

Glossary/Glosario

English	Español
<p>1.5 × IQR rule for outliers An observation is called an outlier if it falls more than $1.5 \times IQR$ above the third quartile or below the first quartile. (p. 56)</p>	<p>regla 1.5 × la gama entre cuartiles para valores atípicos Se le dice valor atípico a una observación si cae a más de $1.5 \times$ la gama entre cuartiles por encima del tercer cuartil o por debajo del primer cuartil. (pág. 56)</p>
<p>10% condition When taking an SRS of size n from a population of size N, check that $n \leq \frac{1}{10}N$. (pp. 401, 494)</p>	<p>condición del 10% Cuando se toma una muestra aleatoria sencilla de tamaño n de una población de tamaño N, se verifica que $n \leq \frac{1}{10}N$. (págs. 401, 494)</p>
<p>68–95–99.7 rule (also known as the empirical rule) In the Normal distribution with mean μ and standard deviation σ, (a) approximately 68% of the observations fall within σ of the mean μ, (b) approximately 95% of the observations fall within 2σ of μ, and (c) approximately 99.7% of the observations fall within 3σ of μ. (p. 110)</p>	<p>regla 68–95–99.7 A la que también se le dice la “regla empírica”. En la distribución normal con media μ y desviación estándar σ, (a) aproximadamente el 68% de las observaciones caen dentro de σ de la media μ, (b) aproximadamente el 95% de las observaciones caen dentro de 2σ de μ, y (c) aproximadamente el 99.7% de las observaciones caen dentro de 3σ de μ. (pág. 110)</p>
A	
<p>addition rule for mutually exclusive events If A and B are mutually exclusive events, $P(A \text{ or } B) = P(A) + P(B)$. (p. 308)</p>	<p>regla de suma para eventos que se excluyen mutuamente Si A y B son eventos que se excluyen entre sí, $P(A \text{ o } B) = P(A) + P(B)$. (pág. 308)</p>
<p>alternative hypothesis H_a The claim that we are trying to find evidence <i>for</i> in a significance test. (p. 540)</p>	<p>hipótesis H_a alternativa La proposición de que en una prueba de significancia estadística estamos tratando de hallar evidencia que esté <i>a favor</i>. (pág. 540)</p>
<p>anonymity The names of individuals participating in a study are not known even to the director of the study. (p. 271)</p>	<p>anonimato Cuando se desconocen los nombres de las personas que participan en un estudio; inclusive el director del estudio los ignora. (pág. 271)</p>
<p>association Knowing the value of one variable helps predict the value of the other. If knowing the value of one variable does not help predict the value of the other, there is no association between the variables. (p. 18)</p>	<p>asociación Saber el valor de una variable facilita la predicción del valor de la otra. Si saber el valor de una variable no facilita la predicción del valor de la otra, entonces no existe ninguna asociación entre las variables. (pág. 18)</p>
B	
<p>back-to-back stemplot (also called back-to-back stem-and-leaf plot) Plot used to compare the distribution of a quantitative variable for two groups. Each observation in both groups is separated into a stem, consisting of all but the final digit, and a leaf, the final digit. The stems are arranged in a vertical column with the smallest at the top. The values from one group are plotted on the left side of the stem and the values from the other group are plotted on the right side of the stem. Each leaf is written in the row next to its stem, with the leaves arranged in increasing order out from the stem. (p. 32)</p>	<p>diagrama de tallos contiguos (también se le dice diagrama de tallos y hojas contiguos) Se utiliza para comparar la distribución de una variable cuantitativa en dos grupos. Cada observación efectuada en ambos grupos se separa en un tallo, que consiste de todos los dígitos salvo el último, y una hoja, que consta del último dígito. Los tallos se organizan en una columna vertical con las cifras más pequeñas arriba. Los valores de un grupo se diagraman al lado izquierdo del tallo y los valores del otro grupo se diagraman al lado derecho del tallo. Cada hoja se coloca en el renglón que está al lado de su tallo, y las hojas dispuestas en orden ascendente extendiéndose hacia fuera a partir del tallo. (pág. 32)</p>

G-2 Glossary/Glosario

<p>bar graph Graph used to display the distribution of a categorical variable or to compare the sizes of different quantities. The horizontal axis of a bar graph identifies the categories or quantities being compared. The graph is drawn with blank spaces between the bars to separate the items being compared. (p. 8)</p>	<p>gráfico de barras (pág. 8) Se usa para ilustrar la distribución de una variable categorizada o para comparar el tamaño de diferentes cantidades. El eje horizontal del gráfico de barras identifica las categorías o las cantidades que se han de comparar. Se puede dibujar con espacios en blanco entre las barras a fin de separar las diversas categorías que se desea comparar. (pág. 8)</p>
<p>bias The design of a statistical study shows bias if it would consistently underestimate or consistently overestimate the value you want to know. (p. 212)</p>	<p>sesgo Al diseñar un estudio estadístico se demuestra un sesgo si de manera constante se subestima o sobrestima el valor que se desea saber. (pág. 212)</p>
<p>biased estimator A statistic used to estimate a parameter is biased if the mean of its sampling distribution is not equal to the true value of the parameter being estimated. (p. 431)</p>	<p>calculador sesgado La estadística que se usa para computar un parámetro está sesgada si la media de la distribución de su muestreo no equivale al valor real del parámetro que se está computando. (pág. 431)</p>
<p>bimodal A graph of quantitative data with two clear peaks. (p. 29)</p>	<p>bimodal Gráfico de datos cuantitativos con dos picos bien definidos. (pág. 29)</p>
<p>binomial coefficient The number of ways of arranging k successes among n observations is given by the binomial coefficient $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ for $k = 0, 1, 2, \dots, n$ where $n! = n(n-1)(n-2) \cdot \dots \cdot 3 \cdot 2 \cdot 1$ and $0! = 1$. (p. 392)</p>	<p>coeficiente binomial La cantidad de maneras de organizar k aciertos entre n observaciones se representa con el coeficiente binomial $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ para $k = 0, 1, 2, \dots, n$ en el que $n! = n(n-1)(n-2) \cdot \dots \cdot 3 \cdot 2 \cdot 1$ y $0! = 1$. (pág. 392)</p>
<p>binomial distribution In a binomial setting, suppose we let X = the number of successes. The probability distribution of X is a binomial distribution with parameters n and p, where n is the number of trials of the chance process and p is the probability of a success on any one trial. The possible values of X are the whole numbers from 0 to n. (p. 388)</p>	<p>distribución binomial En un entorno binomial, supongamos que se permite que X = la cantidad de aciertos. La distribución de la probabilidad de X es una distribución binomial con los parámetros n y p, en la que n es la cantidad de ensayos del proceso de probabilidad y p es la probabilidad de un acierto en cualquiera de los ensayos. Los posibles valores de X son los números enteros de 0 a n. (pág. 388)</p>
<p>binomial probability formula If X has the binomial distribution with n trials and probability p of success on each trial, the possible values of X are 0, 1, 2, ..., n. If k is any one of these values, $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. (p. 409)</p>	<p>fórmula de probabilidad binomial Si X tiene la distribución binomial con n ensayos y la probabilidad p de acierto en cada ensayo, los posibles valores de X son 0, 1, 2, ..., n. Si k es cualquiera de estos valores, $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ (pág. 409)</p>
<p>binomial random variable The count X of successes in a binomial setting. (p. 388)</p>	<p>variable aleatoria binomial La cuenta X de aciertos en un entorno binomial. (pág. 388)</p>
<p>binomial setting Arises when we perform several independent trials of the same chance process and record the number of times that a particular outcome occurs. The four conditions for a binomial setting:</p> <ul style="list-style-type: none"> • Binary? The possible outcomes of each trial can be classified as “success” or “failure.” • Independent? Trials must be independent; that is, knowing the result of one trial must not tell us anything about the result of any other trial. • Number? The number of trials n of the chance process must be fixed in advance. • Success? There is the same probability p of success on each trial. (p. 388) 	<p>entorno binomial Surge cuando se realizan varios ensayos independientes del mismo proceso de probabilidad y se anota la cantidad de veces que se produce un resultado dado. Las cuatro condiciones que definen un entorno binomial son:</p> <ul style="list-style-type: none"> • Binario? Los resultados posibles de cada ensayo se pueden clasificar como “acierto” o “fracaso”. • Independiente? Los ensayos han de ser independientes; es decir, saber el resultado de un ensayo no debe indicar nada acerca del resultado de otro ensayo. • Número? La cantidad de ensayos n del proceso de probabilidad se tiene que fijar con anticipación. • Acierto? Existe la misma probabilidad p de lograr un acierto en cada ensayo. (pág. 388)
<p>block Group of experimental units that are known before the experiment to be similar in some way that is expected to affect the response to the treatments. (p. 252)</p>	<p>bloque Grupo de unidades experimentales que antes del experimento se sabe son similares de alguna manera previsible que afecta la respuesta a los tratamientos. (pág. 252)</p>

<p>boxplot Graph of the five-number summary. The box spans the quartiles and shows the spread of the central half of the distribution. The median is marked within the box. Lines extend from the box to the smallest and largest observations that are not outliers. Outliers are marked with a special symbol such as an asterisk (*). (p. 57)</p>	<p>diagrama de caja y bigotes Un gráfico del resumen de cinco cifras. La caja abarca los cuartiles y muestra el alcance de la mitad central de la distribución. Dentro de la caja se marca la media. Las líneas se extienden a partir de la caja a las observaciones más pequeña y más grande que no son valores atípicos. Los valores atípicos se marcan con un símbolo especial tal como un asterisco (*). (pág. 57)</p>
<p>C</p>	
<p>categorical variable Variable that places an individual into one of several groups or categories. (p. 3)</p>	<p>variable categorizada Coloca a un individuo en uno o varios grupos o categorías. (pág. 3)</p>
<p>census Study that attempts to collect data from every individual in the population. (p. 210)</p>	<p>censo Un estudio en el que se trata de recoger datos acerca de cada individuo en la población. (pág. 210)</p>
<p>central limit theorem (CLT) In an SRS of size n from any population with mean μ and finite standard deviation σ, when n is large, the sampling distribution of the sample mean \bar{x} is approximately Normal. (p. 457)</p>	<p>teorema del límite central Traza una muestra aleatoria sencilla de tamaño n a partir de una población con la media μ y una desviación estándar finita de σ. El teorema del límite central manifiesta que cuando n es grande, la distribución de muestreo de la media de la muestra \bar{x} es aproximadamente normal. (pág. 457)</p>
<p>Chebyshev's inequality In any distribution, the proportion of observations falling within k standard deviations of the mean is at least $1 - \frac{1}{k^2}$. (p. 112)</p>	<p>desigualdad de Chebychov En cualquier distribución, la proporción de observaciones que yacen dentro de k desviaciones estándar de la media es al menos $1 - \frac{1}{k^2}$. (pág. 112)</p>
<p>chi-square distribution Family of distributions that take only nonnegative values and are skewed to the right. A particular chi-square distribution is specified by giving its degrees of freedom. (p. 685)</p>	<p>distribución de ji cuadrado Familia de distribuciones que acepta solo valores no negativos y que está sesgada hacia la derecha. Se especifica una distribución de ji cuadrado dada citando sus grados de libertad. (pág. 685)</p>
<p>chi-square statistic Measure of how far the observed counts are from the expected counts. The formula is</p> $\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$ <p>where the sum is over all possible values of the categorical variable or all cells in the two-way table. (p. 682)</p>	<p>estadística de ji cuadrado Una medición de la distancia entre las cuentas observadas y las cuentas previstas. La fórmula es</p> $\chi^2 = \sum \frac{(\text{Observadas} - \text{Previstas})^2}{\text{Previstas}}$ <p>en la que la suma está sobre todos los valores posibles de la variable categorizada o sobre todas las celdas en la tabla de doble vía. (pág. 682)</p>
<p>chi-square test for goodness of fit Suppose the Random, 10%, and Large Counts conditions are met. To determine whether a categorical variable has a specified distribution in the population of interest, expressed as the proportion of individuals falling into each possible category, perform a test of</p> <p>H_0: The specified distribution of the categorical variable in the population of interest is correct.</p> <p>H_a: The specified distribution of the categorical variable in the population of interest is not correct.</p> <p>Start by finding the expected count for each category assuming that H_0 is true. Then calculate the chi-square statistic</p> $\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$	<p>prueba de ji cuadrado para confirmar el cuadro Supongamos que se cumplen las condiciones de aleatorio, del 10% y de cuentas grandes. Para determinar si una variable categorizada tiene una distribución específica en la población de interés, expresada como la proporción de individuos que se encuentran dentro de cada categoría posible, se realiza una prueba de</p> <p>H_0: La distribución especificada de la variable categorizada en la población de interés es correcta</p> <p>H_a: La distribución especificada de la variable categorizada en la población de interés no es correcta.</p> <p>Se comienza hallando la cuenta prevista para cada categoría, asumiendo que H_0 es verdad. Luego, se calcula la estadística de ji cuadrado</p> $\chi^2 = \sum \frac{(\text{Observadas} - \text{Previstas})^2}{\text{Previstas}}$

