capítulo

10 Análisis de datos cuantitativos

Al analizar los datos cuantitativos debemos recordar dos cuestiones: primero, que los modelos estadísticos son representaciones de la realidad, no la realidad misma; y segundo, los resultados numéricos siempre se interpretan en contexto, por ejemplo, un mismo valor de presión arterial no es igual en un bebé que en una persona de la tercera edad.

Roberto Hernández-Sampieri

Proceso de investigación cuantitativa

Paso 9 Analizar los datos

- Decidir el programa de análisis de datos que se utilizará.
- Explorar los datos obtenidos en la recolección.
- Analizar descriptivamente los datos por variable.
- Visualizar los datos por variable.
- Evaluar la confiabilidad, validez y objetividad de los instrumentos de medición utilizados.
- Analizar e interpretar mediante pruebas estadísticas las hipótesis planteadas (análisis estadístico inferencial).
- Realizar análisis adicionales.
- Preparar los resultados para presentarlos.

Objetivos de aprendizaje

Al terminar este capítulo, el alumno será capaz de:

1. Revisar el proceso para analizar los datos cuantitativos.
2. Reforzar los conocimientos estadísticos fundamentales.
3. Comprender las principales pruebas o métodos estadísticos desarrollados, así como sus aplicaciones y la forma de interpretar sus resultados.
4. Diferenciar la estadística descriptiva y la inferencial, la paramétrica y la no paramétrica.
5. Analizar la interrelación entre distintas pruebas estadísticas.

Síntesis

En el capítulo se presentan brevemente los principales programas computacionales de análisis estadístico que emplea la mayoría de los investigadores, así como el proceso fundamental para efectuar análisis cuantitativo. Asimismo, se comentan, analizan y ejemplifican las pruebas estadísticas más utilizadas. Se muestra la secuencia de análisis más común, con estadísticas descriptivas, análisis paramétricos, no paramétricos y multivariados. En la mayoría de estos análisis, el enfoque del capítulo se centra en los usos y la interpretación de los métodos, más que en los procedimientos de cálculo, debido a que los análisis se realizan con ayuda de una computadora.

Se realiza mediante programas computacionales como:

- SPSS®
- Minitab
- SAS
- STATS

Análisis de datos cuantitativos

Cuyo procedimiento es:

Fases

1. Seleccionar el programa estadístico para el análisis de datos
2. Ejecutar el programa
3. Explorar los datos: analizarlos y visualizarlos por variable del estudio
4. Se evalúa la confiabilidad y validez del o de los instrumentos escogidos
5. Se lleva a cabo análisis estadístico descriptivo de cada variable del estudio
6. Se realizan análisis estadísticos inferenciales respecto a las hipótesis planteadas
7. Se efectúan análisis adicionales
8. Se preparan los resultados para presentarlos

El análisis se realiza tomando en cuenta los niveles de medición de las variables y mediante la estadística, que puede ser

Descriptiva

- Distribución de frecuencias
- Medidas de tendencia central
  - Media
  - Mediana
  - Moda
- Medidas de variabilidad
  - Rango
  - Desviación estándar
  - Varianza
- Gráficas
- Puntuaciones *z* (en centro de recursos en línea)

Inferencia

- Sirve para estimar parámetros y probar hipótesis
- Se basa en la distribución muestral

- Análisis paramétrico
  - Coeficientes de correlación
  - Regresión lineal
  - Prueba *t*
  - Prueba de la diferencia de proporciones
  - Análisis de varianza
  - Análisis de covarianza (en centro de recursos en línea)
- Análisis no paramétrico
  - *Chi* cuadrada
  - Coeficientes de Spearman y Kendall
  - Coeficientes para tabulaciones cruzadas
- Análisis multivariados (centro de recursos en línea)

Nota: Este capítulo se complementa con uno adicional que se puede descargar del centro de recursos en línea, en: Material complementario → Capítulos → Capítulo 8, "Análisis estadístico: segunda parte"; junto con el documento 2, "Fórmulas y procedimientos estadísticos" y el apéndice 4 "Tablas anexas", que también pueden descargarse del mismo sitio .

Centro de recursos en línea

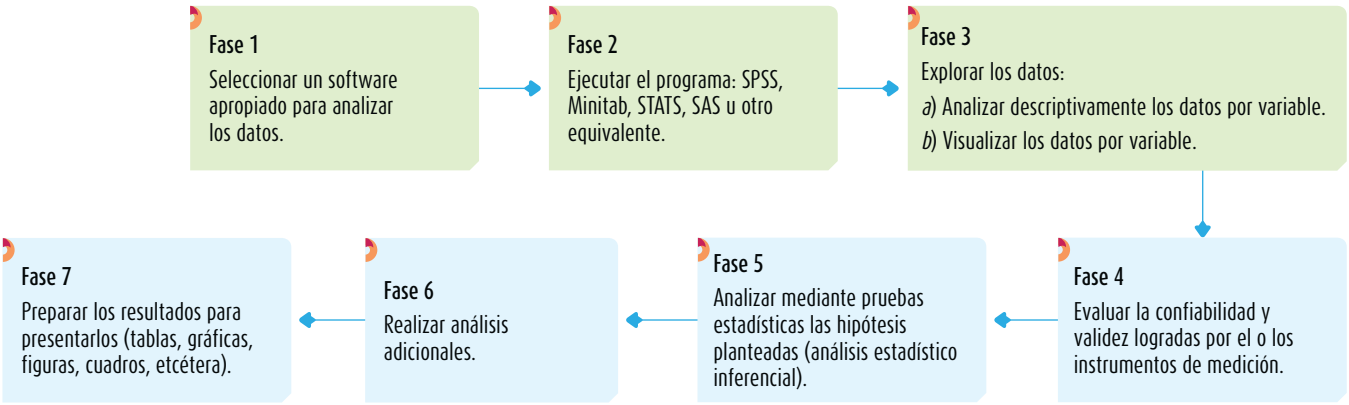
¿Qué procedimiento se sigue para analizar cuantitativamente los datos?

1 Una vez que los datos se han codificado, transferido a una matriz, guardado en un archivo y "limpiado" los errores, el investigador procede a analizarlos.

En la actualidad, el análisis cuantitativo de los datos se lleva a cabo *por computadora u ordenador*. Ya casi nadie lo hace de forma manual ni aplicando fórmulas, en especial si hay un volumen considerable de datos. Por otra parte, en la mayoría de las instituciones de educación media y superior, centros de investigación, empresas y sindicatos se dispone de sistemas de cómputo para archivar y analizar datos. De esta suposición parte el presente capítulo. Por ello, se centra en la *interpretación de los resultados de los métodos de análisis cuantitativo* y no en los procedimientos de cálculo.

El análisis de los datos se efectúa sobre la *matriz de datos* utilizando un *programa computacional*. El proceso de análisis se esquematiza en la figura 10.1. Posteriormente lo veremos paso a paso.

Figura 10.1 Proceso para efectuar análisis estadístico.



Paso 1: seleccionar un programa de análisis

Hay diversos programas para analizar datos. En esencia su funcionamiento es muy similar e incluyen las dos partes o segmentos que se mencionaron en el capítulo anterior: una parte de definiciones de las variables, que a su vez explican los datos (los elementos de la codificación ítem por ítem o indicador por indicador), y la otra parte, la matriz de datos. La primera parte es para que se comprenda la segunda. Las definiciones, desde luego, las prepara el investigador. Lo que éste hace, una vez recolectados los datos, es precisar los parámetros de la matriz de datos en el programa (nombre de cada variable en la matriz —que equivale a un ítem, reactivo, indicador, categoría o subcategoría de contenido u observación—, tipo de variable o ítem, ancho en dígitos, etc.) e introducir o capturar los datos en la matriz, la cual es como cualquier hoja de cálculo. Asimismo, recordemos que la matriz de datos tiene columnas (variables, ítems o indicadores), filas o renglones (casos) y celdas (intersecciones entre una columna y un renglón). Cada celda contiene un dato (que significa un valor de un caso en una variable). Supongamos que tenemos cuatro casos o personas y tres variables (género, color de cabello y edad); la matriz se vería como se muestra en la tabla 10.1.

Tabla 10.1 Ejemplo de matriz de datos con tres variables y cuatro casos

Caso	Columna 1 (género)	Columna 2 (color de pelo)	Columna 3 (edad)
1	1	1	35
2	1	1	29
3	2	1	28
4	2	4	33

La codificación (especificada en la parte de las definiciones de las variables o columnas que corresponden a ítems) sería:

- Género (1 = masculino y 2 = femenino).
- Color de cabello (1 = negro, 2 = castaño, 3 = pelirrojo, 4 = rubio).
- Edad (dato "bruto o crudo" en años).

De esta forma, si se lee por renglón o fila (caso), de izquierda a derecha, la primera celda indica un hombre (1); la segunda, de cabello negro (1); y la tercera, de 35 años (35). En el segundo renglón, un hombre de cabello negro y 29 años. La tercera fila, una mujer de cabello color negro, con 28 años. La cuarta fila (caso número cuatro) nos señala una mujer (2), rubia (4) y de 33 años (33). Pero, si leemos por columna o variable de arriba hacia abajo, tendríamos en la primera (género) dos hombres y dos mujeres (1, 1, 2, 2).

Por lo general, en la parte superior de la matriz de datos aparecen las opciones de los comandos para operar el programa de análisis estadístico como cualquier otro software (Archivo, Edición o Editar datos, etc.). Una vez que estamos seguros que no hay errores en la matriz, procedemos a realizar el análisis de la misma, el análisis estadístico. En cada programa tales opciones varían, pero en cuestiones mínimas.

Ahora, comentaremos brevemente los programas más importantes y de dos de ellos señalaremos sus comandos generales.

Statistical Package for the Social Sciences o Paquete Estadístico para las Ciencias Sociales (IBM® SPSS)

El SPSS (Paquete Estadístico para las Ciencias Sociales), desarrollado en la Universidad de Chicago, es uno de los más difundidos y actualmente es propiedad de IBM®. Contiene todos los análisis estadísticos que se describirán en este capítulo. En Iberoamérica, algunas instituciones educativas tienen versiones antiguas del SPSS; otras, versiones más recientes, ya sea en español o inglés y en los distintos sistemas operativos como Windows, Macintosh y UNIX.

Centro de recursos en línea

Como ocurre con otros programas, el IBM® SPSS se actualiza constantemente con versiones nuevas en varios idiomas.[1]

Asimismo, cada año surgen textos o manuales acordes con estas nuevas versiones. Sin embargo, en el centro de recursos el lector encontrará un manual que abarca las cuestiones esenciales de este paquete de análisis. Lo mejor para mantenerse al día en materia de SPSS® es consultar el sitio de IBM®que corresponda a su país (con las palabras clave: "IBM SPSS" puede encontrarlo mediante cualquier motor de búsqueda como Google). Asimismo, se puede "bajar" o "descargar" a la computadora una demostración del programa por tiempo limitado. Para la actualización de manuales, las palabras clave serían: "SPSS manuals" o "SPSS manuales".

La empresa IBM® afirma que se pueden solucionar diversos problemas de investigación con la suite de software IBM® SPSS Statistics, de la cual la versión "Base" contiene alrededor de 80% de los análisis. Para cuestiones más específicas se pueden adquirir diversos módulos que la compañía vende por separado, los cuales se integran a la versión *Base* con otros títulos o nombres.

Hasta agosto de 2013, las versiones más recientes tenían mejoras como mayor facilidad de uso, nuevas técnicas de análisis, mayor rendimiento y mejor integración con los demás productos IBM® (IBM SPSS®, 2013)

Como dijimos, el IBM® SPSS contiene las dos partes citadas que se denominan: *a*) vista de variables (para definiciones de las variables y consecuentemente, de los datos) y *b*) vista de los datos (matriz de datos). En ambas vistas se observan los comandos para operar en la parte superior.

[1] En el otoño de 2013, la familia de IBM SPSS Statistics contaba con alrededor de 30 productos con diferentes grados de complejidad. A éstos habría que añadirles los productos de las familias Collaboration and Deployment Services, Analytical Decision Management, Data Collection y Modeler.

El paquete IBM® SPSS trabaja de una manera muy sencilla: éste abre la matriz de datos y el investigador usuario selecciona las opciones más apropiadas para su análisis, tal como se hace en otros programas.

A continuación se describen de manera general algunas funciones principales del programa. Para profundizar y aprender su uso recomendamos revisar el manual de IBM SPSS®, que puede encontrar en el centro de recursos en línea (en el apartado de "Manuales auxiliares").

File (archivo): este botón tiene las siguientes funciones:

- Nuevo. Sirve para construir un archivo nuevo.
- Abrir. Se utiliza para abrir un archivo de datos previamente generado, de sintaxis, resultados o de proceso.
- Abrir base de datos. Tiene la función de generar, editar y ejecutar consultas en bases de datos previamente creadas.
- Leer datos de texto. Se usa para abrir archivos de texto.
- Cerrar. Como es común, sirve para cerrar el archivo en el que se está trabajando.
- Guardar. De igual forma, su función es guardar el archivo que se encuentra en uso.
- Guardar como. Se utiliza para guardar el archivo en el que se está trabajando con un nombre distinto al que ya tiene.
- Mostrar información del archivo de datos. Se trata de un archivo de trabajo o archivo externo.
- Caché de los datos. Es una copia temporal de los datos, la cual ayuda a mejorar el rendimiento cuando los archivos grandes de datos se leen desde una fuente externa. Si bien el archivo activo virtual puede ayudar a disminuir la cantidad de espacio en disco temporal, el no tener una copia temporal del archivo en activo provoca que la fuente de datos original se tenga que leer en cada procedimiento.
- Repositorio. Sirve para conectar, almacenar desde SPSS Statistics, publicar en la web, añadir un archivo, recuperar en SPSS Statistics o descargar un archivo.
- Presentación preliminar. Muestra en pantalla completa la tarea que se está llevando a cabo.
- Imprimir. Se utiliza para imprimir la tarea actual.
- Datos usados recientemente. Indica los datos que se ocuparon recientemente.
- Archivos usados recientemente. Muestra los últimos archivos utilizados.
- Salir. Cierra el programa IBM SPSS®.

Edit (edición): se emplea para modificar archivos, manipular la matriz (deshacer y rehacer acciones; cortar, copiar y pegar datos; pegar, borrar e insertar variables o casos), buscar y reemplazar datos e ir a casos o variables particulares entre otras acciones de edición.

View (ver): como su nombre lo dice es para visualizar la barra de estado, editor de menús, fuentes, cuadrícula, etiquetas de valor, marcar datos imputados, personalizar la vista de variables, etcétera.

Data (datos): en este menú se pueden definir las propiedades de las variables así como el nivel de medición, copiar propiedades de datos, establecer un nuevo atributo personalizado, definir fechas y conjuntos de respuestas múltiples, validación, identificación de casos duplicados y atípicos, ordenar casos y variables, trasponer, fundir archivos, reestructurar, agregar, diseño ortogonal, copiar conjunto de datos, segmentar archivos, seleccionar y ponderar casos.

Transform (transformar): con este botón se despliegan las opciones de calcular variable (crear variables compuestas por varios ítems o indicadores), contar valores dentro de los casos, valores de cambio, recodificar en las mismas o en distintas variables de manera personalizada o automática, agrupación visual, intervalos óptimos, preparar datos para modelado, asignar rangos a casos, asistente para fecha y hora, crear serie temporal, reemplazar valores perdidos y generar números aleatorios.

Analyze (analizar): por medio de esta opción se pueden solicitar análisis estadísticos que básicamente serían:

1. Informes (resúmenes de casos, información de columnas y renglones)
2. Estadísticos descriptivos (tablas de frecuencias, medidas de tendencia central y dispersión, razones, tablas de contingencia)

3. Tablas (personalización de las tablas)
4. Comparar medias (prueba *t* y análisis de varianza —ANOVA— unidireccional)
5. Modelo lineal general (independiente o factor y dependiente, con covariable)
6. Modelos lineales generalizados
7. Modelos mixtos
8. Correlaciones (bivariada —dos— y multivariadas —tres o más—) para cualquier nivel de medición de las variables
9. Regresión (lineal, curvilineal y múltiple)
10. Loglineal
11. Redes neuronales
12. Clasificación (conglomerados y análisis discriminante)
13. Reducción de dimensiones (análisis de factores)
14. Escala (fiabilidad y escalamiento multidimensional)
15. Pruebas no paramétricas
16. Predicciones
17. Supervivencia
18. Respuesta múltiple (escalas)
19. Análisis de valores perdidos
20. Imputación múltiple
21. Muestras complejas
22. Control de calidad
23. Curva COR

Direct marketing (marketing directo): mediante esta nueva función se pueden clasificar y agrupar los datos de, en el caso de las empresas, sus clientes para obtener una comprensión más profunda de éstos.

Graphs (gráficos): con esta función se solicitan gráficos (barras en formato unidimensional y 3D, líneas, áreas, de sectores o pastel, máximos y mínimos, diagramas de caja, barras de error, pirámide de población, dispersión, histograma, etcétera).

Utilities (utilidades o herramientas): se definen ambientes, conjuntos, información sobre variables, etcétera.

Window (ventana): sirve para moverse a través de archivos y hacia otros programas.

Help (ayuda): cuenta con contenidos de ayuda, cómo utilizar SPSS, comandos, guías, "asesor estadístico" y demás elementos aplicados al paquete (con índice).

Minitab®

Minitab es un paquete que goza de popularidad por su relativo bajo costo. Incluye un considerable número de pruebas estadísticas y cuenta con un tutorial para aprender a utilizarlo y practicar; además, es muy sencillo de manejar.

Minitab tiene un sitio web (http://www.minitab.com/) en el cual se puede descargar una versión de prueba gratuita por tiempo limitado.

Para comenzar a utilizar Minitab, se abre una sesión (la cual se define con nombre y fecha) y se abre una matriz u hoja de trabajo (en la parte superior de la pantalla aparece la sesión y en la parte inferior se presenta la matriz). Se definen las variables (*C* —columnas—): nombre, formato (numérico, texto, fecha/tiempo), ancho (en dígitos), su descripción y orden de los valores. Los renglones o filas son casos. Los análisis realizados aparecen en la sesión (parte o pantalla superior) y las gráficas se reproducen en recuadros.

Entre sus comandos están los siguientes:

Archivo: sirve para construir un nuevo archivo, localizar uno ya construido, guardar o abrir archivos, abrir una gráfica de Minitab, especificar impresora, imprimir, cerrar, entre otras funciones.

Editar: útil para modificar archivos, buscar datos, copiar, cortar y eliminar celdas, conectar Minitab con otras aplicaciones, etcétera.

Datos: Se utiliza para ajustar o combinar columnas, incluye dividir la matriz, copiar o eliminar columnas y renglones o filas, establecer rangos, recodificar, cambiar el tipo de datos, desplegar datos, mostrar los datos de la hoja de trabajo en la ventana de sesión, entre otros.

Calcular: calcula las estadísticas de columnas y filas, distribuciones de probabilidad, matrices, estandarizaciones, operaciones aritméticas, etcétera.

Estadísticas: de manera fundamental, ejecuta los siguientes tipos de estadísticas:

1. Básicas: descriptiva e inferencial como distribución normal, prueba *t*, prueba de hipótesis acerca de la media poblacional, correlación, covarianza y *Chi* cuadrada.
2. Regresión lineal y múltiple.
3. Análisis de varianza (ANOVA) unidireccional y factorial.
4. DOE (análisis para diseños experimentales, análisis de respuestas).
5. Gráficas de control: de atributos, multivariados, de tiempo, individuales y grupales.
6. Herramientas de calidad: diagramas de dispersión, Pareto, causa-efecto, entre otros.
7. Confiabilidad: análisis de distribución, planes de prueba, análisis de garantía, prueba acelerada de vida útil, etcétera.
8. Análisis multivariado: conglomerados, análisis de factores (validación), análisis discriminante, análisis de conglomerados, de correspondencia simple o múltiple.
9. Series de tiempos: autocorrelación, correlación parcial, correlación cruzada, entre otras.
10. Tablas: tabulación cruzada, C*hi* cuadrada.
11. Estadística no paramétrica.
12. EDA (análisis exploratorio de datos, diagramas de caja, fotograma, etcétera).
13. Poder y tamaño de muestra (1-muestra *z*, 1-muestra-*t*, 2-muestra-*t*, ANOVA y otras. Sirve para determinar si el tamaño de muestra es apropiado para varias pruebas estadísticas).

Gráfica: sirve para solicitar gráficos, histogramas, barras de pastel, diagramas de dispersión, Pareto, series de tiempos, etcétera.

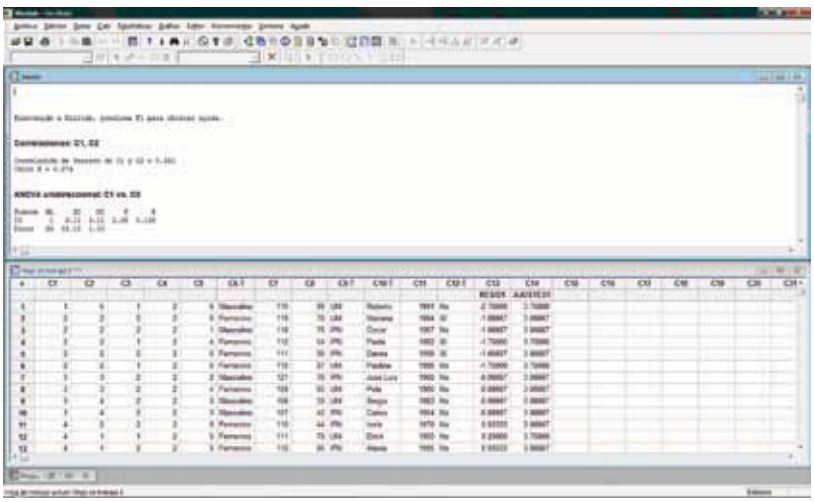
Editor: Se usa para mover, redefinir e insertar columnas, buscar o ir a un caso particular, entre otras acciones.

Herramientas: es útil para definir ambientes, conjuntos, información sobre variables, conexión a internet, consultas, etcétera.

Ventana: sirve para moverse a través de archivos y hacia otros programas, minimizar ventanas y demás funciones similares a otros programas.

Ayuda: cuenta con contenidos de ayuda, cómo utilizar Minitab, comandos, guías y demás elementos de Windows aplicados al paquete. En la figura 10.2 se muestra una vista de la pantalla de Minitab.

Figura 10.2 Pantalla de Minitab.



Otro programa de análisis sumamente difundido es el SAS (Sistema de Análisis Estadístico), que fue diseñado en la Universidad de Carolina del Norte. Es muy poderoso y su utilización se ha incrementado notablemente. Es un paquete muy completo para computadoras personales que contiene una variedad considerable de pruebas estadísticas (análisis de varianza, regresión, análisis de datos categóricos, análisis no paramétricos, etc.). Su página es: http://www.sas.com/technologies/analytics/statistics/stat/.

En el centro de recursos en línea se incluye un vínculo al sitio de Decision Analyst, donde se puede descargar una versión de prueba del software STATS®, el cual ha sido incluido desde la segunda edición de esta obra y contiene los análisis bivariados más elementales para comenzar a comprenderlos y practicarlos. Por otro lado, en internet se encuentran diversos programas gratuitos de análisis estadístico para cualquier ciencia o disciplina.

Centro de recursos en línea

Por lo general se elige el programa de análisis que está disponible en nuestra institución educativa, centro de investigación u organización de trabajo, o el que podamos comprar u obtener en internet. Todos los programas mencionados son excelentes opciones. Cualquiera nos sirve, solamente que debemos seleccionar uno. Recomendamos que en el centro de cómputo de su institución soliciten información respecto de los programas disponibles.

Paso 2: ejecutar el programa

La mayoría de los programas son fáciles de usar, pues lo único que hay que hacer es solicitar los análisis requeridos seleccionando las opciones apropiadas.

Paso 3: explorar los datos

En esta etapa, inmediata a la ejecución del programa, se inicia el análisis. Cabe señalar que si hemos llevado a cabo la investigación reflexionando paso a paso, la fase analítica es relativamente sencilla, porque: 1) formulamos las preguntas de investigación que pretendemos contestar, 2) visualizamos un alcance (exploratorio, descriptivo, correlacional o explicativo), 3) establecimos nuestras hipótesis (o estamos conscientes de que no las tenemos), 4) definimos las variables, 5) elaboramos un instrumento (conocemos qué ítems o indicadores miden qué variables y qué nivel de medición tiene cada variable: nominal, ordinal, de intervalos o razón) y 6) recolectamos los datos. Sabemos qué deseamos hacer, es decir, tenemos claridad.

La exploración típica se muestra en la figura 10.3 (que se hizo con base en el programa SPSS, pues, insistimos, puede variar de programa a programa en cuanto a comandos o instrucciones, pero no en lo referente a las funciones implementadas). Algunos conceptos pueden, por ahora, no significar nada para el lector que se inicia en los menesteres de la investigación, pero se irán explicando a lo largo del capítulo.

Veamos ahora los conceptos estadísticos que se aplican a la exploración de datos, pero antes es necesario realizar un par de apuntes, uno sobre las *variables del estudio* y las *variables de la matriz de datos*, y el otro sobre los factores de los que depende el análisis.

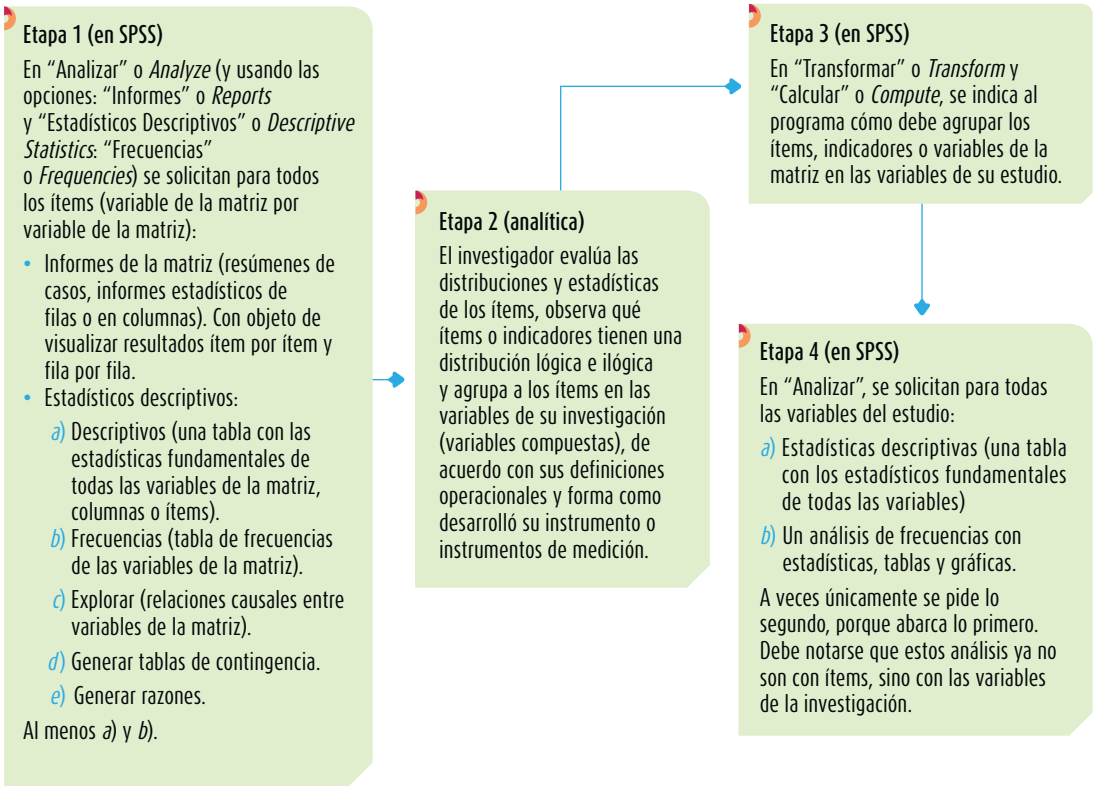
Apunte 1

Al final del capítulo anterior se introdujo el concepto de **variable de la matriz de datos**, que es distinto del concepto **variable de la investigación**. Las variables de la matriz de datos son columnas que constituyen indicadores o ítems. Las variables de la investigación son las propiedades medidas y que forman parte de las hipótesis o que se pretenden describir (género, edad, actitud hacia el presidente municipal, inteligencia, duración de un material, presión arterial, etc.). En ocasiones, las *variables de la investigación* requieren un único ítem, lectura o indicador para ser medidas (como en la tabla 10.2 con la variable "tipo de escuela a la que asiste"), pero en otras se necesitan varios ítems para tal finalidad. Cuando sólo se precisa de un ítem o indicador, las variables de la investigación

Variables de la matriz de datos Son columnas constituidas por ítems.

Variables de la investigación Son las propiedades medidas y que forman parte de las hipótesis o que se pretenden describir.

Figura 10.3 Secuencia más común para explorar datos en SPSS.



ocupan una columna de la matriz (una variable de la matriz). Pero si están compuestas de varios ítems, ocuparán tantas columnas como ítems (o variables en la matriz) las conformen. Esto se ejemplifica en la tabla 10.2 con los casos de las variables "presión arterial", "satisfacción respecto al superior" y "moral de los empleados".

