Este plan requiere que el consumo de botellas de vino comience en un nivel muy alto y decrezca a un índice continuo de 5 por ciento anual. Dado que el consumo es continuamente decreciente, debemos usar integración para calcular el consumo de vino en un año particular (x) como sigue:

$$\begin{aligned}\text{consumo en un año } x &\approx \int_{x-1}^{x} c(t)\,dt = \int_{x-1}^{x} 79e^{-0.05t}\,dt = -1\,580e^{-0.05t}\Big|_{x-1}^{x} \\ &= 1\,580(e^{-0.05(x-1)} - e^{-0.05x}).\end{aligned} \quad \textbf{(2.165)}$$

Si $x = 1$, el consumo es aproximadamente de 77 botellas en el primer año. Luego decrece de manera uniforme, terminando en aproximadamente 30 botellas consumidas en el año 20.

PREGUNTA: Nuestra primera ilustración fue sólo un ejemplo de la segunda en la que $\delta = \gamma = 0$. Explica cómo valores alternos de estos parámetros afectarán la trayectoria de consumo óptimo de vino. Explica tus resultados intuitivamente (para más información sobre consumo óptimo en el tiempo, véase el capítulo 17).

ESTADÍSTICA MATEMÁTICA

En los últimos años la teoría microeconómica se ha centrado cada vez más en problemas ofrecidos por la incertidumbre y la información imperfecta. Para comprender gran parte de esa bibliografía es importante tener firmes bases de estadística matemática. Así, el propósito de esta sección es resumir algunos principios estadísticos que encontraremos en distintos apartados de este libro.

Variables aleatorias y funciones de densidad de probabilidad

Una *variable aleatoria* describe (en forma numérica) los resultados de un experimento sujeto al azar. Por ejemplo, podríamos lanzar una moneda y observar si sale cara o cruz. Si llamamos a esta variable aleatoria x, podemos denotar los resultados posibles ("realizaciones") de la variable como:

$$x = \begin{cases} 1 & \text{si la moneda es cara,} \\ 0 & \text{si la moneda es cruz.} \end{cases}$$

Nótese que antes de lanzar la moneda x puede ser 1 o 0. Sólo después de despejada la incertidumbre (es decir, después de lanzada la moneda) sabremos cuál es el valor de x.[23]

Variables aleatorias discretas y continuas

Los resultados de un experimento aleatorio pueden ser un número finito de posibilidades o un continuo de posibilidades. Por ejemplo, el número que resulta de lanzar un dado es una variable aleatoria con seis resultados posibles. Si se lanzan dos dados podría registrarse la suma de sus caras (en cuyo caso hay 12 resultados, algunos de los cuales tienen más probabilidades que otros) o un número de dos dígitos, uno para el valor de cada dado (en cuyo caso habría 36 resultados igualmente probables). Estos son ejemplos de variables aleatorias discretas.

Por otro lado, una variable aleatoria continua puede adoptar cualquier valor en un rango dado de números reales. Por ejemplo, la temperatura a la intemperie del día de mañana podría verse

[23] A veces las variables aleatorias se denotan con $\tilde{x}$ para distinguir entre variables cuyo resultado está sujeto al azar y variables algebraicas (no aleatorias). Este recurso de notación puede ser útil para seguir la pista de lo que es aleatorio y lo que no en un problema particular, y lo usaremos en algunos casos. Cuando no haya ambigüedad, sin embargo, no emplearemos esta notación especial.

como una variable continua (suponiendo que las temperaturas pueden medirse con precisión) que va de, digamos, -50 °C a $+50$ °C. Desde luego que algunas de estas temperaturas serían de ocurrencia improbable, pero en principio la temperatura medida de manera precisa podría estar en cualquier punto entre estos dos límites. De igual forma, el cambio porcentual del día de mañana en el valor de un índice bursátil particular podría verse como susceptible de adoptar todos los valores entre -100% y, digamos, $+1\,000\%$. De nueva cuenta, por supuesto, serían mucho más probables de ocurrir las variaciones porcentuales alrededor de 0% que los valores extremos.

Funciones de densidad de probabilidad

La *función de densidad de probabilidad* (FDP) para cualquier variable aleatoria indica la probabilidad de que ocurra cada resultado específico. Para a una variable aleatoria discreta definir esa función no plantea ninguna dificultad particular. En el caso del lanzamiento de una moneda, por ejemplo, la FDP [denotada por $f(x)$] estaría dada por

$$\begin{aligned} f(x=1) &= 0.5, \\ f(x=0) &= 0.5. \end{aligned} \tag{2.166}$$

Para el caso del lanzamiento de un dado la FDP sería:

$$\begin{aligned} f(x=1) &= 1/6, \\ f(x=2) &= 1/6, \\ f(x=3) &= 1/6, \\ f(x=4) &= 1/6, \\ f(x=5) &= 1/6, \\ f(x=6) &= 1/6. \end{aligned} \tag{2.167}$$

Nótese que en ambos casos las probabilidades especificadas por la FDP suman 1.0. Esto se debe a que, por definición, uno de los resultados del experimento aleatorio debe ocurrir. De modo más general, si todos los resultados de una variable aleatoria discreta se denotan como x_i para $i = 1, \ldots, n$, debemos tener:

$$\sum_{i=1}^{n} f(x_i) = 1. \tag{2.168}$$

En una variable aleatoria continua debemos tener cuidado al definir el concepto de FDP. Dado que esa variable aleatoria adopta un continuo de valores, si le asignamos cualquier valor diferente de cero como la probabilidad de un resultado específico (es decir, una temperatura de $+25.53470$ °C) pronto podríamos tener sumas de probabilidades infinitamente grandes. De ahí que para una variable aleatoria continua definamos la FDP $f(x)$ como una función con la propiedad de que la probabilidad de que x ocurra en un intervalo reducido particular dx está dada por el área de $f(x)dx$. Usando esta convención la propiedad de que las probabilidades de un experimento aleatorio sumen 1.0 se enuncia como sigue:

$$\int_{-\infty}^{+\infty} f(x)\, dx = 1.0. \tag{2.169}$$

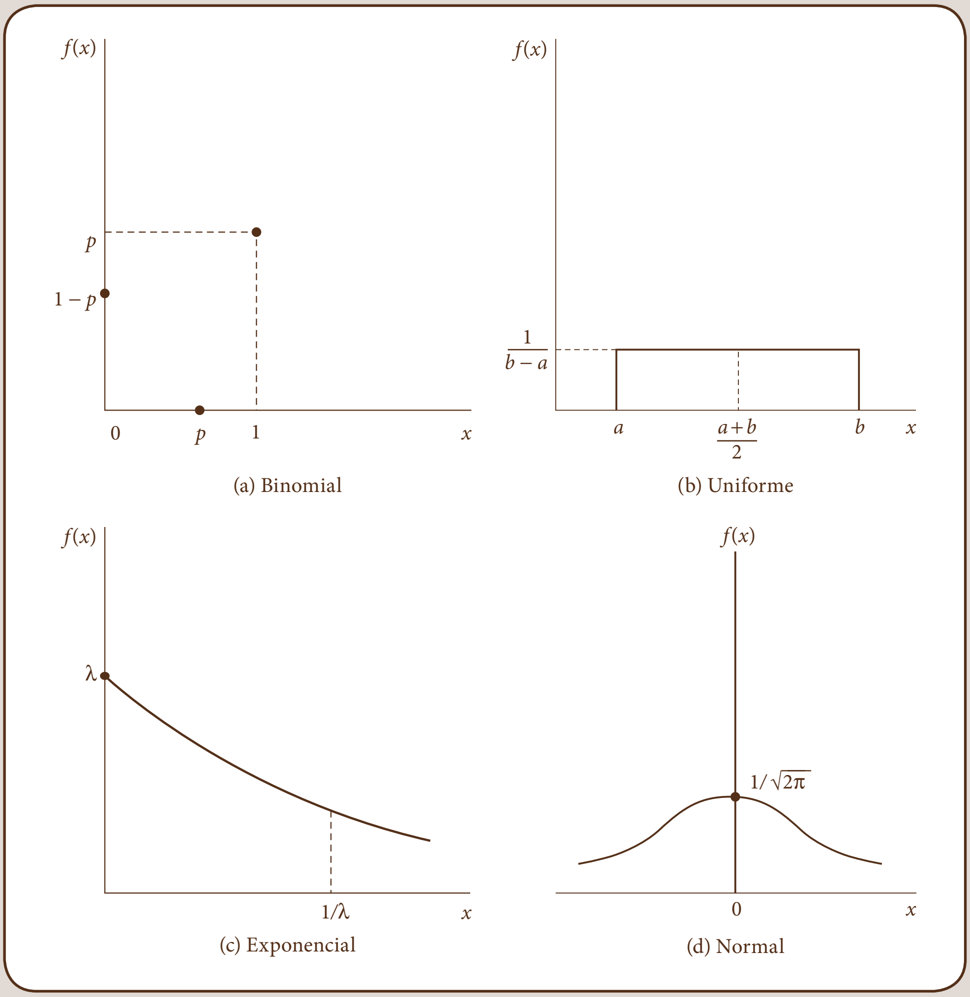
Algunas FDP importantes

La mayoría de las funciones operará como una FDP, siempre y cuando $f(x) \geq 0$ y la función sumen (o integren) 1.0. El truco, desde luego, es encontrar funciones que reflejen experimentos aleatorios que ocurren en la realidad. Aquí nos ocuparemos de cuatro de dichas funciones las cuales nos serán útiles en varias secciones de este libro. Las gráficas de esas cuatro funciones aparecen en la figura 2.6.

FIGURA 2.6

Cuatro funciones comunes de densidad de probabilidad.

Las variables aleatorias con estas FDP son de amplio uso. Cada gráfica indica el valor esperado de la FDP mostrada.



1. ***Distribuición binomial.*** Esta es la distribución discreta básica. Usualmente se supone que x adopta sólo dos valores, 1 y 0. La FDP de la binomial está dada por:

$$\begin{aligned} &f(x = 1) = p, \\ &f(x = 0) = 1 - p, \\ &\text{donde} \quad 0 < p < 1. \end{aligned} \tag{2.170}$$

   El ejemplo de lanzar una moneda es obviamente un caso especial de la binomial donde $p =$ 0.5.

2. ***Distribución uniforme.*** Esta es la FDP continua más simple. Supone que los valores posibles de la variable x ocurren en un intervalo definido y que cada valor es igualmente probable. Es decir:

$$\begin{aligned} f(x) &= \frac{1}{b-a} \quad \text{para } a \le x \le b; \\ f(x) &= 0 \quad \text{para } x < a \text{ o } x > b. \end{aligned} \tag{2.171}$$

Obsérvese aquí que las probabilidades integran 1.0:

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_a^b \frac{1}{b-a}\,dx = \frac{x}{b-a}\bigg|_a^b = \frac{b}{b-a} - \frac{a}{b-a} = \frac{b-a}{b-a} = 1.0. \tag{2.172}$$

3. ***Distribución exponencial.*** Esta es una distribución continua para la cual las probabilidades decrecen a un índice exponencial uniforme al incrementarse x. Formalmente:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0, \\ 0 & \text{si } x \le 0, \end{cases} \tag{2.173}$$

donde λ es una constante positiva. De nuevo, es fácil demostrar que esta función integra 1.0:

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_0^{\infty} \lambda e^{-\lambda x}\,dx = -e^{-\lambda x}\bigg|_0^{\infty} = 0 - (-1) = 1.0. \tag{2.174}$$

4. ***Distribución normal.*** La distribución normal (o de Gauss) es la más importante en la estadística matemática. Su importancia se deriva sobre todo del *teorema del límite central* el cual establece que la distribución de cualquier suma de variables aleatorias independientes aproximará cada vez más la distribución normal conforme el número de esas variables aumente. Dado que los promedios muestrales pueden considerarse sumas de variables aleatorias independientes, este teorema señala que cualquier promedio muestral tendrá una distribución normal sin importar cuál sea la distribución de la población de la cual se seleccionó la muestra. De ahí que suela ser apropiado suponer que una variable aleatoria tiene una distribución normal si se le puede concebir como una especie de promedio.

   La forma matemática de la FDP normal es

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \tag{2.175}$$

y esto se define para todos los valores reales de x. Aunque esta función podría parecer complicada, algunas de sus propiedades pueden describirse fácilmente. Primero, esta función es simétrica en torno a cero (a causa del término x^2). Segundo, es asintótica a cero al aumentar o reducir x. Tercero, alcanza su valor óptimo en $x = 0$. Este valor es $1/\sqrt{2\pi} \approx 0.4$. Por último, la gráfica de esta función tiene una "forma de campana" regular de uso común en la estadística. La integración de esta función es relativamente complicada (aunque fácil en coordenadas polares). La presencia de la constante $1/\sqrt{2\pi}$ es necesaria para que la función integre 1.0.

Valor esperado

El *valor esperado* de una variable aleatoria es el valor numérico que se puede esperar que tenga en promedio la variable aleatoria.[24] Se trata del "centro de gravedad" de la FDP. Para una variable aleatoria discreta que adopta los valores $x_1, x_2, \ldots, x_n$, el valor esperado se define como

$$E(x) = \sum_{i=1}^{n} x_i f(x_i) \tag{2.176}$$

[24] El valor esperado de una variable aleatoria también se conoce como *media* de esa variable. En ocasiones en el estudio del muestreo esto puede causar confusión entre el valor esperado de una variable aleatoria y el concepto diferente del promedio aritmético de la muestra.