G-4 Glossary/Glosario

<p>where the sum is over the k different categories. The P-value is the area to the right of χ^2 under the density curve of the chi-square distribution with $k - 1$ degrees of freedom. (p. 680)</p>	<p>en la que la suma está sobre las k categorías diferentes. El valor P es el área a la derecha de χ^2 bajo la curva de densidad de la distribución ji cuadrado con $k - 1$ grados de libertad. (pág. 680)</p>
<p>chi-square test for homogeneity Suppose the Random, 10%, and Large Counts conditions are met. You can use the chi-square test for homogeneity to test</p> <p>H_0: There is no difference in the distribution of a categorical variable for several populations or treatments.</p> <p>H_a: There is a difference in the distribution of a categorical variable for several populations or treatments.</p> <p>Start by finding the expected counts. Then calculate the chi-square statistic</p> $\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$ <p>where the sum is over all cells (not including totals) in the two-way table. If H_0 is true, the χ^2 statistic has approximately a chi-square distribution with degrees of freedom = (number of rows - 1)(number of columns - 1). The P-value is the area to the right of χ^2 under the corresponding chi-square density curve. (p. 708)</p>	<p>prueba de ji cuadrado de homogeneidad Supongamos que se han cumplido las condiciones de aleatorio, del 10% y de cuentas grandes. Se puede usar la prueba de ji cuadrado de homogeneidad para verificar que</p> <p>H_0: No hay diferencia en la distribución de una variable categorizada entre varias poblaciones o tratamientos.</p> <p>H_a: Sí hay diferencia en la distribución de una variable categorizada entre varias poblaciones o tratamientos.</p> <p>Se comienza hallando las cuentas previstas. Luego se computa la estadística de ji cuadrado</p> $\chi^2 = \sum \frac{(\text{Observadas} - \text{Previstas})^2}{\text{Previstas}}$ <p>en la que la suma está por sobre todas las celdas (sin incluir los totales) en la tabla de doble vía. Si H_0 es verdad, la estadística χ^2 tiene una distribución de aproximadamente ji cuadrado con grados de libertad = (número de renglones - 1)(número de columnas - 1). El valor P es el área a la derecha de χ^2 bajo la curva de densidad de ji cuadrado correspondiente. (pág. 708)</p>
<p>chi-square test for independence Suppose the Random, 10%, and Large Counts conditions are met. You can use the chi-square test for independence to test</p> <p>H_0: There is no association between two categorical variables in the population of interest.</p> <p>H_a: There is an association between two categorical variables in the population of interest.</p> <p>Or, alternatively,</p> <p>H_0: Two categorical variables are independent in the population of interest</p> <p>H_a: Two categorical variables are not independent in the population of interest.</p> <p>Start by finding the expected counts. Then calculate the chi-square statistic</p> $\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$ <p>where the sum is over all cells in the two-way table. If H_0 is true, the χ^2 statistic has approximately a chi-square distribution with degrees of freedom = (number of rows - 1)(number of columns - 1). The P-value is the area to the right of χ^2 under the corresponding chi-square density curve. (p. 697)</p>	<p>prueba de ji cuadrado de independencia Supongamos que se han cumplido las condiciones de aleatorio, del 10% y de cuentas grandes. Se puede usar la prueba de ji cuadrado de independencia para verificar que</p> <p>H_0: No hay ninguna asociación entre dos variables categorizadas en la población de interés.</p> <p>H_a: Sí hay una asociación entre dos variables categorizadas en la población de interés.</p> <p>O alternativamente,</p> <p>H_0: Dos variables categorizadas son independientes en la población de interés.</p> <p>H_a: Dos variables categorizadas no son independientes en la población de interés.</p> <p>Se comienza encontrando las cuentas previstas. Luego se computa la estadística de ji cuadrado</p> $\chi^2 = \sum \frac{(\text{Observadas} - \text{Previstas})^2}{\text{Previstas}}$ <p>en la que la suma está sobre todas las celdas en la tabla de doble vía. Si H_0 es verdad, la estadística χ^2 tiene una distribución de aproximadamente ji cuadrado con grados de libertad = (número de renglones - 1)(número de columnas - 1). El valor P es el área a la derecha de χ^2 bajo la curva de densidad de ji cuadrado correspondiente. (pág. 697)</p>
<p>cluster sample Sample obtained by classifying the population into groups of individuals that are located near each other, called <i>clusters</i>, and then choosing an SRS of the clusters. All individuals in the chosen clusters are included in the sample. (p. 221)</p>	<p>muestra de clúster Muestra que se obtiene clasificando la población en grupos de individuos que están ubicados uno cerca del otro, llamados <i>clústers</i>, y luego escogiendo una muestra aleatoria sencilla de los clústers. Todos los individuos en los clústers escogidos se incluyen en la muestra. (pág. 221)</p>

<p>coefficient of determination r^2 Fraction of the variation in the values of y that is accounted for by the least-squares regression line of y on x. We can calculate r^2 using the formula</p> $r^2 = 1 - \frac{\sum \text{residuals}^2}{\sum (y_i - \bar{y})^2}$ <p>(p. 179)</p>	<p>coeficiente de determinación r^2 La fracción de la variación en los valores y que se tiene en cuenta por la línea de regresión de mínimos cuadrados de y sobre x. Se puede computar r^2 utilizando la fórmula</p> $r^2 = 1 - \frac{\sum \text{residuales}^2}{\sum (y_i - \bar{y})^2}$ <p>(pág. 179)</p>
<p>comparison Experimental design principle. Use a design that compares two or more treatments. (p. 240)</p>	<p>comparación Principio de diseño experimental. Se usa un diseño que compara dos o más tratamientos. (pág. 240)</p>
<p>complement of an event A^C Event “not A”. (p. 307)</p>	<p>complemento de un evento A^C Se refiere al evento “que no es A”. (pág. 307)</p>
<p>complement rule The probability that an event does not occur is 1 minus the probability that the event does occur. In symbols, $P(A^C) = 1 - P(A)$. (pp. 308, 314)</p>	<p>regla del complemento La probabilidad de que no suceda un evento es 1 menos la probabilidad de que el evento sí suceda. En representación simbólica, $P(A^C) = 1 - P(A)$. (págs. 308, 314)</p>
<p>completely randomized design Design in which the experimental units are assigned to the treatments completely by chance. (p. 245)</p>	<p>diseño completamente aleatorizado Cuando las unidades experimentales se les asignan a los tratamientos de manera completamente al azar. (pág. 245)</p>
<p>components Individual terms $\frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$ that are added together to produce the test statistic χ^2. (p. 690)</p>	<p>componentes Los términos individuales $\frac{(\text{Observadas} - \text{Previstas})^2}{\text{Previstas}}$ que se suman para producir la estadística de prueba χ^2. (pág. 690)</p>
<p>conditional distribution Term that describes the values of one variable among individuals who have a specific value of another variable. There is a separate conditional distribution for each value of the other variable. (p. 15)</p>	<p>distribución condicional Describe los valores de una variable entre individuos que tienen un valor específico de otra variable. Hay una distribución condicional separada para cada valor de la otra variable. (pág. 15)</p>
<p>conditional probability Probability that one event happens given that another event is already known to have happened. Suppose we know that event A has happened. Then the probability that event B happens given that event A has happened is denoted by $P(B A)$. To find the conditional probability $P(B A)$, use the formula</p> $P(B A) = \frac{P(A \cap B)}{P(A)}$ <p>(p. 320)</p>	<p>probabilidad condicional La probabilidad de que un evento suceda a la luz de que se sabe que otro evento ya sucedió. Supongamos que nos consta que el evento A ya sucedió. Entonces la probabilidad de que el evento B suceda en vista de que el evento A ya sucedió, se denota con $P(B A)$. Para hallar la probabilidad condicional $P(B A)$, se usa la fórmula</p> $P(B A) = \frac{P(A \cap B)}{P(A)}$ <p>(pág. 320)</p>
<p>conditions for regression inference Suppose we have n observations on an explanatory variable x and a response variable y. Our goal is to study or predict the behavior of y for given values of x.</p> <ul style="list-style-type: none"> • Linear: The actual relationship between x and y is linear. For any fixed value of x, the mean response y falls on the population (true) regression line $\mu_y = \alpha + \beta x$. • Independent: Individual observations are independent. When sampling is done without replacement, check the 10% condition. • Normal: For any fixed value of x, the response y varies according to a Normal distribution. 	<p>condiciones para la inferencia de regresión Supongamos que tenemos n observaciones en una variable explicativa x y una variable de respuesta y. Nuestra meta consiste en estudiar o predecir el comportamiento de y ante los valores dados de x.</p> <ul style="list-style-type: none"> • Lineal (<i>linear</i>): La relación real entre x y y es lineal. Para todo valor fijo de x, la respuesta media y cae en la línea de regresión de la población (verdadera) $\mu_y = \alpha + \beta x$. • Independiente (<i>independent</i>): Las observaciones individuales son independientes. Cuando el muestreo se hace sin reemplazo, se verifica la condición del 10%. • Normal (<i>normal</i>): Para cualquier valor fijo de x, la respuesta y varía según una distribución normal.

G-6 Glossary/Glosario

<ul style="list-style-type: none"> • Equal SD: The standard deviation of y (call it σ) is the same for all values of x. • Random: The data are produced from a well-designed random sample or randomized experiment. (p. 743) 	<ul style="list-style-type: none"> • Desviación estándar equivalente: La desviación estándar de y es la misma para todos los valores de x. • Aleatorio (<i>random</i>): Los datos son producidos a partir de una muestra aleatoria o un experimento aleatorio, ambos bien diseñados. (pág. 743)
<p>confidence interval Gives an interval of plausible values for a parameter. The interval is calculated from the data and has the form</p> $\text{point estimate} \pm \text{margin of error}$ <p>or, alternatively,</p> $\text{statistic} \pm (\text{critical value}) \cdot (\text{standard deviation of statistic})$ <p>(p. 480)</p>	<p>intervalo de confianza Ofrece un intervalo de valores plausibles para un parámetro. El intervalo se computa a partir de los datos y tiene la forma</p> $\text{Estimado de punto} \pm \text{margen de error}$ <p>o alternativamente,</p> $\text{estadística} \pm (\text{valor crítico}) \cdot (\text{desviación estándar de la estadística})$ <p>(pág. 480)</p>
<p>confidence level C Success rate of the method for calculating the confidence interval. In $C\%$ of all possible samples, the method would yield an interval that captures the true parameter value. (p. 480)</p>	<p>nivel de confianza C La tasa de aciertos del método con el que se computa el intervalo de confianza. En el $C\%$ de todas las muestras posibles, el método produciría un intervalo que capta el valor verdadero del parámetro. (pág. 480)</p>
<p>confidential A basic principle of data ethics that requires that individual data be kept private. (p. 270)</p>	<p>confidencial Principio básico de la ética de la gestión de datos. Requiere que los datos individuales se mantengan en reserva. (pág. 270)</p>
<p>confounding When two variables are associated in such a way that their effects on a response variable cannot be distinguished from each other. (p. 236)</p>	<p>confuso Cuando dos variables se asocian de tal manera que sus efectos en una variable de respuesta no se pueden distinguir el uno del otro. (pág. 236)</p>
<p>continuous random variable Variable that takes all values in an interval of numbers. The probability distribution of a continuous random variable is described by a density curve. The probability of any event is the area under the density curve and above the values of the variable that make up the event. (p. 356)</p>	<p>variable aleatoria continua Emplea todos los valores en un intervalo de cifras. La distribución de la probabilidad de una variable aleatoria continua se describe con una curva de densidad. La probabilidad de cualquier evento en el área debajo de la curva de densidad y encima de los valores de la variable que componen el evento. (pág. 356)</p>
<p>control Experimental design principle that mandates keeping other variables that might affect the response the same for all groups. (p. 242)</p>	<p>control Principio del diseño experimental. Se mantienen otras variables que podrían afectar la respuesta iguales para todos los grupos. (pág. 242)</p>
<p>control group Experimental group whose primary purpose is to provide a baseline for comparing the effects of the other treatments. Depending on the purpose of the experiment, a control group may be given a placebo or an active treatment. (p. 246)</p>	<p>grupo de control Grupo experimental cuyo fin primario es establecer una línea base mediante la cual se comparan los efectos de otros tratamientos. Según el objeto del experimento, a un grupo de control se le puede dar un placebo o un tratamiento activo. (pág. 246)</p>
<p>convenience sample Sample selected by taking from the population individuals that are easy to reach. (p. 212)</p>	<p>muestra de conveniencia Muestra escogida de individuos de la población con quienes es fácil hacer contacto. (pág. 212)</p>
<p>correlation Measures the direction and strength of the linear relationship between two quantitative variables. Correlation is usually written as r. We can calculate r using the formula</p> $r = \frac{1}{n-1} \sum \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right). \text{ (pp. 150, 154)}$	<p>correlación Mide el sentido y la fuerza de la relación lineal entre dos variables cuantitativas. La correlación generalmente se denomina con una r. Calculamos la r con la fórmula</p> $r = \frac{1}{n-1} \sum \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right). \text{ (págs. 150, 154)}$
<p>critical value Multiplier that makes the interval wide enough to have the stated capture rate. The critical value depends on both the confidence level C and the sampling distribution of the statistic. (pp. 486, 497)</p>	<p>valor crítico Multiplicador que amplía el intervalo lo suficiente para retener la tasa de captación indicada. El valor crítico depende de tanto el nivel de confianza C como de la distribución de muestreo de la estadística. (págs. 486, 497)</p>

