INCLUYE VERSIÓN DIGITAL

Macchi

Introducción a la Estadística en Ciencias de la Salud

3.ª EDICIÓN

EDITORIAL MEDICA panamericana

Introducción a la Estadística en Ciencias de la Salud

Introducción a la Estadística en Ciencias de la Salud

3.ª EDICIÓN

RICARDO LUIS MACCHI

Odontólogo y Doctor en Odontología, Universidad de Buenos Aires
Master of Science, Universidad de Michigan, Estados Unidos
Profesor Emérito, Cátedra de Materiales Dentales, Facultad de Odontología, Universidad de Buenos Aires
Miembro de Número, Academia Nacional de Odontología, Buenos Aires, Argentina

—

EDITORIAL MEDICA panamericana

BUENOS AIRES - BOGOTÁ - MADRID - MÉXICO
e-mail: info@medicapanamericana.com
www.medicapanamericana.com

ISBN: 978-950-06-

Esta edición se terminó de imprimir en los talleres de
, Buenos Aires, Argentina
en el mes de noviembre de 2019

IMPRESO EN LA ARGENTINA

Las ciencias de la salud están en permanente cambio. A medida que las nuevas investigaciones y la experiencia clínica amplían nuestro conocimiento, se requieren modificaciones en las modalidades terapéuticas y en los tratamientos farmacológicos. Los autores de esta obra han verificado toda la información con fuentes confiables para asegurarse de que ésta sea completa y acorde con los estándares aceptados en el momento de la publicación. Sin embargo, en vista de la posibilidad de un error humano o de cambios en las ciencias de la salud, ni los autores, ni la editorial o cualquier otra persona implicada en la preparación o la publicación de este trabajo, garantizan que la totalidad de la información aquí contenida sea exacta o completa y no se responsabilizan por errores u omisiones o por los resultados obtenidos del uso de esta información. Se aconseja a los lectores confirmarla con otras fuentes. Por ejemplo, y en particular, se recomienda a los lectores revisar el prospecto de cada fármaco que planean administrar para cerciorarse de que la información contenida en este libro sea correcta y que no se hayan producido cambios en las dosis sugeridas o en las contraindicaciones para su administración. Esta recomendación cobra especial importancia con relación a fármacos nuevos o de uso infrecuente.

Imagen de tapa:
??????????????

EDITORIAL MEDICA panamericana

Visite nuestra página web:
http://www.medicapanamericana.com

ARGENTINA
Marcelo T. de Alvear 2145
(C1122AAG) Buenos Aires, Argentina
Tel.: (54-11) 4821-5520 / 2066 /
Fax (54-11) 4821-1214
e-mail: info@medicapanamericana.com

COLOMBIA
Carrera 7a A Nº 69-19 - Bogotá D.C., Colombia
Tel.: (57-1) 345-4508 / 314-5014 /
Fax: (57-1) 314-5015 / 345-0019
e-mail: infomp@medicapanamericana.com.co

ESPAÑA
Calle Sauceda 10, 5a planta (28050) - Madrid, España
Tel.: (34-91) 1317800 / Fax: (34-91) 4570919
e-mail: info@medicapanamericana.es

MÉXICO
Av. Miguel de Cervantes Saavedra Nº 233 piso 8, Oficina 801
Colonia Granada, Delegación Miguel Hidalgo -
C.P. 11520 - México, Distrito Federal
Tel.: (52-55) 5250-0664 / 5262-9470 / Fax: (52-55) 2624-2827
e-mail: infomp@medicapanamericana.com.mx

Prefacio de la tercera edición

La tercera edición de este texto, como las anteriores, está dedicada a quienes se forman o desempeñan en las ciencias de la salud. Tratamos así de poner a su disposición una sencilla introducción a los conocimientos básicos de las herramientas utilizadas en los procedimientos estadísticos.

La experiencia recogida en el desarrollo de actividades docentes en carreras de grado y posgrado nos ha permitido realizar algunos cambios e incorporar ciertos conceptos que complementan los incluidos en las ediciones anteriores, aunque manteniendo el formato y el criterio original.

Las palabras de los prefacios de la primera y segunda edición siguen vigentes y reflejan el espíritu que nos ha guiado en esta tarea.

Espero que la respuesta de los lectores continúe siendo la que hemos constatado hasta ahora.

Ricardo L. Macchi

Prefacio de la segunda edición

Alentados por la respuesta recibida, presentamos una nueva edición de este texto. No son muchas las modificaciones incorporadas y hemos mantenido el enfoque de considerar la obra como una manera de introducir al lector en el tema. Lo hemos mantenido porque nuestra experiencia en la docencia nos muestra que, en las ciencias de la salud, entender y analizar los resultados estadísticos que se encuentran en los documentos científicos e informativos generales continúa siendo una dificultad frecuente.

Creemos que la forma de desarrollo de los temas en el texto posibilitará la capacitación inicial para superar esa dificultad. Fue pensado para los profesionales que trabajan en distintas áreas: asistencial, docente y de investigación, y en las distintas ramas de las ciencias de la salud.

El objetivo general planteado es que el lector llegue a ser capaz de identificar los principios que justifican la utilización de técnicas estadísticas en la evaluación de los resultados obtenidos en un proceso de investigación en ciencias de la salud. No pretende capacitar en la aplicación de técnicas de procesamiento estadístico, sino generar una situación que ayude, a quien lo necesite, a encarar exitosamente el estudio más detallado del tema mediante la consulta de textos más avanzados y la participación en cursos específicos.

Como en nuestra intención original, deseamos brindarle al lector un acceso ágil a información que le facilitará su inserción paulatina en el mundo de la estadística y la investigación científica.

Ricardo L. Macchi
Marzo, 2005

Prefacio de la primera edición

En el ejercicio de la labor asistencial, docente o de investigación en ciencias de la salud es de rutina la consulta y el manejo de información en forma de datos que requieren de un procesamiento numérico.

Para la interpretación y valoración de la información presentada de esa manera y, cuando se hace necesario, para interactuar con los correspondientes expertos, el profesional que se desempeña en esas áreas debe identificar los fundamentos de las técnicas estadísticas.

En las páginas de este texto se analizan conceptos que pretenden poner al alcance del lector esos conocimientos básicos, sin cubrir con profundidad la descripción de las técnicas de procesamiento estadístico de datos.

El objetivo es que su lectura contribuya a la formación en la interpretación de la información de resultados estadísticos que se encuentran en los documentos científicos e informativos generales.

Además, se espera que el lector que lo necesite pueda posteriormente encarar exitosamente el estudio más detallado del tema mediante la consulta de textos más avanzados, la participación en cursos específicos y, fundamentalmente, mediante la aplicación de los procedimientos estadísticos en su tarea cotidiana.

Ricardo L. Macchi
Marzo, 2001

Índice

CAPÍTULO

1

DEFINICIONES Y APLICACIONES

FENÓMENOS Y SU DESCRIPCIÓN

Las ciencias de la salud se encuadran dentro de las denominadas ciencias fácticas, puesto que en ellas el objeto de estudio es un conjunto de hechos o fenómenos implícitos en el concepto de salud.

Al igual que en las demás ciencias que se incluyen bajo esa denominación, son de particular interés los hechos o fenómenos que varían al cambiar las circunstancias bajo las cuales se producen. Por ejemplo, el comportamiento, que puede variar según el individuo (paciente) al que se trate o de la situación ante la cual se encuentre o el resultado de la administración de un medicamento, que también puede variar según el paciente, la dosis que se utilice y, seguramente, por muchas otras causas.

Por estas razones, los hechos de interés son definidos como variables, por lo cual para el trabajo en el campo científico se hace necesario identificarlas y diferenciarlas, a fin de poder analizarlas, evaluar las condiciones en que se producen y así intentar prever, prevenir o modificar su ocurrencia.

En el campo de las ciencias de la salud esto significa la posibilidad de realizar acciones preventivas, diagnósticas o terapéuticas.

La capacitación en cuanto a las técnicas de valoración de variables es, entonces, una parte constituyente de la formación profesional.

Estas consideraciones se aplican en cualesquiera de las actividades que se consideren dentro de las que realiza un profesional de la salud: asistenciales, de investigación o docentes.

En la tarea asistencial, por la necesidad de evaluar las variables que presente el objeto de su atención (un paciente o una comunidad); en la investigación, porque una variable es precisamente el objeto para investigar; y en la docencia, para poder analizar la forma en que se produce la variable aprendizaje o cómo se ve afectada ante diferentes circunstancias.

METODOLOGÍAS CUALITATIVA Y CUANTITATIVA

La tarea de descripción y valoración de las variables de interés en una investigación, en la labor asistencial o en la labor docente puede ser realizada de dos maneras. En todos los casos se busca, como ya se indicó, identificar y diferenciar esos hechos para luego poder analizarlos y así llegar a conclusiones relacionadas con las causas que los producen o sobre la forma en que se puede prever o modificar su ocurrencia.

En la primera manera, el hecho o fenómeno, la variable, se detalla mediante una descripción narrativa; es decir, se utilizan palabras para la elaboración de un texto. Esta forma de trabajo se identifica como **metodología cualitativa.**

En la segunda manera, la variable es descrita mediante un dato que puede luego ser considerado, en conjunto con otros similares, y analizado mediante técnicas de procesamiento numérico. En este caso, el trabajo se identifica como de **metodología cuantitativa**.

La metodología cuantitativa aplica técnicas de procesamiento de números, las cuales constituyen el objeto de interés de lo que se conoce como **estadística**.

—

Teniendo en cuenta que muchas de las variables que son de interés en las ciencias de la salud se prestan al trabajo con metodología cuantitativa, surge la necesidad de que el profesional que se dedica a ellas conozca los principios fundamentales de esta ciencia y técnica dedicada al procesamiento de datos numéricos. Solo así estará capacitado para evaluar convenientemente la información sobre hechos que hacen a su labor asistencial, de investigación o docente y, cuando surja la necesidad, podrá interactuar con profesionales de la estadística para llegar, en una tarea interdisciplinaria, a la generación y aplicación de conocimientos.

ESTADÍSTICA DESCRIPTIVA Y ESTADÍSTICA INFERENCIAL

Ya se indicó que la estadística se ocupa del procesamiento numérico de datos. Esta disciplina incluye dos grandes capítulos en función del objetivo final de su aplicación.

En uno de esos capítulos, las técnicas estadísticas se utilizan para resumir los datos obtenidos en un conjunto de situaciones que tienen algo en común. Por ejemplo, para resumir el resultado obtenido en un grupo de individuos con una determinada enfermedad y que fueron sometidos a un tratamiento específico, o ante la presencia de casos de una determinada condición en los habitantes de una región geográfica específica.

Las técnicas que se utilizan para obtener una valoración numérica de la manifestación de una variable dentro de un conjunto de individuos están dentro de lo que se denomina **estadística descriptiva**.

—

Es habitual que el interés científico esté centrado en la totalidad de los hechos que se producen en condiciones determinadas. Siguiendo los ejemplos del párrafo anterior, el resultado del tratamiento en la totalidad

de los pacientes con esa determinada enfermedad o la totalidad de los habitantes de esa región geográfica específica. Es decir, que el objetivo es describir la manera en que se producen los hechos y la forma que toma la variable en una población. Se indica con este término a un conjunto de elementos, individuos o, más genéricamente, a unidades experimentales (unidades a partir de las cuales se lleva a cabo un experimento) o de observación (unidades en la que el fenómeno se observa o analiza), que tienen por lo menos una característica observable en común. Siguiendo los ejemplos, padecer una misma enfermedad o habitar en una misma región geográfica.

Las poblaciones de interés son generalmente demasiado grandes como para que los datos puedan ser registrados en cada uno de sus integrantes. La forma de trabajo y las técnicas de investigación significan, por ello, registrar datos solo en un subconjunto de la población denominado **muestra,** en la cual deben estar representadas las características o condiciones que definen al conjunto total.

Las técnicas de lo que se conoce como **estadística inferencial** permiten, mediante el procesamiento numérico de los datos registrados en una muestra, realizar inferencias sobre la forma que asume la variable de interés en la población respectiva.

—

Las técnicas de la estadística inferencial incluyen la **estimación de parámetros** con "intervalos de confianza" y la **prueba de hipótesis formuladas como punto de partida** de una investigación.

Los siguientes capítulos incluyen la presentación básica de los procedimientos de la estadística descriptiva y de los principios en los que se fundamenta la estadística inferencial.

SÍNTESIS CONCEPTUAL

Los hechos de interés en el campo de las ciencias fácticas se definen como **variables y es necesario diferenciarlas** para poder analizarlas.

Cuando se emplea la metodología cuantitativa, esa diferenciación se hace a partir de datos que permiten posteriormente su procesamiento numérico mediante las técnicas estadísticas.

La estadística descriptiva permite resumir información sobre la manifestación de una variable a partir de un conjunto de datos.

La estadística inferencial permite, a partir de una muestra, realizar inferencias sobre la forma que asume la variable de interés en la población respectiva.

EJEMPLO 1-1

Con la finalidad de planificar estrategias preventivas en una comunidad, se hizo necesario conocer el grado de información sobre el cuidado de la salud que tienen sus integrantes.

Para ello, la variable de interés, la información sobre el cuidado de la salud, puede tratar de valorarse con técnicas que permitan el procesamiento numérico a partir de una muestra de individuos de esa población.

La estadística inferencial permite, a partir de la información obtenida, estimar la situación de la población y concretar la tarea de planificación sobre una base de certidumbre razonable.

EJEMPLO 1-2

El objetivo de una investigación fue tratar de establecer si puede aceptarse o no la hipótesis de que la administración de ácido acetilsalicílico (AAS) a pacientes de un determinado nivel de edad y condición basal modifica la aparición de enfermedades coronarias, en comparación con lo observado al administrar un placebo.

En esta situación, las variables en análisis son la administración de un determinado medicamento, AAS o placebo, y la manera, magnitud o forma en que se produce la aparición de la enfermedad.

Si la segunda de estas variables se evalúa con la posibilidad de aplicación de técnicas de procesamiento numérico, podrá utilizarse la estadística inferencial para fundamentar la decisión de rechazar o no la hipótesis formulada a partir de los resultados obtenidos en una muestra de pacientes con las citadas características.

CAPÍTULO

2

DATOS: TIPOS Y CARACTERÍSTICAS

INTRODUCCIÓN

En el capítulo anterior se manifestó que el trabajo en las ciencias fácticas, dentro de las cuales se ubican las ciencias de la salud, se lleva a cabo tratando de comprender y explicar fenómenos de interés o para estimar cómo se puede modificar la forma en que estos se producen. Esos fenómenos constituyen las variables que deben ser observadas y de las que se debe registrar la forma en que se manifiestan.

Cuando se utiliza la metodología cuantitativa (**cap. 1, Definiciones y aplicaciones**) se trabaja con recolección de datos a través de mediciones fisiológicas o de otra índole, observación de comportamientos, toma de encuestas o mediante otras técnicas. Los datos así obtenidos representan una información que permite describir los hechos o fenómenos, es decir, las variables de interés.

Los datos son una forma de evaluar un atributo de una unidad experimental –sujeto experimental, en el caso de la investigación clínica– o una unidad de observación, si es que se actúa sobre ella para tratar de generar una modificación en ese atributo en una situación espacial y temporal determinada.

—

En un experimento, los datos que evalúan la variable independiente (tratamientos) permiten conformar los grupos en los que se evaluará la variable dependiente (respuesta). El análisis de los datos que evalúan a esta última, el desenlace o la respuesta al tratamiento, permite tomar decisiones sobre hipótesis formuladas, elaborar teorías explicativas o ambas.

En la investigación con metodología cuantitativa los datos pueden, en última instancia, evaluarse numéricamente y someterse a procedimientos de análisis estadístico.

—

Existen varias formas posibles de datos que permiten el procesamiento estadístico, y en cada circunstancia (asistencial, docente o de investigación) es necesario seleccionar la más conveniente.

DATOS NUMÉRICOS

Una posibilidad es describir cada hecho en particular con un número que permita identificarlo y diferenciarlo de otros hechos registrados en condiciones similares.

—

Por ejemplo, identificar lo que sucede en un integrante de una población y diferenciar la forma que la variable asume en él, en comparación de cómo lo hace en otro integrante de la misma población.

Con frecuencia se utiliza la palabra numéricos para hacer referencia a este tipo de datos, y es la que se utilizará en este texto, ya que es la denominación que generalmente se utiliza en programas de computación para estadísticas. Sin embargo, es importante tener presente que también se emplean otras denominaciones, como **datos cuantitativos** o **datos de medición.**

El número que describe la variable puede ser obtenido de varias maneras, lo que da lugar a diferentes formas de datos numéricos.

De relación o proporción

En este caso, el número que permite identificar el hecho o variable se obtiene al relacionarlo con una forma de la variable tomada como patrón o referencia.

En términos numéricos, "relacionar" significa aplicar la operación matemática conocida como división. Esto indica que la variable se describe al dividir la forma en la que se manifiesta en una unidad experimental o de observación por la manera en la que se produce en el patrón o referencia.

Un ejemplo permite comprender mejor esta idea. Supóngase que la variable de interés es la estatura de los individuos, definida como la longitud de la distancia entre la cabeza y los pies en posición erguida. La manera de obtener este tipo de dato consistiría en registrar esa distancia en cada individuo y ver cuántas veces cabe un patrón dentro de esa longitud, por ejemplo, una varilla cualquiera; es decir, dividir la longitud problema por la longitud patrón. Así se obtendría un número, como 4, 4,23, 3,42, etc., que es una valoración de la estatura, variable de interés, en cada individuo.

El patrón empleado puede ser cualquiera que se considere conveniente, pero, si existiera, resulta preferible emplear uno que sea reconocido de manera generalizada como tal. De este modo, se simplifica la comparación entre datos obtenidos para una misma variable en diferentes condiciones. Así, para la estatura, que ya fue definida como una longitud, resulta apropiado tomar como patrón o referencia la longitud "metro", cuya aceptación es prácticamente universal.

En última instancia, se registrará la estatura en forma de: 1,65 m, 1,72 m, etcétera. En la práctica, es probable que la división mencionada no se realice, sino que se emplee un instrumento, una regla u otro dispositivo, que permita registrar el dato en forma simple.

Nótese que cada hecho se identifica con un número y que ese número puede asumir cualquier valor entre dos límites. Ambos límites, en teoría, son los límites de la escala de números naturales que se extiende desde infinito negativo hasta infinito positivo.

Así, el valor de la estatura podrá ser cualquier número entre esos dos límites y en una escala continua. Se indica continua porque no existe ningún intervalo vacío entre dos números, cualesquiera que se tomen. De este modo, la estatura puede ser 1,70 o 1,73 m, pero entre ambos puede ser 1,725 o 1,7248 m, y así sucesivamente. Obviamente, en una situación real se debe resolver hasta dónde "redondear" el registro, que en el caso de la estatura de seres humanos es probable que solo se registren datos al centímetro. Distinta sería la situación al evaluar la longitud del diámetro de un microorganismo, que se redondeará posiblemente a décimas de micrómetro, o de la distancia entre dos ciudades, que se redondeará al kilómetro.

De la misma manera, los valores de estatura, así como los del diámetro de microorganismos o la distancia entre ciudades, se ubicarán entre límites reales que no son el infinito positivo o negativo. Estas situaciones son solo derivadas de razones de practicidad, pero el dato no deja de ser un **dato numérico continuo, lo cual debe ser tenido en** cuenta en el procesamiento ulterior de los datos.

Interválicos

Otra manera de llegar a datos numéricos continuos es establecer un intervalo numérico entre dos formas de la variable de interés y describir una situación, en particular por su ubicación dentro de ese intervalo.

Un ejemplo típico es la evaluación de la variable temperatura. En la escala centígrada o de Celsius se definen dos situaciones de temperatura, en las cuales una se considera como 0, temperatura de congelación del agua en condiciones normales de presión, y otra como 100, temperatura de ebullición del agua en las mismas condiciones. Una temperatura corporal de 36,8 °C representa la posición del individuo dentro de ese intervalo.

A diferencia de lo que sucede con los datos numéricos obtenidos de la forma descrita en el acápite anterior, en el caso de este tipo de datos el valor 0 no indica la ausencia de manifestación del fenómeno variable, sino únicamente un estado particular arbitrariamente definido.

Nótese que también en este caso los valores pueden ser infinitos (continuos), aunque en una situación particular se los redondee en función de la necesidad y de las posibilidades de los instrumentos que se empleen para el registro del dato.

Discretos

En ocasiones, el número que describe la situación o variable se obtiene al contar cuánto de algo tiene la unidad experimental. Por ejemplo, la cantidad de dientes faltantes en su boca o la cantidad de respuestas correctas en un cuestionario.

Si bien en este caso el dato también es numérico, no es continuo, sino **discreto,** con lo que se indica así que entre uno y otro valor existe un "vacío". Esta situación debe ser tenida en cuenta en algunas situaciones de procesamiento estadístico de datos.

DATOS OBTENIDOS POR CATEGORIZACIÓN

Otra manera de evaluar las variables y registrar los datos consiste en definir categorías en función de determinadas condiciones o atributos –numéricos o de cualidad– de la unidad en la que se manifieste el fenómeno.

—

Las categorías se deben definir de manera tal que, para la variable, cada situación pueda ser incluida siempre en una de ellas y que la ubicación en una no permita su ubicación en otra: las categorías deben ser **exhaustivas** y **excluyentes**.

En lo que respecta a la variable, pueden distinguirse categorizaciones ordinales y nominales según si esas categorías representan una graduación o no.

Datos ordinales

En esta situación, las categorías establecidas representan una graduación u ordenamiento en lo que a la variable se refiere. Considérese como ejemplo la variable estatura, que más arriba se indicó que podría describirse a través de un dato numérico. Podrían definirse categorías, como "estatura baja", "estatura media baja", "estatura media elevada" y "estatura elevada". Los criterios para definirlas pueden surgir de diversas formas: cantidad mínima y máxima de centímetros de longitud cabeza-pie, superar determinadas marcas en una pared u otras.

Puede verse que la ubicación en una categoría significa establecer una situación de comparación de orden o grado respecto de la ubicación en otra. Las unidades experimentales ubicadas en la categoría "estatura baja" tienen menor estatura que las ubicadas en la de "estatura alta".

Es frecuente asignar letras o números a las categorías definidas. Así, en la evaluación de ciertas condiciones patológicas se establecen categorías que indican el grado de enfermedad y se las numera de 0 o 1 en adelante. Por ejemplo, si se observa ausencia de inflamación, se establece un valor 0; si se detecta una ligera inflamación con algún cambio de color y escasa tumefacción, sería 1, y así sucesivamente.

Los valores numéricos obtenidos de esta manera se denominan, en ocasiones, con el nombre de **puntajes** o **grados**. Si bien en estos casos se utilizan números, debe tenerse presente que estos son solo una forma de identificar una categoría y no son datos numéricos. Esta diferencia es sustancial, ya que en los datos numéricos un valor doble indica el doble en la variable (dos metros de longitud es el doble de un metro de longitud), mientras que en los datos ordinales no es así. Tener una inflamación de grado 2 significa tener una mayor inflamación que la que se presenta con un grado 1, pero no necesariamente el doble.

Esta situación también indica que con los datos ordinales no se debe, en principio, hacer operaciones matemáticas que sí es posible hacer con los datos numéricos. Como ejemplo, véase que el desempeño de un alumno en un curso se estima usualmente con un puntaje, por lo general, en una escala de 0 a 10. Este puntaje es un dato ordinal que indica que el alumno que obtuvo calificación 8 "sabe más" que aquel que obtuvo calificación 4, pero no necesariamente el doble. Asimismo, si se juntan o suman los aprendizajes de dos alumnos que obtuvieron 4, no necesariamente se obtiene el aprendizaje del que obtuvo un 8.

También es posible establecer un ordenamiento en la totalidad de los integrantes de un conjunto. Por ejemplo, ordenar a cada uno de los individuos de un grupo en función de su estatura, del más bajo al más alto, semejante a formar una fila ordenada de menor a mayor. A partir de ello es posible asignar números a cada uno, ordenándolos de menor a mayor o de mayor a menor, de manera tal que indiquen la posición en la

serie ordenada. Este tipo de dato a veces se denomina dato de **seriación**.

Repitiendo conceptos anteriores, es de importancia reconocer si se está ante datos numéricos u ordinales, antes de proceder a su procesamiento estadístico.

Datos nominales

En este caso, las categorías que se establecen no representan graduación alguna en la variable, sino tan solo diferencias en atributos de cualidad. Por este motivo, a veces se hace referencia a estos datos como datos cualitativos.

Un ejemplo podría estar en la categorización de los integrantes de una comunidad en función de la religión que profesa cada uno de ellos: cristiano no católico, católico, judío, musulmán, otra creencia religiosa, no creyente. La ubicación en cada una de las categorías no indica un ordenamiento, sino tan sólo una condición diferente frente a la variable.

Cuando se establecen solo dos categorías, se hace referencia a la presencia de datos dicótomos. Por ejemplo: género masculino o femenino, éxito o fracaso de un tratamiento, sano o enfermo. En estos casos de situaciones dicotómicas los datos se consideran nominales, aunque se pueda pensar que, por ejemplo, el sano tiene mejor salud que el enfermo. Dicho de otra manera, para poder definir datos ordinales deben conformarse, por lo menos, tres categorías.

EXACTITUD, SENSIBILIDAD, CONFIABILIDAD Y VALIDEZ DE LOS DATOS

La aplicación de un procedimiento estadístico presupone que los datos describen de forma satisfactoria la variable de interés. En la bibliografía sobre técnicas de investigación puede encontrarse información pertinente sobre las condiciones que deben reunir los datos para cumplir con ese requisito.

Como indicación general, solo se hará aquí mención a algunas de esas consideraciones.

Un dato debe ser exacto en el sentido de registrar la variable tal como es. En una situación real, un dato representa la valoración de la variable con el agregado del error que se comete al registrarlo. Este error puede surgir de la falta de calibración del instrumento utilizado (aparato) o del usuario del instrumento. Por este motivo, los aparatos y los encargados del registro de los datos deben ser adecuadamente "calibrados" antes de comenzar con la tarea de registro.

Los datos deben tener una adecuada **sensibilidad, esto significa que puedan distinguir** los hechos que resultan de interés para diferenciar. Por ejemplo, si para evaluar la masa corporal de los integrantes de un grupo de seres humanos se utiliza la balanza que se emplea en las carreteras para pesar camiones, seguramente no se podrán establecer las diferencias entre esas personas, ya que el instrumento es sensible para registrar pesos cercanos a media o a una tonelada. De la misma manera, la balanza con la que es posible pesar a esas personas no cuenta con la sensibilidad suficiente para registrar la cantidad de fármaco presente en la cápsula de un medicamento.

Nótese que los datos numéricos permiten obtener una mayor sensibilidad que los que se obtienen agrupando en categorías. Esto es así porque en una misma categoría pueden estar incluidas situaciones (individuos) que, en realidad, son distintas. Por ejemplo, al indicar la categoría "estatura elevada" pueden incluirse en ella individuos que

no necesariamente tienen igual estatura. Un dato numérico obtenido por relación sí permitiría diferenciarlos.

Por otro lado, un ordenamiento en seriación permitiría la diferenciación, pero no la cuantificación de esa diferencia. Por ejemplo, podría diferenciarse al más alto del segundo en una serie ordenada de estaturas, pero no se tendría información de cuál es la diferencia entre ellos.

Por estos motivos se prefiere, siempre que sea posible, evaluar las variables mediante datos numéricos.

Por otro lado, los datos se deben registrar de manera tal que su **confiabilidad** esté asegurada. Este concepto permite repetir el resultado del registro cuando una misma situación para una variable es evaluada de manera repetida. La presencia de confiabilidad da lugar a la obtención del mismo dato; es decir, el mismo número o la ubicación en la misma categoría, según el tipo de dato del que se trate en cada una de las veces en las que valore el mismo atributo variable en la misma unidad. Nuevamente, es necesario preparar de modo adecuado a los instrumentos y a sus usuarios para evitar la ausencia de confiabilidad, lo cual lleva al error en los datos obtenidos.

Por último –o quizás en primer lugar– los datos deben tener **validez**. Esta condición se refiere al grado en que el dato valora el fenómeno en el que está centrado el interés del investigador. Si valora un atributo diferente del que se refiere la variable definida, el dato no es considerado válido.

Por ejemplo, si la variable de interés estuviera representada por la estatura de un sujeto experimental, un dato como el que se ha mencionado, y que es difícil de cuestionar en cuanto a su validez, es el obtenido a partir de la valoración de la distancia en centímetros entre la cabeza y los pies del sujeto en posición erguida. Si en un estudio sobre la misma variable se utilizara una balanza para registrar la masa corporal en kilogramos, se estaría frente a un dato no válido para la finalidad buscada.

No siempre la validez de un dato o su ausencia surgen con tanta claridad como en el ejemplo. Cuando las variables en juego son atributos, como "simpatía", "capacidad diagnóstica", "angustia frente a una enfermedad", no resulta tan fácil encontrar una forma de dato con validez incuestionable.

No se debe iniciar la aplicación de un procesamiento estadístico a datos sin considerar si cumplen con estos requisitos necesarios.

El procesamiento estadístico adecuado aplicado a datos inadecuados lleva a conclusiones cuestionables o inaceptables.

—

A lo largo de este texto se partirá de la suposición de que los datos con los que se trabaja reúnen las condiciones exigibles.

SÍNTESIS CONCEPTUAL

- **Un dato valora un atributo de una unidad** en una situación espacial y temporal determinada.
- **Los datos que permiten ser procesados** estadísticamente son numéricos o de categorización.
- **Los datos de categorización pueden ser** ordinales o nominales, según si las categorías representan un ordenamiento o no para el atributo variable.
- **La técnica de procesamiento estadístico** debe estar acorde con el tipo de dato que se debe procesar.
- **No se debe iniciar la aplicación de un** procesamiento estadístico a datos sin considerar si se cumple con los requisitos de validez, sensibilidad, exactitud y confiabilidad.

EJEMPLO 2-1

En las siguientes situaciones se presentan datos con los que se ha tratado de describir el estado para una variable en una unidad experimental. En cada caso se indica qué tipo de dato ha sido seleccionado.

a) El número de sesiones de radioterapia necesario para producir la remisión de un tumor: numérico discreto.
b) El tiempo, redondeado en días, trascurrido desde el inicio de un tratamiento hasta la desaparición del síntoma: numérico continuo.
c) Etapas de la evolución de un cáncer, como I, II, III o IV: ordinal.
d) Diagnóstico del estado psicológico patológico, como psicosis, neurosis, psicopatía, no determinado: nominal.
e) Disminución de la presión arterial sistólica o no luego de la administración de un fármaco: nominal dicótomo.
f) Presión diastólica en mm Hg: numérico continuo.
g) Calidad de la atención recibida durante la internación en una escala de siete puntos: ordinal.

CAPÍTULO

3

ALMACENAMIENTO Y RECUPERACIÓN DE LOS DATOS

INTRODUCCIÓN

Los datos, los cuales se ha resuelto emplear para describir las variables de interés, se recolectan con procedimientos que aseguren su exactitud y confiabilidad. Todos esos datos deben almacenarse en un soporte que permita su recuperación para el análisis y el procesamiento estadístico.

Los datos se pueden almacenar, inicialmente, en un soporte de papel (anotados en planillas).

Sin embargo, resulta conveniente que esos datos sean finalmente ingresados o "cargados" en soportes informáticos, como bancos de datos y planillas de cálculos, que permiten no solo almacenarlos, sino también procesarlos.

BANCO DE DATOS

Así como se denomina banco a una institución en la cual se depositan dinero o valores, del mismo modo se designa con el nombre de banco de datos a un "depósito" de datos en forma ordenada y que permita su fácil recuperación. Con frecuencia se utiliza la denominación "base de datos" con el mismo significado.

