad $t(n)$ efectivamente lo es.

435. Sea X una v.a. con distribución $t(n)$. Demuestre que

a) $E(X) = 0$ si $n > 1$.

$$b) E(X^2) = \frac{n}{n-2} \quad \text{si } n > 2.$$

$$c) \text{Var}(X) = \frac{n}{n-2} \quad \text{si } n > 2.$$

436. **Fórmula recursiva para momentos pares.** Sea X_n una v.a. con distribución $t(n)$ para cada $n > 2$. Demuestre que si m es un número par tal que $0 < m < n$, entonces se cumple la siguiente fórmula recursiva

$$E(X_n^m) = \left(\frac{n}{n-2}\right)^{m/2} (m-1) E(X_{n-2}^{m-2}).$$

En consecuencia,

$$\begin{aligned} E(X_n^m) &= \left(\frac{n}{n-2}\right)^{m/2} (m-1) \left(\frac{n-2}{n-4}\right)^{(m-2)/2} (m-3) \cdots \\ &\quad \cdots \left(\frac{n-m+2}{n-m}\right) (1) \\ &= \frac{(m-1)(m-3)\cdots 1}{(n-2)(n-4)\cdots(n-m)} n^{m/2}. \end{aligned} \quad (1)$$

437. **Momentos.** Sea X una v.a. con distribución $t(n)$ con $n > 2$ y sea $m \geq 1$ un entero. Demuestre que el m -ésimo momento de X es

$$E(X^m) = \begin{cases} 0 & \text{si } m \text{ es impar y } 2 \leq m < n, \\ \frac{\Gamma((m+1)/2) \Gamma((n-m)/2) n^{m/2}}{\sqrt{\pi} \Gamma(n/2)} & \text{si } m \text{ es par y } 2 \leq m < n, \\ \text{"no existe"} & \text{si } m \geq n. \end{cases}$$

438. **Moda y mediana** Demuestre que la mediana de una variable aleatoria con distribución $t(n)$ es cero y que tiene una única moda también en cero.

439. **No existencia de la f.g.m.** Demuestre que no existe la función generadora de momento para la distribución $t(n)$.

440. **Convergencia a la distribución normal estándar.** Sea $f(x)$ la función de densidad $t(n)$. Demuestre que

$$\lim_{n \rightarrow \infty} f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

3.16. Distribución F

Se dice que la variable aleatoria continua X tiene una distribución F de Fisher⁶ - Snedecor⁷ con parámetros $a > 0$ y $b > 0$ si su función de densidad está dada por la siguiente expresión.

$$f(x) = \begin{cases} \frac{\Gamma(\frac{a+b}{2})}{\Gamma(\frac{a}{2})\Gamma(\frac{b}{2})} \left(\frac{a}{b}\right)^{a/2} x^{a/2-1} \left(1 + \frac{a}{b}x\right)^{-(a+b)/2} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

En este caso se escribe $X \sim F(a, b)$. Una gráfica de esta función de densidad aparece en la Figura 3.21 y sus valores pueden ser calculados en el paquete R usando el siguiente comando.

```
# df(x,n) evalúa f(x) de la distribución F(a,b)
> df(0.5,4,10)
[1] 0.669796
```

La función de distribución no tiene una expresión reducida y la indicaremos simplemente como la integral correspondiente, es decir, para $x > 0$,

$$F(x) = \int_0^x \frac{\Gamma(\frac{a+b}{2})}{\Gamma(\frac{a}{2})\Gamma(\frac{b}{2})} \left(\frac{a}{b}\right)^{a/2} u^{a/2-1} \left(1 + \frac{a}{b}u\right)^{-(a+b)/2} du,$$

cuyos valores en R pueden encontrarse usando el siguiente comando.

```
# pf(x,n) evalúa F(x) de la distribución F(a,b)
> pf(0.5,4,10)
[1] 0.2632245
```

Para esta distribución se puede demostrar que su esperanza y varianza son

$$\text{a) } E(X) = \frac{b}{b-2}, \quad \text{si } b > 2.$$

$$\text{b) } \text{Var}(X) = \frac{2b^2(a+b-2)}{a(b-2)^2(b-4)}, \quad \text{si } b > 4.$$

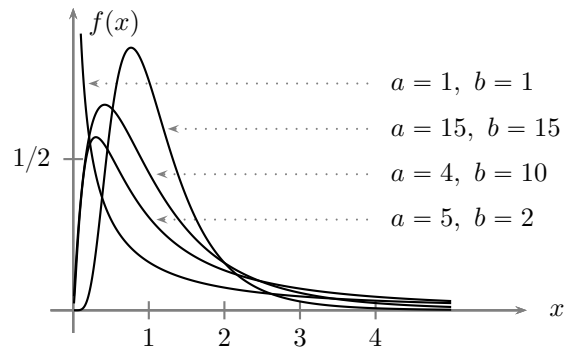


Figura 3.21

La distribución F aparece como resultado de la siguiente operación entre variables aleatorias.

Proposición 3.12 Sean $X \sim \chi^2(a)$ y $Y \sim \chi^2(b)$ dos variables aleatorias independientes. Entonces

$$\frac{X/a}{Y/b} \sim F(a, b).$$

Por último, mencionaremos que se pueden generar valores al azar de la distribución F en el paquete R usando un comando similar a los anteriores.

```
# rf(k,a,b) genera k valores al azar de la distribución F(a,b)
> rf(5,4,10)
[1] 0.57341208 0.36602858 1.05682859 0.08009087 3.80035154
```

⁶Ronald Aylmer Fisher (1890-1962), estadístico inglés.

⁷George Waddel Snedecor (1881-1974), matemático y estadístico estadounidense.

Ejercicios

441. Demuestre que la función de densidad $F(a, b)$ es, efectivamente, una función de densidad.

442. Sea X una v.a. con distribución $F(a, b)$. Demuestre que

$$a) E(X) = \frac{b}{b-2}, \quad \text{si } b > 2.$$

$$b) E(X^2) = \frac{b^2(a+2)}{a(b-2)(b-4)}, \quad \text{si } b > 4.$$

$$c) \text{Var}(X) = \frac{2b^2(a+b-2)}{a(b-2)^2(b-4)}, \quad \text{si } b > 4.$$

443. **Moda.** Sea X una v.a. con distribución $F(a, b)$. Demuestre que si $a > 2$ entonces X tiene una única moda dada por

$$x^* = \frac{b(a-2)}{a(b+2)}.$$

444. **Momentos.** Sea X una v.a. con distribución $F(a, b)$. Demuestre que el n -ésimo momento de X , para $2n < b$, es

$$E(X^n) = \left(\frac{b}{a}\right)^n \frac{\Gamma(a/2 + n)}{\Gamma(a/2)} \frac{\Gamma(b/2 - n)}{\Gamma(b/2)}.$$

445. **No existencia de la f.g.m.** Demuestre que la función generadora de momentos de la distribución $F(a, b)$ no existe.

446. Demuestre que si $X \sim t(n)$ entonces

$$X^2 \sim F(1, n).$$

447. Demuestre que si $X \sim F(a, b)$ entonces

$$\frac{1}{X} \sim F(b, a).$$

§

Con esto concluimos una revisión elemental de algunas distribuciones de probabilidad continuas. Se debe recordar que existen muchas más distribuciones de este tipo, algunas más conocidas que otras, pero todas ellas útiles como modelos probabilísticos en las muy diversas áreas de aplicación de la probabilidad. En el siguiente capítulo se estudiaremos varias variables aleatorias a la vez. Generalmente se refiere a ellas como variables aleatorias multidimensionales, o bien como vectores aleatorios. Sus correspondientes distribuciones adquieren el nombre de distribuciones multivariadas.

Capítulo 4

Vectores aleatorios

Este capítulo contiene una breve introducción al tema de variables aleatorias multidimensionales, también llamadas vectores aleatorios. Para hacer la escritura corta se consideran únicamente vectores aleatorios de dimensión dos, aunque las definiciones y resultados que se mencionan pueden extenderse fácilmente, en la mayoría de los casos, para vectores de dimensión superior. Para el material que se presenta a continuación sería provechoso contar con algunos conocimientos elementales del cálculo diferencial e integral en varias variables o por lo menos mantener la calma cuando parezca que los símbolos matemáticos no tienen ningún sentido.

4.1. Vectores aleatorios

Definición 4.1 Un vector aleatorio de dimensión dos es un vector de la forma (X, Y) en donde cada coordenada es una variable aleatoria. De manera análoga se definen vectores aleatorios multidimensionales (X_1, \dots, X_n) .

Se dice que un vector aleatorio es discreto o continuo, si todas las variables aleatorias que lo conforman son de ese mismo tipo. Por simplicidad, en la

mayoría de los casos consideraremos vectores aleatorios cuyas coordenadas son variables aleatorias todas discretas o continuas, y en pocos casos, combinaciones de ellas. Un vector aleatorio (X, Y) puede considerarse como una función de Ω en \mathbb{R}^2 como se muestra en la Figura 4.1.

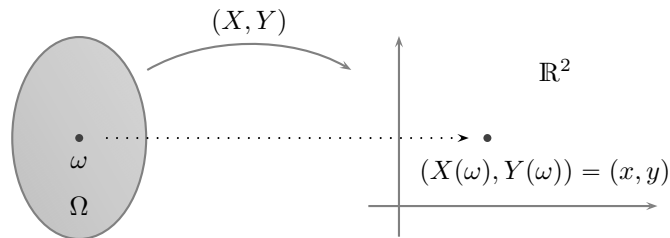


Figura 4.1

Es decir, el vector (X, Y) evaluado en ω es $(X, Y)(\omega) = (X(\omega), Y(\omega))$ con posible valor (x, y) . Nuevamente observe que el vector con letras mayúsculas (X, Y) es el vector aleatorio, mientras que el vector con letras minúsculas (x, y) es un punto en el plano. Así, el vector $(X(\omega), Y(\omega))$ representa la respuesta conjunta de dos preguntas o mediciones efectuadas a un mismo elemento ω del espacio muestral Ω . A veces la información de la que se dispone acerca de un fenómeno está agrupada de esta forma. En nuestro caso hemos mencionado únicamente dos variables aleatorias, pero vectores de dimensión mayor son posibles.

Ejemplo 4.1 Suponga que tenemos una población de mujeres y que la variable X toma el valor 1 cuando la mujer es fumadora y cero cuando no lo es. Sea Y la variable que registra el número de hijos de una mujer dada. Entonces el vector (X, Y) puede tomar los valores

$$(0, 0), (0, 1), (0, 2), \dots$$

$$(1, 0), (1, 1), (1, 2), \dots$$

Para una población particular de 43 mujeres podríamos conformar una tabla con la frecuencia de cada uno de estos valores del vector, por ejemplo,

$x \backslash y$	0	1	2	3	4	5
0	2	5	8	3	2	0
1	10	9	2	0	1	1

Con ayuda de esta tabla puede definirse una distribución de probabilidad para el vector (X, Y) dividiendo cada entrada de la tabla por el número 43. De esta manera, por ejemplo, $P(X = 0, Y = 0) = 2/43$, $P(X = 0, Y = 1) = 5/43$, etcétera. •

Ejemplo 4.2 Suponga que se cuenta con una población de personas que participan en un proceso de elección. Se escoge a uno de los votantes al azar y el vector $(X(\omega), Y(\omega))$ puede representar el nivel económico y la preferencia electoral del votante ω . Varios estudios pueden llevarse a cabo con la información recabada sobre un conjunto representativo de personas en este proceso de elección. Por ejemplo, ¿existe alguna relación predictiva entre estas variables aleatorias o podemos considerar que son independientes? Existen métodos estadísticos que ayudan a responder a este tipo de preguntas. Estos procedimientos tienen como elemento base los modelos de probabilidad que estamos estudiando. •

Definiremos a continuación algunas funciones asociadas a vectores aleatorios. Estas funciones son análogas al caso unidimensional estudiado antes.

4.2. Función de probabilidad conjunta

Estudiaremos primero el caso de vectores aleatorios discretos. La situación es muy similar al caso unidimensional.

Definición 4.2 La función de probabilidad del vector aleatorio discreto (X, Y) , en donde X toma los valores x_1, x_2, \dots y Y toma los valores y_1, y_2, \dots , es la función $f(x, y) : \mathbb{R}^2 \rightarrow [0, 1]$ dada por

$$f(x, y) = \begin{cases} P(X = x, Y = y) & \text{si } (x, y) \in \{x_1, x_2, \dots\} \times \{y_1, y_2, \dots\}, \\ 0 & \text{en otro caso.} \end{cases}$$

Es decir, la función $f(x, y)$ es la probabilidad de que la variable X tome el valor x y, al mismo tiempo, la variable Y tome el valor y . Tal función se llama también función de probabilidad conjunta de las variables X y Y , y para enfatizar este hecho a veces se escribe $f_{X,Y}(x, y)$, pero en general omitiremos los subíndices para hacer la notación más corta, aunque siempre asociando el valor x a la variable X y el valor y a la variable Y . Haremos uso de los subíndices únicamente cuando sea necesario especificar las variables aleatorias en estudio. Toda función $f(x, y)$ de la forma anterior cumple las siguientes dos propiedades.

a) $f(x, y) \geq 0$.

b) $\sum_{x,y} f(x, y) = 1$.

La suma indicada en realidad es una doble suma, una sobre todos los posibles valores x y otra sobre todos los posibles valores y , no importando el orden en el que se efectúan las sumas. Inversamente, toda función definida sobre \mathbb{R}^2 que sea cero, excepto en un conjunto discreto de parejas (x, y) que cumpla estas propiedades, se llama función de probabilidad bivariada o conjunta, sin necesidad de contar con dos variables aleatorias previas que la definan.

Otra forma equivalente de presentar a la función de probabilidad de un vector discreto (X, Y) es a través de una tabla como la siguiente.

$x \setminus y$	y_1	y_2	\cdots
x_1	$f(x_1, y_1)$	$f(x_1, y_2)$	\cdots
x_2	$f(x_2, y_1)$	$f(x_2, y_2)$	\cdots
\vdots	\vdots	\vdots	

También se pueden elaborar gráficas en \mathbb{R}^3 de las funciones de probabilidad bivariadas y su aspecto general se muestra en la Figura 4.2 en el caso discreto. Ejemplos en el caso continuo aparecerán más adelante. Observe que sería difícil graficar funciones de probabilidad de vectores aleatorios de dimensión 3 y superiores.

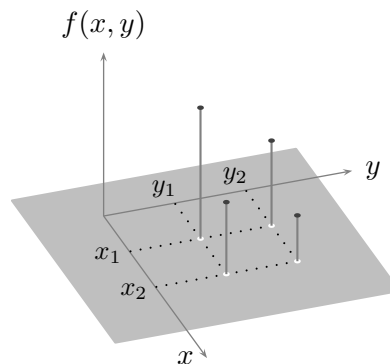


Figura 4.2

El cálculo de probabilidades de eventos relativos a un vector aleatorio discreto (X, Y) con función de probabilidad $f(x, y)$ se lleva a cabo de la siguiente forma: si A y B son dos conjuntos borelianos, entonces la probabilidad del evento $(X \in A) \cap (Y \in B)$ se calcula como sigue.

$$P(X \in A, Y \in B) = \sum_{x \in A} \sum_{y \in B} f(x, y).$$

Las sumas son además sobre aquellos valores de x y y en donde $f(x, y)$ es estrictamente positiva. Esta es la probabilidad de que la variable aleatoria X tome un valor en el conjunto A y al mismo tiempo la variable Y tome un valor en el conjunto B .

Ejemplo 4.3 Considere el vector aleatorio discreto (X, Y) con función de probabilidad dada por la siguiente tabla y cuya gráfica se muestra en la Figura 4.3.

$x \backslash y$	0	1
-1	0.3	0.1
1	0.4	0.2

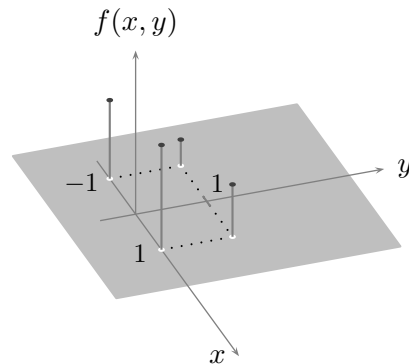


Figura 4.3

De la tabla anterior se entiende que la variable X toma valores en el conjunto $\{-1, 1\}$ mientras que Y toma valores en $\{0, 1\}$. Además, las probabilidades conjuntas están dadas por las entradas de la tabla, por ejemplo, $P(X = -1, Y = 0) = 0.3$, esto es, la probabilidad de que X tome el valor -1 y, al mismo tiempo, Y tome el valor 0 es 0.3 . La misma información puede

escribirse de la siguiente manera.

$$f(x, y) = \begin{cases} 0.3 & \text{si } x = -1, y = 0, \\ 0.1 & \text{si } x = -1, y = 1, \\ 0.4 & \text{si } x = 1, y = 0, \\ 0.2 & \text{si } x = 1, y = 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Como todos estos valores son probabilidades, naturalmente son no negativos y todos ellos suman uno. Por lo tanto, $f(x, y)$ es, efectivamente, una función de probabilidad bivariada. •

Ejemplo 4.4 Encontraremos la constante c que hace que la siguiente función sea de probabilidad conjunta.

$$f(x, y) = \begin{cases} cxy & \text{si } (x, y) \in \{1, 2\} \times \{1, 2\}, \\ 0 & \text{en otro caso.} \end{cases}$$

Los posible valores del vector (X, Y) son $(1, 1)$, $(1, 2)$, $(2, 1)$ y $(2, 2)$, con probabilidades respectivas c , $2c$, $2c$ y $4c$. Como la suma de estas probabilidades debe ser uno, se llega a la ecuación $9c = 1$, de donde se obtiene que $c = 1/9$. •

Veamos ahora la situación en el caso de vectores aleatorios continuos.

Definición 4.3 Sea (X, Y) un vector aleatorio continuo. Se dice que la función integrable y no negativa $f(x, y) : \mathbb{R}^2 \rightarrow [0, \infty)$ es la función de densidad del vector (X, Y) o bien que es la función de densidad conjunta de las variables X y Y si para todo par (x, y) en \mathbb{R}^2 se cumple la igualdad

$$P(X \leq x, Y \leq y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) dv du. \quad (4.1)$$

La doble integral que aparece en (4.1) representa el volumen bajo la superficie dada por la función $f(u, v)$ sobre la región que se encuentra a la

izquierda y abajo del punto (x, y) . Toda función de densidad $f(x, y)$ de estas características satisface las siguientes dos propiedades.

a) $f(x, y) \geq 0$.

b) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$.

Recíprocamente, decimos que una función $f(x, y) : \mathbb{R}^2 \rightarrow [0, \infty)$ es una función de densidad conjunta o bivariada si cumple con las dos condiciones arriba señaladas. La doble integral indicada se lleva a cabo efectuando una integral a la vez considerando que la otra variable es constante. Como el integrando es una función no negativa y la doble integral es finita, por un resultado del cálculo integral en varias variables llamado teorema de Fubini, el orden en el que se llevan a cabo las integrales no es relevante, de modo que resultado siempre es el mismo. El aspecto general de una función de densidad conjunta de dos variables aleatorias continuas es el de una superficie en \mathbb{R}^3 como la que se muestra en la Figura 4.4.

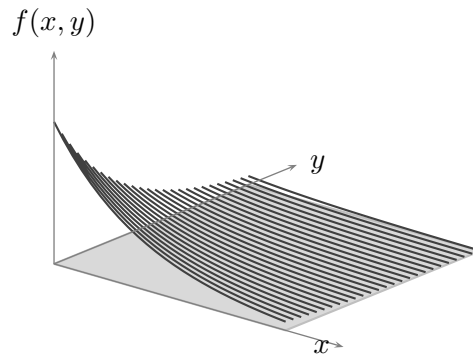


Figura 4.4

El cálculo de probabilidades de eventos relativos a un vector aleatorio continuo (X, Y) con función de densidad $f(x, y)$ se lleva a cabo de la siguiente forma. Si $a < b$ y $c < d$, entonces la probabilidad del evento $(a < X < b) \cap (c < Y < d)$ se calcula como el volumen bajo la superficie $f(x, y)$ en el rectángulo $(a, b) \times (c, d)$, es decir,

$$P(a < X < b, c < Y < d) = \int_a^b \int_c^d f(x, y) dy dx.$$

Ejemplo 4.5 La siguiente función es una de las funciones de densidad conjunta más sencillas, se trata de la distribución uniforme continua bidimensional. Sean $a < b$, $c < d$, y defina la función

$$f(x, y) = \begin{cases} \frac{1}{(b-a)(d-c)} & \text{si } a < x < b, c < y < d, \\ 0 & \text{en otro caso,} \end{cases}$$

cuya gráfica aparece en la Figura 4.5. Esta es una función constante positiva en el rectángulo $(a, b) \times (c, d)$ y es de densidad pues es no negativa e integra uno sobre \mathbb{R}^2 . La doble integral sobre \mathbb{R}^2 es simplemente el volumen del paralelepípedo que se muestra en la Figura 4.5.

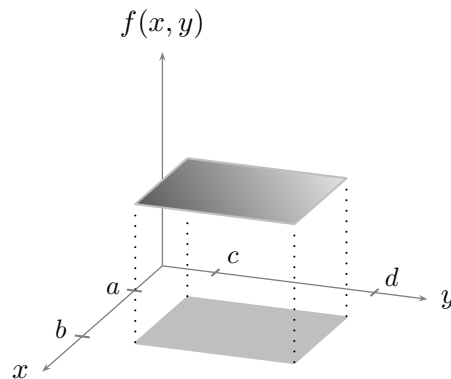


Figura 4.5

Se pueden comprobar fácilmente las siguientes probabilidades.

- a) $P(-\infty < X < b, -\infty < Y < d) = 1.$
- b) $P(X \geq a, Y \geq c) = 1.$

$$c) P(a < X < b, Y \geq d) = 0.$$

$$d) P(X \leq (a + b)/2, Y \leq (c + d)/2) = 1/4.$$

•

Ejemplo 4.6 Comprobaremos que la siguiente función es de densidad.

$$f(x, y) = \begin{cases} x + y & \text{si } 0 < x, y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

La gráfica de esta función se muestra en la Figura 4.6. Claramente $f(x, y) \geq 0$ para cualquier $(x, y) \in \mathbb{R}^2$. Resta verificar que la función integra uno sobre el plano. Se puede comprobar que

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dx dy = \int_0^1 \int_0^1 (x + y) \, dx dy = \frac{1}{2} + \frac{1}{2} = 1.$$

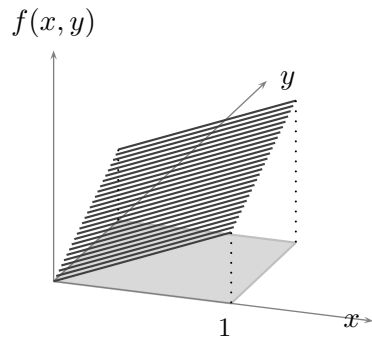


Figura 4.6

Como un ejemplo calcularemos $P(X < 1/2, Y < 1/2)$. Observe con atención

las dos primeras líneas de estos cálculos.

$$\begin{aligned}
 P(X < 1/2, Y < 1/2) &= \int_{-\infty}^{1/2} \int_{-\infty}^{1/2} f(x, y) \, dx dy \\
 &= \int_0^{1/2} \int_0^{1/2} (x + y) \, dx dy \\
 &= \int_0^{1/2} (x^2/2 + xy) \Big|_{x=0}^{x=1/2} dy \\
 &= \int_0^{1/2} (1/8 + y/2) \, dy \\
 &= 1/8.
 \end{aligned}$$

•

Ejemplo 4.7 Encontraremos la constante c para que la siguiente función sea de densidad.

$$f(x, y) = \begin{cases} cxy & \text{si } 0 < x < y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

La constante c debe ser tal que la función $f(x, y)$ es no negativa y que su integral sobre todo el plano sea uno. De esta última condición obtenemos que

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dx dy = \int_0^1 \int_0^y cxy \, dx dy = \int_0^1 \frac{c}{2} y^3 \, dy = \frac{c}{8}.$$

Por lo tanto, $c = 8$. La gráfica de la función $f(x, y)$ se muestra en la Figura 4.7. Como un ejemplo calcularemos $P(X < 1/2, Y < 1/2)$. Observe con cuidado los siguientes cálculos.

$$\begin{aligned}
 P(X < 1/2, Y < 1/2) &= \int_{-\infty}^{1/2} \int_{-\infty}^{1/2} f(x, y) \, dx dy \\
 &= \int_0^{1/2} \int_0^y 8xy \, dx dy \\
 &= \int_0^{1/2} 4y^3 \, dy \\
 &= 1/16.
 \end{aligned}$$

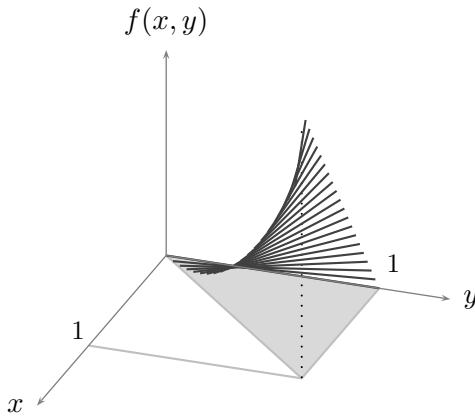


Figura 4.7

Otros ejemplos de distribuciones conjuntas aparecen en la sección de ejercicios, principalmente para el caso de dos dimensiones. En la siguiente sección se definirá la función de distribución para vectores aleatorios.

Ejercicios

448. Sean X y Y dos variables aleatorias discretas con función de probabilidad conjunta dada por la siguiente tabla.

$x \setminus y$	0	1	2
0	1/30	2/30	3/30
1	4/30	0	6/30
2	5/30	4/30	5/30

Encuentre

- | | |
|------------------------------|-----------------------------------|
| a) $P(X > 0, Y \geq 1)$. | f) $P(Y \leq 1 X = 1)$. |
| b) $P(X \leq 1, Y \geq 1)$. | g) $P(XY = 0)$. |
| c) $P(X = 1)$. | h) $P(XY \geq 2)$. |
| d) $P(Y = 2)$. | i) $P(Y \geq 2X)$. |
| e) $P(X = 0 Y = 2)$. | j) $P(X + Y \text{ sea impar})$. |

449. Sean X y Y dos variables aleatorias continuas con función de densidad conjunta dada por la siguiente expresión y cuya gráfica se muestra en la Figura 4.8.

$$f(x, y) = \begin{cases} 6x^2y & \text{si } 0 \leq x, y \leq 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Encuentre

- | | |
|-----------------------------------|----------------------------|
| a) $P(X \leq 1/2, Y \geq 1/2)$. | f) $P(X - Y \leq 1/2)$. |
| b) $P(Y \geq 1/2)$. | g) $P(XY < 1)$. |
| c) $P(X \leq 1/2 Y \geq 1/2)$. | h) $P(Y \geq X^2)$. |
| d) $P(X + Y > 1)$. | i) $P(X^2 + Y^2 \leq 1)$. |
| e) $P(Y > X)$. | j) $P(Y \leq 4X(1 - X))$. |

450. Demuestre que las siguientes funciones son de probabilidad.

- a) $f(x, y) = \begin{cases} 2^{-(x+y)} & \text{si } x, y = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$
- b) $f(x, y) = \begin{cases} 16(1/3)^{x+2y} & \text{si } x, y = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$
- c) Sea n un número natural y sean p_1 y p_2 dos probabilidades distintas de cero tales que $p_1 + p_2 < 1$. Para valores de x y y en el conjunto $\{0, 1, \dots, n\}$ tales que $0 \leq x + y \leq n$ se define

$$f(x, y) = \frac{n!}{x! y! (n - x - y)!} p_1^x p_2^y (1 - p_1 - p_2)^{n-x-y}.$$

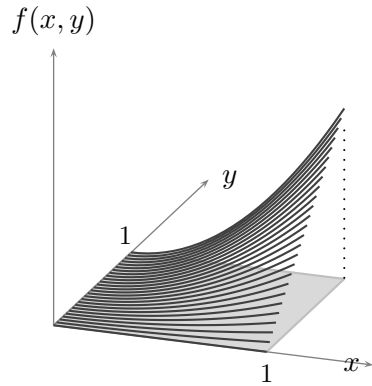


Figura 4.8

451. Demuestre que las siguientes funciones son de densidad.

$$a) f(x, y) = \begin{cases} e^{-(x+y)} & \text{si } x, y > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x, y) = \begin{cases} 6y^2 e^{-2x} & \text{si } 0 < y < 1, x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f(x, y) = \begin{cases} 3xy(1-x) & \text{si } 0 < x < 1, 0 < y < 2, \\ 0 & \text{en otro caso.} \end{cases}$$

452. Encuentre el valor de la constante c en cada caso, para que la siguiente función sea de probabilidad.

$$a) f(x, y) = \begin{cases} \frac{c}{x!y!} & \text{si } x, y = 0, 1, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x, y) = \begin{cases} cxy & \text{si } 0 < x, y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f(x, y) = \begin{cases} cx(x-y) & \text{si } 0 < x < 1, -x < y < x, \\ 0 & \text{en otro caso.} \end{cases}$$

$$d) f(x, y) = \begin{cases} c \min \{x, y\} & \text{si } 0 < x, y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$e) f(x, y) = \begin{cases} c \max \{x, y\} & \text{si } 0 < x, y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$f) f(x, y, z) = \begin{cases} c(x^2 + y^2 + z^2) & \text{si } -1 < x, y, z < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$g) f(x, y, z) = \begin{cases} c(x + y + z) & \text{si } 0 < x, y, z < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$h) f(x_1, \dots, x_n) = \begin{cases} c(x_1 + \dots + x_n) & \text{si } 0 < x_1, \dots, x_n < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$i) f(x_1, \dots, x_n) = \begin{cases} c x_1 \cdots x_n & \text{si } 0 < x_1, \dots, x_n < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

453. Sea (X, Y) un vector aleatorio discreto con función de probabilidad

$$f(x, y) = \begin{cases} e^{-2} \frac{c^{x+y}}{x! y!} & \text{si } x, y = 0, 1, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

- a) Encuentre el valor de la constante c .
- b) Calcule $P(X + Y = n)$, $n = 0, 1, \dots$

454. Sea (X, Y) un vector aleatorio continuo con función de densidad

$$f(x, y) = \begin{cases} c x^2 y^3 & \text{si } 0 < x, y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

- a) Encuentre el valor de la constante c .
- b) Calcule $P(X < Y)$.
- c) Calcule $P(X + Y > 1)$.

455. Sea (X, Y) un vector aleatorio continuo con función de densidad

$$f(x, y) = \begin{cases} c e^{-(x+y)} & \text{si } 0 < x < y, \\ 0 & \text{en otro caso.} \end{cases}$$

- a) Encuentre el valor de la constante c .
- b) Calcule $P(|X| + |Y| \leq r)$, para $r \geq 0$.
- c) Encuentre $\lim_{r \rightarrow \infty} P(|X| + |Y| \leq r)$.
- d) Calcule $P(X \leq \theta Y)$, para $0 < \theta < 1$.
- e) Encuentre $\lim_{\theta \rightarrow 0} P(X \leq \theta Y)$.

456. Sea (X, Y) un vector aleatorio continuo con función de densidad

$$f(x, y) = \begin{cases} 2 & \text{si } x > 0, y > 0, x + y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Grafique $f(x, y)$ y demuestre que es una función de densidad. Encuentre además:

- a) $P(X \leq 1/2, Y \leq 1/2)$.
- b) $P(X + Y > 2/3)$.
- c) $P(X > 2Y)$.
- d) $P(Y > 2X^2)$.

457. Un dado equilibrado se lanza dos veces. Sea X la variable aleatoria que denota el menor de estos resultados. Encuentre la función de probabilidad de X .

458. Se lanza un dado equilibrado dos veces consecutivas. Sea X el resultado del primer lanzamiento y sea Y el resultado del segundo lanzamiento. Encuentre la función de probabilidad del vector $(X, X + Y)$.

459. Sea (X, Y) un vector aleatorio discreto. Proporcione ejemplos de distribuciones de este vector de tal forma que para algunos valores de x y y se cumplan las afirmaciones de los siguientes incisos. Esto muestra que no existe una relación general de orden entre las probabilidades $P(X = x | Y = y)$ y $P(X = x)$.

- a) $P(X = x | Y = y) < P(X = x)$.

$$b) P(X = x | Y = y) = P(X = x).$$

$$c) P(X = x | Y = y) > P(X = x).$$

460. **Distribución normal bivariada.** Se dice que el vector aleatorio continuo (X, Y) tiene distribución normal de parámetros $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \rho)$, en donde μ_1 y μ_2 son cualesquiera números reales, $\sigma_1^2 > 0$, $\sigma_2^2 > 0$ y $-1 < \rho < 1$, si su función de densidad es

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho}{\sigma_1\sigma_2}(x-\mu_1)(y-\mu_2) + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right).$$

En el caso cuando $\mu_1 = \mu_2 = 0$ y $\sigma_1^2 = \sigma_2^2 = 1$, se obtiene la distribución normal bivariada estándar y la expresión de la función de densidad adquiere la expresión simple que aparece a continuación, y cuya gráfica aparece en la Figura 4.9 cuando $\rho = 0$.

$$f(x, y) = \frac{1}{2\pi\sqrt{(1-\rho^2)}} \exp\left(-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right).$$

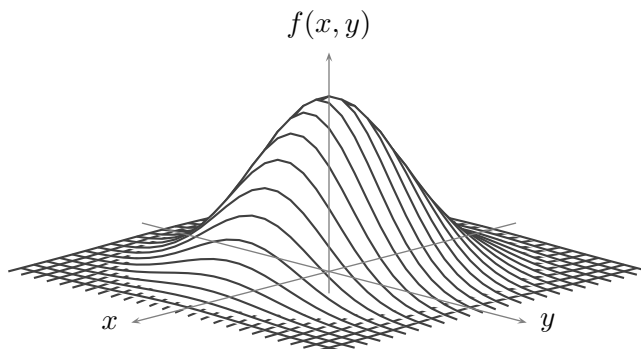


Figura 4.9

Demuestre que

- a) $f(x, y)$ es, efectivamente, una función de densidad bivariada.
 b) $f(x, y)$ se puede escribir de la siguiente forma

$$f(x, y) = \frac{1}{2\pi\sqrt{\det \Sigma}} \exp \left[-\frac{1}{2} (x - \mu_1, y - \mu_2) \Sigma^{-1} \begin{pmatrix} x - \mu_1 \\ y - \mu_2 \end{pmatrix} \right],$$

en donde Σ es la matriz cuadrada

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

4.3. Función de distribución conjunta

Además de la función de densidad o de probabilidad, existe también la función de distribución para un vector (X, Y) , sea éste discreto o continuo. Su definición aparece a continuación y es muy semejante al caso unidimensional.

Definición 4.4 La función de distribución del vector (X, Y) , denotada por $F(x, y) : \mathbb{R}^2 \rightarrow [0, 1]$, se define de la siguiente manera

$$F(x, y) = P(X \leq x, Y \leq y).$$

La pequeña coma que aparece en el lado derecho de esta igualdad significa la intersección de los eventos $(X \leq x)$ y $(Y \leq y)$, es decir, el número $F(x, y)$ es la probabilidad del evento $(X \leq x) \cap (Y \leq y)$. Más precisamente, esta función debe escribirse como $F_{X,Y}(x, y)$, pero recordemos que omitiremos los subíndices para mantener la notación simple. También aquí asociaremos la variable X con el valor x , y la variable Y con el valor y . A esta función se le conoce también con el nombre de función de acumulación de probabilidad del vector (X, Y) , y también se dice que es la función de distribución conjunta de las variables X y Y .

Se enuncian a continuación algunas propiedades que cumple toda función de distribución conjunta. Omitiremos la demostración de estas propiedades pues siguen el mismo tipo de ideas que en el caso unidimensional.

Proposición 4.1 La función de distribución $F(x, y)$ del vector aleatorio (X, Y) satisface las siguientes propiedades

1. $\lim_{x \rightarrow \infty} \lim_{y \rightarrow \infty} F(x, y) = 1.$
2. $\lim_{x \rightarrow -\infty} F(x, y) = \lim_{y \rightarrow -\infty} F(x, y) = 0.$
3. $F(x, y)$ es continua por la derecha en cada variable.
4. $F(x, y)$ es una función monótona no decreciente en cada variable.
5. Para cualesquiera números $a < b$, y $c < d$, se cumple la desigualdad

$$F(b, d) - F(a, d) - F(b, c) + F(a, c) \geq 0.$$

Observe que las primeras cuatro propiedades son completamente análogas al caso unidimensional. Por otro lado, puede comprobarse geoméricamente que la expresión que aparece en la quinta propiedad es idéntica a la probabilidad

$$P[(a < X \leq b) \cap (c < Y \leq d)],$$

que corresponde a la probabilidad de que el vector (X, Y) tome un valor dentro del rectángulo $(a, b] \times (c, d]$, como se muestra en la Figura 4.10.

Recíprocamente, se dice que una función $F(x, y) : \mathbb{R}^2 \rightarrow [0, 1]$ es una función de distribución conjunta o bivariada si satisface las cinco propiedades anteriores. En el caso continuo supondremos que la función de distribución bivariada $F(x, y)$ puede expresarse de la siguiente forma.

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) dv du,$$

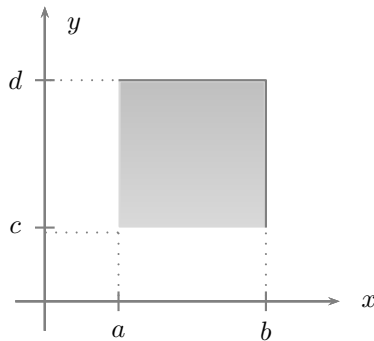


Figura 4.10

en donde $f(u, v)$ es una función no negativa y corresponde a la función de densidad bivariada asociada. El concepto de función de distribución bivariada puede extenderse al caso de vectores multidimensionales sin mayor dificultad como se muestra a continuación.

Definición 4.5 La función de distribución del vector aleatorio (X_1, \dots, X_n) es la función $F(x_1, \dots, x_n) : \mathbb{R}^n \rightarrow [0, 1]$ dada por

$$F(x_1, \dots, x_n) = P(X_1 \leq x_1, \dots, X_n \leq x_n).$$

Regresemos ahora al caso bidimensional. Las funciones $F(x, y)$ y $f(x, y)$ son equivalentes y, en nuestro caso, es siempre posible encontrar una a partir de la otra. Explicaremos este procedimiento a continuación.

De la función de densidad a la función de distribución

Conociendo la función de densidad $f(x, y)$ se puede encontrar la función de distribución $F(x, y)$ simplemente integrando en el caso continuo o sumando en el caso discreto. Para el caso continuo tenemos que

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) \, dv \, du.$$

En el caso discreto se suman todos los valores de $f(u, v)$ para valores de u menores o iguales a x , y valores de v menores o iguales a y , es decir,

$$F(x, y) = \sum_{u \leq x} \sum_{v \leq y} f(u, v).$$

Ejemplo 4.8 Sea (X, Y) un vector aleatorio discreto con función de probabilidad $f(x, y)$ dada por la siguiente tabla.

$x \backslash y$	0	1
0	1/4	1/4
1	1/4	1/4

Para encontrar la función $F(x, y)$ se necesita calcular la probabilidad $P(X \leq x, Y \leq y)$ para cada par de números reales (x, y) . El plano cartesiano \mathbb{R}^2 puede ser dividido en cinco regiones y en cada una de ellas calcular la función de distribución. Estas regiones se muestran en la Figura 4.11. La función de distribución es entonces

$$F(x, y) = \begin{cases} 0 & \text{si } x < 0 \text{ o } y < 0, \\ 1/4 & \text{si } 0 \leq x < 1 \text{ y } 0 \leq y < 1, \\ 1/2 & \text{si } 0 \leq x < 1 \text{ y } y \geq 1, \\ 1/2 & \text{si } 0 \leq y < 1 \text{ y } x \geq 1, \\ 1 & \text{si } x \geq 1 \text{ y } y \geq 1. \end{cases}$$

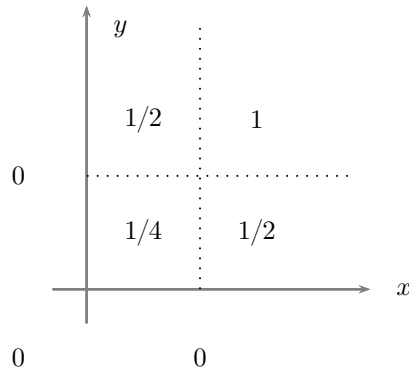


Figura 4.11

De la función de distribución a la función de densidad

Recíprocamente, puede encontrarse la función de densidad $f(x, y)$ a partir de la función de distribución $F(x, y)$ de la siguiente forma. En el caso continuo sabemos que $f(x, y)$ y $F(x, y)$ guardan la relación

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) \, dv \, du,$$

y, por el teorema fundamental del cálculo, tenemos entonces que en los puntos (x, y) en donde $f(x, y)$ es continua,

$$f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y).$$

En el caso discreto la situación no es tan sencilla, pero puede demostrarse que

$$f(x, y) = F(x, y) - F(x-, y) - F(x, y-) + F(x-, y-),$$

en donde, por ejemplo, $F(x-, y)$ es el límite de la función $F(x, y)$ en el punto (x, y) considerando que y es constante y la aproximación a x es por la izquierda.

Ejercicios

461. Sea (X, Y) un vector aleatorio discreto con función de probabilidad como indica la tabla de abajo. Encuentre y grafique la correspondiente función de distribución.

$x \backslash y$	0	1
a)	0	0
	1	1/2
	1/2	0

$x \backslash y$	0	1	2	3
b)	0	0	1/4	1/4
	1	1/4	0	0
	1/4	0	0	1/4

$x \backslash y$	0	1	2
c)	0	1/9	1/9
	1	1/9	1/9
	2	1/9	1/9

462. Sea (X, Y) un vector aleatorio continuo con función de probabilidad como se indica abajo. Encuentre y grafique la correspondiente función de distribución.

$$a) f(x, y) = \begin{cases} 1 & \text{si } 0 < x, y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x, y) = \begin{cases} 2(1-x) & \text{si } 0 < x, y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f(x, y) = \begin{cases} 2ye^{-x} & \text{si } 0 < y < 1, x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

463. Sean (X, Y) un vector aleatorio discreto con función de probabilidad dada por la siguiente tabla.

$x \backslash y$	0	1	2	3
0	1/12	1/4	1/8	1/120
1	1/6	1/4	1/20	0
2	1/24	1/40	0	0

Encuentre

- a) $P(X = 1, Y = 2)$.
 b) $P(X = 0, 1 \leq Y < 3)$.
 c) $P(X + Y \leq 1)$.
 d) $P(X > Y)$.
 e) $F_{X,Y}(1.2, 0.9)$.
 f) $F_{X,Y}(-3, 1.5)$.
 g) $F_{X,Y}(2, 0)$.
 h) $F_{X,Y}(4, 2.7)$.
464. Sea $F(x, y)$ la función de distribución conjunta de dos variables aleatorias discretas. Demuestre que la función de probabilidad conjunta asociada $f(x, y)$ puede calcularse a partir de $F(x, y)$ como muestra la siguiente fórmula.

$$f(x, y) = F(x, y) - F(x-, y) - F(x, y-) + F(x-, y-).$$

465. Sean $F(x, y)$ y $G(x, y)$ dos funciones de distribución bivariadas. Demuestre que, para cualquier constante $\lambda \in [0, 1]$, la siguiente función es de distribución.

$$(x, y) \mapsto \lambda F(x, y) + (1 - \lambda)G(x, y).$$

466. Sea (X, Y) un vector aleatorio continuo con distribución uniforme en el cuadrado $(-2, 2) \times (-2, 2)$. Calcule $P(|Y| > |X| + 1)$. Encuentre y grafique las siguientes funciones.

- a) $f_{X,Y}(x, y)$.
 b) $f_X(x)$.
 c) $f_Y(y)$.
 d) $F_{X,Y}(x, y)$.

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx.$$

En general, las funciones de probabilidad marginales $f_X(x)$ y $f_Y(y)$ son distintas, aunque hay ocasiones en que pueden ser iguales. Es inmediato verificar que estas funciones son, efectivamente, funciones de densidad univariadas, pues son no negativas e integran uno. Veamos un ejemplo.

Ejemplo 4.9 Sea (X, Y) un vector aleatorio continuo con función de densidad dada por

$$f(x, y) = \begin{cases} 4xy & \text{si } 0 < x, y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Es sencillo verificar que esta función es, efectivamente, una función de densidad bivariada pues es no negativa e integra uno.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = \int_0^1 \int_0^1 4xy dx dy = 4(1/2)(1/2) = 1.$$

Calcularemos ahora las funciones de densidad marginales $f_X(x)$ y $f_Y(y)$. Esto debe hacerse para cada valor de x y y en \mathbb{R} . Para $x \notin (0, 1)$, $f(x, y) = 0$ y por lo tanto $f_X(x) = 0$. Para $x \in (0, 1)$,

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_0^1 4xy dy = 2x.$$

Por lo tanto,

$$f_X(x) = \begin{cases} 2x & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

De manera análoga, o por simetría,

$$f_Y(y) = \begin{cases} 2y & \text{si } 0 < y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Es inmediato comprobar que estas funciones son funciones de densidad univariadas. Observe que en este caso particular se cumple la identidad $f(x, y) = f_X(x)f_Y(y)$, lo cual expresa el importante concepto de independencia de v.a.s que hemos mencionado antes. •

La definición de función de probabilidad marginal para vectores discretos involucra una suma en lugar de la integral, por ejemplo,

$$f_X(x) = \sum_y f(x, y),$$

de manera análoga se define la función de probabilidad marginal $f_Y(y)$. Veamos un ejemplo de este procedimiento.

Ejemplo 4.10 Sea (X, Y) un vector aleatorio discreto con función de probabilidad dada por

$$f(x, y) = \begin{cases} (x + 2y)/30 & \text{si } (x, y) \in \{1, 2, 3\} \times \{1, 2\}, \\ 0 & \text{en otro caso.} \end{cases}$$

No es difícil comprobar que esta función es una función de probabilidad bivariada, es decir, es no negativa y suma uno.

$$\sum_{x=1}^3 \sum_{y=1}^2 \frac{x + 2y}{30} = 1.$$

Las funciones de probabilidad marginales $f_X(x)$ y $f_Y(y)$ son

$$f_X(x) = \sum_{y=1}^2 f(x, y) = \begin{cases} 8/30 & \text{si } x = 1, \\ 10/30 & \text{si } x = 2, \\ 12/30 & \text{si } x = 3, \\ 0 & \text{en otro caso.} \end{cases}$$

y

$$f_Y(y) = \sum_{x=1}^3 f(x, y) = \begin{cases} 12/30 & \text{si } y = 1, \\ 18/30 & \text{si } y = 2, \\ 0 & \text{en otro caso.} \end{cases}$$

Claramente estas funciones son funciones de probabilidad univariadas. •

Un poco más generalmente, la función de densidad marginal de la variable X_1 a partir de la función de densidad del vector (X_1, \dots, X_n) es, en el caso continuo,

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_2 \cdots dx_n.$$

De manera análoga se puede obtener la función de densidad marginal de cualquiera de las variables que componen el vector multidimensional. Y también, de manera similar, se pueden calcular estas densidades marginales de vectores que son subconjuntos del vector original. Por ejemplo, la función de densidad marginal del vector (X_1, X_2) a partir de (X_1, \dots, X_n) es, en el caso continuo,

$$f_{X_1, X_2}(x_1, x_2) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) dx_3 \cdots dx_n.$$

Otro aspecto interesante sobre estas funciones es que pueden existir distintas funciones de densidad conjuntas que producen las mismas funciones de densidad marginales. En el Ejercicio 469 se muestra esta situación.

Ejercicios

468. Sea (X, Y) un vector aleatorio con función de probabilidad como aparece abajo. En cada caso encuentre las funciones de probabilidad marginales $f_X(x)$ y $f_Y(y)$.

$x \backslash y$	0	1
a)	0	1/16
	1	5/16
	4/16	6/16

$$b) f(x, y) = \begin{cases} e^{-x} & \text{si } 0 < y < x, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f(x, y) = \begin{cases} 2e^{-x-y} & \text{si } 0 < x < y, \\ 0 & \text{en otro caso.} \end{cases}$$

$$d) f(x, y) = \frac{n!}{x! y! (n - x - y)!} p_1^x p_2^y (1 - p_1 - p_2)^{n-x-y},$$

en donde $n \in \mathbb{N}$, p_1 y p_2 son dos probabilidades estrictamente positivas tales que $p_1 + p_2 < 1$ y $x, y = 0, 1, \dots, n$ son tales que $0 \leq x + y \leq n$.

469. **Distintas conjuntas, mismas marginales.** Sea (X, Y) un vector aleatorio discreto con función de probabilidad dada por la tabla que aparece abajo. Compruebe que para cualquier valor de los parámetros θ y p tales que $0 \leq p \leq 1/2$ y $(1 - 2p)/(1 - p) \leq \theta \leq 1$, las correspondientes funciones de probabilidad marginales son $\text{Ber}(p)$.

$x \backslash y$	0	1
0	$\theta(1 - p)$	$(1 - \theta)(1 - p)$
1	$(1 - \theta)(1 - p)$	$p - (1 - \theta)(1 - p)$

470. Sea (X, Y) un vector aleatorio continuo con distribución normal de parámetros $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \rho)$. Esta distribución se definió en el Ejercicio 460, en la página 307. Demuestre que

- a) la distribución marginal de X es $N(\mu_1, \sigma_1^2)$.
 b) la distribución marginal de Y es $N(\mu_2, \sigma_2^2)$.

4.5. Función de distribución marginal

Ahora veremos la forma de obtener la función de distribución individual de una variable aleatoria a partir de la función de distribución de un vector aleatorio. Nuevamente consideraremos primero el caso bidimensional y después explicaremos la situación para el caso de dimensión mayor.

Definición 4.7 Sea (X, Y) un vector aleatorio, continuo o discreto, con función de distribución $F(x, y)$. La función de distribución marginal de la variable X se define como la función de una variable

$$F_X(x) = \lim_{y \rightarrow \infty} F(x, y).$$

Análogamente, la función de distribución marginal de la variable Y se define como la función

$$F_Y(y) = \lim_{x \rightarrow \infty} F(x, y).$$

Observemos que los límites anteriores siempre existen pues la función de distribución conjunta es acotada y no decreciente en cada variable. No es difícil comprobar que estas funciones de distribución marginales son, efectivamente, funciones de distribución univariadas. Para un vector de dimensión tres (X, Y, Z) , a partir de $F_{X,Y,Z}(x, y, z)$ y tomando los límites necesarios, pueden obtenerse, por ejemplo, las funciones de distribución marginales $F_{X,Y}(x, y)$, $F_{X,Z}(x, z)$, $F_X(x)$. En efecto,

$$\begin{aligned} F_{X,Y}(x, y) &= \lim_{z \rightarrow \infty} F(x, y, z), \\ F_{X,Z}(x, z) &= \lim_{y \rightarrow \infty} F(x, y, z), \\ F_X(x) &= \lim_{y \rightarrow \infty} \lim_{z \rightarrow \infty} F(x, y, z). \end{aligned}$$

Más generalmente, a partir de la función de distribución de un vector (X_1, \dots, X_n) se puede obtener, de manera análoga, la función de distribución de cualquier subvector.

