



Bioestadística:

principios, métodos y aplicaciones

Coordinadores:

Miriam Janet Cervantes López
Jaime Cruz Casados
Laura Nelly Cruz Casados
Raúl Sadrach Obando Carmona



Bioestadística: principios, métodos y aplicaciones

Bioestadística: principios, métodos y aplicaciones / Miriam Janet Cervantes López, Jaime Cruz Casados, Laura Nelly Cruz Casados y Raúl Sadrach Obando Carmona, coordinadores .—Cd. Victoria, Tamaulipas: Universidad Autónoma de Tamaulipas; 2026.

208 págs. ; 17 x 23 cm

1. Investigación

LC: RA408.5 B5.6 2026 DEWEY: 001.4 MBGR

Universidad Autónoma de Tamaulipas
Matamoros SN, Zona Centro
Ciudad Victoria, Tamaulipas C.P. 87000
D. R. © 2026

Consejo de Publicaciones UAT
Centro Universitario Victoria
Centro de Gestión del Conocimiento. Segundo Piso
Ciudad Victoria, Tamaulipas, México. C.P. 87149
Tel. (52) 834 3181-800 • extensión: 2905
cpublicaciones@uat.edu.mx • www.uat.edu.mx • https://libros.uat.edu.mx/

Libro aprobado por el Consejo de Publicaciones UAT
ISBN UAT: 978-607-69439-1-5

Se prohíbe la reproducción total o parcial de esta obra incluido el diseño tipográfico y de portada, sea cual fuera el medio, electrónico o mecánico, sin el consentimiento del Consejo de Publicaciones UAT.
Libro digital

Esta obra y sus capítulos fueron sometidos a una revisión de pares a doble ciego, la cual fue realizada por especialistas pertenecientes al Sistema Nacional de Investigadoras e Investigadores. Así mismo, fueron aprobados para su publicación por el Consejo de Publicaciones de la Universidad Autónoma de Tamaulipas y el Comité Interno de la editorial Fontamara.



Bioestadística:

principios, métodos y aplicaciones

Coordinadores:

Miriam Janet Cervantes López

Jaime Cruz Casados

Laura Nelly Cruz Casados

Raúl Sadrach Obando Carmona



MVZ MC Dámaso Leonardo Anaya Alvarado
PRESIDENTE

Dr. Fernando Leal Ríos
VICEPRESIDENTE

Dra. Dora María Lladó Lárraga
SECRETARIA TÉCNICA

Mtro. Eduardo García Fuentes
VOCAL

Dra. Rosa Issel Acosta González
VOCAL

CP Jesús Francisco Castillo Cedillo
VOCAL

MVZ Rogelio de Jesús Ramírez Flores
VOCAL

Comité Editorial del Consejo de Publicaciones de la Universidad Autónoma de Tamaulipas

