CAPÍTULO

2

ANÁLISIS ESTADÍSTICO: MANEJO Y CONTROL DEL ERROR

Introducción

Así sea realizada para la investigación científica, la mercadotecnia de un producto, un pronóstico meteorológico o una simple apuesta, la predicción del futuro es un pasatiempo común. Los científicos realizan predicciones empíricas para probar la exactitud de sus ideas. Por ejemplo, ¿cuál es la probabilidad de que sea víctima de un delito en su área de trabajo? Madriz (1996) encontró tres factores de predicción basados en la idea de que el riesgo de ser víctima de un delito puede reducirse por medio del estudio cuidadoso de actividades rutinarias. Un primer factor de riesgo es la exposición —la vulnerabilidad circunstancial— como trabajar solo por la noche en una tienda. Un segundo factor es la proximidad a delincuentes potenciales, como trabajar en una tienda ubicada en un área de alto índice delictivo. Un tercero es el atractivo del objetivo, es decir, la deseabilidad de la propiedad de una víctima, como tener grandes cantidades de dinero disponible.

Si el dueño de una tienda pusiera a sus empleados en riesgo innecesario, un robo o asesinato no serían un suceso aleatorio o una equivocación; sería un error. En el capítulo 1 notamos que los **errores** son *grados conocidos de imprecisión*. Conocer la relación entre las circunstancias y la probabilidad de un robo permite realizar mediciones preventivas que reduzcan las oportunidades calculadas para que los "errores" ocurran. Las mediciones para la reducción del riesgo podrían incluir tener al menos dos empleados presentes, cerrar a las 11:00 P.M., ubicarse en un área de alto tráfico, manejar pequeñas cantidades de dinero disponible e instalar sistemas de alarma. La reducción del error depende de la comprensión de las relaciones predictivas entre variables.

Como brevemente notamos en el capítulo 1, la estadística trata sobre la comprensión exacta y el **control del error estadístico**, los *grados conocidos de imprecisión en los procedimientos utilizados para reunir y procesar información*. Los errores no son

equivocaciones. Los errores son cantidades conocidas de imprecisión que pueden calcularse y reducirse por una selección cuidadosa e informada de diseños de muestreo, instrumentos de medición y fórmulas estadísticas.

> **Error estadístico**
>
> Grados conocidos de imprecisión en los procedimientos utilizados para reunir y procesar información.

El análisis estadístico comúnmente implica un muestreo: analizar sólo una pequeña parte del grupo que se estudia. Por ejemplo, para aprender acerca de todas las tiendas pequeñas, podríamos estudiar una muestra de 20 tiendas. ¿Pueden los datos de la muestra de 20 tiendas revelar con precisión cómo funcionan todas ellas? La investigación también involucra la observación y la medición. ¿Podemos asumir que nuestras mediciones son completamente exactas? El muestreo y la medición son dos fuentes potenciales de error al obtener conclusiones en la investigación. El **error de muestreo** representa la *inexactitud en las predicciones sobre una población que resulta del hecho de que no observamos a todos los sujetos en la población.* El **error de medición** es la *inexactitud en la investigación que se deriva de instrumentos de medición imprecisos, de las dificultades en la clasificación de las observaciones y de la necesidad de redondear los números.* Después de discutir cada uno de estos tipos de error, mostraremos cómo están relacionados.

Control del error de muestreo

Analizar significa descomponer algo y examinarlo con detalle de manera organizada. Al realizar trabajo estadístico, analizamos grupos de personas, objetos o eventos y medimos variables para obtener promedios, tendencias o porcentajes. La *medición de una sola persona*, por ejemplo, registrar como 19 años la edad de María López, no proporciona una estadística; simplemente es una **observación**. Sin embargo, determinar que la edad promedio en una clase de 30 estudiantes es de 19.5 años es calcular un estadístico con base en un conjunto de observaciones. El campo de la estadística implica **cálculos resumidos** de muchas observaciones; esto es, la *adición de un grupo de mediciones.* Nuestros intereses se enfocan en observar muchos casos, recabar información precisa sobre ellos y hacer declaraciones concisas sobre el grupo, no sobre los individuos.

El grupo de sujetos que observamos a menudo es bastante pequeño. Nuestro propósito es estudiar el número pequeño de sujetos para obtener conclusiones sobre la población más grande a la cual esos sujetos pertenecen. Estudiar cada caso de un fenómeno es impráctico, costoso e innecesario. Por ejemplo, no tenemos que encuestar a cada votante probable para determinar el apoyo al candidato A. En cambio, podemos encuestar una muestra representativa de votantes probables,

quizás 500. Este grupo más pequeño se llama muestra, mientras el grupo más grande, completo, al que pertenece se denomina población o universo.

La figura 2-1 ejemplifica la noción del muestreo. La **población** (o universo) es *un grupo grande de personas de interés particular que deseamos estudiar y entender.* Con frecuencia las poblaciones estudiadas incluyen a las personas de un país, estado o comunidad; los presos en las instalaciones correccionales de un estado; los estudiantes actualmente inscritos en una universidad; las familias con hijos en edad escolar; los pacientes de un hospital; los jefes de cocina en restaurantes de la ciudad de Nueva York; y los ejecutivos de corporaciones. Una **muestra** es *un subgrupo pequeño de la población; la muestra se observa y se mide y después se utiliza para obtener conclusiones sobre la población.*

Población (o universo)

Grupo grande de personas de interés particular que deseamos estudiar y entender

Muestra

Subgrupo pequeño de la población; la muestra se observa y se mide y después se utiliza para obtener conclusiones sobre la población.

El muestreo es algo que hacemos todo el tiempo. Probamos una cucharada (una muestra) para decidir si agregamos más picante en polvo a la olla (la población). Para explorar una carrera académica, digamos, sociología, tomaríamos uno o dos cursos (una muestra) para determinar si el universo de ideas y actividades de la sociología nos agrada. Una primera cita con alguien es un muestreo de la personalidad del individuo, una primera exposición al universo de sus tendencias de conducta y actitudes. El muestreo es una conducta humana común y eficaz.

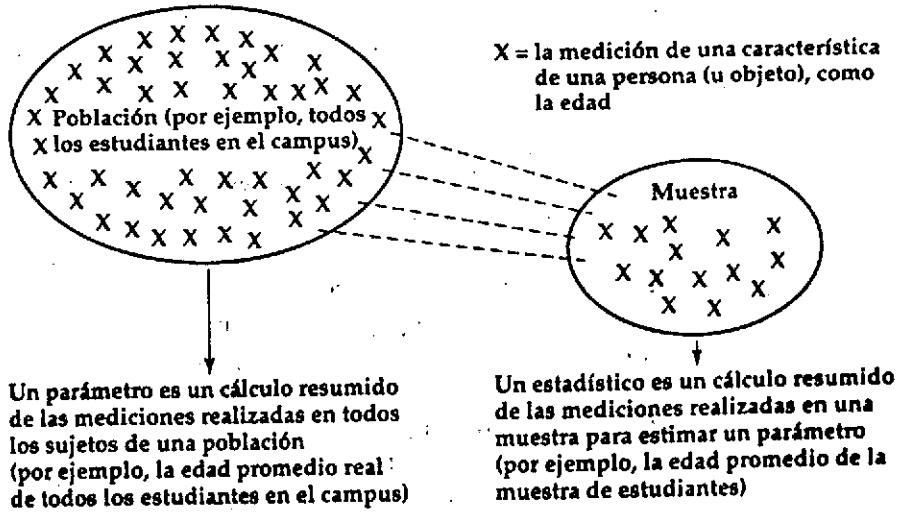


FIGURA 2-1
La relación de una población (universo) de mediciones con una muestra de mediciones

Nuestro interés, sin embargo, no está en la muestra por sí misma. En cambio, queremos aprender sobre la población entera. Para adquirir información completamente correcta respecto de una población entera, mediríamos *todos* sus miembros y resumiríamos los resultados en términos matemáticos, reportando porcentajes, tasas y promedios. *Al cálculo resumido de mediciones realizadas en todos los sujetos en una población* se le llama **parámetro**. Por ejemplo, el promedio de edad de presos en la Prisión de Sharpwire es un parámetro. El porcentaje de ejecutivos mujeres en la Menrule Plastics Corporation es un parámetro. Por desgracia, la mayoría de las poblaciones son tan grandes que no podemos invertir el tiempo y los recursos necesarios para medir a todos los miembros. Por ejemplo, sería absurdo medir las estaturas de todos los adultos en un país. A causa de los altos costos para medir a cada sujeto en una población, los verdaderos valores de los parámetros comúnmente son desconocidos.

Por fortuna, el muestreo nos permite *estimar* parámetros con precisión. Con las muestras calculamos estadísticos en vez de parámetros. Un **estadístico** es *un cálculo resumido de mediciones realizadas en una muestra para estimar un parámetro de la población*. Por ejemplo, en una muestra de 800 republicanos registrados en New Jersey encontraríamos que 74 por ciento, apoyan al gobernador. Este porcentaje constituye un estadístico: sólo una estimación del verdadero apoyo al gobernador. Una muestra y las estadísticas calculadas acerca de ésta son simples herramientas para obtener *conclusiones sobre una población en general* —la población como un todo—. Tales conclusiones, si fueron realizadas siguiendo procedimientos estadísticos adecuados, se llaman generalizaciones estadísticas.

Parámetro

Cálculo resumido de mediciones realizadas en todos los sujetos en una población.

Estadístico

Cálculo resumido de mediciones realizadas en una muestra para estimar un parámetro de la población.

Nunca debemos perder de vista el hecho de que la población es lo que nos preocupa. Por ejemplo, una muestra de votantes en "encuesta de salida" (tomada cuando las personas salen de las casillas) sugeriría que el candidato A es el ganador. Ésta, sin embargo, es una estimación, una aproximación del nivel de apoyo real. El verdadero ganador sólo se conocerá después de que se cuenten todos los votos, es decir, cuando la población entera de votantes haya sido medida.

Una manera de recordar que una muestra sólo proporciona estimaciones es comparar los resultados de varias muestras de la misma población. Si un profesor de estadística mandara a cada uno de los 30 miembros de la clase a reunir una muestra de 10 compañeros estudiantes y estimara el promedio de edad de los estudiantes, cada miembro de la clase obtendría un resultado ligeramente diferente. (Si no está convencido, consiga usted mismo dos muestras.) Esta variabilidad en los

resultados de las muestras sólo refleja el hecho de que el estadístico de una muestra única es sólo una estimación del verdadero parámetro de la población.

Entonces, ¿cómo confiaremos en los resultados de una sola muestra? La respuesta a esta pregunta implica buenas y malas noticias. Las malas son que el estadístico debe reconocer que las conclusiones de una muestra no son totalmente correctas; tales estadísticas son sólo estimaciones de parámetros. Las buenas noticias son que los procedimientos estadísticos y la lógica de la teoría de probabilidad permite a los estadísticos especificar un grado de error conocido en las predicciones y, por consiguiente, estipular el grado de confianza que tendríamos en una conclusión basada en estadísticas. En pocas palabras, aunque las estimaciones estadísticas no son perfectas, sabemos qué tan cerca están de la perfección.

Estimación estadística cuidadosa contra adivinación o estimación apresurada

La imaginación estadística enfatiza el entendimiento de un detalle en su contexto apropiado, teniendo cuidado de no emitir conclusiones simplistas o fantásticas. La estimación estadística es diferente del sentido común de la "adivinación o estimación apresurada", que es a menudo tendencioso. Una **estimación estadística** es *el informe de una medición resumida basada en el muestreo sistemático y en mediciones precisas e informadas con grados conocidos de error y confianza.* Una adivinación o **estimación apresurada** es *un informe de una medición sumaria basada en las experiencias personales limitadas y comúnmente subjetivas, evidencia anecdótica u observaciones casuales apresuradas.*

La adivinación podría ocurrir cuando un reportero de noticias elige al candidato A como el seguro ganador porque el informe de las encuestas de salida lo apoya con 52 por ciento de probables votantes. En contraste, tomando en cuenta el tamaño de la muestra, un estadístico sería más cauto y destacaría el hecho de que 52 por ciento significa 52 más y menos 5 puntos porcentuales; por consiguiente, el apoyo se encuentra entre 47 y 57 por ciento. La victoria del candidato A no está asegurada porque el apoyo podría ser *tan bajo como* 47 por ciento. Además, el estadístico mantiene un grado de confianza para la estimación como 95 por ciento. (No podemos exigir 100 por ciento de confianza hasta que todos los votos se contabilicen.) La estimación estadística cuidadosa es diferente incluso de una buena suposición. El estadístico difiere de otros "pronosticadores" en dos maneras importantes: El estadístico (1) controla y maneja el grado de error en las estadísticas reportadas y (2) señala de forma precisa la confianza en sus conclusiones.

