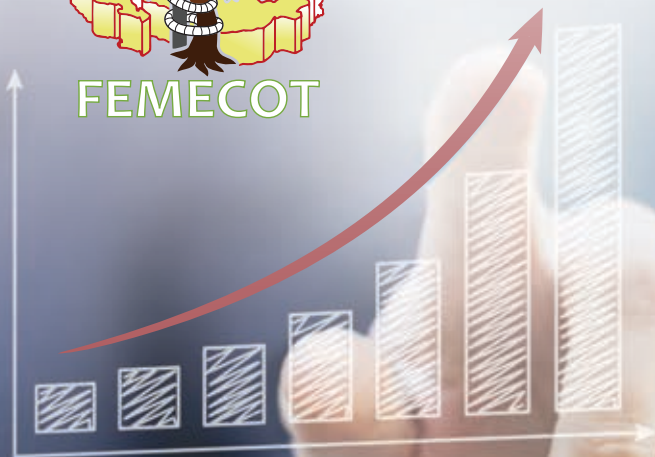
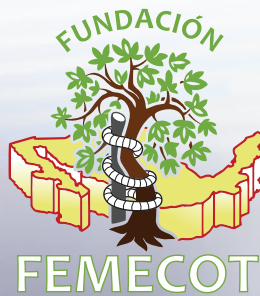
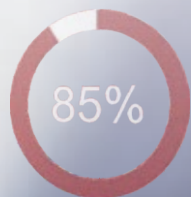


Bioestadística sin Matemáticas

El impacto de la evidencia



Bioestadística sin Matemáticas

El impacto de la evidencia

Bioestadística sin Matemáticas

El impacto de la evidencia



Acad. D. en C. José María Jiménez Ávila



En México, todos somos

FEMECOT



Bioestadística sin Matemáticas. El impacto de la evidencia
© 2021 Federación Mexicana de Colegios
de Ortopedia y Traumatología, A.C.
ISBN 978-607-98618-1-0

Ninguna parte de esta obra puede ser reproducida
o transmitida mediante ningún sistema o método,
electrónico o mecánico, sin consentimiento
por escrito del titular del copyright.
La información contenida en esta publicación es
responsabilidad de los autores y no necesariamente
refleja la opinión de la FEMECOT.

Esta obra comprende 1,000 ejemplares
más sobrantes para reposición.

Editado e impreso en México por:



Tel.: 55 8589-8527 al 32.
Julio de 2021.

Índice

- 11 **Colaboradores**
- 13 **Agradecimientos**
Acad. D. en C. José María Jiménez Ávila
- 14 **Nota del presidente de la FEMECOT**
Leonardo López Almejo
- 16 **Bioestadística sin Matemáticas. El impacto de la evidencia**
Dr. Sergio Quintero Hernández
- 17 **Nota del editor**
Acad. D. en C. José María Jiménez Ávila
- 18 **Prefacio**
Dr. José Máximo Gómez Acevedo
- Unidad 1. Principios fundamentales para
iniciar con la bioestadística**
- 20 **Introducción**
Leonardo López Almejo,
Luis Gerardo Padilla Rojas,
Israel Gutiérrez Mendoza
- 25 **La importancia de la interpretación de la
bioestadística en la práctica clínica**
Luis Gerardo Padilla Rojas,
Leonardo López Almejo
- 28 **La variable. Eje integrador en la bioestadística**
Marlene Vanessa Salcido Reyna
- 32 **El estadista. Figura elemental en el proceso del análisis**
Arelhi Catalina González Cisneros
- 34 **Bases fundamentales para la elaboración de una base de datos**
Drusso López Estrada,
Moctezuma Hernández Olive

- 37 **Elementos para elegir una muestra**
José María Jiménez Ávila,
Arelhi Catalina González Cisneros
- 41 **¿Cómo se selecciona una prueba estadística?**
Claudia Alejandra Iguual Félix,
José María Jiménez Ávila
- 45 **Bioestadística descriptiva**
Aldo Fidel Izaguirre Hernández,
Guadalupe González Gutiérrez
- 49 **Bioestadística analítica o inferencial**
Gabriela Vázquez Armenta

Unidad 2. ¿Cómo se inicia un análisis estadístico?

- 52 **Frecuencias y proporciones**
Ignacio Méndez Gómez Humarán,
Israel Gutiérrez Mendoza,
Leonardo López Almejo
- 56 **Medidas de tendencia central**
Romina L Montes de Oca Dávila,
Aldo Fidel Izaguirre Hernández
- 59 **Medidas de dispersión: rango y varianza**
José Alfredo Penagos Paniagua
- 61 **Pruebas de normalidad: Kolmogorov-Smirnov y Shapiro-Wilk**
José María Jiménez Ávila
- 64 **Pruebas de hipótesis (valor de «p») e intervalo de confianza**
José Manuel Pérez Atanasio
- 67 **¿Cómo se presentan los datos?**
Arelhi Catalina González Cisneros

Unidad 3. Pruebas paramétricas en la ortopedia

- 69 **Prueba de t de Student**
Arelhi Catalina González Cisneros
- 72 **Prueba de análisis de varianza (ANOVA) de una vía**
Laura Cortés Sanabria
- 76 **Correlación de Pearson**
Adriana Margarita Leal Moctezuma,
Aldo Fidel Izaguirre Hernández

Unidad 4. Pruebas no paramétricas en la ortopedia

- 79 **Chi cuadrada, χ^2 , ji cuadrada**
Drusso López Estrada,
María de la Luz Arenas Sordo
- 82 **Prueba de Wilcoxon**
Sergio Ramírez González,
Israel Gutiérrez Mendoza,
Leonardo López Almejo
- 88 **Prueba U de Mann-Whitney**
José Manuel Pérez Atanasio
- 91 **Prueba de Kruskal-Wallis**
José Alfredo Penagos Paniagua
- 93 **Correlación de rho de Spearman**
Laura Cortés Sanabria
- 96 **Prueba de McNemar**
José Manuel Pérez Atanasio
- 98 **Regresión lineal simple**
Víctor Toledo Infanson
- 101 **Distribución de Poisson**
Drusso López Estrada,
Carla Lisette García Ramos
- 104 **Odds ratio y riesgo relativo**
José Manuel Pérez Atanasio,
Silvestre Fuentes Figueroa

Unidad 5. ¿Qué debe saberse cuando se tienen los datos?

- 107 **Elaboración de una revisión sistemática y metaanálisis**
Marlene Vanessa Salcido Reyna
- 111 **La fuerza de la evidencia**
Víctor Toledo Infanson
- 114 **Sesgos y enigmas de la bioestadística: reflexiones**
José Luis Osma Rueda

Colaboradores

Adriana Margarita Leal Moctezuma

Médico. Facultad de Medicina «Dr. Alberto Romo Caballero» de la Universidad Autónoma de Tamaulipas.
admlealm@hotmail.com

Aldo Fidel Izaguirre Hernández

Médico ortopedista. Subespecialista en cirugía articular y artroscopia y lesiones deportivas. D. en C. Investigador de tiempo completo de la Facultad de Medicina «Dr. Alberto Romo Caballero» de la Universidad Autónoma de Tamaulipas. Hospital Naval de Tampico.
aldoizag@gmail.com

Arelhi Catalina González Cisneros

Médico. Ciencias clínicas. Instituto Tecnológico y de Estudios Superiores de Monterrey. Guadalajara, Jalisco, México.
arelhiux@hotmail.com

Carla Lisette García Ramos

Maestra en Ciencias Médicas de la UNAM. Coordinadora de investigación. División de Cirugía de Columna del Instituto Nacional de Rehabilitación (INR).
carla0608@hotmail.com

Claudia Igual Félix

Médico. Ciencias médicas. Instituto Tecnológico y de Estudios Superiores de Monterrey. Guadalajara, Jalisco, México.
a01250598@itesm.mx

Drusso López Estrada

Médico ortopedista. Cirugía de columna. Instituto Nacional de Rehabilitación. Maestría en Ciencias Médicas de la UNAM. Villahermosa, Tabasco, México.
drusso_drusso@hotmail.com

Gabriela Vázquez Armenta

Médico otorrinolaringóloga. Doctor en Ciencias Biomédicas con orientación en inmunología. Decana de la Escuela de Medicina y Ciencias de la Salud del Instituto Tecnológico y de Estudios Superiores de Monterrey en la región Occidente.
gaby@tec.mx

Guadalupe González Gutiérrez

Médico ortopedista y subespecialista en cirugía articular. Adscrita al Centro de Rehabilitación e Inclusión Infantil Tamaulipas.
lupita.gonzalez.gtz@gmail.com

Ignacio Méndez Gómez Humarán

Centro de Investigación en Matemáticas, A.C. Unidad Aguascalientes. México.
mgh2000@gmail.com

Israel Gutiérrez Mendoza

Médico ortopedista. Maestría en Ciencias Médicas. Profesor Investigador de la Universidad Cuauhtémoc Campus Aguascalientes.
drgutierrez1@yahoo.com.mx

José Alfredo Penagos Paniagua

Médico ortopedista. Maestro en Ciencias Médicas. Adscrito a la Clínica de Artroscopia y Reemplazo Articular de Rodilla. Hospital General de Zona No. 30 del Instituto Mexicano del Seguro Social. Mexicali, Baja California, México.
penagosalf@hotmail.com

José Luis Osma Rueda

Médico ortopedista. Presidente de la SCCOT. Presidente del Comité de Ética e Investigación Científica de la Universidad Industrial de Santander. Director

de Grupos de Investigación GRICES-UIS y CLEMI-Sociedad Colombiana de Cirugía Ortopedia y Traumatología (SCCOT). Bogotá, Colombia.
osmarueda@gmail.com

José María Jiménez Avila

Médico ortopedista. Director regional de servicio social del Instituto Tecnológico y de Estudios Superiores de Monterrey. Centro Médico Nacional de Occidente. IMSS. Clínica de Columna. Guadalajara, Jalisco, México.
josemajimeneza@hotmail.com

Laura Cortés Sanabria

Médico nefrólogo. Directora de Educación e Investigación del Centro Médico Nacional de Occidente. Instituto Mexicano del Seguro Social. Guadalajara, México.
cortes_sanabria@yahoo.com.mx

Leonardo López Almejo

Médico ortopedista. Cirujano de mano. Presidente de la Federación Mexicana de Colegios de Ortopedia y Traumatología, A.C. (FEMECOT). Profesor titular del Postgrado en Cirugía de Plexo Braquial y Nervio Periférico. Universidad Autónoma de Coahuila. Escuela de Medicina. Torreón, Coahuila.
lyon77777@hotmail.com

Luis Gerardo Padilla Rojas

Médico ortopedista. Cirugía articular y artroscopia. Hospital Puerta de Hierro Andares. Zapopan, Jalisco. Miembro del Cuerpo Directivo FEMECOT. Miembro fundador y secretario de la Fundación FEMECOT, A.C.
lupadilla@gmail.com

Manuel Pérez Atanasio

Médico ortopedista con adiestramiento en remplazos articulares. Maestro en Ciencias Médicas. Hospital General Regional No. 2 «El Marqués», IMSS. Querétaro, México.
drmanuelperez@yahoo.com

María de la Luz Arenas Sordo

Médico especialista en genética médica. Doctora en Ciencias Médicas de la UNAM. Investigadora del Instituto Nacional de Rehabilitación (INR). Docente de postgrado de la UNAM.
mlarenassordo@hotmail.com

Marlene Vanessa Salcido Reyna

Médico ortopedista. Maestría en Gestión Directiva en Salud. Diplomado en Investigación Científica. Mexicali, B.C., México.
drasalcidoreyna@gmail.com

Máximo Gómez Acevedo

Médico ortopedista. Cirujano articular. Presidente de la Fundación FEMECOT. Guadalajara, México.
maxgomez@hotmail.com

Moctezuma Hernández Olive

Médico ortopedista. Doctor en Educación de la Universidad de Morelos (UM). Jefe del Servicio de Ortopedia y Traumatología del Instituto de Seguridad Social del Estado de Tabasco (ISSET).
moctezuma_61@hotmail.com

Romina L Montes de Oca Dávila

Médico. Escuela de Medicina del Instituto de Estudios Superiores de Tamaulipas - Anáhuac. Tamaulipas.
rominamodavila@gmail.com

Sergio Quintero Hernández

Médico ortopedista. Expresidente de la Federación Mexicana de Colegios de Ortopedia y Traumatología, A.C. Guadalajara, México.
quinher@hotmail.com

Sergio Ramírez González

Investigador en la Universidad Autónoma de Aguascalientes. México.
sramirez@correo.uaa.mx

Silvestre Fuentes Figueroa

Médico ortopedista pediatra. Editor de la Revista Mexicana de Ortopedia Pediátrica. Médico adscrito del Hospital Shriners para niños, Ciudad de México.
fuentessilvestre@gmail.com

Víctor Toledo Infanson

Médico ortopedista. Coordinador del Servicio de Ortopedia del Hospital General de Agua Prieta. Servicios de Salud Sonora. Editor de la Revista Orthotips AMOT. Sonora, México.
dr.victor.toledo@gmail.com



Agradecimientos

Acad. D. en C. José María Jiménez Ávila

En la vida siempre hay que ser agradecidos, y particularmente con todos los autores y coautores, que han participado de manera activa, con el único interés de compartir sus experiencias con el objeto de que los cirujanos ortopédicos, se motiven a compartir sus experiencias, llevándolas a un artículo científico.

Gracias por su valiosa contribución en la elaboración de cada uno de los capítulos, donde han logrado plasmar las ideas complejas que tiene la bioestadística en conceptos más sencillos, que se acercan al sentido común.

Igualmente, se agradece a la FEMECOT por la confianza y por el apoyo incondicional en la elaboración de este volumen.

Tener un segundo libro, creado desde las entrañas de lo más valioso que tiene la Federación, que son todos sus miembros, es motivo de una gran satisfacción, ya que se tiene la seguridad de que vendrá a fortalecer las herramientas que tiene el cirujano ortopeda al realizar su actividad quirúrgica diaria y que permitirá aportar ese extra que se necesita, a fin de transformar los conceptos científicos que incrementan el área del conocimiento.



Nota del presidente de la FEMECOT

Leonardo López Almejo

Federación Mexicana de Colegios de Ortopedia y Traumatología, A.C. (FEMECOT).

La investigación en medicina y en el área de la ortopedia y traumatología es un gran reto. Generar las herramientas que faciliten el camino del investigador es una satisfacción.

Para la Federación, este libro representa el esfuerzo de un grupo de personas cuyo objetivo es permear la experiencia a través del trabajo coordinado, que genera la posibilidad infinita de crear nuevas áreas del conocimiento y solucionar problemas prácticos con base en una metodología y un proceso científico.

El libro *Bioestadística sin Matemáticas. El impacto de la evidencia* tiene como iniciativa esencial otorgar fundamentos sólidos a los cirujanos ortopedistas interesados en la materia, que son y serán parte del futuro en la investigación en México y América Latina y que, al contar con la experiencia plena, ofrecerán nuevas alternativas en la atención médica de alta calidad, otorgando un gran valor en la preparación de escritos médicos que soporten el rigor de un análisis crítico y puedan aportar a la ortopedia mexicana.

La riqueza de este volumen permitirá al lector entender el concepto básico de la bioestadística, la que se

puede definir como el método objetivo, racional y matemático a través del cual una hipótesis científica se puede comprobar, mediante aquellos datos numéricos debidamente capturados, validados, elaborados, analizados e interpretados, que se requieren dentro de las acciones de salud. Y que facilitará transformar lo difícil en algo sencillo, fomentando estrategias efectivas que garanticen la creación y el buen desarrollo de la indagación, favoreciendo un mejor nivel de evidencia en los escritos científicos realizados.

La estadística posee un sinnúmero de técnicas que dan pie para ahondar en el estudio de la variabilidad y distinguir si su valor corresponde a un individuo normal o si, por el contrario, pueda representar una enfermedad.

No se debe olvidar que a través de la historia se han hecho esfuerzos por validar el conocimiento y quitar paradigmas milenarios en el manejo de enfermedades del sistema musculoesquelético.

Es por ello que se agradece al grupo de profesores, que han tenido a bien apoyar con su conocimiento y experiencia, ética y humanismo, la creación de este libro con horas de esfuerzo, y así transmitir la enseñanza sistematiza-

da y consciente, con principios y valores, que ayudarán a la exploración integral de las enfermedades, su evolución y efecto, en beneficio de muchos, por el bienestar de los pacientes.

Finalmente, se concluye mencionando que al investigar se genera sa-

lud, y al publicar, se aprende a desempeñarse de una mejor manera con los propios pacientes y con la sociedad, facilitando el proceso educativo del conocimiento.

En México, todos somos FEMECOT.



Bioestadística sin Matemáticas. El impacto de la evidencia

Dr. Sergio Quintero Hernández
Presidente de la FEMECOT 2018-2020.

Para la Federación, y para un servidor en lo personal, esta obra es una gran satisfacción, porque se van cumpliendo los objetivos que se ha planteado en relación con la investigación, en el área de la ortopedia y traumatología.

El libro *Bioestadística sin Matemáticas. El impacto de la evidencia* será una gran herramienta, que sin duda orientará a los ortopedistas a realizar un mejor análisis estadístico en sus indagaciones, logrando interpretar los resultados obtenidos de sus datos y tener evidencias fuertes y confiables, invitándolos a que sus experiencias, que son muy valiosas, brinquen ese gran obstáculo que es escribir y poder dejar plasmado su conocimiento en un manuscrito y posteriormente publicarlas, con el propósito de que formen parte de los conocimientos que puede aportar la ortopedia mexicana a la ciencia a nivel local, nacional e internacional.

El contenido de este libro les facilitará entender los principios fundamentales de las «variables», así como

crear una base de datos; en la medida en que conozcan y sepan utilizar la bioestadística, mediante la aplicación de un análisis con las respectivas pruebas, tanto paramétricas como no paramétricas, podrá además conducirlos a darle un valor diferente a sus resultados, lo cual les permitirá contar con la evidencia con validez científica del manejo de sus pacientes; por ello los invito a aprovechar esta obra, que ha sido preparada pensando en convertir lo difícil en algo *sencillo* y con *sentido común*.

Por último, quiero agradecer a todos los profesores altamente reconocidos y personas relacionadas con la salud en el campo profesional y de la investigación, que colaboraron en la creación de este libro, haciendo a su comunidad participe de sus conocimientos y experiencias, para hacer que la investigación parezca ser más fácil y que seguramente será un gran avance en la profesionalización de los ortopedistas agremiados a la FEMECOT.

Todos somos FEMECOT.



Nota del editor

Acad. D. en C. José María Jiménez Ávila

“Objetividad y verdad” conceptos probabilísticos que pueden dar una falsa seguridad a los profesionales en sus decisiones.

Acad. D. en C. José María Jiménez Ávila

La evidencia se define como una certeza clara y manifiesta de lo que no se puede dudar, y la profesión médica se encuentra ligada a un proceso metodológico denominado «medicina basada en evidencia», situación que enciende una alerta, ya que la metodología, como concepto, hace referencia al proceso de exploración que permite cumplir ciertos objetivos en el marco de la ciencia, por lo tanto es una pieza esencial de toda investigación, o bien del proceso que se utiliza para hacer un diagnóstico y, en este caso, se ha denominado propeuéutica, lo que facilita sistematizar los procedimientos y técnicas que se requieren con el fin de concretar el desafío de realizar un diagnóstico etiológico preciso.

Esta reflexión conduce a concluir que, desde el momento en que se escoge la medicina como profesión, se ha hecho investigación y se han utilizado las herramientas del método científico y se han aplicado sistemáticamente.

Pero el problema se genera cuando aparece la bioestadística, con sus

grandes pruebas e interpretaciones basadas en números y en palabras tan complejas como la famosa «significancia estadística» y que sirven de base para confirmar las hipótesis formuladas en los protocolos de investigación.

La bioestadística es la ciencia de las probabilidades y tiene por objetivo cuantificar y analizar o inferir los resultados, en el entendido de que éstos no dan certeza de ningún evento, simplemente indican la probabilidad de que algo pueda suceder y da paso a tomar decisiones lo más cercanas a la realidad, lo cual le da un valor predictivo.

La vida es probabilística, «cruzar la calle tiene un riesgo y lo tienes que hacer», «es una probabilidad que tú aceptas», ya que le apuestas a las mejores probabilidades; los resultados de la investigación así son, siempre existirá 95% de que las cosas salgan como se esperan y siempre existirá un 5% de que las cosas no ocurran como se planean y de ahí aparece el *santo grial* en bioestadística y es la famosa $p < 0.05$, que lo hace estadísticamente significativo, pero eso no es sinónimo de clínicamente importante.

No olvidar tampoco que los números son crudos y no entienden de razones, por lo que los datos, para ser interpretados adecuadamente, deben contener un sentido de *plausibilidad*, lo cual le da esa respuesta que se asemeja al *sentido común*.



Prefacio

Dr. José Máximo Gómez Acevedo
Presidente FEMECOT 2016-2018.

Ninguna investigación humana puede ser llamada ciencia real si no puede demostrarse matemáticamente.
Leonardo da Vinci

El mundo encontrará un nuevo sentido de las cosas y nunca más volverá a ser el mismo, ya que las experiencias novedosas generarán conocimientos inéditos y destrezas renovadas, la clave es darle utilidad a este naciente conocimiento de forma que los errores y las desgracias no se repitan.

¿Qué hacer con el conocimiento nuevo, cómo usar los datos, se pueden medir los resultados e inferir las consecuencias de los mismos?, en su primera entrega de un libro 100% FEMECOT, la Federación Mexicana de Colegios de Ortopedia y Traumatología, A.C. proporcionó al lector las bases y la metodología con la meta de obtener indicadores y conocimiento válido y las formas de discriminar entre cifras útiles e información intrascendente.

La diferencia entre una moda intelectual y el aprendizaje asertivo es que la primera es fútil y pasajera, sin enriquecer en ningún sentido a quien la practica, en cambio, el aprendizaje asertivo genera cambios de conducta y conocimiento trascendente que inciden tanto individual como colectivamente en el desarrollo de las per-

sonas y las sociedades, obteniéndose de esta manera beneficios comunes a todos los núcleos de avance profesional y social.

En su segunda entrega, la Federación Mexicana de Colegios de Ortopedia y Traumatología, A.C. ha querido proporcionar al lector las herramientas necesarias para darle forma y sentido a los nuevos conocimientos adquiridos con la metodología expuesta en el primer tomo, publicación con impacto en ortopedia y traumatología.

En esta segunda edición, el comité editorial y el comité científico de la FEMECOT hacen una disquisición extensa pero muy práctica de los métodos de medición de los datos obtenidos en una investigación, a modo de con ellos hacer conclusiones valiosas y aprovechables.

Presentado en cinco unidades que inician con los conceptos fundamentales de la bioestadística y cómo emprender un análisis estadístico, así como las pruebas paramétricas y no paramétricas que se pueden hacer en ortopedia, hasta, finalmente, qué hacer con todo ese conocimiento y qué se debe saber de esos resultados, del mismo modo que se pueda entender qué hacer con ellos, este segundo volumen lleva de la mano al lector, en forma llana y sencilla, a recordar conocimientos adquiridos durante su forma-

ción básica, que pudieron haber quedado en suspenso por la falta de uso, incompatibilidad con sus actividades cotidianas, inmersión en actividades laborales que obstaculizan el desarrollo de investigación o simplemente el alejamiento por inercia del registro e investigación, que son la herramienta indispensable para la realización del análisis estadístico y la obtención de conclusiones.

Ahora, más que nunca, es necesario que el ortopedista registre, pondere, analice y publique su quehacer diario, compartir las experiencias propias; ubica al médico en una participación recíproca que tiene un gran valor, produciendo una riqueza incommensurable que se traduce en beneficios personales, gremiales y generales, dándole un sentido de trascendencia diferente a la actividad diaria y posicionando tanto al individuo como colec-

tivamente ante una sociedad cada vez más participativa, informada y ávida de prácticas profesionales sustentadas en evidencia científica.

La medición de esta práctica y su confrontación con otros, abre la puerta a encontrar similitudes y discrepancias que aportan información veraz y eficaz para las buenas prácticas en ortopedia y traumatología.

Esta obra es producto del esfuerzo de trabajo continuo y perseverante de los comités científico y editorial de la Federación Mexicana de Colegios de Ortopedia y Traumatología, en ella se plasma el amor por la profesión, la inquietud por la evolución y la convicción por el trabajo solidario en beneficio de la cofradía ortopédica, de la sociedad mexicana y de la humanidad en general, es justo reconocer su labor desinteresada y su entrega al desarrollo académico de la institución.

Unidad

1

Principios fundamentales para iniciar con la bioestadística

Introducción

Leonardo López Almejo,
Luis Gerardo Padilla Rojas,
Israel Gutiérrez Mendoza

En un mundo globalizado, donde minuto a minuto se tienen puntos de vista, opiniones, resultados, análisis y conclusiones sobre el conocimiento que se aplica, es pertinente recordar y valorar los esfuerzos específicos que se han hecho a lo largo de la historia moderna y así validar el conocimiento que se usa cotidianamente y, en algunos casos, para desechar paradigmas milenarios en el enfoque de las enfermedades musculoesqueléticas. En la evolución humana ha existido siempre una inquietud básica que es la búsqueda de la verdad, reflexionando sobre el conocimiento, qué lo hace verdadero y cómo es posible asegurarse de ello.

Vale recordar que la palabra escrita permanece, pues no se puede hacer un análisis crítico y científico de los dichos; la investigación genera conocimiento, facilita la solución de problemas y contribuye a mejorar la calidad de vida.

Cuando se decide generar nuevos saberes mediante un proyecto de investigación, se requiere escribir un protocolo, recabar y organizar los datos obtenidos de los propios pacientes o de variables estudiadas, siendo el siguiente paso el análisis de dichos números, ahí es donde tiene cabida la *estadística*, que se define como una

ciencia que se interesa por el acopio, presentación y resumen de las cifras y la obtención de información a partir de ellas, con el propósito de estudiar posibles relaciones entre incógnitas de interés para el ser humano.

Una primera clasificación de esta disciplina, acorde con la definición anterior, distingue entre el diseño de experimentos, la estadística *descriptiva*, y la *analítica o inferencial*, considerando además la estadística matemática y la estadística aplicada como una segunda clasificación.

También es importante mencionar la definición dada por Barlett: «*La Estadística es la Ciencia que nos indica el proceso a seguir en el tratamiento de la información en aquellas circunstancias que envuelva la incertidumbre*».

Con base en lo anterior, se tendrá la necesidad de aplicar estos conocimientos a proyectos de investigación relacionados con las ciencias de la salud, es decir la bioestadística, la cual se podrá definir como el método objetivo, racional y matemático a través del cual una hipótesis científica puede ser comprobada, mediante todos aquellos datos numéricos debidamente capturados, validados, elaborados, analizados e interpretados que se requieren para las acciones de salud.

Las estadísticas de salud se pueden agrupar en:

1. Estadística de poblaciones o demográficas.
2. De hechos biológicos que tienen trascendencia sanitaria, como los nacimientos y defunciones: estadísticas vitales.
3. De la enfermedad, que se intenta prevenir y tratar: estadísticas de morbilidad.
4. De los medios tanto específicos como inespecíficos que buscan proteger, fomentar, detectar y recuperar la salud: estadísticas de recursos.
5. De las acciones e intervenciones, que desarrollan los recursos básicos para la evaluación: estadísticas de servicios.

Para todo ello, es fundamental conocer algunos conceptos y términos básicos:

Población: es el objeto del estudio. Se trata de un concepto bastante abstracto, aunque en el caso de las ciencias de la salud se seguirá normalmente la acepción común del término, es decir, un amplio colectivo de individuos. Además entendiendo como grupo de estudio a personas, animales, órganos, células, etc.

Carácter y variables: sobre la población se estudiarán uno o varios caracteres. No se dará una definición de carácter, sino que se entenderá como una noción común. Son ejemplos de caracteres: el sexo, la edad, el peso, la talla, el nivel de colesterol, etc. La expresión de un carácter en cada individuo da lugar a una función o aplicación matemática que, en el contexto estadístico, se denomina variable aleatoria.

Muestra: se va a estudiar un cierto carácter que dará lugar a una variable, denótese por «X», sobre una población, que suele ser demasiado grande. Ello obliga a contentarse con estudiar el carácter sobre un subconjunto de

«n» individuos de la población. Dicho subconjunto se dice que es una muestra de tamaño «n». Es posible entender por muestra tanto a los «n» individuos, como a las «n» cifras correspondientes a la medición de la variable. En todo caso, la letra «n» queda reservada para denotar el tamaño de muestra.

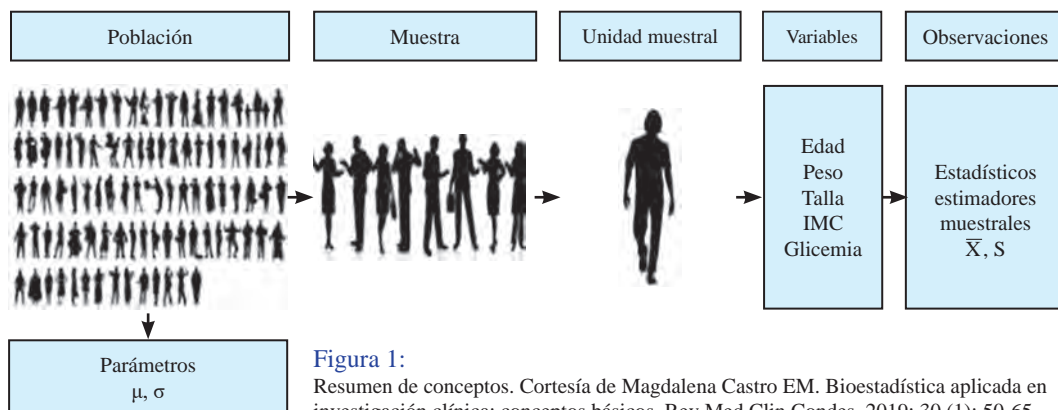
Parámetro: es cualquier número, resumen de los elementos de una población. En otra definición de Villarroya, LA 2018, el parámetro es una función de los datos calculada en la población. Los parámetros se denotan con letra griega, tales como « μ », que es la media de una variable que ha sido calculada en la población total; es la varianza de esa cantidad calculada en la población total. Como la mayoría de la investigación se realiza en una muestra, no es factible obtener un parámetro, sólo es posible calcular una estimación, a través de los estadísticos.

Estadísticos: es cualquier número resumen de una muestra. También se le puede llamar parámetro estimado, que es el término correcto cuando se habla de un número resumen que deriva de una muestra. Como es sólo un estimador del parámetro, se denota con letras en latín; «X» es la media de una muestra; «S», es la varianza de una muestra.

Observar: es medir. Lo observado es lo medido. Las observaciones son las mediciones que se realizan en las unidades muestrales.

Unidad de análisis o unidad muestral: es el objeto con la característica de interés que será observado o que será medido, pueden ser personas, cortes histológicos, crecimiento de colonias, entre otros.

Atributo: es la característica de la unidad de análisis que se va a observar. Por ejemplo: si se desea observar antropométricamente al recién nacido (RN), se deben consignar los atributos a medir: edad gestacional (EG), peso al nacimiento (PN) y talla de nacimiento (TN).



Al analizar esa información, se hace un resumen enlazando estos conceptos de una manera práctica (*Figura 1*).

Media: es la medida de tendencia central utilizada para describir las variables (continuas) ya comentadas. Se obtiene sumando todos los valores en una muestra y dividiendo entre el número de valores sumados, siendo así como se conforma la media aritmética conocida como promedio. La media siempre viene acompañada de su medida de dispersión, es decir, si la edad promedio en una muestra fue de 60 años, es necesario saber en torno a qué edades oscilaba la mayoría de los sujetos estudiados; esta oscilación de valores es la dispersión. Y la medida de dispersión de los datos sobre su media es la desviación estándar. Tras considerar el ejemplo anterior, donde la edad promedio es de 60 años y presenta una desviación estándar de cinco años, significa que alrededor de 70% de los sujetos tienen entre 55 y 65.

Valor de p: representa la probabilidad de que dichos resultados se deban al azar y no a la variable que se considera como causa. En términos generales, en los estudios clínicos es aceptado que el azar intervenga en no más de 5% y expresado con base en 100, es el 0.05, derivando de aquí la palabra «estadísticamente significativo», que en realidad debería de ser: existe

menos de 5% de probabilidad de que los resultados se deban al azar.

Odds ratio (razón de momios): es una medida de asociación entre variables de causa y efecto, ejemplo: el número de veces que se incrementa la posibilidad de padecer cáncer (variable de resultado) si se es fumador (factor de riesgo).

En los últimos años, las ciencias de la salud han experimentado un importante proceso de cuantificación: además del uso tradicional de información cualitativa, como puede ser el aspecto de una herida, o el estado general del enfermo, se ha aprovechado el desarrollo de la tecnología para la determinación de ciertas cantidades numéricas que pudieran tener alguna relación con el bienestar del paciente, como pueden ser la presión sanguínea, el nivel de glucosa en suero, etcétera. Si, salvo anomalías, algunas de esas cantidades permanecen invariantes de un individuo a otro (número de ojos, o de brazos, por ejemplo), otras reflejan lo que se suele conocer como variabilidad biológica, es decir, el hecho de que una persona es siempre diferente de otra persona (e, incluso, un ser humano es diferente a sí mismo en distintas etapas de su vida); como consecuencia, dos personas diferentes tendrán, en general, diferentes niveles de glucosa (o de colesterol) en su suero sanguíneo.

La estadística posee un conjunto de técnicas que habilitan al investigador para profundizar en el estudio de la variabilidad que pueda presentar un conjunto de datos de esta naturaleza, y distinguir, por ejemplo, si el valor de una cierta variable para un individuo concreto se puede considerar normal o si, por el contrario, podría ser un indicio de la presencia de una cierta enfermedad. Conscientes de ese hecho, las revistas de investigación en el campo biomédico exigen un tratamiento estadístico riguroso de los números; algo análogo se puede decir de la investigación en biología.

Otro campo de aplicación de la estadística en las ciencias de la salud tiene que ver con su dimensión social, pues las autoridades sanitarias de un país necesitan tener una idea clara de las características de la población a la que se quiere aplicar una cierta política sanitaria y, siendo usualmente imposible o muy costoso estudiar a todos y cada uno de los individuos de la población, se puede hacer uso de la estadística para conseguir esa información a partir de una muestra representativa de la población.

A fin de interpretar la evidencia, lo primero que se debe de leer en un artículo es el título, después el resumen. Se debe identificar si existe congruencia entre la metodología, los resultados y las conclusiones. Cuando se analiza la tabla de resultados, se tienen que observar los valores de «p». Los conceptos básicos que se han de comprender para la interpretación serán:

1. Tipos de variables.
2. Medidas de tendencia central y dispersión.
3. Tablas demográficas.
4. Valor de «p».
5. *Odds ratio*.

El concepto de la medición a la interpretación se define como un proceso esencial que vincula la in-

vestigación del alumno cotidiano y la realidad.

Debido a ello, el método estadístico se hace cada día más necesario para el profesional de la salud, tanto en su dimensión clínica, como en la administrativa y en la investigativa.

La investigación científica se ha consolidado como el mejor modo de generar conocimiento confiable, válido, verificable y reproducible; en el ámbito de la salud humana, la investigación biomédica ha permitido ampliar la comprensión de los padecimientos y avanzar en su prevención y en el tratamiento y rehabilitación de las personas en situación de enfermedad.

El método científico es un sistema ordenado de generación de conocimiento mediante pasos secuenciales, que brindan mayor certidumbre a un resultado y ayudan a obtener conclusiones válidas, que aporten verosimilitud al nuevo conocimiento generado.

Finalmente, se debe considerar que las conclusiones estadísticas son probabilidades, porque la medicina no es exacta y es sólo una parte de la verdad.

A lo anterior debe sumarse el criterio y la experiencia clínica del profesional de la salud y la necesidad del enfermo, siendo los elementos que completan el escenario en la toma de decisiones en beneficio de los pacientes.