En este tipo de sistemas se reconocen **campos y registros, dentro de los cuales** se almacenan los datos. Un campo representa una variable que puede evaluarse en un individuo o unidad experimental. En el banco de datos de los alumnos de una institución educativa, los campos podrían estar representados por: apellido y nombres, domicilio, edad, calificaciones, entre otros. De manera similar, es fácil imaginar los posibles campos en un banco de datos

de pacientes de un hospital: datos filiatorios, estado actual, tratamiento recibido, situación de pago, etcétera.

En una investigación, los campos podrían representar las variables que se tienen en cuenta: dosis de un medicamento, cantidad de una sustancia en sangre, resultado de un tratamiento, entre otras.

Los registros, por otro lado, corresponden a cada individuo o elemento sobre el cual se registra el dato que evalúa la variable identificada en cada campo. Así, cada registro representa a un alumno en el caso de la institución educativa, a un paciente en el caso del hospital y a una unidad experimental (paciente, animal de laboratorio, tubo de ensayo, probeta, etc.) en una investigación.

La carga de los datos consiste en insertar, para cada registro, la valorización correspondiente a cada campo; es decir, a cada variable. En el caso de los datos para procesamiento estadístico, dicha valoración puede realizarse en cualquier forma o tipo de datos analizados en el capítulo anterior.

Existen diversos programas informáticos o ***softwares*** que permiten construir bancos de datos de estas características y recuperar la información cuando y como se la necesite. Así, puede recuperarse la información sobre los datos correspondientes a un determinado registro, los datos de un alumno o un paciente, o los valores que cumplen requisitos específicos en un determinado campo, pacientes con una enfermedad específica o alumnos con determinadas calificaciones.

Si bien estos programas también pueden utilizarse para realizar algunos procedimientos de análisis, como suma de valores o algún otro cálculo similar, para esta finalidad se utilizan con mayor frecuencia las planillas de cálculo.

PLANILLA DE CÁLCULOS

Una planilla de cálculos es una tabla con columnas y filas, en cuyas intersecciones –denominadas celdas– se ingresa la información en forma de datos de alguna naturaleza.

—

En el caso de los programas informáticos, es habitual que las columnas se identifiquen con letras y las filas, con números, como se muestra en el **cuadro 3-1**. Cada celda se puede identificar con una letra y un número; estos indicarán, respectivamente, la columna y la fila a la que pertenecen.

Los programas incluidos en los paquetes utilitarios más comunes (***Excel*** en el paquete ***Office*** de Microsoft®, por ejemplo) permiten trabajar con más de un centenar de columnas y decenas de miles de registros, lo que significa la posibilidad de ingresar una cantidad muy grande de datos.

Las planillas de cálculos también permiten procesar los datos ingresados, realizar diversas operaciones matemáticas y aplicar muchos de los procedimientos estadísticos que se describirán en los siguientes capítulos.

Además de estos programas genéricos, existen otros que, a partir de un formato inicial similar, permiten aplicar una mayor cantidad de procedimientos estadísticos y de mayor complejidad que los que aquí se analizan. En el último capítulo se hará referencia a algunos de ellos. La mayor parte de ellos permiten identificar a las columnas no solo con letras, sino también con palabras o abreviaturas que pueden estar asociadas con la denominación de las variables, y así poder identificar el significado de los datos ingresados con facilidad.

CUADRO 3-1. FORMA DE PRESENTACIÓN DE LA HOJA DE UNA PLANILLA DE CÁLCULO

	A	B	C	D	E	F	G	H	I
1									
2									
3									
4									
5									
6									
7									

Así como los bancos de datos complejos frecuentemente son diseñados por profesionales de la informática, el trabajo con una planilla de cálculos puede ser organizado por cualquier investigador o profesional que necesite almacenar y procesar datos.

Para ello, una vez abierta (en la pantalla de una computadora) la "hoja" de una planilla de cálculos, la primera fila (la número 1) estará destinada a incluir, en cada columna, la identificación de cada una de las variables de las cuales se almacenarán datos. Es decir, que cada columna será el equivalente a un campo de un banco de datos.

Cada fila subsiguiente (número 2 en adelante) se utilizará para ubicar los datos obtenidos en cada registro, individuo o unidad experimental, en la celda de la columna que corresponda a la variable evaluada.

Algunas consideraciones generales pueden hacerse sobre estos procedimientos. En primer lugar, la identificación de la variable se puede hacer con su descripción completa. Por ejemplo, podría escribirse "Presión arterial sistólica", "Resultado de la administración del medicamento", "Calificación obtenida en el examen" u otras similares. No obstante, cuando se prevé realizar procedimientos estadísticos es conveniente no emplear más de ocho caracteres para esa identificación. Esto ocurre porque, en algún momento, puede ser necesario "exportar" los datos a otros programas que tienen esa restricción. Por motivos similares conviene evitar el uso de espacios en blanco, guiones o símbolos en esa identificación, puesto que pueden significar órdenes determinadas para algunos programas informáticos. En los ejemplos que se incluyen un poco más adelante se podrá apreciar cómo se tienen en cuenta estas recomendaciones.

La organización de la planilla se puede realizar de dos maneras. Una se presenta en el ejemplo del **cuadro 3-2** y es aplicable cuando se registran datos sobre una sola variable, aunque esa variable pueda evaluarse en dos o más circunstancias o en condiciones distintas. Por ejemplo, el resultado de la administración de diversos medicamentos sobre la presión arterial sistólica, o la opinión sobre la calidad de la atención de la salud en cada uno de los diversos centros hospitalarios.

En estos casos, en cada columna se ingresan los datos obtenidos en individuos o unidades experimentales que hayan sido incluidos en cada una de esas condiciones de evaluación de la variable. En las situaciones descritas, en una misma columna se deberían ubicar los datos obtenidos de pacientes o animales de laboratorio que recibieron un mismo medicamento o de pacientes que recibieron atención en una misma unidad hospitalaria.

En cambio, cuando por cada individuo o unidad experimental se obtienen datos para más de una variable (p. ej., edad, género, enfermedad, tratamiento administrado, dosis, resultado obtenido, etc.), resulta conveniente, y aun necesario para el ulterior procesamiento, emplear el esquema del **cuadro 3-3**.

CUADRO 3-2. ORGANIZACIÓN DEL ALMACENAMIENTO DE DATOS CORRESPONDIENTE A VALORES OBTENIDOS CON LA ADMINISTRACIÓN DE DISTINTOS MEDICAMENTOS

	A	B	C	D	E	F
1	**MED_A**	**MED_B**	**MED_C**	**MED_D**	**MED_E**	**MED_F**
2	14	50	23	16	35	24
3	13	48	22	17	34	25
4	18	47	21	14	33	24
5	16	45	25	18	37	24
6	13		27	14	32	
7			24	16		

MED_X, medicamento X.

CUADRO 3-3. ORGANIZACIÓN DEL ALMACENAMIENTO DE DATOS CORRESPONDIENTE A DIVERSAS VARIABLES EVALUADAS EN CADA UNIDAD EXPERIMENTAL

	A	B	C	D	E	F	G
1	Trat.	Sexo	Edad	Dolor	IND_A	Fieb.	Sist.
2	Cir.	M	45	0	4	SÍ	130
3	Cir.	M	42	0	4	SÍ	135
4	Med.	F	48	0	2	NO	120
5	Cir.	F	51	1	3	NO	140
6	Med.	F	40	1	3	SÍ	120
7	Cir.	M	47	0	4	SÍ	130
8	Med.	M	47	1	2	NO	150
9	Cir.	F	45	0	3	SÍ	130
10	Med.	M	41	1	2	NO	140
11	Med.	M	46	1	2	NO	140
12	Cir.	M	48	1	3	SÍ	130
13	Cir.	F	49	0	4	SÍ	120
14	Med.	F	50	0	1	NO	140

Trat., tratamiento aplicado; Cir., cirugía; Med., medicación; Edad, años desde el último cumpleaños; Sexo: M, masculino / F, femenino; Dolor: 0, ausencia / 1, presencia; IND_A, índice utilizado para evaluar la evolución; Fieb., fiebre; Sist., presión sanguínea sistólica en mm Hg.

En este cuadro, cada columna se reserva para cada una de las variables incluidas, y cada fila para incluir los datos obtenidos de cada individuo o unidad. Así, una vez cargados los datos es posible recorrer la tabla por fila para visualizar todo lo relativo a un registro (paciente, tubo de ensayo, animal de laboratorio, etc.), o por columna para visualizar qué es lo que se registró para una determinada variable en cada uno de los registros.

Los programas que utilizan planillas de cálculos permiten incluir números o caracteres alfanuméricos, letras y números en cada celda. Cuando se trata de datos numéricos, obviamente deben ingresarse números para luego poder procesarlos. Cuando se trata de datos ordinales o nominales es posible incluir letras; por ejemplo: sí, no; nulo, leve, moderado, grave; masculino, femenino. Sin embargo, si se prevé "exportar" los datos a algún programa de procesamiento estadístico, debe tenerse presente que algunos de ellos requieren números en las celdas para el procesamiento. Esto significa que será necesario establecer alguna codificación numérica para representar al dato ordinal o nominal obtenido. Así, podrá resolverse considerar "0" a la ausencia de dolor y "1" a su presencia; "1" al ciudadano nativo, "2" al naturalizado" y "3" al extranjero. Debe entenderse que esto representa solo una codificación y no la cuantificación de un dato nominal.

Una vez finalizado el ingreso de los datos es útil realizar alguna verificación que permita detectar errores cometidos en la tarea, por lo menos los más relevantes.

—

Por ejemplo, cuando se cargan valores de edad en años de seres humanos, es posible observar los valores más altos y más bajos. Si aparece un valor de 376, es fácil deducir que es consecuencia de un error de carga, lo mismo sucede si se detecta la presencia de un valor negativo. De la misma manera, si se detecta un valor "3" para una variable en la que se codificó "1" = género femenino y "2" = género masculino, quedará resaltada la presencia de un error de carga.

Esta tarea de control es fácil de hacer con los programas informáticos que utilizan planillas de cálculos y se debe tomar como una rutina antes del procesamiento de los datos, especialmente cuando el volumen de la información (la cantidad de datos) es muy grande.

DATOS ESTADÍSTICOS

Los datos se obtienen a partir de cada uno de los individuos o unidades experimentales que son parte de una población. Las planillas de cálculos y los programas de estadística permiten procesar de diversa forma los datos cargados.

Dentro de esas formas se destaca la obtención de valores (números), que se conocen como **datos estadísticos** y sirven para resumir el conjunto de datos.

Los datos estadísticos permiten expresar cómo se manifiesta un atributo –una variable– en un conjunto de individuos a partir de los datos individuales registrados para cada uno de ellos.

El valor obtenido a partir de los datos individuales de todos los integrantes de una población es el parámetro para una variable determinada.

—

Un parámetro es, por lo tanto, un valor; en última instancia, un dato estadístico que describe el comportamiento de una variable no en un individuo o unidad experimental, sino en la totalidad de individuos o unidades experimentales que constituyen una población.

De esto surge que el objetivo de una investigación es obtener un parámetro que valore la situación de una población para una variable específica; por ejemplo, el estado de su salud, el nivel educativo, etcétera.

Se habrá notado que, para obtener el valor de un parámetro, se debe disponer de un banco de datos o de una planilla de cálculos en donde estén incluidos la totalidad de los registros correspondientes a la población. Esta situación no es usual, sino, por el contrario, prácticamente inexistente debido al tamaño de las poblaciones de interés científico.

Por lo tanto, los cálculos que usualmente se realizan culminan con la obtención del resumen de solo una parte de los datos de la población: los de una muestra tomada de ella. Ese resumen no es un parámetro, sino tan solo un valor que lo estima. Frecuentemente se utiliza el término estadístico para hacer referencia a un valor que describe el comportamiento de una variable en una muestra y que, en consecuencia, es una estimación del correspondiente parámetro.

En los próximos capítulos se introducirán las técnicas de obtención de resúmenes descriptivos de datos. Cuando se trata de resúmenes numéricos, esas técnicas llevan a la obtención de parámetros o estadísticos, según se procese la totalidad de los datos de una población o una parte de estos.

Posteriormente, y en capítulos subsiguientes, se introducirán los principios de procesamiento de datos de las muestras y se presentarán algunas técnicas que, mediante cálculos estadísticos, permiten hacer inferencias respecto de los respectivos parámetros.

SÍNTESIS CONCEPTUAL

- **Los datos obtenidos a partir de la valo**ración de variables se ingresan en bancos de datos y planillas de cálculos.
- **A**ntes de iniciar el procesamiento estadístico es útil realizar alguna verificación que permita detectar errores cometidos durante el ingreso de los datos.
- **Un primer resultado del procesamien**to estadístico es la obtención de lo que se conoce como datos estadísticos, que permiten expresar cómo se manifiesta una variable en un conjunto de individuos a partir de los datos individuales registrados para cada uno de ellos.
- **Cuando se han procesado todos los da**tos de una población, el dato estadístico obtenido es un parámetro.
- **Cuando se han procesado solo los da**tos de una muestra, se obtiene un dato estadístico a partir del cual se pueden aplicar técnicas para hacer inferencias sobre el respectivo parámetro.

CAPÍTULO

4

RESUMEN DE DATOS NOMINALES

INTRODUCCIÓN

Cuando la recolección y el almacenamiento de datos nominales se ha completado, el análisis del conjunto, población o muestra puede iniciarse al contar la cantidad de registros, individuos o unidades experimentales que se encuentran incluidos en cada categoría.

Supóngase, como ejemplo, que se han evaluado 1200 individuos (registrados en el correspondiente banco de datos) y se ha ubicado a cada uno de ellos en una de dos categorías (dato dicótomo): "sano" o "enfermo". Recuérdese que en el procesamiento estadístico se presupone que esa categorización se ha realizado de manera tal que el dato obtenido esté razonablemente libre de error, sea válido y confiable, y que la sensibilidad sea suficiente para los objetivos planteados.

El paso inicial para llegar a describir un conjunto de datos nominales es contar cuántos de esos individuos se encuentran en cada una de las categorías.

La tarea que conduce a establecer la frecuencia con la que aparecen los datos en cada categoría es muy rápida y sencilla cuando estos se encuentran en bancos de datos o planillas de cálculos que permiten realizar el conteo mediante funciones preestablecidas en el programa.

Como resultado de esa labor, podría conocerse que el conjunto incluye a 300 enfermos y 900 sanos; esto constituye ya una primera información que permite obtener una imagen del conjunto que se evalúa. Puede decirse que los enfermos aparecen con una "frecuencia" de 300 y los sanos con una de 900.

PRESENTACIÓN EN GRÁFICOS

La información obtenida a través del conteo (p. ej., 300 enfermos y 900 sanos) puede presentarse en forma de gráfico para facilitar su interpretación. En la **figura 4-1** se muestran tres gráficos obtenidos a partir de esos datos. Los dos primeros son de columnas o de barras, aunque algunos programas informáticos reservan esta última denominación cuando la orientación es

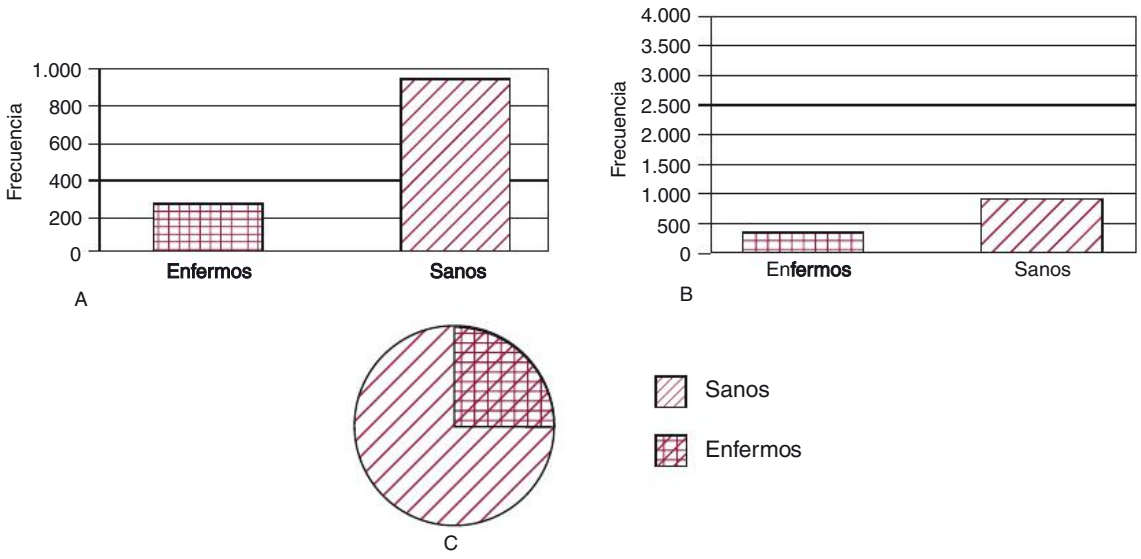

Fig. 4-1. Frecuencias de individuos sanos y enfermos.

horizontal en vez de vertical. Estos gráficos se utilizan para presentar las frecuencias en que aparecen los datos en las distintas categorías dentro del conjunto, población o muestra. La relación de la altura de las columnas o la longitud de las barras brinda una idea de la relación entre las frecuencias para cada categoría.

Si bien esta imagen es útil, debe tenerse en cuenta que puede inducir a errores de apreciación. Efectivamente, los gráficos A y B de la figura fueron construidos con los mismos datos. Sin embargo, la escala utilizada en el eje vertical en cada uno de ellos genera una imagen de mayor diferencia de frecuencia en A que en B. En A parece haber una gran diferencia entre la cantidad de enfermos y sanos, mientras que en B el contraste parece ser menor. Cuando se analizan datos en gráficos de esta naturaleza siempre debe prestarse atención a los valores de la escala utilizada para evitar extraer conclusiones apresuradas.

El gráfico C, de sectores circulares o "gráfico torta", resulta más "aséptico" en la presentación de la información. No influye el tamaño del círculo, porque siempre la relación entre los sectores que representan la frecuencia para cada categoría será la misma.

La presentación de la información que resume los datos nominales en forma de gráficos es aceptable y útil, pero está limitada por la obtención de una imagen algo subjetiva y no permite una elaboración matemática posterior para comparar con exactitud conjuntos distintos. Mucho menos pueden realizarse con ellos inferencias sobre las poblaciones de las cuales se obtuvieron los datos, cuando se trata de muestras.

RAZONES Y PROPORCIONES

Para permitir un análisis más acabado del resumen de un conjunto de datos nominales, y posteriormente encarar la tarea inferencial, se hace necesario resumir los datos en forma numérica, o sea, generar un dato estadístico que permita el análisis de la variable en el conjunto.

Una manera de hacerlo, especialmente cuando se trata de datos dicótomos, es establecer la relación entre las frecuencias de los datos en cada categoría. Esto significa dividir la cantidad de datos presentes en una categoría por la cantidad de datos presentes en la otra.

En el ejemplo, si se quisiera describir la situación en cuanto a la presencia de enfermedad, esto podría traducirse en la división del valor 300, frecuencia de enfermos, por el valor 900, frecuencia de sanos. El resultado, un tercio o 0,33, es la **razón** entre ambas categorías y permite obtener información sobre la presencia de enfermedad –el numerador de la razón– en ese conjunto. En palabras puede expresarse diciendo que: "En el conjunto evaluado existe un enfermo por cada tres sanos".

Nótese que de esta manera es posible no solo apreciar la situación, sino compararla con otras similares. Así, si en otro conjunto la razón enfermos/sanos fuera 0,50 (un enfermo por cada dos sanos), sería posible visualizar que en el primero existe un menor nivel de presencia de enfermos.

El cálculo de razones se utiliza, aunque con mayor asiduidad, en especial cuando se trabaja con más de dos categorías; para resumir la situación de un conjunto de datos nominales es calculada la **proporción** correspondiente a esas diferentes categorías.

Para el cálculo de una proporción se relaciona (divide) la frecuencia correspondiente a una determinada categoría –la que corresponde a la expresión de la variable sobre la que se quiere generar información– por el total de datos integrantes del conjunto. Así, en el ejemplo anterior la proporción sería el resultado de dividir 300 (frecuencia de enfermos) por 1200 (total del conjunto); es decir, 0,25.

En el cálculo de una proporción, los datos que se incluyen en el numerador –enfermos en el ejemplo– están incluidos en el denominador, ya que es el total; esto no sucede en el cálculo de una razón.

—

La **figura 4-2** muestra los símbolos y ecuaciones o fórmulas que se emplean para calcular las proporciones en el caso de datos nominales obtenidos de poblaciones y de muestras. Obsérvese que, si bien la manera de realizar el cálculo es igual, el significado del resultado es distinto. Cuando se trata de poblaciones, se obtiene un parámetro; mientras que, cuando el conjunto considerado es una muestra, se obtiene un dato estadístico que permite su estimación. Por este motivo, los símbolos utilizados son diferentes.

En el ejemplo de los párrafos anteriores, la proporción constituye un resumen de los datos y puede interpretarse en palabras al indicar que existe 0,25 (o sea, 1/4) de enfermos por cada integrante del conjunto. Esto es así desde el punto de vista matemático y la proporción es el valor utilizado para el

Población

$$p = \frac{f(x)}{N}$$

Muestra

$$\hat{p} = \frac{f(x)}{n}$$

Donde:

- p : **proporción en una población**
- N : **tamaño de la población**
- $f(x)$: **frecuencia en una categoría**
- $\hat{p}$: **proporción en una muestra**
- n : **tamaño de la muestra**

Fig. 4-2. Fórmulas para el cálculo de proporciones para la descripción de conjuntos de datos nominales.

procesamiento estadístico, por ejemplo, al realizar inferencias. Sin embargo, su apreciación resulta dificultosa, ya que es problemático imaginar el significado de "un cuarto de enfermo".

Para obviar esto último y facilitar la transmisión y comprensión de la situación en el conjunto de datos, es habitual multiplicar la proporción por un valor que la transforme en un número entero. El valor utilizado con mayor frecuencia es 100, y en el ejemplo esto significa multiplicar 0,25 por ese valor. El valor o **porcentaje** resultante (25%) indica que en el conjunto existen 25 enfermos por cada 100 individuos. Esto resulta más comprensible y permite una fácil comparación entre distintos conjuntos; si hay 50% de enfermos, hay más enfermos que en el conjunto del ejemplo.

Si bien el valor 100 es el más utilizado, cuando la frecuencia en una categoría es muy baja puede utilizarse una constante mayor. Por ejemplo, la tasa de mortalidad (frecuencia en la categoría "muerte") se expresa generalmente en un valor por mil (p. ej., 5‰).

Algunas prevenciones deben tomarse al pretender extraer conclusiones a partir de la observación de proporciones y los porcentajes correspondientes. En primer lugar, el análisis del valor respectivo debe hacerse teniendo en cuenta cuáles fueron los datos a partir de los cuales se los calculó. Así, por ejemplo, una tasa (proporción o porcentaje referido a un momento o período determinado) mayor de mortalidad en un grupo de individuos respecto de otro puede indicar situaciones distintas, puede significar una mayor cantidad de enfermedad o mayor edad; solo con información adicional sobre los individuos se pueden extraer conclusiones. Estas consideraciones muestran, desde ya, que la estadística genera números, pero que las conclusiones a las que se arriba a partir de ellos no son siempre directas.

Otro aspecto para tener en cuenta es que, ante la presencia de un porcentaje, siempre debe tomarse la precaución de evaluar la cantidad total sobre la que fue calculado. Cuando el porcentaje es obtenido a partir de un número reducido de datos puede dar una imagen sesgada de la realidad.

Un ejemplo de lo anterior es el informe de un autor sobre la situación en un conjunto de individuos que habían sido evaluados con datos nominales. El resultado indicaba que, en el conjunto, un 2% de los varones estaba casado con el 50% de las mujeres. La primera imagen que podríamos generar en nuestra mente a partir de estos datos cambia por completo si nos enteramos de que en el conjunto había 50 varones y 2 mujeres, y que uno de los varones, el 2%, estaba casado con una de las mujeres, el 50%.

Puede concluirse que solo tiene sentido calcular porcentajes cuando el conjunto de datos tiene un tamaño considerable, por ejemplo, de más de 100 datos.

Una proporción, o su expresión en porcentaje, puede ser indicativa de diversas situaciones según sea el origen de los valores que se hayan empleado para su cálculo. Es la forma habitual de indicar **probabilidad** de ocurrencia de un evento, ya que se lo calcula al relacionar una forma en que se produce ese evento con la cantidad de formas en que podría producirse. Así, la probabilidad de que al arrojar una moneda esta caiga con una de sus caras expuesta es 0,5 o 50%, valor que surge al relacionar esa forma con las dos posibles formas en que podría ocurrir el evento (1/2 = 0,5).

Prevalencia e incidencia

Los valores estadísticos generados a partir de datos nominales se emplean en las ciencias de la salud para describir diferentes situaciones.

Dos porcentajes, cuyos usos son muy habituales para evaluar la situación en cuanto a un estado patológico, son las tasas de prevalencia y de incidencia.

—

Como proporciones expresadas en porcentajes, ambas tasas se calculan al dividir la frecuencia de datos en una categoría por la cantidad total de datos y, por lo general, al multiplicar la proporción así obtenida por un valor constante, generalmente 100.

La diferencia entre ambas tasas radica en cuáles son los datos que se toman en cuenta para obtener la frecuencia. En la tasa de prevalencia se cuenta la cantidad de datos en la categoría en un momento determinado, mientras que en la tasa de incidencia se cuenta la cantidad de datos que aparecieron en la categoría durante un lapso determinado; por ejemplo, un año.

Esto significa que en la tasa de incidencia no se tienen en cuenta los datos existentes en la categoría desarrollados en períodos anteriores. La situación puede determinar que en el caso de enfermedades crónicas (el paciente no se cura ni se muere) la tasa de prevalencia aumente a pesar de que a partir de medidas preventivas se logre disminuir la tasa de incidencia.

Proporciones o porcentajes para la valoración de pruebas diagnósticas

Cuando se quiere establecer la utilidad de un procedimiento para detectar la presencia de una situación determinada (enfermedad, potencial de fracaso); es decir, evaluar las posibilidades de una prueba diagnóstica, lo que se hace es comparar el resultado de su aplicación con lo que muestra la situación realmente existente.

Esto último presupone que existe alguna forma incuestionable, o por lo menos aceptada como válida, para detectar esa situación. Es habitual denominar a esta forma **prueba de referencia o patrón de oro.** Por ejemplo, en la evaluación de una prueba que pretende diagnosticar la presencia de un tumor maligno, podría aceptarse como prueba de referencia el diagnóstico al que se ha llegado a partir del estudio anatomopatológico de una biopsia.

En definitiva, el procedimiento experimental consiste en seleccionar un conjunto de individuos o unidades y separarlos en dos grupos, según tengan la situación problema o no, mediante la prueba de referencia.

Luego, en cada uno de los integrantes de ambos grupos se utiliza la prueba diagnóstica en evaluación y se registra si el resultado es positivo (presencia de la situación) o negativo (ausencia de ella). Si la prueba funciona a la perfección, es de esperar que en todos los que tengan la situación problema, según la referencia, el resultado sea positivo y que sea negativo en los restantes.

Los resultados que podrían obtenerse en una experiencia de ese tipo pueden verse en el **cuadro 4-1**, en el que se observa que la situación ideal no se ha dado. En efecto, en

CUADRO 4-1. EVALUACIÓN DE LAS PRUEBAS DIAGNÓSTICAS

	Con enfermedad	Sin enfermedad	Total
Prueba positiva	80	100	180
Prueba negativa	20	400	420
Totales	100	500	600

Sensibilidad: 80/100 = 0,80 (80%); Especificidad: 400/500 = 0,80 (80%); Valor predictivo positivo: 80/180 = 0,44 (44%); Valor predictivo negativo: 400/420 = 0,96 (96%).

algunos de los individuos con la situación (enfermedad) el resultado fue negativo (falsos negativos), mientras que en algunos sin la situación el resultado fue positivo (falsos positivos).

A partir de los datos pueden calcularse varias tasas que brindan diferente información sobre la prueba en evaluación.

Al calcular la tasa porcentual a partir de la frecuencia de resultados positivos y la cantidad total de casos con enfermedad, 80 y 100, respectivamente, se obtiene la denominada **sensibilidad**, que en este caso es del 80%. Este valor indica que al utilizar la prueba en un conjunto de individuos que tienen la situación (enfermos) se puede esperar detectar 80 de cada 100, mientras que 20 quedarán sin ser detectados y, por ende, quizá sin la indicación de tratamiento.

Por otro lado, se puede calcular la proporción o porcentaje al dividir la frecuencia de resultados negativos por la cantidad total de individuos sin la situación (no enfermos). El valor así calculado es la **especificidad; esto indica que, al aplicar la prueba** diagnóstica, 80 de cada 100 individuos sanos se detectan con esa condición, aunque 20 de cada 100 se considerarán enfermos y es muy probable que se los someta a un tratamiento innecesario.

Los valores de sensibilidad y especificidad orientan en cuanto a la selección de pruebas diagnósticas, especialmente en su aplicación a grandes grupos de individuos.

—

Así, se utilizan pruebas de alta sensibilidad para evitar dejar sin tratamiento a individuos que lo necesiten, aun con riesgo de aplicarlo innecesariamente a algunos, si su tasa de especificidad es baja.

También pueden combinarse pruebas diagnósticas, al utilizar una de alta sensibilidad al inicio para asegurar la detección de prácticamente la totalidad de enfermos y luego emplear, en los así detectados, una prueba de alta especificidad para confirmar el diagnóstico y evitar la aplicación innecesaria de tratamiento.

Cuando una prueba diagnóstica se aplica a un individuo en particular, se obtiene una mayor información sobre sus posibilidades a partir de otros valores de proporciones o porcentajes.

Al utilizar como numerador la frecuencia de resultados positivos verdaderos y como denominador la cantidad total de positivos (80 y 180, respectivamente, en el ejemplo) se puede calcular el **valor predictivo positivo, que es del 44% en este caso. Esto** indica que solo 44 de cada 100 veces que se obtiene un resultado positivo se está frente a un individuo realmente enfermo. Desde este punto de vista, la detección de un caso positivo no da, con esta prueba hipotética, ninguna confianza diagnóstica.

Sin embargo, y en el mismo ejemplo, el valor calculado a partir de la frecuencia de negativos verdaderos y la cantidad total de negativos (400 y 420) es del 96% y constituye el denominado **valor predictivo negativo**. Esto indica que la detección de un caso negativo permite aseverar con bastante confianza que se está frente a la ausencia de enfermedad.

Puede visualizarse que la selección de una determinada prueba diagnóstica debe realizarse en función de estos valores, a fin de aplicar la más conveniente a una situación en particular.

Téngase presente que los valores de evaluación de una prueba diagnóstica, calculados a partir de los datos obtenidos de una muestra, no se deben tomar como parámetros que describen su comportamiento real, sino como parámetros estadísticos que la estiman. Con ellos, deben aplicarse los procedimientos de estadística inferencial para extraer conclusiones aplicables a la respectiva población.

VALORACIÓN DEL RIESGO

Las proporciones y razones permiten evaluar el riesgo que representa una determinada condición para que aparezca un hecho definido y generalmente no deseado.

—

En los aspectos más frecuentes de las ciencias de la salud, esto significa evaluar si la presencia de una situación o un factor determinado, como el hábito de fumar o ejercer una determinada profesión, significa una posibilidad definida de desarrollar una afección específica, por ejemplo, enfermedad pulmonar o alteraciones en la columna vertebral, respectivamente.

La evaluación de ese riesgo puede realizarse al comparar los hechos que se producen en conjuntos de individuos o unidades experimentales (en los que el factor está presente) respecto de los que se producen en conjuntos de individuos o unidades experimentales en donde no lo está, como fumadores y no fumadores, por ejemplo.

Los procedimientos numéricos que se emplean varían según si los datos son obtenidos a partir de diseños experimentales prospectivos (de cohorte) o retrospectivos (de caso y testigo).