Un tipo particularmente insidioso de la estimación apresurada es un **estereotipo** prejuicioso, es decir, *una generalización falsa que implica que todos los individuos en una categoría comparten ciertas características, normalmente indeseables.* Existe un estereotipo racista, por ejemplo, en creer que los afroamericanos son muy ignorantes, perezosos o inmorales para mantener a sus familias y que ésta es la causa de la pobreza en Estados Unidos. De hecho, casi 7 de cada 10 estadunidenses pobres son blancos y la mayoría de la gente pobre tiene empleo. Las estimaciones apresuradas a menudo se guían por sentimientos que refuerzan estereotipos y sentimientos como odio, temor y superioridad. En contraste, las generalizaciones estadísticas se inter-

pretan con cautela y dentro del contexto más grande de comprobación científica con sus resguardos contra la subjetividad. La tabla 2-1 compara las estimaciones apresuradas con las estimaciones estadísticas.

Error de muestreo y su manejo con la teoría de la probabilidad

Como la única manera de conocer un parámetro verdadero es mediante el sondeo de la población entera, cada estadístico calculado de una muestra es una estimación. Por casualidad, los estadísticos de algunas muestras están más cerca del valor del parámetro verdadero que otros. La **teoría de la probabilidad** (capítulo 6) consiste en *el análisis y la comprensión de las probabilidades de ocurrencia.* Nos brinda un conjunto de reglas para determinar la exactitud de los estadísticos de la muestra y calcular los grados de confianza que tenemos en las conclusiones sobre una población.

Para manejar exitosamente el error de muestreo debemos enfocarnos en sus fuentes específicas: el tamaño y la representatividad de la muestra. El **tamaño de la muestra** se refiere al *número de casos u observaciones que constituyen una muestra: el número de personas u objetos observados.* De manera general, cuanto mayor sea la muestra, menor será el rango del error. Suponga que un investigador envía a dos asistentes para determinar la edad promedio del conjunto de estudiantes. Uno les preguntó sus edades a 3 estudiantes, mientras que el segundo les preguntó a 1 000. La intuición nos lleva a tener mayor confianza en los resultados de la muestra mayor, porque la muestra más pequeña pudiera reunir más fácilmente sólo estudiantes jóvenes o sólo mayores. En un capítulo posterior aprenderemos a calcular e informar estadísticos con un "intervalo de confianza" de más menos una cantidad exacta de error para cualquier tamaño de muestra dada. Con una muestra de 1 000 encontraríamos que la edad promedio en el campus es de 22.4 años más menos 0.3 años, lo que sugiere que el promedio de edad se ubica entre 22.7 años (esto es, 22.4 + 0.3) y 22.1 años (esto es, 22.4 – 0.3). El cálculo de "más menos algún error de muestreo" está basado en probabilidades matemáticas o en la teoría de la probabilidad.

TABLA 2-1 "Estimación apresurada" del sentido común contra estimación estadística

Estimación apresurada del sentido común	*Estimación estadística*
La idea se basa en experiencias personales limitadas y comúnmente subjetivas, evidencia anecdótica u observaciones apresuradas.	La idea se basa en el muestreo sistemático en la medición.
Produce conjeturas y conclusiones equivocadas.	Produce estimaciones confiables con grados conocidos de error y confianza.
Genera y refuerza estereotipos.	Produce generalizaciones estadísticas.
Usualmente un asunto de opinión.	Usualmente un asunto de hecho.

La teoría de la probabilidad también nos permite señalar exactamente qué tan a menudo un estadístico predecirá el parámetro incorrectamente, es decir, qué tan a menudo los errores pueden causar una respuesta incorrecta. Por ejemplo, podemos advertir que 5 por ciento de las veces nuestros procedimientos generan una conclusión falsa. Al notar este nivel de error, sin embargo, estamos percibiendo también nuestro nivel de confianza. Si nuestra estimación es incorrecta sólo 5 por ciento de las veces, entonces es correcta el 95 por ciento del total; así, tenemos 95 por ciento de certeza.

Un segundo factor que afecta la exactitud del muestreo es hasta qué punto todos los segmentos de una población realmente están incluidos en la muestra: la representatividad de la muestra. Una **muestra representativa** es *aquella en la que todos los segmentos de la población están incluidos en la muestra en sus proporciones correctas de la población*. Por ejemplo, si una población del campus realmente es 54 por ciento hombres y 46 por ciento mujeres, una muestra representativa tendrá que acercarse a esos porcentajes.

Muestra representativa

Muestra en la que todos los segmentos de la población están incluidos en la muestra en sus porciones correctas de la población.

Una **muestra no representativa** es aquella *en la que algunos segmentos de la población están sobre o sub representados en la muestra*. Éste es un tipo peligroso de error de muestreo porque puede generar resultados totalmente engañosos. Suponga, por ejemplo, que la administración del campus desea encuestar a los estudiantes sobre su apoyo para ampliar el estadio de futbol. Los voluntarios de la Asociación Estudiantil de Enfermería llevan a cabo la encuesta y se les pide registrar el voto de cada décimo estudiante; pero en cambio ellos registran los votos de cada décimo estudiante que sale del edificio de enfermería. Sin sorpresa, los resultados muestran que sólo 23 por ciento de los estudiantes están a favor de la ampliación. ¿Por qué? Porque los miembros de la asociación en realidad encuestaron a la población de estudiantes de enfermería, que en su gran mayoría son mujeres y por lo tanto no es representativa del campus en conjunto. Diríamos que esta muestra está sesgada por una porción muy desproporcionada de mujeres. Tal muestra no representativa permitió que un segmento de la población tuviera más "votos" de lo que les correspondía sobre una cuestión.

Hay una variedad de diseños de muestreo; pero uno de los más empleados es la muestra aleatoria simple. Una **muestra aleatoria** es *aquella en la cual cada persona (u objeto) en la población tiene la misma oportunidad de ser seleccionada(o) para la muestra*. (En términos técnicos, decimos que todos en la población tienen *una misma probabilidad de inclusión* en la muestra.) Este diseño es como una rifa o lotería, donde cada persona en la población sólo entraría una vez. Una muestra aleatoria de tamaño suficiente producirá normalmente una muestra representativa.

Muestra aleatoria

Muestra en la cual cada persona (u objeto) en la población tiene la misma oportunidad de ser seleccionada(o) para la muestra.

El tamaño y la representatividad de la muestra, sin embargo, constituyen aspectos separados. Una muestra grande no garantiza una muestra representativa. Un error sistemático y repetitivo en el muestreo puede producir una muestra grande pero sesgada. Un caso típico de error de muestreo sistemático ocurrió en la campaña presidencial de 1936, en la cual la revista *Literary Digest* seleccionó una muestra grande de directorios telefónicos y propietarios de automóviles. Los resultados mostraron un apoyo aplastante para el candidato republicano, Alf Landon, sobre Franklin D. Roosevelt, el candidato demócrata. Cuando llegó el día de la elección, no fue Landon sino Roosevelt quien ganó la elección —¡y con un triunfo abrumador!—. La encuesta de *Literary Digest* ignoró sistemáticamente a los votantes sin teléfonos ni automóviles, y por tanto, no registró los votos de los electores pobres, quienes constituyeron el grueso del apoyo de Roosevelt (Babbie, 1992: 192-193). Existen métodos para verificar la representatividad de una muestra, y los cubriremos en el capítulo 10. Basta decir aquí que una pequeña muestra representativa es mejor que una grande no representativa. Una cucharadita del guiso con todos los ingredientes constituye una mejor prueba de sabor que una taza obtenida sólo de la superficie de la olla.

Control del error de medición

Además de evitar los errores de muestreo, debemos definir con precisión cómo se harán las mediciones y cómo se codificarán las respuestas una vez que se recopilen los datos. *El conjunto de procedimientos u operaciones para medir una variable* se llama **definición operacional**. Por ejemplo, suponga que utilizamos datos del censo de Estados Unidos para dirigir un estudio sobre la pobreza urbana, con una muestra de 300 ciudades. Existen varias formas de *operacionar* una medición de la pobreza. El desafío consiste en seleccionar la manera que represente con mayor precisión cuántos hogares en una ciudad están habitados por familias pobres. Una medición es el porcentaje de hogares que reciben vales de alimentos. Una segunda es la tasa de desempleo en la ciudad. Una tercera sería el porcentaje de hogares que viven debajo del nivel de pobreza que se define en el ámbito federal (ingreso específico justo para el tamaño de la familia). De hecho, la tercera opción generalmente se reconoce como la mejor aproximación hacia la pobreza para una comunidad, y por ello la escogeríamos como nuestra definición operacional. Una guía eficaz para la elección de una definición operacional consiste en identificar los tipos comunes de error de medición, y hacer todo lo posible para minimizarlos. En general, el error de medición tiene que ver con la precisión, validez y confiabilidad de una definición operacional.

Una **medición precisa** es *aquella en la que el grado del error de medición es suficientemente pequeño para la tarea en cuestión*. La precisión depende de las circunstancias prácticas y es controlada especificando el error de redondeo (un tema que se discutirá brevemente). Por ejemplo, al cortar troncos de 50 centímetros de longitud para la chimenea, bastará un grado pequeño de precisión, porque podemos operar con un amplio grado de error, por decir, "unos 15 centímetros de más o de menos". En contraste, para una prueba de calidad de un circuito de microcomputadora, puede requerirse una precisión de un milésimo de centímetro. El grado de precisión es una cuestión de tolerancia. Nos preguntamos: ¿qué tanto error de medición podemos tolerar sin encontrar problemas prácticos ni conclusiones científicas fallidas?

Una **medición válida** es *la que mide lo que se requiere*. La validez atañe a la pregunta: una definición operacional ¿mide lo que se supone? Por ejemplo, medir el tamaño de la cintura como una definición operacional de la variable peso no sería válida, pero el uso de una báscula propiamente calibrada sí lo sería. Las mediciones absolutamente válidas son difíciles de lograr. Por ejemplo, la educación comúnmente es operacionalizada en años de escolaridad, pero existen otras maneras de educarse, como el entrenamiento en el ejército o la capacitación en el trabajo. La validez de una medición rara vez se conoce con certeza absoluta.

La **confiabilidad** tiene que ver con *la consistencia de las mediciones de un sujeto a otro y de un tiempo a otro*. Con una báscula confiable, dos sujetos de la muestra que realmente pesen lo mismo tendrán las mismas mediciones. Por otro lado, si la báscula no es confiable, no está midiendo realmente el peso. Es decir, una báscula no confiable es inútil y no válida.

En estudios sociales, la confiabilidad depende principalmente de que una pregunta sea formulada de manera que todos los encuestados la interpreten de la misma forma. Por ejemplo, para medir el *ingreso anual* en una encuesta realizada de puerta en puerta, sería inválido y no confiable sólo preguntar: "¿Cuál es su ingreso total?" ¿Se alude al solo ingreso de los encuestados o al de todos aquellos que trabajan en la casa? ¿Incluye los dividendos y otros ingresos además del sueldo? ¿Podrían mentir los encuestados sobre este asunto privado? Aunque no sea perfecta, una mejor forma de medir el ingreso familiar consiste en mostrar un tabulador con categorías de ingreso y preguntar: "¿En cuál de estos grupos se ubica su ingreso familiar total, de todas las fuentes, del año pasado, antes de descontar los impuestos? Simplemente dígame la letra." (National Opinion Research Center, 1994; *www.icpsr.umich.edu/gss/codebook/income.htm*).

Niveles de medición: selección cuidadosa de los procedimientos estadísticos

Medición

La **medición** es *la asignación de símbolos, tanto nombres como números, a las diferencias que observamos en las cualidades o cantidades de una variable. La medición de un sujeto particular de la muestra en una sola variable* es la **puntuación** del sujeto para esa variable o, para usar terminología computacional, un código. Suponga por un momento

que su clase de estadística constituye una muestra. Podríamos registrar las variables de edad, semestre, género, promedio y raza. Para una estudiante, Juana, estas puntuaciones son 20 en edad, primer ingreso en semestre, femenino en género, 3.25 en promedio y blanca en raza; para otro, Rubén, las codificaciones respectivas son 19, estudiante de segundo semestre, masculino, 3.48 en promedio y afroamericano. Utilizaremos los términos *puntuación* y *código* indistintamente. El *valor* de una puntuación es su cantidad.