Ahora que conocemos la forma de obtener las funciones de densidad y de distribución marginales a partir de las correspondientes funciones conjuntas, podemos enunciar con precisión el concepto de independencia entre variables aleatorias. Veremos en la siguiente sección este importante concepto, el cual resulta ser una hipótesis recurrente en los procedimientos de la estadística matemática y otras áreas de aplicación de la probabilidad.

Ejercicios

471. Sean X y Y dos variables aleatorias discretas con función de probabilidad conjunta dada por la tabla que aparece abajo. Encuentre la función de distribución conjunta $F_{X,Y}(x, y)$ y a partir de ella encuentre las funciones de distribución marginales $F_X(x)$ y $F_Y(y)$.

$x \backslash y$	0	1
0	1/8	2/8
1	2/8	3/8

472. Encuentre las funciones de distribución marginales $F_X(x)$ y $F_Y(y)$ para cada una de las siguientes funciones de distribución conjuntas. En cada caso grafique $F(x, y)$, $F_X(x)$ y $F_Y(y)$.

a) Si $a < b$ y $c < d$, entonces

$$F(x, y) = \begin{cases} 0 & \text{si } x < a \text{ o } y < c, \\ 1/2 & \text{si } a \leq x < b, c \leq y < d, \\ 3/4 & \text{si } a \leq x < b, y \geq d, \\ 3/4 & \text{si } x \geq b, c \leq y < d, \\ 1 & \text{si } x \geq b, y \geq d. \end{cases}$$

$$b) F(x, y) = \begin{cases} (1 - e^{-x})(1 - e^{-y}) & \text{si } x, y > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

4.6. Independencia de variables aleatorias

Sea X_1, \dots, X_n una colección de variables aleatorias con función de distribución conjunta $F(x_1, \dots, x_n)$. Suponga que las respectivas funciones de distribución marginales son $F_{X_1}(x_1), \dots, F_{X_n}(x_n)$.

Definición 4.8 Se dice que las variables aleatorias X_1, \dots, X_n son independientes si para cualesquiera números reales x_1, \dots, x_n se cumple la igualdad

$$F(x_1, \dots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n).$$

Alternativamente, puede definirse la independencia en términos de la función de densidad, suponiendo su existencia, como sigue.

Definición 4.9 Se dice que las variables aleatorias X_1, \dots, X_n con función de probabilidad conjunta $f(x_1, \dots, x_n)$ son independientes si para cualesquiera números reales x_1, \dots, x_n se cumple la igualdad

$$f(x_1, \dots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n).$$

Puede demostrarse que las dos condiciones anteriores son equivalentes. En el caso particular cuando las variables aleatorias son discretas, la condición de independencia se escribe de la forma siguiente. Para cualesquiera números x_1, \dots, x_n ,

$$P(X_1 = x_1, \dots, X_n = x_n) = P(X_1 = x_1) \cdots P(X_n = x_n).$$

Ejemplo 4.11 Sean X y Y dos variables aleatorias continuas con función de densidad conjunta dada por

$$f(x, y) = \begin{cases} e^{-x-y} & \text{si } x, y > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

Pueden calcularse las funciones de densidad marginales y encontrar que

$$f_X(x) = \begin{cases} e^{-x} & \text{si } x > 0, \\ 0 & \text{en otro caso,} \end{cases} \quad y \quad f_Y(y) = \begin{cases} e^{-y} & \text{si } y > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

Se verifica entonces que $f(x, y) = f_X(x)f_Y(y)$ para cualesquiera números reales x y y . Esto demuestra la independencia de las variables X y Y . •

Ejemplo 4.12 Sean X y Y dos variables aleatorias discretas con función de probabilidad $f(x, y)$ dada por

$$f(x, y) = \begin{cases} 1/4 & \text{si } x, y \in \{0, 1\}, \\ 0 & \text{en otro caso.} \end{cases}$$

Las funciones de probabilidad marginales son

$$f_X(x) = \begin{cases} 1/2 & \text{si } x \in \{0, 1\}, \\ 0 & \text{en otro caso,} \end{cases} \quad y \quad f_Y(y) = \begin{cases} 1/2 & \text{si } y \in \{0, 1\}, \\ 0 & \text{en otro caso.} \end{cases}$$

Por lo tanto, $f(x, y) = f_X(x)f_Y(y)$ para cualesquiera números reales x y y . Se concluye entonces que X y Y son independientes. •

Adicionalmente tenemos la siguiente extensión del concepto de independencia de variables aleatorias.

Definición 4.10 Se dice que un conjunto infinito de variables aleatorias es independiente si cualquier subconjunto finito de él lo es.

Este es el sentido en el que una sucesión infinita de variables aleatorias independientes debe entenderse. Tal hipótesis aparece, por ejemplo, en el enunciado de la ley de los grandes números y el teorema central del límite que estudiaremos más adelante.

Ejercicios

473. Sean X y Y dos variables aleatorias discretas cuya función de probabilidad conjunta admite la factorización que aparece abajo para ciertas funciones $g(x)$ y $h(y)$.

$$P(X = x, Y = y) = g(x)h(y).$$

- a) Expresar $P(X = x)$ y $P(Y = y)$ en términos de $g(x)$ y $h(y)$.
- b) Demuestre que X y Y son independientes.
- c) Demuestre que $\sum_x g(x) = \sum_y h(y) = 1$.

474. Sean X y Y dos variables aleatorias independientes con distribución geométrica de parámetros p y q , respectivamente. Calcule

- a) $P(X = Y)$.
- b) $P(X \leq Y)$.

475. Sea X_1, X_2, \dots una sucesión infinita de variables aleatorias independientes, cada una con la misma distribución $\text{Ber}(p)$, e independientes de otra variable aleatoria N con distribución $\text{Poisson}(\lambda)$. Demuestre que

$$S_N := \sum_{i=1}^N X_i \sim \text{Poisson}(\lambda p).$$

Cuando $N = 0$, la suma es vacía y se define como cero. Si N representa el número de delitos ocurridos, de los cuales sólo la fracción p son reportados a la autoridad, entonces S_N representa el número de delitos reportados.

476. El siguiente ejemplo muestra que la condición $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ no es suficiente para concluir que X y Y son independientes. Sea (X, Y) un vector aleatorio discreto con función de probabilidad dada por la tabla que aparece abajo. Compruebe que se cumple la igualdad $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ y que, sin embargo, X y Y no son independientes.

$x \backslash y$	-1	0	1
-1	1/8	0	1/8
0	0	1/2	0
1	1/8	0	1/8

477. Sea (X, Y) un vector aleatorio discreto con función de probabilidad $f(x, y)$ dada por la siguiente tabla.

$x \setminus y$	0	1	2
0	1/10	2/10	1/10
1	1/10	0	1/10
2	1/10	1/10	2/10

- Grafique $f(x, y)$ y demuestre que efectivamente se trata de una función de probabilidad.
 - Calcule y grafique las densidades marginales $f_X(x)$ y $f_Y(y)$. Verifique que ambas son, efectivamente, funciones de probabilidad.
 - Calcule y grafique la función de distribución conjunta $F(x, y)$.
 - Calcule y grafique las distribuciones marginales $F_X(x)$ y $F_Y(y)$. Verifique que ambas son, efectivamente, funciones de distribución.
 - ¿Son X y Y independientes?
478. Sea (X, Y) un vector aleatorio continuo con distribución uniforme en el cuadrado $(-1, 1) \times (-1, 1)$.
- Grafique $f(x, y)$ y demuestre que efectivamente se trata de una función de densidad.
 - Calcule y grafique las densidades marginales $f_X(x)$ y $f_Y(y)$. Verifique que ambas son, efectivamente, funciones de densidad.
 - Calcule y grafique la función de distribución conjunta $F(x, y)$.
 - Calcule y grafique las distribuciones marginales $F_X(x)$ y $F_Y(y)$. Verifique que ambas son, efectivamente, funciones de distribución.
 - ¿Son X y Y independientes?
479. Determine si las siguientes funciones de probabilidad corresponden a variables aleatorias independientes.

$x \backslash y$	0	1
0	1/4	1/4
1	1/4	1/4

$x \backslash y$	0	1	2
0	1/30	2/30	3/30
1	4/30	5/60	6/30
2	3/30	3/30	3/30

$x \backslash y$	0	1	2
0	1/60	2/60	3/60
1	2/60	4/60	6/60
2	3/60	6/60	9/60
3	4/60	8/60	12/60

$$d) \quad f(x, y) = \begin{cases} 4xy & \text{si } 0 < x, y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$e) \quad f(x, y) = \begin{cases} x + y & \text{si } 0 < x, y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$f) \quad f(x, y, z) = \begin{cases} 8xyz & \text{si } 0 < x, y, z < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$g) \quad f(x, y, z) = \begin{cases} x^2 + y^2 + z^2 & \text{si } 0 < x, y, z < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$h) \quad f(x_1, \dots, x_n) = \begin{cases} 2^n x_1 \cdots x_n & \text{si } 0 < x_i < 1, \quad i = 1, \dots, n, \\ 0 & \text{en otro caso.} \end{cases}$$

$$i) \quad f(x, y) = \begin{cases} e^{-x} & \text{si } 0 < y < x, \\ 0 & \text{en otro caso.} \end{cases}$$

$$j) \quad f(x, y) = \begin{cases} 2e^{-x-y} & \text{si } 0 < x < y, \\ 0 & \text{en otro caso.} \end{cases}$$

480. Sean X y Y dos variables aleatorias independientes, a partir de ellas defina las siguientes variables aleatorias

$$\begin{aligned} U &= \text{máx}\{X, Y\}, \\ V &= \text{mín}\{X, Y\}. \end{aligned}$$

Demuestre que

$$a) \quad F_U(u) = F_X(u) F_Y(u).$$

$$b) \quad F_V(v) = 1 - [1 - F_X(v)][1 - F_Y(v)].$$

Encuentre las distribuciones individuales de U y V cuando X y Y tienen ambas distribución

$$c) \quad \exp(\lambda).$$

$$d) \quad \text{geo}(p).$$

4.7. Distribución condicional

Recordemos que la probabilidad condicional de un evento A , dado un evento B , está dada por

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Esta definición puede extenderse al caso de funciones de probabilidad o de densidad y también para el caso de funciones de distribución. En esta sección enunciaremos tales extensiones.

Definición 4.11 Sea (X, Y) un vector aleatorio discreto o continuo con función de probabilidad o de densidad $f_{X,Y}(x, y)$. Sea y un valor de la variable Y tal que $f_Y(y) \neq 0$. A la función $x \mapsto f_{X|Y}(x|y)$, definida a continuación, se le llama la función de probabilidad o densidad condicional de X dado que $Y = y$.

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}. \quad (4.2)$$

Observe que a la función dada por (4.2) se le considera como una función de x y que el valor de y es fijo y puede considerársele como un parámetro de dicha función, es decir, para cada valor fijo de y se tiene una función diferente. En el caso discreto la expresión (4.2) es, efectivamente, la definición de probabilidad condicional

$$f_{X|Y}(x|y) = \frac{P(X = x, Y = y)}{P(Y = y)},$$

sin embargo, recordemos que en el caso continuo las expresiones $f_{X,Y}(x, y)$ y $f_Y(y)$ no son necesariamente probabilidades.

Sumando o integrando sobre los posible valores x , es inmediato comprobar que la función dada por (4.2) es, efectivamente, una función de probabilidad o de densidad. Observe además que cuando X y Y son independientes, para cualquier valor de y se tiene que

$$f_{X|Y}(x|y) = f_X(x).$$

Ejemplo 4.13 Sea (X, Y) un vector aleatorio discreto con función de probabilidad dada por la siguiente tabla.

$x \backslash y$	0	1	2	3
0	1/10	1/10	2/10	1/10
1	1/10	2/10	1/10	1/10

Calcularemos la función de probabilidad condicional $f_{X|Y}(x|y)$ para $y = 1$. Sumando las probabilidades de la columna correspondiente a $y = 1$ se encuentra que $f_Y(1) = 3/10$. Por lo tanto, aplicando la fórmula (4.2), tenemos que

$$f_{X|Y}(x|1) = \frac{f_{X,Y}(x,1)}{f_Y(1)} = \begin{cases} 1/3 & \text{si } x = 0, \\ 2/3 & \text{si } x = 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Lo cual corresponde a una función de probabilidad univariada. De manera análoga pueden calcularse $f_{X|Y}(x|0)$, $f_{X|Y}(x|2)$, $f_{X|Y}(x|3)$, y también $f_{Y|X}(y|0)$ y $f_{Y|X}(y|1)$. ¿Puede usted encontrar estas funciones de probabilidad? •

Ejemplo 4.14 Sea (X, Y) un vector aleatorio continuo con función de densidad dada por

$$f_{X,Y}(x,y) = \begin{cases} x+y & \text{si } 0 < x, y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Calcularemos la función de probabilidad condicional $f_{X|Y}(x|y)$ para cada y en el intervalo $(0, 1)$. Integrando sobre x , tenemos que la función de densidad marginal de Y es

$$f_Y(y) = \begin{cases} (1+2y)/2 & \text{si } 0 < y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Por lo tanto, para cada $y \in (0, 1)$ fijo, la función de densidad condicional de X dado $Y = y$ está dada por

$$x \mapsto f_{X|Y}(x|y) = \begin{cases} 2(x+y)/(1+2y) & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

De manera similar, para cada $x \in (0, 1)$ fijo,

$$y \mapsto f_{Y|X}(y|x) = \begin{cases} 2(x+y)/(1+2x) & \text{si } 0 < y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

•

La fórmula (4.2) puede extenderse, de manera análoga, al caso de vectores de dimensión mayor. Por ejemplo, para un vector aleatorio de dimensión tres (X, Y, Z) , pueden calcularse funciones de densidad condicionales como $f_{X|Y,Z}(x|y, z)$ o $f_{X,Z|Y}(x, z|y)$. Veamos ahora la extensión al caso de funciones de distribución.

Definición 4.12 Sea (X, Y) un vector aleatorio con función de probabilidad o densidad $f_{X,Y}(x, y)$. Sea y un valor de Y tal que $f_Y(y) \neq 0$. La función de distribución condicional de X , dado $Y = y$, es la función

$$x \mapsto F_{X|Y}(x|y) = \begin{cases} \sum_{u \leq x} f_{X|Y}(u|y) & \text{en el caso discreto,} \\ \int_{-\infty}^x f_{X|Y}(u|y) du & \text{en el caso continuo.} \end{cases}$$

De esta forma, la función de distribución condicional se calcula como la suma o integral de la correspondiente función de probabilidad o densidad condicional. De acuerdo a esta definición, uno podría escribir a la función $F_{X|Y}(x|y)$ como $P(X \leq x | Y = y)$, sin embargo tal expresión puede causar confusión pues al aplicar la definición de probabilidad condicional tendríamos el término $P(Y = y)$ como denominador. Este término puede ser cero y por lo tanto la probabilidad condicional indicada podría no estar definida. Preferiremos entonces conservar las expresiones completas que aparecen en la definición. Por otro lado, observamos nuevamente que cuando X y Y son independientes,

$$F_{X|Y}(x|y) = F_X(x).$$

Veremos a continuación algunos ejemplos de la forma en la que son calculadas estas funciones de distribución condicionales.

Ejemplo 4.15 Consideremos nuevamente el vector aleatorio discreto (X, Y) con función de probabilidad dada por la siguiente tabla.

$x \setminus y$	0	1	2	3
0	1/10	1/10	2/10	1/10
1	1/10	2/10	1/10	1/10

Habíamos encontrado que la función de probabilidad condicional $f_{X|Y}(x|y)$ para $y = 1$ es

$$f_{X|Y}(x|1) = \begin{cases} 1/3 & \text{si } x = 0, \\ 2/3 & \text{si } x = 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Por lo tanto,

$$F_{X|Y}(x|1) = \begin{cases} 0 & \text{si } x < 0, \\ 1/3 & \text{si } 0 \leq x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

Esta es una función de distribución univariada. •

Ejemplo 4.16 En un ejemplo anterior habíamos considerado el vector aleatorio continuo (X, Y) con función de densidad

$$f_{X,Y}(x, y) = \begin{cases} x + y & \text{si } 0 < x, y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Habíamos encontrado que la función de probabilidad condicional $f_{X|Y}(x|y)$ para cada y en el intervalo $(0, 1)$ es

$$f_{X|Y}(x|y) = \begin{cases} 2(x+y)/(1+2y) & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Por lo tanto,

$$F_{X|Y}(x|y) = \begin{cases} 0 & \text{si } x \leq 0, \\ (x^2 + 2xy)/(1+2y) & \text{si } 0 < x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

Esta es una función de distribución univariada en donde y es considerada como una constante.

•

Ejercicios

481. Sea (X, Y) un vector aleatorio discreto o continuo y sea y tal que $f_Y(y) \neq 0$. Demuestre que la función de densidad o de probabilidad condicional

$$f_{X|Y}(x|y)$$

es efectivamente una función de densidad o de probabilidad univariada.

482. Sea (X, Y) un vector aleatorio discreto o continuo y sea y tal que $f_Y(y) \neq 0$. Demuestre que la función de distribución condicional

$$F_{X|Y}(x|y)$$

es efectivamente una función de distribución univariada.

483. Sea (X, Y) un vector aleatorio discreto con función de probabilidad dada por la siguiente tabla.

$x \setminus y$	0	1	2
0	0.1	0.05	0.1
1	0.05	0.2	0.1
2	0.05	0.05	0.3

Calcule las siguientes funciones.

- | | |
|---------------------|---------------------|
| a) $f_{X Y}(x 0)$. | d) $F_{X Y}(x 0)$. |
| b) $f_{X Y}(x 1)$. | e) $F_{X Y}(x 1)$. |
| c) $f_{X Y}(x 2)$. | f) $F_{X Y}(x 2)$. |

484. Sea (X, Y) un vector aleatorio continuo con función de densidad

$$f_{X,Y}(x, y) = \begin{cases} e^{-x} & \text{si } 0 < y < x, \\ 0 & \text{en otro caso.} \end{cases}$$

Calcule las siguientes funciones.

- a) $x \mapsto f_{X|Y}(x|y)$, $y > 0$ fijo.
- b) $y \mapsto f_{Y|X}(y|x)$, $x > 0$ fijo.
- c) $x \mapsto F_{X|Y}(x|y)$, $y > 0$ fijo.
- d) $y \mapsto F_{Y|X}(y|x)$, $x > 0$ fijo.

485. Se lanza un dado equilibrado dos veces consecutivas. Sea X el resultado del primer lanzamiento y sea Y el resultado del segundo lanzamiento. Calcule la distribución condicional de X dado que

- a) $Y = 2$.
- b) $X + Y = 5$.
- c) $X + Y \geq 5$.

486. Se lanza un dado equilibrado dos veces. Sea X el resultado del primer lanzamiento y sea Y el mayor de los dos resultados.

- a) Encuentre la función de probabilidad conjunta de X y Y .
- b) Calcule las funciones $f_{Y|X}(y|x=3)$ y $f_{X|Y}(x|y=3)$.

4.8. Esperanza condicional

Definiremos a continuación el valor esperado de una variable aleatoria dado que un evento ha ocurrido para otra variable aleatoria, cuando se conoce la distribución conjunta de las dos variables.

Definición 4.13 Sea (X, Y) un vector aleatorio continuo con función de densidad $f_{X,Y}(x, y)$ y sea y un valor tal que $f_Y(y) \neq 0$. La esperanza condicional de X , dado $Y = y$, es la esperanza de la función de densidad condicional $f_{X|Y}(x | y)$, cuando existe, es decir,

$$E(X | Y = y) = \int_{-\infty}^{\infty} x f_{X|Y}(x | y) dx.$$

Efectuando un cambio en el orden de las integrales, es inmediato comprobar que

$$E(X) = \int_{-\infty}^{\infty} E(X | Y = y) f_Y(y) dy,$$

cuya expresión recuerda el teorema de probabilidad total, pero esta vez en términos de esperanzas. En el caso cuando el vector (X, Y) es discreto, la definición es análoga.

$$\begin{aligned} E(X | Y = y) &= \sum_x x f_{X|Y}(x | y) \\ &= \sum_x x P(X = x | Y = y), \end{aligned}$$

suponiendo nuevamente que $f_Y(y) \neq 0$ y que la suma indicada es absolutamente convergente. Nuevamente, efectuando un cambio en el orden de las sumas se encuentra la expresión

$$E(X) = \sum_y E(X | Y = y) P(Y = y).$$

En cualquier caso, observe además que cuando X y Y son independientes,

$$E(X | Y = y) = E(X).$$

Ejercicios

487. Un experimento consiste en lanzar un dado equilibrado, repetidas veces, hasta obtener alguno de los resultados por segunda vez. Encuentre el número esperado de lanzamientos en este experimento.

488. Sea (X, Y) un vector aleatorio discreto con función de probabilidad como aparece en la siguiente tabla.

$x \backslash y$	0	1	2
0	1/8	0	1/8
1	1/8	1/2	1/8

Calcule $E(X | Y = y)$ para $y = 0, 1, 2$, y verifique además que se cumple la identidad

$$E(X) = \sum_{y=0}^2 E(X | Y = y) P(Y = y).$$

489. Se lanza un dado equilibrado dos veces consecutivas. Sea X el resultado del primer lanzamiento y sea Y el resultado del segundo lanzamiento. Calcule y compare las siguientes cantidades.

- a) $E(X)$.
 b) $E(X | X + Y = 6)$.

490. ¿Es $E(X | Y = y)$ menor, igual o mayor a $E(X)$? En general, no existe una relación de orden entre estas cantidades. Proporcione ejemplos de distribuciones para (X, Y) en donde se cumpla cada una de las relaciones de orden indicadas.

491. Para cada y en $(0, 1)$ calcule la esperanza condicional $E(X | Y = y)$ cuando X y Y tienen función de densidad conjunta

$$f_{X,Y}(x, y) = \begin{cases} 12x^2 & \text{si } 0 < x < y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

492. Sea X con distribución Poisson(Λ), en donde Λ es una variable aleatoria con distribución unif $\{1, \dots, n\}$. Condicionando sobre el valor de Λ encuentre la esperanza de X .

493. Sea X con distribución bin($N + 1, p$), en donde N es una variable aleatoria con distribución geo(q). Condicionando sobre el valor de N encuentre la esperanza de X .

494. Sea X_1, X_2, \dots una sucesión de variables aleatorias independientes e idénticamente distribuidas con media y varianza finitas, e independientes de otra variable N con valores $0, 1, \dots$ y con esperanza y varianza finita. Defina la variable aleatoria

$$S_N = \sum_{i=1}^N X_i.$$

Condicionando sobre el valor de N y usando las hipótesis de independencia demuestre que:

- a) $E(S_N) = E(N) E(X_1)$.
 b) $\text{Var}(S_N) = \text{Var}(X_1) E(N) + \text{Var}(N) E^2(X_1)$.
495. En una población pequeña ocurren cada día 0, 1, 2, o 3 accidentes automovilísticos con probabilidades $1/6, 1/3, 1/3$ y $1/6$, respectivamente. En un accidente cualquiera se requiere el uso de una ambulancia con probabilidad $2/3$. Calcule el número de veces promedio que se requiere el uso de una ambulancia por accidentes automovilísticos en un día cualquiera en esta población.

4.9. Esperanza, varianza y covarianza

En esta sección se definirá la esperanza y varianza de un vector aleatorio, así como la covarianza entre dos variables aleatorias.

Definición 4.14 La esperanza de un vector aleatorio (X, Y) , compuesto por dos variables aleatorias con esperanzas finitas, es el vector de las esperanzas, es decir,

$$E(X, Y) = (E(X), E(Y)).$$

De esta manera, encontrar la esperanza de un vector aleatorio se reduce al cálculo de la esperanza de cada una de las variables del vector. Es claro que

esta definición puede extenderse, sin ninguna dificultad, para dimensiones mayores. Se verá ahora la definición de varianza de un vector aleatorio en dimensión dos.

Definición 4.15 La varianza de un vector aleatorio (X, Y) , compuesto por dos variables aleatorias con varianzas finitas, es la matriz cuadrada

$$\text{Var}(X, Y) = \begin{pmatrix} \text{Var}(X) & \text{Cov}(X, Y) \\ \text{Cov}(Y, X) & \text{Var}(Y) \end{pmatrix},$$

en donde $\text{Cov}(X, Y)$ es la covarianza de X y Y y se define como sigue

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))].$$

Como veremos más adelante, la covarianza está estrechamente relacionada con otro concepto que se define para dos variables aleatorias, llamado coeficiente de correlación, para el cual se cuenta con una interpretación bastante clara. Dejaremos entonces la interpretación de la covarianza en términos del coeficiente de correlación.

Explicaremos ahora la forma de calcular la covarianza según la definición anterior. Cuando X y Y son variables aleatorias discretas, la covarianza se calcula de la forma siguiente.

$$\text{Cov}(X, Y) = \sum_{x,y} (x - E(X))(y - E(Y)) f(x, y),$$

en donde la suma es doble, es decir, se suma sobre todos los posibles valores x y también sobre todos los posibles valores y . En el caso cuando las variables aleatorias son continuas se tiene que

$$\text{Cov}(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - E(X))(y - E(Y)) f(x, y) dx dy.$$

Desarrollando el producto que aparece en la definición de covarianza y aplicando la linealidad de la esperanza se encuentra que la covarianza puede calcularse también como indica la siguiente fórmula.

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y).$$

Por otro lado, a partir de la definición misma de covarianza, o a partir de la fórmula recién enunciada, es inmediato observar que la covarianza es simétrica, es decir,

$$\text{Cov}(X, Y) = \text{Cov}(Y, X).$$

Esto tiene como consecuencia que la matriz $\text{Var}(X, Y)$, que aparece en la Definición 4.15, es simétrica. Otra propiedad interesante y fácil de obtener se encuentra cuando se calcula la covarianza entre una variable aleatoria X y ella misma. En este caso la covarianza se reduce a la varianza de X , en símbolos,

$$\text{Cov}(X, X) = \text{Var}(X).$$

Por lo tanto, la definición de varianza del vector (X, Y) se puede escribir como sigue.

$$\text{Var}(X, Y) = \begin{pmatrix} \text{Cov}(X, X) & \text{Cov}(X, Y) \\ \text{Cov}(Y, X) & \text{Cov}(Y, Y) \end{pmatrix}.$$

De esta expresión se desprende una forma natural de definir la varianza de un vector de cualquier dimensión. La varianza de (X_1, \dots, X_n) es aquella matriz cuadrada de $n \times n$, cuya entrada (i, j) está dada por el número $\text{Cov}(X_i, X_j)$, es decir,

$$\text{Var}(X_1, \dots, X_n) = (\text{Cov}(X_i, X_j))_{i,j}.$$

Es posible demostrar también que si X y Y son independientes, entonces $\text{Cov}(X, Y) = 0$. El recíproco es, en general, falso, es decir, el hecho de que la covarianza sea cero no implica necesariamente que las variables aleatorias en cuestión sean independientes. Por último, recordemos que hemos mencionado que la varianza de la suma de dos variables aleatorias no es, en general, la suma de las varianzas, sin embargo se cuenta con la siguiente fórmula general, la cual puede ser encontrada a partir de la definición de varianza y se deja como ejercicio al lector.

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \text{Cov}(X, Y). \quad (4.3)$$

A partir de los resultados que se verán en la siguiente sección, puede comprobarse además la siguiente relación general entre la covarianza y la varianza.

$$-\sqrt{\text{Var}(X) \text{Var}(Y)} \leq \text{Cov}(X, Y) \leq +\sqrt{\text{Var}(X) \text{Var}(Y)}.$$

Ejercicios

496. Calcule la covarianza entre X y Y cuando estas variables tienen distribución conjunta como se indica en cada inciso.

$x \backslash y$	0	1
a)	0	1/4
	1	1/4

$$b) f(x, y) = \begin{cases} 3 \min\{x, y\} & \text{si } 0 < x, y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f(x, y) = \begin{cases} (3/2) \max\{x, y\} & \text{si } 0 < x, y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$d) f(x, y) = \begin{cases} e^{-x} & \text{si } 0 < y < x, \\ 0 & \text{en otro caso.} \end{cases}$$

$$e) f(x, y) = \begin{cases} 2e^{-x-y} & \text{si } 0 < x < y, \\ 0 & \text{en otro caso.} \end{cases}$$

497. **Propiedades de la covarianza.** Demuestre con detalle las siguientes propiedades de la covarianza.

$$a) \text{Cov}(X, Y) = E(XY) - E(X)E(Y).$$

$$b) \text{Cov}(X, Y) = \text{Cov}(Y, X), \quad (\text{simetría}).$$

- c) $\text{Cov}(X, c) = \text{Cov}(c, X) = 0$, c constante.
- d) $\text{Cov}(cX, Y) = \text{Cov}(X, cY) = c \text{Cov}(X, Y)$, c constante.
- e) $\text{Cov}(X + c, Y) = \text{Cov}(X, Y + c) = \text{Cov}(X, Y)$, c constante.
- f) $\text{Cov}(X_1 + X_2, Y) = \text{Cov}(X_1, Y) + \text{Cov}(X_2, Y)$,
Esta propiedad, junto con la anterior y la simetría, establecen que la covarianza es una función lineal en cada variable.
- g) $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \text{Cov}(X, Y)$.
- h) Si X y Y son independientes entonces $\text{Cov}(X, Y) = 0$.
En consecuencia, cuando se cumple la hipótesis de independencia, se tiene que $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$. Es útil observar también que la propiedad enunciada proporciona un mecanismo para comprobar que dos variables aleatorias no son independientes pues, si sabemos que $\text{Cov}(X, Y) \neq 0$, podemos entonces concluir que X y Y no son independientes.
- i) En general, $\text{Cov}(X, Y) = 0 \not\Rightarrow X, Y$ independientes.

498. Sean X y Y dos variables aleatorias con valores en el intervalo $[a, b]$.

- a) Demuestre que $|\text{Cov}(X, Y)| \leq (b - a)^2/4$.
- b) Encuentre dos variables aleatorias X y Y tales que

$$|\text{Cov}(X, Y)| = (b - a)^2/4.$$

499. Sea (X, Y) un vector aleatorio continuo con distribución normal de parámetros $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \rho)$. Esta distribución se definió en el Ejercicio 460, en la página 307. Demuestre que

- a) $E(X, Y) = (\mu_1, \mu_2)$.
- b) $\text{Cov}(X, Y) = \rho\sigma_1\sigma_2$.
- c) $\text{Var}(X, Y) = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$.

4.10. Coeficiente de correlación

Habiendo definido la covarianza, podemos ahora dar la definición del coeficiente de correlación entre dos variables aleatorias. Supondremos que tales variables aleatorias tienen esperanza y varianza finitas.

Definición 4.16 El coeficiente de correlación entre las variables aleatorias X y Y con varianzas finitas distintas de cero, se define como el número

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$$

Al número $\rho(X, Y)$ se le denota también por $\rho_{X, Y}$, en donde ρ es la letra griega ro. El lector puede observar inmediatamente que la diferencia entre la covarianza y el coeficiente de correlación radica únicamente en que este último se obtiene al dividir la covarianza por el producto de las desviaciones estándares de las variables aleatorias. Puede demostrarse que este cambio de escala tiene como consecuencia que el coeficiente de correlación tome como valor máximo 1, y como valor mínimo -1 , es decir, se tiene el siguiente resultado.

Proposición 4.2 El coeficiente de correlación entre dos variables aleatorias X y Y con varianzas finitas distintas de cero satisface las desigualdades

$$-1 \leq \rho(X, Y) \leq 1.$$

Véanse los ejercicios 501 y 502 para una demostración de este resultado. Explicaremos ahora la interpretación del coeficiente de correlación. Cuando X y Y son tales que $\rho(X, Y) = 1$, entonces existen constantes a y b , con a positiva tales que $Y = aX + b$, es decir, se puede establecer una dependencia lineal directa entre las dos variables aleatorias. En el otro caso extremo, cuando $\rho(X, Y) = -1$, entonces nuevamente existen constantes a y b , pero

ahora con a negativa, tales que $Y = aX + b$. De nuevo, se trata de una relación lineal entre las dos variables aleatorias, pero ahora tal relación es inversa en el sentido de que cuando una de las variables aleatorias crece, la otra decrece. De esta forma, el coeficiente de correlación es una medida del grado de dependencia lineal entre dos variables aleatorias.

Existen varias formas en que dos variables aleatorias pueden depender una de otra, el coeficiente de correlación no mide todas estas dependencias, únicamente mide la dependencia de tipo lineal. Así, hemos mencionado que cuando el coeficiente de correlación es $+1$, ó -1 , la dependencia lineal es exacta. Como en el caso de la covarianza, puede demostrarse que si dos variables aleatorias son independientes, entonces el coeficiente de correlación es cero, y nuevamente, el recíproco es, en general, falso, es decir, la condición de que el coeficiente de correlación sea cero no es suficiente para garantizar que las variables aleatorias sean independientes, excepto en el caso en el que las variables tienen distribución conjunta normal. Esta distribución se ha definido en el Ejercicio 460, en la página 307.

Ejercicios

500. Calcule el coeficiente de correlación entre X y Y cuando estas variables tienen distribución conjunta como se indica en cada inciso.

a) $f(x, y)$ está dada por la siguiente tabla.

$x \backslash y$	0	1
0	1/4	1/4
1	1/4	1/4

$$b) f(x, y) = \begin{cases} 3x & \text{si } 0 < x < y < 1, \\ 3y & \text{si } 0 < y < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

501. Siga las siguientes indicaciones para demostrar que el coeficiente de correlación únicamente toma valores en el intervalo $[-1, 1]$.

- a) Sean U y V dos variables aleatorias, ambas con esperanza nula y con segundo momento finito. Claramente, para cualquier número real t ,

$$E[(tU - V)^2] \geq 0. \quad (4.4)$$

- b) Considerando la ecuación cuadrática en t al desarrollar el cuadrado en (4.4) y observando el signo del discriminante, obtenga la desigualdad

$$E^2(UV) \leq E(U^2) E(V^2). \quad (4.5)$$

- c) Sustituya $U = X - E(X)$ y $V = Y - E(Y)$ en (4.5) y obtenga el resultado buscado.

502. Demuestre las siguientes dos identidades y, a partir de ellas, demuestre nuevamente que el coeficiente de correlación únicamente toma valores en el intervalo $[-1, 1]$.

- a) $\text{Var}\left(\frac{X}{\sqrt{\text{Var}(X)}} + \frac{Y}{\sqrt{\text{Var}(Y)}}\right) = 2(1 + \rho(X, Y)).$
 b) $\text{Var}\left(\frac{X}{\sqrt{\text{Var}(X)}} - \frac{Y}{\sqrt{\text{Var}(Y)}}\right) = 2(1 - \rho(X, Y)).$

503. **Otras propiedades del coeficiente de correlación.** Demuestre las siguientes propiedades.

- a) $\rho(X, X) = 1.$
 b) $\rho(X, -X) = -1.$
 c) $\rho(X, Y) = \rho(Y, X).$
 d) $\rho(cX, Y) = \rho(X, cY) = \text{signo}(c) \rho(X, Y), \quad c \neq 0 \text{ constante.}$
 e) $\rho(cX, cY) = \rho(X, Y), \quad c \neq 0 \text{ constante.}$
 f) $\rho(X + c, Y) = \rho(X, Y + c) = \rho(X, Y), \quad c \text{ constante.}$
 g) $\rho(X, aX + b) = \text{signo}(a), \quad a \neq 0, b \text{ constantes.}$
 h) $\rho(X + a, X + b) = 1, \quad a, b \text{ constantes.}$

504. Sea (X, Y) un vector aleatorio continuo con distribución normal de parámetros $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \rho)$. Esta distribución se definió en el Ejercicio 460, en la página 307. Demuestre que

- a) $\rho(X, Y) = \rho.$
 b) X y Y son independientes si y sólo si $\rho = 0.$

Capítulo 5

Teoremas límite

En este último capítulo se revisarán brevemente dos de los teoremas límite más importantes en la probabilidad: la ley de los grandes números y el teorema central del límite. Las primeras secciones contienen material preparatorio para el entendimiento y demostración de estos teoremas. Se mostrarán también algunos ejemplos del uso y aplicación de estos resultados.

5.1. Desigualdad de Chebyshev

Este resultado es de carácter teórico y proporciona una cota superior para la probabilidad de que una variable aleatoria tome un valor que diste de su media en más de una cierta cantidad ϵ arbitraria.

Proposición 5.1 (Desigualdad de Chebyshev¹) Sea X una variable aleatoria con media μ y varianza finita σ^2 . Para cualquier número real $\epsilon > 0$,

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}. \quad (5.1)$$

¹Pafnuty Lvovich Chebyshev (1821–1894), matemático ruso.

Demostración. Supongamos primero que X es continua con función de densidad $f(x)$. Entonces

$$\begin{aligned}\sigma^2 &= \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \\ &\geq \int_{|x - \mu| \geq \epsilon} (x - \mu)^2 f(x) dx \\ &\geq \epsilon^2 \int_{|x - \mu| \geq \epsilon} f(x) dx \\ &= \epsilon^2 P(|X - \mu| \geq \epsilon).\end{aligned}$$

Despejando la probabilidad encontrada se obtiene el resultado. La demostración es enteramente análoga en el caso cuando X es discreta, para ello debe reemplazarse la integral por el símbolo de suma. •

Una demostración alternativa de este resultado aparece en el Ejercicio 506. Observemos que el parámetro ϵ que aparece en la desigualdad de Chebyshev debe ser, en realidad, estrictamente mayor a σ pues de lo contrario, si $0 < \epsilon \leq \sigma$, entonces $\sigma^2/\epsilon^2 \geq 1$ y tal cantidad no proporciona ninguna información útil como cota superior para una probabilidad.

Es también interesante observar que la desigualdad de Chebyshev es óptima en el sentido de que, sin hipótesis adicionales, se puede alcanzar la cota superior. En el Ejercicio 511 se pide comprobar este hecho en un caso particular. Haremos uso de la desigualdad de Chebyshev más adelante para demostrar la ley débil de los grandes números.

Ejercicios

505. Bajo las mismas condiciones y notación del enunciado de la desigualdad de Chebyshev, demuestre que:

- $P(|X - \mu| \leq \epsilon) \geq 1 - \frac{\sigma^2}{\epsilon^2}$.
- $P(|X - \mu| \geq \epsilon\sigma) \leq \frac{1}{\epsilon^2}$.
- $P(|X - \mu| \leq \epsilon\sigma) \geq 1 - \frac{1}{\epsilon^2}$.

506. **Demostración alternativa de la desigualdad de Chebyshev.** El siguiente procedimiento hace uso del método de truncamiento de una variable aleatoria para demostrar la desigualdad de Chebyshev. Bajo las mismas condiciones y notación del enunciado de esta desigualdad, lleve a cabo los siguientes pasos.

a) Defina la variable aleatoria

$$Z = \begin{cases} 0 & \text{si } (X - \mu)^2 < \epsilon^2, \\ \epsilon^2 & \text{si } (X - \mu)^2 \geq \epsilon^2. \end{cases}$$

b) Observe que $0 \leq Z \leq (X - \mu)^2$ y por lo tanto $E(Z) \leq E(X - \mu)^2$.

c) Calcule $E(Z)$ y obtenga la desigualdad de Chebyshev del inciso anterior.

507. **Desigualdad de Markov².** Demuestre cada uno de los siguientes resultados. A cualquier de ellos se le llama desigualdad de Markov.

a) Sea X una variable aleatoria no negativa y con esperanza finita μ . Demuestre que para cualquier constante $\epsilon > 0$,

$$P(X \geq \epsilon) \leq \frac{\mu}{\epsilon}. \quad (5.2)$$

b) Sea X una variable aleatoria no negativa y con n -ésimo momento finito. Demuestre que para cualquier constante $\epsilon > 0$,

$$P(X \geq \epsilon) \leq \frac{E(X^n)}{\epsilon^n}.$$

c) Sea X una variable aleatoria y sea $\varphi : \mathbb{R} \rightarrow [0, \infty)$ una función monótona no decreciente tal que $E(\varphi(X)) < \infty$ y su inversa φ^{-1} existe. Demuestre que para cualquier constante $\epsilon > 0$,

$$P(X \geq \epsilon) \leq \frac{E(\varphi(X))}{\varphi(\epsilon)}.$$

²Andrey Andreyevich Markov (1856–1922), matemático ruso.

508. Sea X una variable aleatoria con f.g.m. $M(t)$. Demuestre que para $t > 0$ y $x > 0$,

$$P(X \geq x) \leq e^{-tx} M(t).$$

509. Sea X una variable aleatoria tal que $\text{Var}(X) = 0$. Demuestre que existe una constante c tal que $P(X = c) = 1$.

Nota. Compare este enunciado con el resultado en el caso discreto que aparece en el Ejercicio 247, en la página 177.

510. **Markov \Rightarrow Chebyshev.** Sea X una variable aleatoria con varianza finita. Demuestre la desigualdad de Chebyshev (5.1) para esta variable aleatoria, a partir de la desigualdad de Markov (5.2).

511. **La desigualdad de Chebyshev es óptima.** Este resultado demuestra que, sin hipótesis adicionales, la cota superior dada por la desigualdad de Chebyshev es óptima, es decir, no puede establecerse una cota superior más pequeña. Sea X una variable aleatoria discreta con función de probabilidad

$$f(x) = \begin{cases} 1/18 & \text{si } x = -1, 1, \\ 16/18 & \text{si } x = 0, \\ 0 & \text{en otro caso.} \end{cases}$$

- a) Calcule la esperanza μ y la varianza σ^2 de esta v.a.
- b) Ahora calcule exactamente $P(|X - \mu| \geq 3\sigma)$ y compruebe que esta cantidad coincide con la cota superior dada por la desigualdad de Chebyshev.
512. Sea X una variable aleatoria con distribución $N(\mu, \sigma^2)$.
- a) Use la desigualdad de Chebyshev para estimar el valor mínimo del número real $k > 0$ de tal modo que la probabilidad de que X tome un valor entre $\mu - k\sigma$ y $\mu + k\sigma$ sea, al menos, 0.95.
- b) Use la tabla de la distribución normal para encontrar el valor de k que cumpla la condición del inciso anterior.
513. Sea $\Phi(x)$ la función de distribución $N(0, 1)$. Use la desigualdad de Chebyshev para demostrar que para cualquier $x > 0$,

$$a) \Phi(x) \geq 1 - \frac{1}{2x^2}.$$

$$b) \Phi(-x) \leq \frac{1}{2x^2}.$$

514. Use la distribución $\exp(\lambda)$ y la desigualdad de Chebyshev para demostrar que para cualquier número real $x \geq 1$,

$$e^{-(x+1)} \leq 1/x^2.$$

515. Sea X una variable aleatoria con función de densidad $f(x) = (1/2)e^{-|x|}$ para $-\infty < x < \infty$. Sea μ su media y σ^2 su varianza. Calculando la probabilidad $P(|X - \mu| \geq x)$ y la cota superior de Chebyshev dada por σ^2/x^2 , demuestre que para cualquier $x > 0$,

$$e^{-x} \leq 2/x^2.$$

5.2. Convergencia de variables aleatorias

Sabemos que una sucesión numérica x_1, x_2, \dots es convergente a un número x si para cualquier $\epsilon > 0$ existe un número natural N a partir del cual los elementos de la sucesión se encuentran cercanos al número x , es decir, para $n \geq N$,

$$|x_n - x| < \epsilon.$$

Si en lugar de la sucesión numérica tenemos una sucesión de variables aleatorias, ¿cómo se puede definir el concepto de convergencia en este caso? Veremos a continuación que puede responderse a esta pregunta de varias maneras. Consideremos entonces que tenemos una sucesión infinita de variables aleatorias X_1, X_2, \dots y un espacio de probabilidad en donde todas estas variables aleatorias están definidas. La variedad de formas en las que puede definirse la convergencia de variables aleatorias estará dada por las formas en las que se decida medir la cercanía de la sucesión con el límite a través de la medida de probabilidad.

Convergencia puntual

Para cada ω fijo, la sucesión $X_1(\omega), X_2(\omega), \dots$ es una sucesión de números reales, por lo tanto podemos definir la convergencia de las variables aleatorias cuando esta sucesión numérica es convergente para cada ω fijo. En este caso, la variable aleatoria límite se define de forma puntual:

$$X(\omega) := \lim_{n \rightarrow \infty} X_n(\omega).$$

A este tipo de convergencia se le llama convergencia puntual y se escribe

$$X_n \rightarrow X, \quad \text{para cada } \omega \in \Omega.$$

Convergencia casi segura (o fuerte)

Un tipo de convergencia menos estricta que la anterior ocurre cuando se permite que la convergencia puntual se observe sobre un conjunto de probabilidad uno, es decir, se dice que la sucesión X_1, X_2, \dots converge casi seguramente, o casi dondequiera, a la variable X si para casi toda ω , $X_n(\omega)$ converge a $X(\omega)$, en símbolos,

$$P(\omega \in \Omega : X_n(\omega) \rightarrow X(\omega)) = 1,$$

y se escribe

$$X_n \xrightarrow{\text{c.s.}} X.$$

De este modo se permite que exista un subconjunto de Ω en donde no se verifique la convergencia, pero tal subconjunto debe tener medida de probabilidad cero. A este tipo de convergencia también se le llama convergencia fuerte. Es claro que si una sucesión de v.a.s es convergente puntualmente, entonces es también convergente en el sentido casi seguro. El recíproco es falso.

Convergencia en probabilidad

Otra forma aún menos restrictiva que la convergencia casi segura es la siguiente. La sucesión de variables aleatorias X_1, X_2, \dots converge en probabilidad a la variable aleatoria X si para cualquier $\epsilon > 0$,

$$P(\omega \in \Omega : |X_n(\omega) - X(\omega)| > \epsilon) \rightarrow 0,$$

cuando n tiende a infinito. Es decir, la probabilidad del conjunto en donde X_n y X distan en más de ϵ converge a cero conforme $n \rightarrow \infty$. En este caso se escribe

$$X_n \xrightarrow{p} X.$$

Convergencia en distribución (o débil)

El último tipo de convergencia que consideraremos hace uso de las funciones de distribución de las variables aleatorias. Se dice que la sucesión de variables X_1, X_2, \dots converge en distribución, o que converge débilmente, a la variable aleatoria X si cuando $n \rightarrow \infty$,

$$F_{X_n}(x) \rightarrow F_X(x)$$

para todo punto x en donde $F_X(x)$ es continua. Es decir, para aquellos valores reales x que cumplan la condición mencionada, debe verificarse que

$$\lim_{n \rightarrow \infty} P(X_n \leq x) = P(X \leq x).$$

En este caso se escribe

$$X_n \xrightarrow{d} X.$$

El teorema de continuidad de la función generadora de momentos que enunciamos en la página 199 se refiere precisamente a este tipo de convergencia. Existen otros tipos de convergencia para variables aleatorias, pero los que hemos mencionado son suficientes para poder enunciar algunos teoremas límite importantes en probabilidad. Antes de pasar al estudio de estos resultados, responderemos a la pregunta que posiblemente se habrá hecho el lector respecto a las posibles relaciones que pudieran existir entre los tipos de convergencia de variables aleatorias:

¿existe alguna relación entre los diferentes tipos de convergencia de variables aleatorias?

La respuesta a esta pregunta se muestra gráficamente en la Figura 5.1, en donde un punto dentro de alguna región representa una sucesión de v.a.s que es convergente en el sentido indicado. La contención de una región en otra significa que el tipo de convergencia de la región contenida implica el

tipo de convergencia de la región que contiene, así, por ejemplo, la convergencia casi segura implica la convergencia en probabilidad, y ésta, a su vez, implica la convergencia en distribución. El diagrama establece además que las contenciones son propias, es decir, existen elementos en una región que no pertenecen a los subconjuntos contenidos, por ejemplo, existen sucesiones de v.a.s que convergen en probabilidad, pero no convergen en el sentido casi seguro. El lector interesado en la demostración de las afirmaciones anteriores y en ejemplos de sucesiones de variables aleatorias convergentes en un sentido u otro, puede consultar textos más avanzados de probabilidad como [9].

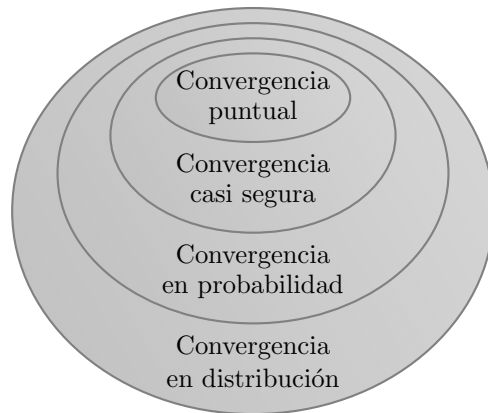


Figura 5.1

Los teoremas límite que estudiaremos en las siguientes secciones tratan sobre la convergencia de sucesiones de variables aleatorias de la forma

$$\frac{1}{n} \sum_{i=1}^n X_i.$$

Estamos listos ahora para enunciar y demostrar la ley de los grandes números y el teorema central del límite.

Ejercicios

516. Para cada número natural n suponga que X_n es una variable aleatoria con distribución $\text{unif}(0, 1/n)$. Sea X la v.a. constante cero. Demuestre que

$$X_n \xrightarrow{d} X.$$

517. Este es un caso particular en donde la convergencia en probabilidad es equivalente a la convergencia en distribución. Sea c una constante. Demuestre que

$$X_n \xrightarrow{p} c \quad \text{si y sólo si} \quad X_n \xrightarrow{d} c.$$

5.3. La ley de los grandes números

El teorema conocido como la ley de los grandes números es un resultado muy interesante que puede observarse en la naturaleza. Constituye uno de los resultados más importantes de la teoría de la probabilidad y tiene mucha relevancia en las aplicaciones tanto teóricas como prácticas. Este teorema establece que, bajo ciertas condiciones, el promedio aritmético de variables aleatorias converge a una constante cuando el número de sumandos crece a infinito. Ya desde el siglo XVI, el matemático Gerolamo Cardano (1501-1576) había hecho la observación de que la precisión de las estadísticas empíricas mejoraban conforme se incrementaba el número de observaciones. Pero fue Jacobo Bernoulli quien, en 1713 y después de muchos años de trabajo, logró formalizar por primera vez el enunciado del teorema y dar una demostración rigurosa para el caso de variables aleatorias dicotómicas. Debido a este gran éxito en la carrera de Jacobo Bernoulli, a este resultado se le conoce también como teorema de Bernoulli. Sin embargo, fue Simone D. Poisson quien usó y popularizó el término *ley de los grandes números*. Otros matemáticos han contribuido notablemente a la generalización y extensión del teorema de Bernoulli, entre ellos están Chebyshev, Markov, Borel, Cantelli, Kolmogorov y Khinchin.

Teorema 5.1 (Ley de los grandes números) Sea X_1, X_2, \dots una sucesión infinita de variables aleatorias independientes e idénticamente distribuidas con media finita μ . Entonces, cuando $n \rightarrow \infty$,

$$\frac{1}{n} \sum_{i=1}^n X_i \longrightarrow \mu.$$

en donde la convergencia se verifica en el sentido casi seguro (ley fuerte) y también en probabilidad (ley débil).