Dra. Lourdes Arizpe Slogher • Universidad Nacional Autónoma de México | **Dr. Amalio Blanco** • Universidad Autónoma de Madrid, España | **Dra. Rosalba Casas Guerrero** • Universidad Nacional Autónoma de México | **Dr. Francisco Díaz Bretones** • Universidad de Granada, España | **Dr. Rolando Díaz Lowing** • Universidad Nacional Autónoma de México | **Dr. Manuel Fernández Ríos** • Universidad Autónoma de Madrid, España | **Dr. Manuel Fernández Navarro** • Universidad Autónoma Metropolitana, México | **Dra. Juana Juárez Romero** • Universidad Autónoma Metropolitana, México | **Dr. Manuel Marín Sánchez** • Universidad de Sevilla, España | **Dr. Cervando Martínez** • University of Texas at San Antonio, E.U.A. | **Dr. Darío Páez** • Universidad del País Vasco, España | **Dra. María Cristina Puga Espinosa** • Universidad Nacional Autónoma de México | **Dr. Luis Arturo Rivas Tovar** • Instituto Politécnico Nacional, México | **Dr. Aroldo Rodrigues** • University of California at Fresno, E.U.A. | **Dr. José Manuel Valenzuela Arce** • Colegio de la Frontera Norte, México | **Dra. Margarita Velázquez Gutiérrez** • Universidad Nacional Autónoma de México | **Dr. José Manuel Sabucedo Cameselle** • Universidad de Santiago de Compostela, España | **Dr. Alessandro Soares da Silva** • Universidad de São Paulo, Brasil | **Dr. Akexandre Dorna** • Universidad de CAEN, Francia | **Dr. Ismael Vidales Delgado** • Universidad Regiomontana, México | **Dr. José Francisco Zúñiga García** • Universidad de Granada, España | **Dr. Bernardo Jiménez** • Universidad de Guadalajara, México | **Dr. Juan Enrique Marciano Medina** • Universidad de Puerto Rico-Humacao | **Dra. Ursula Oswald** • Universidad Nacional Autónoma de México | **Arq. Carlos Mario Yori** • Universidad Nacional de Colombia | **Arq. Walter Debenedetti** • Universidad de Patrimonio, Colonia, Uruguay | **Dr. Andrés Piqueras** • Universitat Jaume I, Valencia, España | **Dra. Yolanda Troyano Rodríguez** • Universidad de Sevilla, España | **Dra. María Lucero Guzmán Jiménez** • Universidad Nacional Autónoma de México | **Dra. Patricia González Aldea** • Universidad Carlos III de Madrid, España | **Dr. Marcelo Urrea** • Revista Latinoamericana de Psicología Social | **Dr. Rubén Ardila** • Universidad Nacional de Colombia | **Dr. Jorge Gissi** • Pontificia Universidad Católica de Chile | **Dr. Julio F. Villegas †** • Universidad Diego Portales, Chile | **Ángel Bonifaz Ezeta †** • Universidad Nacional Autónoma de México

Índice

Prólogo	15
Unidad 1. Fundamentos de la estadística	17
1.1. Introducción a la estadística	19
1.1.1. Definición de estadística	19
1.1.2. División de la estadística	20
1.1.3. Utilidad e importancia de la estadística	21
1.1.4. Uso de la estadística en el campo de las ciencias de la salud	22
1.1.5. Fuentes de información estadística	23
1.2. Conceptos de estadística	26
1.2.1. Definiciones básicas	27
1.2.2. Tipos de variables	29
1.2.2.1. Variables cuantitativas	29
1.2.2.2. Variables cualitativas	30
1.2.3. Proceso de una investigación estadística	31
1.2.4. Identificación de conceptos básicos en un contexto de realidad	36
1.2.5. Aplicaciones cotidianas de la estadística en el área médica	37
Unidad 2. Representación numérica y gráfica de datos	47
2.1. Recolección de datos	49
2.2. Organización de datos y distribución de frecuencias	51
2.2.1. Datos no agrupados	51
2.2.2. Datos agrupados	53
2.3. Representación gráfica de datos	69
2.4. Análisis de datos	72
Unidad 3. Medidas de tendencia central y de dispersión	75
3.1. Medidas de tendencia central	77
3.1.1. Media aritmética	78
3.1.2. Mediana	79
3.1.3. Moda	80

3.2. Medidas de dispersión	81
3.2.1. Recorrido	82
3.2.2. Varianza	83
3.2.3. Desviación estándar	85
Unidad 4. Principios de muestreo y estimación	87
4.1. Conceptos básicos de muestreo	89
4.2. Muestreo probabilístico	90
4.2.1. Muestreo aleatorio simple	91
4.2.2. Muestreo Sistemático	92
4.2.3. Muestreo estratificado	94
4.2.4. Muestreo por conglomerados	96
4.3. Estimación en el muestreo aleatorio simple	97
4.3.1. Metodología del muestreo aleatorio simple	99
4.3.2. Teorema del Límite Central	103
Unidad 5. Inferencia estadística	107
5.1. Parámetros y estimadores	109
5.1.1. Intervalos de confianza	111
5.2. Pruebas de hipótesis	117
5.2.1. Hipótesis estadísticas	117
5.2.2. Significancia alfa y significancia estadística (valor p)	118
5.2.3. Errores tipo I y II	119
5.2.4. Estadística paramétrica y no paramétrica	120
Unidad 6. Inferencia sobre una muestra	125
6.1. Datos cuantitativos	127
6.1.1. Prueba z y t para una muestra	127
6.2. Datos cualitativos	131
6.2.1. Kolmogórov-Smirnov	132
6.2.2. Chi-cuadrada	134
6.2.3. Binomial	137
Unidad 7. Inferencia sobre muestras independientes	143
7.1. Dos muestras	145
7.1.1. t de Student	146
7.1.2. U de Mann-Whitney	149