El contexto de este libro está enfocado en el momento en que se tienen los resultados, lo cual permite fortalecer las conclusiones del proyecto, y para dar un mayor peso es preciso conocer la fuerza de las pruebas bioestadísticas, que es un proceso esencial donde las hipótesis, rodeadas de probabilidad, buscan encontrar el valor más objetivo.

La finalidad del libro es mostrar al lector ortopedista o profesional de la salud, una manera sencilla de entender y realizar el análisis estadístico de los datos obtenidos de tales trabajos de investigación, sin tener la nece-

sidad de introducirse en fórmulas o conceptos estadísticos complicados; así como revelar de manera didáctica una lógica y pasos a seguir para obtener eficientemente el análisis, obteniendo con ello una puerta abierta hacia la escritura de los resultados, conclusión, discusión y difusión de los futuros proyectos de investigación.

Conviene recordar que publicar es un signo de salud científica, y los médicos que investigan tienen un mejor desempeño con sus pacientes y su comunidad.

Lecturas complementarias

1. López ED. La investigación en la historia de la ortopedia. En: Toledo-Infanson V, Jiménez Avila JM. Publicación con impacto en ortopedia y traumatología. México: Graphimed S.A. de C.V.; 2019. pp. 23-26.
2. Padilla RLG. ¿Por qué el ortopedista debe hacer investigación? En: Toledo-Infanson V, Jiménez Avila JM. Publicación con impacto en ortopedia y traumatología. México: Graphimed S.A. de C.V.; 2019. pp. 27-30.
3. García NA. Elementos de bioestadística. 3a ed. España: Universidad de Extremadura, Servicio de Publicaciones; 2011.
4. Montanero FJ. Manual de bioestadística. Grado de enfermería. España: Universidad de Extremadura, Servicio de Publicaciones; 2016.
5. De la Loza A. Estadística médica y de la salud. En: Álvarez AR, Kuri-Morales PA. Salud pública y medicina preventiva. 4a ed. Distrito Federal: Editorial El Manual Moderno; 2012. pp. 101-46.
6. Castro M. Bioestadística aplicada en investigación clínica: conceptos básicos. Rev Med Clin Condes. 2019; 30 (1): 50-65.
7. Montanero FJ, Minues AC. Estadística básica para ciencias de la salud. España: Universidad de Extremadura, Servicio de Publicaciones; 2018.
8. Cavada G. Conceptos básicos de bioestadística. En: Burdiles P. Fundamentos del cuidado quirúrgico. Chile: Mediterráneo; 2011.
9. Pérez AJM, Miguel PA. Bioestadística básica e indispensable para la interpretación de la evidencia. En: Toledo-Infanson V, Jiménez Avila JM. Publicación con impacto en ortopedia y traumatología. México: Graphimed S.A. de C.V.; 2019. pp. 83-86.
10. Burdiles P, Castro M, Simian D. Planificación y factibilidad de un proyecto de investigación clínica. Rev Med Clin Condes. 2019; 30 (1): 8-18.

Unidad

1

Principios fundamentales para iniciar con la bioestadística

La importancia de la interpretación de la bioestadística en la práctica clínica

Luis Gerardo Padilla Rojas,
Leonardo López Almejo

Toda Ciencia es medición,
toda medición es estadística.
Hermann von Helmholtz

Como médicos ortopédicos y personas de ciencia, el análisis de los datos y la interpretación de éstos, tanto en la investigación como en la práctica clínica diaria, son necesarios para basar dicho quehacer en forma sustentable.

La medicina es una ciencia estadística, pues siempre se trabaja con la probabilidad de enfermar y con la oportunidad incierta del diagnóstico y del tratamiento. El galeno debe adecuar el conocimiento científico y tecnológico a la situación clínica personal y social del paciente del que se ocupa en ese momento, ya que «no existen enfermedades, sino enfermos». El buen juicio médico logra un equilibrio entre el riesgo que comporta toda intervención diagnóstica y/o terapéutica y el beneficio esperable de dicha intervención en el paciente concreto, lo que obliga a combinar ciencia y arte, pues no es fácil trabajar con la incertidumbre, y sólo la estadística proporciona el instrumento adecuado que concede

convivir con el azar y con las limitaciones del conocimiento médico.

Es necesario que como ortopedistas se conozcan los principios que guían la aplicación de los métodos estadísticos, porque es el método objetivo, racional y matemático a través del cual una hipótesis científica puede ser comprobada.

Al tomar en cuenta que la investigación clínica se realiza con una muestra que debe representar la población de estudio, es necesario que los datos sean matemáticamente analizados, utilizando las pruebas estadísticas apropiadas, para que se puedan extraer conclusiones científicamente válidas.

La evidencia en salud está construida con base en la bioestadística, así de relevante es esta disciplina.

La estadística es el método que se encarga de recopilar, organizar, resumir, analizar e interpretar datos numéricos, con el objetivo de llegar a conclusiones y ayudar a la toma de decisiones; asimismo, estudia la variabilidad, así como el proceso aleatorio que la genera, siguiendo las leyes de probabilidad.

La bioestadística es la rama que se enfoca en los problemas planteados

dentro de la biología, genética o medicina, con esa meta, pone en práctica los métodos de recolección e interpretación de cifras propias de la estadística y los rigurosos procedimientos del método científico. Sus procesos son aplicados a estudios también de ecología y genómica, así como salud pública, nutrición, epidemiología.

Por medio de la bioestadística se pueden obtener cifras importantes de forma sistematizada de modo que, luego de ser ordenados e interpretados, puedan ser utilizados para elaborar o contrastar hipótesis de trabajo.

Indudablemente, el pensamiento estadístico ha permitido establecer un sistema organizado de investigación, desde el diseño del mismo: el muestreo, el control de calidad, el análisis y la presentación de la información; de esta manera, ha concedido resolver y optimizar la metodología que logra dar respuesta a las diversas hipótesis que se manejan en el mundo de las ciencias de la vida.

En todo trabajo, proceso o investigación dirigido a obtener información cuantitativa, la estadística es aplicable.

El primer médico que aplicó las metodologías de la estadística en el campo de la salud fue el francés Pierre Charles-Alexandre Louis. En el siglo XVIII, él aplicó el «método numérico» para estudiar la tuberculosis; su trabajo se convirtió en la base para posteriores investigaciones que permitieron moldear el concepto de bioestadística.

Los fundamentos teóricos de la bioestadística deben ser cuidadosamente aplicados al realizar, evaluar e interpretar un trabajo científico, ya que el desconocimiento o utilización incorrecta, puede conducir a conclusiones erróneas y a malas decisiones.

Con el propósito de verificar si una diferencia encontrada en un ensayo diagnóstico es estadísticamente significativa, se debe aquilatar en el valor de «p», denominado el valor descriptivo,

que es directamente relacionado al poder de la prueba. Puede ser definido como la «mínima posibilidad de error, concluyendo que existe un significado estadístico».

Un resultado, para ser considerado estadísticamente significativo, se tiene cuando el valor de «p» es menor a un valor establecido considerado «aceptable», que generalmente es 0.05, 5% de posibilidad de error.

«Estadísticamente significativo» no necesariamente es igual a «clínicamente importante». El valor de «p» es influido por el tamaño de la muestra, grandes muestras tienden a tener valores de «p» más bajos y sus resultados tienden a ser con menor significado clínico, por el contrario, con muestras pequeñas tiende a existir un mayor valor de «p».

Es necesario aprender a diferenciar entre lo *estadísticamente significativo* y lo *clínicamente relevante*, dos términos que hacen referencia a situaciones totalmente distintas.

Hallazgos estadísticos:

1. Reflejan la influencia del azar en el resultado.
2. Derivan de grupos de individuos.
3. Requieren cantidades adecuadas de datos para ser válidas.
4. Las respuestas estadísticas son probabilísticas.
5. El análisis estadístico siempre requiere medición.

Hallazgos clínicos:

1. Reflejan el valor biológico del resultado.
2. Se practica en individuos específicos.
3. Las decisiones clínicas deben tomarse a menudo con datos insuficientes.
4. Los tratamientos médicos requieren decisiones oportunas comprometidas.
5. La práctica clínica no puede medir fácilmente cada función, cada sistema, cada signo o síntoma clínico.

Vale recordar que, en vez de hablar de «estadísticamente significativo», es más apropiado utilizar el concepto de «relevancia clínica»; porque la relevancia clínica de un fenómeno va más allá de cálculos matemáticos y depende de varios factores como la gravedad del problema, la morbilidad y mortalidad generada por el mismo, la magnitud de la diferencia, la vulnerabilidad y los costos implícitos.

La siguiente aseveración es un mito: «Una buena correlación estadística (coeficiente de correlación «r» elevado) ($p < 0.05$) siempre significa que existe una relación entre las variables evaluadas y por tanto, existe una relación causa-efecto».

En un artículo publicado por Antoja Ribó, *La inocencia de las cigüeñas. Rev Lab Clin. 2010*, se describe el estudio de Bartels en Alemania, en que obtuvieron como resultado, en un mismo año, un descenso de forma significativa en la tasa de nacimientos de la población humana y el número de ejemplares de cigüeñas. La correlación entre ambos valores era alta, lo que demostraba científicamente su relación.

La conclusión fue: la baja natalidad alemana en el año de estudio se explica por la disminución del número de cigüeñas repartidoras de bebés; «La culpa era de las cigüeñas».

Esto demuestra cómo un estudio con una buena correlación estadística ($p < 0.05$) NO siempre significa que existe una relación entre las variables analizadas y por lo tanto NO implica una relación causa efecto.

Las conclusiones que se obtienen de aplicar las pruebas estadísticas sólo tienen sentido cuando las variables (datos) están bien definidas y tienen plausibilidad biológica.

Como bien menciona Pocock, los métodos estadísticos NO son un sustituto del sentido común y la objetividad, nunca deberían estar dirigidos a impresionar o confundir al lector,

sino que deben ser una contribución importante a la claridad de los conocimientos científicos.

Se debe entonces ejercer la práctica clínica apegada al método científico, estudios realizados con bases metodológicas adecuadas, estudio de variables y análisis estadísticos rigurosos sin ser reemplazados por la experiencia, la perceptividad, el sentido común y la intuición.

La estadística es un instrumento que, si no se usa adecuadamente, puede demostrar cualquier mentira. Las mediciones y afirmaciones de probabilidad pueden ser de gran ayuda en la práctica médica, sin embargo, no hay que olvidar ser críticos en el análisis y la interpretación que todo estudio rigurosamente científico y con fundamentos estadísticos merece.

Lecturas complementarias

1. Lopes B, Ramos IC, Ribeiro G, Correa R, Valbon B, Luz A, et al. Biostatistics: fundamental concepts and practical applications. *Rev Bras Oftalmol.* 2014; 73 (1): 16-22.
2. Cantú Martínez PC, Gómez Guzmán LG. El valor de la estadística para la salud pública. *Rev Salud Publica Nutr.* 2003; 4 (1). Disponible en: <http://respyn2.uanl.mx/iv/1/ensayos/bioestadistica.html>
3. Castro M. Bioestadística aplicada en investigación clínica: conceptos básicos. *Revista Médica Clínica Las Condes.* 2019; 30 (1): 50-65.
4. Díaz Portillo J. Guía práctica del curso de bioestadística aplicada a las ciencias de la salud. Madrid: INGESA; 2011.
5. Pocock SJ. The role of statistics in medical research. *Br J Psychiatry.* 1980; 137: 188-190.
6. Antoja Ribó F. The innocence of the storks. *Rev Lab Clin.* 2010; 3 (2): 51.
7. Neyman J, Pearson ES. On the problem of the most efficient tests of statistical hypotheses. *Philos Trans R Soc Lond Ser A.* 1933; 231: 289-337.
8. Jiménez Avila JM, Falavigna A. Educación en investigación: de la medición a la interpretación-bioestadística quirúrgica. Curitiba: TRAÇO Diferencial; 2018.
9. Forthofer RN, Lee ES, Hernandez M. *Biostatistics: a guide to design, analysis, and discovery.* San Diego, CA, United States: Elsevier Academic Press; 2006.

Unidad

1

Principios fundamentales para iniciar con la bioestadística

La variable. Eje integrador en la bioestadística

Marlene Vanessa Salcido Reyna

Quien no sabe lo que busca, no entiende lo que encuentra.

Claude Bernard

El objetivo más importante de una variable es definirla, identificarla, operacionalizarla e incorporarla a la vida diaria, así, desde el momento en que nace la idea del proyecto, se estará trabajando en él.

La variable es lo que se va a medir, entre más sencilla esté definida y más exacta sea la medición, cada estudio tendrá validez interna y externa, es decir que los resultados serán fiables y reproducibles.

El diseño del estudio, la recolección de mediciones, la creación de las tablas y el análisis estadístico requieren que estas cifras sean claras y precisas.

1. Definición de variable

Es una característica del sujeto o fenómeno que se estudia, es diferente en momentos distintos y puede medirse.

Las variables en un estudio de investigación constituirán la información recolectada para responder las preguntas de investigación y deberán encontrarse en la estructura de la hipótesis, así como la relación entre ellas (causa-efecto) (*Figura 1*).

Existen dos definiciones de variable:

Definición conceptual: corresponde al significado descrito en el diccionario.

Definición operacional: se refiere a la manera que será medida la variable.

Ejemplo: variable **obesidad**.

Definición conceptual: exceso de grasa, adiposidad.

Definición operacional: índice de masa corporal > 30.

2. Clasificación de las variables

De acuerdo con su relación:

a. **Variable independiente:** explica, condiciona y determina el cambio en los valores de la variable dependiente, también conocida como factor de riesgo, es manipulada o medida por el investigador.

b. **Variable dependiente:** describe el fenómeno o situación, es la cifra afectada, también conocida como variable de resultado, de interés o de desenlace.

c. **Variable interviniente o condicionante:** es aquella que se interpone entre la variable independiente y la dependiente, no es objeto de estudio, pero puede alterar el resultado.

De acuerdo a su naturaleza:

- a. **Cualitativos:** Su elemento de variación es cualitativo, no tienen representación numeral, no pueden ser medidas en términos de cantidad como el sexo, estado civil o jerarquía de un empleo, pueden ser nominales u ordinales.
- b. **Cuantitativos:** los elementos tienen un carácter cuantitativo o numérico, tienen representación numeral, pueden ser continuas, la unidad de medición puede ser fraccionada, como la talla, o bien puede ser discontinua, se trata de un número entero que no puede ser fraccionado como número de hijos o de embarazos.

De acuerdo con su escala o valor de medición:

- a. **Nominales:** da lugar a una clasificación o categoría (aquellas que no tienen un orden).

Ejemplos de variables nominales	Valor de medición
Sexo	Femenino/masculino
Estado civil	Casado/soltero/viudo

- b. **Ordinales:** presenta un orden en sus categorías, pero no implican grados de distancia iguales entre ellas, se utilizan para clasificar sujetos, hecho en forma jerárquica (se da un orden).

Ejemplos de variables ordinales	Valor de medición
Estado nutricional	Desnutrido/nutrido/sobrepeso/obeso
Grado de satisfacción	Muy satisfecho/satisfecho/insatisfecho
Estadios de la enfermedad	I/II/III/IV

- c. **De intervalo:** es una cantidad numérica y se puede expresar en grados y magnitudes, clasifican y ordenan categorías indicando

grados de distancia que existen entre las mismas.

Ejemplos de variables de intervalo	Valor de medición
Días de hospitalización	Días (número)
Número de embarazos	Número
Temperatura	Grados Celsius
Grasa corporal	% de grasa

- d. **De razón:** poseen un cero absoluto, medición de una variable respecto de otra.

Ejemplos de variables razón	Valor de medición
Peso	Kilogramos
Estatura	Centímetros

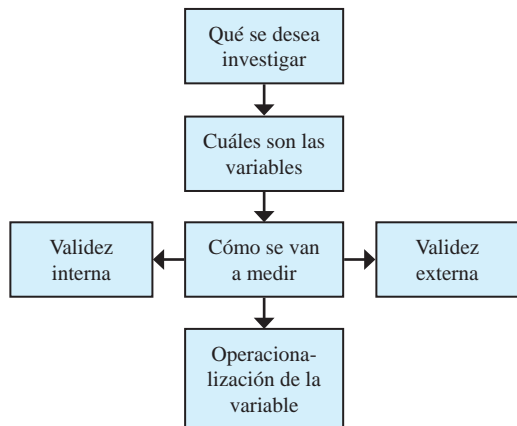


Figura 1: Estrategia para la definición de las variables.

Operacionalización de una variable

Se define la variable y cómo se va a medir durante el estudio; al momento en que se realiza el diseño del estudio, la operacionalización de las variables permitirá establecer desde la recolección de datos hasta la evaluación de los resultados.

«Proceso mediante el cual se lleva a la variable de un plano abstracto a un plano funcional.»

Indicador: es aquel que refleja la medida de la variable. Expresa razón, proporciona tasas e índices, permite hacer una valoración medible. Ejemplo: variable pobreza, el indicador: tasa de migración o desempleo.

Unidad de medida: respuesta después de la medición, puede ser cuantitativa en kilogramos, metros, o cualitativa; mucho, regular, poco.

Índice: es la expresión del indicador, ejemplo: *índice* ocupacional, porcentaje de camas desocupadas (Figura 2).

Valor: resultado que se obtiene de una medición.

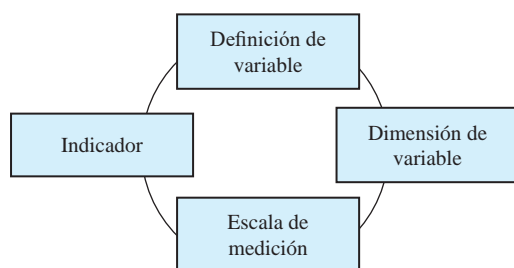


Figura 2: Organigrama en la definición de las variables.

Ejemplos de operacionalización de variables:

Denominación	Nombre que se le da a la variable
Tipo	De acuerdo con la localización en la hipótesis
Naturaleza	Cuantitativa o cualitativa
Medición	Escala de medición
Indicador	Parámetro en que se basa el inicio de la medición
Unidad de medida	De acuerdo con el indicador

Instrumento	Referir el instrumento de recolección de la información
Dimensión	Dimensión clínica, social, geográfica, biológica, diagnóstica
Definición operacional	Definiciones de trabajo de la variable
Definición conceptual	Definición con que se trata o se conoce la variable de acuerdo con el diccionario

Denominación	Edad
Tipo	Independiente
Naturaleza	Cuantitativa
Medición	De razón
Indicador	Fecha de nacimiento
Unidad de medida	Años
Instrumento	Ficha de datos
Dimensión	Biológica
Definición operacional	
Definición conceptual	Tiempo que ha vivido una persona, o animales o vegetales

Denominación	Sexo
Tipo	Independiente
Naturaleza	Cualitativa
Medición	Nominal
Indicador	Identidad sexual
Unidad de medida	% de hombres, % de mujeres
Instrumento	Ficha de datos
Dimensión	Biológica
Definición operacional	Masculino o femenino
Definición conceptual	Condición orgánica masculino o femenino

Denominación	Estado civil
Tipo	Independiente
Naturaleza	Cualitativa
Medición	Nominal
Indicador	Partida de matrimonio
Unidad de medida	% de acuerdo con la condición
Instrumento	Ficha de datos
Dimensión	Social
Definición operacional	Soltero, casado, viudo
Definición conceptual	Condición de unión social de la pareja

Lecturas complementarias

1. Jiménez Avila JM, Falavigna A. Educación en investigación: de la medición a la interpretación-bioestadística quirúrgica. Curitiba: TRAÇO Diferencial; 2018. pp. 292.
2. Núñez Flores MI. Las variables: estructura y función en la hipótesis. Invest Educ. 2007; 11 (20): 163-179.

3. Hernández Sampierí R, Fernández Collado C, Baptista Lucio P. Metodología de la investigación. 4a edición. México, D.F.: McGraw Hill Interamericana; 2010. pp. 93.
4. Jiménez AJM. Tipos de publicaciones científicas. Ortho-tips. 2015; 11 (2): 58-67.
5. Falavigna A, Jiménez-Avila JM. AOSpine. Educación en investigación: de la idea a la publicación. 2014. ISBN 978-85-7061-742-2.

Unidad

1

Principios fundamentales para iniciar con la bioestadística

El estadista. Figura elemental en el proceso del análisis

Arelhi Catalina González Cisneros

Se considera útil conocer un poco acerca de lo que es la bioestadística y su importancia. Esta especialidad es vista como aquella disciplina que se enfoca en los problemas planteados dentro de la medicina y otras áreas de ciencias de la vida como la biología y la genética.

Para realizar dicha actividad se ponen en práctica los métodos de recolección e interpretación de cifras, característicos de la estadística, y los procedimientos del método científico.

Tras tener ya la idea de la importancia de la bioestadística, hay que comentar acerca del estadista; aquella persona que materializa la base de datos, el proceso del análisis y su interpretación.

El estadista es un investigador con un perfil característico, ya que debe prepararse y conocer, como mínimo, las bases fundamentales de la terminología estadística de la investigación.

Desde dominar la pregunta PICOT, hasta conocer la diferencia entre qué y cuáles son las variables y sus tipos y cómo se lleva a cabo el proceso de operacionalización de las mismas. Poseer el conocimiento de destrezas básicas sobre análisis.

Además, el estadista conoce cómo obtener datos importantes sistematizados, luego puede ordenarlos e inter-

pretarlos y a su vez utilizarlos para elaborar o contrastar hipótesis de trabajo.

Al agregar a la parte de la preparación de un estadista, se puede incluir el conocimiento y dominio de programas como Epi Info y SPSS, que ayudan en la elaboración de la base de datos. Actividad que, al ser realizada por un experto, como el estadista, economiza tiempo de la investigación, además de disminuir sesgos. Ya que estos programas son herramientas de ayuda en el proceso de análisis de detalles cualitativos, mas no son el instrumento de análisis, tarea en la que es indispensable un estadista.

Los programas de análisis de datos sólo funcionan para llevar un control y una sistematización del proceso de análisis numérico, por lo que le corresponde al estadista asignar los significados a esas cifras. Dicho de otra manera, independientemente de la existencia de los programas para análisis de esos números, estos no identifican datos mal recogidos o sesgados, cosa en la que es fundamental un estadista.

El estadista además sabe identificar las dificultades más comunes al desarrollar un análisis de pormenores cualitativos, como: resaltar el carácter de los hechos cualitativos y todos los significados que pueden encontrarse.

El análisis realizado por el estadista es una tarea compleja con exigencia de preparación y planificación acerca del significado del análisis realizado.

Como se mencionó, el análisis numérico es una tarea laboriosa y compleja, ante esta dificultad, el estadista preparado y calificado para dicha tarea tiene, además de lo ya mencionado, la habilidad de establecer algún procedimiento usando los datos y así facilitar su comprensión.

Por lo que surge esta pregunta ¿Se necesita un estadístico como parte del equipo de investigación clínica? La respuesta es Sí, ya que los estadísticos participan en la selección apropiada del diseño de estudio, son parte sustancial en el cálculo del tamaño de la muestra, ayudan a realizar el análisis e interpretación de los resultados y son elemento importante en la obtención de conclusiones sustentadas en fuentes confiables.

¿Y en qué momento debe uno ponerse en contacto con el estadístico?; nunca es demasiado temprano, las estadísticas no se pueden realizar en un protocolo mal diseñado, después de que el estudio haya comenzado, por lo que se sugiere enviar un borrador de la propuesta de estudio al estadístico al principio o bien durante la fase de planeación.

Las consecuencias de un enfoque estadístico mal desarrollado pueden dar lugar a un estudio clínico poco sólido, que no alcance a probar adecuadamente las hipótesis.

Por lo que, cuando se tenga contacto con el estadístico, existen preguntas que esta persona debe saber siempre, como: ¿Cuál es la hipótesis de investigación?, ¿Cuál es el tipo de variable y unidad de medición?, ¿Cuál es la medida más importante? (variable de resultado primario), ¿Cuál es el tipo de diseño de estudio? y ¿Cuál es la diferencia clínicamente significativa para el resultado primario?

No es de esperarse que tenga todas las respuestas en la primera reunión, pero las conversaciones en curso con el estadístico pueden servir en el desarrollo de ideas más concretas, que en ese momento permitirá posiblemente cambiar o modificar algunos aspectos relacionados con la metodología y la estrategia de búsqueda.

Por tal motivo, si el médico está pensando en realizar un proyecto de investigación, mismo que desea publicar, debe tener en mente a un buen estadístico, ese es el mayor secreto de los grandes investigadores, todos tienen uno que termina siendo el mejor de sus aliados.

Lecturas complementarias

1. Jiménez Avila JM, Falavigna A. Educación en investigación: de la medición a la interpretación-bioestadística quirúrgica. Curitiba: TRAÇO Diferencial; 2018.
2. Rodríguez Sabiote C, Lorenzo Quiles O, Herrera Torres L. Teoría y práctica del análisis de datos cualitativos. Proceso general y criterios de calidad. Revista Internacional de Ciencias Sociales y Humanidades, SOCIOTAM. 2005; 15 (2): 133-154.

Unidad

1

Principios fundamentales para iniciar con la bioestadística

Bases fundamentales para la elaboración de una base de datos

Drusso López Estrada,
Moctezuma Hernández Olive

Actualmente, las tecnologías de la información han evolucionado para que los procesos sean cada vez más rápidos y eficientes en menor tiempo, debido a que los sistemas de computación han solucionado los mecanismos de control y recopilación de datos. Por esta razón, se tuvieron que diseñar medios para transportar los números de manera eficiente.

Dentro de los fundamentos de bases de datos es posible encontrar la teoría que incluye los principios formales para precisar y emplear cifras estructuradas e interrelacionadas, de esta manera, en la definición de los datos, se utiliza un modelo que sirva como manipulación en su lenguaje, es así como diferentes modelos analíticos se han propuesto buscando un mayor nivel expresivo que logre representar el mundo real.

Por esta razón, en los medios de comunicación y particularmente para los comunicadores, la tecnología y especialmente Internet representan un gran avance y desafío, ya que es muy útil para ellos, puesto que es una herramienta que cada día avanza y evoluciona en forma extraordinaria, de acuerdo con los desafíos que impone la sociedad.

Con todo esto, se deduce que la comunicación es vital para los seres

humanos, ya que desde que se nace hasta que se muere se siente la necesidad de comunicarse y expresar todo aquello que hay alrededor.

Con base en lo anterior, la potencia y limitaciones de los modelos que se implementan en las instituciones o empresas se pueden evaluar desde un punto de vista teórico y se evidencian desde un punto de vista práctico cuando se trata de implementarlos en aplicaciones tradicionales y modernas, estas últimas generalmente requieren tipos de datos complejos.

Los lenguajes de manipulación de números tienen como propósito ofrecer facilidad, simplicidad y flexibilidad a la hora de utilizarlos para actualizar y recuperar información desde la base de datos, estos lenguajes son en su gran mayoría declarativos, lo que reduce significativamente el tiempo de desarrollo y mantenimiento de las aplicaciones.

Dada la naturaleza declarativa de los lenguajes de consulta, el desempeño del sistema depende fundamentalmente del proceso de optimización, que garantiza la generación del mejor plan de ejecución para una consulta dada.

El optimizador de consultas utiliza algoritmos especializados a fin de evaluar e implementar las diferentes

operaciones que permiten expresar las consultas.

Reglas de transformación lógicas y físicas se aplican para producir el mejor plan de ejecución, un importante polo de investigación ha girado en torno a los sistemas de bases de datos.

Modelos y lenguajes de manipulación de cifras u optimización de consultas son, entre otros, temas permanentes de investigación, diferentes modelos de bases de datos se han propuesto en lógica, dando lugar a las denominadas bases de datos deductivas.

De esta manera, y para poder comprender o profundizar el tema, es importante mencionar los conceptos básicos de las bases de datos; como parte de ellos se mencionan los siguientes:

Base de datos (BD): conjunto de información relacionada entre sí.

Sistema gestor de base de datos (SGBD): programa de creación y manejo de bases de datos, puede manejar una o muchas bases de datos.

Modelo de datos: conjunto de conceptos y reglas de composición de estos que, combinados de alguna forma, son capaces de representar un sistema de información, tanto en su parte estática como dinámica.

Sistema de información (SI): conjunto de números, ordenadamente conectados entre sí siguiendo unas ciertas reglas, que aporta al sistema objeto (la organización a la que sirve y marca las directrices de funcionamiento) la información necesaria para el cumplimiento de sus fines, debido a lo cual tendrá que recoger, procesar y almacenar los datos, facilitando la recuperación, elaboración y presentación de estos.

Desarrollo de un SI mecanizado:

1. Análisis-investigación y modelización.
2. Diseño: lógico y físico.
3. Implementación: programas, carga de valores, pruebas.

Análisis:

Descripción lo más fiel del sistema físico.

Recogida de requerimientos:

1. Obtiene un modelo no influido por un sistema mecanizado concreto (sistema final)

Diseño:

1. Traslación de las ideas a un modelo comprensible por el sistema mecanizado (ordenador).
2. Diseño conceptual, diseño lógico, diseño físico

Implementación:

1. Creación del modelo en el sistema mecanizado (ordenador).
2. Carga de datos.
3. Pruebas de fiabilidad y rendimiento.

Hasta aquí una breve descripción de los conceptos básicos, no sin antes mencionar los elementos o características principales de los sistemas de una base de datos.

Entre ellas se encuentra como principal característica la integridad, ésta se refiere a la validez y la consistencia de los números almacenados, la cual se expresa mediante reglas o restricciones que no se pueden vulnerar.

Así mismo, la seguridad de la base de datos es la protección frente a usuarios no autorizados cuando no existen buenas medidas de seguridad, la base de datos se vuelve más vulnerable que los sistemas de archivos.

Otra característica que hace referencia a la base de datos es la concurrencia, es decir, si hay varios usuarios accediendo simultáneamente es posible que haya interferencia en el acceso, de modo que se pierda la información y al mismo tiempo la integridad.

La recuperación existe cuando el usuario es quien proporciona las medidas necesarias para proteger los datos ante cualquier alteración del sistema o en las aplicaciones, es decir que los usuarios tienen que hacer copia de

seguridad día con día y restaurarla de ser preciso.

En la actualidad, contar con una base de datos es de vital importancia porque permite un rol fundamental en las tecnologías de la información y comunicación para organizar el conocimiento de artículos científicos y conferencias, etc.

Esto también facilita agrupar y almacenar los contenidos en un lugar específico y dejarlos accesibles de modo que las personas interesadas puedan hacer uso de ellas. La base de datos en la salud tiene un papel preponderante, pues aporta de manera didáctica al ámbito de la investigación y en la formación de los profesionales, ya que facilita la consulta de la información con respecto a los problemas sanitarios, temas enfocados a las enfermedades mismas que permite la mayor utilidad y además crea un mecanismo de control en los datos para poder analizar y tomar decisiones en beneficio de la institución.

La importancia en el uso de la informática de la medicina sirve de base para el desarrollo del conocimiento científico en las diferentes disciplinas médicas, tomando como referencia la

actualización que todo médico debe poseer.

Por último, el principal objetivo de los registros de una base de datos es proporcionar información útil sobre la prevalencia, evolución, resultados y necesidades, los cuales deben cumplir con una serie de características como la validez, la exactitud, la confianza y la calidad.

A fin de que se puedan utilizar los valores sin limitaciones, es preciso mantener una correcta estrategia y fuentes interactivas de conocimientos.

Lecturas complementarias

1. Luján MS. Conceptos básicos de bases de datos. Gestión de datos. Curso 2001-2002. Universidad de Alicante.
2. Millán ME. Fundamentos de bases de datos. Santiago de Cali: Programa Editorial Universidad del Valle; 2012. p. 154. Serie. 005.74 cd 21 ed. A1335388.
3. Planas M, Rodríguez T, Lecha M. La importancia de los datos. Nutr Hosp. 2004; 19 (1): 11-13.
4. Ramírez Robledo LE, Arcila A, Buriticá LE, Castrillón J. Paradigmas y modelos de investigación. Guía Didáctica y Módulo. 2a ed. Colombia: Fundación Universitaria Luis Amigó, Facultad de Educación; 2004.
5. Hernández-Sampieri R, Fernández-Collado C, Baptista-Lucio P. Metodología de la investigación. 6a ed. México: McGraw-Hill; 2004.

Unidad

1

Principios fundamentales para iniciar con la bioestadística

Elementos para elegir una muestra

José María Jiménez Ávila,
Arelhi Catalina González Cisneros

Cuando se inicia una investigación, uno espera que esa parte que se toma de la población (conjunto de individuos con una serie de características específicas) sea representativa y eso es lo que se denomina como muestra, la cual se define como el subconjunto de la población, seleccionada de tal manera, que personifique a la población.

Y ¿por qué se debe muestrear?, la respuesta es porque las muestras pueden estudiarse con más rapidez que las poblaciones, es menos costoso y una parte importante es que en la mayor parte de las situaciones el estudio de una población entera es casi imposible.

Una vez que el investigador ha decidido qué va a estudiar y a quién va a estudiar, debe resolver cuántos individuos ha de contener la muestra (*Figura 1*).

Del tamaño de la muestra, ésta mostrará su potencia y servirá para determinar cuántos sujetos se necesitan para responder a la pregunta de investigación (PICOT). Así como el número de personas necesarias para evitar errores de tipo I y II, por lo tanto, en investigación clínica la muestra tiene que ser sólo lo suficientemente grande.

Demasiado pequeña, sin resultados concluyentes.

Demasiado grande, riesgo para los pacientes, las diferencias clínicamente no relevantes se pueden encontrar como estadísticamente significativas.

Por lo tanto, conocer el tamaño de la muestra ayuda a planificar la investigación y concede planificar, porque regularmente se tiene un presupuesto (pacientes, más caro) y un tiempo límite (más tiempos, prolongación del estudio) (*Figura 2*).

Muestra pequeña: ¿Contesta la pregunta?

Muestra grande: ¡Costosa y difícil!



Figura 1.

Para elegir una muestra representativa, se requiere incluir individuos, los cuales van a entrar al estudio y van a contribuir con su información a cumplir con los objetivos, por lo que se deben tener claros los criterios de inclusión, de exclusión y los de eliminación.

El objetivo del reclutamiento es englobar el número suficiente de individuos, de forma que para el estudio se cumplan los requerimientos del ta-

maño de la muestra y que la misma no sea sesgada.

La forma de seleccionar la muestra se puede dividir en diferentes tipos. (Figura 2):

Muestreo probabilístico: utiliza un proceso aleatorio para garantizar que cada unidad de la población tenga una probabilidad específica de ser seleccionada.

Muestreo no probabilístico: no se basan en la probabilidad, y el objetivo será producir una muestra lo más parecida posible a una muestra probabilística.

Muestreo probabilístico

1. Muestreo aleatorio simple: proceso de enumerar cada unidad de la población accesible, y a continuación seleccionar la muestra al azar. Utilizar una tabla de números aleatorios, por ejemplo: tarjetas o el *volado*.
2. Muestreo sistemático: consiste en una selección según un proceso periódico, es susceptible de errores ocasionados por posibles periodicidades naturales presentes en la población y se toma a los individuos alternativamente a partir de una lista y se elige al azar: cada 3º, 5º, 10º, etc.
3. Muestreo estratificado: consiste en dividir a la población primero en estratos o subgrupos (sexo, raza, gravedad de la enfermedad, etc.) y luego de cada estrato se selecciona una muestra aleatoria, requiriendo de una muestra más pequeña.
4. Muestreo por conglomerados: se obtiene en un proceso de dos etapas, en donde la población se divide en conglomerados (agrupaciones naturales de individuos), y de éstos se selecciona un subconjunto. En general los conglomerados se basan en áreas geográficas o distritos.

Muestreo probabilístico	Muestreo no probabilístico
Muestreo aleatorio simple	Muestreo de casos consecutivos
Muestreo sistemático	Muestreo de conveniencia
Muestreo estratificado	Muestreo a criterio
Muestreo en conglomerados	



Figura 2: Tipos de muestreo.

Muestreo no probabilístico

1. **Muestreo de casos consecutivos:** consiste en elegir a cada paciente que cumpla con los criterios de selección dentro de un intervalo de tiempo específico, o hasta alcanzar un número definido de pacientes.
2. **Muestreo por conveniencia:** proceso de seleccionar a los miembros de la población accesible de los que se pueda disponer con facilidad, se decide una característica para poder seleccionar, por ejemplo: letra del apellido, número de afiliación, fecha de nacimiento, etc.
3. **Muestra a criterio:** consiste en escoger «a dedo» de entre la población accesible a los individuos que se supone son más apropiados para participar en el estudio.