Demostración. (Ley débil, es decir, convergencia en probabilidad, suponiendo segundo momento finito). Recordemos que la desigualdad de Chebyshev para una variable aleatoria X con media μ y varianza σ^2 establece que para cualquier $\epsilon > 0$,

$$P(|X - \mu| > \epsilon) \leq \frac{\sigma^2}{\epsilon^2}.$$

Aplicando este resultado a la variable aleatoria $S_n = (X_1 + \dots + X_n)/n$, cuya esperanza es μ y varianza es σ^2/n , tenemos que para cualquier $\epsilon > 0$,

$$P(|S_n - \mu| > \epsilon) \leq \frac{\sigma^2}{n\epsilon^2}.$$

De modo que, al hacer $n \rightarrow \infty$, se obtiene que

$$P(|S_n - \mu| > \epsilon) \rightarrow 0,$$

lo cual significa que

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{p} \mu.$$

•

Así, sin importar la distribución de las variables aleatorias, el promedio aritmético converge a la media μ conforme n tiende a infinito. Como se ha mencionado, únicamente se ha presentado la demostración en el caso cuando la convergencia es en probabilidad y suponiendo adicionalmente que el segundo momento existe. Demostraciones más generales de este resultado

pueden encontrarse, por ejemplo, en el texto de Gut [9].

La siguiente demostración alternativa de la ley débil (convergencia en probabilidad) es también bastante directa, aunque tiene la desventaja de suponer la existencia de la función generadora de momentos. Tal hipótesis adicional garantizará además la convergencia en distribución de los promedios parciales.

Demostración. Sea nuevamente $S_n = (X_1 + \cdots + X_n)/n$ y sea $M_X(t)$ la f.g.m. de cualquiera de las variables X_i . Haremos uso de la expansión (2.25) de la página 197 y de la notación o -pequeña, que puede consultarse en el apéndice en la página 381. La función generadora de momentos de la variable S_n es

$$\begin{aligned} M_{S_n}(t) &= E(e^{tS_n}) \\ &= E(e^{\frac{t}{n}(X_1 + \cdots + X_n)}) \\ &= (M_X(\frac{t}{n}))^n \\ &= (1 + \frac{t}{n}E(X) + o(\frac{t}{n}))^n \\ &= (1 + \frac{t}{n}E(X))^n + o(\frac{t}{n}), \end{aligned}$$

en donde se ha escrito la n -ésima potencia de un trinomio en dos sumandos: uno en donde el primer término tiene exponente n y el segundo $o(\frac{t}{n})$ desaparece pues tiene exponente cero, el segundo sumando tiene potencias positivas de $o(\frac{t}{n})$ y todos estos términos se agrupan en una misma expresión escrita como $o(\frac{t}{n})$. Por lo tanto,

$$\lim_{n \rightarrow \infty} M_{S_n}(t) = e^{tE(X)},$$

en donde esta última expresión corresponde a la f.g.m. de la variable aleatoria constante igual a $E(X)$. Por la Proposición 2.1 de la página 199 sobre la continuidad de las funciones generadoras de momentos, tenemos que

$$S_n \xrightarrow{d} E(X).$$

En el Ejercicio 517 se pide comprobar que la convergencia en distribución a una constante es equivalente a la convergencia en probabilidad a la misma constante. Por lo tanto, también tenemos que

$$S_n \xrightarrow{p} E(X).$$

•

Se verán ahora algunos ejemplos de aplicación de este resultado.

Ejemplo 5.1 (Probabilidad frecuentista) Consideremos un experimento aleatorio cualquiera, y sea A un evento con probabilidad desconocida p . Nos interesa encontrar este valor de p . Suponga que se efectúan varias realizaciones sucesivas e independientes del experimento y se observa, en cada ensayo, la ocurrencia o no ocurrencia del evento A . Para cada entero $i \geq 1$ defina la variable aleatoria

$$X_i = \begin{cases} 1 & \text{si ocurre el evento } A \text{ en el } i\text{-ésimo ensayo,} \\ 0 & \text{si no ocurre el evento } A \text{ en el } i\text{-ésimo ensayo.} \end{cases}$$

Entonces las variables X_1, X_2, \dots son independientes, cada una con distribución $\text{Ber}(p)$, en donde p es la probabilidad del evento A . Por lo tanto, $E(X_i) = p$ y $\text{Var}(X_i) = p(1-p)$. La ley de los grandes números asegura que la fracción de ensayos en los que se observa el evento A converge a la constante desconocida p cuando el número de ensayos crece a infinito, es decir,

$$\frac{n_A}{n} = \frac{1}{n} \sum_{i=1}^n X_i \rightarrow p.$$

Esta es la definición frecuentista de la probabilidad que habíamos estudiado antes en la página 31, de la cual hemos podido ahora corroborar su validez con la ayuda de la teoría desarrollada a partir de los axiomas de la probabilidad de Kolmogorov y, en particular, de la ley de los grandes números.

•

Ejemplo 5.2 (Método de Montecarlo) Supongamos que se desea calcular la integral

$$\int_0^1 g(x) dx,$$

para una cierta función $g(x)$ integrable en el intervalo $(0, 1)$. Esta integral puede escribirse como

$$\int_0^1 g(x) f(x) dx,$$

en donde $f(x)$ es la función de densidad de la distribución uniforme continua en $(0, 1)$, es decir, es idénticamente uno en este intervalo. De este modo la integral anterior puede identificarse como la esperanza de la v.a. $g(X)$, en donde $X \sim \text{unif}(0, 1)$, es decir,

$$\int_0^1 g(x) dx = E[g(X)],$$

suponiendo que $g(X)$ es, efectivamente, una variable aleatoria. Si X_1, X_2, \dots son v.a.s i.i.d. con distribución $\text{unif}(0, 1)$, entonces $g(X_1), g(X_2), \dots$ también son v.a.s i.i.d. (no necesariamente con distribución uniforme) y, por la ley de los grandes números, tenemos que, cuando $n \rightarrow \infty$,

$$\frac{1}{n} \sum_{i=1}^n g(X_i) \rightarrow \int_0^1 g(x) dx.$$

Así, un conjunto de observaciones x_1, \dots, x_n de la distribución $\text{unif}(0, 1)$ puede usarse para resolver, de manera aproximada, el problema de integración originalmente planteado:

$$\frac{1}{n} \sum_{i=1}^n g(i) \approx \int_0^1 g(x) dx.$$

A este procedimiento se le conoce como el método de Montecarlo. En la literatura científica, tal término se refiere a una amplia gama de métodos computacionales que hacen uso de muestras aleatorias para estimar un valor numérico de interés. Por ejemplo, en el Ejercicio 370, en la página 253, se explica una forma de aproximar el valor de π a través de muestras de una variable aleatoria con distribución $\text{unif}(0, 1)$. •

Simulación 5.1 En este ejemplo se utiliza el paquete R para ilustrar la ley de los grandes números. Si se define la variable

$$S_n = \frac{1}{n}(X_1 + \dots + X_n),$$

**Programa en R para ilustrar la ley de los grandes números
en el caso de variables aleatorias $\text{Ber}(p)$ con $p = 1/2$**

```

N <- 200                                     # Número de ensayos
s <- rep(1,N)                                # Vector s de tamaño N, cada entrada es 1
p <- 0.5                                     # Parámetro p de la distribución Bernoulli
s[1] <- rbinom(1,1,p)                        # Primera entrada del vector s
for(n in 2:N){                               # Se calcula s[n] a partir de s[n-1]
  s[n] <- ((n-1)*s[n-1]+rbinom(1,1,p))/n
}
plot(s,type="l")                             # Graficación, la letra es ele, no uno
abline(h=p)                                  # Graficación línea de referencia horizontal

```

Figura 5.2

entonces podemos expresar S_n en términos de S_{n-1} de la siguiente forma:
para $n \geq 2$,

$$S_n = \frac{1}{n}((n-1)S_{n-1} + X_n). \quad (5.3)$$

Esta expresión es muy útil para analizar numéricamente el comportamiento de S_n a lo largo del tiempo. En particular, se simularán 200 lanzamientos independientes de una moneda equilibrada, es decir, se generarán 200 valores al azar de una variable aleatoria con distribución $\text{Ber}(p)$, con $p = 1/2$. El código R se muestra en la Figura 5.2. Los resultados obtenidos en R se exportaron después a $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ para generar la gráfica en $\text{P}^{\text{S}}\text{Tricks}$ que se muestra en la Figura 5.3. En esta gráfica puede apreciarse el comportamiento oscilante de los promedios $S_n = (X_1 + \cdots + X_n)/n$ conforme n crece. Se observa un comportamiento inicial errático y su eventual estabilización en el valor $1/2$.

Los puntos graficados fueron unidos por una línea continua para una mejor visualización. En este ejemplo se ha utilizado la distribución Bernoulli, pero cualquier otra distribución puede ser usada para observar el interesante acercamiento del promedio S_n a la media de la distribución. •

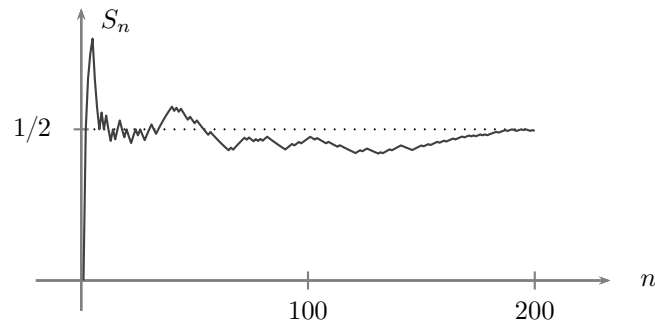


Figura 5.3

Ejercicios

518. Simulación de la ley de los grandes números.

- a) Escoja usted una distribución de probabilidad discreta de su preferencia, especificando valores numéricos para sus parámetros. Genere $n = 1000$ valores independientes al azar x_1, \dots, x_n de esta distribución y calcule los promedios parciales

$$s_n = \frac{1}{n} \sum_{i=1}^n x_i \quad n = 1, 2, \dots, 1000.$$

Grafique la función $n \mapsto s_n$ uniendo con una línea recta sus valores. Denote por μ la media de la distribución. Trace en la misma gráfica la función constante μ y compruebe gráficamente que la función $n \mapsto s_n$ oscila y se aproxima al valor μ conforme n crece. Esta es una comprobación experimental de la ley de los grandes números.

- b) Haga lo mismo que en el inciso anterior, ahora con una distribución continua de su preferencia.

519. Estimaciones.

Sea x_1, \dots, x_n una colección finita de observaciones de una v.a. con distribución normal con media μ y varianza σ^2 desconocidas. Con base en la ley de los grandes números, proporcione

expresiones en función de x_1, \dots, x_n que puedan servir para estimar los valores de μ y σ^2 .

520. **Convergencia de la media geométrica.** Sea X_1, X_2, \dots una sucesión de v.a.s positivas, independientes, idénticamente distribuidas y tales que $E(\ln X_1) = \mu < \infty$. Demuestre que, cuando $n \rightarrow \infty$,

$$\sqrt[n]{X_1 \cdots X_n} \xrightarrow{c.s.} e^\mu.$$

521. **Teorema de equipartición asintótica.** La entropía en base 2 de una variable aleatoria discreta X con función de probabilidad $p(x)$ se define como el número

$$H(X) := - \sum_x p(x) \log_2 p(x).$$

Demuestre que si X_1, X_2, \dots es una sucesión de v.a.s discretas, independientes, idénticamente distribuidas, con función de probabilidad $p(x)$ y con entropía finita, entonces, cuando $n \rightarrow \infty$,

$$-\frac{1}{n} \log_2 p(X_1, \dots, X_n) \xrightarrow{c.s.} H(X_1),$$

en donde $p(x_1, \dots, x_n)$ también denota la función de probabilidad conjunta de las variables X_1, \dots, X_n . En la teoría de la información, a este resultado se le conoce como el teorema de equipartición asintótica.

522. **Una extensión de la ley débil de los grandes números.** Sea X_1, X_2, \dots una sucesión de variables aleatorias independientes, no necesariamente idénticamente distribuidas, pero con media común μ y varianzas acotadas, es decir, $\text{Var}(X_i) \leq c$, en donde c es una constante. Use la desigualdad de Chebyshev para demostrar que, cuando $n \rightarrow \infty$,

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{p} \mu.$$

523. **Otra extensión de la ley débil de los grandes números.** Sea X_1, X_2, \dots una sucesión de variables aleatorias, no necesariamente idénticamente distribuidas ni independientes, pero con media común

μ y varianzas tales que $\text{Var}(\frac{1}{n} \sum_{i=1}^n X_i) \rightarrow 0$ cuando $n \rightarrow \infty$. Use la desigualdad de Chebyshev para demostrar que, cuando $n \rightarrow \infty$,

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{p} \mu.$$

5.4. El teorema central del límite

Este teorema es muy importante y tiene una gama amplia de aplicaciones. En nuestro caso, nos ayudará a simplificar el cálculo de ciertas probabilidades y a aproximar algunas distribuciones.

Teorema 5.2 (Teorema central del límite) Sea X_1, X_2, \dots una sucesión infinita de variables aleatorias independientes e idénticamente distribuidas, con media μ y varianza finita σ^2 . Entonces la función de distribución de la variable aleatoria

$$Z_n = \frac{(X_1 + \dots + X_n) - n\mu}{\sqrt{n\sigma^2}}$$

tiende a la función de distribución normal estándar cuando n tiende a infinito.

Demostración. Por simplicidad, supondremos que, a diferencia del enunciado del teorema, las variables aleatorias tienen todos sus momentos finitos de tal manera que su f.g.m. es finita y tiene la expresión de la serie de potencias (2.25) de la página 197. Nuevamente haremos uso de la notación o -pequeña. Podemos calcular la f.g.m. de la variable Z_n de la siguiente ma-

nera.

$$\begin{aligned}
 M_{Z_n}(t) &= E(e^{tZ_n}) \\
 &= E\left(e^{\frac{t}{\sqrt{n}}\left(\frac{X_1-\mu}{\sigma}+\dots+\frac{X_n-\mu}{\sigma}\right)}\right) \\
 &= \left(M_{\frac{X-\mu}{\sigma}}\left(\frac{t}{\sqrt{n}}\right)\right)^n \\
 &= \left(1 + \frac{t}{\sqrt{n}}E\left(\frac{X-\mu}{\sigma}\right) + \frac{t^2}{2n}E\left(\frac{X-\mu}{\sigma}\right)^2 + o\left(\frac{t^2}{n}\right)\right)^n \\
 &= \left(1 + \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right)^n \\
 &= \left(1 + \frac{t^2}{2n}\right)^n + o\left(\frac{t^2}{n}\right),
 \end{aligned}$$

en donde se ha escrito la n -ésima potencia de un trinomio en dos sumandos: uno en donde el primer término tiene exponente n y el segundo $o\left(\frac{t^2}{n}\right)$ desaparece pues tiene exponente cero, el segundo sumando tiene potencias positivas de $o\left(\frac{t^2}{n}\right)$ y todos estos términos se agrupan en una misma expresión escrita como $o\left(\frac{t^2}{n}\right)$. Por lo tanto,

$$\lim_{n \rightarrow \infty} M_{Z_n}(t) = e^{t^2/2},$$

en donde esta última expresión corresponde a la f.g.m. de una variable aleatoria con distribución $N(0, 1)$. Por la continuidad de las funciones generadoras de momentos,

$$Z_n \xrightarrow{d} N(0, 1).$$

•

En términos matemáticos, este resultado establece que para cualquier número real x ,

$$\lim_{n \rightarrow \infty} F_{Z_n}(x) = \Phi(x),$$

sin importar la distribución de las variables X_1, X_2, \dots , así es que éstas pueden tener distribución Bernoulli, binomial, exponencial, gamma, etc., en general, pueden ser discretas o continuas, y este resultado sorprendente asegura que la variable Z_n tiene siempre una distribución aproximada normal estándar para valores grandes de n .

Una forma de expresar el teorema central del límite de manera informal es a través de la siguiente afirmación: para cualesquiera números reales $a < b$,

$$P\left(a < \frac{(X_1 + \cdots + X_n) - n\mu}{\sqrt{n\sigma^2}} < b\right) \approx \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

Esto nos permitirá aproximar probabilidades de eventos que involucran sumas de variables aleatorias en términos de probabilidades de la distribución normal estándar. Observe que, dividiendo el numerador y denominador entre n , y definiendo $S_n = (X_1 + \cdots + X_n)/n$, la variable Z_n puede escribirse de la siguiente forma

$$Z_n = \frac{S_n - \mu}{\sqrt{\sigma^2/n}}. \quad (5.4)$$

Es interesante observar también que la ley de los grandes números asegura que el numerador de (5.4) converge a cero conforme n tiende a infinito, sin embargo, el denominador de esta expresión también converge a cero y estos límites ocurren de tal manera que el cociente no es constante, sino una variable aleatoria con distribución normal estándar.

Una de las primeras versiones demostradas del teorema central del límite es aquella en donde las variables aleatorias tienen distribución Bernoulli. Se enuncia a continuación este resultado en el contexto de ocurrencias o no ocurrencias de un evento en una sucesión de ensayos independientes de un experimento aleatorio cualquiera.

Teorema 5.3 (Teorema de De Moivre-Laplace)

Suponga que se tiene una sucesión infinita de ensayos independientes de un experimento aleatorio. Sea A un evento de este experimento aleatorio con probabilidad de ocurrencia $p > 0$. Sea n_A el número de ocurrencias del evento de interés en los primeros n ensayos del experimento. Entonces para cualesquiera números reales $a < b$,

$$\lim_{n \rightarrow \infty} P\left(a < \frac{n_A/n - p}{\sqrt{p(1-p)/n}} < b\right) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

**Programa en R para ilustrar el teorema central del límite
en el caso de variables aleatorias $\text{Ber}(p)$ con $p = 0.7$**

```
# Número de valores simulados de la v.a. suma s[1],...,s[k]:
k <- 1000
s <- rep(0,k)
# Número de sumandos x[1],...,x[n]:
n <- 5
x <- rep(0,n)
# Parámetro(s):
p <- 0.7
# Generación al azar de n sumandos x[i] y k sumas s[i]:
for (i in 1:k){
  x <- rbinom(n,1,p)
  s[i] <- sum(x)
}
# Cálculo de media, varianza y estandarización
media <- n*p
var <- n*p*(1-p)
s <- (s-media)/sqrt(var)
# Graficación de la función de densidad
par(mfrow=c(1,2))
curve(dnorm(x),from=-3,to=3,ylim=c(0,0.6),ylab="F. de
densidad",lwd=2,col="blue")
hist(s,freq=FALSE,breaks=50,add=T,xlim=c(-3,3),ylim=c(0,5))
# Graficación de la función de distribución
distempirica <- ecdf(s)
curve(distempirica,from=-3,to=3,cex=0.1,ylab="F. de dist.")
curve(pnorm(x),from=-3,to=3,add=TRUE,col="blue")
```

Figura 5.4

Simulación 5.2 En el código que aparece en la Figura 5.4 se muestra la forma en la que puede usarse el paquete `R` para comprobar, mediante simulación, el teorema central del límite. Como en el enunciado del teorema, el parámetro n corresponde al número de sumandos en $X_1 + \dots + X_n$. El parámetro $k = 1000$ se usa para generar k valores al azar de la variable aleatoria suma $X_1 + \dots + X_n$, y de esa forma aproximar su función de distribución. En estas simulaciones se ha utilizado la distribución $\text{Ber}(p)$ con $p = 0.7$ para los sumandos, lo cual puede modificarse con facilidad. Los resultados aparecen en la Figura 5.5, en la página 366, para $n = 5, 50, 100, 200, 500$ y 1000 . Puede apreciarse con claridad la forma sorprendente en la que la función de distribución de la variable

$$Z_n = \frac{(X_1 + \dots + X_n) - n\mu}{\sqrt{n\sigma^2}}$$

se hace cada vez más parecida a la función de distribución normal estándar. Para obtener las gráficas mostradas en la Figura 5.5 de la página 366 se obtuvieron primero los datos en `R` usando el ambiente gráfico `RStudio`, se trasladaron después estos datos a `LATEX` para su graficación a través del paquete `PSTricks`. A fin de que aparecieran únicamente líneas horizontales y verticales en la función de distribución aproximante, se llevó a cabo un proceso de interpolación en los subintervalos en donde aparecían líneas inclinadas. Las últimas dos gráficas requirieron un refinamiento en el número de puntos a graficar.

Para este mismo ejemplo se muestran en la Figura 5.6, en la página 367, varios histogramas que, paulatinamente, adquieren la forma de la función de densidad normal estándar conforme el parámetro n crece. En este tipo de gráficas, y para hacer comparaciones entre dos histogramas, es importante escoger de manera adecuada el tamaño de la base de los rectángulos. •

Ejemplo 5.3 Se lanza una dado repetidas veces y se definen las variables aleatorias X_1, X_2, \dots como los resultados de estos lanzamientos. Es razonable suponer que estas variables aleatorias son independientes y con idéntica distribución uniforme en el conjunto $\{1, 2, 3, 4, 5, 6\}$. En particular, la esperanza es $\mu = 3.5$ y la varianza es $\sigma^2 = 2.91\bar{6}$. Por la ley de los grandes

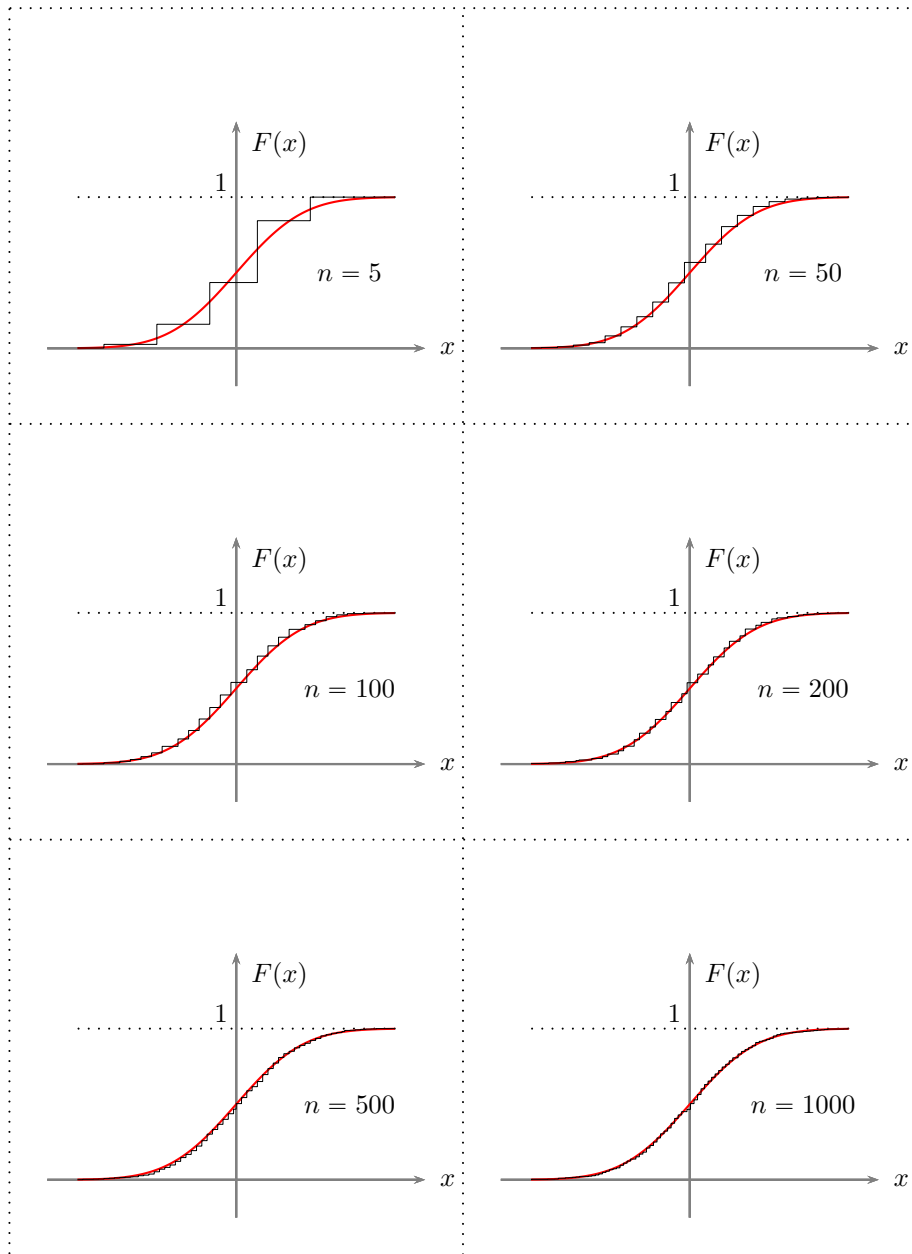


Figura 5.5

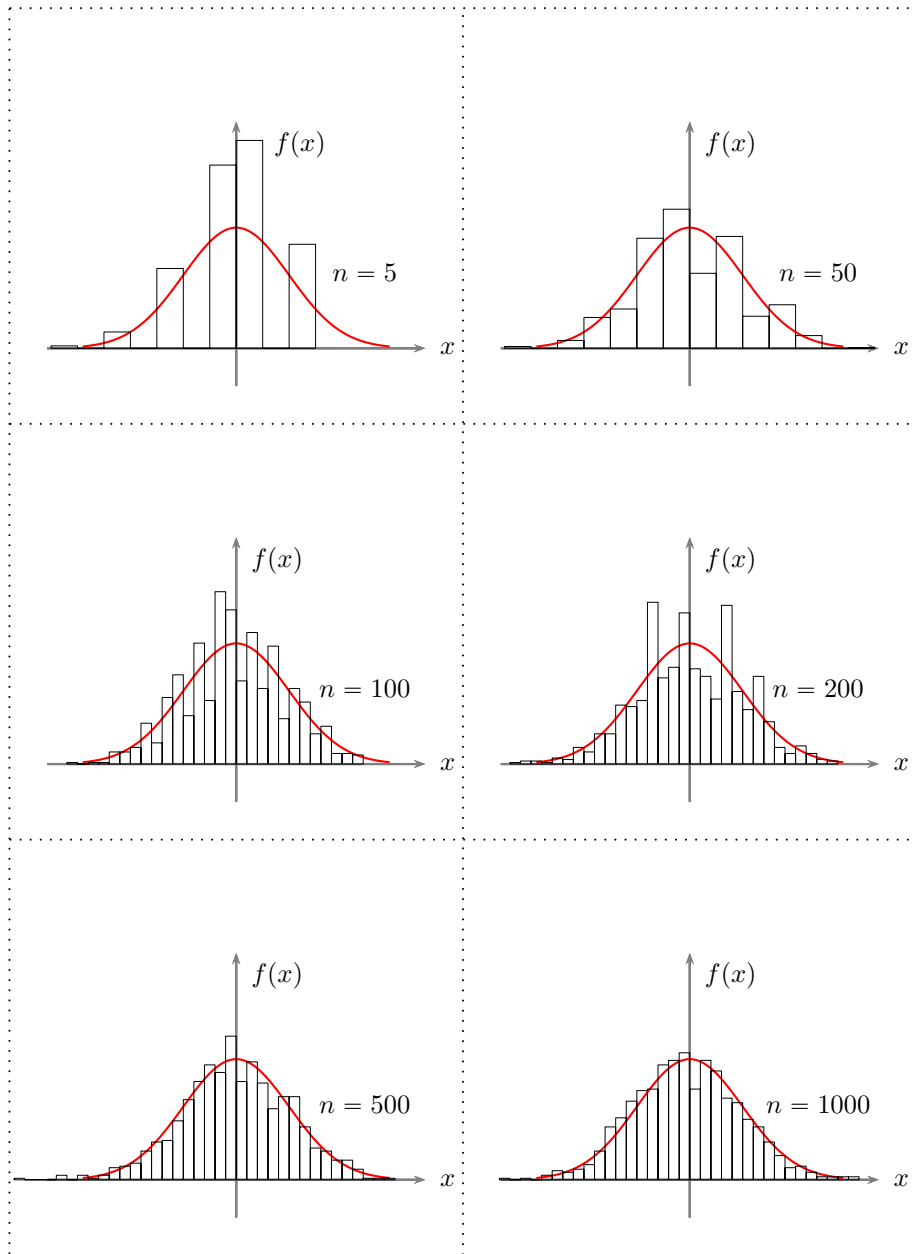


Figura 5.6

números, sabemos que el promedio parcial $S_n = (X_1 + \cdots + X_n)/n$ se aproxima a la media 3.5 conforme n crece. ¿Cuántas veces debe lanzarse el dado de tal forma que S_n se encuentre entre 3 y 4 con una probabilidad de 0.99?

Solución. Se busca el valor de n tal que

$$P(3 \leq S_n \leq 4) = 0.99.$$

Restando en cada lado de las desigualdades la media μ y dividiendo entre $\sqrt{\sigma^2/n}$, la igualdad anterior es equivalente a la expresión

$$P\left(\frac{3 - 3.5}{\sqrt{\sigma^2/n}} \leq \frac{S_n - 3.5}{\sqrt{\sigma^2/n}} \leq \frac{4 - 3.5}{\sqrt{\sigma^2/n}}\right) = 0.99.$$

Por el teorema central del límite, la probabilidad indicada es aproximadamente igual a $P(-0.5/\sqrt{\sigma^2/n} \leq Z \leq 0.5/\sqrt{\sigma^2/n})$, en donde Z es una variable aleatoria con distribución normal estándar. Es decir, tenemos ahora la ecuación de aproximación

$$\Phi\left(\frac{0.5}{\sqrt{\sigma^2/n}}\right) - \Phi\left(\frac{-0.5}{\sqrt{\sigma^2/n}}\right) = 0.99.$$

De tablas de la distribución normal estándar, puede verificarse que el valor de x tal que $\Phi(x) - \Phi(-x) = 0.99$ es $x = 2.58$. De este modo, se tiene que $0.5/\sqrt{\sigma^2/n} = 2.58$, de donde se obtiene $n = 226.5$. •

Ejemplo 5.4 Se desea diseñar un estacionamiento de coches para un conjunto de 200 departamentos que se encuentran en construcción. Suponga que para cada departamento, el número de automóviles será de 0, 1 o 2, con probabilidades 0.1, 0.6 y 0.3, respectivamente. Se desea que, con una certeza del 95 %, haya espacio disponible para todos los coches cuando los departamentos se vendan. ¿Cuántos espacios de estacionamiento deben construirse?

Solución. Sean X_1, \dots, X_{200} las variables aleatorias que denotan el número de automóviles que poseen los futuros dueños de los departamentos. Podemos suponer que estas variables aleatorias discretas son independientes unas

de otras y todas ellas tienen la misma distribución de probabilidad:

$$\begin{aligned}P(X = 0) &= 0.1, \\P(X = 1) &= 0.6, \\P(X = 2) &= 0.3.\end{aligned}$$

De esta forma, la variable aleatoria suma $X_1 + \cdots + X_{200}$ denota el total de automóviles que habrá en el complejo de departamentos. Se desconoce la distribución de esta variable aleatoria, sin embargo se desea encontrar el valor de n tal que $P(X_1 + \cdots + X_{200} \leq n) = 0.95$. Haremos uso del teorema central del límite para resolver este problema, y para ello se necesita calcular la esperanza y varianza de X . Puede comprobarse que $E(X) = 1.2$ y $\text{Var}(X) = 0.36$, cantidades que denotaremos por μ y σ^2 , respectivamente. La ecuación planteada es entonces

$$P(X_1 + \cdots + X_{200} \leq n) = 0.95,$$

en donde la incógnita es el valor de n . Restando en ambos lados de la desigualdad 200μ y dividiendo entre $\sqrt{200\sigma^2}$, la ecuación anterior es equivalente a

$$P\left(\frac{X_1 + \cdots + X_{200} - 200\mu}{\sqrt{200\sigma^2}} \leq \frac{n - 200\mu}{\sqrt{200\sigma^2}}\right) = 0.95. \quad (5.5)$$

Por el teorema central del límite, la probabilidad indicada es aproximada a $\Phi((n - 200\mu)/\sqrt{200\sigma^2})$. De este modo, tenemos ahora la ecuación

$$\Phi\left(\frac{n - 200\mu}{\sqrt{200\sigma^2}}\right) = 0.95.$$

Observe que la variable aleatoria que aparece en la ecuación (5.5) y cuya distribución de probabilidad se desconoce y en general es difícil encontrar, se ha aproximado por una variable aleatoria normal estándar, y allí radica la utilidad del teorema central del límite. De la tabla de la distribución normal, podemos ahora verificar que el valor de x tal que $\Phi(x) = 0.95$ es $x = 1.65$. De este modo se llega a la igualdad $(n - 200\mu)/\sqrt{200\sigma^2} = 1.65$, de donde se obtiene que $n = 253.99$. Es decir, el tamaño del estacionamiento debe ser de aproximadamente 254 lugares. •

Ejercicios

524. Simulación del teorema central del límite.

- a) Escoja usted una distribución de probabilidad discreta de su preferencia, especificando valores numéricos para sus parámetros. Denote por μ a la media de la distribución y sea σ^2 su varianza. Lleve a cabo las indicaciones de los siguientes incisos para:

$$n = 20, 40, 60, 80, 100.$$

$$N = 50, 100.$$

Esta es una comprobación del teorema central del límite.

- 1) Genere n valores independientes al azar x_1, \dots, x_n y calcule el promedio

$$s_n = \frac{1}{n} \sum_{i=1}^n x_i.$$

- 2) Repita N veces el inciso anterior calculando los promedios centrados

$$\underbrace{\frac{s_n - \mu}{\sigma/\sqrt{n}}, \dots, \frac{s_n - \mu}{\sigma/\sqrt{n}}}_N.$$

- 3) Elabore un histograma de estos N valores trazando, en la misma gráfica, la función de densidad $N(0, 1)$. Utilice un tamaño adecuado para la base de los rectángulos.
- 4) Elabore una gráfica de la frecuencia relativa acumulada (función de distribución empírica) de los N valores uniendo los puntos con una línea recta y en la misma gráfica dibuje la función de distribución $N(0, 1)$.

- b) Haga lo mismo que en el inciso anterior, ahora con una distribución continua de su preferencia.

525. Sea X con distribución $\text{bin}(n, p)$ y sea Z con distribución normal estándar. Demuestre que cuando $n \rightarrow \infty$,

$$\frac{X - np}{\sqrt{np(1-p)}} \xrightarrow{d} Z.$$

526. Sea X con distribución $\chi^2(n)$ y sea Z con distribución normal estándar. Demuestre que cuando $n \rightarrow \infty$,

$$\frac{X - n}{\sqrt{2n}} \xrightarrow{d} Z.$$

527. Sean X_1, \dots, X_n v.a.s independientes, cada una de ellas con distribución Poisson(λ). Encuentre una aproximación para las siguientes probabilidades, en términos de la función de distribución $\Phi(x)$ de la distribución normal estándar.

- a) $P(a < X_1 + \dots + X_n < b)$.
 b) $P(X_1 + \dots + X_n \geq n\lambda)$.

528. Sea A un evento de un experimento aleatorio cuya probabilidad es $p \in (0, 1)$. Suponga que se efectúan n ensayos independientes del experimento y denote por n_A el número de veces que se observa la ocurrencia del evento A . Demuestre que, para n suficientemente grande y para cualquier $\epsilon > 0$,

$$P(p - \epsilon \leq \frac{n_A}{n} \leq p + \epsilon) \approx \Phi\left(\frac{\epsilon\sqrt{n}}{\sqrt{p(1-p)}}\right) - \Phi\left(\frac{-\epsilon\sqrt{n}}{\sqrt{p(1-p)}}\right).$$

529. La probabilidad de que un componente electrónico falle durante ciertas pruebas de control de calidad es 0.05. Use el teorema de De Moivre-Laplace para encontrar una aproximación de la probabilidad de que, al probar 100 componentes, el número de fallas sea

- a) al menos 5.
 b) menor a 5.
 c) entre 5 y 10, inclusive.

530. La probabilidad de un cierto evento en un ensayo de un experimento aleatorio es 0.7. Encuentre una aproximación para la probabilidad de que este evento aparezca en la mayoría de una sucesión de 200 ensayos independientes.

531. La probabilidad de ocurrencia de un cierto evento en un ensayo es 0.3. Encuentre una aproximación para la probabilidad de que la frecuencia relativa de este evento en 100 ensayos independientes

- a)* se encuentre entre 0.2 y 0.5, inclusive.
 - b)* sea por lo menos 0.4 .
 - c)* sea, a lo sumo, igual a 0.35 .
532. La probabilidad de que un condensador falle durante el periodo de garantía es 0.1 . Encuentre el valor exacto y una aproximación usando el teorema central del límite, de la probabilidad de que, al revisar el funcionamiento de 50 condensadores durante su periodo de garantía, fallen
- a)* entre 3 y 6 condensadores, inclusive.
 - b)* por lo menos 7 condensadores.
 - c)* hasta 2 condensadores.

Jacobo Bernoulli

J. Bernoulli (Suiza, 1654–1705) fue uno de los matemáticos más sobresalientes de la familia Bernoulli. Sus padres tuvieron una posición económica cómoda, siendo Jacobo Bernoulli el primero de los hijos. Por instrucciones de sus padres, estudió filosofía y teología. Se graduó de la Universidad de Basilea en Suiza con el grado de maestro en filosofía en 1671 y obtuvo también una *licenciante* en teología en 1676. Durante sus estudios universitarios y contrario al deseo de sus padres,



J. Bernoulli

Jacobo Bernoulli también estudió matemáticas y astronomía. Con el paso del tiempo, fue claro que el verdadero amor de Jacobo Bernoulli no era la teología ni la filosofía, sino las matemáticas y la física teórica, pues en estas disciplinas impartió clases y desarrolló trabajos científicos de primer nivel. Cronológicamente, Jacobo Bernoulli fue el primer matemático de la familia Bernoulli y posiblemente haya sido de cierta influencia para que algunos de los siguientes miembros de la familia decidieran seguir una carrera científica. Después de graduarse de la Universidad de Basilea, se trasladó a Ginebra, en donde trabajó como tutor. En los años siguientes viajó a Francia, Holanda e Inglaterra, en donde conoció y mantuvo comunicación con renombrados matemáticos y científicos de estos países. En 1683 regresó a Suiza y empezó a trabajar en la Universidad de Basilea, habiendo ya empezado a publicar sus trabajos un año antes. En dicha universidad fue contratado como profesor de matemáticas en 1687. Produjo trabajos de mucha trascendencia en las áreas del cálculo infinitesimal, el álgebra, la teoría de series, el cálculo de variaciones, la mecánica y, particularmente, la teoría de la probabilidad. En 1713, ocho años después de la muerte de Jacobo Bernoulli, se publicó una de sus obras más preciadas: *Ars Conjectandi* (El arte de conjeturar). A este trabajo de Bernoulli se le considera como una obra fundacional del cálculo combinatorio y la teoría de la probabilidad. En este trabajo aparece por primera vez una demostración rigurosa de la ley de los grandes números en el caso cuando las variables aleatorias tienen distribución Bernoulli. A este importante resultado se le conoce ahora justamente como teorema de

Bernoulli. Posiblemente uno de los sucesos más relevantes en la vida de Jacobo Bernoulli fue la penosa rivalidad que mantuvo con su hermano menor, Johann Bernoulli (1667-1748). Era el año de 1687 cuando Johann le pidió a su hermano Jacobo que le enseñara matemáticas. Así, los dos hermanos empezaron a estudiar juntos el difícil cálculo diferencial e integral de Leibnitz y otros trabajos relacionados. Ambos empezaron a resolver problemas importantes en el área, pero eventualmente surgió entre ellos la rivalidad por buscar cada uno un mayor reconocimiento que el otro. Jacobo, como hermano mayor y mentor, sentía que Johann debía sus éxitos a él. Johann, por su parte, y posiblemente mejor matemático que Jacobo, sentía que sus méritos eran propios. Esta rivalidad persistió hasta la muerte de Jacobo Bernoulli, acaecida el 16 de agosto de 1705 a la temprana edad de 50 años. El puesto de profesor de matemáticas que Jacobo tenía en la Universidad de Basilea lo ocupó su hermano Johann.

Jacobo Bernoulli tuvo fascinación por la espiral logarítmica. Ésta es una curva que aparece frecuentemente en la naturaleza y está dada por la ecuación en coordenadas polares $\theta = \log_b(r/a)$. Jacobo Bernoulli pensaba que sus propiedades eran casi mágicas y que representaba un símbolo de permanencia eterna y de constante restauración exacta al ser perfecto. Por ello es que dejó instrucciones para que

en su lápida fuera grabada la inscripción en latín *Eadem Mutata Resurgo*, que significa “Resurgiré nuevamente aunque cambiado”. En honor a la enorme contribución a la ciencia por parte de la extensa familia Bernoulli, y a sugerencia del matemático y estadístico Jerzy Neyman, la agrupación internacional más importante en probabilidad y estadística lleva el nombre de *Bernoulli Society* y en su logo aparece la espiral logarítmica y la frase del epitafio de Jacobo Bernoulli.



Fuente: Archivo MacTutor, Universidad de St. Andrews [24].

Apéndice A

Fórmulas varias

Notación

\mathbb{N}	Conjunto de números naturales $1, 2, 3, \dots$
\mathbb{Z}	Conjunto de números enteros $0, \pm 1, \pm 2, \pm 3, \dots$
\mathbb{Q}	Conjunto de números racionales a/b en donde $a, b \in \mathbb{Z}$ con $b \neq 0$.
\mathbb{R}	Conjunto de números reales.
x^+	$\text{máx}\{x, 0\}$.
x^-	$\text{mín}\{x, 0\}$.
$f(x^+)$	Límite por la derecha de la función f en el punto x .
$f(x^-)$	Límite por la izquierda de la función f en el punto x .
$x \mapsto f(x)$	Función $f(x)$.
$:=$	Se define como.

El alfabeto griego

A α	alfa	I ι	iota	P ρ, ϱ	ro
B β	beta	K κ	kapa	$\Sigma \sigma, \varsigma$	sigma
$\Gamma \gamma$	gamma	$\Lambda \lambda$	lambda	T τ	tau
$\Delta \delta$	delta	M μ	mu	$\Upsilon \upsilon$	upsilon
E ϵ, ε	epsilon	N ν	nu	$\Phi \phi, \varphi$	fi
Z ζ	zeta	$\Xi \xi$	xi	X χ	ji
H η	eta	O \omicron	omicron	$\Psi \psi$	psi
$\Theta \theta, \vartheta$	teta	$\Pi \pi$	pi	$\Omega \omega$	omega

Exponentes

- a) $x^1 = x$.
- b) $x^0 = 1, \quad x \neq 0$.
- c) $x^{-1} = \frac{1}{x}, \quad x \neq 0$.
- d) $x^n x^m = x^{n+m}$.
- e) $\frac{x^n}{x^m} = x^{n-m}$.
- f) $(x^n)^m = x^{nm}$.
- g) $(xy)^n = x^n y^n$.
- h) $\left(\frac{x}{y}\right)^n = \frac{x^n}{y^n}$.
- i) $x^{-n} = \frac{1}{x^n}, \quad x \neq 0$.
- j) $x^{m/n} = \sqrt[n]{x^m}$.

Logaritmos

- a) $\log ab = \log a + \log b.$
- b) $\log \frac{a}{b} = \log a - \log b.$
- c) $\log a^n = n \log a.$
- d) $\log \sqrt[n]{a} = \frac{1}{n} \log a.$
- e) $\log 1 = 0.$
- f) $\log_a a = 1.$

Identidades trigonométricas

- a) $\operatorname{sen}^2 x + \operatorname{cos}^2 x = 1.$
- b) $\operatorname{sen}(x \pm y) = \operatorname{sen} x \operatorname{cos} y \pm \operatorname{cos} x \operatorname{sen} y.$
- c) $\operatorname{cos}(x \pm y) = \operatorname{cos} x \operatorname{cos} y \mp \operatorname{sen} x \operatorname{sen} y.$
- d) $\operatorname{cos}(\operatorname{arc} \operatorname{sen} x) = \operatorname{sen}(\operatorname{arc} \operatorname{cos} x) = \sqrt{1 - x^2}$ si $-1 \leq x \leq 1.$

Fórmulas para sumas

- a) $\sum_{k=m}^n x_k = x_m + x_{m+1} + \cdots + x_n, \quad m \leq n.$
- b) $\sum_{k=1}^n c = nc, \quad c \text{ constante.}$
- c) $\sum_{k=1}^n k = \frac{n(n+1)}{2}.$
- d) $\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}.$

$$e) \quad \sum_{k=1}^n k^3 = \left[\frac{n(n+1)(2n+1)}{2} \right]^2.$$

$$f) \quad \sum_{k=m}^n a^k = \frac{a^m - a^{n+1}}{1-a}, \quad a \neq 1.$$

$$g) \quad \sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x, \quad x \in \mathbb{R}.$$

$$h) \quad \sum_{k=0}^n \binom{n}{k} a^k b^{n-k} = (a+b)^n, \quad a, b \in \mathbb{R}, n \in \mathbb{N}.$$

$$i) \quad \sum_{k=1}^{\infty} \frac{1}{k} \text{ es divergente.}$$

$$j) \quad \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} = \ln 2.$$

$$k) \quad \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6} \quad (\text{Fórmula de Euler}).$$

$$l) \quad \sum_{x=0}^{\infty} \binom{a}{x} t^x = (1+t)^a, \quad |t| < 1, a \in \mathbb{R}.$$

Fórmulas de derivación

$$a) \quad \frac{d}{dx} c = 0, \quad c \text{ constante.}$$

$$b) \quad \frac{d}{dx} x = 1.$$

$$c) \quad \frac{d}{dx} x^n = nx^{n-1}.$$

$$d) \quad \frac{d}{dx} e^x = e^x.$$

- e) $\frac{d}{dx} \ln x = \frac{1}{x}$.
- f) $\frac{d}{dx} \operatorname{sen} x = \cos x$.
- g) $\frac{d}{dx} \cos x = -\operatorname{sen} x$.
- h) $\frac{d}{dx} \tan x = \sec^2 x$.
- i) $\frac{d}{dx} \operatorname{arc} \operatorname{sen} x = \frac{1}{\sqrt{1-x^2}}$.
- j) $\frac{d}{dx} \operatorname{arc} \cos x = -\frac{1}{\sqrt{1-x^2}}$.
- k) $\frac{d}{dx} [f(x) \pm g(x)] = f'(x) \pm g'(x)$.
- l) $\frac{d}{dx} [f(x)g(x)] = f(x)g'(x) + f'(x)g(x)$.
- m) $\frac{d}{dx} \frac{f(x)}{g(x)} = \frac{g(x)f'(x) - f(x)g'(x)}{g^2(x)}$.
- n) $\frac{d}{dx} f(g(x)) = f'(g(x))g'(x)$ (Regla de la cadena).

A.1. Fórmulas de integración

- a) $\int df(x) = \int f'(x) dx = f(x) + c$.
- b) $\int c dx = c \int dx$, c constante.
- c) $\int x^n dx = \frac{x^{n+1}}{n+1} + c$, $n \neq -1$.
- d) $\int \frac{dx}{x} = \ln x + c$.

$$\text{e)} \quad \int e^{ax} dx = \frac{1}{a} e^{ax} + c.$$

$$\text{f)} \quad \int \ln x dx = x \ln x - x + c.$$

$$\text{g)} \quad \int \text{sen } x dx = -\cos x + c.$$

$$\text{h)} \quad \int \cos x dx = \text{sen } x + c.$$

$$\text{i)} \quad \int u dv = uv - \int v du \quad (\text{Integración por partes}).$$

El lema de Abel

Sea a_0, a_1, \dots una sucesión de números reales o complejos tal que $\sum_{n=0}^{\infty} a_n$ es convergente. Entonces la función real $G(t) = \sum_{n=0}^{\infty} a_n t^n$ es continua por la izquierda en $t = 1$, es decir,

$$\lim_{t \nearrow 1} G(t) = \sum_{n=0}^{\infty} a_n.$$

Fórmula de Stirling

Para n grande,

$$n! \approx \sqrt{2\pi} n^{n+1/2} e^{-n}.$$

n	$n!$	Stirling
1	1	0.92
2	2	1.91
3	6	5.83
4	24	23.50
5	120	118.01
6	720	710.07
7	5040	4980.39
8	40320	39902.39
...

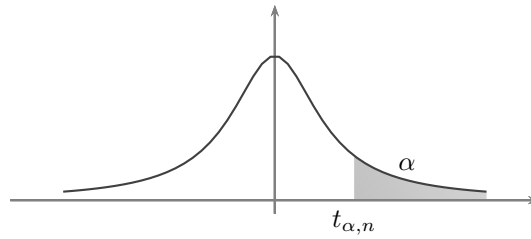
Notación o -pequeña

Se dice que una función $f(x)$, definida en un intervalo no trivial alrededor del cero, es o -pequeña de x cuando $x \rightarrow 0$ si

$$\lim_{x \rightarrow 0} \frac{f(x)}{x} = 0.$$

Esto significa que la función $f(x)$ tiende a cero cuando $x \rightarrow 0$ más rápidamente de lo que lo hace $x \rightarrow 0$. Las funciones $f(x) = x^k$ con $k \geq 2$ son ejemplos de funciones $o(x)$ cuando $x \rightarrow 0$, y se escribe $f(x) = o(x)$ cuando $x \rightarrow 0$.

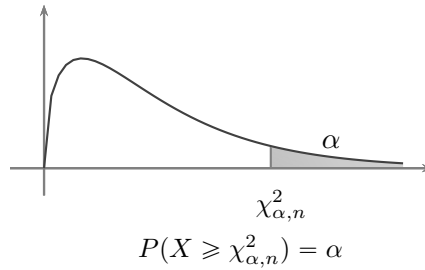
Tabla de la distribución $t(n)$



$$P(X \geq t_{\alpha,n}) = \alpha$$

$n \setminus \alpha$	0.005	0.01	0.025	0.05	0.1
1	63.657	31.821	12.706	6.314	3.078
2	9.925	6.965	4.303	2.920	1.886
3	5.841	4.541	3.182	2.353	1.638
4	4.604	3.474	2.776	2.132	1.533
5	4.032	3.365	2.571	2.015	1.476
6	3.707	3.143	2.447	1.943	1.440
7	3.499	2.998	2.365	1.895	1.415
8	3.355	2.896	2.306	1.860	1.397
9	3.250	2.821	2.262	1.833	1.383
10	3.169	2.764	2.228	1.812	1.372
11	3.106	2.718	2.201	1.796	1.363
12	3.055	2.681	2.179	1.782	1.356
13	3.012	2.650	2.160	1.771	1.350
14	2.977	2.624	2.145	1.761	1.345
15	2.947	2.602	2.131	1.753	1.341
16	2.921	2.583	2.120	1.746	1.337
17	2.898	2.567	2.110	1.740	1.333
18	2.878	2.552	2.101	1.734	1.330
19	2.861	2.539	2.093	1.729	1.328
20	2.845	2.528	2.086	1.725	1.325
21	2.831	2.518	2.080	1.721	1.323
22	2.819	2.508	2.074	1.717	1.321
23	2.807	2.500	2.069	1.714	1.319
24	2.797	2.492	2.064	1.711	1.318
25	2.787	2.485	2.060	1.708	1.316
26	2.779	2.479	2.056	1.706	1.315
27	2.771	2.473	2.052	1.703	1.314
28	2.763	2.467	2.048	1.701	1.313
29	2.756	2.462	2.045	1.699	1.311
∞	2.576	2.326	1.960	1.645	1.282

Tabla de la distribución $\chi^2(n)$



$n \setminus \alpha$	0.995	0.990	0.975	0.950	0.900	0.100	0.050	0.025	0.010	0.005
1	0.0	0.0	0.0	0.0	0.02	2.71	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.10	0.21	4.61	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	0.58	6.25	7.81	9.35	11.34	12.84
4	0.21	0.30	0.48	0.71	1.06	7.78	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	1.61	9.24	11.07	12.83	15.09	16.75
6	0.68	0.87	1.24	1.64	2.20	10.65	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.27	7.26	8.55	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.87	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	34.28	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	36.74	40.11	43.19	46.96	49.65
28	12.46	13.57	15.31	16.93	18.94	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	63.17	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	85.53	90.53	95.02	100.42	104.22
80	51.17	53.54	57.15	60.39	64.28	96.58	101.88	106.63	112.33	116.32
90	59.20	61.75	65.65	69.13	73.29	107.57	113.14	118.14	124.12	128.30
100	67.33	70.06	74.22	77.93	82.36	118.50	124.34	129.56	135.81	140.17

Apéndice B

Introducción a R

El paquete estadístico R es un paquete computacional muy poderoso y muy completo. Es de distribución libre y está disponible para MacOSX, Windows y Linux. Se puede obtener y encontrar mayor información en la página

<http://www.r-project.org>

En esta sección se presenta una breve introducción a algunos comandos y estructuras básicas de R. Para conocer de manera más completa los comandos a los que se hará referencia en este apéndice, incluyendo sus muchas variaciones y opciones, pueden consultarse libros dedicados a R, como [1] o [4], o bien los múltiples manuales y páginas web disponibles en internet.