<p>cumulative relative frequency graph Graph used to examine location within a distribution. Cumulative relative frequency graphs begin by grouping the observations into equal-width classes. The completed graph shows the accumulating percent of observations as you move through the classes in increasing order. (p. 87)</p>	<p>gráfico de la frecuencia relativa acumulada Se usa para examinar la ubicación dentro de una distribución. Los gráficos de la frecuencia relativa acumulada se inician agrupando las observaciones en clases del mismo ancho. El gráfico completado muestra el porcentaje de observaciones que se van acumulando a medida que se desplaza por las clases en orden ascendente. (pág. 87)</p>
D	
<p>data analysis Process of describing data using graphs and numerical summaries. (p. 2)</p>	<p>análisis de los datos Proceso que describe los datos haciendo uso de gráficos y resúmenes numéricos. (pág. 2)</p>
<p>density curve Curve that (a) is always on or above the horizontal axis and (b) has area exactly 1 underneath it. A density curve describes the overall pattern of a distribution. The area under the curve and above any interval of values on the horizontal axis is the proportion of all observations that fall in that interval. (p. 105)</p>	<p>curva de densidad Curva que (a) siempre está sobre o por encima del eje horizontal y (b) tiene 1 área exactamente debajo. La curva de densidad describe el patrón general de una distribución. El área debajo de la curva y encima de todo intervalo de valores en el eje horizontal es la proporción de todas las observaciones que caen en dicho intervalo. (pág. 105)</p>
<p>describing a distribution In any graph of data, look for the overall pattern and for striking departures from that pattern. <i>Shape, center, and spread</i> describe the overall pattern of the distribution of a quantitative variable. (p. 26)</p>	<p>descripción de una distribución En un gráfico de datos, se observa cuál es el patrón general y se busca también valores atípicos que no se ajusten al patrón. <i>Forma, centro y amplitud</i> describen el patrón general de la distribución de una variable cuantitativa. (pág. 26)</p>
<p>describing a scatterplot In any graph of data, look for the overall pattern and for striking departures from that pattern. <i>Direction, form, and strength</i> describe the overall pattern of a scatterplot. (p. 147)</p>	<p>descripción de un gráfico de dispersión En todo gráfico de datos, se observa cuál es el patrón general y se busca también valores atípicos que no se ajusten al patrón. <i>Dirección, forma y fuerza</i> describen el patrón general de la distribución de un gráfico de dispersión. (pág. 147)</p>
<p>discrete random variable Takes a fixed set of possible values with gaps between. The probability distribution of a discrete random variable gives its possible values and their probabilities. The probability of any event is the sum of the probabilities for the values of the variable that make up the event. (p. 348)</p>	<p>variable aleatoria discreta Emplea un conjunto fijo de valores posibles entre los cuales hay brechas. La distribución de la probabilidad de una variable aleatoria discreta arroja valores posibles y sus probabilidades. La probabilidad de cualquier evento es la suma de las probabilidades de los valores de la variable que compone el evento. (pág. 348)</p>
<p>distribution Tells what values a variable takes and how often it takes these values. (p. 4)</p>	<p>distribución Indica qué valores adopta una variable y con qué frecuencia adopta dichos valores. (pág. 4)</p>
<p>distribution of sample data Gives the values of the variable for all the individuals in the sample. (p. 428)</p>	<p>distribución de los datos de la muestra Indica los valores de la variable que les corresponden a todos los individuos en la muestra. (pág. 428)</p>
<p>dotplot Simple graph that shows each data value as a dot above its location on a number line. (p. 25)</p>	<p>gráfico de puntos Un gráfico sencillo que muestra el valor de cada dato encima de su ubicación a lo largo de una línea de cifras. (pág. 25)</p>
<p>double-blind An experiment in which neither the subjects nor those who interact with them and measure the response variable know which treatment a subject received. (p. 248)</p>	<p>doble ciego Experimento en el que ninguno de los sujetos ni aquellos que interactúan con los sujetos y que miden la variable de respuesta saben qué tratamiento recibió el sujeto. (pág. 248)</p>
E	
<p>event Any collection of outcomes from some chance process. An event is a subset of the sample space. Events are usually designated by capital letters, like A, B, C, and so on. (p. 306)</p>	<p>evento Cualquier colección de los resultados de un proceso de probabilidad. Es decir, un evento es un subconjunto del espacio de muestras. Los eventos generalmente se designan con mayúsculas tales como A, B, C, y así sucesivamente. (pág. 306)</p>

G-8 Glossary/Glosario

expected counts Expected numbers of individuals in the sample that would fall in each cell of the one-way or two-way table if H_0 were true. (p. 681)	cuentas previstas Las cantidades previstas de individuos en la muestra que caerían en cada celda en la tabla, sea de una vía o de dos vías, si H_0 fuera verdad. (pág. 681)
experiment A study in which researchers deliberately impose treatments on individuals to measure their responses. (p. 235)	experimento Estudio en el que los investigadores deliberadamente les imponen tratamientos a individuos, con el fin de medir sus respuestas. (pág. 235)
experimental units Smallest collection of individuals to which treatments are applied. (p. 237)	unidades experimentales La colección más pequeña de individuos a quienes se les aplican los tratamientos. (pág. 237)
explanatory variable Variable that may help explain or predict changes in a response variable. (pp. 143, 236)	variable explicativa Variable que puede ayudar a explicar o predecir cambios en una variable de respuesta. (págs. 143, 236)
exponential model Relationship of the form $y = ab^x$. If the relationship between two variables follows an exponential model and we plot the logarithm (base 10 or base e) of y against x , we should observe a straight-line pattern in the transformed data. (p. 775)	modelo exponencial Relación de la forma $y = ab^x$. Si la relación entre dos variables se ajusta a un modelo exponencial, y trazamos el logaritmo (de base 10 o de base e) de y con respecto a x , se debe observar un patrón en línea recta en los datos transformados. (pág. 775)
extrapolation Use of a regression line for prediction far outside the interval of values of the explanatory variable x used to obtain the line. Such predictions are often not accurate. (p. 168)	extrapolación Uso de una línea de regresión para hacer predicciones muy por fuera del intervalo de valores de la variable explicativa x que se utiliza para obtener la línea. Tales predicciones a menudo carecen de precisión. (pág. 168)
F	
factorial For any positive whole number n , its factorial $n!$ is $n! = n \cdot (n - 1) \cdot (n - 2) \cdot \dots \cdot 3 \cdot 2 \cdot 1$. In addition, we define $0! = 1$. (p. 409)	factorial Para cualquier número entero positivo n , su factorial $n!$ es $n! = n \cdot (n - 1) \cdot (n - 2) \cdot \dots \cdot 3 \cdot 2 \cdot 1$. Además, definimos $0! = 1$. (pág. 409)
factors Explanatory variables in an experiment. (p. 238)	factores Las variables explicativas en un experimento. (pág. 238)
fail to reject H_0 If the observed result is not very unlikely to occur when the null hypothesis is true, we should fail to reject H_0 and say that we do not have convincing evidence for H_a . (p. 544)	no rechazar H_0 Si no es muy improbable que el resultado observado suceda cuando es verdad la hipótesis nula, no se debe rechazar H_0 y se ha de indicar que no contamos con evidencia convincente de H_a . (pág. 544)
first quartile Q_1 If the observations in a data set are ordered from lowest to highest, the first quartile Q_1 is the median of the observations whose position is to the left of the median. (p. 54)	primer cuartil Q_1 Si las observaciones del conjunto de datos se organizan en orden ascendente (del más bajo al más alto), el primer cuartil Q_1 es la media de las observaciones cuya posición se encuentra a la izquierda de la media. (pág. 54)
five-number summary Smallest observation, first quartile, median, third quartile, and largest observation, written in order from smallest to largest. In symbols: Minimum Q_1 Median Q_3 Maximum (p. 57)	resumen de cinco cifras Consta de la observación más pequeña, el primer cuartil, la media, el tercer cuartil y la observación más grande, enumeradas en orden ascendente, desde la más pequeña hasta la más grande. Representado en forma simbólica, el resumen de cinco cifras es Mínimo Q_1 Media Q_3 Máximo (pág. 57)
frequency table Table that displays the count (frequency) of observations in each category or class. (p. 8)	tabla de frecuencias Muestra la cuenta (frecuencia) de observaciones en cada categoría o clase. (pág. 8)
G	
general addition rule If A and B are any two events resulting from some chance process, then the probability that event A or event B (or both) occur is $P(A \text{ or } B) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$. (p. 310)	regla general de adición Si A y B son dos eventos cualquiera que resulten de algún proceso de probabilidad, la probabilidad de que el evento A o el evento B (o ambos) suceda es $P(A \text{ o } B) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$. (pág. 310)

<p>general multiplication rule The probability that events A and B both occur can be found using the formula $P(A \cap B) = P(A) \cdot P(B A)$ where $P(B A)$ is the conditional probability that event B occurs given that event A has already occurred. (p. 321)</p>	<p>regla general de multiplicación La probabilidad de que sucedan los eventos A y B se puede determinar utilizando la fórmula $P(A \cap B) = P(A) \cdot P(B A)$ en la que $P(B A)$ es la probabilidad condicional de que suceda el evento B a la luz de que el evento A ya sucedió. (pág. 321)</p>
<p>geometric distribution In a geometric setting, suppose we let $Y =$ the number of trials it takes to get a success. The probability distribution of Y is a geometric distribution with parameter p, the probability of a success on any trial. The possible values of Y are 1, 2, 3, (p. 405)</p>	<p>distribución geométrica En un entorno geométrico, supongamos que se permite que $Y =$ la cantidad de ensayos que se precisan para lograr un acierto. La distribución de la probabilidad de Y es una distribución geométrica con el parámetro p, la probabilidad de lograr n acierto en cualquier ensayo. Los valores posibles de Y son 1, 2, 3, (pág. 405)</p>
<p>geometric probability formula If Y has the geometric distribution with probability p of success on each trial, the possible values of Y are 1, 2, 3, If k is any one of these values, $P(Y = k) = (1 - p)^{k-1}p$. (p. 406)</p>	<p>fórmula de probabilidad geométrica Si Y tiene una distribución geométrica con la probabilidad p de acierto en cada ensayo, los posibles valores de Y son 1, 2, 3, Si k es uno cualquiera de estos valores, $P(Y = k) = (1 - p)^{k-1}p$. (pág. 406)</p>
<p>geometric random variable The number of trials Y that it takes to get a success in a geometric setting. (p. 405)</p>	<p>variable aleatoria geométrica La cantidad de ensayos Y que se precisan para lograr un acierto en un entorno geométrico. (pág. 405)</p>
<p>geometric setting Arises when we perform independent trials of the same chance process and record the number of trials it takes to get one success. On each trial, the probability p of success must be the same. (p. 404)</p>	<p>entorno geométrico Surge un entorno geométrico cuando se realizan ensayos independientes del mismo proceso de probabilidad y se graban la cantidad de ensayos que se precisan para lograr un acierto. En cada ensayo, la probabilidad p de lograr un acierto tiene que ser la misma. (pág. 404)</p>
<p>H</p>	
<p>histogram Graph that displays the distribution of a quantitative variable. The horizontal axis is marked in the units of measurement for the variable. The vertical axis contains the scale of counts or percents. Each bar in the graph represents an equal-width class. The base of the bar covers the class, and the bar height is the class frequency or relative frequency. (p. 33)</p>	<p>histograma Muestra la distribución de una variable cuantitativa. En el eje horizontal se denotan las unidades de medición de la variable. El eje vertical contiene la escala de cuentas o porcentajes. Cada barra del gráfico representa una clase de ancho equivalente. La base de la barra abarca la clase, y la altura de la barra es la frecuencia o la frecuencia relativa de la clase. (pág. 33)</p>
<p>I</p>	
<p>independent events Two events are independent if the occurrence of one event does not change the probability that the other event will happen. In other words, events A and B are independent if $P(A B) = P(A)$ and $P(B A) = P(B)$. (p. 327)</p>	<p>eventos independientes Dos eventos son independientes si el hecho de que uno suceda no cambia la probabilidad de que el otro suceda. Es decir, los eventos A y B son independientes si $P(A B) = P(A)$ y $P(B A) = P(B)$. (pág. 327)</p>
<p>independent random variables If knowing whether any event involving X alone has occurred tells us nothing about the occurrence of any event involving Y alone and vice versa, then X and Y are independent random variables. (p. 371)</p>	<p>variables aleatorias independientes Saber que ha sucedido un evento que implique el valor X solo, no nos indica nada respecto al hecho de que suceda un evento que implique el valor Y solo, y viceversa. En tal caso, tanto X como Y son variables aleatorias independientes. (pág. 371)</p>
<p>individuals Objects described by a set of data. Individuals may be people, animals, or things. (p. 2)</p>	<p>individuos Objetos descritos por un conjunto de datos. Los individuos pueden ser personas, animales o cosas. (pág. 2)</p>
<p>inference Drawing conclusions that go beyond the data at hand. (pp. 5, 223)</p>	<p>inferencia Llegar a conclusiones que van más allá de los datos que están a la mano. (págs. 5, 223)</p>
<p>inference about cause and effect Conclusion from the results of an experiment that the treatments caused the difference in responses. Requires a well-designed experiment in which the treatments are randomly assigned to the experimental units. (p. 266)</p>	<p>inferencia sobre causa y efecto Uso de los resultados de un experimento para llegar a la conclusión de que son los tratamientos los que marcan la diferencia en las respuestas. Exige un experimento bien diseñado en el que los tratamientos se asignan de manera aleatoria a las unidades experimentales. (pág. 266)</p>

G-10 Glossary/Glosario

<p>inference about a population Conclusion about the larger population based on sample data. Requires that the individuals taking part in a study be randomly selected from the population of interest. (p. 266)</p>	<p>inferencia sobre una población Conclusión sobre una población en general con base en datos muestrales. Se precisa que los participantes del estudio sean escogidos de manera aleatoria a partir de la población de interés. (pág. 266)</p>
<p>influential observation An observation is influential for a statistical calculation if removing it would markedly change the result of the calculation. Points that are outliers in the x direction of a scatterplot are often influential for the least-squares regression line. (p. 189)</p>	<p>observación influyente La observación es influyente en un cómputo estadístico si al retirarla se notaría un cambio sustancial en el resultado del cómputo. Los puntos que son valores atípicos en el sentido x en un gráfico de dispersión, a menudo son influyentes en al menos una línea de regresión de mínimos cuadrados. (pág. 189)</p>
<p>informed consent Basic principle of data ethics that states that individuals must be informed in advance about the nature of a study and any risk of harm it may bring. Participating individuals must then consent in writing. (p. 270)</p>	<p>autorización informada Principio básico de ética en la gestión de los datos. A los individuos se les ha de informar con antelación acerca de la naturaleza de un estudio y de los riesgos o perjuicios que podría conllevar. Los individuos que participen luego tendrán que dar su autorización por escrito. (pág. 270)</p>
<p>institutional review board Board charged with protecting the safety and well-being of the participants in advance of a planned study and with monitoring the study itself. (p. 270)</p>	<p>junta de revisión institucional Principio básico de la ética de la gestión de datos. Todos los estudios planificados tienen que contar con aprobación anticipada y tienen que contar con un monitoreo por una junta de revisión institucional cuya función consiste en salvaguardar la seguridad y el bienestar de los participantes. (pág. 270)</p>
<p>interquartile range $IQR = Q_3 - Q_1$. (p. 54)</p>	<p>gama entre cuartiles $IQR = Q_3 - Q_1$. (pág. 54)</p>
<p>intersection The intersection of events A and B, denoted by $A \cap B$, refers to the occurrence of both of two events at the same time. (p. 311)</p>	<p>intersección El punto de cruce de los eventos A y B, designado con $A \cap B$, se refiere a la situación en la que ambos eventos suceden simultáneamente. (pág. 311)</p>
L	
<p>lack of realism When the treatments, the subjects, or the environment of an experiment are not realistic. Lack of realism can limit researchers' ability to apply the conclusions of an experiment to the settings of greatest interest. (p. 268)</p>	<p>falta de realismo Cuando los tratamientos, los sujetos o el entorno de un experimento no son realistas. La carencia de realismo puede limitar la capacidad de los investigadores de aplicar las conclusiones de un experimento a los entornos de gran interés. (pág. 268)</p>
<p>Large Counts condition It is safe to use Normal approximation for performing inference about a proportion p if $np \geq 10$ and $n(1 - p) \geq 10$. (p. 403)</p>	<p>condición de cuentas grandes Se puede utilizar sin problemas la aproximación normal para realizar la inferencia de una proporción p si $np \geq 10$ y $n(1 - p) \geq 10$. (pág. 403)</p>
<p>Large Counts condition for a chi-square test It is safe to use a chi-square distribution to perform calculations if all expected counts are at least 5. (p. 687)</p>	<p>condición de cuentas grandes en la prueba de ji cuadrado Se puede utilizar sin problemas la distribución de ji cuadrado para realizar cómputos si todas las cuentas previstas son de al menos 5. (pág. 687)</p>
<p>law of large numbers If we observe more and more repetitions of any chance process, the proportion of times that a specific outcome occurs approaches a single value, which we call the probability of that outcome. (p. 291)</p>	<p>ley de las cifras grandes Si se observan más y más repeticiones en cualquier proceso de probabilidad, la proporción de veces que se da un resultado específico se aproxima a un valor sencillo, al cual se le denomina la probabilidad de dicho resultado. (pág. 291)</p>
<p>least-squares regression line The line that makes the sum of the squared vertical distances of the data points from the line as small as possible. (p. 169)</p>	<p>línea de regresión de mínimos cuadrados La línea que reduce al mínimo posible la suma de las distancias verticales cuadráticas de los puntos de datos a partir de la línea. (pág. 169)</p>
<p>level Specific value of an explanatory variable (factor) in an experiment.</p>	<p>nivel Valor específico de una variable explicativa (factor) en un experimento. (pág. 238)</p>