7.2. Tres o más muestras	152
7.2.1. Anova de un factor	153
7.2.2. Anova de k factores	158
7.2.3. Kruskal-Wallis	163
Unidad 8. Inferencia sobre muestras relacionadas	169
8.1. Dos muestras	171
8.1.1. t de Student	171
8.1.2. Wilcoxon	175
8.1.3. McNemar	178
8.2. Tres o más muestras	183
8.2.1. Friedman	184
8.2.2. Q de Cochran	189
Unidad 9. Correlación y regresión	195
9.1. Correlación	197
9.2. Correlaciones bivariadas	198
9.3. Modelo de regresión lineal simple y múltiple	201
9.3.1. Regresión lineal simple	201
9.3.2. Regresión lineal múltiple	204
9.4. Modelo de regresión logística	209
Referencias	215
Glosario de términos	223

Índice de figuras

Figura 1. Definición de estadística	19
Figura 2. Clasificación de la estadística	20
Figura 3. Conceptos de estadística	26
Figura 4. Tipos de variables	29
Figura 5. Aplicaciones de la estadística en el área médica	38
Figura 6. Técnicas para recolección de datos	49
Figura 7. Organización de datos y distribución de frecuencias	51
Figura 8. Conjunto de calificaciones	52
Figura 9. Gráfica de tallo y hojas	52
Figura 10. Frecuencia	53
Figura 11. Frecuencia acumulada	54
Figura 12. Frecuencia relativa	55
Figura 13. Frecuencia relativa acumulada	55
Figura 14. Componentes fundamentales sobre intervalos en una distribución de frecuencias	58
Figura 15. Patrones típicos de distribuciones de frecuencia	74
Figura 16. Medidas de tendencia central	77
Figura 17. Media aritmética	78
Figura 18. Cálculo de media aritmética	79
Figura 19. Mediana	80
Figura 20. Moda	81
Figura 21. Medidas de dispersión	82
Figura 22. Recorrido	82
Figura 23. Cálculo de recorrido	83
Figura 24. Varianza	84
Figura 25. Cálculo de la media	84
Figura 26. Cálculo de la varianza	85
Figura 27. Desviación estándar	86
Figura 28. Cálculo de la desviación estándar	86
Figura 29. Población y muestra	89
Figura 30. Análisis muestral	90
Figura 31. Fórmulas para cálculo de la muestra	100
Figura 32. Cálculo del tamaño de muestra óptimo	102
Figura 33. Nivel de confianza	112
Figura 34. Intervalo de confianza para una media poblacional	113