Al dar inicio al paquete R se presenta una ventana como la que aparece en la Figura B.1. Los comandos se ingresan donde aparece el caracter `>`, como si se tratara de una calculadora, por ejemplo,

```
> 1+1
[1] 2
```

Para aquellas personas con pocos conocimientos en computación, es posible tener una primera experiencia con R sin necesidad de instalar el paquete, visitando sitios web como

<http://www.compileonline.com>

en donde la interacción es a través de un navegador de internet. Este sitio

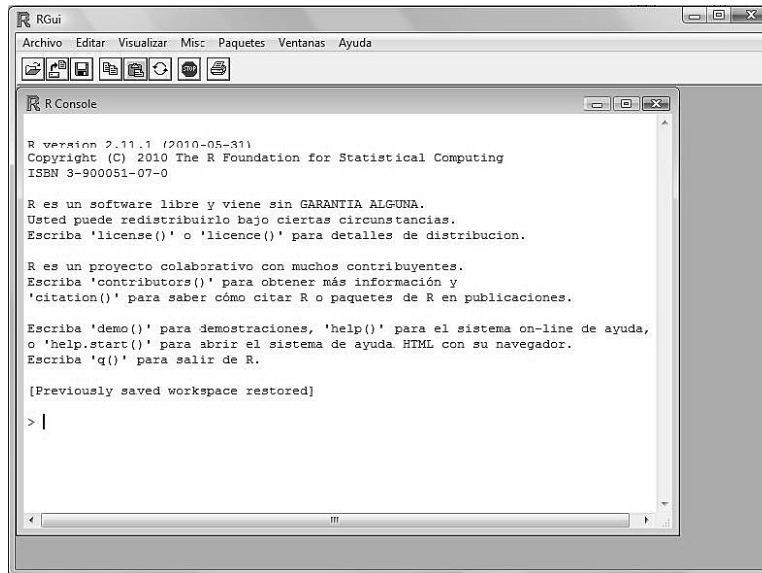


Figura B.1

permite compilar y ejecutar en línea varios lenguajes de programación. En el caso de R, se debe seleccionar la opción **R programming** dentro de la tabla **Programming languages**. Esta forma de trabajar con R es limitada y no enteramente confiable, pero es muy útil para iniciarse en la programación en R y para experimentar y hacer programas sencillos. En la Figura B.2 se muestra la ventana de un navegador en el sitio <http://www.compileonline.com> presentando el programa en R desarrollado en la página 358 para ilustrar la ley de los grandes números.

Para un uso más serio de R, es recomendable la instalación tanto del paquete como de alguna interfase gráfica (GUI *Graphic Unit Interface*) para una experiencia más amigable. Una de las varias opciones de interfase gráfica para R es RStudio, cuyo sitio web es

<http://www.rstudio.com>

En la Figura B.3 se muestra esta interfase. En la ventana superior izquierda se pueden ingresar series de comandos, los cuales se ejecutan haciendo click

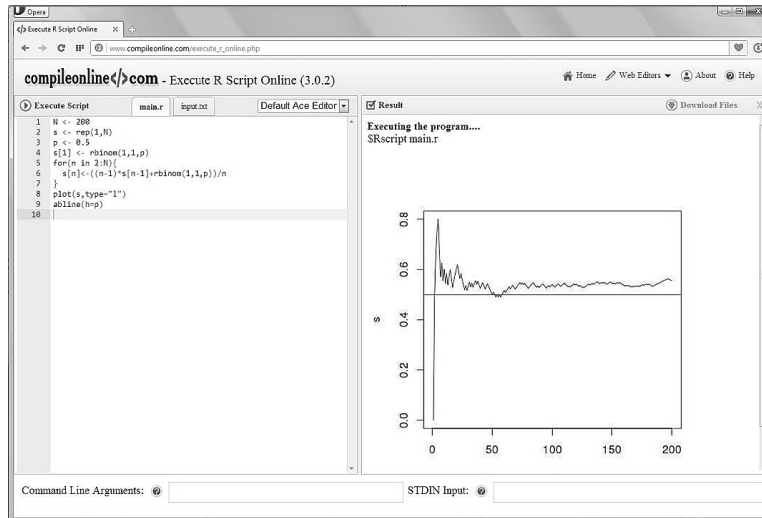


Figura B.2

en el ícono superior con el nombre **Source**. Los resultados se muestran en las dos ventanas inferiores.

Comentarios e impresión en R

Una de las prácticas más recomendables y necesarias al elaborar programas de cómputo es la de insertar comentarios para explicar lo que se lleva a cabo en una instrucción o conjunto de instrucciones. En R, los caracteres que aparezcan a la derecha del símbolo `#` sobre la misma línea, son tomados como comentarios. Por otro lado, se puede usar el comando `cat` para imprimir un comentario o el valor de una variable. Por ejemplo, el código

```

# Esta línea es un comentario
x <- 5    # Se asigna el valor 5 a la variable x
cat ("El valor de la variable x es", x, "\n")
```

produce como resultado el mensaje `El valor de la variable x es 5.`

Operaciones aritméticas

Algunas de las siguientes funciones se pueden aplicar para números, vectores y matrices.

<code>x+y</code>	suma
<code>x-y</code>	resta
<code>x*y</code>	producto
<code>x/y</code>	cociente
<code>x^y</code> o <code>x**y</code>	potencia
<code>x%%y</code>	x módulo y
<code>x%/%y</code>	división entera (parte entera del cociente)

Funciones numéricas

<code>abs(x)</code>	valor absoluto
<code>sqrt(x)</code>	raíz cuadrada
<code>ceiling(x)</code>	$\lceil x \rceil$ entero mayor o igual a x más pequeño
<code>floor(x)</code>	$\lfloor x \rfloor$ entero menor o igual a x más grande
<code>trunc(x)</code>	parte entera de x
<code>sin(x)</code>	seno
<code>cos(x)</code>	coseno
<code>tan(x)</code>	tangente
<code>log(x)</code>	logaritmo en base e
<code>log(x, base=a)</code>	logaritmo en base a
<code>exp(x)</code>	exponencial
<code>factorial(x)</code>	factorial
<code>max(x,y)</code>	máximo
<code>min(x,y)</code>	mínimo

Constantes

<code>pi</code>	π	(3.141593...)
<code>exp(1)</code>	e	(2.718282...)

Operadores lógicos

<code>x<y</code>	menor que
<code>x>y</code>	mayor que
<code>x<=y</code>	menor o igual a
<code>x>=y</code>	mayor o igual a
<code>x==y</code>	igual a
<code>x!=y</code>	distinto a
<code>x&y</code>	y
<code>x y</code>	o
<code>!x</code>	negación de x

Pregunta if

La estructura general es

```
if (condición) { expresión }
```

Por ejemplo,

```
if (x>=0) { cat("El valor de x es mayor o igual a cero.\n") }
```

Pregunta if-else

La estructura general es

```
if (condición) { expresión1 } else {expresión2 }
```

Por ejemplo,

```
if (x>=0) { cat("El valor de x es mayor o igual a cero.\n") }  
else      { cat("El valor de x es negativo.\n") }
```

Ciclo for

La estructura general es

```
for (variable in secuencia) { expresión }
```

Por ejemplo, el código

```
for (x in 1:3){  
  cat("x es", x, "\n") }
```

produce los resultados

```
x es 1
x es 2
x es 3
```

Ciclo while

La estructura general es

```
while (condición) { expresión }
```

Por ejemplo, el código

```
x <- 0
while (x<3) {
  cat("x es", x, "\n")
  x <- x+1 }
cat("x ya es", x, "\n")
```

produce los resultados

```
x es 0
x es 1
x es 2
x ya es 3
```


Apéndice C

Solución a los ejercicios

Esta sección contiene algunas sugerencias de solución a los ejercicios planteados. Para algunos ejercicios es necesario ser más explícito al dar una solución o al justificar una respuesta, considere por tanto que este material contiene simplemente ideas para generar una solución completa, correcta y bien escrita. La mayoría de las gráficas han sido omitidas. Recuerde además que los métodos empleados o sugeridos para llegar a una solución no son necesariamente únicos.

Capítulo 1: Probabilidad elemental

1.1 Experimentos aleatorios

1. Las respuestas no son únicas. Las condiciones del experimento y la información con la que cuenta el observador pueden determinar si el experimento es de un tipo o de otro.

<i>a)</i> A.	<i>d)</i> D.	<i>g)</i> A.	<i>j)</i> D.
<i>b)</i> D.	<i>e)</i> A.	<i>h)</i> D.	<i>k)</i> A.
<i>c)</i> A.	<i>f)</i> A.	<i>i)</i> A.	<i>l)</i> A.

2. Este es un intento: el azar es una medida de la falta de información. Pero, ¿qué es la información?

1.2 Espacio muestral

3. Las respuestas no son únicas. Dependen de lo que le interese medir al observador y de la forma en la que puede o desea medir las variables.

- a) \mathbb{R}^3 .
 - b) \mathbb{N} .
 - c) Conjunto de sucesiones finitas de elementos del conjunto $\{1, 2, 3, 4, 5, 6\}$ tal que el último elemento es “6”, junto con la sucesión infinita en donde nunca aparece “6”.
 - d) Si \mathbb{D} es el conjunto de días de un año, entonces un espacio muestral para este experimento puede ser el conjunto producto \mathbb{D}^n .
 - e) Conjunto de vectores de la forma (x_1, x_2, \dots, x_n) , en donde cada coordenada es un entero mayor o igual a cero y la suma de todas ellas es r .
 - f) $(0, \infty)$.
 - g) $\{0, 1, 2, \dots\}$.
4. a) $\Omega = \{(a, b, c) : a, b, c = \text{“cara”}, \text{“cruz”}\}$.
- b) $\Omega = \{(a, 3 - a) : a = 0, 1, 2, 3\}$.
5. a) $A = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$, $\#A = 6$.
- b) $B = \{(1, 2), (1, 4), (1, 6), (3, 2), (3, 4), (3, 6), (5, 2), (5, 4), (5, 6), \dots\}$,
 $\#B = 18$.
- c) $C = \{(1, 1), (1, 2), (2, 1), (2, 2), (2, 3), (3, 2), (3, 3), (3, 4), \dots\}$, $\#C = 16$.
- d) $D = \{(1, 5), (1, 6), (2, 6), (5, 1), (6, 1), (6, 2)\}$, $\#D = 6$.
- e) $E = A$, $\#E = 6$.
- f) $F = \{(1, 1), (1, 3), (1, 5), (2, 2), (2, 4), (2, 6), \dots\}$, $\#F = 18$.
- g) C y D son disjuntos, por lo tanto G consta de la lista completa de C y de D , $\#G = 22$.

1.3 Operaciones con conjuntos

6. Para cada identidad, tome una de las expresiones y, mediante el uso de las definiciones de las operaciones involucradas o sus propiedades, reproduzca la otra expresión. Intente seguir una escritura estructurada como en la siguiente

solución para el inciso (b).

$$\begin{aligned}
 A^c - B^c &= A^c \cap (B^c)^c \\
 &= A^c \cap B \\
 &= B \cap A^c \\
 &= B - A.
 \end{aligned}$$

7. Únicamente se provee la validez o invalidez de las afirmaciones.

- | | | |
|---------------|---------------|---------------|
| a) Verdadero. | c) Verdadero. | e) Verdadero. |
| b) Verdadero. | d) Verdadero. | f) Falso. |

8. Utilice la definición de diferencia simétrica y un diagrama de Venn, en algunos casos, para obtener alguna idea de la validez o invalidez de las afirmaciones. En caso de que la afirmación sea falsa es necesario dar un ejemplo particular. Si la afirmación es válida, intente escribir una demostración siguiendo una escritura estructurada como la siguiente para el inciso (f).

$$\begin{aligned}
 A \Delta A^c &= (A - A^c) \cup (A^c - A) \\
 &= (A \cap (A^c)^c) \cup (A^c \cap A^c) \\
 &= (A \cap A) \cup (A^c \cap A^c) \\
 &= A \cup A^c \\
 &= \Omega.
 \end{aligned}$$

9. En cada caso es útil apoyarse en un diagrama de Venn. Otras expresiones pueden ser equivalentes.

- $(A \cup B) \cap C^c$.
- $A^c \cap B^c \cap C^c$.
- $(A \cap B^c \cap C^c) \cup (A^c \cap B \cap C^c) \cup (A^c \cap B^c \cap C)$.
- $(A \cap B \cap C^c) \cup (A \cap B^c \cap C) \cup (A^c \cap B \cap C)$.
- $A \cup B \cup C$.
- $(A \cap B \cap C)^c$.

10. De 100 personas, 66.6 son mujeres y 33.3 son hombres. El porcentaje de personas que son mayores a 50 años es $(0.42)(66.6) + (0.38)(33.3) = 40.626$

11. Toda función indicadora separa su dominio en dos regiones disjuntas. Verifique estas identidades en cada una de estas partes.

12. a) $(D_1 \cap D_2^c \cap D_3^c \cap D_4^c) \cup (D_1^c \cap D_2 \cap D_3^c \cap D_4^c) \cup \dots$
 b) $(D_1 \cap D_2 \cap D_3^c \cap D_4^c) \cup (D_1 \cap D_2^c \cap D_3 \cap D_4^c) \cup \dots$
 c) $(D_1 \cap D_2 \cap D_3^c \cap D_4^c) \cup (D_1^c \cap D_2 \cap D_3 \cap D_4^c) \cup \dots$
 d) $(D_1 \cap D_2^c \cap D_3 \cap D_4^c) \cup (D_1^c \cap D_2 \cap D_3^c \cap D_4)$
 $\cup (D_1 \cap D_2^c \cap D_3^c \cap D_4^c) \cup \dots \cup (D_1^c \cap D_2^c \cap D_3^c \cap D_4)$
 $\cup (D_1^c \cap D_2^c \cap D_3^c \cap D_4^c)$.
 e) $(D_1 \cap D_2 \cap D_3^c \cap D_4^c) \cup \dots \cup (D_1^c \cap D_2^c \cap D_3 \cap D_4)$
 $\cup (D_1 \cap D_2 \cap D_3 \cap D_4^c) \cup \dots \cup (D_1^c \cap D_2 \cap D_3 \cap D_4)$
 $\cup (D_1 \cap D_2 \cap D_3 \cap D_4)$.
13. Dibuje en un plano cartesiano el cuadrado unitario $(0, 1) \times (0, 1)$. Dentro de este cuadrado, identifique el conjunto de puntos que pertenecen a los eventos A , B , C y D .
14. La persona escogida
- es hombre y tiene empleo.
 - es mujer y no tiene empleo.
 - es hombre o tiene empleo.
 - es hombre y no tiene empleo.
 - es hombre, tiene empleo y está casado.
 - es hombre, está casado y no tiene empleo.
 - es hombre, no tiene empleo y no está casado.
 - no está casada y tiene empleo.
15. El número escogido
- es par y termina en 0.
 - termina en 5 ó en 0.
 - no es ninguno.
 - es par pero no termina en 0.
16. a) $F = A \cap ((B_1 \cap B_2) \cup B_3) \cap C$.
 b) $F^c = A^c \cup ((B_1^c \cup B_2^c) \cap B_3^c) \cup C^c$.
17. El conjunto A corresponde a un círculo de radio 1 y el conjunto B es el interior de un rombo centrado en el origen. A partir de la identificación de estos conjuntos pueden encontrarse los otros conjuntos observando que $B \subset A$.

$$35. p = \begin{cases} t(2T - t)/T^2 & \text{si } 0 \leq t < T, \\ 1 & \text{si } t \geq T. \end{cases}$$

$$36. p = \begin{cases} \frac{2\ell}{L} & \text{si } 2\ell < L, \\ 1 & \text{si } 2\ell \geq L. \end{cases}$$

$$37. p = 3/4.$$

$$38. p = \begin{cases} \frac{L^2 - 2(a^2 + (L - b)^2)}{L^2} & \text{si } a < L/2 < b, \\ \frac{2(b + a)(b - a)}{L^2} & \text{si } b \leq L/2, \\ \frac{2(2L - a - b)(b - a)}{L^2} & \text{si } a \geq L/2. \end{cases}$$

$$39. p = 1/6.$$

$$40. p = 1/2.$$

41. En un plano cartesiano dibuje la curva de nivel $z = z_0$, por ejemplo, para $z_0 = \ell/2$. La respuesta es $1/2$.

$$42. a) p = (1/2) \ln(e/2). \quad b) p = 1/8.$$

$$43. p = 1/6.$$

1.6 Probabilidad frecuentista

44. Establezca primero cada afirmación para el cociente n_A/n y después tome el límite cuando $n \rightarrow \infty$.

45. A través de la instrucción `mean` se puede hacer que `R` calcule también el promedio de los resultados. Por ejemplo, el siguiente código produce lo requerido.

```
# Ejecución de 100 lanzamientos y asignación de resultados al vector x:
x = sample(0 : 1, 100, replace = TRUE)
x          # Muestra de resultados
mean(x)    # Cálculo del promedio de los resultados
```

46. Siga los lineamientos como en la solución del ejercicio anterior.

1.9 Sigmas álgebras

60. Exprese cada uno de estos eventos como el resultado de aplicar algunas de las operaciones de unión, intersección o complemento de los eventos A y B .
61. Al dibujar un diagram de Venn general para estos dos eventos, uno puede identificar cuatro regiones distintas y ajenas. Cada evento de la σ -álgebra en cuestión es una unión de algunas de estas cuatro regiones.
62. Tomando como últimos elementos al conjunto vacío, puede comprobarse que la condición de que la unión numerable de eventos debe estar en la σ -álgebra se reduce a la misma condición para uniones finitas. Como un ejemplo de una álgebra que no es σ -álgebra, considere $\Omega = \mathbb{N}$, la colección $\mathcal{F} = \{A \subseteq \mathbb{N} : A \text{ es finito o } A^c \text{ es finito}\}$ y la sucesión $A_n = \{2n\}$ para $n = 1, 2, \dots$. Compruebe que \mathcal{F} es una álgebra pero no es una σ -álgebra pues no contiene a $\bigcup_{n=1}^{\infty} A_n = \{2, 4, 6, \dots\}$.
63. Como el espacio muestral es finito, al considerar una sucesión infinita de eventos, sólo un número finito de ellos es distinto. Así, cualquier unión infinita de eventos se reduce a una unión finita de ellos.
64. Compruebe que la colección definida abajo cumple las tres condiciones de la definición para que sea σ -álgebra.

$$\mathcal{F}_1 \cap \mathcal{F}_2 = \{A \subseteq \Omega : A \in \mathcal{F}_1 \text{ y } A \in \mathcal{F}_2\}.$$

Para el tercer inciso, considere por ejemplo $\mathcal{F}_1 = \{\emptyset, \Omega, A, A^c\}$ y $\mathcal{F}_2 = \{\emptyset, \Omega, B, B^c\}$.

65. Todo elemento de esta σ -álgebra es la unión de algunos de estos cuatro eventos.

1.11 Espacios de probabilidad

66. En algunos de estos cálculos conviene observar que $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$.
- a) $P(\{1, \dots, n\}) = n/(n+1)$. c) $P(\{1, 3, 5, \dots\}) = \ln 2$.
- b) $P(\{n, n+1, \dots\}) = 1/n$. d) $P(\{2, 4, 6, \dots\}) = 1 - \ln 2$.
67. a) $P((0, 1/2)) = 1/4$. c) $P((1/3, 2/3)) = 1/3$.
- b) $P((-1, 1)) = 1$. d) $P((1/2, \infty)) = 3/4$.
68. a) $\Omega = \{1, \dots, 6\}$, $\mathcal{F} = 2^\Omega$, $P(i) = 1/6$.
- b) $\Omega = \{(i, j) : i, j = 0, 1, 2, \dots\}$, $\mathcal{F} = 2^\Omega$, $P((i, j)) = (1/2)^{i+j}/4$.

- c) $\Omega = \{1, 2, \dots\}$, $\mathcal{F} = 2^\Omega$, $P(i) = 1/2^i$.
 d) $\Omega = (0, 1)$, $\mathcal{F} = (0, 1) \cap \mathcal{B}(\mathbb{R})$, $P((a, b)) = b - a$ si $(a, b) \subseteq (0, 1)$.
 e) $\Omega = \{(x, y) : x^2 + y^2 \leq 1\}$, $\mathcal{F} = \Omega \cap \mathcal{B}(\mathbb{R}^2)$, $P(A) = \text{“Área de } A\text{”}$.

1.12 Análisis combinatorio

69. a) $p = 1/(3 \cdot 6^5)$. c) $p = 1/2^6$.
 b) $p = 5/(3^2 6^2)$. d) $p = 1/2^5$.
70. Hay 4 números positivos de un dígito que son divisibles por 2, $(9)(5) = 45$ de dos dígitos, $(9)(10)(5) = 450$ de tres dígitos, $(9)(10)(10)(5) = 4500$ de cuatro dígitos, y $(9)(10)(10)(10)(5) = 45000$ de cinco dígitos. El total es 49,999. De éstos, 5,555 empiezan con el dígito 1.
71.
$$p = \begin{cases} 1 - \frac{365!}{(365-n)!(365)^n} & \text{si } 2 \leq n \leq 365, \\ 1 & \text{si } n \geq 366. \end{cases}$$
72. $p = n!(L/n)^n$.
73. $p = 5/54$.
74. El evento $A \cup B \cup C$ genera 7 subregiones no vacías. Cada una de estas subregiones debe ser parte de alguno de los tres componentes de la descomposición. Así, puede haber 3^7 posibles descomposiciones, suponiendo relevante el orden de los componentes.
75. $2^n - 1$.
76. $n(n-1)(n-2) = n!/(n-3)!$
77. $(n-1)!$
78. Suponiendo que los automóviles llegan al estacionamiento uno por uno, hay $12!/4!$ formas distintas en que se pueden estacionar. De éstas, 9 son favorables al evento en cuestión. Por lo tanto, la respuesta es $p = 9 \cdot 4!/12!$.
79. $p = \frac{n(n-1)^{r-1}}{n^r}$.
80. a) $p = 5/36$. b) $p = 3/21$.
81. a) m^n . b) $m!/(m-n)!$ c) $\sum \binom{n}{k_1 k_2 \dots k_m}$, en donde la suma se efectúa sobre todos los posibles enteros k_1, \dots, k_m tales que son mayores o iguales a uno y $k_1 + \dots + k_m = n$.

$$82. p = \binom{10}{1} \binom{90}{4} / \binom{100}{5}.$$

83. Observe que $(1+x)^n(1+x)^m = (1+x)^{n+m}$. Ahora iguale el coeficiente del término x^k de cada expresión.

84. Simplifique el lado derecho en ambos casos.

85. Use el método de inducción sobre el valor de m . Cuando $m = 1$, se trata de un identidad evidente y cuando $m = 2$, se trata del teorema del binomio. Para la demostración, considere la identidad

$$(x_1 + \cdots + x_m)^n = ((x_1 + \cdots + x_{m-1}) + x_m)^n.$$

Desarrolle el lado derecho usando el teorema del binomio y la hipótesis de inducción. Se debe definir también la nueva variable $k_m = n - k$.

$$86. \binom{n+m-1}{n}.$$

$$87. \binom{n+m-1}{m}.$$

$$88. n(n+1)/2.$$

$$89. (n-k)(n-k+1).$$

$$90. a) \binom{n}{k}. \quad b) \binom{n+k-1}{k}.$$

$$91. a) \binom{n+k-1}{n}. \quad b) \sum_{m=0}^n \binom{m+k-1}{m}. \quad c) \infty.$$

$$92. \binom{2n-r}{n-r} (1/2)^{n+1} (1/2)^{n-r}.$$

93. La respuesta en ambos casos es $15/6^3$.

94. La probabilidad buscada aparece abajo. Se comprueba que la suma de estas probabilidades es uno.

$$P(X = x) = \begin{cases} (x-1)/36 & \text{si } x = 2, 3, \dots, 7, \\ (13-x)/36 & \text{si } x = 8, 9, \dots, 12, \\ 0 & \text{en otro caso.} \end{cases}$$

$$95. a) \text{ Prob.} = \binom{n}{2r} 2^{2r} / \binom{2n}{2r} \quad \text{si } 2r \leq n, \text{ de lo contrario es cero.}$$

$$b) \text{ Prob.} = \binom{n}{1} \binom{n-1}{2r-2} 2^{2r-2} / \binom{2n}{2r}.$$

$$c) \text{ Prob.} = \binom{n}{r} / \binom{2n}{2r}.$$

$$96. \quad a) 1/n.$$

$$b) ((n-1)/n)^{n-1} (1/n).$$

$$97. \quad \binom{n}{2}.$$

98. En número máximo de regiones es:

$$a) n(n+1)/2 + 1.$$

$$b) n(n-1) + 2.$$

$$c) n(n^2 + 5)/6 + 1.$$

$$d) n(n^2 - 3n + 8)/3.$$

$$99. (n+1)(m+1).$$

$$100. \quad a) (n-1)! n!$$

$$b) n!$$

$$c) \binom{2n+1}{n} (n+1)!$$

$$d) \binom{2n+1}{n} n!$$

$$101. 2^n - 1.$$

1.13 Probabilidad condicional

102. Aplique la definición de probabilidad condicional y use la misma propiedad que en el caso no condicional o el hecho de que cualquier evento C admite la descomposición ajena $(C \cap B) \cup (C \cap B^c)$. Por ejemplo, para el inciso (a),

$$\begin{aligned} P(A|B) &= \frac{P(B \cap A)}{P(B)} \\ &= \frac{P(B) - P(B \cap A^c)}{P(B)} \\ &= 1 - P(A^c|B). \end{aligned}$$

103. Se provee únicamente la validez o invalidez de la afirmación.

- a) Falso. d) Verdadero.
 b) Falso. e) Verdadero.
 c) Verdadero. f) Verdadero.

104. Únicamente se señala la validez o invalidez de cada afirmación y es necesario dar una demostración o un contraejemplo en cada caso.

- a) Falso. Tómese $A = \Omega$ y $0 < P(B) < 1$.
 b) Falso. Tómese $B \subseteq A$ con $P(A - B) > 0$.
 c) Falso. Tómese $B = A^c$ con $P(A) > 0$.
 d) Falso. Tómese $A = B$ con $0 < P(B) < 1$.
 e) Falso. Tómese cualquier $C \subseteq B$.
 f) Falso. Tómese A y B ajenos.
 g) Ambas implicaciones son falsas:
 (\Rightarrow) Tómese A y B ajenos con $P(A) = P(B) > 0$ y $C = B$.
 (\Leftarrow) Tómese $C \subseteq A \subseteq B$ con $P(B - A) > 0$.
 h) (\Rightarrow) Verdadero.
 (\Leftarrow) Falso. $P(C) = 0$ no implica $C = \emptyset$.
 i) Ambas implicaciones son falsas:
 (\Rightarrow) Tómese A y B ajenos y $C = A$.
 (\Leftarrow) Tómese $C = B$ con $P(A) > P(B)$.
 j) (\Rightarrow) Verdadero.
 (\Leftarrow) Falso. Tómese $C = B \subset A = \Omega$.
 k) Ambas implicaciones son ciertas, de hecho son idénticas, pues partiendo de una de ellas y cambiando los papeles de A y B se obtiene la otra. El siguiente análisis muestra que a partir de la primera igualdad se encuentra que $P(A) = P(A|B)$, y de aquí se demuestra la segunda igualdad.

$$\begin{aligned}
 P(A) &= P(A \cap B) + P(A \cap B^c) \\
 &= P(A \cap B^c)(P(B)/P(B^c) + 1) \\
 &= P(A \cap B^c)/P(B^c) \\
 &= P(A|B^c) \\
 &= P(A|B).
 \end{aligned}$$

- l) Falso. Tómese $B = C \neq \emptyset$ con $P(A|B) > 0$.

105. En cada caso proporcione un ejemplo concreto.

- a) Considere $A \cap B = \emptyset$ con $P(A) > 0$.
 b) Considere $A = B^c$ con $0 < P(B) < 1$.
 c) Considere $A = B$ con $0 < P(B) < 1$.
 d) Considere $A = B = \Omega$.
 e) Considere $A = B^c$ con $0 < P(B) < 1$.

106. a) $p = 0.18$. c) $p = 0.08$. e) $p = 18/26$.
 b) $p = 0.42$. d) $p = 0.32$. f) $p = 32/74$.

107. $p = 15/24$.

108. Usando la hipótesis de independencia, el lado derecho es

$$\begin{aligned} & P(B) \frac{P(A \cap B \cap C)}{P(A \cap B)} + P(B^c) \frac{P(A \cap B^c \cap C)}{P(A \cap B^c)} \\ &= \frac{P(A \cap B \cap C)}{P(A)} + \frac{P(A \cap B^c \cap C)}{P(A)} \\ &= \frac{P(A \cap C)}{P(A)} = P(C|A). \end{aligned}$$

109.

$$\begin{aligned} P(A | \bigcup_{i=1}^n B_i) &= \sum_{i=1}^n P(A | B_i) / P(\bigcup_{i=1}^n B_i) \\ &= p \sum_{i=1}^n P(B_i) / P(\bigcup_{i=1}^n B_i) = p. \end{aligned}$$

110. Aplique la definición de probabilidad condicional y simplifique el lado derecho.

111. Claramente $P(R_1) = r/(r+b)$. Suponga $P(R_n) = r/(r+b)$. Para calcular $P(R_{n+1})$, condicione sobre el resultado de la primera extracción. Después de dicha extracción, la configuración de la urna ha cambiado y restan n extracciones adicionales, y allí es donde se usa la hipótesis de inducción convenientemente interpretada.

112. Claramente debe cumplirse que $p_0 = 0$ y $p_N = 1$. En cambio, para $n = 1, 2, \dots, N-1$, condicionando sobre el resultado de la primera apuesta y denotando por la letra A_1 el evento cuando el jugador A gana esa primera apuesta y B_1 cuando B gana la apuesta, tenemos que

$$P(E_n) = P(E_n | A_1) P(A_1) + P(E_n | B_1) P(B_1) = P(E_{n+1})p + P(E_{n-1})q.$$

Esto es, $p_n = p_{n+1}p + p_{n-1}q$. Multiplicando el lado izquierdo por $p + q = 1$ y reacomodando términos se llega a la ecuación en diferencias buscada. Esta ecuación en diferencias puede resolverse directamente mediante algún método, sin embargo, aquí se pide simplemente comprobar que la respuesta presentada efectivamente es solución de la ecuación. Verifique esto en los dos casos: $p = 1/2$ y $p \neq 1/2$.

1.14 Teorema de probabilidad total

113. Sean M_1 y M_2 los eventos cuando la persona escogida es mujer en la primera y segunda selección, respectivamente. De manera similar se definen los eventos H_1 y H_2 , en el caso de que la persona sea hombre. Entonces

$$\begin{aligned} a) \quad P(M_2) &= P(M_2 | M_1) P(M_1) + P(M_2 | H_1) P(H_1) \\ &= \frac{m-1}{m+n-1} \frac{m}{m+n} + \frac{m}{m+n-1} \frac{n}{m+n} = \frac{m}{m+n}. \end{aligned}$$

$$\begin{aligned} b) \quad P(M_1 | M_2) &= P(M_2 | M_1) P(M_1) / P(M_2) \\ &= \frac{m-1}{m+n-1} \frac{m}{m+n} \frac{m+n}{m} = \frac{m-1}{m+n-1}. \end{aligned}$$

$$c) \quad P(H_2) = \frac{n}{m+n}.$$

$$d) \quad P(H_1 | M_2) = \frac{n}{m+n-1}.$$

$$\begin{aligned} e) \quad P(M_1 \cap M_2) + P(H_1 \cap H_2) &= P(M_2 | M_1) P(M_1) + P(H_2 | H_1) P(H_1) \\ &= \frac{(m-1)m + (n-1)n}{(m+n-1)(m+n)}. \end{aligned}$$

$$\begin{aligned} f) \quad P(M_1 \cap H_2) + P(H_1 \cap M_2) &= P(H_2 | M_1) P(M_1) + P(M_2 | H_1) P(H_1) \\ &= \frac{2nm}{(m+n-1)(m+n)}. \end{aligned}$$

114. Sea M_k el evento en donde la k -ésima persona escogida es mujer. Se puede usar el método de inducción sobre k para demostrar que $P(M_k) = m/(m+n)$, siempre y cuando $k \leq m, n$. Para $k = 1$ tenemos, claramente, que $P(M_1) = m/(m+n)$. Supondremos válido este resultado para $k-1$. Para obtener el resultado para k se condiciona sobre el resultado de la primera selección, quedando $k-1$ selecciones por hacer dentro de un grupo de $m+n-1$

personas:

$$\begin{aligned}
 P(M_k) &= P(M_k | M_1)P(M_1) + P(M_k | M_1^c)P(M_1^c) \\
 &= \frac{m-1}{m+n-1} \frac{m}{m+n} + \frac{m}{m+n-1} \frac{n}{m+n} \\
 &= \frac{m}{m+n}.
 \end{aligned}$$

115. Sean los eventos B_1 y R_1 correspondientes a obtener bola blanca o roja de la urna A . Sea R_2 el evento de interés, esto es, obtener bola roja de la urna B . Entonces

$$\begin{aligned}
 P(R_2) &= P(R_2 | B_1)P(B_1) + P(R_2 | R_1)P(R_1) \\
 &= (1/3)(2/6) + (2/3)(4/6) = 5/9.
 \end{aligned}$$

116. Sea A_i el evento de obtener una bola azul de la caja i . Entonces

$$\begin{aligned}
 P(A_3) &= P(A_3 | A_1 \cap A_2)P(A_2 | A_1)P(A_1) \\
 &\quad + P(A_3 | A_1^c \cap A_2)P(A_2 | A_1^c)P(A_1^c) \\
 &\quad + P(A_3 | A_1 \cap A_2^c)P(A_2^c | A_1)P(A_1) \\
 &\quad + P(A_3 | A_1^c \cap A_2^c)P(A_2^c | A_1^c)P(A_1^c) \\
 &= (3/4)(2/3)(1/2) + (2/4)(1/3)(1/2) \\
 &\quad + (2/4)(2/3)(1/2) + (1/4)(2/3)(1/2) \\
 &= 1/2.
 \end{aligned}$$

117. Sean los eventos M_i escoger la moneda i y C obtener cruz al lanzar la moneda. Entonces

$$P(C) = \sum_{i=1}^4 P(C | M_i)P(M_i) = (1/4) \sum_{i=1}^4 (1 - 0.2i) = 1/2.$$

118. Sea N el resultado del primer lanzamiento y sean X_1, X_2, \dots los resultados de los siguientes lanzamientos y cuya suma se denota por S . Entonces

$$\begin{aligned}
 P(N = S) &= P(N = 1, X_1 = 1) + P(N = 2, X_1 = 1, X_2 = 1) + \dots \\
 &\quad \dots + P(N = 6, X_1 = 1, \dots, X_6 = 1) \\
 &= (1/6)^2 + (1/6)^3 + \dots + (1/6)^7 \\
 &= (1 - (1/6)^6)/30.
 \end{aligned}$$

119. Sean E_0 y E_1 los eventos correspondientes a enviar un 0 y un 1, y sean R_0 y R_1 los eventos recibir un 0 y un 1. Entonces

- a) $P(R_0) = P(R_0 | E_0)P(E_0) + P(R_0 | E_1)P(E_1)$
 $= (0.8)(0.45) + (0.1)(0.55) = 0.415.$
- b) $P(R_1) = 1 - P(R_0) = 0.585.$
- c) $P(E_0 \cap R_0) + P(E_1 \cap R_1) = P(R_0 | E_0)P(E_0) + P(R_1 | E_1)P(E_1)$
 $= (0.8)(0.45) + (0.9)(0.55) = 0.855.$
- d) $P(E_0 \cap R_1) + P(E_1 \cap R_0) = 1 - 0.855 = 0.145.$

120. Sea D el resultado del dado y B el evento cuando todas las bolas escogidas son blancas. Entonces

$$P(B) = \sum_{d=1}^4 P(B | D = d)P(D = d) = (1/6) \sum_{d=1}^4 \binom{4}{x} / \binom{10}{x}.$$

121. Sea S el evento cuando el estudiante sabe la respuesta a la pregunta y sea C el evento cuando la respuesta proporcionada es correcta.

- a) Por el teorema de probabilidad total,

$$P(C) = P(C | S)P(S) + P(C | S^c)P(S^c) = (1)(0.6) + (0.25)(0.4) = 0.7.$$

Esto significa que la probabilidad de tener una respuesta correcta se incrementa de 0.6 a 0.7 cuando se adopta la estrategia de escoger una respuesta al azar en el caso de no saber la respuesta.

- b) La pregunta planteada es

$$P(S | C) = P(C | S)P(S) / P(C) = 0.6 / 0.7 = 0.8571.$$

- c) Sea A el evento de acreditar el examen. Por lo anterior, sabemos que $P(C) = 0.7$. Entonces

$$P(A) = \sum_{k=6}^{10} \binom{10}{k} (0.7)^k (0.3)^{10-k} = 0.6331.$$

122. Sean los eventos B_i y N_i obtener bola blanca y negra en la i -ésima extracción, respectivamente, para $i = 1, 2$. Entonces:

- a) $P(N_2 | N_1) = 3/6.$
- b) $P(B_1 \cap B_2) + P(N_1 \cap N_2) = P(B_2 | B_1)P(B_1) + P(N_2 | N_1)P(N_1)$
 $= (2/6)(3/7) + (3/6)(4/7) = 3/7.$
- c) $P(B_2) = P(B_2 | B_1)P(B_1) + P(B_2 | N_1)P(N_1)$
 $= (2/6)(3/7) + (3/6)(4/7) = 3/7.$

$$d) P(B_1 | B_2) = P(B_2 | B_1)P(B_1)/P(B_2) = (2/6)(3/7)/(3/7) = 1/3.$$

123. Sea N el evento de interés, esto es, cuando el nombre CAROLINA permanece sin cambio después de efectuar el experimento aleatorio. Sea A el evento seleccionar las letras AA. Denotando por C_2^8 a $\binom{8}{2}$, tenemos que

$$\begin{aligned} P(N) &= P(N | A)P(A) + P(N | A^c)P(A^c) \\ &= (1)(1/C_2^8) + (1/2)(C_2^8 - 1)/C_2^8 \\ &= 1/28 + (1/2)(27/28) = 29/56. \end{aligned}$$

124. Conocer la información solicitada al celador no cambia la probabilidad de ser ejecutado para el prisionero A . Para verificar esta afirmación, defina los eventos A , B y C , en donde el correspondiente preso es ejecutado. En particular, $P(A) = 1/3$. Sea el evento B^* cuando el celador menciona al prisionero B como uno de los dos que serán puestos en libertad. Supongamos que A sabe cuál de sus dos compañeros saldrá en libertad, por ejemplo, B . Entonces, con esta información, la probabilidad de que A sea ejecutado es

$$\begin{aligned} P(A | B^*) &= \frac{P(B^* | A)P(A)}{P(B^* | A)P(A) + P(B^* | B)P(B) + P(B^* | C)P(C)} \\ &= \frac{(1/2)(1/3)}{(1/2)(1/3) + (0)(1/3) + (1)(1/3)} \\ &= 1/3. \end{aligned}$$

Observe que B^* no es necesariamente B^c , pues el primer evento incluye la situación en donde el celador escoge al azar entre B^c o C^c cuando A es el que será ejecutado. Compare la similitud de situaciones de este problema y el problema de Monty Hall.

125. Se tienen dos razonamientos.

- Inicialmente la probabilidad de que el concursante gane es de $1/3$. Después de que se abre una de las puertas, la probabilidad de ganar es de $1/2$. Esta probabilidad no cambia si se decide cambiar de puerta. Así es que no hay diferencia entre cambiar de puerta o no hacerlo.
- Inicialmente la probabilidad de ganar es de $1/3$ y la probabilidad de que el premio se encuentre en una de las dos puertas no escogidas es $2/3$. Si se toma la decisión de cambiar de puerta una vez que se ha abierto una puerta sin premio, la probabilidad de ganar se incrementa a $2/3$. Por lo tanto, el concursante debe siempre hacer el cambio.

¿Cuál de estas soluciones es la correcta? Sean A , B y C las tres puertas. Denote por las mismas letras al evento cuando el premio se encuentre en

la puerta correspondiente. Sin pérdida de generalidad, suponga que A es la puerta escogida por el jugador. Entonces $P(A) = 1/3$. Sea B^* la puerta que abre el presentador. Entonces, si el jugador no cambia de puerta, la probabilidad de ganar es

$$\begin{aligned} P(A|B^*) &= \frac{P(B^*|A)P(A)}{P(B^*|A)P(A) + P(B^*|B)P(B) + P(B^*|C)P(C)} \\ &= \frac{(1/2)(1/3)}{(1/2)(1/3) + (0)(1/3) + (1)(1/3)} \\ &= 1/3. \end{aligned}$$

Por otro lado, si cambia a la puerta no abierta, a la cual llamaremos C , la probabilidad de ganar es

$$\begin{aligned} P(C|B^*) &= \frac{P(B^*|C)P(C)}{P(B^*|A)P(A) + P(B^*|B)P(B) + P(B^*|C)P(C)} \\ &= \frac{(1)(1/3)}{(1/2)(1/3) + (0)(1/3) + (1)(1/3)} \\ &= 2/3. \end{aligned}$$

Por lo tanto, el concursante debe siempre cambiar de puerta, pues al hacerlo se incrementa la probabilidad de ganar de $1/3$ a $2/3$. Numéricamente, la solución puede corroborarse mediante un programa de cómputo como el que aparece abajo.

```
puertas <- c(1:3)                                # Puertas
sincambio <- 0                                   # Variable auxiliar
concambio <- 0                                   # Variable auxiliar
n <- 1000                                        # Numero de simulaciones
for(i in 1:n){
  x <- c(sample(puertas,1))                      # Selección del jugador
  premio <- c(sample(puertas,1))                 # Puerta con premio
  psinpremio <- c(setdiff(puertas,premio))
  psinpremio <- c(setdiff(psinpremio,x))
  if (length(psinpremio)==1) { pabierta <- psinpremio }
  else { pabierta <- c(sample(psinpremio,1)) }
  pcambio <- setdiff(puertas,x)
  pcambio <- setdiff(pcambio,pabierta)
  if (x==premio) { sincambio <- sincambio+1 }
  if (premio==pcambio) { concambio <- concambio+1 }
}
```



```
cat("Sin cambio la prob. de ganar es", sincambio/n, "\n")
cat("Con cambio la prob. de ganar es", concambio/n, "\n")
```

1.15 Teorema de Bayes

126. a) Dado que la moneda obtenida es de oro, ésta sólo pudo provenir de la primera o de la segunda caja. Dado que las cajas se escogen al azar, la probabilidad de que la caja de procedencia sea la primera es $1/2$.
- b) Sean los eventos

$$\begin{aligned} C_i &= \text{“Escoger la caja } i\text{”}, \quad i = 1, 2, 3, \\ O &= \text{“Obtener moneda de oro”}. \end{aligned}$$

Entonces la pregunta planteada consiste en encontrar $P(C_1 | O)$. Por el teorema de Bayes,

$$\begin{aligned} P(C_1 | O) &= \frac{P(O | C_1)P(C_1)}{P(O | C_1)P(C_1) + P(O | C_2)P(C_2) + P(O | C_3)P(C_3)} \\ &= \frac{(1)(1/3)}{(1)(1/3) + (1/2)(1/3) + (0)(1/3)} = 2/3. \end{aligned}$$

127. a) $p = 11/108$. b) $p = 1/4$.

128. Sean los eventos

$$\begin{aligned} B &= \text{“Se obtiene bola blanca de la urna } I\text{.”} \\ B_i &= \text{“Se obtienen } i \text{ bolas blancas de la urna } II\text{”}, \quad i = 0, 1, 2. \end{aligned}$$

Entonces

$$\begin{aligned} a) \quad P(B) &= P(B | B_0)P(B_0) + P(B | B_1)P(B_1) + P(B | B_2)P(B_2) \\ &= (1/4)(2/4)(1/3) + (2/4)2(2/4)(2/3) + (3/4)(2/4)(1/3) \\ &= 1/2. \end{aligned}$$

$$\begin{aligned} b) \quad 1 - P(B_0 | B) &= 1 - P(B | B_0)P(B_0)/P(B) \\ &= 1 - (1/4)(2/4)(1/3)(2/1) = 11/12. \end{aligned}$$

129. Sean nuevamente E_0 y E_1 los eventos correspondientes a enviar un 0 y un 1, y sean R_0 y R_1 los eventos recibir un 0 y un 1. Antes se había encontrado que $P(R_0) = 0.415$ y $P(R_1) = 0.585$. Entonces

$$a) \quad P(E_0 | R_0) = P(R_0 | E_0)P(E_0)/P(R_0) = (0.8)(0.45)/(0.415) = 0.867.$$

$$b) \quad P(E_1 | R_1) = P(R_1 | E_1)P(E_1)/P(R_1) = (0.9)(0.55)/(0.585) = 0.846.$$

130. Defina el evento A_i como aquel en el que se obtiene una bola azul en la i -ésima extracción, para $i = 1, 2$. Entonces

$$a) P(A_2 | A_1) = 3/6.$$

$$b) P(A_1 \cap A_2) + P(A_1^c \cap A_2^c) = P(A_2 | A_1)P(A_1) + P(A_2^c | A_1^c)P(A_1^c) \\ = (3/6)(4/7) + (2/6)(3/7) = 3/7.$$

$$c) P(A_1 \cap A_2^c) + P(A_1 \cap A_2) = 1 - 3/7 = 4/7.$$

$$d) P(A_2) = P(A_2 | A_1)P(A_1) + P(A_2 | A_1^c)P(A_1^c) \\ = (3/6)(4/7) + (4/6)(3/7) = 4/7.$$

$$e) P(A_1 | A_2) = P(A_2 | A_1)P(A_1)/P(A_2) = (3/6)(4/7)(7/4) = 1/2.$$

131. Defina los eventos R_I y R_{II} como aquel en donde se obtiene bola roja de las urnas I y II , respectivamente. Entonces

$$P(R_{II}) = P(R_{II} | R_I)P(R_I) + P(R_{II} | R_I^c)P(R_I^c) \\ = (2/3)(1/2) + (1/3)(1/2) = 1/2.$$

Por lo tanto,

$$P(R_I | R_{II}) = P(R_{II} | R_I)P(R_I)/P(R_{II}) = (2/3)(1/2)(2/1) = 2/3.$$

1.16 Independencia de eventos

132. a) $P(B) = 1/2$. b) $P(B) = 3/10$. c) $P(B) = 3/5$.

133. a) Los eventos A y B no son independientes pues $1/36 = P(A \cap B) \neq P(A)P(B) = (5/36)(4/6)$.

b) Los eventos A y B son independientes pues $1/36 = P(A \cap B) = P(A)P(B) = (6/36)(1/6)$.

134. Únicamente se establece la validez o invalidez de la afirmación.

a) Cierto.

b) Cierto.

c) Cierto.

d) Cierto.

e) Cierto.

f) Falso, puede ocurrir que $P(A) = 1$.

g) Falso, puede ocurrir que $P(A) = 0$.

h) Cierto, pues $1/4 = P(A \cap B) = P(A)P(B) = (1/2)(1/2)$.

135. Se muestra únicamente la implicación $a) \Rightarrow b)$. Las otras implicaciones pueden demostrarse de manera similar.

$$\begin{aligned} P(A \cap B^c) &= P(A) - P(A \cap B) \\ &= P(A) - P(A)P(B) \\ &= P(A)(1 - P(B)) \\ &= P(A)P(B^c). \end{aligned}$$

136. \Rightarrow) Evidente.

$$\begin{aligned} \Leftarrow) P(A \cap B) &= P(A \cap B^c)P(B)/P(B^c) = (P(A) - P(A \cap B))P(B)/P(B^c). \\ \text{De donde se obtiene,} \\ P(A \cap B) &= (1 + P(B)/P(B^c))^{-1}P(A)P(B)/P(B^c) = P(A)P(B). \end{aligned}$$

137. Es inmediato comprobar las afirmaciones de la tabla.

138. a) $P(A \cup B) = P(A) + P(B) - P(A)P(B)$.
 b) $P(A \cup B^c) = 1 - P(B) + P(A)P(B)$.
 c) $P(A - B) = P(A) - P(A)P(B)$.
 d) $P(A \Delta B) = P(A) + P(B) - 2P(A)P(B)$.
 e) $P(A - (A \cap B)) = P(A) - P(A)P(B)$.
 f) $P(A^c \cup B^c) = 1 - P(A)P(B)$.

139. a) $P(A^c \cap B^c) = (1 - a)(1 - b)$.
 b) $P(A \cap B^c) + P(A^c \cap B) = a + b - 2ab$.
 c) $1 - P(A^c \cap B^c) = a + b - ab$.
 d) $P(A \cap B) = ab$.
 e) $1 - P(A \cap B) = 1 - ab$.

140. Las cuatro afirmaciones son falsas. Puede tomarse como ejemplo el espacio muestral equiprobable $\Omega = \{1, 2, 3, 4\}$ y los eventos $A = \{1, 2\}$ y $B = \{2, 3\}$. Se comprueba que A y B son independientes y sin embargo las parejas de eventos dadas en los incisos no son independientes.