Como esta simple ejemplificación revela, no todas las variables se miden de la misma forma. Algunas se registran con nombres o categorías que identifican diferencias en tipo o calidad, como afroamericano y blanco para la variable raza. Otras variables permiten distinciones de grado o distancia entre cantidades, como las variables de edad y promedio. Estas variables tienen una **unidad de medición**, un *intervalo determinado o distancia entre las cantidades de las variables*. Las anotamos numéricamente, como las marcas numeradas en una regla o en una cinta métrica. La unidad de medición para una escala de temperatura es un grado; para el peso, un kilogramo; para la altura, un centímetro y así sucesivamente.

Para comprender las finas distinciones entre las propiedades de medición de las variables, usamos un esquema llamado niveles de medición. El **nivel de medición de una variable** *identifica sus propiedades de medición, las cuales determinan el tipo de operaciones matemáticas (suma, multiplicación, etcétera) que puede usarse apropiadamente con dicho nivel, así como las fórmulas estadísticas que se utilizan para probar las hipótesis teóricas.* Estos niveles se llaman nominal, ordinal, de intervalo y de razón. El nivel de medición de una variable es una guía importante para seleccionar fórmulas estadísticas y procedimientos.

Nivel de medición de una variable

Identifica las propiedades de medición de la variable, y determina el tipo de operaciones matemáticas (suma, multiplicación, etcétera) que puede usarse apropiadamente con dicho nivel, así como las fórmulas estadísticas que utiliza para probar las hipótesis teóricas.

Variables nominales

Las **variables nominales** son aquellas donde los códigos *sólo indican una diferencia en categoría, clase, calidad o tipo.* La palabra *nominal* viene del vocablo latín para *nombre*, y estas variables tienen categorías de nombre. Algunos ejemplos incluyen lugar de nacimiento (Chicago, Atlanta, Monterrey, etc.), sabor favorito de helado (vainilla, chocolate, frambuesa), marca de automóvil (Ford, Lexus, Pontiac) y carrera académica (psicología, química, ingeniería eléctrica).

Las variables nominales no admiten puntuaciones numéricas ordenadas significativamente. No obstante, gracias a que las computadoras procesan números con mayor eficacia, a veces numeramos las categorías de estas variables en códigos computacionales. Por ejemplo, para la variable *género* asignaríamos los códigos como 0 = hombre y 1 = mujer. La elección de números para tales códigos es

arbitraria; también hubiéramos podido codificar 0 = mujer y 1 = hombre. Además, las categorías de una variable nominal no pueden clasificarse significativamente en orden de magnitud (de elevado a bajo) aun cuando se asignen códigos a los números ordenados. Por ejemplo, codificar mujer como 1 y hombre como 0 no implica que las mujeres tengan una puntuación de 1 más que los hombres. No existe ningún sentido de grado con las variables nominales. Una persona es hombre o es mujer, y en cualquier caso no tiene un grado. Incluso algunas puntuaciones numéricas en apariencia; son realmente variables nominales. Por ejemplo, el número del seguro social es, de hecho, una categoría, y no tiene sentido calcular su promedio.

Puesto que muchas variables nominales tienen sólo dos categorías, existe un nombre especial para ellas. Una **variable dicotómica** tiene *sólo dos categorías*. Una variable dicotómica común en las encuestas es cualquiera con las respuestas "sí" y "no" y, en diseños de investigación de laboratorio, aquella que distingue la "presencia" (el grupo experimental) o la "ausencia" (el grupo de control). Por ejemplo, al probar la efectividad de un nuevo medicamento contra la fiebre de heno, al grupo experimental se le administra la nueva droga y se registra como 1. Al grupo de control se le da una droga de imitación (o placebo) y se anota como cero. En la computadora nombramos esta variable GRUPO. Cuando deseamos aislar al grupo experimental para el análisis, damos instrucciones a la computadora para que busque los códigos de GRUPO y saque dichos casos con el código 1.

Variables ordinales

Como las variables nominales, las **variables ordinales** *designan categorías, pero tienen la propiedad adicional de permitir clasificar las categorías* desde la mayor hasta la menor, de la mejor a la peor o de la primera a la última. Las variables ordinales comunes incluyen clasificación de clase social (alta, media, baja, indigente), nivel de clase educativa (último año, primer ingreso, etc.), y calidad de vivienda (estándar, insuficiente, en ruinas). Las preguntas de estudio que miden actitudes y opiniones a menudo emplean puntuaciones ordenadas. Por ejemplo, la variable "actitud hacia el aborto legal" podría ordenar el grado de acuerdo mediante el uso de categorías de respuesta: *totalmente de acuerdo, de acuerdo, no sabe, en desacuerdo, totalmente en desacuerdo*. Este conjunto de códigos ampliamente utilizado se denomina escala de Likert, en honor a su creador, Rensis Likert (1932).

Variables de intervalo

Las **variables de intervalo** tienen las características de las variables nominales y ordinales, y además *una unidad numérica de medición definida*. Las variables de intervalo identifican las diferencias en monto, cantidad, grado o distancia, y se les asignan puntuaciones numéricas muy útiles. Los ejemplos incluyen la temperatura (registrada al grado térmico más cercano) y el coeficiente de inteligencia (CI), que va desde cero hasta 200 puntos. Con las variables de intervalo, los intervalos o distancias entre las puntuaciones son las mismas entre cualquier par de puntos en la escala de medición. Por ejemplo, con la variable temperatura, la diferencia entre 10

y 11 grados Fahrenheit es la misma que entre 40 y 41 grados. Un conjunto de unidades de medición ordenadas proporciona la habilidad para sumar, restar, multiplicar y dividir puntuaciones y calcular promedios.

Las variables de intervalo dan un sentido de "cuánto" o "de qué tamaño": qué tan caliente, qué tan obstinado, qué tan conservador, qué tan deprimido, qué tan largo y qué tan pesado. Con las variables de intervalo pensamos en términos de distancia entre las puntuaciones sobre una línea recta. Por ejemplo, si el promedio de las calificaciones de un grupo en una prueba es 80, y Carlos obtuvo 85 y Berta 90, entonces la puntuación de Berta estuvo dos veces más arriba del promedio que la puntuación de Carlos. Además, los márgenes de error con las variables de intervalo están más definidos y son más fáciles de manejar porque las puntuaciones numéricas pueden redondearse.

Comparar las propiedades de variables de intervalo y variables ordinales es informativo. A diferencia de las variables de intervalo, las variables ordinales carecen de una unidad de medición determinada aun cuando las categorías ordenadas sean numeradas. Por ejemplo, la ubicación final en una carrera de caballos (1, 2, 3, etcétera) es sólo ordinal; simplemente indica qué caballos cruzaron la línea final en primero, segundo y tercer lugar, y así sucesivamente; pero no aclara qué tan separados terminaron unos de otros. Para que esta variable ordinal se trate como una variable de intervalo, los caballos tendrían que terminar en un solo orden y distancias iguales entre ellos. Además, la resta entre números de posición de una variable ordinal proporciona sólo diferencias entre rangos, no la distancia entre las puntuaciones. Por ejemplo, si los caballos llamados "Piernas Largas" y "Problemas en el Puente" terminan en tercero y sexto lugar respectivamente, entonces "Piernas Largas" llegó tres posiciones adelante. Estos caballos podrían haber llegado a la meta separados por unas cuantas pulgadas o cientos de yardas. Mientras las variables ordinales permiten algunos cálculos, como diferencias en rango y rango promedio, tienen utilidad matemática limitada. Las variables de intervalo poseen utilidad matemática mucho mayor que las variables ordinales.

Variables ordinales de tipo intervalo. A causa de que los procedimientos estadísticos de nivel de intervalo son más detallados e informativos que los procedimientos ordinales, a menudo se realizan esfuerzos para forzar las variables ordinales en variables de intervalo. Esta excepción a veces se justifica porque las estadísticas del nivel de intervalo son muy sólidas (véase el capítulo 16). ¿Cuándo podemos tratar una variable ordinal como si tuviera un nivel de medición de intervalo? Primero, la variable ordinal debe tener al menos siete categorías o puntuaciones de clasificación; cuantas más, mejor. Segundo, la distribución de las puntuaciones no puede estar sesgada o ser bimodal. Esto significa que las puntuaciones se distribuyen normalmente (véase el capítulo 5) o de forma rectangular (esto es, uniformemente distribuidos a lo largo de su rango con una frecuencia parecida en cada punto de la escala). Cuando estas circunstancias se aplican, podemos referirnos al nivel de medición de una variable ordinal como de "tipo intervalo". Sin embargo, debemos ser cautelosos en la interpretación de los cálculos con tales variables. La aplicación de esta excepción a los cuatro niveles convencionales de medición se hará evidente en los capítulos posteriores.

Variables de razón

Las **variables de razón** poseen las características de las variables de intervalo y *un punto cero verdadero, donde una puntuación cero significa ninguno.* Peso, altura, edad, distancia, tamaño de la población, duración de tiempo y promedio son ejemplos de variables de razón.

Comparar variables de razón con variables de intervalo resulta informativo porque ambas tienen intervalos establecidos en su unidad de medición; pero sólo las variables de razón incluyen un punto cero con significado. Algunas variables de *intervalo* pueden tener una puntuación de cero, pero el punto cero es arbitrario; es decir, podría colocarse en cualquier punto dentro del rango posible de una variable porque el cero no significa ninguno. Por ejemplo, la temperatura cero no significa ausencia de temperatura. Así, en la escala Fahrenheit está ubicado en 32 grados debajo del punto de congelación; mientras en la escala Celsius se encuentra en el punto de congelación.

Los puntos cero verdaderos de las variables de razón permiten incluso mayor flexibilidad en los cálculos y el análisis estadístico. Como las variables de intervalo, las variables de razón pueden multiplicarse y dividirse; pero también podemos calcular razones, de ahí el nombre. Una **razón** es *la cantidad de una observación con respecto a otra.* Por ejemplo, si Juan come tres rebanadas de pizza y Esther come una, la razón es tres a uno, que se escribe como 3:1. Con una variable de nivel de razón la respuesta para una razón calculada tiene sentido; mientras con una variable de nivel de intervalo no lo tiene. Por ejemplo, un joven de 40 kilogramos es dos veces más pesado que uno de 20 kilogramos, una razón de 2:1. Pero no tiene sentido afirmar para la variable de intervalo *temperatura* que Miami es cuatro veces más caluroso, donde hay 80 grados, que Nueva York, donde hay 20 grados. ¡Nueva York no es caluroso en absoluto! Entonces, una manera para determinar si una variable tiene un cero verdadero es intentar interpretarlo como una razón. Si la razón no tiene sentido, la variable está, si acaso, en un nivel de intervalo y su punto cero es arbitrario.

Debido a las similitudes de los procedimientos estadísticos aplicados a las variables de intervalo y a las de razón, a menudo agrupamos estas distinciones refiriéndonos a estas variables como intervalo/razón. De igual forma, nos referimos a las variables nominales/ordinales. La tabla 2-2 resume las propiedades de los cuatro niveles convencionales de medición.

Para aprovechar las unidades determinadas de medición, frecuentemente los científicos crean maneras de cambiar variables nominales/ordinales en las formas de intervalo/razón. La tabla 2-3 presenta diversas variables nominales que se transformaron en una variable de nivel de razón llamada Índice de Comportamiento de Riesgo para la Salud. Las variables nominales se registran como cero para *no* y 1 para *sí*. Esto se llama codificación prototipo (conocida como *dummy* en inglés) porque las puntuaciones numéricas son artificiales; cero y 1 no distinguen montos o cantidades. En cambio, cero significa que el factor de riesgo no está presente, y 1 que sí está presente. La variable de razón RIESGO es el número total de factores de riesgo de un individuo, esta variable tiene un punto cero verdadero. Si Jeremías fuma, bebe, conduce en estado de ebriedad y consume drogas; mientras que Adán sólo fuma, entonces Jere-

TABLA 2-2 Características de los cuatro niveles de medición

Nivel de medición	*Ejemplos*	*Cualidades*	*Operaciones matemáticas permitidas*
Nominal	Género, raza, preferencia religiosa, estado civil	Clasificación en categorías; denominación de categorías	Conteo del número de casos (es decir, la frecuencia) de cada categoría de la variable; comparación de los tamaños de las categorías
Ordinal	Rango de clase social, preguntas de actitud y opinión	Clasificación de categorías; ordenamiento de los rangos de categorías de la menor a la mayor	Todas las anteriores y los juicios de mayor que y menor que y el cálculo de diferencias y promedios de rangos
De intervalo	Temperatura, índices resumidos, escalas de actitud y opinión	Todas las anteriores y las distancias entre las puntuaciones tienen una unidad determinada de medición	Todas las anteriores y las operaciones matemáticas como suma, resta, multiplicación, división, raíz cuadrada
De razón	Peso, ingreso, edad, años de educación, tamaño de la población	Todo lo anterior y un punto cero real	Todas las anteriores y el cálculo de razones con significado

mías experimenta cuatro veces más conducta riesgosa que Adán, una razón de 4:1. Esperamos que su propia puntuación RIESGO sea más baja.