Es decir, cada resultado es ponderado por la probabilidad de que ocurra, y se suma a todos los resultados posibles. Para una variable aleatoria continua, la ecuación 2.176 se generaliza fácilmente como

$$E(x) = \int_{-\infty}^{+\infty} xf(x)\,dx. \tag{2.177}$$

De nuevo, en esta integración cada valor de x es ponderado por la probabilidad de que ocurra dicho valor.

El concepto de valor esperado puede generalizarse para incluir el valor esperado de cualquier función de una variable aleatoria [digamos, $g(x)$]. En el caso continuo, por ejemplo, escribiríamos

$$E[g(x)] = \int_{-\infty}^{+\infty} g(x)f(x)\,dx. \tag{2.178}$$

Como caso especial, consideremos una función lineal $y = ax + b$. Entonces,

$$\begin{aligned} E(y) = E(ax+b) &= \int_{-\infty}^{+\infty} (ax+b)f(x)\;dx \\ &= a\int_{-\infty}^{+\infty} xf(x)\;dx + b\int_{-\infty}^{+\infty} f(x)\;dx = aE(x) + b. \end{aligned} \tag{2.179}$$

A veces los valores esperados se formulan en términos de la *función de distribución acumulativa* (FDA) $F(x)$, definida como

$$F(x) = \int_{-\infty}^{x} f(t)\;dt. \tag{2.180}$$

Es decir, $F(x)$ representa la probabilidad de que la variable aleatoria t sea menor que o igual a x.

Usando esta notación el valor esperado de x puede escribirse como

$$E(x) = \int_{-\infty}^{+\infty} xdF(x) \tag{2.181}$$

Debido al teorema fundamental del cálculo, la ecuación 2.181 y la ecuación 2.177 significan exactamente lo mismo.

EJEMPLO 2.15 Valores esperados de algunas variables aleatorias

Los valores esperados de cada una de las variables aleatorias con las FDP simples que ya hemos presentado son fáciles de calcular. Todos estos valores esperados están indicados en las gráficas de las FDP de las funciones en la figura 2.6.

1. ***Binomial.*** En este caso:

$$E(x) = 1 \cdot f(x=1) + 0 \cdot f(x=0) = 1 \cdot p + 0 \cdot (1-p) = p. \tag{2.182}$$

Para el caso del lanzamiento de la moneda (donde $p = 0.5$), esto indica que $E(x) = p = 0.5$; el valor esperado de esta variable aleatoria es, como quizá lo supusiste, de una mitad.

2. ***Uniforme.*** Para esta variable aleatoria continua,

$$E(x) = \int_a^b \frac{x}{b-a}\,dx = \frac{x^2}{2(b-a)}\bigg|_a^b = \frac{b^2}{2(b-a)} - \frac{a^2}{2(b-a)} = \frac{b+a}{2}. \quad \textbf{(2.183)}$$

De nuevo, como quizá conozcas, el valor esperado de la distribución uniforme está a justo medio camino entre a y b.

3. ***Exponencial.*** Para este caso de probabilidades decrecientes:

$$E(x) = \int_0^\infty x\lambda e^{-\lambda x}\,dx = -xe^{-\lambda x} - \frac{1}{\lambda}e^{-\lambda x}\bigg|_0^\infty = \frac{1}{\lambda}, \quad \textbf{(2.184)}$$

donde la integración se desprende del ejemplo de la integración por partes que ya hemos mostrado (ecuación 2.137). Nótese aquí que cuanto más rápido decrecen las probabilidades, menor es el valor esperado de x. Por ejemplo, si $\lambda = 0.5$ entonces $E(x) = 2$, mientras que si $\lambda = 0.05$, entonces $E(x) = 20$.

4. ***Normal.*** Dado que la FDP normal es simétrica en torno a cero, parece claro que $E(x) = 0$. Una prueba formal usa una integración de cambio de variable, concediendo que $u = x^2/2$ $(du = xdx)$:

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} xe^{-x^2/2}\,dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-u}\,du = \frac{1}{\sqrt{2\pi}}[-e^{-x^2/2}]\bigg|_{-\infty}^{+\infty} = \frac{1}{\sqrt{2\pi}}[0-0] = 0. \quad \textbf{(2.185)}$$

Claro que el valor esperado de una variable aleatoria distribuida normalmente (o de cualquier variable aleatoria) puede ser alterado por una transformación lineal, como se muestra en la ecuación 2.179.

PREGUNTA: Una transformación lineal cambia el valor esperado de una variable aleatoria en forma predecible; si $y = ax + b$, entonces $E(y) = aE(x) + b$. De ahí que para esta transformación [digamos, $h(x)$] tengamos $E[h(x)] = h[E(x)$. Supongamos, en cambio, que x es transformada por una función cóncava, digamos $g(x)$ con $g' > 0$ y $g'' < 0$. ¿Qué resultaría de comparar $E[g(x)]$ con $g[E(x)]$?

Nota: Esta es una ilustración de la desigualdad de Jensen, concepto que se describirá en detalle en el capítulo 7. Véase también el problema 2.14.

Varianza y desviación estándar

El valor esperado de una variable aleatoria es una medida de tendencia central. En cambio, la *varianza* de una variable aleatoria [denotada por σ_x^2 o $\mathrm{Var}(x)$] es una medida de dispersión. Específicamente, la varianza se define como la "desviación cuadrada esperada" de una variable aleatoria respecto de su valor esperado. Formalmente:

$$\mathrm{Var}(x) = \sigma_x^2 = E[(x - E(x))^2] = \int_{-\infty}^{+\infty} (x - E(x))^2 f(x)\,dx. \quad \textbf{(2.186)}$$

Algo imprecisa, la varianza mide la desviación cuadrada "típica" respecto al valor central de una variable aleatoria. Al hacer este cálculo, las desviaciones respecto al valor esperado se elevan al cuadrado para que las desviaciones positivas y negativas respecto al valor esperado contribuyan por igual a esta medida de dispersión. Una vez hecho el cálculo, el proceso de elevación al cuadrado puede revertirse para producir una medida de dispersión que esté en las unidades originales en las que se midió la variable aleatoria. Esta raíz cuadrada de la varianza se llama *desviación estándar* y se denota con $\sigma_x(= \sqrt{\sigma_x^2})$. El nombre de este término transmite eficazmente su significado: σ_x es en efecto la desviación típica ("estándar") de una variable aleatoria respecto a su valor esperado.

Cuando una variable aleatoria está sujeta a la transformación lineal, su varianza y desviación estándar cambiarán en forma muy obvia. Si $y = ax + b$, entonces

$$\sigma_y^2 = \int_{-\infty}^{+\infty} [ax + b - E(ax + b)]^2 f(x)\, dx = \int_{-\infty}^{+\infty} a^2 [x - E(x)]^2 f(x) dx = a^2 \sigma_x^2. \tag{2.187}$$

De ahí que la adición de una constante a una variable aleatoria no cambie su varianza, mientras que la multiplicación por una constante multiplica la varianza por el cuadrado de la constante. Así, resulta claro que multiplicar una variable por una constante multiplica su desviación estándar por esa constante: $\sigma_{ax} = a\sigma_x$.