Cualquier estrategia a escoger tiene relevancia y deberá ajustarse a las características del proyecto, tiempo, factibilidad; lógicamente, entre más condiciones aleatorias se tengan, la validez y la potencia del proyecto será mayor.

Errores de diseño al elegir a los individuos y cómo prevenirlos

Error: la población diana no es adecuada para la pregunta.

Prevención: trazar criterios de inclusión que especifiquen la edad, sexo y las características clínicas de los individuos apropiados y diseñar una lista concreta de criterios de ex-

clusión que eliminen a los individuos inadecuados.

Error: la población accesible no representa suficientemente a la población diana.

Prevención: diseñar criterios de inclusión que especifiquen el marco temporal y las características geográficas de la población accesible adecuada.

Error: la muestra seleccionada no es suficientemente representativa de la población accesible, debido a la presencia de sesgos en el diseño de muestreo.

Prevención: utilizar una muestra probabilística o de casos consecutivos siempre que resulte práctico, en caso contrario, tener buen criterio al diseñar una muestra por conveniencia y obtener una muestra piloto para confirmar las características y la disponibilidad de los individuos.

Errores de ejecución al elegir a los individuos y cómo prevenirlos

Error: la muestra que se ha obtenido en realidad no representa suficientemente a la muestra inicialmente diseñada, debido a errores aleatorios de muestreo (azar).

Prevención: incrementar el número de individuos o emplear otras estrategias y considerar el muestreo estratificado para ampliar subgrupos específicos.

Error: sesgo por falta de respuesta: por imposibilidad de contactar al individuo, que el individuo se niegue a participar o por equivocaciones no intencionales.

Prevención: considerar incentivos, como el pago de los transportes o la comunicación de los resultados o pruebas, considerar actuaciones especiales para recoger información sobre los individuos que no respondieron, comprobar los métodos de muestreo antes de iniciar el estudio y controlar la

calidad del proceso de reclutamiento durante el estudio.

Fórmula para calcular el tamaño de la muestra

El tamaño de la muestra es de suma importancia para darle validez y potencia a los resultados de la investigación, existen actualmente calculadoras electrónicas que facilitan el trabajo matemático, pero en ocasiones las cifras que aparecen no suelen ser alentadoras por el número de pacientes que se requiere, eso no debe generar preocupación, ya que como todo en la vida puede negociarse.

Por lo tanto, ésta es la fórmula más utilizada y en algunos programas computacionales es importante conocer el dato de la prevalencia de la enfermedad (Figura 3).

$$n = \frac{N \sigma^2 Z^2}{(N - 1)e^2 + \sigma^2 Z^2}$$

n = tamaño de la muestra
N = tamaño de la población



Desviación estándar de la población. Generalmente, cuando no se tiene su valor, suele utilizarse una valfa constante de 0.5. Z = Valor obtenido mediante niveles de confianza. Es una cifra constante que, si no se tiene su valor, se lo toma en relación al 95% de confianza, que equivale a 1.96 (como más usual), o bien en relación al 99% de confianza, cuando equivale a 2.58, elección que queda a criterio del investigador. e = Límite aceptable de error muestral que, generalmente, cuando no se tiene su valor, suele utilizarse un número que varía entre el 1% (0.01) y el 9% (0.09), elección que queda a criterio del investigador.

Figura 3: Fórmula para el tamaño de la muestra.

Conclusión

La muestra contiene elementos que son de suma importancia para darle validez a cada investigación y por eso debe siempre calcularse antes de iniciar el estudio, ya que se requiere utilizar una estimación de la diferencia entre los grupos de estudio que son considerados como parte de la relevancia

Elementos para elegir una muestra

clínica (literatura, opinión de expertos, estudios piloto).

Al tamaño de la muestra se debe agregar la posible tasa de deserción, que muchas veces es por falta en el seguimiento.

Uno de los valores importantes que se debe tener en cuenta es la prevalencia de la enfermedad, ya que hasta los programas estadísticos requieren de ese dato para otorgar un resultado matemático.

Lecturas complementarias

1. Salazar Pinto C, Del Castillo Galarza S. Análisis y descripción de los datos. En:

Salazar Pinto C, Del Castillo Galarza S. Fundamentos básicos de estadística. Quito; 2018. pp. 67-71.

2. Deshpande S, Gogtay NJ, Thatte UM. Measures of central tendency and dispersion. *J Assoc Physicians India*. 2016; 64 (7): 64-66.
3. Guerra-Dávila T. Estadística descriptiva. En: Guerra-Dávila T. Bioestadística. México, D.F.: UNAM-FES Zaragoza; 2014. pp. 86-88.
4. Manikandan S. Measures of dispersion. *J Pharmacol Pharmacother*. 2011; 2 (4): 315-316.
5. Mishra P, Pandey CM, Singh U, Gupta A, Sahu C, Keshri A. Descriptive statistics and normality tests for statistical data. *Ann Card Anaesth*. 2019; 22 (1): 67-72.
6. Guzmán Pantoja JE, Jiménez Avila JM. Manual de análisis estadístico y uso de base de datos. México: Universidad de Guadalajara; 2010.

¿Cómo se selecciona una prueba estadística?

Claudia Alejandra Igual Félix,
José María Jiménez Ávila

Un análisis estadístico es básico en una investigación, por lo que siempre debe estar presente (Flores-Ruiz y cols., 2017). De acuerdo con Tetitla (2017), la importancia del análisis estadístico estriba en tomar la decisión de si la hipótesis se acepta o se rechaza; de esa dimensión es la importancia de elegir correctamente la exploración estadística que se va a utilizar.

Existen diferentes pruebas estadísticas, por lo que el objetivo de la revisión será exponer al lector de este texto la forma correcta y, de manera simplificada, la elección de una prueba adecuada, acorde con los objetivos que debe cumplir la investigación.

Las indagaciones estadísticas se utilizan buscando poder determinar una conclusión a partir de la información recolectada, para probar como verdadera o errónea la hipótesis planteada. Es importante seguir una secuencia ordenada en el proceso de investigación, donde se clasifique el diseño de la investigación, la clase de variable y su método de interpretación adecuado correspondiente.

El esquema proporcionado por Flores (2017), en su publicación *Selección de la prueba estadística adecuada (Figura 1)*, proporciona al investigador una forma ordenada en la elección de

la misma; a continuación, se describirán los elementos necesarios para llegar a la prueba estadística correcta.

El comienzo de la selección debe iniciar reconociendo si la clase de la investigación será estadística descriptiva o bien inferencial. La primera de éstas se dedica a la descripción, visualización y resumen de los datos que se originan del único grupo estudiado; por otro lado, la estadística inferencial se concentra en generar modelos, inferencias y predicciones asociados a los grupos estudiados, tomando en cuenta la aleatoriedad de las observaciones (Bazán P, 2018).

Una vez encontrado el diseño de cada estudio, el segundo paso es clasificar las variables, es decir, considerando en ellas «la cualidad, propiedad o característica de personas u objetos estudiados, que puede ser clasificada como cualitativa o cuantitativa» (Jiménez J, 2018).

La clasificación de las variables es necesaria para la selección de la prueba estadística apropiada, el investigador debe saber la manera de categorizar la información, mucho antes de ser obtenida para el estudio, y así disminuir la posibilidad de errores. Estas mediciones pueden clasificarse en cuantitativas, cuando describen una cantidad y se subdividen en con-

¿Cómo se selecciona una prueba estadística?

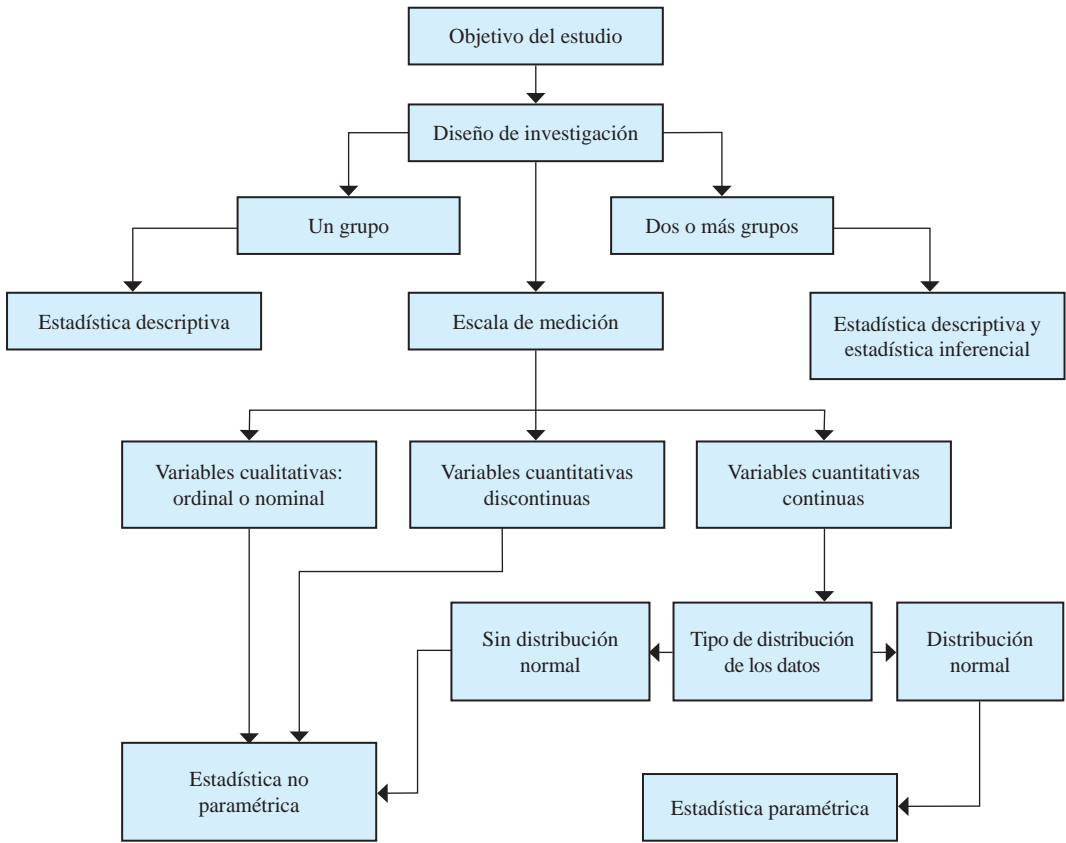


Figura 1: Proceso de selección de una prueba estadística.

tinuas, que pueden expresarse con números decimales, y discretas cuando definen números enteros. Por otro lado, las variables pueden clasificarse en cualitativas, que a su vez se subdividen en nominales y ordinales; las nominales agrupan las características en las que no hay diferencia entre una y otra, tales como el sexo (masculino/femenino), mientras las variables cualitativas ordinales se clasifican por dimensión, como el estadio o gravedad de una enfermedad (leve/moderada/grave), o la consolidación de una fractura (Tabla 1).

La clasificación de las variables es indispensable para saber cuál prueba estadística es la indicada y más eficiente al ordenar tal información. Antes de continuar con la elección de la misma, es necesario que los inves-

Tabla 1: Clasificación de las variables.

Cuantitativas (intervalares)	
Continua	Discretas (discontinuas)
Ejemplos: presión arterial, peso, edad, talla, IMC	Ejemplos: número de hijos, número de fracturas
Categorías (cualitativas)	
Ordinales	Nominales
Ejemplos: etapa-ficación de tumores, consolidación de fracturas	– Dicotómicas (p. ej.: vivo/muerto) – Policotómicas (p. ej.: grupo sanguíneo)

Proceso inspirado por: Tetitla Munive JM. Capítulo 9: Cómo elegir la prueba estadística. En: Martínez Montaña MLC, Briones Rojas R, Cortés Riveroll JGR. Metodología de la investigación para el área de la salud. México: McGraw-Hill Interamericana; 2013. pp. 71-78.

tigadores organicen su información (variables) en medidas de tendencia central y dispersión. Se continuará con dichas cantidades una vez ordenadas y será posible definir si se utilizará estadística paramétrica o no paramétrica, que son los dos grupos de pruebas estadísticas.

A los sondeos estadísticos paramétricos les corresponden las variables cuantitativas continuas; mientras que las de tipo cualitativo (ya sean nominales u ordinales), y las cuantitativas discontinuas, se incluyen en las pruebas estadísticas no paramétricas (Flores-Ruiz y colegas, 2017). De acuerdo con Nayak (2011), para los datos numéricos, es importante concretar si siguen los parámetros de la curva de distribución normal (curva de Gauss), en cuyo caso se aplican las pruebas paramétricas.

En ambos grupos, se debe identificar si las muestras son entre ellas independientes o relacionadas. Un ejemplo sería, si se quisiera comparar la recuperación de la fractura de fémur en un grupo que sigue un tratamiento específico, para evaluar la efectividad de éste, con otro grupo que padece de la misma fractura, el cual no siguiera el tratamiento. Entonces se hablaría de un estudio clínico, y tales grupos serían independientes entre ellos, evaluando el desarrollo de cada grupo, se llegaría a la conclusión de cual de esos dos grupos presenta una mejor recuperación.

Una vez completados los pasos anteriores, se continuará con la elección del ensayo estadístico, sabiendo qué clase de información se tiene, ya sea cuantitativa o cualitativa, si se sigue una distribución normal o no, y de esta manera se especificará la prueba correcta, disminuyendo lugar para errores en la investigación, y no evaluar esa información con una prueba que no contribuya a verificar o contradecir la hipótesis planteada.

La *Figura 2* es un esquema elaborado con el propósito de guiar al investigador a determinar la prueba. Si los datos presentados son cuantitativos, y se habla de dos grupos, es decir una comparación de promedios, el análisis t de Student es el más aceptable. Si se presentan tres o más grupos, se determinará si son independientes o dependientes entre ellos, si son dependientes, la prueba ideal sería ANOVA de dos vías, si son independientes, se utilizaría ANOVA de una sola vía. Es importante mencionar que las siguientes especificaciones deberán seguirse si los números obtenidos siguen la forma normal.

En el caso de una distribución que no sea normal, los datos cuantitativos deberán evaluarse con análisis como U Mann-Whitney, Wilcoxon, Kruskal-Wallis, y Friedman, seleccionando cada una en relación del número de grupos y la relación o independencia entre los grupos evaluados (*Figura 3*).

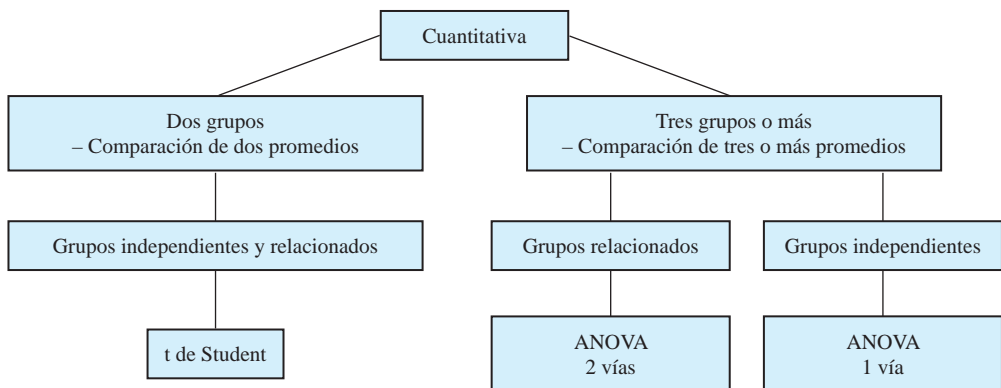


Figura 2: Selección de prueba estadística con distribución normal.

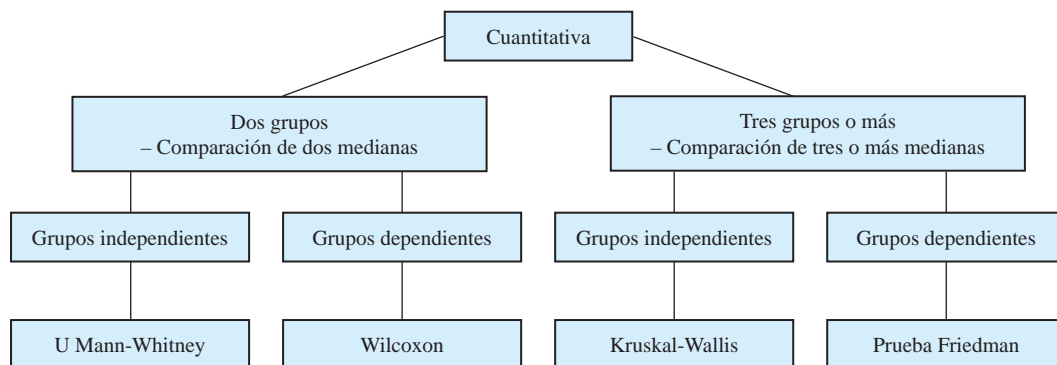


Figura 3: Selección de prueba con una distribución no normal.

Los valores que se clasifiquen como cualitativos siguen pruebas diferentes. Si se quisiera contrastar dos grupos independientes se utilizaría la prueba de χ^2 , este análisis también se utilizará cuando se cotejen tres grupos independientes, es decir una comparación de porcentajes. Si se tienen dos grupos dependientes y están relacionados, la verificación ideal será la de McNemar, si los grupos son tres o más, la prueba será la Q de Cochran.

Los sondeos estadísticos deben decidirse antes de recolectar la información necesaria en el estudio, éstas son la conclusión del manejo de información y sirven de base para comprobar o negar la pregunta de investigación, en este capítulo se manejaron las pruebas más utilizadas y una guía básica para llegar a ellas de forma correcta.

Lecturas complementarias

1. Bazán P. How to choose statistical tests. Education in research: from measurement to interpretation - surgical biostatistics. Asociación AOSpine Latino América. 2018. pp. 53-61.
2. Flores-Ruiz E, Miranda-Navales MG, Villasís-Keever MA. El protocolo de investigación VI: cómo elegir la prueba estadística adecuada. Estadística inferencial. Rev Alerg Méx. 2017; 64 (3): 364-370.
3. Jiménez Avila J. The main actor: the variable. Education in research: from measurement to interpretation - surgical biostatistics. Asociación AOSpine Latino América. 2018. pp. 29-32.
4. Nayak BK, Hazra A. How to choose the right statistical test? Indian J Ophthalmol. 2011; 59 (2): 85-86.
5. Tetitla Munive JM. Capítulo 9: Cómo elegir la prueba estadística. En: Martínez Montaña MLC, Briones Rojas R, Cortés Riveroll JGR. Metodología de la investigación para el área de la salud. México: McGraw-Hill Interamericana; 2013. pp. 71-78.

Unidad

7

Principios fundamentales para iniciar con la bioestadística

Bioestadística descriptiva

Aldo Fidel Izaguirre Hernández,
Guadalupe González Gutiérrez

Para promover la práctica basada en evidencias, los ortopedistas preparados requieren leer y criticar la investigación relevante, casi toda investigación reporta resultados descriptivos acerca de la muestra y cifras obtenidas.

Existe una gran variedad de estadísticas descriptivas utilizadas en la presentación de resultados, pero el ortopedista debe saber cuáles estadísticos son útiles para los diferentes tipos de datos que se presentan.

Es especialmente importante conocer el tipo de estadísticos apropiados para las diferentes mediciones realizadas, las estadísticas descriptivas de una sola variable más comunes caen en cuatro categorías: estadísticos de forma de distribución, medidas de tendencia central, medidas de dispersión y medidas de percentiles (Tabla 1).

Estadísticas de forma y normalidad

La distribución de una variable es una forma en que la información acumula-

da se agrupa o dispersa en una gráfica, es el aspecto de esta característica de los números graficados en el nivel de la variable en el eje de las x y el número de casos encontrado en cada punto que constituye el eje y .

De acuerdo con la manera en que se distribuyan los datos de variables cuantitativas podrán presentar mayor sesgo o curtosis, y tendrán o no una distribución gaussiana (Figura 1), estas

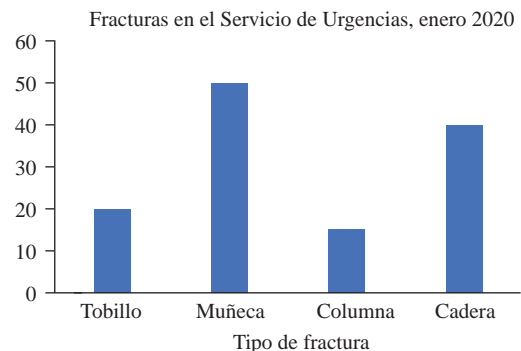


Figura 1: Distribución de una variable nominal de tipos de fractura.

Tabla 1: Estadísticas descriptivas.

Estadísticas de forma y normalidad	Medidas de tendencia central	Medidas de dispersión	Percentiles
Sesgo – simetría de la distribución	Media	Rango	Percentiles
Curtosis (apuntalamiento de la distribución)	Mediana	Varianza	Cuartiles
	Moda	Desviación estándar	Rango intercuartilar

características determinarán el tipo de exámenes estadísticos a utilizar.

Pruebas de sesgo

En estadística descriptiva, el sesgo mide si las dos mitades de la distribución son simétricas, una forma con una cola larga a la derecha se llama sesgada positivamente, lo opuesto, con una cola larga a la izquierda se llama sesgo negativo, cualquier tipo de sesgo, positivo o negativo, indica que la distribución es diferente a una curva normal o gaussiana (Figuras 2 y 3).

Los análisis de tal asimetría en los programas de estadística, como SPSS, requieren el sesgo y el error estándar del mismo, para eso el investigador debe dividir el sesgo por su error estándar.

Una distribución normal se representa por el valor que va de ± 2 . Valores que van de -2 o +2 implican una cualidad sesgada, más que normal (Figura 4).

Pruebas de curtosis

La distribución de los datos tiene muchas formas y la altura o apuntalamiento tiene que ver con las medidas de curtosis, una ordenación que tiene una forma muy delgada de lado a lado y picuda es denominada leptocúrtica, aquella que está plana, se llama platocúrtica, ambas se alejan de la normal,

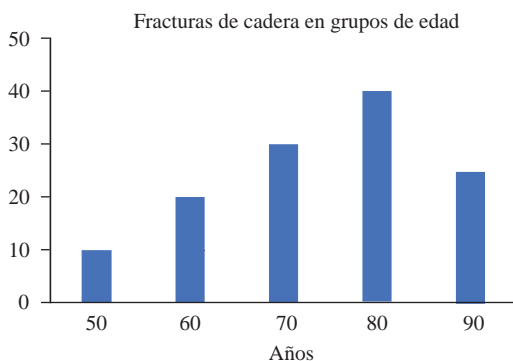


Figura 2: Curva con sesgo negativo, en donde se observa la distribución de fracturas por rango de edad.

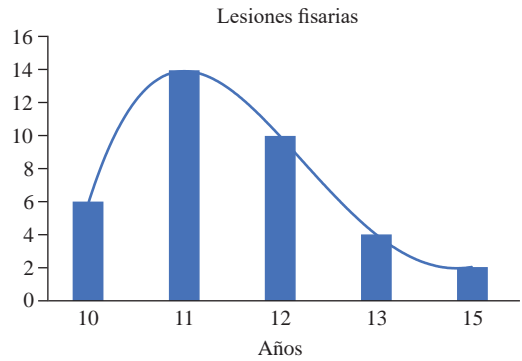


Figura 3: Curva de con sesgo positivo, en donde se observa la distribución de lesiones fisarias por rango de edad.

que es designada mesocúrtica; de igual manera que el sesgo, la curtosis puede ser dividida por su error estándar y cuando se encuentra en ± 2 presenta una distribución normal.

Además de las exploraciones de este tipo, normal o no normal, de una variable dada, existen las medidas de tendencia central como son media, mediana y moda, que se verán en un capítulo aparte.

Medidas de dispersión o de variación

Otro aspecto importante para los investigadores es la consideración de la variabilidad de los resultados entre la medición de los sujetos, la variabilidad describe cómo los casos tienden a distribuirse en el rango posible de una variable estudiada.

Para una cualitativa, la única manera de ver como se distribuyen es determinar la frecuencia absoluta o relativa en los rangos de dicha variable, pero las características matemáticas de una cuantitativa (o dimensional) permiten hacer mediciones precisas de variabilidad, dichas métricas de dispersión incluyen al rango, varianza y desviación estándar.

Rango

Es la distancia desde el valor más bajo al más alto de una variable estudiada, se

calcula restando el valor más bajo de las mediciones, del más alto, esto concede al investigador ver cómo se distribuyen los valores estudiados de los sujetos, sin embargo, no permite ver si esta distribución es uniforme o si se acumula en los extremos de la muestra.

Varianza y desviación estándar

Si de la media se resta cada valor de una variable, se tendrá una indicación de qué tanto cada resultado se desvía de dicha media, si se intenta sumar esos valores, siempre resultará cero, lo que no sirve matemáticamente; si se eleva al cuadrado cada desviación de la media, ya no se obtiene cero como resultado (por ejemplo $-4^2 = 16$). La suma de dichas desviaciones elevadas al cuadrado, y divididas entre el tamaño de la muestra, da un promedio de desviación, llamada varianza.

Una mejor opción es tomar la varianza y aplicarle una raíz cuadrada para eliminar el efecto de la previa elevación al cuadrado y así se obtendrá la desviación estándar, que usualmente se representa con la letra griega delta: « δ ». Esta medida es extremadamente útil en estadística, ya que se expresa en la unidad original de medida, da un estimado confiable del grado en que los números se dispersan de la media, un número de desviación estándar

bajo significa que la mayoría de las mediciones se acerca a la media.

También es útil saber que una desviación estándar por abajo y por arriba de la media concentra al 68% de una muestra y dos desviaciones estándar por abajo y por arriba de la media concentran al 95% de la muestra, así la desviación estándar es una medición de variabilidad que aplica para cualquier distribución, independientemente de la unidad medida.

Percentilas y cuartilas

En ocasiones es útil conocer cómo el valor medido de un sujeto se relaciona con otros sujetos, como cuando se aplica al examen de residencias. Una estadística excelente para este propósito es la percentila.

Una percentila ordena todas las medidas de la más alta a la más baja y calcula el porcentaje de resultados que cae por debajo de un valor individual, por ejemplo, si un aspirante a cursar la residencia en ortopedia cae en la percentila 99, quiere decir que 99% de los aspirantes tuvieron un resultado por debajo de él o ella.

Una cuartila es simplemente la división de todos los resultados en cuatro secciones, los resultados de la percentila 75 a la 99 son la cuartila más alta, los resultados de la percentila 50 a 74 pertenecen a la segunda cuartila más alta.

Los resultados de la percentila 26 a 49 pertenecen a la tercera cuartila más alta y los resultados de la percentila 1 a 25 pertenecen a la cuartila más baja, otro valor de dispersión en este tipo de medidas es contar los valores de la segunda a tercera cuartila y esto es conocido como rango intercuartilar y muchas veces esta información es representada en diagramas de cajas y bigotes.

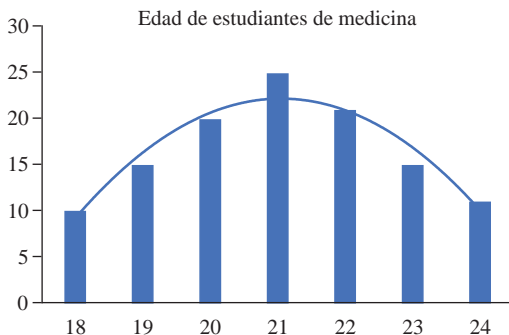


Figura 4: Curva con aparente distribución normal o gaussiana. Describe la cantidad de estudiantes de medicina en segundo año de la carrera de medicina en una universidad de Tamaulipas. Observe su forma mesocúrtica, sin sesgos.

Conclusión

Las estadísticas descriptivas permiten resumir y explorar los datos ob-

tenidos e iniciar las respuestas de las hipótesis planteadas, se pueden utilizar medidas de tendencia central y de dispersión, así como diferentes gráficos de acuerdo a las variables estudiadas.

Lecturas complementarias

1. Petrie A. Medical statistics at a glance. 3rd ed. Chichester: Wiley-Blackwell; 2009.
2. Sprinthall R. Basic statistical analysis. 7th ed. Boston: Allyn and Bacon; 2003.
3. Altman DG. Practical statistics for medical research. London: Chapman & Hall/CRC; 1999. pp. 10-45.

Unidad

1

Principios fundamentales para iniciar con la bioestadística

Bioestadística analítica o inferencial

Gabriela Vázquez Armenta

Una de las ramas de las matemáticas que estudia los resultados del azar es la que se conoce como la «teoría de la probabilidad» y ésta es la base de la denominada inferencia estadística.

El análisis analítico o inferencial se realiza después de haber acabado el análisis descriptivo, donde se evidencia cómo son y cuáles son las características del grupo de la población a estudiar, así como las características particulares de la muestra.

La inferencia estadística es el procedimiento que permite establecer conclusiones teniendo en cuenta los efectos de probabilidad (resultados), y regularmente es en esta etapa donde se aceptan o se rechazan las hipótesis, así como los patrones y la intensidad de la asociación entre las variables, destacando el resultado dentro del concepto: «existen o no diferencias» entre los grupos, esto es: «son iguales o son diferentes».

Cuando se llevan a cabo experimentos, se pueden obtener una serie de probabilidades con resultados posibles y, mediante la probabilidad del evento, se puede expresar la confianza de que el evento ocurra al observar el experimento y esto se puede representar de manera numérica del «0» al «1», de tal forma que el «1» expresa que el evento

ocurrirá con toda seguridad, mientras que el «0» corresponde a un evento que con toda seguridad NO ocurrirá.

Tipos de pruebas

Los procedimientos estadísticos pueden clasificarse en tres grupos:

1. Aquellos que suponen la distribución normal, que en conjunto se reconocen como pruebas paramétricas.
2. Aquellos que no consideran la distribución normal de los datos, que se identifican como procedimientos libres de distribución (binomial y de Poisson).
3. Aquellos que no se refieren a parámetros de población, que se denominan pruebas no paramétricas (χ^2 , prueba exacta de Fisher) (*Figura 1*).

Nota: en algunos textos de estadística, los datos se pueden agrupar, tanto en los procesos no paramétricos como en los libres de distribución, y se presentan con el título de «no paramétricos».

Se cuenta con una gran cantidad de exploraciones que se pueden clasificar en dos tipos: pruebas paramétricas y no paramétricas, las cuales deben

considerarse al iniciar el análisis estadístico, por lo que deberá tomarse en cuenta el tipo de variable, el número de participantes por grupo y el tipo de análisis (medias, medianas o frecuencias) (Tabla 1).

Paramétricas

Las pruebas paramétricas son las que se consideran de mayor potencia estadística y las que proporcionan más información, lo que permite establecer estimaciones de los parámetros de la población a través de muestras estadísticas.

Requisitos:

1. Variables con distribución normal, misma que se puede determinar con la prueba de Kolmogorov-Smirnov y la prueba de Shapiro-Wilk.
2. No se usa para variables ordinales.
3. Se requiere un mínimo de 30 personas por grupo en la muestra, ya que mientras más grande es la muestra, más exacta será la estimación probabilística.
4. La hipótesis se realiza sobre los valores numéricos, especialmente el promedio de la población.

No paramétricas

Las pruebas no paramétricas o de distribución libre son de fácil aplicación y éstas se pueden utilizar para realizar un análisis de variables ordinales o normales, esto cuando no se requiere establecer o plantear una «inferencia»

Tabla 1: Tipos de pruebas estadísticas.

Pruebas paramétricas en cirugía
t de Student ANOVA Correlación de Pearson
Pruebas no paramétricas en cirugía
χ^2 Prueba de Wilcoxon U de Mann-Whitney Kruskal-Wallis Correlación de Spearman McNemar Regresión lineal Probabilidad de Poisson Odds Ratio

sobre cuáles son los parámetros de la población.

Requisitos:

1. Variables de tipo ordinal o nominal.
2. Cuando dos series de observaciones se derivan de distintas poblaciones.
3. La muestra puede ser menor o igual a 20 personas.
4. La hipótesis se realiza sobre los rangos, mediana o frecuencia de los datos.

Es importante conocer estos requisitos y familiarizarse con todos y cada uno de ellos, ya que la selección de la prueba estadística dependerá del tipo de variable, de la presencia o ausencia de normalidad de la distribución de los grupos que se van a analizar y de cómo se formuló la pregunta de investigación (PICOT) (Tabla 2).

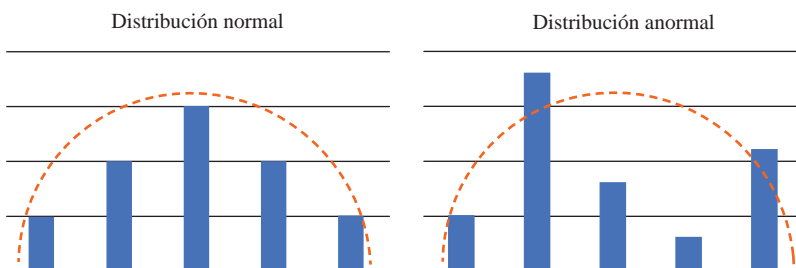


Figura 1: Distribución normal y anormal de los datos.

Tabla 2: Clasificación de acuerdo con el objetivo y la variable.

Tipo de variable	Tipo de muestra	Diferencia		Mostrar relación	Predecir una variable*
		2 grupos	3 grupos	2 variables	Variabes desenlace
Cuantitativa (distribución normal)	NR	t Student**	ANOVA 1 factor	Pearson	Regresión lineal
	R	t Student***	ANOVA 1 factor	Spearman	
Cualitativa ordinal	NR	U Mann-Whitney	Kruskal-Wallis	Spearman	
	R	Wilcoxon	Friedman		
Cualitativa dicotómica	NR	χ^2 (Fisher)	χ^2	Coeficiente phi	Regresión logística
	R	McNemar			

NR = no relacionada, R = relacionada.

* La variable predictiva puede ser cuantitativa, dicotómica u ordinal (transformar a variables tipo *dummy*).

** t de Student para muestras independientes.

*** t de Student para muestras relacionadas.

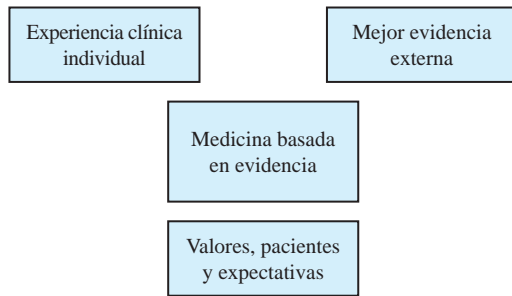


Figura 2: ¿Qué es hacer medicina basada en evidencia?

Interpretación

Estos análisis permiten contar con instrumentos en los que el cirujano, ante los pacientes, pueda pronosticar la sobrevida o las posibilidades de que un procedimiento médico o quirúrgico se llegará a infectar o bien la probabilidad de tener cualquier tipo de complicación; en este caso, el médico basará su estimación en el conocimiento que tenga de lo escrito en la literatura, de las características del paciente así como su expectativa y su evolución, considerando su propia experiencia y los recursos de que dispone para la atención del enfermo y el posible desenlace.

Esto es lo que se conoce como «medicina basada en la evidencia», la cual

se define como el uso consciente, explícito y juicioso de la «mejor evidencia científica» disponible, para tomar decisiones (Figura 2).

Lecturas complementarias

1. Jiménez Avila JM, Falavigna A. Educación en investigación: de la medición a la interpretación-bioestadística quirúrgica. Curitiba: TRAÇO Diferencial; 2018.
2. Siegel S, Castellan NJ. Nonparametric statistics for the behavioral sciences. 2a ed. New York: McGraw-Hill; 1988.
3. Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. JAMA. 1992; 268 (17): 2420-2425.
4. Gómez-Gómez M, Danglot-Banck C, Vega-Franco L. Sinopsis de pruebas estadísticas no paramétricas. Cuando usarlas. Rev Mex Pediatr. 2003; 70 (2): 91-99.
5. Celis de la Rosa A. Bioestadística, en: elementos de probabilidad para la inferencia estadística. México: Manual Moderno; 2004.
6. Daniel WW. Estadística con aplicaciones a las ciencias sociales y a la educación. México: McGraw-Hill; 1988.
7. Gómez-Gómez M, Danglot-Banck C, Velásquez-Jones L. Bases para la revisión crítica de artículos médicos. Rev Mex Pediatr. 2001; 68 (4): 152-159.
8. Pita Fernández S. Tipos de estudios clínico-epidemiológicos. Atención Primaria de la Red. 2001.