141. Demostraremos $(a) \Leftrightarrow (b)$, en donde los eventos de cualquiera de estos incisos son los mismos que los del otro inciso, excepto uno de ellos que ha sido reemplazado por su complemento. Esta misma situación se presenta en las afirmaciones $(b) \Leftrightarrow (c)$ y $(c) \Leftrightarrow (d)$. De modo que la demostración que presentaremos es la misma para las afirmaciones indicadas. Además, debido

a la misma observación, es suficiente demostrar $(a) \Rightarrow (b)$. Primeramente tenemos que

$$\begin{aligned} P(A^c \cap B \cap C) &= P(B \cap C) - P(A \cap B \cap C) \\ &= P(B)P(C) - P(A)P(B)P(C) \\ &= (1 - P(A))P(B)P(C) \\ &= P(A^c)P(B)P(C). \end{aligned}$$

Para comprobar la independencia dos a dos, el procedimiento modelo es análogo.

$$\begin{aligned} P(A^c \cap B) &= P(B) - P(A \cap B) \\ &= P(B) - P(A)P(B) \\ &= (1 - P(A))P(B) \\ &= P(A^c)P(B). \end{aligned}$$

142. Todas las afirmaciones son falsas. Tomando como espacio muestral el conjunto equiprobable $\Omega = \{1, 2, 3, 4\}$, considere los eventos:

- a) $A = \{1, 2\}$, $B = \{1, 3\}$ y $C = \{2, 4\}$.
- b) Como en el inciso anterior.
- c) $A = \{1, 2\}$, $B = \{2, 3\}$ y $C = \{2, 4\}$.
- d) Como en el inciso anterior.

143. a)

$$\begin{aligned} P(A \cap (B \cap C)) &= P(A) - P(A \cap (B \cap C)^c) \\ &= P(A) - P(A \cap (B^c \cup C^c)) \\ &= P(A) - P((A \cap B^c) \cup (A \cap C^c)) \\ &= P(A) - P(A \cap B^c) - P(A \cap C^c) + P(A \cap B^c \cap C^c) \\ &= P(A) - P(A \cap B^c) - P(A \cap C^c) + P(A \cap (B \cup C)^c) \\ &= P(A) - P(A)P(B^c) - P(A)P(C^c) + P(A)P((B \cup C)^c) \\ &= P(A)(P(B) + P(C) - P(B \cup C)) \\ &= P(A)P(B \cap C). \end{aligned}$$

b)

$$\begin{aligned}
P(A \cap (B \cup C)) &= P((A \cap B) \cup (A \cap C)) \\
&= P(A \cap B) + P(A \cap C) - P(A \cap B \cap C) \\
&= P(A)P(B) + P(A)P(C) - P(A)P(B \cap C) \\
&= P(A)(P(B) + P(C) - P(B \cap C)) \\
&= P(A)P(B \cup C).
\end{aligned}$$

144. Sea D_i el evento en el cual el avión es detectado por el radar i , $i = 1, 2, 3$.
Entonces

$$a) P(D_1^c \cap D_2^c \cap D_3^c) = P(D_1^c)P(D_2^c)P(D_3^c) = (0.15)^3.$$

$$b) 1 - P(D_1^c \cap D_2^c \cap D_3^c) = 1 - (0.15)^3.$$

$$c) 3P(D_1 \cap D_2 \cap D_3^c) + P(D_1 \cap D_2 \cap D_3) = 3(0.85)^2(0.15) + (0.85)^3.$$

145. a) $P(A \cap (B \cap C)) = P(A)P(B)P(C) = P(A)P(B \cap C)$.

$$\begin{aligned}
b) P(A \cap (B \cup C)) &= P((A \cap B) \cup (A \cap C)) \\
&= P(A)P(B) + P(A)P(C) - P(A)P(B)P(C) \\
&= P(A)(P(B) + P(C) - P(B \cap C)) = P(A)P(B \cup C).
\end{aligned}$$

c) Idéntico al inciso (a).

d) Idéntico al inciso (b).

$$e) P(B \cap (A - C)) = P(B)P(A)P(C^c) = P(B)P(A \cap C^c).$$

$$\begin{aligned}
f) P(B \cap (A^c \cup C^c)) &= P((B \cap A^c) \cup (B \cap C^c)) \\
&= P(B)P(A^c) + P(B)P(C^c) - P(A^c)P(B)P(C^c) \\
&= P(B)(P(A^c) + P(C^c) - P(A^c \cap C^c)) \\
&= P(B)P(A^c \cup C^c).
\end{aligned}$$

146. a) $P(A \cup B \cup C) = 1 - P(A^c)P(B^c)P(C^c)$.

$$b) P(A \cup B \cup C^c) = 1 - P(A^c)P(B^c)P(C).$$

$$c) P(A^c \cap B^c \cap C^c) = P(A^c)P(B^c)P(C^c).$$

$$d) P(A - B - C) = P(A)P(B^c)P(C^c).$$

$$e) P(A \cap (B \cup C)) = P(A)(P(B) + P(C) - P(B)P(C)).$$

$$f) P(A \cap B \cap C^c) = P(A)P(B)P(C^c).$$

$$g) P(A \cup (B \cap C)) = P(A) + P(B)P(C) - P(A)P(B)P(C).$$

$$h) P(A \Delta B) = P(A)P(B^c) + P(A^c)P(B).$$

$$\begin{aligned}
i) P(A \Delta B \Delta C) &= P(A)P(B^c)P(C^c) + P(A^c)P(B)P(C^c) \\
&\quad + P(A^c)P(B^c)P(C).
\end{aligned}$$

147. a) $P(A^c \cap B^c \cap C^c) = (1-p)(1-q)(1-r)$.
 b) $P(A \cap B^c \cap C^c) + P(A^c \cap B \cap C^c) + P(A^c \cap B^c \cap C)$
 $= p(1-q)(1-r) + (1-p)q(1-r) + (1-p)(1-q)r$.
 c) $1 - P(A^c \cap B^c \cap C^c) = 1 - (1-p)(1-q)(1-r)$.
 d) $P(A \cap B \cap C^c) + P(A \cap B^c \cap C) + P(A^c \cap B \cap C)$
 $= pq(1-r) + p(1-q)r + (1-p)qr$.
 e) $P(A \cap B \cap C^c) + P(A \cap B^c \cap C) + P(A^c \cap B \cap C) + P(A \cap B \cap C)$
 $= pq(1-r) + p(1-q)r + (1-p)qr + pqr$.
 f) $P(A^c \cap B^c \cap C^c) + P(A \cap B^c \cap C^c) + P(A^c \cap B \cap C^c) + P(A^c \cap B^c \cap C)$
 $= (1-p)(1-q)(1-r) + p(1-q)(1-r) + (1-p)q(1-r) + (1-p)(1-q)r$.
 g) $1 - P(A \cap B \cap C) = 1 - pqr$.
 h) $P(A \cap B \cap C) = pqr$.

148. Para verificar que n eventos son independientes, se deben tomar éstos por pares, por tres, por cuatro, etcétera. El total de identidades por verificar es entonces

$$\binom{n}{2} + \binom{n}{3} + \cdots + \binom{n}{n} = \sum_{k=0}^n \binom{n}{k} - \binom{n}{0} - \binom{n}{1} = 2^n - 1 - n.$$

149. Considere el espacio muestral equiprobable $\Omega = \{1, 2, 3, 4\}$ y los eventos $A = \{1, 2\}$, $B_1 = \{1, 3\}$ y $B_2 = \{1, 4\}$. Es inmediato verificar que A es independiente de B_1 y también de B_2 . Sin embargo, A no es independiente de $B_1 \cup B_2$, $B_1 \cap B_2$, $B_1 - B_2$, ni tampoco de $B_1 \triangle B_2$.

150. Simplemente tome complemento y use el hecho de que los eventos A_1^c, \dots, A_n^c también son independientes.

151. a) $P(\text{"Al menos un evento ocurra"}) = 1 - (1-p)^n$.
 b) $P(\text{"Exactamente } m \text{ eventos ocurran"}) = \binom{n}{m} p^m (1-p)^{n-m}$.
 c) $P(\text{"Al menos } m \text{ eventos ocurran"}) = \sum_{k=m}^n \binom{n}{k} p^k (1-p)^{n-k}$.
 d) $P(\text{"A lo sumo } m \text{ eventos ocurran"}) = \sum_{k=0}^m \binom{n}{k} p^k (1-p)^{n-k}$.

152. Considere el espacio muestral equiprobable $\Omega = \{1, 2, 3, 4\}$.

- a) Defina los eventos $A = \{1, 2\}$, $B = \{1, 3\}$ y $C = \{2, 3\}$. Entonces A y B son independientes, pero no son condicionados al evento C pues $0 = P(A \cap B | C) \neq P(A | C)P(B | C) = (1/2)(1/2)$.

b) Defina los eventos $A = B = C = \{1\}$. Entonces $P(A \cap B | C) = P(A | C)P(B | C) = 1$, pero claramente A y B no son independientes.

153. Si el arreglo en paralelo tiene n componentes, entonces el evento cuando el arreglo funciona es $C_1 \cup \dots \cup C_n$, en donde C_i denota el evento en donde el i -ésimo componente funciona. Así, por la hipótesis de independencia,

$$P(C_1 \cup \dots \cup C_n) = 1 - P(C_1^c) \dots P(C_n^c) = 1 - (0.05)^n \geq 0.99.$$

De donde se obtiene $n \geq 1.53$. Es decir, basta con dos componentes en paralelo para satisfacer la condición solicitada.

154. Defina los eventos

$$\begin{aligned} A &= \text{“El jugador 1 encesta primero”}, \\ A_n &= \text{“El jugador 1 encesta su } n\text{-ésimo tiro”}. \end{aligned}$$

De manera similar defina los eventos B y B_n para el segundo jugador. Entonces, por la hipótesis de independencia,

$$\begin{aligned} P(A) &= P(A_1) + P(A_1^c \cap B_1^c \cap A_2) + P(A_1^c \cap B_1^c \cap A_2^c \cap B_2^c \cap A_3) + \dots \\ &= p + (1-p)(1-q)p + (1-p)^2(1-q)^2p + \dots \\ &= p \sum_{k=0}^{\infty} ((1-p)(1-q))^k \\ &= \frac{p}{1 - (1-p)(1-q)}. \end{aligned}$$

Análogamente,

$$\begin{aligned} P(B) &= P(A_1^c \cap B_1) + P(A_1^c \cap B_1^c \cap A_2^c \cap B_2) \\ &\quad + P(A_1^c \cap B_1^c \cap A_2^c \cap B_2^c \cap A_3^c \cap B_3) + \dots \\ &= (1-p)q + (1-p)^2(1-q)p + (1-p)^3(1-q)^2q + \dots \\ &= q(1-p) \sum_{k=0}^{\infty} ((1-p)(1-q))^k \\ &= \frac{q(1-p)}{1 - (1-p)(1-q)}. \end{aligned}$$

1.17 Continuidad de la probabilidad

155. En cada caso considere la sucesión de complementos de los eventos originales. Esto ayuda a demostrar una proposición a partir de la otra.

2.2 Función de probabilidad

166. Es inmediato verificar que las funciones son no negativas. Comprobaremos que suman o integran uno.

$$\begin{aligned} a) \sum_{x=1}^{\infty} (x-1)(1/2)^x &= \sum_{x=1}^{\infty} x(1/2)^{x+1} = \sum_{x=1}^{\infty} \left(\sum_{y=1}^x 1 \right) (1/2)^{x+1} \\ &= \sum_{y=1}^{\infty} \sum_{x=y}^{\infty} (1/2)^{x+1} = 1. \end{aligned}$$

$$b) \frac{1-p}{1-p^n} \sum_{x=1}^n p^{x-1} = \frac{1-p}{1-p^n} \frac{1-p^n}{1-p} = 1.$$

$$c) (1/2) \int_0^1 x^{-1/2} dx = (1/2) x^{1/2} / (1/2) \Big|_0^1 = 1.$$

$$d) \int_{-1}^1 f(x) dx = 2 \int_0^1 (3/2)(1-x)^2 dx = -(1-x)^3 \Big|_0^1 = 1.$$

167.

- | | |
|----------------------------------|----------------------------|
| a) $\theta = 2/(n(n+1))$. | h) $\theta = 6/5$. |
| b) $\theta = 6/(n(n+1)(2n+1))$. | i) $\theta = \ln 2$. |
| c) $\theta = 1/(2(2^n - 1))$. | j) $\theta = 3/4$. |
| d) $\theta = 2^n/(2^n - 1)$. | k) $\theta = 1$. |
| e) $\theta = 120/53$. | l) $\theta = 1/2$. |
| f) $\theta = 1/2$. | m) $\theta = (n+1)(n+2)$. |
| g) $\theta = 3/2$. | n) $\theta = (n+1)(n+2)$. |

168. a) No puede ser función de densidad pues es negativa sobre una parte del intervalo donde está definida.

b) Sí puede ser una función de densidad cuando se toma $c = 3/4$.

- | | |
|-------------------------------|-----------------------------|
| 169. a) $P(X \leq 1) = 7/8$. | d) $P(X^2 \geq 1) = 1/2$. |
| b) $P(X \leq 1) = 3/4$. | e) $P(X - 1 < 2) = 3/4$. |
| c) $P(-1 < X \leq 2) = 3/4$. | f) $P(X - X^2 < 0) = 1/4$. |

170. El valor de c es 3.

$$\begin{array}{ll} a) P(X \leq 1/2) = 1/8. & c) P(1/4 < X < 3/4) = 13/32. \\ b) P(X \geq 1/4) = 63/64. & d) a = 1/\sqrt[3]{2}. \end{array}$$

$$171. \quad \begin{array}{ll} a) P(X \leq 1/3) = 14/27. & e) P(-1/4 < X < 2) = 65/128. \\ b) P(X \geq 0) = 1/2. & f) P(2X + 1 < 2) = 9/16. \\ c) P(|X| > 1) = 0. & g) P(X^2 < 1/9) = 1/27. \\ d) P(|X| < 1/2) = 1/8. & h) P(X^2 - X < 2) = 1. \end{array}$$

172. Claramente $g(x)$ es no negativa. Además

$$\begin{aligned} 2 \int_0^{\infty} f(a+x) dx &= \int_0^{\infty} f(a-x) dx + \int_0^{\infty} f(a+x) dx \\ &= \int_{-\infty}^a f(u) du + \int_a^{\infty} f(u) du \\ &= 1. \end{aligned}$$

173. Únicamente se proporciona la validez o invalidez de la afirmación. En cada caso es necesario producir una demostración o un contraejemplo.

$$\begin{array}{ll} a) \text{ Falso.} & e) \text{ Verdadero.} \\ b) \text{ Falso.} & f) \text{ Verdadero.} \\ c) \text{ Verdadero.} & g) \text{ Verdadero.} \\ d) \text{ Verdadero.} & h) \text{ Falso.} \end{array}$$

$$174. \quad \begin{array}{l} a) f_Y(y) = \begin{cases} f_X(y/2) & \text{si } y = 0, 2, 4, \dots \\ 0 & \text{en otro caso.} \end{cases} \\ b) f_Y(y) = \begin{cases} f_X(y-n) & \text{si } y = n, n+1, \dots \\ 0 & \text{en otro caso.} \end{cases} \\ c) f_Y(y) = \begin{cases} \sum_{x=0}^{\infty} f_X(4x+2) & \text{si } y = -1, \\ \sum_{x=0}^{\infty} f_X(2x+1) & \text{si } y = 0, \\ \sum_{x=0}^{\infty} f_X(4x) & \text{si } y = 1, \\ 0 & \text{en otro caso.} \end{cases} \\ d) f_Y(y) = \begin{cases} \sum_{x=0}^{\infty} f_X(2x) & \text{si } y = 0, \\ \sum_{x=0}^{\infty} f_X(2x+1) & \text{si } y = 1, \\ 0 & \text{en otro caso.} \end{cases} \end{array}$$

$$175. f(x) = \begin{cases} (6 - |x|)/36 & \text{si } x = -5, \dots, 5, \\ 0 & \text{en otro caso.} \end{cases}$$

$$176. a) \theta = 4.$$

b) Debido a la forma de la función de densidad, el intervalo (a, b) de longitud mínima tiene que ser simétrico respecto del valor $1/2$. Haciendo el cálculo de las integrales (o áreas de triángulos) correspondientes se encuentra que $a = 3/8$ y $b = 5/8$.

$$177. f(x) = \begin{cases} 19/25 & \text{si } x = 1, \\ (19/(26 - x)) \prod_{i=0}^{x-2} (6 - i)/(25 - i) & \text{si } x = 2, \dots, 7, \\ 0 & \text{en otro caso.} \end{cases}$$

178. a) Denotando por las mismas letras A y B los eventos en donde el jugador respectivo gana, tenemos que $P(A) = 2/3$ y $P(B) = 1/3$.

$$b) f(x) = \begin{cases} (1/2)^x & \text{si } x = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

$$179. f(x) = \begin{cases} (1/6)(5/6)^{x-1} & \text{si } x = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

$$180. f(x) = \begin{cases} 6(5!/(7-x)!)\binom{x-1}{1}/6^x & \text{si } x = 2, \dots, 7, \\ 0 & \text{en otro caso.} \end{cases}$$

$$181. f(x) = \begin{cases} 2\binom{x-1}{2}/2^x & \text{si } x = 3, 4, 5, \\ 0 & \text{en otro caso.} \end{cases}$$

$$182. f(x) = \begin{cases} \binom{10}{x}(1/4)^x(3/4)^{n-x} & \text{si } x = 0, 1, \dots, 10, \\ 0 & \text{en otro caso.} \end{cases}$$

183. Se muestran únicamente las expresiones analíticas de las funciones de probabilidad condicionales. La distribución de probabilidad original se transforma en una nueva distribución de probabilidad, concentrada en el respectivo conjunto A .

$$a) f(x) = \begin{cases} 1/4 & \text{si } x = 2, 3, 4, 5, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x) = \begin{cases} 1 & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f(x) = \begin{cases} e^{-(x-1)} & \text{si } x > 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$d) f(x) = \begin{cases} (4-2x)/3 & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

2.3 Función de distribución

184. Únicamente se proporcionan las expresiones de estas funciones.

$$f(x) = \begin{cases} p & \text{si } x = 3, \\ 1-p & \text{si } x = 5, \\ 0 & \text{en otro caso.} \end{cases}$$

$$F(x) = \begin{cases} 0 & \text{si } x < 3, \\ p & \text{si } 3 \leq x < 5, \\ 1 & \text{si } x \geq 5. \end{cases}$$

185. Puede usted escoger desde una variable aleatoria constante hasta una que toma seis valores distintos. Elija usted los valores numéricos que toma la variable aleatoria.

186. Las funciones de probabilidad y de distribución son:

$$f(x) = \begin{cases} 1 & \text{si } x = c, \\ 0 & \text{en otro caso.} \end{cases}$$

$$F(x) = \begin{cases} 0 & \text{si } x < c, \\ 1 & \text{si } x \geq c. \end{cases}$$

187. Para el primer inciso, considere por ejemplo la sucesión no decreciente de eventos $A_n = (X \leq x - 1/n)$ cuyo límite es el evento $(X < x)$ cuando $n \rightarrow \infty$. Proceda como en la demostración de la propiedad de continuidad por la derecha de $F(x)$. Los otros incisos se siguen de este primero y de la definición $P(X \leq x) = F(x)$.

188. Para cada número natural n , defina el evento $A_n = (x - 1/n < X \leq x)$. Entonces A_1, A_2, \dots es una sucesión decreciente de eventos cuyo límite es el evento $(X = x)$. Por la propiedad de continuidad de las medidas de probabilidad para sucesiones monótonas decrecientes, tenemos que

$$\begin{aligned} P(X = x) &= \lim_{n \rightarrow \infty} P(x - 1/n < X \leq x) \\ &= \lim_{n \rightarrow \infty} F(x) - F(x - 1/n) \\ &= F(x) - F(x-). \end{aligned}$$

189. En cada caso compruebe que se cumplen las siguientes cuatro propiedades.

-) $\lim_{x \rightarrow \infty} F(x) = 1$.
-) $\lim_{x \rightarrow -\infty} F(x) = 0$.
-) Si $x_1 \leq x_2$, entonces $F(x_1) \leq F(x_2)$.
-) $F(x) = F(x+)$.

190. La función de probabilidad es

$$f(x) = \begin{cases} 1/4 & \text{si } x = -1, 1, 3, 5, \\ 0 & \text{en otro caso.} \end{cases}$$

Las probabilidades buscadas son

- a) $P(X \leq 3) = 3/4$.
- b) $P(X = 3) = 1/4$.
- c) $P(X < 3) = 2/4$.
- d) $P(X \geq 1) = 3/4$.
- e) $P(-1/2 < X < 4) = 2/4$.
- f) $P(X = 5) = 1/4$.

191. En cada caso observe que se trata de una función no negativa que integra o suma uno, por lo tanto es una función de probabilidad. Las correspondientes funciones de distribución son:

$$a) F(x) = \begin{cases} 0 & \text{si } x < 1, \\ 1 - (1/2)^k & \text{si } k \leq x < k + 1; k = 1, 2, \dots \end{cases}$$

$$b) F(x) = \begin{cases} 0 & \text{si } x < 0, \\ x^2 & \text{si } 0 \leq x \leq 1, \\ 1 & \text{si } x > 1. \end{cases}$$

$$c) F(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - (1 - x)^2 & \text{si } 0 \leq x \leq 1, \\ 1 & \text{si } x > 1. \end{cases}$$

$$d) F(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 - e^{-4x} & \text{si } x > 0. \end{cases}$$

192. Gráficamente puede observarse que las funciones dadas son no negativas e integran uno. Por lo tanto, son funciones de probabilidad. Las correspondientes funciones de distribución son:

$$\begin{aligned}
 a) \quad F(x) &= \begin{cases} 0 & \text{si } x < 0, \\ x^2/2 & \text{si } 0 \leq x < 1, \\ 1 - (2-x)^2/2 & \text{si } 1 \leq x < 2, \\ 1 & \text{si } x > 2. \end{cases} \\
 b) \quad F(x) &= \begin{cases} 0 & \text{si } x < 0, \\ x^2/a^2 & \text{si } 0 \leq x < a, \\ 1 & \text{si } x \geq a. \end{cases} \\
 c) \quad F(x) &= \begin{cases} 0 & \text{si } x < -\sqrt{2/\pi}, \\ (r^2/2)[\pi/2 + \arcsen(x/r) & \text{si } -\sqrt{2/\pi} \leq x < \sqrt{2/\pi}, \\ + (x/r)\sqrt{1 - (x/r)^2}] & \\ 1 & \text{si } x \geq \sqrt{2/\pi}. \end{cases} \\
 d) \quad F(x) &= \begin{cases} 0 & \text{si } x < -a, \\ (1 - x^2/a^2)/2 & \text{si } -a \leq x < 0, \\ 1 - (a-x)^2/(2a^2) & \text{si } 0 \leq x < a, \\ 1 & \text{si } x \geq a. \end{cases} \\
 e) \quad F(x) &= \begin{cases} 0 & \text{si } x < -2a, \\ (x+2a)/(3a) & \text{si } -2a \leq x < -a, \\ 1/3 + (x+a)/(6a) & \text{si } -a \leq x < a, \\ 2/3 + (x-a)/(3a) & \text{si } a \leq x < 2a, \\ 1 & \text{si } x \geq 2a. \end{cases} \\
 f) \quad F(x) &= \begin{cases} 0 & \text{si } x < 0, \\ 1/2 - (a-x)^2/(2a^2) & \text{si } 0 \leq x < a, \\ 1/2 + (x-a)^2/(2a^2) & \text{si } a \leq x < 2a, \\ 1 & \text{si } x \geq 2a. \end{cases}
 \end{aligned}$$

193. En cada caso compruebe que la función satisface las cuatro propiedades para que sea una función de distribución. Sólo el inciso (e) corresponde a una variable aleatoria discreta, en los otros casos se trata de v.a.s continuas. Las correspondientes funciones de probabilidad son las siguientes:

$$a) \quad f(x) = \begin{cases} 1 & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f(x) = \begin{cases} e^{-x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f(x) = \begin{cases} \text{sen } x & \text{si } 0 < x < \pi/2, \\ 0 & \text{en otro caso.} \end{cases}$$

$$d) f(x) = \begin{cases} 1/(2a) & \text{si } -a < x < a, \\ 0 & \text{en otro caso.} \end{cases}$$

$$e) f(x) = \begin{cases} 1/5 & \text{si } x = 0, \\ 2/5 & \text{si } x = 1, 2, \\ 0 & \text{en otro caso.} \end{cases}$$

$$f) f(x) = \begin{cases} 2x & \text{si } 0 < x < 1/2, \\ 3/2 & \text{si } 1/2 \leq x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$g) f(x) = \begin{cases} 1/2 & \text{si } -2 < x < -1, \\ 1/2 & \text{si } 1 < x < 2, \\ 0 & \text{en otro caso.} \end{cases}$$

194. a) La función $f_X(x)$ es una función de densidad pues es no negativa e integra uno. Ambos hechos son inmediatos de comprobar.
 b) Integrando la función de densidad, desde menos infinito hasta un valor x cualquiera, se obtiene que la función de distribución es

$$F_X(x) = \begin{cases} 0 & \text{si } x < 1, \\ 1 - 1/x & \text{si } x \geq 1. \end{cases}$$

$$c) F_Y(y) = \begin{cases} 0 & \text{si } y < 1/e, \\ -1/\ln y & \text{si } y \geq 1/e. \end{cases}$$

$$d) f_Y(y) = \begin{cases} (\ln y)^{-2}/y & \text{si } y \geq 1/e, \\ 0 & \text{en otro caso.} \end{cases}$$

195. Es inmediato comprobar que $f(x)$ es no negativa e integra uno. Por lo tanto, se trata de una función de densidad. La correspondiente función de distribución es

$$F(x) = \begin{cases} 0 & \text{si } x \leq -1, \\ (1 + x^3)/2 & \text{si } -1 < x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

Las probabilidades buscadas son:

- a) $P(|X| < 1/2) = (1/2)^3$. c) $P(|X| < 1/n) = (1/n)^3$.
 b) $P(X < 0) = 1/2$. d) $P(|X| > 1/2) = 1 - (1/2)^3$.

196. Es inmediato comprobar que $f(x)$ es no negativa e integra uno. Por lo tanto, se trata de una función de densidad. La correspondiente función de distribución es

$$F(x) = \begin{cases} -1/(4x) & \text{si } x < -1, \\ (2+x)/4 & \text{si } -1 \leq x \leq 1, \\ 1 - 1/(4x) & \text{si } x > 1. \end{cases}$$

Las probabilidades buscadas son:

- a) $P(|X| < 3/2) = 2/3$. c) $P(1/2 < X < 3/2) = 5/24$.
 b) $P(X > 0) = 1/2$. d) $P(|X| \leq 1) = 1/2$.

197. La función $f(x)$ es no negativa y un ejercicio simple de integración muestra que esta función integra uno. Por lo tanto, se trata de una función de densidad. La correspondiente función de distribución es

$$F(x) = \begin{cases} 0 & \text{si } x \leq -1, \\ (1+x)^3/2 & \text{si } -1 < x < 0, \\ 1 - (1-x)^3/2 & \text{si } 0 \leq x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

Las probabilidades buscadas son:

- a) $P(|X| < 1/2) = 7/8$. c) $P(1/2 < X < 3/2) = 1/16$.
 b) $P(X < 0) = 1/2$. d) $P(|X| > 1/2) = 1/8$.

198. a) Estas probabilidades están dadas por los saltos de la función de distribución en los puntos indicados y son:

$$\begin{array}{ll} P(X = 0) = 0. & P(X = 2) = 1/20. \\ P(X = 1/2) = 0. & P(X = 3) = 1/5. \\ P(X = 1) = 1/4. & P(X = 4) = 0. \end{array}$$

- b) $P(1/2 < X < 5/2) = 27/40$.

199. Sea $H(x) = \lambda F(x) + (1 - \lambda) G(x)$. Es sencillo comprobar que se cumplen las cuatro propiedades que determinan que esta función es de distribución:

$$\cdot) \lim_{x \rightarrow \infty} H(x) = 1.$$

-) $\lim_{x \rightarrow -\infty} H(x) = 0$.
-) Si $x_1 \leq x_2$, entonces $H(x_1) \leq H(x_2)$.
-) $H(x) = H(x+)$.

200. Se omiten las gráficas de estas funciones, pero no es difícil dibujarlas a partir de sus expresiones analíticas.

$$a) F_U(u) = \begin{cases} 0 & \text{si } u \leq 0, \\ 1 - e^{-u} & \text{si } 0 < u < c, \\ 1 & \text{si } u \geq c. \end{cases}$$

$$b) F_V(v) = \begin{cases} 0 & \text{si } v \leq c, \\ 1 - e^{-v} & \text{si } v > c. \end{cases}$$

201. Para cualquier número real x ,

$$F_{cX}(x) = P(cX \leq x) = P(X \leq x/c) = F_X(x/c).$$

Ahora derive respecto de x y utilice la regla de la cadena.

202. Observemos primero que la variable aleatoria X^2 únicamente puede tomar valores no negativos. Así, para cualquier $x > 0$,

$$F_{X^2}(x) = P(X^2 \leq x) = P(-\sqrt{x} \leq X \leq \sqrt{x}) = F_X(\sqrt{x}) - F_X(-\sqrt{x}).$$

Ahora derive respecto de x y utilice la regla de la cadena.

203. a) Falso. Como ejemplo pueden tomarse las variables constantes $X = 0$ y $Y = 1$. Entonces, para cualquier $x \in (0, 1)$, $1 = F_X(x) > F_Y(x) = 0$.
- b) Falso. Considere el espacio muestral $\Omega = (0, 1)$ y la variable aleatoria continua $X(\omega) = \omega$ con función de densidad como aparece abajo. Sea $Y = 1 - X$. Puede comprobarse que Y tiene la misma distribución que X y por lo tanto $F_X(x) \leq F_Y(x)$ para todo $x \in \mathbb{R}$. Sin embargo, para $1/2 < \omega < 1$, $X(\omega) = \omega > Y(\omega) = 1 - \omega$.

$$f(x) = \begin{cases} 1 & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

- c) Cierto. Como X y Y son funciones idénticas, para cualquier $\omega \in \Omega$, $X(\omega) = Y(\omega)$ y por lo tanto los eventos $(X(\omega) \leq x)$ y $(Y(\omega) \leq x)$ son idénticos para cualquier número real x . Tomando probabilidades, $F_X(x) = F_Y(x)$.

- d) Falso. Sea X con función de probabilidad como aparece abajo. Puede comprobarse que la variable $Y = 1 - X$ tiene la misma distribución que X , y sin embargo son distintas; cuando una toma el valor cero, la otra toma el valor uno y viceversa.

$$f(x) = \begin{cases} 1/2 & \text{si } x = 0, 1, \\ 0 & \text{en otro caso.} \end{cases}$$

204. Sea X discreta con función de probabilidad como aparece abajo y sea $Y = -X$. Entonces es claro que X y Y tienen la misma distribución de probabilidad. Sin embargo, $XY = -X^2 \neq X^2$.

$$f(x) = \begin{cases} 1/3 & \text{si } x = -1, 0, 1, \\ 0 & \text{en otro caso.} \end{cases}$$

205. Las funciones son:

$$f(x) = \begin{cases} (x-1)/36 & \text{si } x = 2, 3, 4, 5, 6, 7, \\ (13-x)/36 & \text{si } x = 8, 9, 10, 11, \\ 0 & \text{en otro caso.} \end{cases}$$

$$F(x) = \begin{cases} 0 & \text{si } x < 2, \\ k(k-1)/72 & \text{si } k \leq x < k+1; \quad k = 2, 3, 4, 5, 6, 7, \\ (42 + (k-7)(18-k))/72 & \text{si } k \leq x < k+1; \quad k = 8, 9, 10, 11, \\ 1 & \text{si } x \geq 12. \end{cases}$$

206. Para $x = 1, 2, \dots$

$$F(x) = P(X \leq x) = 1 - P(X > x) = 1 - (1/2)^x.$$

Por lo tanto,

$$F(x) = \begin{cases} 0 & \text{si } x < 1, \\ 1 - (1/2)^k & \text{si } k \leq x < k+1; \quad k = 1, 2, \dots \end{cases}$$

2.4 Teorema de cambio de variable

207. a) $f_Y(y) = f_X((y-b)/a)/a$.
 b) $f_Y(y) = f_X(-y)$.
 c) $f_Y(y) = (f_X(\sqrt[3]{y}) + f_X(-\sqrt[3]{y})) y^{-2/3}/2$.

$$d) f_Y(y) = \begin{cases} f_X(\ln y)/y & \text{si } y > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

208. $\theta = 1/6$.

$$a) f_Y(y) = \begin{cases} (y-b)^2/a^3 + (y-b)/a^2 + 1/(6a) & \text{si } b < x < a+b, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f_Y(y) = \begin{cases} 2y^5 + 2y^3 + y/3 & \text{si } 0 < y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f_Y(y) = \begin{cases} (1+y)^2/8 + (1+y)/4 + 1/12 & \text{si } -1 < y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

209. $\theta = 2$.

$$a) f_Y(y) = \begin{cases} (1 - |2\sqrt{y} - 1|)/\sqrt{y} & \text{si } 0 < y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f_Y(y) = \begin{cases} 2(y - |2 - y|)/y^3 & \text{si } y > 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f_Y(y) = \begin{cases} 2e^y(1 - |2e^y - 1|) & \text{si } y < 0, \\ 0 & \text{en otro caso.} \end{cases}$$

2.5 Independencia de variables aleatorias

210. a) Este resultado se obtiene de la siguiente identidad de eventos

$$\begin{aligned} (X = x, Y = y) &= (X \leq x, Y \leq y) - (X \leq x-1, Y \leq y) \\ &\quad - (X \leq x, Y \leq y-1) + (X \leq x-1, Y \leq y-1), \end{aligned}$$

lo cual puede verificarse a partir de la Figura C.1.

b) Por el inciso anterior y la hipótesis de independencia,

$$\begin{aligned} P(X = x, Y = y) &= P(X \leq x)P(Y \leq y) - P(X \leq x-1)P(Y \leq y) \\ &\quad - P(X \leq x)P(Y \leq y-1) \\ &\quad + P(X \leq x-1)P(Y \leq y-1) \\ &= [P(X \leq x) - P(X \leq x-1)] \\ &\quad \times [P(Y \leq y) - P(Y \leq y-1)] \\ &= P(X = x)P(Y = y). \end{aligned}$$

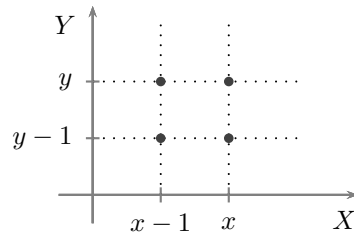


Figura C.1

211. a) La variable U toma valores en el conjunto $\{1, \dots, n\}$. Sea u uno de estos valores. Entonces

$$\begin{aligned}
 P(U \leq u) &= P(\text{máx}\{X_1, \dots, X_m\} \leq u) \\
 &= P(X_1 \leq u, \dots, X_m \leq u) \\
 &= P(X_1 \leq u) \cdots P(X_m \leq u) \\
 &= (u/n)^m.
 \end{aligned}$$

Por lo tanto, $P(U = u) = (u/n)^m - ((u-1)/n)^m$, $u = 1, \dots, n$.

- b) La variable V toma valores en el conjunto $\{1, \dots, n\}$. Sea v uno de estos valores. Entonces

$$\begin{aligned}
 P(V \geq v) &= P(\text{mín}\{X_1, \dots, X_m\} \geq v) \\
 &= P(X_1 \geq v, \dots, X_m \geq v) \\
 &= P(X_1 \geq v) \cdots P(X_m \geq v) \\
 &= ((n-v+1)/n)^m.
 \end{aligned}$$

Por lo tanto, $P(V = v) = ((n-v+1)/n)^m - ((n-v)/n)^m$, $v = 1, \dots, n$.

2.6 Esperanza

212. a) $E(X) = (n+1)/2$. c) $E(X) = (2n+1)/3$.
 b) $E(X) = 1$.

213.

217.

$$\begin{aligned}
 E(X) &= \int_0^{\infty} x f(x) dx \\
 &= \int_0^{\infty} \left(\int_0^x 1 dy \right) f(x) dx \\
 &= \int_0^{\infty} \int_y^{\infty} f(x) dx dy \\
 &= \int_0^{\infty} P(X > y) dy \\
 &= \int_0^{\infty} (1 - F(y)) dy.
 \end{aligned}$$

$$218. \quad a) \quad E(X) = \int_0^{\infty} (1 - F(x)) dx = \int_0^2 (1 - x/2) dx = 1.$$

$$b) \quad E(X) = \int_0^{\infty} (1 - F(x)) dx = \int_0^{\infty} e^{-x} dx = 1.$$

219. Únicamente se proporciona la validez o invalidez de cada afirmación.

a) Falso.

g) Falso, pero sí se puede concluir que $P(X = 0) = 1$.

b) Falso.

h) Falso.

c) Falso.

i) Falso.

d) Verdadero.

j) Verdadero.

e) Verdadero.

f) Falso.

k) Falso.

220. Claramente la función $f(x)$ es no negativa. Para demostrar que suma uno, utilice fracciones parciales. La esperanza de X no existe pues

$$\sum_{x=1}^{\infty} |x| f(x) = \sum_{x=1}^{\infty} 1/(x+1) = \infty.$$

221. a) La función $f(x)$ es no negativa y es sencillo comprobar que integra uno, por lo tanto es una función de densidad. La esperanza de esta distribución no existe pues

$$\int_1^{\infty} |x| f(x) dx = \int_1^{\infty} (1/x) dx = \infty.$$

- b) Claramente la función $f(x)$ es no negativa y puede comprobarse que integra uno al hacer el cambio de variable $x = \tan u$ y recordar que $dx = \sec^2 u du$ y $1 + \tan^2 u = \sec^2 u$. Por otro lado, la esperanza de esta distribución no existe pues nuevamente, al hacer el cambio de variable $x = \tan u$, se tiene que

$$\int_{-\infty}^{\infty} |x| f(x) dx = (2/\pi) \int_0^{\infty} x/(1+x^2) dx = (2/\pi) \int_0^{\pi/2} \tan u du = \infty.$$

222. El valor promedio del premio es infinito.
223. El número promedio de días es 2.
224. El valor promedio del premio es $161/36$.
225. a) Use inducción sobre n .
b) Observe que la función $\varphi(x) = 1/x$ es convexa. Utilice el inciso anterior.
226. $a = 1$.
227. a) $E(Y) = 1/8$.
b) $E(Y) = 1/2$.
c) $E(Y) = 1$.
d) $E(Y) = 0$.
228. a) $E(Y) = e(1-p)/(1-ep)$.
b) $E(Y) = \lambda/(\lambda-1)$.
229. Puede comprobarse que la variable XY tiene la misma distribución que X . Se comprueba entonces que $E(XY) = E(X)E(Y) = 0$. Sin embargo, X y Y no son independientes pues $1/3 = P(X = -1, Y = 1) \neq P(X = -1)P(Y = 1) = (1/3)(2/3)$.

230. Claramente ambas funciones son no negativas. Además,

$$a) \int_0^{\infty} 2(1-F(x))f(x) dx = -(1-F(x))^2 \Big|_0^{\infty} = 1.$$

$$b) \int_0^{\infty} (1-F(x))/\mu dx = 1.$$

231. La demostración es similar al caso demostrado en la Proposición 2.7, en la

página 162. Consideraremos únicamente el caso discreto.

$$\begin{aligned}
 E[g(X)h(Y)] &= \sum_{x,y} g(x)h(y)P(X=x, Y=y) \\
 &= \sum_x \sum_y g(x)h(y)P(X=x)P(Y=y) \\
 &= \left(\sum_x g(x)P(X=x) \right) \left(\sum_y h(y)P(Y=y) \right) \\
 &= E[g(X)]E[h(Y)].
 \end{aligned}$$

2.7 Varianza

232. a) $E(X) = (n+1)/2$, $\text{Var}(X) = (n^2-1)/12$.

b) $E(X) = 2$, $\text{Var}(X) = 2$.

233. a) $E(X) = 0$, $\text{Var}(X) = 5/18$.

b) $E(X) = 1$, $\text{Var}(X) = 1$.

c) $E(X) = 0$, $\text{Var}(X) = 1/2$.

234. a) $E(X) = 0$, $\text{Var}(X) = 2$.

b) $E(X) = 0$, $\text{Var}(X) = 1/6$.

c) $E(X) = 1/2$, $\text{Var}(X) = 1/20$.

d) $E(X) = 1/2$, $\text{Var}(X) = 1/12$.

e) $E(X) = 0$, $\text{Var}(X) = a^2/3$.

f) $E(X) = 1$, $\text{Var}(X) = \pi - 3$.

235. $P(X=1) = (1+\sqrt{5})/4$, $P(X=-1) = (3-\sqrt{5})/4$.

236. $\theta = \sqrt{2}$.

237. La función $f(x) = ax^2 + bx$ tiene raíces en $x_0 = 0$ y $x_1 = -b/a$, suponiendo $a \neq 0$. Por otro lado, las condiciones de que esta función debe ser no negativa e integrar uno en el intervalo $(0, 1)$ se cumplen en las siguientes cuatro situaciones.

·) $a = 0$ y $b > 0$. En este caso $b = 2$.

·) $a > 0$ y $b = 0$. En este caso $a = 3$.

·) $a > 0$ y $b > 0$. En este caso $b = (2/3)(3-a)$ y $0 < a < 3$.

·) $a < 0$ y $b > 0$. En este caso $-b/a \geq 1$ y $b = (2/3)(3-a)$.

- a) La esperanza es $E(X) = a/4 + b/3$. El valor mínimo es $1/2$ y se alcanza cuando $a = -6$ y $b = 6$.
- b) La varianza es $\text{Var}(X) = (a/5 + b/4) - (a/4 + b/3)^2$. El valor mínimo es $3/80$ y se alcanza cuando $a = 3$ y $b = 0$.

238. Usando las propiedades demostradas, tenemos que:

- a) $\text{Var}(X) = E(X^2) - E^2(X) \leq E(X^2)$.
- b) $\text{Var}(a - X) = \text{Var}(-X + a) = \text{Var}(-X) = \text{Var}(X)$.
- c) $\text{Var}(aX + b) = \text{Var}(aX) = a^2 \text{Var}(X)$.
- d) $\text{Var}(X + Y) = E[(X + Y)^2] - E^2(X + Y) = E(X^2) + 2E(XY) + E(Y^2) - E^2(X) - 2E(X)E(Y) - E^2(Y) = \text{Var}(X) + \text{Var}(Y) + 2E(XY) - 2E(X)E(Y) = \text{Var}(X) + \text{Var}(Y) + 2E[(X - E(X))(Y - E(Y))]$.

239. a)

$$\begin{aligned} g(x) &= E[(X - x)^2] \\ &= E[(X - E(X) - (x - E(X)))^2] \\ &= E[(X - E(X))^2] - 2(x - E(X))E[X - E(X)] + (x - E(X))^2 \\ &= \text{Var}(X) + (x - E(X))^2. \end{aligned}$$

- b) Esto es una consecuencia inmediata del inciso anterior pues $g(x)$ es una función cuadrática en x y cuyo valor mínimo se alcanza en $x = E(X)$.

240. a) Sea x cualquier valor de la variable aleatoria. Como $a \leq x \leq b$, multiplicando por $f(x)$ y sumando o integrando, según sea el caso, se obtiene el resultado buscado.
- b) La función $g(x) = E[(X - x)^2]$ tiene un mínimo en $x = E(X)$. Ese valor mínimo es $\text{Var}(X)$. En particular,

$$\text{Var}(X) \leq g((a+b)/2) = E[(X - (a+b)/2)^2] \leq (b - (a+b)/2)^2 = (b-a)^2/4.$$

- c) Tome X discreta tal que $P(X = a) = P(X = b) = 1/2$.

241. Únicamente se provee la validez o invalidez de la afirmación.

- | | |
|---------------|--|
| a) Verdadero. | f) Verdadero. |
| b) Falso | g) Falso, aunque puede concluirse que $P(X = 0) = 1$. |
| c) Verdadero. | h) Falso. |
| d) Falso. | i) Falso. |
| e) Falso. | |

$$242. \quad a) \quad E(\bar{X}) = E\left(\frac{1}{n} \sum_{k=1}^n X_k\right) = \frac{1}{n} \sum_{k=1}^n E(X_k) = \frac{1}{n} \sum_{k=1}^n \mu = \mu.$$

$$b) \quad \text{Var}(\bar{X}) = \text{Var}\left(\frac{1}{n} \sum_{k=1}^n X_k\right) = \frac{1}{n^2} \sum_{k=1}^n \text{Var}(X_k) = \frac{1}{n^2} \sum_{k=1}^n \sigma^2 = \sigma^2/n.$$

243.

$$\begin{aligned} \text{Var}(X + Y) &= \text{Var}(X) + \text{Var}(Y) \\ \Leftrightarrow E(X + Y)^2 - (E(X) + E(Y))^2 &= E(X^2) - E^2(X) + E(Y^2) - E^2(Y) \\ \Leftrightarrow E(X^2) + 2E(XY) + E(Y^2) - E^2(X) - 2E(X)E(Y) - E(Y^2) \\ &= E(X^2) - E^2(X) + E(Y^2) - E^2(Y) \\ \Leftrightarrow E(XY) &= E(X)E(Y). \end{aligned}$$

$$244. \quad E(X) = a/\lambda_1 + (1 - a)/\lambda_2, \quad \text{Var}(X) = a^2/\lambda_1^2 + (1 - a)^2/\lambda_2^2.$$

$$245. \quad E(X) = a, \quad \text{Var}(X) = \ell^2/3.$$

$$246. \quad E(XY) = E((Z - 1)(Z + 1)) = E(Z^2 - 1) = E(Z^2) - 1 = \text{Var}(Z) - 1 = 0.$$

247. Suponga que $\text{Var}(X) = \sum_x (x - E(X))^2 f(x) = 0$, en donde la suma se efectúa sobre los valores x tales que $f(x) > 0$. Como esta suma consta de términos no negativos, todos estos términos deben ser iguales a cero. Es decir, para cualquier valor x tal que $f(x) > 0$,

$$(x - E(X))^2 f(x) = 0.$$

Sólo existe un valor x que cumple estas dos condiciones y es $x = E(X)$. Por lo tanto, X es constante igual a $E(X)$.

248. Esta distribución surge de la expansión $\log(1+x) = x - x^2/2 + x^3/3 - x^4/4 + \dots$ para $-1 < x < 1$. De modo que

$$-\log(1 - p) = p + p^2/2 + p^3/3 + p^4/4 + \dots$$

$$a) \quad \text{Claramente } f(x) \geq 0 \text{ y } \sum_{x=1}^{\infty} f(x) = -\frac{1}{\log(1-p)} \sum_{x=1}^{\infty} p^x/x = 1.$$

$$b) \quad E(X) = -\frac{1}{\log(1-p)} \sum_{x=1}^{\infty} p^x = -\frac{1}{\log(1-p)} \frac{p}{1-p}.$$

$$\begin{aligned} c) \quad E(X^2) &= -\frac{1}{\log(1-p)} \sum_{x=1}^{\infty} x p^x = -\frac{1}{\log(1-p)} \sum_{x=1}^{\infty} \left(\sum_{y=1}^x 1\right) p^x \\ &= -\frac{1}{\log(1-p)} \sum_{y=1}^{\infty} \sum_{x=y}^{\infty} p^x = -\frac{1}{\log(1-p)} \sum_{y=1}^{\infty} \frac{p^y}{1-p} = -\frac{1}{\log(1-p)} \frac{p}{(1-p)^2}. \end{aligned}$$

- d) Esta expresión se obtiene al aplicar la fórmula $\text{Var}(X) = E(X^2) - E^2(X)$.

2.8 Momentos

249. a) $E(X^n) = \frac{1 + (-1)^{n+2}}{n + 2}$.

b) $E(X^n) = \frac{6}{(n + 2)(n + 3)}$.

c) $E(X^n) = \frac{2^{n+1}}{n + 2}$.

250. $E(X - \mu)^n = \frac{1 + (-1)^n}{(n + 1)(n + 2)}$.

251. a) $E(X^n) = \frac{b^{n+1} - a^{n+1}}{(n + 1)(b - a)}$.

b) $E(X - \mu)^n = \frac{(b - a)^n(1 + (-1)^{n+2})}{(n + 1)2^{n+1}}$.

252. a) $E(X) = a$.

b) $E(X - a)^n = 0$ para n impar.

253. La primera desigualdad es evidente. La segunda desigualdad se sigue del hecho de que $\text{Var}(X) = E(X^2) - E^2(X) \geq 0$.

254. Ambos incisos son, en general, falsos. Tome por ejemplo la variable aleatoria continua X con valores en el intervalo $(0, 1)$ y con función de densidad como aparece abajo. Se verifica que $E(X) = 1/2 > E(X^2) = 1/3 > E(X^3) = 1/4$.

$$f(x) = \begin{cases} 1 & \text{si } 0 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

255. Siendo una variable aleatoria X una función, se dice que ésta es acotada si existe una constante finita c tal que $|X| \leq c$. Por lo tanto,

$$|E(X^n)| \leq E(|X|^n) = E(|X|^n) \leq c^n.$$

256. En cada inciso utilice el método de integración por partes para resolver la integral.

2.9 Cuantiles

257. a) $c_{0.7} = 1$. $c_{0.8} = 2$ $c_{0.9} = 2$.
 b) $c_{0.7} = 0$. $c_{0.8} = 1$ $c_{0.9} = 1$.
 c) $c_{0.7} = 2$. $c_{0.8} = 3$ $c_{0.9} = 4$.

258. a) $c_{0.25} = -0.5$,
 $c_{0.50} = 0$,
 $c_{0.75} = 0.5$,
 $c_{1.00} = 1$.
 b) $c_{0.25} = 0$,
 $c_{0.50} = 0$,
 $c_{0.75} = 0$,
 $c_{1.00} = 0$.
 c) $c_{0.25} = 0.2876$,
 $c_{0.50} = 0.6931$,
 $c_{0.75} = 1.3862$,
 $c_{1.00} = \infty$.
- d) $c_{0.25} = 0.25$,
 $c_{0.50} = 0.5$,
 $c_{0.75} = 0.75$,
 $c_{1.00} = 1$.
 e) $c_{0.25} = -a/2$,
 $c_{0.50} = 0$,
 $c_{0.75} = a/2$,
 $c_{1.00} = a$.
 f) $c_{0.25} = 0.7227$,
 $c_{0.50} = 1.0471$,
 $c_{0.75} = 1.3181$,
 $c_{1.00} = \pi/2$.

259. $P(X = 2) = 0.3$, $P(X = 4) = 0.7$.

260. El número c_{p_2} es el más pequeño tal que $F(c_{p_2}) \geq p_2$. Entonces $F(c_{p_2}) \geq p_1$. Pero c_{p_1} es el número más pequeño tal que $F(c_{p_1}) \geq p_1$. Por lo tanto, $c_{p_1} \leq c_{p_2}$.

2.10 Moda

261. a) En el caso $p = 1/2$, la distribución es bimodal.

$$x^* = \begin{cases} 0 & \text{si } 0 \leq p \leq 1/2, \\ 1 & \text{si } 1/2 \leq p \leq 1. \end{cases}$$

- b) La distribución tiene dos modas: $x^* = 2, 3$.
 c) $x^* = 1$.
 d) Cada punto del intervalo $(0,1)$ es una moda.

262. Suponga que los valores de la variable aleatoria son x_1, x_2, \dots . Como $f(x)$ es una función de probabilidad, existe un número natural N tal que el valor máximo de $f(x)$ se encuentra en el conjunto $\{f(x_1), f(x_2), \dots, f(x_N)\}$. El valor x correspondiente es una moda.