EJEMPLO 2.16 Varianzas y desviaciones estándar de variables aleatorias simples

En aplicaciones económicas a veces puede ser útil conocer las varianzas y las desviaciones estándar de las cuatro variables aleatorias simples que hemos venido analizando.

1. ***Binomial.*** La varianza de la binomial puede calcularse aplicando la definición en su análoga discreta:

$$\begin{aligned} \sigma_x^2 &= \sum_{i=1}^{n} (x_i - E(x))^2 f(x_i) = (1 - p)^2 \cdot p + (0 - p)^2 (1 - p) \\ &= (1 - p)(p - p^2 + p^2) = p(1 - p). \end{aligned} \tag{2.188}$$

   De ahí que $\sigma_x = \sqrt{p(1 - p)}$. Una implicación de este resultado es que una variable binomial alcanza su más grande varianza y desviación estándar cuando $p = 0.5$, en cuyo caso $\sigma_x^2 = 0.25$ y $\sigma_x = 0.5$. Dada la forma parabólica relativamente plana de $p(1 - p)$, las desviaciones modestas de p respecto a 0.5 no cambian sustancialmente esta varianza.
2. ***Uniforme.*** Calcular la varianza de la distribución uniforme produce un resultado algo interesante:

$$\begin{aligned} \sigma_x^2 &= \int_a^b \left(x - \frac{a+b}{2}\right)^2 \frac{1}{b-a}\, dx = \left(x - \frac{a+b}{2}\right)^3 \cdot \frac{1}{3(b-a)}\Bigg|_a^b \\ &= \frac{1}{3(b-a)} \left[\frac{(b-a)^3}{8} - \frac{(a-b)^3}{8}\right] = \frac{(b-a)^2}{12}. \end{aligned} \tag{2.189}$$

   Este es uno de los pocos casos en los que el número 12 tiene un uso en matemáticas, aparte del de medir cantidades de naranjas o donas.
3. ***Exponencial.*** Integrar la fórmula de la varianza para la exponencial es relativamente laborioso. Por fortuna, el resultado es simple; para la exponencial, resulta que $\sigma_x^2 = 1/\lambda^2$ y $\sigma_x = 1/\lambda$. De ahí que la media y la desviación estándar sean iguales para la distribución exponencial; esta es una "distribución de un parámetro".
4. ***Normal.*** La integración también puede ser difícil en este caso. Pero, de nuevo, el resultado es simple: para la distribución normal, $\sigma_x^2 = \sigma_x = 1$. Las áreas bajo la curva normal pueden calcularse fácilmente, y las tablas de estas pueden consultarse en cualquier texto de estadística. Dos hechos útiles acerca de la FDP normal son:

$$\int_{-1}^{+1} f(x)\, dx \approx 0.68 \quad \text{y} \quad \int_{-2}^{+2} f(x)\, dx \approx 0.95. \tag{2.190}$$

Es decir, existe una probabilidad de aproximadamente dos tercios de que una variable normal esté a ± 1 desviación estándar del valor esperado, aunque "las más de las veces" (es decir, con una probabilidad de 0.95) estará a ± 2 desviaciones estándar.

Estandarización de la normal. Si la variable aleatoria x tiene una FDP normal estándar, tendrá un valor esperado de 0 y una desviación estándar de 1. Sin embargo, una transformación lineal simple puede usarse para darle a esta variable aleatoria el valor esperado (μ) y la desviación estándar (σ) deseados. Considérese la transformación $y = \sigma x + \mu$. Ahora,

$$E(y) = \sigma E(x) + \mu = \mu \quad \text{y} \quad \operatorname{Var}(y) = \sigma_y^2 = \sigma^2 \operatorname{Var}(x) = \sigma^2. \tag{2.191}$$

Revertir este proceso puede servir para "estandarizar" una variable aleatoria normalmente distribuida (y) con un valor esperado (μ) y desviación estándar (σ) arbitrarios (lo que a veces se denota como $y \sim N(\mu, \sigma)$ usando $z = (y - \mu)/\sigma$. Por ejemplo, los resultados del Scholastic Aptitude Test (SAT) (y) están normalmente distribuidos con un valor esperado de 500 puntos y una desviación estándar de 100 puntos (es decir, $y \sim N(500, 100)$). De ahí que $z = (y - 500)/100$ tenga una distribución normal estándar con valor esperado de 0 y desviación estándar de 1. La ecuación 2.190 señala que aproximadamente 68 por ciento de los resultados se ubica entre los 400 y los 600 puntos, mientras que 95 por ciento de ellos se ubica entre los 300 y los 700 puntos.

PREGUNTA: Supongamos que la variable aleatoria x está uniformemente distribuida en el intervalo [0, 12]. ¿Cuáles son la media y desviación estándar de x? ¿Qué fracción de la distribución de x está a ± 1 desviación estándar de la media? ¿Qué fracción de la distribución está a ± 2 desviaciones estándar del valor esperado? Explica por qué esto difiere de las fracciones estimadas para la distribución normal.

Covarianza

Algunos problemas económicos implican dos o más variables aleatorias. Por ejemplo, un inversionista podría considerar repartir su patrimonio entre varios activos, los rendimientos de los cuales se dan como aleatorios. Aunque los conceptos de valor esperado, varianza, etcétera, se aplican en forma más o menos directa al analizar una variable aleatoria en estos casos, también es necesario considerar la relación entre las variables para obtener un cuadro completo. El concepto de covarianza se usa para cuantificar esta relación. Pero antes de ofrecer una definición debemos desarrollar algunas bases.

Consideremos un caso con dos variables aleatorias continuas, x y y. La FDP de estas dos variables, denotada por $f(x, y)$, tiene la propiedad de que la probabilidad asociada con un conjunto de resultados en un área reducida (con dimensiones $dxdy$) está dada por $f(x, y)dxdy$. Para ser una FDP apropiada debe darse el caso de que:

$$f(x, y) \geq 0 \text{ y } \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x, y)\, dx\, dy = 1. \tag{2.192}$$

Las medidas de una sola variable que ya hemos presentado pueden desarrollarse en este contexto de dos variables "desintegrando" la otra variable. Es decir,

$$\begin{aligned} E(x) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x f(x, y)\, dy\, dx \quad \text{y} \\ \operatorname{Var}(x) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} [x - E(x)]^2 f(x, y)\, dy\, dx. \end{aligned} \tag{2.193}$$

De esta manera, los parámetros que describen la variable aleatoria x se miden en todos los resultados posibles de y después de tomar en cuenta la probabilidad de esos diversos resultados.