Unidad

2

¿Cómo se inicia un análisis estadístico?

Frecuencias y proporciones

Ignacio Méndez Gómez Humarán,
Israel Gutiérrez Mendoza,
Leonardo López Almejo

En todo proceso de investigación, y en particular en temas de ortopedia y traumatología, el proceso de atención médica y consultas permite la recolección de gran cantidad de cifras. Es común que, ante tal cúmulo de información, no exista claridad en cuanto a cómo resumirla e interpretarla y para ello es el análisis estadístico. Inicialmente, el proceso de la información consiste en organizar los datos recabados de los pacientes o de expedientes, identificar sus características y resumirlos por medio de tablas de frecuencias y diversas estadísticas descriptivas. Posteriormente, se pueden utilizar técnicas estadísticas inferenciales que habilitan al investigador para comparar los valores entre variables y así establecer conclusiones y tomar decisiones debidamente sustentadas.

Organización de datos

Cuando se tiene una gran cantidad de números, se requiere establecer una tabla electrónica de datos, comúnmente se puede utilizar Microsoft Excel o algún programa estadístico donde guardar los registros de los pacientes. La estructura de estas tablas normalmente debe contener un renglón para cada paciente y una columna dedicada a cada variable medida.

En la construcción de las tablas de datos y en el análisis de esas cifras, es necesario identificar el tipo de variables que se tienen registradas. Existen básicamente dos tipos de variables, que son:

1. Variables cualitativas o categóricas: son las que representan una clasificación de los sujetos de estudio, por ejemplo, el sexo del paciente, el tipo de lesión, el grado de la lesión, etc.
2. Las cuantitativas o numéricas: son las que se representan por mediciones numéricas, por ejemplo, la edad en años, la estatura en centímetros, el peso en kilogramos, etc.

Las variables cualitativas se pueden codificar por medio de números para facilitar su registro en la computadora, por ejemplo, el grado de severidad de una lesión (leve = 1, moderado = 2,

Tabla 1: Estadística descriptiva de pacientes con condromalacia patelofemoral operados mediante artroscopia de rodilla con anestesia intraarticular.

Grado de condromalacia patelofemoral	Frecuencias simples (pacientes)	Porcentaje
I	5	25
II	11	55
III	4	20
Total	20	100

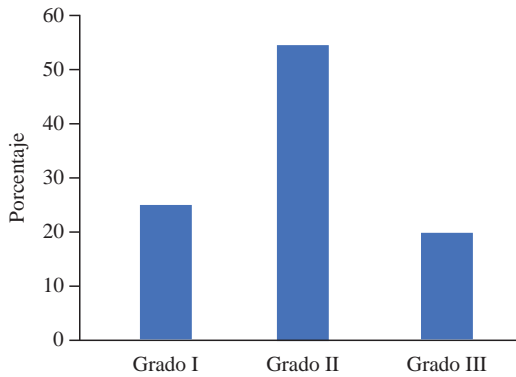


Figura 1: Porcentaje de pacientes con condromalacia patelofemoral operados mediante artroscopia de rodilla con anestesia intraarticular.

severo = 3). Las de tipo cuantitativo se pueden categorizar y convertirlas a cualitativas cuando se agrupan los datos observados en grupos o clases, por ejemplo, la edad de adultos mayores en categorías según el número de años cumplidos (60 a 64, 65 a 69, 70 a 74, 75 a 79 años y 80 y más).

Frecuencias y proporciones

El análisis inicial para las variables medidas en un estudio es **la estadística descriptiva**, para variables cualitativas o cuantitativas categorizadas, se recomienda la exploración mediante las frecuencias simples y porcentajes de las obtenidas en la práctica; una manera de representar el conjunto es mediante una **tabla de frecuencias**. Estas caracterizaciones constan de una tabla con tres columnas; la primera de ellas contiene el nombre de las variables medidas; la segunda contiene el número de casos que corresponden a cada una de las ellas, llamadas **frecuencias simples**; y la tercera contiene las proporciones de casos de cada categoría con respecto al total, llamadas **proporciones o porcentajes**, obtenidas de dividir el número de casos observados de una de las subvariables del estudio entre el número total de casos, obtenidos de la sumatoria de todas las subvariables y multiplicando

el resultado por 100; también se puede obtener este análisis mediante cualquier programa estadístico computacional. Ejemplo de un estudio donde el total de una población con fractura de muñeca es de 127 pacientes, de ellos 100 son hombres y 27 mujeres; para obtener el porcentaje o proporción de mujeres con fractura de muñeca se divide la cantidad de mujeres (27) entre la población total (127) y se multiplica por 100, obteniendo que las mujeres representan 21.2% de la población del estudio.

En la siguiente **Tabla 1** se presentan las frecuencias simples y porcentajes para la variable cualitativa llamada grados de condromalacia patelofemoral, encontradas en pacientes con síndrome de hiperpresión lateral patelofemoral en el Servicio de Rodilla Ortopédica de la UMAE Magdalena de las Salinas durante un estudio de anestesia intraarticular en artroscopia de rodilla.

Una manera alterna de representarlo es mediante una gráfica de barras (**Figura 1**) e histogramas de frecuencias, entre otras.

Las frecuencias relativas o proporciones son muy útiles cuando se

Tabla 2: Estadística descriptiva de pacientes con diferentes concentraciones de colesterol sérico de acuerdo con su grupo etario expresado en frecuencias relativas.

Colesterol sérico mg/100 ml	Frecuencias relativas		
	Grupo de edad (años)		Total
	(25-34)	(55-64)	
80-119	13	5	18
120-159	150	48	198
160-199	442	265	707
200-239	299	458	757
240-279	115	281	396
280-319	34	128	162
320-359	9	35	44
360-399	5	7	12
Total	1,067	1,227	2,294

Tabla 3: Estadística descriptiva de pacientes con diferentes concentraciones de colesterol sérico de acuerdo con su grupo etario expresado en porcentajes.

Colesterol sérico mg/100 ml	Distribución porcentual		Total
	Grupo de edad (años)		
	(25-34)	(55-64)	
80-119	1.2	0.4	0.8
120-159	14.1	3.9	8.6
160-199	41.4	21.6	30.8
200-239	28.0	37.3	33.0
240-279	10.8	22.9	17.3
280-319	3.2	10.4	7.1
320-359	0.8	2.9	1.9
360-399	0.5	0.6	0.5
Total	100.0	100.0	100.0

quiere comparar distribuciones de frecuencias con distintos tamaños de muestra. En la siguiente *Tabla 2* se muestran las frecuencias observadas de una muestra de 2,294 hombres adultos, clasificados según ocho grupos de concentración de colesterol sérico y dos grupos de edad.

Con el fin de comparar las distribuciones porcentuales según los grupos de concentración de colesterol sérico, se calcularon las frecuencias relativas expresadas como porcentajes para cada grupo de edad (tabla de porcentajes por columna). En la siguiente *Tabla 3* se presentan las distribuciones porcentuales por grupo de edad y la total, esta última conocida como marginal.

Se puede observar que, para el grupo de 25 a 34 años, el grupo de colesterol sérico de 160-199 mg/100 ml presenta la mayor proporción con 41.4%, mientras que la mayor proporción en el grupo de 55 a 64 años está en el grupo de 200-239, con 37.3%. Nótese que si se ignora el grupo de edad (distribución marginal), ambos grupos de colesterol sérico tienen 30.8 y 33.0% respectivamente. Esto posibilita reconocer que la distribución de la concentración de colesterol sérico está

recorrida a grupos de mayor concentración cuando la edad es mayor. En la siguiente *Figura 2*, se presenta gráficamente la distribución porcentual de los hombres adultos según grupo de edad y concentración de colesterol sérico.

Esta gráfica permite observar claramente el desplazamiento de la concentración de colesterol sérico a niveles mayores en hombres de edades de 55 a 64 años.

Las proporciones o porcentajes acumulados también se utilizan para estudiar y comparar las distribuciones de frecuencias de casos observados en la muestra. En la siguiente *Tabla 4* se despliegan las distribuciones porcentuales de colesterol sérico para los grupos de edad estudiados.

Como se puede observar, 56.7% de los hombres en edades de 25 a 34 años registran niveles de colesterol sérico inferiores a 200 mg/100 ml, mientras que sólo 25.9% de los hombres de 55 a 64 años de edad están por debajo de 200 mg/100 ml. Si se utiliza el complemento de cada porcentaje, se puede decir que el (100-56.7) 43.3% de los hombres de 25 a 34 años registran niveles de colesterol de 200 mg/100 ml o mayores, mientras que el (100-25.9)

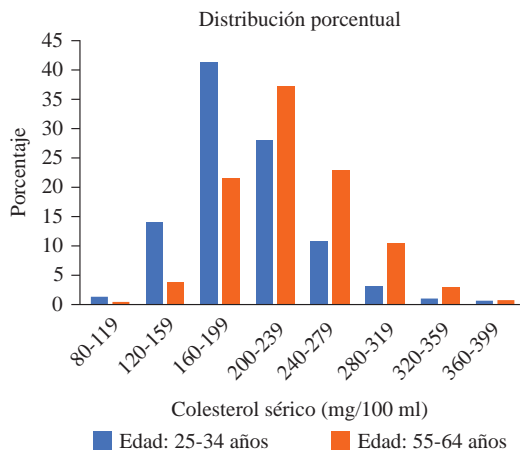


Figura 2: Distribución porcentual de pacientes con diferentes niveles de colesterol sérico de acuerdo con dos grupos etarios.

Tabla 4: Distribución porcentual acumulada de pacientes con diferentes concentraciones de colesterol sérico de acuerdo con su grupo etario.

Colesterol sérico mg/100 ml	Distribución porcentual acumulada		
	Grupo de edad (años)		Total
	(25-34)	(55-64)	
80-119	1.2	0.4	0.8
120-159	15.3	4.3	9.4
160-199	56.7	25.9	40.2
200-239	84.7	63.2	73.2
240-279	95.5	86.1	90.5
280-319	98.7	96.6	97.6
320-359	99.5	99.4	99.5
360-399	100.0	100.0	100.0

74.1% de los hombres de 55 a 64 años registran niveles de colesterol sérico por arriba de 200 mg/100 ml.

Conclusión

Dentro de las cifras de resumen de datos obtenidas de la medición de las variables de un estudio, el análisis

estadístico descriptivo para las cualitativas o cuantitativas categorizadas está representado por las frecuencias y proporciones que han sido explicados en este capítulo, las medidas de tendencia central y dispersión para las variables cuantitativas es el siguiente paso de dicho análisis.

Lecturas complementarias

1. Gutiérrez-Mendoza I, Pérez-Correa JJ, Serna-Vela F, Góngora-Ortega J, Vilchis-Huerta V, Pérez-Guzmán C, et al. Efecto analgésico del fentanilo intraarticular en artroscopía de rodilla para el síndrome de hiperpresión lateral patelofemoral. *Acta Ortop Mex.* 2009; 23 (6): 336-341.
2. Daniel W. *Bioestadística: base para el análisis de las ciencias de la salud.* 4a edición. México: Editorial Limusa; 2010.
3. Dawson-Saunders B, Trapp GR. *Bioestadística médica.* México: Editorial Manual Moderno; 1993.
4. Walpole RE, Myers RH, Myers SL, Ye K. *Probabilidad y estadística para ingeniería y ciencias.* 8a edición. México: Pearson Prentice Hall; 2007.
5. Pagano M, Gauvreau K. *Principles of biostatistics.* 2nd edition. Pacific Grove, CA, USA: Duxbury Press; 2000. p. 525.

Unidad

2

¿Cómo se inicia un análisis estadístico?

Medidas de tendencia central

Romina L Montes de Oca Dávila,
Aldo Fidel Izaguirre Hernández

Las medidas de tendencia central ayudan al investigador a determinar dónde se ubican más frecuentemente los casos estudiados, son las medidas estadísticas que identifican un valor único como representativo de una distribución entera; existen diferentes niveles de medición que requieren distintas medidas de tendencia central, también es importante notar cómo tal forma de ordenación afecta a esta estadística descriptiva.

1. Describen qué tan cerca una población de datos o un grupo de valores se encuentran de un punto central numérico.
2. Son elementos básicos para describir el comportamiento de un fenómeno bajo estudio; herramientas esenciales para iniciar la comprobación de hipótesis.
3. Las tres medidas centrales más comunes son: media (o promedio), mediana, moda y son usadas como punto de partida en la estadística descriptiva e inferencial.

Media

Describe un grupo completo de observaciones o mediciones, con un único valor representando el centro aritmético de los números, para calcular la media se suman todas las observa-

ciones y se dividen entre el total de las observaciones de una población (N) o muestra (n).

Mediana

Describe el punto medio de un rango de datos ordenados ascendente o descendentemente.

1	2	3	21	35
---	---	---	----	----

Si el número de elementos de un grupo de observaciones son impares entonces la mediana es simplemente el valor de la posición de en el medio del arreglo de datos.

1	2	3	21	35	42
---	---	---	----	----	----

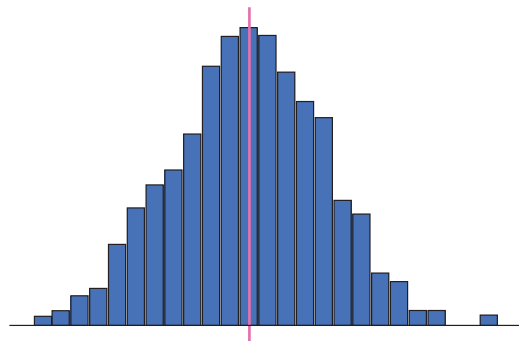


Figura 1: El valor de media y mediana son iguales siempre y cuando el grupo de datos obedezca una **distribución normal**.

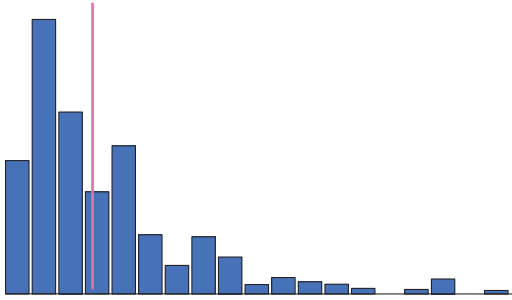


Figura 2: El valor de media y mediana son diferentes cuando el grupo de datos obedezca una **distribución anormal**.

En un grupo de elementos donde existe un número par de observaciones, entonces se deberá tomar los dos valores de en medio y posteriormente calcular la media (promedio aritmético). En el ejemplo de arriba serían los valores 3 y 21, la media será 12.

Media vs mediana

Tras compararla con la media, la mediana no considera los valores extremos de la población o muestra, por lo tanto, frecuentemente la media es sólo usada de manera informativa y no conclusiva para datos con distribución sesgada, por ejemplo, la media podría no ser una buena medida al describir las edades de los jefes de servicio, adscritos y residentes de un hospital.

Es relevante mencionar que cuando la media = mediana se considera una disposición normal y que obedece a una distribución de Gauss.

En una ordenación simétrica, la media, mediana y moda son idénticas, en la variante asimétrica a la derecha, la moda se localiza en el pico, la media se encuentra desplazada a la derecha y la mediana se encuentra entre las dos medidas, en las distribuciones sesgadas a la izquierda, ocurre lo mismo, pero en espejo.

Los datos simétricos (*Figura 1*) (línea rosa) muestran una distribución normal, lo cual tiene una buena represen-

tación del centro de las observaciones en contraste, en la *Figura 2* se tiene la concentración de cifras fuertemente a la izquierda representando un sesgo de observaciones, por lo cual la mediana no coincide en posición con la media de esta muestra.

Moda

Describe el número de repeticiones de una observación, es decir, representa un patrón en un grupo de datos, una distribución puede ser unimodal, bimodal, trimodal, etc. La moda puede ser usada con la media y mediana para dar una caracterización general de una ordenación de cifras; mientras las otras dos medidas de tendencia central requieren ser calculadas, la moda puede encontrarse simplemente contando el número de veces que una observación ocurre en una disposición de datos de un estudio.

Distribución unimodal: en este caso la observación 8 ocurre en el mismo valor 9 veces, sobre toda la población de datos (*Figura 3*).

Bimodal: en este caso existen dos modas, 4 y 16, una distribución con más de una moda puede indicar que las mediciones de las observaciones describen dos fenómenos diferentes y éstos fueron tratados como un solo comportamiento, los que significa que los datos fueron incluidos en el estudio de esta forma (*Figura 4*).

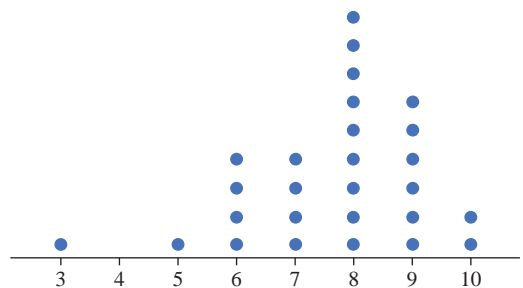


Figura 3: Representación de población unimodal de datos.

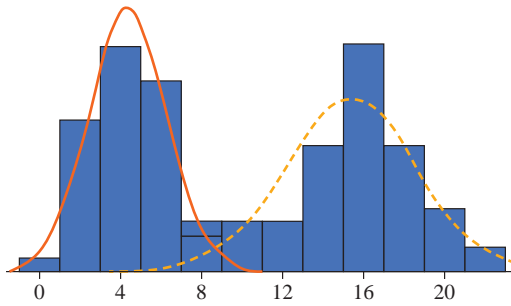


Figura 4: Representación de población bimodal de datos.

Conclusiones

Para realizar una exploración analítica de los datos precisa, se requiere la habilidad de interpretar la estadística descriptiva básica, cualquier error o mala interpretación, en este nivel de análisis, puede llevar a resultados contradictorios y contribuir a una confusión estadística.

El uso de la medida de tendencia central utilizada debe ser la correcta para demostrar que realmente se representen los números, es de especial importancia que no haya errores, sobre todo si se implica comparación entre grupos, en donde es aún más importante que las medidas de tendencia central representen a las muestras obtenidas.

Lecturas complementarias

1. Petrie A. Medical statistics at a glance. 3rd ed. Chichester: Wiley-Blackwell; 2009.
2. Sprinthall R. Basic statistical analysis. 7th ed. Boston: Allyn and Bacon; 2003.
3. Altman DG. Practical statistics for medical research. London: Chapman & Hall/CRC; 1999. pp.10-45.
4. Gonzales VA, Ottenbacher KJ. Measures of central tendency in rehabilitation research: what do they mean? Am J Phys Med Rehabil. 2001; 80 (2): 141-146.

Unidad

2

¿Cómo se inicia un análisis estadístico?

Medidas de dispersión: rango y varianza

José Alfredo Penagos Paniagua

Estas medidas son necesarias para la mejor comprensión de la distribución de un conjunto de observaciones realizadas en un estudio estadístico y se complementan con las medidas de tendencia central, toda vez que proporcionan conjuntamente una descripción numérica más completa de los datos.

Hay que recordar que las medidas de tendencia central localizan generalmente a un valor que se halla ubicado en el centro de la distribución, pero no informa sobre el grado de dispersión o variabilidad del conjunto de números. Las medidas de dispersión proporcionan información sobre la posición relativa de otros puntos de datos en la muestra con respecto a la medida de tendencia central que se utiliza, es decir que son índices que dan una idea sobre la homogeneidad o heterogeneidad de los datos.

Las medidas de dispersión comúnmente utilizadas son rango, desviación estándar (DE), desviación estándar geométrica y varianza, en lo que compete a este capítulo sólo se hablará del rango y la varianza.

Rango

Es una medida de la variabilidad porque representa sólo la distancia entre

el extremo superior o valor más alto y el y el inferior o valor más bajo de la distribución de datos.

$$R = \text{valor máximo} - \text{valor mínimo} = X_n - X_1.$$

Su ventaja radica en la facilidad de cálculo, expresión y aplicación, por otro lado, tiene muchas desventajas, su uso puede estar influenciado por valores atípicos o valores extremos en algunas lecturas, lo que puede conducir a interpretaciones erróneas, ya que no utiliza todas las observaciones en un conjunto de datos, su utilización está más ligada al control estadístico de procesos y no es muy utilizada como medida de dispersión.

Ejemplo:

Determinar el rango de las siguientes edades de los pacientes que se enumeran a continuación:

32 21 38 22 77 49 50 19 28 43 56 20.

Valor máximo = 56,
valor mínimo = 19,
rango = $56 - 19 = 37$.

Este valor facilita también comparar con estudios de la misma naturaleza y establecer con buen criterio que el

conjunto de datos que tenga la menor amplitud de variación, será el que tenga menos variabilidad o menor dispersión.

Varianza

Es la variación o dispersión cuadrática de una distribución de datos, es decir, el cuadrado de la desviación estándar, se considera una medida absoluta de la variación, porque sólo tiene significado cuando va acompañada de las dimensiones, al cuadrado, de la variable que se analiza.

Se define como el promedio corregido de las distancias cuadradas de cada valor de la variable en la distribución, con respecto a su media aritmética. Matemáticamente se define, en su forma más sencilla, como:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

Donde el término, $n - 1$ se denomina grados de libertad, que representa el número de observaciones realmente aleatorias, que intervienen en el cálculo de esta medida.

Si en la distribución hay valores repetidos, es conveniente introducir el término de frecuencia en el cálculo de la varianza:

$$s^2 = \frac{\sum_{i=1}^k f_i (x_i - \bar{x})^2}{n - 1}$$

f_i , representa el número de veces que cada distancia cuadrática se repite y k es el número de categorías diferentes que se presentan en la distribución.

Ejemplo:

Determinar la varianza en el peso corporal de los siguientes pacientes:

Paciente	1	2	3	4	5
Peso (kg)	28	26	27.30	25.75	27.50

Con la media aritmética calculada previamente $x = 26.91$ se puede obtener la varianza como sigue:

$$s^2 = \frac{(28-26.91)^2 + (26-26.91)^2 + (27.30-26.91)^2 + (25.75-26.91)^2 + (27.50-26.91)^2}{5-1}$$

$$s^2 = 0.9655 \text{ (kg)}^2$$

Conclusión

Las medidas de dispersión indican, por un lado, el grado de variabilidad que hay en la muestra, por el otro lado, indican la representación del valor central, ya que, si obtiene un valor pequeño, significa que los valores se concentran en ese centro, esto significa que hay poca variabilidad en los datos y el centro representa bien a todos. En cambio, si obtiene un valor grande, significa que los valores no están concentrados sino dispersos, esto manifiesta que hay mucha variabilidad y el centro no será muy representativo; por otro lado, a la hora de hacer inferencias se necesitará un tamaño de muestra más grande si se quiere reducir el error, aumentará considerablemente por el incremento de la variabilidad.

Lecturas complementarias

1. Salazar Pinto C, Del Castillo Galarza S. Análisis y descripción de los datos. En: Salazar Pinto C, Del Castillo Galarza S. Fundamentos básicos de estadística. Quito; 2018. pp. 67-71.
2. Deshpande S, Gogtay NJ, Thatte UM. Measures of central tendency and dispersion. J Assoc Physicians India. 2016; 64 (7): 64-66.
3. Guerra-Dávila T. Estadística descriptiva. En: Guerra-Dávila T. Bioestadística. México, D.F.: UNAM-FEZ Zaragoza; 2014. pp. 86-88.
4. Manikandan S. Measures of dispersion. J Pharmacol Pharmacother. 2011; 2 (4): 315-316.
5. Mishra P, Pandey CM, Singh U, Gupta A, Sahu C, Keshri A. Descriptive statistics and normality tests for statistical data. Ann Card Anaesth. 2019; 22 (1): 67-72.

Pruebas de normalidad: Kolmogorov-Smirnov y Shapiro-Wilk

José María Jiménez Ávila

Cuando se observan los datos es importante considerar si existe una distribución normal, con el objeto de calcular la probabilidad de que una variable aleatoria subyacente al conjunto de cifras se encuentre distribuida normalmente. Tal forma normal es una de las más importantes distribuciones de probabilidades de la estadística, conocida también como de Gauss o Gaussiana.

La normalidad de la distribución de la muestra es de utilidad para discernir entre el uso de la media, mediana, desviación estándar o rango para el análisis estadístico descriptivo y el uso de pruebas paramétricas o no paramétricas en el análisis estadístico inferencial.

Para poder valorar si una muestra se distribuye de forma normal, se deben tomar en cuenta los siguientes datos:

1. Presencia o ausencia de simetría en la curva de distribución.
2. Tipo de curtosis presente en la curva de distribución.
3. Resultado de la prueba de Kolmogorov-Smirnov o de Shapiro-Wilk.

Se habla de asimetría en las ocasiones en que la muestra se distribu-

ye de forma anormal en relación con la media aritmética o promedio. Existen dos tipos de asimetría (*Figura 1*):

1. Asimetría positiva o derecha: cuando la minoría de la muestra se encuentra a la derecha de la media aritmética.
2. Asimetría negativa o izquierda: cuando la minoría de la muestra se encuentra a la izquierda de la media aritmética.

Una asimetría igual a cero corresponde a una curva simétrica o normal, sin embargo, no siempre el valor de la asimetría será perfecto, por lo tanto, se acepta un margen de error de ± 0.5 , ya sea positivo o negativo, por otro lado, al obtener una asimetría mayor a cero equivale a una curva con asimetría positiva y menor a cero a una curva con asimetría negativa.

Se entiende como curtosis al grado de concentración de la muestra en relación con la media aritmética o promedio. Existen tres tipos de curtosis (*Figura 2*):

1. Concentración mesocúrtica o normal: se ve al obtenerse una curtosis

- igual a cero, con un margen de error de ± 0.5 ya sea positiva o negativa.
- 2. Concentración leptocúrtica o positiva: al obtener una curtosis mayor a cero.
- 3. Concentración platicúrtica o negativa: al obtener una curtosis menor a cero.

La prueba de Kolmogorov-Smirnov o K-S es la más utilizada y la más potente para evaluar la normalidad de distribución de tamaños de muestra mayores de 30 ($n \geq 30$).

Cuando el tamaño de la muestra es menor a 30 ($n \leq 30$), el análisis de Shapiro-Wilk es el indicado para el análisis de la normalidad de la distribución.

El valor de la significancia que se obtenga en las pruebas de normalidad determinará la presencia o ausencia de normalidad:

- 1. Una significancia > 0.05 indica una muestra con distribución normal y por lo tanto el uso de media y desviación estándar para su análisis descriptivo y se sugiere el uso de

pruebas paramétricas en su análisis inferencial.

- 2. Una significancia ≤ 0.05 indica una muestra con ausencia de normalidad en ese ordenamiento, y por lo tanto indica el uso de mediana y rango en su análisis descriptivo y se sugiere el uso de pruebas no paramétricas para su análisis inferencial.

En relación con la distribución de la muestra se aceptará o rechazará la hipótesis:

- 1. Hipótesis nula (H_0): establece que la muestra sigue una distribución normal.
- 2. Hipótesis alterna (H_1): establece que la muestra no sigue una distribución normal.

Por lo tanto, de acuerdo con la distribución de la muestra, será el tipo de análisis estadístico descriptivo e inferencial a realizar (*Figura 3*).

Cuando el tamaño de la muestra es menor a 50, existen programas

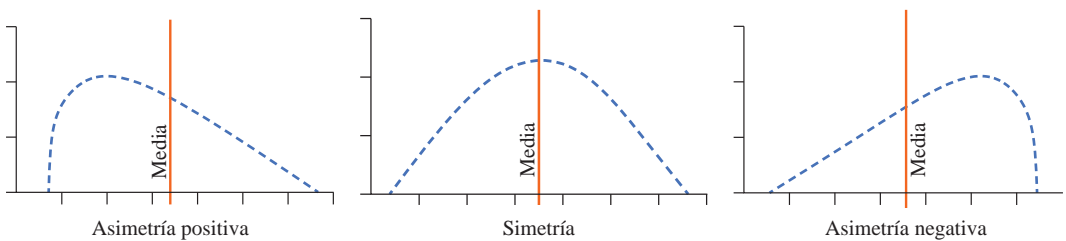


Figura 1: Tipos de distribución.

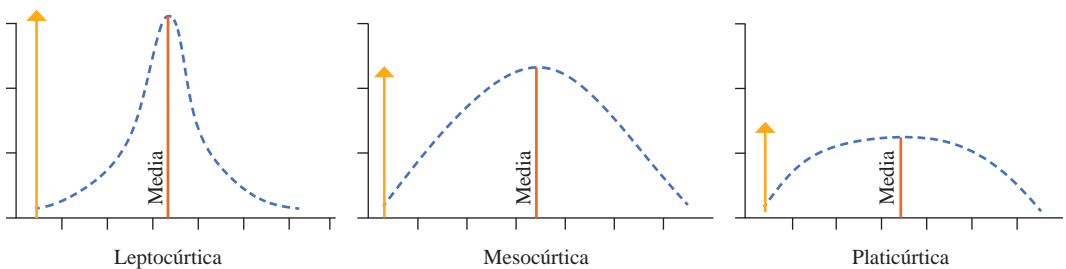


Figura 2: Tipos de asimetría.

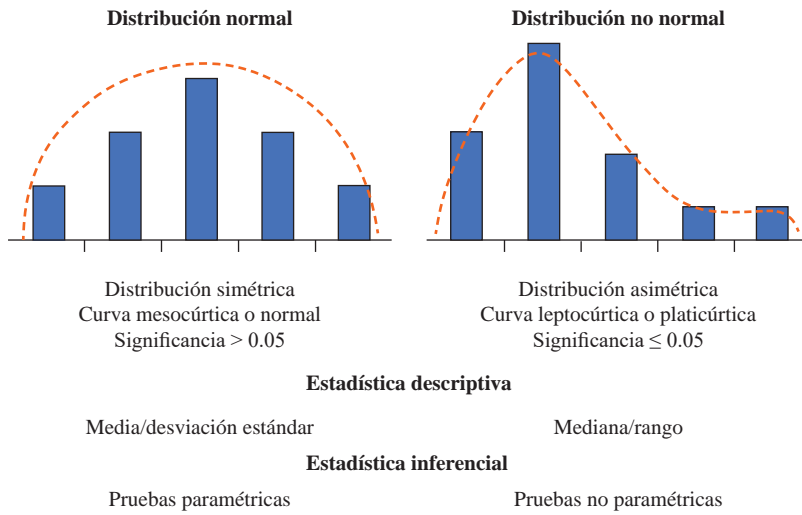


Figura 3:

Tipo de distribución: simétrica y asimétrica.

computacionales como el SPSS que calcula automáticamente ambas pruebas.

Interpretación de los resultados

Con la meta de valorar la presencia o ausencia de normalidad de la distribución de las variables y discernir el tipo de prueba estadística a utilizar, deber tomarse en cuenta el valor de la asimetría, de la curtosis y de la significancia de las variables analizadas.

Conviene recordar que una asimetría igual a cero, con un margen de error de ± 0.5 positivo o negativo, corresponde a una curva simétrica o normal, una asimetría mayor a cero equivale a una curva con asimetría positiva y menor a cero a una curva con asimetría negativa.

Una curtosis igual a cero, con un margen de error de ± 0.5 positivo o negativo corresponde a una concentración mesocúrtica o normal, una curtosis mayor a cero a una concentración leptocúrtica o positiva y una curtosis menor a cero a una concentración platicúrtica o negativa.

Una significancia con valor de $p > 0.05$ indica una muestra con distribución normal, por lo tanto, el uso de media y desviación estándar para su análisis descriptivo y el uso de pruebas paramétricas en su análisis inferencial, mientras que un valor de $p \leq 0.05$ indica ausencia de normalidad en el comportamiento de la muestra, por lo tanto, el uso de mediana y rango en su análisis descriptivo y el uso de sondeos no paramétricos para su análisis inferencial.

Lecturas complementarias

1. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). *Biometrika*. 1965; 52 (3-4): 591-611.
2. López MM, Castelo Branco VTF, Soares JB. Utilización de test estadísticos de Kolmogorov-Smirnov y Shapiro Wilk para verificación de la normalidad. *Transportes*. 2013; 21 (1): 59-66.
3. Talavera JO, Rivas-Ruiz R. Investigación clínica IV. Pertinencia de la prueba estadística. *Rev Med Inst Mex Seguro Soc*. 2011; 49 (4): 401-405.
4. Gómez-Gómez M, Danglot-Banck C, Vega-Franco L. Sinopsis de pruebas estadísticas no paramétricas. Cuándo usarlas. *Rev Mex Pediatr*. 2003; 70 (2): 91-99.

Pruebas de hipótesis (valor de «p») e intervalo de confianza

José Manuel Pérez Atanasio

Cuando se presentan los resultados de una investigación, habitualmente se quiere demostrar que la hipótesis lanzada es cierta; que el implante, la técnica quirúrgica o algún procedimiento es superior en comparación a otros.

Para que tal aseveración se fundamente en pruebas estadísticas se utiliza el valor de «p» y los intervalos de confianza.

Sin embargo, no siempre existe un conocimiento claro del significado de estos conceptos y habitualmente sólo son utilizados para demostrar que se hizo un análisis matemático o para sorprender a un colega menos adiestrado en bioestadística, por lo que es frecuente que en la presentación de trabajos de investigación se escuche «es estadísticamente significativo» o que el expositor haga énfasis en valores infinitamente pequeños de «p» asumiendo que esto le otorga validez científica.

El objetivo de este capítulo es exponer de manera clara qué son y cómo se interpretan estos valores.

Descripción de la prueba

Pruebas de hipótesis

Al realizar un estudio comparativo entre dos procedimientos siempre se

tiene un favorito, cierto candidato; por ejemplo, al comparar un clavo contra una placa, para el tratamiento de las fracturas diafisarias de fémur, la efectividad del clavo es superior a la placa; los dos elementos que se van a comparar se llaman variables de causa o independientes y la observación que se espera se considera la variable de resultado o dependiente, la cual representa el cómo se va a medir que es superior uno de otro, siguiendo con el argumento, el tiempo de consolidación de la fractura puede ser un buen parámetro de resultado.

Una vez que se tienen definidas las variables de causa y la de resultado, se debe plantear esta relación en una hipótesis, es decir, en un posible desenlace.

En estadística existen dos tipos de hipótesis: la nula y la alterna.

Estos conceptos suelen ser confusos por las distintas formas de explicarlos en los libros de bioestadística o por las formas matemáticas de expresarlas, o incluso por las traducciones que se hagan del idioma original de redacción.

Para facilitar su diferenciación, lo primero que se tiene que hacer en un protocolo es definir claramente qué es lo que se cree, a qué se le apuesta, en este ejemplo se piensa que el cla-

vo es superior a la placa, por lo que el tiempo de consolidación será menor, también es importante proponer una magnitud, es decir, que tan menor será el tiempo de consolidación.

Ahora se debe plantear en una oración afirmativa: el tiempo de consolidación con el clavo será menor a seis meses. Esta afirmación que representa lo que se defiende se llamará hipótesis alternativa. La hipótesis nula será entonces la negación de la primera.

Hipótesis nula: el tiempo de consolidación con el clavo será igual o mayor a seis meses.

Hipótesis alterna: el tiempo de consolidación con el clavo será menor a seis meses.

La necesidad de tener dos hipótesis es para poder aplicarles pruebas estadísticas y entonces afirmar con el análisis de los datos si se acepta o no la hipótesis planteada.

El valor de «p»

Este valor representa la probabilidad de que dichos resultados sean debidos al azar o a otras variables no estudiadas.

En estudios clínicos este valor se considera como significativo cuando es menor al 5%, que expresado en valores del 0 al 1 es de 0.05.