Los campos de los métodos de investigación y análisis psicométrico se dedican a construir índices de intervalo/razón y escalas de medición de grandes conjuntos de preguntas de estudio en el nivel nominal/ordinal. Un **índice** es la adición *de eventos objetivos, comportamientos, conocimiento y circunstancias.* Una **escala de medición** constituye *una adición de respuestas subjetivas sobre actitudes, sentimientos y opiniones.* Entre las escalas comunes de actitud y de fenómenos sociales están la espiritualidad (qué tan religiosa es una persona), la autoestima, la distancia social (diferencia percibida en el estatus social entre las personas) y el continuo liberalismo-conservadurismo.

Codificación y conteo de las observaciones

Una vez completada la recolección de datos, el siguiente paso en el manejo de datos consiste en codificar y registrar todas las mediciones en una hoja de vaciado o en un archivo de datos en la computadora. La tabla 2-4 presenta un ejemplo de un **registro de codificación**, *una descripción concisa de símbolos que describen el significado de cada puntuación para cada variable.* Tales datos vienen de un estudio sobre estudiantes en el Instituto Apple Pond, que es ficticio. En este registro de codificación

TABLA 2-3 Elaboración de un índice para transformar diversas variables nominales en una variable de razón

Número y nombre de la variable	*Definición operacional (cómo se mide la variable)*	*Nivel de medición*	*Código (cómo se registra)*
1. FUMA	¿Es fumador habitual?	Nominal	0 = no 1 = sí
2. ALCOHOL	Ha consumido cinco o más bebidas alcohólicas en una sola ocasión en el último mes	Nominal	0 = no 1 = sí
3. EJERCICIO	No se ejercita regularmente	Nominal	0 = no 1 = sí
4. DROGAS	Ha usado una droga ilícita en el último mes	Nominal	0 = no 1 = sí
5. CONDEBRI	Ha manejado mientras se encuentra en estado de ebriedad	Nominal	0 = no 1 = sí
6. RIESGO	Reporte del número de conductas de riesgo	Razón	Suma de las respuestas "sí" para las variables 1 a la 5

sustituimos símbolos numéricos por las categorías de masculino y femenino, y por niveles de estudios. Esto se hace porque a las computadoras se les facilita contar y seleccionar números (en lenguaje computacional, símbolos *numéricos*), que palabras (*caracteres* o símbolos de cadena).

En un registro de codificación, se tiene cuidado en ser muy preciso, porque la codificación de respuestas podría generar error de medición. Cada variable se codifica siguiendo dos principios básicos: inclusividad y exclusividad. El principio de inclusividad establece que para una variable dada *debe haber una puntuación o*

TABLA 2-4 Registro de codificación para respuestas de cuestionarios de especialistas en justicia delictiva en el Instituto Apple Pond

Nombre de la variable	*Descripción de los códigos de la variable*
NOMBRE DEL ESTUDIANTE	Registre el nombre, apellido paterno y apellido materno; espacio en blanco = faltante.
EDAD	Registre la edad informada hasta 97 años; 98 = 98 años o más; 99 = faltante.
GÉNERO	0 = hombre, 1 = mujer, 9 = faltante.
PROMEDIO	Promedio en una escala de cuatro puntos = número de puntos de calidad ganados por hora crédito (redondeado a dos posiciones decimales); 9.99 = faltante.
NIVEL DE CLASE	1 = nuevo ingreso, 2 = intermedio inicial, 3 = intermedio avanzado, 4 = avanzado, 8 = otro, 9 = faltante.

un código para cada observación realizada. Dicho de manera más sencilla, ¿incluimos una categoría de respuesta o una puntuación para cada respuesta posible? Por ejemplo, para la variable nominal *raza*, podríamos indicar las categorías blanco, afroamericano, nativo americano, asiaticoamericano, hispano y otra(s). La categoría de respuesta *otra(s)* evita la necesidad de ocupar espacio en el cuestionario para las categorías que se espera que tengan pocas respuestas en el lugar del estudio (como esquimal en Kansas). El código otra(s) es una *categoría residual* que abarca los remanentes (piense en la palabra *residuo*).

Aun cuando ignoramos la cuestión de cómo codificar a las personas de herencia mixta, si sólo usamos blanco y afroamericano, la categoría de raza no será inclusiva de, por decir, asiaticoamericano. Sin su propia categoría u otra(s), no es factible asumir que todos los asiaticoamericanos se registrarán de la misma forma. Algunos anotarán blanco, pero otros quizá dejen el espacio de la pregunta sin contestar. Después de calcular los totales quizá no seamos capaces de señalar exactamente cuántos tenemos de *cualquier* raza. Perdimos a los asiaticoamericanos que no contestaron la pregunta, y algunos de nuestros "blancos" son asiaticoamericanos; sin embargo no tenemos noción de cuántos. Semejante descuido de pérdida de control sobre el error de medición puede hacer que los datos de raza sean inservibles.

El **principio de exclusividad** sostiene que para una variable dada *cada observación es asignada a una y sólo una puntuación.* Así, cada categoría debe excluir puntuaciones que no le pertenezcan, y cualquiera de las dos categorías no debe compartir una respuesta. Por ejemplo, con la variable *afiliación religiosa en la niñez*, las categorías de respuesta protestante, bautista, católico, judío y otra(s) no serían mutuamente excluyentes porque un bautista quizá se anote como bautista, mientras otro lo haga como protestante. Cuando se sumen los totales de cada categoría, no podremos decir cuántos bautistas había en la muestra. Algunos quizá se registraron como "protestante", pero no tenemos forma de especificar quiénes y cuántos lo hicieron. La tabla 2-5 muestra los resultados del Estudio Social General de 1994 para esta variable. La exclusividad se asegura formulando a los protestantes las preguntas adicionales necesarias para conocer sus denominaciones específicas. (Para ayudar a su comprensión sobre la información de las tablas en este texto, modificamos las tablas para diferenciar claramente las "Especificaciones" de los datos disponibles y los "Cálculos". En los reportes publicados de estas tablas dichos términos no aparecerían.)

Regrese al registro de codificación de la tabla 2-4 y observe que el principio de inclusividad se completa proporcionando un *código para los datos perdidos*, llamados **valores perdidos**. Por ejemplo, decimos que género y nivel de estudios tiene un valor perdido de 9. En algunos estudios, los valores perdidos ocurren cuando por accidente el entrevistador se salta una pregunta o un encuestado no contesta. Al calcular las estadísticas para una variable, pasamos por alto los casos que resultan en un valor perdido.

La tabla 2-6 es una tabla desglosada de los resultados de la aplicación del cuestionario empleado en una encuesta aplicada a 10 especialistas en justicia delictiva del Instituto Apple Pond. Una tabla desglosada es una matriz que muestra las puntuaciones de todas las variables organizadas en columnas; y todos los casos, en filas.

TABLA 2-5 Distribución de la afiliación religiosa en la niñez, respuesta a las preguntas: ¿en qué religión fue educado? Si fue protestante: ¿en qué denominación específica, si la hay?

Especificaciones		Cálculos
Categoría de respuesta	*a) Número*	*b) Porcentaje de la muestra total*
Protestante		
Bautista	706	23.73%
Metodista	319	10.72
Luterana	220	7.39
Presbiteriana	139	4.67
Episcopal	68	2.29
Otra	309	10.39
Ninguna denominación		
o iglesia sin denominación	69	2.32
Total protestante	1 830	61.51%
Católica	882	29.65
Judía	55	1.85
Ninguna	127	4.27
Otra	74	2.49
Sin respuesta	7	0.23
Total	2 975	100.00%

FUENTE: National Opinion Research Center, Estudio Social General 1994.
www.icpsr.umich.edu/gss/codebook/relig16.htm.
www.icpsr.umich.edu/gss/codebook/denom16.htm.

TABLA 2-6 Tabla desglosada de respuestas al cuestionario de 10 especialistas en justicia delictiva en el Instituto Apple Pond (datos ficticios)

	Especificaciones			
Nombre del estudiante	*Edad*	*Género*	*Promedio*	*Nivel de clase*
Jessica A. Cortland	19	1	3.21	2
Mark E. Pippin	22	0	2.75	4
Stayman V. Winesap	19	0	2.43	1
Barry D. McIntosh	21	0	3.39	3
Harriet G. Smith	20	1	3.87	3
Antonio B. Rome	22	0	2.32	3
Robert J. Cox	18	0	3.25	1
Rodney I. Greening	20	0	9.99	2
Thomas R. York	22	0	2.47	4
Goldie D. Licious	19	1	3.68	2

Una tabla desglosada es útil para resumir datos de una forma eficaz. Por ejemplo, si contamos rápidamente el número de mujeres en la muestra sumando las unidades listadas bajo "Género". Mediante esta simple tabla desglosada podemos ver rápidamente que la muestra se compone de siete hombres y tres mujeres; existen dos estudiantes de primer año, tres de segundo año, tres de tercer año, y dos de último año; el rango de edades oscila entre los 18 y los 22 años; y el rango del promedio va de 2.43 a 3.87 con un caso no reportado. Por supuesto, para una muestra grande, un procedimiento eficaz implica tanto la especificación de estos códigos de la tabla desglosada en un archivo de datos de la computadora como hacer que el programa computacional se encargue de los cálculos. Los archivos de datos de computadora están organizados como estas tablas desglosadas.

Distribuciones de frecuencias

Una vez que todos los datos están organizados en una tabla desglosada o en un archivo de datos de computadora, el siguiente paso en el análisis consiste en enfocarse por separado en cada variable y contestar la pregunta: ¿cuántos sujetos caen en cada categoría o puntuación? Organizamos los datos de cada variable en una **distribución de frecuencias,** *que es una lista de todas las puntuaciones observadas de una variable y la frecuencia (f) de cada puntuación (o categoría).* Utilizamos letras mayúsculas para representar una variable. Si *X* se define como la variable *género,* la distribución de frecuencias de *X* simplemente muestra cuántos hombres y mujeres hay en la muestra. La tabla 2-5 mostrada anteriormente proporciona la distribución de frecuencias para la afiliación religiosa en la niñez.

> **Distribución de frecuencias**
>
> Listado de todas las puntuaciones observadas de una variable y la frecuencia (*f*) de cada puntuación (o categoría).

Estandarización de distribuciones de puntuaciones

El conocimiento de la frecuencia de una categoría no resulta muy informativo por sí mismo. Por ejemplo, alguien nota que existen cinco millonarios viviendo en una ciudad. Cinco no son muchos para la ciudad de Nueva York; pero lo serían para un pueblo de 800 personas. Así, es más informativo reportar la frecuencia de una categoría como una proporción o porcentaje con respecto al número total de sujetos en la muestra. La imaginación estadística nos impele a expresar la frecuencia de una categoría en un contexto mayor, como una parte en relación con un todo. Planteamos la pregunta: cinco millonarios, ¿de cuántas personas? Como observamos en el capítulo 1, las fracciones, las proporciones y los porcentajes ofrecen denominadores comunes o "medidas estándar" para facilitar la comparación de categorías y muestras. Para una muestra como un todo, la **distribución de frecuencias con pro-**

porciones consiste en *un listado de la* proporción *de respuestas para cada categoría o puntuación de una variable.* La **distribución de frecuencias de porcentajes** es *un listado del porcentaje de respuestas para cada categoría o puntuación de una variable.*

Distribución de frecuencias con proporciones

Listado de la proporción de respuestas para cada categoría o puntuación de una variable.

Distribución de frecuencias de porcentajes

Listado del porcentaje de respuestas para cada categoría o puntuación de una variable.