Riesgo relativo

En un diseño prospectivo se conforman dos grupos de individuos, según la presencia del posible factor de riesgo o no. Ambos grupos se siguen a través del tiempo y en cada uno de sus integrantes se registra la aparición del desenlace o no, desarrollo de la enfermedad o no.

Al cabo del lapso previsto para la experiencia, se pueden haber recolectado datos como los que se muestran en el **cuadro 4-2**. A partir de ellos se puede evaluar en cada grupo el **riesgo**, la relación porcentual entre la frecuencia de enfermedad y el total de integrantes del grupo. En el ejemplo, esos valores son 20 y 10% para los grupos con factor de riesgo y sin él, respectivamente. Estos valores indican la **probabilidad** de contraer la condición indeseable en presencia o ausencia del factor de interés.

La relación entre ambas proporciones –o entre los porcentajes (40 / 20)–, que en este caso es 2, se denomina **riesgo relativo**. Un valor 1 en el riesgo relativo indica que el factor no constituye un riesgo; un valor mayor de 1, como en el ejemplo, indica que el riesgo es mayor con la presencia del factor; y un valor menor de 1 indicaría que el fac-

tor no solo no es un riesgo, sino que podría ser un factor beneficioso para disminuir la posibilidad de desarrollo de la enfermedad.

Odds ratio o razón de productos cruzados

En los diseños retrospectivos, los grupos se conforman según se haya producido el desenlace o no, presencia de enfermedad o su ausencia. Luego, se valora la exposición de los integrantes de esos grupos al factor de riesgo en el pasado.

Los datos podrían ser los del ejemplo del **cuadro 4-3**. Nótese que en este caso no se conoce el total de individuos expuestos al factor de riesgo, ya que ellos fueron seleccionados una vez producido el desenlace o no.

Por este motivo, no es posible calcular la incidencia que indica el riesgo (recuérdese que, en este caso, el denominador es la cantidad total de individuos del conjunto). En cambio, es posible calcular razones al relacionar las frecuencias de la presencia del factor de riesgo en los grupos de enfermos y no enfermos.

En el ejemplo, esa razón, que se describe como **chance** u ***odds*** en inglés, es 2 (40 / 20) y 0,89 (160 / 180) en los grupos con enfermedad y sin ella, respectivamente.

Para valorar el factor de riesgo, se establece la razón entre las dos razones, que en este caso es 2,25 (2 / 0,89) y se la designa con el nombre de **razón de chances, razón de productos cruzados** o, con mucha asiduidad, con las palabras inglesas ***odds ratio***.

Un valor mayor de 1 (2,25 en el ejemplo) indica una mayor frecuencia de individuos con el factor de riesgo en el grupo con enfermedad y, por ende, la posible contribución que este tiene en su desarrollo.

Al igual que con lo que sucede en la evaluación de pruebas diagnósticas, debe tenerse presente que si los cálculos de riesgo relativo o de ***odds ratio*** se realizan a partir de muestras, solo deben servir de base para la aplicación de la estadística inferencial en la estimación de la situación en las correspondientes poblaciones.

CUADRO 4-2. EVALUACIÓN DE LOS FACTORES DE RIESGO (DISEÑO PROSPECTIVO)

	Con enfermedad	Sin enfermedad	Total
Con factor de riesgo	40	160	200
Sin factor de riesgo	20	180	200

Riesgo con factor: 40 / 200 = 0,20 (20%); Riesgo sin factor: 20 / 200 = 0,10 (10%); Riesgo relativo: 0,20 / 0,10 = 2.

CUADRO 4-3. EVALUACIÓN DE LOS FACTORES DE RIESGO (DISEÑO RETROSPECTIVO)

	Con enfermedad	Sin enfermedad
Con factor de riesgo	40	160
Sin factor de riesgo	20	180
Total	60	340

Chance (*odds*) con enfermedad: 40 / 20 = 2; *Chance* (*odds*) sin enfermedad: 160 / 180 = 0,89; *Odds ratio*: 2 / 0,89 = 2,25.

SÍNTESIS CONCEPTUAL

- **El procesamiento descriptivo inicial de** un conjunto de datos de categorización consiste en contar cuántos de ellos corresponden a cada una de las categorías consideradas.
- **Para resumir los datos de categorización** en forma numérica se calculan razones o proporciones.
- **En las ciencias de la salud, las razones** o proporciones se usan de manera habitual para el cálculo de porcentajes a fin de establecer las tasas de prevalencia y de incidencia de una patología, así como para la evaluación de pruebas diagnósticas mediante el cálculo de porcentajes de sensibilidad, especificidad y valor predictivo.
- **Las proporciones y razones también** permiten evaluar el riesgo que representa una determinada condición para que aparezca un hecho definido y, por lo general, no deseado, mediante los valores de riesgo relativo y de ***odds ratio***.

EJEMPLO 4-1

En un grupo de 2520 reclusos de una unidad penitenciaria, se observó que 625 de ellos tenían manifestaciones de estados depresivos en el mes de enero.

Durante el período transcurrido desde esa observación y hasta diciembre del mismo año, la población de reclusos se mantuvo constante y se recibieron en el consultorio psiquiátrico de la unidad 323 consultas por nuevos casos de depresión.

Puede considerarse que la tasa de prevalencia de depresión al comenzar el período considerado era de 24,8% (625 × 100/2520), mientras que la tasa de incidencia de la enfermedad durante el período fue de 12,8% (323 × 100/2520).

Si los casos iniciales y los que se produjeron no hubieran remitido, la tasa de prevalencia al final del período sería de 37,6%; o sea, la relación porcentual entre el total de casos, los iniciales más los nuevos, y el total de la población.

EJEMPLO 4-2

Se desea analizar la utilidad de una prueba colorimétrica simplificada para evaluar la presencia o ausencia de actividad cariogénica. Se resuelve utilizar como referencia la categorización de individuos como positivos o negativos, según que el recuento de unidades formadoras de colonias microbianas (UFC) generadas a partir de muestras tomadas de su cavidad bucal supere un valor prefijado o no.

Los resultados fueron los siguientes:

	Referencia positiva	Referencia negativa	Total
Nueva prueba positiva	500	120	620
Nueva prueba negativa	100	450	550
Total	600	570	1170

A partir de estos resultados se pueden establecer las siguientes tasas porcentuales para la valoración de la prueba:

Sensibilidad: 83,3%
Especificidad: 78,9%
Valor predictivo positivo: 80,6%
Valor predictivo negativo: 81,8%

EJEMPLO 4-3

En un estudio llevado a cabo para evaluar el riesgo de aparición de xerostomía en pacientes que recibían una determinada medicación antidepresiva o no, se obtuvieron los siguientes resultados:

	Con xerostomía	Sin xerostomía
Con medicación	910	454
Sin medicación	150	550

En un diseño prospectivo, se habría comenzado con la selección de 1364 pacientes con medicación y 700 sin medicación. Los resultados permitirían calcular un riesgo relativo de 3,1, el cual permite valorar el efecto en estudio.

En un diseño retrospectivo, se habrían seleccionado 1060 individuos con xerostomía y 1004 sin xerostomía que sirvieran de grupo testigo. En este caso, se calcularía con la misma finalidad un ***odds ratio*** de 7,3.

CAPÍTULO

5

RESUMEN DE DATOS NUMÉRICOS

INTRODUCCIÓN

Al igual que en el caso de los datos nominales, una vez que se han completado la recolección y el almacenamiento de los datos numéricos se hace necesario proceder a su análisis. De esta manera, podrán realizarse estimaciones e inferencias posteriores sobre la situación en el conjunto, población o muestra en cuanto a la variable que esos datos evalúan.

Un primer análisis puede consistir en la evaluación de la forma en la que los datos están distribuidos, es decir, con qué frecuencia (en qué cantidad) aparecen en el conjunto, individuos o unidades experimentales con un determinado valor para el dato.

Los aspectos relacionados con la distribución de la frecuencia de los datos se considerarán en el capítulo siguiente. Con ello, se obtiene información de utilidad, aunque las técnicas estadísticas –al igual que en el caso de los datos nominales– requieren la obtención de resúmenes numéricos. Estos son los parámetros que permiten la descripción de una población y, en el caso de las muestras, los estadísticos a partir de los cuales se pueden realizar inferencias sobre las poblaciones de las cuales se obtuvieron esas muestras.

MEDIDAS DE TENDENCIA CENTRAL: MEDIA ARITMÉTICA, MEDIANA Y MODA

En capítulos anteriores se indicó que en los datos numéricos continuos, y desde un encuadre puramente matemático, la escala comienza en infinito negativo y se extiende hasta el infinito positivo e incluye, entre ellos, una infinita cantidad de valores.

Considérese, a modo de ejemplo, una pequeña población de cinco individuos en la cual se obtuvieron los siguientes datos numéricos continuos en la evaluación de una variable: 3; 2; 3; 1; 6.

Un resumen de los datos numéricos consiste en obtener un valor que permita establecer en qué lugar de la escala de valores posibles tiende a ubicarse el conjunto de datos en consideración, que es lo que se denomina una medida de tendencia central.

—

En este caso, ese valor es un parámetro, ya que se ha supuesto estar frente a una población. En un lenguaje menos técnico, una medida de tendencia central se conoce como el promedio de un conjunto de datos numéricos.

Una manera para obtener esta medida de tendencia central, que por ser la más común se asocia habitualmente con el cálculo de un promedio, es sumar todos los valores y dividir el resultado por la cantidad de valores, es decir, por el tamaño de la población.

El parámetro, o estadístico en el caso de una muestra, así obtenido se denomina **media aritmética y se simboliza con la** letra griega **μ** para el caso de las poblaciones o, habitualmente, $\bar{x}$ en el caso de que se haya trabajado con muestras.

En la población del ejemplo, la suma de los datos da como resultado 15 y, como el tamaño de la población es 5, el valor de la media aritmética es 3 (15/5).

En la **figura 5-1** se muestran las ecuaciones y símbolos que se utilizan para el cálculo de estos y otros parámetros y estadísticos.

Otra medida de tendencia central es la denominada **mediana**, que está dada por el valor que divide al conjunto en dos partes iguales, una vez ordenados los datos en forma ascendente o descendente según sus valores. Es decir que quedan separados igual cantidad de datos con un valor inferior y superior al valor del dato mediano.

En el ejemplo, los datos ordenados en forma ascendente quedarían así: 1, 2, 3, 3, 6. El tercero de los datos (3) es la mediana del conjunto, ya que separa a dos datos con valores superiores y a dos con valores inferiores a él. Si el conjunto tuviera una cantidad par de datos, se consideraría como valor de la mediana al promedio (media aritmética) de los datos centrales de la serie ordenada (el mayor de la mitad inferior y el menor de la superior, si el ordenamiento fuera ascendente).

Una tercera forma de obtener una medida de tendencia central es considerar como tal al valor que se repite con mayor frecuencia en el conjunto, si es que existe alguno.

Población	Muestra
$\mu = \frac{\Sigma(x)}{N}$	$\bar{x} = \frac{\Sigma(x)}{n}$
$\sigma^2 = \frac{\Sigma(x-\mu)^2}{N}$	$S^2 = \frac{\Sigma(x-\bar{x})^2}{n-1}$
$\sigma = \sqrt{\frac{\Sigma(x-\mu)^2}{N}}$	$S = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$

- μ : **Media aritmética de una población**
- Σ : **Sumatoria**
- x : **Datos (valores numéricos)**
- N : **Tamaño de la población**
- $\bar{x}$: **Media aritmética de una muestra**
- n : **Tamaño de la muestra**
- σ^2 : **Variancia de una población**
- σ : **Desviación estándar de una población**
- S^2 : **Variancia de una muestra**
- S : **Desviación estándar de una muestra**

Fig. 5-1. Fórmulas para el cálculo de la media aritmética, la variancia y la desviación estándar.

En el ejemplo, el dato "3" aparece dos veces (frecuencia 2), mientras que los restantes solo una. Por lo tanto, 3 (por su mayor frecuencia) es la medida de tendencia central conocida como **moda** para este conjunto.

En todo conjunto de datos numéricos se puede registrar un solo valor de media aritmética y uno de mediana. En cambio, puede no registrarse una moda, si no existe un dato que aparezca con una frecuencia mayor; o encontrarse varias modas, si más de un valor aparece con una misma frecuencia mayor que la del resto: la distribución de los datos puede ser **bimodal, trimodal** o **polimodal.**

Existen otras medidas de tendencia central (como la media geométrica) que en algunas situaciones específicas son de aplicación, pero que no se considerarán aquí.

De las tres analizadas, la media aritmética es la de mayor aplicación, especialmente en la estadística inferencial. Una razón para ello deriva del hecho de que su determinación se hace, en términos no precisamente matemáticos, de manera "democrática". Efectivamente, todos y cada uno de los datos integrantes del conjunto se "consultan" para obtener el valor (suma) que luego se divide por el total. Esto significa que cualquier cambio que se produzca en un dato se traduce, necesariamente, en un cambio pequeño o grande en el valor de la media aritmética.

Esta situación "democrática" no ocurre en el caso de la mediana, ya que un dato –el del medio– asume la responsabilidad de "representar" al conjunto. Los cambios en los demás valores pueden no cambiar el valor de la mediana. En el ejemplo, si el dato 6 se modificara y pasara a ser 5 o 7, la mediana seguiría siendo 3.

Por otro lado, la moda es algo más "democrática", porque es la "mayoría"–el dato con mayor frecuencia– la que asume la "representación", aunque sin que la "minoría" tenga oportunidad de "opinar". Asimismo, en este caso, los cambios en algunos datos no necesariamente hacen que cambie la moda.

Nótese que, en el ejemplo considerado, los valores de la media aritmética, mediana y moda son los mismos. No en todos los conjuntos de datos se verifica esta condición y en el próximo capítulo se analizarán algunas de sus consecuencias en la interpretación de los datos.

MEDIDAS DE DISPERSIÓN: RANGO, VARIANCIA Y DESVIACIÓN ESTÁNDAR

Un solo valor –razón, proporción– es suficiente para resumir la situación en un conjunto de datos nominales.

En el caso de los datos numéricos, las medidas de tendencia central no brindan la totalidad de la información necesaria.

—

Considérese otra población del mismo tamaño que la del apartado anterior; es decir, cinco individuos o unidades experimentales, aunque con los siguientes datos obtenidos en cada uno de ellos: 3, 3, 3, 3, 3.

En este caso, la media aritmética (15/5) es 3; la mediana (el dato "del medio" en la serie ordenada) es 3; y la moda (el dato con mayor frecuencia) es 3. Es decir, que este conjunto es igual en términos de tendencia central al anteriormente considerado.

No obstante, es fácil visualizar que ambos conjuntos no son iguales en cuanto a otra característica. En el segundo caso, no solo

es 3 la tendencia central, sino que todos los datos son 3; es decir, que no existe ninguna **dispersión** o **variación** entre los datos en el conjunto. En el primero, en cambio, la tendencia central es 3, aunque existen datos con valores mayores y menores de 3, lo que indica que, en este conjunto, existe una dispersión determinada.

Surge de esta observación que, al intentar describir conjuntos de datos numéricos, no es suficiente resumirlos en términos de una medida de la tendencia central.

Es necesario calcular alguna medida de dispersión o variación (parámetro o estadístico) para complementar la información que brinda la tendencia central.

Una manera sencilla y rápida de obtener información sobre la dispersión es establecer la diferencia entre los datos de mayor y menor valor. Esta medida de dispersión se conoce como **rango** o **recorrido**, y un valor 0 en él indica ausencia de dispersión.

En el primer ejemplo el rango es 5 (6 – 1), mientras que en el segundo es 0 (3 – 3).

El rango cumple con la finalidad buscada de valorar la dispersión y permite apreciarla, aunque no constituye un parámetro o estadístico que permita realizar análisis o inferencias más elaboradas. Al seguir el concepto no matemático utilizado con la comparación entre formas de evaluación de la tendencia central, puede considerarse que el rango no es "democrático". Efectivamente, para su cálculo solo se toman dos datos (el mayor y el menor) y, por lo tanto, cualquier cambio en los restantes no se registra mientras no superen en más o en menos los dos valores extremos.

Para lograr una medida "democrática" de la dispersión, el procedimiento utilizado es el que se muestra en el **cuadro 5-1**, que parte de la población del primer ejemplo de este capítulo.

En la primera columna de la tabla (encabezada con "x") se encuentran los valores de los datos. En un primer paso, se "consulta" a cada dato sobre qué "aporte" de dispersión hace, entendiéndose por ello cuánto está "desviado" o "qué variación o dispersión tiene" respecto del valor "democrático" que los representa (la media aritmética). En términos matemáticos, esto significa establecer la desviación (diferencia) de cada dato respecto de la media aritmética. Los resultados para el ejemplo se muestran en la segunda columna encabezada con $(x - \mu)$. De esta manera, se obtienen los valores: 0 para el primer dato (no está desviado respecto de la media); –1 para el segundo (está desviado una unidad hacia abajo); ……; 3 para el último (está desviado tres unidades hacia arriba).

Al tenerse ahora información sobre las desviaciones de cada dato respecto de la media aritmética, puede pensarse en calcular el promedio de desviación (o dispersión) en el conjunto de los datos. Para ello, es posible intentar sumar esas desviaciones y dividir el resultado por la cantidad de datos mediante el procedimiento habitual de obtención de un "promedio".

CUADRO 5-1. MEDIDA DE LA DISPERSIÓN EN UNA POBLACIÓN DE DATOS CON MEDIA ARITMÉTICA (μ) = 3

x	$(x - \mu)$	$(x - \mu)^2$
3	0	0
2	– 1	1
3	0	0
1	– 2	4
6	3	9
Suma	0	14

Sin embargo, la dificultad que se presenta es que la suma de las desviaciones es 0, como se ve en la fila inferior de la tabla. Por consiguiente, el promedio sería 0 (0/5), lo que indica ausencia de dispersión; esto, obviamente, no concuerda con la realidad del conjunto.

Puede demostrarse empíricamente y con deducciones matemáticas que siempre y en cualquier conjunto, independientemente de su tamaño y valores, la suma de las desviaciones de cada valor respecto de la media es 0.

El procedimiento útil consiste, por este motivo, en obtener un valor de dispersión de cada dato que sea más alto cuanto mayor sea la desviación respecto de la media, aunque siempre en valores positivos. La forma matemática de hacerlo no es tener en cuenta la desviación, sino su cuadrado. El resultado será más alto cuanto mayor sea la desviación, aunque siempre positivo, ya que el cuadrado de un número negativo es positivo (negativo por negativo es igual a positivo).

En la tercera columna de la tabla se muestra el resultado de la operación y está encabezada con $(x - \mu)^2$. Los valores, que son los **cuadrados de las desviaciones de cada valor respecto de la media,** para el ejemplo son: 0 para el primer dato (0^2); 1 para el segundo (-1^2); .....; 9 para el último (3^2).

La suma de esta columna es 14 y constituye la suma de los **cuadrados de las desviaciones de cada valor respecto de su media.** En la terminología estadística se denomina a este valor **suma de los cuadrados y queda implícito a qué cuadrados hace** referencia.

De este valor sí es posible calcular un promedio al dividirlo por la cantidad de datos involucrados. En el ejemplo, el valor obtenido es 2,8 (14/5) y es el promedio de los cuadrados de las desviaciones de cada valor respecto de su media. Esto representa una medida de la dispersión, ya que el valor es tanto mayor cuanto mayor sea la dispersión de los datos en el conjunto y es 0 cuando no existe dispersión. Véase esto último en el segundo ejemplo presentado en este capítulo, en el cual los cinco datos tenían valor 3. Al ser 3 la media, la diferencia de cada uno respecto de la media es 0 (3 – 3). Como el resultado de 0^2 **es 0 y la suma de los cinco** ceros es 0 y 0/5 da como resultado 0, el promedio de los cuadrados de las desviaciones de cada valor respecto de la media, es decir que la medida de la dispersión es 0.

Para simplificar la nomenclatura, esta medida "democrática" de la dispersión se denomina **variancia** o **varianza**, aunque también puede identificarse como media cuadrática o cuadrado medio.

La ecuación o fórmula para el cálculo de la variancia (que es solo la simbolización del procedimiento que se describió) se muestra en la **figura 5-1**. En esas ecuaciones puede verse que, para el caso de las poblaciones, la variancia se calcula al dividir la **suma de los cuadrados por el tamaño de la población.**

En caso de que el conjunto para describir sea una muestra no solo cambia el símbolo para identificar a la variancia (σ^2 para una población y s^2 **para una muestra), sino** que el denominador no será el tamaño de la muestra, sino ese valor menos uno. Este valor del denominador (n – 1) se denomina **grados de libertad. En resumen, en el caso** de las muestras, la variancia es el resultado de la división de la suma de los cuadrados por los grados de libertad.

Si el ejemplo sobre el que se trabajó se hubiera considerado como una muestra, el resultado del cálculo de la variancia sería

3,5, ya que la suma de los cuadrados (14) se debería haber divido por 4 (5 – 1), que son los grados de libertad (n – 1) para esa situación. Las razones por las cuales cambia el denominador están más allá de lo que se abordará en este capítulo, por lo que no serán consideraradas. Sí es necesario tener presente que en la práctica no se emplean las fórmulas mostradas para el cálculo de la variancia, sino fórmulas derivadas de ellas, que hacen más rápido el procedimiento.

Por otro lado, hoy en día los datos se almacenan en bancos de datos o planillas de cálculos informáticos. Estos programas incluyen funciones que permiten el cálculo de la variancia (y de otros parámetros y estadísticos, como las medidas de tendencia central) mediante el empleo de funciones prediseñadas.

De esta manera, solo es necesario seleccionar en el programa, o en una calculadora electrónica científica, la correspondiente función. Es en este caso y en algunos programas (Microsoft Excel®, por ejemplo) se presenta la opción de cálculo de variancia de una población o de una muestra, ya que el programa no puede, por sí solo, reconocer si los datos que debe procesar corresponden a una población o a una muestra.

Los programas específicos para tareas estadísticas, como se los utiliza habitualmente para hacer inferencias a partir de muestras, por lo general calculan la variancia mediante los grados de libertad como denominador.

La variancia es de uso altamente frecuente en el análisis de conjuntos de datos numéricos y la realización de inferencias.

—

Sin embargo, no es fácil visualizar su significado en términos de poder relacionar la dispersión de uno o varios datos específicos y, además, tampoco puede relacionarse su valor con el de la media aritmética.

Esta situación se produce porque los datos y la media aritmética están en una escala –teóricamente la de los números naturales, de infinito negativo a infinito positivo– y la variancia en otra –sin números negativos, ya que es resultado de operaciones de potenciación (elevación al cuadrado)–.

Para disponer de una medida de dispersión que se pueda relacionar con los datos y su media, resulta útil volver al valor obtenido en la escala original. La manera de lograrlo es aplicar, en el valor de la variancia, la operación inversa a la potenciación, la radicación. El valor se obtiene, en consecuencia, al extraer la raíz cuadrada de la variancia. En el caso del ejemplo, el resultado redondeado en centésimas es 1,67 (raíz cuadrada de 2,8), si el conjunto que se considera es una población; y 1,87 (raíz cuadrada de 3,5), si se considera una muestra.

El valor de la raíz cuadrada del promedio de los cuadrados de las desviaciones de cada valor respecto de la media se denomina **desviación estándar** o **desviación típica**.

—

Como en el caso del cálculo de la variancia, en la práctica no se emplean las fórmulas de definición que se muestran en la figura. Asimismo, al emplear herramientas informáticas puede ser necesario especificar si los datos son los de una población o los de una muestra, para que así se aplique el denominador apropiado, N o n – 1.

POSICIÓN DE UN DATO RESPECTO DE LA MEDIA

Al conocer el valor de la desviación estándar de una población se hace posible establecer qué desviación tiene un dato en particular respecto de la media aritmética, en valores de esa medida de la dispersión.

—

Por ejemplo, supóngase que en una población (y en este caso se está considerando una población de un gran tamaño, como los habitantes de una nación o los pacientes que padecen una afección determinada) la media aritmética para una variable evaluada en forma de datos numéricos es 195 y la desviación estándar es 6; si el dato para un integrante de esa población es 204, puede decirse que ese dato está 9 unidades (de las utilizadas para la evaluación de la variable) por encima de la media, que significa 1,5 desviaciones estándar, según surge del siguiente cálculo:

$$(x - \mu)/\sigma = (204 - 195)/6 = 1{,}5$$

Al resultado de la ecuación se lo designa habitualmente con el símbolo "z" e indica la ubicación de un dato respecto de la media de la población a la que pertenece en términos de desviaciones estándar.

En el **cuadro 5-2** pueden verse algunas otras situaciones para la misma población hipotética ($\mu = 204$; $\sigma - 6$).

El cálculo del valor de "z" de la forma descrita [$z = (x - \mu/\sigma)$] se emplea en diversas aplicaciones prácticas de los procedimientos estadísticos. Algunas de ellas se analizarán en el próximo capítulo.

Como agregado, es de interés mencionar en este momento que los datos ordinales pueden resumirse al considerarlos en categorías en la forma analizada para los datos nominales. No obstante, y especialmente en el caso de los puntajes e índices estandarizados, puede ser aceptable calcular parámetros y estadísticos como los de los datos numéricos.

CUADRO 5-2. VALORES DE Z PARA DATOS DE UNA POBLACIÓN CON $\mu = 204$ Y $\sigma = 6$

x	y
187	–2,8
207	0,5
199	–0,8
212	1,3
196	–1,3
209	0,8

SÍNTESIS CONCEPTUAL

- **Para obtener información numérica** sobre un conjunto de datos numéricos se calculan inicialmente medidas de tendencia central o promedio, como la media aritmética, la mediana y la moda.
- **Además de la medida de tendencia** central, es necesario calcular alguna medida de dispersión, como el rango, la variancia o la desviación estándar, para complementar la información que brinda la tendencia central.
- **Puede calcularse la cantidad de desvia-**ciones estándar que separan a un dato de la media aritmética del conjunto al que pertenece.
- **El valor "z" representa la cantidad de** desviaciones estándar que separa a un dato de la media aritmética.

EJEMPLO 5-1

En una muestra de 50 pacientes obesos a quienes se les indicó una dieta hipocalórica se registraron las siguientes pérdidas de peso en kg al cabo de 30 días:

4,3	3,4	4,6	3,2	4
3,6	4	3,9	3,9	3,8
4,3	3,6	3	2,7	3,7
3,2	3,5	3,1	2,5	4,2
3,7	3,7	3,3	4,4	3,4
3,1	3,3	4,5	3	4,5
2,9	3,5	3,6	4,1	4,9
3	2,3	4,8	4,1	4,2
4,2	3,1	3,4	3,4	4,4
3,7	3,3	3,2	3,7	3,6

Estos datos permiten calcular una media aritmética (x, que es el estadístico que estima el parámetro en la población respectiva) de 3,66 kg y una desviación estándar de 0,59 kg.

La mediana en el mismo conjunto de datos es 3,60 kg y la moda 3,70 kg.

EJEMPLO 5-2

En una población de niños de 12 años de edad con una determinada condición general y social se encontró que el valor de la media aritmética (parámetro) de los resultados de la administración de una prueba para la evaluación de su capacidad intelectual es de 96 con una desviación estándar de 2,3.

El resultado de 101 obtenido por un niño de esas características está 2,2 desviaciones estándar (valor "z") por arriba del valor de la media aritmética de la población, mientras que uno de 95 está 0,4 desviaciones estándar por abajo de ella.

EJEMPLO 5-3

En una población de alumnos que rinden un examen con puntaje posible entre 0 y 100, la media aritmética es 78 y la desviación estándar es 6.

El alumno que quiera obtener un puntaje que esté una desviación estándar y media por encima de la media aritmética (z = 1,5) deberá obtener 87.

CAPÍTULO

6

DISTRIBUCIÓN DE FRECUENCIAS

INTRODUCCIÓN

En las poblaciones descritas con datos numéricos es de interés analizar la forma en que esos datos están distribuidos.

Este enunciado significa analizar la frecuencia con la que se manifiesta la presencia de cada valor de dato.

El **cuadro 6-1** incluye un ejemplo de la distribución de frecuencias de datos en un conjunto de ese tipo. Obsérvese que, para facilitar la visualización e interpretación, los datos se han agrupado en intervalos. Luego se procedió a contabilizar la cantidad de individuos o unidades experimentales con datos incluidos en cada intervalo.

La misma información puede ser presentada gráficamente en lo que se denomina **histograma, como se muestra en la figura 6-1.** Su aspecto es similar al de un gráfico de barras, aunque en este caso las barras se ubican una a continuación de la otra, dado que se trata de representaciones de datos numéricos continuos y no de categorías nominales.

En realidad, la representación de la distribución de datos continuos podría hacerse directamente en un sistema de coordenadas cartesianas ortogonales. En el eje de las abscisas (el eje horizontal) pueden representarse los valores correspondientes a los datos. Si estos son de tipo continuo existen, teóricamente, infinitos valores posibles entre el mayor y el menor, que pueden ser hasta infinito positivo y negativo, respectivamente.

En el eje de las ordenadas (el eje vertical) pueden representarse las frecuencias correspondientes a cada uno de esos valores para los datos, teóricamente, infinitos.

Al marcar en el sistema el punto de intersección del valor de cada dato y su frecuencia, se obtiene una serie de puntos. Esa serie teóricamente sería infinita y, por lo tanto, se visualizaría como una línea. Esta es la representación gráfica de la distribución de

CUADRO 6-1. DISTRIBUCIÓN DE FRECUENCIAS EN UN CONJUNTO DE DATOS

Intervalo	Frecuencia
50,0-54,9	3
55,0-59,9	3
60,0-64,9	18
65,0-69,9	16
70,0-74,9	17
75,0-79,9	10
80,0-84,9	10
85,0-89,9	5
90,0-94,9	7
95,0-99,9	3
> 100	0
Total	**92**

los datos en el conjunto (población) en consideración. Como en una situación real no es posible obtener la frecuencia que corresponde a una cantidad infinita de datos, sino solo de algunos de ellos, la representación gráfica que se genera no es una línea continua, sino lo que se denomina **polígono de frecuencias.**

FORMA DE DISTRIBUCIÓN

Según sea la variable de interés y la manera en que se la haya evaluado con un dato numérico continuo, la línea que representa gráficamente la distribución de frecuencias asume diferentes formas.

Así, por ejemplo, cuando se consideran los ingresos mensuales de los habitantes de un país en unidades monetarias suele generarse un gráfico como el de la **figura 6-2.** La línea indica que, cuando el valor del dato es bajo (ingresos cercanos a 0), el valor de la ordenada –es decir, la frecuencia– es bajo. Cuando los valores aumentan algo se visualiza un aumento de la frecuencia, ya que es común que una cantidad grande de habitantes perciba ingresos mayores que 0, aunque no muy elevados. A medida que aumentan más los valores, la frecuencia disminuye y solo queda una muy baja, correspondiente a los individuos con ingresos muy altos.

Como puede verse, la forma indica la presencia de una distribución marcadamente asimétrica o con un sesgo hacia un lado. Se acostumbra a hablar de sesgo positivo en un caso como el del ejemplo, y negativo si la asimetría se manifiesta en el sentido

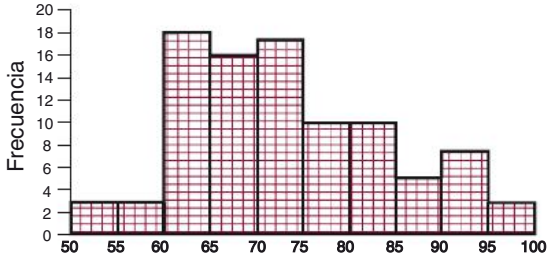

Fig. 6-1. Histograma correspondiente a los datos del **cuadro 6.1**.