No en todos los estudios este nivel de corte de significancia es de 5%, por ejemplo, en estudios de genética o de ingenierías moleculares se esperan valores menores al 1%, por el grado de precisión con el que se realizan las pruebas; sin embargo, en los estudios clínicos es muy complicado controlar todas las variables que pueden influir en el resultado por la condición propia de los seres humanos.

Con el objetivo de que el lector pueda comprender mejor el significado del valor de «p», se presentarán dos resultados hipotéticos del caso anterior con su interpretación.

Ejemplos:

1. Si se utiliza un clavo centromedular, el tiempo de consolidación de las fracturas diafisarias de fémur es menor en comparación con el lapso obtenido si se utiliza una placa. Resultado en las pruebas estadísticas con un valor de $p < 0.05$.
2. Al utilizar un clavo centromedular, el tiempo de consolidación de las fracturas diafisarias de fémur es inferior en comparación con el lapso observado si se utiliza una placa. Resultado en las pruebas estadísticas con un valor de $p > 0.05$.

Nótese que los dos resultados sólo difieren en el valor de «p» obtenido en los análisis estadísticos, en el primer ejercicio es menor a 0.05 y en el segundo es mayor a 0.05.

Para su interpretación se debe asumir, en el ejemplo 1, que el azar u otras variables no estudiadas intervinieron en menos de 5%, por lo que se puede considerar que efectivamente el clavo es superior a la placa, por otra parte, en el ejemplo 2 el azar está presente en más de 5% de los resultados, por lo que no es posible afirmar la superioridad del clavo debido a que muy probablemente otras variables no estudiadas influyeron en el resultado.

No existe un consenso claro de cómo reportar el valor de «p» obtenido, se puede expresar con la cifra exacta mostrada en el software de análisis estadístico o simplemente como menor o mayor a 0.05.

En estudios experimentales de ciencias básicas es recomendable reportar el valor exacto, y en estudios clínicos como menor o mayor a 0.05.

Intervalo de confianza

El intervalo de confianza es una aproximación matemática de los resultados que se obtendrían si se repitiera ese experimento muchas veces.

Este intervalo puede ser definido por el investigador, pero de manera habitual es de 95%, lo que se puede interpretar como que si se repitiera tal experimento 100 veces, en 95% de las ocasiones el resultado estaría entre los valores de los intervalos de confianza.

El intervalo de confianza puede ser amplio o estrecho, es decir cercano o muy disperso con respecto al valor obtenido, tiene mayor fortaleza estadística tener un intervalo estrecho, debido a que habitualmente expresa que el resultado se realizó en una muestra lo suficientemente grande para ser representativa del fenómeno, en la otra cara de la moneda, si se tiene un intervalo muy amplio significa que hizo falta analizar a una mayor población y que aquellos resultados deben ser tomados con cautela al tratar de extrapolarlos hacia una población general.

El intervalo de confianza es muy útil cuando se expresan los resultados como factores de riesgo mediante el *odds ratio* o el riesgo relativo. Por ejemplo, cuando se tiene un resultado de *odds ratio* de 5 y su intervalo de confianza al 95% es de 5-7 esto se puede entender como que el factor de riesgo es de cinco veces y si el experimento se repitiera 100 veces el riesgo estaría entre cinco y siete veces.

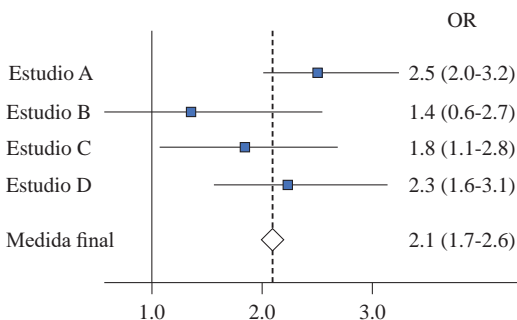


Figura 1: Ejemplo de una gráfica tipo *forest plot*. Se muestran los resultados de cuatro estudios con sus valores de *odds ratio* (OR) y sus intervalos de confianza, el rombo representa al valor integrador de los cuatro estudios.

En los estudios de revisión sistemática con metaanálisis, se realiza una prueba para estudiar los intervalos de confianza reportados en distintos estudios que habitualmente difieren uno de otro y se trata de conformar un solo resultado de mayor precisión, el resultado del análisis se muestra en una gráfica conocida como *forest plot*, en la que se presentan los distintos resultados obtenidos con sus intervalos de confianza con un rombo al final que expresa el conjunto de todos los estudios. En la *Figura 7* se muestra un modelo hipotético de un metaanálisis, en el que se exploraron cuatro estudios con distintos valores de *odds ratio* y se realizó una medición final que representa el valor integrando los cuatro estudios.

Conclusión

El valor de «p» representa la probabilidad de que el desenlace obtenido se deba al azar. El intervalo de confianza es un estimado de cómo se comportaría el resultado si se repitiera varias veces el estudio.

Lecturas complementarias

1. Falavigna A, Jiménez Avila JM. Education in research: from the idea to the publication. Caxias do Sul, Brazil: Educus; 2015.
2. Jiménez Avila JM, Falavigna A. Education in research: from measurement to interpretation - surgical biostatistics. Educus AO Spine Latin America; 2018.
3. Toledo-Infanson V, Jiménez Avila JM. Publicación con impacto en ortopedia y traumatología. México: Graphimed S.A. de C.V.; 2019.
4. Pérez AJM. Principios de bioestadística para comprender los resultados reportados en los artículos científicos. Ortho-tips. 2015; 11 (2): 91-95.
5. Higgins JPT, Green S (editors). Cochrane handbook for systematic reviews of interventions. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. 2011.
6. O'Brien SF, Yi QL. How do I interpret a confidence interval? Transfusion. 2016; 56 (7): 1680-1683.

Unidad

2

¿Cómo se inicia un análisis estadístico?

¿Cómo se presentan los datos?

Arelhi Catalina González Cisneros

Una de las partes más importantes del manuscrito y más utilizadas en la estadística, es la presentación de los datos. Al tener diferentes tipos de gráficos, se logra evidenciar los resultados de forma rápida y efectiva.

Elementos gráficos de gran utilidad son los cuadros o las figuras las cuales, al ser correctamente realizadas, además de presentar la mayor cantidad de ideas, demuestran orden y economizan espacio. Los gráficos deben tener el mismo formato, enumerarlos consecutivamente, agregar el pie de foto, cuadro o figura, con su título y descripción correspondiente, de manera breve y precisa.

Una recomendación importante es no olvidar el objetivo de su investigación. Los gráficos deben mostrar los factores que por sí solos reflejen el resultado del mensaje que se desea dar a la audiencia, de una manera fácil,

comprensible y que permita al lector vislumbrar la relación que guardan estos datos entre sí. Se recomienda incluir: título, cuerpo de figura y la descripción, así como la fuente de donde se obtuvieron los datos.

Cada revista de investigación posee sus normas dirigidas a los autores, en las cuales se mencionan los lineamientos y el contenido estricto para cada artículo y, por lo general en relación con los gráficos, la mayoría de las revistas acepta un total de seis figuras.

Recuerde no repetir información en el texto si ya se mencionó en las figuras.

Existen diversos tipos de representación gráfica. Éstos son:

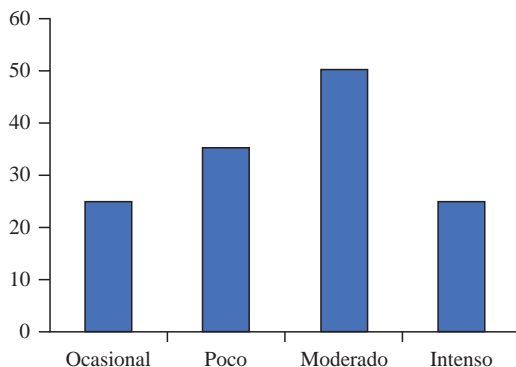


Figura 1: Ejemplo de diagrama de barras.

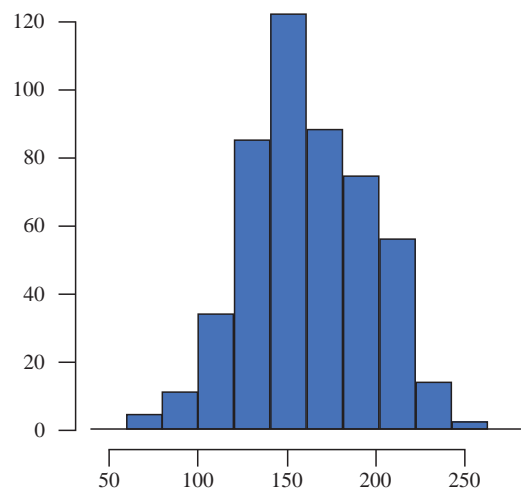


Figura 2: Ejemplo de histograma.

¿Cómo se presentan los datos?



Figura 3: Ejemplo de diagrama circular.

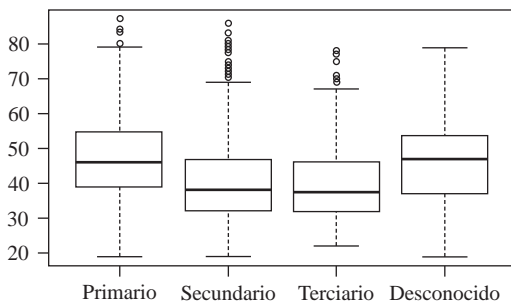


Figura 4: Ejemplo de diagrama de caja.

1. **Diagrama de barras**, el cual ayuda a evidenciar variables cuantitativas discretas y cualitativas en escala nominal. Se utiliza para mostrar los cambios en cantidades durante un periodo. Se recomienda la misma anchura en las barras y que se presenten de mayor a menor, a menos de que exista una secuencia pre-determinada que deba respetarse. También limitarse a cuatro u ocho barras (Figura 1).
2. **Histograma**, este gráfico ayuda cuando se trata de variables cuantitativas continuas, ya que agrupa frecuencias de presentación de las variables en intervalos (Figura 2).
3. **Diagrama de pizza**, circular o de pastel, adecuado para variables cualitativas en escala nominal. Muestra los datos porcentuales. Se recomienda limitar las rebanadas a máximo seis, ya que muchas porciones provocarán que el contenido

Frecuencia	Tallo & hojas
1	2* 4
1	2. 6
2	3* 33
6	3. 789999
8	4* 12333444
12	4. 56666678899
18	5* 00000111111222344
21	5. 666667888889999999999
21	6* 0000011111111122223444
19	6. 5555566666788888899
20	7* 001111222223333334444
16	7. 5555566777788999
4	8* 1234
1	8. 6

Figura 5: Ejemplo de diagrama de tallos y hojas. Observamos que el tallo corresponde a las decenas (en este ejemplo, cada decena se reparte en dos posiciones de tallo, el asterisco y el punto), y la hoja a las unidades. Este tipo de gráfico presenta la información original de datos pero más compacta y manipulable.

- deseado a mostrar no se aprecie correctamente (Figura 3).
4. **Diagrama de caja**, que funciona adecuadamente para las cuantitativas, cuyo objetivo es demostrar la desviación de la distribución normal. Estos diagramas poseen un extremo superior y uno inferior; el superior corresponde al percentil 75, mientras que el inferior corresponde al 25; y la línea horizontal, la cual divide a la caja, corresponde a la media (Figura 4).
 5. **Diagrama de tallos y hojas**, para variables cuantitativas que poseen pocos valores, debido a que, en este tipo de presentación, cada uno de los valores de las variables representadas se agrupan en relación con su distribución (Figura 5).

Lecturas complementarias

1. Jiménez Avila JM, Falavigna A. Educación en investigación: de la medición a la interpretación-bioestadística quirúrgica. Curitiba: TRAÇO Diferencial; 2018.
2. Socorro-Márquez F. ¿Cómo presentar datos estadísticos? Seminario de Investigación. 2015.
3. Martínez-González MA, de Irala J. Procedimientos descriptivos. En: Martínez-González MA. Bioestadística amigable. 2a ed. Madrid: Diaz de Santos; 2006.

Unidad

3

Pruebas paramétricas en la ortopedia

Prueba de t de Student

Arelhi Catalina González Cisneros

Antes de comenzar a hablar de la prueba t de Student, es esencial conocer que existen dos grandes grupos de sondeos estadísticos, las pruebas paramétricas y no paramétricas.

Estas se basan en la distribución de la muestra o de la población. En general, la diferencia entre estos dos conjuntos es que las paramétricas asumen los datos del comportamiento de una variable (media o varianza) y poseen un comportamiento normal, mientras que las no paramétricas no asumen nada acerca de la distribución, sólo toman en cuenta el ordenamiento y recuento de los valores de la variable sin importar esta última. Para saber qué tipo de prueba utilizar, debemos localizar ciertas características. Para utilizar las pruebas paramétricas, las variables suelen ser cuantitativas, se utilizan escalas de intervalo o razón, sus variables siguen una distribución normal, las varianzas son iguales y las muestras son grandes (> 30); por otro lado, para las pruebas no paramétricas, las variables son de distribución libres, variables cuantitativas, ordinales o nominales, grupos de varianzas grandes y sus muestras son pequeñas ($n < 30$ o < 10).

Cuando se tienen muestras pequeñas, es necesario encontrar un procedimiento alternativo para construir

intervalos de confianza. La prueba t de Student, desarrollada por William Sealy Gosset, ideada en un principio para valorar la distribución de muestras pequeñas, se basa en el teorema de pequeñas muestras y en el teorema del límite central.

Al hablar de bioestadística, una pregunta frecuente es valorar la significancia estadística de las diferencias entre muestras de individuos, ya sea que pertenezcan al mismo grupo o no. A esto se le conoce como muestras dependientes o independientes, respectivamente.

La t de Student es un tipo de estadística deductiva. Se utiliza para comparar muestras continuas distribuidas de forma normal entre dos grupos, por lo que existen dos tipos de t de Student, aquélla para muestras independientes (ejemplo: dos tratamientos) y otra para muestras dependientes (ejemplo: antes y después). Se utiliza frecuentemente en publicaciones médicas indizadas, nacionales e internacionales. Fue diseñada para examinar las diferencias entre dos muestras independientes y pequeñas, con una característica normal y varianzas homogéneas, es decir que se compara una variable de respuesta entre dos o más muestras independientes, siempre y cuando dicha cifra siga una distribución normal, por ejemplo, para saber de algún tratamiento nuevo, si

éste tiene mayor porcentaje de mejoría que otro ya utilizado.

La prueba t de Student, es una herramienta importante, ya que además de centrarse en observar las diferencias entre las variables, permite comparar las medias de dos mediciones.

Sus fundamentos son:

1. Las medias de las dos mediciones pueden ser diferentes.
2. Medias iguales para mediciones que provengan de la misma población.
3. Mientras más grandes sean las diferencias entre las dos medias, mayor será la certeza de que las diferencias no son al azar.

Por otra parte, sus supuestos son:

1. La variable medida se distribuye normalmente.
2. El nivel de medición de la variable es de intervalo.

La metodología de la t de Student comienza probando que cada una de las muestras tiene un comportamiento normal, obteniendo para cada una de éstas su tamaño (n_1 y n_2), sus medias (m_1 y m_2) y sus varianzas (v_1 y v_2). Continúa probando que las varianzas sean homogéneas, estableciendo la disparidad entre las medias y calculando la varianza común (vc) de las dos muestras.

$$vc = \frac{((n_1 - 1) \cdot v_1 + (n_2 - 1) \cdot v_2)}{(n_1 + n_2 - 2)}$$

Esto significa que la vc es igual a un promedio ponderado de las varianzas de las dos muestras, donde los pesos para ese promedio son iguales al tamaño, menos uno ($n - 1$) en cada una de las muestras.

Una vez obtenida la vc, se calcula el error estándar de la diferencia de las medias.

$$ESM = \frac{\sqrt{((vc) (n_1 + n_2))}}{(n_1 n_2)}$$

Y así resulta la t de Student como el cociente de la diferencia de medias divididas entre el ESM anterior.

Al ver concordancia con la hipótesis nula y alterna, es necesario demostrar que existe diferencia entre las medias de las muestras, para eso, se consulta una tabla de t de Student con grado de libertad igual a $n_1 + n_2 - 2$ y se calcula el valor de «p». Si los datos no se apartan significativamente de la realidad y la muestra es pequeña (menor de 30 individuos), la t de Student se considera un buen método con el cual comparar poblaciones, además de evaluar si la media de dos grupos es estadísticamente diferente una de otra, siendo apropiada de utilizar cada vez que se compara la media de dos poblaciones.

Finalmente, existen diversos factores que contribuyen para indicar si la diferencia entre dos medias de los grupos se puede considerar significativa:

1. Cuanto mayor es la diferencia entre ambas medias, mayor es la probabilidad de que una diferencia estadística significativa existe.
2. Cuanto menor sean las variaciones entre los dos grupos, mayor es la probabilidad de que exista una diferencia estadísticamente significativa.
3. El tamaño de la muestra es importante en la determinación de la significación de la diferencia entre las medias, ya que al aumentar el tamaño de la muestra, las medias tienden a ser más estables y más representativas.

La prueba t de Student, como se mencionó, se utiliza para comparar una variable de respuesta entre dos o

más muestras independientes. Si los datos no se alejan de la normalidad y la muestra es pequeña (< 30), la *t* de Student se considera una buena, sencilla y fácil herramienta para comparar poblaciones; además de evaluar si la media de dos grupos es estadísticamente significativa (diferente) una de la otra.

Lecturas complementarias

1. Falavigna A, Jiménez-Avila JM. Educación en investigación: de la idea a la publicación. 2a ed. Curitiba: TRAÇO Diferencial; 2015.
2. Jiménez Avila JM, Falavigna A. Educación en investigación: de la medición a la interpretación-bioestadística quirúrgica. Curitiba: TRAÇO Diferencial; 2018.
3. Sánchez Turcios RA. *t*-Student: usos y abusos. *Rev Mex Cardiol*. 2015; 26 (1): 59-61.

Prueba de análisis de varianza (ANOVA) de una vía

Laura Cortés Sanabria

La elección de una prueba estadística o prueba de hipótesis depende del tipo de variable dependiente (cuantitativa, cualitativa), del tamaño de la muestra (grande [distribución normal], pequeña [distribución anormal]), la dispersión de los datos (baja dispersión [paramétrica], amplia dispersión [no paramétrica]), número de grupos que se están comparando (dos grupos, más de dos grupos); si son conjuntos independientes (p. ej.; casos vs controles) o relacionados (p. ej.: mediciones basal vs final en un mismo grupo). Finalmente, el análisis estadístico también debe ser elegido correctamente, con base en elegir si sólo se busca comparar grupos, analizar una correlación, una concordancia, una asociación bivariada o multivariada, o la validez de una prueba diagnóstica o pronóstica.

En el caso del análisis de varianza de una vía, mejor conocida como de análisis de varianza (ANOVA), se utiliza para comparar tres o más muestras o conjuntos, en el caso de que la variable dependiente sea cuantitativa y tenga una distribución normal o paramétrica, lo anterior debido a esta, por lo que la t de Student sólo es útil para hacer inferencias mediante la comparación de dos muestras, pero no es adecuada cuando se tienen más de dos. La

prueba de ANOVA es un procedimiento estadístico que ayudará al médico a aceptar o rechazar la hipótesis nula de que no existen diferencias entre tres o más medias poblacionales.

Descripción de la prueba

Numerosos estudios de investigación en medicina comparan las medias de más de dos grupos, en este caso, la prueba de t de Student no es de utilidad, puesto que esta última prueba, sólo puede utilizarse cuando se comparan las medias de dos grupos. La elección recomendada para analizar datos, si existen más de dos grupos, se denomina análisis de varianza (ANOVA), la cual protege al investigador contra la «inflación» del error estándar, ya que si no se efectúa un examen global entre grupos, al ir comparando por pares de grupos alteraran este error estándar, no en cada comparación, sino en el experimento como un todo, entonces se estaría cometiendo una equivocación, al aceptar una diferencia de los resultados entre grupo como significativa cuando no lo es (error tipo 1), es decir, considerar la hipótesis nula como falsa cuando la hipótesis nula es verdadera, por ej.: asumir que el nuevo tratamiento es mejor que el tratamiento tradicional siendo la realidad

que el nuevo tratamiento no es mejor que el tratamiento tradicional.

El ANOVA proporciona un medio para dividir en dos partes la variación total del resultado, por ejemplo, cuando se realiza un estudio de investigación, en el cual se compara el componente físico de la calidad de vida relacionada con la salud (CVRS) de acuerdo con tres técnicas quirúrgicas para resolver una hernia de disco. El ANOVA toma en cuenta la variación de todas las observaciones, y la divide en: 1) la variación entre cada sujeto y la media de su grupo, y 2) la variación entre la media de cada grupo y la media global. Si las medias de los grupos son bastante diferentes unas de otras, habrá variación considerable entre éstas y la gran media, en comparación con la variación dentro de cada grupo. No obstante, si las medias de los grupos no difieren mucho, la variación entre éstas y la media global no será mucho mayor que la variación entre sujetos dentro de cada grupo.

Puntos importantes

La prueba estadística de ANOVA se va a utilizar cuando:

1. El tipo de variable dependiente que se va a analizar es de tipo cuantitativa.
2. Se comparan las medias entre tres o más grupos independientes.
3. La distribución de la muestra debe ser normal (paramétrica o en campana de Gauss).
4. Debe haber homogeneidad en las varianzas de cada grupo.
5. Se utiliza para evitar incrementar el error estándar al efectuar el análisis pareja por pareja de grupos, es decir, mantiene el error tipo 1 original al comparar varios grupos.

Ejemplo:

Supóngase un ensayo clínico en el cual se incluyen tres grupos de pa-

cientes que fueron asignados al azar, para ser intervenidos por tres determinadas técnicas quirúrgicas que intenten corregir una hernia de disco (técnica 1, 2, 3) y los investigadores midieron tres meses después dimensión física y mental de la calidad de vida relacionada a la salud (CVRS) mediante el cuestionario SF-36 (escala de 0 [peor resultado] a 100 [mejor resultado]). La *Tabla 7* muestra los resultados obtenidos en la dimensión física de la CVRS en cada uno de los grupos.

El promedio total de la muestra es 51 ± 19 . A continuación se va a explicar

Tabla 1: Comparación de la dimensión física de la CVRS de acuerdo con el tipo de técnica quirúrgica en el tratamiento de la hernia de disco.

	Técnica quirúrgica		
	Tipo 1 (N97)	Tipo 2 (N53)	Tipo 3 (N93)
	20	61	48
	13	27	36
	11	44	53
	48	79	53
	36	83	34
	21	73	44
	90	44	54
	30	15	34
	16	45	38
	72	27	19
	20	57	52
	49	74	11
	79	90	53
	62	53	17
	51	90	34
	43	79	85
	73	89	44
	30	36	55
	38	59	26
	33	94	51
	—	—	—
	—	—	—
Media	48	56	50
DE	21	21	19

Con el objetivo de optimizar espacio, se muestran resultados sólo de algunos pacientes de cada grupo, sin embargo, el valor de la media y desviación estándar corresponde a los resultados del total de pacientes incluidos en cada grupo (N). DE = desviación estándar.

cómo trabaja la prueba de ANOVA los datos, para poder comprender mejor la descripción de la prueba. El sujeto 1 en el grupo que fue operado de hernia de disco, mediante la técnica quirúrgica tipo 1, tiene un resultado en la dimensión física de la CVRS de 20, la gran media para todos los pacientes incluidos en el estudio es de 51, de modo que el sujeto 1 difiere de la gran media por $20 - 51 = -31$. Esta diferencia puede dividirse en dos partes: la diferencia entre 20 y la media de su propio grupo (48); así como la disparidad entre la media de su propio grupo y la media global. Por tanto:

$$(20 - 48) + (48 - 51) = -28 + -3 = -31$$

Al parecer, el sujeto 1 es «más diferente» de la media del grupo que se sometió a la técnica quirúrgica tipo 1, que lo que la media del grupo que se sometió a la técnica quirúrgica tipo 1 lo es respecto de la media global de los tres grupos. Una vez que se ha determinado que existen diferencias entre las medias, las pruebas de rango post hoc (Bonferroni, Tukey, Sidak, etc.) se identifican subconjuntos homogéneos de medias que no se diferencian unas de otras. Afortunadamente, se cuenta con paquetes estadísticos que realizan las estimaciones necesarias para obtener el nivel de significancia y, por lo tanto, se puede identificar si existen diferencias significativas entre grupos. Por ejemplo, utilizando el paquete estadístico SPSS para Windows, tomando en consideración el argumento anterior se obtienen las *Tablas 2 y 3* que muestran los resultados del ANOVA.

En la *Tabla 2*, el valor del estadístico de prueba, $F = 3,089$ es significativamente distinto de 1 para cualquier nivel de significación, por lo tanto, se rechaza la hipótesis de igualdad de medias y el valor de la significancia es menor a 0.05% ($p = 0.047$). Sin embargo, la significancia corresponde a todo el modelo, por lo que, para determinar la diferencia significativa entre cada uno de los grupos, los resultados del análisis de ANOVA se muestran en la *Tabla 3*.

Interpretación

Los resultados se pueden mostrar en una tabla como el valor de las diferencias de medias entre grupos o la media con su respectiva desviación estándar de cada grupo (*Tabla 4*). Al tomar en consideración el mismo ejemplo, en la *Tabla 4*, se muestran los resultados de la dimensión física y los diferentes dominios que la integran (percepción de salud general, funcionamiento físico, rol físico, dolor) en cada grupo, de acuerdo con cada método quirúrgico.

Para una adecuada interpretación de los valores y diferencias encontradas en la *Tabla 4*, se deberá describir de la siguiente manera: con un nivel de confianza de 95%, los resultados en la dimensión física de la CVRS, fueron significativamente mejores en pacientes con hernia de disco que fueron sometidos a cirugía mediante la técnica quirúrgica 2 en comparación con los pacientes en quienes se utilizó la técnica quirúrgica 1 ($p = 0.043$). Los resultados de la dimensión

Tabla 2: Resultados del ANOVA de un factor obtenidos para el modelo global.

Dimensión física	Suma de cuadrados	gl	Media cuadrática	F	Sig
Intergrupos	2,598.368	2	1,299.184	3.089	0.047
Intragrupos	152,683.181	363	420.615		
Total	155,281.549	365			

Tabla 3: Resultados del ANOVA de un factor obtenidos para cada uno de los grupos.

Tipo de técnica quirúrgica	Diferencia de medias	Error típico	Valor de p*	Intervalo de confianza al 95%		
				Límite inferior	Límite superior	
1	2	-8.17	3.31	0.043	-11.5	-4.85
	3	-2.17	2.52	1.00	-4.71	0.36
2	1	8.17	3.31	0.043	4.85	11.50
	3	5.99	3.69	0.317	2.29	9.70
3	1	2.17	2.52	1.00	-0.36	4.71
	2	-5.99	3.69	0.317	-9.70	-2.29

Tabla 4: Comparación de la dimensión física y los dominios que la componen de acuerdo con el tipo de técnica quirúrgica.

Dimensión	Técnica quirúrgica (media ± DE)		
	1	2	3
Dimensión física	48 ± 21	56 ± 21*	50 ± 19
Percepción de salud general	43 ± 23	42 ± 24	44 ± 21
Funcionamiento físico	52 ± 19	64 ± 18*	58 ± 18
Rol físico	26 ± 8	29 ± 15	23 ± 16
Dolor	70 ± 25	87 ± 19*	74 ± 17**

* p < 0.05 técnica quirúrgica 2 vs técnica quirúrgica 1.

** vs técnica quirúrgica 3.

física en pacientes en quienes se utilizó la técnica quirúrgica 3 no fueron diferentes en comparación con los grupos de la técnica 2 y 1 respectivamente. El dominio de funcionamiento físico fue mejor en técnica 2 vs técnica 1 (p = 0.043). Mientras que en el dominio dolor, los resultados fueron

mejores en la técnica 2 que en la técnica 1 y 3 respectivamente.

En percepción de salud general y rol físico no se observaron diferencias significativas entre grupos.

Conclusión

La prueba de ANOVA de una vía se debe utilizar cuando se comparan las medias de una variable cuantitativa entre tres grupos o más, siempre y cuando la distribución de la muestra sea normal.

Lecturas complementarias

1. Dawson-Saunders B, Trapp GR. Bioestadística médica. 2a edición, 3a reimpresión. México: Editorial Manual Moderno; 1999.
2. Celis de la Rosa AJ. Bioestadística. 2a ed. México: Editorial Manual Moderno; 2008.
3. Daniel WW. Bioestadística base para el análisis de las ciencias de la salud. 4a ed. México: Editorial Limusa Wiley; 2002.
4. Pérez C. Técnicas estadísticas con SPSS. España: Editorial Prentice Hall; 2001.

Correlación de Pearson

Adriana Margarita Leal Moctezuma,
Aldo Fidel Izaguirre Hernández

Los investigadores usualmente estudian si existe asociación entre dos variables para estimar la fuerza de dicha asociación, los análisis de correlación pueden estimar cuantitativamente la fuerza y la dirección de la asociación entre dos variables, los dos coeficientes de correlación más comunes son el de Pearson y el de Spearman.

Coefficiente de correlación lineal de Pearson

La correlación es una medida de asociación entre dos variables en donde una de ellas incrementa y la otra también; o una se incrementa y la otra disminuye, así, en datos correlacionados, el cambio en la magnitud de una variable se asocia con un cambio de magnitud de otra, ya sea en una dirección similar o dirección opuesta, en otras palabras, a más valor de una variable se asocian valores más altos (correlación positiva), o valores más bajos (correlación negativa) de otra variable y viceversa.

Una relación lineal entre dos variables es un tipo de asociación especial, de manera común, el término correlación se usa en el contexto de relaciones lineales entre dos variables continuas, aleatorias, y se conoce como correlación de Pearson y se abrevia con una «r».

El grado en que una variable continua se asocia con un cambio de otra continua se describe matemáticamente o en términos de covarianza de las variables. La covarianza es similar a la varianza, pero mientras que la varianza describe la variabilidad de una variable, la covarianza es una medida de cómo dos variables se relacionan y varían juntas; dado que la covarianza es una medida de dos variables y su magnitud absoluta es de difícil interpretación, el coeficiente de correlación de Pearson explica esta asociación en una escala sin dimensiones que va de -1 a +1.

Con los valores 0 y 1 se forma una relación perfecta positiva, con los valores -1 y 1, perfecta negativa, si se habla de la relación entre las variables X e Y, es perfecta positiva cuando exactamente en la medida que aumenta una de ellas, aumenta la otra (*Figura 1*).

Se dice que la relación es perfecta negativa cuando exactamente en la medida que aumenta una variable disminuye la otra, por ejemplo, la relación entre presión y volumen se ajusta a este caso (*Figura 2*).

En los fenómenos de investigación en ortopedia, usualmente se encuentran componentes aleatorios que no permiten relaciones perfectas (*Figura 3*).

Por ejemplo, a más horas de estudio diarias para el examen de consejo, mayor será la calificación en dicho examen.

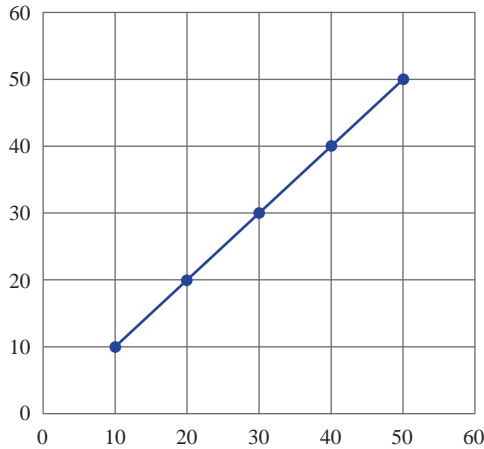


Figura 1: Ejemplo de relación perfecta positiva.

Por último, si los puntos no adoptan una configuración clara, puede ser que no haya alguna correlación (*Figura 4*).

Fórmula utilizada

El coeficiente de correlación de Pearson se define mediante la expresión:

$$r = \frac{\sum (x - \bar{x}) (y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

Hace referencia a la media de los productos cruzados de las puntuaciones estandarizadas de X y de Y, es un índice libre de escala de medida, las puntuaciones estandarizadas muestran la posición en desviaciones de un individuo respecto a su media.

Preceptos

Los datos deben provenir de muestras aleatorias, representativas de la población estudiada.

Ambas variables son continuas, tienen una distribución normal.

Interpretación del coeficiente de correlación

De acuerdo con la rama de estudio en donde se utilice la correlación, los lineamientos empíricos que describen

la fuerza de asociación serán diferentes. Por ejemplo, en psicología, Cohen sugiere una fuerza de asociación fuerte cuando el valor de r es mayor a 0.5.

En ortopedia se aceptan de manera convencional los siguientes valores:

Magnitud del coeficiente de correlación	Interpretación
0.00-0.10	Sin correlación
0.10-0.39	Correlación débil
0.40-0.69	Correlación moderada
0.70-0.89	Correlación fuerte
Mayor a 0.9	Correlación muy fuerte

Las correlaciones frecuentemente no se entienden, y se malinterpretan, es importante notar que la correlación observada (la fuerza de asociación) no implica que existe causalidad entre dos variables.

Además, las correlaciones no determinan acuerdo o fiabilidad entre dos mediciones, o dos observadores, dos mediciones pueden correlacionar, pero diferir, como cuando una observación mide consistentemente mayor que otra.

Conclusiones

Los coeficientes de correlación describen la fuerza de asociación y dirección de asociación de un par de variables. La correlación de Pearson es una me-

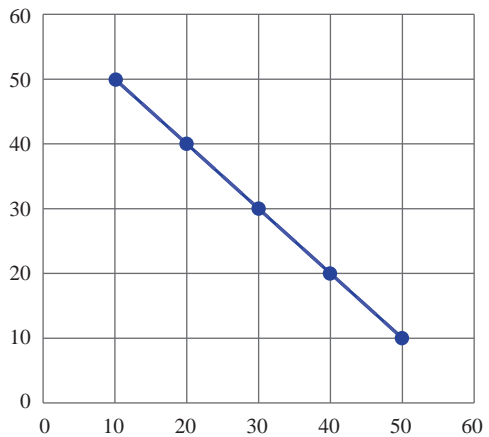


Figura 2: Ejemplo de relación perfecta negativa.

Correlación de Pearson

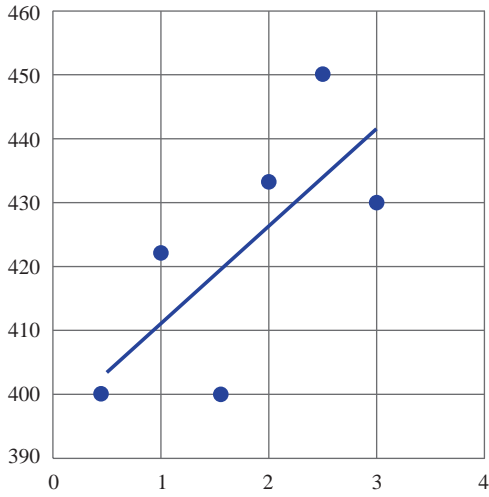


Figura 3: Ejemplo de relación no perfecta.

didada lineal de asociación entre dos variables aleatorias distribuidas de forma gaussiana.

Los ortopedistas deberán evitar inferir causalidad por correlación y no es una forma apropiada para medir concordancia entre mediciones u observaciones.

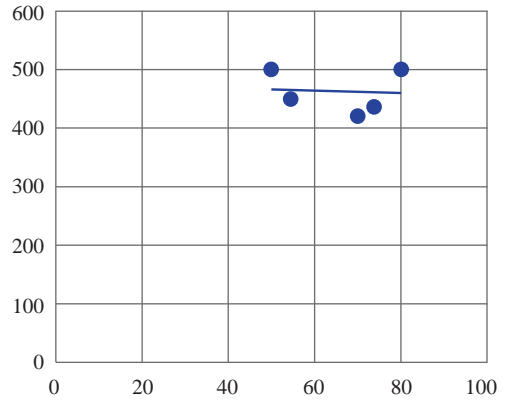


Figura 4: Ejemplo de correlación.

