Introduction to Probability

Charles M. Grinstead
Swarthmore College

J. Laurie Snell
Dartmouth College

Version 3.0.1, 22 July 2003[1]

To our wives
and in memory of
Reese T. Prosser

Contents

Copying

As indicated in the copyright notice above, this work is freely redistributable under the terms of the GNU General Public License as published by the Free Software Foundation. This work comes with ABSOLUTELY NO WARRANTY, to the extent permitted by applicable law. For a discussion of what all this means, please see the appendix on 'GNU Math'.

Preface

La teoría de la probabilidad comenzó en Francia del siglo XVII cuando los dos grandes matemáticos franceses, Blaise Pascal y Pierre de Fermat, concordaron sobre dos problemas en juegos de azar. Problemas como aquellos resueltos por Pascal y Fermat, continuaron influenciando a investigadores tempranos tales como Huygens, Bernoulli, y DeMoivre para establecer una teoría matemática de probabilidades. Hoy, la teoría de la probabilidades es una rama de la matemática bien establecida que encuentra aplicaciones en todas las áreas de la actividad académica desde la música a la física, y en experiencias cotidianas que van desde el pronóstico meteorológico hasta la predicción de riesgos en los nuevos tratamientos médicos.

Este texto esta diseñado a modo de curso de introducción a la Probabilidad para estudiantes que cursan los últimos tres años de programas en matemática, física, ciencias sociales, ingeniería y ciencias de la computación en universidades americanas. Presenta un tratamiento completo de ideas y técnicas de la Probabilidad necesarias para una firme comprensión de la misma. Los contenidos pueden utilizarse en cursos de distintas extensiones y niveles, enfatizando en áreas particulares.

Para cursos estándar de un único período, donde se cubran probabilidad discreta y continua, los estudiantes deberán haber cursado, como requisito previo, dos período de cálculo, incluyendo introducción a integrales múltiples. Para cubrir el Capítulo 11, el cual contiene material relacionado con cadenas de Markov, es necesario conocimiento de teoría de matrices.

Este libro también puede usarse en cursos de probabilidad discreta. El material ha sido organizado de manera tal, que los tratamientos de probabilidad discreta y continua se encuentran presentes por separado, aunque en forma paralela. Esta organización disipa un enfoque excesivamente formal o riguroso de la materia y ofrece fuertes valores pedagógicos en donde discusiones sobre probabilidad discreta pueden, ocasionalmente, motivar las discusiones de probabilidad continua más abstractas. Estudiantes de probabilidad discreta, deberán haber cursado un período de cálculo como requisito previo.

Un conocimiento muy básico de computación es necesario para obtener todos los beneficios que otorga el material informático incluido. Cada uno de los programas mencionados en el texto ha sido escrito en los lenguajes TrueBASIC, Maple y Mathematica

Este libro esta disponible en la Web como parte del Chance Project, dedicado a proporcionar material para cursos elementales en probabilidad y estadística. Los programas de computación, soluciones a los ejercicios de números impares y errata actual también se encuentran disponibles en este sitio Web. Los instructores pueden obtener todas las soluciones de los autores escribiendo a jlsnell@dartmouth edu o cgrinst@swarthmore edu.

CARACTERISTICAS

Nivel de rigor y énfasis: La Probabilidad es un campo de la Matemática maravillosamente intuitivo y aplicable. Hemos tratado de no estropear su belleza presentando demasiada formalidad matemática. En cambio, hemos tratado de desarrollar las ideas claves en un estilo algo ocioso, para presentar una variedad de aplicaciones interesantes y para mostrar algunos de los ejemplos menos intuitivos que hacen a la Probabilidad una materia viva.

Ejercicios: Hay más de 600 ejercicios en este texto que presentan muchas oportunidades de practicar técnicas y desarrollar un entendimiento contundente de las ideas. En la lista hay ejercicios de rutina para resolver con o sin el uso de una computadora y ejercicios más teóricos para mejorar la comprensión de los conceptos básicos. Los ejercicios más dificiles están indicados con un asterisco. Un manual de soluciones para todos los ejercicios está disponible para los instructores.

Reseñas históricas: La Probabilidad Introductoria es una materia en la cual las ideas fundamentales aún se encuentran intimamente ligadas a aquellas de sus fundadores. Por esta razón, existen numeroso comentarios históricos en el texto, que tratan especialmente con el desarrollo de la probabilidad discreta.

Uso pedagógico de los programas de computadora: La teoría de la Probabilidad hace predicciones hacerca de experimentos cuyos resultados dependen del azar. En consecuencia, ésta se presta con belleza, como una herramienta matemática, al uso de las computadoras para simular y analizar experimentos de azar.

En el texto la computadora se usa de muchas maneras. Primero, proporciona un laboratorio donde los experimentos de azar pueden ser simulados y los estudiantes pueden tomar conciencia de la variedad de tales experimentos. Este uso de la computadora en probabilidades ya fué perfectamente ilustrado por William Feller en la segunda edición de su famoso texto *An Introduction to Probability Theory and Its Applications* (New York: Wiley, 1950). En el prefacio, Feller describió su tratamiento de la fluctuación al arrojar una moneda: “Los resultados son tan asombrosos y tan en desacuerdo con la intuición común que aún colegas sofisticados dudaron que las monedas realmente se comportaran de la manera de que la teoría predice. Por consiguiente, se incluye el registro.”

Además de proporcionar un laboratorio para el estudiante, la computadora es una herramienta poderosa para comprender los resultados básicos de la teoría de probabilidades. Por ejemplo, la representación gráfica de la aproximación de las distribuciones binomiales estandarizadas a la curva normal es una demostración más convincente del Teorema Central del Límite que muchas de las pruebas formales para este resultado fundamental.

Finalmente, la computadora permite que el estudiante resuelva problemas que no se prestan para fórmulas closed-form tales como tiempo de espera en colas. En realidad, la introducción de la computadora cambia la forma en que miramos muchos de los problemas de probabilidades. Por ejemplo, pudiendo calcular las probabilidades binomiales exactas para experimentos de hasta 1000 pruebas cambia nuestra manera de ver las aproximaciones normales y de Poisson.

RECONOCIMIENTOS

Quienquiera que escriba un libro de probabilidad en la actualidad, le debe un gran reconocimiento a William Feller, que nos enseñó a todos de qué manera dar vida a la Probabilidad como materia de estudio. Si encuentras un ejemplo, una aplicación o un ejercicio que realmente te agrade, probablemente tenga su origen en el texto clásico de Feller, *An Introduction to Probability and Its Applications.*

Estamos endeudados con mucha gente por su colaboración en esta tarea. El acercamiento a las Cadenas de Markov presentado en el libro fue desarrollada por John Kemeny y el segundo autor. Reese Prosser fue un coautor silencioso del material de probabilidades continuas en una versión previa de este libro. Mark Kernighan contribuyó con 40 páginas de comentarios acerca de la versión anterior. Muchos de estos comentarios provocaron nuestro pensamiento; además, proporcionaron la perspectiva del estudiante con respecto a este trabajo. Muchos de los cambios importantes que aparecen en esta versión tienen su génesis en dichas notas.

Fuxing Hou y Lee Nave dieron su colaboración extensiva en la composición y las figuras. John Finn proporcionó consejos pedagógicos de gran valor acerca del texto y los programas de computadora. Kart Knaub y Jessica Sklar son responsables de la implementación de los programas en Mathematica y Maple. Jessica y Gang Wang asistieron con las soluciones.

Finalmente, agradecemos a la American Mathematical Society, y en particular a Sergei Gelfang y John Swing, por su interés en este libro; su colaboración en la producción y su disposición a hacer el trabajo libremente redistribuible.

Chapter 1

Discrete Probability Distributions

1.1 Simulation of Discrete Probabilities

Probabilidad

En este capítulo, comenzaremos por considerar experimentos de probabilidad con un número finito de resultados posibles $\omega_1, \omega_2, \ldots, \omega_n$. Por ejemplo, arrojamos un dado con resultados posibles 1, 2, 3, 4, 5 y 6 según que cara acabe hacia arriba. Arrojamos una moneda con resultados posibles cara y cruz (H y T).

Con frecuencia, poder referirse a un resultado de un experimento es de utilidad. Por ejemplo, si quisiéramos escribir la expresión matemática que devuelve la suma de los valores obtenidos al arrojar un dado cuatro veces, podríamos enunciar que X_i, $i = 1, 2, 3, 4$, representan los valores de los resultados, y entonces tendríamos

$$X_1 + X_2 + X_3 + X_4$$

para la suma de las cuatro tiradas. Las X_i's se denominan *variables aleatorias*. Una variable aleatoria es simplemente una expresión cuyo valor es el resultado de un experimento en particular. Tal como es el caso de otros tipos de variables matemáticas, las variables aleatorias pueden adoptar diferentes valores.

Sea X la variable aleatoria que representa el valor obtenido al arrojar un dado. Estipularemos probabilidades para los resultados posibles de este experimento. Esto se hace asignando a cada resultado ω_j un número no negativo $m(\omega_j)$ de forma tal que

$$m(\omega_1) + m(\omega_2) + \cdots + m(\omega_6) = 1 \ .$$

La función $m(\omega_j)$ se denomina *función de distribución* de la variable aleatoria X. Para el caso de arrojar un dado, daremos probabilidades iguales o probabilidades $1/6$ a cada uno de los resultados posibles. Con esta asignación de probabilidades, se puede enunciar

$$P(X \leq 4) = \frac{2}{3}$$

para expresar que la probabilidad de que al arrojar un dado no se obtengan valores mayores que 4 es $2/3$

Sea Y la variable aleatoria que representa el valor obtenido al arrojar una moneda. En este caso, hay dos resultados posibles, que podemos etiquetar como H y T. A menos que haya razones para sospechar que una de las superficies sale con más frecuencia que la otra, es natural que asignemos probabilidad $1/2$ a cada uno de los dos resultados posibles.

En los dos experimentos anteriores, a cada resultado se le asigna probabilidad igual. En general, éste no será el caso. Por ejemplo, si se descubre que una droga es efectiva el 30 por ciento de las veces que se la utiliza, podríamos asignar probabilidad 0,3 a que la droga sea efectiva la próxima vez que se la utilice y 0,7 a que sea inefectiva. Este último ejemplo ilustra el *concepto de frecuencia de la Probabilidad* intuitivo. Esto es, si tenemos probabilidad p de que el resultado de un experimento sea A, entonces si lo repetimos muchísimas veces deberíamos esperar que la fracción de las ocurrencias de A sea cercana a p. Para chequear ideas intuitivas como ésta, veremos que es de gran ayuda observar algunos de estos problemas experimentalmente. Por ejemplo, podríamos arrojar una moneda una gran cantidad de veces y ver si la fracción de caras que se obtiene es de alrededor de $1/2$. Asimismo, podríamos simular este experimento en una computadora.

Simulación

Queremos ser capaces de llevar a cabo un experimento que corresponde a un conjunto de probabilidades dado; por ejemplo, $m(\omega_1) = 1/2$, $m(\omega_2) = 1/3$, y $m(\omega_3) = 1/6$. En este caso, uno podría marcar tres caras de un dado de seis lados con un ω_1, dos con un ω_2, y una con un ω_3.

En general, asumimos que $m(\omega_1), m(\omega_2), \ldots, m(\omega_n)$ son todos números racionales, con denominador común menor n. Si $n > 2$, podemos imaginar un largo dado cilíndrico con una sección transversal que es un n-gono regular. Si $m(\omega_j) = n_j/n$, entonces podemos etiquetar n_j de las caras largas del cilindro con un ω_j, y si sale alguna de las caras base, simplemente tiramos de nuevo. Si $n = 2$, podemos usar una moneda para efectuar el experimento.

Estaremos particularmente interesados en repetir un experimento de azar una gran cantidad de veces. Aunque el dado cilíndrico sería conveniente para llevar a cabo unas pocas repeticiones, dificultaría la realización de grandes cantidades de pruebas. Dado que una computadora moderna puede ejecutar muchas operaciones en un período de tiempo muy corto, es natural que nos volquemos a ella para esta tarea.

Números Aleatorios

Primero debemos encontrar una analogía informática a la acción de arrojar un dado. Esto se realiza en la computadora por medio de un *generador de números aleatorios.* Dependiendo del paquete de software en particular, podemos preguntar a la computadora por un número real entre 0 y 1, o por un entero dentro de un rango de enteros consecutivos. En el primer caso, los números reales se eligen de forma

.203309	.762057	.151121	.623868
.932052	.415178	.716719	.967412
.069664	.670982	.352320	.049723
.750216	.784810	.089734	.966730
.946708	.380365	.027381	.900794

Table 1.1: Muestra de la salida del programa **RandomNumbers**.

tal que la probabilidad de que el resultado pertenezca a cualquier subintervalo de este intervalo unidad es igual a la longitud de dicho subintervalo. En el segundo caso, cada entero tiene la misma probabilidad de ser elegido.

Sea X una variable aleatoria con función de distribución $m(\omega)$, donde ω pertenece al conjunto $\{\omega_1, \omega_2, \omega_3\}$, y $m(\omega_1) = 1/2$, $m(\omega_2) = 1/3$, y $m(\omega_3) = 1/6$. Si nuestro paquete de software puede devolver un entero aleatorio del conjunto $\{1, 2, ..., 6\}$, simplemente le pedimos que lo haga, y hacemos corresponder 1, 2, y 3 a ω_1, 4 y 5 a ω_2, y 6 a ω_3. Si nuestro paquete de software devuelve un número real aleatorio r en el intervalo $(0,\ 1)$, entonces la expresión

$$\lfloor 6r \rfloor + 1$$

será un entero aleatorio entre 1 y 6. (La notación $\lfloor x \rfloor$ significa el mayor entero que no excede a x, y se lee "piso x.")

El método por el cual se generan números aleatorios en computadora se trata en la discusión histórica al final de esta sección. El siguiente ejemplo es una muestra de la salida del programa **RandomNumbers**.

Ejemplo 1.1 (Random Number Generation) El programa **RandomNumbers** genera n números reales aleatorios en el intervalo $[0, 1]$, donde n es elegido por el usuario. Cuando ejecutamos el programa con $n = 20$, obtuvimos los resultados que se muestran en la Tabla 1.1. □

Ejemplo 1.2 (Coin Tossing) Como hemos observado, nuestra intuición sugiere que la probabilidad de obtener una cara al arrojar la moneda una única vez es $1/2$. Para que la computadora haga esta prueba, le pedimos que elija un número al azar en el intervalo $[0, 1]$ y verifique si el mismo es menor que $1/2$. De ser así, llamaremos al resultado *cara*; y en caso contrario lo llamaremos *cruz.* Otra manera de proceder sería solicitando a la computadora que elija un número entero al azar del intervalo $\{0, 1\}$. El programa **CoinTosses** lleva a cabo el experimento de arrojar una moneda n veces. Ejecutando este programa, con $n = 20$, obtuvimos el siguiente resultado:

THTTTHTTTTHTTTTTHHTT.