Y cuando las *variables de la investigación* se integran de varios ítems o variables en la matriz, las columnas pueden ser continuas o no (estar ubicadas de manera seguida o en distintas partes de la matriz). En el cuarto ejemplo (variable "moral de los empleados"), las preguntas podrían ser las número 1, 2, 3, 4 y 5 del cuestionario; entonces, las primeras cinco columnas de la matriz representarán a estos ítems. Pero pueden ubicarse en distintos segmentos del cuestionario (por ejemplo, ser las preguntas 1, 5, 17, 22 y 38); entonces, las columnas que las representen se ubicarán de forma discontinua (serán las columnas o variables de la matriz 1, 5, 17, 22 y 38); porque regularmente la secuencia de las columnas corresponde a la secuencia de los ítems en el instrumento de medición.

Esta explicación se hace porque hemos visto estudiantes que confunden las variables de la matriz de datos con las variables del estudio. Son cuestiones vinculadas pero distintas.

Cuando una variable de la investigación está integrada por diversas variables de la matriz o ítems, suele llamarse *variable compuesta* y su puntuación total es el resultado de adicionar los valores de los reactivos que la conforman. Tal vez el caso más claro es la escala de Likert, en la que se suman las puntuaciones de cada ítem y se logra la calificación final. A veces la adición es una sumatoria, otras ocasiones es multiplicativa, un promedio o de otras formas, según se haya desarrollado el instrumento. Al ejecutar el programa y durante la fase exploratoria, se toman en cuenta todas las *variables de la investigación* e *ítems* y se considera a las *variables compuestas*, entonces se indica en el programa cómo están constituidas, mediante algunas instrucciones (en cada programa son distintas en cuanto al nombre, pero su función es similar). Por ejemplo, en SPSS se crean nuevas variables compuestas en la matriz de datos con el comando "Transformar" y luego con el comando "Calcular variable", de este modo, se construye la variable compuesta mediante una expresión numérica. Revisemos un ejemplo.

Tabla 10.2 Ejemplos de variables de investigación y formulación de ítems

Variable: tipo de escuela a la que asiste (con un ítem)	Variable: presión arterial (con dos indicadores)	Variable: satisfacción respecto al superior (con tres ítems)	Variable: moral del departamento donde se trabaja (con cinco ítems)
¿Asiste a una escuela pública o privada? [1] Escuela pública [2] Escuela privada	Lectura de la presión arterial sistólica:	1. ¿En qué medida está usted satisfecho con su superior inmediato? [1] Sumamente insatisfecho [2] Más bien insatisfecho [3] Ni insatisfecho ni satisfecho [4] Más bien satisfecho [5] Sumamente satisfecho	1. "En el departamento donde trabajo nos mantenemos unidos". [5] Totalmente de acuerdo [4] De acuerdo [3] Ni de acuerdo ni en desacuerdo [2] En desacuerdo [1] Totalmente en desacuerdo
	Lectura de la presión arterial diastólica:	2. ¿Qué tan satisfecho está usted con el trato que recibe de parte de su superior inmediato? [1] Sumamente insatisfecho [2] Más bien insatisfecho [3] Ni insatisfecho ni satisfecho [4] Más bien satisfecho [5] Sumamente satisfecho	2. "La mayoría de las veces en mi departamento compartimos la información más que guardarla para nosotros". [5] Totalmente de acuerdo [4] De acuerdo [3] Ni de acuerdo ni en desacuerdo [2] En desacuerdo [1] Totalmente en desacuerdo
		3. ¿Qué tan satisfecho está con la orientación que le proporciona su superior inmediato para que usted realice su trabajo? [1] Sumamente insatisfecho [2] Más bien insatisfecho [3] Ni insatisfecho ni satisfecho [4] Más bien satisfecho [5] Sumamente satisfecho	3. "En mi departamento nos mantenemos en contacto permanentemente". [5] Totalmente de acuerdo [4] De acuerdo [3] Ni de acuerdo ni en desacuerdo [2] En desacuerdo [1] Totalmente en desacuerdo
			4. "En mi departamento nos reunimos con frecuencia para hablar tanto de asuntos de trabajo como de cuestiones personales". [5] Totalmente de acuerdo [4] De acuerdo [3] Ni de acuerdo ni en desacuerdo [2] En desacuerdo [1] Totalmente en desacuerdo
			5. "En mi trabajo todos nos llevamos bien". [5] Totalmente de acuerdo [4] De acuerdo [3] Ni de acuerdo ni en desacuerdo [2] En desacuerdo [1] Totalmente en desacuerdo
Esta variable es medida por una sola pregunta y ocupa una columna o variable de la matriz.	Esta variable es medida por dos indicadores o lecturas y ocupa dos columnas o variables de la matriz.	Esta variable es medida por tres preguntas y ocupa tres columnas o variables de la matriz.	Esta variable es medida por cinco preguntas y ocupa cinco columnas o variables de la matriz.

En el caso de la variable "moral en el departamento donde se trabaja", podríamos asignar las siguientes columnas (en el supuesto de que fueran continuas) a los cinco ítems, tal como se muestra en la tabla 10.3.

Y tener la siguiente matriz (ejemplo):

Ejemplo

Tabla 10.3 Ejemplo con la variable moral

Variable de la investigación: moral	Variable de la matriz que corresponde a la variable de la investigación	Ubicación en la matriz
1. "En el departamento donde trabajo nos mantenemos unidos" 5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo	Frase 1 (fr1)	Columna 1
2. "La mayoría de las veces en mi departamento compartimos la información más que guardarla para nosotros" 5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo	Frase 2 (fr2)	Columna 2
3. "En mi departamento nos mantenemos en contacto permanentemente" 5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo	Frase 3 (fr3)	Columna 3
4. "En mi departamento nos reunimos con frecuencia para hablar tanto de asuntos de trabajo como de cuestiones personales" 5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo	Frase 4 (fr4)	Columna 4
5. "En mi trabajo todos nos llevamos muy bien" 5 Totalmente de acuerdo 4 De acuerdo 3 Ni de acuerdo ni en desacuerdo 2 En desacuerdo 1 Totalmente en desacuerdo	Frase 5 (fr5)	Columna 5

Casos	fr1	fr2	fr3	fr4	fr5
1	1	2	2	4	3
2	2	2	2	2	2
K	2	3	2	2	3

En las opciones "Transformar" y "Calcular" o "Computar" el programa nos pide que indiquemos el nombre de la nueva variable (en este caso la compuesta por cinco frases): *moral*. Y nos solicita que desarrollemos la expresión numérica que corresponda a esta variable compuesta: *fr1+fr2+fr3+fr4+fr5* (automáticamente el programa realiza la operación y agrega la nueva variable compuesta "moral" a la matriz de datos y realiza los cálculos, y ahora sí, la *variable del estudio* es una variable más de la matriz de datos). La matriz se modificaría de la siguiente manera:

Ejemplo

Casos	fr1	fr2	fr3	fr4	fr5	Moral
1	1	2	2	4	3	12
2	2	2	2	2	2	10
K	2	3	2	2	3	12

Desde luego, para mantener esta variable debemos demostrar que fue medida de forma confiable y válida, así como evaluar si todos los ítems aportan favorablemente a ambos elementos o algunos no. Y en lugar de una suma, la variable *moral* podría ser un promedio de las cinco frases o variables de la matriz (como ya se mencionó en el tema de la escala de Likert). Entonces, la expresión en "Calcular" es: (*fr1+fr2+fr3+fr4+fr5*)/*5*, y los valores en "moral" serían:

Ejemplo

Casos	fr1	fr2	fr3	fr4	fr5	Moral
1	1	2	2	4	3	2.4
2	2	2	2	2	2	2.0
K	2	3	2	2	3	2.4

Por último, las variables de la investigación son las que nos interesan, ya sea que estén compuestas por uno, dos, diez, 50 o más ítems. El primer análisis es sobre los ítems, únicamente para explorar; el análisis descriptivo final es sobre las *variables del estudio*.

Apunte 2

Los análisis de los datos dependen de tres factores:[2]

- ***a)*** El *nivel de medición* de las variables.
- ***b)*** La manera como se hayan formulado las *hipótesis*.
- ***c)*** El *interés analítico del investigador* (que depende del planteamiento del problema).

[2] Babbie (2014), O'Leary (2014), Hollander, Wolfe y Chicken (2013); Jarman (2013), Kon y Rai (2013), Hernández-Sampieri *et al.* (2013), Collier, LaPorte y Seawright (2012); Martin y Bridgmon (2012), Garson (2012), Howell (2011), Mertens (2010), Gershkoff (2008), Scott y Albaum (2006), Reynolds (1984) y Hildebbrand, Laing y Rosenthal (1977).

Por ejemplo, los análisis que se aplican a una variable nominal son distintos a los de una variable por intervalos. Se sugiere repasar los niveles de medición vistos en el capítulo anterior.

El investigador busca, en primer término, describir sus datos y posteriormente efectuar análisis estadísticos para relacionar sus variables. Es decir, realiza análisis de estadística descriptiva para cada una de las variables de la matriz (ítems o indicadores) y luego para cada una de las variables del estudio, finalmente aplica cálculos estadísticos para probar sus hipótesis. Los tipos o métodos de análisis cuantitativo o estadístico son variados y se comentarán a continuación; pero cabe señalar que el análisis no es indiscriminado, sino que cada método tiene su razón de ser y un propósito específico; por ello, no deben hacerse más análisis de los necesarios. La estadística no es un fin en sí misma, sino una herramienta para evaluar los datos.

Estadística descriptiva para cada variable

2 La primera tarea es describir los datos, los valores o las puntuaciones obtenidas para cada variable. Por ejemplo, si aplicamos a 2 112 niños el cuestionario sobre los usos y las gratificaciones que la televisión tiene para ellos, ¿cómo pueden describirse estos datos? Esto se logra al describir la distribución de las puntuaciones o frecuencias de cada variable.

¿Qué es una distribución de frecuencias?

Distribución de frecuencias Conjunto de puntuaciones de una variable ordenadas en sus respectivas categorías.

Una **distribución de frecuencias** es un conjunto de puntuaciones respecto de una variable ordenadas en sus respectivas categorías y generalmente se presenta como una tabla (O'Leary, 2014 y Nicol, 2006).

La tabla 10.4 muestra un ejemplo de una distribución de frecuencias.

Ejemplo

En un estudio entre 200 personas latinas que viven en el estado de California, Estados Unidos,[3] se les preguntó: ¿cómo prefiere que se refieran a usted en cuanto a su origen étnico? Las respuestas fueron:

Tabla 10.4 Ejemplo de una distribución de frecuencias

Variable: preferencias al referir el origen étnico (nombrada en SPSS: prefoe)		
Categorías	**Códigos (valores)**	**Frecuencias**
Hispano	1	52
Latino	2	88
Latinoamericano	3	6
Americano	4	22
Otros	5	20
No respondieron	6	12
Total		200

A veces, las *categorías* de las distribuciones de frecuencias son tantas que es necesario resumirlas. Por ejemplo, examinemos detenidamente la distribución de la tabla 10.5. Esta distribución podría compendiarse como en la tabla 10.6.

[3] Encuesta con 7% de margen de error (University of Southern California y Bendixen and Associates, 2002).

Tabla 10.5 Ejemplo de una distribución que necesita resumirse

Variable: calificación en la prueba de motivación			
Categorías	Frecuencias	Categorías	Frecuencias
48	1	74	1
55	2	75	4
56	3	76	3
57	5	78	2
58	7	80	4
60	1	82	2
61	1	83	1
62	2	84	1
63	3	86	5
64	2	87	2
65	1	89	1
66	1	90	3
68	1	92	1
69	1	**Total**	63
73	2		

Tabla 10.6 Ejemplo de una distribución resumida

Variable: calificación en la prueba de motivación	
Categorías	Frecuencias
55 o menos	3
56-60	16
61-65	9
66-70	3
71-75	7
76-80	9
81-85	4
86-90	11
91-96	1
Total	63

¿Qué otros elementos contiene una distribución de frecuencias?

Las distribuciones de frecuencias pueden completarse agregando los porcentajes de casos en cada categoría, los porcentajes válidos (excluyendo los valores perdidos) y los porcentajes acumulados (porcentaje de lo que se va acumulando en cada categoría, desde la más baja hasta la más alta).

La tabla 10.7 muestra un ejemplo con las frecuencias y porcentajes en sí, los porcentajes válidos y los acumulados. El *porcentaje acumulado* constituye lo que aumenta en cada categoría de manera porcentual y progresiva (en orden de aparición de las categorías), tomando en cuenta los *porcentajes válidos*. En la categoría "sí se ha obtenido la cooperación", se ha acumulado 74.6%. En la categoría "no se ha obtenido la cooperación", se acumula 78.7% (74.6% de la categoría anterior y 4.1% de la categoría en cuestión). En la última categoría siempre se acumula el total (100%).[4]

Tabla 10.7 Ejemplo de una distribución de frecuencias con todos sus elementos

Variable: cooperación del personal con el proyecto de calidad de la empresa				
Categorías	Códigos	Frecuencias	Porcentaje válido	Porcentaje acumulado
Sí se ha obtenido la cooperación	1	91	74.6	74.6
No se ha obtenido la cooperación	2	5	4.1	78.7
No respondieron	3	26	21.3	100.0
Total		**122**	**100.0**	

Las columnas *porcentaje* y *porcentaje válido* son iguales (mismas cifras o valores) cuando *no* hay valores perdidos; pero si tenemos valores perdidos, la columna *porcentaje válido* presenta los cálculos sobre el total menos tales valores. En la tabla 10.8 se muestra un ejemplo con valores perdidos en el caso de un estudio exploratorio sobre los motivos de los niños celayenses para elegir su personaje televisivo favorito (García y Hernández-Sampieri, 2005).

[4] En variables nominales el porcentaje acumulado es relativo porque no hay orden o jerarquía entre categorías, pero se buscó un ejemplo simple para entender más fácilmente el concepto.

Al elaborar el informe de resultados, una distribución se presenta con los elementos más informativos para el lector y la descripción de los resultados o un comentario, tal como se muestra en la tabla 10.9.

Tabla 10.8 Ejemplo de tabla con valores perdidos (en SPSS)[5]

Motivos de la preferencia de su personaje favorito					
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	Divertidos	142	72.1	73.2	73.2
	Buenos	10	5.1	5.2	78.4
	Tienen poderes	23	11.7	11.9	90.2
	Son fuertes	19	9.6	9.8	100.0
	Total	194	98.5	100.0	
Perdidos	No contestaron	3	1.5		
Total		**197**	**100.0**		

Tabla 10.9 Ejemplo de una distribución de frecuencias para presentar a un usuario

¿Se ha obtenido la cooperación del personal para el proyecto de calidad?		
Obtención	**Núm. de organizaciones**	**Porcentajes**
Sí	91	74.6
No	5	4.1
No respondieron	26	21.3
Total	122	100.0
COMENTARIO: Prácticamente tres cuartas partes de las organizaciones sí han obtenido la cooperación del personal. Llama la atención que poco más de una quinta parte no quiso comprometerse con su respuesta. Las empresas que no han logrado la cooperación del personal mencionaron como factores el absentismo, rechazo al cambio y conformismo.		

En los programas de análisis se solicita la distribución de frecuencias de cada variable de la investigación (por ejemplo, en SPSS: Analizar → Estadísticos descriptivos → Frecuencias).[6]

¿De qué otra manera pueden presentarse las distribuciones de frecuencias?

3 Las distribuciones de frecuencias, especialmente cuando utilizamos los porcentajes, pueden presentarse en forma de histogramas o gráficas de otro tipo (por ejemplo: de pastel). Algunos ejemplos se muestran en la figura 10.4.

SPSS, Minitab y SAS producen tales gráficas, o bien, los datos pueden exportarse a otros programas o paquetes que las generan (de cualquier tipo, a colores, utilizando efectos de movimiento y en tercera dimensión, como por ejemplo: Power Point).

Centro de recursos en línea

Para obtener las gráficas en SPSS no olvide consultar el centro de recursos en línea de esta obra el *manual* "Introducción al IBM SPSS®".

Polígonos de frecuencias Relacionan las puntuaciones con sus respectivas frecuencias por medio de gráficas útiles para describir los datos.

Las distribuciones de frecuencias también se pueden graficar como polígonos de frecuencias

Los **polígonos de frecuencias** relacionan las puntuaciones con sus respectivas frecuencias. Es más bien propio de un nivel de medición por intervalos o razón. Los

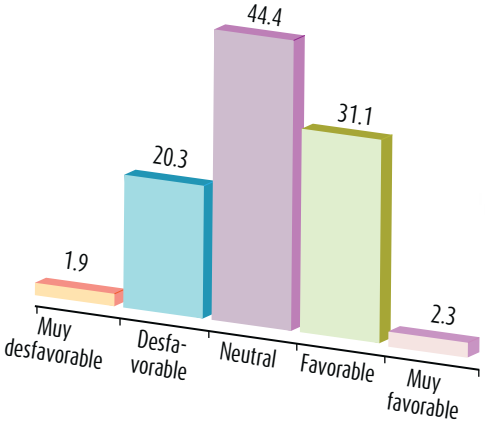
[5] En todos los casos, cifras redondeadas por el programa.

[6] Esta secuencia en SPSS para obtener los análisis de frecuencias requeridos, al igual que el resto de análisis (valores, tablas y gráficas), se incluyen en el manual "Introducción al IBM SPSS®", que puede descargarse del centro de recursos en línea.

Figura 10.4 Ejemplos de gráficas para presentar distribuciones.

Histogramas

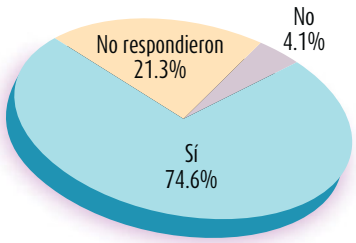
Opinión acerca del actual alcalde del municipio de San Martín Aurelio



Solamente la tercera parte de los ciudadanos expresa una opinión positiva respecto al alcalde (favorable o muy favorable).

Gráficas circulares

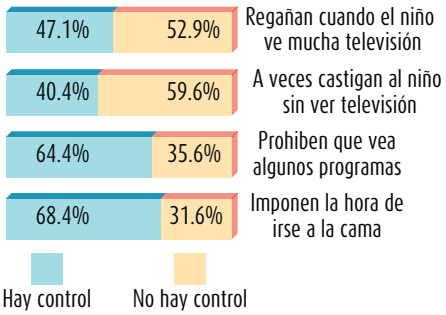
Cooperación de todo el personal (o la mayoría) para el proyecto de calidad (122 = 100%)



Prácticamente tres cuartas partes han obtenido la cooperación de todo el personal para el proyecto de la empresa. Pero llama la atención que poco más de una quinta parte no quiso comprometerse con su respuesta. Los cinco motivos de no cooperación con dicho proyecto fueron: absentismo, falta de interés, rechazo al cambio, falta de concientización y conformismo.

Otros tipos de gráficas

Control paterno sobre el uso que los niños hacen de la televisión.



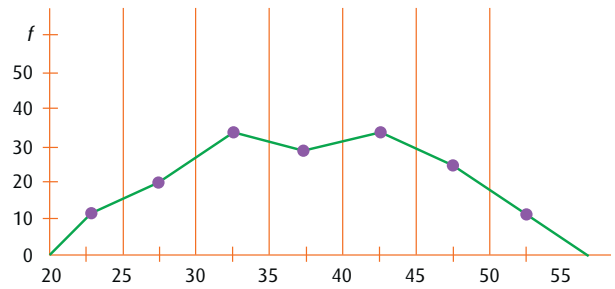
polígonos se construyen sobre los puntos medios de los intervalos. Por ejemplo, si los intervalos fueran 20-24, 25-29, 30-34, 35-39, y siguientes; los puntos medios serían 22, 27, 32, 37, etc. SPSS o Minitab realizan esta labor en forma automática. Un ejemplo de un polígono de frecuencias se muestra en la figura 10.5.

El polígono de frecuencias obedece a la siguiente distribución:

Categorías/intervalos	Frecuencias absolutas
20-24.9	10
25-29.9	20
30-34.9	35
35-39.9	33
40-44.9	36
45-49.9	27
50-54.9	8
TOTAL	169

Figura 10.5 Ejemplo de un polígono de frecuencias.

Variable: satisfacción en el trabajo



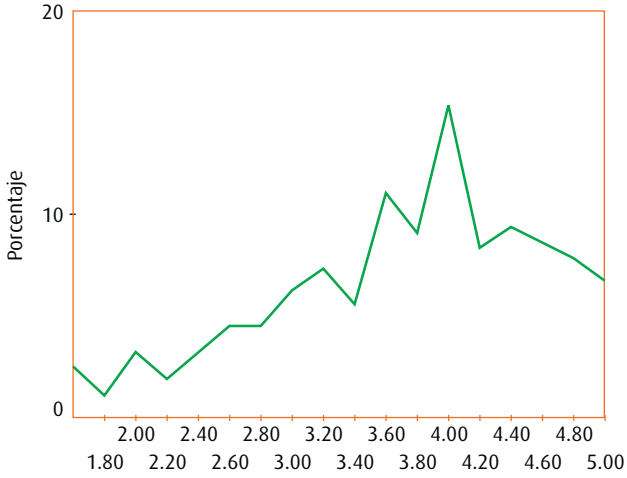
Los polígonos de frecuencias representan curvas útiles para describir los datos. Nos indican hacia dónde se concentran los casos (personas, organizaciones, segmentos de contenido, mediciones de polución, datos de presión arterial, etc.) en la escala de la variable; más adelante se hablará de ello.

En resumen, para cada una de las variables de la investigación se obtiene su distribución de frecuencias y se grafica (histograma, gráfica de barras, gráfica circular o polígono de frecuencias) (Huck, 2006).

En la figura 10.6 se muestra otro ejemplo.

Figura 10.6 Ejemplo de un polígono de frecuencias con la variable innovación.

Variable: innovación
Con respecto a la innovación en la empresa, que es la percepción del apoyo a las iniciativas tendientes a introducir mejoras en la manera como se realiza el trabajo, a nivel organizacional y departamental, la mayoría de los individuos tienden a estar en altos niveles de la escala.



El polígono puede presentarse con frecuencias como en la figura 10.5 o con porcentajes como con este segundo ejemplo. Pero además de la distribución o polígono de frecuencias, deben calcularse las *medidas de tendencia central* y de *variabilidad o dispersión.*

¿Cuáles son las medidas de tendencia central?

Medidas de tendencia central Valores medios o centrales de una distribución que sirven para ubicarla dentro de la escala de medición de la variable.

Las **medidas de tendencia central** son puntos en una distribución obtenida, los valores medios o centrales de ésta, y nos ayudan a ubicarla dentro de la escala de medición de la variable analizada. Las principales medidas de tendencia central son tres: *moda, mediana* y *media*. El nivel de medición de la variable determina cuál es la medida de tendencia central apropiada para interpretar (Graham, 2013, Kwok, 2008a y Platt, 2003a).

Moda Categoría o puntuación que se presenta con mayor frecuencia.

La **moda** es la categoría o puntuación que ocurre con mayor frecuencia. En la tabla 10.7, la moda es "1" (sí se ha obtenido la cooperación). Se utiliza con cualquier nivel de medición.

La mediana es el valor que divide la distribución por la mitad. Esto es, la mitad de los casos caen por debajo de la mediana y la otra mitad se ubica por encima de ésta. La mediana refleja la posición intermedia de la distribución (Hempel, 2006). Por ejemplo, si los datos obtenidos fueran:

24 31 35 35 38 43 45 50 57

La mediana es 38, porque deja cuatro casos por encima (43, 45, 50 y 57) y cuatro casos por debajo (35, 35, 31 y 24). Parte a la distribución en dos mitades. En general, para descubrir la puntuación que constituye la mediana de una distribución, simplemente se aplica la fórmula:

$$\frac{N+1}{2}$$

Si tenemos nueve casos, $\frac{9+1}{2}$ entonces buscamos el quinto valor y éste es la mediana. Note que la mediana es el valor observado que se localiza a la mitad de la distribución, no el valor de cinco. La fórmula no nos proporciona directamente el valor de la mediana, sino el número de caso en donde está la mediana.

La mediana es una medida de tendencia central propia de los niveles de medición ordinal, por intervalos y de razón. No tiene sentido con variables nominales, porque en este nivel no hay jerarquías ni noción de encima o debajo. Asimismo, la mediana es particularmente útil cuando hay valores extremos en la distribución. No es sensible a éstos. Si tuviéramos los siguientes datos:

24 31 35 35 38 43 45 50 248

la mediana seguiría siendo 38.

Para la interpretación de la media y la mediana, se incluye un comentario al respecto en el siguiente ejemplo.[7]

Ejemplo

¿Qué edad tiene? Si teme contestar no se preocupe, los perfiles de edad difieren de un país a otro.[8]

A mediados de 2013, la población mundial superó los 7 100 millones de habitantes y se espera que en 2015 seamos más de 7 300 millones de humanos y en 2050 pasemos de 9 000 millones (United States Census Bureau, 2013, Alberich, 2013 y Organización de las Naciones Unidas, 2011).

En 2012, la mediana de edad mundial fue de 29 años, lo que significa que la mitad de los habitantes del globo terrestre sobrepasa esta edad y el otro medio es más joven. Cabe señalar que la mediana varía de un lugar a otro. Por ejemplo, en 2012 por bloque de países fue de 28.1 años en Asia, 39.9 en Europa, 26.9 en Oceanía, 27.8 en Sudamérica, 30.2 en América Central y El Caribe (incluyendo México), 20.1 en África y 38.4 en Norteamérica (Canadá y Estados Unidos). Actualmente, países con población muy joven son Uganda y Níger, con una edad mediana de prácticamente 15 años, y entre los más viejos podemos incluir a Japón (44.8), Alemania (45.3) y Mónaco (49.9). En América Latina tenemos algunos ejemplos como: Perú (26.5), México (27.4), Colombia (28.3), Chile (32.8), República Dominicana (26.5), Brasil (29.6), Costa Rica (29.2), Panamá (27.8), Ecuador (26), Paraguay (25.9), Uruguay (33.8), Honduras (21.3), El Salvador (24.7), Argentina (30.7) y Bolivia (22.8).

La mediana de edad ha ido en constante ascenso durante el siglo pasado y el actual. Se estima que para 2050 la edad mediana mundial habrá aumentado a más de 36 años. Buena noticia para el actual ciudadano global medio, porque parece ser que se encuentra en la situación de "envejecer más lentamente".

Media Promedio aritmético de una distribución. Es la medida de tendencia central más utilizada.

La **media** es tal vez la medida de tendencia central más utilizada (Graham, 2013, Kwok, 2008b y Leech, Onwuegbuzie y Daniel, 2006) y puede definirse como el promedio aritmético de una distribución. Se simboliza como $\overline{X}$, y es la suma de todos los valores dividida entre el número de casos. Es una medida solamente aplicable a mediciones por intervalos o de razón. Carece de sentido para variables medidas en un nivel nominal u ordinal. Resulta sensible a valores extremos. Si tuviéramos las siguientes puntuaciones:

8 7 6 4 3 2 6 9 8

El promedio sería igual a 5.88. Pero bastaría una puntuación extrema para alterarlo de manera notoria:

8 7 6 4 3 2 6 9 **20** (promedio igual a 7.22).

La mediana puede ser una medida de interpretación más útil que la media si la distribución está más cargada hacia puntuaciones extremas (Kwok, 2008a y Hempel, 2006).

El cálculo de la media lo podrá encontrar el lector en el centro de recursos en línea de la obra: Material complementario → Documentos → Documento 2, "Fórmulas y procedimientos estadísticos".

Centro de recursos en línea

¿Cuáles son las medidas de la variabilidad?

Medidas de la variabilidad Intervalos que indican la dispersión de los datos en la escala de medición de la variable.

Las **medidas de la variabilidad** indican la dispersión de los datos en la escala de medición de la variable considerada y responden a la pregunta: ¿dónde están diseminadas las puntuaciones o los valores obtenidos? Las medidas de tendencia central son valores en una distribución y las medidas de la variabilidad son intervalos que designan distancias o un número de unidades en la escala de medición (Kon y Rai, 2013 y O'Brien, 2007). Las medidas de la variabilidad más utilizadas son *rango*, *desviación estándar* y *varianza*.

[7] Basado en una idea de Leguizamo (1987).

[8] Datos obtenidos de Getamap (2013), Kaiser Family Foundation (2013) y Worldstat (2013). Son estimaciones con un margen de error de 1% y se basan en información de 2010-2011 y proyecciones a 2012.

Rango Extensión total de los datos en la escala.

El **rango**, también llamado *recorrido*, es la diferencia entre la puntuación mayor y la puntuación menor, e indica el número de unidades en la escala de medición que se necesitan para incluir los valores máximo y mínimo. Se calcula así: $X_M - X_m$ (puntuación mayor menos puntuación menor). Si tenemos los siguientes valores:

17 18 20 20 24 28 28 30 33

El rango será: 33 – 17 = 16.