Lecturas complementarias

1. Petrie A. Medical statistics at a glance. 3rd ed. Chichester: Wiley-Blackwell; 2009.
2. Sprinthall R. Basic statistical analysis. 7th ed. Boston: Allyn and Bacon; 2003.
3. Altman DG. Practical statistics for medical research. London: Chapman & Hall/CRC; 1999. pp. 10-45.
4. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. Anesth Analg. 2018; 126 (5): 1763-1768.

Unidad

4+

Pruebas no paramétricas en la ortopedia

Chi cuadrada, χ^2 , ji cuadrada

Drusso López Estrada,
María de la Luz Arenas Sordo

La chi cuadrada tiene varios sinónimos, como se puede apreciar en el título. Cualquiera de estas formas es aceptable.

Antes de escoger una prueba estadística, como la χ^2 , se debe considerar que se necesita tener la pregunta de investigación adecuada y haber recabado la información correcta, ya que de otro modo, ningún análisis podrá corregir los errores que se hayan cometido.

A continuación se expondrá cuándo elegir este ensayo, y su utilidad de acuerdo con el tipo de investigación que se está llevando a cabo.

Descripción de la prueba

Esta prueba facilita buscar diferencias entre dos o más grupos. Es decir, revelará si los conjuntos se parecen o son diferentes. Pero sólo puede aplicarse en las variables en las que se obtienen frecuencias. En otras palabras, se usa para ver si existen discrepancias en proporciones o también para buscar independencia entre dos factores. Por esta razón también se le conoce como χ^2 de independencia.

Tipo de variables

Como se recuerda, las variables pueden, en forma inicial, clasificarse en **cuantitativas (numéricas)** y **cualitativas (frecuencias)**, la χ^2 es una prueba que com-

para frecuencias, entonces se utiliza en las cualitativas, por ejemplo la variable sexo. Estas variables también se conocen como categóricas, porque establecen categorías. Hay que hacer notar que las variables son exclusivas y exhaustivas, no se puede tener sexo masculino y femenino a la vez, se tiene uno u otro sexo.

También es necesario que existan frecuencias esperadas no tan pequeñas, en general mayores a 5, de no ser así, la más recomendable será la prueba exacta de Fisher.

Puntos importantes

1. Es un sondeo para comparar frecuencias entre dos o más grupos.
2. En todas las casillas de las tablas de 2×2 el número de casos deberá ser de cinco o más.
3. Si se trata sólo de cotejar hacia un extremo (por ejemplo, es mejor o es peor) se denomina de una sola cola. El valor de $p = 0.025$.
4. Si se contrasta considerando ambos extremos (es mejor o peor) se denomina de dos colas. Valor de $p = 0.05$.
5. Si los grupos se parecen la p es > 0.05 .

Ejemplo:

La tabla de contingencia clásica (tablas de 2×2) dividirá a los resultados

en dos categorías, prácticamente sí o no (Tabla 1).

Se calculan las frecuencias esperadas, para considerar si la prueba correcta que deba aplicarse es la χ^2 . Éstas se obtienen tomando en cuenta el número total (renglones = 18) y total (columnas = 79 o 75) y se divide entre el total general de pacientes (154) y así sucesivamente, cada renglón con cada grupo, recuérdese que puede haber más de un grupo. El resultado mostró que ninguna de las frecuencias esperadas es menor a 5, por lo que se puede utilizar χ^2 para comparar los grupos. En el ejemplo presentado, la p es < a 0.05, lo que exhibe que sí hay diferencia entre ambos.

Aplicación en la ortopedia

Cuando se realizan estudios en los que se investigan o se comparan por lo menos dos grupos, algunas variables son cualitativas, por ejemplo, el sexo o alguna otra de percepción. Y aunque en general se busca que las variables de importancia sean de tipo cuantitativo, siempre habrá algunas cualitativas. En general, cuando se comparan grupos, se desea que esos conjuntos se parezcan excepto en lo que se está estudiando. Se quiere que tengan aproximadamente el mismo número de hombres y mujeres, mismas edades, tallas y pesos similares etc., así se controlan algunas variables y los resultados pueden ser más confiables, de lo contrario se tendrían que realizar pruebas esta-

dísticas para controlar esas variables que podrían dar resultados equivocados.

En los estudios de ortopedia, la gran mayoría de las variables son de tipo cuantitativo, por ejemplo, días de recuperación, tamaño de las lesiones, grados de desviación, etc., sin embargo, como se mencionó previamente, también habrá aquellas cualitativas. Éstas pueden incluso modificar algunos de los resultados, por ejemplo, si se quiere estudiar presencia o no de dolor en grupos con mismos tratamientos quirúrgicos y se dividen esos conjuntos por sexo, es probable que se obtenga que en general sí hay diferencia porque las mujeres tienen umbrales más altos para la percepción del dolor, pero si esos colectivos son diferentes no por el sexo, sino por edad, por ejemplo, probablemente no exista discrepancia. Siempre es importante conocer las características de los grupos bajo estudio y el confrontar ciertas variables cualitativas facilitará conocerlas.

Si las variables cualitativas que están incluidas en el proyecto no son las que se están estudiando de manera especial, lo que se desearía es que esas no sean diferentes, es decir que en eso los grupos se parezcan de modo que puedan compararse las verdaderas variables de estudio.

Interpretación

El límite tradicional que se impone, para considerar que sí hay discordancias entre los grupos estudiados, es p menor de 0.05 (< 0.05) y así decir que

Tabla 1: Ejemplo de tabla de contingencia clásica.

Trombosis profunda	Grupo 1. Cirugía cadera (reemplazo) y medicamento antitrombótico	Grupo 2. Cirugía cadera (reemplazo) y no medicamento antitrombótico	Total
Sí	10	8	18
No	69	67	136
Total	79	75	154

sí hay divergencias entre los conjuntos estudiados.

¿Por qué ese valor?

Se considera que con esa cuantía se puede tener sólo 5% de error, lo cual es realmente bajo. Esto es que en cinco casos de 100 el resultado fue al azar. Es poder expresar que sí hay diferencia y que no la haya o al revés. Esto en ocasiones puede suceder debido al tamaño de la muestra, por eso la insistencia en que muestras de tamaños grandes siempre son mejores, aunque en ese cálculo matemático se obtengan cifras bajas.

Pero no olvidar que, cuando se busca que en los grupos de estudio no haya discordancia, la p deberá ser mayor a 0.05.

Cuando el valor que se utiliza es 0.05 implica que la prueba es de dos colas. Este concepto indica que el resultado puede ser tanto para un extremo como para el otro. Por ejemplo, cuando se desea saber si una complicación es menor o mayor a lo informado en la literatura, es decir que el resultado puede ser hacia un extremo o hacia el otro. Si sólo se desea saber si la complicación es menos frecuente, el estudio sería de una cola y la p tendría que ser igual o menor a 0.025 (< 0.025).

Conclusión

La chi cuadrada (χ^2) es un método estadístico que permite comparar frecuencias.

Éstas son las que se obtienen de las variables cualitativas.

Si las comparaciones entre dos o más grupos son a una sola dirección (más o menos) se llaman de una cola y la chi cuadrada deberá ser menor a 0.025, si son hacia ambos lados, se utiliza el límite de 0.05.

Si se analizan frecuencias de estos dos colectivos, para saber que los grupos se parecen, la p deberá ser mayor a 0.05 y si por el contrario las que se analizan son el eje de la investigación y se quiere o se piensa que habrá diferencia entre ellos, la p deberá ser menor a 0.05.

A fin de poder aplicar una prueba de χ^2 se debe recordar que las celdas de la tabla de 2×2 o mayor, deberán contener valores mayores a 5 en cada una.

Lecturas complementarias

1. Dawson-Saunders B, Trapp RG. Bioestadística médica. 4a edición. México: Manual Moderno; 2005.
2. Pagano M, Gauvreau K. Fundamentos de bioestadística. 2a edición. México: Ed. Thomson Learning; 2001.
3. Triola MF. Estadística. México: Ed. Pearson; 2004.

Prueba de Wilcoxon

Sergio Ramírez González,
Israel Gutiérrez Mendoza,
Leonardo López Almejo

Fue publicada en 1945 por Frank Wilcoxon, es un sondeo no paramétrico, que no tiene distribución normal y que posee las siguientes características: fácil de aplicar, se utiliza en variables ordinales, con observaciones de diferentes poblaciones. En una opción orientada a muestras pequeñas, es un análisis estadístico equivalente a la *t* de Student para muestras dependientes o relacionadas y para variables cuantitativas de distribución no normal o bien cualitativas ordinales. En el caso de muestras dependientes o emparejadas se han propuesto dos pruebas: la de los signos y la de rangos signados de Wilcoxon, se aplica a muestras reducidas, mayores de 6 y menores de 25, la experiencia consiste en sumar los rangos de signo frecuente.

Actúa sobre rangos de orden, es decir que utiliza las posiciones que ocupan los datos una vez ordenados. Por lo tanto, sólo es aplicable a variables cuyos valores se pueden ordenar. Tiene menos poder estadístico (menor probabilidad de rechazar la hipótesis nula cuando realmente es falsa) ya que se ignoran valores extremos.

El principio del análisis de signos se basa en la comparación del número de casos en que la primera muestra es mayor que la segunda.

Cuando el número de casos en que la segunda muestra es mayor que la primera muestra, la prueba no considera el tamaño de la diferencia entre cada par.

Wilcoxon propuso un análisis que tiene en cuenta el tamaño de la discrepancia dentro de los pares. Esta exploración se llama de rangos signados de Wilcoxon.

La prueba del signo de Wilcoxon puede aplicarse para una muestra o bien dos muestras emparejadas, en el caso de una muestra, la prueba del signo permite decidir si la mediana de la distribución de la variable en la población de la cual provienen los datos coincide o no con cierto valor; es la técnica más sencilla para una población pequeña que no tenga una clasificación normal y está diseñada para probar hipótesis relacionadas con la media de cualquier población. La lógica de la exploración de los rangos con signo de Wilcoxon es similar a la de *t* pareada. Si no hay disparidad en el antes y después, por ejemplo, las divergencias entre las observaciones deberían tender a cero.

Se basa en los signos que genera la diferencia al contrastar los números en una muestra con respecto a la mediana o con respecto a otros datos

tomados de la misma muestra o población, presentándose así dos casos:

1. Una muestra.
2. Una muestra o una población en pares.

La prueba de los signos es aquella que utiliza signos positivos y negativos para probar diferentes afirmaciones: acerca de la mediana, que implican datos nominales, que implican datos muestrales o poblacionales apareadas.

La idea primordial del método del signo es el análisis de las frecuencias de los signos positivos y de los negativos, para determinar si son significativamente diferentes, por razones de consistencia y simplicidad se emplea un estadístico de prueba con base en el número de veces que ocurre el signo menos frecuente.

Esta modalidad es utilizada para probar hipótesis sobre una mediana poblacional, en el caso de muchos de los procedimientos no paramétricos, la media es reemplazada por la mediana como el parámetro de ubicación pertinente bajo análisis:

Supuesto de la prueba:

$$P(X > \mu) = P(X < \mu) = 0.5$$

- Elección de datos de manera aleatoria.
- Sin el requisito de que las cifras muestrales provengan de una población con una distribución en específico.

Estadístico de prueba:

- Para $n \leq 25$ se emplea S (el número de veces que ocurre el signo, positivo o negativo, con menor frecuencia).
- Para $n > 25$ se emplea el estadístico Z .

$$Z = \frac{(x + 0.5) - (n/2)}{\sqrt{n/2}}$$

Estadístico:

La prueba de los signos utiliza el estadístico S , donde:

S = número de observaciones correspondiente al signo con menor frecuencia.

S sólo depende del signo (positivo o negativo) referente a la diferencia entre cada observación y el valor de la mediana.

Valores críticos:

Para $n \leq 25$ los valores críticos de S se encuentran en la tabla de valores críticos para la prueba de los signos.

Para $n > 25$ los valores críticos de Z se encuentran en la tabla de valores de Z para la distribución normal.

Cuando se utiliza la exploración del signo con datos apareados, se convierten los números en observaciones con signos positivos y negativos, considerando restar cada cuantía de la segunda variable del valor correspondiente de la primera variable y registrar sólo el signo de la diferencia que se encontró en el paso anterior, excluyendo los valores que son iguales.

Prueba de los signos para números apareados, la clave de la aplicación de esta modalidad es: si dos conjuntos de datos tienen medianas iguales, el número de signos positivos debe ser aproximadamente igual al número de signos negativos.

En una base de datos con el programa SPSS para realizar la prueba de Wilcoxon se va a la pestaña de **Análisis, Pruebas no paramétricas, Cuadros de diálogos antiguos, Dos muestras relacionadas** (Figura 1). Enseguida aparece el cuadro de diálogo en la que eligen las variables a contrastar, y se marca la de Wilcoxon (Figura 2).

Se seleccionan las variables o muestras a contrastar (antes y después), el programa presenta los espacios en pares y se selecciona el

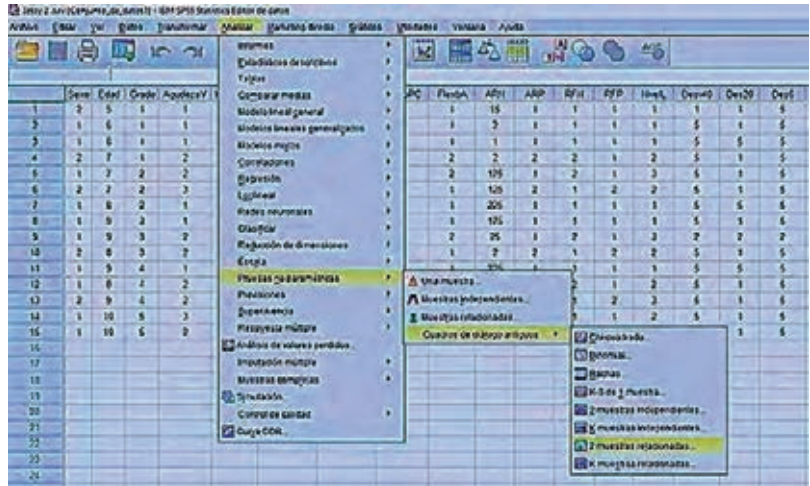


Figura 1:

Pasos para seleccionar la prueba de Wilcoxon en el programa SPSS.

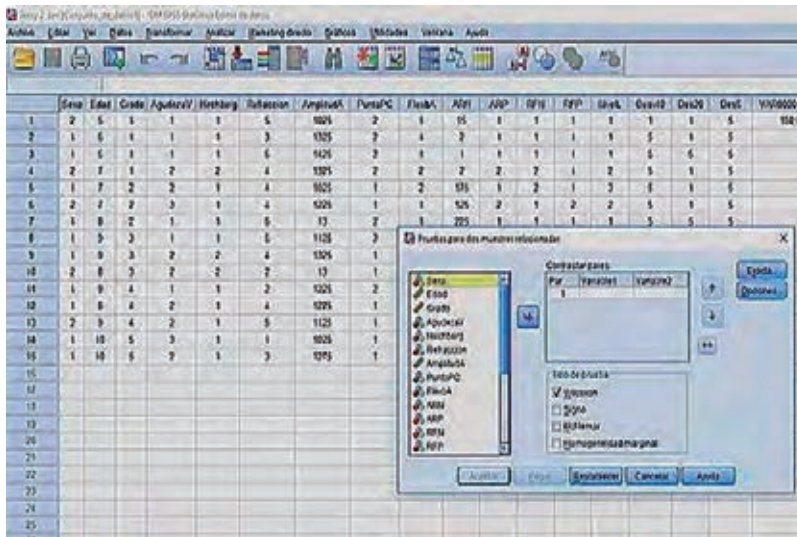


Figura 2:

Selección de muestras relacionadas y elección de la prueba.

análisis que se quiere hacer, en este caso Wilcoxon.

Ejemplo:

La aptitud clínica de los residentes de ortopedia y traumatología en el manejo del paciente politraumatizado (dependiendo del mecanismo de la lesión).

Se diseñó una estrategia educativa que permitiera la reflexión, la emisión de juicios y la elaboración de propuestas para la construcción del conocimiento. Con una estrategia de seis semanas, se efectuó con 100% de los médicos resi-

dentos de ortopedia y traumatología. El hecho de haber incluido a todos los residentes de ortopedia, de los cuatro años, fue con el fin de percibir la influencia de la experiencia clínica. Dos veces a la semana, durante 90 minutos en el aula, se procedía al debate y discusión de casos clínicos. Se utilizó como parte de la estrategia el caso clínico real problematizado, con una guía de lectura, apoyo visual consistente por lo general en fotografías relacionadas al percance, más la hoja de recolección de datos entregada por los paramédicos, así como bibliografía básica.

El programa se llevó a cabo en 11 residentes: tres del primer curso, cuatro del segundo, uno de tercero y tres del cuarto, se inició aplicando el instrumento de medición, se desarrolló el programa educativo y después de las 10 sesiones los residentes contestaron de nuevo el instrumento de medición.

En la base de datos se observa la evaluación previa a la intervención y la calificación posterior, así como los datos de edad y grado de residente (Figura 3).

Se selecciona la prueba de Wilcoxon de acuerdo al proceso de la Figura 4.

Al dar un clic, el programa conduce a la prueba para seleccionar las variables y análisis estadístico (Figura 5).

Los resultados obtenidos se observan en la Figura 6, en la que se muestra que existe una diferencia estadísticamente significativa ($p = 0.008$) entre las evaluaciones antes y después de la intervención educativa.

Al analizar los resultados se observó una mejoría en relación al número de aciertos contestados en cada uno de los 11 participantes, siendo significativa en los tres residentes del primer grado, los cuatro del segundo y los tres del cuarto.

Conclusión

La prueba de Wilcoxon es empleada cuando se requiere comparar los resultados de dos variables cuantitati-

	Residente	Edad	Sexo	Eval1	Eval2	var	var	var
1	1	26	2	69	99			
2	1	28	1	45	67			
3	1	27	1	61	94			
4	2	29	2	105	112			
5	2	29	2	57	110			
6	2	30	2	66	108			
7	2	29	1	113	103			
8	3	31	2	93	97			
9	4	32	2	61	109			
10	4	35	2	87	94			
11	4	33	1	109	123			
12								
13								

Figura 3:
Base de datos.

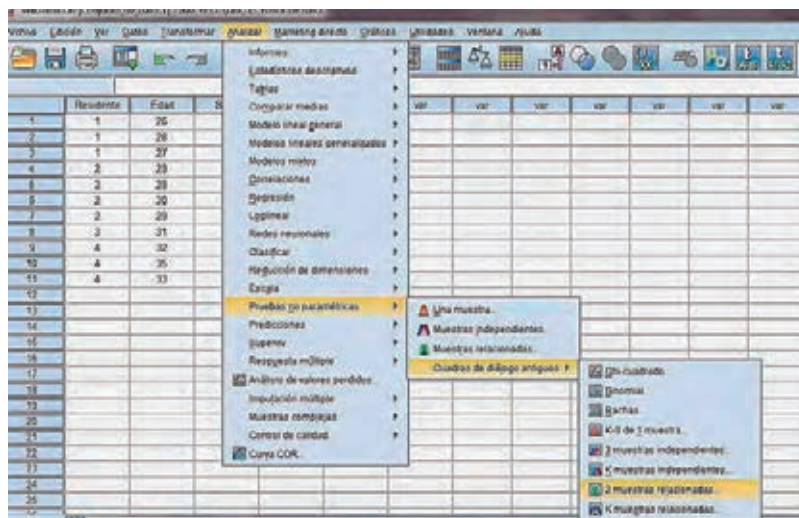


Figura 4:

Proceso para seleccionar la prueba de Wilcoxon.

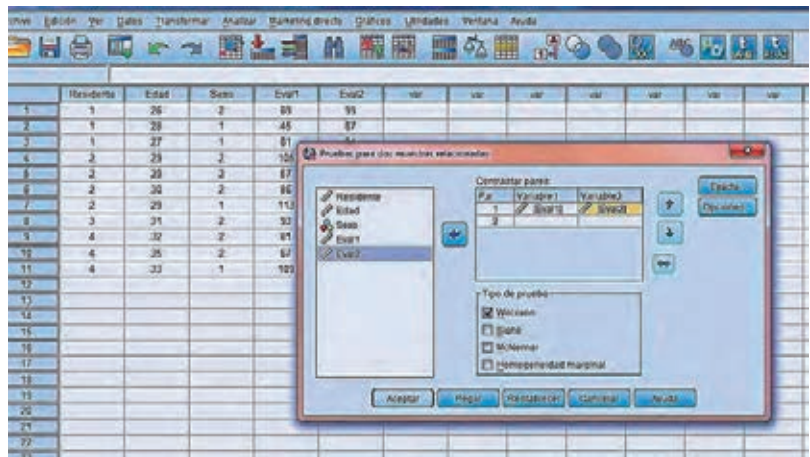


Figura 5:

Selección de las variables pareadas a contrastar para la prueba de Wilcoxon.

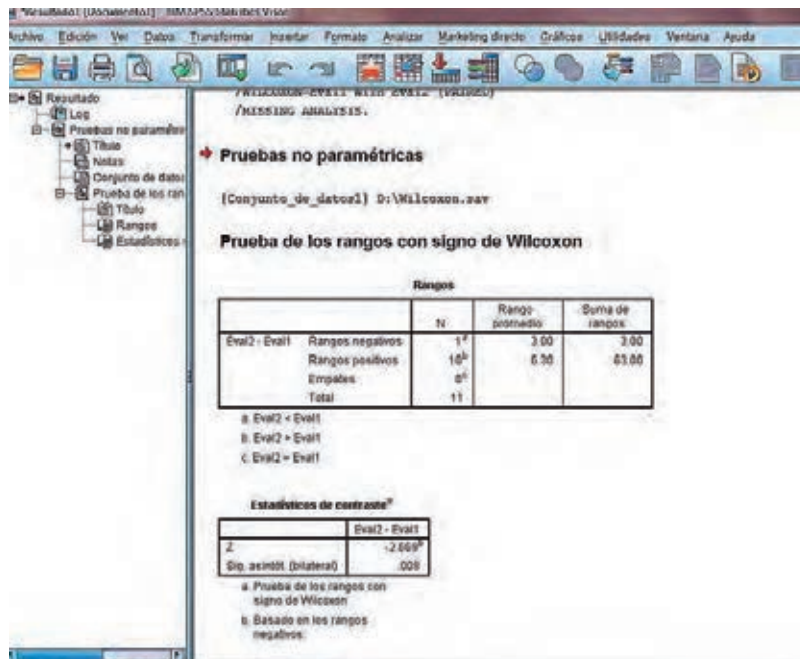


Figura 6:

Resultado de la prueba con el valor p (ver el último cuadro donde dice: Sig. asintót. (bilateral) con valor de 0.008).

vas que no tuvieron una distribución normal o para cualitativas ordinales. Se emplea en muestras relacionadas o pareadas, como por ejemplo una medición antes y otra después de la aplicación de un procedimiento o tratamiento en una población elegida por el investigador. Esta forma resulta de fácil aplicación en un programa estadístico y arroja como resultado el valor p de la comparación de ambas variables con la finalidad

de conocer si existen diferencias estadísticamente significativas entre ellas.

Lecturas complementarias

1. Berlanga Silvente V, Rubio Hurtado MJ. Clasificación de pruebas no paramétricas. Cómo aplicarlas en SPSS. REIRE, Revista d'Innovació i Recerca en Educació. 2012; 5 (2): 101-113.
2. Luis CS. Estadística simplificada. 2a ed. México: Trillas; 2002.

3. Luis CS. Manual práctico de estadística para ciencias de la salud. México: Trillas; 2011.
4. Armitage GB. Estadística para la investigación biomédica. España: Harcourt Brace; 1997.
5. Walpole ER, Myers HR, Myers LS. Probabilidad y estadística para ingenieros. 6ta ed. México: Prentice-Hall Hispanoamericana; 1999: 620-623.
6. Celis de la Rosa J. Bioestadística. México: Manual Moderno; 2004.
7. Arredondo GE, Castillo SEJ, Sánchez GF. La aptitud clínica de los residentes de ortopedia y traumatología en el manejo del paciente politraumatizado (dependiendo del mecanismo de la lesión). Acta Ortop Mex. 2006; 20 (2): 64-71.

Unidad

4

Pruebas no paramétricas en la ortopedia

Prueba U de Mann-Whitney

José Manuel Pérez Atanasio

La prueba estadística conocida como U de Mann-Whitney busca identificar si existe diferencia entre las medias de dos muestras independientes, también es llamada Mann-Whitney-Wilcoxon, prueba de suma de rangos Wilcoxon, o bien de Wilcoxon-Mann-Whitney. Se aplica para muestras que presenten una distribución distinta a la normal, por lo que se considera la versión no paramétrica del método t de Student.

Para comprender cómo se interpreta este análisis se debe partir de conceptos elementales como la curva de distribución de la normalidad o de Gauss, esta curva muestra al centro la media de los valores de la población y a cada lado sus desviaciones estándar como se observa en la *Figura 1*.

La utilidad de esta curva es que a partir de la media se puede saber qué porcentaje de la población se tendrá si se suman o restan las desviaciones estándar. Por ejemplo, si en una población de pacientes la media del tiempo de consolidación de la fractura es de 45 días y su desviación estándar es de cinco días, se sabe que 68% de los pacientes consolidarán en 45 ± 5 días, en otros términos, desde 40 y hasta 50 días.

El 95% de la población entre $45 \pm 5 \times 2$ días, es decir entre 35 y 55 días, para obtener esta cuantía se suman y se

restan a la media dos desviaciones estándar, que en este caso corresponde a 5×2 , o sea 10 días.

A partir de tres desviaciones estándar se incluye al 99% de la población, se pueden sumar y restar cuatro o más desviaciones estándar y se mantendrá en 99% de la población con decimales infinitos, de hecho, por definición no se puede llegar al 100% de la población. Debido a esta limitación y en condiciones clínicas es habitual tomar como representativa una población de 95%.

En la *Figura 2* se puede ver una representación gráfica de una población en la que a partir de su media se suman o restan las desviaciones estándar.

De manera personal, cuando descubrí la curva de distribución de Gauss,

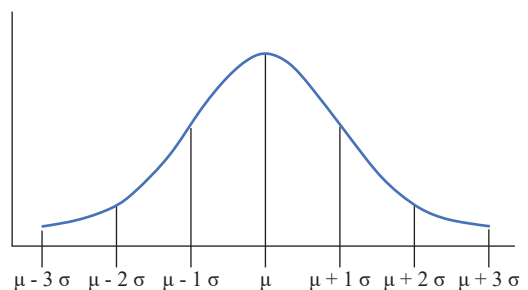


Figura 1: Curva de distribución de la normalidad o de Gauss. Se muestra la media (μ) al centro de la curva y hacia cada lado las desviaciones estándar (σ). Obsérvese que hacia la derecha se suman a la media las desviaciones estándar y hacia la izquierda se restan.
 μ = media; σ = desviación estándar.

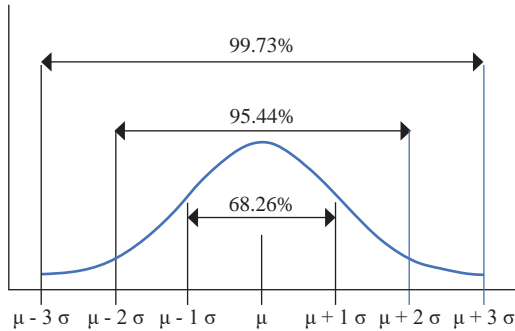


Figura 2: Porcentaje que se incluye del total de la muestra si se suman bilateralmente una, dos o tres desviaciones estándar a partir de la media, en la curva de distribución de la normalidad o de Gauss.
 μ = media; σ = desviación estándar.

en la que siempre ± 2 desviación estándar a partir de la media se incluye al 95% de la población estudiada, me pareció de una belleza prístina por su sencillez y gran utilidad. Fue entonces cuando investigué quién la había descrito y supe que fue el alemán Johann Carl Friedrich Gauss nacido en 1777, y para mi sorpresa no fui el único que se había maravillado con su descubrimiento, el gobierno alemán en reconocimiento a su aportación plasmó su imagen en un billete de 10 marcos, en el que se observa al fondo la fórmula matemática de la función para obtenerla, como se muestra en la *Figura 3*.

Descripción de la prueba

Al retomar el objetivo de este capítulo, se debe mencionar que las incursiones estadísticas que buscan identificar diferencia entre las medias de dos muestras, como la prueba U de Mann-Whitney, en esencia lo que tratan de saber es si las medias están lo suficientemente separadas para ser un hecho estadísticamente significativo, como se muestra en la *Figura 4*.

La prueba U de Mann-Whitney, que es de tipo no-paramétrica, se utiliza cuando no se cumplen las condiciones exigidas para la aplicación de las

paramétricas o cuando las muestras son pequeñas.

Aplicación en la ortopedia

Se quiere saber si el tiempo de consolidación de las fracturas diafisarias de fémur es menor cuando se utiliza una placa o con un clavo, en dos grupos reducidos de pacientes.

Lo primero que se tiene que hacer es obtener la media y la desviación estándar del tiempo de consolidación de la fractura si se utiliza una placa o un clavo.

Dígase que con la placa los pacientes consolidaron en 200 ± 30 días ($\mu = 200, \sigma = 30$) y con el clavo en 140 ± 20 días ($\mu = 140, \sigma = 20$). Parece claro que con el clavo consolidan más rápido, pero para afirmarlo estadísticamente, se decide aplicar la prueba U de Mann-Whitney buscando evaluar la disparidad entre estas medias. El resultado de esta acción muestra un valor de p menor a 0.05, por lo que se puede interpretar que, en la muestra estudiada, existe una divergencia estadísticamente significativa en el tiempo de consolidación de las fracturas si se usa una placa versus un clavo.



Figura 3: Imagen de Johann Carl Friedrich Gauss en un fragmento del billete de 10 marcos alemanes, actualmente fuera de circulación. Nótese al fondo la curva de distribución y su fórmula matemática.

Prueba U de Mann-Whitney

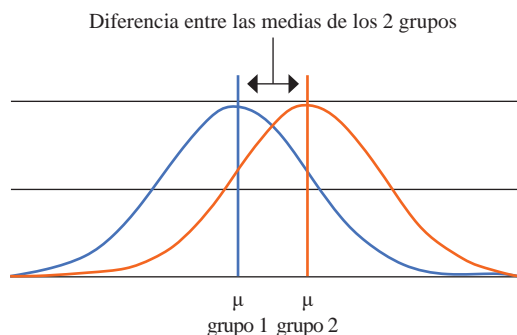


Figura 4: Representación esquemática de las pruebas estadísticas que buscan identificar si existe discrepancia entre las medias (μ) de dos grupos.

Conclusión

La prueba U de Mann-Whitney es de tipo no paramétrica que busca identificar si existe diferencia entre las

medias de dos muestras independientes.

Lecturas complementarias

1. Rivas-Ruiz R, Moreno-Palacios J, Talavera JO. Investigación clínica XVI. Diferencias de medianas con la U de Mann-Whitney. Rev Med Inst Mex Seguro Soc. 2013; 51 (4): 414-419.
2. Falavigna A, Jiménez Avila JM. Education in research: from the idea to the publication. Caxias do Sul, Brazil: Educ; 2015.
3. Jiménez Avila J, Falavigna A. Education in research: from measurement to interpretation - surgical biostatistics. Asociación AOSpine Latin America; 2018.
4. Toledo-Infanson V, Jiménez Avila JM. Publicación con impacto en ortopedia y traumatología. México: Ed. Graphimedic; 2019.
5. Pérez AJM. Principios de bioestadística para comprender los resultados reportados en los artículos científicos. Ortho-tips. 2015; 11 (2): 91-95.

Prueba de Kruskal-Wallis

José Alfredo Penagos Paniagua

La prueba de Kruskal-Wallis es de modalidad no paramétrica, que evalúa las diferencias entre tres o más grupos muestreados independientemente en una sola variable continua no distribuida normalmente. Los datos no distribuidos normalmente (p. ej., los ordinales o de rango) son adecuados para la exploración de Kruskal-Wallis. En contraste, el análisis de varianza unidireccional (ANOVA), que es una forma paramétrica, puede usarse para una variable continua normalmente distribuida, la prueba de Kruskal-Wallis es una extensión de la prueba U Mann-Whitney de dos grupos, por lo tanto, Kruskal-Wallis es una forma más generalizada del método U de Mann-Whitney y es la versión no paramétrica del ANOVA unidireccional.

Una aplicación típica del ensayo de Kruskal-Wallis es evaluar si tres o más conjuntos difieren en una sola variable que no cumple con los supuestos de normalidad de ANOVA, como la variable de interés no cumple con los supuestos de normalidad, es posible que no se comparen las medias grupales; en cambio, se comparan rangos. La hipótesis nula específica que los grupos son subconjuntos de la misma población (es decir, $H_0: (a, b, c, \dots, N) \subseteq p$), para probar la hipótesis nula, se combinan los grupos en un solo grupo y se clasifica la variable de interés, los nuevos puntajes de rango se suman por gru-

po (T_a, T_b, \dots, T_n) y junto con los tamaños de muestra de grupo, se pueden usar para calcular el estadístico H.

H refleja la varianza en los rangos entre grupos y se parece mucho a la distribución de χ^2 , al probar el resultado nulo, se puede usar H y referirse a una tabla de χ^2 con grados de libertad igual a n (número de grupos) menos 1. Si H excede un valor crítico, se puede concluir que los colectivos no provienen de la misma población; muchos investigadores a menudo concluyen que es una prueba general para las diferencias entre los grupos, pero otros tantos se oponen a esta interpretación.

Prueba estadística

La fórmula computacional para la prueba estadística H, es:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1)$$

donde N es el tamaño total de la muestra, n_i es el tamaño del grupo i -ésimo, k es el número de grupos y R_i es la suma de rango del grupo i -ésimo. Rechace H_0 cuando $H \geq$ valor crítico.

Tamaños de muestra grandes

Al trabajar con muestras de gran tamaño, la distribución nula es aproxi-

mada por la calculada con χ^2 con $k - 1$ grados de libertad. Por lo tanto, la regla de decisión es rechazar H_0 si $H \geq \chi^2_{\alpha, k - 1}$, donde $\chi^2_{\alpha, k - 1}$ es el valor de χ^2 a α nominal con $k - 1$ grados de libertad.

Ejemplo:

En el estudio publicado en 2016 por Vanlommel L y colaboradores, en *Knee Surgery, Sports Traumatology, Arthroscopy*, titulado *Predictores de resultados después de la manipulación bajo anestesia en pacientes con artroplastia total de rodilla rígida*, se analizaron retrospectivamente los datos de 158 pacientes (138 rodillas) con una rodilla rígida después de artroplastia total de rodilla (ATR) que recibieron una manipulación bajo anestesia (MBA) entre 2004 y 2014. Se identificaron y examinaron las variables pre, peri y postoperatorias para determinar su influencia en la flexión después de la manipulación. Con la meta de determinar la flexión y extensión, se utilizaron mediciones goniométricas con una alta confiabilidad intra e interobservador. En este estudio, la prueba de Kruskal-Wallis se utilizó para comparar la mejoría en la flexión posterior a la MBA para tres diferentes diseños de prótesis de rodilla. Después de la MBA, se observó una mejora media en la flexión de 30.3° en el seguimiento final. La flexión preoperatoria de la ATR, el diseño de la ATR y el intervalo entre el

procedimiento de ATR y la MBA fueron positivos asociados con un aumento de la flexión después de la MBA ($p = 0.005$).

En conclusión, la forma de Kruskal-Wallis no requiere el cumplimiento de supuestos de distribución normal, datos de intervalo y homogeneidad de la varianza del grupo. Esta prueba es una técnica más flexible, conveniente, fácil de usar y potente similar a un ANOVA paramétrico de una vía.