Observar que en 20 tiradas, obtuvimos 5 caras y 15 cruces. Arrojemos una moneda n veces, donde n sea mucho más grande que 20, y veamos si obtenemos una proporción de caras más cercana a nuestra estimación intuitiva de $1/2$. El programa **CoinTosses** sigue la cuenta del número de caras. Cuando ejecutamos este programa con $n = 1000$, obtuvimos 494 caras. Cuando lo ejecutamos con $n = 10000$, obtuvimos 5039 caras.

Notamos que cuando arrojamos la moneda 10,000 veces, la proporción de caras fue más cercana al valor real, es decir, 0,5. Un modelo matemático para este experimento se denomina Pruebas de Bernoulli (Bernoulli Trials, ver Capítulo 3). La *Ley de los Números Grandes,* que estudiaremos en el más adelante (ver Capítulo 8), mostrará que en el modelo de las Pruebas de Bernoulli, la proporción de caras debería ser cercana a 0,5; consistente con nuestra idea intuitiva de la interpretación de frecuencia de la Probabilidad.

Por supuesto, nuestro programa podría modificarse fácilmente para simular monedas para las cuales la probabilidad de obtener caras sea p, donde p, es un número real entre 0 y 1. □

En el caso de arrojar una moneda, ya conocemos la probabilidad del evento que ocurre en cada experimento. El verdadero poder de la simulación proviene de la habilidad de estimar probabilidades cuando éstas no se conocen de antemano. Este método se ha utilizado en descubrimientos recientes de estrategias que hacen que el juego de BlackJack sea favorable al jugador. Ilustramos esta idea en una situación simple, en la cual podemos computar la probabilidad real y ver que tan efectiva es la simulación.

Ejemplo 1.3 (Tirada de dados) Consideramos un juego de dados que jugó un papel importante en el desarrollo histórico de la Probabilidad. Las famosas cartas entre Pascal y Fermat, que muchos creen que dieron inicio a un estudio serio de esta rama de la matemática, fueron instigadas por una solicitud de ayuda del noble francés y jugador, Chevalier de Méré. Se cuenta que de Méré venía apostando a que al arrojar un dado cuatro veces, se obtendría al menos un seis. Como ganaba consistentemente, para conseguir más oponentes, cambió el juego apostando a que al arrojar dos dados 24 veces, se obtendría un par se seis. Se afirma que de Méré perdió con 24 y pensó que 25 tiradas serían necesarias para que el juego le fuera favorable. Fue *un grand scandale* que la matemática estuvo equivocada.

Intentaremos ver que de Méré está en lo cierto al simular sus varias apuestas. El programa **DeMere1** simula un gran número de experimentos viendo, en cada caso, si sale un seis al arrojar un dado cuatro veces. Cuando ejecutamos este programa para 1000 jugadas, obtuvimos un seis dentro de las primeras cuatro tiradas el 48,6 por ciento de las oportunidades. Cuando lo ejecutamos para 10.000 jugadas, ocurrió en el 51,98 por ciento de los casos.

Observamos que el resultado de la segunda ejecución sugiere que de Méré tenía razón en creer que su apuesta con un dado era favorable; sin embargo, si hubiéramos basado nuestra suposición en la primera ejecución, hubiéramos concluido que estaba equivocado. *Resultados precisos por simulación requieren una gran cantidad de experimentos.* □

El programa **DeMere2** simula la segunda apuesta de de Méré, que un par de seis saldrá al arrojar un par de dados n veces. La simulación anterior muestra que es importante saber cuantas pruebas deberíamos simular para esperar cierto nivel de precisión en nuestra aproximación. Veremos más adelante que en estos tipos de experimentos, una tosca regla del pulgar dice que, al menos en el 95% de los casos, el error no excede el recíproco de la raíz cuadrada del número de pruebas. Por fortuna, para este juego de dados, será sencillo computar las probabilidades exactas. Demostraremos en la siguiente sección que para la primera apuesta, la probabilidad de que de Méré gane es $1 - (5/6)^4 = .518$.

Se puede entender este cálculo de la siguiente manera: la probabilidad de que no salga un 6 en la primera tirada es $(5/6)$. La probabilidad de que no salga un 6 en ninguna de las dos primeras tiradas es $(5/6)^2$. Razonando de la misma manera, la probabilidad de que no salga un 6 en ninguna de las cuatro primeras tiradas es $(5/6)^4$. Así, la probabilidad de obtener al menos un 6 en las primeras cuatro tiradas es $1-(5/6)^4$. Asimismo, para la segunda apuesta, con 24 tiradas, la probabilidad de que de Méré gane es $1 - (35/36)^{24} = .491$, y con 25 tiradas es $1 - (35/36)^{25} = .506$.

Utilizando la regla del pulgar mencionada anteriormente, se requeriría de 27.000 tiradas para tener una chance razonable de determinar estas probabilidades con suficiente precisión como para afirmar que están en lados opuestos del 0,5. Es interesante ponderar si un jugador puede detectar tales probabilidades con la exactitud requerida, basado en su experiencia en el juego. Algunos escritores de la historia de la probabilidad sugieren que, en realidad, de Méré sólo estaba interesado en estas cuestiones porque eran intrigantes problemas de probabilidad.

Ejemplo 1.4 (Heads or Tails) Para nuestro ejemplo siguiente, consideremos un problema donde la respuesta correcta es difícil de obtener, pero para el cual la simulación fácilmente nos da resultados cualitativos. Peter y Paul juegan un juego llamado *caras o cruces.* En este juego, se arroja una moneda regular una secuencia de veces—elegimos 40. Cada vez que sale cara Peter gana un centavo de Paul, y cada vez que sale cruz Peter entrega un centavo a Paul. Por ejemplo, si los resultados de 40 tiradas son

$$\text{THTHHHHTTTHTHHTTHHTTTTHHHTHHTHHHTHHHTTTHH.}$$

las victorias de Peter se pueden graficar como en la Figura 1.1.

Peter ha ganado 6 centavos en este juego en particular. Es natural preguntarse por la probabilidad de que gane j centavos; aquí j podría ser cualquier número par desde -40 hasta 40. Es razonable suponer que el valor de j con la mayor probabilidad es $j = 0$, ya que esto ocurre cuando el número de caras es igual al número de cruces. Igualmente, supondríamos que los valores de j con la menor probabilidad son $j = \pm 40$.

La segunda pregunta interesante relacionada con este juego es la siguiente: En cuántas oportunidades, de las 40 tiradas, irá ganando Peter? Observando la gráfica de sus victorias (Figura 1.1), vemos que Peter va ganando cuando sus victorias son positivas, pero tenemos que hacer algún tipo de convención para cuando éstas sean 0 si queremos que todas las tiradas contribuyan al número de veces que él se encuentra al frente. Convenimos que cuando las victorias de Peter sean 0, él estará al frente si es que iba ganando en la tirada anterior y no lo estará si venía perdiendo en dicha instancia. Con esta convención, Peter está al frente en 34 oportunidades

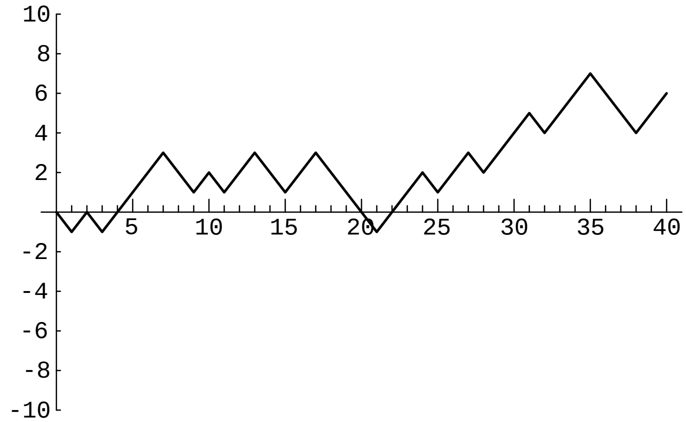


Figure 1.1: Victorias de Peter en 40 juegos de caras y cruces.

en nuestro ejemplo. Nuevamente, nuestra intuición podría sugerir que el número de veces al frente más probable es 1/2 de 40, ó 20, y los números menos probables son los casos extremos de 40 ó 0.

Esto es fácil de establecer simulando el juego una gran cantidad de veces y siguiendo el rastro al número de veces en que las victorias finales de Peter sean j, y al número de oportunidades en que él termine estando al frente por una diferencia de k. Las proporciones sobre todos los juegos dan, entonces, estimaciones para las probabilidades correspondientes. El programa **HTSimulation** lleva a cabo esta simulación. Notar que cuando el número de tiradas de un juego es par, sólo es posible estar al frente una cantidad par de oportunidades. Hemos simulado este juego 10.000 veces. Los resultados se muestran en las Figuras 1.2 y 1.3. Estos gráficos, que llamamos gráficos de puntas , fueron generados utilizando el programa **Spikegraph**. La línea vertical, o punta (spike), en la posición x sobre el eje horizontal, tiene una altura igual a la proporción de resultados iguales a x. Nuestra intuición acerca de las victorias finales de Peter fue bastante correcta, pero al intuir la cantidad de veces en que Peter iría ganando estuvimos completamente equivocados. La simulación sugiere que el número de veces menos probable de ir al frente es 20 y el más probable es 0 ó 40. La explicación para esto, que de hecho es correcto, se sugiere jugando el juego de caras y cruces con un número de tiradas muy grande y mirando la gráfica de las victorias de Peter. En la Figura 1.4 mostramos los resultados de una simulación del juego para 1000 tiradas, y en la Figura 1.5 para 10.000.

En el segundo ejemplo Peter estuvo al frente casi todo el tiempo. Es un hecho a remarcar, sin embargo, que si el juego se prolonga por el tiempo suficiente, las victorias de Peter continuarán volviendo a 0, pero habrá largos períodos de tiempo entre las veces en que esto ocurrirá. Estos y otros resultados relacionados se discutirán en Chapter 12. □

En todos los ejemplos que vimos hasta ahora, hemos simulado resultados igualmente probables. Ilustramos a continuación un ejemplo en el cual los resultados no

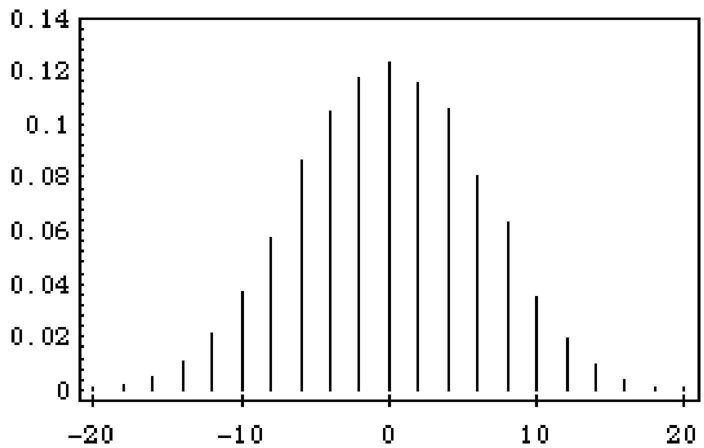


Figure 1.2: Distribución de las victorias.

son igualmente probables.

Ejemplo 1.5 (Carreras de caballos) Cuatro caballos (Acorn, Balky, Chesnut y Dolby) han competido muchas veces. Se estima que Acorn gana el 30 por ciento de las veces, Balky el 40, Chesnut el 20, y Dolby el 10.

Podemos hacer que nuestra computadora lleve a cabo una carrera de la siguiente forma: Se elige un número aleatorio x. Si $x < 0,3$ entonces decimos que ganó Acorn. Si $0,3 \leq x < 0,7$ entonces gana Balky. Si $0,7 \leq x < 0,9$ entonces gana Chesnut. Finalmente, si if $0,9 \leq x$ entonces gana Dolby.

El programa **HorseRace** usa este método de simulación de resultados de n carreras. Ejecutando este programa para $n = 10$ hallamos que Acorn ganó el 40 por ciento de las veces, Balky el 20 por ciento, Chesnut el 10 por ciento y Dolby el 30 por ciento. Sería necesaria una mayor cantidad de carreras para obtener un resultado más acorde con la experiencia pasada. Por lo tanto, ejecutamos el programa para simular 1000 carreras con nuestros cuatro caballos. Aunque quedaron muy cansados de tanto correr, se comportaron de forma bastante consistente con nuestras estimaciones de sus habilidades. Acorn ganó el 29,8 por ciento; Balky el 39,4 por ciento; Chestnut el 19,5 por ciento y Dolby 11,3 por ciento de las oportunidades.

El programa **GeneralSimulation** utiliza este método para simular repeticiones de un experimento arbitrario con un número finito de resultados que ocurren con probabilidades conocidas. $\square$

Resea histórica

Cualquiera que juega al mismo juego de azar una y otra vez, ciertamente está llevando a cabo una simulación, y en este sentido el proceso de simulación se ha venido dando durante siglos. Como hemos remarcado, muchos de los problemas de

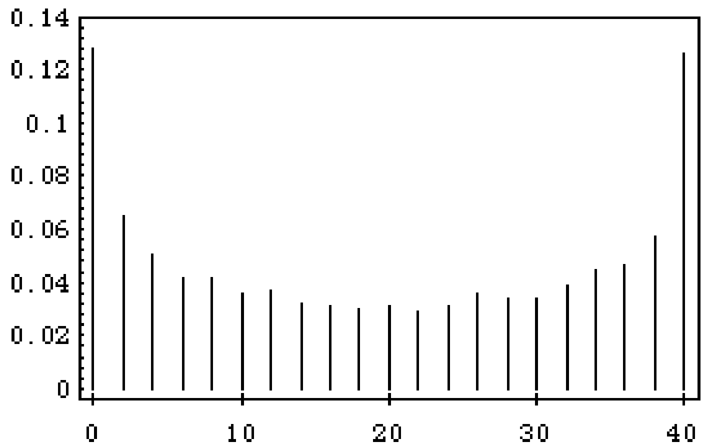


Figure 1.3: Distribución de la cantidad de veces al frente.