2.11 Función generadora de probabilidad

263. El resultado se obtiene al derivar y evaluar en cero, término a término, la serie infinita $G(t) = \sum_{x=0}^{\infty} t^x P(X = x)$. Esta serie es convergente por lo menos para t en el intervalo $[-1, 1]$.

264. a) La f.g.p. es $G(t) = 2t(3 - t)^{-1}$, para $|t| < 3$. Calculando $G'(1-)$ y $G''(1-)$ se encuentra que $E(X) = 3/2$ y $\text{Var}(X) = 3/4$.

b) La f.g.p. es $G(t) = 3t(4 - t)^{-1}$, para $|t| < 4$. Calculando $G'(1-)$ y $G''(1-)$ se encuentra que $E(X) = 4/3$ y $\text{Var}(X) = 4/9$.

265. a) Por el lema de Abel,

$$G(1-) = \lim_{t \nearrow 1} \sum_{x=0}^{\infty} t^x P(X = x) = \sum_{x=0}^{\infty} P(X = x) = 1.$$

b) Suponiendo que la esperanza es finita, por el lema de Abel,

$$G^{(1)}(1-) = \lim_{t \nearrow 1} \sum_{x=1}^{\infty} x t^{x-1} P(X = x) = \sum_{x=1}^{\infty} x P(X = x) = E(X).$$

c) Suponiendo que el segundo momento es finito, por el lema de Abel,

$$\begin{aligned} G^{(2)}(1-) &= \lim_{t \nearrow 1} \sum_{x=2}^{\infty} x(x-1)t^{x-2} P(X = x) \\ &= \sum_{x=2}^{\infty} x(x-1) P(X = x) \\ &= E(X(X-1)). \end{aligned}$$

Por lo tanto,

$$\begin{aligned} G^{(2)}(1-) + G^{(1)}(1-) - [G^{(1)}(1-)]^2 \\ &= E(X(X-1)) + E(X) - E^2(X) \\ &= \text{Var}(X). \end{aligned}$$

$$266. G(t) = \sum_{x=0}^{n-1} t^x (1/n) = (1/n)(1 - t^n)/(1 - t).$$

2.12 Función generadora de momentos

267. Las f.g.m.s tienen las expresiones que aparecen abajo. Derive estas funciones dos veces y utilice la fórmula $M^{(n)}(0) = E(X^n)$ para hallar la esperanza y la varianza.

$$a) M(t) = \frac{e^t}{2 - e^t}, \quad t < \ln 2.$$

$$b) M(t) = \frac{2e^t}{3 - e^t}, \quad t < \ln 3.$$

268. Las f.g.m.s tienen las expresiones que aparecen abajo. Derive estas funciones dos veces y utilice la fórmula $M^{(n)}(0) = E(X^n)$ para hallar la esperanza y la varianza.

$$a) M(t) = \frac{2te^t - 2e^t + 2}{t^2}, \quad t \neq 0.$$

$$b) M(t) = \frac{1}{1 - t}, \quad t < 1.$$

$$c) M(t) = \frac{1}{1 - t^2}, \quad -1 < t < 1.$$

$$d) M(t) = \frac{e^{-t} + e^t - 2}{t^2}, \quad t \neq 0.$$

$$e) M(t) = \frac{6}{t^3} (te^t + t - 2e^t + 2), \quad t \neq 0.$$

$$f) M(t) = \frac{1}{4t} (e^{4t} - e^{2t}) + \frac{1}{2t} (e^{5t} - e^{4t}), \quad t \neq 0.$$

269. $M_{aX+b}(t) = E(e^{t(aX+b)}) = e^{bt} E(e^{atX}) = e^{bt} M_X(at)$.

270. Para cualquier $t > 0$,

$$\begin{aligned} M(t) &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\pi(1+x^2)} dx \\ &\geq \int_0^{\infty} e^{tx} \frac{1}{\pi(1+x^2)} dx \\ &= \infty, \end{aligned}$$

pues el integrando $e^{tx} \frac{1}{\pi(1+x^2)}$ tiende a infinito cuando $x \rightarrow \infty$. Un análisis similar puede llevarse a cabo cuando $t < 0$. Así, la integral es infinita para cualquier $t \neq 0$.

Capítulo 3: Distribuciones de probabilidad

3.1 Distribución uniforme discreta

271. a) $P(X \leq x) = x/n$.
 b) $P(X \geq x) = (n - x + 1)/n$.
 c) $P(x_1 \leq X \leq x_2) = (x_2 - x_1 + 1)/n$.
 d) $P(x_1 < X \leq x_2) = (x_2 - x_1)/n$.
 e) $P(x_1 \leq X < x_2) = (x_2 - x_1)/n$.
 f) $P(x_1 < X < x_2) = (x_2 - x_1 - 1)/n$.

272. Tanto X^3 como $-X$ toman los mismos valores que X y lo hacen con las mismas probabilidades. Esto es, para $x = -1, 0, 1$,

$$P(X^3 = x) = P(-X = x) = P(X = x) = 1/3.$$

273. a) $E(X) = \sum_{x=1}^n x \frac{1}{n} = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}$.
 b) $E(X^2) = \sum_{x=1}^n x^2 \frac{1}{n} = \frac{1}{n} \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$.
 c) $\text{Var}(X) = E(X^2) - E^2(X) = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n^2-1}{12}$.

274. $E(X^n) = 0^n P(X=0) + 1^n P(X=1) = 1/2$.

275. $c_{0.25} = n$, $c_{0.5} = 2n$, $c_{0.75} = 3n$, $c_1 = 4n$.

276. Usando la fórmula para sumas geométricas, la f.g.p. de esta distribución se puede calcular como se muestra abajo. Derive esta función dos veces y utilice las fórmulas $G^{(1)}(1-) = E(X)$ y $G^{(2)}(1) = E(X(X-1))$ para hallar la esperanza y la varianza.

$$G(t) = \sum_{x=1}^n t^x \frac{1}{n} = \frac{1}{n} \frac{t - t^{n+1}}{1-t} = \frac{t(1-t^n)}{n(1-t)}.$$

277. La f.g.m. se calcula como aparece abajo, usando la fórmula para sumas geométricas. Derive esta función dos veces y utilice la fórmula $M^{(n)}(0) = E(X^n)$ para hallar la esperanza y la varianza.

$$M(t) = \sum_{x=1}^n e^{tx} \frac{1}{n} = \frac{1}{n} \sum_{x=1}^n (e^t)^x = \frac{e^t(1-e^{nt})}{n(1-e^t)}.$$

278. Sea U una v.a. con distribución unif(0, 1). Entonces para $i = 1, \dots, n$,

$$P(X = x_i) = P((i-1)/n < U \leq i/n) = 1/n.$$

279. Sean X y Y los valores al azar a y b , respectivamente. Entonces

$$\begin{array}{ll} a) P(X = Y) = 10/100. & c) P(X > Y + 1) = 36/100. \\ b) P(X < Y) = 45/100. & d) P(|X - Y| \geq 2) = 72/100. \end{array}$$

$$280. \quad \begin{array}{ll} a) (24/36)^5. & c) \binom{n}{k} (12/36)^k (24/36)^{n-k}. \\ b) (24/36)^3 (12/36). & \end{array}$$

281. $3/5$.

$$282. \quad a) f_U(u) = \begin{cases} 1/n & \text{si } u = 1, \dots, m-1; m \neq 1, \\ (n-m+1)/n & \text{si } u = m, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f_V(v) = \begin{cases} m/n & \text{si } v = m, \\ 1/n & \text{si } v = m+1, \dots, n; m \neq n, \\ 0 & \text{en otro caso.} \end{cases}$$

3.2 Distribución Bernoulli

283. a) Ber(p). b) Ber($1-p$). c) Ber($1-p$).

$$284. \quad a) f(y) = \begin{cases} p^{(y-b)/a} (1-p)^{1-(y-b)/a} & \text{si } y = b, a+b, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) E(Y) = ap + b.$$

$$c) \text{Var}(X) = a^2 p(1-p).$$

$$d) E(Y^n) = b^n(1-p) + (a+b)^n p, \quad n \geq 1.$$

285. $E(X^n) = 0^n P(X=0) + 1^n P(X=1) = p$.

286. El máximo de la función $g(p) = p(1-p)$ se alcanza en $p = 1/2$.

287. $c_{0.25} = 0$, $c_{0.5} = 0$, $c_{0.75} = 1$, $c_1 = 1$.

288. La f.g.p. se calcula como aparece abajo. Derive esta función dos veces y utilice las fórmulas $G^{(1)}(1-) = E(X)$ y $G^{(2)}(1-) = E(X(X-1))$ para hallar la esperanza y la varianza.

$$G(t) = E(t^X) = t^0 P(X=0) + t^1 P(X=1) = 1-p+pt.$$

289. La f.g.m. se calcula como aparece abajo. Derive esta función dos veces y utilice la fórmula $M^{(n)}(0) = E(X^n)$ para hallar la esperanza y la varianza.

$$M(t) = E(e^{tX}) = e^{t \cdot 0}P(X = 0) + e^{t \cdot 1}P(X = 1) = 1 - p + pe^t.$$

290. Claramente $P(X = 0) = 1 - p$ y $P(X = 1) = p$.

$$291. \quad a) \quad f_{X+Y}(u) = \begin{cases} (1-p)^2 & \text{si } u = 0, \\ 2p(1-p) & \text{si } u = 1, \\ p^2 & \text{si } u = 2, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) \quad f_{X-Y}(u) = \begin{cases} p(1-p) & \text{si } u = -1, \\ (1-p)^2 + p^2 & \text{si } u = 0, \\ p(1-p) & \text{si } u = 1, \\ 0 & \text{en otro caso.} \end{cases}$$

c) $XY \sim \text{Ber}(p^2)$.

d) $X(1-Y) \sim \text{Ber}(p(1-p))$.

e) La variable $X(1-X)$ es constante igual a cero.

$$f) \quad f_{X+Y-1}(u) = \begin{cases} (1-p)^2 & \text{si } u = -1, \\ 2p(1-p) & \text{si } u = 0, \\ p^2 & \text{si } u = 1, \\ 0 & \text{en otro caso.} \end{cases}$$

292. La variable aleatoria producto $X_1 \cdots X_n$ sólo toma los valores cero y uno. Sus probabilidades son $P(X_1 \cdots X_n = 1) = P(X_1 = 1) \cdots P(X_n = 1) = p^n$ y, por complemento, $P(X_1 \cdots X_n = 0) = 1 - p^n$. Por lo tanto, $X_1 \cdots X_n \sim \text{Ber}(p^n)$.

3.3 Distribución binomial

293. Claramente $f(x)$ es no negativa. Por otro lado, usando el teorema del binomio,

$$\sum_{x=0}^n \binom{n}{x} p^x (1-p)^{n-x} = (p + (1-p))^n = 1.$$

294. a) $n = 12, \quad p = 1/2$.

b) $n = 24, \quad p = 1/2$.

295. La fórmula iterativa se obtiene como se muestra abajo. Los otros resultados se derivan de un análisis del factor que relaciona $f(x+1)$ con $f(x)$:

$$\begin{aligned} f(x+1) &= \binom{n}{x+1} p^{x+1} (1-p)^{n-(x+1)} \\ &= \frac{p}{1-p} \frac{n-x}{x+1} \binom{n}{x} p^x (1-p)^{n-x} \\ &= \frac{p}{1-p} \frac{n-x}{x+1} f(x). \end{aligned}$$

296. Se muestra únicamente el cálculo de la esperanza. El segundo momento puede calcularse de manera similar, escribiendo $x^2 = x(x-1) + x$.

$$\begin{aligned} E(X) &= \sum_{x=1}^n x \binom{n}{x} p^x (1-p)^{n-x} \\ &= np \sum_{x=1}^n \binom{n-1}{x-1} p^{x-1} (1-p)^{(n-1)-(x-1)} \\ &= np. \end{aligned}$$

297. La función $g(x) = np(1-p)$ tiene un máximo global en $p = 1/2$.

298. La f.g.p. se calcula como se muestra abajo, usando el teorema del binomio. Derive esta función dos veces y utilice las fórmulas $G^{(1)}(1-) = E(X)$ y $G^{(2)}(1-) = E(X(X-1))$ para hallar la esperanza y la varianza.

$$\begin{aligned} G(t) &= \sum_{x=0}^n t^x \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=0}^n \binom{n}{x} (pt)^x (1-p)^{n-x} \\ &= (1-p + pt)^n. \end{aligned}$$

299. La f.g.m. se calcula como se muestra abajo usando nuevamente el teorema del binomio. Derive esta función dos veces y utilice la fórmula $M^{(n)}(0) = E(X^n)$ para hallar la esperanza y la varianza.

$$\begin{aligned} M(t) &= \sum_{x=0}^n e^{tx} \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=0}^n \binom{n}{x} (pe^t)^x (1-p)^{n-x} \\ &= (1-p + pe^t)^n. \end{aligned}$$

300. Haremos uso de la f.g.m. El procedimiento para usar la f.g.p. es completamente análogo. Por la hipótesis de independencia,

$$M_{X_1+\dots+X_n}(t) = M_{X_1}(t) \cdots M_{X_n}(t) = (1-p+pe^t) \cdots (1-p+pe^t) = (1-p+pe^t)^n.$$

301. a) Sea u cualquier valor en el conjunto $\{0, 1, \dots, n+m\}$. Las siguientes sumas se efectúan para todos los valores enteros x y y tales que $0 \leq x \leq n$, $0 \leq y \leq m$ y $x+y = u$. Por la hipótesis de independencia,

$$\begin{aligned} P(X+Y = u) &= \sum_{x,y} P(X=x) P(Y=y) \\ &= \sum_{x,y} \binom{n}{x} p^x (1-p)^{n-x} \binom{m}{y} p^y (1-p)^{m-y} \\ &= p^u (1-p)^{n+m-u} \sum_{x,y} \binom{n}{x} \binom{m}{y} \\ &= \binom{n+m}{u} p^u (1-p)^{n+m-u}, \end{aligned}$$

en donde la última identidad se obtiene a partir del resultado del Ejercicio 83 en la página 67.

$$b) G_{X+Y}(t) = G_X(t) G_Y(t) = (1-p+pt)^n (1-p+pt)^m = (1-p+pt)^{n+m}.$$

$$c) M_{X+Y}(t) = M_X(t) M_Y(t) = (1-p+pe^t)^n (1-p+pe^t)^m = (1-p+pe^t)^{n+m}.$$

302. Para $x = 0, 1, \dots, n$,

$$P(n-X = x) = P(X = n-x) = \binom{n}{n-x} p^{n-x} (1-p)^x = \binom{n}{x} (1-p)^x p^{n-x}.$$

303.

$$\begin{aligned} \lim_{n \rightarrow \infty} P(X_n > k) &= 1 - \lim_{n \rightarrow \infty} P(X_n \leq k) \\ &= 1 - \lim_{n \rightarrow \infty} \sum_{x=0}^k \binom{n}{x} p^x (1-p)^{n-x} \\ &= 1 - \sum_{x=0}^k \frac{1}{x!} p^x (1-p)^{-x} \lim_{n \rightarrow \infty} \frac{n!}{(n-x)!} (1-p)^n. \end{aligned}$$

El límite indicado es cero pues

$$\begin{aligned} \frac{n!}{(n-x)!}(1-p)^n &= (n-x+1) \cdots n (1-p)^n \\ &\leq n^x (1-p)^n \\ &= e^{x \ln n + n \ln(1-p)} \\ &= e^x n e^{n \ln(1-p)} \\ &\rightarrow 0, \quad \text{pues } \ln(1-p) < 0. \end{aligned}$$

304. Debe observarse que, conforme m crece, s_m tiende a la media de la distribución $\text{bin}(n, p)$, es decir, al valor np .
305. Sea X el número de semillas que germinarán en un paquete de 20 semillas. Entonces X tiene distribución $\text{bin}(n, p)$ con $n = 20$ y $p = 0.9$. Se pide encontrar $P(X \leq 17)$. Esta probabilidad es

$$P(X \leq 17) = 1 - P(X \geq 18) = 1 - \sum_{x=18}^{20} \binom{20}{x} (0.9)^x (0.1)^{20-x} = 0.3230.$$

Por lo tanto, el 32.30 % de los paquetes no cumplirán la garantía.

306. Sea X la v.a. que registra el número de asegurados que presentan el accidente en cuestión en el periodo de un año. Entonces X tiene distribución $\text{bin}(n, p)$ con $n = 10$ y $p = 0.15$. Se pide encontrar $P(X > 5)$. Esta probabilidad es

$$\sum_{x=6}^{10} \binom{10}{x} (0.15)^x (0.85)^{10-x} = 0.00138.$$

307. Sea X la v.a. que registra el número de unos que se obtienen en este experimento. Por lo tanto, X tiene distribución $\text{bin}(10, 1/6)$ y la información es que $X \geq k$. Entonces

$$a) P(X = k | X \geq k) = \binom{10}{k} (1/6)^k (5/6)^{10-k} / \sum_{j=k}^{10} \binom{10}{j} (1/6)^j (5/6)^{10-j}.$$

$$b) P(X \geq k+1 | X \geq k) = \frac{\sum_{j=k+1}^{10} \binom{10}{j} (1/6)^j (5/6)^{10-j}}{\sum_{j=k}^{10} \binom{10}{j} (1/6)^j (5/6)^{10-j}}.$$

$$c) P(X \leq k+1 | X \geq k) = \frac{\binom{10}{k} (1/6)^k (5/6)^{10-k} + \binom{10}{k+1} (1/6)^{k+1} (5/6)^{10-k-1}}{\sum_{j=k}^{10} \binom{10}{j} (1/6)^j (5/6)^{10-j}}.$$

$$d) P(X = 10 | X \geq k) = (1/6)^{10} / \sum_{j=k}^{10} \binom{10}{j} (1/6)^j (5/6)^{10-j}.$$

308. a) En cada oportunidad, la bola puede ir a la izquierda o a la derecha y, convenientemente, se puede asignar el valor 0 y el valor 1 a cada una de estas situaciones. El número total de trayectorias distintas es entonces 2^5 , cada una con probabilidad $(1/2)^5$.
- b) El número de trayectorias que llegan a la casilla x es $\binom{5}{x}$ para $x = 0, 1, \dots, 5$.
- c) Se trata de la distribución $\text{bin}(5, 1/2)$. La probabilidad de que la bola caiga en la urna x es $\binom{5}{x}(1/2)^5$.
- d) El número total de trayectorias es 2^n y esta vez no todas ellas tienen la misma probabilidad de ocurrir. El número de trayectorias que llegan a la casilla x es $\binom{n}{x}$ para $x = 0, 1, \dots, n$. La probabilidad de que la bola caiga en la urna x es $\binom{n}{x}p^x(1-p)^{n-x}$. Esta es la distribución $\text{bin}(n, p)$.

3.4 Distribución geométrica

309. Para demostrar que la suma de las probabilidades es uno, utilice la fórmula para sumas geométricas que aparece en el apéndice. Por otro lado, es inmediato comprobar que para $x \geq 0$ entero, $f(x) \geq f(x+1)$.
310. La esperanza se calcula como aparece abajo. Para la varianza, calcule primero el segundo momento y use una técnica similar: exprese $x^2 = 2(x(x+1)/2 - x/2)$ en donde $x(x+1)/2 = \sum_{y=1}^x y$.

$$\begin{aligned}
 E(X) &= \sum_{x=1}^{\infty} x(1-p)^x p \\
 &= \sum_{x=1}^{\infty} \left(\sum_{y=1}^x 1 \right) (1-p)^x p \\
 &= \sum_{y=1}^{\infty} \sum_{x=y}^{\infty} (1-p)^x p \\
 &= \sum_{y=1}^{\infty} (1-p)^y \\
 &= (1-p)/p.
 \end{aligned}$$

311. No es difícil demostrar que para $x \geq 0$ entero, $F(x) = 1 - (1-p)^{x+1}$. Por lo tanto, $1 - F(x) = (1-p)^{x+1}$ y, en consecuencia,

$$\sum_{x=0}^{\infty} (1 - F(x)) = \sum_{x=0}^{\infty} (1-p)^{x+1} = (1-p)/p.$$

312. Para cada entero $n \geq 0$, el evento $(X = n)$ es idéntico al evento $(X_0 = 0, X_1 = 0, \dots, X_{n-1} = 0, X_n = 1)$. Por lo tanto,

$$\begin{aligned} P(X = n) &= P(X_0 = 0, X_1 = 0, \dots, X_{n-1} = 0, X_n = 1) \\ &= P(X_0 = 0)P(X_1 = 0) \cdots P(X_{n-1} = 0)P(X_n = 1) \\ &= (1-p)^{n-1}p. \end{aligned}$$

313. La f.g.p. se calcula como aparece abajo. Derive esta función dos veces y utilice las fórmulas $G^{(1)}(1-) = E(X)$ y $G^{(2)}(1-) = E(X(X-1))$ para hallar la esperanza y la varianza. Para $|t| < 1/(1-p)$,

$$G(t) = E(t^X) = \sum_{x=0}^{\infty} t^x (1-p)^x p = p \sum_{x=0}^{\infty} (t(1-p))^x = p/(1 - (1-p)t).$$

314. La f.g.m. se calcula como aparece abajo. Derive esta función dos veces y utilice la fórmula $M^{(n)}(0) = E(X^n)$ para hallar la esperanza y la varianza. Para $|t| < -\ln(1-p)$,

$$M(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx} (1-p)^x p = p \sum_{x=0}^{\infty} (e^t(1-p))^x = p/(1 - (1-p)e^t).$$

315. Aplique la definición de probabilidad condicional y observe que se cumple la contención de eventos que aparece abajo. Substituya las expresiones de las probabilidades resultantes.

$$(X \geq n+m) \subseteq (X \geq m).$$

316. Use las expresiones de la esperanza y la función de probabilidad geométrica para verificar que la suma de las siguientes probabilidades es uno.

$$\begin{aligned} P(X+Y=k) &= \sum_{x=0}^k P(X=x, Y=k-x) \\ &= \sum_{x=0}^k P(X=x)P(Y=k-x) \\ &= \sum_{x=0}^k (1-p)^x p (1-p)^{k-x} p \\ &= (k+1)(1-p)^k p^2. \end{aligned}$$

317. $p_1 = 2/3$, $p_2 = 1/3$.

318. Sea p la probabilidad de obtener “cara”. Para $x \geq 10$ entero,

$$P(X = x) = \binom{x-1}{9} (1-p)^{x-10} p^{10}.$$

$$319. P(Y = y) = \begin{cases} (1-p)^y p & \text{si } y = 0, 1, \dots, n-1, \\ (1-p)^n & \text{si } y = n, \\ 0 & \text{en otro caso.} \end{cases}$$

320. a) La variable $Y = 1 + X$ toma los valores $1, 2, \dots$ con probabilidades

$$f_Y(y) = P(Y = y) = P(1 + X = y) = P(X = y - 1) = (1-p)^{y-1} p.$$

b) Para $y = 1, 2, \dots$

$$F_Y(y) = P(Y \leq y) = P(1 + X \leq y) = P(X \leq y - 1) = 1 - (1-p)^y.$$

c) $E(Y) = E(1 + X) = 1 + E(X) = 1 + (1-p)/p = 1/p$.

d) $\text{Var}(Y) = \text{Var}(1 + X) = \text{Var}(X) = (1-p)/p^2$.

321. Use la fórmula para sumas geométricas para demostrar que la suma de los valores de $f(x)$ es uno. Por otro lado, $P(X > 10) = (1-p)^{10}$.

3.5 Distribución binomial negativa

322. Este no es un ejercicio sencillo. Se puede usar la identidad (3.3), que aparece en la página 229 y la siguiente expansión válida para $|t| < 1$ y cualquier número real a ,

$$(1+t)^a = \sum_{x=0}^{\infty} \binom{a}{x} t^x.$$

Claramente $f(x) \geq 0$ y

$$\begin{aligned} \sum_{x=0}^{\infty} \binom{r+x-1}{x} p^r (1-p)^x &= p^r \sum_{x=0}^{\infty} (-1)^x \binom{-r}{x} (1-p)^x \\ &= p^r \sum_{x=0}^{\infty} \binom{-r}{x} (p-1)^x \\ &= p^r (1+p-1)^{-r} = 1. \end{aligned}$$

323. Aplique la definición $\binom{a}{x} := a(a-1)\cdots(a-x+1)/x!$, para cualquier $a \in \mathbb{R}$, en la expresión del lado derecho.

324. La fórmula iterativa para la función de probabilidad es sencilla de verificar y de ella se desprenden el resto de las afirmaciones.
325. La variable X puede tomar los valores $0, 1, 2, \dots$. Para cualquiera de estos valores,

$$\begin{aligned} P(X = x) &= P(\text{mín}\{n \geq r : \sum_{k=1}^n X_k = r\} = r + x) \\ &= P(X_{r+x} = 1 \text{ y en } (X_1, X_2, \dots, X_{r+x-1}) \text{ hay } r - 1 \text{ unos}) \\ &= \binom{r+x-1}{r-1} (1-p)^x p^r. \end{aligned}$$

326. La f.g.p. se calcula como aparece abajo. Derive esta función dos veces y utilice las fórmulas $G^{(1)}(1-) = E(X)$ y $G^{(2)}(1-) = E(X(X-1))$ para hallar la esperanza y la varianza.

$$\begin{aligned} G(t) &= \sum_{x=0}^{\infty} t^x \binom{x+r-1}{r-1} (1-p)^x p^r \\ &= \sum_{x=0}^{\infty} \binom{x+r-1}{r-1} ((1-p)t)^x p^r \\ &= \frac{p^r}{(1-(1-p)t)^r} \sum_{x=0}^{\infty} \binom{x+r-1}{r-1} (1-(1-p)t)^x (1-(1-p)t)^r \\ &= (p/(1-(1-p)t))^r \quad \text{si } |t| < 1/(1-p). \end{aligned}$$

327. La f.g.m. se calcula como aparece abajo. Derive esta función dos veces y utilice la fórmula $M^{(n)}(0) = E(X^n)$ para hallar la esperanza y la varianza.

$$\begin{aligned} M(t) &= \sum_{x=0}^{\infty} e^{tx} \binom{x+r-1}{r-1} (1-p)^x p^r \\ &= \sum_{x=0}^{\infty} \binom{x+r-1}{r-1} ((1-p)e^t)^x p^r \\ &= \frac{p^r}{(1-(1-p)e^t)^r} \sum_{x=0}^{\infty} \binom{x+r-1}{r-1} (1-(1-p)e^t)^x (1-(1-p)e^t)^r \\ &= (p/(1-(1-p)e^t))^r \quad \text{si } |t| < -\ln(1-p). \end{aligned}$$

328. Haremos uso de la f.g.m. El procedimiento para usar la f.g.p. es completa-

mente análogo. Por la hipótesis de independencia,

$$\begin{aligned} M_{X_1+\dots+X_r}(t) &= M_{X_1}(t) \cdots M_{X_r}(t) \\ &= (p/(1 - (1 - p)e^t)) \cdots (p/(1 - (1 - p)e^t)) \\ &= (p/(1 - (1 - p)e^t))^r. \end{aligned}$$

329. a) Utilice la expresión $P(X + Y = k) = \sum_{u=0}^k P(X = u)P(Y = k - u)$. Substituya las probabilidades que aparecen dentro de la suma y lleve a cabo la suma para encontrar la función de probabilidad bin neg($r + s, p$).
- b) Por la hipótesis de independencia, $G_{X+Y}(t) = G_X(t) \cdot G_Y(t)$. Substituya las dos expresiones del lado derecho y simplifique encontrando la f.g.p. de la distribución bin neg($r + s, p$).
- c) Por la hipótesis de independencia, $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$. Substituya las dos expresiones del lado derecho y simplifique encontrando la f.g.m. de la distribución bin neg($r + s, p$).
330. El número de personas consultadas hasta obtener 20 que cumplan las características es la v.a. $X + 20$, en donde X tiene distribución bin neg(r, p) con $r = 20$ y $p = 1/100$. Entonces $E(X + 20) = E(X) + 20 = r((1 - p)/p) + 20 = 200$. Esta es una aproximación a la respuesta buscada cuando el tamaño de la población es grande.

331. Prob. = $\binom{n-1}{5} (1/6)^6 (5/6)^{n-6}$, para $n = 6, 7, \dots$

332. a) La variable $Y = r + X$ toma los valores $y = r, r + 1, \dots$ con probabilidades

$$f_Y(y) = P(Y = y) = P(r + X = y) = P(X = y - r) = \binom{y-1}{y-r} (1-p)^{y-r} p^r.$$

b) $E(Y) = E(r + X) = r + E(X) = r + r(1 - p)/p = r/p$.

c) $\text{Var}(Y) = \text{Var}(r + X) = \text{Var}(X) = r(1 - p)/p^2$.

333. a) La probabilidad es $\binom{n-1}{k-1} (1/2)^{n-1}$ para $n = k, k + 1, \dots, 2k - 1$.

- b) Puede procederse por inducción sobre k . Sea $g(k)$ la suma de las probabilidades, es decir,

$$g(k) = \sum_{n=k}^{2k-1} \binom{n-1}{k-1} (1/2)^{n-1}.$$

Se busca demostrar que $g(k)$ es constante igual a uno. Se puede verificar directamente que $g(1)$, $g(2)$ y $g(3)$ es, efectivamente, uno. Usando

la fórmula general $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$ puede calcularse $g(k+1)$ en términos de $g(k)$ y resulta que ambas cantidades son idénticas.

$$\begin{aligned} g(k+1) &= \sum_{n=k+1}^{2k+1} \binom{n-1}{k} (1/2)^{n-1} \\ &= (1/2)^k + \sum_{n=k+2}^{2k+1} \left[\binom{n-2}{k-1} + \binom{n-2}{k} \right] (1/2)^{n-1} \\ &= (1/2)^k + (1/2) \sum_{m=k+1}^{2k} \left[\binom{m-1}{k-1} + \binom{m-1}{k} \right] (1/2)^{m-1}. \end{aligned}$$

Separe las sumas e identifique los términos $g(k)$ y $g(k+1)$, añadiendo o restando los sumandos que sean necesarios. Esto lleva a una ecuación entre $g(k)$ y $g(k+1)$ que indica que son iguales.

- c) La probabilidad es $\binom{n-1}{k-1} [p^k(1-p)^{n-k} + (1-p)^k p^{n-k}]$ para $n = k, k+1, \dots, 2k-1$. Cuando $p = 1/2$, esta expresión se reduce a la del primer inciso.
- d) Proceda como en el segundo inciso. La escritura es más elaborada, pues en lugar del término $(1/2)^{n-1}$ ahora aparece $p^k(1-p)^{n-k} + (1-p)^k p^{n-k}$, pero los cálculos son semejantes.

334. Mediante el cambio de variable $n = x + r$, el problema se reduce al del ejercicio anterior.

$$F_X(r-1) = \sum_{x=0}^{r-1} \binom{r+x-1}{x} (1/2)^{r+x} = (1/2) \sum_{n=r}^{2r-1} \binom{n-1}{r-1} (1/2)^{n-1} = 1/2.$$

3.6 Distribución hipergeométrica

335. Iguale los coeficientes de la identidad $(a+b)^N = (a+b)^K (a+b)^{N-K}$.

336. Para el inciso (a),

$$\begin{aligned} E(X) &= \sum_{x=1}^n x \binom{K}{x} \binom{N-K}{n-x} / \binom{N}{n} \\ &= n \frac{K}{N} \sum_{x=0}^{n-1} \binom{K-1}{x} \binom{(N-1)-(K-1)}{(n-1)-x} / \binom{N-1}{n-1} \\ &= n \frac{K}{N}. \end{aligned}$$

Para el inciso (b), exprese x^2 como $x(x-1) + x$ y proceda como en el primer inciso.

337. Cuando $n = 1$ la v.a. X sólo puede tomar los valores 0 y 1. La función de probabilidad se reduce a $f(0) = (N - K)/N$ y $f(1) = K/N$.
338. La fórmula iterativa para la función de probabilidad es sencilla de verificar, y de ella se desprenden el resto de las afirmaciones.
339. Se descompone la expresión de la función de probabilidad como se muestra a continuación. En particular, se escribe $N! = (N-x)!N(N-1)\cdots(N-x+1)$.

$$\begin{aligned} f(x) &= \binom{K}{x} \binom{N-K}{n-x} / \binom{N}{n} \\ &= \frac{K!(N-K)!n!(N-n)!}{x!(K-x)!(n-x)!(N-K-n+x)!N!} \\ &= \binom{n}{x} \frac{K!(N-K)!}{N!} \frac{(N-n)!}{(K-x)!(N-K-n+x)!} \\ &= \binom{n}{x} \frac{K(K-1)\cdots(K-x+1)}{N(N-1)\cdots(N-x+1)} \\ &\quad \frac{(N-K)(N-K-1)\cdots(N-K-n+x+1)}{(N-x)(N-x-1)\cdots(N-n+1)}. \end{aligned}$$

El primer cociente tiende a p^x y el segundo a $(1-p)^{n-x}$.

340. hipergeo($N_1 + N_2 + N_3, N_1, n$).

341. 0.9231433.

3.7 Distribución Poisson

342. Recuerde la expansión $e^x = \sum_{k=0}^{\infty} x^k/k!$

343. Para el inciso (a) tenemos que

$$E(X) = \sum_{x=1}^{\infty} x e^{-\lambda} \frac{\lambda^x}{x!} = \lambda \sum_{x=1}^{\infty} e^{-\lambda} \frac{\lambda^{x-1}}{(x-1)!} = \lambda.$$

Para el inciso (b), calcule primero el segundo momento haciendo uso de la expresión $x^2 = x(x-1) + x$.

344. $\lambda = (n!/m!)^{1/(n-m)}$.

345. La fórmula iterativa para la función de probabilidad es sencilla de verificar, y de ella se desprende el resto de las afirmaciones.
346. Use la fórmula de Stirling: $n! \approx \sqrt{2\pi} n^{n+1/2} e^{-n}$, n grande.
347. La f.g.p. se calcule como aparece abajo. Derive esta función dos veces y utilice las fórmulas $G^{(1)}(1-) = E(X)$ y $G^{(2)}(1-) = E(X(X-1))$ para hallar la esperanza y la varianza.

$$G(t) = \sum_{x=0}^{\infty} t^x e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda t)^x}{x!} = e^{\lambda(t-1)}.$$

348.

$$\begin{aligned} E(X^n) &= \sum_{x=1}^{\infty} x^n e^{-\lambda} \frac{\lambda^x}{x!} \\ &= \lambda \sum_{x=1}^{\infty} x^{n-1} e^{-\lambda} \frac{\lambda^{x-1}}{(x-1)!} \\ &= \lambda \sum_{x=0}^{\infty} (1+x)^{n-1} e^{-\lambda} \frac{\lambda^x}{x!} \\ &= \lambda \sum_{k=0}^{n-1} \binom{n-1}{k} \sum_{x=0}^{\infty} x^k e^{-\lambda} \frac{\lambda^x}{x!} \\ &= \lambda \sum_{k=0}^{n-1} \binom{n-1}{k} E(X^k). \end{aligned}$$

349. La f.g.m. se calcule como aparece abajo. Derive esta función dos veces y utilice la fórmula $M^{(n)}(0) = E(X^n)$ para hallar la esperanza y la varianza.

$$M(t) = \sum_{x=0}^{\infty} e^{tx} e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} = e^{\lambda(e^t-1)}.$$

350. El resultado se sigue de la hipótesis de independencia y el teorema de caracterización que aparece en la página 199.

$$M_{X+Y}(t) = M_X(t) M_Y(t) = e^{\lambda_1(e^t-1)} e^{\lambda_2(e^t-1)} = e^{(\lambda_1+\lambda_2)(e^t-1)}.$$

351. Condicione sobre el valor de N .

$$352. E(X!) = \sum_{x=0}^{\infty} x! e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \lambda^x = e^{-\lambda}/(1-\lambda), \quad \text{si } |\lambda| < 1.$$

353. Si X denota el número de errores por página, entonces

- a) $P(X = 0) = e^{-1}$.
- b) $P(X = 2) = (1/2)e^{-1}$.
- c) $P(X \geq 3) = 1 - (5/2)e^{-1}$.

354. Si X denota el número de semillas por naranja, entonces

- a) $P(X = 0) = e^{-3}$.
- b) $P(X \geq 2) = 1 - 4e^{-3}$.
- c) $P(X \leq 3) = 13e^{-3}$.

355. Si X denota el número de accidentes al día, entonces

- a) $P(X \geq 2) = 1 - 4e^{-3}$.
- b) $P(X \geq 2 | X \geq 1) = (1 - 4e^{-3}) / (1 - e^{-3})$.

3.8 Distribución uniforme continua

356. a) $c = 1/2$. c) $c = 1$.
 b) $c = 1$. d) $c = 3$.

357. $E(X) = \int_a^b x/(b-a) dx = (a+b)/2$. De manera análoga, calcule el segundo momento y use estas dos expresiones para calcular la varianza.

358. $E(X^n) = \int_{-1}^1 x^n/2 dx = x^{n+1}/(2(n+1))|_{-1}^1$. Al hacer estas evaluaciones se encuentra que esta cantidad se anula cuando n es impar y es $1/(n+1)$ cuando n es par.

359. $E(X^n) = \int_a^b x^n/(b-a) dx = x^{n+1}/((b-a)(n+1))|_a^b = (b^{n+1} - a^{n+1})/((b-a)(n+1))$.

360. $c_p = a + (b-a)p$.

361. La f.g.m. se calcula como aparece abajo. Derive esta función dos veces y utilice la fórmula $M^{(n)}(0) = E(X^n)$ para hallar la esperanza y la varianza.

$$M(t) = \int_a^b e^{tx} \frac{1}{b-a} dx = \frac{1}{b-a} \frac{1}{t} e^{tx} \Big|_a^b = \frac{e^{bt} - e^{at}}{t(b-a)}.$$

362. Como U toma valores en $(0, 1)$, la variable $X = a + (b-a)U$ toma valores en (a, b) . Así, para $x \in (a, b)$,

$$P(X \leq x) = P(a + (b-a)U \leq x) = P(U \leq (x-a)/(b-a)) = (x-a)/(b-a).$$

363. a)

$$\begin{aligned}
 P(|X - 1/2| > \epsilon) &= 1 - F(1/2 + \epsilon) + F(1/2 - \epsilon) \\
 &= (1 - 2\epsilon)^+ \\
 &= \begin{cases} 1 - 2\epsilon & \text{si } \epsilon < 1/2, \\ 0 & \text{si } \epsilon \geq 1/2. \end{cases}
 \end{aligned}$$

b)

$$\begin{aligned}
 P((2X - 1/2)^2 \leq \epsilon) &= F(1/4 + \sqrt{\epsilon}/2) - F(1/4 - \sqrt{\epsilon}/2) \\
 &= \begin{cases} \sqrt{\epsilon} & \text{si } \epsilon < 1/4, \\ 1/4 + \sqrt{\epsilon}/2 & \text{si } 1/4 \leq \epsilon < 9/4, \\ 1 & \text{si } \epsilon \geq 9/4. \end{cases}
 \end{aligned}$$

364. Si $\alpha = 0$, la variable Y es constante igual a uno. Para $\alpha \neq 0$, la función de densidad de Y es

$$f(y) = \begin{cases} (1/\alpha)y^{(1-\alpha)/\alpha} & \text{si } 0 < y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

365. a) $Y \sim \text{unif}(0, 2)$.b) $Y \sim \text{unif}(0, 1)$.

$$c) F_Y(y) = \begin{cases} 0 & \text{si } y \leq 0, \\ \sqrt{y} & \text{si } 0 < y < 1, \\ 1 & \text{si } y \geq 1. \end{cases}$$

$$d) F_Y(y) = \begin{cases} 0 & \text{si } y \leq -1, \\ (\sqrt[3]{y} + 1)/2 & \text{si } -1 < y < 1, \\ 1 & \text{si } y \geq 1. \end{cases}$$

366. a) $P(X^2 < 1/4) = 1/2$.b) $P(|X + 1| \leq |X - 1|) = 1/2$.

367. a) Sea x un valor de X . Entonces $a \leq x \leq b$. Por lo tanto, $af(x) \leq xf(x) \leq bf(x)$. Ahora se puede sumar o integrar sobre todos los posibles valores x para así obtener el resultado.

b) Sea x un valor de X y sea μ su media. Por el inciso anterior, también μ se encuentra entre a y b . Puede comprobarse que $(x - \mu)^2 \leq (b - a)^2$. Por lo tanto, $(x - \mu)^2 f(x) \leq (b - a)^2 f(x)$, y nuevamente sumando o integrando sobre x se obtiene la segunda desigualdad.

368. $a = \alpha - \sqrt{3\beta}$, $b = \alpha + \sqrt{3\beta}$.

369. Se lleva X a una distribución $\text{unif}(0, 1)$ y después ésta a una distribución $\text{unif}(c, d)$. Como $X \sim \text{unif}(a, b)$, $(X - a)/(b - a) \sim \text{unif}(0, 1)$. Entonces $Y = c + (d - c)(X - a)/(b - a) \sim \text{unif}(c, d)$. De aquí se obtiene que para $y \in (c, d)$:

a) $F_Y(y) = F_X(a + (b - a)(y - c)/(d - c))$.

b) $f_Y(y) = f_X(a + (b - a)(y - c)/(d - c)) (b - a)/(d - c)$.

370. Mediante el siguiente código en R puede encontrarse una aproximación para π siguiendo el procedimiento planteado. Revise la sección sobre la ley de los grandes números y el código que aparece en la Figura 5.2, en la página 358, para una explicación de estos comandos.

```
N <- 100
s <- rep(0,N)
x <- runif(1,0,1)
y <- runif(1,0,1)
if (y<=sqrt(1-x^2)){s[1]=1}
for (n in 2:N){
  x <- runif(1,0,1)
  y <- runif(1,0,1)
  if (y<=sqrt(1-x^2)){
    s[n] <- ((n-1)*s[n-1]+1)/n }
  else{ s[n] <- (n-1)*s[n-1]/n }
}
cat('El valor aproximado de pi es',4*s[n],'\n')
plot(4*s,type="l")
abline(h=pi)
```

3.9 Distribución exponencial

371. Claramente la función $f(x)$ es no negativa e integra uno, pues

$$\int_0^{\infty} \lambda e^{-\lambda x} dx = -e^{-\lambda x} \Big|_0^{\infty} = 1.$$

Integrando $f(x)$ en el intervalo $(-\infty, x]$ se encuentra que $F(x) = 0$ para $x \leq 0$ y para $x > 0$,

$$F(x) = \int_{-\infty}^x f(u) du = \int_0^x \lambda e^{-\lambda u} du = 1 - e^{-\lambda x}.$$

372. a) $P(X < 1) = F(1) = 1 - e^{-2}$.

- b) $P(X \geq 2) = 1 - F(2) = e^{-4}$.
 c) $P(X < 1 | X < 2) = (1 - e^{-2}) / (1 - e^{-4})$.
 d) $P(1 \leq X \leq 2 | X > 0) = F(2) - F(1) = e^{-2} - e^{-4}$.

373. Para la esperanza tenemos que, definiendo $u = x$ y $dv = \lambda e^{-\lambda x} dx$,

$$E(X) = \int_0^{\infty} x \lambda e^{-\lambda x} dx = x(-e^{-\lambda x}) \Big|_0^{\infty} - \int_0^{\infty} (-e^{-\lambda x}) dx = -\frac{1}{\lambda} e^{-\lambda x} \Big|_0^{\infty} = \frac{1}{\lambda}.$$

De manera similar, calcule los otros momentos.

374. Mostramos únicamente el caso del primer momento. Los otros momentos se pueden calcular de manera análoga.

$$E(X) = \int_0^{\infty} P(X > x) dx = \int_0^{\infty} (1 - F(x)) dx = \int_0^{\infty} e^{-\lambda x} dx = \frac{1}{\lambda}.$$

375. Use el método de inducción sobre el valor de n . Para $n = 1$ tenemos que $E(X) = 1/\lambda$. Suponga válido el resultado para $n - 1$. Calcule $E(X^n)$ usando el método de integración por partes encontrando que

$$E(X^n) = \frac{n}{\lambda} E(X^{n-1}).$$

376. Para $x > 0$, $F_{cX}(x) = P(cX \leq x) = P(X \leq x/c) = F_X(x/c) = 1 - e^{-(\lambda/c)x}$. Esto significa que cX tiene distribución $\exp(\lambda/c)$.

377. La f.g.m. se calcula como aparece abajo. Derive esta función dos veces y utilice la fórmula $M^{(n)}(0) = E(X^n)$ para hallar la esperanza y la varianza.

$$M(t) = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx = \int_0^{\infty} \lambda e^{-(\lambda-t)x} dx = \frac{\lambda}{\lambda-t} \quad \text{si } t < \lambda.$$

378. $c_p = -(1/\lambda) \ln(1 - p)$.

379. Aplique la definición de probabilidad condicional y observe que

$$(X > x + y) \cap (X > y) = (X > x + y).$$

380. Para $= 0, 1, 2, \dots$,

$$P(Y = y) = P(y < X \leq y + 1) = e^{-\lambda y} - e^{-\lambda(y+1)} = (e^{-\lambda})^y (1 - e^{-\lambda}).$$

381. Se observa que si $u \in (0, 1)$ entonces $x = (-1/\lambda) \ln(1 - u) \in (0, \infty)$. Así, la v.a. X toma valores en $(0, \infty)$. Para cualquier x en este intervalo,

$$P(X \leq x) = P\left(-\frac{1}{\lambda} \ln(1 - U) \leq x\right) = P(U \leq 1 - e^{-\lambda x}) = 1 - e^{-\lambda x}.$$

382. a) Sea N el número de ofertas recibidas hasta vender el coche. Entonces, para $n = 1, 2, \dots$, $P(N = n) = (1 - e^{-\lambda 50,000})^{n-1} (e^{-\lambda 50,000})$.

b) Sea V el precio de venta. Entonces $P(V > 55,000)$ es igual a

$$\sum_{n=1}^{\infty} P(V > 55,000, N = n) \sum_{n=1}^{\infty} (1 - e^{-\lambda 50,000})^{n-1} (e^{-\lambda 55,000}) = e^{-\lambda 5,000}.$$

$$c) E(V) = \int_0^{\infty} P(V > v) dv = \int_0^{50} 1 dv + \int_{50}^{\infty} e^{-\lambda(v-50)} dv = 50 + 1/\lambda.$$

383. Sea X el tiempo de reparación. Entonces

$$a) P(X > 2) = e^{-1}.$$

$$b) P(X \leq 4) = 1 - e^{-2}.$$

$$c) P(X \leq 4 | X > 2) = P(2 < X \leq 4) / P(X > 2) = 1 - e^{-1}.$$

3.10 Distribución gamma

384. Claramente $f(x) \geq 0$. Para demostrar que $\int_0^{\infty} f(x) dx = 1$ lleve a cabo el cambio de variable $t = \lambda x$ y reduzca la integral a la definición de la función gamma.

$$385. a) P(X < 1) = 1 - 4e^{-3}.$$

$$b) P(X \geq 2) = 7e^{-6}.$$

$$c) P(X < 1 | X < 2) = P(X < 1) / P(X < 2) = (1 - 4e^{-3}) / (1 - 7e^{-6}).$$

$$d) P(1 \leq X \leq 2 | X > 0) = P(1 \leq X \leq 2) = 4e^{-3} - 7e^{-6}.$$

386. Se muestra únicamente la forma de calcular la esperanza. Un método similar puede usarse para encontrar los otros momentos.

$$E(X) = \int_0^{\infty} x \frac{(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} \lambda e^{-\lambda x} dx = \frac{\alpha}{\lambda} \int_0^{\infty} x \frac{(\lambda x)^{\alpha}}{\Gamma(\alpha+1)} \lambda e^{-\lambda x} dx = \frac{\alpha}{\lambda}.$$

387. Derive la función de densidad $f(x)$ e iguale a cero. Resuelva la ecuación observando que se debe cumplir la condición $\alpha > 1$ para que la solución sea positiva. Analizando la expresión de $f'(x)$ compruebe que el punto encontrado es un máximo y es único.

388. Para $n \geq 1$,

$$\begin{aligned} E(X^n) &= \int_0^\infty x^n \frac{(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} \lambda e^{-\lambda x} dx \\ &= \frac{\alpha(\alpha+1) \cdots (\alpha+n-1)}{\lambda^n} \int_0^\infty \frac{(\lambda x)^{\alpha+n-1}}{\Gamma(\alpha+n)} \lambda e^{-\lambda x} dx \\ &= \frac{\alpha(\alpha+1) \cdots (\alpha+n-1)}{\lambda^n}. \end{aligned}$$

389. Para $x > 0$,

$$\begin{aligned} P(cX \leq x) &= P(X \leq x/c) \\ &= \int_0^{x/c} \frac{(\lambda u)^{\alpha-1}}{\Gamma(\alpha)} \lambda e^{-\lambda u} du \\ &= \int_0^x \frac{((\lambda/c)u)^{\alpha-1}}{\Gamma(\alpha)} (\lambda/c) e^{-(\lambda/c)v} dv. \end{aligned}$$

390. La f.g.m. se calcula como aparece abajo. Derive esta función dos veces y utilice la fórmula $M^{(n)}(0) = E(X^n)$ para hallar la esperanza y la varianza.

$$\begin{aligned} M(t) &= \int_0^\infty e^{tx} \frac{(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} \lambda e^{-\lambda x} dx \\ &= \frac{\lambda^\alpha}{(\lambda-t)^\alpha} \int_0^\infty \frac{((\lambda-t)x)^{\alpha-1}}{\Gamma(\alpha)} (\lambda-t) e^{-(\lambda-t)x} dx \\ &= \left(\frac{\lambda}{\lambda-t} \right)^\alpha \quad \text{si } t < \lambda. \end{aligned}$$

391. Por independencia, para $t < \lambda$,

$$M_{X+Y}(t) = M_X(t)M_Y(t) = \left(\frac{\lambda}{\lambda-t} \right)^{\alpha_1} \left(\frac{\lambda}{\lambda-t} \right)^{\alpha_2} = \left(\frac{\lambda}{\lambda-t} \right)^{\alpha_1+\alpha_2}.$$

392. Por independencia, para $t < \lambda$,

$$M_{X_1+\dots+X_n}(t) = M_{X_1}(t) \cdots M_{X_n}(t) = \left(\frac{\lambda}{\lambda-t} \right) \cdots \left(\frac{\lambda}{\lambda-t} \right) = \left(\frac{\lambda}{\lambda-t} \right)^n.$$

393. Para obtener el inciso (a), utilice integración por partes con $u = (\lambda y)^{n-1}/\Gamma(n)$ y $dv = \lambda e^{-\lambda y} dy$ en la expresión que aparece abajo. El inciso (b) se obtiene aplicando el inciso (a) repetidas veces.

$$F_n(x) = \int_0^x \frac{(\lambda y)^{n-1}}{\Gamma(n)} \lambda e^{-\lambda y} dy.$$

394. Para obtener el inciso (a), utilice integración por partes con $u = t^\alpha$ y $dv = e^{-t} dt$ en la expresión que aparece abajo. El inciso (b) se obtiene aplicando el inciso (a) repetidas veces, y puede usarse el método de inducción para una demostración formal, previa comprobación de los incisos (c) y (d), los cuales son ejercicios simples de integración.

$$\Gamma(\alpha + 1) = \int_0^\infty t^\alpha e^{-t} dt.$$

Para obtener el inciso (d) lleve a cabo el cambio de variable $t = x^2/2$ en la integral dada por $\Gamma(1/2)$ y use la siguiente identidad que será evidente una vez que revisemos la distribución normal estándar.

$$\int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \frac{1}{2}.$$

3.11 Distribución beta

395. a) Haga el cambio de variable $u = 1 - x$ en la integral.
 b) Este es un ejercicio simple de integración.
 c) Esto es consecuencia de los incisos (a) y (b).
 d) Use integración por partes con $u = x^a$ y $dv = (1 - x)^{b-1} dx$.
 e) Use integración por partes con $u = x^a(1 - x)$ y $dv = (1 - x)^{b-2} dx$.
 f) Esto es consecuencia de los incisos (a) y (e).
 g) Haga el cambio de variable $x = \text{sen}^2 \theta$ en la integral.
396. Para $x \in (0, 1)$, y al hacer el cambio de variable $v = 1 - u$,

$$P(1-X \leq x) = P(X \geq 1-x) = \int_{1-x}^1 u^{a-1}(1-u)^{b-1} du = \int_0^x v^{b-1}(1-v)^{a-1} dv.$$

397. Se muestra únicamente un procedimiento para encontrar la esperanza. Los otros momentos pueden hallarse de manera similar, usando las propiedades de la función beta.

$$E(X) = \frac{1}{B(a, b)} \int_0^1 x^a(1-x)^{b-1} dx = \frac{B(a+1, b)}{B(a, b)} = \frac{a}{a+b}.$$

398. Para $n \geq 0$ entero, y usando las propiedades de la función beta, el n -ésimo momento se puede calcular como aparece abajo. Reduciendo varias veces la

función beta en el numerador se llega al resultado.

$$\begin{aligned} E(X^n) &= \frac{1}{B(a, b)} \int_0^1 x^{a+n-1} (1-x)^{b-1} dx \\ &= \frac{B(a+n, b)}{B(a, b)} \\ &= \frac{a+n-1}{a+b+n-1} \frac{B(a+n-1, b)}{B(a, b)}. \end{aligned}$$

$$399. \quad a) \quad F(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ x^a & \text{si } 0 < x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

$$b) \quad F(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ 1 - (1-x)^b & \text{si } 0 < x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

400. Derive la función de densidad $f(x)$ e iguale a cero. Resuelva la ecuación observando que se deben cumplir las condiciones $a > 1$ y $b > 1$ para garantizar que la solución sea positiva. Analizando la expresión de $f'(x)$ compruebe que el punto encontrado es un máximo y es único.