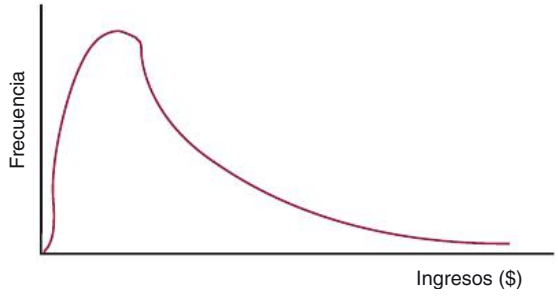

Fig. 6-2. Gráfico posible de la distribución de ingresos anuales ($) en una población.

contrario. La situación puede resumirse de manera numérica mediante el cálculo del **coeficiente de asimetría.** Diversas planillas de cálculos y programas de estadística permiten hacerlo.

El conocimiento de la presencia de una distribución de este tipo, asimétrica, orienta en la selección de una medida de tendencia central. Si se utiliza la media aritmética, se obtiene un valor que no condice con lo que usualmente se espera de un promedio, un valor que indica en dónde está la mitad o la mayoría de la serie de datos. Esta situación se produce porque unos pocos datos elevados, unas pocas personas con ingresos muy altos, modifican sustancialmente el valor de ese parámetro.

Por otro lado, al utilizar la moda se obtiene solo una imagen del valor para la mayoría, aunque sin que se represente la situación real de una cantidad significativa de datos. Por ende, puede considerarse a la mediana como una forma más acertada de descripción al indicar el valor que realmente separa el conjunto en dos mitades: la de mayores ingresos y la de menores ingresos.

Si se tomara como dato la valoración en años de la edad en el momento del fallecimiento, el gráfico que representa la distribución de frecuencias tomaría la forma de la **figura 6-3**. La frecuencia relativamente elevada con baja edad está representada por la mortalidad infantil y en los primeros años de vida. A partir de esos primeros años, la frecuencia de individuos que mueren es menor, aunque aumenta notoriamente a partir de cierta edad. La distribución es, en este caso, razonablemente simétrica; sin embargo, la ubicación del valor de la media aritmética corresponde a una frecuencia relativamente baja, como se indica en el gráfico. En este caso, la información sobre la moda o las modas (cuando la distribución es bimodal o multimodal) ayuda a obtener una imagen más cercana a la situación real.

PERCENTILES, CUARTILES Y QUINTILES

Es posible ubicar a cada uno de los datos numéricos dentro de la distribución del conjunto del que forman parte. Esta posibilidad se utiliza para establecer qué posición ocupa el individuo –paciente, alumno, etc.– en el dato que se registró dentro del conjunto al que pertenece. A partir de ello se pueden tomar decisiones sobre la forma de actuar frente a la situación que se presenta. Por ejemplo, la ubicación del dato que indique el nivel de desempeño de un alumno dentro del conjunto de alumnos que pasaron por la misma experiencia docente permite establecer si el alumno merece una distinción especial o si debe ser considerado un alumno "término medio o normal". De la misma manera se actúa cuando se registra un dato numérico en un paciente para evaluar el contenido de una sustancia en sangre. Según la ubicación del valor registrado, dentro de lo que se espera en el conjunto de pacientes al que pertenece, se establece si el valor es normal, elevado o bajo y se instituye un criterio terapéutico específico.

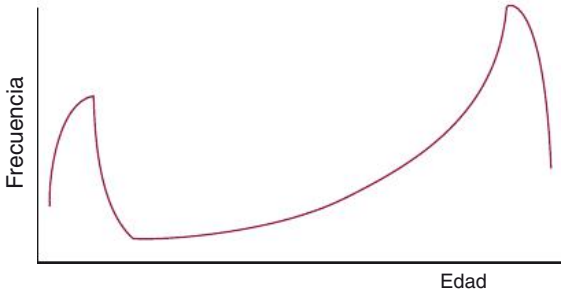

Fig. 6-3. Gráfico posible de la distribución de edad en el momento de la muerte en una población.

En términos de procesamiento estadístico, la posición de un dato numérico dentro de un conjunto se establece al determinar el **percentil correspondiente a ese dato.**

El percentil de un dato indica qué porcentaje –proporción multiplicada por cien– de datos del conjunto tiene un valor menor que él.

Así, por ejemplo, a un dato con un valor igual a la mediana del conjunto le corresponde el percentil 50, ya que, como se definió en el capítulo anterior, la mediana separa al conjunto en mitades. Un dato con percentil 5 está dentro de los valores más bajos del conjunto –su valor es solo mayor que una "minoría" del conjunto–, mientras que uno con percentil 95 está dentro de los más altos –su valor es solo superado por una pequeña proporción de los datos del conjunto–.

Para el trabajo clínico habitual es frecuente repartir al conjunto de datos en cuatro partes denominadas cuartiles: del percentil 0 al 25, del 25 al 50, del 50 al 75 y del 75 al 100. A partir de esa información no es infrecuente considerar valores "normales" a los que pertenecen al segundo y tercer cuartil, lo que significa que el percentil correspondiente está entre 25 y 75. Al intervalo comprendido entre estos percentiles se lo conoce como **rango o intervalo intercuartil. En pediatría es frecuente disponer** de una tabla o gráfico que permite ubicar la talla o el peso de un niño dentro de estos percentiles o cuartiles, y así tener información sobre su situación de crecimiento y desarrollo.

En el campo de las ciencias sociales es bastante habitual repartir la población en **quintiles** (cada uno corresponde a un intervalo de 20 en percentiles), a partir de datos como los que reflejan los ingresos del grupo familiar.

DISTRIBUCIÓN NORMAL O GAUSSIANA

La forma que asumen las representaciones gráficas de las distribuciones de frecuencias puede ser diversa. Sin embargo, una gran cantidad de datos numéricos con los que se valoran las variables de interés en el campo de las ciencias de la salud lleva al empleo de gráficos de distribución de frecuencias similares al que se muestra en la **figura 6-4**.

También es de interés tener presente que esa misma forma de distribución de frecuencias se observa cuando se representa la distribución de los errores que se cometen al registrar los datos. Si una misma situación se evalúa de manera repetida (p. ej., si se pesa repetidas veces una masa determinada en una balanza), la mayoría de los datos tienden a ser registrados con un determinado valor. Con menos frecuencia se registran datos que se alejan de ese valor central y más común, y esta situación es la que aparece representada gráficamente.

A esta forma de distribución de datos numéricos se la conoce como **distribución normal. Algunas de sus características se** deducen de la observación de la **figura 6-4**.

a) **La distribución normal es simétrica respecto** de un valor que corresponde a la media aritmética de los datos considerados. Esto significa que los valores de la media aritmética y de la mediana son coincidentes.

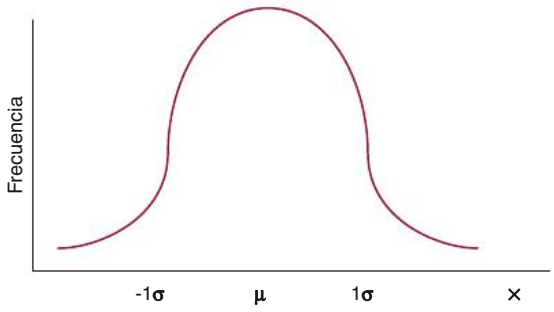

Fig. 6-4. Gráfico de la distribución gaussiana (normal).

b) **El dato con mayor frecuencia, es decir,** la moda representada por el punto más alto de la línea corresponde al valor de la media aritmética y de la mediana. Las tres medidas de tendencia central más comunes son coincidentes en esta forma de distribución.

c) **La forma de la línea puede ser semejante** a la del corte de una campana con dos puntos ubicados en forma simétrica respecto de la media, en los cuales la línea cambia de dirección. Esos dos puntos de inflexión corresponden a los representados con los datos ubicados a una desviación estándar por abajo y por arriba de la media aritmética. En símbolos, esto significa que esos puntos están en los valores de la abscisa **$(\mu - 1\sigma)$ y $(\mu + 1\sigma)$.**

Una línea de características así definidas puede obtenerse también a partir de la representación gráfica de la resolución de una determinada ecuación. El matemático Gauss, varios siglos atrás, trabajó con una ecuación que, una vez resuelta, genera una línea de las características de la distribución normal.

Se dice que un conjunto de datos numéricos tiene una distribución gaussiana o normal cuando la forma que asume se corresponde con la que genera la resolución de lo que se conoce como ecuación de Gauss.

—

Esta ecuación se muestra en la **figura 6-5** y de su observación se pueden extraer algunas consecuencias que explican su utilidad práctica. En la ecuación se encuentran dos incógnitas: "x" e "y", y como es habitual en la simbología matemática, "y" es el valor de la ordenada, la frecuencia; y "x" el de la abscisa, el valor de un dato.

También puede observarse que para resolver la ecuación en el caso de una población en particular se necesita conocer el valor de cuatro constantes. Dos de ellas, "**π**" y "e", son las mismas para cualquier población de la cual se trate. La primera **(π) es** la relación entre diámetro y circunferencia, y la segunda, la base de los logaritmos naturales. Las otras dos constantes, "**μ" y "σ",** son específicas para cada población, ya que corresponden a sus parámetros de media aritmética y de desviación estándar, respectivamente.

Todo esto significa que, al conocer el valor de la media aritmética y la desviación estándar de una población, es posible establecer la frecuencia relativa y, por ende, el percentil correspondiente a un determinado dato. En una situación real esto implica, por ejemplo, poder calcular qué fracción –la cual se puede expresar en porcentajes si se lo desea– de individuos en una población tiene datos de un determinado valor o qué porcentaje de individuos tiene valores inferiores a él (el percentil que corresponde a ese dato o individuo).

$$Y = \frac{1}{\sigma \quad (2\Pi)^{-2}} \; e^{-1/2\,((x-\mu)/\sigma)}$$

Fig. 6-5. Ecuación de la distribución gaussiana.

La resolución de la ecuación permite establecer que, en todo conjunto con distribución gaussiana, el 95% del área debajo de la línea (95% de los individuos de la población) tiene valores para el dato entre 1,96 desviaciones estándar por abajo y 1,96 desviaciones estándar por arriba del valor de la media aritmética.

—

Esto ocurre porque para la posición del valor z = –1,96 (véase en el capítulo anterior el análisis de la posición de un dato respecto del conjunto en términos de desviaciones estándar) corresponde el percentil 2,5, y para la posición del valor z = 1,96, el percentil 97,5. Entre ambos queda comprendido el 95% del área y el conjunto de datos que esa área representa.

Cuando se toma el intervalo entre $\mu \pm 2{,}5\ \sigma$ (dos y media desviaciones estándar por arriba y por debajo de la media aritmética) se incluye prácticamente a la totalidad del área. Esto es así porque esas cantidades de desviaciones estándar corresponden, aproximadamente, a los percentiles 0,5 y 99,5. Sin embargo, en la resolución matemática solo se cubre la totalidad del área cuando el intervalo se extiende desde infinito negativo hasta infinito positivo.

Si esto se traslada a la situación de una población real con distribución gaussiana puede decirse que, de los datos incluidos en ella, el 95% tiene valores entre $\mu \pm 1{,}96\ \sigma$.

APLICACIONES DE LA DISTRIBUCIÓN NORMAL

El conocimiento anteriormente analizado puede aplicarse en distintas situaciones prácticas. Entre ellas se encuentra la posibilidad de ubicar a un individuo dentro de la población a la que pertenece en función del dato que se obtuvo en él para la evaluación de una variable. Esto es posible siempre que ese dato esté distribuido en forma gaussiana en esa población y se conozcan los parámetros de la media aritmética y de la desviación estándar correspondientes a esa distribución.

Un ejemplo de esas aplicaciones consiste en el uso de procedimientos de diagnóstico de capacidades o alteraciones del comportamiento que se emplean, entre otras técnicas, en psicometría. Con frecuencia se utilizan pruebas para evaluar la "inteligencia" en los integrantes de una población definida (rango de edad específico).

La administración de esas pruebas a un número grande de individuos permite calcular, con un grado de certeza razonable, los parámetros de la población. Es usual procesar los datos para generar una situación en la que la media aritmética asume un valor 100 y la desviación estándar, un valor 10.

Si a un individuo en particular se le administra la prueba y se obtiene un resultado determinado, 109, por ejemplo, es posible determinar si su comportamiento es semejante al de la mayor parte de sus "compañeros" o si difiere de lo que se espera en la mayoría.

Para ello, se aplica el procedimiento descrito en el capítulo anterior por medio del cual se calcula el valor de "z", la ubicación del dato respecto de la media en términos de desviación estándar – $z = (x - \mu) / \sigma$. En el ejemplo, z sería igual a 0,9 = (109 – 100) / 10.

Este individuo está dentro de una desviación estándar del valor de la media, por lo que se puede estimar que es un integrante del área central (véanse los últimos párrafos del apartado anterior) y, por lo tanto, puede considerárselo un individuo "normal", si se acepta que lo más frecuente es lo "normal".

En cambio, un individuo que en la misma prueba genere un dato 73 puede ser considerado "anormal" en términos de deficiencia de inteligencia (suponiendo, por supuesto, que esta variable se haya evaluado en forma válida por esta prueba). Efectivamente, para este caso, el valor de "z" es −2,7 = (73 − 100)/10, lo que significa una ubicación por debajo (como indica el signo negativo) del 99% central, que está entre $\mu \pm 2{,}5\ \sigma$.

En función de las mismas consideraciones, un resultado 124 (z = 2,4) identificaría a un "genio en potencia" por su ubicación por encima del 95% central.

En algunos procesos de control de calidad y en otras aplicaciones se hace uso de las propiedades de la distribución gaussiana. En todos los casos, se parte de conceptos básicos que pueden ser resumidos en el enunciado siguiente, referido a la situación en una población de datos con esa forma: la "mayoría" (alrededor del 95%) tiene valores de datos entre casi dos desviaciones estándar a la izquierda y a la derecha del valor de la media aritmética (más exactamente, 1,96).

La palabra "mayoría" es una denominación arbitraria y, por lo tanto, discutible, aunque su significado del 95% surge de la aplicación matemática de una ecuación y, por ello, es más fácil de aceptar.

De la misma manera, si se selecciona en forma aleatoria, al azar, un integrante de una población con distribución gaussiana puede estimarse cuál es la probabilidad de obtener un dato dentro de un determinado rango de valores.

La probabilidad, que se simboliza con la letra ***P***, está representada por la relación entre el resultado buscado y la totalidad de los resultados posibles. Por ejemplo, la probabilidad de que al dejar caer una moneda el lado denominado "cara" quede hacia arriba es 1/2 (0,5 o 50%), ya que 1 es el resultado buscado y 2 son los resultados posibles.

El valor de la probabilidad es un número que se ubica dentro del rango entre 0 y 1 (o 0 y 100%, si se lo expresa porcentualmente), en el cual el primer valor corresponde a imposibilidad y el segundo, a un resultado seguro.

En una distribución gaussiana, los resultados posibles son infinitos y la fracción cubierta por un determinado rango o intervalo de valores indica la probabilidad de su ocurrencia.

Puede decirse que, al seleccionar al azar a un integrante de una población con distribución gaussiana, es "poco probable" ($P < 0{,}05$) que el dato que lo describe esté alejado de la media más de dos desviaciones estándar.

Otra vez, en este caso la expresión "poco probable" es arbitraria, aunque no el valor de *P*. Este valor es menor que 0,05 (o menor que 5%) en el enunciado, ya que los valores a los que se hace referencia son los que están por fuera del rango central que abarca al 95%.

Los temas tratados en este capítulo están referidos al análisis de poblaciones que se presuponían conocidas o razonablemente conocidas en sus parámetros y forma de distribución. En los capítulos próximos se

utilizarán los conceptos adquiridos para sentar las bases para la interpretación de los procedimientos de la **estadística inferencial, es decir, los principios del trabajo** estadístico a partir de muestras tomadas de una población.

SÍNTESIS CONCEPTUAL

- **Es de interés analizar la forma en la que** los datos están distribuidos en un conjunto.
- **El percentil de un dato indica qué** porcentaje de datos del conjunto tienen un valor inferior a él.
- **Cuando la forma de la distribución** de los datos numéricos puede ser asimilada a una distribución normal o gaussiana, es posible aplicar la ecuación correspondiente para conocer su percentil, a partir del conocimiento de la posición de un dato respecto de la media en términos de desviaciones estándar (valor "z" del dato).
- **La resolución de la ecuación permite** establecer que, en todo conjunto con distribución gaussiana, el 95% del área debajo de la línea (95% de los individuos de la población) tiene valores para el dato entre 1,96 desviaciones estándar por debajo y 1,96 desviaciones estándar por arriba del valor de la media aritmética.

EJEMPLO 6-1

Un alumno obtuvo un puntaje de 72 en una prueba estandarizada de biología, en la cual los parámetros para la población a la que él pertenece son $\mu - 63$ y $\sigma = 5$. En una prueba del mismo tipo, aunque sobre química, en la cual $\mu = 74$ y $\sigma = 8$, obtuvo un puntaje de 82. ¿En cuál de las dos disciplinas es un alumno más "destacado"?

En biología, ya que su puntaje 72 está 1,8 desviaciones estándar por encima de la media (valor "z"), mientras que el 82 que obtuvo en química está solo a 1. Si se supone una distribución aproximadamente normal de ese puntaje, esto significa que superó a una mayor cantidad de compañeros en biología que en química.

EJEMPLO 6-2

En una población de 200 000 personas de género masculino, la edad a la que sus integrantes quedan totalmente desdentados está distribuida en forma aproximadamente normal, con $\mu = 58$ años y $\sigma = 12$ años. Si se decide brindar un servicio de prótesis completa a los menores de 46 años, ¿para qué cantidad de individuos deben asegurarse recursos?

Aproximadamente para 32 000. Esto es así porque la edad 46 está una desviación estándar por debajo de la media aritmética de la población (valor "z"). Si entre una desviación estándar por encima y por debajo de este valor (58) se encuentra el 68% de la población, por fuera queda el 32%. De estos últimos, la mitad (16%) estarán por debajo, y el 16% de 200 000 es 32 000.

EJEMPLO 6-3

El costo de los tratamientos que se ofrecen en una clínica son distribuidos en forma razonablemente normal, con $\mu = \$\ 2250$ y $\sigma = \$\ 150$. El tratamiento que necesita un paciente tiene un costo mayor que el de la mayoría de los pacientes atendidos. ¿Cuál es el costo del tratamiento para este paciente?

Si se acepta como mayoría el 95% más frecuente en la distribución, puede estimarse un costo superior a $ 2550, ya que entre este valor y $ 1950 (valores que están dos desviaciones estándar por encima y por debajo de la media aritmética) se encuentra ese porcentaje, según surge de la ecuación de Gauss.

CAPÍTULO

7

MUESTREO

INTRODUCCIÓN

La manera usual de realizar un estudio del comportamiento de variables en una población es tomando muestras de individuos o unidades experimentales pertenecientes a ella.

A partir de los datos registrados en estas unidades es posible, luego, realizar inferencias sobre el conjunto total: la población.

Una condición que debe reunir una muestra para realizar esas inferencias es ser "representativa". Esto significa que en ella deben estar representadas todas las condiciones presentes en la población y que pueden influir en el dato a partir del cual se evalúa la variable de interés.

La representatividad de una muestra se garantiza por la forma de selección de sus componentes. La aplicación de las técnicas estadísticas presupone que una muestra es representativa.

Una vez decidido cómo se asegurará la representatividad, se debe garantizar que –dentro de la población definida o dentro de un estrato o subconjunto de ella– la selección se realice en forma aleatoria. Esta condición significa que, durante el procedimiento, cada uno de los integrantes de la población tiene la misma probabilidad o una probabilidad conocida de ser seleccionado.

A lo largo de este capítulo se analizará el comportamiento de las muestras tomadas en esas condiciones, en su relación con los parámetros de la población de origen. En primer lugar, se examinará la situación para variables descritas mediante datos numéricos y luego se harán algunas apreciaciones para el caso de los datos nominales. Como en otras situaciones, los datos ordinales pueden considerarse como categóricos, al igual que los nominales; o, cuando se trata de puntajes o índices razonablemente estandarizados, tratarlos como numéricos, aun cuando en realidad no lo sean.

MUESTRAS CON DATOS NUMÉRICOS

Supóngase que se está frente a una población hipotética y pequeña de cuatro individuos (n = 4), en la cual quienes la componen tienen los siguientes datos numéricos para una determinada variable: a = 4; b = 3; c = 3; d = 2. Esos valores se incluyen en la primera fila del **cuadro 7-1**. La media aritmética (μ) en ese conjunto es 3.

Si se supone que alguien está interesado en el valor de ese parámetro, aunque no tenga acceso a la población, sino solo a algunos de sus integrantes, necesitaría trabajar a partir de una muestra. Se verá a continuación la situación que se plantearía al utilizar una muestra de tamaño 2 (n = 2) para la tarea. Se entenderá que esta situación se presenta a manera de ejemplo, ya que las poblaciones que presentan un interés real tienen tamaños notoriamente mayores.

Para seleccionar en forma aleatoria a dos de las cuatro unidades de la población podrían colocarse cuatro bolillas identificadas con las respectivas letras en un bolillero y retirar dos por sorteo.

Debe destacarse que en este caso no se está cumpliendo el requisito de aleatoriedad en su totalidad.

En efecto, la posibilidad de selección de la primera bolilla ha sido de uno en cuatro ($P = 1/4$) y en la segunda, de uno en tres ($P = 1/3$). Para trabajar de manera realmente aleatoria es necesario obtener la muestra "con reemplazo". Esto significa que cada unidad se debe seleccionar, registrar el dato en ella y luego incorporarla nuevamente a la población para mantener constantes las posibilidades de selección. Este mecanismo, que hace posible que una misma unidad sea seleccionada más de una vez, no es el que se aplica en las situaciones reales, y obliga a algunas modificaciones en los procedimientos que se describirán más adelante. No obstante, esas modificaciones tienen un peso significativo en los resultados solo cuando el tamaño de la muestra supera el 10% del tamaño de la población, lo que rara vez sucede en las investigaciones reales. Por este motivo, se trabajará aquí asumiendo la aleatoriedad aun cuando no sea real por seleccionar una muestra que, en el ejemplo, tiene un tamaño (2) que representa la mitad del tamaño de la población (4).

Según el ejemplo, en la primera columna del **cuadro 7-1** puede verse la composición de las seis posibles muestras que pueden obtenerse en las condiciones planteadas.

CUADRO 7-1. RESULTADOS EN LAS MUESTRAS TOMADAS DE UNA POBLACIÓN HIPOTÉTICA DE DATOS NUMÉRICOS

Muestra (n = 2)	Media	Muestra (n = 3)	Media
a = 4 b = 3	3,50	a = 4 b = 3 c = 3	3,33
a = 4 c = 3	3,00	a = 4 b = 3 d = 2	3,00
a = 4 d = 2	3,00	a = 4 c = 3 d = 2	3,00
b = 3 c = 3	3,00	b = 3 c = 3 d = 2	2,67
b = 3 d = 2	2,50	Suma	12,00
c = 3 d = 2	2,50	Media	3,00
Suma	18,00		
Media	3,00		

Población: a = 4; b = 3; c = 3; d = 2; μ = 3.

Al calcularse en cada una de esas muestras el dato estadístico de tendencia central media aritmética (recuérdese que se acostumbra a hablar de un estadístico cuando el valor calculado es en una muestra, mientras que se emplea el término parámetro cuando lo es en una población), se obtienen los resultados que se muestran en la segunda columna del cuadro.

De esos resultados surge que, en dos de las muestras, el valor del estadístico $\bar{x}$ coincide con el parámetro de la población ($\mu = 3$); en otras dos al estadístico le correspondió un valor mayor que el del parámetro; y en otras dos, un valor menor.

Una primera conclusión que es posible extraer es que puede producirse una estimación correcta, una sobreestimación o una subestimación, al estimar la media aritmética de una población a través de la media aritmética de una muestra. Esto no depende de una forma de trabajar correcta, sino tan solo de la mayor o menor "suerte" que se tenga en la selección aleatoria de la muestra.

Se puede reconocer que la media aritmética de las muestras tomadas de una población varía.

—

Por otro lado, en la última fila del cuadro puede verse que la media aritmética –el promedio– de las medias aritméticas de las muestras obtenidas es 3 (el valor de su suma, 18, dividido por la cantidad de muestras totales, 5); este valor corresponde al del parámetro de la población ($\mu = 3$).

Una segunda conclusión es que debido a que el valor de la media aritmética de la muestra tomada de una población varía –aunque a veces se lo estima bien y otras se lo sobreestima o se lo subestima–, al parámetro, en promedio, se lo estima bien.

Puede expresarse que, en promedio, la media aritmética del conjunto de medias aritméticas de muestras de una población es igual a la media aritmética –parámetro– de la población de la cual se tomaron las muestras.

—

Véase ahora, en las columnas tercera y cuarta del cuadro, lo que sucede al tomar muestras de tamaño 3 (n = 3) de la misma población hipotética. También en este caso, en promedio, se lo estima bien, aunque la magnitud de la sobreestimación o de la subestimación es menor que cuando las muestras son de menor tamaño. En efecto, cuando n = 2 el error de estimación fue de 0,50 en más o en menos, mientras que cuando n = 3, este fue de 0,33.

Una tercera conclusión es que la magnitud del error que puede cometerse al estimar la media aritmética de una población a partir del correspondiente estadístico disminuye al aumentar el tamaño de la muestra utilizada.

Por último, véanse en el **cuadro 7-2** los resultados que se obtuvieron al repetir el procedimiento en otra población de tamaño 4, aunque con integrantes: a = 5; b = 3; c = 3; d = l. La media aritmética (μ) en este conjunto es también 3, aunque su dispersión es mayor. Esto puede visualizarse a partir del rango o recorrido que es 4 (5 – 1), mientras que es 2 (4 – 2) en la población del primer ejemplo de este capítulo.

CUADRO 7-2. RESULTADOS EN LAS MUESTRAS TOMADAS DE UNA POBLACIÓN HIPOTÉTICA DE DATOS NUMÉRICOS

Muestra (n = 2)	Media	Muestra (n = 3)	Media
a = 5 b = 3	4,00	a = 4 b = 3 c = 3	3,67
a = 5 c = 3	4,00	a = 4 b = 3 d = 2	3,00
a = 5 d = 1	3,00	a = 4 c = 3 d = 2	3,00
b = 3 c = 3	3,00	b = 3 c = 3 d = 2	2,33
b = 3 d = 1	2,00	Suma	12,00
c = 3 d = 1	2,00	Media	3,00
Suma	18,00		
Media	3,00		

Población: a = 5; b= 3; c = 3; d = 1; μ = 3.

Las conclusiones extraídas se aplican a esta nueva situación, aunque al comparar los resultados de las dos tablas puede observarse que para un mismo tamaño de muestra la magnitud del error que puede cometerse en la estimación es mayor en este caso.

Una cuarta conclusión es, por lo tanto, que la magnitud del error que puede cometerse al estimar la media aritmética de una población a partir del correspondiente estadístico aumenta al incrementarse la dispersión de la población de la que se toma la muestra.

ERROR ESTÁNDAR

Los resultados de la supuesta experiencia descrita en los ejemplos planteados llevan a la conclusión que se enuncia en el siguiente párrafo.

La magnitud del error posible, al estimar la media aritmética de una población a partir de la media aritmética de una muestra tomada aleatoriamente de ella, aumenta al incrementar la dispersión de la población y al disminuir el tamaño de la muestra, y disminuye al reducir la dispersión de la población y aumentar el tamaño de la muestra.

—

Expresado en términos matemáticos, puede decirse que la magnitud del error es directamente proporcional a la dispersión de la población de origen de la muestra e inversamente proporcional al tamaño de esta.

Así, al ser la variancia la medida democrática de la dispersión, este enunciado puede resumirse mediante la siguiente fórmula:

$$\text{Magnitud del error} = \sigma^2 / n$$

El resultado de la fórmula está en una escala diferente de la de la media aritmética (recuérdese lo analizado en el **cap. 5, Resumen de datos numéricos**) por lo que es útil extraer la correspondiente raíz cuadrada y así llegar al valor de lo que se denomina **error estándar,** cuya fórmula es la siguiente:

$$\text{Error estándar} = \sigma / \sqrt{n}$$

Es decir, el error estándar puede calcularse al dividir el valor de la desviación estándar

de la población por la raíz cuadrada del tamaño de la muestra utilizada.

Obsérvese que existen dos situaciones en las cuales la posibilidad de error es nula (error estándar igual a 0). Una de ellas se produce cuando no existe dispersión en la población original, es decir que todos sus datos son iguales. Al ser el numerador 0, el cociente también lo es, ya que este valor dividido por cualquier otro arroja ese resultado.

La segunda situación se verifica cuando la muestra tomada es infinitamente grande, es decir, cuando se evalúa a la totalidad de la población; en este caso, el denominador es infinito y el resultado de dividir cualquier valor por infinito es 0.

Como se comprenderá, se trata de dos situaciones inexistentes en la realidad de la investigación. En los datos numéricos es prácticamente imposible evitar alguna dispersión, porque no todos los individuos de una población se comportan exactamente igual, o porque es prácticamente imposible no cometer algún error en la recolección de datos. Por otro lado, las poblaciones de interés tienen un tamaño demasiado grande como para que sea posible trabajar con todos sus integrantes.

Debe hacerse una consideración adicional. Cuando, como en los ejemplos con los que se ha trabajado, las muestras se obtuvieron **sin reemplazo**, la fórmula para el cálculo del error estándar debe modificarse al multiplicarla por un factor de corrección. Sin embargo, ese factor de corrección genera un valor de error estándar, que puede considerarse que afecta los resultados de análisis posteriores solo cuando el tamaño de la muestra supera alrededor del 10% del volumen de la población respectiva. Esta situación es prácticamente inexistente en las investigaciones en las ciencias de la salud, por lo que en el trabajo habitual esto no es tenido en cuenta y las técnicas estadísticas se aplican como si las muestras hubieran sido obtenidas **con reemplazo**.

DISTRIBUCIÓN DE MEDIAS ARITMÉTICAS DE LAS MUESTRAS

El error estándar representa una medida de la dispersión de la distribución de los valores de las medias de las muestras tomadas de una población, de la misma manera que la desviación estándar lo es de la dispersión de los datos originales.

Para que este valor adquiera significado en su relación con la medida de tendencia central –la media aritmética– es necesario establecer, de manera empírica o matemática, cuál es la forma de distribución de la variable: en el caso que nos ocupa, el valor de la media aritmética de las distintas muestras tomadas de la población. Al hacerlo es posible verificar lo que se enuncia a continuación.

La distribución de los valores de las medias aritméticas de las muestras tomadas en una población es gaussiana, aun cuando la distribución de los datos de la población no tenga esa característica.

—

Esto permite aplicar, a la distribución de las medias de las muestras, los conceptos y procedimientos basados en la ecuación correspondiente a esa distribución –analizados en el capítulo anterior– con la salvedad de que, en lugar del valor de la desviación estándar, debe tenerse en cuenta el del error estándar.

En función de lo enunciado, puede decirse que, de todas las muestras tomadas aleatoriamente a partir de una población, el 95% tiene valores de media aritmética comprendidos entre poco menos de dos errores estándar (1,96 exactamente) por debajo y por arriba de la media aritmética de la correspondiente población.

Por ejemplo, si de una población con $\mu = 1000$ y $\sigma = 40$ se toman muestras con $n = 25$, puede esperarse que, de modo aproximado, el 95% de ellas tengan valores de media aritmética de entre 984 y 1016. Esto es porque el error estándar en esta situación es 8 ($40 / \sqrt{25}$) y dos veces 8 es 16.

De lo expresado también se puede deducir que, al tomar una muestra al azar, es "poco probable" ($P < 0,05$) que su media aritmética esté alejada de la media de la población más de dos errores estándar.

Es necesario tener presente estos conceptos para encarar la tarea que se plantea en los capítulos siguientes. Asimismo, manténgase presente también que, desde la ecuación matemática, cualquier valor de media aritmética de una muestra es posible, ya que la ecuación gaussiana genera una línea –en forma de campana– que cubre un área que se extiende desde el valor de infinito negativo hasta el infinito positivo.

MUESTRAS CON DATOS NOMINALES

Al tomar muestras de poblaciones de datos nominales, la situación es equivalente a la que se ha descrito para los datos numéricos.