<p>linear transformation A transformation of a random variable that involves adding a constant a, multiplying by a constant b, or both. We can write a linear transformation of the random variable X in the form $Y = a + bX$. The shape, center, and spread of the probability distribution of Y are as follows: Shape: Same as the probability distribution of X unless b is negative. Center: $\mu_Y = a + b\mu_X$ Spread: $\sigma_Y = b \sigma_X$ (p. 368)</p>	<p>transformación lineal La transformación de una variable aleatoria que implica agregar una a constante, multiplicada por una b constante, o ambas. La transformación lineal de la variable aleatoria X se puede escribir en la forma $Y = a + bX$. La forma, el centro y la amplitud de la distribución de la probabilidad de Y son como sigue: Forma: Igual que la distribución de la probabilidad de X a menos que b sea negativo. Centro: $\mu_Y = a + b\mu_X$ Amplitud: $\sigma_Y = b \sigma_X$ (pág. 368)</p>
M	
<p>margin of error The difference between the point estimate and the true parameter value will be less than the margin of error in $C\%$ of all samples, where C is the confidence level. (p. 480)</p>	<p>margen de error La diferencia entre el estimado del punto y el valor real del parámetro será menor que el margen de error en $C\%$ de todas las muestras, en el que C es el nivel de confianza. (pág. 480)</p>
<p>marginal distribution The distribution of one of the categorical variables in a two-way table of counts among all individuals described by the table. (p. 12)</p>	<p>distribución marginal La distribución de una de las variables categorizadas en una tabla de doble vía de cuentas entre todos los individuos descritos por la tabla. (pág. 12)</p>
<p>matched pairs design Common form of blocking for comparing just two treatments. In some matched pairs designs, each subject receives both treatments in a random order. In others, the subjects are matched in pairs as closely as possible, and each subject in a pair is randomly assigned to receive one of the treatments. (p. 255)</p>	<p>diseño de pares coincidentes Forma común de crear bloques para efectos de comparación de tan solo dos tratamientos. En algunos diseños de pares coincidentes, cada tema se somete a ambos tratamientos en un orden aleatorio. En otros, los temas se ponen en pares que coincidan lo más posible y cada tema en un par se asigna de manera aleatoria a fin de que reciba uno de los tratamientos. (pág. 255)</p>
<p>mean \bar{x} Arithmetic average. To find the mean of a set of observations, add their values and divide by the number of observations. In symbols, $\bar{x} = \frac{\sum x_i}{n}$ (p. 49)</p>	<p>media \bar{x} El promedio aritmético. Para hallar la media de un conjunto de observaciones, se suman todos los valores y se divide entre el número de observaciones. En símbolos, $\bar{x} = \frac{\sum x_i}{n}$ (pág. 49)</p>
<p>mean of a density curve Point at which a density curve would balance if made of solid material. (p. 107)</p>	<p>media de una curva de densidad El punto en el cual la curva se equilibraría si estuviera elaborada de un material macizo. (pág. 107)</p>
<p>mean (expected value) of a discrete random variable To find the mean (expected value) of X, multiply each possible value by its probability, then add all the products: $\mu_X = E(X) = x_1p_1 + x_2p_2 + x_3p_3 + \dots = \sum x_i p_i$ (p. 351)</p>	<p>media (valor previsto) de una variable aleatoria discreta Para hallar la media (un valor previsto) de X, se multiplica cada valor posible por su probabilidad y luego se suman todos los productos: $\mu_X = E(X) = x_1p_1 + x_2p_2 + x_3p_3 + \dots = \sum x_i p_i$ (pág. 351)</p>
<p>mean (expected value) of a geometric random variable If Y is a geometric random variable with probability of success p on each trial, then its mean (expected value) is $\mu_Y = E(Y) = \frac{1}{p}$. That is, the expected number of trials required to get the first success is $1/p$. (p. 408)</p>	<p>media (valor previsto) de una variable aleatoria geométrica Si Y es una variable aleatoria geométrica con probabilidad de acierto p en cada ensayo, entonces su media (un valor previsto) es $\mu_Y = E(Y) = \frac{1}{p}$. Es decir, la cifra de ensayos prevista para lograr el primer acierto es $1/p$. (pág. 408)</p>
<p>mean and standard deviation of a binomial random variable If a count X of successes has the binomial distribution with number of trials n and probability of success p, the mean and standard deviation of X are $\mu_X = np$ and $\sigma_X = \sqrt{np(1-p)}$. (p. 398)</p>	<p>media y desviación estándar de una variable binomial aleatoria Si una cuenta X de aciertos tiene una distribución binomial con la cantidad de ensayos n y la probabilidad de aciertos p, la media y la desviación estándar de X son $\mu_X = np$ and $\sigma_X = \sqrt{np(1-p)}$ (pág. 398)</p>

G-12 Glossary/Glosario

<p>mean and standard deviation of the sampling distribution of a sample mean \bar{x} Suppose that \bar{x} is the mean of an SRS of size n from a large population with mean μ and standard deviation σ. Then</p> <ul style="list-style-type: none"> • The mean of the sampling distribution of \bar{x} is $\mu_{\bar{x}} = \mu$. • The standard deviation of the sampling distribution of \bar{x} is $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ as long as the <i>10% condition</i> is satisfied: $n \leq \frac{1}{10}N$. (p. 452) 	<p>media y desviación estándar de la distribución de muestreo de la media de una muestra Supongamos que \bar{x} es la media de una muestra aleatoria sencilla de tamaño n a partir de una población grande con media μ y una desviación estándar σ. Así</p> <ul style="list-style-type: none"> • La media de la distribución de muestreo de \bar{x} es $\mu_{\bar{x}} = \mu$. • La desviación estándar de la distribución de muestreo de \bar{x} es $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ siempre y cuando se cumpla con la <i>condición del 10%</i>: $n \leq \frac{1}{10}N$. (pág. 452)
<p>median The midpoint of a distribution; the number such that about half the observations are smaller and about half are larger. To find the median of a distribution: (1) Arrange all observations in order of size, from smallest to largest. (2) If the number of observations n is odd, the median is the center observation in the ordered list. (3) If the number of observations n is even, the median is the average of the two center observations in the ordered list. (p. 51)</p>	<p>media El punto intermedio de una distribución, con una cifra tal que aproximadamente la mitad de las observaciones son más pequeñas y la mitad son más grandes. Para hallar la media de una distribución: (1) Se organizan todas las observaciones en orden de su tamaño, de las más pequeñas a las más grandes. (2) Si la cantidad de observaciones n es impar, la media es la observación central en la lista organizada. (3) Si la cantidad de observaciones n es par, la media es el promedio de las dos observaciones centrales en la lista organizada. (pág. 51)</p>
<p>median of a density curve The point with half the area under the curve to its left and the remaining half of the area to its right. (p. 106)</p>	<p>media de una curva de densidad El punto en el que la mitad del área que está debajo de la curva está a la izquierda y la otra mitad del área está a la derecha. (pág. 106)</p>
<p>mode Value or class in a statistical distribution having the greatest frequency. (p. 26)</p>	<p>modo En una distribución estadística, el valor o clase que tiene la mayor frecuencia. (pág. 26)</p>
<p>multimodal A graph of quantitative data with more than two clear peaks. (p. 29)</p>	<p>multimodal Gráfico de datos cuantitativos que tiene más de dos picos claros. (pág. 29)</p>
<p>multiple comparisons Problem of how to do many comparisons at once with an overall measure of confidence in all our conclusions. (p. 700)</p>	<p>comparaciones múltiples El problema de cómo hacer muchas comparaciones a la vez con una medida de confianza general en todas las conclusiones a las que se llega. (pág. 700)</p>
<p>multiplication rule for independent events If A and B are independent events, then the probability that A and B both occur is $P(A \cap B) = P(A) \cdot P(B)$. (p. 328)</p>	<p>regla de multiplicación de eventos independientes Si A y B son eventos independientes, la probabilidad de que sucedan ambos, tanto A como B es $P(A \cap B) = P(A) \cdot P(B)$. (pág. 328)</p>
<p>mutually exclusive (disjoint) Two events that have no outcomes in common and so can never occur together. (p. 307)</p>	<p>exclusivos mutuamente (desencajamiento) Dos eventos que no tienen resultados en común y por lo tanto nunca pueden suceder a la vez. (pág. 307)</p>
N	
<p>negative association When above-average values of one variable tend to accompany below-average values of the other. (p. 148)</p>	<p>asociación negativa Cuando los valores por encima del promedio de una variable tienden a acompañar a los valores por debajo del promedio de la otra. (pág. 148)</p>
<p>nonresponse Occurs when an individual chosen for the sample can't be contacted or refuses to participate. (p. 225)</p>	<p>no respondió Sucede cuando a un individuo escogido para la muestra no se le puede contactar o el sujeto se niega a participar. (pág. 225)</p>
<p>Normal approximation to a binomial distribution Suppose that a count X of successes has the binomial distribution with n trials and success probability p. When n is large, the distribution of X is approximately Normal with mean np and standard deviation $\sqrt{np(1-p)}$. We use this approximation when $np \geq 10$ and $n(1-p) \geq 10$. (p. 403)</p>	<p>aproximación normal hacia una distribución binomial Supongamos que una cuenta X de aciertos tiene la distribución binomial con n ensayos y una probabilidad de acierto p. Cuando n es grande, la distribución de X es aproximadamente normal con media np y desviación estándar $\sqrt{np(1-p)}$. Se hace uso de esta aproximación cuando $np \geq 10$ and $n(1-p) \geq 10$. (pág. 403)</p>

<p>Normal curves Important class of density curves that are symmetric, single-peaked, and bell-shaped. (p. 109)</p>	<p>curvas normales Clase importante de curvas de densidad que son simétricas, de un solo pico y con la forma de curva de campana. (pág. 109)</p>
<p>Normal distribution Distribution described by a Normal density curve. Any particular Normal distribution is completely specified by two numbers, its mean μ and standard deviation σ. The mean of a Normal distribution is at the center of the symmetric Normal curve. The standard deviation is the distance from the center to the change-of-curvature points on either side. We abbreviate the Normal distribution with mean μ and standard deviation σ as $N(\mu, \sigma)$. (p. 109)</p>	<p>distribución normal Según la describe una curva de densidad normal. Cualquier distribución normal dada se especifica completamente con dos cifras, su media μ y la desviación estándar σ. La media de una distribución normal yace en el centro de la curva normal simétrica. La desviación estándar es la distancia del centro a los puntos a ambos lados en los que cambia la curva. La distribución normal con la media μ y la desviación estándar σ se abrevia $N(\mu, \sigma)$. (pág. 109)</p>
<p>Normal/Large Sample condition for inference about a mean The population has a Normal distribution or the sample size is large ($n \geq 30$). If the population distribution has unknown shape and $n < 30$, use a graph of the sample data to assess the Normality of the population. Do not use t procedures if the graph shows strong skewness or outliers. (p. 515)</p>	<p>condición de muestra normal/grande para inferir sobre una media La población tiene una distribución normal o la muestra es de tamaño grande ($n \geq 30$). Si la distribución de la población tiene una forma desconocida y $n < 30$, se usa un gráfico de los datos de la muestra para evaluar la normalidad de la población. No se han de usar los procedimientos t si en el gráfico se aprecian valores atípicos o un sesgo marcado. (pág. 515)</p>
<p>Normal probability plot Plot used to assess whether a data set follows a Normal distribution. To make a Normal probability plot, (1) arrange the data values from smallest to largest and record the percentile of each observation, (2) use the standard Normal distribution to find the z-scores at these same percentiles, and (3) plot each observation x against the corresponding z. If the points on a Normal probability plot lie close to a straight line, the plot indicates that the data are approximately Normal. (p. 122)</p>	<p>gráfico de probabilidad normal Se usa para evaluar si un conjunto de datos se ciñe a una distribución normal. Para trazar un gráfico de probabilidad normal, (1) se dispone de los valores de los datos del más pequeño al más grande y se anota el percentil de cada observación, (2) se usa la distribución normal estándar para hallar los puntos z en esos mismos percentiles, y (3) se traza cada observación x con la z correspondiente. Si los puntos en un gráfico de probabilidad normal yacen cerca de una línea recta, el gráfico indica que los datos son aproximadamente normales. (pág. 122)</p>
<p>null hypothesis H_0 Claim we weight evidence against in a significance test. Often the null hypothesis is a statement of “no difference.” (p. 540)</p>	<p>hipótesis nula H_0 Contrapeso de la evidencia en una prueba de significancia. A menudo la hipótesis nula es una declaración de “no hay diferencia.” (pág. 540)</p>
<p>O</p>	
<p>observational study Study that observes individuals and measures variables of interest but does not attempt to influence the responses. (p. 235)</p>	<p>estudio de observación Se observan los individuos y se miden las variables de interés pero no se trata de influir en las respuestas. (pág. 235)</p>
<p>observed counts Actual numbers of individuals in the sample that fall in each cell of the one-way or two-way table. (p. 681)</p>	<p>cuentas observadas Las cifras reales que corresponden a individuos en la muestra que caen en cada celda de la tabla de una vía o en la de dos vías. (pág. 681)</p>
<p>one-sample t interval for a mean When the Random, 10%, and Normal/Large Sample conditions are met, a $C\%$ confidence interval for μ is</p> $\bar{x} \pm t^* \frac{s_x}{\sqrt{n}}$ <p>where t^* is the critical value for the t distribution with $df = n - 1$, with $C\%$ of the area between $-t^*$ and t^*. (p. 518)</p>	<p>intervalo t de una sola muestra para una media Cuando se cumplen las condiciones de aleatorio, del 10% y de cuentas normales/grandes, un intervalo de confianza $C\%$ para es</p> $\bar{x} \pm t^* \frac{s_x}{\sqrt{n}}$ <p>en el que t^* es el valor crítico para la distribución t con $df = n - 1$, con $C\%$ del área entre $-t^*$ y t^*. (pág. 518)</p>