3.12 Distribución Weibull

401. Claramente $f(x) \geq 0$. Para demostrar que la integral es uno, lleve a cabo el cambio de variable $\lambda u := (\lambda x)^\alpha$. Ello reduce la integral a la integral de la función de densidad $\exp(-u)$.
402. Haga el cambio de variable $\lambda u := (\lambda y)^\alpha$.
403. En cada integral lleve a cabo el cambio de variable $u = (\lambda x)^\alpha$ y reduzca la integral a la función gamma.
404. Derive la función de densidad $f(x)$ e iguale a cero. Resuelva la ecuación observando que se deben cumplir la condición $\alpha > 1$ para garantizar que la solución sea positiva. Analizando la expresión de $f'(x)$ compruebe que el punto encontrado es un máximo y es único.
405. Nuevamente lleve a cabo el cambio de variable $u = (\lambda x)^\alpha$ y reduzca la integral a la función gamma.
406. $x = \frac{1}{\lambda} (-\ln(1-p))^{1/\alpha}$.

407. Si $u \in (0, 1)$ entonces $x = (1/\lambda)(-\ln(1-u))^{1/\alpha} \in (0, \infty)$, y existe una correspondencia biunívoca entre estos dos intervalos a través de esta función. Sea $X = (1/\lambda)(-\ln(1-U))^{1/\alpha}$. Para cualquier $x \in (0, \infty)$,

$$P(X \leq x) = P\left(\frac{1}{\lambda}(-\ln(1-U))^{1/\alpha} \leq x\right) = P(U \leq 1 - e^{-(\lambda x)^\alpha}) = 1 - e^{-(\lambda x)^\alpha}.$$

3.13 Distribución normal

408. a) Este no es un ejercicio sencillo. El siguiente método requiere de un poco de conocimiento de integración sobre el plano y hace uso del cambio de coordenadas cartesianas a coordenadas polares. Sea I la integral en cuestión. Mediante el cambio de variable $y = (x - \mu)/\sigma$, el integrando se reduce a la función de densidad de la distribución normal estándar. Entonces

$$\begin{aligned} I^2 &= \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx\right) \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy\right) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-(x^2+y^2)/2} dx dy. \end{aligned}$$

Ahora haga el cambio de variable $(x, y) \mapsto (r \cos \theta, r \sin \theta)$, en donde $dx dy$ se reemplaza por $r dr d\theta$.

b) Compruebe que $f'(x) = 0$ ssi $x = \mu$, con $f''(\mu) < 0$.

c) Compruebe que $f''(x) = 0$ ssi $x = \mu \pm \sigma$, con $f''(x) > 0$ si $|x - \mu| > \sigma$ y $f''(x) < 0$ si $|x - \mu| < \sigma$.

409. $c = 2.576$. Esto significa que el 99% de la probabilidad en la distribución $N(\mu, \sigma^2)$ se encuentra alrededor de la media μ , entre $\mu - 2.576\sigma$ y $\mu + 2.576\sigma$.

410. Para la segunda desigualdad observe que

$$x \int_x^{\infty} e^{-u^2/2} du \leq \int_x^{\infty} u e^{-u^2/2} du.$$

Para la primera desigualdad defina

$$g(x) = \int_x^{\infty} e^{-u^2/2} du - \frac{x}{1+x^2} e^{-x^2/2}.$$

Demuestre que $g(0) > 0$, $g'(x) < 0$ y $\lim_{x \rightarrow \infty} g(x) = 0$.

411. a) Lleve a cabo el cambio de variable $u = (x - \mu)/\sigma$.
 b) Lleve a cabo el cambio de variable $y = (x - \mu)/\sigma$ y después use integración por partes con $u = y$ y $dv = y(1/\sqrt{2\pi})e^{-y^2/2} dy$.
 c) Use los dos incisos anteriores.

412. Tanto la moda como la mediana son iguales a μ .
413. Para llevar a cabo la integral $E(e^{tX})$ se puede hacer el cambio de variable $u = (x - \mu)/\sigma$, después unir los exponentes de la función exponencial y completar el cuadrado. Derivando esta función dos veces y utilizando la fórmula $M^{(n)}(0) = E(X^n)$ se puede hallar la esperanza y la varianza.
414. Si n es impar, entonces la función $x \mapsto x^n f(x)$ es una función impar con integral finita en las partes positiva y negativa del eje. Por lo tanto, su integral sobre \mathbb{R} es cero. Si n es par, la integral $E(X^n)$ puede llevarse a cabo por el método de integración por partes con $u = x^{n-1}$ y $dv = x/\sqrt{2\pi\sigma^2}e^{-(x-\mu)^2/2\sigma^2} dx$, encontrando que $E(X^n) = (n-1)\sigma^2 E(X^{n-2})$. Procediendo de manera iterada,

$$E(X^n) = (n-1)(n-2)\cdots 4 \cdot 2 (\sigma^2)^{n/2} = \frac{n!}{(n/2)!} \left(\frac{\sigma^2}{2}\right)^{n/2}.$$

415. Por independencia,

$$M_{X_1+X_2}(t) = M_{X_1}(t) M_{X_2}(t) = e^{\mu_1 t + \frac{1}{2}\sigma_1^2 t^2} e^{\mu_2 t + \frac{1}{2}\sigma_2^2 t^2} = e^{(\mu_1 + \mu_2)t + \frac{1}{2}(\sigma_1^2 + \sigma_2^2)t^2}.$$

416. a) $F_Z(z) = P(Z \leq z) = P(X \leq \mu + \sigma z) = F_X(\mu + \sigma z)$. Derivando se obtiene $f_Z(z) = f_X(\mu + \sigma z) \cdot \sigma$. Esta expresión se reduce a la función de densidad normal estándar.
- b) $F_X(x) = P(X \leq x) = P(Z \leq (x - \mu)/\sigma) = F_Z((x - \mu)/\sigma)$. Derivando se obtiene $f_X(x) = f_Z((x - \mu)/\sigma) \cdot (1/\sigma)$. Esta expresión es la función de densidad $N(\mu, \sigma^2)$.

- 417.

$$\begin{aligned} P(a < Z^2 < b) &= P(\sqrt{a} < Z < \sqrt{b}) + P(-\sqrt{b} < Z < -\sqrt{a}) \\ &= \Phi(\sqrt{b}) - \Phi(\sqrt{a}) + \Phi(-\sqrt{a}) - \Phi(-\sqrt{b}) \\ &= \Phi(\sqrt{b}) - \Phi(\sqrt{a}) + (1 - \Phi(\sqrt{a})) - (1 - \Phi(\sqrt{b})) \\ &= 2(\Phi(\sqrt{b}) - \Phi(\sqrt{a})). \end{aligned}$$

418. a) $P(X \leq 7) = P((X - 5)/\sqrt{10} \leq (7 - 5)/\sqrt{10}) = \Phi(2/\sqrt{10})$.
- b) $P(X > 4) = P((X - 5)/\sqrt{10} > (4 - 5)/\sqrt{10}) = 1 - \Phi(-1/\sqrt{10})$.
- c) $P(|X - 2| \leq 3) = P(-3 \leq X - 2 \leq 3) = P(-1 \leq X \leq 4) = P(-6/\sqrt{10} \leq (X - 5)/\sqrt{10} \leq -1/\sqrt{10}) = \Phi(-1/\sqrt{10}) - \Phi(-6/\sqrt{10})$.
- d) $P(|X - 6| > 1) = 1 - P(|X - 6| \leq 1) = 1 - P(-1 \leq X - 6 \leq 1) = 1 - P(5 \leq X \leq 7) = 1 - P(0 \leq (X - 5)/\sqrt{10} \leq 2/\sqrt{10}) = 1 - \Phi(2/\sqrt{10}) + \Phi(0) = 3/2 - \Phi(2/\sqrt{10})$.

- 419.

- a) $P(X > 5) = 0.79767$. d) $P(X \geq 6) = 0.1587$.
 b) $P(4 < X < 16) = 0.6825$. e) $P(|X - 4| \leq 6) = 0.4772$.
 c) $P(X \leq 8) = 0.3693$. f) $P(|X - 6| > 3) = 0.6879$.

420. a) $P(X > 21,000) = 0.0228$.

b) $P(X > 21,500 | X > 21,000) = 0.0570$.

421. Se busca x tal que $P(X \leq x) = 0.95$, es decir x es tal que $P(Z \leq (x - 30)/5) = 0.95$. Entonces, de la tabla de la distribución normal, se encuentra que $(x - 30)/5 = 1.645$, esto es, $x = 38.225$.

422. a) $E(X^{2001}) = 0$.

b) $F_Y(y) = \begin{cases} 0 & \text{si } y \leq 0, \\ 2F_X(y) - 1 & \text{si } y > 0. \end{cases}$

c) $f_Y(y) = 2f_X(y) 1_{(0,\infty)}(y) = \begin{cases} 2/\sqrt{2\pi\sigma^2} e^{-y^2/2\sigma^2} & \text{si } y > 0, \\ 0 & \text{en otro caso.} \end{cases}$

d) $E(Y) = \sqrt{2\sigma^2/\pi}$.

e) $E(Y^2) = \sigma^2$.

f) $\text{Var}(Y) = (1 - 2/\pi)\sigma^2$.

423. La variable Y toma valores en el intervalo $(0, \infty)$. Para cualquiera de estos valores y , $P(Y \leq y) = P(e^X \leq y) = P(X \leq \ln y) = F_X(\ln y)$. Derivando respecto de y , $f_Y(y) = (1/y) f_X(\ln y)$, de donde se obtiene el resultado.

3.14 Distribución ji-cuadrada

424. Claramente $f(x) \geq 0$. Además, haciendo el cambio de variable $t = x/2$, tenemos que

$$\begin{aligned} \frac{1}{\Gamma(n/2)} \left(\frac{1}{2}\right)^{n/2} \int_0^\infty x^{n/2-1} e^{-x/2} dx &= \frac{1}{\Gamma(n/2)} \left(\frac{1}{2}\right)^{n/2} \int_0^\infty (2t)^{n/2-1} e^{-t} 2dt \\ &= \frac{1}{\Gamma(n/2)} \left(\frac{1}{2}\right)^{n/2} 2^{n/2} \Gamma(n/2) = 1. \end{aligned}$$

425. a) $F_{X^2}(x) = P(X^2 \leq x) = P(-\sqrt{x} \leq X \leq \sqrt{x}) = F_X(\sqrt{x}) - F_X(-\sqrt{x})$.

b) Del inciso anterior, en el caso cuando X es absolutamente continua, $f_{X^2}(x) = (f_X(\sqrt{x}) + f_X(-\sqrt{x}))(1/\sqrt{x})$. Ahora sólo resta simplificar esta expresión cuando $f_X(x)$ es la función de densidad normal estándar para encontrar la función de densidad $\chi^2(1)$.

426. Substituya las expresiones en la fórmula $f_{cX}(x) = (1/c) f_X(x/c)$ y encuentre la función de densidad gamma(α, λ).
427. Este es el caso particular $c = 1$ del ejercicio anterior.
428. Se muestra el procedimiento para el caso de la esperanza. Se reconstruye en el integrando la función de densidad $\chi^2(n+2)$. Los otros momentos se pueden calcular de manera similar.

$$\begin{aligned}
 E(X) &= \int_0^{\infty} x \frac{1}{\Gamma(n/2)} \left(\frac{1}{2}\right)^{n/2} x^{n/2-1} e^{-x/2} dx \\
 &= \frac{2\Gamma(n/2+1)}{\Gamma(n/2)} \int_0^{\infty} \frac{1}{\Gamma(n/2+1)} \left(\frac{1}{2}\right)^{n/2+1} x^{(n/2+1)-1} e^{-x/2} dx \\
 &= \frac{2(n/2)\Gamma(n/2)}{\Gamma(n/2)} \\
 &= n.
 \end{aligned}$$

429. Derive la función de densidad $f(x)$ e iguale a cero. Resuelva la ecuación observando que se deben cumplir la condición $n > 2$ para garantizar que la solución sea positiva. Analizando la expresión de $f'(x)$ compruebe que el punto encontrado es un máximo y es único.
430. El procedimiento es similar al presentado para el cálculo de la esperanza.

$$\begin{aligned}
 E(X^m) &= \int_0^{\infty} x^m \frac{1}{\Gamma(n/2)} \left(\frac{1}{2}\right)^{n/2} x^{n/2-1} e^{-x/2} dx \\
 &= \frac{2^m \Gamma(n/2+m)}{\Gamma(n/2)} \int_0^{\infty} \frac{1}{\Gamma(n/2+m)} \left(\frac{1}{2}\right)^{n/2+m} x^{(n/2+m)-1} e^{-x/2} dx \\
 &= \frac{2^m \Gamma(n/2+m)}{\Gamma(n/2)}.
 \end{aligned}$$

431. La f.g.m. se calcula como aparece abajo, efectuando el cambio de variable $u = x(1-2t)$ para $t < 1/2$. Derive la f.g.m. dos veces y utilice la fórmula

$M^{(n)}(0) = E(X^n)$ para hallar la esperanza y la varianza.

$$\begin{aligned}
 M(t) &= \int_0^\infty e^{tx} \frac{1}{\Gamma(n/2)} \left(\frac{1}{2}\right)^{n/2} x^{n/2-1} e^{-x/2} dx \\
 &= \int_0^\infty \frac{1}{\Gamma(n/2)} \left(\frac{1}{2}\right)^{n/2} x^{n/2-1} e^{-x(1-2t)/2} dx \\
 &= \left(\frac{1}{1-2t}\right)^{n/2} \int_0^\infty \frac{1}{\Gamma(n/2)} \left(\frac{1}{2}\right)^{n/2} u^{n/2-1} e^{-u/2} du \\
 &= \left(\frac{1}{1-2t}\right)^{n/2}.
 \end{aligned}$$

432. Por la hipótesis de independencia, para $t < 1/2$,

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t) = \left(\frac{1}{1-2t}\right)^{(n+m)/2}.$$

433. Sea $X = -2\ln(U)$. Se observa que X es una v.a. continua que toma valores en $(0, \infty)$. Para x en este intervalo,

$$P(X \leq x) = P(-2\ln(U) \leq x) = P(U \geq e^{-x/2}) = 1 - e^{-x/2}.$$

3.15 Distribución t

434. Claramente $f(x) \geq 0$. Para demostrar que $f(x)$ integra uno, haga el cambio de variable $u = (1 + x^2/n)^{-1}$. La integral se reduce a la función $B(a, b)$ con $a = n/2$ y $b = 1/2$. Recuerde la identidad $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$.

435. a) La función $x \mapsto xf(x)$ es una función impar cuya integral es finita, cuando $n > 1$, en las partes negativa y positiva del eje. Por lo tanto, la integral sobre todo \mathbb{R} es cero.

b) Use integración por partes con $u = x$ y $dv = x(1 + x^2/n)^{-(n+1)/2} dx$. Después haga el cambio de variable $y = x\sqrt{(n-2)/n}$ para reconstruir la función de densidad $t(n-2)$. De aquí surge la condición $n > 2$.

c) En este caso la varianza es el segundo momento.

436. Use integración por partes con $u = x^{m-1}$ y $dv = x(1 + x^2/n)^{-(n+1)/2} dx$. Después haga el cambio de variable $y = x\sqrt{(n-2)/n}$ para reconstruir el momento $m-2$ de la función de densidad $t(n-2)$.

437. Si m es impar con $2 \leq m < n$, entonces la función $x \mapsto xf(x)$ es una función impar con integral finita en los intervalos $(-\infty, 0)$ y $(0, \infty)$. Por lo tanto, la

integral sobre todo \mathbb{R} es cero. Si m par con $2 \leq m < n$, entonces el m -ésimo momento es finito y está dado por la expresión que aparece en el ejercicio anterior. Se comprueba que esta fórmula se puede escribir en términos de la función gamma, como aparece en el enunciado. Por último, cuando $m \geq n$, el m -ésimo momento no está definido pues el integrando no es absolutamente convergente, en efecto, denotando por c a las constantes involucradas,

$$E(|X|^m) = c \int_{-\infty}^{\infty} |x|^m (1 + x^2/n)^{-\frac{n+1}{2}} dx.$$

El integrando es un polinomio cuyo grado es $m - (n + 1)$. De modo que la integral es divergente cuando $m - (n + 1) \geq -1$, es decir, cuando $m \geq n$.

438. La función de densidad $f(x)$ de la distribución $t(n)$ es una función par, estrictamente creciente en $(-\infty, 0)$ y estrictamente decreciente en $(0, \infty)$. Tiene por lo tanto un máximo absoluto en $x = 0$. Esto significa que la moda y la mediana es $x = 0$.

439. Demuestre que para cualquier $t > 0$,

$$a) \lim_{x \rightarrow \infty} e^{tx} (1 + x^2/n)^{-(n+1)/2} = \infty.$$

$$b) \int_{-\infty}^0 e^{tx} (1 + x^2/n)^{-(n+1)/2} dx < \infty.$$

440. Cuando $n \rightarrow \infty$, tenemos que

$$\begin{aligned} (1 + x^2/n)^{-(n+1)/2} &= (1 + x^2/n)^{-1/2} (1 + x^2/n)^{-n/2} \\ &\rightarrow e^{-x^2/2}. \end{aligned}$$

Por otro lado,

$$\begin{aligned}
 \frac{1}{\sqrt{n}} \frac{\Gamma((n+1)/2)}{\pi \Gamma(n/2)} &= \frac{1}{\sqrt{n}} \frac{\Gamma((n+1)/2)}{\Gamma(1/2) \Gamma(n/2)} \\
 &= \frac{1}{\sqrt{n}} \frac{1}{B(1/2, n/2)} \\
 &= \left(\sqrt{n} \int_0^1 x^{-1/2} (1-x)^{n/2-1} dx \right)^{-1} \\
 &= \left(\int_0^n y^{-1/2} (1-y/n)^{n/2-1} dy \right)^{-1} \quad (x = y/n) \\
 &\rightarrow \left(\int_0^\infty y^{-1/2} e^{-y/2} dy \right)^{-1} \\
 &= \left(\sqrt{2} \int_0^\infty t^{-1/2} e^{-t} dt \right)^{-1} \quad (t = y/2) \\
 &= \left(\sqrt{2} \Gamma(1/2) \right)^{-1} = \frac{1}{\sqrt{2\pi}}.
 \end{aligned}$$

3.16 Distribución F

441. Claramente $f(x) \geq 0$. Por otro lado, haciendo el cambio de variable $y = (1 + \frac{a}{b}x)^{-1}$ se tiene que

$$\begin{aligned}
 \int_0^\infty x^{a/2-1} \left(1 + \frac{a}{b}x\right)^{-(a+b)/2} dx &= \left(\frac{b}{a}\right)^{a/2} \int_0^1 y^{b/2-1} (1-y)^{a/2-1} dy \\
 &= \left(\frac{b}{a}\right)^{a/2} B\left(\frac{b}{2}, \frac{a}{2}\right) \\
 &= \left(\frac{b}{a}\right)^{a/2} \frac{\Gamma(\frac{b}{2})\Gamma(\frac{a}{2})}{\Gamma(\frac{a+b}{2})}.
 \end{aligned}$$

442. Se muestra únicamente el cálculo de la esperanza y el procedimiento es el mismo que en la demostración de que la función de densidad lo es. Los otros momentos se pueden calcular de manera similar. Nuevamente haciendo el

cambio de variable $y = (1 + \frac{a}{b}x)^{-1}$ y suponiendo $b > 2$, se tiene que

$$\begin{aligned}
 E(X) &= \frac{\Gamma(\frac{a+b}{2})}{\Gamma(\frac{a}{2})\Gamma(\frac{b}{2})} \left(\frac{a}{b}\right)^{a/2} \int_0^\infty x x^{a/2-1} (1 + \frac{a}{b}x)^{-(a+b)/2} dx \\
 &= \frac{\Gamma(\frac{a+b}{2})}{\Gamma(\frac{a}{2})\Gamma(\frac{b}{2})} \left(\frac{b}{a}\right) \int_0^1 y^{(b/2-1)-1} (1-y)^{(a/2+1)-1} dy \\
 &= \frac{\Gamma(\frac{a+b}{2})}{\Gamma(\frac{a}{2})\Gamma(\frac{b}{2})} \left(\frac{b}{a}\right) B\left(\frac{b-2}{2}, \frac{a+2}{2}\right) \\
 &= \frac{\Gamma(\frac{a+b}{2})}{\Gamma(\frac{a}{2})\Gamma(\frac{b}{2})} \left(\frac{b}{a}\right) \frac{\Gamma(\frac{b-2}{2})\Gamma(\frac{a+2}{2})}{\Gamma(\frac{a+b}{2})} \\
 &= \frac{b}{b-2}.
 \end{aligned}$$

443. Derive la función de densidad $f(x)$ e iguale a cero. Resuelva la ecuación observando que se debe cumplir la condición $a > 2$ para garantizar que la solución sea positiva. Analizando la expresión de $f'(x)$ compruebe que el punto encontrado es un máximo y es único.

444. El cálculo del n -ésimo momento es similar al cálculo de la esperanza. Suponiendo $b > 2n$ y haciendo el cambio de variable $y = (1 + \frac{a}{b}x)^{-1}$ tenemos nuevamente que

$$\begin{aligned}
 E(X^n) &= \frac{\Gamma(\frac{a+b}{2})}{\Gamma(\frac{a}{2})\Gamma(\frac{b}{2})} \left(\frac{a}{b}\right)^{a/2} \int_0^\infty x^n x^{a/2-1} (1 + \frac{a}{b}x)^{-(a+b)/2} dx \\
 &= \frac{\Gamma(\frac{a+b}{2})}{\Gamma(\frac{a}{2})\Gamma(\frac{b}{2})} \left(\frac{b}{a}\right)^n \int_0^1 y^{(b/2-n)-1} (1-y)^{(a/2+n)-1} dy \\
 &= \frac{\Gamma(\frac{a+b}{2})}{\Gamma(\frac{a}{2})\Gamma(\frac{b}{2})} \left(\frac{b}{a}\right)^n B\left(\frac{b-2n}{2}, \frac{a+2n}{2}\right) \\
 &= \frac{\Gamma(\frac{a+b}{2})}{\Gamma(\frac{a}{2})\Gamma(\frac{b}{2})} \left(\frac{b}{a}\right)^n \frac{\Gamma(\frac{b-2n}{2})\Gamma(\frac{a+2n}{2})}{\Gamma(\frac{a+b}{2})} \\
 &= \left(\frac{b}{a}\right)^n \frac{\Gamma(\frac{b-2n}{2})\Gamma(\frac{a+2n}{2})}{\Gamma(\frac{a}{2})\Gamma(\frac{b}{2})}.
 \end{aligned}$$

445. Demuestre que para cualquier $t > 0$,

$$\lim_{x \rightarrow \infty} e^{tx} x^{a/2-1} \left(1 - \frac{a}{b}x\right)^{-(a+b)/2} = \infty.$$

446. Use la fórmula $f_{X^2}(x) = (f_X(\sqrt{x}) + f_X(-\sqrt{x}))(1/(2\sqrt{x}))$, para $x > 0$. Substituyendo la expresión de la función de densidad $t(n)$ se obtiene, para $x > 0$,

$$f_{X^2}(x) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)} x^{-1/2} (1+x/n)^{-(n+1)/2},$$

correspondiente a la función de densidad $F(1, n)$.

447. Para cualquier $x > 0$,

$$F_{1/X}(x) = P(1/X \leq x) = P(X \geq 1/x) = 1 - P(X \leq 1/x) = 1 - F_X(1/x).$$

Por lo tanto, $f_{1/X}(x) = f_X(1/x)(1/x^2)$. Substituyendo tenemos, para $x > 0$,

$$\begin{aligned} f_{1/X}(x) &= \frac{\Gamma(\frac{a+b}{2})}{\Gamma(\frac{a}{2})\Gamma(\frac{b}{2})} \left(\frac{a}{b}\right)^{a/2} x^{1-a/2} \left(1 + \frac{a}{b} \frac{1}{x}\right)^{-(a+b)/2} \frac{1}{x^2} \\ &= \frac{\Gamma(\frac{a+b}{2})}{\Gamma(\frac{a}{2})\Gamma(\frac{b}{2})} \left(\frac{b}{a}\right)^{b/2-1} x^{b/2} \left(1 + \frac{b}{a} x\right)^{-(a+b)/2}. \end{aligned}$$

Capítulo 4: Vectores aleatorios

4.2 Función de probabilidad conjunta

448. a) $P(X > 0, Y \geq 1) = 15/30$. f) $P(Y \leq 1 | X = 1) = 4/6$.
 b) $P(X \leq 1, Y \geq 1) = 11/30$. g) $P(XY = 0) = 15/30$.
 c) $P(X = 1) = 10/30$. h) $P(XY \geq 2) = 15/30$.
 d) $P(Y = 2) = 14/30$. i) $P(Y \geq 2X) = 12/30$.
 e) $P(X = 0 | Y = 2) = 3/14$. j) $P(X + Y \text{ sea impar}) = 16/30$.
449. a) $P(X \leq 1/2, Y \geq 1/2) = 3/32$. f) $P(|X - Y| \leq 1/2) = 7/8$.
 b) $P(Y \geq 1/2) = 3/4$. g) $P(XY < 1) = 1$.
 c) $P(X \leq 1/2 | Y \geq 1/2) = 1/8$. h) $P(Y \geq X^2) = 4/7$.
 d) $P(X + Y > 1) = 9/10$. i) $P(X^2 + Y^2 \leq 1) = 2/5$.
 e) $P(Y > X) = 2/5$. j) $P(Y \leq 4X(1 - X)) = 16/35$.

450. Claramente las funciones son no negativas. Además,

$$a) \sum_{x,y=1}^{\infty} 2^{-(x+y)} = \left(\sum_{x=1}^{\infty} 2^{-x}\right) \left(\sum_{y=1}^{\infty} 2^{-y}\right) = 1.$$

$$\begin{aligned}
 b) \quad \sum_{x,y=1}^{\infty} 16(1/3)^{x+2y} &= 16 \left(\sum_{x=1}^{\infty} (1/3)^x \right) \left(\sum_{y=1}^{\infty} (1/9)^y \right) = 1. \\
 c) \quad \sum_{x=0}^n \sum_{y=0}^{n-x} f(x,y) &= \sum_{x=0}^n \binom{n}{x} p_1^x \sum_{y=0}^{n-x} \binom{n-x}{y} p_2^y (1-p_1-p_2)^{(n-x)-y}. \\
 &= \sum_{x=0}^n \binom{n}{x} p_1^x (1-p_1)^{n-x} = 1.
 \end{aligned}$$

451. Claramente las funciones son no negativas. Además,

$$\begin{aligned}
 a) \quad \int_0^{\infty} \int_0^{\infty} f(x,y) dx dy &= \left(\int_0^{\infty} e^{-x} dx \right) \left(\int_0^{\infty} e^{-y} dy \right) = 1. \\
 b) \quad \int_0^{\infty} \int_0^1 6y^2 e^{-2x} dy dx &= \left(\int_0^1 3y^2 dy \right) \left(\int_0^{\infty} 2e^{-2x} dx \right) = 1. \\
 c) \quad \int_0^1 \int_0^2 3xy(1-x) dy dx &= \left(\int_0^2 y dy \right) \left(\int_0^1 3x(1-x) dx \right) = 1.
 \end{aligned}$$

452. a) $c = e^{-2}$. d) $c = 3$. g) $c = 2/3$.
 b) $c = 4$. e) $c = 3/2$. h) $c = 2/n$.
 c) $c = 2$. f) $c = 1/8$. i) $c = 2^n$.

453. a) $c = 1$.
 b) La variable $X + Y$ tiene distribución Poisson de parámetro $\lambda = 2$, es decir, para $n = 0, 1, \dots$

$$P(X + Y = n) = e^{-2} 2^n / n!$$

454. a) $c = 12$.
 b) $P(X < Y) = 4/7$.
 c) $P(X + Y > 1) = 34/35$.

455. a) $c = 2$.
 b) $P(|X| + |Y| \leq r) = 1 - e^{-r} - r e^{-r}$.
 c) $\lim_{r \rightarrow \infty} P(|X| + |Y| \leq r) = 1$.
 d) $P(X \leq \theta Y) = 2\theta / (1 + \theta)$.
 e) $\lim_{\theta \rightarrow 0} P(X \leq \theta Y) = 0$.

456. La función es no negativa y puede verificarse que integra uno pues se trata de la función constante 2 sobre el triángulo $x > 0, y > 0, x + y < 1$. Las probabilidades solicitadas son:

- a) $P(X \leq 1/2, Y \leq 1/2) = 1/2.$ c) $P(X > 2Y) = 1/3.$
 b) $P(X + Y > 2/3) = 5/9.$ d) $P(Y > 2X^2) = 7/12.$

$$457. f_X(x) = \begin{cases} (1 + 2(6 - x))/36 & \text{si } x = 1, \dots, 6, \\ 0 & \text{en otro caso.} \end{cases}$$

458. Sea i un posible valor de X y sea j un posible valor de $X + Y$. Para valores i y j tales que $1 \leq j - i \leq 6$,

$$P(X = i, X + Y = j) = P(X = i, Y = j - i) = 1/36.$$

459. Las probabilidades $P(X = 0 | Y = 0)$ y $P(X = 0)$ satisfacen la relación buscada cuando (X, Y) tiene distribución:

$a) \begin{array}{ccc} x \backslash y & 0 & 1 \\ \hline 0 & 0 & 1/4 \\ 1 & 1/4 & 1/2 \end{array}$	$b) \begin{array}{ccc} x \backslash y & 0 & 1 \\ \hline 0 & 1/4 & 1/4 \\ 1 & 1/4 & 1/4 \end{array}$	$c) \begin{array}{ccc} x \backslash y & 0 & 1 \\ \hline 0 & 1/2 & 0 \\ 1 & 1/4 & 1/4 \end{array}$

460. a) Claramente $f(x, y) \geq 0$. La integral de $f(x, y)$ sobre \mathbb{R}^2 se puede calcular integrando primero sobre la variable x y después sobre la variable y . Al considerar la integral sobre x , nos fijamos en los términos que involucran a esta variable en el exponente y se puede completar el cuadrado de la siguiente forma

$$\begin{aligned} & -\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho}{\sigma_1\sigma_2}(x-\mu_1)(y-\mu_2) \right] \\ = & -\frac{1}{2} \left(\frac{(x - (\mu_1 + (\rho\sigma_1/\sigma_2)(y - \mu_2)))^2}{\sigma_1^2(1-\rho^2)} - \frac{\rho^2}{1-\rho^2} \frac{(y - \mu_2)^2}{\sigma_2^2} \right). \end{aligned}$$

Ahora la integral sobre x corresponde a la integral de una función de densidad normal univariada con media $\mu_1 + (\rho\sigma_1/\sigma_2)(y - \mu_2)$ y varianza $\sigma_1^2(1 - \rho^2)$. Al incorporar la constante adecuada, esta integral vale uno, y al reducir las expresiones se obtiene la función de densidad normal en la variable y , con media μ_2 y varianza σ_2^2 , cuya integral es uno.

- b) La matriz inversa de Σ aparece abajo. Haciendo las multiplicaciones del vector $(x - \mu_1, y - \mu_2)$ por la izquierda y por la derecha de Σ^{-1} , se obtiene la expresión buscada.

$$\Sigma^{-1} = \frac{1}{(1-\rho^2)\sigma_1^2\sigma_2^2} \begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}.$$

4.3 Función de distribución conjunta

461. Dibuje los ejes coordenados en un plano cartesiano. Marque los valores x que aparecen en la tabla y dibuje una línea vertical en cada uno de estos valores. En el eje vertical marque los valores y y sus correspondientes líneas horizontales. De esta manera se han formado varios rectángulos. Todos los puntos dentro de cada rectángulo tienen la misma probabilidad acumulada si se tiene cuidado de especificar la inclusión o exclusión de las orillas. Escriba sobre cada rectángulo esta probabilidad acumulada. A partir de este esquema se puede especificar de manera completa a la función de distribución conjunta.
462. Identificando en un plano cartesiano la región en donde la función de densidad es distinta de cero e integrando esta función en la región $(-\infty, x] \times (-\infty, y]$ para distintos valores de x y y , se obtienen las expresiones que aparecen abajo. La graficación de estas funciones requiere cierta habilidad de visualización geométrica.

$$a) F(x, y) = \begin{cases} 0 & \text{si } x \leq 0 \text{ o } y \leq 0, \\ xy & \text{si } 0 < x, y < 1, \\ x & \text{si } 0 < x < 1, y \geq 1, \\ y & \text{si } 0 < y < 1, x \geq 1, \\ 1 & \text{si } x, y \geq 1. \end{cases}$$

$$b) F(x, y) = \begin{cases} 0 & \text{si } x \leq 0 \text{ o } y \leq 0, \\ 2xy(1 - x/2) & \text{si } 0 < x, y < 1, \\ 2x(1 - x/2) & \text{si } 0 < x < 1, y \geq 1, \\ y & \text{si } 0 < y < 1, x \geq 1, \\ 1 & \text{si } x, y \geq 1. \end{cases}$$

$$c) F(x, y) = \begin{cases} 0 & \text{si } x \leq 0 \text{ o } y \leq 0, \\ y^2(1 - e^{-x}) & \text{si } 0 < y < 1, x > 0, \\ 1 - e^{-x} & \text{si } y \geq 1, x > 0. \end{cases}$$

463. a) $P(X = 1, Y = 2) = 1/20.$ e) $F_{X,Y}(1.2, 0.9) = 3/12.$
 b) $P(X = 0, 1 \leq Y < 3) = 3/8.$ f) $F_{X,Y}(-3, 1.5) = 0.$
 c) $P(X + Y \leq 1) = 6/12.$ g) $F_{X,Y}(2, 0) = 7/24.$
 d) $P(X > Y) = 28/120.$ h) $F_{X,Y}(4, 2.7) = 119/120.$

464. El evento $(X \leq x)$ se puede descomponer en la unión disjunta $(X < x) \cup (X = x)$ y, análogamente, $(Y = y) = (Y < y) \cup (Y = y)$. Descomponga adecuadamente el evento $(X \leq x, Y \leq y)$.

465. Demuestre que la función indicada satisface las cinco propiedades que aparecen en la Proposición 4.1 de la página 309.

466. $P(|Y| > |X| + 1) = 1/8$. Por otro lado,

$$a) f_{X,Y}(x, y) = \begin{cases} 1/16 & \text{si } -2 < x, y < 2, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f_X(x) = \begin{cases} 1/4 & \text{si } -2 < x < 2, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f_Y(y) = \begin{cases} 1/4 & \text{si } -2 < y < 2, \\ 0 & \text{en otro caso.} \end{cases}$$

$$d) F_{X,Y}(x, y) = \begin{cases} 0 & \text{si } x \leq -2 \text{ o } y \leq -2, \\ (x+2)(y+2)/16 & \text{si } -2 < x, y < 2, \\ (x+2)/4 & \text{si } -2 < x < 2, y \geq 2, \\ (y+2)/4 & \text{si } -2 < y < 2, x \geq 2, \\ 1 & \text{si } x, y \geq 2. \end{cases}$$

$$e) F_X(x) = \begin{cases} 0 & \text{si } x \leq -2, \\ (x+2)/4 & \text{si } -2 < x < 2, \\ 1 & \text{si } x \geq 2. \end{cases}$$

$$f) F_Y(y) = \begin{cases} 0 & \text{si } y \leq -2, \\ (y+2)/4 & \text{si } -2 < y < 2, \\ 1 & \text{si } y \geq 2. \end{cases}$$

$$g) F_{X+Y}(u) = \begin{cases} 0 & \text{si } u \leq -4, \\ (4+u)^2/32 & \text{si } -4 < u \leq 0, \\ 1 - (4-u)^2/32 & \text{si } 0 < u < 4, \\ 1 & \text{si } u \geq 4. \end{cases}$$

$$h) f_{X+Y}(u) = \begin{cases} (4+u)/16 & \text{si } -4 < u < 0, \\ (4-u)/16 & \text{si } 0 \leq u < 4, \\ 0 & \text{en otro caso.} \end{cases}$$

$$i) F_{X-Y}(u) = \begin{cases} 0 & \text{si } u \leq -4, \\ (4+u)^2/32 & \text{si } -4 < u \leq 0, \\ 1 - (4-u)^2/32 & \text{si } 0 < u < 4, \\ 1 & \text{si } u \geq 4. \end{cases}$$

$$j) f_{X-Y}(u) = \begin{cases} (4+u)/16 & \text{si } -4 < u < 0, \\ (4-u)/16 & \text{si } 0 \leq u < 4, \\ 0 & \text{en otro caso.} \end{cases}$$

Los últimos dos incisos pueden obtenerse haciendo el análisis correspondiente y también a partir de los dos incisos previos al observar que las variables $X+Y$ y $X-Y$ tienen la misma distribución de probabilidad.

$$467. \quad a) F_{XY}(u) = \begin{cases} 0 & \text{si } u \leq -1, \\ (1+u-u \ln(-u))/2 & \text{si } -1 < u \leq 0, \\ 1 - (1-u+u \ln u)/2 & \text{si } 0 < u < 1, \\ 1 & \text{si } u \geq 1. \end{cases}$$

$$b) f_{XY}(u) = \begin{cases} -(1/2) \ln(-u) & \text{si } -1 < u < 0, \\ -(1/2) \ln u & \text{si } 0 \leq u < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

4.4 Función de probabilidad marginal

$$468. \quad a) f_X(x) = \begin{cases} 6/16 & \text{si } x = 0, \\ 10/16 & \text{si } x = 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$f_Y(y) = \begin{cases} 5/16 & \text{si } y = 0, \\ 11/16 & \text{si } y = 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f_X(x) = \begin{cases} x e^{-x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

$$f_Y(y) = \begin{cases} e^{-y} & \text{si } y > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f_X(x) = \begin{cases} 2 e^{-2x} & \text{si } x > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

$$f_Y(y) = \begin{cases} 2(e^{-y} - e^{-2y}) & \text{si } y > 0, \\ 0 & \text{en otro caso.} \end{cases}$$

$$d) f_X(x) = \begin{cases} \binom{n}{x} p_1^x (1 - p_1)^{n-x} & \text{si } x = 0, \dots, n, \\ 0 & \text{en otro caso.} \end{cases}$$

$$f_Y(y) = \begin{cases} \binom{n}{y} p_2^y (1 - p_2)^{n-y} & \text{si } y = 0, \dots, n, \\ 0 & \text{en otro caso.} \end{cases}$$

469. Sumando las cuatro entradas de la tabla se comprueba que esta es una función de probabilidad bivariada. Sumando las entradas en el mismo renglón, o en la misma columna, se encuentran las funciones de probabilidad marginales. Ambas son $\text{Ber}(p)$. Esto muestra que existen distribuciones conjuntas distintas que pueden producir las mismas distribuciones marginales.
470. El procedimiento es similar al utilizado para demostrar que $f(x, y)$ es una función de densidad bivariada en el Ejercicio 460. Por ejemplo, para encontrar la función de densidad marginal de Y es necesario calcular la integral de $f(x, y)$ respecto de x . Al fijarnos en los términos que involucran a esta variable en el exponente, se puede completar el cuadrado de la siguiente forma

$$-\frac{1}{2(1 - \rho^2)} \left[\frac{(x - \mu_1)^2}{\sigma_1^2} - \frac{2\rho}{\sigma_1\sigma_2} (x - \mu_1)(y - \mu_2) \right]$$

$$= -\frac{1}{2} \left(\frac{(x - (\mu_1 + (\rho\sigma_1/\sigma_2)(y - \mu_2)))^2}{\sigma_1^2(1 - \rho^2)} - \frac{\rho^2}{1 - \rho^2} \frac{(y - \mu_2)^2}{\sigma_2^2} \right).$$

Entonces la integral respecto de x corresponde a la integral de una función de densidad normal univariada con media $\mu_1 + (\rho\sigma_1/\sigma_2)(y - \mu_2)$ y varianza $\sigma_1^2(1 - \rho^2)$. Al incorporar la constante adecuada, esta integral vale uno, y al reducir las expresiones se obtiene la función de densidad normal en la variable y , con media μ_2 y varianza σ_2^2 . Por simetría, se obtiene el resultado correspondiente a la variable X .

4.5 Función de distribución marginal

$$471. F_{X,Y}(x, y) = \begin{cases} 0 & \text{si } x < 0 \text{ ó } y < 0, \\ 1/8 & \text{si } 0 \leq x < 1 \text{ y } 0 \leq y < 1, \\ 3/8 & \text{si } 0 \leq x < 1 \text{ y } y \geq 1, \\ 3/8 & \text{si } 0 \leq y < 1 \text{ y } x \geq 1, \\ 1 & \text{si } x \geq 1 \text{ y } y \geq 1. \end{cases}$$

$$F_X(u) = F_Y(u) = \begin{cases} 0 & \text{si } u < 0, \\ 3/8 & \text{si } 0 \leq u < 1, \\ 1 & \text{si } u \geq 1. \end{cases}$$

472. Se presentan únicamente las expresiones analíticas de estas funciones y se omiten las gráficas.

$$a) F_X(x) = \begin{cases} 0 & \text{si } x < a, \\ 3/4 & \text{si } a \leq x < b, \\ 1 & \text{si } x \geq b. \end{cases}$$

$$F_Y(y) = \begin{cases} 0 & \text{si } y < c, \\ 3/4 & \text{si } c \leq y < d, \\ 1 & \text{si } y \geq d. \end{cases}$$

$$b) F_X(u) = F_Y(u) = \begin{cases} 0 & \text{si } u < 0, \\ 1 - e^{-u} & \text{si } u \geq 0. \end{cases}$$

4.6 Independencia de variables aleatorias

473. a) $P(X = x) = g(x)$ y $P(Y = y) = h(y)$.

b) $P(X = x, Y = y) = g(x)h(y) = P(X = x)P(Y = y)$.

c) Las funciones $g(x)$ y $h(y)$ deben ser no negativas para aquellos valores x y y que las variables X y Y pueden tomar. El resultado a demostrar se sigue de lo siguiente:

$$1 = \sum_{x,y} P(X = x, Y = y) = \sum_{x,y} g(x)h(y) = \left(\sum_x g(x)\right) \left(\sum_y h(y)\right).$$

Este producto es uno cuando cada factor es uno.

$$\begin{aligned} 474. \quad a) \quad P(X = Y) &= \sum_{x=0}^{\infty} P(X = x, Y = x) \\ &= \sum_{x=0}^{\infty} P(X = x)P(Y = x) \\ &= \sum_{x=0}^{\infty} (1-p)^x p (1-q)^x q \\ &= pq \sum_{x=0}^{\infty} ((1-p)(1-q))^x \\ &= \frac{pq}{1 - (1-p)(1-q)}. \end{aligned}$$

$$\begin{aligned}
b) \quad P(X \leq Y) &= \sum_{x=0}^{\infty} P(X = x, Y \geq x) \\
&= \sum_{x=0}^{\infty} P(X = x) P(Y \geq x) \\
&= \sum_{x=0}^{\infty} (1-p)^x p \sum_{y=x}^{\infty} (1-q)^y q \\
&= \sum_{x=0}^{\infty} (1-p)^x p (1-q)^x \\
&= p \sum_{x=0}^{\infty} ((1-p)(1-q))^x \\
&= \frac{p}{1 - (1-p)(1-q)}.
\end{aligned}$$

475. Como cada sumando puede tomar los valores 0 o 1, la suma S_N toma los valores 0, 1, 2, ... Sea x cualquiera de estos valores. Por el teorema de probabilidad total, y usando la hipótesis de independencia,

$$\begin{aligned}
P(S_N = x) &= \sum_{n=0}^{\infty} P(S_N = x | N = n) P(N = n) \\
&= \sum_{n=x}^{\infty} P(S_n = x | N = n) P(N = n) \\
&= \sum_{n=x}^{\infty} P(S_n = x) P(N = n) \\
&= \sum_{n=x}^{\infty} \binom{n}{x} p^x (1-p)^{n-x} e^{-\lambda} \frac{\lambda^n}{n!} \\
&= e^{-\lambda p} \frac{(\lambda p)^x}{x!} e^{-\lambda(1-p)} \sum_{n=x}^{\infty} \frac{(\lambda(1-p))^{n-x}}{(n-x)!} \\
&= e^{-\lambda p} \frac{(\lambda p)^x}{x!}.
\end{aligned}$$

476. Sumando las entradas de la tabla por renglones o por columnas se pueden encontrar las funciones de probabilidad individuales de las variables X y Y . Estas funciones de probabilidad son idénticas pues la tabla de probabilidades es simétrica. Puede verificarse que $E(X) = E(Y) = 0$ y $\text{Var}(X) = \text{Var}(Y) = 1/2$. También puede encontrarse la función de probabilidad de la variable $X + Y$ y de allí encontrar que $\text{Var}(X + Y) = 1$. Se verifica entonces que $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) = 1$. Sin embargo, X y Y no son independientes

pues, por ejemplo, $P(X = 0, Y = 0) = 1/2$, lo cual no coincide con $P(X = 0)P(Y = 0) = (1/2)(1/2)$.

477. a) Claramente la función es no negativa y suma uno.
 b) Sumando los renglones o columnas se encuentra que:

$$f_X(x) = \begin{cases} 4/10 & \text{si } x = 0, \\ 2/10 & \text{si } x = 1, \\ 4/10 & \text{si } x = 2, \\ 0 & \text{en otro caso.} \end{cases}$$

$$f_Y(y) = \begin{cases} 3/10 & \text{si } y = 0, \\ 3/10 & \text{si } y = 1, \\ 4/10 & \text{si } y = 2, \\ 0 & \text{en otro caso.} \end{cases}$$

- c) Sumando los valores de la tabla se encuentra que:

$$F(x, y) = \begin{cases} 0 & \text{si } x < 0 \text{ ó } y < 0, \\ 1/10 & \text{si } 0 \leq x < 1, 0 \leq y < 1, \\ 3/10 & \text{si } 0 \leq x < 1, 1 \leq y < 2, \\ 4/10 & \text{si } 0 \leq x < 1, y \geq 2, \\ 2/10 & \text{si } 1 \leq x < 2, 0 \leq y < 1, \\ 4/10 & \text{si } 1 \leq x < 2, 1 \leq y < 2, \\ 6/10 & \text{si } 1 \leq x < 2, y \geq 2, \\ 3/10 & \text{si } x \geq 2, 0 \leq y < 1, \\ 6/10 & \text{si } x \geq 2, 1 \leq y < 2, \\ 10/10 & \text{si } x \geq 2, y \geq 2. \end{cases}$$

- d) A partir de $F(x, y)$ se puede encontrar que:

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0, \\ 4/10 & \text{si } 0 \leq x < 1, \\ 6/10 & \text{si } 1 \leq x < 2, \\ 10/10 & \text{si } x \geq 2. \end{cases}$$

$$F_Y(y) = \begin{cases} 0 & \text{si } y < 0, \\ 3/10 & \text{si } 0 \leq y < 1, \\ 6/10 & \text{si } 1 \leq y < 2, \\ 10/10 & \text{si } y \geq 2. \end{cases}$$

e) Las variables X y Y no son independientes pues, por ejemplo, $F(0, 0) = 1/10$ y no coincide con $F_X(0)F_Y(0) = (4/10)(3/10)$.

478. a) Por definición, la función $f(x, y)$ toma el valor $1/4$ sobre el cuadrado $(-1, 1) \times (-1, 1)$. Por lo tanto es no negativa e integra uno.

b) Integrando sobre todos los posibles valores x o y se encuentra que:

$$f_X(x) = \begin{cases} 1/2 & \text{si } -1 < x < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

$$f_Y(y) = \begin{cases} 1/2 & \text{si } -1 < y < 1, \\ 0 & \text{en otro caso.} \end{cases}$$

c) Integrando la función de densidad $f(x, y)$ hasta un punto (x, y) cualquiera se encuentra que:

$$F(x, y) = \begin{cases} 0 & \text{si } x < -1 \text{ o } y < -1, \\ (x+1)(y+1)/4 & \text{si } -1 \leq x < 1, -1 \leq y < 1, \\ (x+1)/2 & \text{si } -1 \leq x < 1, y \geq 1, \\ (y+1)/2 & \text{si } x \geq 1, -1 \leq y < 1, \\ 1 & \text{si } x \geq 1, y \geq 1. \end{cases}$$

d) A partir de $F(x, y)$ se puede encontrar que:

$$F_X(x) = \begin{cases} 0 & \text{si } x < -1, \\ (x+1)/2 & \text{si } -1 \leq x < 1, \\ 1 & \text{si } x \geq 1. \end{cases}$$

$$F_Y(y) = \begin{cases} 0 & \text{si } y < -1, \\ (y+1)/2 & \text{si } -1 \leq y < 1, \\ 1 & \text{si } y \geq 1. \end{cases}$$

e) Las variables X y Y son independientes pues se puede verificar que $F(x, y) = F_X(x)F_Y(y)$ para cualquier valor de (x, y) . Equivalentemente, $f(x, y) = f_X(x)f_Y(y)$ para todo (x, y) .