Considérese una población hipotética de 8 individuos, de los cuales 4 ($p = 0,5$ o 50%) están en la categoría "enfermos".

Aunque en forma simplificada, ya que las secuencias en las que pueden resultar seleccionadas las unidades no son tenidas en cuenta, los resultados posibles al tomar muestras de tamaño cuatro ($n = 4$) se muestran en el **cuadro 7-3.** Tal como en los casos anteriores, al estimar el parámetro con el valor del estadístico en ocasiones se "acierta" y en otras se lo sobreestima o subestima, aunque "en promedio" se lo estima bien.

CUADRO 7-3. RESULTADOS EN LAS MUESTRAS TOMADAS DE UNA POBLACIÓN HIPOTÉTICA DE DATOS NOMINALES

Muestra	% enfermos
A 4 enfermos 0 sanos	100,0
B 3 enfermos 1 sano	75,0
C 2 enfermos 2 sanos	50,0
D 1 enfermo 3 sanos	25,0
E 0 enfermo 4 sanos	0,0
Suma	**250,0**
% promedio	**50,0**

Población: enfermos = 4; sanos = 4; $P = 0,5$; prevalencia = 50%.

También en este caso la magnitud del error posible en la estimación es inversamente proporcional al tamaño de la muestra: a mayor tamaño de muestra, menor error posible.

La diferencia estriba en que la distribución no es en este caso gaussiana, sino que puede ser descrita con otro tipo de ecuación, conocida como **binomial**, y el valor del error estándar es la raíz cuadrada del valor obtenido de:

$$p\,(1 - p) / n$$

Esto es la raíz cuadrada del resultado del producto de la proporción que corresponde a una categoría (0,5 en la categoría "enfermos" en el ejemplo) por la proporción que no está en esa categoría ($1 - p$; es decir 0,5 en el ejemplo) dividido por el tamaño de la muestra (4 en el ejemplo).

Nótese que, también en este caso, el tamaño de la muestra es el denominador para el cálculo del error estándar. Por ello, al igual que con las muestras de datos numéricos, la magnitud del error posible aumenta al disminuir el tamaño de la muestra o disminuye con su aumento.

SÍNTESIS CONCEPTUAL

- **Cuando se toman muestras de un conjunto** de datos numéricos, la media aritmética varía entre las muestras, aunque el dato estadístico del conjunto de todas las muestras posibles es, en promedio, igual al parámetro de la población de la que fueron obtenidas.
- **La distribución de las medias aritméticas** de esas muestras toma una forma semejante a la distribución gaussiana, con una medida dispersión cuantificable mediante el error estándar.
- **Los valores estadísticos de las muestras** de conjuntos de datos nominales varían con una distribución descrita por la denominada distribución binomial.
- **Tanto en el caso de los datos numéricos** como en el de los datos nominales, el valor del error estándar es inversamente proporcional al tamaño de las muestras.

EJEMPLO 7-1

En una población de adultos sin manifestaciones de presencia de cálculos sobre sus superficies dentarias, el contenido de calcio en saliva tiene un valor de media aritmética de 5,6 mg/100 mL, con una desviación estándar de 0,9 mg/100 mL.

¿Es "poco probable" ($P < 0,05$) que la media aritmética de una muestra de tamaño 100 tenga un valor de 5,3 mg/100 mL o no?

Es poco probable, ya que este valor está alejado de la media de la población, 0,30, más de dos errores estándar. El error estándar en este caso es 0,09 (0,9 / $\sqrt{100}$), que multiplicado por 2 es 0,18.

¿Y si la muestra hubiera tenido un tamaño igual a 20? El valor obtenido no sería poco probable, ya que en este caso el error estándar sería de 0,20 (0,9 / $\sqrt{20}$), que multiplicado por 2 es 0,40, un valor menor de 0,30.

EJEMPLO 7-2

En una población de adultos jóvenes, la estatura media (media aritmética) es de 1,70 m y la desviación estándar es de 0,24 m.

¿Menor o mayor a qué valor deberá ser la media aritmética de una muestra de tamaño 64 tomada de esa población para poder considerarse que se está frente a una situación poco probable ($P < 0,05$)?

El error estándar de la distribución de las medias de muestras de ese tamaño tomadas de esa población es 0,03 (0,24 / $\sqrt{64}$). Los valores 1,64 y 1,76 están dos errores estándar alejados de la media. Por lo tanto, cuando la media de la muestra obtenida sea menor o mayor, respectivamente, de esos dos valores, se estará frente a una situación poco probable.

CAPÍTULO

8

ESTIMACIÓN DE PARÁMETROS

INTRODUCCIÓN

Cuando se llevan a cabo investigaciones descriptivas mediante metodología cuantitativa, el objetivo se centra en la obtención del valor del parámetro que permita describir a una población en relación con la variable de interés. Ese parámetro suele estar representado por una proporción o una media aritmética, según se empleen datos nominales o numéricos, respectivamente, para la evaluación de la variable.

Al trabajar con muestras tomadas de una población, los valores de sus parámetros se estiman a partir de los correspondientes estadísticos mediante técnicas estadísticas inferenciales.

Estas últimas se basan en los conceptos que se analizaron en el capítulo anterior, en el cual se concluyó que la media aritmética o la tasa de frecuencia en una categoría registrada en una muestra, en promedio, estiman de forma correcta los correspondientes valores de la población.

Esto significa que se podría intentar estimar la proporción (o porcentaje) de datos en una determinada categoría, o la media aritmética en una población, al tomar como base el conocimiento de que tiene alguna relación con el valor de la proporción o de la media aritmética ($\bar{x}$), calculada en la muestra que se utiliza en la labor de investigación.

Sin embargo, al proceder de esta manera no es posible tener mucha "confianza" en la estimación realizada. Puede haberse tenido la "suerte" suficiente para extraer un subconjunto de los integrantes de una población, una muestra, en la que se manifieste esa situación; aunque, a menos que en ella no haya dispersión o la muestra haya sido infinitamente grande, también puede haberse tenido la "mala suerte" de que esos estadísticos sobreestimen o subestimen los parámetros de la población.

La situación podría asemejarse a la "confianza" que se puede tener de "ganar un sorteo" mediante la adquisición de un número entre todos los que se sortearán.

Si el total es 100 y tenemos en nuestro poder uno, podríamos decir que tenemos una "confianza" de uno en cien (0,01 o 1%) de "ganar el premio". Al conseguir dos o más números podremos duplicar o aumentar nuestra "confianza", aunque para transformarla en "seguridad" de ganar sería necesario disponer en nuestro poder de la totalidad de los números.

Las técnicas de la estadística inferencial que se emplean en la investigación descriptiva se basan en el principio de estimación de los valores del parámetro de una población dentro de un intervalo.

—

Ese intervalo numérico se calcula de tal forma que el investigador puede tener una determinada confianza, aunque no la seguridad, de que el valor buscado se encuentra dentro de él.

En el próximo apartado se analizarán y fundamentarán los procedimientos a partir de los cuales se calculan los denominados intervalos de confianza para la estimación de la media aritmética de una población.

INTERVALOS DE CONFIANZA: DATOS NUMÉRICOS

Fundamentos

El **cuadro 8-1** incluye una serie de números que pueden considerarse los datos numéricos para una determinada variable, que corresponden a los integrantes de una población.

Al procesar estos datos se pueden obtener los valores 150 para la media aritmética **(μ) y 15 para la desviación estándar (σ), respectivamente.**

Si se toman muestras aleatorias de tamaño (n) 25 a partir de esa población, se puede estimar que en ellas el valor de $\bar{x}$ tenderá a ubicarse cercano a 150 la mayor parte de las veces. También puede estimarse que en ocasiones superará el valor 150 y en otras estará por debajo de él, situación que está dada por la distribución de los valores de x alrededor del valor de la media de la población original.

Tal como se mencionó en el **capítulo 7, Muestreo, la forma de esa distribución es** gaussiana y la medida de su dispersión está dada por el "error estándar" (desviación estándar dividida por la raíz cuadrada del tamaño de la muestra), que en este caso es 3 ($15 / \sqrt{25}$).

Esto significa que, de las muchas muestras de tamaño 25 que pueden obtenerse, es de esperar que el 95% tenga valores de $\bar{x}$ entre 144 y 156 aproximadamente. Estos valores corresponden a dos errores estándar ($2 \times 3 = 6$) por debajo y por arriba del valor de la media de la población, y dentro de los que la resolución de la ecuación indica que se ubica aproximadamente el 95% del área total. Recuérdese que el valor exacto para la ecuación gaussiana en este caso sería 1,96 (valor de z que corresponde a los percentiles 2,5 y 97,5 cuando tiene signo negativo o positivo, respectivamente).

En resumen, si se hiciera una "apuesta" en la cual se indica que el valor que se obtendrá al tomar una muestra de tamaño 25 de la población del **cuadro 8-1** estará entre 144 y 156, puede tenerse una "confianza"

CUADRO 8-1. DATOS DE UNA POBLACIÓN CON $\mu = 150$ Y $\sigma = 15$

149	185	133	165	160	169	149	174	143	136	148	131
154	162	134	148	153	155	178	152	145	130	181	131
150	144	126	141	143	173	150	137	140	156	148	136
147	150	157	133	161	131	141	131	164	133	136	158
141	169	137	155	115	142	164	148	147	149	125	147
140	134	147	169	127	166	143	124	144	145	170	142
144	123	156	159	147	166	157	124	152	128	153	179
158	155	145	160	128	127	157	147	170	144	140	154
135	161	140	189	147	157	160	149	149	144	166	131
127	158	154	164	139	147	150	153	164	133	144	170
161	141	146	132	169	166	150	137	183	145	163	131
158	175	146	148	150	160	152	164	153	128	160	131
150	144	134	157	126	153	151	152	156	157	160	139
143	133	168	118	159	120	158	154	170	173	172	142
161	133	147	164	154	123	174	166	142	139	168	133
132	155	134	149	160	150	144	136	146	154	149	140
155	154	148	151	158	114	169	156	150	173	154	147
139	133	149	176	147	164	156	161	191	143	143	135
144	132	141	147	138	157	148	145	143	159	167	164
155	165	143	153	157	150	131	159	145	161	171	169
157	144	187	162	158	125	130	165	145	167	168	145
155	144	136	145	161	129	136	142	143	163	146	126
169	164	142	173	158	146	155	111	168	159	153	144
152	156	141	172	145	163	138	142	140	132	159	154
116	137	148	154	136	179	172	153	144	127	168	144
144	182	138	144	171	142	173	149	165	132	162	144
149	175	129	140	154	145	140	131	157	141	140	171
170	183	127	159	147	149	156	152	146	160	142	139
143	146	150	132	160	148	167	143	128	168	174	130
144	163	166	182	141	128	143	167	176	173	165	144

(continúa)

CUADRO 8-1. DATOS DE UNA POBLACIÓN CON $\mu = 150$ Y $\sigma = 15$ (*CONTINUACIÓN*)

145	144	149	183	148	141	134	139	133	131	144	148
137	164	163	154	136	157	165	138	134	141	174	169
184	155	178	126	166	135	136	144	137	154	174	166
164	151	155	136	168	153	145	135	160	150	134	130
154	154	156	122	145	129	171	151	163	147	151	162
142	140	170	149	147	153	174	149	164	147	139	153
127	146	151	131	134	141	168	168	157	141	170	156
130	140	142	136	131	138	146	153	131	123	160	163
158	129	136	123	146	110	142	128	163	173	127	124
162	160	168	160	141	147	166	151	140	153	155	149
149	138	165	149	160	164	161	179	136	142	157	157
153	134	144	152	135	175	152	140	140	157	155	172
134	157	151	185	150	160	123	152	141	145	143	147
152	158	156	151	132	178	145	143	156	180	141	141
146	132	138	175	136	156	125	138	135	158	110	174
170	158	168	150	164	149	154	111	139	143	150	115
165	153	152	140	159	146	153	164	140	134	135	167
161	127	144	152	157	148	138	166	180	147	125	174
121	165	159	162	131	177	155	152	134	157	166	163
167	137	141	159	137	163	121	171	183	151	114	136

del 95% de ganarla. Se estaría apostando a un resultado que se produce el 95% de las veces, si se repitiera el procedimiento.

Si se quisieran probar empíricamente estas consideraciones, se deberían seleccionar varias veces y en forma aleatoria 25 valores del **cuadro 8-1** (p. ej., al ubicar el extremo de un lápiz sobre ese cuadro) y calcular el valor de la media aritmética de cada una de las muestras obtenidas. Podrá verse que la mayoría de las veces (casi todas) esos valores estarán dentro del intervalo indicado.

Como ya se habrá notado, la situación de este ejemplo es distinta a la que se plantea al realizar una investigación real. En ella, el investigador no toma muestras repetidas de la población, sino una sola y de un determinado tamaño. Sin embargo, puede tener una "confianza" del 95% de que esa muestra seleccionada tiene un valor de media aritmética que está dentro de dos errores estándar en menos y en más del valor de la media aritmética de la población. Si se repitiera el procedimiento muchas veces, 95

de cada 100 veces el intervalo construido incluiría el valor del parámetro que se quiere estimar.

Si procede, entonces, a restar y a sumar el equivalente a dos errores estándar al valor de la x de su muestra, obtendrá un intervalo dentro del que podrá decir con un "95% de confianza" que estima que se encuentra la media aritmética de la población que se quería describir.

Realice este procedimiento con los resultados que haya obtenido al tomar muestras de la población del cuadro. Es decir, sume y reste 6 (dos errores estándar) a cada valor de x que haya calculado. Los intervalos obtenidos incluirán el valor 150 (media aritmética de la población), excepto cuando por "mala suerte" haya obtenido una muestra con $\bar{x}$ menor que 144 o mayor que 156.

Para esta situación planteada, el denominado **margen de error** calculado es 6: el resultado de multiplicar 2 (el valor de z o la cantidad de errores estándar asociada con la confianza fijada en 95%) por 3 (el valor de error estándar calculado para el tamaño de muestra, que se fijó en 25).

Un intervalo de confianza se calcula al sumar y al restar el margen de error al valor del estadístico de la muestra obtenida. Esta es una cantidad (valor z) de errores estándar asociada con la confianza que se desee tener en la estimación.

—

Valor de "t" de *Student*

Es posible que ya se haya notado que al intentar aplicar el procedimiento descrito para establecer un intervalo de confianza para estimar la media aritmética de una población en una situación real surge una dificultad que parece ser insalvable.

En esa situación se toma una muestra de una población de la que no se conoce ninguno de sus parámetros. Así, para obtener un intervalo de confianza se debe sumar y restar al valor de la media aritmética de esa muestra una cantidad determinada de errores estándar, 1,96 (o de modo aproximado 2), si se desea trabajar con una confianza del 95%.

La dificultad surge porque, para obtener el valor del error estándar, es necesario dividir el valor de la desviación estándar de la población por la raíz cuadrada del tamaño de la muestra. Esta segunda cifra es conocida por quien tomó la muestra y surge del número de datos disponibles; el numerador, en cambio, es desconocido.

La única forma de salvar este inconveniente consiste en trabajar en forma exclusiva con lo único que se dispone: los datos de la muestra. Se puede calcular la desviación estándar de estos (recuérdese que el denominador en este caso está dado por los grados de libertad, n – 1), pero el valor resultante no es el parámetro que mide la dispersión en la población, sino un estadístico que lo estima.

Al ser lo único disponible, no parece irracional calcular una estimación del error estándar real mediante la división del valor de esa desviación estándar de la muestra por la raíz cuadrada del tamaño de la muestra ($s \sqrt{n}$).

Así, un intervalo puede calcularse al sumar y al restar una cierta cantidad de esa estimación del error estándar a la media aritmética de la muestra, aunque con el reconocimiento de que la "confianza" que se puede tener en que ese intervalo incluya al parámetro de tendencia central de la

población no es la misma que cuando se dispone del valor real de la dispersión.

Para recalcar el concepto: si se suman y restan dos estimaciones del error estándar (calculadas a partir de la desviación estándar de la muestra), se obtiene un intervalo que nos brinda una "confianza" menor que el 95% de incluir en él la media aritmética de la población.

A fin de compensar esta pérdida de confianza, el procedimiento que se emplea al trabajar solo con datos de una muestra consiste en sumar y restar una cantidad mayor de errores estándar estimados que la que se utilizaría si se conociera su valor real; es decir, que se ajusta el valor de z empleado en el cálculo.

El matemático W. S. Gosset, interesado en la estadística, profundizó el estudio del *plus* de conocimientos que surgía del análisis de la distribución de Gauss. Por alguna razón denominó "t" al valor que surge de una distribución derivada de la gaussiana, pero que es aplicable al trabajo con muestras y, por ende, con la estimación del error estándar.

No publicó sus conclusiones con su nombre, sino mediante el seudónimo ***Student*** y, por ello, hoy se hace referencia a esos valores con la denominación de **"t" de *Student***.

El valor de "t" que debe utilizarse depende del tamaño de la muestra que se haya tomado de la población y de la confianza con la que se desee trabajar.

El **cuadro 8-2** tiene tres columnas. La primera tiene el encabezado de "grados de

CUADRO 8-2. ALGUNOS VALORES DE LA DISTRIBUCIÓN DE "T" DE *STUDENT*

Grados de libertad	*P* = 0,05	*P* = 0,01
1	12,706	63,657
2	4,303	9,925
3	3,182	5,841
4	2,776	4,604
5	2,571	4,032
6	2,447	3,707
7	2,365	3,499
8	2,306	3,355
9	2,262	3,250
10	2,228	3,169
11	2,201	3,106
12	2,179	3,055
13	2,160	3,012
14	2,145	2,977
15	2,131	2,947
16	2,120	2,921
17	2,110	2,898
18	2,101	2,878
19	2,093	2,861
20	2,086	2,845
21	2,080	2,831
22	2,074	2,819
23	2,069	2,807
24	2,064	2,797
25	2,060	2,787
30	2,042	2,750
60	2,000	2,660
Infinito	1,960	2,576

libertad". Esto significa que de las diferentes filas que se incluyen será necesario buscar aquella que corresponda a los grados de libertad de la muestra con la que se esté trabajando. Para el ejemplo del apartado anterior, significa que se debería buscar la información en la fila 24, ya que las muestras con las que se había trabajado eran de n = 25.

Las otras dos columnas están encabezadas por la letra ***P***, que indica probabilidad. En la segunda se indica 0,05 y en la tercera, 0,01, lo que es equivalente a 5 y 1%, respectivamente. Esto significa que los valores en ellas corresponden a la posibilidad de error que se está dispuesto a aceptar. Como se deducirá, esto representa buscar el valor en la columna 0,05 si se desea tener una "confianza" del 95%.

En definitiva, si se tomara una muestra de tamaño 25 de la población del **cuadro 8-1** y solo se dispusiera de los datos de esa muestra (es decir, que no se conociera la desviación estándar de la población), la cantidad de errores estándar estimados para sumar y restar a la media de la población –para el cálculo del margen de error– sería 2,064 para estimar la media aritmética de la población con una confianza de 95%. Este número es el que aparece en el **cuadro 8-2** en la intersección de la fila correspondiente a 24 grados de libertad (25 – 1) y la columna encabezada por ***P*** = 0,05.

En resumen, para calcular un intervalo de confianza para la media aritmética de una población a partir de una muestra, el procedimiento consiste en:

a) **Calcular la media aritmética y la desviación** estándar de la muestra.
b) **Calcular la estimación del error estándar** a partir de la desviación estándar de la muestra y su tamaño.
c) **Buscar el valor de "t" correspondiente,** según los grados de libertad que da la muestra y la confianza deseada en la estimación.
d) **Calcular el margen de error al multiplicar** el valor del error estándar estimado a partir de la muestra por el valor de "t" encontrado en (c).
e) **Calcular los límites inferior y superior** del intervalo al restar y sumar al valor de la media aritmética de la muestra el valor del margen de error.

Como ejemplo, al trabajar con redondeo a dos cifras decimales para el caso de una muestra de tamaño 15 con los siguientes valores para cada dato: 656, 631, 613, 635, 656, 618, 624, 613, 618, 615, 587, 666, 639, 612 y 645.

a) **Calcular la media aritmética = 628,53 y** la desviación estándar = 21,13.
b) **Calcular el error estándar estimado: 5,45** (21,13 / $\sqrt{15}$).
c) **De la tabla surge que el valor para 95%** de confianza y 14 grados de libertad es 2,145.
d) **Calcular el margen de error = 5,45 ×** 2,145 = 11,69.
e) Calcular:
Límite inferior = 628,53 – 11,69 = 616,84.
Límite superior = 628,53 + 11,69 = 640,22.

En resumen, se puede decir que se estima con un 95% de confianza que el valor del parámetro media aritmética de la población de la que se tomó la muestra está entre 616,84 y 640,23, o bien que se estima que el parámetro está entre 628,53 ± 11,69.

En la práctica, estas operaciones se hacen en forma automatizada mediante programas informáticos para cálculos estadísticos

y algunas planillas de cálculo. En estos casos solo es necesario ingresar los correspondientes datos e indicar el nivel de confianza con el que se quiere calcular el intervalo. Lo usual es trabajar con un nivel del 95%, pero es posible utilizar otro si el investigador lo desea.

INTERVALOS DE CONFIANZA: DATOS NOMINALES

En el caso de que se utilicen datos nominales para la descripción de la variable, el objetivo es estimar la proporción "p" o el porcentaje correspondiente a una determinada categoría (proporción o porcentaje de enfermos, de mujeres, de posibles votantes por un candidato, etc.).

Los fundamentos del procedimiento para calcular un intervalo de confianza son los mismos que aquellos en los que se basó el trabajo con datos numéricos.

Los procedimientos inferenciales, en el caso de datos nominales, cambian en función de que la distribución de la estimación del valor del parámetro de la población que se obtiene de las muestras no es gaussiana, sino binomial.

—

El trabajo se simplifica si se dispone de tablas en las que los límites inferior y superior de los intervalos de confianza para las distintas situaciones que se pueden presentar ya han sido calculados. El **cuadro 8-3**, por ejemplo, muestra los intervalos calculados para situaciones que se pueden presentar al tomar muestras de tamaño 40.

La columna "f(x)" muestra la frecuencia con que se pueden encontrar datos en una categoría en esas condiciones y se muestran valores entre 0 y 20 (de 0 a 20 enfermos, mujeres, posibles votantes por un candidato, etc., en la muestra).

En la columna "% en muestra" se incluye la tasa porcentual correspondiente a la frecuencia observada, mientras que las dos columnas restantes se refieren a los límites inferior y superior del correspondiente intervalo con 95% de confianza.

CUADRO 8-3. LÍMITES PARA INTERVALOS DE CONFIANZA (95%) PARA ESTIMAR UNA TASA PORCENTUAL A PARTIR DE MUESTRAS CON N = 40

f(x)	% en muestra	Límite inferior	Límite superior
0	0,00	0,00	8,81
1	2,50	0,06	13,16
2	5,00	0,61	16,92
3	7,50	1,57	20,39
4	10,00	2,79	23,66
5	12,50	4,19	26,80
6	15,00	5,71	29,84
7	17,50	7,34	32,78
8	20,00	9,05	35,65
9	22,50	10,84	38,45
10	25,00	12,69	41,20
11	27,50	14,60	43,83
12	30,00	16,56	46,53
13	32,50	18,57	49,13
14	35,00	20,63	51,68
15	37,50	22,73	54,20
16	40,00	24,86	56,67
17	42,50	27,04	59,11
18	45,00	29,26	61,51
19	47,50	31,51	63,87
20	50,00	33,80	66,20

Si en una muestra de ese tamaño (n = 40) se registrara la presencia de 14 enfermos (o cualquier otra condición de interés), la lectura de la tabla indicaría que se puede estimar con 95% de confianza que la tasa de "presencia de enfermedad" (o de la condición en estudio) en la población está entre el 20,63 y el 51,68%.

Diversos programas informáticos de cálculo estadístico permiten calcular también los intervalos de confianza para proporciones o tasas.

ESTIMACIÓN DEL TAMAÑO DE LA MUESTRA

Si se analizan las operaciones numéricas que se siguen para el cálculo de los intervalos de confianza, explicitadas especialmente para el caso de datos numéricos, puede deducirse cuáles son los factores que determinan la amplitud de ese intervalo o el margen de error en la estimación.

Ese margen de error se refiere al valor de por cuánto en más o menos se estima el valor del parámetro. En el ejemplo de cálculo que se planteó este era 11,69, valor que surge al multiplicar el valor de "t" correspondiente a la confianza de la estimación (2,145) por la estimación del error estándar (5,45). Como este, a su vez, depende de la medida de la dispersión (desviación estándar) y del tamaño de la muestra, puede decirse que el margen de error con el que se estima la media aritmética de una población a partir de una muestra y mediante un intervalo de confianza depende de la confianza deseada en la estimación, de la dispersión de los datos y del tamaño de la muestra.

Por lo tanto, y en una operación del tipo resolución de ecuación despejando incógnitas, puede calcularse el tamaño de la muestra que se necesita para realizar una investigación cuyo objetivo sea estimar el parámetro de una población.

El tamaño de muestra conveniente está en función de: la confianza deseada en la estimación, el margen de error que se desea en la estimación y la medida de la dispersión esperada en los datos.

A modo de ejemplo de un trabajo con datos numéricos, supóngase que se desea estimar la media aritmética de la cantidad en gramos por litro de una sustancia en sangre en una población de pacientes con determinadas características.

Para el primero de los factores –la confianza deseada– puede seleccionarse un valor de 95%, que es el usual. Esto significa que en su momento se multiplicará el valor del error estándar, alrededor de 2, según lo que indique la tabla de "t".

Para el segundo factor habrá que considerar cuál es la precisión –cuánto más o cuanto menos– que permite obtener una información de utilidad. Consideremos como ejemplo ± 0,10.

Por último, será necesario contar con alguna estimación sobre la dispersión que puede esperarse en los datos. Esta puede surgir de la consulta de trabajos realizados, con anterioridad y en condiciones similares, por el propio investigador u otros; si no estuviera disponible solo queda la opción de realizar lo que se denomina una "prueba piloto" para obtenerla.

En el ejemplo, supóngase que es posible esperar un valor de 0,32 para la desviación estándar. En ese caso, el resultado será aproximadamente 41. Este valor surge al multiplicar el cuadrado de la desviación estándar

esperada (0,32) por el cuadrado del valor de "t" que se supone se utilizará (en principio, 2 para el 95% de confianza) y dividir por el cuadrado de la precisión deseada (0,10).

El resultado debe ser considerado una aproximación, y en el trabajo real seguramente se utilizará una muestra algo superior, para mayor tranquilidad en el futuro logro del objetivo buscado.

Como es de suponer, hoy en día se dispone de programas informáticos que, una vez ingresada toda esta información, realizan los cálculos que arrojan como resultado el tamaño de muestra necesario para la situación en particular.

En el caso de los datos nominales, el procedimiento es similar, aunque con una simplificación. Los valores necesarios para el cálculo son: la confianza deseada en la estimación, la precisión que se desea para la estimación (el margen de error) y la proporción esperada en la población.

Este último valor es el que determina la dispersión de la distribución de la proporción en las muestras (véase **cap. 7, Muestreo**). Si no se dispone de información al respecto, puede considerarse que la proporción esperada es 0,5 (50%), ya que esta representa la situación más desfavorable y la que obliga a trabajar con las muestras de mayor tamaño. Al realizar los cálculos en función de este dato, el resultado corresponderá a una muestra que puede resultar algo más grande de lo necesario, pero nunca más pequeña.

De nuevo, en este caso de cálculo de tamaños de muestra para la estimación de proporciones o porcentajes, se dispone de programas informáticos que procesan la información de manera automática.

Es de interés mencionar en este momento que se han descrito aquí principios y procedimientos para las dos situaciones más frecuentes en la investigación descriptiva: estimación de media aritmética y proporciones (porcentajes). Es posible calcular también intervalos de confianza para estimar otros parámetros que describen otras propiedades de las poblaciones mediante los mismos principios: las muestras tienden, en promedio, a reproducir lo que pasa en la población de la que provienen.

SÍNTESIS CONCEPTUAL

- **La estadística inferencial se emplea en** la investigación descriptiva para estimar los valores de parámetros de una población dentro de un intervalo, conocido como intervalo de confianza, a partir de los datos de una muestra.
- **Un intervalo de confianza se calcula al** sumar y restar el margen de error al valor del estadístico de la muestra obtenida.
- **El margen de error es una cantidad de** errores estándar asociada con la confianza que se desea tener en la estimación.
- **El valor del margen de error está determinado** por la confianza deseada en la estimación, la dispersión esperada en los datos y el tamaño de la muestra seleccionada.
- **Puede calcularse el tamaño de muestra** conveniente a partir de la dispersión estimada en los datos, la confianza con la que se desea realizar la estimación del parámetro y el margen de error que se considera apropiado.

EJEMPLO 8-1

En una muestra de 350 mujeres se evaluó la edad en la que se presentaron los primeros síntomas de osteoporosis.

En esa muestra se obtuvieron los siguientes estadísticos: media aritmética 48,2 años y desviación estándar 10,2 años.

A partir de estos datos, ¿qué estimación con 95% de confianza puede hacerse respecto del parámetro media aritmética de esa población?

El correspondiente intervalo de confianza tiene como límite inferior 47,1 y como límite superior 49,3 años, valores obtenidos al sumar y restar dos errores estándar (el margen de error, o sea: $2 \times 0{,}55$) a la media aritmética de la muestra. En resumen, puede estimarse con 95% de confianza que el parámetro de la población está entre 47,1 y 49,3.

EJEMPLO 8-2

¿Cuál hubiera sido la estimación si el tamaño de la muestra hubiese sido 25?

En este caso, el error estándar sería 2,04, y debe sumarse y restarse a la media aritmética de la muestra 2,064, según lo que indica la distribución de "t" para 24 grados de libertad (véase cuadro 8-2).

En definitiva, el intervalo para 95% de confianza indicaría que puede estimarse que el parámetro de la población está entre 44 y 52,4.

Este intervalo tiene una amplitud mayor (margen de error mayor) que el calculado en el ejemplo 8.1, dado el menor tamaño de muestra utilizado.

CAPÍTULO

9

PRUEBA DE HIPÓTESIS: GENERALIDADES

INTRODUCCIÓN

La investigación científica busca la generación de nuevos conocimientos y, en el campo de las ciencias tácticas, se lleva a cabo, principalmente, mediante la aplicación del método hipotético deductivo.

Este método trabaja mediante el planteo de un enunciado que se supone que constituye la respuesta a un interrogante determinado. Este enunciado de veracidad supuesta, aunque no conocida, denominado hipótesis, debe permitir la deducción de consecuencias que puedan ser contrastadas de manera empírica.

Esto significa que, a partir de una hipótesis planteada, se debe poder deducir cómo se producirían determinados hechos si esta hipótesis es verdadera.

La contrastación empírica consiste en generar una situación que haga posible observar la forma en la que en realidad se producen esos hechos.

Si el resultado de ese "experimento" muestra que los hechos se producen de la forma en la que se los dedujo en la hipótesis, esta se acepta como verdadera; si, por el contrario, los hechos observados no se corresponden con los esperados, la hipótesis se rechaza.

Una dificultad que surge en la aplicación de esta metodología en el campo de las ciencias de la salud consiste en que las hipótesis que se formulan se refieren a situaciones que se plantean en poblaciones de gran tamaño. Por ejemplo, conciernen a las causas que motivan un determinado estado patológico o los resultados de la aplicación de una técnica diagnóstica, preventiva o

terapéutica, en grandes grupos de individuos con determinadas características.

Esto hace que la valoración de cómo se produce la totalidad de los hechos sea imposible, puesto que no es posible concretar un experimento que incluya a la totalidad de los individuos de una población.

Al igual que en el caso de la investigación descriptiva, el trabajo se realiza al evaluar los resultados que se pueden observar en una muestra. Esto crea un grado de incertidumbre en el momento de tomar la decisión de aceptar o rechazar la hipótesis planteada.