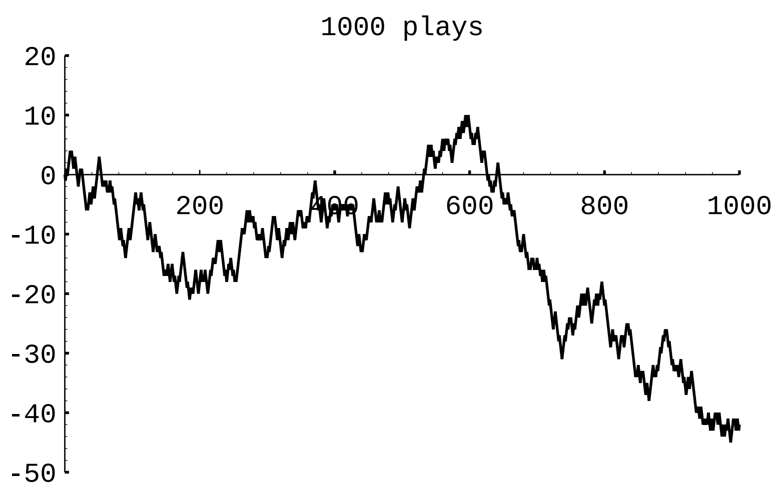


Figure 1.4: Victorias de Peter en 1000 juegos de caras y cruces.

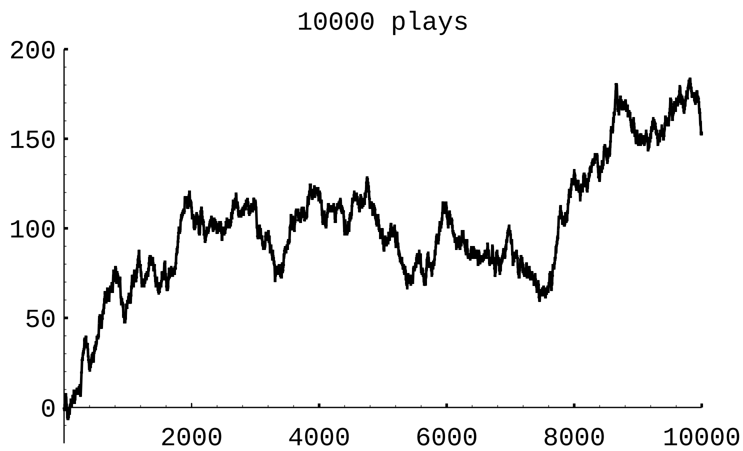


Figure 1.5: Victorias de Peter en 10.000 juegos de caras y cruces.

probabilidad tempranos bien pudieron haber sido sugeridos por jugadores experimentados.

Es natural para quien intenta comprender la teoría de la Probabilidad ensayar experimentos simples arrojando monedas, dados y demás. El naturalista Buffon arrojó una moneda 4040 veces, obteniendo como resultado 2048 caras y 1992 cruces. También estimó el número π tirando agujas sobre una superficie rayada y registrando cuántas veces las agujas cruzaban una línea (see Section 2.1). El biólogo inglés W. F. R. Weldon[1] registró 26.306 tiradas de 12 dados, y el científico suizo Rudolf Wolf[2] registró 100.000 tiradas de un sólo dado sin una computadora. Tales experimentos consumen mucho tiempo y puede que no representen con precisión el fenómeno de probabilidad en estudio. Por ejemplo, para los experimentos de Weldon y Wolf, un análisis posterior de los datos registrados mostró una tendencia sospechosa en los dados. El estadístico Karl Pearson analizó una gran cantidad de resultados en ciertas mesas de ruleta y sugirió que las ruedas estaban predispuestas. El escribió en 1894:

> Claramente, dado que el Casino no está al valioso servicio de la estadística probabilística, no posee una *raison d'être* (razón de ser) científica. Los hombres de la ciencia no pueden dejar que se desconsideren sus más refinadas teorías de una manera tan desvergonzada! El gobierno francés debe ser instado por la jerarquía de la ciencia a cerrar los salones de juego; esto sería, por supuesto, un acto cortés para entregar los recursos restantes del Casino a la Académie des Sciences para la dotación de un laboratorio de probabilidad ortodoxa; en particular, de la nueva rama de este campo de estudio, la aplicación de la teoría de la probabilidad a los problemas biológicos de la evolución, que es de esperar que ocupe

[1] T. C. Fry, *Probability and Its Engineering Uses,* 2nd ed. (Princeton: Van Nostrand, 1965).
[2] E. Czuber, *Wahrscheinlichkeitsrechnung,* 3rd ed. (Berlin: Teubner, 1914).

gran parte del pensamiento de los hombres en un futuro cercano.[3]

Sin embargo, estos experimentos tempranos fueron sugerentes y llevaron a importantes descubrimientos en Probabilidad y Estadística. Éstos condujeron a Pearson al *chi-squared test,* que es de gran importancia para testear si los datos observados se ajustan a una distribución de probabilidad dada.

Para comienzos del siglo XX, estaba en claro la necesidad de mejorar la forma de generar números aleatorios. En 1927, L. H. C. Tippett publicó una lista de 41.600 dígitos obtenidos seleccionando números al azar de los informes del censo. En 1955, RAND Corporation imprimió una tabla de 1.000.000 de números aleatorios generados a partir de ruido electrónico. El advenimiento de procesadores de alta velocidad incrementó la posibilidad de generar números aleatorios directamente en la computadora, y a finales de la década del 40 John von Neumann sugirió que esto fuera hecho de la siguiente manera: Suponga que quiere una secuencia de números aleatorios de cuatro dígitos. Elija un número de cuatro dígitos cualquiera, digamos 6235, para comenzar. Eleve este número al cuadrado para obtener 38.875.225. Para el segundo número seleccione los cuatro dígitos del medio de este cuadrado (es decir, 8752). Haga el mismo proceso comenzando con 8752 para obtener el tercer número, y así sucesivamente.

Métodos más modernos involucran el concepto de aritmética modular. Si a es un entero y m es un entero positivo, entonces decimos que $a \pmod{m}$ es el resto de la división en la que a es el dividendo y m el divisor. Por ejemplo, $10 \pmod 4 = 2$, $8 \pmod 2 = 0$, y así sucesivamente. Para generar una secuencia de números aleatorios $X_0, X_1, X_2, \ldots$ elija un número inicial X_0 y luego obtenga los números X_{n+1} de X_n usando la fórmula

$$X_{n+1} = (aX_n + c) \pmod{m} ,$$

donde a, c, y m son constantes cuidadosamente seleccionadas. Entonces, $X_0, X_1, X_2, \ldots$ es una secuencia de enteros entre 0 y $m-1$. Para obtener una secuencia de números reales en el intervalo $[0, 1)$, dividimos cada X_j por m. La secuencia resultante consiste en números racionales de la forma j/m, donde $0 \leq j \leq m-1$. Como m suele ser un entero muy grande, consideramos a los números en la secuencia como números reales aleatorios del intervalo $[0, 1)$.

Tanto para la el método de cuadrados de von Neumann como para la técnica de aritmética modular, la secuencia de números resultante está completamente determinada por el primer número. Así, no hay nada realmente aleatorio en ellas. Sin embargo, ambas producen números que se comportan muy de acuerdo a lo que la teoría predeciría para experimentos aleatorios. Para obtener secuencias diferentes para experimentos distintos, el número inicial X_0 se elige por medio de algún otro procedimiento que podría involucrar, por ejemplo, la hora del día.[4]

Durante la Segunda Guerra Mundial, físicos del Los Alamos Scientific Laboratory necesitaban saber, por cuestiones de blindaje, qué tan lejos los neutrones

[3]K. Pearson, "Science and Monte Carlo," *Fortnightly Review*, vol. 55 (1894), p. 193; cited in S. M. Stigler, *The History of Statistics* (Cambridge: Harvard University Press, 1986).

[4]Para una discusión en detalle acerca de los números aleatorios, ver D. E. Knuth, *The Art of Computer Programming,* vol. II (Reading: Addison-Wesley, 1969).

viajan a través de diversos materiales. Esta pregunta estaba más allá del alcance de los cálculos teóricos. Daniel McCracken, al escribir para el *Scientific American*, declara:

> Los físicos tenían la mayoría de los datos necesarios: conocían la distancia promedio que un neutrón viajaría a una velocidad dada en una substancia antes de que este colisionara con un núcleo atómico, que las probabilidades decían que el neutrón rebotaría en lugar de ser absorbido por el núcleo, cuanta energía era de esperar que el neutrón perdiera luego de una colisión dada, etcétera.[5]

John von Neumann y Stanislas Ulam sugirieron que el problema fuese resuelto modelando el experimento con dispositivos de probabilidad en una computadora. Como su trabajo era secreto, fue necesario darle un nombre codificado. Von Neumann eligió el nombre "Monte Carlo." Desde entonces, a este método de simulación se lo conoce como el *Método Monte Carlo.*

William Feller indicó las posibilidades de utilizar simulaciones por computadora para ilustrar conceptos básicos de probabilidad en su libro *An Introduction to Probability Theory and Its Applications.* Al discutir acerca del problema de la cantidad de veces que un jugador está al frente en el juego de "caras y cruces" Feller escribe:

> Los resultados concernientes con fluctuaciones en tiradas de monedas muestran que las creencias ampliamente sostenidas acerca de la ley de los números grandes son falacias. Estos resultados son tan asombrosos y discrepan tanto con la intuición común que incluso colegas sofisticados dudaron que las monedas realmente se portarían mal como la teoría predice. El registro de un experimento simulado es, por tanto, incluido.[6]

Feller proporciona un argumento mostrando el resultado de 10.000 juegos de *caras y cruces* similar al de la Figura 1.5.

El sistema de apuestas de la martingala descripto en el Ejercicio 10 tiene una historia larga e interesante. Russell Barnhart sealó al los autores que su uso puede rastrearse al menos hasta 1754, cuando Casanova, al redactar sus memorias, *History of My Life,* escribe:

> Ella [la amante de Casanova] me hizo prometer que iría al casino [el Ridotto en Venecia] a jugar dinero en sociedad con ella. Fui, entonces, llevando todo el oro que había encontrado, y determinadamente duplicando las apuestas de acuerdo son el sistema conocido como la martingala, gané tres o cuatro veces al día durante el resto del Carnaval. Nunca perdía la sexta carta. De haberlo hecho, me hubiera quedado sin fondos, lo que equivalía a dos mil monedas de oro venecianas.[7]

---

[5] D. D. McCracken, "The Monte Carlo Method," *Scientific American,* vol. 192 (May 1955), p. 90.

[6] W. Feller, *Introduction to Probability Theory and its Applications,* vol. 1, 3rd ed. (New York: John Wiley & Sons, 1968), p. xi.

[7] G. Casanova, *History of My Life,* vol. IV, Chap. 7, trans. W. R. Trask (New York: Harcourt-Brace, 1968), p. 124.

Aun si no hubiera habido ceros en la rueda de la ruleta de manera que el juego fuera perfectamente parejo, el sistema de la martingala, o cualquier otro sistema con el mismo fin, no puede hacer que el juego se vuelva favorable. La idea de que un juego parejo permanece parejo y que juegos desparejos continúan desparejos bajo los sistemas del juego, ha sido explotada por matemáticos para obtener resultados importantes en el estudio de la probabilidad. Introduciremos el concepto general de la martingala en el Capítulo 6.

La palabra *martingala* en sí misma posee también una historia muy interesante. Su origen es oscuro. El *Oxford English Dictionary* da ejemplos de su uso a principios del siglo XVII y menciona como procedencia probable la referencia en el Libro Uno de Rabelais Libro Uno, Capítulo 19:

> Todo se hizo según lo planeado, la única cuestión fue que Gargantua dudó si serían capaces de conseguir, rápidamente, pantalones adecuados para las piernas de su viejo amigo; hesitó, también, acerca del corte más apropiado para el orador—la martingala, que tiene un efecto de puente levadizo en las nalgas, para permitirle a uno hacer sus negocios más fácilmente; el estilo marinero, que proporciona más confort a los riones; el suizo, que es más tibio sobre el abdomen; o la cola de bacalao, que es más fresca en el lomo.[8]

En el inglés moderno, *martingale* tiene varios significados diferentes, todos relacionados con *mantener bajo* o *holding down* además de su uso en el juego. Por ejemplo, es una correa del arnés de un caballo que mantiene baja la cabeza del animal, y también parte del aparejo de navegación que se utiliza para mantener la proa baja.

El sistema Labouchere descripto en el Ejercicio 9 toma su nombre de Henry du Pre Labouchere (18311912), un periodista inglés y miembro del Parlamento. Labouchere atribuyó el sistema a Condorcet. Condorcet (1743–1794) fue un líder político en los tiempos de la Revolución Francesa que estaba interesado en aplicar la teoría de la probabilidad a la política y la economía. Por ejemplo, calculó la probabilidad de que un jurado usando el voto de la mayoría habría de tomar la decisión correcta si cada uno de sus miembros tuviera la misma probabilidad de decidir correctamente. Sus escritos proporcionaron una riqueza de ideas acerca de cómo aplicar la probabilidad a asuntos humanos.[9]

Ejercicios

1 Modificar el programa **CoinTosses** para arrojar una moneda n veces, y cada 100 tiradas, imprimir la proporción de caras menos 1/2. Aparentan estos números acercarse a 0 a medida que n aumenta? Modificar el programa una vez más para imprimir, cada 100 veces, las siguientes cantidades: la proporción de caras menos 1/2, y el número de caras menos la mitad del número de tiradas. Aparentan estos números acercarse a cero a medida que n aumenta?

[8] Quoted in the *Portable Rabelais,* ed. S. Putnam (New York: Viking, 1946), p. 113.

[9] Le Marquise de Condorcet, *Essai sur l'Application de l'Analyse à la Probabilité dès Décisions Rendues a la Pluralité des Voix* (Paris: Imprimerie Royale, 1785).

2 Modificar el programa **CoinTosses** de manera tal que arroje una moneda n veces y registre si la proporción de caras está entre 0,1 de 0,5 ó no (es decir, entre 0,4 y 0,6). Repetir este experimento 100 veces. Qué tan grande debe ser n para que aproximadamente 95 de 100 veces la proporción de caras esté entre 0,4 y 0,6?

3 A principios del siglo XVII, le pidieron a Galileo que explicara el hecho de que, aunque el número de ternas de enteros del 1 al 6 con suma igual a 9 es el mismo que el número de tales ternas cuya suma es igual a 10, cuando se arrojan tres dados, un 9 parece salir con menos frecuencia que un 10—según la experiencia de los jugadores.

(a) Escribir un programa que simule la tirada de tres dados una gran cantidad de veces y seguir el paso de la proporción de veces que la suma es 9 y que la suma es 10.

(b) Se puede concluir de la simulación anterior que los jugadores estaban en lo cierto?