Lecturas complementarias

1. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. *J Am Statistic Assoc.* 1952; 47: 583-621.
2. McKnight PE, Najab J. Kruskal Wallis test. In: Weiner IB, Craighead WE. *The Corsini encyclopedia of psychology.* 4th ed. Vol 3. New York: John Wiley & Sons Inc; 2010. p. 904.
3. Ostertagová E, Ostertag O, Kovác J. Methodology and application of the Kruskal-Wallis test. *Applied Mechanics and Materials.* 2014; 611: 115-120.
4. Sawilowsky S, Fahoome G. Kruskal-Wallis test. In: Everitt BS, Howell DC. *Encyclopedia of statistics in behavioral science.* Vol 2. Chichester: John Wiley & Sons, Ltd.; 2005. pp. 1026-1028.
5. Guo S, Zhong S, Zhang A. Privacy-preserving Kruskal-Wallis test. *Computer Methods and Programs in Biomedicine.* 2013; 112 (1): 135-145.
6. Vanlommel L, Luyckx T, Vercruyse G, Bellemans J, Vandenuecker H. Predictors of outcome after manipulation under anaesthesia in patients with a stiff total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc.* 2017; 25 (11): 3637-3643.

Correlación de rho de Spearman

Laura Cortés Sanabria

Una correlación se refiere al grado de dependencia mutua entre dos variables, p. ej. el comportamiento de las cifras de tensión arterial de acuerdo con la edad de una persona. El coeficiente de correlación lineal intenta medir la magnitud o intensidad de esta dependencia, es decir la calidad del ajuste al analizar las dos variables relacionadas que se denominan Y (p. ej. presión arterial) y X (p. ej. edad).

La forma como se interpreta el resultado de la correlación lineal entre las variables X y Y es la siguiente (*Tabla 1*).

El grado de asociación entre X y Y será mayor conforme la cuantía de «r» se acerque más al valor de uno, y será menor a medida que «r» se acerque más a cero. Es importante también interpretar el tamaño de «r» mediante el coeficiente de determinación, que se estima elevando al cuadrado el coeficiente de correlación (r) en el caso de que la correlación entre la edad y la presión arterial sea de (0.56) o 0.31; este resultado revela que 31% de la variación en la cifra de tensión arterial puede explicarse por los valores de la variable edad.

Existen diferentes pruebas para medir la correlación simple que une dos variables (Pearson, Spearman, gamma de Goodman y Kruskal, Tau-b de Kendall), sin embargo, las más utilizadas

son la correlación de Pearson y la rho de Spearman.

A la rho de Spearman también se le conoce como correlación del rango, y se usa con datos ordinales o en situaciones en donde las variables numéricas no tienen distribución normal de probabilidades (campana de Gauss) (*Figura 1A*), en otras palabras, esta variante no se ajusta a los llamados criterios paramétricos (la media, mediana y moda coinciden al centro de la curva) debido a un pequeño número de muestra (menor de 30 observaciones) o existe variabilidad importante en el valor de los datos que ocasiona una distribución no paramétrica (*Figura 1B*).

Descripción de la prueba

Cuando se tiene una distribución anormal por la presencia de valores extremos, la medida de tendencia central que se estima es la mediana que corresponde a la valía de la observación central; es decir, la mitad de los valores de la variable es menor y la otra mitad mayor con relación al monto central.

Las medidas de dispersión que se utilizan en este caso son el rango (amplitud), que es la diferencia entre el valor máximo y mínimo de una variable. Por ejemplo, la valoración más pequeña de presión arterial sistólica en un

Tabla 1: Ejemplo de la interpretación del resultado de una correlación.

Valor del coeficiente (r)	Interpretación	Ejemplo
1	Existe correlación perfecta positiva y la relación funcional entre las variables X, Y es exacta y positiva, evolucionando ambas variables en el mismo sentido	A mayor edad, mayor es la cifra de presión arterial. A menor edad, menor es la cifra de tensión arterial
-1	Existe correlación perfecta negativa y la relación funcional de las variables X, Y es exacta pero negativa, es decir, mudan ambas variables en sentido opuesto	Al aumentar la edad disminuye la tensión arterial Al disminuir la edad aumenta la presión arterial
0	La correlación es nula, o sea no hay asociación entre las variables X, Y, siendo imposible encontrar una relación funcional entre ellas	No existe una asociación entre la edad con el comportamiento de la cifra de presión arterial

total de 20 pacientes es 90 mmHg y el más alto es 170 mmHg, por tanto, el rango es $170 - 90 = 80$.

El monto del coeficiente de correlación «r» se afecta de manera importante debido a la presencia de valores extremos de una variable y, por tanto, el resultado de la asociación entre las dos está sesgado o incluye cifras distantes. Un método fácil y sencillo para afrontar el problema de cuando se tienen estos valores extremos en la correlación consiste en transformar los datos a rangos y en seguida calcular la correlación de los rangos.

Este procedimiento estadístico es lo que se denomina el coeficiente de correlación por rangos de Spearman conocido también como coeficiente de correlación ordinal.

Por lo tanto, la rho de Spearman se utiliza para determinar una correlación entre dos variables (X, Y) cuando la variable es cuantitativa u ordinal, en presencia de una distribución no paramétrica de los datos.

Puntos importantes

1. La rho de Spearman es una correlación bivariada, es decir, sólo se busca analizar la asociación entre una variable dependiente (Y [ej. presión arterial]) y una independiente (X [ej. edad]), o sea qué tanto cambia la variable Y (tensión arterial) por efecto de la variable X (edad).

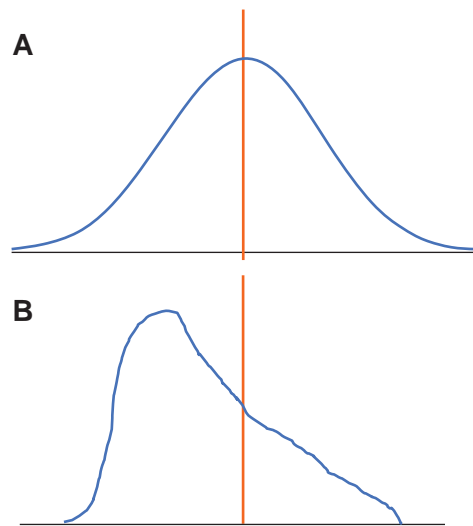


Figura 1: A) Distribución normal. B) Distribución anormal.

2. Debe utilizarse para bases de datos en las que existan valores extremos, pues si se calcula la correlación de Pearson, los resultados se verán afectados.
3. El grado de asociación obtenido no es el de los valores de las variables, sino el de las clasificaciones de los rangos de dichos valores.
4. La interpretación del resultado del coeficiente de correlación de Spearman se encuentra en el rango de valores de -1 y 1.
5. La significancia estadística de un coeficiente debe tenerse en cuenta conjuntamente con la relevancia clínica del fenómeno que se estudia.

Ejemplo:

Se realizó un estudio descriptivo para determinar la asociación que existe entre la edad (variable X) sobre la cifra de tensión arterial (variable Y) en adultos de ≥ 20 años de edad aparentemente sanos.

La siguiente tabla indica las edades en años y las cuantías de la presión arterial sistólica (PAS) de los 20 adultos (Tabla 2).

Al utilizar el paquete estadístico SPSS versión 20 para Microsoft Windows, se estimaron los coeficientes de correlación de Pearson (0.72, $p < 0.0001$) y rho de Spearman (0.47, $p = 0.037$), en donde se puede observar una clara discrepancia entre ambos escenarios, debido a que el resultado del coeficiente en la correlación de Pearson se encuentra afectado por los últimos cuatro individuos de la lista, que por edad y PAS se alejan del resto de la población.

Tabla 2: Ejemplo de estudio para determinar la asociación entre dos variables.

Edad (variable X)	PAS (variable Y)	Edad (variable X)	PAS (variable Y)
20	100	46	128
20	144	53	113
28	130	53	100
30	120	56	135
31	134	58	124
33	120	58	130
34	126	80	165
35	128	81	180
43	140	90	190
46	132	92	190
Media \pm desviación estándar		49 \pm 22	135 \pm 5
Mediana (mínimo-máximo)		46 (20-92)	130 (100-190)
Correlación de Pearson		0.72, $p < 0.0001$	
Rho de Spearman		0.47, $p = 0.037$	

PAS = presión arterial sistólica.

El valor de «r» de la rho de Spearman es mucho menor que el de la correlación de Pearson, por lo tanto, cuando se quiere analizar la asociación que existe entre dos variables y se tienen valores extremos en una o ambas es mejor calcular la correlación de los rangos o rho de Spearman.

Interpretación

Por ser un coeficiente de correlación, el resultado de la correlación por rangos de Spearman se interpreta de la siguiente manera:

Valor del coeficiente =

1: concordancia perfecta.

-1: discordancia perfecta.

0: no existe concordancia ni discordancia.

Conclusión

La rho de Spearman es de utilidad cuando un investigador desea determinar la correlación o asociación entre:

1. Dos variables ordinales.
2. Dos variables cuantitativas pero que presentan algunos montos alejados del resto de los valores de la variable (valores extremos u *outliers*).
3. Cuando la muestra es pequeña y no tiene una distribución en campana de Gauss.

Lecturas complementarias

1. Dawson-Saunders B, Trapp GR. Bioestadística médica. 2a edición, 3a reimpression. México: Editorial Manual Moderno; 1999.
2. Celis de la Rosa AJ. Bioestadística. 2a edición. México: Editorial Manual Moderno; 2008.
3. Daniel WW. Bioestadística base para el análisis de las ciencias de la salud. 4a edición. México: Editorial Limusa Wiley; 2002.

Prueba de McNemar

José Manuel Pérez Atanasio

La prueba estadística de McNemar se utiliza para determinar si existe diferencia entre dos proporciones de una variable dicotómica en grupos relacionados.

Antes de empezar a describir esta modalidad y su utilidad, es conveniente partir de entender qué es una variable dicotómica y qué son los grupos relacionados.

Se podría expresar el resultado de un examen con la calificación obtenida en un salón de clases en una escala del 0 al 100 con las unidades y dos decimales, en otros términos: 66.34, 72.58, 89.72, etcétera, por lo que cada calificación de los alumnos podría tener valores continuos, el expresar las calificaciones de esta manera se considera como una variable continua.

Por otra parte, si se expresa la calificación obtenida sólo con dos posibilidades: aprobado o reprobado, se considera una variable dicotómica que es aquella que sólo puede tomar dos valores y ningún otro.

El concepto de grupos relacionados se refiere a muestras que son obtenidas en una misma población, pero en distinto momento, por ejemplo, las calificaciones obtenidas en dos exámenes aplicados a residentes de ortopedia sobre principios biomecánicos, uno realizado durante su primer mes de ingreso y otro al finalizar el curso. Los residentes son el mismo grupo

evaluado, pero en dos momentos distintos.

Ejemplo:

Para comprender la prueba de McNemar se va a continuar con el ejemplo del grupo de médicos residentes contestando un examen de principios biomecánicos en dos momentos de su residencia.

Lo primero que se debe hacer es calificar los exámenes con dos posibilidades: aprobado o reprobado, con esto se expresa la calificación como una variable dicotómica, lo que es un requisito para aplicar aplicación estadística. Posteriormente se agrupan los resultados en la *Tabla 1*.

Ahora se calculan las proporciones de aprobados y reprobados en los dos momentos.

En el primer examen sólo aprobaron dos de los 12 residentes, es decir, una proporción de 2/12 que equivale al 16.6%.

En el segundo examen, probablemente derivado del entusiasmo de los profesores y al esfuerzo de los alumnos aprobaron nueve de los 12 residentes, con una proporción de 9/12 equivalente al 75%.

Ahora bien, al ver esta diferencia en la proporción de aprobados en los dos exámenes es evidente que hubo una mejoría, pero para demostrar que esta

disparidad es estadísticamente significativa se debe aplicar la prueba de McNemar. Los datos se pueden incorporar a una tabla de Microsoft Excel o directamente al software de análisis estadístico de su preferencia, pero es importante que sea de forma numérica, en otros términos, asignando un número a cada valor posible, por lo que se representará con un «1» cuando esté aprobado el alumno y con un «0» cuando este reprobado.

Una vez que se tienen los montos, se selecciona la prueba para identi-

ficar la discordancia de proporciones correspondiente a la de McNemar, el valor que se muestra en los resultados se expresará con un valor de «p» que es el que se tiene que interpretar, en este ejemplo la cifra mostrada para «p» es menor de 0.05, por lo que se debe interpretar que existe una discrepancia estadísticamente significativa entre la proporción de aprobados en el examen realizado en los dos momentos distintos, y que es poco probable (menos de 5% de probabilidad) que esta diferencia se deba al azar.

Tabla 1: Ejemplo de tabla de resultados usando una variable dicotómica.

Alumno	Resultados del examen aplicado en el primer mes de la residencia	Resultados del mismo examen, pero aplicado al finalizar el curso
1 Juan	Reprobado	Aprobado
2 Pedro	Reprobado	Aprobado
3 Antonio	Reprobado	Aprobado
4 Ramón	Reprobado	Aprobado
5 José	Reprobado	Reprobado
6 Severiano	Reprobado	Aprobado
7 María	Aprobada	Aprobada
8 Adela	Reprobada	Reprobada
9 Adriana	Aprobada	Reprobada
10 Ruth	Reprobada	Aprobada
11 Vanesa	Reprobada	Aprobada
12 Jimena	Reprobada	Aprobada

Conclusión

La prueba estadística de McNemar se utiliza para determinar si existe desigualdad entre dos proporciones de una variable dicotómica en grupos relacionados.

Lecturas complementarias

1. McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*. 1947; 12 (2): 153-157.
2. Jiménez Avila J, Falavigna A. Education in research: from measurement to interpretation - surgical biostatistics. Asociación AOSpine Latin America; 2018.
3. Toledo-Infanson V, Jiménez Avila JM. Publicación con impacto en ortopedia y traumatología. México: Ed. Graphimedic; 2019.

Regresión lineal simple

Víctor Toledo Infanson

Introducción

El modelo de regresión lineal simple es una técnica paramétrica tan simple que hay quienes consideran que ni siquiera vale la pena estudiarla, sin embargo, es su misma simplicidad la que da su utilidad.

Este método permite encontrar el valor que se espera de una variable aleatoria (x) considerando que otra que se conoce (y) tiene un monto específico. La aplicación de este método conlleva la suposición de una linealidad, cuando la demanda presenta un comportamiento creciente o decreciente. Por esta razón se hace indispensable que previo a la selección de este método, exista un análisis de regresión que determine la intensidad de las relaciones entre las variables que participan.

Descripción de la prueba

El análisis de regresión se define como «el estudio de la dependencia» o de cómo una respuesta o variable depende de uno o más predictores o variables independientes.

En este punto es importante aclarar que se considera regresión lineal simple si se relacionan sólo dos variables, de las cuales la dependiente es cuantitativa.

Cuando se utilizan dos o más variables para predecir una cuantitativa

se considera regresión lineal múltiple que sería un tema aparte.

Por medio de esta prueba se estará buscando una línea o función matemática que emita, de manera simple, una dependencia causal entre una variable independiente que será llamada (x) y otra dependiente que se llamará (y) y a continuación expresar matemáticamente esta respuesta mediante una ecuación.

Esta dependencia causal se puede expresar de dos maneras: puede ser una dependencia estadística, que consistiría en que ambas se relacionan fuertemente, pero las observaciones poseen un componente aleatorio que impiden que la nube de puntos aparezca exactamente distribuida a lo largo de una línea (recta o curva).

La dependencia funcional o exacta posee una línea o función matemática que indica exactamente cómo están relacionadas ambas variables.

El análisis de regresión consiste en obtener la línea «teórica» hacia la que tienden los puntos de un diagrama de dispersión.

Se trata, por tanto, de averiguar la dependencia estadística mediante la eliminación de los factores aleatorios.

Se podría citar como ejemplo que cuando se conoce la variable dependiente (cifras de tensión arterial) se podría determinar con base en este análisis la independiente (hipertrofia ventricular).

La prueba de regresión lineal busca dos objetivos, siendo el primero anticipar el valor que tomará la respuesta y además (segundo objetivo) cuantificar la precisión de esta predicción.

Para hacer un estudio de regresión lineal se requiere contar con una variable dependiente (y), la cual se debe obtener por muestreo (aleatorio), luego las dos, la independiente (x) y la dependiente (y) deben estar asociadas linealmente, de tal manera que la expresión gráfica de éstas debe visualizarse como una línea recta, así mismo se espera que la varianza de la variable dependiente (y) sea igual a pesar de tener diferentes valores de la variable independiente (x).

Fórmula

La representación más simple de una regresión lineal es en la siguiente ecuación $y = ax + b$ (Figura 1).

Se llama simple porque sólo hay una variable independiente involucrada, que vendría siendo « x ».

En esta ecuación:

y : es la variable dependiente o la que se busca predecir.

x : es la variable independiente o la que se usa para hacer una predicción.

a : es la pendiente o el valor que debe ser determinado, se le conoce como coeficiente y es una especie de magnitud de cambio que pasa por y cuando x cambia.

b : es la constante que debe ser determinada, se le conoce como intercepto porque cuando $x = 0$, entonces $y = b$.

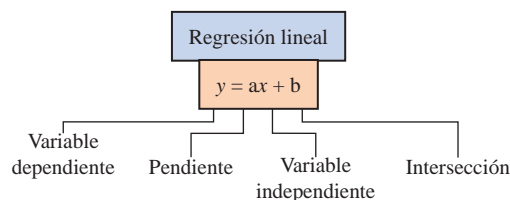


Figura 1: Ecuación que representa una regresión lineal simple.

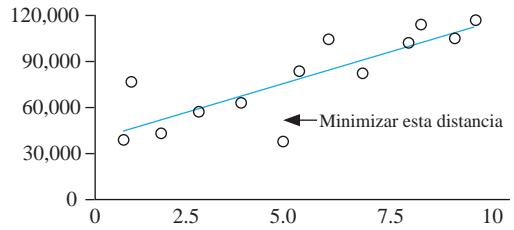


Figura 2: Gráfica donde se expresa mediante una línea de puntos la ecuación de una regresión lineal simple.

La expresión gráfica de la regresión lineal simple está representada por el diagrama de dispersión o diagrama de puntos (Figura 2).

Aquí la ecuación se representa por la línea recta que expresa la relación lineal entre el monto medio de y para los distintos valores correspondientes de x .

En una regresión lineal simple se busca minimizar la distancia vertical entre todos los datos y la línea, de tal manera que, intentando determinar la mejor línea, se debe minimizar la distancia desde todos los puntos y la distancia hacia dicha línea.

Existe más de un método que logra cumplir con este objetivo, pero todos estos métodos tienen un solo objetivo; minimizar la distancia.

Se puede lograr que el modelo de regresión encuentre la mejor línea de ajustes, utilizando el criterio de mínimos cuadrados para reducir el error, es una parte inevitable del proceso de predicción, no importa cuán poderoso sea el algoritmo que se elija, siempre habrá un error irreductible.

No es posible eliminar el error por completo, sin embargo, se puede intentar reducirlo al nivel más bajo usando la técnica antes mencionada.

La ventaja del error cuadrado sobre otras técnicas se basa en que utiliza una magnitud que tiene buenas propiedades matemáticas, por lo que es más sencillo diferenciar y calcular el descenso del gradiente.

Es fácil de analizar y computacionalmente más rápido, es decir, puede

aplicarse velozmente a conjuntos de datos que tienen miles de características y, por último, la interpretación es mucho más fácil que otras técnicas de regresión.

Ejemplo aplicado a la cirugía ortopédica

Un ejemplo práctico de esta técnica paramétrica sería predecir el tiempo de consolidación de una fractura de fémur (y) en un universo de pacientes adultos jóvenes con un tipo específico de fractura, diafisaria del tercio medio con trazo oblicuo largo con clasificación AO/OTA 32A2b (a) operados con la misma técnica: clavo centromedular bloqueado (x). En este ejemplo, de acuerdo con la ecuación expuesta previamente $y = ax + b$ (y), esta predicción podría establecerse si se aplica ese modelo a un programa estadístico computarizado creado con este fin, el cual reportaría seguramente una linealidad en el comportamiento del resultado, siempre y cuando se cumpla estrictamente con las condiciones establecidas para un análisis de regresión lineal.

Conclusiones

El análisis de regresión lineal simple se utiliza cuando se busca prever una variable en relación con la variación de otra, en la práctica, se identifica primero la existencia de una correlación significativa entre dos variables conti-

nuas y a continuación se realiza el análisis de regresión lineal para generar la ecuación que determina la variable de resultado (y) en relación con los valores de la predictora (x).

La regresión es una técnica paramétrica, por lo que hace suposiciones, sin embargo, se deben mantener las condiciones de modo que estas suposiciones no se violen pues si no se cuida esta condición, la regresión hace predicciones tendenciosas y erráticas, por lo que se debe tener en cuenta cuando se esté trabajando con este algoritmo.

Lecturas complementarias

1. Jiménez Avila J, Falavigna A. Education in research: from measurement to interpretation - surgical biostatistics. Asociación AOSpine Latin America; 2018.
2. Flores-Ruiz E, Miranda-Navales MG, Villasís-Keever MA. El protocolo de investigación VI: cómo elegir la prueba estadística adecuada. Estadística inferencial. Rev Alerg Méx. 2017; 64 (3): 364-370.
3. Calduch Cervera R. Métodos y técnicas de investigación internacional. Madrid: Universidad Complutense de Madrid; 2014. pp.100-102.
4. Montgomery DC, Peck EA, Vining GG. Introduction to linear regression analysis. Vol. 821. Hoboken: John Wiley & Sons; 2012.
5. Godfrey K. Simple linear regression in medical research. N Engl J Med. 1985; 313 (26): 1629-1636.
6. Novales A. Análisis de regresión. Madrid, España: Universidad Complutense de Madrid; 2010. p. 116.
7. Cardona Madariaga DF, González Rodríguez JL, Rivera Lozano M, Cárdenas Vallejo EH. Aplicación de la regresión lineal en un problema de pobreza. Interacción. 2014; 12: 73-84.

Unidad

4

Pruebas no paramétricas en la ortopedia

Distribución de Poisson

Drusso López Estrada,
Carla Lisette García Ramos

Definición

La de Poisson es una distribución de probabilidad discreta que, partiendo de una frecuencia de ocurrencia media λ , enuncia la probabilidad de que se presente un determinado número de eventos $k \in x$ durante un intervalo de tiempo dado o en una región específica. La probabilidad de que más de un resultado ocurra en un intervalo de tiempo tan corto o en esa región tan minúscula es imperceptible, por lo cual se le asigna el valor de cero.

La forma de Poisson describe la probabilidad como un acontecimiento fortuito, ocurrido en un tiempo o intervalo de espacio, suponiendo que la probabilidad de que dicha eventualidad suceda es muy pequeña, pero el número de intentos es muy grande, entonces el evento real aparece algunas ocasiones.

Fue llamada así en honor al astrónomo francés Siméon-Denis Poisson, quien la describió en 1837 dentro de su texto: *Recherchés sur la probabilité des jugements en matières criminelles et matière civile* (Investigación sobre la probabilidad de los juicios en materias criminales y civiles).

Se representa mediante la fórmula:

$$f(k, \lambda) = \frac{e^{-\lambda} \lambda^k}{k!}$$

Notación	Poisson (λ)
Parámetro	$\lambda > 0$
Distribución	$k = 1, 2, \dots$
Función de densidad de probabilidad (FDP)	$\frac{\lambda^k e^{-\lambda}}{k!}$
Función de distribución acumulativa (FDA)	$\sum_{i=1}^k \frac{\lambda^i e^{-\lambda}}{i!}$
Promedio	λ
Varianza	λ
Asimetría	$\lambda^{-1/2}$
Curtosis	λ^{-1}

Donde x sea una variable aleatoria que representa el número de eventos aleatorios independientes, que ocurren a una rapidez constante sobre el tiempo o el espacio. Se dice entonces que la variable aleatoria tiene una distribución de Poisson con función de probabilidad.

- k es el número de ocurrencias del evento o fenómeno (la función da la probabilidad de que el evento suceda precisamente k veces).
- λ es la media de éxitos por unidad de tiempo, área o producto.
- e es la base de los logaritmos naturales ($e = 2,71828\dots$).
- x señala un valor específico que la variable puede tomar (el número de éxitos que se desea que ocurran).
- Por definición, el valor esperado (media en el intervalo o región de interés) de una distribución de pro-

babilidad de Poisson es igual a la media de la distribución. $E(x) = \lambda$.

La varianza del número de eventos de una distribución de probabilidad de Poisson también es igual a la media de la distribución lambda, de este modo, la desviación estándar es la raíz cuadrada de lambda $V(x) = \lambda, \Omega = \sqrt{\lambda}$.

Tanto el valor esperado como la varianza de una variable aleatoria con distribución de Poisson son iguales a λ . Los momentos de orden superior son polinomios de Touchard en λ cuyos coeficientes tienen una interpretación combinatoria.

Según la fórmula de Dobinski, cuando el valor esperado de la distribución de Poisson es uno, entonces el n -ésimo momento iguala al número de particiones de tamaño n .

La moda de una variable aleatoria de distribución de Poisson con un λ no entero es igual a $[\lambda]$, el mayor de los enteros menores que λ . Cuando λ es un entero positivo, la moda es λ (Figura 1).

Las variables aleatorias de Poisson tienen la propiedad de ser infinitamente divisibles (Figura 2).

Propiedades

La variante de Poisson desempeña un papel importante, por derecho propio, como un modelo probabilístico apro-

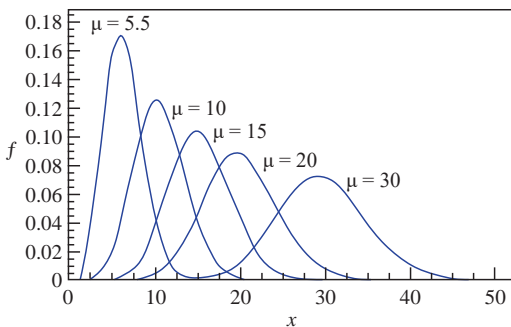


Figura 1: Distribución de Poisson con un λ no entero es igual a $[\lambda]$, el mayor de los enteros menores que λ . Cuando λ es un entero positivo, la moda es λ .

Distribución de Poisson

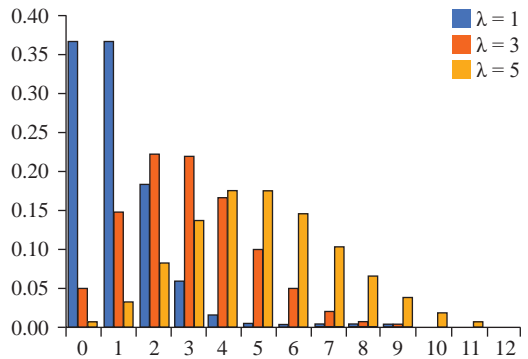


Figura 2: Gráfico de las variables aleatorias de Poisson tienen la propiedad de ser infinitamente divisibles.

piado para un gran número de fenómenos aleatorios.

Y los usos más comunes son:

1. Distribución de las llamadas telefónicas que llegan a un conmutador.
2. Demanda de los servicios en un hospital por parte de los pacientes.
3. Número de accidentes en un cruce.
4. Número de bacterias por cm^2 .

Ejemplo:

Los camiones que llegan a una terminal de transporte, y que se sabe que siguen un proceso de Poisson, con tasa de ocho camiones por hora, de modo que el número de llegadas por un periodo de horas es una variable de Poisson con parámetro.

$$\lambda = 8t$$

1. ¿Cuál es la probabilidad de que exactamente cinco camiones lleguen durante un periodo de una hora?
2. ¿Cuántos camiones se pueden esperar que lleguen durante 90 minutos?

Solución:

La llegada de los camiones a la terminal de transporte se distribuye según

Poisson. Sea x una variable que representa el número de camiones que lleguen a la terminal de transporte durante un periodo de tiempo t .

$$\lambda = 8 \text{ camiones} * \text{ tiempo} = 8 * 1 = 8$$

1. Se pide calcular la probabilidad de que lleguen cinco camiones durante una hora.

$$P(x = 5) = \frac{e^{-\lambda} \lambda^k}{k!} = \frac{e^{-8} 8^5}{5!} = 0.091$$

Por tanto, existe una probabilidad de 9.1% de que lleguen exactamente cinco camiones a la terminal durante una hora.

2. Se pide calcular la cantidad de camiones que podrían llegar en un tiempo de hora y media.

$$E(x) = \lambda = 8 * 1.5 = 12 \text{ camiones}$$

Por propiedad de la distribución de Poisson, $Var(x) = 12$. Con lo cual se tendría que la desviación estándar (DE), para este caso es $DE = \sqrt{12} \approx 3.46$.

Se espera entonces que lleguen, en una hora y media, en promedio 12 camiones a la terminal de transporte, con una desviación estándar de tres, esto quiere decir que, en realidad, se espera que lleguen entre nueve y 15 camiones.

Aproximación de la distribución binomial por una de Poisson

A fin de evitar calcular estas variantes binomiales se puede usar Poisson, pero debe cumplir con ciertas características, tales como $n > 30$, $np > 5$ o $nq < 5$; en los casos en que se cumplen tales condiciones, se podría sustituir la media de la forma binomial en lugar de la media de la distribución de Poisson de modo que $\lambda = np$.

Lecturas complementarias

1. Arroyo I, Bravo LC, Linas H, Muñoz FL. Distribuciones Poisson y gamma: una discreta y continua relación. *Prospect*. 2014; 12 (1): 99-107.
2. Palmisano A. *The encyclopedia of archaeological sciences*. USA: Ed. Wiley; 2018.
3. Kissell R, Poserina J. *Optimal sports math statistics and fantasy*. USA: Editorial Academic Press; 2017.

Odds ratio y riesgo relativo

José Manuel Pérez Atanasio,
Silvestre Fuentes Figueroa

Cuando se realiza una investigación se identifican variables que se consideran como un factor asociado para que se presenten ciertos resultados en los propios pacientes; por ejemplo, en pacientes con diabetes y obesidad el índice de infecciones posterior a una artroplastia total de rodilla parece ser más alta.

Se detecta entonces una posible asociación entre una variable predisponente (factor de riesgo) y un desenlace esperado. Pero inmediatamente surgen las siguientes preguntas: ¿Qué tan fuerte es esta asociación?, ¿se le puede medir matemáticamente? ¿En un futuro podría servir para predecir consecuencias similares en condiciones análogas?

A fin de contestar estas preguntas la bioestadística propone dos valores: el *odds ratio* y el riesgo relativo.

Estos indicadores estadísticos revelan el número de veces que se incrementa la posibilidad de observar un resultado si existe el factor de riesgo, evidentemente estos valores resultan de un análisis matemático y siempre son aproximaciones a la realidad, por lo que no expresan una certeza de ocurrencia de 100% y sólo son válidos si se reproducen en condiciones similares al estudio de donde se obtuvieron.

Descripción de la prueba

El *odds ratio* o razón de momios se obtiene mediante la tabla de 2×2 o de contingencia, que se elabora con los datos obtenidos de un estudio observacional en los que se analizan los factores de riesgo y presencia o no de la enfermedad, como se muestra a continuación (Tabla 1).

A fin de calcular el valor de OR, se aplica la siguiente fórmula:

$$\text{odds ratio} = (a \times d) / (c \times b)$$

El riesgo relativo se obtiene matemáticamente de manera similar, pero con diseño de estudios prospectivos, es decir, cuando el investigador observa un factor de riesgo y sigue de manera prospectiva a la población a ver si desarrollan o no la enfermedad.

Tabla 1: Tabla de 2×2 o de contingencia para obtener el cálculo de *odds ratio*.

Factor de riesgo	Presentan la enfermedad		
	Sí	No	
Expuestos	a	b	a + b
No expuestos	c	d	c + d
	a + c	b + d	a + b + c + d

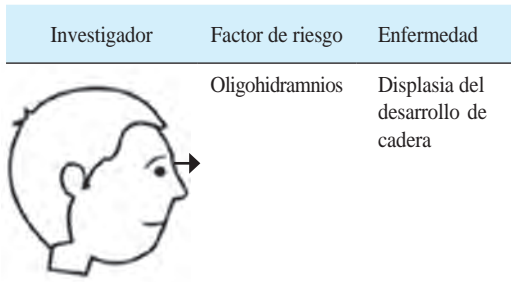


Figura 1: Posición del investigador en un estudio de diseño prospectivo.

Para poner como ejemplo un estudio en el que se analice la displasia del desarrollo de cadera, se puede identificar al oligohidramnios como un factor de riesgo en su desarrollo (Figura 1). En un ensayo prospectivo, el investigador estudia a los niños que durante el embarazo presentaron oligohidramnios y los evalúa periódicamente buscando identificar si se presenta la displasia del desarrollo de cadera durante la infancia; en el análisis estadístico de esta muestra se obtendrá un **riesgo relativo**.

Por otra parte, en un estudio retrospectivo el investigador analiza los casos que tienen displasia del desarrollo de cadera e indaga la presencia o no de oligohidramnios durante el periodo gestacional (Figura 2); como producto del análisis estadístico en este diseño se obtendrá un *odds ratio*.

Puntos importantes

El *odds ratio* y el riesgo relativo son estimaciones de asociación entre una variable de causa y una de resultado.

El riesgo relativo se calcula en estudios prospectivos como diseños de cohorte o ensayos clínicos controlados aleatorizados y el *odds ratio* en diseños retrospectivos como casos y controles.

El riesgo relativo toma valores desde cero y hasta uno, que podría ser interpretado en porcentaje, por lo que

se puede expresar en valores que van de cero a 100% de posibilidad de que ocurra la enfermedad.

Los valores del *odds ratio* van de cero hasta el infinito, entre más alto sea el valor, mayor será la probabilidad de ocurrencia de la enfermedad.

Ejemplo:

Presencia de oligohidramnios como factor de riesgo en el desarrollo de displasia del desarrollo de cadera, en un estudio retrospectivo de casos y controles, considerando casos aquellos pacientes que tienen displasia del desarrollo de cadera (Tabla 2).

Para calcular el *odds ratio* se sustituyen los valores de la Tabla 2 en la siguiente fórmula:

$$\text{Odds ratio} = (a \times d) / (c \times b) \\ (28 \times 53) / (7 \times 12) = 17.6$$

Esta prueba se puede realizar en cualquier programa informático de estadística, lo importante es comprender el contenido de la tabla de contingencia y el llenado adecuado de los valores en las casillas correspondientes.

Interpretación

En la muestra estudiada, el tener oligohidramnios incrementó **17.6** veces la posibilidad de desarrollar displasia del desarrollo de cadera.

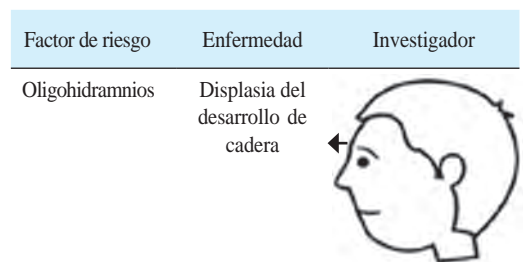


Figura 2: Posición del investigador en un diseño de estudio retrospectivo.

Tabla 2: Tabla de contingencia en un estudio hipotético en el que se analizaron 100 pacientes de manera retrospectiva intentando identificar el oligohidramnios como factor de riesgo para presentar displasia del desarrollo de cadera.

Factor de riesgo (oligohidramnios)	Presentan la enfermedad (displasia del desarrollo de cadera)		
	Sí	No	
Expuestos	28	7	35
No expuestos	12	53	65
	40	60	100

Aplicación en la Ortopedia

El *odds ratio* es una de las pruebas estadísticas de mayor fortaleza en la identificación de los factores de riesgo asociados con la presencia de una enfermedad. Por la facilidad de realizarla en estudios retrospectivos es muy utilizada en tesis de alumnos que no tienen el tiempo suficiente para el seguimiento de pacientes en estudios prospectivos como una cohorte, siempre que se identifique alguna complicación de una cirugía, en los resultados postquirúrgicos o en la evolución de los pacientes; se presta para estudiar las posibles causas que pudieron generarlas y el *odds ratio* facilita dar un valor matemático a esta probabilidad.

Conclusión

El *odds ratio* y el riesgo relativo representan el número de veces que se incrementa la posibilidad de padecer una enfermedad si existe el factor de riesgo involucrado. El *odds ratio* es útil en diseños de estudio retrospectivo y el riesgo relativo en los estudios de tipo prospectivo.