Los motivos de esa incertidumbre se encuentran en conceptos cubiertos en los dos últimos capítulos. El comportamiento de una muestra tiende a reproducir al de la población, que en promedio lo reproduce, pero en ocasiones lo subestima y en otras lo sobreestima, a menos que la población sea totalmente homogénea, sin dispersión, o que la muestra incluya a la totalidad de los integrantes de la población.

Para ejemplificar estas consideraciones se planteará un ejemplo simple. Supóngase que se dispone de una moneda y es necesario decidir si esta se encuentra correctamente balanceada o no en lo que respecta al conjunto de hechos que se producen al arrojarla al aire. Los hechos que se mencionan se refieren al resultado que se produce al caer la moneda sobre una superficie plana.

Este hecho, o variable de interés, puede evaluarse con un dato nominal dicótomo: cara o ceca. En función de ello puede formularse una hipótesis que describa, en términos de ese dato, lo que se supone que debe encontrarse en la correspondiente población de hechos: la proporción de datos "cara" (o "ceca") es igual a 0,5 ($p = 0{,}5$) o, lo que es lo mismo, el porcentaje de resultados cara es 50%.

De esa hipótesis pueden extraerse consecuencias deductivas: si la moneda se arroja al aire la totalidad de las veces que es posible hacerlo, es decir si se genera la población de interés, en ella se encontrará un valor de 0,5 para la proporción del dato "cara". Expresado de otra manera, en el 50% de la totalidad de las veces que se arroje la moneda, el resultado será "cara".

Se entenderá que no es factible realizar la contrastación empírica de esta hipótesis en la población, ya que esto significaría arrojar la moneda infinitas veces, cosa que no es posible hacer.

Ante esa dificultad, la alternativa consiste en realizar el experimento con una muestra: arrojar la moneda una cantidad determinada de veces. Esto significa generar una muestra de un determinado tamaño, n = 40, por ejemplo.

Concluido el experimento se tendrá un conjunto de 40 datos nominales, cara o ceca, a partir de los cuales será necesario tomar la decisión de aceptar o rechazar la hipótesis planteada: $P = 0{,}5$.

Si el "experimento" arrojara un resultado de 21 datos cara y 19 ceca, un análisis –por el momento intuitivo– orientará con un alto grado de probabilidad a no arriesgar un rechazo a esa hipótesis. Lo que sucede en este tipo de "juegos" indica que, aun cuando la moneda no esté "cargada", no es tan infrecuente que se observe ese resultado, que no se aleja mucho de lo esperado. Por lo tanto, proceder al rechazo de lo planteado inclina a no "confiar" en la decisión y pensar en que es posible que se esté cometiendo el error de rechazar una hipótesis que podría ser verdadera.

Si, en cambio, el resultado registrado fuera de 40 datos cara y ninguno ceca, con ese mismo análisis intuitivo no habría muchas

vacilaciones antes de declarar el rechazo de la hipótesis. La decisión sería tomada con bastante "confianza", aunque podría significar que se está cometiendo el error de rechazar una hipótesis verdadera: el resultado observado en este caso no es imposible, pero sí poco frecuente.

Como se ve, la decisión se ha tomado en cada caso según la "razonabilidad" de la correspondencia entre el resultado observado y el esperado, deducido a partir de la hipótesis.

Ninguna decisión tomada a partir de una muestra está exenta de error: la hipótesis no rechazada podría ser falsa y la hipótesis rechazada podría ser verdadera.

—

ERRORES DE TIPO I Y II

El planteo frente a los datos del supuesto experimento con una moneda es el mismo que hace un investigador frente a los resultados de cualquier experimento en el que obtuvo datos a partir de muestras.

Frente a esos datos se encuentra en la situación que se resume en el **cuadro 9-1.**

CUADRO 9-1. SITUACIÓN ANTE RESULTADOS DE UN EXPERIMENTO PARA LA CONTRASTACIÓN EMPÍRICA DE UNA HIPÓTESIS

	Hipótesis verdadera	Hipótesis falsa
Hipótesis aceptada		Error de tipo II $P = \beta$
Hipótesis rechazada	Error de tipo II $P = \beta$	

En las columnas de ese cuadro de doble entrada se incluyen las dos condiciones que pueden darse para la hipótesis planteada respecto de una variable de interés científico: puede ser verdadera o falsa.

En las filas se incluyen las dos decisiones que puede tomar el investigador luego del análisis de los datos obtenidos: aceptar o rechazar la hipótesis.

Las cuatro celdas del cuadro muestran las cuatro situaciones que pueden generarse a partir de lo anterior.

Si la decisión es de rechazo y la hipótesis es falsa, se estará frente a una situación de ausencia de error; aunque, si es verdadera, se habrá cometido un error consistente en rechazar una hipótesis que es verdadera, lo que se denomina **error de tipo I.**

Si la decisión es de aceptación y la hipótesis es verdadera, se estará frente a una situación de ausencia de error; aunque, si es falsa, se habrá cometido un error consistente en aceptar una hipótesis que es falsa, lo que se denomina **error de tipo II.**

No es posible tener la confianza absoluta –seguridad– de no haber cometido un error, ya que la única forma de lograrlo sería tener el conocimiento real de la población, una situación imposible en las poblaciones de interés científico.

—

SIGNIFICADOS DE ALFA Y BETA

Si bien no es posible tener la seguridad absoluta de una ausencia de error en la decisión tomada respecto de una hipótesis, sí es posible fijar algún criterio para orientar en cuanto a su aceptación o su rechazo. En realidad, esos criterios ya se siguieron en la

toma intuitiva de decisión en el experimento realizado con la moneda.

La decisión de rechazo de la hipótesis se tomó con el segundo de los resultados: 40 cara y 0 ceca. La intuición indicó que el resultado observado es muy "poco probable", aunque no imposible, si la hipótesis es verdadera. Por lo tanto, no es imposible tampoco que se esté cometiendo un error con el rechazo, aunque sí es poco probable.

La probabilidad de cometer un error de tipo I se simboliza con la letra griega α (alfa).

Cuando el resultado fue 21 caras y 19 cecas, la decisión fue de "no rechazo"; esto lleva implícita la aceptación de la hipótesis. Esta decisión no está libre de error, ya que ese resultado podría ser consecuencia de una moneda "algo cargada" y no de la mala suerte.

La probabilidad de que se esté cometiendo un error de tipo II se simboliza con la letra griega β (beta).

Las técnicas estadísticas permiten calcular los valores de alfa y beta en las distintas situaciones experimentales.

NIVEL DE SIGNIFICACIÓN Y PODER DE UN EXPERIMENTO

El criterio para seguir en la decisión de "rechazo" o "no rechazo" de una hipótesis es patrimonio de quien deba tomarla. Sin embargo, se comparten algunos principios generales que no difieren de los ya utilizados de manera intuitiva.

La decisión de rechazo solo se toma cuando los hechos observados se alejan sustancialmente de los esperados. Esto significa que la probabilidad que tienen esos hechos de producirse es baja y, por ende, alfa –la probabilidad de error de tipo I– es baja en esas circunstancias.

Esta situación se asemeja a la que se presenta en el ámbito de la justicia, cuando un juez o un jurado deben tomar una decisión respecto de un procesado por un delito. El juez o el jurado parten de la hipótesis de inocencia del acusado: toda persona es inocente hasta que se demuestre su culpabilidad.

Proceden, entonces, a analizar "las pruebas", que son los hechos que equivalen a los resultados de un experimento. Si esos hechos, "pruebas", indican que el acusado fue encontrado en la escena del delito y en condiciones que están más allá de lo que se espera en un inocente, se rechaza la hipótesis de inocencia y se lo condena.

Es probable que no se pueda tener la seguridad absoluta de que la condena haya sido correcta, aunque la decisión habrá sido tomada de manera correcta, ya que se confía en la baja probabilidad de que se esté cometiendo un error de tipo I; es decir, el valor de alfa es bajo.

Si, en cambio, existieran dudas sobre algunas de las "pruebas", no se toma la decisión de condena, sino que se acepta la inocencia, aunque esto signifique suponer que pueda estar cometiéndose un error de tipo II.

Los principios aceptados en la justicia consideran que es preferible cometer el error de dejar libre a un culpable (error de tipo II) que cometer el error de condenar a un inocente (error de tipo I).

Con ese mismo principio del orden jurídico se trabaja en la investigación científica. Solo se rechaza la hipótesis cuando el aná-

lisis de los resultados indica que la probabilidad de su ocurrencia, si fuera verdadera, es "suficientemente" baja; solo se rechaza la hipótesis cuando alfa es "suficientemente" baja.

Queda por responder qué se entiende por alfa "suficientemente" baja. La experiencia acumulada en el campo de las ciencias fácticas, y las de la salud en particular, indica que es razonable trabajar con un nivel de probabilidad de error de tipo I (alfa) de 0,05 o 5%.

Es usual tomar la decisión de rechazar una hipótesis cuando los resultados encontrados tienen una probabilidad de presentarse inferior a 0,05. Este valor es el "nivel de significación" que con mayor frecuencia se establece para la toma de la decisión de rechazo.

—

Entonces, puede decirse que, al concluir un experimento, se aplican técnicas estadísticas para establecer si la probabilidad (*P*) de obtener el resultado observado es inferior o no a un valor "crítico" preestablecido si la hipótesis planteada es verdadera (generalmente 0,05 y denominado nivel de significación).

Si el cálculo indica que *P* es inferior a 0,05, la hipótesis se rechaza y se establece que la diferencia es estadísticamente significativa (la diferencia entre lo observado y lo esperado). Si, en cambio, *P* es igual o superior a 0,05, la hipótesis no es rechazada y se considera que esa diferencia no es estadísticamente significativa.

—

En todos los casos, *P* es el valor de **alfa,** o sea, el valor de la probabilidad de cometer un error de tipo I (rechazar una hipótesis verdadera) en la toma de la decisión.

En el segundo caso, cuando la hipótesis no se rechaza, puede considerarse conveniente establecer cuál es la probabilidad de cometer un error de tipo II (aceptar una hipótesis falsa), si la hipótesis real tuviera una determinada diferencia respecto de la formulada.

Esta última probabilidad representa el valor de **beta** y su complemento (1 – p) es el denominado **poder** del experimento para declarar significativa una diferencia entre lo observado y lo esperado que se considere de interés.

En la prueba de hipótesis, las técnicas estadísticas se utilizan para calcular el valor de alfa y así resolver sobre el "rechazo" o el "no rechazo" de una hipótesis, en función de que esté por debajo o que supere el nivel de significación.

En el caso de "no rechazo", se pueden emplear para el cálculo del poder del experimento realizado para detectar situaciones de diferencia que puedan ser de interés.

—

Si se detecta que ese poder es muy bajo (como referencia se puede indicar inferior a 0,8 u 80%), la decisión debe considerarse provisoria y se debe analizar la necesidad de ampliar o modificar el experimento realizado.

En los próximos capítulos se presentarán los principios de aplicación de las técnicas estadísticas más frecuentemente utilizadas en la prueba de hipótesis.

SÍNTESIS CONCEPTUAL

- **La prueba de hipótesis se basa en la** aplicación de las técnicas estadísticas para evaluar la probabilidad de encontrar el resultado de un "experimento" si la hipótesis planteada es verdadera.
- **Si esa probabilidad es baja (gene-**ralmente inferior a 0,05), se toma la decisión de rechazo al considerar que la probabilidad de que se esté cometiendo un error de tipo I (alfa) es baja.
- **Cuando una hipótesis no se rechaza,** puede ser conveniente calcular la probabilidad de que se esté cometiendo un error de tipo II (beta).
- **Un valor elevado de beta puede indicar** que el diseño de la investigación no tiene poder para encontrar diferencias de interés.

CAPÍTULO

10

PRUEBA DE "T"

INTRODUCCIÓN

Cuando la variable sobre la que se formula una hipótesis se evalúa a través de datos numéricos, su formulación puede representarse mediante alguna aseveración respecto de un parámetro de la población.

Como ejemplo, considérese esta situación específica: se necesita establecer si un proceso de fabricación permite la obtención de lotes de comprimidos que contengan 500 mg de un determinado fármaco. Este valor, 500, debe considerarse como la media aritmética de los comprimidos que componen el lote, ya que es razonable esperar una cierta y aceptable variación en el proceso.

Para resolver el interrogante planteado (¿el proceso permite obtener una población con una media aritmética igual a 500 mg para la variable "contenido de medicamento"?), siguiendo los pasos del método hipotético deductivo, debe formularse una hipótesis que permita extraer consecuencias deductivas.

En este caso, la hipótesis puede formularse diciendo –en símbolos–: $\mu = 500$ mg. La decisión sobre su aceptación o rechazo se tomará en función de su contrastación empírica. Este enunciado permite, como lo requiere el mencionado método, una deducción unívoca: de ser cierta, solo puede deducirse que la media tiene exactamente ese valor. Si la hipótesis planteara que $\mu \neq 500$ mg, este principio no se verificaría, ya que la deducción sería que la media es cualquier valor excepto el indicado.

Como es de suponer, para la contrastación empírica se realizará un "experimento" con una muestra de comprimidos tomada de manera aleatoria del lote, que es la población de interés.

Una situación posible sería tomar 14 comprimidos, es decir, una muestra con $n = 14$, en los cuales podrían registrarse los siguientes valores de contenido de medicamento en mg: 498, 497, 500, 501, 496, 500, 497, 496, 496, 499, 498, 498, 499 y 501.

La media aritmética ($\bar{x}$) de esa muestra es 498,29, que es menor que el valor para la media aritmética de la población que se planteó en la hipótesis.

En función de los criterios que se analizaron en el capítulo anterior, esa falta de concordancia llevará al rechazo de la hipótesis si, y solo si, la probabilidad de que se obtenga un resultado como el del experimento es inferior a un nivel crítico o nivel de significación que podría establecerse en 0,05. Este valor, a su vez, corresponde a la probabilidad de error de tipo I (α) que se está dispuesto a aceptar en caso de rechazo.

Para decidir si la situación observada corresponde o no a una probabilidad inferior a ese nivel, se tiene en cuenta que, si la hipótesis es verdadera, los valores de las medias de las muestras de tamaño 14 tomadas de la población tienen, en promedio, un valor de 500 mg. Por otro lado, en alguna de esas muestras la media aritmética es menor y en otras, mayor que ese valor, aunque en el 95% de ellas la diferencia con respecto a 500 no es mayor que 1,96 errores estándar. De nuevo, esta situación está determinada porque la distribución de las medias aritméticas de las muestras tomadas de una población sigue una distribución gaussiana y, en función de ello, el valor z 1,96 negativo corresponde al percentil 2,5 y el 1,96 positivo, al percentil 97,5.

Esto significa que, si la diferencia entre la media aritmética de la muestra obtenida, 498,29 mg, y la de hipótesis, 500 mg, supera los 1,96 errores estándar podrá considerarse que se está ante una situación que tiene una probabilidad de ocurrir inferior a 0,05 o 5%. Entonces, podrá procederse a rechazar la hipótesis con un nivel de probabilidad de error de tipo I (α) inferior a 0,05, que es el nivel de significación fijado.

—

En el ejemplo planteado, como en los casos de investigación habitual, surge la dificultad del desconocimiento del error estándar, ya que este se establece a partir de la desviación estándar de la población, un valor desconocido a menos que se pueda evaluar a todos los integrantes de la población, y la raíz cuadrada del tamaño de la muestra.

Al igual que en el caso del cálculo de intervalos de confianza, la única alternativa posible es estimar el error estándar a partir de los únicos datos disponibles, que son los de la muestra. Esto obligará a no tomar en consideración para el rechazo el valor de 1,96 errores estándar, sino "un poco más", lo que está establecido por la distribución de "t" de Student.

En el ejemplo, la desviación estándar de la muestra (s) es 1,77, que dividida por la raíz cuadrada de 14 determina el valor 0,47 para el error estándar estimado.

En función de ello, se puede estimar que la diferencia entre la media aritmética de la muestra y la de la hipótesis, –1,71 (498,29 – 500) está –3,62 (–1,71 / 0,47) errores estándar estimado de la media de la hipótesis.

El **cuadro 8-2** (**cap. 8, Estimación de parámetros**) muestra los valores de "t" para diferentes grados de libertad. En el experimento realizado, los grados de libertad fueron 13 (14 – 1); esto corresponde al denominador de la fórmula de cálculo de la variancia y la desviación estándar de la muestra.

El citado cuadro muestra que: para 13 grados de libertad, el valor t = 2,160 es el que

debe superarse en valor absoluto –o sea, negativo o positivo– para que la situación observada esté por fuera del 95% central más frecuente o, lo que es lo mismo, tenga una probabilidad de ocurrir inferior a 0,05.

El valor obtenido, que se considera el valor de "t" obtenido y que en el ejemplo fue –3,62, supera en valor absoluto al indicado en el cuadro para $P = 0,05$.

Esto significa que la situación observada tiene una menor probabilidad de ocurrencia en ella y, según las condiciones establecidas, debe rechazarse la hipótesis y decirse que la diferencia entre lo observado y lo esperado a partir de la hipótesis es estadísticamente significativa.

COMPARACIÓN ENTRE DOS GRUPOS

En una investigación es frecuente encontrar la necesidad de establecer, a partir de los resultados obtenidos en dos grupos que constituyen otras tantas muestras, si existen diferencias entre las poblaciones de las que fueron tomadas.

Esta situación se produce por la necesidad de comparar, por ejemplo, los resultados de un tratamiento determinado con los resultados que se obtienen al no aplicar ninguno, o por administrar un placebo o un tratamiento de efecto ya conocido. Otra de las situaciones análogas sería la que se produce por la necesidad de comparar la manera en la que se presenta una variable en dos poblaciones que difieren en alguna característica, como género, edad, condición social u otra.

Cuando la citada variable –que es la **variable dependiente en la situación planteada**– se evalúa mediante datos numéricos, una posible hipótesis que permite la posterior contrastación empírica consiste en asumir que no existen diferencias entre las medias aritméticas que describen la situación en ambas poblaciones.

Expresado en otros términos, en esa hipótesis se indica que la diferencia entre ambas medias aritméticas es nula y puede expresarse en símbolos de la siguiente forma: H_q: $\mu_A - \mu_8 = 0$; donde H_0 simboliza la hipótesis nula, μ_A representa la media aritmética de una de las poblaciones y μ_B, la de la otra.

Una hipótesis formulada en términos de la ausencia de diferencia entre los parámetros de las poblaciones para comparar se conoce con el nombre de hipótesis nula, y constituye el punto de partida de los experimentos más frecuentes.

—

Con esto se trata de establecer si la situación observada se produce o no conforme sería probable esperar si esa hipótesis fuera verdadera; a partir de ello, se toma la decisión de rechazo o aceptación.

Un procedimiento para realizar ese análisis es la **prueba de "t"** que es, en sus principios, igual a la ya analizada.

Véase el ejemplo que se presenta en el **cuadro 10-1**. Se muestran los resultados obtenidos en dos grupos de unidades experimentales. Pueden representar muestras de dos poblaciones que difieren en una característica, que es la variable cuyo efecto se pretende estudiar: la que en la terminología de la investigación se conoce como **variable independiente**. Esta puede estar representada por la aplicación o no de una determinada medida preventiva o terapéutica, por pertenecer a un grupo social o a otro o por alguna otra característica.

CUADRO 10-1. RESULTADOS NUMÉRICOS PARA PROBAR UNA HIPÓTESIS NULA REFERIDA A LA DIFERENCIA ENTRE LA MEDIA ARITMÉTICA DE DOS POBLACIONES

Grupo A	Grupo B	
133	126	
135	129	
131	132	
130	127	
129	128	
133	130	
132	128	
134	127	
134	131	
130	128	
	129	
Media aritmética	A = 132,1	B = 128,6
Tamaño de la muestra	A = 10	B = 11
Diferencia observada	3,5	
Suma de cuadrados	A =36,9	B = 32,5
Grados de libertad	A = 9	B = 10
Variancia ponderada	3,66	
Error estándar	A = 0,60	B = 0,58
Error estándar de la diferencia	0,84	
Valor de T	4,19	
Probabilidad	< 0,05	

Los resultados muestran que la diferencia entre las medias aritméticas de los dos grupos no fue 0, como era de esperar a partir de la hipótesis nula, sino que fue 3,5 (132,1 – 128,6).

Así, debe relacionarse la diferencia de 3,5 entre el resultado observado y el esperado a partir de la hipótesis (0), con la medida de la dispersión para obtener la magnitud de esa diferencia en cantidad de errores estándar estimados, ya que se trabajará solo con los datos de las muestras.

En los casos de comparación entre grupos, el error estándar no se calcula para cada una de las muestras en función de la desviación estándar de cada una de ellas, sino a partir de la dispersión del experimento en su conjunto, que incluye a ambas.

—

Para ello, se calcula la suma de cuadrados (suma de los cuadrados de las desviaciones de cada valor respecto de su media, como se describió en el **cap. 5, Resumen de datos numéricos**) en cada grupo. En el cuadro se muestra el resultado correspondiente.

La suma del resultado para cada muestra (36,9 + 32,5 = 69,4) es la suma de cuadrados para el conjunto del experimento, lo que, dividido por la totalidad de los grados de libertad, permitirá obtener el valor de la **variancia "ponderada"** (ya que en su cálculo tiene más "peso" la muestra de mayor tamaño) o "agrupada". En este caso, los grados de libertad son 19, 9 (10 – 1) y corresponden al grupo A y 10 (11 – 1), al grupo B. El resultado es 3,66.

La raíz cuadrada de esa variancia "ponderada" –en este caso 1,9– permite calcular la estimación del error estándar para el grupo A (1,9 / $\sqrt{10}$, el tamaño del grupo A) y para el grupo B (1,9 / $\sqrt{11}$, el tamaño del grupo B). Los valores son 0,60 y 0,58 para los grupos A y B, respectivamente.

Para obtener la medida del error estándar de la diferencia entre las medias aritméticas de ambas muestras, primero se obtiene la suma de los cuadrados de ambos errores estándar y luego se extrae la raíz cuadrada de ella.

—

En el ejemplo, $\sqrt[2]{0{,}60^2 + 0{,}58^2} = 0{,}84$. A partir de este último valor se puede calcular el valor de "t" para establecer la probabilidad de observación del resultado obtenido. La diferencia observada fue 3,5; valor que dividido por el error estándar de la diferencia (0,84) es igual a 4,19.

En el **cuadro 8-2** puede observarse que, para una situación de 19 grados de libertad (en el experimento, 9 del grupo A + 10 del grupo B), el valor 2,093 es el que separa a los valores extremos que tienen una probabilidad de ocurrir inferior a 0,05. Al ser el valor obtenido en el experimento (4,19), puede rechazarse la hipótesis nula, ya que la probabilidad de cometer, en este caso, un error de tipo I es menor que 0,05; también puede decirse que α es, en este caso, inferior a 0,05.

Esta conclusión también puede expresarse en palabras, al decir que se encontró una diferencia estadísticamente significativa entre lo observado y lo esperado, lo que en este caso corresponde a una diferencia estadísticamente significativa entre los grupos A y el B.

Por supuesto, todo el procedimiento descrito puede automatizarse en programas estadísticos o en algunas planillas de cálculos. Una vez dada la orden de procesamiento, se visualiza en la pantalla el resultado del valor de "t" final y la probabilidad asociada con él. El investigador solo tiene, entonces, que verificar si esa probabilidad es mayor o menor que el valor crítico que ha establecido y, en función de esto, rechazar la hipótesis nula planteada o no.

SIGNIFICACIÓN ESTADÍSTICA Y RELEVANCIA

Es importante destacar que la decisión tomada se relaciona con la estadística y no deben, a partir de ella, extraerse en forma directa conclusiones de aplicación práctica o clínica.

Una diferencia puede ser estadísticamente significativa y, sin embargo, no ser suficiente para tener relevancia clínica o práctica.

—

En el ejemplo que hemos presentado, la diferencia observada entre las dos muestras fue 4,29. Esto significa que, a este valor se le debe adjuntar el **margen de error** correspondiente para evaluarlo en el contexto de aplicación del conocimiento. Este, como se analizó en el **capítulo 8, Estimación de parámetros**, es una cantidad de errores estándar acordes con la confianza que se plantea para la estimación del parámetro. En los cálculos realizados, el error estándar estimado fue 0,84 (**cuadro 10-1**) y el valor de "t" de ***Student*** para 95% de confianza y 19 grados de libertad, de 2,093 (**cuadro 8-2**).

En consecuencia, puede estimarse con un 95% de confianza que se puede esperar, entre las medias aritméticas de las dos formas de manifestación de la variable independiente (entre las medias aritméticas de las poblaciones que las dos muestras representan), una diferencia entre 1,7 y 5,3

(al redondear las cifras a un decimal). Dicho de otra manera, la diferencia que se puede estimar con esa confianza es 3,5 ± 1,8.

Para establecer la relevancia clínica (o de aplicación, si no se trata de una situación clínica) se debe tener información sobre cuál es la diferencia que produce un efecto que "valga la pena" desde el punto de vista clínico o de aplicación. Si el valor de esa diferencia "clínicamente relevante" es menor que el límite inferior del intervalo de confianza calculado, se deduce que es estadísticamente significativa y clínicamente relevante. Si, en cambio, ese valor supera al del límite superior, la diferencia se considerará estadísticamente significativa y clínicamente no relevante. Por último, si la diferencia clínicamente relevante está incluida entre los límites del intervalo, se concluirá que: con los resultados obtenidos en la investigación, pudo tomarse la decisión estadística, aunque no es posible llegar a conclusiones definitivas sobre su relevancia clínica.

CONSIDERACIONES ADICIONALES

La prueba de "t" realizada de la forma en que se la describió más arriba es aplicable a situaciones de experimentos con dos grupos independientes. Si se utilizó un diseño experimental de grupos emparejados (p. ej., en el mismo paciente se registran datos en sus hemisectores derecho e izquierdo para constituir así los grupos A y B), el procedimiento es algo distinto, aunque el principio general no cambia.

Otro aspecto para tener en cuenta, que deriva en modificaciones al procedimiento, se produce cuando se plantea una hipótesis direccional. Esto significa plantear como hipótesis que en una población la media aritmética es igual o mayor (o menor) de 0. En este caso, la prueba se realiza a "una cola" en lugar de a "dos colas". Esta expresión hace referencia al extremo de la distribución en la que debe ubicarse el resultado del experimento para rechazar la hipótesis.

Asimismo, la hipótesis de la cual se parte puede no consistir en una diferencia nula, en el sentido estricto, sino en una diferencia de una magnitud determinada.

También debe tenerse en cuenta que la prueba de "t" presupone que ocurre una situación de homocedasticidad en el experimento.

Este término significa que la medida de la dispersión en ambas poblaciones, la variancia, es igual. Por lo tanto, puede ser conveniente –en especial, cuando los grupos tienen un tamaño diferente– analizar si se puede aceptar una hipótesis de igualdad de variancia entre ambas poblaciones ante los resultados del experimento (en el capítulo siguiente se explicará la manera de hacerlo).

En el caso de rechazarse la hipótesis de diferencia nula en las variancias, deberán realizarse algunas modificaciones en los cálculos para evitar el aumento de α, la probabilidad de error de tipo I. De nuevo, los programas informáticos realizan esta corrección en forma automática.

PODER Y TAMAÑO DE LA MUESTRA

En el **cuadro 10-2** pueden observarse los datos obtenidos en un experimento con dos muestras y el resultado del análisis estadístico realizado mediante la prueba de "t".

Se decidió no rechazar la hipótesis nula e indicar que la diferencia no es significa-

tiva, ya que el valor de "t" obtenido (1,73) es menor en valor absoluto que el que se toma como referencia del **cuadro 8-2** para un nivel de significación de 0,05 y 6 grados de libertad.

En una situación en la que no se estableció una diferencia estadísticamente significativa, se debe considerar que, al no rechazarse la hipótesis, se la acepta y, en consecuencia, se puede estar cometiendo un error de tipo II (aceptar una hipótesis que, en realidad, es falsa).

—

Frente a esto, debe definirse cuál es la diferencia entre los grupos que tiene un significado práctico o clínico, no estadístico. Por ejemplo, si se plantea establecer la presencia de diferencias de media aritmética de masa corporal entre dos poblaciones de seres humanos, podría ser de importancia práctica detectar diferencias de medio kilogramo. Si esas diferencias de masa ocurren entre ratas de laboratorio, es muy probable que sea necesario detectar una diferencia menor.

Considérese que en la situación del **cuadro 10-2** sería conveniente llegar a establecer diferencias significativas de, por lo menos, 2,5 unidades entre las medias aritméticas de las poblaciones a las que pertenecen ambas muestras.

En este caso, puede calcularse que, si esa diferencia de 2,5 realmente existiera, el poder del experimento para detectarla (1 – 3) es menor que 0,5 o 50%.

Esto significa que, si esa diferencia de importancia práctica existe, el experimento no la detectará el 50% de las veces que se lleva a cabo, o que puede esperarse cometer un error de tipo II ese porcentaje de veces.

Para aumentar el poder del experimento se debe aumentar el tamaño de la muestra. De esta manera, se disminuye el valor de la estimación del error estándar y se disminuye el valor de β.

—

Se entiende que esta situación se genera porque para el cálculo del error estándar se utiliza como divisor a la raíz cuadrada del tamaño de la muestra.

El **cuadro 10-3** muestra los resultados de la ampliación del experimento del **cuadro 10-2**, con el aumento del tamaño de muestra a 8 para cada grupo. Puede verse que, en este caso, la misma diferencia antes observada (2) se encontró estadísticamente significativa.

CUADRO 10-2. PRUEBA DE "T" DE DATOS SIN DIFERENCIA SIGNIFICATIVA

Grupo A	Grupo B	
14,0	11,0	
14,0	12,0	
17,0	14,0	
15,0	15,0	
Media aritmética	A = 15,0	B = 13,0
Tamaño de la muestra	A = 4	B = 4
Diferencia observada	2,0	
Variancia ponderada	2,7	
Error estándar de la diferencia	1,2	
Valor de T	1,73	
Probabilidad	> 0,05	

CUADRO 10-3. AMPLIACIÓN DEL EXPERIMENTO DEL **CUADRO 10-2**

Grupo A	Grupo B	
14,0	11,0	
14,0	12,0	
17,0	14,0	
15,0	15,0	
15,0	11,0	
15,0	12,0	
14,0	14,0	
17,0	15,0	
Media aritmética	A = 15,0	B = 13,0
Tamaño de la muestra	A = 8	B = 8
Diferencia observada	2,0	
Variancia ponderada	2,3	
Error estándar de la diferencia	0,8	
Valor de t	2,65	
Probabilidad	< 0,05	

Al diseñar una investigación, puede estimarse con anterioridad el tamaño de la muestra conveniente de manera similar a la que se indicó para la investigación descriptiva y el cálculo de intervalos de confianza.

Las fórmulas para el cálculo de ese tamaño estimado de muestra y que procesan algunos programas informáticos requieren el ingreso de los siguientes datos:

a) El nivel de α con que el se pretende trabajar (p. ej., 0,05).
b) La diferencia que se desea llegar a establecer como significativa, si en realidad existe.
c) El poder que se desea en el experimento que, por lo general, se fija en por lo menos 0,8 u 80% (o sea, un valor de $\beta \leq 0,2$ en esas condiciones).
d) La dispersión (desviación estándar o variancia) que se espera encontrar en los grupos, dato que se obtiene de experimentos previos del propio investigador, de la bibliografía o de un experimento "piloto".

SÍNTESIS CONCEPTUAL

- **Cuando se plantea una hipótesis** (hipótesis nula) relacionada con la media aritmética de una población o con una diferencia entre las medias aritméticas de dos poblaciones (comparación entre dos grupos), puede utilizarse la prueba de "t" de *Student*.
- **La prueba de "t" de *Student* permite** calcular, a partir de valores obtenidos en muestras con datos numéricos, la probabilidad de cometer un error de tipo I (α) al rechazar una hipótesis nula.
- **Luego de declarar estadísticamente** significativa una diferencia, es necesario realizar una evaluación de su relevancia práctica o clínica si se quieren tomar decisiones, no solo estadísticas, sino de aplicación de resultados.
- **Ante la ausencia de una diferencia es**tadísticamente significativa, no debe asumirse necesariamente que esto implique igualdad sin antes estimar el poder del diseño de la investigación para detectar diferencias que sean relevantes en el nivel clínico o práctico.