4 El raquetball, un jugador continúa sirviendo siempre y cuando gane los tantos; se anota un punto sólo cuando el jugador que tiene el servicio gana la jugada. El primer contendiente que alcance los 21 puntos gana el partido. Asuma que usted sirve primero y que tiene una probabilidad 0,6 de ganar tantos con su servicio y probabilidad 0,5 cuando sirve su oponente. Estimar, por simulación, la probabilidad de que usted gane un partido.

5 Considere apostar a que salga triple seis, al menos una vez, al arrojar tres dados n veces. Calcule $f(n)$, la probabilidad de obtener al menos un triple seis en n tiradas de tres dados. Determine el menor valor de n necesario para que esta apuesta sea favorable. (DeMoivre diría que es aproximadamente $216 \log 2 = 149.7$ y entonces contestaría 150—ver Ejercicio 1.2.17. Está usted de acuerdo con él?)

6 En Las Vegas, una rueda de ruleta posee 38 ranuras numeradas 0; 00; 1; 2; $\ldots$; 36. Las ranuras 0 y 00 son verdes y de las 36 ranuras restantes la mitad son rojas y la otra mitad son negras. Un *crupier* gira la rueda y arroja sobre ella una bolilla de marfil. Si usted apuesta 1 dólar al rojo, ganará 1 dólar si la bolilla se detiene sobre una ranura roja, de lo contrario perderá un dólar. Escribir un programa para encontrar todas las victorias de un jugador que realiza 1000 apuestas al rojo.

7 Otra forma de jugar a la ruleta es apostar a un número específico (digamos 17). Si la bolilla se detiene en su número, usted se llevará 35 dólares además del dólar que apostó. Si no, perderá su dólar. Escribir un programa que trace sus victorias al apostar 500 veces a una ruleta de Las Vegas, primero al apostar siempre al rojo (ver Ejercicio 6), y luego para una segunda visita a esta ciudad, cuando usted juega 500 veces apostando siempre al 17. Qué diferencias se evidencian en las gráficas de sus victorias en estas dos ocasiones?

8 Un estudiante astuto notó que en nuestra simulación del juego de caras y cruces (ver Ejemplo 1.4), la proporción de veces que el jugador está al frente es muy cercana a la proporción de veces que sus ganancias netas terminan siendo 0. Halle la solución para estas probabilidades enumerando todos los casos para dos y para cuatro tiradas de monedas, y vea si piensa que estas probabilidades son, en realidad, iguales.

9 El *sistema Labouchere* para la ruleta se juega de la siguiente manera: Escriba una lista de números, usualmente 1; 2; 3; 4. Apueste la suma del primero y el último, $1 + 4 = 5$, al rojo. Si gana, borre el primer y el último número de su lista. Si pierde, agregue la cantidad de su última apuesta al final de la lista. Luego, use la nueva lista y apueste la suma del primer y el último número (si sólo queda un número, apueste esa cantidad). Continúe hasta vaciar la lista. Probar que si esto ocurriera, usted ganaría la suma de $1 + 2 + 3 + 4 = 10$, de su lista original. Simular este sistema y ver si siempre se vacía la lista, por lo tanto, siempre se gana. De ser así, por qué no es este un sistema de juego a prueba de tontos?

10 Otro sistema de apuestas bien conocido es el *sistema de duplicación de martingala*. Suponga que usted está apostando al rojo en la ruleta. Cada vez que gane, apueste 1 dólar en la ronda siguiente. Cada vez que gane, duplique su apuesta anterior. Continúe jugando hasta que haya ganado al menos 5 dólares o haya perdido más de 100. Escriba un programa que simule este sistema y juéguelo una cantidad de veces para ver cómo le va. En su libro *The Newcomes,* W. M. Thackeray "No ha jugado hasta ahora? No lo hagas; sobre todo evite una martingala si lo hace.[10] Fue este un buen consejo?

11 Modifique el programa **HTSimulation** para que siga la pista al máximo de las victorias de Peter en cada juego de 40 tiradas. El programa deberá imprimir la proporción de veces que sus victorias totales tomen los valores $0, 2, 4, \ldots, 40$. Calcule la probabilidad exacta correspondiente para juegos de dos y cuatro tiradas.

12 En la próximas elecciones nacionales del Presidente de los Estados Unidos, una encuestadora planea predecir el ganador del voto popular tomando muestras aleatorias de 1000 votantes y declarando que el ganador será quien obtenga la mayoría de votos en su encuesta. Suponga que el 48 por ciento de los electores planea votar al candidato republicano y el 52 por ciento restante al demócrata. Para tener una idea de cuán razonable es la idea de la encuestadora, escriba un programa que haga esta predicción por simulación. . Repita la simulación 100 veces vea en cuántas oportunidades la predicción se vuelve realidad. Repita su experimento, asumiendo ahora que el 49 por ciento de la población planea votar al candidato republicano; primero con una muestra de 1000 y luego con una de 3000. (La Encuesta de Gallup usa alrededor de 3000.) (Esta idea se discute más detalladamente en el Capítulo 9, Section 9.1.)

[10] W. M. Thackerey, *The Newcomes* (London: Bradbury and Evans, 1854–55).

13 El psicólogo Tversky y sus colegas[11] dicen que cuatro de cinco personas responderán (a) a la siguiente pregunta:

Cierto pueblo posee dos hospitales. En el más grande, 45 bebés nacen en un día, y en el más chico nacen 15 bebés diariamente. Aunque la proporción total de varones es de alrededor de 50 por ciento, la proporción real en cada hospital puede ser mayor o menor que 50 para un día. A fin de ao, qué hospital tendrá la mayor cantidad de días en los cuales más del 60 por ciento de los bebés fueron varones?

(a) el hospital grande

(b) el hospital chico

(c) ninguno—el cantidad de días será prácticamente la misma.

Asuma que la probabilidad de que un bebé sea varón es 0,5 (estimaciones reales apuntan más bien a 0,513). Decida, por simulación, cuál es la respuesta correcta a la pregunta. Puede sugerir por qué tanta gente responde mal?

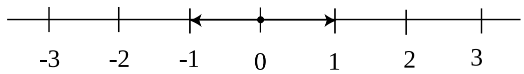
14 Le proponen el siguiente juego: Se arrojará una moneda hasta obtener la primera cara. Si esto ocurre en la $j$va tirada, le pagarán 2^j dólares. Usted ganará al menos 2 dólares, así que debería estar dispuesto a pagar para jugar este juego.—pero cuánto? Muy poca gente pagaría tanto como 10 dólares. Intente encontrar, por simulación, una cantidad que usted considere razonable para pagar, por juego, si es que le permiten jugarlo muchas veces. La cantidad que usted estaría dispuesto a pagar por juego, depende del número de veces que le permitan jugar?

15 Tversky y sus colegas[12] estudiaron registros de 48 partidos de básquet de los Philadelphia 76ers en la temporada 1980–81 para ver si algún jugador había tenido períodos en los que fue bueno y todos sus tiros entraban, y otros en los que fue malo y apenas capaz de darle al panel. Los jugadores estimaron que tirar después de un acierto era aproximadamente 25 por ciento más probable que tirar después de haber errado. En realidad, pasaba lo contrario–los 76ers tenían 6 por ciento más chances de anotar después de un tiro fallido que de uno acertado. Tversky informó que la cantidad de rachas buenas y malas fue lo que uno esperaría por efectos puramente aleatorios. Asumiendo que un jugador tiene chances 50 y 50 de tirar y hacen 20 tiros por juego, estimar por simulación la proporción de juegos en los que el jugador tendrá una racha de 5 o más aciertos.

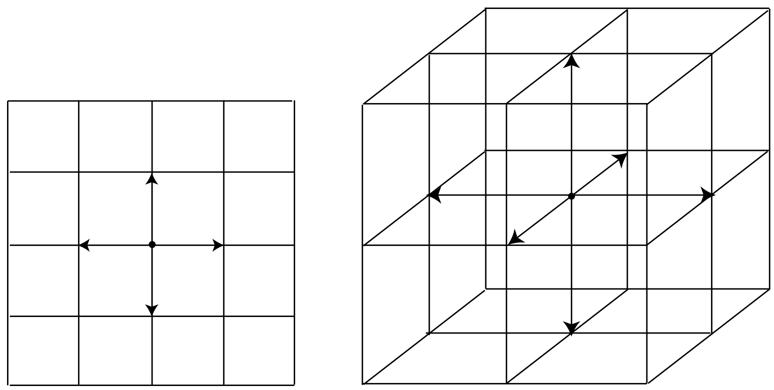
16 Estimar, por simulación, el número promedio de nios que habría en una familia si toda la gente tuviera hijos hasta tener un varón. Hacer lo mismo para el caso de que toda la gente tuviera hijos hasta tener al menos un varón y una

[11] Ver K. McKean, "Decisions, Decisions," *Discover,* Junio 1985, pp. 22–31. Kevin McKean, Discover Magazine, ©1987 Family Media, Inc. Reimpreso con permiso. Este popular artículo habla acerca del trabajo de Tverksy et. al. en *Judgement Under Uncertainty: Heuristics and Biases* (Cambridge: Cambridge University Press, 1982).

[12] ibid.



a. Random walk in one dimension.



b. Random walk in two dimensions.   c. Random walk in three dimensions.

Figure 1.6: Random walk.

nena. Cuántos nios más esperaría encontrar según el segundo esquema, en comparación con el primero para 100.000 familias? (Asumir que los varones y las nenas son igualmente probables.)

17 Los matemáticos tienen la fama de tener algunas de sus mejores ideas sentados en la mesa de un bar, viajando en ómnibus o paseando por el parque. A principios del siglo XX, el famoso matemático George Pólya vivía en un hotel cerca del bosque en Zurich. Le gustaba caminar por el bosque y pensar acerca de cuestiones matemáticas. Pólya describe el siguiente incidente:

> En el hotel también vivían algunos estudiantes con los que normalmente compartía mis comidas y guardaba una relación amistosa. Un día uno the ellos esperaba la visita de su novia, lo cual yo sabía, pero no preveía que él y su novia también saldrían a pasear por el parque, y de repente allí los encontré. Y los seguí encontrando repetidas veces esa maana, no recuerdo cuantas veces, pero ciertamente demasiado seguido y me sentí avergonzado: Parecía como si hubiera estado espiando, lo que puedo asegurar, no era el caso.[13]

[13] G. Pólya, "Two Incidents," *Scientists at Work: Festschrift in Honour of Herman Wold,* ed. T. Dalenius, G. Karlsson, and S. Malmquist (Uppsala: Almquist & Wiksells Boktryckeri AB, 1970).

Esto lo puso a pensar si lo caminantes que deambulan estaban destinados a encontrarse.

Pólya consideró a los caminantes sin rumbo fijo en una, dos y tres dimensiones. En una dimensión, concibió al caminante en una calle muy larga. En cada intersección, éste arroja una moneda para decidir en qué dirección seguir caminando (ver Figura 1.6a). En dos dimensiones, la persona camina en una grilla de calles, y en cada intersección elige una de las cuatro direcciones a seguir posibles con igual probabilidad (ver Figura 1.6b). En tres dimensiones (deberíamos hablar de un escalador sin rumbo), el caminante se mueve en una grilla de tres dimensiones, y en cada intersección tenemos ahora seis direcciones diferentes a elegir, todas con igual probabilidad. (ver Figura 1.6c).

Refererirse a la Sección 12.1, donde se discuten este y otros problemas relacionados.

(a) Escriba un programa que simule un caminante sin rumbo fijo en una dimensión comenzando desde 0. Su programa deberá imprimir las longitudes de tiempo entre regresos al punto de partida (retornos a 0). Vea si puede adivinar por simulación la respuesta a la siguiente pregunta: El caminante volverá siempre al punto de partida eventualmente, o podría alejarse para siempre?

(b) Las trayectorias de dos caminantes en dos dimensiones que se encuentran después de n pasos puede considerarse una única trayectoria que comienza en $(0;0)$ y regresa a $(0,0)$ después de $2n$ pasos. Esto significa que la probabilidad de que dos caminantes que deambulan en dos dimensiones se encuentren es la misma que la probabilidad de que un caminante, en dos dimensiones, regrese a su punto de partida. De esta manera, preguntar si dos caminantes se encontrarán siempre, es lo mismo que preguntar si un caminante siempre llegará a su punto de partida.

Escribir un programa que simule una caminata aleatoria en dos dimensiones y vea si usted cree que el caminante regresará siempre a $(0,0)$. De ser así, Pólya estaría seguro de seguir encontrándose a sus amigos en el parque. Tal vez ya haya conjeturado una respuesta a la pregunta: Un caminante sin rumbo en una o dos dimensiones vuelve siempre al punto de partida? Pólya respondió a esta questión para una, dos y tres dimensiones. Estableció el remarcable resultado que la respuesta es *sí* en una y dos dimensiones y *no* en tres dimensiones.

(c) Escriba un programa que simule una caminata aleatoria en tres dimensiones y vea, a partir de esta simulación y los resultados de (a) y (b), si podría haber adivinado los resultados que obtuvo Pólya.

1.2 Distribuciones de Probabilidad Discreta

En este libro estudiaremos muchos experimentos diferentes desde el punto de vista probabilístico. Lo que está envuelto en este estudio se pondrá evidente cuando la teoría se desarrolle y se analizen los ejemplos. Sin embargo, la idea global puede describirse e ilustrarse como sigue: a cada experimento que nosotros consideremos, se asociará una variable aleatoria, que representa el resultado de cualquier experimento particular. El conjunto de posibles resultados se llama el *espacio muestral*. En la primera parte de esta sección, nosotros consideraremos el caso dónde el experimento sólo tiene finitamente muchos resultados posibles, es decir, el espacio muestral es finito. Entonces generalizaremos al caso en que el espacio muestral es finito o contablemente infinito. Esto nos lleva a la siguiente definición.

Variables aleatorias y espacios muestrales

Definición 1.1 Suponga que tenemos un experimento cuyo resultado depende del azar. Representamos al resultado del experimento por una letra romana mayúscula, como X, a la que llamo *variable aleatoria*. El *espacio muestral* del experimento es el conjunto de todos los resultados posibles. Si el espacio muestral es finito o contablemente infinito, se dice que la variable aleatoria es *discreta*. □

Generalmente denotamos un espacio muestral por la letra griega mayúscula Ω. Como declaramos anteriormente, en la correspondencia entre un experimento y la teoría matemática que lo estudia, el espacio muestral Ω corresponde al conjunto de resultados posibles del experimento.

Hacemos dos definiciones adicionales ahora. Éstas son subsidiarias a la definición de espacio muestral y sirven para precisar alguna terminología común usada junto con los espacios muestrales. En primer lugar, definimos a los elementos de un espacio muestral como *resultados*. Segundo, definimos a cada subconjunto de un espacio muestral como un *suceso*. Normalmente, denotaremos a los resultados por letras minúsculas y a sucesos por letras mayúsculas.