Cuanto *más grande* sea el *rango*, *mayor* será la *dispersión de los datos* de una distribución.

Desviación estándar Promedio de desviación de las puntuaciones con respecto a la media que se expresa en las unidades originales de medición de la distribución.

La **desviación estándar** o característica es el promedio de desviación de las puntuaciones con respecto a la media (Jarman, 2013 y Levin, 2003). Esta medida se expresa en las unidades originales de medición de la distribución. Se interpreta en relación con la media. Cuanto mayor sea la dispersión de los datos alrededor de la media, mayor será la desviación estándar. Se simboliza como: *s* o la sigma minúscula σ, o bien mediante la abreviatura DE. Su cálculo lo podrá encontrar el lector en el centro de recursos, en: Material complementario → Documentos → Documento 2, "Fórmulas y procedimientos estadísticos".

Centro de recursos en línea

La desviación estándar se interpreta como *cuánto se desvía, en promedio, de la media un conjunto de puntuaciones*.

Supongamos que un investigador obtuvo para su muestra una media (promedio) de ingreso familiar anual de 6 000 unidades monetarias y una desviación estándar de 1 000. La interpretación es que los ingresos familiares de la muestra se desvían, en promedio, mil unidades monetarias respecto a la media.

La desviación estándar sólo se utiliza en variables medidas por intervalos o de razón.

La varianza

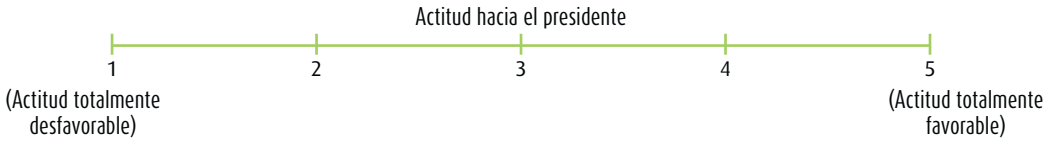
Varianza Desviación estándar elevada al cuadrado.

La **varianza** es la desviación estándar elevada al cuadrado y se simboliza como s^2. Es un concepto estadístico muy importante, ya que la mayoría de las pruebas cuantitativas se fundamentan en él. Diversos métodos estadísticos parten de la descomposición de la varianza (Zhang, 2013; Beins y McCarthy, 2009; Wilcox, 2008; y Jackson, 2008). Sin embargo, con fines descriptivos se utiliza preferentemente la desviación estándar.

¿Cómo se interpretan las medidas de tendencia central y de la variabilidad?

Cabe destacar que al describir nuestros datos, respecto a cada *variable del estudio*, interpretamos las medidas de tendencia central y de la variabilidad en conjunto, no aisladamente. Consideramos todos los valores. Para interpretarlos, lo primero que hacemos es tomar en cuenta el rango potencial de la escala. Supongamos que aplicamos una escala de actitudes del tipo Likert para medir la "actitud hacia el presidente" de una nación (digamos que la escala tuviera 18 ítems y se promediaran sus valores). El rango potencial es de uno a cinco (véase la figura 10.7).

Figura 10.7 Ejemplo de escala con rango potencial.



Si obtuviéramos los siguientes resultados:

Variable: actitud hacia el presidente
Moda: 4.0

Mediana: 3.9
Media ($\overline{X}$): 4.2
Desviación estándar: 0.7
Puntuación más alta observada (máximo): 5.0
Puntuación más baja observada (mínimo): 2.0
Rango: 3

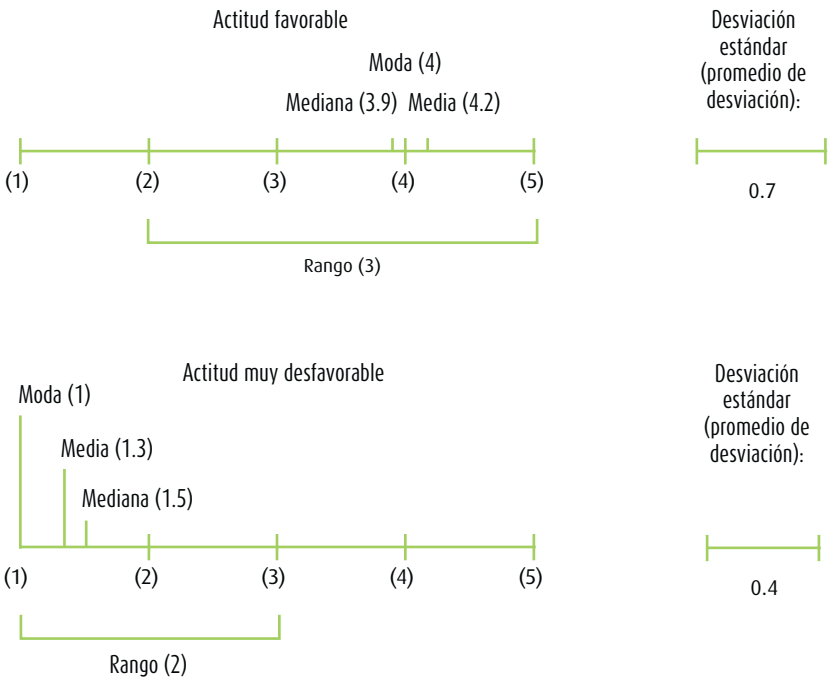
podríamos hacer la siguiente interpretación descriptiva: la actitud hacia el presidente es favorable. La categoría que más se repitió fue 4 (favorable). Cincuenta por ciento de los individuos está por encima del valor 3.9 y el restante 50% se sitúa por debajo de este valor (mediana). En promedio, los participantes se ubican en 4.2 (favorable). Asimismo, se desvían de 4.2, en promedio, 0.7 unidades de la escala. Ninguna persona calificó al presidente de manera muy desfavorable (no hay "1"). Las puntuaciones tienden a ubicarse en valores medios o elevados.

En cambio, si los resultados fueran:

Variable: actitud hacia el presidente
Moda: 1
Mediana: 1.5
Media ($\overline{X}$): 1.3
Desviación estándar: 0.4
Máximo: 3.0
Mínimo: 1.0
Rango: 2.0

La interpretación es que la actitud hacia el presidente es muy desfavorable. En la figura 10.8 vemos gráficamente la comparación de resultados. La variabilidad también es menor en el caso de la actitud muy desfavorable (los datos se encuentran menos dispersos).

Figura 10.8 Ejemplo de interpretación gráfica de las estadísticas descriptivas.



Otro ejemplo de interpretación de los resultados de una medición respecto a una variable es el que ahora se presenta.

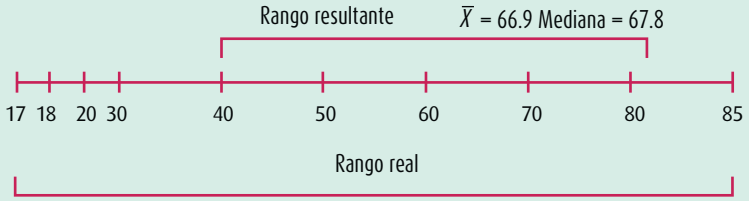
Ejemplo

Hernández-Sampieri y Cortés (1982) aplicaron una prueba de motivación intrínseca sobre la ejecución de una tarea a 60 participantes de un experimento. La escala contenía 17 ítems (con cinco opciones cada uno, uno a cinco) y los resultados fueron los siguientes:[9]

n: 60	Rango: 41	Mínimo: 40	Máximo: 81
Media: 66.9	Mediana: 67.8	Moda: 61	DE: 9.1
Varianza: 83	Curtosis: 0.6	Asimetría: −0.8	EE: 1.18
Sumatoria: 4 013			

¿Qué podríamos decir sobre la motivación intrínseca de los participantes?

El nivel de motivación intrínseca exhibido por los participantes tiende a ser elevado, como lo indican los resultados. El rango real de la escala iba de 17 a 85. El rango resultante para esta investigación varió de 40 a 81. Por tanto, es evidente que los individuos se inclinaron hacia valores elevados en la medida de motivación intrínseca. Además, la media de los participantes es de 66.9 y la mediana de 67.8, lo cual confirma la tendencia de la muestra hacia valores altos de la escala. A pesar de que la dispersión de las puntuaciones de los sujetos es considerable (la desviación estándar es igual a 9.1 y el rango es de 41), esta dispersión se manifiesta en el área más elevada de la escala. Veámoslo gráficamente.



Escala de motivación intrínseca (datos ordinales, supuestos como datos en nivel de intervalo).

En resumen, la tarea resultó intrínsecamente motivante para la mayoría de los participantes; sólo que para algunos resultó muy motivante, para otros, relativamente motivante, y para los demás, medianamente motivante. Esto es, que la tendencia general es hacia valores superiores.

Ahora bien, ¿qué significa un alto nivel de motivación intrínseca exhibido con respecto a una tarea? Implica que la tarea fue percibida como atractiva, interesante, divertida y categorizada como una experiencia agradable. Asimismo, involucra que los individuos, al ejecutarla, derivaron de ella sentimientos de satisfacción, goce y realización personal. Por lo general, quien se encuentra intrínsecamente motivado hacia una labor, disfrutará la ejecución de ésta, ya que obtendrá de la labor *per se* recompensas internas, como sentimientos de logro y autorrealización. Además de ser absorbido por el desarrollo de la tarea y, al tener un buen desempeño, la opinión de sí mismo mejorará o se verá reforzada.

¿Hay alguna otra estadística descriptiva?

Sí, *la asimetría y la curtosis.* Los *polígonos de frecuencia* son *curvas*, por ello se representan como tales (figura 10.9), para que puedan analizarse en términos de probabilidad y visualizar su grado de dispersión. Estos dos elementos resultan esenciales para analizar estas curvas o polígonos de frecuencias.

Asimetría y **curtosis** Estadísticas que se usan para conocer cuánto se parece una distribución a la distribución teórica llamada *curva normal* o campana de Gauss y dónde se concentran las puntuaciones.

La **asimetría** es una estadística necesaria para conocer cuánto se parece nuestra distribución a una distribución teórica llamada *curva normal* (la cual se representa también en la figura 10.9) y constituye un indicador del lado de la curva donde se agrupan las frecuencias. Si es cero (asimetría = 0), la curva o distribución es simétrica. Cuando es positiva, quiere decir que hay más valores agrupados hacia la izquierda de la curva (por debajo de la media). Cuando es negativa, significa que los valores tienden a agruparse hacia la derecha de la curva (por encima de la media) (Hume, 2011; Taylor, 2007a; Salkind, 2006; y Burkhart, 2003).

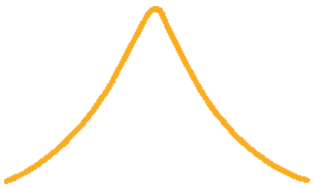
La **curtosis** es un indicador de lo plana o "picuda" que es una curva. Cuando es cero (curtosis = 0), significa que puede tratarse de una *curva normal*. Si es positiva, quiere decir que la curva, la distri-

[9] EE significa "error estándar".

bución o el polígono es más "picudo" o elevado. Si la curtosis es negativa, indica que es más plana la curva (Hume, 2011, Taylor, 2007b, Field, 2006 y Cameron, 2003).

La asimetría y la curtosis requieren al menos un nivel de medición por intervalos. En la figura 10.9 se muestran ejemplos de curvas con su interpretación.

Figura 10.9 Ejemplos de curvas o distribuciones y su interpretación.



Distribución simétrica (asimetría = 0), con curtosis positiva, y una desviación estándar y varianza medias.

Distribución con asimetría negativa, curtosis positiva, y desviación estándar y varianza mayores.

Distribución con asimetría positiva, curtosis negativa, y desviación estándar y varianza considerables.

Distribución con asimetría negativa, curtosis positiva, y desviación estándar y varianza menores.

Distribución simétrica, curtosis positiva, y una desviación estándar y varianza bajas.

Curva normal, curtosis = 0, asimetría = 0, y desviación estándar y varianza promedios.

¿Cómo se traducen las estadísticas descriptivas al inglés?

Algunos programas y paquetes estadísticos computacionales pueden realizar el cálculo de las estadísticas descriptivas, cuyos resultados aparecen junto al nombre respectivo de éstas, muchas veces en inglés.

A continuación se indican las diferentes estadísticas y su equivalente en inglés.

Estadística	Equivalente en inglés
• Moda	• *Mode*
• Mediana	• *Median*
• Media	• *Mean*
• Desviación estándar	• *Standard deviation*
• Varianza	• *Variance*
• Máximo	• *Maximum*
• Mínimo	• *Minimum*
• Rango	• *Range*
• Asimetría	• *Skewness*
• Curtosis	• *Kurtosis*

Nota final

Debe recordarse que en una investigación se obtiene una distribución de frecuencias y se calculan las estadísticas descriptivas para cada *variable*, las que se necesiten de acuerdo con los propósitos de la investigación y los niveles de medición.

Ejemplo

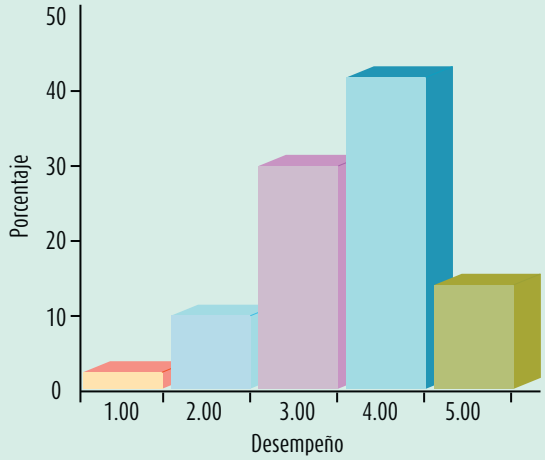
Hernández-Sampieri (2005), en su investigación sobre el clima organizacional, obtuvo las siguientes estadísticas fundamentales de sus variables en una de las muestras:

Variable	*n*	Mínimo	Máximo	Media	Desviación estándar
Moral	390	1.00	5.00	3.3818	0.91905
Dirección	393	1.00	5.00	2.7904	1.08775
Innovación	396	1.00	5.00	3.4621	0.91185
Identificación	383	1.00	5.00	3.6584	0.91283
Comunicación	397	1.00	5.00	3.2519	0.87446
Desempeño	403	1.00	5.00	3.6402	0.86793
Motivación intrínseca	401	2.00	5.00	3.9111	0.73900
Autonomía	395	1.00	5.00	3.2025	0.85466
Satisfacción	399	1.00	5.00	3.7249	0.90591
Liderazgo	392	1.00	5.00	3.4532	1.10019
Visión	391	1.00	5.00	3.7341	0.89206
Recompensas	381	1.00	5.00	2.4528	1.14364

Notas: Todas las variables son compuestas (integradas de varios ítems). La columna *n* representa el número de casos válidos para cada variable. El *n* total de la muestra es de 420, pero como podemos ver en la tabla, el número de casos es distinto en las diferentes variables, porque SPSS elimina de toda la variable los casos que no hayan respondido a un ítem o más reactivos. La variable con mayor promedio es la *motivación intrínseca* y la más baja es *recompensas*.

Posteriormente, obtuvo las tablas y distribuciones de frecuencias de todas sus 12 variables, de las cuales solamente incluimos la variable "desempeño" por cuestiones de espacio.

Desempeño				
	Valores	Frecuencia	Porcentaje válido	Porcentaje acumulado
	1	2	0.5	0.5
	2	35	8.7	9.2
	3	133	33.0	42.2
	4	169	41.9	84.1
	5	64	15.9	100.0
Total *n* = 420 Perdidos = 17		403	100.0	



Para el cálculo de estadísticas descriptivas (tendencia central y dispersión) en SPSS, se sugiere descargar la versión de prueba del sitio de SPSS y consultar el manual respectivo.

Puntuaciones *z*

Las puntuaciones *z* son transformaciones que se pueden hacer a los valores o las puntuaciones obtenidas, con el propósito de analizar su distancia respecto a la media, en unidades de desviación estándar. Una puntuación *z* nos indica la dirección y el grado en que un valor individual obtenido se aleja de la media, en una escala de unidades de desviación estándar. El lector puede conocer más sobre las puntuaciones *z* en el capítulo 8 adicional que puede descargarse del centro de recursos en línea, en: Material complementario → Capítulos → Capítulo 8, "Análisis estadístico: segunda parte".

Centro de recursos en línea

Razones y tasas

Una razón es la relación entre dos categorías. Por ejemplo:

Categorías	Frecuencia
Masculino	60
Femenino	30

La razón de hombres a mujeres es de $\frac{60}{30} = 2$. Es decir, por cada dos hombres hay una mujer.

Una **tasa** es la relación entre el número de casos, frecuencias o eventos de una categoría y el número total de observaciones, multiplicada por un múltiplo de 10, generalmente 100 o 1 000. La fórmula es:

Tasa Relación entre el número de casos de una categoría y el número total de observaciones.

$$\text{Tasa} = \frac{\text{Número de eventos}}{\text{Número total de eventos posibles}} \times 100 \text{ o } 1\,000$$

$$\text{Ejemplo} \quad \frac{\text{Número de nacidos vivos en la ciudad}}{\text{Número de habitantes en la ciudad}} \quad 1\,000$$

$$\text{Tasa de nacidos vivos en Santa Lucía: } \frac{10\,000}{300\,000} \times 1\,000 = 33.33$$

Es decir, hay 33.33 nacidos vivos por cada 1 000 habitantes en Santa Lucía.

Corolario

Hemos analizado descriptivamente los datos por *variable del estudio* y los visualizamos gráficamente. En caso de que alguna distribución resulte ilógica, debemos cuestionarnos si la variable debe ser excluida, sea por errores del instrumento de medición o en la recolección de los datos, ya que la codificación puede ser verificada. Supongamos que en una investigación en empresas, al medir la satisfacción laboral, resulta que 90% se encuentra "sumamente satisfecho" (¿es lógico?); u otro caso sería que, en ingresos anuales el promedio fuera de 15 000 dólares por familia (¿resulta creíble en tal municipio?). La tarea es revisar la información descriptiva de todas las variables y verificar su veracidad.

Asimismo, si nos encontramos un porcentaje alto de valores perdidos (por ejemplo, de 20%),[10] debemos preguntarnos: ¿por qué tantos participantes no respondieron o contestaron erróneamente?, ¿por qué no se tienen registros completos de todos los casos, eventos o unidades de análisis? (como datos de laboratorio en un estudio clínico, mediciones de ciertas reacciones químicas, etcétera).

Ahora, debemos demostrar la confiabilidad y validez de nuestro instrumento, sobre la base de los datos recolectados.

Paso 4: evaluar la confiabilidad o fiabilidad y validez lograda por el instrumento de medición

La confiabilidad se calcula y evalúa para todo el instrumento de medición utilizado, o bien, si se administraron varios instrumentos, se determina para cada uno de ellos. Asimismo, es común que el instrumento contenga varias escalas para diferentes variables o dimensiones, entonces la fiabilidad se establece para cada escala y para el total de escalas (si se pueden sumar, si son aditivas).[11]

Tal como se mencionó en el capítulo 9, existen diversos procedimientos para calcular la confiabilidad de un instrumento conformado por una o varias escalas que miden las variables de la investigación; cuyos ítems, variables de la matriz o indicadores pueden sumarse, promediarse o correlacionarse. Todos utilizan fórmulas que producen coeficientes de fiabilidad que pueden oscilar entre cero y uno, donde recordemos que un coeficiente de cero significa nula confiabilidad y uno representa un máximo de fiabilidad. Cuanto más se acerque el coeficiente a cero (0), mayor error habrá en la medición (Garson, 2013; Franzen, Robbins y Sawicki, 2010; así como Lauriola, 2003). Los coeficientes expresan la intercorrelación (consistencia) entre los distintos ítems, indicadores o componentes de la prueba (Knapp, 2013; Cervantes, 2005; Cortina, 1993; y Carmines y Zeller, 1991).

Los procedimientos más utilizados para determinar la confiabilidad mediante un coeficiente son:[12]

1. *Medida de estabilidad* (confiabilidad por test-retest). En este procedimiento un mismo instrumento de medición se aplica dos o más veces a un mismo grupo de personas o casos, después de cierto periodo. Si la correlación entre los resultados de las diferentes aplicaciones es muy positiva, el instrumento se considera confiable (Rodríguez, 2006a y Krauss y Chen, 2003). Se trata de una especie de diseño de panel. Desde luego, el periodo entre las mediciones es un factor que hay que considerar. Si el periodo es largo y la variable o el contexto son susceptibles de cambios, ello suele confundir la interpretación del coeficiente de fiabilidad obtenido por este procedimiento. Y si

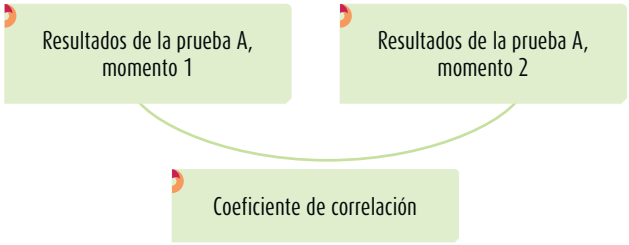
[10] Un porcentaje de valores perdidos (*missing data*) no debe ser mayor de 15%, no es razonable (Creswell, 2005). Cuando tenemos valores perdidos, podemos ignorarlos o sustituirlos por el valor promedio obtenido del total de puntuaciones válidas (esto lo hacen muchos programas de análisis) si así lo deseamos, y puede ser una solución (McKnight *et al.*, 2007). Para profundizar en el tema, véase Leeuw y Hox (2008), Allison (2001) y Enders (2010), en este orden. En las áreas de ingeniería a Latini y Passerini (2003) y en el caso de ciencias de la salud a Allison (2007) y a O'Kelly y Ratitch (2014).

[11] En aparatos o sistemas se verifica la calibración o consistencia entre diversas mediciones.

[12] Un coeficiente de fiabilidad es una medida de la proporción de varianza verdadera en relación con la varianza total observada a través de las puntuaciones o valores resultantes de la administración de un instrumento o protocolo de medición (prueba, escala, calificación de experto, etc.) a una muestra de individuos (Lauriola, 2003). Se aplican más bien a mediciones que involucran respuestas de personas, aunque pueden adaptarse a otros casos que impliquen correlacionar resultados (por ejemplo, a datos de organismos biológicos, protocolos de procesos o aparatos, etcétera).

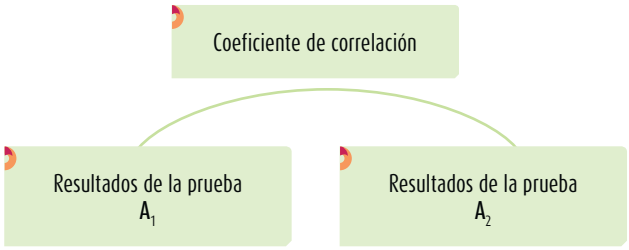
el periodo es corto las personas pueden recordar cómo respondieron en la primera aplicación del instrumento, para aparecer como más consistentes de lo que en realidad son (Bohrnstedt, 1976). El proceso de cálculo con dos aplicaciones se representa en la figura 10.10.

Figura 10.10 Medida de estabilidad.



2. *Método de formas alternativas o paralelas.* En este esquema no se administra el mismo instrumento de medición, sino dos o más versiones equivalentes de éste. Las versiones (casi siempre dos) son similares en contenido, instrucciones, duración y otras características, y se administran a un mismo grupo de personas simultáneamente o dentro de un periodo corto. El instrumento es confiable si la correlación entre los resultados de ambas administraciones es positiva de manera significativa (Rodríguez, 2006b). Los patrones de respuesta deben variar poco entre las aplicaciones. Una variación de este método es el de las formas alternas prueba-posprueba (Creswell, 2005), cuya diferencia reside en que el tiempo que transcurre entre la administración de las versiones es mucho más largo, que es el caso de algunos experimentos. El método se representa en la figura 10.11.

Figura 10.11 Método de formas alternativas o paralelas.



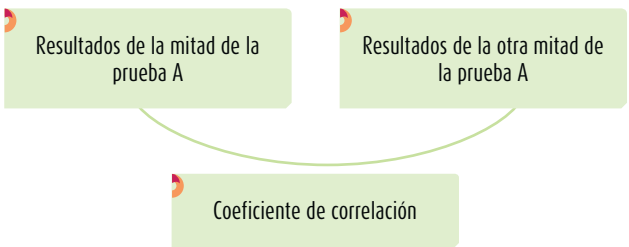
Estos dos métodos (estabilidad y formas alternas) también pueden aplicarse cuando utilizamos dos instrumentos distintos para medir las mismas variables en las unidades de análisis (por ejemplo, dos sistemas para medir propiedades eléctricas o dos protocolos para medir la presión arterial).

3. *Método de mitades partidas (split-halves).* Los procedimientos anteriores requieren cuando menos dos administraciones de la medición en la muestra. En cambio, el **método de mitades partidas** necesita sólo una aplicación de la medición. Específicamente, el conjunto total de ítems o reactivos se divide en dos mitades equivalentes y se comparan las puntuaciones o resultados de ambas. Si el instrumento es confiable, las puntuaciones de las dos mitades deben estar muy correlacionadas (Rodríguez, 2006c y McKelvie, 2003). Un individuo con baja puntuación en una mitad tenderá a mostrar también una baja puntuación en la otra mitad. El procedimiento se diagrama en la figura 10.12.

Métodos de mitades partidas y de consistencia interna Se aplican a instrumentos que implican medidas compuestas o escalas, es decir, están constituidas por varios ítems, indicadores o mediciones.

4. *Medidas de coherencia o consistencia interna*. Éstos son coeficientes que estiman la confiabilidad: *a*) *el alfa de Cronbach* (desarrollado por J.L. Cronbach) y *b*) *los coeficientes KR-20 y KR-21* de Kuder y Richardson (1937). El método de cálculo de éstos requiere una sola administración del instrumento de medición. Su ventaja reside en que no es necesario dividir en dos mitades a los ítems del instrumento, simplemente se aplica la medición y se calcula el coeficiente. La mayoría de los programas estadísticos como SPSS y Minitab los determinan y solamente deben interpretarse.

Figura 10.12 Método de mitades partidas.



Respecto a la interpretación de los distintos coeficientes mencionados cabe señalar que no hay una regla que indique "a partir de este valor no hay fiabilidad del instrumento". Más bien, el investigador calcula su valor, lo declara y lo somete a escrutinio de los usuarios del estudio u otros investigadores, explicitando el método utilizado (Chen y Krauss, 2003; McKelvie, 2003; Lauriola, 2003; y Carmines y Zeller, 1991). Algunos autores consideran que el coeficiente debe estar entre 0.70 y 0.90 (Tavakol y Dennick, 2011; DeVellis, 2003; Streiner, 2003; Nunnally y Bernstein, 1994; Petterson, 1994). Nunnally (1987) por encima de 0.80. Lauriola (2003) sugiere un valor mínimo de 0.70 para la comparación entre grupos y 0.90 para escalas. Garson (2013) establece que 0.60 es aceptable para

propósitos exploratorios y 0.70 para fines confirmatorios, resultando 0.80 "bueno" en un alcance explicativo. Ahora bien, también un coeficiente mayor de 0.90 puede implicar redundancia de ítems o indicadores y la necesidad de reducir el instrumento (Tavakol y Dennick, 2011).

Con respecto a los métodos basados en coeficientes de correlación, usted se formará una idea más clara después de revisar el apartado de correlación que se presenta más adelante en este capítulo. Pero hay una consideración importante que hacer ahora. El coeficiente que elijamos para determinar la confiabilidad debe ser apropiado al nivel de medición de la escala de nuestra variable (por ejemplo, si la escala de mi variable es por intervalos, puedo utilizar el coeficiente de correlación de Pearson; pero si es ordinal podré usar el coeficiente de Spearman o de Kendall; y si es nominal, otros coeficientes). El *alfa* de *Cronbach* trabaja con variables de intervalos o de razón, KR-20 y KR-21 con ítems dicotómicos (por ejemplo: sí-no)[13] y $\rho xx'$ con reactivos tricotómicos (Knapp, 2013; Alkharusi, 2010; Vittengl, White, McGovern y Morton, 2006; y Feldt, 2005). Además, existen otros coeficientes como el alfa estratificado, la confiabilidad máxima, los coeficientes de Raju, Kristof, Angoff-Feldt, Feldt-Gilmer, Guttman λ_2, λ_4 maximizado y el análisis de Hoyt.

El cálculo del coeficiente *alfa* y algunas consideraciones sobre los factores que lo afectan se incluyen en el capítulo 8 adicional: "Análisis estadístico: segunda parte", que se encuentra en "Material complementario".

Centro de recursos en línea

Con la finalidad de comprender mejor los métodos para determinar la confiabilidad, véase la tabla 10.10.