Para obtener estas distribuciones para cada categoría de respuesta o puntuación de una variable, escribimos una fracción, y después la dividimos para obtener la proporción y el porcentaje. Para facilitar la interpretación, la distribución de frecuencias de porcentajes es la que comúnmente se reporta. Por ejemplo, en los datos de la tabla desglosada del Instituto Apple Pond en la tabla 2-6, podemos observar que el porcentaje de hombres es

$$p\text{ [de la muestra de estudiantes que son hombres]} = \frac{\#\text{hombres}}{n} = \frac{7}{10} = .7000$$

$$\%\text{ [de la muestra de estudiantes que son hombres]} = (p)\,(100) = (.7000)\,(100) = 70.00\%$$

donde p significa la proporción y n es el tamaño de la muestra. Después de hacer lo mismo para las mujeres, tenemos la distribución de frecuencias de los porcentajes de la variable *género* en la columna de la derecha en la tabla 2-7. En ésta también se

TABLA 2-7 Frecuencia, frecuencia proporcional y distribuciones de frecuencias porcentuales de la variable género para una muestra de 10 estudiantes del Instituto Apple Pond

Especificaciones		Cálculos	
Género (X)	*Frecuencia (f)*	*Frecuencia proporcional*	*Frecuencia porcentual*
Hombre	7	.7000	70.00%
Mujer	3	.3000	30.00
Total	10	1.0000	100.00%

incluye la frecuencia y las distribuciones de frecuencias proporcionales. El total de todas las proporciones y porcentajes para una distribución será igual a 1.0000 y 100.00 por ciento, respectivamente, considerando el error de redondeo.

Cálculo de las frecuencias proporcionales y porcentuales de una categoría

$$p\text{ [de la muestra total (}n\text{) en una categoría]} = \frac{f\text{ de una categoría}}{n} = \frac{\#\text{ en una categoría}}{n}$$

% [de la muestra total (n) en una categoría] = (p [de la muestra total (100) (n) en una categoría]) (100)

donde

p [de la muestra total (n) en una categoría] = proporción de todos los casos que caen en la categoría,

f = frecuencia de casos (o número de casos) en la categoría,

n = tamaño de la muestra

La tabla 2-5 muestra la frecuencia y la distribución de frecuencias con porcentajes para la variable afiliación religiosa en la niñez.

Comparación de las frecuencias de dos variables nominales/ordinales

Otro dispositivo común para presentar datos es una **tabulación cruzada** (o clasificación cruzada) que *compara dos variables nominales/ordinales simultáneamente.* Dichas tablas de "tabulación cruzada" son esenciales para probar una hipótesis sobre la relación entre esas variables (véase capítulo 13). ¿Existe una relación, por ejemplo, entre género (X) y asistencia a la iglesia (Y) entre los estudiantes universitarios? Los resultados (ficticios) de un estudio en una universidad se presentan en la tabla 2-8.

Para los datos de la tabla 2-8, la tabla 2-9 muestra el lenguaje convencional que se utiliza para describir partes de una tabulación cruzada. Los números en las casillas representan la frecuencia de ocurrencias conjuntas (o simplemente "frecuencia conjunta") de categorías de las dos variables. Una ocurrencia conjunta es la combinación de categorías *para un solo individuo.* Por ejemplo, Carlos asistió a la iglesia; por ende, él se cuenta como parte de la frecuencia conjunta de la casilla "hombres asistentes", donde hay 66 sujetos. La frecuencia conjunta de hombres que no asistieron es de 134; de mujeres que asistieron, 94, y de mujeres que no asistieron, 146. Las sumas en los márgenes de la tabla se denominan totales marginales. El total de la columna para los hombres nos indica que en la muestra hay 200 hombres en total; de igual forma, hay 240 mujeres. La suma para asistencia y no asistencia es de 160 y 280, respectivamente. El gran total es el tamaño de la muestra (n) que se

TABLA 2-8 Asistencia a la iglesia en el último mes por género entre estudiantes universitarios, $n = 440$ (datos ficticios)

	Especificaciones		
	Género (X)		
Asistencia a la iglesia (Y)	*Hombres*	*Mujeres*	*Total*
Asistió	66	94	160
No asistió	134	146	280
Total	200	240	440

presenta en la parte inferior derecha. Observe que tanto el total de las columnas como el total de las filas coinciden con el gran total.

Si existe una relación entre género y asistencia a la iglesia, cualquier género tiene una frecuencia mucho mayor de asistencia a la iglesia. Los números brutos en la tabla pueden ser confusos, puesto que hay más mujeres que hombres en la muestra. Debemos preguntarnos: ¿de cuántos? Así, utilizamos los totales marginales para calcular los *porcentajes* de hombres y mujeres que asistieron a la iglesia:

TABLA 2-9 Lenguaje de la tabulación cruzada: asistencia a la iglesia en el último mes por género entre estudiantes universitarios, $n = 440$

	Género (X)		
Asistencia a la iglesia (Y)	*Hombres*	*Mujeres*	*Total*
Asistió	66	94	160
No asistió	134	146	280
Total	200	240	440

Casillas (contiene frecuencias conjuntas)

Totales (marginales) de fila

Totales (marginales) de columna

Gran total (n)

$$p\ [\text{de hombres que asistieron a la iglesia durante el mes pasado}] = \frac{\#\text{ hombres que asisten}}{\#\text{ total de hombres}} = \frac{66}{200} = .3300$$

$$\%\ [\text{de hombres que asistieron a la iglesia durante el mes pasado}] = (p)\ (100) = (.3300)\ (100) = 33.00\%$$

Asimismo, el porcentaje de mujeres que asistieron a la iglesia durante el mes pasado es 94/240 (100) = 39.17 por ciento. Después de calcular estos porcentajes, observamos claramente que un porcentaje mayor de mujeres asistió a la iglesia durante el mes pasado. Estos porcentajes se denominan porcentajes de columna porque se basan en totales marginales de columna. Es decir, un **porcentaje de columna** es *la frecuencia de una casilla como un porcentaje del total marginal de columna.* De igual manera, un **porcentaje de fila** es *la frecuencia de una casilla como un porcentaje del total marginal de fila.* Por ejemplo, el porcentaje de fila de los asistentes a la iglesia que son hombres es 66/160 (100) = 41.25 por ciento. Estas simples tablas de tabulación cruzada proporcionan estadísticas descriptivas importantes y ayudan a determinan si existe una relación entre dos variables nominales/ordinales.

Cálculo de porcentajes de columna y de fila de casilla

$$\%\ \text{columna [de ocurrencia conjunta]} = \frac{\#\text{ en una casilla}}{\#\text{ total en una columna}}\ (100)$$

$$\%\ \text{de fila [de ocurrencia conjunta]} = \frac{\#\text{ en una casilla}}{\#\text{ total de fila}}\ (100)$$

Codificación y conteo de datos de intervalo/razón

Las variables con niveles de medición de intervalo/razón se distinguen de las variables nominales/ordinales por sus cualidades numéricas, sobre todo por sus unidades de medición, como millas, kilómetros, pulgadas, segundos y kilogramos. Tales variables "cuantitativas" nos permiten imaginar una regla y pensar linealmente, en términos de la distancia entre puntos sobre una línea recta. Es más, podemos hacer mediciones muy precisas, con el nivel de precisión determinado por las circunstancias prácticas. Por ejemplo, ¿es práctico medir la distancia a la escuela en términos de cientos de metros o en centésimas de milímetro? La precisión también está confinada por las limitaciones de los instrumentos de medición. Por ejemplo, la medida de distancia en centésimas de milímetro fue posible sólo después de que se inventó el microscopio. Con las variables de intervalo/razón, el grado de precisión se especifica según hasta dónde redondeamos los números.

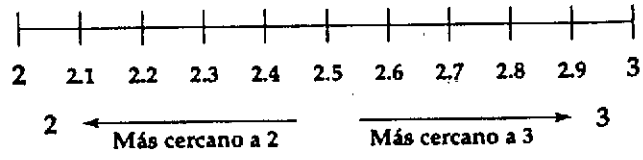
Redondeo de las observaciones de intervalo/razón

La observación de una variable de intervalo/razón tal vez no nos ofrezca la puntuación real porque sus mediciones a menudo pueden hacerse infinitamente de manera más precisa. Por ejemplo, podemos medir distancias hasta la unidad más cercana en kilómetros, metros, centímetros y así sucesivamente. Por lo tanto, redondeamos las puntuaciones de intervalo/razón hasta cierto grado de precisión elegido y especificado. De esta manera, reconocemos que el código registrado para la puntuación tiene algún error de medición. El **error de redondeo** es *la diferencia entre la puntuación real o perfecta (qué quizá nunca conozcamos) y nuestra puntuación observada y redondeada.* El error de redondeo depende de qué *posición decimal elegimos como nuestro nivel de precisión* —nuestra unidad de redondeo—. (Si es necesario, revise la ubicación de las posiciones decimales en el apéndice A.) Si decidimos medir el tiempo a la centésima de segundo más cercana, como se efectúa en los eventos olímpicos de pista, entonces nuestra unidad de redondeo es la posición de las centésimas.

El procedimiento para redondear una puntuación de una variable de intervalo/razón es como sigue:

1. Especifique la unidad de redondeo según su posición decimal.
2. Observe el número *a la derecha* de la unidad de redondeo y siga estas reglas.
   A. Si es 0, 1, 2, 3 o 4, redondee hacia el entero inferior.
   B. Si es 6, 7, 8 o 9, redondee hacia el entero superior.
   C. Si es 5, observe la siguiente posición decimal a la derecha y, si el número es 5 o mayor, redondee hacia el entero superior. Si no existe algún número en esa siguiente posición decimal, deje el redondeo en ese número.

Piense en el redondeo como un movimiento hacia el punto más cercano sobre una línea. Por ejemplo, si redondeamos al *entero más cercano* (la posición de las unidades), simplemente estamos moviendo al entero más cercano.



Por lo tanto, aquí 2.1, 2.2, 2.3 y 2.4 se redondean hacia abajo a 2 sólo porque están más cerca de 2 que de 3. Se ofrecen ejemplos adicionales en el apéndice A.

Los límites reales de puntuaciones redondeadas

Una vez que conocemos las puntuaciones redondeadas, *los números en la posición decimal de la unidad de redondeo* se consideran estimaciones. El valor real de una puntuación podría ser cualquiera de las puntuaciones que se redondean para obtener la puntuación registrada. Por ejemplo, suponga que medimos la altura de Jonathan a la pulgada más cercana y registramos 69 pulgadas. Varias horas des-

pués, con Jonathan ausente, Alan nos pregunta qué tan alto es Jonathan. Observamos nuestra tabla desglosada de datos y vemos el código de 69 pulgadas. En este momento, podemos afirmar que la altura real de Jonathan está entre $68\frac{1}{2}$ y $69\frac{1}{2}$ pulgadas, o 69 más menos media pulgada. Este *rango de posibles valores reales de una puntuación (ya) redondeada* se llama **límite real** o **límite verdadero** de la puntuación.

Los límites reales de una puntuación redondeada especifican el rango de números que podrían redondearse para obtener la puntuación registrada. En este sentido, calcular los límites reales es la inversa del redondeo. Por ejemplo, suponga que registramos cuánto tiempo toma a cada uno de 150 estudiantes completar un proyecto de laboratorio de química y lo redondeamos a la hora más cercana. También considere que 56 de estos estudiantes reciben una puntuación de dos horas. Algunos de ellos tomaron un poco menos de dos horas, y otros un poco más. Precisamente, un estudiante que registra dos horas pudo haber tomado tan poco como 90 minutos ($1\frac{1}{2}$ horas), el *límite real inferior*, o tanto como 150 minutos ($2\frac{1}{2}$ horas), el *límite real superior*. Una puntuación de dos horas *en realidad* significa entre $1\frac{1}{2}$ y $2\frac{1}{2}$ horas, por eso llamamos a este rango de tiempos los límites reales.

Límite real inferior de dos horas		Límite real superior de dos horas
$1\frac{1}{2}$ horas ←	2 horas	→ $2\frac{1}{2}$ horas
(90 minutos)	(120 minutos)	(150 minutos)

Calculamos los límites reales moviéndonos media unidad de redondeo en cada dirección, utilizando el siguiente procedimiento:

1. Enfóquese en la "unidad de redondeo", la posición decimal a la cual se redondeó la puntuación. Divida esta unidad de redondeo entre 2. (Precaución: No divida entre 2 el número en la posición decimal de la unidad de redondeo.)
2. Reste el resultado del paso anterior de la puntuación redondeada observada para obtener el límite real inferior.
3. Sume el resultado del paso (1) a la puntuación redondeada observada para obtener el límite real superior.

Por ejemplo, para los 56 estudiantes que registraron dos horas en el proyecto del laboratorio de química, redondeamos a la hora más cercana (la posición de las unidades). Dividimos *esta unidad de redondeo de una hora* entre 2 para obtener hora y media. Entonces restamos este resultado de la puntuación redondeada observada de dos horas para obtener el límite real inferior ($1\frac{1}{2}$ horas) y lo sumamos a la puntuación observada de dos horas para obtener el límite real superior ($2\frac{1}{2}$ horas). Incluso es improbable que uno de estos 56 estudiantes tomara exactamente dos horas para completar el proyecto; dos horas es una estimación redondeada. Podemos

tener la certeza, sin embargo, de que cada uno de los 56 terminaron entre $1\frac{1}{2}$ y $2\frac{1}{2}$ horas. Nuestro grado de precisión es la unidad de redondeo de una hora.