Ejemplo 1.6 Un dado se arroja una vez. Permitamos a X denotar el resultado de este experimento. Entonces el espacio muestral para este experimento es el conjunto de 6 elementos

$$\Omega = \{1, 2, 3, 4, 5, 6\} ,$$

donde cada resultado i, para $i = 1, \ldots, 6$, corresponde al número de puntos en la cara volteada. El suceso

$$E = \{2, 4, 6\}$$

corresponde a la declaración que el resultado de la tirada es un número par. El suceso E también puede describirse diciendo que X es par. A menos que haya un razón para creer que el dado está cargado, la asunción natural es que cada resultado es igualmente probable. Adoptando esta convención significa que asignamos una probabilidad de $1/6$ a cada uno de los seis resultados, es decir, $m(i) = 1/6$, para $1 \le i \le 6$. □

Funciones de Distribución

Ahora describiremos la asignación de probabilidades. Las definiciones se motivan por el ejemplo anterior, en el cual asignamos a cada resultado del espacio muestral un número no negativo tal que la suma de los números asignados es igual a 1.

Definición 1.2 Sea X una variable aleatoria qué denota el valor del resultado de un cierto experimento, y asumiendo que este sólo tiene finitamente muchos resultados posibles. Sea Ω el espacio muestral del experimentob (es decir, el conjunto de todos los posibles valores de X, o equivalente, el conjunto de todos los posibles resultados del experimento.) Una *función de distribución* para X es una función a valor real m cuyo dominio es Ω y que satisface:

1. $m(\omega) \geq 0$, para todo $\omega \in \Omega$, y
2. $\sum_{\omega \in \Omega} m(\omega) = 1$.

Para cualquier subconjunto E de Ω, definimos la *probabilidad* de E como el número $P(E)$ dado por

$$P(E) = \sum_{\omega \in E} m(\omega) \ .$$

□

Ejemplo 1.7 Considere un experimento en que una moneda se tira dos veces. Sea X la variable aleatoria que corresponde a este experimento. Notamos que hay varias maneras de registrar los resultados de este experimento. Por ejemplo, podríamos registrar las dos tiradas, en el orden en que ocurrieron. En este caso, tenemos Ω ={HH,HT,TH,TT}. También simplemente podríamos registrar los resultados anotando el número de caras que aparecían. En este caso, tenemos Ω ={0,1,2}. Finalmente, podríamos registrar los dos resultados, sin tener en cuenta el orden en que ocurrieron. En este caso, tenemos Ω ={HH,HT,TT}.

Usaremos, por el momento, el primero de los espacios muestrales dado antes. Asumiremos que los cuatro resultados son igualmente probables, y definiremos la función de distribución $m(\omega)$ por

$$m(\mathrm{HH}) = m(\mathrm{HT}) = m(\mathrm{TH}) = m(\mathrm{TT}) = \frac{1}{4} \ .$$

Sea E ={HH,HT,TH} el suceso que al menos sale una cara. Entonces, la probabilidad de E puede calcularse como sigue:

$$\begin{aligned} P(E) &= m(\mathrm{HH}) + m(\mathrm{HT}) + m(\mathrm{TH}) \\ &= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4} \ . \end{aligned}$$

Similarmente, si F ={HH,HT} es el suceso que salga cara en la primera tirada, entonces tenemos

$$\begin{aligned} P(F) &= m(\mathrm{HH}) + m(\mathrm{HT}) \\ &= \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \ . \end{aligned}$$

□

Ejemplo 1.8 (Ejemplo 1.6 continuado) El espacio muestral para el experimento en que el dado se arroja es el conjunto de 6 elementos $\Omega = \{1, 2, 3, 4, 5, 6\}$. Asumimos que el dado es equilibrado, y escogemos la función de distribución definida por

$$m(i) = \frac{1}{6}, \qquad \text{para } i = 1, \ldots, 6 \ .$$

Si E es el suceso que el resultado de la tirada es un número par, entonces $E = \{2, 4, 6\}$ y

$$\begin{aligned} P(E) &= m(2) + m(4) + m(6) \\ &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2} \ . \end{aligned}$$

□

Notar que es una consecuencia inmediata de las definiciones anteriores, que para cada $\omega \in \Omega$,

$$P(\{\omega\}) = m(\omega) \ .$$

Es decir, la probabilidad del suceso elemental $\{\omega\}$, compuesto de un solo resultado ω, es igual al valor $m(\omega)$ asignado al resultado ω por la función de distribución.

Ejemplo 1.9 Tres personas, A, B, and C, están corriendo para la misma oficina, y asumimos que uno y sólo uno de ellos ganará. El espacio muestral puede tomarse como el conjunto de 3 elementos Ω ={A,B,C} dónde cada elemento corresponde al resultado de que ese candidato gane. Suponga que A y B tienen la misma oportunidad de ganar, pero C solo tiene $1/2$ de la oportunidad de A o B. Entonces asignamos

$$m(\mathrm{A}) = m(\mathrm{B}) = 2m(\mathrm{C}) \ .$$

dado

$$m(\mathrm{A}) + m(\mathrm{B}) + m(\mathrm{C}) = 1 \ ,$$

vemos que

$$2m(\mathrm{C}) + 2m(\mathrm{C}) + m(\mathrm{C}) = 1 \ ,$$

que implica $5m(\mathrm{C}) = 1$. De aquí,

$$m(\mathrm{A}) = \frac{2}{5} \ , \qquad m(\mathrm{B}) = \frac{2}{5} \ , \qquad m(\mathrm{C}) = \frac{1}{5} \ .$$

Sea E el suceso que cualquiera A o C gane. Entonces E ={A,C}, y

$$P(E) = m(\mathrm{A}) + m(\mathrm{C}) = \frac{2}{5} + \frac{1}{5} = \frac{3}{5} \ .$$

□

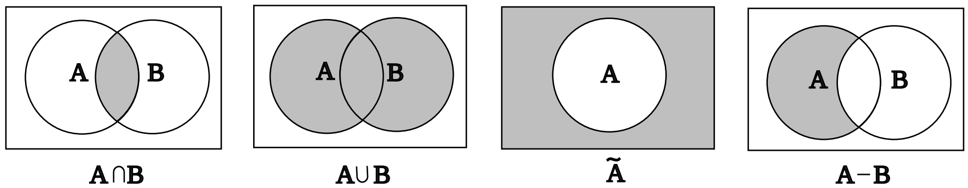


Figure 1.7: Operaciones de conjunto básicas.

En muchos casos, los sucesos pueden describirse en términos de otros sucesos a traves del uso de construcciones estandares de la teoría de conjuntos. Repasaremos las definiciones de estas construcciones brevemente. El lector es referido a la Figura 1.7 para los diagramas de Venn que ilustran estas construcciones.

Sean A y B dos conjuntos. Entonces la unión de A y B es el conjunto

$$A \cup B = \{x \,|\, x \in A \text{ o } x \in B\} \; .$$

La intersección de A y B es el conjunto

$$A \cap B = \{x \,|\, x \in A \text{ y } x \in B\} \; .$$

La diferencia de A y B es el conjunto

$$A - B = \{x \,|\, x \in A \text{ y } x \notin B\} \; .$$

El conjunto A es un subconjunto de B, escrito $A \subset B$, si cada elemento de A es también un elemento de B. Finalmente, el complemento de A es el conjunto

$$\tilde{A} = \{x \,|\, x \in \Omega \text{ y } x \notin A\} \; .$$

La razón por la que estas construcciones son importantes es que típicamente el caso de sucesos complicados descriptos en español pueden separarse en sucesos más simples usando estas construcciones. Por ejemplo, si A es el suceso "nevará mañana y lloverá el próximo día," B es el suceso "nevará mañana," y C es el suceso "lloverá dentro de dos días," entonces A es la intersección de los sucesos B y C. Similarmente, si D es el suceso "nevará mañana o lloverá el próximo día," entonces $D = B \cup C$. (Note el cuidado que debe tomarse aquí, porque a veces la palabra "o" en español significa exactamente que una de las dos alternativas ocurrirá. El significado está normalmente claro en el contexto. En este libro, usaremos siempre la palabra "o" en el sentido inclusivo, es decir, A o B significa que al menos uno de los dos sucesos A, B es verdadero.) El suceso $\tilde{B}$ es el suceso "no nevará mañana." Finalmente, si E es el suceso "nevará mañana pero no lloverá el próximo día," entonces $E = B - C$.

Propiedades

Teorema 1.1 Las probabilidades asignadas a sucesos por una función de distribución en el espacio muestral Ω satisfacen las siguientes propiedades:

1. $P(E) \geq 0$ para cada $E \subset \Omega$.
2. $P(\Omega) = 1$.
3. Si $E \subset F \subset \Omega$, entonces $P(E) \leq P(F)$.
4. Si A y B son subconjuntos disjuntos de Ω, entonces $P(A \cup B) = P(A) + P(B)$.
5. $P(\tilde{A}) = 1 - P(A)$ para cada $A \subset \Omega$.

Dem. Para cualquier suceso E la probabilidad $P(E)$ es determinada de la distribución m por

$$P(E) = \sum_{\omega \in E} m(\omega) ,$$

para cada $E \subset \Omega$. Puesto que la función m es no negativa, sigue que $P(E)$ es también no negativa. Así, la Propiedad 1 es verdadera.

La Propiedad 2 es demostrada por las equaciones

$$P(\Omega) = \sum_{\omega \in \Omega} m(\omega) = 1 .$$

Suponga que $E \subset F \subset \Omega$. Entonces cada elemento de ω que pertenece a E también pertenece a F. Por lo tanto,

$$\sum_{\omega \in E} m(\omega) \leq \sum_{\omega \in F} m(\omega) ,$$

ya que que cada término en la suma a la izquierda está en la suma a la derecha, y todos los términos en ambas sumas son no negativos. Esto implica que

$$P(E) \leq P(F) ,$$

y se demuestra la Propiedad 3.

Suponga ahora que A y B son subconjuntos disjuntos de Ω. Entonces cada elemento ω de $A \cup B$ está en A y no en B o en B y no en A. Sigue que

$$\begin{aligned} P(A \cup B) &= \textstyle\sum_{\omega \in A \cup B} m(\omega) = \sum_{\omega \in A} m(\omega) + \sum_{\omega \in B} m(\omega) \\ &= P(A) + P(B) , \end{aligned}$$

y la Propiedad 4 es demostrada.

Finalmente, para demostrar la Propiedad 5, consideremos la unión disjunta

$$\Omega = A \cup \tilde{A} .$$

Dado $P(\Omega) = 1$, la propiedad de la aditividad disjunta (Propiedad 4) implica que

$$1 = P(A) + P(\tilde{A}) ,$$

de donde $P(\tilde{A}) = 1 - P(A)$. □

Es importante comprender que la Propiedad 4 en el Teorema 1.1 pueda extenderse a más de dos conjuntos. La propiedad general de aditividad finita está dada por el teorema siguiente.

Teorema 1.2 Si $A_1, \ldots, A_n$ son subconjuntos disjuntos de a pares de Ω (es decir, dos de los A_i no tienen un elemento en común), entonces

$$P(A_1 \cup \cdots \cup A_n) = \sum_{i=1}^{n} P(A_i) \ .$$

Dem. Sea ω un elemento en la unión

$$A_1 \cup \cdots \cup A_n \ .$$

Entonces $m(\omega)$ ocurre exactamente una vez en cada lado de la igualdad en la declaración del teorema. □

Usaremos a menudo la siguiente consecuencia del teorema anterior.

Teorema 1.3 Sean $A_1, \ldots, A_n$ sucesos disjuntos de a pares con $\Omega = A_1 \cup \cdots \cup A_n$, y sea E cualquier suceso. Entonces

$$P(E) = \sum_{i=1}^{n} P(E \cap A_i) \ .$$

Dem. Los conjuntos $E \cap A_1, \ldots, E \cap A_n$ son disjuntos de a pares, y su unión es el conjunto E. El resultado sigue ahora del Teorema 1.2. □

Corolario 1.1 Para dos sucesos cualquiera A y B,

$$P(A) = P(A \cap B) + P(A \cap \tilde{B}) \ .$$

□

La Propiedad 4 puede generalizarse de otra manera. Suponga que A y B son subconjuntos de Ω que no necesariamente son disjuntos. Entonces:

Teorema 1.4 Si A y B son subconjuntos de Ω, entonces

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \ . \qquad (1.1)$$

Dem. El lado izquierdo de la Ecuación 1.1 es la suma de $m(\omega)$ para ω en A o B. Debemos mostrar que el lado derecho de la Ecuación 1.1 también suma $m(\omega)$ para

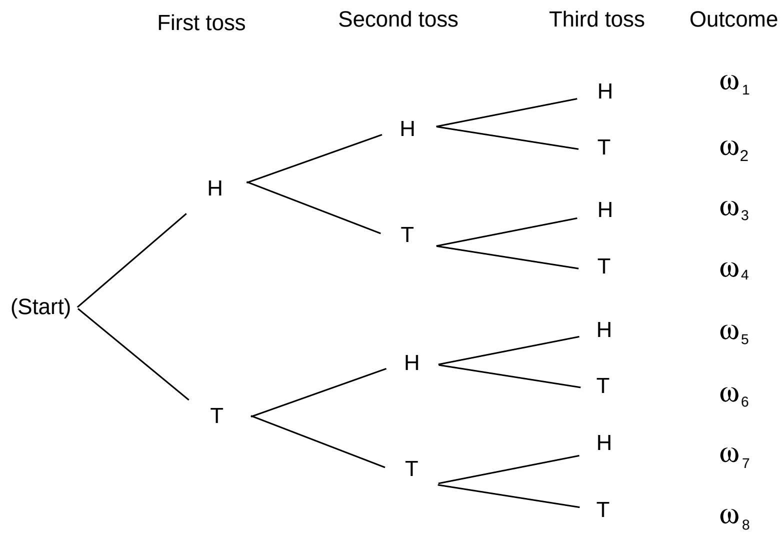


Figure 1.8: Diagrama de árbol para tres tiradas de una moneda.

ω en A o B. Si ω está exactamente en uno de los dos conjuntos, entonces es contado en uno solo de loas tres términos del lado derecho de la Ecuación 1.1. Si está en ambos A y B, se suma dos veces de los cálculos de $P(A)$ y $P(B)$ y restado una vez para $P(A \cap B)$. Asi se está contando exactamente una vez por el lado derecho. Por supuesto, si $A \cap B = \emptyset$, entonces la Ecuación 1.1 se reduce a la Propiedad 4. (La Ecuación 1.1 también puede generalizarse; vea el Teorema 3.8.) $\square$

Diagramas de Arbol

Ejemplo 1.10 Permítanos ilustrar las propiedades de probabilidades de sucesos en términos de tres tiradas de una moneda. Cuando tenemos un experimento que tiene lugar en etapas como éste, encontramos a menudo conveniente representar los resultados por un *diagrama de árbol* como el mostrado en la Figura 1.8.