G-14 Glossary/Glosario

<p>one-sample t test for a mean Suppose that the Random, 10%, and Normal/Large Sample conditions are met. To test the hypothesis $H_0: \mu = \mu_0$, compute the one-sample t statistic</p> $t = \frac{\bar{x} - \mu_0}{\frac{s_x}{\sqrt{n}}}$ <p>Find the P-value by calculating the probability of getting a t statistic this large or larger in the direction specified by the alternative hypothesis H_a in a t distribution with $df = n - 1$. (p. 580)</p>	<p>intervalo t de una sola muestra para una media Supongamos que se han cumplido las condiciones de aleatorio, del 10% y de muestra normal/grande. Para probar la hipótesis $H_0: \mu = \mu_0$, se computa la estadística t de una sola muestra</p> $t = \frac{\bar{x} - \mu_0}{\frac{s_x}{\sqrt{n}}}$ <p>Se halla el valor P computando la probabilidad de obtener una estadística t de este tamaño o más grande en el sentido especificado por la hipótesis alternativa en una distribución t con $df = n - 1$. (pág. 580)</p>
<p>one-sample z interval for a proportion When the Random, 10%, and Large Counts conditions are met, a $C\%$ confidence interval for the unknown proportion p is</p> $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$ <p>where z^* is the critical value for the standard Normal curve with $C\%$ of its area between $-z^*$ and z^*. (p. 498)</p>	<p>intervalo z de una sola muestra para una proporción Cuando se cumplen las condiciones de aleatorio, del 10% y de cuentas grandes, un intervalo de confianza $C\%$ para la proporción desconocida p es</p> $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$ <p>en el que z^* es el valor crítico para la curva normal estándar con $C\%$ de su área entre $-z^*$ y z^*. (pág. 498)</p>
<p>one-sample z test for a proportion Suppose that the Random, 10%, and Large Counts conditions are met. To test the hypothesis $H_0: p = p_0$, compute the z statistic</p> $z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$ <p>Find the P-value by calculating the probability of getting a z statistic this large or larger in the direction specified by the alternative hypothesis H_a. (p. 559)</p>	<p>prueba z de una sola muestra para una proporción Supongamos que se han cumplido las condiciones de aleatorio, del 10% y de muestras grandes. Para probar la hipótesis $H_0: p = p_0$, se computa la estadística</p> $z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$ <p>Se halla el valor P computando la probabilidad de obtener una estadística z de este tamaño o más grande en el sentido especificado por la hipótesis alternativa H_a. (pág. 559)</p>
<p>one-sided alternative hypothesis An alternative hypothesis that states that a parameter is larger than the null hypothesis value or that states that the parameter is smaller than the null value. (p. 541)</p>	<p>hipótesis alternativa unilateral Hipótesis alternativa que indica que un parámetro es más grande que el valor de la hipótesis nula o que indica que el parámetro es más pequeño que el valor nulo. (pág. 541)</p>
<p>one-way table Table used to display the distribution of a single categorical variable. (p. 680)</p>	<p>tabla de una vía Se usa para mostrar la distribución de una sola variable categorizada. (pág. 680)</p>
<p>outlier Individual value that falls outside the overall pattern of a distribution. (p. 26)</p>	<p>valor atípico Un valor individual que cae por fuera del patrón general de la distribución. (pág. 26)</p>
<p>outlier in regression Observation that lies outside the overall pattern of the other observations. Points that are outliers in the y direction but not the x direction of a scatterplot have large residuals. Other outliers may not have large residuals. (p. 189)</p>	<p>valor atípico en regresión Observación que yace por fuera del patrón general de las otras observaciones. Los puntos que son valores atípicos en el sentido y pero no en el sentido x de un gráfico de dispersión tienen residuales grandes. Es posible que otros valores atípicos no tengan residuales grandes. (pág. 189)</p>
P	
<p>P-value The probability, computed assuming H_0 is true, that the statistic would take a value as extreme as or more extreme than the one actually observed, in the direction specified by H_a. The smaller the P-value, the stronger the evidence against H_0 and in favor of H_a provided by the data. (p. 543)</p>	<p>valor P La probabilidad, computada suponiendo que H_0 es verdad, de que la estadística tomaría un valor tan extremo o más extremo que el que de hecho se observa, en el sentido especificado por H_a. Cuanto menor sea el valor P, más fuerte será la evidencia contra H_0 y en favor de H_a que proporcionan los datos. (pág. 543)</p>

<p>paired data Study designs that involve making two observations on the same individual or one observation on each of two similar individuals result in paired data. (p. 586)</p>	<p>datos apareados Se estudian diseños que implican hacer dos observaciones del mismo individuo, o una observación de cada uno de dos individuos parecidos, resultando en datos apareados. (pág. 586)</p>
<p>paired <i>t</i> procedures When paired data result from measuring the same quantitative variable twice, we can make comparisons by analyzing the differences in each pair. If the conditions for inference are met, we can use one-sample <i>t</i> procedures to perform inference about the mean difference μ_d. These methods are sometimes called paired <i>t</i> procedures. (p. 586)</p>	<p>procedimientos <i>t</i> apareados Cuando la misma variable cuantitativa se mide dos veces pueden resultar datos apareados, con los cuales se pueden hacer comparaciones que analizan las diferencias en cada par. Si se cumplen las condiciones para la inferencia, se pueden usar procedimientos <i>t</i> de una sola muestra para realizar una inferencia acerca de la diferencia media μ_d. A estos métodos a veces se les denomina procedimientos <i>t</i> apareados. (pág. 586)</p>
<p>parameter A number that describes some characteristic of the population. (p. 424)</p>	<p>parámetro Número que describe algunas de las características de la población. (pág. 424)</p>
<p>percentile The pth percentile of a distribution is the value with p percent of the observations less than it. (p. 85)</p>	<p>percentil El percentil pavo de una distribución es el valor cuyo porcentaje de las observaciones es menor que la cifra. (pág. 85)</p>
<p>pie chart Chart that shows the distribution of a categorical variable as a “pie” whose slices are sized by the counts or percents for the categories. A pie chart must include all the categories that make up a whole. (p. 8)</p>	<p>gráfico circular Muestra la distribución de una variable categorizada n la forma de un círculo subdividido según las cuentas o los porcentajes de las categorías. El gráfico circular tiene que incluir todas las categorías que componen la totalidad. (pág. 8)</p>
<p>placebo Inactive (fake) treatment. (p. 244)</p>	<p>placebo Tratamiento inactivo (falso). (pág. 244)</p>
<p>placebo effect Describes the fact that some subjects respond favorably to any treatment, even an inactive one (placebo). (p. 247)</p>	<p>efecto placebo Describe el hecho de que algunos sujetos responden de manera favorable a cualquier tratamiento, incluso uno inactivo (con placebo). (pág. 247)</p>
<p>point estimate Specific value of a point estimator from a sample. (p. 477)</p>	<p>estimado de punto El valor específico de un estimador de punto tomado de una muestra. (pág. 477)</p>
<p>point estimator Statistic that provides an estimate of a population parameter. (p. 477)</p>	<p>estimador de punto Estadística que nos da un estimado de un parámetro de la población. (pág. 477)</p>
<p>pooled or combined sample proportion The overall proportion of successes in the two samples is</p> $\hat{p}_C = \frac{\text{count of successes in both samples combined}}{\text{count of individuals in both samples combined}} = \frac{X_1 + X_2}{n_1 + n_2}$ <p>(p. 621)</p>	<p>proporción combinada de la muestra La proporción total de aciertos en las dos muestras es</p> $\hat{p}_C = \frac{\text{cuenta de aciertos en ambas muestras combinadas}}{\text{cuenta de individuos en ambas muestras combinadas}} = \frac{X_1 + X_2}{n_1 + n_2}$ <p>(pág. 621)</p>
<p>population In a statistical study, the entire group of individuals we want information about. (p. 210)</p>	<p>población En un estudio estadístico, la población es el grupo completo de individuos sobre el cual deseamos contar con información. (pág. 210)</p>
<p>population distribution Gives the values of the variable for all the individuals in the population. (p. 428)</p>	<p>distribución de la población Presenta los valores de la variable para todos los individuos en la población. (pág. 428)</p>
<p>population regression line Regression line $\mu_y = \alpha + \beta x$ based on the entire population of data. (p. 739)</p>	<p>línea de regresión de la población La línea de regresión $\mu_y = \alpha + \beta x$ basada en la totalidad de la población de datos. (pág. 739)</p>
<p>positive association When above-average values of one variable to accompany above-average values of the other and also of below-average values to occur together. (p. 148)</p>	<p>asociación positiva Cuando los valores por encima del promedio de una variable tienden a acompañar a los valores por encima del promedio de la otra, y los valores por debajo del promedio también tienden a suceder juntos. (pág. 148)</p>
<p>power The probability that a test will reject H_0 at a chosen significance level α when a specified alternative value of the parameter is true. The power of a test against any alternative is 1 minus the probability of a Type II error for that alternative; that is, power = $1 - \beta$. (p. 565)</p>	<p>poder La probabilidad de que una prueba rechace H_0 en un nivel de significancia dado cuando un valor alternativo especificado del parámetro es verdad. El poder de una prueba con respecto a cualquier alternativa es 1 menos la probabilidad de un error Tipo II para dicha alternativa; es decir, poder = $1 - \beta$. (pág. 565)</p>

<p>power model Relationship of the form $y = ax^p$. When experience or theory suggests that the relationship between two variables is described by a power model, you can transform the data to achieve linearity in two ways: (1) raise the values of the explanatory variable x to the p power and plot the points (x^p, y), or (2) take the pth root of the values of the response variable y and plot the points $(x, \sqrt[p]{y})$. If you don't know what power to use, taking the logarithms of both variables should produce a linear pattern. (p. 767)</p>	<p>modelo de poder Relación de la forma $y = ax^p$. Cuando la experiencia o una teoría sugiere que la relación entre dos variables la describe un modelo de poder, se pueden transformar los datos para que logren la linealidad de dos maneras: (1) elevar los valores de la variable explicativa x a la potencia p y trazar los puntos (x^p, y), o (2) tomar la raíz p de los valores de la variable de respuesta y y trazar los puntos $(x, \sqrt[p]{y})$. Si no se sabe qué potencia se ha de usar, se toman los logaritmos de ambas variables para producir un patrón lineal. (pág. 767)</p>
<p>predicted value \hat{y} (read “y hat”) is the predicted value of the response variable y for a given value of the explanatory variable x. (p. 166)</p>	<p>valor proyectado \hat{y} es el valor proyectado de la variable de respuesta y para un valor dado de la variable explicativa x. (pág. 166)</p>
<p>probability A number between 0 and 1 that describes the proportion of times an outcome of a chance process would occur in a very long series of repetitions. (p. 291)</p>	<p>probabilidad Cifra entre 0 y 1 que describe la proporción de veces que un resultado de un proceso aleatorio sucedería en una serie muy prolongada de repeticiones. (pág. 291)</p>
<p>probability distribution Gives the possible values of a random variable and their probabilities. (p. 348)</p>	<p>distribución de la probabilidad Presenta los valores posibles de una variable aleatoria y sus posibles probabilidades. (pág. 348)</p>
<p>probability model Description of some chance process that consists of two parts: a sample space S and a probability for each outcome. (p. 305)</p>	<p>modelo de probabilidad Descripción de un proceso de probabilidad que consta de dos partes: un espacio de muestra s y una probabilidad para cada resultado. (pág. 305)</p>
<p>Q</p>	
<p>quantitative variable Variable that takes numerical values for which it makes sense to find an average. (p. 3)</p>	<p>variable cuantitativa Toma valores numéricos para los cuales tiene sentido hallar un promedio. (pág. 3)</p>
<p>R</p>	
<p>random assignment Experimental design principle. Use chance to assign experimental units to treatments. Doing so helps create roughly equivalent groups of experimental units by balancing the effects of other variables among the treatment groups. (p. 241)</p>	<p>asignación aleatoria Principio de diseño experimental. Se usa el azar para asignar unidades experimentales a los tratamientos a fin de ayudar a formar grupos de unidades experimentales más o menos equivalentes al equilibrar los efectos de otras variables entre los grupos de tratamiento. (pág. 241)</p>
<p>random condition The data come from a well-designed random sample or randomized experiment. (p. 493)</p>	<p>condición aleatoria Los datos provienen de una muestra aleatoria bien diseñada o de un experimento aleatorizado. (pág. 493)</p>
<p>random sampling Using a chance process to determine which members of a population are included in the sample. (p. 214)</p>	<p>muestreo aleatorio Uso de un proceso de probabilidad para determinar cuáles miembros de la población se han de incluir en la muestra. (pág. 214)</p>
<p>random variable Variable that takes numerical values that describe the outcomes of some chance process. (p. 348)</p>	<p>variable aleatoria Toma valores numéricos que describen los resultados de algún proceso de azar y probabilidad. (pág. 348)</p>
<p>randomization distribution Distribution of a statistic (like $\hat{p}_1 - \hat{p}_2$ or $\bar{x}_1 - \bar{x}_2$) in repeated random assignments of experimental units to treatment groups assuming that the specific treatment received doesn't affect individual responses. When the conditions are met, usual inference procedures based on the sampling distribution of the statistic will be approximately correct. (p. 627)</p>	<p>distribución de la aleatoriedad La distribución de una estadística (como $\hat{p}_1 - \hat{p}_2$ o $\bar{x}_1 - \bar{x}_2$) en designaciones aleatorizadas reiteradas de unidades experimentales a grupos de tratamiento, asumiendo que el tratamiento específico no afecte las respuestas individuales. Cuando se cumplan las condiciones, los procedimientos de inferencia corrientes que se basan en la distribución del muestreo de la estadística serán aproximadamente correctos. (pág. 627)</p>
<p>randomized block design Experimental design begun by forming blocks consisting of individuals that are similar in some way that is important to the response. Random assignment of treatments is then carried out separately within each block. (p. 252)</p>	<p>diseño de bloques aleatorios Se comienza con la formación de bloques compuestos por individuos que son similares de alguna manera que sea importante para la respuesta. La asignación aleatoria de tratamientos luego se realiza separadamente dentro de cada bloque. (pág. 252)</p>