En este contexto, la *covarianza* entre x y y busca medir la dirección de asociación entre las variables. Específicamente, la covarianza entre x y y [que denota $\mathrm{Cov}(x, y)$] se define como

$$\mathrm{Cov}(x, y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} [x - E(x)][y - E(y)]f(x, y)\,dx\,dy. \quad \textbf{(2.194)}$$

La covarianza entre dos variables aleatorias puede ser positiva, negativa o cero. Si valores de x mayores que $E(x)$ tienden a ocurrir relativamente a menudo con valores de y mayores que $E(y)$ (y, de igual forma, si valores bajos de x tienden a ocurrir junto con valores bajos de y), la covarianza será positiva. En este caso los valores de x y y tienden a moverse en la misma dirección. O bien, si valores altos de x tienden a asociarse con valores bajos de y (y viceversa), la covarianza será negativa.

Dos variables aleatorias se definen como *independientes* si la probabilidad de cualquier valor particular de, digamos, x no se ve afectada por el valor particular que podría ocurrir de y (y viceversa).[25] En términos matemáticos esto significa que la FDP debe tener la propiedad de que $f(x, y) = g(x)h(y)$; es decir, la FDP conjunta puede expresarse como el producto de las FDP de dos variables. Si x y y son independientes, su covarianza será cero:

$$\begin{aligned}\mathrm{Cov}(x, y) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} [x - E(x)][y - E(y)]g(x)\,h(y)\,dx\,dy \\ &= \int_{-\infty}^{+\infty} [x - E(x)]g(x)\,dx \cdot \int_{-\infty}^{+\infty} [y - E(y)]h(y)\,dy = 0 \cdot 0 = 0. \quad \textbf{(2.195)}\end{aligned}$$

Sin embargo, lo opuesto de este enunciado no es necesariamente cierto. Una covarianza de cero no necesariamente implica independencia estadística.

Por último, el concepto de covarianza es crucial para comprender la varianza de sumas o diferencias de variables aleatorias. Aunque el valor esperado de una suma de variables aleatorias es (como cabría suponer) la suma de sus valores esperados:

$$\begin{aligned}E(x + y) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x + y)f(x, y)\,dx\,dy \\ &= \int_{-\infty}^{+\infty} xf(x, y)\,dy\,dx + \int_{-\infty}^{+\infty} yf(x, y)\,dx\,dy = E(x) + E(y), \quad \textbf{(2.196)}\end{aligned}$$

la relación para la varianza de esa suma es más complicada. Usar las definiciones que hemos desarrollado produce

[25] Una definición formal se apoya en el concepto de probabilidad condicional. La probabilidad condicional de un evento B dada la ocurrencia de A (lo cual se escribe $P(B|A)$ se define como $P(B|A) = P(A \text{ y } B) = P(A)$; B y A se definen como independientes si $P(B|A) = P(B)$. En este caso, $P(A \text{ y } B) = P(A) \cdot P(B)$.

$$\begin{aligned}\operatorname{Var}(x+y) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} [x+y-E(x+y)]^2 f(x,y)\,dx\,dy \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} [x-E(x)+y-E(y)]^2 f(x,y)\,dx\,dy \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} [x-E(x)]^2+[y-E(y)]^2+2[x-E(x)][y-E(y)]f(x,y)\,dx\,dy \\ &= \operatorname{Var}(x)+\operatorname{Var}(y)+2\operatorname{Cov}(x,y). \end{aligned} \tag{2.197}$$

De ahí que si x y y son independientes, entonces $\operatorname{Var}(x+y)=\operatorname{Var}(x)+\operatorname{Var}(y)$. La varianza de la suma será mayor que la suma de las varianzas si las dos variables aleatorias tienen una covarianza positiva, y será menor que la suma de las varianzas si tienen una covarianza negativa. Los problemas 2.14-2.16 dan más detalles sobre algunos de los resultados estadísticos que se usan en la teoría microeconómica.

Resumen

Pese a la formidable apariencia de algunas secciones de este capítulo, este no es un libro de matemáticas. Más bien, la intención aquí ha sido la de reunir varias herramientas que servirán para desarrollar modelos económicos a lo largo de este texto. El material en este capítulo será útil entonces como referencia práctica.

Una manera de sintetizar las herramientas matemáticas que se han presentado en este capítulo es subrayar de nueva cuenta las lecciones económicas que dichas herramientas ilustran:

- El uso de las matemáticas ofrece a los economistas un medio conveniente y abreviado para desarrollar modelos. Gracias al uso de estas herramientas matemáticas se pueden estudiar en un marco simplificado las implicaciones de varios supuestos económicos.

- El concepto matemático de las derivadas de una función es de amplio uso en los modelos económicos porque a los economistas suele interesarles la forma en que las variaciones marginales en una variable afectan a otra. Las derivadas parciales son especialmente útiles para este propósito porque están definidas para representar dichas variaciones marginales cuando todos los demás factores se mantienen constantes.

- Las matemáticas de la optimización son una herramienta importante para el desarrollo de modelos que asumen que los agentes económicos persiguen racionalmente una meta. En el caso irrestricto las condiciones de primer orden establecen que cualquier actividad que contribuya a la meta del agente debe expandirse hasta el punto en que la contribución marginal de una expansión adicional sea de cero. En términos matemáticos, la condición de primer orden para un óptimo requiere que todas las derivadas parciales sean iguales a cero.

- La mayoría de los problemas de optimización económica implican restricciones a las decisiones que los agentes pueden tomar. En este caso las condiciones de primer orden para un máximo sugieren que cada actividad es operada en un nivel en el cual la razón del beneficio marginal de la actividad con su costo marginal es la misma para todas las actividades utilizadas. Esta razón común de costo marginal-beneficio marginal es igual también al multiplicador de Lagrange, el cual suele introducirse para ayudar a resolver problemas de optimización restringida. El multiplicador de Lagrange puede interpretarse asimismo como el valor implícito (o precio sombra) de la restricción.

- El teorema de la función implícita es un recurso matemático útil para ilustrar la dependencia de las opciones que resultan de un problema de optimización respecto a los parámetros de ese problema (precios de mercado, por ejemplo). El teorema de la envolvente es útil para examinar cómo cambian esas opciones óptimas cuando los parámetros del problema (precios) cambian.

- Algunos problemas de optimización pueden implicar restricciones que son desigualdades más que igualdades. Las soluciones de estos problemas ilustran a menudo la "lasitud complementaria". Es decir, las restricciones se mantienen con la igualdad y sus multiplicadores de Lagrange asociados son diferentes de cero, o las restricciones son desigualdades estrictas y sus multiplicadores de Lagrange asociados son iguales a cero. De nuevo esto ilustra cómo el multiplicador de Lagrange implica algo sobre la "importancia" de las restricciones.