Un *camino* a través del árbol corresponde a un posible resultado del experimento. Para el caso de tres tiradas de una moneda, tenemos ocho caminos $\omega_1, \omega_2, \ldots, \omega_8$ y, asumiendo que cada resultado es igualmente probable, asignamos igual peso, 1/8, a cada camino. Sea E el suceso "al menos sale una cara." Entonces $\tilde{E}$ es el suceso "no salen caras." Este suceso ocurre sólo para un resultado, a saber, $\omega_8 = \mathrm{TTT}$. Así, $\tilde{E} = \{\mathrm{TTT}\}$ y tenemos

$$P(\tilde{E}) = P(\{\mathrm{TTT}\}) = m(\mathrm{TTT}) = \frac{1}{8} \ .$$

Por la Propiedad 5 del Teorema 1.1,

$$P(E) = 1 - P(\tilde{E}) = 1 - \frac{1}{8} = \frac{7}{8} \ .$$

Notar que encontraremos a menudo más fácil de calcular la probabilidad de que un suceso no pase en lugar de la probabilidad de que lo haga. Entonces usamos la Propiedad 5 para obtener la probabilidad deseada.

Sea A el suceso "el primer resultado es cara," y B el suceso "el segundo resultado es cruz." Mirando los caminos en la Figura 1.8, podemos ver que

$$P(A) = P(B) = \frac{1}{2} \ .$$

Es más, $A \cap B = \{\omega_3, \omega_4\}$, y así $P(A \cap B) = 1/4$. Usando el Teorema 1.4, obtenemos

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= \frac{1}{2} + \frac{1}{2} - \frac{1}{4} = \frac{3}{4} \ . \end{aligned}$$

Puesto que $A \cup B$ es el conjunto de 6 elementos,

$$A \cup B = \{\mathrm{HHH,HHT,HTH,HTT,TTH,TTT}\} \ ,$$

vemos que obtenemos el mismo resultado por la enumeración directa. □

En nuestros ejemplos de tirar la moneda y en el ejemplo de arrojar el dado, hemos asignado una probabilidad igual a cada posible resultado del experimento. Correspondiendo a este método de asignar probabilidades, tenemos las definiciones siguientes.

Distribución Uniforme

Definición 1.3 La *distribución uniforme* en un espacio muestral Ω conteniendo n elementos es la función m definida por

$$m(\omega) = \frac{1}{n} \ ,$$

para cada $\omega \in \Omega$. □

Es importante comprender que cuando un experimento se analiza para describir sus posibles resultados, no hay una sola opción correcta del espacio muestral. Para el experimento de tirar una moneda dos veces en el Ejemplo 1.2, seleccionamos el conjunto de 4 elementos $\Omega = \{\mathrm{HH,HT,TH,TT}\}$ como espacio muestral y asignamos la función de distribución uniforme. Estas opciones son ciertamente e intuitivamente naturales. Por otro lado, para algunos propósitos puede ser más útil considerar el espacio muestral de 3 elementos $\bar{\Omega} = \{0, 1, 2\}$ en el que 0 es el resultado "no salieron caras," 1 es el resultado "salió exactamente una cara," y 2 es el resultado "salieron dos caras." La función de distribución $\bar{m}$ en $\bar{\Omega}$ definida por las ecuaciones

$$\bar{m}(0) = \frac{1}{4} \ , \qquad \bar{m}(1) = \frac{1}{2} \ , \qquad \bar{m}(2) = \frac{1}{4}$$

es una correspondiendo a la densidad de probabilidad uniforme en el espacio muestral Ω. Notar que es absolutamente posible escoger una función de distribución diferente. Por ejemplo, consideremos como la función de distribución uniforme en $\bar{\Omega}$, a la función $\bar{q}$ definida por

$$\bar{q}(0) = \bar{q}(1) = \bar{q}(2) = \frac{1}{3} .$$

Aunque $\bar{q}$ es perfectamente una buena función de distribución, no es consistente con los datos observados al tirar la moneda.

Ejemplo 1.11 Considere el experimento que consiste en arrojar un par de dados. Tomamos como el espacio muestral Ω al conjunto de todos los pares ordenados (i, j) de enteros con $1 \leq i \leq 6$ y $1 \leq j \leq 6$. Así,

$$\Omega = \{ (i, j) : 1 \leq i, j \leq 6 \} .$$

(Hay al menos otra opción "razonable" para el espacio muestral, a saber el conjunto de todos los pares de enteros no ordenados, cada uno entre 1 y 6. Para una discusión de por qué no usamos este conjunto, vea el Ejemplo 3.14.) Para determinar el tamaño de Ω, notamos que hay seis opciones para i, y para cada opción de i hay seis opciones para j, llevando a 36 diferentes resultados. Asumamos que el dado no está cargado. En términos matemáticos, significa que asumimos que cada uno de los 36 resultados es igualmente probable, o equivalentemente, que adoptamos la función de distribución uniforme en Ω poniendo

$$m((i, j)) = \frac{1}{36}, \qquad 1 \leq i, j \leq 6 .$$

Cuál es la probabilidad de conseguir una suma de 7 en la tirada de dos dados—o conseguir una suma de 11? El primer suceso, denotado por E, es el subconjunto

$$E = \{(1, 6), (6, 1), (2, 5), (5, 2), (3, 4), (4, 3)\} .$$

Una suma de 11 es el subconjunto F dado por

$$F = \{(5, 6), (6, 5)\} .$$

Consecuentemente,

$$\begin{aligned} P(E) = & \textstyle\sum_{\omega \in E} m(\omega) = 6 \cdot \frac{1}{36} = \frac{1}{6} , \\ P(F) = & \textstyle\sum_{\omega \in F} m(\omega) = 2 \cdot \frac{1}{36} = \frac{1}{18} . \end{aligned}$$

Cuál es la probabilidad de no conseguir *ojos de serpiente* (doble uno) ni *vagón* (doble seis)? El suceso de conseguir cualquiera de los dos resultados es el conjunto

$$E = \{(1, 1), (6, 6)\} .$$

Aquí, la probabilidad de obtener ninguno se da por

$$P(\tilde{E}) = 1 - P(E) = 1 - \frac{2}{36} = \frac{17}{18} .$$

□

En la tirada de moneda anterior y en los experimentos de arrojar dados, hemos asignado una probabilidad igual a cada resultado. Es decir, en cada ejemplo, hemos escogido la función de distribución uniforme. Éstas son las opciones naturales con tal de que la moneda sea equilibrada y los dados no estén cargados. Sin embargo, la decisión acerca de que función de distribución seleccionar para describir un experimento *no* es una parte de la teoría matemática básica de probabilidad. La última sólo comienza cuando el espacio muestral y la función de distribución ya se han definido.

Determinación de Probabilidades

Es importante considerar las maneras en que las distribuciones de probabilidad son determinadas en la práctica. Una manera es por *simetría.* Para el caso de la tirada de una moneda, no vemos ninguna diferencia física entre los dos lados de una moneda que deba afectar la oportunidad de un lado o el otro de salir. Similarmente, con un dado ordinario no hay ninguna diferencia escencial entre cualquiera de dos de sus lados, y así por simetría asignamos la misma probabilidad para cada posible resultado. En general, las consideraciones de simetría sugieren a menudo la función de distribución uniforme. Aquí debe usarse con cuidado. No siempre debemos asumir que, simplemente porque no sabemos ninguna razón para sugerir que un resultado es más probable que otro, es apropiado asignar las probabilidades iguales. Por ejemplo, considere el experimento de suponer el sexo de un niño recién nacido. Se ha observado que la proporción de niños recién nacidos que son varones es aproximadamente .513. Así, es más apropiado asignar una función de distribución qué asigne probabilidad .513 al resultado *varón* y probabilidad .487 al resultado *mujer* que asignar probabilidad 1/2 a cada resultado. Éste es un ejemplo dónde usamos las observaciones estadísticas para determinar probabilidades. Note que estas probabilidades pueden cambiar con los nuevos estudios y pueden variar de país a país. La ingeniería genética incluso podría permitirle a un individuo influir en esta probabilidad para un caso particular.

Apuestas

Las estimaciones estadísticas para las probabilidades están bien si el experimento bajo consideración puede repetirse varias veces bajo circunstancias similares. Sin embargo, asuma que, al principio de una temporada de fútbol, usted quiere asignar una probabilidad al suceso que Dartmouth vencerá a Harvard. Usted realmente no tiene datos que relacionan al equipo de fútbol de este año. Sin embargo, puede determinar su propia probabilidad personal viendo qué tipo de apuesta estaría deseoso de hacer. Por ejemplo, suponga que está dispuesto a hacer una apuesta de 1 dólar dando 2 a 1 posibilidades que Dartmouth ganará. Entonces está dispuesto a pagar 2 dólares si Dartmouth pierde a cambio de recibir 1 dólar si Dartmouth gana. Esto significa que que usted piensa que la probabilidad apropiada para la victoria de Dartmouth es 2/3.

Permítanos mirar más cuidadosamente la relación entre apuestas y probabilidades. Suponga que hacemos una apuesta con r a 1 posibilidades que ocurra un suceso E. Esto significa que pensamos que es r veces probable que E ocurrirá como que E no ocurrirá. En general, r a s posibilidades significará la misma cosa que r/s a 1, es decir, la proporción entre los dos números es la única cantidad de importancia al declarar las apuestas.

Ahora si es r veces probable que E ocurrirá como que E no ocurrirá, entonces la probabilidad que E ocurra debe ser $r/(r+1)$, de aquí tenemos

$$P(E) = r\,P(\tilde{E})$$

y

$$P(E) + P(\tilde{E}) = 1 \ .$$

En general, la declaración que las apuestas son r a s en favor de ocurrir un suceso E es equivalente a la declarición que

$$\begin{aligned} P(E) &= \frac{r/s}{(r/s)+1} \\ &= \frac{r}{r+s} \ . \end{aligned}$$

Si permitimos $P(E) = p$, entonces la ecuación de arriba puede facilmente ser resuelta para r/s en términos de p; obteniendo $r/s = p/(1-p)$. Resumimos la discusión anterior en la siguiente definición.

Definición 1.4 Si $P(E) = p$, las *apuestas* en favor de ocurrir un suceso E son $r : s$ (r a s) donde $r/s = p/(1-p)$. Si r y s son dadas, entonces p puede ser encontrada por el uso de la ecuación $p = r/(r+s)$. □

Ejemplo 1.12 (Ejemplo 1.9 continuado) En el Ejemplo 1.9 asignamos la probabilidad de $1/5$ al suceso que el candidato C gane la carrera. Así las apuestas a favor de ganar C son $1/5 : 4/5$. Estas apuestas podrían ser igualmente bien escritas como $1 : 4$, $2 : 8$, y así en adelante. Una apuesta que C gane es justa si recibimos 4 dólares si C gana y pagamos 1 dólar si C pierde. □

Espacios Muestrales Infinitos

Si un espacio muestral tiene un infinito número de puntos, entonces la manera que una función de distribución es definida depende si el espacio muestral es o no contable. Un espacio muestral es *infinito contable* si los elementos pueden ser contados, es decir, pueden ponerse uno a uno en correspondencia con los enteros positivos, e *infinito no contable* en caso contrario. Los espacios muestrales infinitos en general requieren nuevos conceptos (vea Capítulo 2), pero los espacios infinitos contables no. Si

$$\Omega = \{\omega_1, \omega_2, \omega_3, \ldots\}$$

es un espacio muestral infinito contable, entonces la función de distribución es definida exactamente como en la Definición 1.2, excepto que la suma debe ser una suma infinita *convergente*. El Teorema 1.1 es todavía verdad, como sus extensiones los Teoremas 1.2 y 1.4. Una cosa que no podemos hacer en un espacio muestral infinito contable que podríamos hacer en un espacio muestral finito es definir una función de distribución *uniforme* como en la Definición 1.3. Se pide explicar en el Ejercicio 20 por qué esto no es posible.

Ejemplo 1.13 Una moneda se arroja hasta la primera vez que salga cara. Sea el resultado del experimento, ω, la primera vez que sale cara. Entonces los resultados posibles de nuestros experimento son

$$\Omega = \{1, 2, 3, \ldots\} \ .$$

Note que aunque la moneda pudiera salir cruz en cada vez, no hemos permitido esta posibilidad. Explicaremos el por qué en un momento. La probabilidad que salga cara en la primer tirada es $1/2$. La probabilidad que salga cruz en la primera tirada y cara en la segunda es $1/4$, la probabilidad que tengamos dos cruces seguidas por una cara es $1/8$, y así en adelante. Esto sugiere asignar la función de distribución $m(n) = 1/2^n$ para $n = 1, 2, 3, \ldots$. Para ver que esta es una función de distribución debemos mostrar que

$$\sum_\omega m(\omega) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 1 \ .$$

Que es verdad sigue de la fórmula para la suma de una serie geométrica,

$$1 + r + r^2 + r^3 + \cdots = \frac{1}{1-r} \ ,$$

o

$$r + r^2 + r^3 + r^4 + \cdots = \frac{r}{1-r} \ , \tag{1.2}$$

para $-1 < r < 1$.

Poniendo $r = 1/2$, vemos que tenemos una probabilidad de 1 de que eventualmente la moneda salga cara. El resultado posible de cada vez cruces se la ha asignado probabilidad 0, por lo tanto lo omitimos de nuestro espacio muestral de posibles resultados.