Tabla 10.10 Aspectos básicos de los métodos para determinar la confiabilidad de instrumentos aplicados a personas

Método	Número de veces en que el instrumento es administrado	Número de versiones diferentes del instrumento	Número de participantes que proveen los datos	Inquietud o pregunta que contesta
Estabilidad (*test-retest*)	Dos veces en tiempos distintos	Una versión	Cada participante responde al instrumento dos veces.	¿Responden los individuos de una manera similar a un instrumento si se les administra dos veces?
Formas alternas	Dos veces al mismo tiempo o con una diferencia de tiempo muy corta	Dos versiones diferentes, pero equivalentes	Cada participante responde a cada versión del instrumento.	Cuando dos versiones de un instrumento son similares, ¿hay convergencia o divergencia en las respuestas a ambas versiones?
Formas alternas y prueba-posprueba	Dos veces en tiempos distintos	Dos versiones diferentes, pero equivalentes	Cada participante responde a cada versión del instrumento.	Cuando dos versiones de un instrumento son similares, ¿hay convergencia o divergencia en las respuestas a ambas versiones?
Mitades partidas	Una vez	Una fragmentada en dos partes equivalentes	Cada participante responde a la única versión.	¿Son las puntuaciones de una mitad del instrumento similares a las obtenidas en la otra mitad?
Medidas de consistencia interna (alfa y KR-20 y 21)	Una vez	Una versión	Cada participante responde a la única versión.	¿Las respuestas a los ítems del instrumento son coherentes?

[13] Estos dos coeficientes se usan en el método de "mitades partidas", aunque —como señalan Babbie (2012) y Creswell (2005)— se confía en la mitad de la información del instrumento, por lo que conviene agregar el cálculo de "profecía" Spearman-Brown.

Asimismo, en la tabla 10.11 se presentan ejemplos de estudios con su respectiva confiabilidad.

Tabla 10.11 Ejemplos de confiabilidad

Investigación	Instrumento	Métodos de cálculo y resultados	Comentario
Evaluación de los conocimientos, opiniones, experiencias y acciones en torno al abuso sexual infantil (Kolko *et al.*, 1987).	Escala cognitiva de nueve ítems para infantes en edades preescolares y primeros grados básicos.	Coherencia interna alfa = 0.34.	Confiabilidad extremadamente baja que demuestra incongruencia, atribuida por los autores a lo corto de la escala (pocos ítems).
Desarrollo y validación de una escala autoaplicable para medir la satisfacción sexual en adultos varones y mujeres de México (Álvarez Gayou, Honold y Millán, 2005).	Un inventario para medir la satisfacción sexual que está integrado por 29 reactivos y fue administrado a una muestra de 760 personas, de ambos géneros, cuyas edades fluctuaron entre los 16 y 65 años.	La confiabilidad del inventario establecida por medio de una prueba alfa de Cronbach fue de 0.92.	El valor α indica una fiabilidad sumamente elevada.
Validación de un instrumento para medir la cultura empresarial en función del clima organizacional y vincular empíricamente ambos constructos (Hernández-Sampieri, Méndez y Contreras, 2013).	Cuestionario estandarizado que mide el clima organizacional en función del Modelo de los Valores en Competencia de Quinn y Rohrbaugh, a través de escalas tipo Likert con cuatro opciones de respuesta: dos positivas y dos negativas.	El coeficiente alfa-Cronbach obtenido resultó igual a 0.95 (con 95 ítems). La muestra estuvo conformada por 1 424 empleados de 12 empresas (972 casos válidos completos).	Confiabilidad muy elevada. No hay redundancia de ítems pues se midieron 17 variables del clima organizacional (asociadas entre sí). Los coeficientes para las escalas variaron entre 0.60 y 0.90 (y uno sólo de 0.53).
Actitudes hacia el matrimonio: integración y sus resultados en las relaciones personales (Riggio y Weiser, 2008).	Escalas del Modelo de Inversión (IMS), las cuales a partir de 37 reactivos (cada uno con 9 categorías) miden la entrega, la inversión psicológica y la satisfacción con respecto a una relación romántica actual.	Los coeficientes alfa resultantes de aplicar las escalas a 400 universitarios fueron: 0.94 para entrega y satisfacción, y 0.88 para inversión psicológica.	Coeficientes muy considerables para entrega y satisfacción, y bastante aceptable para inversión.
Fukuda, Saklofske, Tamaoka y Lim, 2012.	Prueba WLEIS de 16 reactivos que mide la inteligencia emocional en adultos jóvenes en cuatro dimensiones: valoración de las propias emociones, valoración de las emociones de los demás, uso de las emociones y control o regulación de las emociones.	Los coeficientes resultantes de aplicar el WLEIS (n=161) fueron por subescala o dimensión: valoración de las propias emociones: 0.80, valoración de las emociones de los demás: 0.74, uso de las emociones: 0.74 y regulación de las emociones: 0.83.	Coeficientes aceptables dentro de los parámetros normales, particularmente porque las escalas tienen pocos ítems.

Otro caso es el ya comentado de Núñez (2001) y su instrumento para medir el sentido de vida, cuya fiabilidad fue de 0.96 en su tercera versión con 99 ítems.

Como podemos observar en la tabla 10.11, cuanta más información se proporcione sobre la confiabilidad, el lector se forma una idea más clara sobre su cálculo y las condiciones en que se demostró. Es indispensable incluir las dimensiones de la variable medida, el tamaño de muestra y el método utilizado. Una cuestión importante es que los coeficientes son sensibles al número de ítems o reactivos: entre más agreguemos, el valor del coeficiente tenderá a ser más elevado.

Además de estimar un coeficiente de correlación o un coeficiente de coherencia entre los ítems del instrumento, es conveniente calcular la correlación ítem-escala completa. Ésta representa la vinculación de cada reactivo con toda la escala. Habrá tantas correlaciones como ítems contenga el instrumento. Corbetta (2003, p. 237) lo ejemplifica adecuadamente de la siguiente manera: si estamos midiendo el autoritarismo, es lógico pensar que quien alcanza altas puntuaciones en esta variable en toda la escala (es muy autoritaria), habrá de tener puntuaciones elevadas en todos los ítems que la conforman. Pero si uno de los reactivos sistemáticamente (en un número considerable de individuos) presenta valores contradictorios con respecto a la escala total, podemos concluir que ese ítem no fun-

ciona correctamente (contradice a los demás reactivos). Los ítems que alcancen coeficientes de correlación bajos con la escala tal vez deban analizarse y, eventualmente, eliminarse.

Asimismo, cada uno de los reactivos puede ser evaluado en su capacidad de discriminación mediante la prueba *t* de Student (paramétrica). Se consideran dos grupos, el primero integrado por 25% de los casos con los puntajes más altos obtenidos en el ítem y el otro grupo compuesto por 25% de los casos con los puntajes más bajos. Los ítems cuya prueba no resulte significativa serán reconsiderados.

Los conceptos estadísticos aquí vertidos (por ejemplo, correlación y prueba *t*) tendrán mayor sentido, una vez que se revisen más ampliamente, lo cual se hará más adelante en este capítulo.

Para determinar la confiabilidad usando los programas estadísticos no olvide consultar los respectivos manuales, descargándolos del centro de recursos en línea.

Validez

Vimos en el capítulo anterior que la evidencia sobre la validez del contenido se obtiene mediante las opiniones de expertos y al asegurarse de que las dimensiones medidas por el instrumento sean representativas del universo o dominio de dimensiones de las variables de interés (a veces mediante un muestreo aleatorio simple). La evidencia de la validez de criterio se produce al correlacionar las puntuaciones de los participantes, obtenidas por medio del instrumento, con sus valores logrados en el criterio. Recordemos que una correlación implica asociar puntuaciones obtenidas por la muestra en dos o más variables.

Por ejemplo, Núñez (2001), además de aplicar su instrumento sobre el sentido de vida, administró otras dos pruebas que teóricamente miden variables similares: el PIL (Propósito de Vida) y el Logo-test de Elizabeth Lukas. El coeficiente de correlación de Pearson entre el instrumento diseñado y el PIL fue de 0.541, valor que se considera moderado. El coeficiente de correlación *rho* de Spearman fue igual a 0.42 entre el Logo-test y su prueba, lo cual indica que los tres instrumentos no miden la misma variable, pero sí conceptos relacionados.

La evidencia de la validez de constructo se obtiene mediante el análisis de factores. Tal método nos indica cuántas dimensiones integran a una variable y qué ítems conforman cada dimensión. Los reactivos que no pertenezcan a una dimensión, quiere decir que están "aislados" y no miden lo mismo que los demás ítems, por tanto, deben eliminarse. Es un método que tradicionalmente se ha considerado complejo, por los cálculos estadísticos implicados, pero que es relativamente sencillo de interpretar y como los cálculos hoy en día los realiza la computadora, está al alcance de cualquier persona que se inicie dentro de la investigación. Este método se revisa —con ejemplos reales— en el capítulo 8 adicional del centro de recursos en línea: "Análisis estadístico: segunda parte".

Para cada escala, una vez que se determina la confiabilidad (de 0 a 1) y se muestra la evidencia sobre la validez, si algunos ítems son problemáticos (no discriminan, no se vinculan a otros ítems, van en sentido contrario a toda la escala, no miden lo mismo, etc.), se eliminan de los cálculos (pero en el reporte de la investigación, se indica cuáles fueron descartados, las razones de ello y cómo alteran los resultados); posteriormente se vuelve a realizar el análisis descriptivo (distribución de frecuencias, medidas de tendencia central y de variabilidad, etcétera).

En el centro de recursos → Material complementario → Ejemplos → Ejemplo 4, "Diseño de una escala autoaplicable para la evaluación de la satisfacción sexual en hombres y mujeres mexicanos" (Álvarez Gayou, Honold y Millán, 2005), se presenta la validación de un instrumento que muestra todos los elementos para ello, paso por paso. Incluye la generación de redes semánticas. Su abordaje es desde el punto de vista de la salud y con propiedad científica. Se recomienda descargarlo y revisarlo.

¿Hasta aquí llegamos?

Cuando el estudio tiene una finalidad puramente exploratoria o descriptiva, debemos interrogarnos: ¿podemos establecer relaciones entre variables? En caso de una respuesta positiva, es factible seguir con la estadística inferencial; pero si dudamos o el alcance se limitó a explorar y describir, el trabajo de análisis concluye y debemos comenzar a preparar el reporte de la investigación.

Paso 5: analizar mediante pruebas estadísticas las hipótesis planteadas (análisis estadístico inferencial)

En este paso se analizan las hipótesis a la luz de pruebas estadísticas que a continuación detallamos.

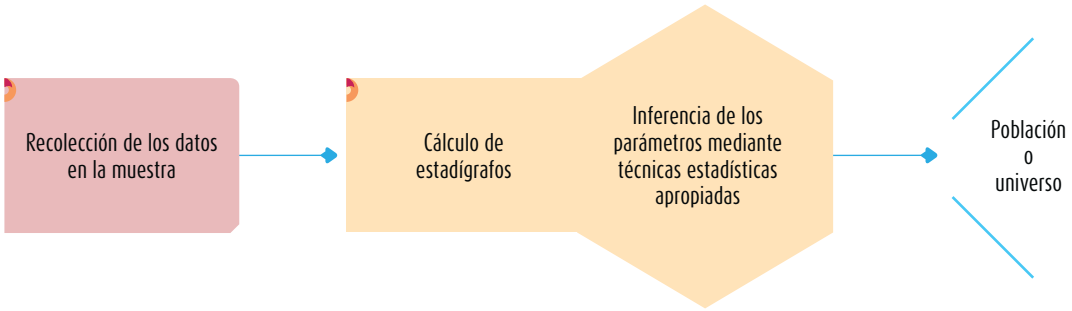
Estadística inferencial: de la muestra a la población

3 y 4

Con frecuencia, el propósito de la investigación va más allá de describir las distribuciones de las variables: se pretende probar hipótesis y generalizar los resultados obtenidos en la muestra a la población o universo. Los datos casi siempre se recolectan de una muestra y sus resultados estadísticos se denominan *estadígrafos*; la media o la desviación estándar de la distribución de una muestra son estadígrafos. A las estadísticas de la población se les conoce como *parámetros*. Éstos no son calculados, porque no se recolectan datos de toda la población, pero pueden ser inferidos de los estadígrafos, de ahí el nombre de **estadística inferencial**. El procedimiento de esta naturaleza de la estadística se esquematiza en la figura 10.13.

Estadística inferencial Estadística para probar hipótesis y estimar parámetros.

Figura 10.13 Procedimiento de la estadística inferencial.



Entonces, la estadística inferencial se utiliza fundamentalmente para dos procedimientos vinculados (O'Leary, 2014; Punch, 2014; Babbie, 2012; Wiersma y Jurs, 2008; Waterman, 2007; Kulikowich y Edwards, 2006; y Maxim, 2003):

a) *Probar hipótesis poblacionales*
b) *Estimar parámetros*

En este capítulo comentaremos la prueba de hipótesis, que se efectúa dependiendo del tipo de hipótesis de que se trate. Existen pruebas estadísticas para diferentes clases de hipótesis como iremos viendo.

La inferencia de los parámetros depende de que hayamos elegido una muestra probabilística con un tamaño que asegure un nivel de significancia o significación adecuado (Jarman, 2013; Lindsay, 2009; y Moriceau, 2009). En el centro de recursos encontrará un ejemplo de inferencia sobre la hipótesis de la media poblacional, en: Material Complementario → Capítulos → Capítulo 8, "Análisis estadístico: segunda parte".

Centro de recursos en línea

¿En qué consiste la prueba de hipótesis?

3

Una hipótesis en el contexto de la estadística inferencial es una proposición respecto de uno o varios parámetros, y lo que el investigador hace por medio de la prueba de hipótesis es determinar si la hipótesis poblacional es congruente con los datos obtenidos en la muestra (Wilcox, 2012; Gordon, 2010; Wiersma y Jurs, 2008; y Stockburger, 2006).

Una hipótesis se retiene como un valor aceptable del parámetro, si es consistente con los datos. Si no lo es, se rechaza (pero los datos no se descartan). Para comprender lo que es la prueba de hipó-

tesis en la estadística inferencial es necesario revisar los conceptos de distribución muestral[14] y nivel de significancia.[15]

¿Qué es una distribución muestral?

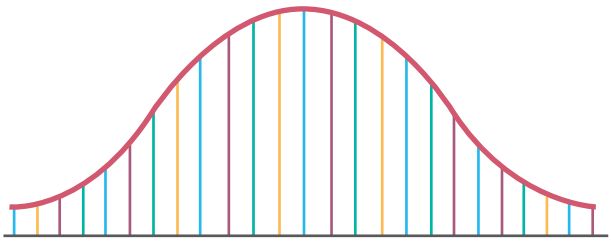
Distribución muestral Conjunto de valores sobre una estadística calculada de todas las muestras posibles de una población.

Una **distribución muestral** es un conjunto de valores sobre una estadística calculada de todas las muestras posibles de determinado tamaño de una población (Bond, 2007a). Las distribuciones muestrales de medias son probablemente las más conocidas. Expliquemos este concepto con un ejemplo. Supongamos que nuestro universo son los automovilistas de una ciudad y deseamos averiguar cuánto tiempo pasan diariamente manejando ("al volante"). De este universo podría extraerse una muestra representativa. Vamos a suponer que el tamaño adecuado de muestra es de 512 automovilistas ($n = 512$). Del mismo universo se podrían extraer diferentes muestras, cada una con 512 personas.

3

Teóricamente, incluso podría elegirse al azar una, dos, tres, cuatro muestras, y las veces que fuera necesario hacerlo, hasta agotar todas las muestras posibles de 512 automovilistas de esa ciudad (todos los individuos serían seleccionados en varias muestras). En cada muestra se obtendría una media del tiempo que pasan los automovilistas manejando. Tendríamos pues, una gran cantidad de medias, tantas como las muestras extraídas ($\overline{X}_1$, $\overline{X}_2$, $\overline{X}_3$, $\overline{X}_4$, $\overline{X}_5$, ... $\overline{X}_k$). Y con éstas elaboraríamos una distribución de medias. Habría muestras que, en promedio, pasaran más tiempo "al volante" que otras. Este concepto se representa en la figura 10.14.

Figura 10.14 Distribución muestral de medias.



Son medias ($\overline{X}$) no se trata de puntuaciones. Cada media representaría una muestra.

Si calculáramos la media de todas las medias de las muestras, prácticamente obtendríamos el valor de la media poblacional. De hecho, casi nunca se obtiene la distribución muestral (la distribución de las medias de todas las muestras posibles). Es más bien un concepto teórico definido por la estadística para los investigadores. Lo que comúnmente hacemos es extraer una sola muestra.

En el ejemplo de los automovilistas, sólo una de las líneas verticales de la distribución muestral presentada en la figura 10.14 es la media obtenida para nuestra única muestra seleccionada de 512 personas. Y la pregunta es: ¿nuestra media calculada se encuentra cerca de la media de la distribución muestral?, debido a que si está cerca podremos tener una estimación precisa de la media poblacional (el parámetro poblacional es prácticamente el mismo que el de la distribución muestral). Esto se expresa en el *teorema central del límite*:

> Si una población (no necesariamente normal) tiene de media m y de desviación estándar s, la distribución de las medias en el muestreo aleatorio realizado en esta población tiende, al aumentar n, a una distribución normal de media m y desviación estándar $\frac{s}{\sqrt{n}}$, donde n es el tamaño de muestra.

El teorema especifica que la distribución muestral tiene una media igual a la de la población, una varianza igual a la varianza de la población dividida entre el tamaño de muestra (su desviación estándar es $\frac{\sigma}{\sqrt{n}}$ y se distribuye normalmente). La desviación estándar (s) es un parámetro normalmente desconocido, aunque es posible estimarlo por la desviación estándar de la muestra. Asimismo, en el capítulo 8 se dijo que cuando las muestras están constituidas por 100 o más elementos tienden a presentar **distribuciones normales** y esto sirve para el propósito de hacer estadística inferencial. La "normalidad" de la distribución en muestras grandes no obedece a la normalidad de la distribución de una población. La distribución de diversas variables a veces es "normal" y en ocasiones está lejos de serlo. Sin embargo, la normalidad no debe confundirse con probabilidad. Mientras lo primero es necesario para efectuar ciertas pruebas estadísticas, lo segundo es requisito indispensable para

Distribución normal Distribución en forma de campana que se logra con muestras de 100 o más unidades muestrales y que es útil y necesaria cuando se hacen inferencias estadísticas.

[14] Distribución muestral y distribución de una muestra son conceptos diferentes: la última es resultado de los datos de nuestra investigación y es por variable.

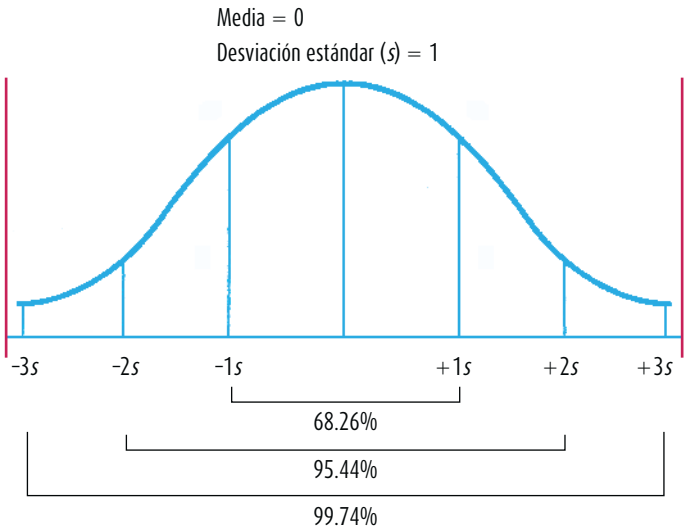
[15] El término *significancia* es un anglicismo, por lo que diversos autores sugieren mejor utilizar "significación" o "significatividad" (por ejemplo: Korniejczuk, 2012).

hacer inferencias correctas sobre una población. El concepto de *distribución normal* es importante otra vez y se ofrece una breve explicación en la figura 10.15.

Una gran cantidad de los fenómenos del comportamiento humano se manifiestan de la siguiente forma: la mayoría de las puntuaciones se concentran en el centro de la distribución, en tanto que en los extremos encontramos sólo algunas puntuaciones (Fu, 2007). Por ejemplo, la inteligencia: hay pocas personas muy inteligentes (genios), pero también hay pocas personas con muy baja inteligencia. La mayoría de los seres humanos somos medianamente inteligentes. Esto podría representarse así:

Figura 10.15 Concepto de curva o distribución normal.

Debido a ello, se creó un modelo de probabilidad llamado curva normal o distribución normal. Como todo modelo es una distribución conceptual que difícilmente se presenta en la realidad tal cual, pero sí se presentan aproximaciones a éste. La curva normal tiene la siguiente configuración:



68.26% del área de la curva normal es cubierta entre $-1s$ y $+1s$, 95.44% del área de esta curva es cubierta entre $-2s$ y $+2s$ y 99.74% se cubre con $-3s$ y $+3s$.

Las principales características de la distribución normal son:

1. Es *unimodal*, una sola moda.
2. La *asimetría es cero*. La mitad de la curva es exactamente igual a la otra mitad. La distancia entre la media y $-3s$ es la misma que la distancia entre la media y $+3s$.
3. *Es una función* particular entre desviaciones con respecto a la media de una distribución y la probabilidad de que éstas ocurran.
4. La base está dada en unidades de desviación estándar (puntuaciones z), destacando las puntuaciones $-1s$, $-2s$, $-3s$, $+1s$, $+2s$ y $+3s$ (que equivalen respectivamente a $-1.00z$, $-2.00z$, $-3.00z$, $+1.00z$, $+2.00z$, $+3.00z$). Las distancias entre puntuaciones z representan áreas bajo la curva. De hecho, la distribución de puntuaciones z es la curva normal.
5. Es *mesocúrtica* (curtosis de cero).
6. La *media*, la *mediana* y la *moda* coinciden en el mismo punto (el centro).

Nivel de significancia Nivel de la probabilidad de equivocarse y que fija de manera *a priori* el investigador.

¿Qué es el nivel de significancia o significación?

Wiersma y Jurs (2008) ofrecen una explicación sencilla del concepto, en la cual nos basaremos para analizar su significado. La probabilidad de que un evento ocurra oscila entre cero (0) y uno (1), donde cero implica la imposibilidad de ocurrencia y uno la certeza de que el fenómeno ocurra. Al lanzar al aire una moneda no cargada, la probabilidad de que salga "cruz" es de 0.50 y la probabilidad de que la moneda caiga en "cara" también es de 0.50. Con un dado, la probabilidad de obtener cualquiera de sus caras al lanzarlo es de 1/6 = 0.1667. La suma de posibilidades siempre es de uno.

Aplicando el concepto de probabilidad a la distribución muestral, tomaremos el área de ésta como 1.00; en consecuencia, cualquier área comprendida entre dos puntos de la distribución corresponderá a la probabilidad de la distribución. Para probar hipótesis inferenciales respecto a la media, el investigador debe evaluar si es alta o baja la probabilidad de que la media de la muestra esté cerca de la media de la distribución muestral. Si es baja, el investigador dudará de generalizar a la población. Si es alta, el investigador podrá hacer generalizaciones. Es aquí donde entra el nivel de significancia o nivel alfa (α),[16] el cual es un nivel de la probabilidad de equivocarse y se fija antes de probar hipótesis inferenciales.

Este concepto fue esbozado en el capítulo 8 con un ejemplo coloquial, pero lo volvemos a recordar: si fuera a apostar en las carreras de caballos y tuviera 95% de probabilidades de atinarle al ganador, contra sólo 5% de perder, ¿apostaría? Obviamente sí, siempre y cuando le aseguraran ese 95% en favor.

Pues bien, algo parecido hace el investigador. Obtiene una estadística en una muestra (por ejemplo, la media) y analiza qué porcentaje tiene de confianza en que dicha estadística se acerque al valor de la distribución muestral (que es el valor de la población o el parámetro). Busca un alto porcentaje de certeza, una probabilidad elevada para estar tranquilo, porque sabe que tal vez haya error de muestreo y, aunque la evidencia parece mostrar una aparente "cercanía" entre el valor calculado en la muestra y el parámetro, tal "cercanía" puede no ser real o deberse a errores en la selección de la muestra.

¿Con qué porcentaje de confianza el investigador generaliza, para suponer que tal cercanía es real y no por un error de muestreo? Existen dos niveles convenidos en las ciencias:

a) *El nivel de significancia de 0.05*, el cual implica que el investigador tiene 95% de seguridad para generalizar sin equivocarse y sólo 5% en contra. En términos de probabilidad, 0.95 y 0.05, respectivamente; ambos suman la unidad. Este nivel es el más común en ciencias sociales.
b) *El nivel de significancia de 0.01*, el cual implica que el investigador tiene 99% en su favor y 1% en contra (0.99 y 0.01 = 1.00) para generalizar sin temor. Muy utilizado cuando las generalizaciones implican riesgos vitales para las personas (pruebas de vacunas, medicamentos, arneses de aviones, resistencia de materiales de construcción al fuego o el peso, etcétera).

A veces el nivel de significancia o significación puede ser todavía más riguroso, por ejemplo, 0.001, 0.00001, 0.00000001 (Liao, 2003), pero al menos debe ser de 0.05. No se acepta un nivel de 0.06 (94% a favor de la generalización confiable), porque se busca hacer ciencia lo más exacta posible.[17]

Tal nivel es un valor de certeza que el investigador fija *a priori*, respecto a no equivocarse (Capraro, 2006). Cuando uno lee en un reporte de investigación que los resultados fueron significativos al nivel de 0.05 ($p < 0.05$), indica lo que se comentó: que existe 5% de posibilidad de error al aceptar la hipótesis, correlación o valor obtenido al aplicar una prueba estadística; o 5% de riesgo de que se rechace una hipótesis nula cuando era verdadera (Babbie, 2012 y Mertens, 2010). Volveremos más adelante sobre este punto.

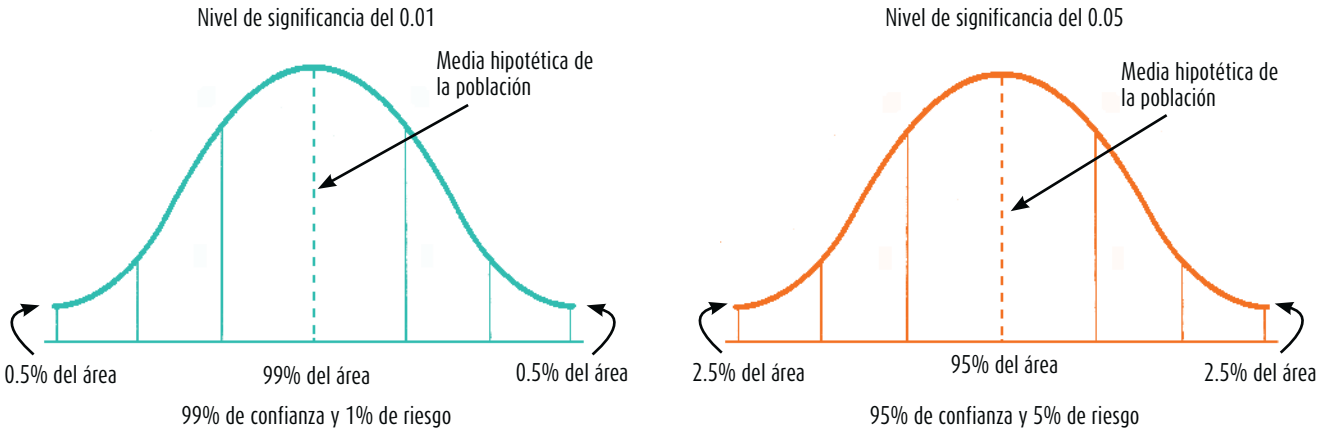
[16] No confundir con el coeficiente alfa de Cronbach, para determinar la confiabilidad.
[17] El nivel de significancia mínimo aceptable es definido por las asociaciones científicas correspondientes al ramo o área en la cual se investiga, incluyendo comités editoriales de revistas académicas.

¿Cómo se relacionan la distribución muestral y el nivel de significancia?

El *nivel de significancia o significación* se expresa en términos de probabilidad (0.05 y 0.01) y la *distribución muestral* también como probabilidad (el área total de ésta como 1.00). Pues bien, para ver si existe o no confianza al generalizar acudimos a la distribución muestral, con una probabilidad adecuada para la investigación. Dicho nivel lo tomamos como un área bajo la distribución muestral, como se observa en la figura 10.16, y depende de si elegimos un nivel de 0.05 o de 0.01. Es decir, que nuestro valor estimado en la muestra no se encuentre en el área de riesgo y estemos lejos del valor de la distribución muestral, que insistimos es muy cercano al de la población.

Así, el nivel de significación representa áreas de riesgo o confianza en la distribución muestral.

Figura 10.16 Niveles de significancia o significación en la distribución muestral.



Notas:

1. Podemos expresarlo en proporciones (0.025, 0.95 y 0.025, respectivamente) o porcentajes como está en la gráfica.
2. Tanto 99% como 95% representan las áreas de confianza de que nuestra estimación se localiza dentro de ellas. La primera al nivel del 0.01 y la segunda al nivel de 0.05. El área de riesgo en el primer caso es de 1% (0.5 + 0.5 = 1%) y en el segundo de 5% (2.5% + 2.5% = 5%) sumando ambos extremos, porque en nuestra estimación de la media poblacional podríamos pasarnos (error) hacia valores más altos o bajos.