Los principios de inclusividad y exclusividad también se aplican a las variables de intervalo/razón. Para una variable como la edad, apegarse al principio de inclusividad parecería razonable; sólo registramos la "edad en el último cumpleaños". No obstante, para garantizar la inclusividad un cuestionario de investigación debe incluir las respuestas "se negó" y "no sabe". La exclusividad es razonable en cuanto que todas las mediciones se realicen de la misma manera, en este caso, la edad en el último cumpleaños. Si un encuestado dice que tiene 26 años, entonces registre 26, no 27 ni 25.

Distribuciones de frecuencias de proporciones y de porcentajes para variables de intervalo/razón

Las distribuciones de frecuencias de proporciones y porcentajes para variables de intervalo/razón se calculan de la misma forma que para variables nominales/ordinales, excepto que en lugar de categorías tenemos puntuaciones. Por ejemplo, si la Universidad Smithville tiene 10 000 estudiantes y 3 000 tienen 19 años, las frecuencias proporcionales y porcentuales para la puntuación de 19 años son

$$p\text{ [de 19 años de edad en la Universidad Smithville]} = \frac{f\text{ de 19 años}}{n} = \frac{3\,000}{10\,000} = .3000$$

$$\%\text{ [de 19 años en la Universidad Smithville]} = (p)\,(100) = 30.00\%$$

Si estos cálculos se realizan para todas las edades, los resultados se presentan como la distribución de frecuencias de porcentaje de la variable *edad* para la población de estudiantes de la Universidad de Smithville.

Distribuciones de frecuencias de porcentajes acumulados

La tabla 2-10 muestra la frecuencia, la frecuencia de porcentaje y las distribuciones de frecuencias de porcentajes acumulados de los niveles de educación de 20 cuidadores —parientes que acompañan a los pacientes con Alzheimer en una clínica— (Clair, Ritchey y Allman, 1993). Estas tres piezas de información son partes típicas de los resultados obtenidos por computadora porque juntos generan respuestas rápidas a una serie de preguntas. Obviamente, la frecuencia de puntuación bruta (f) proporciona una respuesta sobre cuántos sujetos recibieron una puntuación específica y la frecuencia porcentual estandariza la frecuencia de acuerdo al tamaño de la muestra. La información adicional en la tabla 2-10, la frecuencia de porcentajes acumulados, es una valiosa forma para observar las frecuencias de las puntuaciones en una distribución hasta, e inclusive una puntuación de interés. Ésta es la **frecuencia de porcentajes acumulados**, que es *la frecuencia porcentual de una puntuación y además las de todas las puntuaciones que la preceden en la distribución*. Por ejemplo, en el caso de los cuidadores en la tabla 2-10,

TABLA 2-10 Ilustración de una distribución de frecuencias porcentuales acumuladas: Años de educación formal entre cuidadores de pacientes con Alzheimer

Especificaciones		Cálculos	
Años de educación formal (X)	*Frecuencia (f)*	*Frecuencia porcentual*	*Porcentaje acumulado (f)*
5	1	5%	5%
6	1	5	10
7	1	5	15
9	2	10	25
10	1	5	30
11	1	5	35
12	10	50	85
14	2	10	95
16	1	5	100
Total	20	100%	100%

¿qué porcentaje tiene un nivel de educación hasta e inclusive el nivel preparatoria? Para obtener las frecuencias de porcentajes acumulados hacemos una lista con las puntuaciones, de la más baja a la más alta, y calculamos la frecuencia de porcentaje de cada puntuación. Entonces sumamos las frecuencias de porcentaje de la puntuación que nos interesa y todas las puntuaciones menores. En la tabla 2-10, 85 por ciento tenían 12 años de escolaridad o menos. Restando esta frecuencia de porcentajes acumulados del 100 por ciento, rápidamente podemos ver que sólo 15 por ciento de la muestra fueron más allá de la escuela preparatoria.

Percentiles y cuartiles

A menudo visualizamos una distribución de puntuaciones como fraccionada o "fracturada" en grupos que están encima y debajo de una puntuación o en grupos con iguales porcentajes de casos. Las distribuciones de frecuencias acumuladas proporcionan una herramienta para identificar **fractiles (*o cuartiles*)**, *puntuaciones que separan una fracción de los casos de una distribución*. Los rangos percentilares (o, simplemente, percentiles) son un fractil en común. *Entre los casos en una distribución de puntuaciones*, el **rango percentilar** es *el porcentaje de casos que caen en o están debajo de un valor específico de X*. Por ejemplo, en las frecuencias de porcentajes acumulados de la tabla 2-10, observamos que un cuidador con 14 años de educación tiene un nivel educativo igual o superior al 95 por ciento de la muestra, un rango percentilar de 95. Con frecuencia los percentiles se emplean en círculos de educación como una manera de ordenar grados o puntuaciones de pruebas. Por ejemplo, en una prueba de admisión a la universidad, un estudiante con una puntuación del percentil 90 o mayor calificaría para la admisión en una universidad de prestigio.

Los **cuartiles** son *fractiles que identifican los valores de la puntuación que fraccionan una distribución en cuatro grupos del mismo tamaño (es decir, 25 por ciento de los casos en cada grupo).* Cuando una distribución tiene un amplio rango de puntuaciones, los cuartiles se obtienen fácilmente de las distribuciones de frecuencias de porcentajes acumulados. El primer cuartil, Q_1, es el percentil 25; el segundo, Q_2, percentil 50; y el tercero, Q_3, percentil 75. Es común que los programas estadísticos para computadora estén diseñados para identificar cuartiles y otros fractiles, como deciles, los cuales fraccionan una distribución en 10 grupos del mismo tamaño.

La tabla 2-11 presenta la distribución de las calificaciones en un examen parcial (X) e ilustra la utilidad de los cuartiles. En esta clase de 20 estudiantes, el 25 por ciento inferior (o las cinco calificaciones más bajas) van de $X = 69$ a las calificaciones más bajas; el siguiente cuarto de estudiantes, de $X = 72$ a 84; el tercer cuarto, de $X = 85$ a 91, y el cuarto más alto de $X = 93$ a las calificaciones más altas. También vemos que un cuarto de los estudiantes obtuvo 69 o menos, sin conseguir una C; la mitad obtuvo más de 84, tres cuartos obtuvo 91 o menos, la mitad obtuvo entre 72 y 91, y así sucesivamente. Como lo analizaremos en el capítulo 3, un gráfico llamado gráfico de caja usa cuartiles para revelar mucho sobre la manera en que está formada una distribución de puntuaciones.

Fractiles

Puntuaciones que separan una fracción de los casos de una distribución.

Rango percentilar

Entre los casos de una distribución de puntuaciones, el porcentaje de casos que caen en, o están debajo, de un valor específico de X.

Cuartiles

Fractiles que identifican los valores de la puntuación que rompe una distribución en cuatro grupos del mismo tamaño.

Agrupamiento de datos de intervalo/razón

Algunas veces, para mejorar la claridad de la presentación en una tabla o gráfico, se agrupan o comprimen las distribuciones de intervalo/razón en un número menor de categorías ordinales. Por ejemplo, en un estudio sobre adultos en Estados Unidos, la variable edad variará desde 20 hasta cerca de 100, lo cual proporciona 80 puntuaciones. Es confuso presentar una tabla que enliste las frecuencias y las frecuencias de porcentaje de las 80 edades. Para hacerlo más

claro, clasificamos las edades en categorías de 10 años, como se muestra en la tabla 2-12.

Existe una cuestión muy importante respecto del agrupamiento de datos de intervalo/razón. Cuando agrupamos datos, eliminamos detalles, lo cual produce *error de agrupamiento*. Por ejemplo, la tabla 2-12 muestra que hubo 106 encuestados entre 40 y 49 años de edad. ¿Pero la mayoría de ellos estaba más cerca de 40 o de 49? No tenemos manera de saberlo con sólo observar las puntuaciones agrupadas. Si están disponibles los datos no agrupados, podemos usarlos para obtener cálculos de promedios más precisos. En general, en cualquier momento nos movemos de un nivel de medición "superior" a otro "inferior" (esto es, de razón a intervalo, de intervalo a ordinal, o de ordinal a nominal), perdemos información y ello limita lo que matemáticamente puede hacerse.

No obstante, al leer el trabajo de otros, tal vez se nos presenten datos agrupados sin los estadísticos correspondientes. En tales situaciones, intente contactar al autor para obtener las estadísticas descriptivas. Si esto no es posible, entonces pueden calcularse promedios y otros estadísticos a partir de los datos agrupados; pero dichos estadísticos incluirán error de agrupamiento. Véase Freund y Simon (1991: 70) para una discusión acerca del cálculo de estadísticos con datos agrupados.

TABLA 2-11 Cuartiles de una distribución de las calificaciones en un examen parcial

Especificaciones		Cálculos			
Calificación en el examen (X)	*f*	*f porcentual*	*f porcentual acumulada*	*Ubicación de los cuartiles (Q)*	
31	1	5.0%	5.0%		
58	1	5.0	10.0		
63	1	5.0	15.0		
68	1	5.0	20.0		
$Q_1 \longrightarrow$ 69	1	5.0	25.0	←	Q_1 = percentil 25
72	1	5.0	30.0		
76	1	5.0	35.0		
77	1	5.0	40.0		
82	1	5.0	45.0		
$Q_2 \longrightarrow$ 84	1	5.0	50.0	←	Q_2 = percentil 50
85	1	5.0	55.0		
86	2	10.0	65.0		
88	1	5.0	70.0		
$Q_3 \longrightarrow$ 91	1	5.0	75.0	←	Q_3 = percentil 75
93	2	10.0	85.0		
94	1	5.0	90.0		
95	1	5.0	95.0		
97	1	5.0	100.0%		
Total	20	100.0%			

TABLA 2-12 La variable de razón edad, agrupada en categorías ordinales de 10 años

Especificaciones			Cálculos
Código ordinal	*Grupo de edad*	*f*	*Porcentaje*
1	20-29	47	10.49%
2	30-39	68	15.18
3	40-49	106	23.66
4	50-59	96	21.43
5	60-69	53	11.83
6	70-79	45	10.04
7	80–89	24	5.36
8	90 y mayor	9	2.01
	Total	448	100.00%

☹ INSENSATEZ Y FALACIAS ESTADÍSTICAS ☹

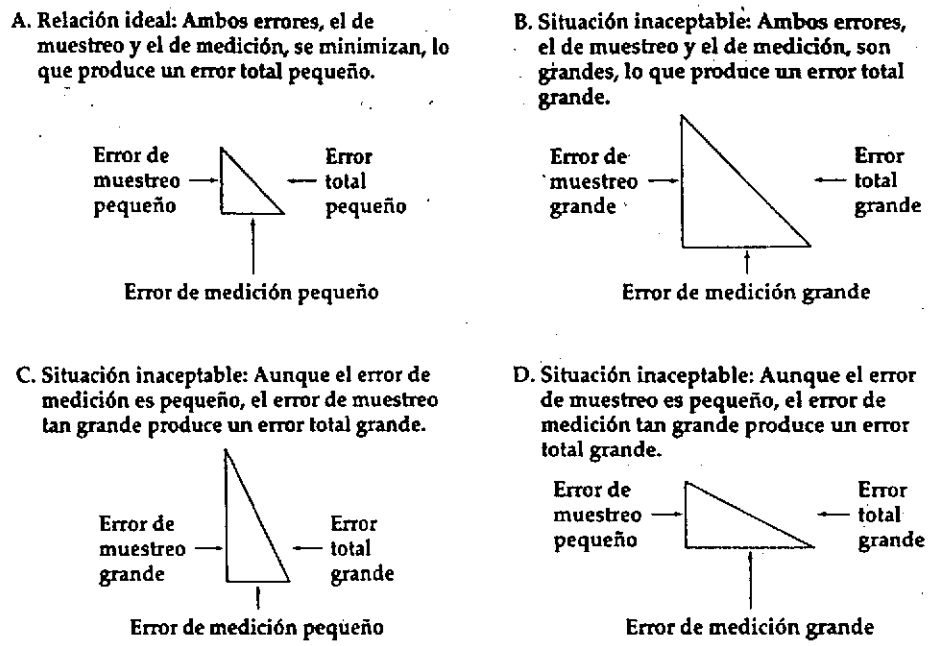
El fracaso para reducir los errores de muestreo y de medición

Para manejar de forma apropiada el error se requiere minimizar tanto el error de muestreo como el error de medición. Aun cuando tengamos una muestra representativa y ésta sea grande, esto por sí mismo no asegura resultados exactos; también debemos tomar buenas mediciones. Por ejemplo, al medir el ingreso en un estudio sobre hogares, una falla para distinguir aquellos hogares de un ingreso, de los de dos, no puede remediarse incrementando el tamaño de la muestra. No importa qué tan grande sea la muestra, y por ende qué tan pequeño sea el error de muestreo, nuestros datos sobre el ingreso no tienen valor. Una falacia estadística común es la creencia de que una muestra grande compensa mediciones pobres.