Lecturas complementarias

1. Cerda J, Vera C, Rada G. Odds ratio: aspectos teóricos y prácticos. Rev Méd Chile. 2013; 141 (10): 1329-1335.
2. Bland JM, Altman DG. Statistics notes. The odds ratio. BMJ. 2000; 320 (7247): 1468.
3. Pérez AJM. Principios de bioestadística para comprender los resultados reportados en los artículos científicos. Ortho-tips. 2015; 11 (2): 91-95.
4. Falavigna A, Jiménez Avila JM. Education in research: from the idea to the publication. Caxias do Sul, Brazil: Educs; 2015.
5. Jiménez Avila J, Falavigna A. Education in research: from measurement to interpretation - surgical biostatistics. Asociación AOSpine Latin America; 2018.
6. Toledo-Infanson V, Jiménez Avila JM. Publicación con impacto en ortopedia y traumatología. México: Graphimedica S.A. de C.V.; 2019.
7. Cálculo de Odds ratio en línea; https://www.medcalc.org/calc/odds_ratio.php.

Unidad

5

¿Qué debe saberse cuando se tienen los datos?

Elaboración de una revisión sistemática y metaanálisis

Marlene Vanessa Salcido Reyna

Las revisiones sistemáticas son investigaciones científicas en las que la unidad de análisis son los ensayos originales primarios, a partir de los cuales se pretende contestar una pregunta de investigación claramente formulada mediante un proceso sistemático y explícito, por eso se les considera una investigación de fuente secundaria (investigación sobre lo investigado).

Por el contrario, a las revisiones que no siguen un proceso sistemático, denominadas revisiones narrativas, no se les puede considerar un proceso formal de investigación, sino simplemente un formato de literatura científica basada sobre todo en opiniones.

Desde un punto de vista formal, las revisiones sistemáticas sintetizan los resultados de investigaciones primarias mediante estrategias que limitan el sesgo y el error aleatorio.

La revisión sistemática es una herramienta de síntesis de información y no siempre es posible presentar resumidamente los productos de los estudios primarios, cuando éstos no se combinan estadísticamente, la revisión se denomina revisión sistemática cualitativa, por el contrario, una revisión sistemática cuantitativa conlleva un metaanálisis.

El metaanálisis es el término utilizado para el método estadístico de com-

binación de los resultados de más de un estudio con el propósito de estimar un efecto promedio o común entre los estudios. Una revisión sistemática puede resumir los resultados con o sin un metaanálisis. Un metaanálisis se puede presentar sin una revisión sistemática, de conjunto de 2 o más estudios.

No debe confundirse una revisión sistemática con un metaanálisis, la primera siempre es posible, es recorrer un camino con tal de llegar a un objetivo donde se debe tener claro cuál es la meta y cuál es la ruta, de tal manera de que si alguien desea hacer una revisión del mismo tema con el mismo objetivo, obtenga exactamente el mismo desenlace, desde la formulación de la pregunta, la búsqueda de las palabras clave, los criterios de inclusión y de exclusión, tipo de artículos y sobre todo las base de datos utilizadas, sin olvidar la fecha en que se realiza tal búsqueda.

En la revisión sistemática, un punto muy importante es que debe ser reproducible.

En el metaanálisis, a fin de que se realice, se requieren artículos que tengan las mismas características, es decir que cumplan las condiciones para realizarlo, los artículos deben poseer características similares en cuanto al

tipo de población, utilizando las mismas escalas de valoración o que en su defecto se puedan homologar, y así tener resultados comparables.

Originalmente, Glass definió al metaanálisis como el «análisis de una colección amplia de análisis de conclusiones de estudios individuales, con el objeto de integrar los hallazgos» y establecer una conclusión basada en diferentes exámenes que contienen información, metodología y resultados similares.

El metaanálisis se realiza con el objetivo de cuantificar el efecto de una intervención y su incertidumbre, incrementar la potencia o precisión de los resultados, explorar las diferencias entre los estudios individuales, resolver controversias para estudios contradictorios o generar nuevas hipótesis.

A veces, la información médica sobre algún tema específico es tan poco clara y contradictoria, que en ocasiones el profesional de la salud no tiene el tiempo o la orientación para poder analizarla en su totalidad y entonces aprovecharla en su real magnitud.

La revisión sistemática y el metaanálisis son herramientas que permiten resumir la información de una manera organizada, y verificada, es decir, con una evaluación de la validez interna (sesgo) y utilizadas de manera correcta son un auxiliar del clínico en la toma de decisiones y en la realización de medicina basada en la evidencia.

Cómo se elabora un metaanálisis

Se debe iniciar con la redacción de un protocolo, originado en una pregunta de investigación basada en una idea, no en datos ni en imágenes, debe contener un objetivo directo, para después continuar con la ejecución del estudio y finalizar con un reporte, artículo o informe de investigación.

En general, el proceso debe contemplar los siguientes cinco aspectos:

1. Formulación de la pregunta.
2. Planteamiento del problema.
3. Formulación de objetivos y una hipótesis.
4. Metodología del estudio:
 - a. Estrategias de búsqueda de la literatura.
 - b. Criterios de inclusión/exclusión de los artículos.
 - c. Registro de los datos de cada estudio y su análisis de calidad.
 - d. Análisis cuantitativo de los trabajos revisados.
5. Ejecución del estudio y elaboración del reporte o informe de investigación metaanalítico a manera de conclusión.

Utilidad de los metaanálisis

El metaanálisis, desde sus orígenes, ha sido una técnica controvertida, sobre todo cuando se ha limitado a un cálculo del efecto global, dado que la metodología estadística no puede «corregir» o «ajustar» las discrepancias entre estudios y mucho menos lograr que resultados contradictorios se combinen en una única medida, sin embargo, un correcto metaanálisis podría ayudar a la interpretación de algún tipo de problema clínico.

La principal utilidad de los metaanálisis es generar una hipótesis, en ocasiones se utilizan para fundamentar la toma de decisiones, amparados en la corriente de la denominada **medicina basada en la evidencia**.

En cualquier caso, el metaanálisis es una herramienta útil en la síntesis del gran volumen de información disponible e incrementar la objetividad a la hora de sistematizar el conocimiento actual, permitiendo la búsqueda de explicaciones a situaciones contradictorias.

Algo importante a tener en cuenta, es que no tiene sentido efectuar un metaanálisis sobre aspectos en los que existe una certeza casi absoluta, por eso será siempre objeto de controversia,

aunque indudablemente constituye la mejor herramienta posible en un mundo cada vez más desbordado por el volumen de información que requiere ser explorada y sintetizada, recordando que habrá ocasiones en que algunos estudios clínicos no soporten la prueba del tiempo, por lo que al ser analizados en el tiempo actual ya no aportarán alguna información relevante.

Limitaciones de los metaanálisis

Existen tres aspectos fundamentales sobre la crítica a los trabajos de metaanálisis:

1. Son estudios secundarios que dependen de los estudios previos. Los revisores no toman en cuenta la calidad de los artículos analizados y se combinan «buenos» con «malos».
2. Enfatiza los sesgos de comparación destemporal y disespacial que, por lo demás son comunes a todos los ensayos multicéntricos y a los que utilizan controles externos históricos.
3. Sesgo de publicación (existen propuestas metodológicas y estadísticas hacia la reducción del sesgo de publicación, algunas se encuentran en proceso de perfeccionamiento).

Muchos de los artículos publicados no incluyen la información necesaria y exacta para realizar un metaanálisis adecuado, lo cual dificulta la unificación de las variables y de los resultados obtenidos a la hora de compararlos entre todos los estudios.

Cochrane y las revisiones sistemáticas

En 1979, el médico británico Archie Cochrane escribió: «Una gran crítica a la profesión, es que no se tengan sumarios clínicos que aparezcan y se actualicen de manera periódica, organizados por especialidad o subespecialidad, de todos los ensayos clínicos que al momento existen», bajo esa visión,

en el año 1993 se fundó la colaboración que ahora lleva su nombre.

En los últimos 20 años, Cochrane ha formado parte de la manera de cómo se toman decisiones en el área de la salud, ya que se ha encargado de recopilar y resumir revisiones sistemáticas en forma, obteniendo datos más confiables derivados de las investigaciones.

Es información de alta calidad y proporciona una herramienta poderosa que logra potenciar los conocimientos y toma de decisiones.

Cochrane es una red internacional e independiente de investigadores profesionales, más de 40,000 colaboradores de más de 140 países, aproximadamente, donde de manera coordinada producen información fiable y accesible, esto sin un apoyo comercial y sin conflicto de interés, representando un modelo de referencia internacional de información fiable y de calidad.

Conclusión

Las revisiones sistemáticas y los metaanálisis tienen un papel relevante en el resumen y la evaluación de validez para conocer un efecto, potenciar o precisar resultados del conocimiento científico y, en consecuencia, se han convertido en un potente instrumento que ayuda en la toma de decisiones que los profesionales de cualquier ámbito tienen que tomar constantemente para optimizar su práctica profesional.

Es por ello que los metaanálisis han posicionado la **práctica basada en la evidencia**, como una herramienta metodológica imprescindible que logra mostrar cuáles son las mejores evidencias o pruebas científicas, acerca de cualquier problema o pregunta social, educativa o clínica.

Al mismo tiempo, alrededor de los estudios primarios y de la investigación en general, la evaluación de la calidad metodológica es una cuestión fundamental que trasciende a las re-



Figura 1: Tipos de evidencia científica.

visiones sistemáticas y al metaanálisis, traslapadas en las propias raíces del método científico.

El número de publicaciones aumenta exponencialmente, y los médicos clínicos y quirúrgicos difícilmente pueden tener acceso al total de los artículos que se publican anualmente sobre su área de trabajo.

Existen limitaciones que los artículos deficientemente elaborados imponen al desarrollo de un metaanálisis, por lo tanto, los comités de las revistas científicas han integrado elementos particulares, con el objeto de perfeccionar los estándares editoriales, siendo cada vez más exigentes en la calidad de las investigaciones publicadas, particularmente en este tipo de manuscritos.

Cabe destacar que según la clasificación de Shekelle, los metaanálisis se encuentran clasificados como trabajos con un nivel de evidencia 1a, encontrándose en la punta de la pirámide (*Figura 1*).

Según la propuesta de la *Scottish Intercollegiate Guidelines Network* (SIGN) y la *Agència d'Avaluació de Tecnologia Mèdica* (AATM), el metaanálisis es considerado dentro de la evidencia más alta (evidencia 1) y su sola

presencia genera grados de recomendación tipo A.

En una conferencia en diciembre de 1997, Sackett, padre de la medicina que se basa en la evidencia expresaba: «Un médico que pretenda diagnosticar y tratar a su paciente con aquello que verdaderamente ha demostrado su valor, necesita leer, todos los días del año, 19 artículos científicos. Algo que es claramente una utopía».

Lecturas complementarias

1. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to meta-analysis*. Chichester, UK: John Wiley & Sons; 2009.
2. Higgins JPT, Green S (Eds.). *Cochrane handbook for systematic reviews of interventions*. Chichester, UK: Wiley-Blackwell; 2008.
3. Konstantopoulos S, Hedges LV. Analyzing effect sizes: fixed-effects models. In: Cooper H, Hedges LV, Valentine JC (Eds.). *The handbook of research synthesis and meta-analysis*. 2nd ed. Nueva York: Russell Sage Foundation; 2010. pp. 279-293.
4. Clark Peralta P. Meta-análisis. En: Moreno Altamirano L. *Epidemiología clínica*. México: Interamericana McGraw-Hill; 1994. pp. 201-215.
5. Dawson-Saunders B, Trapp RG. *Bioestadística médica*. México: Manual Moderno; 1997.
6. Jones DR. El meta-análisis en los estudios epidemiológicos observacionales. *OPS, Bol Ofic Panam*. 1993; 115 (5): 438-445.
7. León Hernández SR, Aguilera Zepeda JM. Metodología de los estudios de metaanálisis en la investigación clínica. *Rev Mex Ortop Traum*. 2001; 15 (2): 86-91.
8. Shekelle PG, Woolf SH, Eccles M, Grimshaw J. Clinical guidelines: developing guidelines. *BMJ*. 1999; 318 (7183): 593-596.
9. Scottish Intercollegiate Guidelines Network (SIGN). *Rev Dic*. 2011.
10. Jovell AJ, Navarro Rubio MD. Evaluación de la evidencia científica. *Med Clin (Barc)*. 1995; 105: 740-743.
11. Higgins JPT, Green S. *Manual Cochrane de revisiones sistemáticas de intervenciones*. March 2011. Disponible en: www.cochrane-handbook.org

Unidad

5

¿Qué debe saberse cuando se tienen los datos?

La fuerza de la evidencia

Víctor Toledo Infanson

Cada estudio que se revisa o se lee y que da una recomendación que atañe al quehacer médico propio, queda sujeto con base al rigor científico de su diseño, a que se le dé un valor de clasificación de evidencia a su contenido.

Existen muchas escalas para clasificar el grado de evidencia de lo que se publica y todas las escalas son muy parecidas aunque con sus particularidades.

La primera escala conocida se utilizó en 1979 por la *Canadian Task Force* (CTF) o Fuerza de Tarea Canadiense, en la valoración periódica de salud que realiza para evaluar medidas de prevención de la salud (www.canadiantaskforce.ca).

Posteriormente, en 1984, esta prueba fue adaptada por la Fuerza de Tarea Preventiva de los EE. UU. y, aunque ha sido actualizada en varias ocasiones y ha sido juzgada por críticos y seguidores, ha logrado mantenerse vigente a lo largo de todo este tiempo.

Según la clasificación de la CTF (*Tabla 1*), la fuerza de la evidencia relacionada con un diagnóstico o con una opción terapéutica en particular depende de los datos disponibles, de tal manera que:

cuando hay cifras procedentes de múltiples ensayos clínicos aleatorizados o de un metaanálisis el nivel de evidencia es **alto o A**,

cuando hay datos de un ensayo clínico aleatorizado único o estudios no aleatorizados, el nivel de evidencia es **intermedio o B** y cuando únicamente se trata de un consenso (opinión de expertos), de casos o métodos de asistencia estándar, se considera un nivel de evidencia **bajo o C**.

Por otro lado, las recomendaciones se basan en la evidencia y se clasifican de la siguiente manera:

Clase I: hay evidencia o bien acuerdo general respecto a que el procedimiento o el tratamiento en su caso resulta beneficioso, útil y eficaz.

Clase II: en este caso la evidencia es contradictoria o bien existen opiniones divergentes respecto a la utilidad o la eficacia de un procedimiento o de un tratamiento.

La clase II se divide en dos subtipos a y b; en la clase II a, el peso de la evidencia o en su caso de la opinión es favorable a la utilidad y eficacia de un tratamiento o procedimiento.

En la clase II b la utilidad o la eficacia de un manejo no está bien establecida por la evidencia.

La clase III es aquella en la que hay evidencia de que lo que se propone no es útil o que incluso puede causar algún perjuicio.

Tabla 1: Niveles de evidencia y fuerzas de la recomendaciones clínicas.

	Clase I Beneficio muy superior al riesgo Se debe proceder	Clase IIa Beneficio superior al riesgo Es razonable proceder	Clase IIb Beneficio parecido al riesgo Se puede proceder	Clase III Beneficio menor al riesgo No proceder
Nivel A	Es efectivo	Recomendaciones a favor	Recomendaciones menos establecidas	No útil o daño
Múltiples estudios poblacionales. Consistencia resultados	Evidencia de múltiples estudios aleatorizados o metaanálisis	Algún conflicto de múltiples estudios aleatorizados o metaanálisis	Conflicto entre múltiples estudios aleatorizados o metaanálisis	Evidencia de múltiples estudios aleatorizados o metaanálisis
Nivel B	Es efectivo	Recomendaciones a favor	Recomendaciones menos establecidas	No útil o daño
2-3 estratos de riesgo evaluados	Limitada evidencia de un estudio aleatorizado o varios no aleatorizados	Algún conflicto de un estudio aleatorizado o no aleatorizado	Conflicto entre un estudio aleatorizado o no aleatorizado	Limitada evidencia de un estudio aleatorizado o varios no aleatorizados
Nivel C	Es efectivo	Recomendaciones a favor	Recomendaciones menos establecidas	No útil o daño
Muy pocos estratos analizados	Opinión de expertos, casos retrospectivos	Algún conflicto de experto o casos retrospectivos	Conflicto de expertos o casos retrospectivos	Opinión de expertos, casos retrospectivos

Modificado de: Fuster V, Rydén LE, Cannom DS, Crijns HJ, Curtis AB, Ellenbogen KA, et al. 2011 ACCF/AHA/HRS focused updates incorporated into the ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation: a report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. *Circulation*. 2011; 123 (10): e269-e367.

Es importante aclarar que el hecho de que una recomendación se clasifique como clase I, no la hace necesariamente la primera opción de manejo o que deba implementarse siempre que se presente una situación en particular, simplemente significa que es una opción adecuada y que se considera mayoritariamente útil.

A fines del siglo pasado, entre 1970 y 1980, David Sackett, David Eddy y Archie Cochrane señalaron la necesidad de reforzar la práctica empírica de la medicina y establecieron las que posteriormente serían las primeras reglas de evidencia con las cuales poder fundamentar las decisiones clínicas con este parámetro.

El término medicina basada en la evidencia (MBE) se introdujo en 1991, enfocado en la formación de los médicos de primer contacto, procurando evaluar la credibilidad de la evidencia de investigación, comprender los re-

sultados de los estudios clínicos y determinar la mejor manera de aplicarlos a su práctica diaria.

Posteriormente, la publicación de la guía detallada en revistas y libros de texto científicos, complementada con herramientas populares como la de evaluación gráfica para epidemiología, dio como consecuencia que la MBE se integrara cada vez más en los planes de estudio médicos en todo el mundo.

Entre los procesos de evolución del nivel de evidencia (NDE), desde el propuesto por la Fuerza de Tarea Canadiense y después de haber sobrevivido a la crítica, el Centro para la Medicina Basada en Evidencia o CEBM, por sus siglas en inglés, en el año 2000 implementó algunas medidas que aceleraron el proceso de localización de evidencias por el médico clínico.

En aquella época, un problema común de la mayoría de los esquemas



Figura 1: Pirámide de grados de comprobación.

de ordenamiento de evidencias era que el puntaje de las evidencias se refería solamente a terapia y prevención, sin incluir pruebas diagnósticas, marcadores pronósticos o daños.

El equipo del CEBM actualizó el NDE proponiendo la inclusión de diferentes aspectos de problemas de salud, incluso terapia, prevención, etiología, pronóstico, diagnóstico, diagnóstico diferencial y análisis económico y de decisión.

Todo esto permitió el desarrollo de la medicina con base en evidencias, como se le conoce en la actualidad, y su desarrollo permite acceder a herramientas de búsqueda de información científica que ha sido sujeta de análisis exhaustivos y revisiones sistemáticas que, en su momento, habilitan su puesta en práctica en la clínica, con la consideración de la individualización del propio universo de práctica profesional.

Conclusión

De las publicaciones a las que se tiene acceso, la fuerza de la evidencia se construye con base en el nivel y grado de recomendación del mismo, sien-

do los ensayos controlados y aleatorizados y los metaanálisis los de mayor jerarquía en la pirámide de grado de recomendación (*Figura 1*), pudiendo afirmarse así que las exploraciones observacionales están por abajo en este sentido de los estudios experimentales.

Se debe tener presente que la medicina basada en la evidencia no está reñida con la práctica clínica ni con las necesidades de los pacientes pues, sin la experiencia clínica, la práctica médica peligra de ser tiranizada por la evidencia eterna.

Lecturas complementarias

1. Enríquez F, Jiménez A. Niveles de evidencia y fuerza de las recomendaciones. *Cir Cardiov*. 2010; 17 (2): 123-125.
2. Fuster V, Rydén LE, Cannon DS, Crijns HJ, Curtis AB, Ellenbogen KA, et al. 2011 ACCF/AHA/HRS focused updates incorporated into the ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation: a report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. *Circulation*. 2011; 123 (10): e269-e367.
3. U.S. Preventive Services Task Force Ratings: Grade Definitions. *Guide to Clinical Preventive Services*. 3rd ed. Periodic Updates, 2000-2003. Agency for Healthcare Research and Quality, Rockville, MD. Available in: <http://www.ahrq.gov/clinic/3rduspstf/ratings.htm>
4. Lizaur A. Cirugía ortopédica y traumatología basada en la evidencia: características y criterios de evaluación de los estudios de investigación. *Revista Española de Cirugía Ortopédica y Traumatología*. 2009; 53 (4): 261-270.
5. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. *Lancet*. 2017; 390 (10092): 415-423. doi: 10.1016/S0140-6736(16)31592-6.
6. Toledo-Infanson V, Jiménez Avila JM. Publicación con impacto en ortopedia y traumatología. México: Graphimedic S.A. de C.V.; 2019.

Unidad

5

¿Qué debe saberse cuando se tienen los datos?

Sesgos y enigmas de la bioestadística: reflexiones

José Luis Osma Rueda

La investigación, en su desarrollo o en sus procesos, requiere de un manejo básico en metodología, no obstante, los fundamentos bioestadísticos se deben conocer o incorporar por los investigadores en el análisis de sus datos para no terminar en resultados sesgados por el abuso o mal uso de las pruebas bioestadísticas; asimismo, no debe olvidarse el pensamiento de Miguel de Cervantes: «Estar preparado es la mitad de la victoria», que de manera filosófica o moral se complementa con el de Séneca: «Todo lo honesto es difícil», y la investigación y el análisis estadístico de sus datos también lo son.

Los sesgos (falta de exactitud) pueden ser sistemáticos o aleatorios y estos sesgos afectan la validez interna y la precisión, respectivamente, los productos de una investigación.

Por lo tanto, es preciso formarse en bioestadística o asesorarse de bioestadísticos, a fin de que los resultados obtenidos sean producto de un proceso integral de investigación (metodología-bioestadística), para obtener la mejor validez interna y externa de esos ensayos.

Guías de informes para diseños epidemiológicos

Un «sesgo», que cada vez es menor, se logra si se conocen las guías de publicación (*reporting guidelines*), pues en cada una de ellas hay apartados que indican parámetros bioestadísticos básicos a tener en cuenta en cada uno de los diseños metodológicos que existen, para responder a la pregunta de investigación (diagnóstico-pronóstico-intervención), facilitando la claridad y transparencia de las publicaciones.

Entre otras, las más utilizadas son: STROBE (*Strengthening the Reporting of Observational Studies in Epidemiology*) para estudios observacionales descriptivos y analíticos, CONSORT (útil en Ensayos Clínicos Aleatorizados [ECA]), PRISMA apropiada en revisiones sistemáticas (RS) de ECA y metaanálisis, y finalmente, para pruebas diagnósticas el STARD.

Es factible consultar las anteriores guías en la red EQUATOR (*Enhancing the Quality and Transparency Of health Research*); disponibles en: <http://www.equator-network.org>. Sumado a lo anterior, Lang y Altman elaboraron una guía en relación con los

métodos estadísticos y presentación de los resultados SAMPL (*Statistical Analyses and Methods in the Published Literature*) que también se encuentra en EQUATOR.

Sesgos al obtener y analizar las variables

Es importante recordar en qué tiempo se estudian las variables, por ejemplo, los estudios de predicción e intervención requieren ser analizados en el tiempo futuro; para estudios de capacidad diagnóstica, en los cuales dos de ellas se estudian al mismo tiempo, se entiende que son estudios transversales. Si una variable precede a otra, el estudio es longitudinal y en los estudios de intervención ésta será la causa.

El tener en cuenta el tiempo es muy importante en las pruebas diagnósticas (validez de criterio), de acuerdo al tipo de muestreo propuesto por Kramer: en el muestreo de tipo transversal a todos se les realiza la prueba y la prueba diagnóstica (estándar de oro) obteniendo de manera **insesgada**: sensibilidad (S), especificidad (E), valor predictivo positivo (VPP), valor predictivo negativo (VPN), LR (+), LR (-), el valor Q (nivel de la prueba), la prevalencia.

Si se utiliza un corte transversal para un análisis predictivo se cometerá un sesgo, lo indicado es hacer un muestreo prospectivo que consiste en aplicar la prueba a todos los sospechosos del diagnóstico (n_1) y, posteriormente, a una muestra aleatoria de los positivos y negativos (n_2), se le aplica el estándar de oro, asimismo, en este muestreo prospectivo o predictivo sólo se puede obtener la prevalencia de la prueba (Q) que es necesaria para calcular las probabilidades (S, E, VPP y VPN).

Además, y quizás un «enigma» y/o «mito» que ha llevado a un sesgo, es el de considerar que la sensibilidad y la especificidad son constantes, pues no

lo son, debido a que las personas y sus estados clínicos tienen variación.

La terminología es difícil de entender en las variables de desenlace, la diferencia entre *endpoint* (efecto final o inmediato de un proceso) y *outcome* (desenlace o resultado medido); como ejemplo: el resultado radiológico posterior al manejo quirúrgico de fractura de húmero proximal sería un *endpoint* y la evaluación final del paciente con esta fractura, aplicando una escala, sería el *outcome* con malo, bueno o excelente resultado, de acuerdo con la puntuación de escala validada utilizada.

Siempre que se comparan dos variables debe haber una hipótesis planteada, y preferiblemente con **la mínima diferencia del tamaño del efecto para ser relevante clínicamente**, lo cual genera un proceso analítico que debe ser contrastado a fin de aceptar o rechazar la hipótesis (de una o dos colas); por lo tanto, se debe plantear un cálculo de muestra que en los estudios descriptivos se debe especificar **cómo se obtuvo** (STROBE) y en los ECA (CONSORT) **cómo se calculó**.

Además, la **calidad** de la muestra es más importante que la dimensión de la muestra (n), no obstante, los intervalos de confianza dependen del tamaño de la muestra y obviamente los valores de «p», pero siempre se debe tener un cálculo de muestra racional y ético, evitando caer en «el azote de los datos». Para obtener valores estadísticamente significativos ($p < 0.05$), o en el aumento de la muestra cayendo en la moda «*p-hacking*», origen de innumerables falsos negativos (error tipo I), reportados en la literatura.

Del sesgo de publicación (sólo se publican resultados estadísticamente significativos), y que han dado origen a la «*p-hacking*», es pertinente comentar lo expresado por Bobenrieth Aste en su artículo: *Mitos y realidades en torno a la publicación científica*.

En su **mito 4** dice: «el investigador científico, por la misma naturaleza obje-

tiva de su quehacer, está libre de prejuicios y de sesgos. Un prejuicio del investigador, si lo hubiera, no afecta el proceso de investigación ni su comunicación, **en el supuesto de que la metodología hay sido adecuada y correcta».**

Realidad 4. «El compromiso intelectual y emocional hacia un conjunto particular de consecuencias anticipadas puede conducir a errores graves. El comprometerse a explicar que una variable explicatoria sea la causa de la modificación en la de resultado podría sesgar en forma importante los corolarios de una investigación.»

En el análisis de las variables se debe evaluar qué metodología estadística o ensayo estadístico podría aportar un análisis no sesgado, antes que pensar en el aumento de la muestra; para finalizar, siempre recordar que el tamaño del efecto, intervalos de confianza (IC de 95%) y la relevancia clínica priman ante los valores de «p» que frecuentemente son presentados en los artículos, algunos de ellos con la «p huérfana», por no tener el «apellido» que es la prueba estadística de donde se obtuvo el valor de «p».

Un ejemplo de «p» huérfana podría ser: en una muestra 20 pacientes (no independientes), la media del DASH antes de la reconstrucción capsular superior en el hombro fue de 30 puntos (IC de 95% 20 a 40), y posterior a dicha reconstrucción fue de 5 (IC de 95% 2 a 7), respectivamente, con una $p = 0.03$ (huérfana).

Es un análisis de antes y después de una variable continua y, en este caso, las pruebas a utilizar serían una t pareada (paramétrica) si las variables continuas son normales, o una no paramétrica (como la de rangos de Wilcoxon) si las variables no son normales; al enunciar el valor de $p = 0.03$ (t Student pareada) o $p = 0.03$ (test Wilcoxon) deja de ser «huérfana» y se puede abstraer que el investigador realizó prueba de normalidad. Si no la realizó y las variables no son normales y utilizó

la t Student pareada hay un sesgo por utilizar la prueba no adecuada en este ejemplo de diferencia de medias de un estudio antes y después, debiendo haber utilizado la prueba no paramétrica de rangos de Wilcoxon.

El valor de «p» no es calidad, no es evidencia, no resuelve la incertidumbre clínica; es un concepto matemático y su abuso o mal uso ha llevado a conclusiones erróneas, olvidando que es el método quien soporta el alcance y validez de los hallazgos.

El análisis estadístico de las variables cualitativas (nominales-categorías-ordinales) también presenta sesgos. Al realizar pruebas de independencia entre dos variables nominales (binarias) la chi cuadrada (χ^2) no está indicada cuando la muestra es menor de 20, o alguna de las frecuencias esperadas es menor a cinco, y en este caso debe utilizarse la prueba estadística exacta de Fisher, de forma que ese valor de «p» no sea sesgado.

En resumen, se deben conocer las características de las variables y qué prueba estadística utilizar (paramétrica o no paramétrica) para obtener un valor de «p» insesgado que indique si se acepta o no la hipótesis planteada, sin olvidar el contexto de la relevancia clínica en la toma de decisiones.

En los estudios de fiabilidad, sinónimo de confiabilidad, hay confusión en la terminología empleada en la literatura en inglés (*agreement* = repetibilidad, *reliability* = confiabilidad, *reproducibility* = reproducibilidad, *concordance* = acuerdo), al no tener claros sus significados, algunas investigaciones en sus publicaciones no tienen el título adecuado, por ejemplo titulan *agreement* (repetibilidad) y de lo que se trata el estudio es de reproducibilidad (*reproducibility*), quizás en este momento se entra en más incertidumbre, pero leyendo y analizando esta terminología cada vez será más claro el concepto.

Para efectos prácticos, la confiabilidad como concepto general se compone de tres aspectos metodológicos: consistencia interna (*internal consistency*), reproducibilidad (*reproducibility*) y acuerdo (*agreement*).

Al igual que las pruebas estadísticas específicas que exploran asociación entre variables, en los estudios de fiabilidad se deben tener en cuenta las características de las mismas, frecuentemente se comenten errores al analizar la confiabilidad intra o interobservador (reproducibilidad) al no considerar si la variable es continua o cualitativa (dicotómica-categoría-ordinal).

En el caso de las continuas se utiliza el coeficiente de correlación intraclass (CCI), nunca utilizar el coeficiente de correlación de Pearson pues éste no mide reproducibilidad; el kappa (k de Cohen) para las nominales dicotómicas (sí/no, por ejemplo), el sesgo aparece cuando la variable es categórica u ordinal y se utiliza el kappa y no es correcto, se debe utilizar el kappa ponderado.

Sesgos de interpretación de los análisis estadísticos

La prevalencia afecta en los estudios diagnósticos el VPP, al igual que en los estudios de fiabilidad la concordancia, por lo tanto, el no tener presente la prevalencia es un sesgo o error en el análisis de los desenlaces en estos dos estudios. Asimismo, el tomar la correlación lineal (r) de Pearson (variables continuas), por ejemplo, como medida de concordancia o acuerdo es erróneo.

En los resultados siempre se debe describir la asociación de interés y si ha realizado un análisis de regresión (logística, lineal múltiple, Cox, binomial), ANOVA o sobrevida, de igual forma se debe referir por qué se realizó, el no hacerlo es un error porque contribuye a la no interpretación de los productos de la investigación.

Los diferentes tipos de sesgos, el azar y las variables de confusión deben siempre analizarse y tenerse en cuenta en el análisis en todos los estudios como explicación posible de la asociación estadística, por esa razón el pensamiento de M. Susser es pertinente en el análisis por causalidad: «cuando hay minas por todas partes, uno no debe aventurarse sin un detector de minas», este detector de minas es el análisis multivariante, además, el único sesgo aparte de su control en el diseño metodológico como en los ECAs es el de confusión, el cual se podría controlar también en la fase de análisis estadístico.

Para finalizar, en una investigación, el comienzo de los sesgos que generarán incertidumbre metodológica inicia con una confusa pregunta de investigación: «si no puede explicar algo en forma sencilla, es que no lo ha entendido suficientemente». Albert Einstein.

Ante la pregunta de «si existe asociación entre las variables ¿es real o ficticia?», la respuesta dependerá del análisis de la lógica subyacente a las pruebas estadísticas empleadas, el método y contexto biológico; por lo tanto, la toma de decisiones debe tener un análisis racional de la relevancia clínica del resultado, habiendo evitado el error sistemático (validez) y mejorando la precisión buscando disminuir el error aleatorio (azar).

Lecturas complementarias

1. The EQUATOR Network | Enhancing the Quality and Transparency of Health Research [Internet]. Equator-network.org. 2020 [cited 28 February 2020].
2. Kramer H. Evaluating medical test. Objective and guidelines. Newbury Park, California, USA: Sage publications, Inc; 1992.
3. Dawson B, Trapp R, Arias Rebatet G. Bioestadística médica. México: El Manual Moderno; 2005.
4. Díaz Portillo J. Guía Práctica del Curso de Bioestadística Aplicada a las Ciencias de la

- Salud. Madrid, España: Instituto Nacional de Gestión Sanitaria; 2011.
5. Cortez J, Cobo E. Un paseo estadístico por las guías de publicación. En: *Publicación científica biomédica*. 2ª edición. Barcelona, España: Elsevier; 2016. pp. 119-137.
 6. Bin Abd Razak HR, Ang JE, Attal H, Howe TS, Allen JC. P-Hacking in orthopaedic literature: a twist to the tail. *J Bone Joint Surg Am*. 2016; 98 (20): e91.
 7. Bobenrieth Astete MA. Mitos y realidades en torno a la publicación científica. *Med Clin*. 2000; 114 (9): 339-341.
 8. Altman DG, Gore SM, Gardner MJ, Pocock SJ. Statistical guidelines for contributors to medical journals. *Br Med J (Clin Res Ed)*. 1983; 286 (6376): 1489-1493.
 9. Lang TA, Altman DG. Basic statistical reporting for articles published in biomedical journals: the "Statistical analyses and methods in the published literature" or the SAMPL guidelines. *Int J Nurs Stud*. 2015; 52 (1): 5-9.
 10. Fletcher RH, Fletcher SW, Wagner EH. *Epidemiología clínica*. 2ª ed. Barcelona: Masson-Williams & Wilkins; 1998.
 11. Susser M. *Conceptos y estrategias en epidemiología. El pensamiento causal en ciencias de la salud*. México: Biblioteca de la Salud; 1991.
 12. Martínez González M, Sánchez Villegas A, Toledo Atucha E, Toledo Atucha E. *Bioestadística amigable*. 3ª ed. Barcelona: Elsevier Health Sciences; 2014.



EN APOYO A
LA EDUCACIÓN
MÉDICA
CONTINUA DE
MÉXICO

La Bioestadística es una disciplina científica que se encarga de la aplicación del análisis estadístico a diferentes áreas vinculadas con las ciencias médicas y que permite, mediante la obtención de valores numéricos, ponderar los resultados de una investigación, lo que permite a una hipótesis redactada en un protocolo generar un juicio de valor y establecer una inferencia, la cual se define como el proceso de sacar conclusiones, teniendo en cuenta los efectos de probabilidad generados por los resultados de la investigación, comprobando las hipótesis de la posible asociación entre dos o más variables y estimando los patrones y la intensidad de la asociación entre las mismas.

“Bioestadística sin Matemáticas. El impacto de la evidencia” es un libro que en cinco unidades muestra los principios fundamentales para saber interpretar las pruebas estadísticas, poder elegir el tipo de prueba apegada a la metodología particular de cada tipo de diseño científico, así como aprender a difundir el impacto que los números arrojan, en el entendido que los números son crudos y no entienden de razones, por lo que requieren de una plausibilidad para que estos tengan sentido al momento de evidenciarlos.

Ahora lo que sigue es aventurarse en la publicación de la experiencia, ya que los ortopedistas que publican tienen un mejor desempeño con sus pacientes y con su comunidad, siendo un signo de salud científica en un país o región, ya que las revistas científicas que muestran artículos nacionales son el reflejo del funcionamiento de las instituciones de salud.

Acad. D. en C. José María Jiménez Ávila

Federación Mexicana de Colegios
de Ortopedia y Traumatología, A.C.
www.femecot.com

ISBN 978-607-98618-1-0



9 786079 861810