range The maximum value minus the minimum value for a set of quantitative data. (p. 54)	gama El valor máximo menos el valor mínimo de un conjunto de datos cuantitativos. (pág. 54)
regression line Line that describes how a response variable y changes as an explanatory variable x changes. We often use a regression line to predict the value of y for a given value of x . (p. 164)	línea de regresión Una línea que describe cómo una variable de respuesta y cambia a medida que cambia una variable explicativa x . A menudo se usa una línea de regresión para predecir el valor de y para un valor dado de x . (pág. 164)
reject H_0 If the observed result is too unlikely to occur just by chance when the null hypothesis is true, we can reject H_0 and say that there is convincing evidence for H_a . (p. 544)	rechazar H_0 Si es demasiado improbable que el resultado observado suceda por simple azar cuando la hipótesis nula es verdad, se puede rechazar H_0 y decir que existe evidencia convincente a favor de H_a . (pág. 544)
relative frequency table Table that shows the percents (relative frequencies) of observations in each category or class. (p. 8)	tabla de frecuencia relativa Permite apreciar los porcentajes (frecuencias relativas) de las observaciones en cada categoría o clase. (pág. 8)
replication Experimental design principle. Use enough experimental units in each group so that any differences in the effects of the treatments can be distinguished from chance differences between the groups. (p. 242)	replicación Principio de diseño experimental. Se usan suficientes unidades experimentales en cada grupo a fin de que todas las diferencias en los efectos de los tratamientos se puedan distinguir de las diferencias de azar entre los grupos. (pág. 242)
residual Difference between an observed value of the response variable and the value predicted by the regression line: $\text{residual} = \text{observed } y - \text{predicted } y = y - \hat{y}$ (p. 169)	residual La diferencia entre un valor observado de la variable de respuesta y y el valor proyectado por la línea de regresión. Es decir, $\text{residual} = \text{observada } y - \text{proyectada } y = y - \hat{y}.$ (pág. 169)
residual plot Scatterplot of the residuals against the explanatory variable. Residual plots help us assess whether a linear model is appropriate. (p. 173)	gráfico residual Gráfico de dispersión de los residuales en comparación con la variable explicativa. Los trazados residuales nos ayudan a evaluar si el modelo lineal es apropiado. (pág. 173)
resistant measure Statistic that is not affected very much by extreme observations. (p. 50)	medida resistente Estadística que no se ve muy afectada por observaciones extremas. (pág. 50)
response bias Systemic pattern of inaccurate answers. (p. 227)	sesgo de la respuesta Patrón sistémico de respuestas imprecisas. (pág. 227)
response variable Variable that measures an outcome of a study. (pp. 143, 236)	variable de respuesta Variable que mide un resultado de un estudio. (págs. 143, 236)
roundoff error Difference between the calculated approximation of a number and its exact mathematical value. (p. 8)	error de redondeo La diferencia entre la aproximación computada y su valor matemático exacto. (pág. 8)
S	
sample Subset of individuals in the population from which we actually collect data. (p. 210)	muestra Subconjunto de individuos en la población a partir de la cual de hecho se recogen datos. (pág. 210)
sample regression line (estimated regression line) Least-squares regression line $\hat{y} = a + bx$ computed from the sample data. (p. 739)	línea de regresión de la muestra (línea de regresión estimada) La línea de regresión de mínimos cuadrados $\hat{y} = a + bx$ computada a partir de los datos de la muestra. (pág. 739)
sample space S Set of all possible outcomes of a chance process. (p. 305)	espacios S de la muestra El conjunto de todos los resultados posibles de un proceso de probabilidad. (pág. 305)
sample survey Study that uses an organized plan to choose a sample that represents some specific population. We base conclusions about the population on data from the sample. (p. 210)	valoración de la muestra Estudio que usa un plan organizado para escoger una muestra que represente una población específica. Basamos las conclusiones sobre la población en datos tomados de la muestra. (pág. 210)

<p>sampling distribution The distribution of values taken by a statistic in all possible samples of the same size from the same population. (p. 427)</p>	<p>distribución del muestreo La distribución de valores tomados por la estadística en todas las muestras posibles del mismo tamaño tomadas de la misma población. (pág. 427)</p>
<p>sampling distribution of a sample mean \bar{x} Suppose that \bar{x} is the mean of an SRS of size n drawn from a large population with mean μ and standard deviation σ. Then</p> <ul style="list-style-type: none"> • The mean of the sampling distribution of \bar{x} is $\mu_{\bar{x}} = \mu$. • The standard deviation of the sampling distribution of \bar{x} is $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ <p>as long as the <i>10% condition</i> is satisfied: $n \leq \frac{1}{10}N$.</p> <ul style="list-style-type: none"> • If the population has a Normal distribution, then the sampling distribution of \bar{x} also has a Normal distribution. Otherwise, the central limit theorem tells us that the sampling distribution of \bar{x} will be approximately Normal in most cases when $n \geq 30$. (p. 452) 	<p>distribución de muestreo de la media de una muestra \bar{x} Supongamos que \bar{x} es la media de una muestra aleatoria sencilla de tamaño n tomada de una población grande con media μ y una desviación estándar σ. Entonces</p> <ul style="list-style-type: none"> • La media de la distribución del muestreo de \bar{x} es $\mu_{\bar{x}} = \mu$. • La desviación estándar de la distribución del muestreo de \bar{x} es $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ <p>siempre y cuando se cumpla con la <i>condición del 10%</i>: $n \leq \frac{1}{10}N$.</p> <ul style="list-style-type: none"> • Si la población tiene una distribución normal, entonces la distribución del muestreo \bar{x} también tiene una distribución normal. De no ser así, el teorema del límite central nos indica que en la mayoría de los casos, la distribución del muestreo \bar{x} será aproximadamente normal en la mayoría de casos cuando $n \geq 30$.
<p>sampling distribution of a sample proportion \hat{p} Choose an SRS of size n from a population of size N with proportion p of successes. Let \hat{p} be the sample proportion of successes. Then</p> <ul style="list-style-type: none"> • The mean of the sampling distribution of \hat{p} is $\mu_{\hat{p}} = p$. • The standard deviation of the sampling distribution of \hat{p} is $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ <p>as long as the <i>10% condition</i> is satisfied: $n \leq \frac{1}{10}N$.</p> <ul style="list-style-type: none"> • As n increases, the sampling distribution of \hat{p} becomes approximately Normal. Before you perform Normal calculations, check that the <i>Large Counts condition</i> is satisfied: $np \geq 10$ and $n(1-p) \geq 10$. (p. 444) 	<p>distribución del muestreo de una proporción de la muestra \hat{p} Se escoge una muestra aleatoria sencilla de tamaño n a partir de una población de tamaño N con proporción p de aciertos. Permita que \hat{p} sea la proporción de aciertos de la muestra. Entonces</p> <ul style="list-style-type: none"> • La media de la distribución del muestreo de \hat{p} es $\mu_{\hat{p}} = p$. • La desviación estándar de la distribución del muestreo de \hat{p} es $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ <p>siempre y cuando que se cumpla con la <i>condición del 10%</i>: $n \leq \frac{1}{10}N$.</p> <ul style="list-style-type: none"> • A medida que aumenta n, la distribución del muestreo de \hat{p} se torna aproximadamente normal. Antes de realizar los cálculos normales, se verifica que se haya cumplido con la <i>condición de cuentas grandes</i>: $np \geq 10$ y $n(1-p) \geq 10$. (pág. 444)
<p>sampling distribution of a slope Choose an SRS of n observations (x, y) from a population of size N with least-squares regression line $\mu_y = \alpha + \beta x$. Let b be the slope of the sample regression line. Then</p> <ul style="list-style-type: none"> • The mean of the sampling distribution of b is $\mu_b = \beta$. • The standard deviation of the sampling distribution of b is $\sigma_b = \frac{\sigma}{\sigma_x \sqrt{n}}$ <p>as long as the <i>10% condition</i> is satisfied: $n \leq \frac{1}{10}N$.</p> <ul style="list-style-type: none"> • The sampling distribution of b will be approximately Normal if the values of the response variable y follow a Normal distribution for each value of the explanatory variable x (the <i>Normal condition</i>). (p. 741) 	<p>distribución del muestreo de una pendiente Se escoge una muestra aleatoria sencilla de n observaciones (x, y) a partir de una población tamaño N con línea de regresión de mínimos cuadrados $\mu_y = \alpha + \beta x$. Se permite que b sea la pendiente de una línea de regresión de la muestra. Entonces</p> <ul style="list-style-type: none"> • La media de la distribución del muestreo de b es $\mu_b = \beta$. • La desviación estándar de la distribución del muestreo de b es $\sigma_b = \frac{\sigma}{\sigma_x \sqrt{n}}$ <p>siempre y cuando se cumpla con la <i>condición del 10%</i>: $n \leq \frac{1}{10}N$.</p> <ul style="list-style-type: none"> • La distribución del muestreo de b será aproximadamente normal si los valores de la variable de respuesta y siguen una distribución normal para cada valor de la variable explicativa x (la <i>condición normal</i>). (pág. 741)

<p>sampling distribution of $\hat{p}_1 - \hat{p}_2$ Choose an SRS of size n_1 from population 1 with proportion of successes p_1 and an independent SRS of size n_2 from population 2 with proportion of successes p_2.</p> <ul style="list-style-type: none"> • Shape: When $n_1 p_1$, $n_1(1-p_1)$, $n_2 p_2$, and $n_2(1-p_2)$ are all at least 10, the sampling distribution of $\hat{p}_1 - \hat{p}_2$ is approximately Normal. • Center: The mean of the sampling distribution is $p_1 - p_2$. • Spread: The standard deviation of the sampling distribution of $\hat{p}_1 - \hat{p}_2$ is $\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$ <p>as long as each sample is no more than 10% of its population. (p. 612)</p>	<p>distribución del muestreo $\hat{p}_1 - \hat{p}_2$ Se escoge una muestra aleatoria sencilla de tamaño n_1 a partir de una población 1 con una proporción de aciertos p_1 y una muestra aleatoria sencilla independiente de tamaño n_2 a partir de la población 2 con una proporción de aciertos p_2.</p> <ul style="list-style-type: none"> • Forma: Cuando $n_1 p_1$, $n_1(1-p_1)$, $n_2 p_2$, and $n_2(1-p_2)$ son por lo menos 10, la distribución del muestreo de $\hat{p}_1 - \hat{p}_2$ es aproximadamente normal. • Centro: La media de la distribución del muestreo es $p_1 - p_2$. • Amplitud: La desviación estándar de la distribución del muestreo de $\hat{p}_1 - \hat{p}_2$ es $\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$ <p>siempre y cuando cada muestra es de no más del 10% de su población. (pág. 612)</p>
<p>sampling distribution of $\bar{x}_1 - \bar{x}_2$ Choose an SRS of size n_1 from population 1 with mean μ_1 and standard deviation σ_1 and an independent SRS of size n_2 from population 2 with mean μ_2 and standard deviation σ_2.</p> <ul style="list-style-type: none"> • Shape: When the population distributions are Normal, the sampling distribution of $\bar{x}_1 - \bar{x}_2$ is Normal. In other cases, the sampling distribution of $\bar{x}_1 - \bar{x}_2$ will be approximately Normal if the sample sizes are large enough ($n_1 \geq 30$ and $n_2 \geq 30$). • Center: The mean of the sampling distribution is $\mu_1 - \mu_2$. • Spread: The standard deviation of the sampling distribution of $\bar{x}_1 - \bar{x}_2$ is $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ <p>as long as each sample is no more than 10% of its population. (p. 638)</p>	<p>distribución de muestreo de $\bar{x}_1 - \bar{x}_2$ Se escoge una muestra aleatoria sencilla de tamaño a partir de una población 1 con una media μ_1 y una desviación estándar σ_1 y una muestra aleatoria sencilla independiente de tamaño n_2 a partir de una población 2 con una media μ_2 y una desviación estándar σ_2.</p> <ul style="list-style-type: none"> • Forma: Cuando las distribuciones de la población son normales, la distribución del muestreo de $\bar{x}_1 - \bar{x}_2$ es normal. En otros casos, la distribución del muestreo de $\bar{x}_1 - \bar{x}_2$ será aproximadamente normal si los tamaños de las muestras son suficientemente grandes ($n_1 \geq 30$ y $n_2 \geq 30$). • Centro: La media de la distribución del muestreo es $\mu_1 - \mu_2$. • Amplitud: La desviación estándar de la distribución del muestreo de $\bar{x}_1 - \bar{x}_2$ es $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ <p>siempre y cuando cada muestra es de no más del 10% de su población. (pág. 638)</p>
<p>sampling variability The value of a statistic varies in repeated random sampling. (p. 425)</p>	<p>variabilidad del muestreo El valor de una estadística varía en muestras aleatorias reiteradas. (pág. 425)</p>
<p>scatterplot Plot that shows the relationship between two quantitative variables measured on the same individuals. The values of one variable appear on the horizontal axis, and the values of the other variable appear on the vertical axis. Each individual in the data appears as a point in the graph. (p. 145)</p>	<p>gráfico de dispersión Permite apreciar la relación entre dos variables cuantitativas midiendo los mismos individuos. Los valores de una variable figuran en el eje horizontal, y los valores de la otra variable figuran en el eje vertical. Cada individuo en los datos figura como un punto en el gráfico. (pág. 145)</p>
<p>segmented bar graph Graph used to compare the distribution of a categorical variable in each of several groups. For each group, there is a single bar with “segments” that correspond to the different values of the categorical variable. The height of each segment is determined by the percent of individuals in the group with that value. Each bar has a total height of 100%. (p. 17)</p>	<p>gráfico de barras segmentado Se usa para comparar la distribución de una variable categorizada en cada uno de varios grupos. Para cada grupo, hay una sola barra que tiene “segmentos” que corresponden a los diferentes valores de la variable categorizada. La altura de cada segmento la determina el porcentaje de individuos en el grupo que tengan ese valor. Cada barra tiene una altura total del 100%. (pág. 17)</p>