479. a) Son independientes, pues $f(x, y) = f_X(x)f_Y(y)$ para todo $x, y \in \mathbb{R}$.

b) No son independientes, pues, por ejemplo, $1/30 = f(0, 0) \neq f_1(0)f_2(0) = (6/30)(8/30)$.

c) Son independientes, pues puede confirmarse que $f(x, y) = f_X(x)f_Y(y)$ para todo $x, y \in \mathbb{R}$.

- d) Son independientes, pues $f(x, y) = f_X(x)f_Y(y)$ para todo $x, y \in \mathbb{R}$.
- e) No son independientes, pues $f_X(x) = x + 1/2$ para $0 < x < 1$ y también $f_Y(y) = y + 1/2$ para $0 < y < 1$. En general, $f(x, y) \neq f_X(x)f_Y(y)$ para $0 < x, y < 1$.
- f) Son independientes, pues puede comprobarse que $f_X(x) = 2x$ para $0 < x < 1$, $f_Y(y) = 2y$ para $0 < y < 1$ y $f_Z(z) = 2z$ para $0 < z < 1$. Se verifica que $f(x, y, z) = f_X(x)f_Y(y)f_Z(z)$ para todo $x, y, z \in \mathbb{R}$.
- g) No son independientes, pues puede comprobarse que $f_X(x) = x + 2/3$ para $0 < x < 1$, $f_Y(y) = y + 2/3$ para $0 < y < 1$, y también $f_Z(z) = z + 2/3$ para $0 < z < 1$. En general, $f(x, y, z) \neq f_X(x)f_Y(y)f_Z(z)$ para $0 < x, y, z < 1$.
- h) Son independientes. pues puede comprobarse que $f_{X_i}(x_i) = 2x_i$ para $0 < x_i < 1, i = 1, \dots, n$. Se verifica que $f(x_1, \dots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n)$ para todo $x_1, \dots, x_n \in \mathbb{R}$.
- i) No son independientes, pues puede comprobarse que $f_X(x) = xe^{-x}$ para $x > 0$ y $f_Y(y) = e^{-y}$ para $y > 0$. En general, $f(x, y) \neq f_X(x)f_Y(y)$ para $0 < y < x$.
- j) No son independientes, pues puede comprobarse que $f_X(x) = 2e^{-2x}$ para $x > 0$ y $f_Y(y) = 2e^{-y}(1 - e^{-y})$ para $y > 0$. En general, $f(x, y) \neq f_X(x)f_Y(y)$ para $0 < x < y$.

$$480. \quad a) \quad F_U(u) = P(\text{máx}\{X, Y\} \leq u) \\ = P(X \leq u, Y \leq u) \\ = P(X \leq u)P(Y \leq u) \\ = F_X(u)F_Y(u).$$

$$b) \quad F_V(v) = P(\text{mín}\{X, Y\} \leq v) \\ = 1 - P(\text{mín}\{X, Y\} > v) \\ = 1 - P(X > v, Y > v) \\ = 1 - P(X > v)P(Y > v) \\ = 1 - (1 - F_X(v))(1 - F_Y(v)).$$

$$c) \quad F_U(u) = \begin{cases} (1 - e^{-\lambda u})^2 & \text{si } u > 0, \\ 0 & \text{si } u \leq 0. \end{cases}$$

$$F_V(v) = \begin{cases} 1 - e^{-2\lambda v} & \text{si } v > 0, \\ 0 & \text{si } v \leq 0. \end{cases}$$

$$d) \quad F_U(u) = \begin{cases} (1 - (1 - p)^{k+1})^2 & \text{si } k \leq u < k + 1; \quad k = 0, 1, \dots \\ 0 & \text{si } u < 0. \end{cases}$$

$$F_V(v) = \begin{cases} 1 - (1-p)^{2(k+1)} & \text{si } k \leq v < k+1; \quad k = 0, 1, \dots \\ 0 & \text{si } v < 0. \end{cases}$$

4.7 Distribución condicional

481. Consideraremos únicamente el caso discreto. Claramente, para cualquier valor de x , $f_{X|Y}(x|y) \geq 0$. Además

$$\sum_x f_{X|Y}(x|y) = \sum_x \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{1}{f_Y(y)} \sum_x f_{X,Y}(x,y) = \frac{f_Y(y)}{f_Y(y)} = 1.$$

482. Consideraremos únicamente el caso discreto. Como $f_{X|Y}(x|y)$ es una función no negativa y que suma uno, tenemos que

$$\cdot) \lim_{x \rightarrow \infty} F_{X|Y}(x|y) = \lim_{x \rightarrow \infty} \sum_{u \leq x} f_{X|Y}(u|y) = \sum_u f_{X|Y}(u|y) = 1.$$

$$\cdot) \lim_{x \rightarrow -\infty} F_{X|Y}(x|y) = \lim_{x \rightarrow -\infty} \sum_{u \leq x} f_{X|Y}(u|y) = 0.$$

$\cdot)$ si $x_1 \leq x_2$ entonces

$$F_{X|Y}(x_1|y) = \sum_{u \leq x_1} f_{X|Y}(u|y) \leq \sum_{u \leq x_2} f_{X|Y}(u|y) = F_{X|Y}(x_2|y).$$

$\cdot)$ Demostraremos que $F_{X|Y}(x+|y) = F_{X|Y}(x|y)$. Sea $D = \{x_1, x_2, \dots\}$ el conjunto de puntos en donde $f_{X|Y}(x|y)$ es estrictamente positiva y sea x cualquier número real. Si x no es un punto de acumulación por la derecha de D , entonces $F_{X|Y}(x+|y) = F_{X|Y}(x+\epsilon|y)$ para $\epsilon > 0$ suficientemente pequeño y por lo tanto $F_{X|Y}(x+|y) = F_{X|Y}(x|y)$. Supongamos ahora que x es un punto de acumulación por la derecha de D . Como la suma de los valores de $f_{X|Y}(x|y)$ es uno, la diferencia $F_{X|Y}(x+\epsilon|y) - F_{X|Y}(x|y)$ puede hacerse tan pequeña como se desee tomando $\epsilon > 0$ suficientemente pequeño. Esto significa nuevamente que $F_{X|Y}(x+|y) = F_{X|Y}(x|y)$.

$$483. \quad a) \quad f_{X|Y}(x|0) = \begin{cases} 1/2 & \text{si } x = 0, \\ 1/4 & \text{si } x = 1, \\ 1/4 & \text{si } x = 2, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) \quad f_{X|Y}(x|1) = \begin{cases} 1/6 & \text{si } x = 0, \\ 2/3 & \text{si } x = 1, \\ 1/6 & \text{si } x = 2, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) f_{X|Y}(x|2) = \begin{cases} 1/5 & \text{si } x = 0, \\ 1/5 & \text{si } x = 1, \\ 3/5 & \text{si } x = 2, \\ 0 & \text{en otro caso.} \end{cases}$$

$$d) F_{X|Y}(x|0) = \begin{cases} 0 & \text{si } x < 0, \\ 1/2 & \text{si } 0 \leq x < 1, \\ 3/4 & \text{si } 1 \leq x < 2, \\ 1 & \text{si } x \geq 2. \end{cases}$$

$$e) F_{X|Y}(x|1) = \begin{cases} 0 & \text{si } x < 0, \\ 1/6 & \text{si } 0 \leq x < 1, \\ 5/6 & \text{si } 1 \leq x < 2, \\ 1 & \text{si } x \geq 2. \end{cases}$$

$$f) F_{X|Y}(x|2) = \begin{cases} 0 & \text{si } x < 0, \\ 1/5 & \text{si } 0 \leq x < 1, \\ 2/5 & \text{si } 1 \leq x < 2, \\ 1 & \text{si } x \geq 2. \end{cases}$$

484. a) Para $y > 0$ fijo,

$$f_{X|Y}(x|y) = \begin{cases} e^{y-x} & \text{si } x > y, \\ 0 & \text{en otro caso.} \end{cases}$$

b) Para $x > 0$ fijo,

$$f_{Y|X}(y|x) = \begin{cases} 1/x & \text{si } 0 < y < x, \\ 0 & \text{en otro caso.} \end{cases}$$

c) Para $y > 0$ fijo,

$$F_{X|Y}(x|y) = \begin{cases} 0 & \text{si } x \leq y, \\ 1 - e^{y-x} & \text{si } x > y. \end{cases}$$

d) Para $x > 0$ fijo,

$$F_{Y|X}(y|x) = \begin{cases} 0 & \text{si } y \leq 0, \\ y/x & \text{si } 0 < y < x, \\ 1 & \text{si } y \geq x. \end{cases}$$

485. a) Por la independencia de los resultados en los lanzamientos, X sigue siendo $\text{unif}\{1, 2, 3, 4, 5, 6\}$.

$$b) P(X = x | X + Y = 5) = \begin{cases} 1/4 & \text{si } x = 1, 2, 3, 4, \\ 0 & \text{en otro caso.} \end{cases}$$

$$c) P(X = x | X + Y \geq 5) = \begin{cases} 3/30 & \text{si } x = 1, \\ 4/30 & \text{si } x = 2, \\ 5/30 & \text{si } x = 3, \\ 6/30 & \text{si } x = 4, 5, 6, \\ 0 & \text{en otro caso.} \end{cases}$$

$$486. \quad a) f_{X,Y}(x, y) = \begin{cases} x/36 & \text{si } x, y = 1, \dots, 6, \\ 1/36 & \text{si } x < y; \quad x, y = 1, \dots, 6, \\ 0 & \text{en otro caso.} \end{cases}$$

$$b) f_{Y|X}(y | x = 3) = \begin{cases} 1/2 & \text{si } y = 3, \\ 1/6 & \text{si } y = 4, 5, 6, \\ 0 & \text{en otro caso.} \end{cases}$$

$$f_{X|Y}(x | y = 3) = \begin{cases} 1/5 & \text{si } x = 1, 2, \\ 3/5 & \text{si } x = 3, \\ 0 & \text{en otro caso.} \end{cases}$$

4.8 Esperanza condicional

487. El número esperado de lanzamientos es $1223/324 = 3.77469$.

488. $E(X | Y = 0) = 1/2$, $E(X | Y = 1) = 1$, $E(X | Y = 2) = 1/2$.

Se comprueba que $E(X) = \sum_{y=0}^2 E(X | Y = y) P(Y = y) = 3/4$.

489. En el segundo caso, los valores de X se restringen a 1, 2, 3, 4, 5 y la esperanza se reduce:

a) $E(X) = 3.5$.

b) $E(X | X + Y = 6) = 3$.

490. Las esperanzas $E(X | Y = 0)$ y $E(X)$ satisfacen la relación buscada cuando (X, Y) tiene distribución:

a)	<table style="border-collapse: collapse; text-align: center;"> <tr> <td style="padding: 0 10px;">$x \backslash y$</td> <td style="padding: 0 10px;">0</td> <td style="padding: 0 10px;">1</td> </tr> <tr> <td style="padding: 0 10px;">0</td> <td style="padding: 0 10px;">1/4</td> <td style="padding: 0 10px;">0</td> </tr> <tr> <td style="padding: 0 10px;">1</td> <td style="padding: 0 10px;">1/2</td> <td style="padding: 0 10px;">1/4</td> </tr> </table>	$x \backslash y$	0	1	0	1/4	0	1	1/2	1/4	b)	<table style="border-collapse: collapse; text-align: center;"> <tr> <td style="padding: 0 10px;">$x \backslash y$</td> <td style="padding: 0 10px;">0</td> <td style="padding: 0 10px;">1</td> </tr> <tr> <td style="padding: 0 10px;">0</td> <td style="padding: 0 10px;">1/4</td> <td style="padding: 0 10px;">1/4</td> </tr> <tr> <td style="padding: 0 10px;">1</td> <td style="padding: 0 10px;">1/4</td> <td style="padding: 0 10px;">1/4</td> </tr> </table>	$x \backslash y$	0	1	0	1/4	1/4	1	1/4	1/4	c)	<table style="border-collapse: collapse; text-align: center;"> <tr> <td style="padding: 0 10px;">$x \backslash y$</td> <td style="padding: 0 10px;">0</td> <td style="padding: 0 10px;">1</td> </tr> <tr> <td style="padding: 0 10px;">0</td> <td style="padding: 0 10px;">0</td> <td style="padding: 0 10px;">1/4</td> </tr> <tr> <td style="padding: 0 10px;">1</td> <td style="padding: 0 10px;">1/4</td> <td style="padding: 0 10px;">1/2</td> </tr> </table>	$x \backslash y$	0	1	0	0	1/4	1	1/4	1/2
$x \backslash y$	0	1																														
0	1/4	0																														
1	1/2	1/4																														
$x \backslash y$	0	1																														
0	1/4	1/4																														
1	1/4	1/4																														
$x \backslash y$	0	1																														
0	0	1/4																														
1	1/4	1/2																														

491. Para $0 < y < 1$,

$$E(X | Y = y) = \int_0^y x (12x^2)/(4y^3) dx = (3/4)y.$$

492.

$$E(X) = \sum_{\lambda=1}^n E(X | \Lambda = \lambda) P(\Lambda = \lambda) = \sum_{\lambda=1}^n \lambda (1/n) = (n+1)/2.$$

493.

$$E(X) = \sum_{n=0}^{\infty} E(X | N = n) P(N = n) = \sum_{n=0}^{\infty} (n+1) p q (1-q)^n = p/q.$$

494. a)

$$\begin{aligned} E(S_N) &= \sum_{n=1}^{\infty} E(S_N | N = n) P(N = n) \\ &= \sum_{n=1}^{\infty} E\left(\sum_{i=1}^N X_i | N = n\right) P(N = n) \\ &= \sum_{n=1}^{\infty} E\left(\sum_{i=1}^n X_i | N = n\right) P(N = n) \\ &= \sum_{n=1}^{\infty} E\left(\sum_{i=1}^n X_i\right) P(N = n) \\ &= \sum_{n=1}^{\infty} \sum_{i=1}^n E(X_i) P(N = n) \\ &= \sum_{n=1}^{\infty} n E(X_1) P(N = n) \\ &= E(N) E(X_1). \end{aligned}$$

b)

$$\begin{aligned}
E(S_N^2) &= \sum_{n=1}^{\infty} E(S_N^2 | N = n) P(N = n) \\
&= \sum_{n=1}^{\infty} E\left(\left(\sum_{i=1}^N X_i\right)^2 | N = n\right) P(N = n) \\
&= \sum_{n=1}^{\infty} E\left(\left(\sum_{i=1}^n X_i\right)^2 | N = n\right) P(N = n) \\
&= \sum_{n=1}^{\infty} E\left(\sum_{i=1}^n X_i^2\right) P(N = n) \\
&= \sum_{n=1}^{\infty} E\left(\sum_{i,j=1}^n X_i X_j\right) P(N = n) \\
&= \sum_{n=1}^{\infty} (nE(X_1^2) + n(n-1)E^2(X_1)) P(N = n) \\
&= E(X_1^2)E(N) + E(N^2)E^2(X_1) - E^2(X_1)E(N) \\
&= \text{Var}(X_1)E(N) + E(N^2)E^2(X_1).
\end{aligned}$$

Por lo tanto,

$$\begin{aligned}
\text{Var}(S_N) = E(S_N^2) - E^2(S_N) &= \text{Var}(X_1)E(N) + E(N^2)E^2(X_1) \\
&\quad - E^2(N)E^2(X_1) \\
&= \text{Var}(X_1)E(N) + \text{Var}(N)E^2(X_1).
\end{aligned}$$

495. El total de veces que se requiere la ambulancia en un día cualquiera puede modelarse como la variable aleatoria

$$S = \sum_{i=1}^N X_i,$$

en donde N es la variable aleatoria que representa el número de accidentes por día y X_1, X_2, \dots es una sucesión de v.a.s independientes con distribución $\text{Ber}(p)$ con $p = 2/3$, e independientes de N . Entonces

$$E(S) = E(N)E(X_1) = (3/2)(2/3) = 1.$$

4.9 Esperanza, varianza y covarianza

496. a) $\text{Cov}(X, Y) = 0$, pues X y Y son independientes.

$$b) \operatorname{Cov}(X, Y) = 2/5 - (5/8)^2 = 0.009375.$$

$$c) \operatorname{Cov}(X, Y) = 3/10 - (9/16)^2 = -0.01640625.$$

$$d) \operatorname{Cov}(X, Y) = 1.$$

$$e) \operatorname{Cov}(X, Y) = 1/4.$$

497. a) $\operatorname{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))]$
 $= E[XY - XE(Y) - YE(X) + E(X)E(Y)]$
 $= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y)$
 $= E(XY) - E(X)E(Y).$
- b) $\operatorname{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))]$
 $= E[(Y - E(Y))(X - E(X))]$
 $= \operatorname{Cov}(Y, X).$
- c) $\operatorname{Cov}(X, c) = E[(X - E(X))(c - E(c))] = E[(c - E(c))(X - E(X))] = 0.$
- d) $\operatorname{Cov}(cX, Y) = E[(cX - E(cX))(Y - E(Y))]$
 $= E[c(X - E(X))(Y - E(Y))]$
 $= E[(X - E(X))(cY - E(cY))]$
 $= \operatorname{Cov}(X, cY)$
 $= cE[(X - E(X))(Y - E(Y))]$
 $= c\operatorname{Cov}(X, Y).$
- e) $\operatorname{Cov}(X + c, Y) = E[(X + c - E(X + c))(Y - E(Y))]$
 $= E[(X - E(X))(Y + c - E(Y + c))]$
 $= \operatorname{Cov}(X, Y + c)$
 $= E[(X - E(X))(Y - E(Y))]$
 $= \operatorname{Cov}(X, Y).$
- f) $\operatorname{Cov}(X_1 + X_2, Y) = E[(X_1 + X_2 - E(X_1 + X_2))(Y - E(Y))]$
 $= E[(X_1 - E(X_1) + X_2 - E(X_2))(Y - E(Y))]$
 $= E[(X_1 - E(X_1))(Y - E(Y)) + (X_2 - E(X_2))(Y - E(Y))]$
 $= E[(X_1 - E(X_1))(Y - E(Y))] + E[(X_2 - E(X_2))(Y - E(Y))]$
 $= \operatorname{Cov}(X_1, Y) + \operatorname{Cov}(X_2, Y).$
- g) $\operatorname{Var}(X + Y) = E[(X + Y)^2] - [E(X + Y)]^2$
 $= E[X^2 + 2XY + Y^2] - [E^2(X) + 2E(X)E(Y) + E^2(Y)]$
 $= E(X^2) - E^2(X) + E(Y^2) - E^2(Y) + 2E(XY) - 2E(X)E(Y)$
 $= \operatorname{Var}(X) + \operatorname{Var}(Y) + 2\operatorname{Cov}(X, Y).$
- h) $\operatorname{Cov}(X, Y) = E(XY) - E(X)E(Y) = E(X)E(Y) - E(X)E(Y) = 0.$

- i) Considere la situación del Ejercicio 229 en la página 169. Las variables X y Y satisfacen la igualdad $E(XY) = E(X)E(Y)$, es decir, $\text{Cov}(X, Y) = 0$, y sin embargo X y Y no son independientes.

498. a) Para cualquier valor real de t se tiene que

$$\begin{aligned} 0 &\leq E[(X - E(X)) + t(Y - E(Y))]^2 \\ &= t^2 \text{Var}(Y) + 2t \text{Cov}(X, Y) + \text{Var}(X). \end{aligned}$$

En consecuencia, el discriminante de esta ecuación cuadrática debe ser menor o igual a cero. Esto lleva a la desigualdad

$$4 \text{Cov}^2(X, Y) - 4 \text{Var}(Y)\text{Var}(X) \leq 0.$$

Recordando que $\text{Var}(X) \leq (b - a)^2/4$, se obtiene el resultado buscado.

$$\begin{aligned} \text{Cov}^2(X, Y) &\leq \text{Var}(Y) \text{Var}(X) \\ &\leq ((b - a)^2/4)^2. \end{aligned}$$

- b) Tome $Y = X$ con $P(X = a) = P(Y = b) = 1/2$. Entonces $\text{Cov}(X, Y) = \text{Var}(X) = (b - a)^2/4$.

499. Recordando que $X \sim N(\mu_1, \sigma_1^2)$ y $Y \sim N(\mu_2, \sigma_2^2)$, tenemos que:

a) $E(X, Y) = (E(X), E(Y)) = (\mu_1, \mu_2)$.

b) $\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dx dy - \mu_1 \mu_2$.

Al considerar la integral respecto de x se completa el cuadrado en esa variable en la expresión del exponente. El procedimiento es similar al presentado en la solución del problema 460.

$$\begin{aligned} &-\frac{1}{2(1 - \rho^2)} \left[\frac{(x - \mu_1)^2}{\sigma_1^2} - \frac{2\rho}{\sigma_1 \sigma_2} (x - \mu_1)(y - \mu_2) \right] \\ &= -\frac{1}{2} \left(\frac{(x - (\mu_1 + (\rho\sigma_1/\sigma_2)(y - \mu_2)))^2}{\sigma_1^2(1 - \rho^2)} - \frac{\rho^2}{1 - \rho^2} \frac{(y - \mu_2)^2}{\sigma_2^2} \right). \end{aligned}$$

Entonces la integral respecto de x corresponde a la integral de x multiplicada por la función de densidad normal univariada con media $\mu_1 + (\rho\sigma_1/\sigma_2)(y - \mu_2)$ y varianza $\sigma_1^2(1 - \rho^2)$. Al incorporar la constante adecuada, esta integral es justamente la media $\mu_1 + (\rho\sigma_1/\sigma_2)(y - \mu_2)$ y la integral respecto de y se reduce a

$$\int_{-\infty}^{\infty} [\mu_1 + (\rho\sigma_1/\sigma_2)(y - \mu_2)] y f(y) dy,$$

en donde $f(y)$ es la función de densidad $N(\mu_2, \sigma_2^2)$. Es inmediato verificar que la integral anterior es $\mu_1\mu_2 + \rho\sigma_1\sigma_2$. Por lo tanto, $\text{Cov}(X, Y) = \rho\sigma_1\sigma_2$.

c) Por el inciso anterior,

$$\text{Var}(X, Y) = \begin{pmatrix} \text{Var}(X) & \text{Cov}(X, Y) \\ \text{Cov}(Y, X) & \text{Var}(Y) \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

4.10 Coeficiente de correlación

500. a) $\text{Cov}(X, Y) = 0$, pues X y Y son independientes.

b) $\text{Cov}(X, Y) = (2/5) - (5/8)^2 = 3/320$.

501. El discriminante de la ecuación cuadrática es necesariamente menor o igual a cero.

502. Ambas identidades son consecuencia inmediata de la aplicación de la fórmula (4.3) en la página 339 y del uso de las propiedades de la varianza y la covarianza. Como estas cantidades deben ser mayores o iguales a cero, cada uno de los paréntesis del lado derecho es mayor o igual a cero. Esto lleva a que $-1 \leq \rho(X, Y) \leq 1$.

503. Estas identidades se siguen de las propiedades de la covarianza y la varianza.

a) $\rho(X, X) = \text{Cov}(X, X)/\text{Var}(X) = \text{Var}(X)/\text{Var}(X) = 1$.

b) $\rho(X, -X) = \text{Cov}(X, -X)/\text{Var}(X) = -\text{Var}(X)/\text{Var}(X) = -1$.

c) $\rho(X, Y) = \text{Cov}(X, Y)/\sqrt{\text{Var}(X)\text{Var}(Y)}$
 $= \text{Cov}(Y, X)/\sqrt{\text{Var}(Y)\text{Var}(X)} = \rho(Y, X)$.

d) $\rho(cX, Y) = \text{Cov}(cX, Y)/\sqrt{\text{Var}(cX)\text{Var}(Y)}$
 $= \text{Cov}(X, cY)/\sqrt{\text{Var}(X)\text{Var}(cY)}$
 $= \rho(X, cY)$
 $= (c/|c|)\text{Cov}(X, Y)/\sqrt{\text{Var}(X)\text{Var}(Y)}$
 $= \text{signo}(c)\rho(X, Y)$.

e) $\rho(cX, cY) = \text{Cov}(cX, cY)/\sqrt{\text{Var}(cX)\text{Var}(cY)}$
 $= (c^2/|c|^2)\text{Cov}(X, Y)/\sqrt{\text{Var}(X)\text{Var}(Y)} = \rho(X, Y)$.

f) $\rho(X + c, Y) = \text{Cov}(X + c, Y)/\sqrt{\text{Var}(X + c)\text{Var}(Y)}$
 $= \text{Cov}(X, Y + c)/\sqrt{\text{Var}(X)\text{Var}(Y + c)}$
 $= \rho(X, Y + c)$

$$= \text{Cov}(X, Y) / \sqrt{\text{Var}(X) \text{Var}(Y)}$$

$$= \rho(X, Y).$$

$$g) \rho(X, aX + b) = \text{Cov}(X, aX + b) / \sqrt{\text{Var}(X) \text{Var}(aX + b)}$$

$$= (a/|a|) \text{Cov}(X, X) / \sqrt{\text{Var}(X) \text{Var}(X)}$$

$$= \text{signo}(a).$$

$$h) \rho(X + a, X + b) = \text{Cov}(X + a, X + b) / \sqrt{\text{Var}(X + a) \text{Var}(X + b)}$$

$$= \text{Cov}(X, X) / \sqrt{\text{Var}(X) \text{Var}(X)}$$

$$= \text{Var}(X) / \text{Var}(X)$$

$$= 1.$$

504. Recordando nuevamente que X y Y tienen distribución $N(\mu_1, \sigma_1^2)$ y $N(\mu_2, \sigma_2^2)$, respectivamente, y que $\text{Cov}(X, Y) = \rho \sigma_1 \sigma_2$, tenemos que:

$$a) \text{ Por definición, } \rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}} = \frac{\rho \sigma_1 \sigma_2}{\sigma_1 \sigma_2} = \rho.$$

b) Es inmediato verificar que si $\rho = 0$, entonces $f(x, y) = f_X(x) f_Y(y)$, en donde $f_X(x)$ y $f_Y(y)$ son las funciones de densidad $N(\mu_1, \sigma_1^2)$ y $N(\mu_2, \sigma_2^2)$, respectivamente. Por otra parte, si suponemos que se cumple la igualdad $f(x, y) = f_X(x) f_Y(y)$ para cualquier x y y , tomando el caso $x = \mu_1$ y $y = \mu_2$, tenemos que $2\pi\sigma_1\sigma_2\sqrt{1-\rho^2} = \sqrt{2\pi\sigma_1^2}\sqrt{2\pi\sigma_2^2}$. Esto implica que $\rho = 0$.

Capítulo 5: Teoremas límite

5.1 Desigualdad de Chebyshev

505. Suponga válida la desigualdad de Chebyshev. Entonces:

$$a) P(|X - \mu| \leq \epsilon) = 1 - P(|X - \mu| > \epsilon) \geq 1 - \sigma^2/\epsilon^2.$$

b) Substituyendo ϵ por $\epsilon\sigma$ en la desigualdad de Chebyshev se obtiene el resultado.

c) Substituyendo ϵ por $\epsilon\sigma$ en el primer inciso se obtiene el resultado.

506. $E(Z) = \epsilon^2 P((X - \mu)^2 \geq \epsilon^2) \leq \text{Var}(X)$, y de aquí se sigue inmediatamente la desigualdad de Chebyshev.

507. a) Utilice la misma técnica que se usó para demostrar la desigualdad de Chebyshev, esta vez empiece escribiendo la definición de esperanza para una variable aleatoria continua no negativa. Véase el siguiente inciso para una mayor ayuda.

b)

$$\begin{aligned}
 E(X^n) &= E(X^n 1_{(X \geq \epsilon)}) + E(X^n 1_{(X < \epsilon)}) \\
 &\geq E(X^n 1_{(X \geq \epsilon)}) \\
 &\geq \epsilon^n E(1_{(X \geq \epsilon)}) \\
 &= \epsilon^n P(X \geq \epsilon).
 \end{aligned}$$

c) Como la función φ es no negativa, la variable aleatoria $\varphi(X)$ es no negativa. Aplicando la desigualdad de Markov, como aparece en el primer inciso, tenemos que

$$P(\varphi(X) \geq \varphi(\epsilon)) \leq \frac{E(\varphi(X))}{\varphi(\epsilon)}.$$

Las condiciones impuestas sobre φ garantizan que la desigualdad $\varphi(X) \geq \varphi(\epsilon)$ sea equivalente a $X \geq \epsilon$.

508. Como e^{tX} es una variable aleatoria no negativa, tomando $\epsilon = e^{tx}$ en la desigualdad de Markov (5.2) se obtiene el siguiente resultado, el cual es equivalente a lo buscado.

$$P(e^{tX} \geq e^{tx}) \leq \frac{E(e^{tX})}{e^{tx}}.$$

509. Por la desigualdad de Chebyshev, para cualquier $\epsilon > 0$,

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2} = 0.$$

Esto es $P(|X - \mu| \geq \epsilon) = 0$ para cualquier $\epsilon > 0$. Equivalentemente, $P(|X - \mu| < \epsilon) = 1$ para cualquier $\epsilon > 0$. Sea $\epsilon_1, \epsilon_2, \dots$ cualquier sucesión monótona decreciente de números no negativos tal que $\epsilon_n \rightarrow 0$ cuando $n \rightarrow \infty$. Entonces la sucesión de eventos $A_n = (|X - \mu| < \epsilon_n)$ es decreciente, cada elemento tiene probabilidad uno y su límite es el evento $(X = \mu)$. Por la propiedad de continuidad de las medidas de probabilidad, tenemos que

$$P(X = \mu) = P(\lim_{n \rightarrow \infty} A_n) = \lim_{n \rightarrow \infty} P(A_n) = \lim_{n \rightarrow \infty} 1 = 1.$$

510. Aplique la desigualdad de Markov (5.2) a la variable aleatoria no negativa $|X - \mu|^2$ y parámetro ϵ^2 para obtener la desigualdad de Chebyshev (5.1).

511. Es inmediato comprobar que $\mu = 0$ y $\sigma^2 = 1/9$. Es decir, $\sigma = 1/3$. Por lo tanto, $P(|X - \mu| \geq 3\sigma) = P(|X| \geq 1) = 1/9$. Por otro lado, la cota superior para esta probabilidad dada por la desigualdad de Chebyshev es $\sigma^2/\epsilon^2 = \sigma^2 = 1/9$.

512. a) La condición mencionada se escribe $P(\mu - k\sigma < X < \mu + k\sigma) \geq 0.95$ y es equivalente a $P(|X - \mu| \geq k\sigma) \leq 0.05$. Por la desigualdad de Chebyshev, lo anterior se cumple si k es tal que $P(|X - \mu| \geq k\sigma) \leq 0.05 \leq \sigma^2/(k\sigma)^2 = 1/k^2$. Esto produce la condición $0.05 \leq 1/k^2$, o bien $k \geq 4.47$.
- b) La condición $P(|X - \mu| \geq k\sigma) = 0.05$ es equivalente a $P(|Z| \geq k) = 0.05$, o bien, $2(1 - \Phi(k)) = 0.05$, de donde se obtiene $\Phi(k) = 0.975$, es decir, $k = 1.96$.
513. a) $\Phi(x) = P(Z \leq x) = 1 - P(Z > x) = 1 - P(|Z| > x)/2 \geq 1 - 1/(2x^2)$.
- b) $\Phi(-x) = P(Z \leq -x) = (1/2)P(|Z| > x) \leq 1/(2x^2)$.
514. Sea X con distribución $\exp(\lambda)$ con $\lambda = 1$ y sea $x \geq 1$. Puede comprobarse que $P(|X - 1| \geq x) = e^{-(1+x)}$. Por otro lado, la cota superior de Chebyshev para esta probabilidad es $1/x^2$.
515. No es difícil comprobar que $\mu = 0$ y $\sigma^2 = 2$. El primer resultado se obtiene al observar que $f(x)$ es una función par y para el segundo resultado se observa que la integral correspondiente es el segundo momento de la distribución $\exp(\lambda)$ con $\lambda = 1$. El cálculo directo de la integral lleva a que $P(|X - \mu| \geq x) = e^{-x}$, y la cota superior de Chebyshev para esta probabilidad es $2/x^2$.

5.2 Convergencia de variables aleatorias

516. La funciones de distribución de las variables X_n y X son

$$F_{X_n}(x) = \begin{cases} 0 & \text{si } x < 0, \\ nx & \text{si } 0 \leq x < 1/n, \\ 1 & \text{si } x \geq 1/n. \end{cases} \quad F_X(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 & \text{si } x \geq 0. \end{cases}$$

Se necesita demostrar que $F_{X_n}(x) \rightarrow F_X(x)$ cuando $n \rightarrow \infty$ para $x \neq 0$. Si $x < 0$, ambas funciones se anulan y por lo tanto coinciden. Si $x > 0$, existe un número natural m , suficientemente grande, tal que $x > 1/m$ y por lo tanto, para $n \geq m$, $F_{X_n}(x) = 1 = F(x)$.

517. Demostraremos primero que la convergencia en probabilidad implica la convergencia en distribución. La función de distribución de la variable aleatoria constante $X = c$ es

$$F_X(x) = \begin{cases} 0 & \text{si } x < c, \\ 1 & \text{si } x \geq c, \end{cases}$$

cuya única discontinuidad aparece en $x = c$. Comprobaremos que $F_{X_n}(x)$ converge a 0 para $x < c$ y converge a 1 para $x > c$. Por hipótesis, cuando

$n \rightarrow \infty$ y para cualquier $\epsilon > 0$, $P(|X_n - c| > \epsilon) \rightarrow 0$. Esto significa que $P(X_n > c + \epsilon) \rightarrow 0$ y $P(X_n < c - \epsilon) \rightarrow 0$. Entonces, para cualquier $x < c$, existe $\epsilon > 0$ tal que $x < c - \epsilon < c$, y por lo tanto $F_{X_n}(x) = P(X_n \leq x) \leq P(X_n \leq c - \epsilon) \leq P(X_n < c - \epsilon/2) \rightarrow 0$. Para $x > c$, existe $\epsilon > 0$ tal que $c < c + \epsilon < x$, y por lo tanto $F_{X_n}(x) = P(X_n \leq x) \geq P(X_n \leq c + \epsilon) = 1 - P(X_n > c + \epsilon) \rightarrow 1$.

Supongamos ahora que X_n converge en distribución a c . Entonces, para cualquier $\epsilon > 0$, $P(|X_n - c| > \epsilon) = P(X_n > c + \epsilon) + P(X_n < c - \epsilon) \leq 1 - P(X_n \leq c + \epsilon) + P(X_n \leq c - \epsilon) = 1 - F_{X_n}(c + \epsilon) + F_{X_n}(c - \epsilon) \rightarrow 1 - 1 + 0 = 0$.

5.3 La ley de los grandes números

518. Puede usted basarse en el código de la Figura 5.2 que aparece en la página 358. En este código se utiliza la instrucción para generar valores al azar de la distribución binomial, pero eso puede fácilmente cambiarse por la distribución discreta o continua de su preferencia. Observe que el código es una implementación de la fórmula recursiva (5.3) para la variable S_n .

519. Si $\hat{\mu}$ y $\hat{\sigma}^2$ denotan los estimadores para μ y σ^2 , respectivamente, entonces se pueden proponer

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2.$$

520. Por la ley fuerte de los grandes números, cuando $n \rightarrow \infty$,

$$\ln \sqrt[n]{X_1 \cdots X_n} = \frac{1}{n} \sum_{i=1}^n \ln X_i \xrightarrow{c.s.} E(\ln X_1) = \mu,$$

y siendo $g(x) = e^x$ una función continua,

$$g(\ln \sqrt[n]{X_1 X_2 \cdots X_n}) \xrightarrow{c.s.} g(\mu).$$

521. Por la hipótesis de independencia, cuando $n \rightarrow \infty$,

$$-\frac{1}{n} \log_2 p(X_1, \dots, X_n) = -\frac{1}{n} \sum_{i=1}^n \log_2 p(X_i) \xrightarrow{c.s.} E(-\log_2 p(X_1)) = H(X_1).$$

522. Para cualquier $\epsilon > 0$, por la desigualdad de Chebyshev,

$$\begin{aligned} P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \mu\right| \geq \epsilon\right) &\leq \frac{1}{\epsilon^2} \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \\ &= \frac{1}{\epsilon^2 n^2} \sum_{i=1}^n \text{Var}(X_i) \\ &\leq \frac{c}{\epsilon^2 n} \rightarrow 0. \end{aligned}$$

Esto es la definición de la convergencia en probabilidad de la variable $(1/n) \sum_{i=1}^n X_i$ a la constante μ .

523. Para cualquier $\epsilon > 0$, por la desigualdad de Chebyshev,

$$P\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \mu\right| \geq \epsilon\right) \leq \frac{1}{\epsilon^2} \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \rightarrow 0.$$

Esta es nuevamente la definición de la convergencia en probabilidad de la variable $(1/n) \sum_{i=1}^n X_i$ a la constante μ .

5.4 El teorema central del límite

524. Puede usted basarse en el código de la Figura 5.4 que aparece en la página 364. En este código se utiliza la instrucción para generar valores al azar de la distribución Bernoulli, pero eso puede fácilmente cambiarse por la distribución discreta o continua de su preferencia.

525. Esta es una aplicación del teorema central del límite. La variable X se puede escribir como la suma $\sum_{i=1}^n X_i$, en donde X_1, \dots, X_n son v.a.s independientes con idéntica distribución $\text{Ber}(p)$. La media y varianza de X es $\mu = np$ y $\sigma^2 = np(1-p)$, respectivamente. Por lo tanto, $(X - np)/\sqrt{np(1-p)} \xrightarrow{d} Z$.

526. La variable X se puede escribir como la suma $\sum_{i=1}^n X_i$, en donde X_1, \dots, X_n son v.a.s independientes con idéntica distribución $\chi^2(k)$ con $k = 1$. La media y varianza de X es $\mu = n(1) = n$ y $\sigma^2 = n(2) = 2n$, respectivamente. Por lo tanto, $(X - n)/\sqrt{2n} \xrightarrow{d} Z$.

527. Sea $S_n = X_1 + \dots + X_n$. Entonces $E(S_n) = n\lambda$ y $\text{Var}(S_n) = n\lambda$.

$$\begin{aligned} a) \quad &P(a < X_1 + \dots + X_n < b) \\ &= P((a - n\lambda)/\sqrt{n\lambda} < (S_n - n\lambda)/\sqrt{n\lambda} < (b - n\lambda)/\sqrt{n\lambda}) \\ &\approx P((a - n\lambda)/\sqrt{n\lambda} < Z < (b - n\lambda)/\sqrt{n\lambda}) \\ &= \Phi((b - n\lambda)/\sqrt{n\lambda}) - \Phi((a - n\lambda)/\sqrt{n\lambda}). \end{aligned}$$

$$b) P(X_1 + \dots + X_n \geq n\lambda) = P((S_n - n\lambda)/\sqrt{n\lambda} \geq 0) \approx P(Z \geq 0) = 1/2.$$

528. Use el teorema de De Moivre-Laplace.

529. Sea X_i con distribución $\text{Ber}(p)$ con $p = 0.05$. Las variables independientes X_1, \dots, X_{100} toman el valor uno cuando el correspondiente componente falla. Por lo tanto, la suma $S_{100} = \sum_{i=1}^{100} X_i$ es el total de componentes que fallan. Entonces

$$\begin{aligned} a) P(S_{100} \geq 5) &= P((S_{100}/100 - p)/\sqrt{p(1-p)/100} \geq (5/100 - p)/\sqrt{p(1-p)/100}) \\ &\approx P(Z \geq 0) = 1/2. \end{aligned}$$

$$\begin{aligned} b) P(S_{100} \leq 4) &= P((S_{100}/100 - p)/\sqrt{p(1-p)/100} \leq (4/100 - p)/\sqrt{p(1-p)/100}) \\ &\approx \Phi(-0.4588) = 0.3264. \end{aligned}$$

$$\begin{aligned} c) P(5 \leq S_{100} \leq 10) &\approx P(0 \leq Z \leq (10/100 - p)/\sqrt{p(1-p)/100}) \\ &= P(0 \leq Z \leq 2.2941) = 0.489. \end{aligned}$$

530. Sean X_1, \dots, X_n los resultados de los $n = 200$ ensayos. Cada X_i tiene distribución $\text{Ber}(p)$ con $p = 0.7$. La totalidad de veces que ocurre el evento de interés en la sucesión de ensayos es la variable aleatoria $S_n = X_1 + \dots + X_n$. Por el teorema central del límite,

$$\begin{aligned} P(S_n \geq 101) &= P((S_n - np)/\sqrt{np(1-p)} \geq (101 - np)/\sqrt{np(1-p)}) \\ &\approx P(Z \geq -4.47) \approx 1. \end{aligned}$$

531. La frecuencia relativa es el cociente S_n/n , en donde $S_n = X_1 + \dots + X_n$, $n = 100$ y X_1, \dots, X_{100} son v.a.s independientes con idéntica distribución $\text{Ber}(p)$, con $p = 0.3$, que indican la ocurrencia o no ocurrencia del evento. Entonces

$$\begin{aligned} a) P(0.2 \leq S_n/n \leq 0.5) &= P(-2.1821 \leq (S_n/n - p)/\sqrt{p(1-p)/n} \leq 4.3643) \\ &\approx \Phi(4.3643) - \Phi(-2.1821) \approx 0.9854. \end{aligned}$$

$$\begin{aligned} b) P(S_n/n \geq 0.4) &= P((S_n/n - p)/\sqrt{p(1-p)/n} \geq 2.1821) \\ &\approx 1 - \Phi(2.1821) \approx 0.0146. \end{aligned}$$

$$\begin{aligned} c) P(S_n/n \leq 0.35) &= P((S_n/n - p)/\sqrt{p(1-p)/n} \leq 1.091) \\ &\approx \Phi(1.091) \approx 0.8621. \end{aligned}$$

532. Sean nuevamente X_1, \dots, X_n las variables aleatorias que indican el funcionamiento de los condensadores, es decir, cada X_i tiene distribución $\text{Ber}(p)$ con $p = 0.1$ y $n = 50$. El total de condensadores con falla durante su periodo de garantía es $S_n = X_1 + \dots + X_n$, cuya distribución exacta es $\text{bin}(n, p)$ si suponemos independencia entre las variables X_i . Por lo tanto,

$$a) P(3 \leq S_n \leq 6) = \sum_{k=3}^6 \binom{n}{k} p^k (1-p)^{n-k} = 0.6584.$$

Por otro lado,

$$P(3 \leq S_n \leq 6) = P(-0.9428 \leq (S_n - np)/\sqrt{np(1-p)} \leq 0.4714) \\ \approx \Phi(0.4714) - \Phi(-0.9428) \approx 0.5084.$$

$$b) P(S_n \geq 7) = \sum_{k=7}^{50} \binom{n}{k} p^k (1-p)^{n-k} = 0.2297.$$

Por otro lado,

$$P(S_n \geq 7) = P((S_n - np)/\sqrt{np(1-p)} \geq 0.9428) \\ \approx 1 - \Phi(0.9428) \approx 0.1729.$$

$$c) P(S_n \leq 2) = \sum_{k=0}^2 \binom{n}{k} p^k (1-p)^{n-k} = 0.1117.$$

Por otro lado,

$$P(S_n \leq 2) = P((S_n - np)/\sqrt{np(1-p)} \leq -1.4142) \\ \approx \Phi(-1.4142) \approx 0.0786.$$

Bibliografía

- [1] Albert J., Rizzo M. *R by example*. Springer, 2012.
- [2] Blake I. F. *An introduction to applied probability*. John Wiley & Sons, 1979.
- [3] Blomm G., Holst L., Sandell D. *Problems and snapshots from the world of probability*. Springer-Verlag, 1994.
- [4] Crawley M. J. *The R book*. Wiley, 2007.
- [5] Esparza Núñez S. *Elementos de probabilidad*. IPN, 1985.
- [6] Garza T. *Elementos de cálculo de probabilidades*. UNAM, 1983.
- [7] Granville W. A. *Cálculo diferencial e integral*. Limusa, 1982.
- [8] Godfrey M. G., Roebuck E. M., Sherlock A. J. *Concise statistics*. Edward Arnold, 1988.
- [9] Gut, A. *Probability: a graduate course*. Springer, 2005.
- [10] Hoel P. G., Port S. C., Stone C. J. *Introduction to probability theory*. Houghton Mifflin, 1971.
- [11] Kolmogorov A. N. *Foundations of the theory of probability*. Chelsea Publishing Company, 1950.
- [12] Mood A. M., Graybill F. A., Boes D. C. *Introduction to the theory of statistics*. McGraw Hill, 1983.
- [13] Miller I., Miller M. *John E. Freund's mathematical statistics*. Prentice Hall, 1999.

- [14] Olofsson P. *Probability, statistics, and stochastic processes*. Wiley, 2005.
- [15] Perero M. *Historia e historias de las matemáticas*. Grupo Editorial Iberoamérica, 1994.
- [16] Ross S. M. *A first course in probability*. Prentice Hall, 2009.
- [17] Ross S. M. *Introduction to probability and statistics for engineers and scientists*. Academic Press, 2009.
- [18] Ross S. M. *Simulation*. Academic Press, 2006.
- [19] Rincón L. *Curso intermedio de probabilidad*. Facultad de Ciencias, UNAM, 2007.
- [20] Rincón L. *Curso elemental de probabilidad y estadística*. Serie Textos, SMM, 2013.
- [21] Székely G. J. *Paradoxes in probability theory and mathematical statistics*. D. Reidel Publishing Company, 2001.
- [22] <http://www.wikipedia.org>
- [23] <http://www.r-project.org/>
- [24] <http://www-history.mcs.st-and.ac.uk/>

Índice analítico

- σ -álgebra, 46
 - de Borel, 50
- Álgebra, 47
- Bernoulli
 - ensayo, 208
- Bernoulli, J., 373
- Borelianos, 51
- Coefficiente
 - binomial, 60
 - de correlación, 341
 - multinomial, 62
- Combinaciones, 60
- Conjunto
 - s Borel medibles, 51
 - s ajenos, 11
 - s de Borel, 51
 - s operaciones, 8
 - potencia, 12
- Convergencia
 - casi segura (o fuerte), 350
 - débil, 351
 - de variables aleatorias, 349
 - en distribución (o débil), 351
 - en probabilidad, 350
 - puntual, 350
- Covarianza, 336
- Cuantiles, 182
- Cuartiles, 183
- De Morgan
 - leyes de, 11
- Densidad
 - conjunta, 297
- Desigualdad
 - de Chebyshev, 345
 - de Markov, 347
- Desviación estándar, 170
- Diagrama de árbol, 13
- Diferencia simétrica, 11, 16
- Distribución
 - Bernoulli, 208
 - beta, 263
 - binomial, 213
 - binomial negativa, 227
 - condicional, 327
 - conjunta, 308
 - de una v.a., 137
 - Erlang, 260
 - exponencial, 253
 - F, 287
 - gamma, 258
 - geométrica, 221
 - hipergeométrica, 234
 - ji-cuadrada, 278

- logarítmica, 178
- lognormal, 278
- marginal, 319
- mixta, 144
- normal, 269
- normal bivariada, 307
- normal estándar, 271
- Poisson, 238
- Rayleigh, 181
- t, 283
- uniforme continua, 248
- uniforme discreta, 204
- Weibull, 266
- Distribución de prob.
 - de una v.a., 137
- Distribuciones
 - condicionales, 333
- Ensayo Bernoulli, 208
- Espacio
 - de probabilidad, 52
 - equiprobable, 21
 - muestral, 5
- Esperanza, 153
 - de un vector, 336
 - de una fn. de una v.a., 156
 - de una v.a., 154
 - propiedades, 161
- Estandarización, 271
- Evento, 5
 - s ajenos, 11
 - compuesto, 6
 - simple, 6
- Experimento
 - aleatorio, 3
 - determinista, 3
- Fórmula
 - s de derivación, 378
 - s de integración, 378
 - s para exponentes, 376
 - s para logaritmos, 377
 - s para sumas, 377
 - de inclusión-exclusión, 45
 - de Stirling, 380
- Función
 - beta, 263
 - convexa, 167
 - de acumulación conjunta, 308
 - de dist. $N(0, 1)$, 272
 - de distribución, 129, 136
 - gamma, 258
 - indicadora, 17, 116, 210
 - signo, 168
- Función de densidad
 - condicional, 328
 - conjunta, 297, 298
 - de un vector, 297
- Función de distribución
 - bivariada, 309
 - condicional, 330
 - conjunta, 308
 - de un vector, 308
 - marginal, 319
 - mixta, 144
- Función de probabilidad, 118
 - condicional, 129, 328
 - conjunta, 293
 - marginal, 315
 - simétrica, 127, 181
- Función generadora
 - de momentos, 195
 - de probabilidad, 187

- Identidades trigonométricas, 377
- Imagen inversa, 108
 - definición, 115
 - propiedades, 115
- Independencia
 - condicional, 97
 - de dos eventos, 89
 - de variables aleatorias, 149, 321
 - de varios eventos, 91
- Kolmogorov, A. N., 104
- Laplace, P.-S., 102
- Lema de Abel, 380
- Ley de los grandes números, 353
- Leyes de De Morgan, 11
- Método de Montecarlo, 356, 357
- Media, 154
- Mediana, 183
- Medida de probabilidad, 36
 - continuidad, 99
 - otras propiedades, 43
- Moda, 186
- Momento
 - s, 178
 - s absolutos, 180
 - s absolutos centrales, 180
 - s centrales, 180
 - s generalizados, 180
- Montecarlo
 - método de, 357
- Notación *o* pequeña, 381
- Ordenaciones
 - con repetición, 57
 - sin repetición, 58
- Pérdida de memoria
 - dist. exponencial, 257
 - dist. geométrica, 225
- Paradoja de
 - la caja de Bertrand, 87
 - San Petersburgo, 166
- Permutaciones, 59
- Potencia de un conjunto, 12
- Principio de multiplicación, 56
- Probabilidad, 36
 - axiomática, 35
 - clásica, 20
 - condicional, 71
 - conjunta, 293
 - de Laplace, 21
 - de un evento, 20
 - frecuentista, 31, 356
 - geométrica, 23
 - marginal, 315
 - subjettiva, 34
- Problema de
 - el ladrón de Bagdad, 166
 - la aguja de Buffón, 27
 - la feria, 24
 - la ruina del jugador, 75
 - la varilla de metal, 29
 - las tres puertas, 83
 - los cumpleaños, 65
 - los dos amigos, 25
 - los tres prisioneros, 82
 - Monty Hall, 83
- Producto
 - cartesiano, 13
 - Regla del, 75
- Regla del producto, 75

- Sigma álgebra, 46
 - de Borel, 50
- Stirling, 380
- Suceso, 5
- Teorema
 - central del límite, 361
 - de Bayes, 83
 - de cambio de variable, 145
 - de De Moivre-Laplace, 363
 - de equipartición asintótica, 360
 - de probabilidad total, 76, 77
 - del binomio, 60
 - extensión, 68
 - del estadístico inconsciente, 156
 - del multinomio, 62
- Triángulo de Pascal, 61
- Urna de Polya, 75
- Valor
 - esperado, 154
 - promedio, 154
- Variable aleatoria, 107
 - continua, 113, 132
 - discreta, 113
 - distribución de una, 137
 - mixta, 144
- Varianza, 169
 - de una v.a., 170
 - de un vector, 337
 - propiedades, 172
- Vector aleatorio, 291
 - continuo, 291
 - discreto, 291