Figura 35. Intervalo de confianza para una proporción poblacional	114
Figura 36. Cálculo de intervalo de confianza para una media poblacional	115
Figura 37. Cálculo de intervalo de confianza para una proporción poblacional	115
Figura 38. Fórmula prueba z	128
Figura 39. Cálculo de estadístico z	129
Figura 40. Fórmula prueba t	130
Figura 41. Cálculo de estadístico t	131
Figura 42. Fórmula de frecuencias esperadas	136
Figura 43. Fórmula estadístico chi-cuadrada	136
Figura 44. Fórmula grados de libertad distribución chi cuadrada	137
Figura 45. Distribución de probabilidad variable aleatoria discreta	138
Figura 46. Expresión esperanza matemática	139
Figura 47. Fórmula varianza $\text{var}(x)$	140
Figura 48. Fórmula distribución binomial	142
Figura 49. Fórmula de la prueba t de Student para muestras independientes	147
Figura 50. Fórmula de la prueba t de Student para muestras pareadas	147
Figura 51. Fórmula de la estadística U	150
Figura 52. Fórmula del estadístico F	154
Figura 53. Estadístico de prueba t Student	172
Figura 54. Fórmula estadística de prueba de McNemar	180
Figura 55. Fórmula para aplicación de la prueba de McNemar	182
Figura 56. Fórmula para cálculo de estadístico de Friedman	186
Figura 57. Cálculo del estadístico Q de Cochran	190
Figura 58. Fórmula general regresión lineal simple	201
Figura 59. Ecuación del modelo de regresión lineal múltiple	204
Figura 60. Modelo matemático del modelo de regresión logística	209
Figura 61. Fórmula de logaritmo de las probabilidades	210
Figura 62. Fórmula del modelo de regresión logística binaria	212

Índice de tablas

Tabla 1. Técnicas para la recolección de los datos	32
Tabla 2. Esquema metodológico para la solución de problemas estadísticos	35
Tabla 3. Presentación tabular de resultados	61
Tabla 4. Frecuencia absoluta	63
Tabla 5. Tabla de distribución de frecuencias	64
Tabla 6. Tabla de datos de temperaturas ambientales	66
Tabla 7. Tabla de frecuencias	67
Tabla 8. Tabla por intervalos de clase	68
Tabla 9. Tabla de doble entrada	69
Tabla 10. Muestreo aleatorio simple	91
Tabla 11. Muestreo sistemático	93
Tabla 12. Muestreo estratificado	94
Tabla 13. Muestreo por conglomerados	96
Tabla 14. Grupos de pacientes adultos	122
Tabla 15. Tabla de contingencia satisfacción de los pacientes	136
Tabla 16. Datos de pacientes diabéticos	149
Tabla 17. Datos de pacientes postoperados	152
Tabla 18. Datos de pacientes hipertensos	156
Tabla 19. Datos de pacientes en recuperación postoperatoria de cirugía ortopédica	161
Tabla 20. Datos recolectados sobre la satisfacción de pacientes sobre calidad de atención médica	165
Tabla 21. Datos recolectados de pacientes con hipertensión arterial	174
Tabla 22. Datos recolectados de pacientes con diagnóstico de hipercolesterolemia	177
Tabla 23. Tabla de contingencia de datos pareados	179
Tabla 24. Datos recolectados de campaña de concienciación sobre seguridad vial	181
Tabla 25. Datos recolectados de pacientes con hipertensión arterial	187
Tabla 26. Datos recolectados de pacientes con trastorno de ansiedad generalizada	192
Tabla 27. Datos recolectados sobre edad y presión arterial sistólica	199

Tabla 28. Datos recolectados sobre altura y peso	203
Tabla 29. Datos recolectados sobre calidad de dieta, nivel de actividad física, glucosa y edad	205
Tabla 30. Datos recolectados de pacientes con riesgo de diabetes	207
Tabla 31. Datos recolectados sobre IMC y PAS de adultos	211

Prólogo

La formación médica exige no solo habilidades clínicas y conocimientos biológicos, sino también una comprensión profunda de los métodos que permiten transformar las observaciones en conocimiento validado. La estadística ocupa un lugar central en el currículo de las ciencias de la salud, pues convierte la experiencia en evidencia y la evidencia en acción sanitaria fundamentada.

Bioestadística: principios, métodos y aplicaciones es una obra concebida como texto guía para estudiantes de medicina, con una clara vocación pedagógica y una rigurosa fundamentación científica. Su enfoque progresivo parte de los conceptos básicos y avanza hacia técnicas de inferencia, regresión y análisis multivariado, con el objetivo de que el lector construya una competencia estadística sólida, crítica y contextualizada para los desafíos reales de la práctica médica.