¿Se pueden cometer errores al probar hipótesis y realizar estadística inferencial?

3 y 4

Nunca estaremos completamente seguros de nuestra estimación. Trabajamos con altos niveles de confianza o seguridad, pero, aunque el riesgo es mínimo, podría cometerse un error. *Los resultados posibles al probar hipótesis son*:[18]

1. Aceptar una hipótesis verdadera (decisión *correcta*).
2. Rechazar una hipótesis falsa (decisión *correcta*).
3. Aceptar una hipótesis falsa (conocido como *error* del *Tipo II o error beta*).
4. Rechazar una hipótesis verdadera (conocido como *error* del *Tipo I o error alfa*).

Ambos tipos de error son indeseables; sin embargo, puede *reducirse sustancialmente la posibilidad* de que se presenten mediante:

a) *Muestras probabilísticas representativas.*
b) *Inspección cuidadosa de los datos.*
c) *Selección de las pruebas estadísticas apropiadas.*
d) *Mayor conocimiento de la población.*

[18] Yaremko *et al.* (2013), Cozby y Bates (2012), Ravid (2011), Mertens (2010), Buskirk (2008) y Wiersma y Jurs (2008).

Prueba de hipótesis

4

Hay dos tipos de análisis estadísticos que pueden realizarse para probar hipótesis: los *análisis paramétricos* y los *no paramétricos.* Cada tipo posee sus características y presuposiciones que lo sustentan; la elección de qué clase de análisis efectuar depende de los supuestos. De igual forma, cabe destacar que en una misma investigación es posible llevar a cabo análisis paramétricos para algunas hipótesis y variables, y análisis no paramétricos para otras. Asimismo, como vimos, los análisis a realizar dependen del planteamiento, tipo de hipótesis y el nivel de medición de las variables que las conforman.

Análisis paramétricos

3 y 4

Para realizar análisis paramétricos debe partirse de los siguientes supuestos:[19]

1. La *distribución poblacional de la variable dependiente es normal*: el universo tiene una distribución normal.
2. El *nivel de medición* de las variables es *por intervalos o razón*.
3. Cuando *dos o más poblaciones son estudiadas, tienen una varianza homogénea*: las poblaciones en cuestión poseen una dispersión similar en sus distribuciones.

Ciertamente estos criterios son tal vez demasiado rigurosos y algunos investigadores sólo basan sus análisis en el tipo de hipótesis y los niveles de medición de las variables. Esto queda a juicio del lector. En la investigación académica y cuando quien la realiza es una persona experimentada, sí debe solicitársele tal rigor.

¿Cuáles son los métodos o las pruebas estadísticas paramétricas más utilizados?

Existen diversas pruebas paramétricas, pero las más utilizadas son:

- Coeficiente de correlación de Pearson y regresión lineal.
- Prueba *t*.
- Prueba de contraste de la diferencia de proporciones.
- Análisis de varianza unidireccional (ANOVA en un sentido).
- Análisis de varianza factorial (ANOVA).
- Análisis de covarianza (ANCOVA).

Centro de recursos en línea

Algunos de estos métodos se tratan aquí en este capítulo y otros se explican en el capítulo 8 adicional, "Análisis estadístico: segunda parte", que puede descargarse del centro de recursos en línea de la obra.

Cada prueba obedece a un tipo de hipótesis de investigación e hipótesis estadística distinta. Las hipótesis estadísticas se comentan en el capítulo 8 del centro de recursos en línea.

¿Qué es el coeficiente de correlación de Pearson?

Es una prueba estadística para analizar la relación entre dos variables medidas en un nivel por intervalos o de razón. Se le conoce también como "coeficiente producto-momento".

Se simboliza: *r*

Hipótesis a probar: correlacional, del tipo de "a mayor *X*, mayor *Y*", "a mayor *X*, menor *Y*", "altos valores en *X* están asociados con altos valores en *Y*", "altos valores en *X* se asocian con bajos valores de *Y*". La hipótesis de investigación señala que la correlación es significativa.

Variables: dos. La prueba en sí no considera a una como independiente y a otra como dependiente, ya que no evalúa la causalidad. La noción de causa-efecto (independiente-dependiente) es posible establecerla teóricamente, pero la prueba no asume dicha causalidad.

[19] O'Leary (2014), Ryan (2013), Babbie (2012), Martin y Bridgmon (2012), Kantor y Kershaw (2010), y Wiersma y Jurs (2008).

El coeficiente de correlación de Pearson se calcula a partir de las puntuaciones obtenidas en una muestra en dos variables. Se relacionan las puntuaciones recolectadas de una variable con las puntuaciones obtenidas de la otra, con los mismos participantes o casos (The SAGE Glossary of the Social and Behavioral Sciences, 2009g; Bagiella, 2007; Onwuegbuzie, Daniel y Leech, 2006a).

Nivel de medición de las variables: intervalos o razón.

Interpretación: el coeficiente *r* de Pearson *puede variar de −1.00 a +1.00,* donde:

−1.00 = *correlación negativa perfecta*. ("A mayor *X*, menor *Y*", de manera proporcional. Es decir, cada vez que *X* aumenta una unidad, *Y* disminuye siempre una cantidad constante). Esto también se aplica "a menor *X*, mayor *Y*".

−0.90 = Correlación negativa muy fuerte.
−0.75 = Correlación negativa considerable.
−0.50 = Correlación negativa media.
−0.25 = Correlación negativa débil.
−0.10 = Correlación negativa muy débil.
0.00 = No existe correlación alguna entre las variables.
+0.10 = Correlación positiva muy débil.
+0.25 = Correlación positiva débil.
+0.50 = Correlación positiva media.
+0.75 = Correlación positiva considerable.
+0.90 = Correlación positiva muy fuerte.
+1.00 = *Correlación positiva perfecta* ("A mayor *X*, mayor *Y*" o "a menor *X*, menor *Y*", de manera proporcional. Cada vez que *X* aumenta, *Y* aumenta siempre una cantidad constante).

El *signo indica la dirección de la correlación* (positiva o negativa); y *el valor numérico, la magnitud de la correlación.* Los principales programas computacionales de análisis estadístico indican si el coeficiente es o no significativo de la siguiente manera:

$r = 0.7831$	(valor del coeficiente)
s o $P = 0.001$	(significancia)
$N = 625$	(número de casos correlacionados)

Si *s* o *P* es menor del valor 0.05, se dice que el coeficiente es *significativo* en el nivel de 0.05 (95% de confianza en que la correlación sea verdadera y 5% de probabilidad de error). Si es menor a 0.01, el coeficiente es *significativo* al nivel de 0.01 (99% de confianza de que la correlación sea verdadera y 1% de probabilidad de error).

O bien, otros programas como IBM SPSS® presentan los coeficientes de correlación en una tabla, donde las filas o columnas son las variables asociadas y se señala con asterisco(s) el nivel de significancia: un asterisco (*) implica que el coeficiente es significativo al nivel del 0.05 y dos asteriscos (**) que es significativo al nivel del 0.01. Esto podemos verlo en el ejemplo de la tabla 10.12:

Tabla 10.12 Correlaciones entre moral y dirección

Correlaciones			
		Moral	**Dirección**
Moral	Correlación de Pearson	1	0.557**
	Sig. (bilateral)		0.000
	N	362	335
Dirección	Correlación de Pearson	0.557**	1
	Sig. (bilateral)	0.000	
	N	335	373

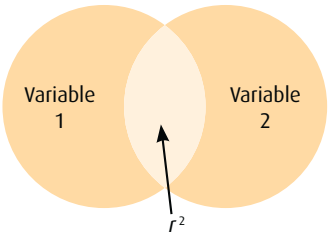
** La correlación es significativa al nivel 0.01 (bilateral, en ambos sentidos entre las variables).

Obsérvese que se correlacionan dos variables: "moral" y "dirección", aunque la correlación aparece dos veces, porque es una tabla que hace todas las comparaciones posibles entre las variables y al hacerlo, genera un eje diagonal (representado por las correlaciones de las variables contra ellas mismas —"moral" con "moral" y "dirección" con "dirección", que carece de sentido porque son las mismas puntuaciones, por eso es perfecta—), y por encima de ese eje aparecen todos los coeficientes, y se repiten por debajo del eje. La correlación es de 0.557 y es significativa en el nivel del 0.000 (menor del 0.01). *N* representa el número de casos correlacionados.

Una correlación de Pearson puede ser significativa, pero si es menor a 0.30 resulta débil, aunque de cualquier manera ayuda a explicar el vínculo entre las variables. Si queremos asociar la *presión arterial* y *el peso* de un grupo de pacientes, la *solubilidad del gas* con la *temperatura* (en ingeniería petrolera) y la *inversión en publicidad* y las *ventas*, es útil este coeficiente.

Consideraciones: cuando el coeficiente *r* de Pearson se eleva al cuadrado (r^2), se obtiene el coeficiente de determinación y el resultado indica la *varianza de factores comunes.* Esto es, el porcentaje de la variación de una variable debido a la variación de la otra variable y viceversa (o cuánto explica o determina una variable la variación de la otra). Veámoslo gráficamente en la figura 10.17.

Figura 10.17 Varianza de factores comunes.



Por ejemplo, si la correlación entre "productividad" y "asistencia al trabajo" es de 0.80.

$$r = 0.80$$
$$r^2 = 0.64$$

"La productividad" constituye a, o explica, 64% de la variación de "la asistencia al trabajo".

"La asistencia al trabajo" explica 64% de "la productividad".

Si *r* es 0.72 y consecuentemente $r^2 = 0.52$, quiere decir que poco más de la mitad de la variabilidad de un constructo o variable está explicada por la otra.

Ejemplo

Hi:	"A mayor motivación intrínseca, mayor productividad".
Resultado:	$r = 0.721$ s o $P = 0.0001$
Interpretación:	se acepta la hipótesis de investigación en el nivel de 0.01. La correlación entre la motivación intrínseca y la productividad es considerable y positiva.
Hi:	"a mayor ingreso, mayor motivación intrínseca".
Resultado:	$r = 0.214$ s o $P = 0.081$
Interpretación:	se acepta la hipótesis nula. El coeficiente no es significativo: 0.081 es mayor que 0.05; recordemos que 0.05 es el nivel mínimo para aceptar la hipótesis.

Nota precautoria: recuerde lo referente a correlaciones espurias que se comentaron en el capítulo 5, "Definición del alcance de la investigación por realizar".

Creswell (2005) señala que un coeficiente de determinación (r^2) entre 0.66 y 0.85 ofrece una buena predicción de una variable respecto de la otra variable; y por encima de 0.85 implica que ambas variables miden casi el mismo concepto subyacente, son "cercanamente" un constructo semejante.

El coeficiente de correlación de Pearson es útil para relaciones lineales, como lo veremos en la regresión lineal, pero no para relaciones curvilineales; en este caso o cuando las variables son ordinales, se suele usar la *rho* de Spearman (ρ) (Onwuegbuzie, Daniel y Leech, 2006b).

Cuando queremos correlacionar simultáneamente más de dos variables, por ejemplo: motivación, satisfacción en el trabajo, moral y autonomía; o como lo hicieron Wood *et al.* (2009) con productividad del médico (medida en unidades de valor relativo McGladrey, MRVU), tiempo médico/paciente, satisfacción y confianza del paciente, se utiliza el coeficiente de correlación múltiple o *R*, el cual se revisa en el capítulo adicional 8 "Análisis estadístico: segunda parte", del centro de recursos.

Centro de recursos en línea

Para el cálculo del coeficiente de correlación de Pearson mediante IBM SPSS®, no olvide consultar el manual respectivo. En Minitab en Estadísticas → Estadísticas básicas.

¿Qué es la regresión lineal?

3

Es un modelo estadístico para estimar el efecto de una variable sobre otra. Está asociado con el coeficiente *r* de Pearson. Brinda la oportunidad de predecir las puntuaciones de una variable a partir de las puntuaciones de la otra variable. Entre mayor sea la correlación entre las variables (covariación), mayor capacidad de predicción.

Hipótesis: correlacionales y causales.

Variables: dos. Una se considera como independiente y otra como dependiente. Pero, para poder hacerlo, debe tenerse un sólido sustento teórico.

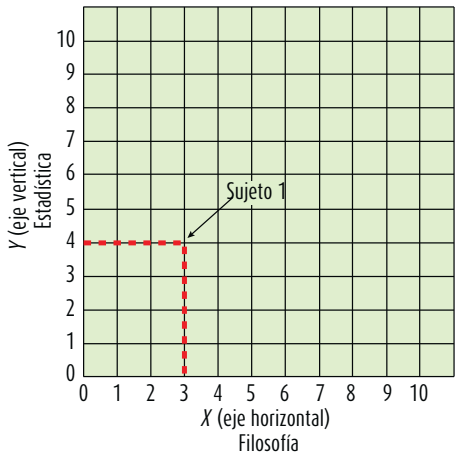
Nivel de medición de las variables: intervalos o razón.

Procedimiento e interpretación: la regresión lineal se determina con base en el diagrama de dispersión. Éste consiste en una gráfica donde se relacionan las puntuaciones de una muestra en dos variables (Martin y Bridgmon, 2012; Bednarczyk y McNutt, 2007; Harrington, 2007; Daniel, Onwuegbuzie y Leech, 2006; y Wood y Park, 2003). Veámoslo con un ejemplo sencillo de ocho casos. Una variable es la calificación en Filosofía y la otra variable es la calificación en Estadística; ambas medidas, hipotéticamente, de 0 a 10.

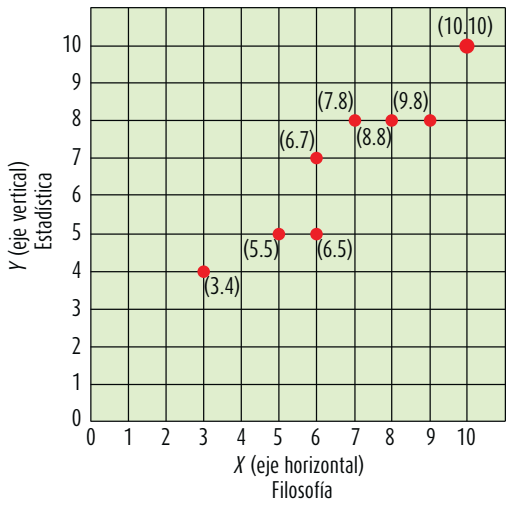
	Puntuaciones	
Sujetos	Filosofía (*X*)	Estadística (*Y*)
1	3	4
2	8	8
3	9	8
4	6	5
5	10	10
6	7	8
7	6	7
8	5	5

Figura 10.18 Ejemplo de gráficas de dispersión.

El *diagrama de dispersión* se construye graficando cada par de puntuaciones en un espacio o plano bidimensional. Sujeto "1" tuvo 3 en *X* (filosofía) y 4 en *Y* (estadística):

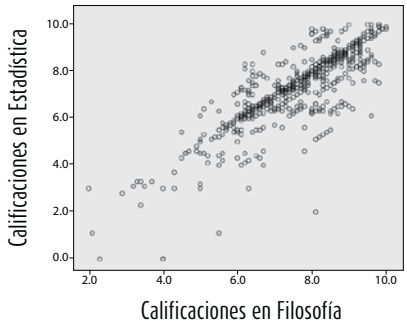


Así se grafican todos los pares:



(*continúa*)

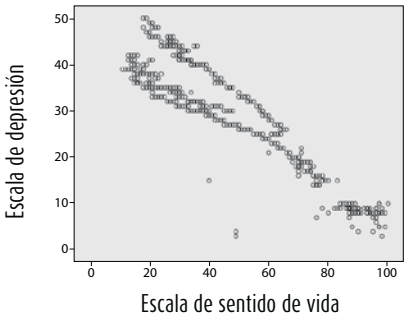
Figura 10.18 (*continuación*)



Los *diagramas de dispersión* son una manera de visualizar gráficamente una correlación. Por ejemplo:

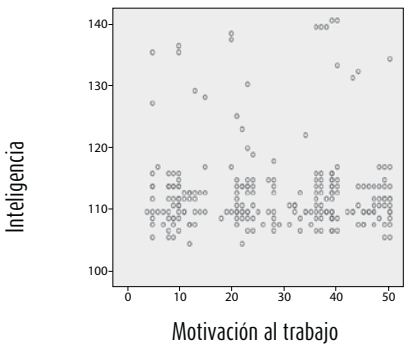
Si aplicáramos los exámenes de Filosofía y Estadística (escala de 0 a 10 en ambas mediciones) a 775 alumnos y obtuviéramos el siguiente resultado: $r = 0.814$** (significativa al nivel del .01). La correlación es considerablemente positiva y el diagrama de dispersión sería el siguiente:[20]

La tendencia es ascendente, altas puntuaciones en *Y*, altas puntuaciones en *X* (mejores calificaciones en Estadística están asociadas con mejores calificaciones en Filosofía).

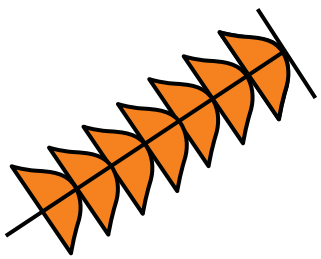


En cambio, si administráramos una prueba sobre la "depresión" (escala de 0 a 50) y una que mida el "sentido de vida" (0 a 100) y el resultado fuera: -0.926** (significativa al nivel del .01). La correlación es sumamente negativa y el diagrama de dispersión sería el siguiente:

La tendencia es descendente, altas puntuaciones en *depresión* se encuentran vinculadas con bajas en *sentido de vida*, y viceversa.



En el caso de que dos variables no estén correlacionadas, por ejemplo: $r = .006$ (no significativa) (digamos entre "inteligencia" −90 a 140− y "motivación al trabajo" −0 a 50−). El diagrama de dispersión no tiene ninguna tendencia:



Así, cada punto representa un caso y un resultado de la intersección de las puntuaciones en ambas variables. El diagrama de dispersión puede ser resumido a una línea, si hay tendencia.

Conociendo la línea y la tendencia, podemos predecir los valores de una variable conociendo los de la otra variable (Shapiro, 2008).

Esta línea es la recta de regresión y se expresa mediante la *ecuación de regresión lineal*:

$$Y = a + bX$$

en donde *Y* es un *valor de la variable dependiente* que se desea predecir, *a* es la *ordenada* en el origen (intersección) y *b* la *pendiente* o inclinación, *X* es el valor que fijamos en la variable independiente o *predictora*.

Los programas de análisis estadístico que incluyen la *regresión lineal*, proporcionan los datos de *a* y *b*.

a o intercepción (*intercept*) y *b* o pendiente (*slope)*

[20] Estos diagramas fueron visualizados a través de SPSS, versión 15.

Para predecir un valor de *Y*, se sustituyen los valores correspondientes en la ecuación.

Ejemplo

a (intercepción) = 1.2

b (pendiente) = 0.8

Entonces podemos hacer la predicción: ¿a un valor de 7 en Filosofía qué valor le corresponde en Estadística?

$$Y = \underset{a}{1.2} + \underset{b}{(0.8)}\,\underset{X}{(7)}$$

$$Y = 6.8$$

Predecimos que a un valor de 7 en *X* le corresponderá un valor de 6.8 en *Y*.

Ejemplo

Regresión lineal

Hi: "La autonomía laboral es una variable que predice la motivación intrínseca en el trabajo. Ambas variables están relacionadas".
Las dos variables fueron medidas en una escala de 1 a 5.

Resultado: a (intercepción) = 0.42
b (pendiente) = 0.65

Interpretación: cuando *X* (autonomía) es 1, la predicción estimada de *Y* es 1.07; cuando *X* es 2, la predicción estimada de *Y* es 1.72; cuando *X* es 3, *Y* será 2.37; cuando *X* es 4, *Y* será 3.02; y cuando *X* es 5, *Y* será 3.67.

$$Y = a + bX$$
$$1.07 = 0.42 + 0.65\ (1)$$
$$1.72 = 0.42 + 0.65\ (2)$$
$$2.37 = 0.42 + 0.65\ (3)$$
$$3.02 = 0.42 + 0.65\ (4)$$
$$3.67 = 0.42 + 0.65\ (5)$$

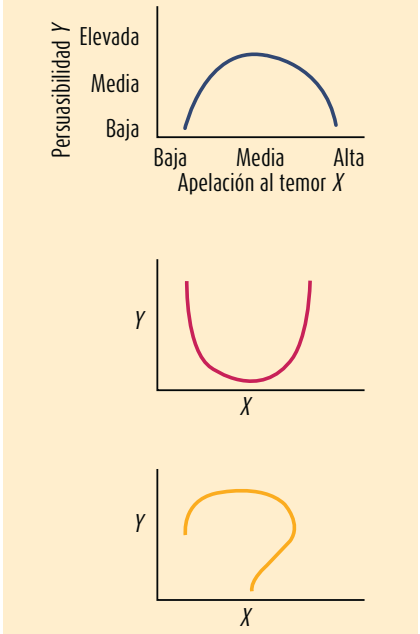
Consideraciones: la regresión lineal es útil con *relaciones lineales,* no con *relaciones curvilineales* (Graham, 2013; Bates y Watts, 2007; y Little, 2003). Porque como señalan León y Montero (2003, p. 191) es un error atribuir a la relación causal una covariación exclusivamente lineal: a mayores valores en la variable independiente, mayores valores en la dependiente. Existen muchas relaciones de causa-efecto que no son lineales, como por ejemplo: la vinculación entre ansiedad y rendimiento. Cierto grado de ansiedad ayuda a conseguir mejores resultados en un examen o la práctica de un deporte; pero, por encima de determinado nivel (nerviosismo extremo), la ejecución empeora. También, la dosis de un medicamento puede tener una relación no lineal con el efecto esperado (cierta dosis no lograrlo por insuficiente y demasiada dosis, provocar otros efectos indeseados con consecuencias muy negativas). Asimismo, determinadas reacciones químicas necesitan una temperatura específica (ni más, ni menos) y lo mismo con la cantidad de riego para una parcela donde se experimenta con cultivos específicos. En la figura 10.19 se muestran ejemplos de estas relaciones.

Las *relaciones curvilineales* son aquellas en las cuales la tendencia varía, por ejemplo: primero es ascendente y luego descendente, o viceversa.

Se ha demostrado que una estrategia persuasiva con niveles altos de apelación al temor, por ejemplo, un comercial televisivo muy dramático, provoca una escasa persuasión, lo mismo que una estrategia persuasiva con niveles muy bajos de apelación al temor.

La estrategia persuasiva más adecuada es la que utiliza niveles medios de apelación al temor. Esta relación es curvilineal; véase figura 10.19.

Figura 10.19 Ejemplos de relaciones curvilineales.



En la práctica, los estudiantes no deben preocuparse por graficar los diagramas de dispersión. Esto lo hace el programa respectivo (SPSS®, Minitab u otro).

¿Qué es la prueba *t*?

Es una prueba estadística para evaluar si dos grupos difieren entre sí de manera significativa respecto a sus medias en una variable.

Se simboliza: *t*.

Hipótesis: de diferencia entre dos grupos. La hipótesis de investigación propone que los grupos difieren entre sí de manera significativa y la hipótesis nula plantea que los grupos no difieren significativamente.

Los grupos pueden ser dos plantas comparadas en su productividad, dos escuelas contrastadas en los resultados a un examen, dos clases de materiales de construcción cotejados en su rendimiento, dos medicamentos comparados en su efecto, etcétera.

Variables: la comparación se realiza sobre una variable (regularmente y de manera teórica: dependiente). Si hay diferentes variables, se efectuarán varias pruebas *t* (una por cada variable), y la razón que motiva la creación de los grupos puede ser una variable independiente. Por ejemplo, un experimento con dos grupos, donde a uno se le aplica el estímulo experimental y al otro no, es de control.

Nivel de medición de la variable de comparación: intervalos o razón.

Cálculo e interpretación: el valor *t* es calculado por el programa estadístico.[21] Los programas, por ejemplo SPSS, arrojan una tabla con varios resultados, de los cuales los más necesarios para interpretar son el valor *t* y su significancia. Veamos un ejemplo y la interpretación.[22]

Ejemplo

Hi: "Los varones le atribuyen mayor importancia al atractivo físico en sus relaciones heterosexuales que las mujeres".

Ho: "Los varones *no* le atribuyen mayor importancia al atractivo físico en sus relaciones heterosexuales que las mujeres".

La variable *atractivo físico* fue medida a través de una escala que varía de 0 a 18. El grupo de mujeres estuvo constituido por 119 personas y el de hombres por 128 (variable que origina el contraste: género). Los resultados fueron:

$\bar{X}_1$ (mujeres) = 12

$\bar{X}_2$ (hombres) = 15

Valor t = 6.698 (significancia menor de 0.01)

n_1 = 119 mujeres

n_2 = 128 hombres

Grados de libertad = 245

Conclusión: se acepta la hipótesis de investigación y se rechaza la hipótesis nula.
Si el valor *t* hubiera sido de 1.05 y no significativo, se aceptaría la hipótesis nula.

La prueba *t* se basa en una distribución muestral o poblacional de diferencia de medias conocida como la distribución *t* de Student que se identifica por los grados de libertad, los cuales constituyen el número de maneras en que los datos pueden variar libremente. Son determinantes, ya que nos indican qué valor debemos esperar de *t*, dependiendo del tamaño de los grupos que se comparan. *Cuanto mayor número de grados de libertad se tengan, la distribución t de Student se acercará más a ser una distribución normal* y usualmente, si los grados de libertad exceden los 120, la distribución nor-

Centro de recursos en línea

[21] Para quienes se interesen en las fórmulas para calcular manualmente el valor de la prueba *t*, se encuentran en el capítulo 8, "Análisis estadístico: segunda parte", en Material complementario → Capítulos.

[22] Se evita aquí la discusión sobre si las percepciones pueden medirse en un nivel de intervalos u ordinal. El ejemplo intenta atraer la atención de los alumnos. Desde luego, si el profesor las considera ordinales, puede cambiar el ejemplo por otro que considere pertinente en su campo.

mal se utiliza como una aproximación adecuada de la distribución *t* de Student (Babbie, 2012, Wiersma y Jurs, 2008; y Godby, 2007).

Los grados de libertad se calculan con la fórmula siguiente, en la que n_1 y n_2 son el tamaño de los grupos que se comparan:

$$gl = (n_1 + n_2) - 2$$

Vogt (1999) señala que los grados de libertad indican cuántos casos fueron usados para calcular un valor estadístico en particular.

Hernández-Sampieri *et al.* (2010) realizaron un análisis mediante la prueba *t* con poco menos de medio millón de alumnos de una institución pública, con la finalidad de comparar el desempeño entre mujeres y hombres respecto al promedio general de la carrera, el valor obtenido fue de 22.802, significancia = 0.000 (menor al 0.01). El promedio de los estudiantes fue de 6.58 ($n = 302\ 272$) y el de las estudiantes de 7.11 ($n = 193\ 436$). Ante la interrogante: ¿se observaron diferencias en el desempeño académico por género? Se puede decir que las mujeres obtienen mayor promedio que los hombres en una diferencia de 0.53 puntos, la cual es significativa al nivel del 0.01.

Así como el caso anterior, esta prueba se utiliza frecuentemente para hacer contrastes por género. Por ejemplo, los efectos de administrar un medicamento (digamos, en la presión arterial u otra variable) o seguir una dieta (en la disminución del peso corporal). Buunk, Castro, Zurriaga y González (2011) llevaron a cabo un estudio en España y Argentina para comparar si los hombres son más celosos que las mujeres ante la presencia de un *rival* en cuatro dimensiones o características: atractivo físico, dominación física, poder social y atributos sociales comunitarios. Resultaron significativos al nivel del 0.01, todos los valores "t"[23] de las dos primeras (77.98 en España y 121.89 en Argentina para el atractivo físico, siendo la media de celos mayor en las mujeres; y 21.67 en España y 42.38 en Argentina para dominación física, siendo la media de celos mayor en los hombres). Esto significa que los hombres experimentaron más celos que las mujeres cuando su rival es físicamente más dominante. En cambio, las mujeres experimentaron más celos que los hombres cuando su rival era más atractiva (a los hombres les preocupa la "musculatura de sus rivales" y a las mujeres "el atractivo físico"). La *n* fue de 388 españoles y 444 argentinos de ambos géneros.

Consideraciones: La prueba *t* se utiliza para comparar los resultados de una preprueba con los resultados de una posprueba en un contexto experimental. Se comparan las medias y las varianzas del grupo en dos momentos diferentes: $\bar{X}_1 \times \bar{X}_2$. O bien, para comparar las prepruebas o pospruebas de dos grupos que participan en un experimento:

G_1 $\bar{X}_1$

G_2 $\bar{X}_2$ $\overset{t}{\bigcirc}$ son las pospruebas

Cuando el valor *t* se calcula mediante un paquete estadístico computacional, la significancia se proporciona como parte de los resultados y debe ser menor a 0.05 o 0.01, lo cual depende del nivel de confianza seleccionado (regularmente se ofrece el resultado en dos versiones, según sea el caso, si se asumen o no varianzas iguales).[24] Lo más importante es visualizar el valor *t* y su significancia; véase la tabla 10.13.