Sea E el suceso que la primera vez que salga cara sea despues de un número par de tiradas. Entonces

$$E = \{2, 4, 6, 8, \ldots\} \ ,$$

y

$$P(E) = \frac{1}{4} + \frac{1}{16} + \frac{1}{64} + \cdots \ .$$

Poniendo $r = 1/4$ en la Ecuación 1.2 vemos que

$$P(E) = \frac{1/4}{1 - 1/4} = \frac{1}{3} \ .$$

Así la probabilidad que salga cara la primera vez después de un número par de tiradas es $1/3$ y después de un número impar de tiradas es $2/3$. □

Reseña Histórica

Una pregunta interesante en la historia de la ciencia es: Por qué la Probabilidad no se desarrolló hasta el siglo XVI? Sabemos que en dicho siglo los problemas relacionados con apuestas y juegos de azar contribuyeron a que la gente empezara a pensar en la Probabilidad. Pero las apuestas y juegos de azar son casi tan antiguos como la misma civilización. En el antiguo Egipto (en los tiempos de la Dinastia I, ca. 3500 B.C.) era popular un juego conocido hoy como "Sabuesos y Chacales". En este juego, el movimiento de los sabuesos y de los chacalaes se basaba en el resultado obtenido al arrojar un dado de cuatro caras, hecho con huesos de animales, llamado astragali. Dados de seis caras de diversos materiales datan del siglo XVI B.C.. Los juegos de apuestas se extendieron en la antigua Grecia y Roma. De hecho, en el Imperio Romano fue necesario en ocasiones sancionar leyes de prohibición del juego. Por qué, entonces, las probabilidades no fueron calculadas hasta el siglo XVI?. Varias explicaciones fueron propuestas para este desarrollo tardío. Una es que no había sido desarrollada la matemática relevante ni era facil de desarrollar. La notación matemática antigua complicaba el cálculo numérico, y nuestra notación algebraica de uso común no fué desarrollada hasta el siglo XVI. Sin embargo, como veremos más adelante, muchos de los conceptos de conbinatoria necesarios para calcular probabilidades habían sido discutidos anteriormente. Dado que muchos de los eventos de azar de esos tiempos tenían que ver con loterias relacionadas a asuntos religiosos, se ha sugerido que barreras religiosas se han impuesto en el estudio de la probabilidad y el juego. Otra sugerencia es que un incentivo más poderoso, como el desarrollo del comercio, era necesario. Sin embargo, ninguna de estas explicaciones parace completamente satisfactoria, y muchos aún se preguntan por qué el esudio de la probabilidad se demoró tanto tiempo. Una discusión interesante al respecto se halla en Hacking.[14]

La primera persona que calculó sitemáticamente probabilidades fue Gerolamo Cardano (1501–1576) en su libro *Liber de Ludo Aleae.* Éste fue traducido del latin por Gould y aparece en el libro *Cardano: The Gambling Scholar* por Ore.[15] Ore presenta una discusión fascinante acerca de la vida de este colorido erudito con relatos acerca de su interés en muchos campos diferentes, incluidas la medicina, la astrología y la matemática. También se encuentra allí una descripción detallada de la famosa batalla entre Cardano y Trataglia por la solución a la ecuación cúbica.

En su libro sobre probabilidad, Cardano trata únicamente el caso especial que hemos llamado función de distribución uniforme. Esta restricción a resultados igualmente probables continuaría por mucho tiempo. En este caso, Cardano comprendió que la probabilidad de que un evento ocurra es la relación entre el número de resultados favorables y el número total de resultados posibles.

Muchos de los ejemplos de Cardano consisten en arrojar dados. Así, él comprendió que los resultados al arrojar dos dados deben ser vistos como 36 pares ordenados (i, j) en lugar de 21 pares desordenados. Este es un detalle sutil que segiría causando problemas a escritores de libros de probabilidad posteriores. Por ejemplo,

[14] I. Hacking, *The Emergence of Probability* (Cambridge: Cambridge University Press, 1975).
[15] O. Ore, *Cardano: The Gambling Scholar* (Princeton: Princeton University Press, 1953).

en el siglo XVIII el famoso matemático francés d'Alembert, autor de varios trabajos sobre probabilidad, sostuvo que al arrojar una moneda dos veces el número de caras resultantes sería 0, 1, or 2, y por lo tanto deberían asignarse iguales probabilidades a los tres resultados posibles.[16] Cardano escogió el espacio muestral correcto para su problema de los dados y calculó las probabilidades correctas para una variedad de eventos.

Concejos para el jugador en potencia se entremesclan dentro del trabajo matemático de Cardano en breves párrafos intitulados, por ejemplo: "Quién debiería jugar y cuando", "Por qué Aristóteles condenó el Juego", "Es verdad que los que enseñan también juegan bien" y así sucesivamente. En un párrafo titulado "El principio fundamental del juego", Cardano escribe:

> El principio más importante de todos en los juegos de azar es simplemente igualdad de condiciones, por ejemplo: de los oponentes, de los observadores, del dinero, de la situación, del cubilete y del propio dado. Cuando te alejas en cierto grado de esta igualdad, si es en favor de tu oponente, eres un tonto; si es a tu favor, eres injusto.[17]

Cardano cometió errores, y si más tarde logr'o advertirlos, nunca se detuvo a corregirlos. Por ejemplo, para un evento que es favorable en tres de cada cuatro casos, Cardano asignó las apuestas correctas $3 : 1$ de que el evento ocurra pero luego asignó apuestas elevando estos números al cuadrado (i.e., $9 : 1$) para que el evento ocurra dos veces seguidas. Luego, considerando el caso donde las apuestas son $1 : 1$, el advirtió que su teoría era errónea y llegó así al resultado correcto cuando f de cada n resultados son favorables, las probabilidades de que el resultado sean favorables dos veces seguidas son $f^2 : n^2 - f^2$. Ore señala que esto es equivalente a comprender que si la probabilidad de que ocurra un evento es p, la probabilidad de que ocurra dos veces es p^2. Cardano procedió a estableces que para tres éxitos la fórmula sería p^3 y para cuatro p^4, aclarando que él entendía que la probabilidad es p^n para n éxitos en n repeticiones independientes de tal experimento. Esto seguirá del concepto de indepencia que se introduce en Section 4.1.

El trabajo de Cardano fue un primer intento extraordinario de escribir las leyes de la probabilidad, pero no fue la chispa que inició el estudio sistemático de la materia. Ésta fue consecuencia de la famosa serie de cartas entre Pascal y Fermat. Dicha correspondencia fue iniciada por Pascal al consultar con Fermat acerca de los problemas que le había asignado Chevalier de Méré,, escritor conocido, figura prominente de la corte de Louis XIV, y ardiente jugador.

El primer problema que Méré presentó fue uno de dados. La historia cuenta que él había estado apostando a que al menos un seis se obtendría al arrojar un dado cuatro veces y ganaba muy seguido. Entonces luego apostó a que se obtendría un par de seis al arrojar dos dados 24 veces. La probabilidad de un seis con un dado es $1/6$ y, por la ley del producto de experimentos independientes, la probabilidad de dos seises al arrojar dos dados es $(1/6)(1/6) = 1/36$. Ore[18] sostiene que las reglas

[16] J. d'Alembert, "Croix ou Pile," in *L'Encyclopédie,* ed. Diderot, vol. 4 (Paris, 1754).

[17] O. Ore, op. cit., p. 189.

[18] O. Ore, "Pascal and the Invention of Probability Theory," *American Mathematics Monthly*,

de juego de la época sugerían que, dado que cuatro repeticiones eran condición favorable para la ocurrencia de un evento con probabilidad $1/6$, para un evento seis veces menos probable, $6 \cdot 4 = 24$ repeticiones serían suficientes para una apuesta favorable. Pascal demostró, mediante cálculo exacto, que 25 tiradas son necesarias para una apuesta favorable para un par de seis.

El segundo problema fue mucho más difícil: era una vieja cuestión que concernía la determinación de una división imparcial de las apuestas en un torneo cuando la serie, por alguna razón, es interrumpida antes de completarse. Hoy se hace referencia a este caso como el problema del puntaje. Éste había sido un problema estándar en textos matemáticos; apareció en el libro de Fray Luca Paccioli *summa de Arithmetica, Geometria, Proportioni et Proportionalità*, impreso en Venecia en 1494,[19] de la siguiente manera:

> Un equipo juega a la pelota de forma tal que un total de 60 puntos se requiere para ganar el juego, y cada turno vale 10 puntos. Las apuestas son 10 ducados. Por algún incidente no se puede continuar el juego cuando un equipo tiene 50 puntos y el otro 20. Se desea saber qué parte del premio monetario corresponde a cada cual. En este caso he descubierto que las opiniones difieren una de otra aunque todas me parecen insuficientes en sus argumentos. Yo expondré la verdad y mostraré la forma correcta.

Soluciones razonables, tales como dividir las apuestas de acuerdo a la relación de turnos ganado por cada jugador, habían sido propuestas, pero ninguna solución correcta se había hallado en el período de la correspondencia entre Pascal y Fermat. Las cartas se tratan mayormente de intentos de resolver esta incógnita por ambos matemáticos. Blaise Pascal (1623–1662) era un niño prodigio, había publicado su tratado de secciones cónicas a la edad de dieciséis, y había inventado una máquina de calcular a la edad de dieciocho. Para el tiempo de las cartas, su demostración del peso de la atmósfera ya había establecido su posición a la vanguardia de los físicos contemporáneos. Pierre de Fermat (1601–1655) era un culto jurista de Toulouse, quien estudiaba matemática en su tiempo libre. Para algunos, él ha sido el príncipe de los aficionados y uno de los más grandes matemáticos puros de todos los tiempos.

Las cartas, traducidas por Maxine Merrington, aparecen en el fascinante informe histórico de Florence David *Games, Gods and Gambling*.[20] En una carta con fecha miércoles 29 de julio de 1654, Pascal escribe a Fermat:

> Señor, Como usted, yo estoy igualmente impaciente, y aunque nuevamente estoy enfermo en cama, no puedo dejar de decirle que ayer en la tarde recibí de manos de M. de Carcavi su carta acerca del problema del puntaje, cosa que admiro más de lo que soy capaz de expresar con palabras. No dispongo de tiempo para explayar este mensaje, pero, en una palabra, usted ha resuelto los dos problemas del puntaje, el de los dados y el otro que consiste en una serie de juegos, con imparcialidad

vol. 67 (1960), pp. 409–419.

[19] ibid., p. 414.

[20] F. N. David, *Games, Gods and Gambling* (London: G. Griffin, 1962), p. 230 ff.

> perfecta; estoy totalmente satisfecho con el resultado pues sin dudas yo estaba equivocado, viendo con admiración que ahora estoy de acuerdo con usted...
>
> Su método es muy consistente y es el primero que tuve en mente en esta investigación; pero dado que la labor combinatoria es excesiva, he descubierto un atajo y en efecto un nuevo método mucho más rápido y prolijo, que me gustaría mencionar en pocas palabras: Porque en adelante quisiera abrir mi corazón a usted, si me permite, ya que me encuentro rebosante de alegría con nuestro acuerdo. Veo que la verdad en Toulouse es la misma que en París.
>
> Aquí, más o menos, es lo que hago para mostrar el valor equitativo de cada juego, cuando dos oponentes juegan, por ejemplo en tres partidas y cada persona ha apostado 32 doblones.
>
> Digamos que el primer hombre había ganado dos veces y el otro sólo una; ahora ellos juegan otra partida, en la que las condiciones son tales que, si el primero gana, se lleva todas las apuestas; esto es 64 doblones; y si el otro gana, cada uno habrá ganado dos partidas, y por lo tanto, si desearan detener el juego, cada cual debería recoger su propia apuesta, esto es, 32 doblones cada jugador.
>
> Entonces considere, señor, que si el primer hombre ganara, se llevaría 64 doblones; y si perdiera se iría con 32. Así, si no quisieran arriesgar esta última partida y pretendieran separarse sin jugarla, el primer hombre debería decir: 'Tengo la certeza de haber obtenido 32 doblones, aun si perdiera serían míos; pero en cuanto a los otros 32, quizás yo los gane, o quizás usted, las chances son las mismas. Dividamos entonces estos 32 doblones en partes iguales y yo tomaré una mitad, además de los otros 32 que tengo asegurados.' él tendrá entonces 48 doblones y los otros 16...

El argumento de Pascal produce la tabla ilustrada en la Figure 1.9 para la cantidad a pagar al jugador A cualquiera sea el punto de abandono.

Cada entrada en la tabla es el promedio de los números por encima y a la derecha del número en cuestión. Este hecho, junto con los valores conocidos al completarse el torneo, determinan todos los valores de esta tabla. Si el jugador A gana la primera partida, él necesitará dos partidas para ganar, mientras que B requerirá tres para tal fin; y entonces, si se suspende el torneo, A debería recibir 44 pistolas.

La carta en la que Fermat presentó su solución se ha perdido; pero por fortuna, Pascal describe el método de Fermat en una carta con fecha Lunes 24 de Agosto de 1654. De la carta de Pascal:[21]

> Este es su procedimiento en caso de que haya dos jugadores: Si dos oponentes, jugando varias partidas, arriban a la situación en la cual el

[21] ibid., p. 239ff.

Number of games B has won				
3	0	0	0	
2	8	16	32	64
1	20	32	48	64
0	32	44	56	64
	0	1	2	3
	Number of games A has won			

Figure 1.9: Tabla de Pascal.

> primero necesita *dos* victorias y el segundo *tres*, para llegar a la división justa de las apuestas, usted afirma que uno debe saber en cuantas partidas el juego estará absolutamente decidido.
>
> Es sencillo calcular que esto ocurrirá en *cuatro* partidas, de lo cual se puede concluir que es necesario ver en cuántas formas cuatro partidas pueden ser arregladas entre dos jugadores, y uno debe entender cuantas combinaciones harían ganador al primer oponente y cuantas al segundo, y cómo distribuir las apuestas en esta proporción. Yo hubiera encontrado este problema difícil de comprender si no lo hubiera sabido de antemano; en realidad usted lo había explicado con esta idea en mente.

Fermat entendió que las distintas maneras en las que el juego podría terminarse no eran siempre igualmente probables. Por ejemplo, cuando A necesita dos partidas y B requiere tres para ganar, dos maneras posibles en las que el torneo terminaría en favor del primero son WLW y LWLW. Estas dos secuencias no tienen la misma probabilidad de ocurrir. Para evitar esta dificultad, Fermat extendió el juego agregando partidas ficticias, de manera tal que cualquier desenlace tendría la misma duración, a saber cuatro partidas. él fue suficientemente astuto como para darse cuenta de que esta extensión no habría de cambiar el ganador, y que en adelante, él podría simplemente contar el número de secuencias favorables para cada jugador ya que ahora eran todas igualmente probables. Si listamos todas las secuencias posibles que se pueden dar en el juego extendido de cuatro partidas, obtendremos los 16 resultados siguientes:

El jugador A gana en caso de que haya al menos dos victorias (los 11 casos subrayados), y B hará lo propio cuando haya al menos tres derrotas (los otros 5 casos). Dado que A triunfa en 11 de cada 16 casos posibles Fermat argumentó que la probabilidad en favor de A es $11/16$. Si las apuestas son 64 doblones, de acuerdo con

WWWW	WLWW	LWWW	LLWW
WWWL	WLWL	LWWL	LLWL
WWLW	WLLW	LWLW	LLLW
WWLL	WLLL	LWLL	LLLL .

el resultado de Pascal A debería recibir 44. Pascal y Fermat desarrollaron métodos más sistemáticos de contar el número de resultados favorables en problemas de este tipo, y éste será uno de nuestros problemas centrales. Tales métodos son parte de los contenidos de *Combinatoria*, tema a tratar en el Chapter 3.