EJEMPLO 10-1

Para comparar los resultados del empleo de dos anestésicos locales diferentes (A y B) se dividió aleatoriamente a 60 voluntarios en dos grupos de 30. Los integrantes de cada uno de ellos fueron anestesiados con uno de los dos fármacos.

La evaluación del tiempo en segundos para lograr el efecto anestésico arrojó un resultado medio de 150, con una desviación estándar de 28 en el grupo que recibió A, mientras que en el que recibió B los respectivos valores fueron 165 y 34.

La hipótesis nula para probar es que no existe una diferencia entre el valor de la media aritmética de ambas poblaciones y puede hacérselo mediante la prueba de "t".

Los cálculos correspondientes, realizados con un programa informático apropiado, permiten establecer que el valor de "t" (en este caso, con 58 grados de libertad) para el experimento fue 1,865. La consulta a una tabla de "t" –que generalmente no es necesario, ya que el programa informático brinda la información– indica que la probabilidad asociada a ese valor es mayor que 0,05 y, por lo tanto, no existe una justificación para el rechazo de la hipótesis nula y puede decirse que no se ha encontrado una diferencia significativa entre ambos fármacos anestésicos.

EJEMPLO 10-2

La diferencia observada en las medias aritméticas en la experiencia del ejemplo anterior fue de 15 (165 en el grupo B y 150 en el A) y no se la encontró de significación estadística. Debe tenerse en cuenta que el no rechazo de la hipótesis nula y su consiguiente aceptación pueden significar la posibilidad de que se esté cometiendo un error de tipo II.

Si se considerara que detectar como significativas diferencias medias de 18 segundos en el efecto de los anestésicos es de importancia "clínica", deberá establecerse el poder del experimento para hacerlo. En este caso, el cálculo –que puede hacerse con un programa informático– indica un poder un poco inferior al 50% para el valor de α seleccionado (0,05), lo que indica que es muy posible que se esté cometiendo ese error, o sea, que es alto el valor de β.

Si se quisiera tener un poder del 80% para detectar esa diferencia con el mismo valor de α (0,05) el tamaño para cada una de las muestras debería ser de alrededor de 64.

EJEMPLO 10-3

En la misma experiencia del ejemplo 10.1 también se registró en cada sujeto la duración de la anestesia en minutos. En este caso, los valores de la media aritmética de las muestras que recibieron A y B fueron 46 y 58, respectivamente, con desviaciones estándar de 12 y 15.

Una prueba de "t" arroja un valor de 3,422 y la correspondiente probabilidad (α) es inferior a 0,05 y también a 0,01. Puede rechazarse la hipótesis nula de igualdad en las medias aritméticas de ambas poblaciones para esta variable y decirse que la diferencia se encontró significativa o "altamente significativa", ya que α fue inferior a 0,01.

EJEMPLO 10-4

Los datos que se muestran a continuación representan valores de eritrosedimentación registrados en los pacientes antes y después de un procedimiento quirúrgico:

Paciente	Antes (a)	Después (b)	Diferencia (B-A)
A	1,0	1,5	0,5
B	11,0	10,0	–1,0
C	3,0	3,5	0,5
D	5,0	5,5	0,5
E	5,5	6,0	0,5
F	2,0	2,5	0,5
G	7,0	8,0	1,0
H	6,0	7,5	1,5
I	5,0	6,5	1,5
J	4,0	6,5	2,5
K	6,0	7,0	1,0
L	5,0	4,0	–1,0
M	1,5	2,0	0,5
N	6,0	7,5	1,5
0	5,0	5,5	0,5
P	2,0	3,0	1,0
Diferencia media			0,7

Los dos grupos de datos, antes y después del procedimiento, no son independientes, ya que fueron registrados por "pares" en un mismo paciente.

La hipótesis nula, en este caso, es enunciar que en la población la media aritmética de la diferencia entre los valores obtenidos antes y después de la intervención quirúrgica es 0.

La prueba de "t" para los datos apareados o emparejados arroja un valor de 3,286 para el que se indica que la probabilidad es inferior a 0,05. Puede rechazarse la hipótesis nula y aceptarse que el procedimiento quirúrgico produce una modificación estadísticamente significativa en el valor medio de la eritrosedimentación.

CAPÍTULO

11

ANÁLISIS DE VARIANCIA

INTRODUCCIÓN

En el **cuadro 11-1** se incluyen datos obtenidos en dos muestras. Puede ser de interés plantear si esos resultados permiten aceptar o no una hipótesis referida a una igualdad de variancia, como medida de dispersión, en las poblaciones de las que fueron tomadas; esto significa que σ^2, el símbolo para la variancia de una población, es igual en ambas.

Si esa hipótesis resulta verdadera, la relación –el cociente– entre ambos valores es 1. En el caso de los datos del citado cuadro, se puede observar que en uno de los grupos la variancia –es decir, lo que corresponde a "V", que estima el correspondiente valor de σ^2– es 61,60, mientras que en el otro es 110,46.

Si se establece la relación al dividir el valor mayor por el menor, se obtiene 1,79 como resultado; es decir que la variancia en la muestra B es 1,79 veces mayor que en la muestra A.

Al igual que en el caso de la media aritmética, el valor de la variancia en las muestras, en promedio, es igual al de la variancia de la población.

Por lo tanto, y en función de la hipótesis planteada, era esperable una relación igual a 1 y no 1,61, como se encontró.

Para establecer si la diferencia entre lo esperado y lo observado en las condiciones de la experiencia tiene una probabilidad de manifestarse menor que 0,05 –el nivel de significación que podría seleccionarse– se debe comparar el valor obtenido con el de una distribución derivada de la normal.

Esa distribución se conoce con el nombre de **distribución de F**, que representa la relación entre dos variancias. El **cuadro 11-2** muestra parcialmente los valores que en esa distribución permiten separar el 95% del área "más probable" del 5% "poco probable". El cuadro tiene dos entradas: en las

CUADRO 11-1. COMPARACIÓN ENTRE DOS VARIANCIAS

Grupo A	Grupo B	
20,8	49,6	
48,0	41,6	
39,7	35,3	
26,0	26,1	
29,3	35,5	
38,3	22,5	
36,5	26,0	
29,9	43,2	
38,3	21,2	
34,1	49,6	
36,3	47,3	
41,0	48,2	
32,9	43,2	
49,0		
29,3		
Variancia	A = 161,60	B = 110,46
Tamaño de la muestra	A = 16	B = 11
Grados de libertad	A = 15	B = 10
Valor de F	1,79	
Probabilidad	> 0,05	

columnas se lee "grados de libertad del numerador" y en las filas, "grados de libertad del denominador".

Para la situación del **cuadro 11-1**, a partir de la cual se calculó la relación entre las variancias de los grupos B y A, los grados de libertad del numerador son 10 (11 – 1) y los del denominador 15 (16 – 1). La lectura indica que el valor 2,54 es el "crítico" para el nivel de significación elegido, lo que indica una probabilidad menor que 0,05, si este es superado.

El cociente obtenido (1,79) es inferior al valor "crítico" de F y, en consecuencia, no se rechaza la hipótesis de igualdad entre las variancias de las poblaciones. No se ha encontrado diferencia estadísticamente significativa entre las variancias de ambos grupos.

Esta prueba de comparación entre variancias permite comparar grupos en cuanto a la influencia de un determinado factor –variable independiente– sobre una variable descrita con datos numéricos, de manera similar a como se lo hizo con la prueba de "t", que tiene la limitación de ser aplicable para situaciones de comparación solo entre dos grupos.

CUADRO 11-2. ALGUNOS VALORES DE LA DISTRIBUCIÓN DE F PARA $P = 0{,}05$

Grados de libertad del numerador	Grados de libertad del denominador					
	1	2	3	4	5	10
5	6,61	5,79	5,41	5,19	5,05	4,74
10	4,96	4,10	3,71	3,48	3,33	2,98
15	4,54	3,68	3,29	3,06	2,90	2,54
20	4,35	3,49	3,10	2,87	2,71	2,35
30	4,17	3,32	2,92	2,69	2,53	2,16
40	4,08	3,23	2,84	2,61	2,45	2,08

La realización de un análisis de variancia permite elaborar comparaciones entre más de dos grupos y establecer si la influencia de diversos factores es estadísticamente significativa o no.

COMPARACIÓN ENTRE VARIOS GRUPOS

El **cuadro 11-3** incluye los datos numéricos obtenidos para la evaluación de una determinada variable en cuatro grupos experimentales: pacientes que recibieron cuatro tratamientos diferentes, animales alimentados con cuatro dietas distintas, o cualquier otra situación equivalente. La razón para realizar un experimento de este tipo es establecer si puede aceptarse o no una hipótesis de igualdad de resultado promedio en las diferentes condiciones. Expresado de otra manera: el objetivo es contrastar una hipótesis, hipótesis nula, en la que se enuncia que el resultado promedio para los datos es igual en las cuatro poblaciones de las cuales se tomaron los grupos.

Puede notarse que los cuatro grupos son del mismo tamaño (n = 10). Esta situación no es necesaria, aunque sí conveniente. La técnica del análisis de la variancia asume homocedasticidad –igual variancia– en las poblaciones. El aumento en la posibilidad de error por no cumplirse este requisito es menor cuando las muestras son de igual tamaño en todos los grupos. Se acostumbra a decir que, en estas condiciones, la prueba estadística es "robusta", resiste bien una situación desfavorable.

El **análisis de variancia** se basa en considerar, en primer lugar, que, si los datos resultantes del experimento –40 en el ejemplo– no fueron iguales, existe una cierta dispersión que puede cuantificarse.

CUADRO 11-3. RESULTADOS EN UN EXPERIMENTO CON CUATRO GRUPOS

A	B	C	D
132	133	135	130
132	133	136	129
133	134	135	130
133	134	137	131
134	135	134	131
131	134	131	132
132	132	130	133
132	133	130	132
131	134	130	132
132	131	131	131

Los 40 datos registrados no fueron iguales, por lo que puede establecerse la variancia que cuantifica la dispersión. Esta se calcula a partir de la suma de los cuadrados (cuadrados de la desviación de cada uno de los 40 valores respecto de la media de esos mismos valores) y los correspondientes grados de libertad, 39 en este caso (40 – 1).

En la última fila del **cuadro 11-4** se indican los valores de suma de los cuadrados y grados de libertad totales para los datos del **cuadro 11-3, en las respectivas columnas.**

Puede considerarse que esa dispersión que se observó en los 40 datos tiene dos orígenes o fuentes posibles. Por un lado, según el grupo al que fuera asignada una unidad experimental puede esperarse que varíe el resultado si la variable de agrupación tiene un efecto detectable.

Una de las columnas del cuadro está encabezado con la expresión "Origen de las

CUADRO 11-4. ANÁLISIS DE VARIANCIA DE LOS DATOS DEL **CUADRO 11 -3**

Origen de las variaciones	Suma de los cuadrados	Grados de libertad	Cuadrado medio	F	Probabilidad
Entre grupos	27,88	3	9,29	3,30	< 0,05
Dentro de los grupos	101,50	36	2,82		
Total	129,38	39			

variaciones" y una de las filas con "Entre grupos". En esta última puede encontrarse al número 3 bajo la columna "Grados de libertad", que corresponde a los grados de libertad para este origen o fuente de variación, y está dado por el número de grupos menos uno (4 – 1). La correspondiente suma de los cuadrados es, para ese mismo origen, 27,88.

Por otro lado, parte de la variación se puede detectar "Dentro de los grupos" y puede estar determinada por diferencias entre las unidades experimentales incluidas en ellos o errores cometidos en el registro de los datos. Como el tamaño de la muestra fue en los cuatro grupos igual a 10, en cada uno de ellos son 9. En la columna respectiva se encuentra el número 36 (9 × 4), que representa la totalidad de los grados de libertad dentro de los grupos del experimento. En la columna "Suma de los cuadrados" de esa misma fila se encuentra el valor 101,50.

Puede observarse que los valores de la fila "Total" en las columnas "Grados de libertad" y "Suma de los cuadrados" corresponden a la suma de los valores en las otras dos filas.

La base del análisis de variancia es separar la variación (dispersión) total del experimento en los componentes que se estima que pueden generarla.

—

Si se dispone de valores de suma de cuadrados y de grados de libertad, es posible relacionarlos para tener una estimación de la variancia para cada uno de los orígenes de la variación.

La columna "Cuadrado medio" o "Media cuadrática", que se recordará que se mencionó en el **capítulo 5, Resumen de datos numéricos** como sinónimo de "variancia", muestra los correspondientes valores.

Se dispone ahora de la variancia, la cual se estima que está originada entre los grupos por efecto de la variable independiente o el factor en análisis (9,29).

Se dispone también de la variancia que se estima originada dentro de los grupos (2,82). Como esta variancia se estima que está originada por todos aquellos factores que no pudieron ser mantenidos bajo control, se la considera, en otras denominaciones, como valoración del **error experimental.**

Puede considerarse que si la hipótesis formulada –la hipótesis nula– es verdadera, es de esperar que la variancia originada entre los grupos sea igual o menor que la originada por el error experimental, o sea, dentro de los grupos.

Los valores obtenidos muestran que la relación entre ambas, que se encuentra bajo la columna F del cuadro, es de 3,30; es decir que la variancia entre los grupos es 3,3 veces mayor que la variancia debida al error experimental.

Queda por establecer la probabilidad que ese valor tiene de producirse si la hipótesis nula fuera verdadera. Para ello, tal como se describió en el comienzo de este capítulo, y dado que la tarea no es más que una comparación entre variancias, solo es necesario buscar en una tabla de distribución de F el valor "crítico" correspondiente a α elegido (p. ej., 0,05).

Para la situación de análisis que corresponde a la relación entre una variancia estimada con 3 grados de libertad (el cuadrado medio entre grupos) y una estimada con 36 grados de libertad (el cuadrado medio dentro de los grupos), ese valor "crítico" es 2,87.

Como el valor de F obtenido (3,30) supera al crítico, puede rechazarse la hipótesis nula, ya que la probabilidad de esta relación en el nivel de variación originada entre los grupos y el error experimental registrado es menor que 0,05, como se indica en la última columna del cuadro. En definitiva, puede decirse que se ha encontrado una diferencia estadísticamente significativa entre los grupos experimentales.

Los programas estadísticos informatizados y algunas planillas de cálculos permiten obtener cuadros como el que se ha mostrado y el investigador solo debe observar el valor de *P*; es decir de α, para tomar una decisión respecto de la hipótesis.

También en este caso, cuando no se encontraron diferencias estadísticamente significativas, debe establecerse si el poder del experimento para detectar diferencias de interés es suficiente (generalmente igual o mayor que 0,8 u 80%). Si es necesario, deberá ampliarse el experimento aumentando el tamaño de la muestra. Este dato, tamaño de la muestra, puede calcularse antes de iniciar la tarea y a partir de la misma información que se citó para el caso de la prueba de "t".

COMPARACIONES MÚLTIPLES

El análisis de variancia realizado para los datos del **cuadro 11-3** establece la influencia significativa del factor o variable utilizada para conformar los grupos. Sin embargo, no permite establecer si entre cada uno de esos grupos la diferencia es estadísticamente significativa o no.

Cuando se encuentra un efecto significativo del factor de agrupamiento, el análisis de variancia debe ser completado con las denominadas pruebas de comparación múltiple.

Estas se basan en establecer cuál es la diferencia mínima entre las medias aritméticas de los grupos, que tiene una probabilidad de ocurrencia menor que un valor crítico –usualmente del 5%– si la hipótesis nula es verdadera. Cuando la diferencia observada entre dos de los grupos del experimento es superior a esa "mínima diferencia significativa", se rechaza la hipótesis de igualdad de media aritmética entre las correspondientes poblaciones.

En el **cuadro 11-5** se muestran las medias aritméticas correspondientes a los cuatro grupos del ejemplo, ordenadas de mayor a menor. Después de haber aplicado una prueba de comparación múltiple, las diferencias no fueron estadísticamente significativas ($P < 0,05$), excepto entre los grupos D y B.

Existen varias formas posibles de realizar esas comparaciones múltiples. Casi todas ellas se conocen por el nombre del investigador que las desarrolló. Dentro de las más utilizadas se encuentran las pruebas de Tukey, Bonferroni, Scheffé y otras. Algunas

CUADRO 11-5. COMPARACIONES MÚLTIPLES ENTRE LAS MEDIAS ARITMÉTICAS DE LOS GRUPOS DEL **CUADRO 11-3**

Grupo	Media aritmética	Desviación estándar
D	131,1	1,2
A	132,2	0,9
C	132,9	2,8
B	133,3	1,2

son de aplicación en situaciones determinadas y específicas, como la prueba de Dunnett, que permite la comparación de cada uno de los diversos grupos experimentales con un grupo control.

De nuevo, los programas informáticos de estadística ofrecen la posibilidad de ejecutar una o varias de estas pruebas y dan la información sobre el resultado correspondiente.

Si se desea evaluar la relevancia práctica o clínica de las diferencias, es necesario establecer el intervalo de confianza para los valores observados y relacionarlos con el conocimiento específico sobre el tema en estudio.

—

ANÁLISIS DE VARIANCIA DE MEDIDAS REPETIDAS Y EN DISEÑOS FACTORIALES

En el caso presentado como ejemplo solo se tomaron dos orígenes de variación dentro del experimento y el análisis de variancia realizado se conoce como de "una vía". Solo se evalúa la significación de un factor o variable independiente.

La técnica de análisis de variancia permite extender más la desagregación de la medida de la variancia total y evaluar la influencia de más de un factor o variable independiente en los resultados obtenidos en una investigación.

—

En los casos de diseños emparejados, o cuando en una misma unidad experimental se hacen mediciones en diferentes momentos (p. ej., mediciones en pacientes en condición basal y luego de diversos períodos de administración de un tratamiento), se puede separar y evaluar la posible variancia originada en las diferencias entre los diversos pacientes y la generada por el tiempo de aplicación del tratamiento.

El **cuadro 11-6** muestra un ejemplo de resultados de un análisis de variancia de "medidas repetidas". En este caso, se tiene un valor de F para cada uno de los orígenes de variación. Cada uno de estos valores de F se obtiene al relacionar, en cada caso, el valor del correspondiente cuadrado medio con el cuadrado medio entre grupos o error experimental. Según sea que ese valor resulte inferior al "crítico" o no, será menor o no que, por ejemplo, 0,05 la probabilidad del resultado encontrado. En función de ello se establecerá como estadísticamente significativa o no la influencia del factor o en la variable respectiva.

Cuando se analizan varios factores, por ejemplo, un fármaco utilizado y el nivel de edad del paciente, el análisis de variancia permite establecer la significación estadística de cada factor y de su interacción.

—

El **cuadro 11-7** es un ejemplo del análisis de un diseño con dos factores. Uno de ellos (A) fue evaluado en dos grupos –por ejemplo, dos fármacos–, lo que se deduce de un único grado de libertad que corresponde a la variación de ese origen. El otro (B), que podría representar el nivel de edad, lo fue en tres, según surge de la presencia de los dos grados de libertad para él.

Los valores que se leen en las columnas F y *P* (probabilidad) indican que es significativo ($P < 0,05$) el efecto de ambos factores y no así el de su interacción ($P > 0,05$). Esto último indica que puede considerarse que el efecto de A es independiente del efecto de B. En el ejemplo supuesto significa que el efecto de los dos fármacos evaluados se produce de la misma manera en todos los niveles de edad. Por lo tanto, pueden realizarse múltiples comparaciones entre medicamentos en forma general.

Si, en cambio, el efecto de la interacción resultara significativo ($P < 0,05$), se debería evaluar el efecto de cada nivel del factor A, de cada fármaco en el ejemplo, dentro de cada uno de los niveles de edad.

CORRELACIÓN Y REGRESIÓN

A partir de los mismos principios de partición de la variancia en sus componentes, es posible realizar otros tipos de análisis. Dentro de este punto merece ser mencionada la evaluación de la posible relación existente entre dos o más datos numéricos registrados en una misma unidad o situación experimental.

Un ejemplo podría ser plantear la evaluación de una posible relación entre el aumento de un dato descriptivo de edad con el descriptivo de la variable "presión arterial", o la relación entre la dosis de un fármaco y el efecto que produce.

CUADRO 11-6. ANÁLISIS DE VARIANCIA DE "DOS VÍAS"

Origen de las variaciones	Suma de los cuadrados	Grados de libertad	Cuadrado medio	F	Probabilidad
Factor A	2136,7	9	237,4	0,60	> 0,05
Factor B	21 740,0	2	10 870,0	26,98	< 0,05
Error experimental	7253,3	18	403,0		
Total	31 130,0	29			

CUADRO 11-7. ANÁLISIS DE VARIANCIA EN UN DISEÑO FACTORIAL

Origen de las variaciones	Suma de los cuadrados	Grados de libertad	Cuadrado medio	F	Probabilidad
Factor A	106,7	1	106,7	16,30	< 0,05
Factor B	703,3	2	351,7	53,74	< 0,05
Interacción	3,3	2	1,7	0,25	> 0,05
Error experimental	353,4	54	6,5		
Total	1166,7	59			

En estos casos puede calcularse el denominado **coeficiente de correlación de Pearson, que es un número con un rango** de entre –1 y 1. Un coeficiente 0 (cero) indica la ausencia de relación entre los datos para cada variable; un coeficiente 1 (uno positivo) indica una relación máxima de aumento de un dato para una variable cuando aumenta el correspondiente a la otra; un coeficiente –1 (uno negativo) indica también una relación máxima, aunque aquí el aumento de uno de los datos se observa acompañado por una disminución en el otro. Los valores intermedios indican graduaciones en la evaluación de la correlación.

Las hipótesis referidas a una correlación entre variables se formulan respecto de poblaciones. Si la determinación del coeficiente de correlación se realiza a partir de los datos de una muestra, se debe realizar un análisis estadístico que establezca la probabilidad de obtener ese coeficiente si la hipótesis fuera verdadera. Según sea ese valor de probabilidad, se rechaza o no se rechaza la hipótesis mediante los criterios habituales.

Ante la existencia de una correlación puede plantearse el interés en describir cómo es la relación entre los datos. Esto significa, por ejemplo, evaluar cuánto aumenta (o disminuye) el valor para una o varias variables cuando aumenta una unidad en un determinado dato. Asimismo, evaluar si el aumento producido sigue una relación lineal o de otro tipo (cuadrática, exponencial, etcétera).

Los procedimientos que se aplican en estos casos constituyen el **denominado análisis de regresión, mediante el cual se** pueden obtener las ecuaciones que describen la relación entre los datos y representar a esta última en gráficos. A partir del análisis realizado con datos de muestras, pueden aplicarse las técnicas inferenciales para estimar el comportamiento en la población o probar una hipótesis respecto de ella.

Las técnicas basadas en el análisis de variancia brindan múltiples posibilidades y se emplean con frecuencia en la investigación científica en las ciencias de la salud.

—

SÍNTESIS CONCEPTUAL

- **El análisis de variancia permite realizar** comparaciones entre más de dos grupos y establecer si la influencia de diversos factores es estadísticamente significativa o no.
- **La base del análisis de variancia consiste** en separar la variación total del experimento en los componentes que pueden generarla y así establecer, mediante el cálculo del valor de F, si la variancia entre grupos no es significativamente mayor de la generada dentro de los grupos.
- Cuando se encuentra un efecto significativo del factor de agrupamiento o diferencias significativas entre grupos, el análisis de variancia se debe completar con las pruebas de comparación múltiple.
- **Las técnicas basadas en el análisis de** variancia brindan múltiples posibilidades y se emplean con frecuencia en la investigación científica en las ciencias de la salud, por ejemplo, en diseños factoriales o estudios de correlación y regresión.

EJEMPLO 11-1

Se realiza un experimento para evaluar ***in vitro*** el efecto que seis diferentes antimicrobianos producen sobre el desarrollo de una cepa específica. El efecto se evaluó con datos numéricos (mm de inhibición registrados en un cultivo) y se hicieron cinco determinaciones (tamaño de la muestra) con cada uno de los fármacos.

Los datos obtenidos se presentan en el siguiente cuadro.

DATOS NUMÉRICOS (MM DE INHIBICIÓN DE UN CULTIVO) DE UN EXPERIMENTO REALIZADO CON SEIS ANTIMICROBIANOS (A-F)

Antimicrobiano					
A	**B**	**C**	**D**	**E**	**F**
19,4	17,7	17,0	20,7	14,3	17,3
32,6	24,8	19,4	21,0	14,4	19,4
27,0	27,9	9,1	20,5	11,8	19,1
32,1	25,2	11,9	18,8	11,6	16,9
33,0	24,3	15,8	18,6	14,2	20,8

Para probar la hipótesis nula de inexistencia de diferencias entre las medias aritméticas que se obtendrían en poblaciones tratadas con los antimicrobianos, es aplicable el análisis de variancia. Los resultados de la aplicación de este procedimiento mediante un programa informático se resumen a continuación.

RESULTADOS DEL ANÁLISIS DE VARIANCIA DEL EXPERIMENTO CON ANTIMICROBIANOS

Origen de las variaciones	Suma de los cuadrados	Grados de libertad	Cuadrado medio	F	*P*
Entre grupos	847,05	5	169,41	14,37	< 0,05
Dentro de los grupos	282,93	24	11,79		
Total	1129,97	29			

El valor de ***P*** indica que es posible rechazar la hipótesis nula para el valor usual de α (0,05), por lo que puede establecerse que el efecto del factor en estudio, tipo de antimicrobiano, es estadísticamente significativo.

Para establecer entre cuáles de los evaluados se puede considerar como significativa la diferencia, se completa el análisis con una prueba de comparación múltiple. El resultado de la prueba de Tukey llevada a cabo con esa finalidad se muestra en el siguiente cuadro.

RESULTADOS DE LA PRUEBA DE TUKEY

Grupo	Medias sin diferencia significativa		
E	13,3		
C	14,6		
F	18,7	18,7	
D	19,9	19,9	
B		24,0	24,0
A			28,8

Los valores son las medias aritméticas de las muestras tratadas con el antimicrobiano que se indica para cada fila. Las diferencias no son estadísticamente significativas ($P > 0,05$) entre las que se muestran en una misma columna, mientras que sí son significativas **($P < 0,05$) las diferencias entre las que están en columnas diferentes.**

EJEMPLO 11-2

Se desea comparar los resultados de resistencia flexural de un material en MPa que se obtiene luego de procesarlo con tres técnicas distintas.

Es por ello que se remite una muestra procesada con cada una de las técnicas a cuatro laboratorios para su ensayo.

Los resultados obtenidos se muestran a continuación.

RESISTENCIA FLEXURAL (EN MPA) DE UN MATERIAL PROCESADO CON TRES TÉCNICAS EN DISTINTOS LABORATORIOS

Laboratorio	Técnica		
	A	B	C
I	660	370	420
II	650	410	380
III	710	480	390
IV	800	510	505

Como se estima que pueden existir diferencias entre los resultados obtenidos por los distintos laboratorios, además de las que podrían existir entre las técnicas, se realiza un análisis de variancia de dos vías que permite separar la variación originada por cada uno de esos dos factores.

Los resultados de ese análisis se muestran en el siguiente cuadro.

ANÁLISIS DE VARIANCIA DE DOS VÍAS CORRESPONDIENTE AL ESTUDIO DE RESISTENCIA DE UN MATERIAL

Origen de las variaciones	Suma de los cuadrados	Grados de libertad	Cuadrado medio	F	*P*
Laboratorio	30 472,9	3	10 157,6	10,9	< 0,05
Técnica	197 812,5	2	98 906,3	106,5	< 0,05
Error	5570,8	6	928,5		
Total	233 856,3	11			

Como puede verse, se encontró significativa la influencia de ambos factores. De no haberse tenido en cuenta la influencia del factor laboratorio, la variación producida por la diferencia entre ellos se hubiera sumado al error y, con ello, se disminuiría la posibilidad de encontrar diferencias significativas entre las técnicas. En otras palabras, hubiera sido menor el poder del experimento para detectarlas.

CAPÍTULO

12

PRUEBA DE CHI-CUADRADO

INTRODUCCIÓN

Cuando se trabaja con datos nominales, la formulación de una hipótesis nula puede hacerse enunciando la igualdad de las proporciones o porcentajes en las poblaciones involucradas.

Para el caso de la comparación de dos grupos, podría enunciarse en símbolos: H_Q: $p_A - p_B = 0$; es decir, que es nula la diferencia entre las proporciones, o los porcentajes, de individuos en una determinada categoría para ambas poblaciones.

La decisión de rechazo o aceptación de esa hipótesis puede escogerse a partir del análisis de la diferencia observada entre proporciones de muestras tomadas de ambas poblaciones. El procedimiento puede ser bastante similar al de una prueba de "t" con datos numéricos, si se dan ciertas condiciones relativas al tamaño de muestra.

Cuando se trabaja con datos de categorización, que se resumen en frecuencias en las distintas categorías, la técnica más frecuente para la prueba de la hipótesis es la prueba de χ^2 ("chi-cuadrado").

COMPARACIÓN EN TABLAS DE 2 × 2

En el **cuadro 12-1** se muestran los posibles resultados de un experimento, en el cual se intentan comparar dos situaciones experimentales respecto de una variable evaluada mediante datos nominales dicótomos.

En el ejemplo, los datos representan la frecuencia de "éxitos" o "fracasos", diferenciados en las filas, obtenidos en cada uno de los dos grupos, que podrían estar representados por los tratamientos A y B, diferenciados en las columnas.

El objetivo es establecer si a partir de estos datos puede estimarse que existe diferencia entre los dos tratamientos, que es la variable independiente, en cuanto al resultado –éxito o fracaso–, que es la variable dependiente.

Teniendo presente que la hipótesis nula es la inexistencia de esa diferencia, puede determinarse cuál es el resultado esperable en el experimento, si esta es verdadera.

Ese resultado esperable se muestra en el **cuadro 12-2**.

Si los dos tratamientos se comportan de la misma manera, es válido esperar que la cantidad total de éxitos y fracasos observados esté repartida en partes iguales entre los dos grupos, si el tamaño de la muestra ha sido igual en ambos.

—

Se verifica, entonces, que se ha encontrado una diferencia entre lo observado y lo esperado, y que se muestra, para cada condición de tratamiento y resultado, en el **cuadro 12-3**.

Puede observarse que la suma de todas esas diferencias es cero, lo que no permite cuantificar la diferencia producida. Para permitir esa valoración, se eleva a cada una de ellas al cuadrado y se lo relaciona con el valor esperado para la correspondiente celda.

Así, para el ejemplo, en el **cuadro 12-4** se muestran los valores 0,46 = (10^2 / 219); .....; 2,44 = (10^2 / 41). La suma del total de esos valores obtenidos (5,79) puede ser ubicada en una distribución que también tiene una vinculación con la distribución gaussiana.

La citada disposición se conoce como **distribución de chi-cuadrado (χ^2) y en** ella se puede encontrar un valor que separa al área "más probable" (95%) de la "poco probable" (5%).

Un valor de chi-cuadrado obtenido de los resultados del experimento que supere un valor crítico acorde con el nivel de significación o α seleccionado indica una situación en la que es posible rechazar la hipótesis nula formulada; un valor menor orienta hacia la decisión contraria.