- Las condiciones de primer orden expuestas en este capítulo son sólo las condiciones necesarias para un máximo o mínimo local. También deben comprobarse las condiciones de segundo orden, que requieren la satisfacción de ciertas condiciones de curvatura.

- Ciertos tipos de funciones ocurren en muchos problemas económicos. Las funciones cuasi cóncavas (aquellas para las cuales las curvas de nivel forman conjuntos convexos) cumplen las condiciones de segundo orden de los problemas de máximo o mínimo restringido cuando las restricciones son lineales. Las funciones homotéticas poseen la útil propiedad de que las disyuntivas implícitas entre las variables de la función sólo dependen de las razones de dichas variables.
- El cálculo integral suele usarse en economía lo mismo como una manera de describir áreas por debajo de las gráficas que como una forma de sumar resultados en el tiempo. Técnicas que implican varios modos de diferenciar integrales desempeñan un papel importante en la teoría del comportamiento de optimización.
- Muchos problemas económicos son dinámicos en cuanto que las decisiones en una fecha afectan a decisiones y resultados en fechas posteriores. Las matemáticas para resolver estos problemas de optimización dinámica suelen ser una generalización simple de los métodos lagrangianos.
- En el estudio de la economía de la incertidumbre y la información suelen usarse conceptos de estadística matemática. El concepto fundamental es la noción de variable aleatoria y su FDP asociada. Parámetros de esta distribución, como su valor esperado o su varianza, también desempeñan papeles importantes en muchos modelos económicos.

Problemas

2.1

Supón que $U(x, y) = 4x^2 + 3y^2$.

a. Calcula $\partial U/\partial x$, $\partial U/\partial y$.
b. Evalúa estas derivadas parciales en $x = 1, y = 2$.
c. Escribe la diferencial total para U.
d. Calcula dy/dx para $dU = 0$; es decir, ¿cuál es la la opción contenida entre x y y manteniendo constante U?
e. Demuestra que $U = 16$ cuando $x = 1, y = 2$.
f. ¿En qué proporción deben cambiar x y y para mantener a U constante en 16 para movimientos respecto a $x = 1, y = 2$?
g. De modo más general, ¿cuál es la forma de la curva de nivel $U = 16$ para esta función? ¿Cuál es la pendiente de esa línea?

2.2

Supón que los ingresos totales de una empresa dependen de la cantidad producida (q) de acuerdo con la función

$$R = 70q - q^2.$$

Los costos totales también dependen de q:

$$C = q^2 + 30q + 100.$$

a. ¿Qué nivel de producción debe generar la empresa para maximizar sus beneficios ($R - C$)? ¿Cuáles serán esos beneficios?
b. Demuestra que las condiciones de segundo orden para un máximo se satisfacen en el nivel de producción determinado en el inciso a).
c. ¿La solución calculada aquí cumple la regla de "ingreso marginal es igual a costo marginal"? Explica.

2.3

Supón que $f(x, y) = xy$. Halla el valor máximo de f si x y y están restringidas a sumar 1. Resuelve este problema de dos maneras: por sustitución y con el método del multiplicador de Lagrange.

2.4

El problema dual descrito en el problema 2.3 es

$$\begin{array}{ll} \text{minimiza} & x + y \\ \text{sujeta a} & xy = 0.25. \end{array}$$

Resuelve este problema usando la técnica de Lagrange. Luego compara el valor que obtuviste para el multiplicador de Lagrange con aquel que obtuviste en el problema 2.3. Explica la relación entre ambas soluciones.

2.5

La altura desde la cual es lanzada una pelota en línea recta con cierta fuerza es una función del momento (t) a partir del cual se le suelta, dado por $f(t) = 0.5gt^2 + 40t$ (donde g es una constante determinada por la gravedad).

a. ¿Cómo depende del parámetro g el valor de t en el que la altura desde la cual se lanza la pelota está en un máximo?
b. Usa tu respuesta del inciso a) para describir cómo la altura máxima cambia al cambiar el parámetro g.
c. Usa el teorema de la envolvente para contestar directamente el inciso b).
d. En la Tierra $g = 32$, aunque este valor varía un poco alrededor del globo. Si dos lugares tuvieran constantes gravitacionales que difirieran en 0.1, ¿cuál sería la diferencia en la altura máxima desde la que se lanza una pelota en ambos lugares?

2.6

Una forma simple de modelar la construcción de un buque petrolero es partir de una hoja de acero grande de forma rectangular de x pies de ancho y $3x$ pies de largo. Corta ahora un pequeño cuadrado de t pies por lado de cada esquina de la hoja; dobla y suelda los lados de la hoja de acero para producir una estructura semejante a una bandeja sin tapa.

a. Demuestra que el volumen de petróleo que esta bandeja puede contener está dado por

$$V = t(x - 2t)(3x - 2t) = 3tx^2 - 8t^2x + 4t^3.$$

b. ¿Cómo debería elegirse t para maximizar V para cualquier valor dado de x?
c. ¿Existe un valor de x que maximice el volumen de petróleo que se puede transportar?
d. Supón que un constructor naval está obligado a usar únicamente 1 000 000 de pies cuadrados de hoja de acero para fabricar el buque petrolero. Esta restricción puede representarse con la ecuación $3x^2 - 4t^2 = 1\,000\,000$ (porque el constructor puede intercambiar los cuadrados recortados por crédito). ¿Qué resultaría de comparar la solución de este problema de máximo restringido con las soluciones descritas en los incisos b) y c)?

2.7

Considera el siguiente problema de maximización restringida:

$$\begin{aligned} &\text{maximiza} \quad && y = x_1 + 5\ln x_2 \\ &\text{sujeta a} \quad && k - x_1 - x_2 = 0, \end{aligned}$$

donde k es una constante a la que se le puede asignar cualquier valor específico.

a. Demuestra que si $k = 10$, este problema puede resolverse como uno que sólo implica restricciones de igualdad.
b. Demuestra que resolver este problema para $k = 4$ requiere que $x_1 = -1$.
c. Si las x en este problema deben ser no negativas, ¿cuál es la solución óptima cuando $k = 4$? (Este problema puede resolverse intuitivamente o usando los métodos que hemos descrito en este capítulo.)
d. ¿Cuál es la solución de este problema cuando $k = 20$? ¿Qué concluyes al comparar esta solución con la del inciso a)?

Nota: Este problema implica lo que se conoce como *función cuasi lineal*. Este tipo de funciones da ejemplos importantes de ciertos tipos de comportamiento en la teoría del consumo tal como veremos.