De la misma forma, la medición precisa no compensará una muestra insuficiente o no representativa, ambas producirán un error de muestreo grande. Por ejemplo, suponga que deseamos determinar el peso promedio de los hombres en un campus y utilizamos una báscula precisa para obtener un error de medición pequeño. Considere, sin embargo, que inadvertidamente obtenemos una muestra no representativa con una porción muy grande de hombres mayores. La medición precisa de peso está en duda por el error de muestreo grande, lo cual nos lleva a creer que el peso promedio de los hombres en el campus es mayor que el verdadero. Los resultados del estudio no tienen valor.

Resulta esencial que ambos errores, el de medición y el de muestreo, se minimicen para obtener conclusiones correctas y fidedignas. La relación entre el error de muestreo, el error de medición y el error total del estudio se ilustran con el teorema pitagórico en la figura 2-2. El matemático griego Pitágoras descubrió que la longitud de la hipotenusa (C), elevada al cuadrado, es decir, el lado del triángulo

FIGURA 2-2
La relación de los errores de muestreo y de medición con el error total



opuesto al ángulo recto, es igual a la suma de los cuadrados de las longitudes de los otros dos lados (A y B); esto es, $C^2 = A^2 + B^2$. Esto significa que la hipotenusa es grande si A o B son grandes. La figura 2-2*A* muestra que el error total es pequeño *sólo cuando ambos errores,* el de medición y el de muestreo, son pequeños. Los ejemplos *B*, *C* y *D* de la figura 2-2, indican cómo el fracaso para minimizar uno o ambos tipos de error produce un error total grande (Blalock, 1979: 573-574).

Para un estadístico verdaderamente cuidadoso, la imaginación estadística y el control del error constituyen desafíos que hacen que los retos de la ciencia y de la estadística sean viajes excitantes en vehículos bien diseñados. La imaginación estadística requiere tener la intuición y la claridad de pensamiento para prever las rutas más grandes que pueden surgir en el viaje. El control y el manejo del error involucran el mantenimiento del vehículo en curso para que se alcance el destino correcto. El viaje es divertido y, si se completa con éxito, es altamente gratificante.

Fórmulas en el capítulo 2

Cálculo de la frecuencia proporcional de una categoría o puntuación:

$$p\text{ [de la muestra total } (n) \text{ en una categoría]} = \frac{f \text{ de la categoría}}{n} = \frac{\#\text{ en la categoría}}{n}$$

Cálculo de la frecuencia porcentual de una categoría o puntuación:

$$\%\text{ [de una muestra total } (n) \text{ en una categoría]} = (p\text{ [de una muestra total } (n) \text{ en una categoría]})\ (100)$$

Cálculo de los porcentajes de columna y de fila de una casilla en una tabulación cruzada:

$$\%\text{ de columna [de ocurrencia conjunta]} = \frac{\#\text{ en una casilla}}{\#\text{ total en una columna}}(100)$$

$$\%\text{ de fila [de ocurrencia conjunta]} = \frac{\#\text{ en una casilla}}{\#\text{ total en una fila}}(100)$$

Cálculo de los límites reales de una puntuación de intervalo/razón:

1. Enfóquese en la "unidad de redondeo", la posición decimal hacia la cual se redondeó la puntuación. Divida esta unidad de redondeo entre 2. (*Precaución*: no divida el número en la posición decimal de la unidad de redondeo entre 2.)
2. Reste el resultado de (1) de la puntuación redondeada observada para obtener el límite real inferior.
3. Sume el resultado de (1) a la puntuación redondeada observada para obtener el límite real superior.

Cálculos para percentiles y cuartiles:

Prepare una tabla desglosada con los siguientes títulos:

Puntuación (X)	*f*	*Porcentaje f*	*Porcentajes acumulados f*

El percentil de una puntuación es su frecuencia acumulada. Los cuartiles se localizan en los percentiles 25, 50 y 75.

Preguntas para el capítulo 2

1. Explique la diferencia entre una observación y un estadístico.
2. Señale la diferencia entre un estadístico y un parámetro.
3. ¿Cómo se puede demostrar satisfactoriamente que los estadísticos de una sola muestra son sólo estimados de los parámetros de una población?
4. Aunque Karen creció lejos de cualquier barrio minoritario, se consideraba una persona abierta y sin prejuicios. Se volvió una trabajadora social y consiguió empleo en el área de recursos humanos que lleva el seguimiento de niños de grupos minoritarios que han sufrido abusos por padres drogadictos. Karen desarrolló abrumadores sentimientos racistas y se convenció de que las minorías son moralmente inferiores. En términos de sesgo de muestreo, ¿qué le sucedió a Karen?
5. Una autoproclamada experta en sexualidad femenina informó que tres cuartas partes de las mujeres declaran que odian a los hombres. El conductor del programa de entrevistas desafía dicho estadístico. La experta responde diciendo que su muestra de más de 6 000 mujeres es la más grande jamás investigada

sobre el tema, y que sus resultados tienen un rango de error de menos de un décimo de 1 por ciento. Su método de muestreo fue un cuestionario enviado por correo dentro de una popular revista femenina. Señale las maneras en las que su muestra podría estar sesgada.

6. Por medio de una encuesta telefónica, usted quiere evaluar la opinión del público sobre la propuesta del gobierno para aumentar los impuestos para construir un nuevo zoológico. Si usted seleccionara aleatoriamente cada 100 números del directorio telefónico del área, ¿obtendría una muestra representativa? Explíquelo en términos de sesgo de la muestra.
7. El Dr. Burnett prueba una nueva droga antidepresiva en un hospital para enfermos mentales. Se encuentra que la droga es bastante eficaz. ¿Tendrá la droga necesariamente la misma efectividad para todas las personas que padecen depresión? Explique en términos de sesgo de la muestra.
8. La esquizofrenia es una enfermedad mental caracterizada por delirios y pérdida general de contacto con la realidad. Contar el número de veces que una persona ataca con violencia a los miembros de su familia, ¿sería una medición válida de esquizofrenia? Explique.
9. La religiosidad es una medición de qué tan religiosa es una persona. En su encuesta, Tomás intenta medirlo preguntando a los encuestados qué tan a menudo acuden a servicios eclesiásticos. ¿Sería ésta una medición válida y confiable? Explique.
10. Suponga que estamos midiendo el peso entre una muestra de fisicoculturistas y registramos que Sam pesa como 114 kilogramos. Él realmente pesa entre 113.5 y 114.5 kilogramos. Estos pesos son los límites _____ y _____ de una puntuación de 114 kilogramos.
11. Jacqueline y Evelin realizaron estudios en ciencias sociales en la universidad, y les preguntaron cuántos esperan ser empleados en la industria después de graduarse. Jacqueline, con una muestra de 650, obtuvo un estimado de 31 por ciento. Evelin, con una muestra de 45, obtuvo un estimado de 23 por ciento. ¿En cuál estimación podemos tener más confianza?, ¿por qué en términos de control del error de muestreo, ¿qué más necesitamos saber sobre las muestras de Jacqueline y Evelin? ¿Por qué?
12. Bob compró leña para su chimenea, la cual mide 90 centímetros a lo ancho. Para asegurarse de que los leños cupieran, tomó una vara de 60 centímetros de largo y se aseguró de que la longitud de cada leño variara, si acaso, 15 centímetros con respecto al palo. Sara diseña circuitos de microcomputadoras. Uno de sus instrumentos de trabajo mide al milímetro más cercano, pero este instrumento es menos preciso que la vara de Bob. Explique.
13. Brian encuestó a 5 000 personas para un estudio sobre la salud. Su definición operacional de nivel de salud es "visitas al médico", es decir, el número de veces que un encuestado fue al médico durante el último año.

    a) ¿Sería ésta una medición válida y confiable para el nivel de salud? Explique.
    b) ¿Sería inteligente que alardeara de su pequeño error de muestreo? Explique.

Ejercicios para el capítulo 2

1. Indique el nivel de medición de las siguientes variables.

Número y nombre de la variable	*Definición operacional y codificación (cómo se mide y registra la variable)*	*Nivel de medición*
a) PROMEDIO (GPA)	Promedio de las calificaciones: número de puntos de calidad académica ganados divididos entre el número de horas-crédito cubiertas	
b) ALTURA	Altura física en centímetros	
c) PENAMUER	Escala de actitud de 10 reactivos acerca del apoyo a la pena de muerte con una suma total que va de 0 a 40	
d) AUTOPLAC	Número de la placa del automóvil: el número de 7 dígitos	
e) ESTLAB	Estatus laboral: 1 = inexperto, 2 = semiexperto, 3 = experto	
f) POBCIUD	Población dentro de los límites de la ciudad	
g) MARCAUTO	Marca de automóvil: Ford, Buick, Toyota, etc.	

2. Indique el nivel de medición de las siguientes variables.

Número y nombre de la variable	*Definición operacional y codificación (cómo se mide y registra la variable)*	*Nivel de medición*
a) PESO	Peso físico en kilogramos	
b) DENSIDAD	Población (número de personas que residen en una área específica) por kilómetro cuadrado	
c) TASAMI	Tasa de mortalidad infantil: número de muertes durante el primer año de vida por 1 000 nacidos vivos	
d) ESTUDIAN	Estatus del estudiante: 1 = estudiante, 2 = graduado, 3 = especial	
e) ESTIMA	Valoración de autoestima: escala sumativa de 15 reactivos con puntuaciones que oscilan entre 0 y 60	
f) SATLAB	Satisfacción laboral: 0 = muy insatisfecho, 1 = insatisfecho, 2 = satisfecho, 3 = muy satisfecho	
g) ESPVIDA	Esperanza de vida: número promedio de años que se espera que los recién nacidos vivan (por lo común ajustado para sexo y edad)	

3. En un estudio sobre cuidadores de miembros familiares ancianos, se entrevistaron 293 mujeres respecto de las tensiones del cuidado y el mantenimiento del empleo al mismo tiempo. Identifique los niveles de medición para cada una de las siguientes variables del estudio:

Variables del cuidador	*Variables de la persona cuidada*
a) Distancia entre la casa y el trabajo	*d*) Género
b) Edad	*e*) Peso
c) Ocupación	*f*) Severidad de la condición

4. En un estudio sobre cuidadores de niños discapacitados, se entrevistaron 141 mujeres respecto de las tensiones del cuidado. Identifique los niveles de medición para cada una de las siguientes variables del estudio:

Variables del cuidador	*Variables del niño*
a) Calidad percibida del cuidado proporcionado (excelente, bueno, suficiente, pobre)	*c*) Diagnóstico o condición médica
b) Ingreso	*d*) ¿Se puede bañar solo?
	e) Temperatura corporal

5. Imagine que mediante un cuestionario a los cuidadores se les solicitan los siguientes datos. Con las categorías de respuesta proporcionadas, ¿cumple cada variable los principios de inclusividad y de exclusividad? Si no, ¿cómo pueden mejorarse para estar de acuerdo con tales principios?

   a) ¿Tiene usted una preferencia religiosa? Si es así, ¿cuál es? (Por favor, marque sólo un espacio.)
   _____ Protestante _____ Católica _____ Judía
   _____ Bautista _____ Ninguna

   b) ¿Cuál es su estado civil? (Por favor, marque sólo un espacio.)
   _____ Soltero _____ Casado

   c) Por favor, señale la categoría que mejor se ajusta a su ingreso familiar anual total de todas las fuentes.
   _____ $0-10 000 _____ $41 000-50 000
   _____ $11 000-$20 000 _____ $50 000 y mayor
   _____ $21 000-30 000

   d) ¿Actualmente está usted empleado fuera de su casa? (Por favor, marque sólo un espacio.)
   _____ Tiempo completo _____ Empleado _____ Medio tiempo
   _____ Desempleado

6. Sibckey *et al.* (1995) estudiaron la preocupación empática, como una motivación para el comportamiento altruista. Con las categorías de respuesta proporcionadas, ¿cada variable cumple los principios de inclusividad y de exclusivi-

dad? Si no, ¿cómo pueden mejorarse para estar de acuerdo con estos principios?