<p>side-by-side bar graph Graph used to compare the distribution of a categorical variable in each of several groups. For each value of the categorical variable, there is a bar corresponding to each group. The height of each bar is determined by the count or percent of individuals in the group with that value. (p. 17)</p>	<p>gráfico de barras contiguas Se usa para comparar la distribución de una variable categorizada en cada uno de varios grupos. Para cada valor de la variable categorizada, hay una barra que corresponde a cada grupo. La altura de la barra la determina el conteo o el porcentaje de individuos en el grupo que tengan ese valor. (pág. 17)</p>
<p>significance level Fixed value α that we use as a cutoff for deciding whether an observed result is too unlikely to happen by chance alone when the null hypothesis is true. The significance level gives the probability of a Type I error. (p. 545)</p>	<p>nivel de significancia Valor fijo α que se usa como punto de corte para decidir si un resultado observado es demasiado improbable para suceder solo al azar cuando la hipótesis nula es verdad. El nivel de significancia genera la probabilidad de un error Tipo I. (pág. 545)</p>
<p>significance test Procedure for using observed data to decide between two competing claims (also called <i>hypotheses</i>). The claims are often statements about a parameter. (p. 539)</p>	<p>prueba de significancia Procedimiento en el que se usan datos observados para decidir entre dos opciones que compiten entre sí (también se les dice <i>hipótesis</i>). Las opciones a menudo son enunciados acerca de un parámetro. (pág. 539)</p>
<p>simple random sample (SRS) Sample chosen in such a way that every group of n individuals in the population has an equal chance to be selected as the sample. (p. 214)</p>	<p>muestra aleatoria sencilla Muestra tomada de tal manera que cada grupo de n individuos en la población tenga la misma oportunidad de ser escogido como la muestra. (pág. 214)</p>
<p>simulation Imitation of chance behavior, based on a model that accurately reflects the situation. (p. 295)</p>	<p>simulación Imitación de conducta de azar, basada en un modelo que refleja la situación con precisión. (pág. 295)</p>
<p>single-blind An experiment in which either the subjects or those who interact with them and measure the response variable, but not both, know which treatment a subject received. (p. 248)</p>	<p>ciego sencillo Experimento en el que ya sea los sujetos o bien aquellos que interactúan con los sujetos y miden la variable de respuesta, pero no ambos, saben cuál fue el tratamiento que recibió un sujeto. (pág. 248)</p>
<p>skewness A distribution is <i>skewed to the right</i> if the right side of the graph (containing the half of the observations with larger values) is much longer than the left side. It is <i>skewed to the left</i> if the left side of the graph is much longer than the right side. (p. 27)</p>	<p>asimetría Distribución que está <i>sesgada hacia la derecha</i> si la derecha del gráfico (que contiene la mitad de las observaciones con valores más grandes) es mucho más larga que el lado izquierdo. Está <i>sesgada hacia la izquierda</i> si el lado izquierdo del gráfico es mucho más largo que el lado derecho. (pág. 27)</p>
<p>slope Suppose that y is a response variable (plotted on the vertical axis) and x is an explanatory variable (plotted on the horizontal axis). A regression line relating y to x has an equation of the form $\hat{y} = a + bx$. In this equation, b is the slope, the amount by which y is predicted to change when x increases by one unit. (p. 166)</p>	<p>pendiente Supongamos que y es una variable de respuesta (trazada en el eje vertical) y x es una variable explicativa (trazada en el eje horizontal). Una línea de regresión que relacione y con x tiene una ecuación en la forma de $\hat{y} = a + bx$. En esta ecuación, b es la pendiente, la cantidad mediante la cual se predice que y cambiará cuando x tenga un aumento de una unidad. (pág. 166)</p>
<p>splitting stems Method for spreading out a stemplot that has too few stems. (p. 32)</p>	<p>tallos separados Método para separar un gráfico de tallos que tiene una carencia de tallos. (pág. 32)</p>
<p>standard deviation s_x Statistic that measures the typical distance of the values in a distribution from the mean. It is calculated by finding an “average” of the squared distances and then taking the square root. In symbols,</p> $s_x = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}$ <p>(p. 61)</p>	<p>desviación estándar s_x Estadística que mide la distancia típica de los valores en una distribución a partir de la media. Se computa hallando un “promedio” de las distancias al cuadrado a las que luego se les computa la raíz cuadrada. En símbolos se representa</p> $s_x = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}$ <p>(pág. 61)</p>
<p>standard deviation of a random variable Square root of the variance of a random variable σ_x^2. The standard deviation measures the typical distance of the values in a distribution from their mean. In symbols,</p> $\sigma_x = \sqrt{\sum (x_i - \mu_x)^2 p_i}$ <p>(p. 353)</p>	<p>desviación estándar de una variable aleatoria La raíz cuadrada de la variación de una variable aleatoria σ_x^2. La desviación estándar mide la distancia típica de los valores en una distribución a partir de su media. En símbolos se represent</p> $\sigma_x = \sqrt{\sum (x_i - \mu_x)^2 p_i}$ <p>(pág. 353)</p>

<p>standard deviation of the residuals (s) If we use a least-squares line to predict the values of a response variable y from an explanatory variable x, the standard deviation of the residuals (s) is given by</p> $s = \sqrt{\frac{\sum \text{residuals}^2}{n - 2}} = \sqrt{\frac{\sum (y_i - \hat{y})^2}{n - 2}}$ <p>This value gives the approximate size of a “typical” prediction error (residual). (p. 177)</p>	<p>desviación estándar de las residuales (s) Si se hace uso de la línea de cuadrados mínimos para predecir los valores de una variable de respuesta y a partir de una variable explicativa x, la desviación estándar de las residuales (s) la da</p> $s = \sqrt{\frac{\sum \text{residuales}^2}{n - 2}} = \sqrt{\frac{\sum (y_i - \hat{y})^2}{n - 2}}$ <p>Este valor ofrece un tamaño aproximado de un error de predicción “típico” (residual). (pág. 177)</p>
<p>standard error When the standard deviation of a statistic is estimated from data, the result is the standard error of the statistic. (p. 497)</p>	<p>error estándar Cuando se computa la desviación estándar de una estadística a partir de datos, el resultado es el error estándar de la estadística. (pág. 497)</p>
<p>standard error of $\hat{p}_1 - \hat{p}_2$ Estimated standard deviation of the statistic $\hat{p}_1 - \hat{p}_2$, given by</p> $\sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}$ <p>(p. 616)</p>	<p>error estándar de $\hat{p}_1 - \hat{p}_2$ La desviación estándar coputada de la estadística $\hat{p}_1 - \hat{p}_2$, dada por</p> $\sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}$ <p>(pág. 616)</p>
<p>standard error of the sample mean $\frac{s_x}{\sqrt{n}}$ where s_x is the sample standard deviation. It describes how far \bar{x} will typically be from μ in repeated SRSs of size n. (p. 518)</p>	<p>error estándar de la media de la muestra $\frac{s_x}{\sqrt{n}}$ en el que s_x es la desviación estándar de la muestra. Describe la distancia típica en la que \bar{x} se separa de μ en muestras aleatorias sencillas reiteradas de tamaño n. (pág. 518)</p>
<p>standard error of the sample proportion $\sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$ where \hat{p} is the sample proportion. It describes how far \hat{p} will typically be from p in repeated SRSs of size n. (p. 496)</p>	<p>error estándar de la proporción de la muestra $\sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$ en la que \hat{p} es la proporción de la muestra. Se describe la distancia típica en la que \hat{p} se separa de p en muestras aleatorias sencillas reiteradas de tamaño n. (pág. 496)</p>
<p>standard error of the slope Formula used to estimate the spread of the sampling distribution of b:</p> $SE_b = \frac{s}{s_x \sqrt{n - 1}}$ <p>(p. 747)</p>	<p>error estándar de la pendiente Se usa para computar la amplitud de la distribución del muestreo de b:</p> $SE_b = \frac{s}{s_x \sqrt{n - 1}}$ <p>(pág. 747)</p>
<p>standard error of $\bar{x}_1 - \bar{x}_2$ Estimated standard deviation of the statistic $\bar{x}_1 - \bar{x}_2$, given by</p> $\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ <p>(p. 640)</p>	<p>error estándar de $\bar{x}_1 - \bar{x}_2$ La desviación estándar computada de la estadística $\bar{x}_1 - \bar{x}_2$, dada por</p> $\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ <p>(pág. 640)</p>
<p>standard Normal distribution Normal distribution with mean 0 and standard deviation 1. (p. 113)</p>	<p>distribución normal estándar La distribución normal con media de 0 y desviación estándar 1. (pág. 113)</p>
<p>standard Normal table (Table A) Table of areas under the standard Normal curve. The table entry for each value z is the area under the curve to the left of z. (p. 113)</p>	<p>tabla normal estándar (Tabla A) Tabla de áreas debajo de la curva normal estándar. La entrada a la tabla de cada valor z es el área debajo de la curva a la izquierda de z. (pág. 113)</p>
<p>standardized score (z-score) If x is an observation from a distribution that has known mean and standard deviation, the standardized value of x is $z = \frac{x - \text{mean}}{\text{standard deviation}}$. A standardized value is often called a z-score. (p. 90)</p>	<p>puntuación estandarizada Si x es una observación a partir de una distribución que tiene una media y desviación estándar conocidas, el valor estandarizado de x es $z = \frac{x - \text{media}}{\text{desviación estándar}}$. Al valor estandarizado a menudo se le dice puntuación z. (pág. 90)</p>

<p>statistic Number that describes some characteristic of a sample. (p. 424)</p>	<p>estadística Número que describe alguna característica de una muestra. (pág. 424)</p>
<p>statistically significant (1) Observed effect so large that it would rarely occur by chance. (p. 249) (2) If the P-value is smaller than alpha, we say that the results of a statistical study are significant at level α. In that case, we reject the null hypothesis H_0 and conclude that there is convincing evidence in favor of the alternative hypothesis H_a. (p. 545)</p>	<p>estadísticamente significativo (1) Efecto observado que es tan grande que raramente sucedería producto del azar. (pág. 249) (2) Si el valor P es menor que alfa, se dice que los resultados de un estudio estadístico son significativos al nivel α. En tal caso, se rechaza la hipótesis nula H_0 y se concluye que hay evidencia convincente a favor de la hipótesis H_a alternativa. (pág. 545)</p>
<p>stemplot (<i>also called stem-and-leaf plot</i>) Simple graphical display for fairly small data sets that gives a quick picture of the shape of a distribution while including the actual numerical values in the graph. Each observation is separated into a stem, consisting of all but the final digit, and a leaf, the final digit. The stems are arranged in a vertical column with the smallest at the top. Each leaf is written in the row to the right of its stem, with the leaves arranged in increasing order out from the stem. (p. 31)</p>	<p>gráfico de tallos al que también se le dice gráfico de tallos y hojas Representación gráfica sencilla de conjuntos de datos relativamente pequeños que dan una imagen rápida de la forma de una distribución, al tiempo que incluyen los valores numéricos mismos en el gráfico. Cada observación se separa en un tallo compuesto de todos menos el último dígito, y una hoja, que es ese último dígito. Los tallos se disponen en una columna vertical en la cual el valor más pequeño está arriba. Cada hoja se escribe en el renglón a la derecha del tallo, con las hojas dispuestas en orden ascendente comenzando a partir del tallo. (pág. 31)</p>
<p>stratified random sample Sample obtained by classifying the population into groups of similar individuals, called <i>strata</i>, then choosing a separate SRS in each stratum and combining these SRSs to form the sample. (p. 219)</p>	<p>muestra aleatoria estratificada Muestra que se obtiene clasificando la población en grupos de individuos parecidos, llamados <i>estratos</i>. Luego, se escoge una muestra aleatoria sencilla separada en cada estrato y se combinan estas muestras aleatorias sencillas para conformar la muestra. (pág. 219)</p>
<p>subjects Experimental units that are human beings. (p. 237)</p>	<p>sujetos Unidades experimentales que son seres humanos. (pág. 237)</p>
<p>symmetric A graph in which the right and left sides are approximately mirror images of each other. (p. 27)</p>	<p>simétrico Si los lados derecho e izquierdo de un gráfico son reflejos aproximados uno del otro, se dice que son simétricos. (pág. 27)</p>
T	
<p>t distribution Draw an SRS of size n from a large population that has a Normal distribution with mean μ and standard deviation σ. The statistic</p> $t = \frac{\bar{x} - \mu}{\frac{s_x}{\sqrt{n}}}$ <p>has the t distribution with <i>degrees of freedom</i> $df = n - 1$. This statistic will have approximately a t_{n-1} distribution if the sample size is large enough. (p. 512)</p>	<p>distribución t Se grafica una muestra aleatoria sencilla de tamaño n a partir de una población grande que tiene una distribución normal con media de μ y desviación estándar σ. La estadística</p> $t = \frac{\bar{x} - \mu}{\frac{s_x}{\sqrt{n}}}$ <p>tiene la distribución t con <i>grados de libertad</i> $df = n - 1$. Esta estadística tendrá una distribución aproximada t_{n-1} si el tamaño de la muestra es suficientemente grande. (pág. 512)</p>
<p>t interval for the slope β When the conditions for regression inference are met, a $C\%$ confidence interval for the slope β of the population (true) regression line is $b \pm t^*SE_b$. In this formula, the standard error of the slope is</p> $SE_b = \frac{s}{s_x\sqrt{n-1}}$ <p>and t^* is the critical value for the t distribution with $df = n - 2$ having $C\%$ of its area between $-t^*$ and t^*. (p. 548)</p>	<p>intervalo t para la pendiente β (pág. 748) Cuando se cumple con las condiciones para lograr una inferencia de regresión, el intervalo de confianza $C\%$ para la pendiente β de la línea de regresión de la población (verdadera) es $b \pm t^*SE_b$. En esta fórmula, el error estándar de la pendiente es</p> $SE_b = \frac{s}{s_x\sqrt{n-1}}$ <p>y t^* es el valor crítico para la distribución t y $df = n - 2$ tiene el $C\%$ de su área entre $-t^*$ y t^*. (pág. 548)</p>