2.8

Supón que una empresa tiene una función de costo marginal dada por $MC(q) = q + 1$.

a. ¿Cuál es la función de costo total de esta empresa? Explica por qué los costos totales sólo se conocen hasta una constante de integración, la cual representa costos fijos.
b. Como has de saber, gracias a un curso previo de economía, si una empresa toma el precio (p) como dado en sus decisiones, entonces generará la producción para la cual $p = MC(q)$. Si la empresa sigue esta regla de maximización de beneficios, ¿cuánto producirá cuando $p = 15$? Suponiendo que en ese precio alcanza el punto de equilibrio, ¿cuáles serán los costos fijos?
c. ¿Cuánto aumentarán los beneficios de esta empresa si el precio aumenta a 20?
d. Suponiendo aún la maximización de beneficios, demuestra que las de esta empresa sólo pueden expresarse como una función del precio que recibe por su producto.
e. Demuestra que el aumento en beneficios de $p = 15$ a $p = 20$ puede calcularse de dos maneras: i) directamente a partir de la ecuación derivada en el inciso d) y ii) integrando la función inversa de costo marginal $[MC^{-1}(p) = p - 1]$ de $p = 15$ a $p = 20$. Explica intuitivamente este resultado usando el teorema de la envolvente.

Problemas analíticos

2.9 Funciones cóncavas y cuasi cóncavas

Demuestra que, si $f(x_1, x_2)$ es una función cóncava, también es una función cuasi cóncava. Hazlo comparando la ecuación 2.114 (definición de cuasi concavidad) con la ecuación 2.98 (definición de concavidad). ¿Puedes dar una razón intuitiva de este resultado? ¿Lo opuesto del enunciado es cierto? ¿Las funciones cuasi cóncavas son necesariamente cóncavas? De no ser así, da un contraejemplo.

2.10 Función de Cobb-Douglas

Una de las funciones más importantes que encontraremos en este libro es la función de Cobb-Douglas:

$$y = (x_1)^{\alpha}(x_2)^{\beta},$$

donde α y β son constantes positivas, cada cual menor que 1.

a. Demuestra que esta función es cuasi cóncava, usando el método de "fuerza bruta" de aplicar la ecuación 2.114.
b. Demuestra que la función de Cobb-Douglas es cuasi cóncava mostrando que cada curva de nivel de la forma $y = c$ (donde c es cualquier constante positiva) es convexa y, por tanto, que el conjunto de puntos para los cuales $y > c$ es un conjunto convexo.
c. Demuestra que si $\alpha + \beta > 1$, la función de Cobb-Douglas no es cóncava (ilustrando otra vez, por tanto, que no todas las funciones cuasi cóncavas son cóncavas).

Nota: En las extensiones de este capítulo se trata nuevamente la función de Cobb-Douglas.

2.11 La función de potencia

Otra función que encontraremos a menudo en este libro es la *función de potencia*:

$$y = x^{\delta},$$

donde $0 \leq \delta \leq 1$ (a veces también examinaremos esta función para casos en los que δ también puede ser negativa, en cuyo caso usaremos la forma $y = x^{\delta}/\delta$ para garantizar que las derivadas tengan el signo adecuado).

a. Demuestra que esta función es cóncava (y, por tanto, también cuasi cóncava, por el resultado del problema 2.9). Nótese que $\delta = 1$ es un caso especial y que la función es "estrictamente" cóncava sólo para $\delta < 1$.
b. Demuestra que la forma multivariada de la función de potencia

$$y = f(x_1, x_2) = (x_1)^{\delta} + (x_2)^{\delta}$$

también es cóncava (y cuasi cóncava). Explica por qué, en este caso, el hecho de que $f_{12} = f_{21} = 0$ vuelve especialmente simple determinar la concavidad.
c. Una manera de incorporar efectos de "escala" en la función descrita en el inciso b) es usar la transformación monótona

$$g(x_1, x_2) = y^{\gamma} = [(x_1)^{\delta} + (x_2)^{\delta}]^{\gamma},$$

donde γ es una constante positiva. ¿Esta transformación preserva la concavidad de la función? ¿Es g cuasi cóncava?

2.12 Prueba del teorema de la envolvente en problemas de optimización restringida

Puesto que en este libro se usará con frecuencia el teorema de la envolvente en los problemas de optimización restringida, probar este teorema en un caso simple ayudará a desarrollar cierta intuición. Así, supongamos que se desea maximizar una función de dos variables y que el valor de esta función también depende de un parámetro, a: $f(x_1, x_2, a)$. Este problema de maximización está sujeto a una restricción que puede escribirse como $g(x_1, x_2, a) = 0$.

a. Escribe la expresión lagrangiana y las condiciones de primer orden de este problema.
b. Suma las dos condiciones de primer orden que implican las x.
c. Diferencia ahora la suma anterior respecto a a; esto muestra cómo deben cambiar las x al cambiar a, al mismo tiempo que se requiere que las condiciones de primer orden sigan siendo válidas.

d. Como ya se indicó en este capítulo tanto la función objetivo como la restricción de este problema pueden enunciarse como funciones de a: $f(x_1(a), x_2(a), a)$, $g(x_1(a), x_2(a), a) = 0$. Diferencia la primera de estas respecto a a. Esto muestra cómo cambia el valor del objetivo al cambiar a mientras las x se mantienen en sus valores óptimos. Deberás tener términos que impliquen las x y un término en $\partial f/\partial a$.
e. Diferencia ahora la restricción formulada en el inciso d) respecto a a. Deberás tener términos en las x y un término en $\partial g/\partial a$.
f. Multiplica los resultados del inciso e) por λ (el multiplicador de Lagrange), y usa esto junto con las condiciones de primer orden del inciso c) para sustituir en la derivada del inciso d). Deberás poder demostrar que

$$\frac{df(x_1(a), x_2(a), a)}{da} = \frac{\partial f}{\partial a} + \lambda \frac{\partial g}{\partial a},$$

justo la derivada parcial de la expresión lagrangiana cuando todas las x están en sus valores óptimos. Esto prueba el teorema de la envolvente. Explica intuitivamente cómo las diversas partes de esta prueba imponen la condición de que las x se ajusten de manera constante para permanecer en sus valores óptimos.
g. Vuelve al ejemplo 2.8 y explica cómo se puede aplicar el teorema de la envolvente a las variaciones en el perímetro de una cerca P; es decir, ¿cómo afectan las variaciones en P a las dimensiones del área que se puede cercar? Demuestra que en este caso el teorema de la envolvente ilustra cómo el multiplicador de Lagrange asigna un valor a la restricción.