Para solicitar en SPSS la prueba *t*, no olvide consultar el manual "Introducción al IBM SPSS®" que se puede descargar del centro de recursos. En Minitab este método se encuentra en: Estadísticas → Estadísticas básicas. En STATS® se denomina: diferencia de dos medias (Difference-Two Means) y simplemente se colocan número de casos o respuestas en cada grupo, medias y desviaciones estándar de los grupos, y automáticamente se calcula el valor *t* y el nivel de significancia expresado en porcentaje.

Centro de recursos en línea

[23] Ellos lo expresan como valores "F", que prácticamente es lo mismo, implicando también análisis de las varianzas.

[24] Cuando se incluyen participantes diferentes en los grupos del experimento, el diseño se considera de "grupos independientes" (León y Montero, 2003) y no se asumen varianzas iguales.

Tabla 10.13 Elementos fundamentales para interpretar los resultados de una prueba *t*

Estadísticos de grupo							
F3	Género	N	Media	Desviación tip.		Error tip. de la media	
	Masculino	86	3.69	1.043		0.113	
	Femenino	88	3.84	1.071		0.114	

Prueba de muestras independientes										
Prueba de Levene para la igualdad de varianzas				Prueba *t* para la igualdad de medias						
									95% intervalo de confianza para la diferencia	
		F	Sig.	*t*	gl	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	inferior	superior
F3	Se han asumido varianzas iguales	0.001	0.970	−0.966	172	0.335	−0.15	0.160	−0.471	0.162
	No se han asumido varianzas iguales			−0.966	171.98	0.335	−0.15	0.160	−0.471	0.162

Valor "F" diferencia entre las varianzas de los grupos (dispersión de los datos)

Valor "t"

Significancia: no es menor al 0.05, mucho menos al 0.01: No hay diferencias entre los grupos en la variable de contraste

¿Qué es el tamaño del efecto?

2 Al comparar grupos, en este caso con la prueba *t*, es importante determinar el tamaño del efecto, que es una medida de la "fuerza" de la diferencia de las medias u otros valores considerados (Creswell, 2013a; Alhija y Levy, 2009; y Cortina, 2003). Resulta ser una medida en unidades de desviación estándar.

¿Cómo se calcula? El tamaño del efecto es justo la diferencia estandarizada entre las medias de los dos grupos. En otras palabras:

$$\text{Tamaño total del efecto} = \frac{\text{Media del grupo 1} - \text{Media del grupo 2}}{\text{Desviación estándar sopesada}}$$

La desviación estándar sopesada es la estimación reunida de la desviación estándar de ambos grupos, basada en la premisa que cualquier diferencia entre sus desviaciones es solamente debida a la variación del muestreo (Rodríguez, 2006 y Creswell, 2005).

La desviación estándar sopesada (denominador en la fórmula) se calcula así:

$$\sqrt{\frac{(N_E - 1)SD_E^2 + (N_C - 1)SD_C^2}{N_E + N_C - 2}}$$

Donde N_E y N_C son el tamaño de los grupos (grados de libertad), respectivamente; en tanto que, SD_E y SD_C son sus desviaciones estándares.

Ejemplo

17.9 − 15.2/3.3 = 0.82 (interpretación: las medias varían menos de una desviación estándar, una respecto de la otra).

Ejemplo

28.5 − 37.5/4.1 = −2.19 (los promedios varían más de dos desviaciones estándar uno sobre otro).

¿Qué es la prueba de diferencia de proporciones?

Es una prueba estadística para analizar si dos proporciones o porcentajes difieren significativamente entre sí.

Hipótesis: de diferencia de proporciones en dos grupos.

Variable: la comparación se realiza sobre una variable. Si hay varias, se efectuará una prueba de diferencia de proporciones por variable.

Nivel de medición de la variable de comparación: cualquier nivel, incluso por intervalos o razón, pero siempre expresados en proporciones o porcentajes.

Procedimiento e interpretación: este análisis puede realizarse muy fácilmente en el programa STATS®, subprograma: Diferencia de dos proporciones (Difference-Two Percentages). Se colocan el número de casos y el porcentaje obtenido para cada grupo y se calcula. Eso es todo. No se necesita de fórmulas y tablas como se hacía anteriormente.

Ejemplo

Hi: "El porcentaje de liberales en la ciudad de Arualm es mayor que en Linderbuck".

En STATS® colocamos los datos que se nos requiere:

Inputs

Grupo uno

Número de respondientes en grupo uno

410

Porcentaje medido en grupo uno

55

Grupo dos

Número de respondientes en grupo uno

301

Porcentaje medido en grupo uno

48

Calculamos **Calculate** y se obtienen los resultados:

Results

Probabilidad de diferencia significativa

94.52%

Valor Z

1.92

Como no se alcanza una significancia de 95% (porque STATS®, al contrario de SPSS® o Minitab, proporciona el porcentaje de ésta a favor), aceptamos la hipótesis nula y rechazamos la de investigación.

Con esta prueba podemos analizar, por ejemplo, si el porcentaje de mujeres con cáncer de mama es significativamente diferente en dos comunidades, si el porcentaje de errores en la producción de arneses automotrices es significativamente distinto en dos plantas, si el porcentaje de reprobados es significativamente desigual entre los alumnos de bachillerato del turno matutino y del vespertino, etc. Desde luego, no es necesario que los grupos por comparar tengan el mismo número de unidades, casos, respondientes o equivalentes (n), salvo las consideraciones de muestreo hechas previamente (tamaño mínimo de grupos).

¿Qué es el análisis de varianza unidireccional o de un factor? (ANOVA *one-way*)

3 Es una prueba estadística para analizar si más de dos grupos difieren significativamente entre sí en cuanto a sus medias y varianzas. La *prueba t* se aplica para *dos grupos* y el *análisis de varianza unidireccional* se usa para *tres, cuatro o más grupos.* Aunque con dos grupos se puede utilizar también.

Hipótesis: de diferencia entre más de dos grupos. La hipótesis de investigación propone que los grupos difieren significativamente entre sí y la hipótesis nula propone que los grupos no difieren significativamente.

Variables: una variable independiente y una variable dependiente.

Nivel de medición de las variables: la variable independiente es categórica y la dependiente es por intervalos o razón.

El hecho de que la variable independiente sea categórica significa que es posible formar grupos diferentes (Martin y Bridgmon, 2012 y Lazar, 2006). Puede ser una variable nominal, ordinal, por intervalos o de razón (pero en estos últimos dos casos la variable debe reducirse a categorías).

Por ejemplo:[25]

- Religión (católica, cristiana, protestante, judía, musulmana, budista, etc.) (puede compararse la satisfacción de los grupos con su religión o el grado de espiritualidad: Soto, 2014).
- Nivel socioeconómico (muy alto, alto, medio, bajo y muy bajo) (contrastarse su lealtad a la marca).
- Antigüedad del empleado en la empresa (de cero a un año, más de un año a cinco años, más de cinco años a 10, más de 10 años a 20 y más de 20 años) (cotejarse su productividad).
- Estadios del cáncer de próstata (I, II, III y IV) (comparar su grado de depresión).
- Obesidad y peso: peso insuficiente, normopeso, sobrepeso, obesidad en grados (I, II, III y IV —extrema—) (cotejar sus niveles de glucosa y presión arterial).
- Giro de la empresa: comercial, industrial y de servicios (comparar los efectos de una medida fiscal en su nivel de tributación).
- Tipo de concreto premezclado (estándar, de fraguado rápido, reforzado con fibras, autocompactante, poroso, antibacteriano, etc.) (contrastar su resistencia).

Análisis de varianza Prueba estadística para analizar si más de dos grupos difieren entre sí de manera significativa en sus medias y varianzas.

Interpretación: el **análisis de varianza** unidireccional produce un valor conocido como *F* o *razón F*, que se basa en una distribución muestral, conocida como *distribución F*, la cual es otro miembro de la familia de distribuciones muestrales. La *razón F* compara las variaciones en las puntuaciones debidas a dos diferentes fuentes: variaciones entre los grupos que se comparan y variaciones dentro de los grupos. Si el valor *F* es significativo implica que los grupos difieren entre sí en sus promedios (Zhang, 2013; The SAGE Glossary of the Social and Behavioral Sciences, 2009h;, Klugkist, 2008; Field, 2006a y 2006b; y Norpoth, 2003). Entonces se acepta la hipótesis de investigación y se rechaza la nula.[26] A continuación se presenta un ejemplo de un estudio en el que el análisis apropiado es el de varianza.

Ejemplo

Hi: "Los niños que se expongan a contenidos de elevada violencia televisiva exhibirán una conducta más agresiva en sus juegos, respecto de los niños que se expongan a contenidos de mediana o baja violencia televisada".

Ho: "Los niños que se expongan a contenidos de elevada violencia televisiva no exhibirán una conducta más agresiva en sus juegos, respecto de los niños que se expongan a contenidos de mediana o baja violencia televisada".

[25] Ejemplos sencillos, simplemente para que el lector que comienza con estos temas tenga una idea de las múltiples aplicaciones del ANOVA. Su profesor puede proporcionar diversas aplicaciones a su área.

[26] El sustento y explicación del análisis de varianza unidireccional que en las primeras cuatro ediciones se incluía en esta parte, ahora la puede encontrar el lector en el centro de recursos: Material Complementario → Capítulos → Capítulo 8, "Análisis estadístico: segunda parte". Le recomiendo descargar y revisar este capítulo.

Centro de recursos en línea

La variable independiente es el grado de exposición a la violencia televisada y la variable dependiente es la agresividad exhibida en los juegos, medida por el número de conductas agresivas observadas (nivel de medición por intervalos).

Para probar la hipótesis se diseña un experimento con cuatro grupos:

G_1X_1 (elevada violencia)
G_2X_2 (mediana violencia)
G_3X_3 (baja violencia)
G_4 – (conducta prosocial)

Número de actos agresivos

En cada grupo hay 25 niños.

La razón *F* fue de 9.89 y resultó significativa en el nivel de 0.05: se acepta la hipótesis de investigación. La diferencia entre las medias de los grupos es admitida, el contenido sumamente violento tiene un efecto sobre la conducta agresiva de los niños en sus juegos. El estímulo experimental tuvo un impacto. Esto se corrobora comparando las medias de las pospruebas de los cuatro grupos, porque el análisis de varianza unidireccional sólo nos señala si la diferencia entre las medias y las distribuciones de los grupos es o no significativa; pero no nos indica en favor de qué grupos lo es. Es posible hacer esto último al visualizar los promedios y compararlos con las distribuciones de sus grupos. Y si adicionalmente queremos cotejar cada par de medias ($\bar{X}_1$ con $\bar{X}_2$, $\bar{X}_1$ con $\bar{X}_3$, $\bar{X}_2$ con $\bar{X}_3$, etc.) y determinar con exactitud dónde están las diferencias significativas, podemos aplicar un contraste posterior, con el cálculo de una prueba *t* para cada par de medias; o bien, por medio de algunas estadísticas que suelen ser parte de los análisis efectuados en los paquetes estadísticos computacionales.

Tales estadísticas se incluyen en la tabla 10.14.

Tabla 10.14 Principales estadísticas para comparaciones posteriores (*post hoc*) en el ANOVA unidireccional o de un factor[27]

Nombre	Siglas
• Diferencia menos significativa	DMS
• Prueba *F* de Ryan-Einot-Gabriel-Welsch	R-E-G-W F
• Prueba de rango de Ryan-Einot-Gabriel-Welsch	R-E-G-W Q
• Prueba de Tukey	
• Otras: Waller-Duncan, T2 de Tamhane, T3 de Dunnett, Games-Howell, *C* de Dunett, Bonferroni, Sidak, Gabriel, Hochberg, Scheffé...	

Ejemplo

Supongamos que por medio de una escala de Likert (1-5)[28] medimos la actitud que tienen hacia el entrenador del equipo de fútbol de una ciudad, las tres porras o grupos de aficionados permanentes: la Ultra, la Central y la de Veteranos. Y queremos analizar si difieren significativamente entre sí. Realizamos el análisis de varianza y los resultados son los que se muestran en la tabla 10.15 con los elementos que suelen incluir los programas de análisis estadístico como SPSS, nada más que éstos abrevian términos.

Tabla 10.15 Ejemplo de análisis de varianza

ANOVA Actitud hacia el entrenador del equipo de fútbol					
Fuente de variación	**Suma de cuadrados**	**Grados de libertad**	**Medias cuadráticas**	**Valor *F***	**Significancia**
Intergrupos	46 768	2	23 384	17.394	0.000
Intragrupos	793 175	590	1 344		
Total	839 943	592			

[27] Algunas pruebas son para cuando se asumen varianzas iguales y otras no, el programa indica cuáles se utilizan en cada caso.

[28] Una vez más, se elude a propósito la polémica de si la escala de Likert es de intervalos u ordinal. Si el profesor la considera ordinal, puede cambiar el ejemplo por el número de veces que han expresado públicamente su apoyo al entrenador o utilizar otro que considere pertinente en su área.

Descriptivos								
Actitud hacia el entrenador del equipo de fútbol								
					Intervalo de confianza para la media a 95%			
	N	Media	Desviación típica	Error típico	Límite inferior	Límite superior	Mínimo	Máximo
Porra Ultra	195	3.61	1.046	0.075	3.46	3.76	1	5
Porra Central	208	3.72	1.090	0.076	3.57	3.87	1	5
Porra Veteranos	190	3.07	1.331	0.097	3.88	3.26	1	5
Total	593	3.48	1.191	0.049	3.38	3.57	1	5

Comentario: la actitud de las diferentes porras hacia el entrenador es significativamente distinta, la más desfavorable es la de los veteranos (su media es de 3.07, cerrando o redondeando a décimas: 3.1).

En el ejemplo tratado en capítulos previos, cuya hipótesis es: "el consumo diario y permanente de selenio como suplemento alimenticio reduce el crecimiento de los tumores cancerígenos en mujeres que se encuentran en la etapa inicial de la enfermedad" y teniendo los tres grupos experimentales: 1) participantes a las que se les suministra un complemento alimenticio de selenio en cápsulas de 200 mg diarios, 2) participantes a las que se les administra un complemento alimenticio de selenio en cápsulas de 100 mg diarios y 3) participantes a las que no se les suministra selenio (grupo de control), se podría hacer al final del periodo experimental un *análisis de varianza* para comparar las tasas de crecimiento de los tumores entre los grupos, así como una estimación de máxima probabilidad.

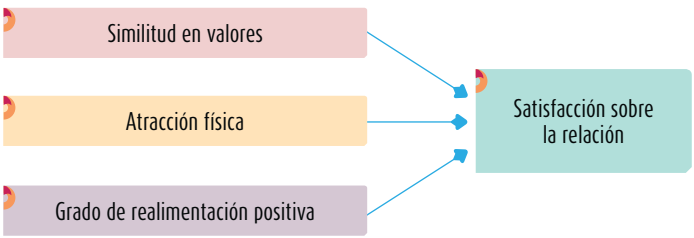
Asimismo, Lee y Guerin (2009) en su estudio para identificar si la satisfacción de la calidad del diseño ambiental del interior de áreas de trabajo u oficinas afecta significativamente la satisfacción general del espacio de trabajo por parte de sus ocupantes y su desempeño laboral, podría efectuarse un análisis de varianza por grupo de edad [30 o menos (1), 31-40 (2), 41-50 (3) y más de 50 (4)] para evaluar si difieren en cuanto a la satisfacción general sobre el espacio de trabajo.

Estadística multivariada

3 y 4 Hasta aquí hemos visto pruebas paramétricas con una sola variable independiente y una dependiente. ¿Pero qué ocurre cuando tenemos diversas variables independientes y una dependiente, varias independientes y dependientes? Se forman esquemas del tipo que se muestra en la figura 10.20.

Figura 10.20 Ejemplos de esquemas con diversas variables tanto dependientes como independientes.

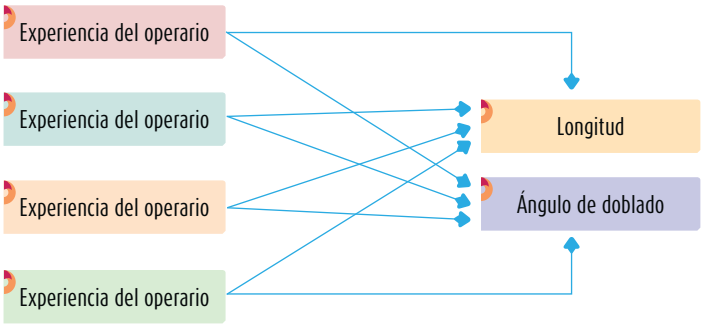
Si queremos probar la hipótesis: "la *similitud en valores*, la *atracción física* y el *grado de realimentación positiva* son factores que inciden en la *satisfacción sobre la relación* en parejas de novios cuyas edades oscilan entre los 24 y los 32 años".



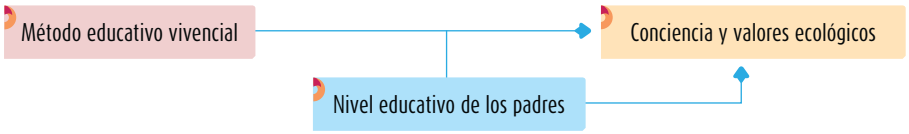
O en el estudio de Pérez, Arango y Agudelo (2009) para determinar el efecto que tienen los factores *experiencia del operario, tipo de dobladora, clase de material* utilizado y su *grosor* sobre la *longitud* y el *ángulo* de *doblado* de las piezas de metal producidas.

(*continúa*)

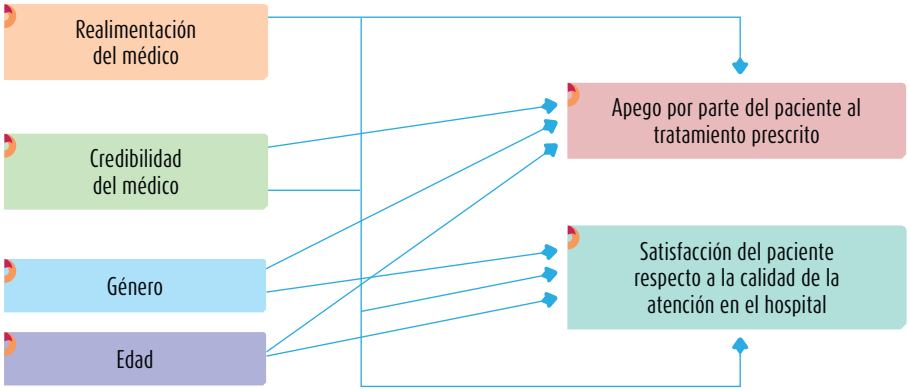
Figura 10.20 (*continuación*)



Asimismo, si pretendemos evaluar si un *método educativo* incrementa la *conciencia* y *valores ecológicos* de los estudiantes de bachillerato, controlando y analizando la influencia de la variable *nivel educativo de los padres*.



Si buscamos conocer la influencia de cuatro variables de los médicos sobre el apego al tratamiento y la satisfacción en torno a la atención por parte de sus pacientes.



Entonces, requerimos otros métodos estadísticos como los que se muestran en la tabla 10.16. Estos métodos se comentan en el capítulo 8 del centro de recursos, "Análisis estadístico: segunda parte", en análisis multivariado (Material Complementario → Capítulos → Capítulo 8).

Centro de recursos en línea

Tabla 10.16 Métodos estadísticos multivariados (se amplía en el capítulo 8 del centro de recursos en línea)

Método	Propósitos fundamentales
Análisis de varianza factorial (ANOVA de varios factores)	Evaluar el efecto de dos o más variables independientes sobre una variable dependiente.
Análisis de covarianza (ANCOVA)	Analizar la relación entre una variable dependiente y dos o más independientes, al eliminar y controlar el efecto de al menos una de estas variables independientes.
Regresión múltiple	Evaluar el efecto de dos o más variables independientes sobre una variable dependiente, así como predecir el valor de la variable dependiente con una o más variables independientes, y estimar cuál es la independiente que mejor predice las puntuaciones de la dependiente. Se trata de una extensión de la regresión lineal.

(*continúa*)

Tabla 10.16 (*continuación*)

Método	Propósitos fundamentales
Análisis multivariado de varianza (MANOVA)	Analizar la relación entre dos o más variables independientes y dos o más variables dependientes.
Análisis lineal de patrones (PATH)	Determinar y representar interrelaciones entre variables a partir de regresiones, así como analizar la magnitud de la influencia de algunas variables sobre otras, influencia directa e indirecta. Es un modelo causal.
Análisis discriminante	Construir un modelo predictivo para pronosticar el grupo de pertenencia de un caso a partir de las características observadas de cada caso (predecir la pertenencia de un caso a una de las categorías de la variable dependiente, sobre la base de dos o más independientes).
Distancias euclidianas	Evaluar la similitud entre variables (en unidades de correlación).

Análisis no paramétricos

3 y 4

Para realizar los análisis no paramétricos debe partirse de las siguientes consideraciones:[29]

1. La mayoría de estos análisis no requieren de presupuestos acerca de la forma de la distribución poblacional. Aceptan distribuciones no normales (distribuciones "libres").
2. Las variables no necesariamente tienen que estar medidas en un nivel por intervalos o de razón; pueden analizar datos nominales u ordinales. De hecho, si se quieren aplicar análisis no paramétricos a datos por intervalos o razón, éstos necesitan resumirse a categorías discretas (a unas cuantas). Las variables deben ser categóricas.

¿Cuáles son los métodos o las pruebas estadísticas no paramétricas más utilizados?

Las *pruebas no paramétricas más utilizadas* son:[30]

1. La ***chi cuadrada*** o χ^2.
2. Los coeficientes de correlación e independencia para tabulaciones cruzadas.
3. Los coeficientes de correlación por rangos ordenados de Spearman y Kendall.

Chi cuadrada Prueba estadística para evaluar hipótesis acerca de la relación entre dos variables categóricas.

¿Qué es la *Chi* cuadrada o χ^2?

Es una prueba estadística para evaluar hipótesis acerca de la relación entre dos variables categóricas.

Se simboliza: χ^2.

Hipótesis por probar: correlacionales.

Variables involucradas: dos. La prueba *Chi* cuadrada no considera relaciones causales.

Nivel de medición de las variables: nominal u ordinal (o intervalos o razón reducidos a ordinales).

Procedimiento: se calcula por medio de una *tabla de contingencia o tabulación cruzada*, que es un cuadro de dos dimensiones y cada dimensión contiene una variable. A su vez, cada variable se subdivide en dos o más categorías.

Un ejemplo de una tabla de contingencia se presenta en la tabla 10.17.

Tabla 10.17 Ejemplo de una tabla de contingencia

		Género		Total
		Masculino	Femenino	
Intención del voto	Candidata A Guadalupe Torres	40	58	98
	Candidata B Liz Almanza	32	130	162
Total		72	188	260

[29] Patrangenaru y Ellingson (2013), Chaubey (2013), Wiersma y Jurs (2008), Pett (2007), Sroka (2006) y Black (2003).
[30] Hollander, Wolfe y Chicken (2013); Howell (2011); Pett (2007), Sroka (2006); Chen y Popovich (2002); y Gibbons (1992).

Esta tabla demuestra el concepto de *tabla de contingencia* o *tabulación cruzada*. Las variables aparecen señaladas a los lados del cuadro (*intención del voto* y *género*), cada una con sus dos categorías. Se dice que se trata de una tabla 2 × 2, donde cada dígito significa una variable y el valor de éste indica el número de categorías de la variable:

2	×	2
↑		↑
Una variable con dos categorías		Otra variable con dos categorías

Un ejemplo de una tabla de contingencia 2 × 3 se muestra en la tabla 10.18. Las dos variables son: *tipo de tratamiento* contra el cáncer gástrico (tres categorías) y resultado final (dos categorías: sobrevivió o falleció). Los números que aparecen en las celdas son frecuencias. Por ejemplo: 20 pacientes que recibieron el tratamiento 1 sobrevivieron. Es necesario hacer notar que no importa cuál variable esté en la parte superior o a la izquierda, porque al final lo fundamental es que todas las categorías de una variable se crucen con todas las categorías de la otra.

Tabla 10.18 Ejemplo de una tabla de contingencia 3 × 2

		Resultado final		Total
		Sobrevivió	Falleció	
Tipo de tratamiento para el cáncer gástrico	Tratamiento 1	20	2	22
	Tratamiento 2	11	8	19
	Tratamiento 3	5	13	18
	Total	36	23	59

Otro ejemplo de relación que podrían analizarse con la *Chi* cuadrada es: *máquina utilizada* en la fabricación de tornillos (cuatro categorías: máquina 1, máquina 2, máquina 3 y máquina 4) y *calidad de la pieza* (dos categorías: defectuosa o sin defectos), para analizar diferencias por máquina.

En esencia, la *Chi* cuadrada es una *comparación* entre la *tabla de frecuencias observadas* y la denominada *tabla de frecuencias esperadas*, la cual constituye la tabla que esperaríamos encontrar si las variables fueran estadísticamente independientes o no estuvieran relacionadas (Wright, 1979). Es una prueba que parte del supuesto de "no relación entre variables" (hipótesis nula) y el investigador evalúa si en su caso esto es cierto o no, analiza si las frecuencias observadas son diferentes de lo que pudiera esperarse en caso de ausencia de correlación (Howell, 2011; Trobia, 2008; Bond, 2007b; Lane, 2006; y Platt, 2003b). La lógica es así: "si no hay relación entre las variables, debe tenerse una tabla así (la de las frecuencias esperadas). Si hay relación, la tabla que obtengamos como resultado en nuestra investigación tiene que ser muy diferente respecto de la tabla de frecuencias esperadas". Es decir, lo que se contó en comparación con lo que se esperaría del azar.

3

La *Chi* cuadrada se puede obtener a través de los programas estadísticos o mediante STATS®.

En SPSS el programa produce un resumen de los casos válidos y perdidos para cada variable (N y porcentaje) y una tabla de contingencia sencilla, como la 10.17, o bien una tabla más compleja con diversos resultados por celda (frecuencias contadas u observadas, frecuencias esperadas, porcentajes respecto a marginales y totales, etc.). Este segundo caso se muestra en el capítulo 8, "Análisis estadístico: segunda parte" (al final), en la parte sobre Chi cuadrada, que como se ha recomendado puede descargarse del centro de recursos.

El programa también proporciona el valor de *Chi* cuadrada junto con otras pruebas (pero recomendamos centrarse en el resultado de ésta), como se muestra en el siguiente ejemplo que corresponde a la tabla 10.17:

	Valor	Grados de libertad (gl)	Significancia
Chi cuadrada de Pearson	13.529	1	0.000

En este caso, el valor de *Chi* cuadrada es significativo al nivel del 0.01, es decir, se acepta la hipótesis de investigación de que existe relación entre las variables *intención del voto* y *género* (Liz Almanza gana, pero sobre todo por el voto femenino).

Algunas veces, solamente se utiliza el valor de *Chi* cuadrada simplemente para ver si hay o no relación entre las variables.

Ejemplo

Hi: "Los tres canales de televisión a nivel nacional difieren en la cantidad de programas de acción social, neutrales y antisociales que difunden". "Hay relación entre la variable *canal de televisión nacional* y la variable *emisión de programas* de acción social*, neutrales y antisociales*".

Resultados:

$$\chi^2 = 7.95$$
$$gl = 4$$

Significancia mayor de 0.05.
Conclusión: se rechaza la hipótesis de investigación y se acepta la nula. No hay relación entre las variables.

Centro de recursos en línea

El lector podrá encontrar el cálculo de la Chi cuadrada en STATS® al final del capítulo 8, "Análisis estadístico: segunda parte", del centro de recursos.

¿Qué son los coeficientes de correlación e independencia para tabulaciones cruzadas?

Además de la *Chi* cuadrada, hay *otros coeficientes* para evaluar si las variables incluidas en la tabla de contingencia o tabulación cruzada están correlacionadas. En la tabla 10.19 se describen los coeficientes más importantes para tal finalidad.[31]

Tabla 10.19 Principales coeficientes para tablas de contingencia

Coeficiente	Para cuadros de contingencia...	Nivel de medición de ambas variables	Interpretación
Phi (φ)	2 × 2	Nominal. Puede utilizarse con variables ordinales reducidas a dos categorías. SPSS lo muestra en cálculos para datos nominales.	En tablas 2 × 2 varía de 0 a 1, donde cero implica ausencia de correlación entre las variables; y uno, que hay correlación perfecta entre las variables. En tablas más grandes, *phi* puede ser mayor de 1.0, pero la interpretación es compleja. Por ello, se recomienda limitar su uso a las tablas 2 × 2.
Coeficiente de contingencia *C* de Pearson	Cualquier tamaño. De hecho es un ajuste de *phi* para tablas con más de dos categorías en las variables. Incluso funciona mejor con tablas de 5 × 5.	Nominal. Puede utilizarse con variables ordinales reducidas a dos categorías.	0 a 1, pero en tablas menores a 5 × 5, se acerca pero nunca alcanza el uno.
V de Cramer (*C*)	Cualquier tamaño.	Cualquier nivel de variables, pero siempre reducidas a categorías. SPSS lo muestra en cálculos para datos nominales.	0 a 1, pero el uno solamente se alcanza si ambas variables tienen el mismo número de categorías (o marginales).