<p>t test for the slope Suppose the conditions for inference are met. To test the hypothesis $H_0: \beta = \beta_0$, compute the test statistic</p> $t = \frac{b - \beta_0}{SE_b}$ <p>Find the P-value by calculating the probability of getting a t statistic this large or larger in the direction specified by the alternative hypothesis H_a. Use the t distribution with $df = n - 2$. (p. 753)</p>	<p>prueba t para la pendiente Supongamos que se cumple con todas las condiciones para la inferencia. Para poner a prueba la hipótesis $H_0: \beta = \beta_0$, se computa la estadística de prueba</p> $t = \frac{b - \beta_0}{SE_b}$ <p>Se halla el valor P computando la probabilidad de obtener una estadística t de este tamaño o más grande en el sentido especificado por la hipótesis alternativa H_a. Se usa la distribución t con $df = n - 2$. (pág. 753)</p>
<p>test statistic Calculation that measures how far a sample statistic diverges from what we would expect if the null hypothesis H_0 were true, in standardized units. That is,</p> $\text{test statistic} = \frac{\text{statistic} - \text{parameter}}{\text{standard deviation of statistic}}$ <p>(p. 556)</p>	<p>estadística de prueba Mide la divergencia entre la estadística de muestra y lo que esperaríamos si la hipótesis nula H_0 fuera verdad, expresado en unidades estandarizadas. Es decir,</p> $\text{estadística de prueba} = \frac{\text{estadística} - \text{parametro}}{\text{desviación estándar de la estadística}}$ <p>(pág. 556)</p>
<p>third quartile Q_3 In a data set in which the observations are ordered from lowest to highest, the median of the observations whose position is to the right of the median. (p. 54)</p>	<p>tercer cuartil Q_3 Si las observaciones en un conjunto de datos se organizan de la más baja a la más alta, el tercer cuartil Q_3 es la media de las observaciones cuya posición está a la derecha de la media. (pág. 54)</p>
<p>transforming Applying a function such as the logarithm or square root to a quantitative variable. (p. 767)</p>	<p>transformación La aplicación de una función tal como el logaritmo o la raíz cuadrada a una variable cuantitativa se denomina transformación del dato. (pág. 767)</p>
<p>treatment Specific condition applied to the individuals in an experiment. If an experiment has several explanatory variables, a treatment is a combination of specific values of these variables. (p. 237)</p>	<p>tratamiento Una condición específica que se les aplica a los individuos en un experimento. Si un experimento tiene varias variables explicativas, el tratamiento es una combinación de los valores específicos de estas variables. (pág. 237)</p>
<p>tree diagram Diagram used to display the sample space for a chance process that involves a sequence of outcomes. (p. 322)</p>	<p>diagrama de árbol Se usa para mostrar el espacio de muestra para un proceso de probabilidad que produce una secuencia de resultados. (pág. 322)</p>
<p>two-sample t interval for a difference between two means When the Random, 10%, and Normal/Large Sample conditions are met, an approximate $C\%$ confidence interval for $\mu_1 - \mu_2$ is</p> $(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ <p>Here t^* is the critical value with area $C\%$ between $-t^*$ and t^* for the t distribution with degrees of freedom from either Option 1 (technology) or Option 2 (the smaller of $n_1 - 1$ and $n_2 - 1$). (p. 641)</p>	<p>intervalo t de dos muestras para obtener la diferencia entre dos medias Cuando se cumple con las condiciones aleatoria, del 10% y de muestra normal/grande, un intervalo de confianza $C\%$ aproximado para $\mu_1 - \mu_2$ es</p> $(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ <p>en el que t^* es el valor crítico con área $C\%$ entre $-t^*$ y t^* para la distribución t con grados de libertad ya sea la Opción 1 (tecnología) o la Opción 2 (el valor menor de $n_1 - 1$ y $n_2 - 1$). (pág. 641)</p>
<p>two-sample t statistic When we standardize the estimate $\bar{x}_1 - \bar{x}_2$, the result is the two-sample t statistic</p> $t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ <p>The statistic t has the same interpretation as any z or t statistic: it says how far $\bar{x}_1 - \bar{x}_2$ is from its mean in standard deviation units. (p. 640)</p>	<p>estadística t de dos muestras Cuando se estandariza el estimado $\bar{x}_1 - \bar{x}_2$, el resultado es una estadística t de dos muestras</p> $t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ <p>La estadística t tiene la misma interpretación de cualquier estadística z o t: indica la distancia de $\bar{x}_1 - \bar{x}_2$ a partir de su media en unidades de desviación estándar. (pág. 640)</p>

<p>two-sample t test for the difference between two means Suppose the Random, 10%, and Normal/Large Sample conditions are met. To test the hypothesis $H_0: \mu_1 - \mu_2 =$ hypothesized value, compute the two-sample t statistic</p> $t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ <p>Find the P-value by calculating the probability of getting a t statistic this large or larger in the direction specified by the alternative hypothesis H_a. Use the t distribution with degrees of freedom approximated by technology or the smaller of $n_1 - 1$ and $n_2 - 1$. (p. 645)</p>	<p>prueba t de dos muestras para obtener la diferencia entre dos medias Supongamos que se ha cumplido con las condiciones aleatoria, del 10% y de muestra normal/grande. Para someter a prueba la hipótesis $H_0: \mu_1 - \mu_2 =$ valor hipotético, se computa la estadística t de dos muestras</p> $t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ <p>Halla el valor P computando la probabilidad de obtener una estadística t de este tamaño o más grande en el sentido especificado por la hipótesis H_a alternativa. Se usa la distribución t con grados de libertad aproximados por la tecnología o el valor menor de $n_1 - 1$ y $n_2 - 1$. (pág. 645)</p>
<p>two-sample z interval for a difference between two proportions When the Random, 10%, and Large Counts conditions are met, an approximate $C\%$ confidence interval for $p_1 - p_2$ is</p> $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}$ <p>where z^* is the critical value for the standard Normal curve with $C\%$ of its area between $-z^*$ and z^*. (p. 617)</p>	<p>intervalo z de dos muestras para obtener la diferencia entre dos proporciones Cuando se cumplen las condiciones de aleatorio, del 10% y de muestras grandes, el intervalo de confianza $C\%$ aproximado para $p_1 - p_2$ es</p> $(\hat{p}_1 - \hat{p}_2) \pm z^* \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}$ <p>en el que z^* es el valor crítico para la curva estándar normal con $C\%$ de su área entre $-z^*$ y z^*. (pág. 617)</p>
<p>two-sample z test for the difference between two proportions Suppose the Random, 10%, and Large Counts conditions are met. To test the hypothesis $H_0: p_1 - p_2 = 0$, first find the pooled proportion \hat{p}_C of successes in both samples combined. Then compute the z statistic</p> $z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\hat{p}_C(1 - \hat{p}_C)}{n_1} + \frac{\hat{p}_C(1 - \hat{p}_C)}{n_2}}}$ <p>Find the P-value by calculating the probability of getting a z statistic this large or larger in the direction specified by the alternative hypothesis H_a. (p. 622)</p>	<p>prueba z de dos muestras para obtener la diferencia entre dos proporciones Supongamos que se cumplen las condiciones aleatoria, del 10% y de muestras grandes. Para someter a prueba la hipótesis $H_0: p_1 - p_2 = 0$, primero se halla la proporción combinada \hat{p}_C de aciertos en ambas muestras combinadas. Luego se computa la estadística z</p> $z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\hat{p}_C(1 - \hat{p}_C)}{n_1} + \frac{\hat{p}_C(1 - \hat{p}_C)}{n_2}}}$ <p>Se halla el valor P computando la probabilidad de obtener una estadística z de este tamaño o más grande en el sentido especificado por la hipótesis H_a alternativa. (pág. 622)</p>
<p>two-sided alternative hypothesis The alternative hypothesis is two-sided if it states that the parameter is different from the null value (it could be either smaller or larger). (p. 541)</p>	<p>hipótesis alternativa bilateral La hipótesis alternativa es bilateral si indica que el parámetro es diferente del valor nulo (podría ser más pequeño o más grande). (pág. 541)</p>
<p>two-way table Table of counts that organizes data about two categorical variables. (p. 12)</p>	<p>tabla de doble vía Una tabla de doble vía de cuentas organiza los datos acerca de dos variables categorizadas. (pág. 12)</p>
<p>Type I error Occurs if we reject H_0 when H_0 is true. (p. 547)</p>	<p>error Tipo I Sucede si rechazamos H_0 cuando H_0 es verdad. (pág. 547)</p>
<p>Type II error Occurs if we fail to reject H_0 when H_a is true. (p. 547)</p>	<p>error Tipo II Sucede si no rechazamos H_0 cuando H_a es verdad. (pág. 547)</p>
<p>U</p>	
<p>unbiased estimator A statistic used for estimating a parameter is unbiased if the mean of its sampling distribution is equal to the true value of the parameter being estimated. (p. 429)</p>	<p>estimador sin sesgo La estadística que se usa para computar un parámetro es un estimador sin sesgo si la media de distribución de su muestreo equivale al valor verdadero del parámetro que se está computando. (pág. 429)</p>

<p>undercoverage Occurs when some members of the population cannot be chosen in a sample. (p. 225)</p>	<p>subcobertura Sucede cuando algunos miembros de la población no pueden ser escogidos para la muestra. (pág. 225)</p>
<p>unimodal A graph of quantitative data with a single peak. (p. 28)</p>	<p>unimodal Gráfico de datos cuantitativos con un solo pico. (pág. 28)</p>
<p>union The union of events A and B, denoted by $A \cup B$, consists of all outcomes in A or B or both. (p. 311)</p>	<p>unión La unión de los eventos A y B, denotados por $A \cup B$, consiste en todos los resultados en A o B, o ambos. (pág. 311)</p>
V	
<p>variability of a statistic Spread of a statistic's sampling distribution. Statistics from larger samples have less variability. (p. 433)</p>	<p>variabilidad de una estadística La variabilidad de una estadística se describe por la amplitud de la distribución de su muestreo. Las estadísticas de muestras más grandes tienen menos variabilidad. (pág. 433)</p>
<p>variable Any characteristic of an individual. A variable can take different values for different individuals. (p. 2)</p>	<p>variable Toda característica de un individuo. Una variable puede tener diferentes valores para diferentes individuos. (pág. 2)</p>
<p>variance of a random variable σ_x^2 Weighted average of the squared deviations of the values of the variable from their mean. In symbols,</p> $\sigma_x^2 = \sum (x_i - \mu_x)^2 p_i$ <p>(p. 352)</p>	<p>variación de una variable aleatoria σ_x^2 Promedio sopesado de las desviaciones cuadráticas de los valores de la variable a partir de su media. En símbolos se expresa</p> $\sigma_x^2 = \sum (x_i - \mu_x)^2 p_i$ <p>(pág. 352)</p>
<p>variance s_x^2 "Average" squared deviation of the observations in a data set from their mean. In symbols,</p> $s_x^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n - 1} = \frac{1}{n - 1} \sum (x_i - \bar{x})^2$ <p>(p. 61)</p>	<p>variación s_x^2 La desviación cuadrática "promedio" de las observaciones en el conjunto de datos según su separación de la media. En símbolos se expresa</p> $s_x^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n - 1} = \frac{1}{n - 1} \sum (x_i - \bar{x})^2$ <p>(pág. 61)</p>
<p>Venn diagrams Used to display the sample space for a chance process. Venn diagrams can also be used to find probabilities involving events A and B. (p. 314)</p>	<p>diagramas Venn Se usan para mostrar el espacio de muestras para un proceso de probabilidad. Los diagramas Venn también se pueden usar para hallar las probabilidades que existen para los eventos A y B. (pág. 314)</p>
<p>voluntary response sample People decide whether to join a sample by responding to a general invitation. (p. 212)</p>	<p>muestra de respuesta voluntaria La gente que decide vincularse a una muestra mediante una respuesta a una invitación generalizada. (pág. 212)</p>
W	
<p>wording of questions Most important influence on the answers given in a survey. Confusing or leading questions can introduce strong bias, and changes in wording can greatly change a survey's outcome. Even the order in which questions are asked matters. (p. 227)</p>	<p>terminología de las preguntas La influencia más importante sobre las respuestas que se dan en un sondeo. Toda pregunta confusa, capciosa o que sugiera una respuesta puede introducirle al proceso un sesgo marcado, y los cambios en la terminología pueden modificar de manera marcada los resultados de tal sondeo. Incluso los resultados se pueden ver afectados por el orden en que se hacen las preguntas. (pág. 227)</p>
Y	
<p>y intercept Suppose that y is a response variable (plotted on the vertical axis of a graph) and x is an explanatory variable (plotted on the horizontal axis). A regression line relating y to x has an equation of the form $\hat{y} = a + bx$. In this equation, the number a is the y intercept, the predicted value of y when $x = 0$. (p. 166)</p>	<p>interceptación y Supongamos que y es una variable de respuesta (graficada en el eje vertical) y x es una variable explicativa (graficada en el eje horizontal). La línea de regresión que relaciona y con x tiene una ecuación $\hat{y} = a + bx$. Según esta ecuación, el número a es la interceptación y, el valor proyectado de y cuando $x = 0$. (pág. 166)</p>

This page intentionally left blank