2.13 Aproximaciones de Taylor

El teorema de Taylor demuestra que cualquier función puede acercarse a cualquier punto conveniente mediante una serie de términos que involucran a la función y sus derivadas. Aquí examinaremos algunas aplicaciones de este teorema a funciones de una y dos variables.

a. Cualquier función continua y diferenciable de una variable, $f(x)$, puede acercarse al punto a mediante la fórmula

$$f(x) = f(a) + f'(a)(x - a) + 0.5f''(a)(x - a)^2 + \text{términos en } f''', f'''', \ldots.$$

El uso de únicamente los tres primeros términos resulta en una aproximación *cuadrática* de Taylor. Emplea esta aproximación junto con la definición de concavidad dada en la ecuación 2.85 para demostrar que toda función cóncava debe estar en o por debajo de la tangente de la función en el punto a.
b. La aproximación cuadrática de Taylor para cualquier función de dos variables, $f(x, y)$, cerca del punto (a, b) está dada por

$$\begin{aligned} f(x, y) = {} & f(a, b) + f_1(a, b)(x - a) + f_2(a, b)(y - b) \\ & + 0.5[f_{11}(a, b)(x - a)^2 + 2f_{12}(a, b)(x - a)(y - b) + f_{22}(y - b)^2]. \end{aligned}$$

Usa esta aproximación para demostrar que toda función cóncava (definida por la ecuación 2.98) debe estar en o por debajo de su plano tangente en (a, b).

2.14 Más sobre valor esperado

Puesto que el concepto de valor esperado desempeña un papel importante en muchas teorías económicas, quizá sea útil resumir algunas propiedades más de esta medida estadística. En este problema se supone que x es una variable aleatoria continua con la FDP $f(x)$.

a. (Desigualdad de Jensen) Supón que $g(x)$ es una función cóncava. Demuestra que $E[g(x)] \leq g[E(x)]$. *Pista*: Construye la tangente de $g(x)$ en el punto $E(x)$. Esta tangente tendrá la forma $c + dx \geq g(x)$ para todos los valores de x y $c + dE(x) = g[E(x)]$ donde c y d son constantes.
b. Usa el procedimiento del inciso a) para demostrar que si $g(x)$ es una función convexa, entonces $E[g(x)] \geq g[E(x)]$.
c. Supón que x adopta sólo valores no negativos; es decir, $0 \leq x \leq \infty$. Usa la integración por partes para demostrar que

$$E(x) = \int_0^\infty [1 - F(x)]\, dx,$$

donde $F(x)$ es la función de distribución acumulativa para x [es decir, $F(x) = \int_0^x f(t)dt$].
d. (Desigualdad de Markov) Demuestra que si x sólo adopta valores positivos, se mantiene la desigualdad siguiente:

$$P(x \geq t) \leq \frac{E(x)}{t}.$$

Pista: $E(x) = \int_0^\infty xf(x)\, dx = \int_0^t xf(x)\, dx + \int_t^\infty xf(x)\, dx$.

e. Considera la FDP $f(x) = 2x^{-3}$ para $x \geq 1$.
   1. Demuestra que esta es una FDP apropiada.
   2. Calcula la $F(x)$ de esta FDP.
   3. Usa los resultados del inciso c) para calcular la $E(x)$ de esta FDP.
   4. Demuestra que la desigualdad de Markov es válida para esta función.

f. El concepto de valor esperado condicional es útil en algunos problemas económicos. Denotamos el valor esperado de x condicionado la ocurrencia de un evento, A, como $E(x|A)$. Para calcular este valor debemos conocer la FDP de x dada la ocurrencia de A [denotada por $f(x|A)$]. Con esta notación, $E(x|A) = \int_{-\infty}^{+\infty} xf(x|A)dx$. Tal vez la manera más fácil de comprender estas relaciones sea mediante un ejemplo. Sea

$$f(x) = \frac{x^2}{3} \quad \text{para} \quad -1 \leq x \leq 2.$$

   1. Demuestra que esta es una FDP apropiada.
   2. Calcula $E(x)$.
   3. Calcula la probabilidad de que $-1 \leq x \leq 0$.
   4. Considera el evento $0 \leq x \leq 2$, y llama a este evento A. ¿Qué es $f(x|A)$?
   5. Calcula $E(x|A)$.
   6. Explica intuitivamente tus resultados.

2.15 Más sobre varianzas

La definición de la varianza de una variable aleatoria puede usarse para demostrar varios resultados adicionales.

a. Demuestra que $\text{Var}(x) = E(x^2) - [E(x)]^2$.

b. Usa la desigualdad de Markov (problema 2.14d) para demostrar que si x sólo puede adoptar valores no negativos,

$$P\,[(x - \mu_x) \geq k] \leq \frac{\sigma_x^2}{k^2}.$$

Este resultado muestra que existen límites a qué tan a menudo una variable aleatoria puede alejarse de su valor esperado. Si $k=h\sigma$, este resultado también indica que

$$P\,[(x - \mu_x) \geq h\sigma] \leq \frac{1}{h^2}.$$

Así, por ejemplo, la probabilidad de que una variable aleatoria esté a más de dos desviaciones estándar de su valor esperado siempre es menor que 0.25. Este resultado teórico se conoce como *desigualdad de Chebyshev*.

c. La ecuación 2.197 demostró que si dos (o más) variables aleatorias son independientes, la varianza de su suma es igual a la suma de sus varianzas. Usa este resultado para demostrar que la suma de n variables aleatorias independientes, cada una de las cuales tiene valor esperado μ y varianza σ^2, posee valor esperado $n\mu$ y varianza $n\sigma^2$. Demuestra también que el promedio de estas n variables aleatorias (que también es una variable aleatoria) tendrá valor esperado μ y varianza σ^2/n. Esto se conoce como *la ley de los números grandes*; es decir, la varianza de un promedio se contrae conforme se incluyen más variables independientes.

d. Usa el resultado del inciso c) para demostrar que si x_1 y x_2 son variables aleatorias independientes, cada cual con el mismo valor esperado y varianza, la varianza de un promedio ponderado de ambas $X = kx_1 + (1 - k)x_2$, $0 \leq k \leq 1$ se minimiza cuando $k = 0.5$. ¿Cuánto se reduce la varianza de esta suma estableciendo apropiadamente k en relación con otros valores posibles de k?

e. ¿Cómo cambiaría el resultado del inciso d) si las dos variables tuvieran varianzas desiguales?

2.16 Más sobre covarianzas

He aquí algunas relaciones útiles asociadas con la covarianza de dos variables aleatorias, x_1 y x_2.

a. Demuestra que $\text{Cov}(x_1, x_2) = E(x_1x_2) - E(x_1)E(x_2)$. Una implicación importante de esto es que si $\text{Cov}(x_1, x_2) = 0$, $E(x_1x_2) = E(x_1)E(x_2)$. Es decir, el valor esperado de un producto de dos variables aleatorias es el producto de los valores esperados de estas variables.

b. Demuestra que $\text{Var}(ax_1 + bx_2) = a^2\text{Var}(x_1) + b^2\text{Var}(x_2) + 2ab\text{Cov}(x_1, x_2)$.

c. En el problema 2.15d se analizó la varianza de $X = kx_1 + (1 - k)x_2$ $0 \leq k \leq 1$. ¿Cambia la conclusión de que esta varianza se minimiza para $k = 0.5$ al considerar casos en los que $\text{Cov}(x_1, x_2) \neq 0$?