(*continúa*)

[31] En SPSS encontrará otros que se incluyen por nivel de medición: Kappa, McNemar, Cochran, Riesgo, etcétera.

Tabla 10.19 (*continuación*)

Coeficiente	Para cuadros de contingencia...	Nivel de medición de ambas variables	Interpretación
Goodman-Kruskal *Lambda* o sólo *Lambda* (λ)	Cualquier tamaño.	Cualquier nivel de variables, pero siempre reducidas a categorías. SPSS lo muestra en cálculos para datos nominales.	Fluctúa entre 0 y 1, asume causalidad, lo que significa que puede predecirse a la variable dependiente definida en la tabla, sobre la base de la independiente. La versión usual de *Lambda* es asimétrica. Sin embargo, SPSS y otros programas presentan tres versiones: una simétrica y dos asimétricas (estas últimas representan a cada una de las variables considerada como dependiente). La versión simétrica es simplemente el promedio de las dos *Lambdas* asimétricas. Una prueba asimétrica presupone que el investigador puede designar cuál es la variable independiente y cuál la dependiente. En una simétrica no se asume tal causalidad.
Coeficiente de incertidumbre o entropía o *U* de Theil	Cualquier tamaño.	Cualquier nivel de variables, pero siempre reducidas a categorías. SPSS lo muestra en cálculos para datos nominales.	Fluctúa entre 0 y 1, asume causalidad, lo que significa que puede predecirse a la variable dependiente definida en la tabla, sobre la base de la independiente. Por razones históricas (de costumbre), el coeficiente se ha computado frecuentemente en términos de predecir la variable de las columnas, sobre la base de la variable de las filas.
Gamma de Goodman y Kruskal	Cualquier tamaño.	Ordinal.	Varía de −1 a +1 (−1 es una relación negativa perfecta, y +1 una relación positiva perfecta).
Tau-a, Tau-b y *Tau-c* (τa, τb, τc)	Cualquier tamaño.	Ordinal.	Varían de −1 a +1. *Tau-a* y *Tau-b* son asimétricas y *Tau-c* es simétrica.
D de Somers	Cualquier tamaño.	Ordinal.	Varía de −1 a +1.
Kappa	Cualquier tamaño.	Datos categorizados por intervalo.	Regularmente de 0 a 1.

Para profundizar en las fórmulas y premisas de estos análisis, véase Chaubey (2013), Hollander *et al.* (2013), Black (2003), Gibbons (1992) y Nie, Hull, Jenkins, Steinbrenner y Bent (1975).

¿Qué otra aplicación tienen las tablas de contingencia?

Las tablas de contingencia, además de servir para el cálculo de *Chi* cuadrada y otros coeficientes, son útiles para describir conjuntamente dos o más variables. Esto se efectúa al convertir las frecuencias observadas en frecuencias relativas o porcentajes. En una tabulación cruzada puede haber tres tipos de porcentajes respecto de cada celda.

- *Porcentaje en relación con el total de frecuencias observadas* ("*N*" *o* "*n*" *de muestra*).
- *Porcentaje en relación con el total marginal de la columna.*
- *Porcentaje en relación con el total marginal del renglón.*

Veamos con un ejemplo hipotético de una tabla 2 × 2 con las variables género y preferencia por un conductor televisivo. Las frecuencias observadas serían:

Ejemplo

		Género		
		Masculino	Femenino	
Preferencia por el conductor	*A*	25	25	50
	B	40	10	50
		65	35	100

Las celdas podrían representarse así:

a	*c*
b	*d*

Tomemos el caso de *a* (celda superior izquierda). La celda *a* (25 frecuencias observadas) con respecto al total ($N = 100$) representa 25%. En relación con el total marginal de columna (cuyo total es 65) representa 38.46% y respecto del total marginal de renglón (cuyo total es 50) significa 50%. Esto puede expresarse así:

Ejemplo

	Frecuencias observadas		
	25		
En relación con *N*	25.00%		
En relación con "*a* + *b*"	38.46%	*c*	*a* + *c* = 50
En relación con "*a* + *c*"	50.00%		
	b	*d*	*b* + *d*
	a + *b* = 65	*c* + *d*	100 = *N*

Así procedemos con cada categoría, como ocurre en la tabla 10.20.

Tabla 10.20 Ejemplo de una tabla de contingencia para describir conjuntamente dos variables

		Género		
		Masculino	Femenino	
Preferencia por el conductor	*A*	25	25	50
		25.0%	25.0%	
		38.5%	71.4%	
		50.0%	50.0%	
	B	40	10	50
		40.0%	10.0%	
		61.5%	28.6%	
		80.0%	20.0%	
		65	35	100

Algunos programas ubican los porcentajes incluidos en las celdas en otro orden. Por ejemplo, el porcentaje respecto al total lo colocan al final, pero las interpretaciones son similares.

Otros coeficientes de correlación

El coeficiente de correlación de Pearson es una estadística apropiada para variables medidas por intervalos o razón y para relaciones lineales. La *Chi* cuadrada y demás coeficientes mencionados son estadísticas adecuadas para tablas de contingencia con variables nominales, ordinales y de intervalos, pero reducidas a categorías; ahora, ¿qué ocurre si las variables de nuestro estudio son ordinales, por intervalos y de razón? O bien, una mezcla de niveles de medición, o los datos no necesariamente los disponemos en una tabla de contingencia. Existen otros coeficientes que comentaremos brevemente.

¿Qué son los coeficientes y la correlación por rangos ordenados de Spearman y Kendall?

Coeficientes *rho* de Spearman y *tau* de Kendall Son medidas de correlación para variables en un nivel de medición ordinal; los individuos o unidades de la muestra pueden ordenarse por rangos.

Los **coeficientes rho de Spearman**, simbolizado como *rs*, y **tau de Kendall**, simbolizado como *t*, son medidas de correlación para variables en un nivel de medición ordinal (ambas), de tal modo que los individuos, casos o unidades de análisis de la muestra pueden ordenarse por rangos (jerarquías). Son coeficientes utilizados para relacionar

estadísticamente escalas tipo Likert por aquellos investigadores que las consideran ordinales. Por ejemplo, supongamos que tenemos para refrescos embotellados o sodas las variables "preferencia en el sabor" y "atractivo del envase", y queremos asociarlas estadísticamente, entonces pedimos a un grupo de personas representativas del mercado que evalúen conjuntamente 10 marcas específicas y las ordenen del 1 al 10; en tanto que, "1" es la categoría o el rango máximo en ambas variables. Finalmente se obtienen los siguientes resultados en la muestra:

Marca[32]	Variable 1 Preferencia en el sabor	Variable 2 Atractivo del envase
Loy	1	2
Wiz Cola	2	5
Fan	3	1
Energizador	4	3
Maron	5	4
Manzanol	6	6
Cold	7	8
Zoda II	8	7
Frutol	9	10
Sabrosol	10	9

Para analizar tales resultados, utilizaríamos los coeficientes *rs* y *t*. Ahora bien, debe observarse que todos los refrescos o sodas tienen que jerarquizarse por rangos que contienen las propiedades de una escala ordinal (se ordena de mayor a menor). Ambos coeficientes varían de –1.0 (correlación negativa perfecta) a +1.0 (correlación positiva perfecta), considerando el 0 como ausencia de correlación entre las variables jerarquizadas. Se trata de estadísticas sumamente eficientes para datos ordinales (Howell, 2011; Khamis, 2008; Abdi, 2006; y Kraemer, 2006). La diferencia entre ellos es explicada por Nie *et al.* (1975, p. 289) de la manera siguiente: el coeficiente de Kendall (*t*) resulta un poco más significativo cuando los datos contienen un número considerable de rangos empatados. El coeficiente de Spearman *rho* parece ser una aproximación cercana al coeficiente *r* de Pearson, cuando los datos son continuos (por ejemplo, no caracterizados por un número considerable de empates en cada rango). De acuerdo con Breen y Luijkx (2010) y Creswell (2005) sirve también para analizar relaciones curvilineales.

También se interpreta su significancia igual que Pearson y otros valores estadísticos.

Otros ejemplos serían: relacionar la opinión de dos médicos respecto a la jerarquización de un grupo de pacientes en cuanto al avance de una enfermedad terminal, a fin de decidir cuáles requieren atención más urgente, o bien, dos epidemiólogos que hicieran una evaluación ordinal de las siete amenazas en salud pública para una comunidad y si hay una asociación significativa entre los dos conjuntos de filas, los funcionarios de salud se sentirán más seguros del diagnóstico de vulnerabilidad y las acciones que deben adoptar (Haug, 2007). Onwuegbuzie *et al.* (2006b) correlacionaron mediante el coeficiente *rho* el porcentaje de juegos ganados con el número de puntos anotados en una temporada para jerarquizar el desempeño de los equipos de La Liga Nacional de Fútbol Americano (NFL). Sülar y Okur (2008) relacionaron evaluaciones subjetivas hechas por 18 expertos a telas sobre los atributos de suavidad, grosor y rugosidad con mediciones objetivas de la resistencia a la tensión, flexión, corte, compresión y las propiedades de la superficie (usando diversos aparatos, entre ellos, una boquilla de pruebas), calculando, entre otras estadísticas, coeficientes de concordancia de Kendall, con el propósito de generar modelos de producción.

[32] Nombres ficticios.

¿Qué otros coeficientes hay?

Un coeficiente muy importante es el *Eta*, que es similar al coeficiente *r de Pearson*, pero con relaciones no lineales, las cuales se comentaron anteriormente. Es decir, *Eta* define la "correlación perfecta" (1.00) como curvilineal y a la "relación nula" (0.0) como la independencia estadística de las variables. Este coeficiente es asimétrico (concepto explicado en la tabla 10.19, cuando se revisa Lambda), y a diferencia de Pearson, se puede obtener un valor diferente para el coeficiente al determinar cuál variable se considera independiente y cuál dependiente. Eta^2 es interpretada como el porcentaje de la varianza en la variable dependiente explicado por la independiente. El investigador puede calcular *Eta* de las dos maneras: al cambiar la definición de la independiente y dependiente, luego promediar los dos coeficientes y obtener uno simétrico. *Eta* puede trabajarse en tablas de contingencia. Otros coeficientes se describen en la tabla 10.21.

Tabla 10.21 Otros coeficientes de correlación

Coeficiente	Nivel de medición de las variables	Ejemplos	Interpretación
Biserial (r_b)	Una ordinal y la otra por intervalos o razón.	Jerarquía en la organización y motivación. Edad y nivel de depresión	−1.00 (correlación negativa perfecta). 0.0 (ausencia de relación). +1.00 (correlación positiva perfecta).
Biserial por rangos (r_{rb})	Una variable nominal y la otra ordinal.	Escuela de procedencia (pública-privada) y rango en una prueba de un idioma extranjero (alto, medio, bajo). Género y jerarquía laboral.	−1.00 (correlación negativa perfecta). 0.0 (ausencia de relación). +1.00 (correlación positiva perfecta).
Biserial puntual (r_{pb})	Una variable por inter-valos o razón y la otra nominal.	Motivación al estudio y licenciatura (Economía, Derecho, Administración, etcétera). Número de cigarrillos fumados diariamente y desarrollo de cáncer pulmonar (presencia-ausencia de la enfermedad).	−1.00 (correlación negativa perfecta). 0.0 (ausencia de relación). +1.00 (correlación positiva perfecta).
Tetracórico o *Tetrachoric*	Las dos dicotómicas, no necesariamente expresadas en tablas. Es utilizado sobre todo cuando las variables son de intervalo o razón y han sido dicotomizadas.	Género y afiliación/no afiliación a un partido político. Decisión de abortar y creencia-no creencia en un ser supremo.	−1.00 (correlación negativa perfecta). 0.0 (ausencia de relación). +1.00 (correlación positiva perfecta).

Existen muchos más coeficientes, pero tal vez los más importantes son los señalados. Lo mejor de todo es que los programas computacionales de análisis estadístico los calculan, lo único que tenemos que hacer es interpretarlos y verbalizar sus resultados con comentarios.

Una vista general a los procedimientos o pruebas estadísticas

3 Ahora, presentamos un par de tablas finales (10.22 y 10.23) sobre los principales métodos estadísticos. En la primera se considera: *a*) *el tipo de pregunta de investigación* (descriptiva, de diferencia de grupos, correlacional o causal), *b*) *el número de variables involucradas*, *c*) *nivel de medición de las variables* o *tipo de datos* y *d*) en comparación de grupos, si son *muestras independientes* o *correlacionadas*. En este último punto, las muestras independientes se seleccionan de manera que no exista ninguna relación entre los casos de las muestras; por ejemplo, un grupo experimental y uno de control en un experimento. No hay ningún emparejamiento de las observaciones entre las muestras. Mientras que en las correlacionadas sí existe una relación entre las unidades o participantes de las muestras; por ejemplo, el mismo grupo antes y después de un tratamiento experimental, preprueba y posprueba. La

segunda, tomando en cuenta la naturaleza de la pregunta de investigación, el número de variables independientes, dependientes y de control, el tipo de variable (categórica-continua) y la distribución. Con ellas, se pretende que el alumno, junto con su profesor o profesora, seleccione las pertinentes para efectuar sus análisis de los datos.

Algunas de las pruebas o métodos estadísticos no fueron desarrollados en el capítulo y varios se encuentran en el capítulo 8 del centro de recursos en línea ("Análisis estadístico: segunda parte"), el cual podrá encontrar en: Material Complementario → Capítulos.

Tabla 10.22 Elección de los procedimientos estadísticos o pruebas[33]

1. Pregunta de investigación: Descriptiva	**Procedimiento o prueba**
• Datos nominales	Moda
• Datos ordinales	Mediana, moda
• Datos por intervalos o razón	Media, mediana, moda, desviación estándar, varianza y rango
2. Pregunta de investigación: diferencias de grupos	
a) Dos variables o grupos	
a.1. Muestras correlacionadas	
• Datos nominales	Prueba de McNemar
• Datos ordinales	Prueba de Wilcoxon para pares de rangos
• Datos por intervalos o razón	Prueba *t* para muestras correlacionadas
a.2. Muestras independientes	
• Datos nominales	*Chi* cuadrada
• Datos ordinales	Prueba Mann-Whitney U o prueba Kolmogorov-Smirnov para dos muestras
• Datos por intervalos o razón	Prueba *t* para muestras no correlacionadas o independientes
b) Más de dos variables o grupos	
b.1. Muestras correlacionadas	
• Datos nominales	Prueba Q de Cochran
• Datos ordinales	Análisis de varianza de Friedman en dos vías
• Datos por intervalos o razón	Análisis de varianza (ANOVA)
• Datos por intervalos o razón, control de efectos de otra variable independiente	Análisis de covarianza (ANCOVA)
b.2. Muestras independientes	
• Datos nominales	*Chi* cuadrada para *k* muestras independientes
• Datos nominales u ordinales (categóricos) y de intervalos-razón	*Chi* cuadrada de Friedman
• Datos ordinales	Análisis de varianza en una vía de Kruskal-Wallis (ANOVA)
• Datos por intervalos o razón	Análisis de varianza (ANOVA)
3. Pregunta de investigación: correlacional	
a) Dos variables	
• Datos nominales	Coeficiente de contingencia o *Phi*
• Datos ordinales	Coeficiente de rangos ordenados de Spearman o coeficiente de rangos ordenados de Kendall
• Datos por intervalos o razón	Coeficiente de correlación de Pearson (producto-momento)
• Una variable independiente y una dependiente (ambas de intervalos o razón)	Regresión lineal
• Datos por intervalos y nominales u ordinales	Coeficiente biserial puntual
• Datos por intervalos y una dicotomía artificial en una escala ordinal (la dicotomía es artificial porque subyace una distribución continua)	Coeficiente biserial
b) Más de dos variables	Análisis discriminante
• Datos nominales	Análisis de correlación parcial por rangos de Kendall
• Datos ordinales	Coeficiente de correlación parcial o múltiple, R^2
• Datos por intervalos o razón	

(*continúa*)

[33] Adaptado de Mertens (2005).

Tabla 10.22 (*continuación*)

4. Pregunta de investigación: causal o predictiva • Diversas independientes y una dependiente (las independientes en cualquier nivel de medición, la dependiente en nivel por intervalos o razón). Cuando las independientes son nominales u ordinales se convierten en variables "dummy" • Diversas independientes y dependientes • Agrupamiento (membresía de todos los datos) • Estructuras y redes causales	Regresión múltiple Análisis multivariado de varianza (MANOVA) Análisis discriminante (en una vía, jerárquico o factorial, de acuerdo con el número de variables involucradas) Análisis de patrones o vías (*path analysis*)
5. Pregunta de investigación: estructura de variables o validación de constructo Las variables deben estar por intervalos o razón	Análisis de factores

Tabla 10.23 Criterios para elegir las pruebas estadísticas[34]

Naturaleza de la pregunta de investigación	Número de variables independientes	Número de variables dependientes	Número de variables de control (covariables)	Tipo de variable: independiente/ dependiente	Distribución	Prueba
Comparación de grupos	1	1	0	Categórica/ continua	Normal	Prueba *t*
Comparación de grupos	1 o más	1	0	Categórica/ continua	Normal	Análisis de varianza
Comparación de grupos	1 o más	1	1	Categórica/ continua	Normal	Análisis de covarianza
Comparación de grupos	1	1	0	Categórica/ continua	No normal	Prueba de Mann-Whitney U
Relación entre grupos o categorías	1	1	0	Categórica/ categórica	No normal	*Chi* cuadrada
Correlación entre variables	La prueba no considera a una variable como independiente y la otra como dependiente, sólo establece el grado de relación. La causalidad la establece el investigador			Continua/ continua	Normal	Correlación de Pearson
Correlación entre variables	La prueba no considera a una variable como independiente y la otra como dependiente, sólo establece el grado de relación. La causalidad la establece el investigador			Categórica/ categórica	No normal	Correlación de Kendall o Spearman
Correlación entre variables	3 o más. La causalidad la establece el investigador. En la regresión múltiple, el coeficiente representa el porcentaje de variabilidad de la variable dependiente que explica el modelo de regresión			Continuas	Normal	Coeficiente de correlación múltiple (R)
Relación causal entre variables	1	1	0	Continua/ continua	Normal	Regresión lineal
Relación causal entre variables	2 o más	1	0	Continua/ continua	Normal	Regresión múltiple

Paso 6: realizar análisis adicionales

5 Este paso implica simplemente que una vez realizados nuestros análisis, es posible que decidamos ejecutar otros análisis o pruebas extras para confirmar tendencias y evaluar los datos desde diferentes ángulos. Por ejemplo, podemos en una tabla de contingencia calcular primero *Chi* cuadrada y luego *Phi*, *Lambda*, *T* de Cramer (*C*) y el coeficiente de contingencia. O después de un ANOVA, efectuar

[34] Adaptado de Creswell (2009) y Pett (2007).

los contrastes posteriores que consideremos apropiados. Resulta este paso un momento clave para verificar que no se nos haya olvidado un análisis pertinente. En esta etapa regularmente se eligen los análisis multivariados.

Paso 7: preparar los resultados para presentarlos

Se recomienda, una vez que se obtengan los resultados de los análisis estadísticos (tablas, gráficas, cuadros, etc.), las siguientes actividades, sobre todo para quienes se inician en la investigación:

1. Revisar cada resultado [análisis general → análisis específico → valores resultantes (incluida la significación) → tablas, diagramas, cuadros y gráficas].
2. Organizar los resultados (primero los descriptivos, por variable del estudio; luego los resultados relativos a la confiabilidad y la validez; posteriormente los inferenciales, que se pueden ordenar por hipótesis o de acuerdo con su desarrollo).
3. Cotejar diferentes resultados: su congruencia y en caso de inconsistencia lógica volverlos a revisar. Asimismo, se debe evitar la combinación de tablas, diagramas o gráficas que repitan datos. Por lo común, columnas o filas idénticas de datos no deben aparecer en dos o más tablas. Cuando éste es el caso, debemos elegir la tabla o elemento que ilustre o refleje mejor los resultados y sea la opción que presente mayor claridad. Una buena pregunta en este momento del proceso es: ¿qué valores, tablas, diagramas, cuadros o gráficas son necesarias?, ¿cuáles explican mejor los resultados?
4. Priorizar la información más valiosa (que es en gran parte resultado de la actividad anterior), sobre todo si se van a producir reportes ejecutivos y otros más extensos.
5. Copiar o "formatear" las tablas en el programa con el cual se elaborará el reporte de la investigación (procesador de textos —como Word— o uno para presentaciones, como Power Point, Flash, Prezi). Algunos programas como SPSS y Minitab permiten que se transfieran los resultados (tablas, por ejemplo) directamente a otro programa (copiar y pegar). Por ello, resulta conveniente usar una versión del programa de análisis que esté en el mismo idioma que se empleará para escribir el reporte o elaborar la presentación. Aunque, de no ser así, el texto de las tablas y gráficas puede modificarse, únicamente es más tardado.
6. Comentar o describir brevemente la esencia de los análisis, valores, tablas, diagramas, gráficas.
7. Volver a revisar los resultados.
8. Y, finalmente, elaborar el reporte de investigación.

En el centro de recursos se encuentran más ejemplos de estudios con diferentes análisis tratados en este capítulo y en el capítulo 8 adicional de la página, "Análisis estadístico: segunda parte". Se incluye al final del presente capítulo una secuencia de análisis en Minitab con la investigación de la televisión y el niño, y en el capítulo 8 adicional una secuencia de análisis en SPSS con un estudio del clima organizacional.

SPSS®
Centro de recursos en línea

Resumen

- El análisis cuantitativo de los datos se efectúa mediante la matriz de datos, la cual está guardada como archivo.
- Los pasos más importantes en el análisis de los datos son:
  - Decidir el programa de análisis de los datos por utilizar.
  - Explorar los datos obtenidos en la recolección:
    - *a*) Analizar descriptivamente los datos por variable del estudio.
    - *b*) Visualizar los datos por variable.
  - Evaluar la confiabilidad y validez del instrumento o instrumentos de medición utilizados.
  - Analizar e interpretar mediante pruebas estadísticas las hipótesis planteadas (análisis estadístico inferencial).
  - Realizar análisis adicionales.
  - Preparar los resultados para presentarlos.
- Los análisis estadísticos se llevan a cabo mediante programas computacionales, los más conocidos son: IBM SPSS, Minitab y SAS.
- El tipo de análisis o pruebas estadísticas depende del nivel de medición de las variables, las hipótesis y el interés del investigador.

- Centro de recursos en línea Los principales análisis estadísticos que pueden hacerse son: estadística descriptiva para cada variable (distribución de frecuencias, medidas de tendencia central y medidas de la variabilidad), la transformación a puntuaciones *z*, razones y tasas, cálculos de estadística inferencial, pruebas paramétricas, pruebas no paramétricas y análisis multivariados. Algunos fueron tratados en este capítulo y otros se comentan en el capítulo 8 del centro de recursos.
- Las distribuciones de frecuencias contienen las categorías, los códigos, las frecuencias absolutas (número de casos), los porcentajes, los porcentajes válidos y los porcentajes acumulados.
- Las distribuciones de frecuencias pueden presentarse en forma gráfica.
- Una distribución de frecuencias puede representarse por medio del polígono de frecuencias o de la curva de frecuencias.
- Las medidas de tendencia central son la moda, la mediana y la media.
- Las medidas de la variabilidad son el rango (diferencia entre el máximo y el mínimo), la desviación estándar y la varianza.
- Otras estadísticas descriptivas de utilidad son la asimetría y la curtosis.
- Centro de recursos en línea Las puntuaciones *z* son transformaciones de los valores obtenidos a unidades de desviación estándar (su explicación se incluye en el capítulo 8 del centro de recursos).
- Una razón es la relación entre dos categorías; una tasa es la relación entre el número de casos de una categoría y el número total de casos, multiplicada por un múltiplo de diez.
- La confiabilidad se calcula mediante coeficientes: de correlación, *alfa* y KR-20 y 21.
- La validez de criterio se obtiene mediante coeficientes de correlación y la de constructo por medio del análisis de factores.
- La estadística inferencial sirve para efectuar generalizaciones de la muestra a la población. Se utiliza para probar hipótesis y estimar parámetros. Se basa en el concepto de distribución muestral.
- La curva o distribución normal es un modelo teórico sumamente útil; su media es cero (0) y su desviación estándar es uno (1).
- El nivel de significancia o significación y el intervalo de confianza son niveles de probabilidad de cometer un error, o de equivocarse en la prueba de hipótesis o la estimación de parámetros. Los niveles más comunes son 0.05 y 0.01.
- Los análisis o las pruebas estadísticas paramétricas más utilizadas son:

Prueba	Tipo de hipótesis
Coeficiente de correlación de Pearson	Correlacional
Regresión lineal	Correlacional/causal
Prueba *t*	Diferencia de grupos
Contraste de la diferencia de proporciones	Diferencia de grupos
Análisis de varianza (ANOVA): unidireccional con una variable independiente y factorial con dos o más variables independientes	Diferencia de grupos/causal
Análisis de covarianza (ANCOVA). Véalo en el centro de recursos en línea del libro	Correlacional/causal

- En todas las pruebas estadísticas paramétricas las variables están medidas en un nivel por intervalos o razón.
- Los análisis o las pruebas estadísticas no paramétricas más utilizadas son:

Prueba	Tipo de hipótesis
Chi cuadrada	Diferencias de grupos para establecer correlación
Coeficientes de correlación e independencia para tabulaciones cruzadas: *phi*, *C* de Pearson, *V* de Cramer, *Lambda*, *Gamma*, *Tau* (varios), Somers, etcétera	Correlacional
Coeficientes de correlación de Spearman y Kendall	Correlacional
Coeficiente *Eta* para relaciones no lineales (ejemplos: curvilineales)	Correlacional

- Las pruebas no paramétricas se utilizan con variables nominales u ordinales o relaciones no lineales.
- Centro de recursos en línea Los análisis multivariados trabajan con más de un par de variables de manera simultánea y se presentan en el capítulo 8 del centro de recursos.
- Una vez analizados los datos, los resultados se preparan para incluirse en el reporte de la investigación.

Conceptos básicos

- Análisis de datos
- Análisis de factores
- Análisis de varianza
- Análisis multivariados
- Asimetría
- Categoría
- *Chi* cuadrada
- Codificación
- Coeficiente de correlación de Pearson
- Coeficiente de Kendall

- Coeficiente de Spearman
- Coeficientes de correlación e independencia para tabulaciones cruzadas
- Contraste de diferencia de proporciones
- Curtosis
- Curva de frecuencias
- Curva o distribución normal
- Desviación estándar
- Distribución de frecuencias
- Estadística
- Estadística descriptiva
- Estadística inferencial
- Estadística no paramétrica
- Estadística paramétrica
- Eta
- Gráficas
- Intervalo de confianza
- Matriz de datos
- Media
- Mediana
- Medidas de tendencia central
- Medidas de variabilidad
- Métodos cuantitativos
- Minitab
- Moda
- Nivel de significación o significancia
- Paquetes estadísticos
- Polígono de frecuencias
- Prueba *t*
- Pruebas estadísticas
- Puntuación *z*
- Rango
- Razón
- Regresión lineal
- SPSS®/IBM
- STATS®
- Tabulación cruzada
- Tasa
- Variable de la matriz de datos
- Variable del estudio
- Varianza
- Vista de las variables
- Vista de los datos

Ejercicios

1. Construya una distribución de frecuencias hipotéticas, con todos sus elementos, e interprétela verbalmente.
2. Localice una investigación científica donde se reporte la estadística descriptiva de las variables y analice las propiedades de cada estadígrafo o información estadística proporcionada (distribución de frecuencias, medidas de tendencia central y medidas de la variabilidad).
3. Un investigador obtuvo, en una muestra, las siguientes frecuencias absolutas para la variable "actitud hacia el director de la escuela":

Categoría	Frecuencias absolutas
Totalmente desfavorable	69
Desfavorable	28
Ni favorable ni desfavorable	20
Favorable	13
Totalmente favorable	6

   a) Calcule las frecuencias relativas o porcentajes.
   b) Grafique los porcentajes mediante un histograma (barras).
   c) Explique los resultados para responder a la pregunta: ¿la actitud hacia el director de la escuela tiende a ser favorable o desfavorable?
4. Un investigador obtuvo, en una muestra de trabajadores, los siguientes resultados al medir el "orgullo por el trabajo realizado". La escala oscilaba entre 0 (nada de orgullo por el trabajo realizado) a 8 (orgullo total).
   Máximo = 5
   Mínimo = 0
   Media = 3.6
   Moda = 3.0
   Mediana = 3.2
   Desviación estándar = 0.6
   ¿Qué puede decirse en esta muestra acerca del orgullo por el trabajo realizado?
5. ¿Qué es la curva normal? ¿Qué son el nivel de significancia o significación y el intervalo de confianza? Responda a estas preguntas en equipo con sus compañeros y coméntelo con su profesor.
6. Dependiendo del campo de estudio que le sea más afín, elija una de las siguientes variables y responda las preguntas: 1) ¿cómo se podría medir?, 2) ¿qué nivel de medición se tendría? y 3) en la ciudad donde vive, ¿esta variable tendería a tener una distribución normal o no? (¿por qué?). Comente las respuestas con su profesor. Las variables serían: presión arterial, motivación de los obreros de las fábricas más grandes en cuanto a número, temperatura ambiental en los meses de noviembre y diciembre (tomando el promedio de cada día), tamaño de los edificios, corrupción de los policías de tránsito o vialidad durante un mes (número de actos de corrupción), peso de los adolescentes (12-15 años), nivel de desempleo en el último año, nivel de deserción escolar en primarias públicas en el semestre más reciente, piezas producidas (tornillos) con defectos durante un mes (mediciones diarias). Desde luego, puede pensar en alguna otra.
7. Centro de recursos en línea Relacione las columnas A y B. En la columna A se presentan hipótesis; y en la columna B, pruebas estadísticas apropiadas para las hipótesis. Se trata de encontrar la prueba que corresponde a cada hipótesis (las respuestas se localizan en el centro de recursos en línea: "Respuestas a los ejercicios").