Vemos que estos dos matemáticos arribaron a dos maneras distintas de resolver el problema del puntaje. El método de Pascal consistía en desarrollar un algoritmo y luego utilizarlo para calcular la división equitativa. El método de Fermat, por otra parte, consistía en cambiar el problema por uno equivalente para el cual se podían utilizar métodos de cálculo o de combinatoria. En el Chapter 3 veremos que, en realidad, Fermat utilizaba lo que se llegó a conocer como Triángulo de Pascal! Hoy en día al estudiar Probabilidad advertimos que ambos enfoques, algorítmico y combinatorio, son igualmente importantes, tal como lo fueron 300 años atrás cuando la probabilidad se iniciaba.

Ejercicios

1 Sea $\Omega = \{a, b, c\}$ un espacio muestral. Sea $m(a) = 1/2$, $m(b) = 1/3$, y $m(c) = 1/6$. Hallar las probabilidades para cada uno de los subconjuntos de Ω.

2 Dar el espacio muestral posible Ω para cada uno de los siguientes experimentos:

(a) Una elección a decidirse entre dos candidatos A y B.

(b) Se arroja una moneda.

(c) Se pregunta a un estudiante el mes del año y el dá de la semana en que cae su cumpleaños.

(d) Se elige un estudiante al azar de una clase de diez alumnos.

(e) Recibes una calificación en este curso.

3 Para cuáles de los casos del Ejercicio 2 sería razonable asignar la función de distribución uniforme?

4 Describir con palabras los eventos especificados por los siguientes subconjuntos de

$$\Omega = \{HHH,\ HHT,\ HTH,\ HTT,\ THH,\ THT,\ TTH,\ TTT\}$$

(ver Ejemplo 1.6).

(a) $E = \{\mathrm{HHH,HHT,HTH,HTT}\}$.

(b) $E = \{\text{HHH,TTT}\}$.

(c) $E = \{\text{HHT,HTH,THH}\}$.

(d) $E = \{\text{HHT,HTH,HTT,THH,THT,TTH,TTT}\}$.

5 Cuáles son las probabilidades de los eventos descriptos en el Ejercicio 4?

6 Un dado se carga de manera tal que la probabilidad de que cada una de las caras salga es proporcional al número de puntos que ésta contenga. (Por ejemplo, la probabilidad de que salga un seis es tres veces la de un dos.) Cuál es la probabilidad de obtener un número par en una tirada?

7 Sea A y B eventos tales que $P(A \cap B) = 1/4$, $P(\tilde{A}) = 1/3$, y $P(B) = 1/2$. Cuál es el valor de $P(A \cup B)$?

8 Un estudiante debe elegir una de los temas, arte, geología, o psicología, como materia optativa. Es igualmente probable que ella elija arte o psicología pero dos veces probable que elija geología. Cuáles son las respectivas probabilidades de que elija arte, geología, y psicología?

9 Un estudiante debe elegir dos entre tres materias optativas: arte, francés, y matemática. Si elige arte con probabilidad 5/8, francés con probabilidad 5/8, y arte y francés con probabilidad 1/4, Cuál es la probabilidad de que elija matemática? Cuál es la probabilidad de que elija arte o francés?

10 Para que un proyecto sea presentado al presidente de los Estados Unidos, debe pasar primero por la Cámara de los Representantes y el Senado. Asumir que, de los proyectos presentados a estas dos entidades, 60 menos una de las dos. Calcular la probabilidad de que el próximo proyecto presentado a los dos grupos llegará al presidente.

11 Qué probabilidad debería dar una persona a favor de los siguientes eventos?

(a) Una carta elegida al azar de un mazo de 52 cartas será un as.

(b) Saldrán dos caras saldrán al arrojar una moneda dos veces.

(c) Cuando se arrojan dos dados saldrán Boxcars (dos seis).

12 Apuestas 3 : 1 a que tu amigo Smith será elegido intendente de tu ciudad Qué probabilidad estás asignando al evento de que Smith gane?

13 En una carrera de caballos, las apuestas se listan 2 : 3 para Romance y 1 : 2 para Downhill. Cómo se listarán las apuestas para el caso de que cualquiera de los dos gane?

14 Sea X una variable aleatoria con función de distribución $m_X(x)$ definida por

$$m_X(-1) = 1/5, \quad m_X(0) = 1/5, \quad m_X(1) = 2/5, \quad m_X(2) = 1/5 \ .$$

(a) Sea Y la variable aleatoria definida por la ecuación $Y = X + 3$. Hallar la función de distribución $m_Y(y)$ de Y.

(b) Sea Z la variable aleatoria definida por la ecuación $Z = X^2$. Hallar la función de distribución $m_Z(z)$ de Z.

***15** John y Mary están tomando un curso de matemática. El sistema de calificación tiene tres notas: A, B y C. La probabilidad de que John saque una B es 0,3. La probabilidad de que Mary saque una B es 0,4. La probabilidad de que ninguno saque una A pero al menos uno saque una B es 0,1. cuál es la probabilidad de que al menos uno saque una B pero ninguno saque una C?

16 En una batalla feroz, no menos del 70 perdieron un ojo, no menos del 75 una mano y no menos del 85para aquellos que simultáneamente perdieron una oreja, un ojo, una mano, y una pierna?[22]

***17** Asumir que la probabilidad de un "éxito" en un único experimento con n resultados es $1/n$. Sea m el número de experimentos necesarios para que sea favorable la apuesta a que al menos un éxito ocurra (ver Ejercicio 1.1.5).

(a) Demostrar que la probabilidad de que en m intentos no haya éxito es $(1 - 1/n)^m$.

(b) (de Moivre) Demostrar que si $m = n \log 2$ entonces

$$\lim_{n\to\infty} \left(1 - \frac{1}{n}\right)^m = \frac{1}{2} \ .$$

Hint:

$$\lim_{n\to\infty} \left(1 - \frac{1}{n}\right)^n = e^{-1} \ .$$

Por lo tanto para grandes valores de n deberíamos elegir m alrededor de $n \log 2$.

(c) Hubiera llegado DeMoivre a la respuesta correcta para las dos apuestas de de Méré si hubiera utilizado esta aproximación?

18 (a) Para los eventos $A_1, \ldots, A_n$, probar que

$$P(A_1 \cup \cdots \cup A_n) \le P(A_1) + \cdots + P(A_n) \ .$$

(b) Para los eventos A y B, probar que

$$P(A \cap B) \ge P(A) + P(B) - 1.$$

19 Si A, B, y C son tres eventos cualquiera, demostrar que

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

[22]See Knot X, in Lewis Carroll, *Mathematical Recreations,* vol. 2 (Dover, 1958).

20 Explicar por qué no es posible definir una función de distribución uniforme (ver Definición 1.3) en un espacio muestral infinito numerable. *Sugerencia*: Asumir que $m(\omega) = a$ para todo ω, donde $0 \leq a \leq 1$. Tiene $m(\omega)$ todas las propiedades de una función de distribución?

21 En el Ejemplo 1.13 encontrar la probabilidad de sacar una cara por primera vez en la décima, onceava, y doceava tirada de la moneda.

22 Se tira un dado hasta que se obtiene un seis por primera vez. Veremos que la probabilidad de que esto ocurra en la enésima tirada es $(5/6)^{n-1} \cdot (1/6)$. Utilizando este dato, describir el espacio de muestral infinito y la función de distribución apropiados para el experimento de tirar un dado hasta que se obtiene un seis por primera vez. Verificar que para tu función de distribución $\sum_\omega m(\omega) = 1$.

23 Sea Ω el espacio muestral

$$\Omega = \{0, 1, 2, \ldots\} ,$$

y sea la función de distribución definida por

$$m(j) = (1 - r)^j r ,$$

para algún r fijo, $0 < r < 1$, y para $j = 0, 1, 2, \ldots$. Demostrar que ésta es una función de distribución para Ω.

24 Nuestro calendario tiene un ciclo de 400 años. B. H. Brown advirtió que el número de veces que el día trece del mes cae en cada uno de los días de la semana en los 4800 meses de un ciclo es como sigue:

Domingo 687

Lunes 685

Martes 685

Miércoles 687

Jueves 684

Viernes 688

Sábado 684

De esto él dedujo que el trece tenía más chances de caer viernes que cualquier otro día de la semana. Explicar que quiso decir con esto.

25 Tversky and Kahneman[23] le pidieron a un grupo de individuos llevar a cabo la siguiente tarea. Se les dice que:

> Linda tiene 31 años, soltera, abierta, y muy brillante. Se licenció en filosofía en la universidad. Como estudiante, se preocupó profundamente por la discriminación racial y otros asuntos sociales, y participó en manifestaciones antinucleares.

[23] K. McKean, "Decisions, Decisions," pp. 22–31.

Luego se pide a los participantes que califiquen la probabilidad de varias alternativas, tales como:
(1) Linda participa activamente en el movimiento feminista.
(2) Linda es una cajera bancaria.
(3) Linda es una cajera bancaria y participa activamente en el movimiento feminista.

Tversky y Kahneman descubrieron que entre 85 más probable, pero la alternativa (3) más probable que la alternativa (2). Es así? Ellos llamaron este fenómeno la *falacia de conjunción*, y notaron que parece no estar afectado por un entrenamiento previo en probabilidad y estadística. Explicar por qué esto es una falacia. Podrías dar una posible explicación para la elección de los participantes?

26 Se escogen dos cartas sucesivamente de un mazo de 52. Encontrar la probabilidad de que la segunda carta sea de mayor valor que la primera. *Sugerencia*: Mostrar que $1 = P(\text{higher}) + P(\text{lower}) + P(\text{same})$ y utilizar el hecho de que $P(\text{higher}) = P(\text{lower})$.

27 Una *tabla de vida* es un diagrama que muestra para un número dado de nacimientos el número estimado de personas que vivirán hasta cierta edad. En el Apéndice C presentamos una tabla de vida basada en 100.000 nacimientos para edades de 0 a 85 años, tanto para hombres como para mujeres. Mostrar de qué manera se puede estimar, basándose en esta tabla, la probabilidad $m(x)$ de que una persona nacida en 1981 viviría hasta la edad x. Escribir un programa que trace $m(x)$ para hombres y para mujeres, y comentar las diferencias que se evidencian en ambos casos.

***28** Este es un intento de justificar el hecho de que no somos capaces de elegir un "entero al azar."

(a) Cuál es, intuitivamente, la probabilidad de que un entero positivo "elegido en forma aleatoria" sea múltiplo de 3?

(b) Sea $P_3(N)$ la probabilidad de que un entero, elegido al azar entre 1 y N, sea múltiplo de 3 (dado que el espacio muestral es finito, ésta es una probabilidad legítima). Mostrar que el límite

$$P_3 = \lim_{N\to\infty} P_3(N)$$

existe y es igual a $1/3$. Esto formaliza la intuición en (a), y nos brinda una forma de asignar "probabilidades" a ciertos eventos que son subconjuntos infinitos de los enteros positivos.

(c) Si A es un conjunto de enteros positivos cualquiera, sea $A(N)$ el número de elementos de A que son menores o iguales a N. Luego, definir la "probabilidad" de A como

$$P(A) = \lim_{N\to\infty} A(N)/N\ ,$$

siempre y cuando este límite exista. Probar que esta definición asignaría probabilidad 0 a cualquier conjunto finito y probabilidad 1 al conjunto de todos los enteros positivos. Así, la probabilidad del conjunto de todos los enteros no es la suma de las probabilidades de los enteros individuales en este conjunto. Esto significa que la definición de probabilidad que aquí se da no es completamente satisfactoria.

(d) Sea A el conjunto de todos los enteros positivos que tienen un número impar de dígitos. Demostrar que $P(A)$ no existe. Esto prueba que bajo la definición de probabilidad dada anteriormente, no todos los conjuntos tienen probabilidad.

29 (from Sholander[24]) En un cruce estándar de hojas de trébol (OJO), hay cuatro rampas para efectuar giros a la derecha, y dentro de éstas, hay cuatro rampas más para giros a la izquierda. Tu auto se aproxima al cruce desde el sur. Se ha instalado un mecanismo que en cada punto donde exista una elección de dirección, de modo que el auto gire a la derecha con probabilidad fija r.

(a) Si $r = 1/2$, Cuál es tu chance de emerger del cruce que va hacia el oeste?

(b) Hallar el valor de r que maximiza tus chances dejar el cruce con dirección oeste.

30 (from Benkoski[25]) Considera un cruce de hojas de trébol "puro", en el cual no hay rampas para giros a la derecha, sino únicamente las dos autopistas rectas que intersecan con hojas de trébol para giros a la izquierda. (Así, para efectuar un giro a la derecha en un cruce como éste, uno debe hacer tres giros a la izquierda.) Como en el problema anterior, tu auto se aproxima al cruce desde el sur. Cuál es el valor de r que maximiza tus chances de dejar el cruce con dirección este?

31 (from vos Savant[26]) Un lector de la columna de Marilyn vos Savant escribió una carta con la siguiente pregunta:

> Mi padre oyó esta historia en la radio. En la Universidad de Duke, dos estudiantes habían recibido notas A en química todo el semestre. Pero en la noche anterior al examen final, estuvieron de fiesta en otro estado y no pudieron volver a Duke a tiempo. Se excusaron con el profesor diciendo que se les había desinflado una rueda, y preguntaron si se les podría otorgar una mesa especial. El profesor accedió, escribió el examen y los ubicó en distintas aulas para que lo resolvieran por separado. La primera pregunta (de un lado de la hoja) valía 5 puntos, y los estudiantes la resolvieron fácilmente. Luego dieron vuelta de página y encontraron la segunda pregunta,

[24] M. Sholander, Problem #1034, *Mathematics Magazine,* vol. 52, no. 3 (May 1979), p. 183.
[25] S. Benkoski, Comment on Problem #1034, *Mathematics Magazine,* vol. 52, no. 3 (May 1979), pp. 183-184.
[26] M. vos Savant, *Parade Magazine*, 3 March 1996, p. 14.

que valía 95 puntos: 'Qué rueda se desinfló?' Cuál era la probabilidad de que ambos estudiantes respondieran lo mismo? Mi padre piensa que es 1 en 16. Es correcto?"

(a) Es la respuesta $1/16$?

(b) La siguiente pregunta se preguntó a los estudiantes de una clase. "Manejaba hoy hacia la escuela, y una de mis cubiertas se desinfló. Qué cubierta crees que fue?" Las respuestas fueron las siguientes: delantera derecha 58 trasera derecha 18 en la población general, y asumamos que los dos que toman el examen son escogidos al azar entre la población general. Cuál es la probabilidad de que den la misma respuesta a la segunda pregunta?