—

CUADRO 12-1. TABLA DE 2 × 2. VALORES OBSERVADOS EN UN EXPERIMENTO

	Grupo A	Grupo B	Total
Éxito	229	209	438
Fracaso	31	51	82
Total	260	260	520

CUADRO 12-2. TABLA DE 2 × 2. VALORES ESPERADOS EN EL EXPERIMENTO DEL **CUADRO 12-1**

	Grupo A	Grupo B	Total
Éxito	219	219	438
Fracaso	41	41	82
Total	260	260	520

CUADRO 12-3. TABLA DE 2 × 2. DIFERENCIA ENTRE VALORES OBSERVADOS Y ESPERADOS EN EL EXPERIMENTO DEL **CUADRO 12-1**

	Grupo A	Grupo B	Total
Éxito	10	–10	0
Fracaso	–10	10	0
Total	0	0	0

CUADRO 12-4. TABLA DE 2 × 2. VALORES DE CHI-CUADRADO PARA EL EXPERIMENTO DEL **CUADRO 12-1**

	Grupo A	Grupo B
Éxito	0,46	0,46
Fracaso	2,44	2,44

Al igual que para las distribuciones de "t" y de "F", los valores críticos para chi-cuadrado dependen de los grados de libertad y el nivel de α que se elija. En tablas de doble entrada, los grados de libertad están dados por el producto del número de filas menos uno por el número de columnas menos uno. Para el caso en análisis, las columnas y las filas son dos, por lo que la situación es de un grado de libertad: $(2 - 1) \times (2 - 1) = 1$.

El **cuadro 12-5** muestra algunos valores de chi-cuadrado para diversos grados de libertad y $P = 0{,}05$. Se observa que para un grado de libertad el valor crítico es 3,84.

Como el valor $\chi^2 = 5{,}79$ obtenido en el experimento supera al "crítico", se puede aceptar que $P < 0{,}05$, rechazar la hipótesis nula y decir que la diferencia entre ambos tratamientos es estadísticamente significativa. Como en otras pruebas de hipótesis, debe establecerse de manera separada si las diferencias con valor estadístico son relevantes para pensar en su traducción en decisiones de aplicación clínica o de otro tipo.

CUADRO 12-5. ALGUNOS VALORES DE LA DISTRIBUCIÓN DE CHI-CUADRADO PARA $P = 0{,}05$

Grados de libertad	Chi-cuadrado
1	3,84
2	5,99
3	7,81
4	9,49
5	11,07
6	12,59
7	14,07
8	15,51
9	16,91
10	18,31

De haberse llegado a una situación contraria, P o $\alpha > 0{,}05$, se debería analizar si el poder del experimento es el adecuado y, en caso contrario, calcular cuánto debe aumentarse el tamaño de la muestra para asegurar un nivel razonable de error de tipo II.

COMPARACIONES EN TABLAS DE F × C

La prueba de chi-cuadrado (χ^2) es aplicable a situaciones de tablas con cualquier número de columnas (c) y cualquier número de filas (f).

En el **cuadro 12-6** se muestran los resultados de un posible experimento, en el cual se comparan cuatro grupos (filas) en función de una variable evaluada con datos nominales con dos categorías posibles (columnas).

El procedimiento de cálculo del valor de chi-cuadrado para el experimento es el que ya se ha descrito para las tablas de 2 × 2. Para cada celda se calcula el valor esperado según la hipótesis, que en cada una se muestra entre paréntesis. Como en este caso las muestras no son de igual tamaño, los valores esperados son proporcionales al tamaño de la correspondiente muestra.

Para cada celda, el valor de chi-cuadrado es igual al cuadrado de la diferencia entre lo observado y lo esperado dividido por el correspondiente valor esperado. La suma

CUADRO 12-6. DATOS Y CHI-CUADRADO PARA UNA TABLA DE F × C

	Columna A	Columna B	Total
Fila A	145 (137,0)	25 (33,0)	170
Fila B	146 (145,1)	34 (34,9)	180
Fila C	108 (128,1)	51 (30,9)	159
Fila D	178 (166,8)	29 (40,2)	207
Total	577	139	716

de todos ellos (22,41) es el valor de chi-cuadrado total, que se compara con el valor crítico según los grados de libertad que, en este segundo ejemplo, es 3 (2 – 1) por las dos columnas multiplicado por (4 – 1) las cuatro filas.

El valor significativo del ejemplo indica que existen diferencias estadísticamente significativas entre las poblaciones de las cuales se obtuvieron los cuatro grupos. Si se quiere avanzar en establecer entre cuáles de ellos es significativa esa diferencia y entre cuáles no lo es, debe continuarse en la partición del valor de chi-cuadrado de manera similar a como se particiona la suma de los cuadrados en el caso del análisis de variancia. Este último procedimiento es conducido por quien practica el análisis, y no se realiza en forma automática con los programas estadísticos, que sí calculan el valor global de chi-cuadrado.

CONSIDERACIONES ADICIONALES

La prueba de chi-cuadrado tiene algunas limitaciones que no permiten su empleo en algunos casos.

En las tablas con un grado de libertad, tablas de 2 × 2, si alguno de los valores esperados es menor que 5, el uso de la prueba de chi-cuadrado debe reemplazarse por la prueba de probabilidad exacta de Fisher.

—

Esta, como su nombre lo indica, permite establecer con exactitud si se está frente a una situación que orienta hacia el rechazo o aceptación de la hipótesis nula, de acuerdo con el nivel de alfa (probabilidad de error de tipo I) elegido.

Algunos autores recomiendan también, para el caso de un grado libertad, realizar una corrección al valor de chi-cuadrado obtenido en el experimento, que se denomina corrección de Yates y que algunos programas estadísticos la hacen de manera automática en esos casos.

Cuando se tratan situaciones con más de un grado de libertad, tablas de f × c, no debe aplicarse la prueba cuando exista alguna celda en la que el valor esperado sea menor que 1 o si en más del 20% de ellas ese valor es menor que 5. En estos casos, se agrupan categorías para cambiar la situación.

Algunas modificaciones al procedimiento básico permiten realizar la prueba de la hipótesis en algunas condiciones diferentes de las ejemplificadas aquí.

Así, por ejemplo, pueden valorarse los datos nominales obtenidos en diseños con grupos emparejados, no independientes, mediante el chi-cuadrado de McNemar o cuando se valoran varios factores –más de una variable independiente– con el uso del chi-cuadrado de Mantel-Haenszel.

SÍNTESIS CONCEPTUAL

- **La prueba de chi-cuadrado es la de uso** más frecuente para la prueba de una hipótesis, cuando se trabaja con datos de categorización que se resumen en frecuencias.
- **A partir de las diferencias entre las frecuencias** observadas y las esperadas, en función de la hipótesis nula, se calcula un valor que se puede ubicar en la distribución conocida como distribución de chi-cuadrado.
- **Según si el valor de chi-cuadrado obtenido** supera o no un valor crítico acorde con el nivel de significación, se rechaza o aprueba la hipótesis nula planteada.
- **La prueba de chi-cuadrado es aplicable** a situaciones de tablas con cualquier número de columnas (c) y cualquier número de filas (f).
- **La prueba de probabilidad exacta de** Fisher es aplicable cuando, en tablas de 2 × 2, alguno de los valores esperados es menor que 5.

EJEMPLO 12-1

Para establecer la conveniencia o no de reemplazar un procedimiento terapéutico ya conocido por uno de desarrollo reciente, se llevó a cabo un experimento con ratas Wistar.

Con ellas se conformaron dos grupos, cada uno fue tratado con uno de los procedimientos para comparar, respectivamente. El resultado se evaluó, registrándose después de un lapso preestablecido si la unidad experimental (rata) había sobrevivido o no.

Los resultados se presentan a continuación.

Procedimiento	Sobrevivieron	Murieron
Conocido	8	12
Reciente	13	7

La prueba de chi-cuadrado permite probar la hipótesis nula de inexistencia de asociación entre el tratamiento aplicado y el resultado obtenido.

Para este caso, el valor de chi-cuadrado calculado es 2,506 (sin corrección). Como la situación es de solo un grado de libertad y el valor es menor que aquel al que le corresponde una probabilidad de 0,05, no debe rechazarse la hipótesis nula. No se pudo encontrar una diferencia estadísticamente significativa entre los resultados obtenidos con ambos tratamientos.

EJEMPLO 12-2

En otro experimento se aplicó uno de dos bactericidas o uno de tres bacteriostáticos en grupos de unidades experimentales. En consecuencia, se constituyeron cuatro grupos en total y en cada uno de ellos se registró si se había logrado un efecto positivo o negativo en las unidades experimentales.

Los resultados fueron:

Sustancia	Positivo	Negativo	Total
Bactericida 1	24	26	50
Bactericida 2	20	30	50
Bacteriostático A	10	40	50
Bacteriostático B	12	38	50
Total	66	134	200

En este caso, los grados de libertad son 3 y el valor de chi-cuadrado que surge de los cálculos es 11,85. La probabilidad asociada con él es inferior a 0,05, por lo que puede declararse que existen diferencias significativas entre lo observado y lo esperado o que las diferentes sustancias evaluadas producen un resultado significativamente diferente.

Un análisis posterior indicaría que la diferencia entre los resultados obtenidos con bactericidas y bateriostáticos es significativa, mientras que no existe significación estadística en las diferencias dentro de cada uno de esos dos tipos de sustancias.

CAPÍTULO

13

ESTADÍSTICA NO PARAMÉTRICA

INTRODUCCIÓN

Las pruebas de "t" y aquellas relacionadas con el análisis de variancia se utilizan para el trabajo con datos numéricos. Ambas técnicas se basan en la suposición de que los datos con los que se trabaja están distribuidos de forma gaussiana en las respectivas poblaciones.

Las hipótesis que se prueban con ellas incluyen alguna suposición respecto de parámetros, como la media aritmética o la variancia; por este motivo, se las conoce como **pruebas paramétricas** y a su estudio y desarrollo, como **estadística paramétrica**.

Existen situaciones en las que esa suposición no se cumple –como ya fue mencionado en el **capítulo 6, Distribución de frecuencias**–, aun cuando se trate de datos numéricos continuos. Cuando los datos son discretos y, más aún, cuando se trata de datos ordinales, como índices o puntajes, es todavía más difícil suponer normalidad en la distribución. Téngase en cuenta que la curva de Gauss es una línea continua, que no se puede obtener en estas últimas situaciones.

Si la distribución se aleja de manera muy significativa de la gaussiana, y especialmente cuando las muestras son relativamente pequeñas en cuanto a tamaño, el empleo de pruebas estadísticas basadas en esa distribución podría no ser conveniente. Su utilización podría llevar a niveles de error superiores a los establecidos, en teoría, en la toma de decisiones.

Una alternativa para esos casos puede ser "transformar" los datos, calculando su logaritmo, raíz cuadrada o mediante alguna otra función matemática. Si con los datos así transformados se obtiene una distribución que no se aleja de manera significativa de la gaussiana, es posible aplicar pruebas paramétricas.

Otra alternativa es formular hipótesis que no incluyan en su enunciado la presencia de parámetros, como la media aritmética o la variancia.

Las pruebas estadísticas que no necesitan analizar la distribución de estadísticos que estimen a los parámetros se conocen como pruebas ajenas a distribuciones, o no paramétricas, y su estudio y desarrollo se denomina estadística no paramétrica.

Las hipótesis que se formulan para estos casos se refieren al ordenamiento, ascendente o descendente de los datos, lo que no significa ninguna suposición sobre la distribución que en ellos se manifiesta.

FUNDAMENTOS

Supóngase la situación más sencilla de investigación en la que se plantea la necesidad de comparar dos grupos independientes de datos ordinales o numéricos sin una distribución gaussiana manifiesta.

Un ejemplo podría estar en la comparación de los resultados de una encuesta tomada a dos grupos –dos muestras– de pacientes atendidos en otros servicios. La información obtenida se puede haber registrado en un índice (p. ej., en una escala de 1 a 5) que valora la opinión de cada paciente sobre la atención recibida.

La hipótesis podría formularse sin incluir ningún parámetro, con la indicación de la inexistencia de diferencias en las respectivas poblaciones y con el enunciado de que es esperable que el orden de los datos registrados en los pacientes esté generado solo por una función aleatoria y no por el hecho de pertenecer a un grupo o al otro. Como se observa, corresponde a la hipótesis nula que es necesario contrastar de manera empírica.

Para visualizar estos aspectos de una manera más fácil, supóngase la hipótesis que podría plantearse frente al hecho que se produce al retirar cartas como las que se usan en juegos, como el póker y otros, de un mazo en las que fueron mezcladas.

Como esas cartas incluyen una mitad de color rojo (R) y otra mitad de color negro (N), al retirar una cierta cantidad es esperable que el azar haga que la distribución esperada corresponda a la que se observa en la columna A del **cuadro 13-1**. Ante una situación como esta, un análisis intuitivo no hace pensar en motivos para rechazar una

CUADRO 13-1. ORDENAMIENTOS POSIBLES DE DOS GRUPOS DE DATOS

Número de orden	Ordenamiento		
	A	B	C
1	N	N	N
2	R	N	N
3	N	N	N
4	R	N	R
5	N	N	N
6	R	N	N
7	N	N	R
8	R	N	N
9	N	N	R
10	R	N	N
11	N	N	R
12	R	R	N
13	N	R	N
14	R	R	R
15	N	R	R
16	R	R	N
17	N	R	R
18	R	R	R
19	N	R	R
20	R	R	N
21	N	R	R
22	R	R	R

hipótesis en la que se enuncie que el orden de aparición de las cartas es aleatorio.

En un experimento científico como el que se citó, el color de la carta estaría sustituido por la identificación del grupo al que pertenece el dato ubicado en una posición de orden específica. La decisión sería no rechazar la hipótesis de inexistencia de diferencia entre las poblaciones de las cuales se obtuvieron los grupos.

Si, por el contrario, el orden observado es el de la columna B del mismo cuadro, un análisis intuitivo similar orienta hacia el rechazo de la hipótesis y a sospechar que "algo más" que el azar está influyendo en ese ordenamiento. En un experimento ese "algo" sería lo que diferencia a ambos grupos, que es la variable independiente y, en el ejemplo planteado, la forma de atención.

Entre esas dos situaciones "extremas" podrían obtenerse otros resultados, como el que se muestra en la columna C del **cuadro 13-1**. En este caso, la simple intuición no alcanza para tomar una decisión, se hace necesario fijar algún nivel de significación y verificar si ese límite se sobrepasa o no para así rechazar la hipótesis nula o no.

Las pruebas no paramétricas permiten calcular la probabilidad de encontrar un determinado ordenamiento, si la hipótesis de ordenamiento aleatorio es verdadera.

—

PRUEBAS NO PARAMÉTRICAS

Estas pruebas se basan en el análisis de los diferentes resultados de ordenamiento que pueden producirse de manera aleatoria en una determinada situación. Esto se hace a partir de principios matemáticos de cálculo de combinaciones y permutaciones.

Con ese conocimiento es posible calcular si un determinado resultado, el obtenido de modo experimental, se ubica dentro de los que son "poco frecuentes" o no cuando solo funciona el azar. El límite para la definición de "poco frecuente" es patrimonio del investigador, aunque, como ya podrá imaginarse, por lo general se fija en el 5%; es decir, una probabilidad de 0,05.

En definitiva, y al seguir criterios comunes con las pruebas estadísticas paramétricas, si el análisis muestra que la probabilidad de obtener el resultado del experimento es menor que 0,05, la hipótesis nula se rechaza por saber que la probabilidad de error de tipo I (alfa) es menor que ese valor. Si es igual o mayor que 0,05 no se la rechaza y será necesario considerar, aunque no calcular en este caso, la posibilidad de que se esté cometiendo un error de tipo II.

Existen distintas pruebas estadísticas no paramétricas que se adecúan a las distintas situaciones experimentales.

—

Para el caso de dos grupos, la prueba de la U de Mann-Whitney es aplicable para el caso de muestras independientes y la de rangos con signo de Wilcoxon, para cuando estas son apareadas.

Cuando la situación presenta tres o más muestras independientes, se indica la prueba de Kruskal-Wallis, que, al igual que el análisis de variancia de la estadística paramétrica, indica si existen globales diferentes o no. De existir, se completa el análisis con una prueba de comparación que permite establecer cuáles son los grupos cuya diferencia es significativa. La prueba de compa-

ración múltiple de Dunn es, con frecuencia, la que se utiliza en ese caso.

Si esos tres o más grupos no son independientes, se utiliza la prueba de Friedman, que equivale al análisis de variancia de medidas repetidas.

También puede evaluarse la posible relación entre dos ordenamientos obtenidos en circunstancias similares. Por ejemplo, evaluar la relación que existe entre la forma en la que dos "jueces" o "árbitros" ordenan unidades experimentales en función de una variable. Para ello, se calcula el coeficiente de correlación de Spearman, que, al igual que el de Pearson para datos numéricos, puede tener valores desde –1 (uno negativo) hasta 1 (uno positivo), pasando por el 0 que indica la ausencia de correlación. Cuando los "jueces" son tres o más, mediante los procedimientos de Kendall es posible evaluar la relación entre todos ellos.

En resumen, las técnicas no paramétricas se utilizan para el trabajo con datos ordinales o numéricos con distribuciones no gaussianas.

Aunque no brindan información de tanta riqueza como lo hacen las pruebas paramétricas, las no paramétricas brindan confianza en la decisión de rechazo de hipótesis nulas en circunstancias en las que los datos no pueden ser asimilados a una distribución específica.

SÍNTESIS CONCEPTUAL

- **Cuando se trabaja con datos ordinales** o numéricos con distribuciones notoriamente alejadas de la gaussiana, no es adecuado formular hipótesis relacionadas con un parámetro de la población de la cual se obtuvieron.
- **Las hipótesis que se formulan ante ese** tipo de datos están referidas a un ordenamiento aleatorio de los datos, o sea, que no existe influencia de la viable dependiente en él.
- **Las pruebas no paramétricas permiten** calcular la probabilidad de encontrar un determinado ordenamiento, si la hipótesis de ordenamiento aleatorio es verdadera.
- **Distintas pruebas estadísticas no paramétricas** se adecúan a las diversas situaciones experimentales que pueden plantearse.
- **Las pruebas no paramétricas no brindan** información de tanta riqueza como las paramétricas, aunque son más confiables cuando se trabaja con datos ordinales o no asimilables a una distribución gaussiana.

EJEMPLO 13-1

Un jurado evaluó el desempeño de alumnos en una guardia hospitalaria, al asignar a cada uno de ellos un puntaje entre 1 y 5. Se plantea establecer si puede considerarse que el género –masculino o femenino– determina diferencias en esa variable.

Los datos se muestran a continuación:

Género	
Masculino	Femenino
3	2
4	2
5	2
4	2
2	3
3	1
4	3
4	3
1	5
5	4
4	3
1	2
1	4
2	5
3	1
1	3

La prueba no paramétrica de Mann-Whitney indica que la probabilidad de observar esta distribución en los datos solo por azar es mayor que 0,05. Por lo tanto, puede establecerse que no hay razones para rechazar la hipótesis nula de inexistencia de diferencias entre los dos grupos. En resumen, no se han encontrado diferencias estadísticamente significativas entre el desempeño de varones y mujeres en ese servicio de guardia hospitalaria.

EJEMPLO 13-2

La calidad de la atención de enfermería recibida se evaluó en muestras de pacientes internados en tres servicios asistenciales (A, B y C). Para la evaluación de la variable se utilizó una escala ordinal de 0 a 3, generada a partir de las respuestas de los pacientes a un cuestionario. El objetivo fue establecer si podía considerarse que la calidad de esa atención difería entre los distintos servicios. Los resultados fueron los siguientes:

A	B	C
1	0	3
1	1	3
2	1	1
2	0	1
1	3	3
1	1	2
2	1	1
1	0	2
1	1	3
0	1	3
1	1	1
1	1	1
1	1	2
2	1	2
1	2	1
2	1	2
1	1	3
0		2
1		3
		3

La prueba de Kruskal-Wallis permite establecer que la probabilidad de obtener por azar la distribución observada en los datos es inferior a 0,05. Por lo tanto, puede decirse que el factor servicio asistencial tiene un efecto estadísticamente significativo en la calidad de la atención de enfermería.

Como los servicios son tres, la prueba de comparación múltiple complementa el análisis e indica que entre los servicios A y B la diferencia no es estadísticamente significativa ($P > 0,05$), **mientras que sí lo es la que existe entre ellos y el servicio C.**

CAPÍTULO

14

SELECCIÓN DE PRUEBAS Y PROGRAMAS

INTRODUCCIÓN

Cada uno de los distintos procedimientos estadísticos que se describieron a lo largo de los diferentes capítulos tiene aplicaciones específicas en la metodología cuantitativa para la descripción de conjuntos de datos, la estimación de parámetros o la prueba de hipótesis.

El investigador debe encarar su trabajo con la selección del procedimiento estadístico más apropiado a la situación que se plantea.

Esto significa que se deben prever estos aspectos del proceso de investigación desde el momento en el que se planifica la tarea y se elabora el correspondiente protocolo, y no dejarlos sin considerar hasta algún momento posterior a la recolección de los datos.

Trabajar de esta forma ayuda a mejorar el diseño del trabajo y a hacerlo más eficiente, en el sentido de alcanzar los objetivos buscados, con el menor costo y el mejor control de la probabilidad de error en la toma de decisiones.

CRITERIOS PARA LA SELECCIÓN

La selección de un procedimiento estadístico se realiza a partir de la evaluación de los distintos aspectos del trabajo de investigación.

Uno de ellos es el tipo de planteo y diseño que está detrás del objetivo de la investigación descriptiva o de la hipótesis formulada. Por ejemplo, se analiza en esta última si se refiere al establecimiento de la posible existencia de diferencias entre poblaciones y, en este caso, si estas se plantean como nulas o de un determinado valor, como direccionales o no, o si la hipótesis se refiere a relaciones entre variables.

Se observa también la cantidad de niveles en la o las variables independientes, ya que esto determina la cantidad de grupos que se armarán para registrar datos en el experimento, así como el tipo de datos utilizados para evaluar las variables.

Con la información sobre las condiciones bajo las cuales se lleva a cabo un proceso de investigación, y al conocer los principios que fundamentan cada procedimiento estadístico, es posible seleccionar el más adecuado, y recolectar y almacenar los datos a fin de optimizar su ejecución.

—

En el **cuadro 14-1** se resume, en parte, el proceso. Las columnas están referidas a la variable independiente e incluyen la situación de inexistencia, que es el caso de la investigación descriptiva, hasta la presencia de dos niveles (dos grupos para comparar) o más. Dentro de cada caso, la conformación de los grupos puede haberse realizado mediante la evaluación de esa variable con datos numéricos o nominales.

En las filas, se diferencian las situaciones dadas por el tipo de dato utilizado para la evaluación de la variable dependiente, mientras que la intersección con cada columna incluye una mención a alguna o algunas de las pruebas que pueden ser de aplicación.

La situación general parece compleja, aunque, en definitiva, no lo es en mayor medida de la que se le presenta a un profesional de la salud que enfrenta a un paciente.

Este profesional debe evaluar lo que el paciente trae y llegar a un diagnóstico. Para ello, debe conocer las distintas condiciones posibles que pueden presentarse en ese paciente y evaluar los signos y síntomas, y todo lo que surja de su historia clínica. Como esta tarea exige una amplia gama de conocimientos básicos y aplicados, el "diagnóstico estadístico" requiere el conocimiento de la metodología y de los procedimientos técnicos de la investigación científica.

CUADRO 14-1. SELECCIÓN DE PROCEDIMIENTOS ESTADÍSTICOS

Variable independiente						
Ninguna			Niveles = 2		Niveles > 2	
			Nominal	Numérico	Nominal	Numérico
Variable dependiente	Numérico	"t"	"t" (indep. o apar.)	Regresión de correlación (Pearson)	ANOVA (Tukey, etc.)	Regresión covariable
	Ordinal	Wilcoxon (signo)	Mann-Whitney	Correlación (Spearman)	Kruskal-Wallis (Dunn)	
	Normal	Binomial	Chi-cuadrado Fisher MacNem	Chi-cuadrado (tendencia)	Mantel-Haenszel	Regresión logística

Una vez logrado el diagnóstico clínico, se selecciona el plan de tratamiento que se considera apropiado. Para ello, se aplican una serie de conocimientos sobre las diferentes alternativas para evaluar las ventajas e inconvenientes de cada una de ellas. En la decisión estadística se aplica el conocimiento de las ventajas e inconvenientes de cada uno de los procedimientos aplicables para seleccionar el más apropiado.

La elección de una prueba estadística es un procedimiento de toma de decisiones que requiere de conocimientos y capacitación para su aplicación. Así como la interconsulta entre profesionales de la salud disminuye la posibilidad de errores terapéuticos, la interconsulta con el experto en estadística ayuda a lograr un diseño más eficiente para la investigación con metodología cuantitativa.

PROGRAMAS INFORMÁTICOS

La selección de un procedimiento estadístico es un proceso lógico que requiere un razonamiento por parte del investigador y de sus colaboradores; la ejecución del procedimiento puede automatizarse mediante el uso de herramientas informáticas.

—

Muchos de los programas de planillas de cálculos en los cuales se ingresan y almacenan datos incluyen funciones estadísticas y, algunos de ellos, procedimientos para análisis.

Para procedimientos estadísticos más avanzados es necesario disponer de programas específicos para estadística. Prácticamente todos ellos no solo permiten ingresar datos, sino también "importar" aquellos que fueron ingresados en programas de bancos de datos o planillas de cálculos, así como "exportar" datos a estos programas.

En algunos sitios de Internet pueden encontrarse páginas que permiten realizar diversos procedimientos estadísticos en línea.

—

Esto significa que se pueden ingresar o copiar datos en un formulario y luego requerir la realización de los cálculos necesarios para arribar al resultado buscado: valores de estadísticos de muestras, márgenes de error e intervalos de confianza, valor de alfa en pruebas de hipótesis, poder estadístico de un determinado diseño experimental, tamaño de muestra conveniente para una investigación, entre otros.

Asimismo, existen programas estadísticos de distribución libre, dentro de los cuales pueden mencionarse: el Epi Info, desarrollado por los Centros para el Control y la Prevención de Enfermedades de los Estados Unidos (https://www.cdc.gov/epiinfo/esp/es_pc.html); el Epidat, que se distribuye por un convenio entre la Organización Panamericana de la Salud y la ***Consellería de Sanidade de la Xunta de Galicia*** **(https://** www.sergas.es/Saude-publica/EPIDAT-4-2?idioma=es); y el OpenEpi (Dean AG, Sullivan KM, Soe MM. OpenEpi: ***Open Source Epidemiologic Statistics for Public Health*****, version 3.01a. www.OpenEpi.com).** Se pueden descargar desde los mencionados sitios web e instalarlos en ordenadores personales.

Además, existe una variedad de programas comerciales de diversa complejidad y cuyas características pueden consultarse en línea, como así también, en algunos casos, es posible descargar versiones de prueba que pueden utilizarse de manera gratuita durante un lapso predeterminado.

Todos los programas informáticos realizan procedimientos que, en el caso de pruebas de hipótesis, finalizan con el informe de un valor de alfa (probabilidad de cometer un error de tipo I al rechazar una hipótesis) al operador.

—

Esto representa la etapa de análisis de los datos. En la investigación descriptiva lo es el cálculo de un intervalo de confianza para un parámetro de la población de interés.

La estadística no interpreta por qué se obtuvieron los datos y por este motivo debe ser considerada solo una herramienta dentro del proceso de investigación.

—

Su empleo, al igual que el de cualquier otra herramienta, como un microscopio, debe estar inserto dentro de un procedimiento metodológico y técnicamente correcto para que el proceso cumpla de manera acertada, o por lo menos con un margen de error razonable y cuantificado, con el objetivo de ampliar el cuerpo de conocimientos de la ciencia en la que se esté trabajando.

SÍNTESIS CONCEPTUAL

- En la planificación de un trabajo de investigación debe considerarse la selección del procedimiento estadístico más apropiado a la situación que se plantea.
- La selección de un procedimiento estadístico es un proceso lógico que requiere un razonamiento por parte del investigador.
- La ejecución del procedimiento estadístico puede automatizarse mediante el uso de herramientas informáticas.
- Los programas informáticos realizan procedimientos estadísticos que, en el caso de pruebas de hipótesis, le informan al operador un valor de probabilidad.
- La estadística no interpreta por qué se obtuvieron los datos; por ello, solo debe ser considerada como una herramienta dentro del proceso de investigación.

BIBLIOGRAFÍA Y SITIOS WEB

BIBLIOGRAFÍA

Bazerque PM. Metodología de y técnicas de la investigación clínica farmacológica. Buenos Aires: Universidad Abierta Interamericana; 2016.

Dawson B, Trapp RG. Bioestadística médica. 4a ed. México: Manual Moderno; 2005.

Iglesias ME. Metodología de la investigación científica: diseño y elaboración de protocolos y proyectos. Buenos Aires: Centro de Publicaciones Educativas y Material Didáctico; 2015.

Norman GR, Streiner DL. Bioestadística. (Trad.) Madrid: Harcourt; 2005.

Polit D, Hungler B. Investigación científica en ciencias de la salud. 6a ed. (Trad.) México: McGraw Hill Interamericana; 2000.

Riegelman RK, Hirsch RP. Cómo estudiar un estudio y probar una prueba: lectura crítica de la literatura médica. 2da ed. (Trad.) Washington: Organización Panamericana de la Salud; 1992. (Publicación OPS No 531). Disponible en: http://iris.paho.org/xmlui/handle/123456789/ [última consulta: 24 jun 2019].

Ruiz A, Morillo LE. Epidemiología clínica. Investigación clínica aplicada. Buenos Aires: Editorial Médica Panamericana; 2004.

Sackett DL, Haynes RB, Guyatt GH, Tugwell P. Epidemiología clínica. Ciencia básica para la medicina clínica. Buenos Aires: Editorial Médica Panamericana; 1994.

SITIOS WEB

El listado siguiente incluye como ejemplos las direcciones URL de varios sitios de Internet que permiten obtener información sobre temas de estadística y, en algunos casos, ofrecen enlaces a páginas que permiten el procesamiento de datos en línea.

Los accesos fueron consultados durante la preparación de esta edición (22 oct 2019).

http://statpages.org/

https://www.fisterra.com/formacion/metodologia-investigacion/

https://www.statisticshowto.datasciencecentral.com/

Índice analítico

Los números de página seguidos de una “c” indican un cuadro, los seguidos de una “f” una figura.

E

F

G

H

I

M

N

O

P

Q

R

S

T

U

V

Z

Macchi

Introducción a la Estadística en Ciencias de la Salud

3.ª EDICIÓN

Una dificultad frecuente para quienes se forman y trabajan en las ciencias de la salud es entender y analizar los resultados estadísticos de los documentos científicos y otros escritos con información general. Esta tercera edición de ***Introducción a la Estadística en Ciencias de la Salud***, **un texto práctico y sencillo, permitirá alcanzar la capacitación** inicial para superar esa dificultad. A partir de su lectura, las secciones de los trabajos que requieran una interpretación y valoración de datos numéricos comenzarán a dejar de ser páginas que solo puede entender un experto en estadística para transformarse en un material pleno de significados comprensibles que el lector podrá incorporar a su caudal de conocimientos.

La experiencia recogida por el autor en el desarrollo de actividades docentes en carreras de grado y posgrado le ha permitido realizar algunos cambios e incorporar conceptos que complementan los incluidos en las ediciones anteriores, aunque manteniendo el formato y criterio originales.

Desarrollada en 14 capítulos, la obra incluye herramientas pedagógicas como textos destacados para jerarquizar aspectos relevantes, descripciones claras y concisas con cuadros que complementan los conceptos explicados, ejemplos ilustrativos al final de la mayoría de los temas y síntesis conceptuales al cierre de los capítulos.

Se incluyen, entre otros, temas como: datos: tipos, características, almacenamiento y recuperación; distribución de frecuencias; muestreo; estimación de parámetros; prueba de hipótesis, prueba de t y de chi-cuadrado; análisis de variancia; estadística no paramétrica, y selección de pruebas y programas. Al final del libro se presenta un listado de textos de consulta más avanzados y enlaces a sitios web relacionados.

Sin duda, un texto de gran utilidad para los profesionales de la salud que requieren una evaluación crítica de la literatura científica, para el mejor desempeño de sus tareas asistenciales, docentes o de investigación.

EDITORIAL MEDICA panamericana