Columna A	Columna B
Hi: "A mayor inteligencia, mayor capacidad de resolver problemas matemáticos" (medidas las variables por intervalos).	— Diferencias de proporciones
Hi: "Los hijos de padres alcohólicos muestran una menor autoestima con respecto a los hijos de padres no alcohólicos" (autoestima medida por intervalos).	— *Chi* cuadrada
Hi: "El porcentaje de delitos por asalto a mano armada, en relación con el total de crímenes cometidos, es mayor en la ciudad de México que en Caracas".	— Spearman
Hi: "El género está relacionado con la preferencia por telenovelas o espectáculos deportivos.	— Coeficiente de correlación de Pearson
Hi: "La intensidad del sabor de productos empacados de pescado está relacionada con la preferencia por la marca" (sabor intenso, sabor medianamente intenso, sabor poco intenso, sabor muy poco intenso) (preferencia = rangos a 12 marcas).	— ANOVA unidireccional
Hi: "Se presentarán diferencias en cuanto al aprovechamiento entre un grupo expuesto a un método de enseñanza novedoso, un grupo que recibe instrucción mediante un método tradicional y un grupo de control que no se expone a ningún método".	— Prueba *t*

8. Desarrolle una hipótesis que requiera analizarse con la prueba *t*, una hipótesis que requiera de analizarse con *Chi* cuadrada y otra con el coeficiente de Spearman o Kendall.
9. Suponga un estudio cuya variable independiente es: años de experiencia del docente, y la dependiente: satisfacción del grupo (ambas medidas por intervalos), ¿qué pruebas y modelo estadístico le servirían para analizar los datos y cómo podrá efectuarse el análisis?
10. ¿Recuerda el estudio de Lee y Guerin (2009)? El objetivo era identificar si la satisfacción de la calidad del diseño ambiental de las zonas de trabajo u oficinas afecta significativamente la satisfacción general del espacio de trabajo por parte de sus ocupantes y su desempeño laboral. A continuación se presentan los coeficientes de las correlaciones que obtuvieron entre cada criterio de la satisfacción de la calidad del diseño del área de trabajo y la satisfacción general sobre el espacio de trabajo y la productividad (p. 299).[35] Interprete y comente con su profesor.

Criterios de satisfacción del empleado (ocupante)	Satisfacción general sobre el espacio	Valor P (Sign.)	Productividad del empleado	Valor P (Sign.)
Diseño de la oficina	0.884	0.116	0.886	0.114
Mobiliario	0.994*	0.006	0.994*	0.006
Temperatura	0.549	0.451	0.083	0.917
Ventilación	0.628	0.372	0.976*	0.024
Iluminación	0.938	0.062	0.936	0.064
Acústica	0.928	0.072	0.784	0.216
Limpieza y mantenimiento	0.656	0.344	0.942	0.058

*Correlación significativa al nivel del 0.05 ($P < 0.05$) (dos colas).

11. Dé un ejemplo hipotético de una razón "*F*" significativa e interprétela.
12. Construya un ejemplo hipotético de una tabulación cruzada y utilícela para fines descriptivos.
13. Busque un artículo de investigación en revistas científicas que contengan resultados de pruebas *t*, ANOVA y χ^2 aplicadas; evalúe la interpretación de los autores.
14. Para interpretar una prueba se requiere evaluar el resultado (valor) y... (complete la frase).
15. Respecto al estudio que ha ido desarrollando a lo largo del proceso cuantitativo, ¿qué pruebas estadísticas le serán útiles para analizar los datos? y ¿qué secuencia de análisis habrá de seguir? (Discútalo con su profesor y sus compañeros).

Ejemplos desarrollados

Comentario: Por cuestiones de espacio, se incluyen unos cuantos resultados de cada ejemplo, con fines ilustrativos.

La relación entre la personalidad y las enfermedades

El promedio de edad de la muestra fue de 53.8 años para los hombres (s = 7.2) y 53 para las mujeres (s = 7.2). No hubo diferencia significativa por género en cuanto a la disposición a participar en el estudio.

Las 19 enfermedades se asociaron entre sí mediante coeficientes de correlación tetracórica (presencia-ausencia), siendo los más elevados los siguientes: infarto agudo de miocardio y angina de pecho (0.82), cardiomiopatía y angina de pecho (0.73), infarto agudo de miocardio y cardiomiopatía (0.71), bronquitis crónica y asma (0.68), cirrosis del hígado y hepatitis (0.82) y

[35] Los términos fueron adaptados.

gastritis crónica y úlcera gástrica duodenal (0.76). En cambio, entre hipertensión e infarto fue de 0.35, y gastritis crónica y angina de pecho de 0.23 (Yousfi *et al.,* 2004, p. 637). En pocas palabras: "estas enfermedades van de la mano".

Se presentaron diferencias de medias notables o muy significativas por género para accidente cerebrovascular, diabetes, hipertensión y enfermedades cardíacas.

Los investigadores efectuaron un análisis de factores para establecer qué dimensiones de la personalidad se agrupan entre sí. Los factores emergentes fueron cinco:

1. Los autores le denominaron "labilidad emocional" (que abarca inhibición, barreras, sentido de coherencia, depresión, neuroticismo, salud/autonomía, ira interna, optimismo, envidia, irritabilidad y psicopatología).
2. Nombrado como "personalidad/conducta del tipo A" (tendencias antisociales, urgencia de tiempo y activación perpetua, control social exagerado y extraversión).
3. Etiquetado como control conductual (racionalismo, deseabilidad social, control de la ira, ira externa y agresión).
4. Locus de control (interno y externo).
5. Psicoticismo (psicoticismo y soporte social).

Finalmente, en relación con el planteamiento del problema, se encontraron correlaciones muy moderadas, aunque estadísticamente significativas de la personalidad con algunas enfermedades, en particular: la urgencia de tiempo y activación perpetua, hostilidad, control social, labilidad emocional, locus de control y psicoticismo, los cuales fueron destacados factorialmente. La labilidad emocional (inhibición, barreras, sentido de coherencia, depresión, neuroticismo, salud/autonomía, ira interna, optimismo, envidia, irritabilidad y psicopatología) parece ser el predictor más sólido de la vulnerabilidad general hacia las enfermedades. También, se encontraron algunas asociaciones pequeñas pero significativas entre enfermedades específicas y el tipo A y el control conductual.

La adiposidad y enfermedades cardíacas presentaron los coeficientes de correlación más altos con la personalidad/conducta de tipo A. A excepción del cáncer, complicaciones de la tiroides y padecimientos urinarios, todas las otras enfermedades se asocian con tal variable, aunque débilmente. Asimismo, salvo el cáncer, lo referente al hígado, las enfermedades pulmonares y de la tiroides, el resto de los padecimientos se correlacionan, aunque mínimamente con el control conductual.

Sin embargo, no se demostró que la personalidad fuera una de las principales causas de las enfermedades crónicas. Solamente se encontraron algunos efectos y relaciones moderadas entre ciertos factores de la misma y los padecimientos. Las preguntas de investigación podrían responderse así: en el estudio se descubrió correlación moderada entre la personalidad y las enfermedades, pero no puede asumirse causalidad. Se requieren más estudios.

La televisión y el niño

El análisis se realizó utilizando el Minitab. Los resultados son diversos para incluirlos en este espacio, se muestra únicamente la secuencia de análisis (véase la figura 10.21) y cabe señalar que el promedio de horas que dedican diariamente a ver televisión es de 3.1. La prueba *t* no reveló diferencias por género en este sentido.

Los programas favoritos de los niños en 2005 fueron: Bob Esponja, las telenovelas infantiles y Los Simpson (vea figura 10.22).

Figura 10.22 Programas preferidos (agrupados aquellos con menos de 4%).

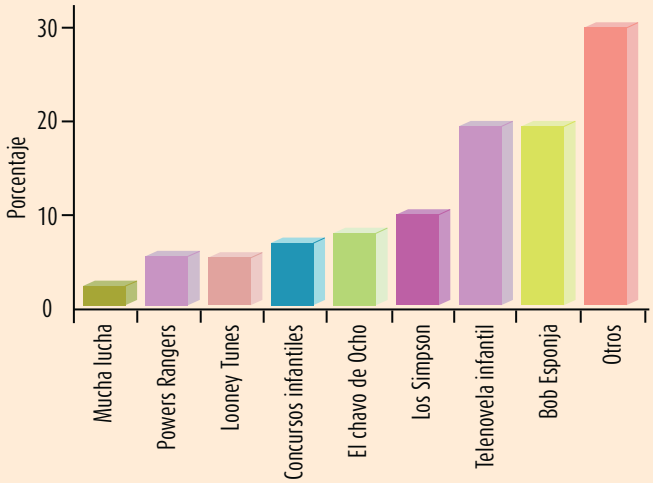
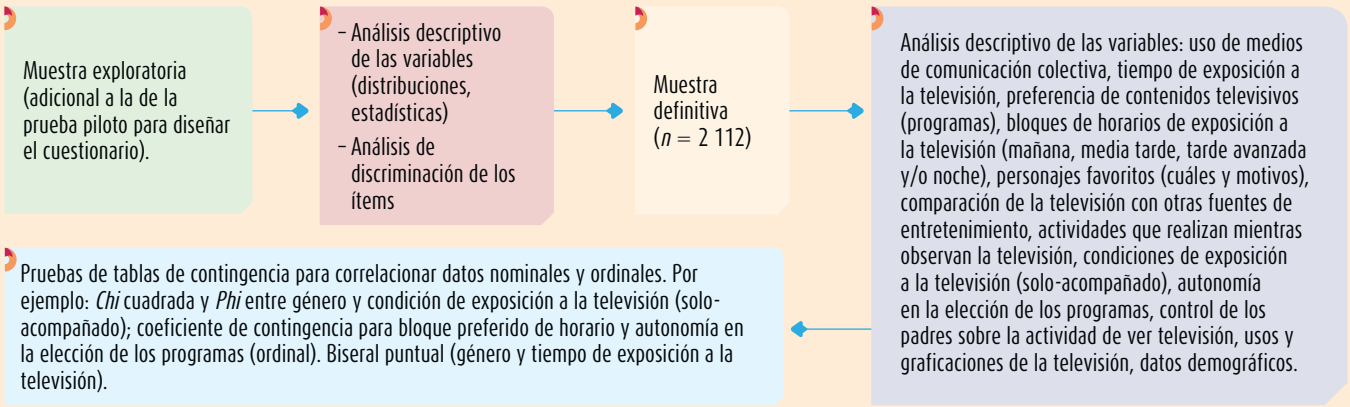


Figura 10.21 La secuencia de análisis con Minitab.

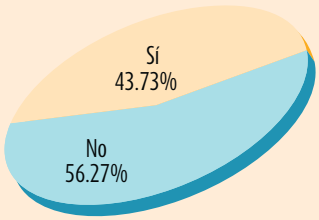


La pareja y relación ideales

Los análisis fueron realizados en el programa SPSS y las gráficas en un programa de ilustración. La $n = 725$ estudiantes. Se presentan únicamente las tablas y gráficas descriptivas de ciertas variables o preguntas en términos de porcentajes, con anotaciones muy breves, para que el lector, a manera de ejercicio —preferentemente grupal— amplíe los comentarios y desarrolle implicaciones de los resultados.

El promedio de edad de la muestra fue de 21 años y la mediana de 20. En cuanto al género: 46% mujeres y 54% hombres de una gran variedad de licenciaturas.

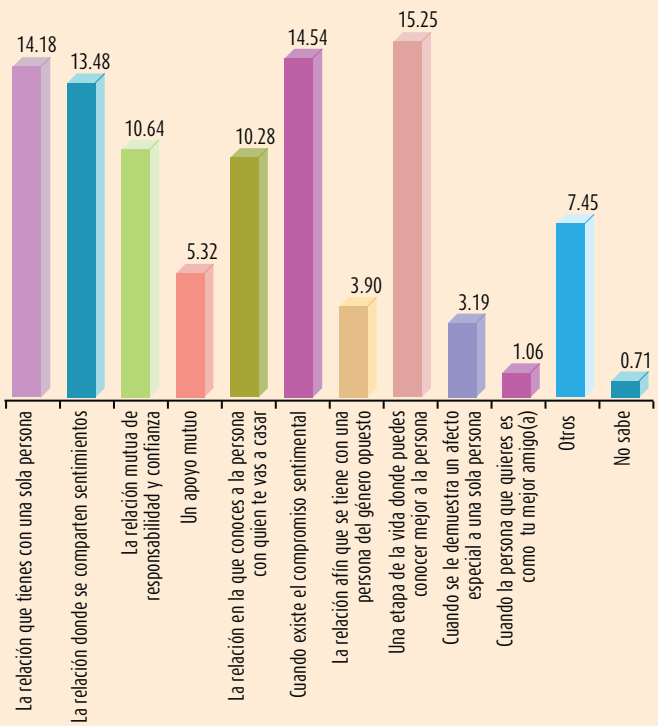
Figura 10.23 ¿Tienes novio/novia?



De la muestra, más de la mitad no tiene novio o novia (figura 10.23).

En cuanto a la definición del noviazgo, solamente uno de cada 10 estudiantes lo concibió bajo la dimensión prematrimonial ("la relación en la que conoces a la persona con quien te vas a casar"). 15.25% señaló que el noviazgo es una "etapa de la vida". ¿Qué más podría comentar de esta gráfica? (Figura 10.24)

Figura 10.24 Definición del noviazgo.



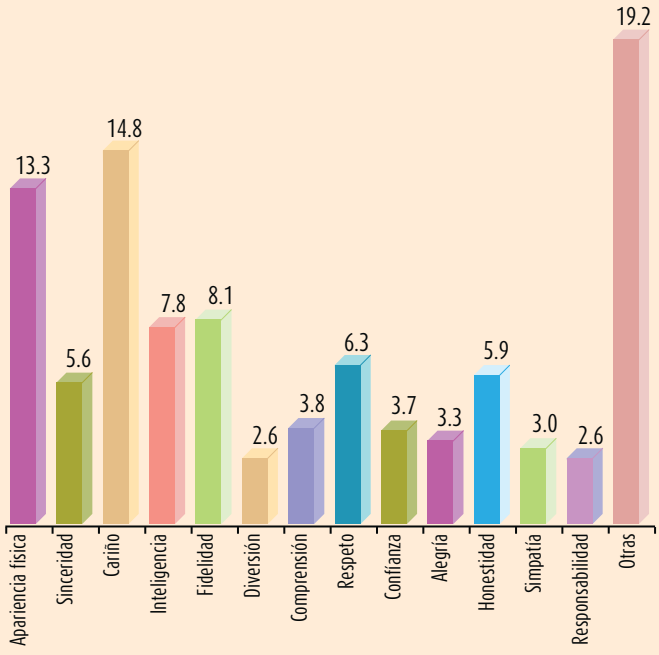
Por lo que respecta a la importancia (tabla 10.24), no hubo quien lo considerara que "no tiene importancia". La media fue de 4.13 y la mediana igual a 4.0 (mínimo 2 y máximo 5, desviación estándar = 0.813). ¿Qué puede decirse de esta tabla de acuerdo con lo expuesto en los apartados de estadística descriptiva de este capítulo? [Esta gráfica puede servir para discutir si la escala es ordinal o de intervalos y lo que nos dice la media].

Tabla 10.24 ¿Qué tan importante es en tu vida tu pareja actual?

Válidos	Categorías	Porcentaje válido	Porcentaje acumulado
5	Sumamente importante	38.4	38.4
4	Importante	37.6	76.0
3	Medianamente importante	22.4	98.4
2	Poco importante	1.6	100.0
	Total	100.0	

Piensa en tu novio(a) ideal y menciona las cualidades que te gustaría que tuviera

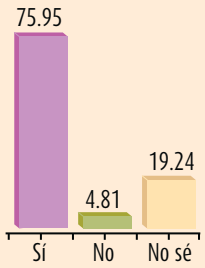
Figura 10.25 Cualidades del novio ideal.



Para la gráfica anterior se tomaron las cinco cualidades mencionadas por todos los estudiantes que integraron la muestra. Estimado lector, ¿qué nos dice la gráfica? Compárela con las cualidades que a usted le gustarían en su novio o novia ideal y discútalas con sus mejores amigos/amigas.

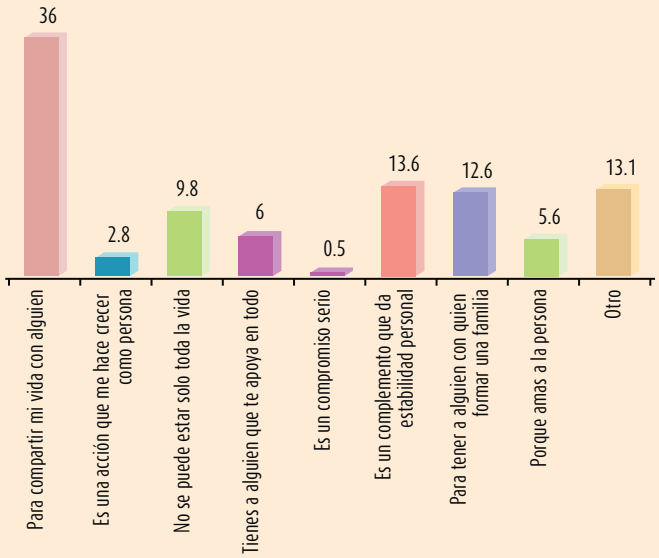
En el futuro, ¿le gustaría o no tener una relación de pareja para toda la vida?

Figura 10.26 Relación de pareja para toda la vida.



Prácticamente una quinta parte de la muestra no sabe si le gustaría o no tener una relación de pareja de por vida. Pero a la enorme mayoría (casi en proporción cuatro a uno) sí le agradaría.

Figura 10.27 Razones del "sí".

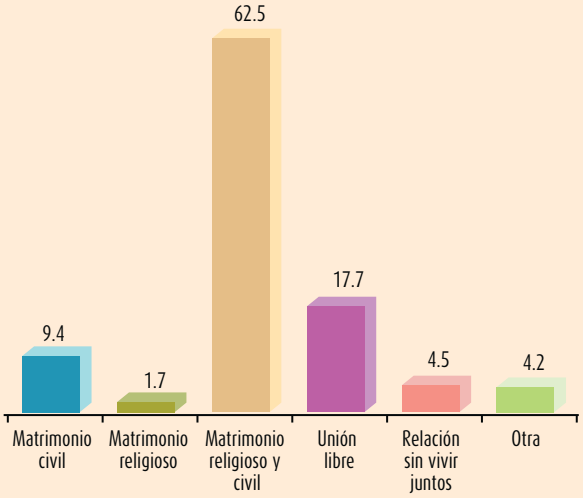


Si la pregunta se hubiera redactado: "En tu futuro, ¿te gustaría o no tener una sola o única relación de pareja para toda la vida?" (agregando "sola o única"), ¿cree usted que las respuestas hubieran cambiado en algo?

La razón principal de quienes respondieron afirmativamente que les gustaría tener una relación de pareja de por vida es el hecho de "compartir una vida" (poco más de una tercera parte). A 22.2% de los respondientes les gustaría en su futuro tener una relación de pareja duradera a largo plazo al margen de los "cánones establecidos" (*unión libre* y *relación sin vivir juntos*). Y 62.5% quisieran un matrimonio tradicional (figura 10.28).

Un último comentario es que al realizar una prueba de diferencia de proporciones entre hombres y mujeres respecto a la apariencia física, no cabe duda que los estudiantes celayenses le dan mayor importancia a ésta que sus compañeras (significancia menor del 0.05).

Figura 10.28 Tipo de relación duradera.



El abuso sexual infantil

En la tabla 10.25 se resume la confiabilidad de los instrumentos para CKAQ-Español ($n = 150$, $\bar{X} = 5.08$, $DS = 3.43$, y rango de 8 a 22 puntos) y RP-México ($n = 150$, $\bar{X} = 11.53$, $DS = 7.97$ y un rango de 0 a 38 puntos).

Tabla 10.25 Confiabilidad de instrumentos

Instrumento	Confiabilidad interna ($p < 0.01$)	Confiabilidad de estabilidad temporal "test-retest" ($p < 0.01$)	Tipo de instrumento
CKAQ-Español	0.69	0.50	Cognitivo
RP-México	0.75	0.75	Conductual
EPA	0.78	0.75	Conductual

Los tres grupos experimentales mostraron que existe un tipo de sensibilidad al instrumento en el CKAK-Español (Kruskal-Wallis $\chi^2 = 78.4$, gl = 2, $p < 0.001$) y RP-México (Kruskal-Wallis $\chi^2 = 83.06$, gl = 2, $p < 0.001$), lo cual indica que los grupos difieren en localización o forma, y reafirma la sensibilidad de las escalas al mostrar un comportamiento diferente entre los grupos, donde quienes más recientemente terminan el Programa de Prevención del Abuso Sexual Infantil (PPASI) mejores puntajes obtienen.

Con el objetivo de indagar el comportamiento de los grupos de seguimiento y control con respecto al grupo que termina un PPASI, se calcularon los porcentajes relativos en las subescalas de hacer, decir, denunciar y el reconocimiento de los contactos positivos y negativos. Los resultados se exhiben en la tabla 10.26, que presenta los porcentajes de aciertos en relación con el grupo que acaba de concluir un PPASI en el RP-México. Se deduce que para las subescalas de reconocimiento de contactos negativos, DECIR y DENUNCIAR, se conserva el cambio esperado, quienes más recientemente hayan participado en un PPASI obtienen un mejor puntaje. En la habilidad de HACER, se observa que el grupo de seguimiento tiene un mejor desempeño en prome-

dio. Esto se puede explicar debido al incremento en la madurez de las niñas y niños, lograda a lo largo de un año aproximadamente; de 5.58 años en el primer grupo y de 6.47 años promedio en el grupo de seguimiento.

En la subescala de contactos positivos se advierte que al terminar el PPASI "Porque me quiero me cuido", tiene un puntaje promedio ligeramente menor que el grupo de control (11.45%), y mucho mejor en el grupo de seguimiento (53.71%), lo que avala que la escala es sensible a medir dicha habilidad y su posible efecto nocivo ante un PPASI. Este resultado si bien acredita parcialmente los resultados de Underwager y Wakefield (1993), que sostiene que al atender a los PPASI, los niños y niñas se muestran desconfiados ante las aproximaciones cotidianas normales. También se constata que al cabo de un año de concluido el programa, los infantes son capaces de superarlo y se muestran mucho más asertivos. Entre los grupos al terminar PPASI y de control se evidenció que solamente la habilidad de reconocer contactos positivos (Mann-Whitney $z = -1.48$, $n = 124$, $p = 0.14$) es la única habilidad que tiene la misma localización, esta evidencia contradice la teoría de Underwager y Wakefield (1993). Se puede concluir en este estudio que, si bien es cierto que al terminar el PPASI las niñas y niños aumentaban ligeramente el recelo ante los contactos positivos, esto no es significativo y con el tiempo, al incremento de la madurez, el fenómeno se supera.

En el caso de los grupos al terminar un PPASI y de seguimiento, se encontró que hay un mismo comportamiento en las subescalas de DECIR (Mann-Whitney $z = -1.20$, $n = 72$, $p = 0.23$) y HACER (Mann-Whitney $z = -1.26$, $n = 72$, $p = 0.21$) por lo que hay permanencia en el tiempo de estas dos habilidades.

La correlación entre las diferentes escalas verifica que hay una vinculación aceptable entre CKAQ-Español y RP-México (Spearman $r = 0.68$, $n = 150$, $p < 0.01$), para el total de casos experimentales. En el caso de los grupos de control (Spearman $r = 0.23$, $n = 79$, $p < 0.05$) y al terminar el PPASI (Spearman $r = 0.35$, $p < 0.05$) las escalas tienen un nivel de correlación aún menor.

Tabla 10.26 Porcentaje de rangos relativos con respecto al grupo que termina un PPASI

Grupo	Contactos negativos (%)	Contactos positivos (%)	Decir	Hacer	Denunciar
Al terminar	100.00	100.00	100.00	100.00	100.00
PPASI	77.82	53.71	90.86	105.71	74.28
Seguimiento					
Control	37.39	115.45	44.14	49.14	41.5 2

El resumen de puntajes por escala y grupo experimental se presenta en la tabla 10.26, donde la columna CKAQ-Español (Transformado) presenta la conversión de una escala con un puntaje de 0 a 22 a una de 0 a 40, para tener un comparativo equivalente entre las escalas cognitiva y conductual. Se observa que la escala cognitiva se encuentra por encima de la conductual en todos los grupos. Sin embargo, el grupo de control observa una mayor diferencia entre las escalas conductuales y cognitivas. El porcentaje relativo, con respecto al puntaje promedio del grupo al terminar PPASI, proporcionalmente el grupo de seguimiento obtiene 87.07% de aciertos y el grupo de control 69.88%, para la escala CKAQ-Español. En el caso del instrumento conductual RP-México, el porcentaje relativo es de 70.85% de aciertos en el grupo de seguimiento y 31.19% en el grupo de control. Lo que evidencia que la sensibilidad al cambio en esta escala conductual es mayor.

Se observa también que la distribución de la escala cognitiva CKAQ-Español, en general, se encuentra por encima de las escalas conductuales, lo que permite deducir que los menores pueden tener cierto grado de conocimiento que no se traduce en habilidades de protección personal.

Tabla 10.27 Resumen descriptivo de puntajes por escala y grupo experimental

Grupo experimental	CKAQ-Español	CKAQ-Español (Transformado)	RP-México
Al terminar PPASI			
Media	18.33	33.32	19.11
Desviación estándar	2.66	4.83	6.48
Seguimiento			
Media	15.96	29.01	13.54
Desviación estándar	2.34	4.26	5.23
Control			
Media	12.81	23.29	5.96
Desviación estándar	2.17	3.94	4.37
Total			
Media	15.08	27.42	11.53
Desviación estándar	3.43	2.24	7.97

Los investigadores opinan

Desde 1990 han disminuido las tensiones entre lo cualitativo y lo cuantitativo, por lo que se buscó establecer una sinergia, así como ser más flexibles y eclécticos, dicho en el buen sentido, en los procedimientos.

La investigación cuantitativa ganó cuando particularizó los instrumentos y tomó en cuenta las características de los grupos a los cuales se dirige el estudio. Lo anterior propició un importante avance en la explicación de los procesos psicológicos, en especial los cognoscitivos; y en los descubrimientos neuropsicológicos, así como en el uso de software para el montaje de experimentos, demostraciones y simulaciones.

En este tipo de investigaciones, destacan las pruebas estadísticas por su utilidad en el análisis de datos categóricos de correspondencia, la ordenación de datos para conocer preferencias, el análisis factorial confirmativo, las correctas estimaciones de conjuntos de datos complejos, el manejo de resultados estadísticos de los experimentos, la validación de datos, la determinación del tamaño de la muestra y el análisis de regresión, entre otros aspectos a considerar.

A pesar de tan importantes avances en la investigación, aún hace falta financiamiento para una promoción significativa y que, además, fomente la especialización de los investigadores, lo cual les permitiría competir de manera efectiva.

Ciro Hernando León Pardo
Coordinador del Área de Investigación
Facultad de Psicología
Universidad Javeriana
Bogotá, Colombia

Para efectuar una buena investigación se requiere plantear de forma correcta el problema, con lo cual tenemos 50%, y también con un rigor metodológico, es decir, incluir todos los pasos del proceso.

Tal apego a la metodología implica el empleo de los recursos pertinentes; por ejemplo, en las investigaciones sociales las pruebas estadísticas proporcionan una visión más precisa del objeto de estudio, ya que apoyan o no las hipótesis para su validación o rechazo.

Los estudiantes pueden concebir una idea de investigación a partir de sus intereses personales, aunque se recomienda que elijan temas íntimamente relacionados con su carrera, y que procuren que sean de actualidad y de interés común.

Para ello, los profesores deben infundir en los alumnos la importancia de la investigación en el terreno académico y en el profesional, destacando su relevancia tanto en la generación de conocimiento como en la búsqueda de soluciones a problemas.

Roberto De Jesús Cruz Castillo
Profesor de tiempo completo
Facultad de Ciencias de la Administración
Universidad Autónoma de Chiapas
Chiapas, México