a) Usted fue motivado para ayudar por (marque sólo una):
_____ Calidez _____ Compasión _____ Simpatía

b) Tiene usted:
_____ Más de 25 años de edad _____ Menos de 25 años de edad

c) ¿Cuántas veces estaría dispuesto a ayudar a la misma persona?
_____ 0-5 _____ 6-10 _____ 11-15 _____ Ninguna

7. Redondee los siguientes números a la unidad de redondeo señalada:

a) 28.349 a la décima más cercana
b) 31.666 a la centésima más cercana
c) 587 al centenar más cercano
d) 25.6388 a la milésima más cercana
e) 25.6388 a la centésima más cercana
f) 25.6388 a la décima más cercana
g) 25.6388 a la decena más cercana

8. Redondee los siguientes números a la unidad de redondeo señalada:

a) 5.456 a la décima más cercana
b) 5.456 a la centésima más cercana
c) 20.821 a la centésima más cercana
d) 381 al centenar más cercano
e) 467 988 al millar más cercano
f) 467 988 a la centena de millar más cercana
g) .00051 a la milésima más cercana

9. Especifique los límites reales de los siguientes números redondeados:

a) 4.0 centímetros
b) 4 centímetros
c) Edad de 5 redondeada al último cumpleaños
d) Edad de 5 redondeada al cumpleaños más cercano
e) 3 300 redondeado al centenar más cercano
f) 3 360 redondeado al 10 más cercano
g) .0030
h) .068

10. Especifique los límites reales de los siguientes números redondeados:

a) 5.00 kilogramos
b) 5 kilogramos
c) Edad de 9 redondeados al último cumpleaños
d) Edad de 9 redondeados al cumpleaños más cercano
e) 71 000 redondeados al centenar más cercano

f) 9 680 redondeado al décimo más cercano
g) .01605
h) .248

11. A continuación se presenta la distribución del parentesco de los familiares cuidadores con pacientes de Alzheimer. Calcule la razón de mujeres a hombres cuidadores.

Parentesco del familiar cuidador con el paciente	*f*
Esposa	114
Esposo	17
Hija	37
Hijo	4
Hermana	8
Hermano	1
Madre	2
Nuera	13
Otro pariente femenino	24
Otro pariente masculino	2

12. A continuación se presenta la distribución de trabajadores por posición en una empresa de comunicaciones (datos ficticios). Calcule la razón de otros empleados a gerentes.

Posición en la empresa	*f*
Posiciones directivas	
Presidente	1
Director ejecutivo	1
Vicepresidente	3
Vicepresidente auxiliar	8
Ayudante administrativo	8
Jefe de personal de piso	12
Otras posiciones	
Secretaria	16
Vendedor	42
Empleado de oficina	18
Técnico/profesional	14
Mantenimiento	3

13. A continuación se presentan datos sobre el número de vehículos registrados para una muestra aleatoria de 20 hogares en el condado de Madison.

a) Compile los datos en una tabla de distribución de frecuencias con columnas para la frecuencia, la frecuencia proporcional, la frecuencia porcentual y la frecuencia de porcentajes acumulados. (No se requiere mostrar las fórmulas.)

b) Si un hogar tiene tres vehículos registrados, ¿cuál es el rango percentilar de la familia? Interprete su respuesta.

2, 1, 2, 4, 2, 3, 4, 2, 1, 4, 2, 1, 0, 3, 2, 4, 3, 4, 2, 2

14. Pearson *et al.* (1997) demostraron que cuando una abuela vive con la familia, probablemente se involucrará en las actividades paternales. Suponga que los siguientes datos representan el número de órdenes dadas a 25 niños por sus abuelas.

a) Compile los datos en una tabla de distribución de frecuencias con columnas para la frecuencia, la frecuencia proporcional, la frecuencia porcentual y la frecuencia de porcentajes acumulados. (No requiere mostrar las fórmulas.)
b) Si una abuela diera dos órdenes, ¿cuál sería su rango percentilar? Interprete su respuesta.

5, 4, 3, 3, 6, 5, 3, 2, 4, 7, 5, 6, 2, 3, 4, 8, 7, 5, 6, 4, 2, 1, 5, 7, 3

15. Se preguntó a una muestra aleatoria de 1 888 profesionales de la salud pública de cuatro países qué tan importante creían que era para la Comunidad Europea formular y vigilar una política de salud nacional europea. Los resultados se presentan en la siguiente tabulación cruzada.

a) En general, ¿qué porcentaje de encuestados considera que una política de salud es muy importante?
b) ¿Qué porcentaje considera que una política de salud es poco importante?
c) Complete la tabla insertando la columna de porcentajes. Muestre la fórmula para el cálculo de una entrada.
d) Haga un comentario sobre cualquier resultado que destaque.

Respuesta	*Bélgica (% col.)*	*Francia (% col.)*	*Alemania (% col.)*	*Países Bajos (% col.)*
Muy importante	88	49	234	194
Importante	226	41	375	355
Poco	16	51	53	56
Nada importante	3	2	3	7
No sabe	47	13	21	54

16. Se estudia una muestra aleatoria de 641 estudiantes de primer año universitario para determinar las formas en las cuales el antecedente económico es una ventaja para alcanzar el éxito académico. La siguiente tabulación cruzada presenta el ordenamiento de las clases sociales, con base en la posesión de una computadora personal (datos ficticios).

a) En general, ¿qué porcentaje de encuestados posee una computadora personal?

b) Complete la tabla insertando los porcentajes por filas. Muestre la fórmula para el cálculo de una entrada
c) Haga un comentario sobre cualquier resultado que destaque.

Nivel de clase social	**Tiene computadora personal propia**	
	No (% fila)	*Sí (% fila)*
Baja (pobreza)	62	2
Trabajadora	79	33
Media baja	83	148
Media alta	36	142
Alta	1	55

17. Como gerente de control de calidad en una nueva cadena de neverías, usted visita dos tiendas locales y aleatoriamente registra las selecciones de sabor del helado de cada quinto cliente, checando (✓) la siguiente lista.

a) Compile los datos en una tabla de distribución de frecuencias con columnas para la frecuencia, la frecuencia proporcional y la frecuencia porcentual. (No requiere mostrar las fórmulas.)
b) ¿Tendría sentido incluir la frecuencia acumulada? En caso negativo, ¿por qué?

Vainilla	✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓
Chocolate	✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓
Fresa	✓✓✓✓✓✓✓✓✓
Galletas y crema	✓✓✓✓✓✓✓✓
Otro	✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓

18. Una funcionaria de encuestas en el lugar de la votación para una elección primaria tabula la participación de los partidos políticos. Cuando los votantes piden sus papeletas electorales, ella registra (✓) de qué partido fue la papeleta que solicitaron, en una lista con el nombre de los partidos.

a) Compile sus datos en una tabla de distribución de frecuencias con columnas para la frecuencia, la frecuencia proporcional y la frecuencia porcentual. (No requiere mostrar las fórmulas.)
b) ¿Tendría sentido incluir la frecuencia acumulada? En caso negativo, ¿por qué?

Partido Demócrata	✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓
Partido Republicano	✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓✓
Nuevo Partido Independiente	✓✓✓✓✓✓

19. Los registros de los siguientes datos proporcionan nombres de sujetos y una indicación de si ellos han consultado un oculista en el último año. Compile los

datos en una tabulación cruzada y conteste las siguientes preguntas (muestre las fórmulas).

a) ¿Qué porcentaje de sujetos son hombres?
b) ¿Qué porcentaje de sujetos consultaron a un oculista en el último año?
c) ¿Qué porcentaje de hombres consultaron a un oculista en el último año?
d) ¿Qué porcentaje de mujeres vieron a un oculista en el último año?

David—Sí	Pedro—No	Cristina—No	María—Sí
Elizabeth—No	Antonio—No	Sean—Sí	Manuel—No
Miguel—No	Samuel—Sí	Arturo—No	Cynthia—No
Eudora—Sí	Leroy—No	Bernabé—Sí	Vanessa—Sí
Enrique—No	Paula—Sí	Ernesto—No	Sabrina—No
Gary—Sí	Elena—No	Marco—Sí	Kevin—No
Bárbara—No	Sonia—Sí	Roberto—No	Cathy—Sí
Andrea—No	Andrés—Sí	Nancy—No	Laura—Sí
Donald—Sí	Carolina—No	Rebeca—Sí	René—No
Kimberly—No	Ginger—Sí	Débora—No	Rafael—No

20. Smetana (1989) sostiene que los conflictos menores pero persistentes entre los adolescentes estadunidenses y sus padres son una parte normal del desarrollo humano. Suponga que los siguientes datos son de un estudio sobre adolescentes, en el cual se les pregunta si tuvieron un conflicto con su padre o madre en las últimas 48 horas. Compile los datos en una tabulación cruzada y conteste las siguientes preguntas (muestre las fórmulas).

a) ¿Qué porcentaje de los encuestados son mujeres?
b) ¿Qué porcentaje de los encuestados tuvieron un conflicto con sus padres?
c) ¿Qué porcentaje de hombres tuvieron un conflicto con sus padres?
d) ¿Qué porcentaje de mujeres tuvieron un conflicto con sus padres?

Peggy—Sí	Jason—Sí	Cristina—No	Marco—Sí
Jeremías—Sí	Paty—Sí	Jaime—Sí	Andrés—No
Anita—No	Kyle—Sí	Julia—Sí	David—No
Linda—Sí	Alejandro—No	Donna—No	María—Sí
Beth—Sí	Wes—No	Janice—No	Rhonda—Sí

21. Para una de sus clases de tamaño moderado, diferente de su clase de estadística, prepare una tabla desglosada con las variables género, raza y color predominante de camisa, blusa o suéter. Elabore las distribuciones de frecuencias y de frecuencia porcentual para tales variables. Construya una tabulación cruzada de género por raza.
22. Para una de sus clases de tamaño moderado, diferente de su clase de estadística, prepare una tabla con las variables género, raza y proximidad hacia la parte

frontal del salón (la mitad de enfrente o la mitad de atrás de las filas). Elabore las distribuciones de frecuencias y de frecuencia porcentual para tales variables. Construya una tabulación cruzada de género por proximidad hacia la parte frontal del salón.

Aplicaciones opcionales en computadora para el capítulo 2

Si su clase usa las aplicaciones de computadora optativas que acompañan este texto, abra los ejercicios del capítulo 2 en el disco compacto de *Computer Applications for The Statistical Imagination*. Los ejercicios involucran 1) crear y salvar archivos de datos (*.SAV) utilizando el *SPSS para Windows*, 2) producir distribuciones de frecuencias, cuartiles y percentiles, y 3) administrar y salvar archivos de resultados.

El control de la calidad de la codificación y entrada de los datos es muy importante para minimizar los errores. Para alcanzar estándares científicos y éticos, un investigador debe ser capaz de mirar a sus colegas a los ojos y asegurarles honestamente que el conjunto de datos no tiene ningún error aleatorio. Aparte del aspecto ético, el cometer errores en el acceso de los datos representa un desperdicio de tiempo y energía en las posteriores fases del análisis. Un investigador ético pero descuidado quizá pase meses analizando un conjunto de datos y después encuentre datos que son incorrectos. El descubrimiento de un solo dato incorrecto requiere de un nuevo y completo análisis de los mismos. Así, debemos verificar diligentemente la exactitud de un conjunto de datos antes de empezar el análisis.

A continuación se ofrecen algunos lineamientos para la detección de errores de codificación y de entrada de datos autocodificados o archivos de datos existentes.

Lineamientos de control de calidad para la entrada de datos

1. Asegúrese de que los valores de los códigos ingresados son consistentes con el registro de codificación y los instrumentos de medición (como cuestionarios).
2. Si posee habilidades limitadas en mecanografía o su vista es mala, solicite que un asistente lea y verifique dos veces los datos. De manera alternativa, use un programa de entrada de datos diseñado especialmente, como *dBase*©. El *SPSS* aceptará sin dificultad archivos *dBase*, así como archivos de datos de hoja de cálculo *Lotus 1-2-3*©, *Excel*©, *Quattro Pro*© y *Multiplan*©.
3. Verifique una vez más los códigos en una copia impresa que enliste todas las variables y sus códigos.
4. Copias impresas de las distribuciones de frecuencias para verificar los códigos raros (es decir, códigos que no deben estar presentes en los datos).

Procedimientos de control de calidad específicos se describen en los ejercicios de aplicación de computadora del capítulo 2 en el disco compacto de *Computer Applications for the Statistical Imagination*.