Este libro, con su amplia cobertura temática, articula teoría, método y aplicación clínica, ya que cada unidad está construida con un enfoque didáctico, respaldado por criterios metodológicos contemporáneos y ejemplos de escenarios médicos reales: desde el análisis de brotes epidémicos hasta el diseño de ensayos clínicos; desde la vigilancia epidemiológica hasta la medicina personalizada. En un entorno académico público, la formación debe responder tanto a estándares de excelencia como a realidades sociales complejas.

En tiempos en que la medicina se redefine por la integración de tecnología, datos y decisiones éticas, formar médicos capaces de leer críticamente la literatura científica, diseñar investigaciones rigurosas y participar en políticas sanitarias basadas en evidencia es una prioridad. Este libro fomenta el pensamiento analítico y el compromiso con una medicina informada, justa y técnicamente solvente.

Unidad

Fundamentos de la estadística

Presentación de la unidad

La estadística es una disciplina fundamental en el campo de las ciencias de la salud, ya que permite transformar datos en información útil para la toma de decisiones clínicas, administrativas y de investigación, con el fin de describir fenómenos, identificar patrones y realizar inferencias sobre una población a partir del estudio de una muestra representativa. En esta unidad se exponen los principios básicos de la estadística, la clasificación, la utilidad y las principales fuentes de información estadística, con un enfoque orientado al ejercicio profesional en salud.

Objetivos

Al finalizar el estudio de esta unidad, el estudiante será capaz de:

- Reconocer los conceptos fundamentales de la estadística.
- Comprender la utilidad de la estadística en el análisis de fenómenos.
- Identificar problemas del área de la salud y abordarlos mediante herramientas estadísticas.

Competencia específica

Organizar información relevante para la delimitación de un problema en materia de ciencias de la salud, empleando conceptos como población, muestra, datos y variables.

1.1. Introducción a la estadística

En el contexto actual de la generación masiva de datos, la estadística es un componente transversal en la formación científica y profesional. Más allá de su definición formal, constituye una herramienta epistemológica para interpretar la información empírica, sustentar la toma de decisiones y formular conclusiones válidas en condiciones de incertidumbre (Frieden, 2017; Pagano et al., 2022). Su valor formativo radica en el desarrollo del pensamiento analítico, la rigurosidad metodológica y la capacidad para abstraer regularidades a partir de datos dispersos.

En disciplinas orientadas al bienestar humano, como las ciencias de la salud, la estadística transforma observaciones individuales en conocimiento generalizable, útil para guiar políticas públicas, diseñar intervenciones clínicas y optimizar procesos de atención (Daniel y Cross, 2018; WHO, 2021).

1.1.1. Definición de estadística

La estadística es una ciencia formal que estudia los métodos adecuados para recolectar, organizar, analizar, interpretar y presentar datos numéricos, con el propósito de extraer conclusiones fundamentadas. Esta disciplina proporciona herramientas esenciales para comprender fenómenos observables, identificar tendencias y formular predicciones en un marco de incertidumbre (Navidi, 2021).

El término estadística proviene del latín status (“estado” o “condición” de algo). En sus inicios, analizaba datos relacionados con el Estado, como censos de población, impuestos y registros militares. Con el desarrollo de la ciencia y la tecnología, su aplicación se ha expandido a la medicina, biología, economía, educación, ingeniería y psicología (Ioannidis, 2019; Wasserstein et al., 2019; Figura 1).

Figura 1

Definición de estadística



Fuente: elaboración propia (2025).

La estadística no solo permite organizar y describir datos, sino también establecer relaciones entre variables, verificar hipótesis, estimar parámetros desconocidos de una población y modelar comportamientos complejos. Su aplicación en las ciencias de la salud resulta imprescindible para evaluar la eficacia de tratamientos, controlar brotes epidémicos, optimizar la asignación de recursos y mejorar la calidad de los servicios sanitarios (Frieden, 2017). En el lenguaje científico, el término estadística puede aludir a la información cuantitativa de una población, los procedimientos para analizar datos o las medidas calculadas a partir de una muestra.

1.1.2. División de la estadística

Para facilitar su estudio y aplicación sistemática, la estadística se clasifica en descriptiva e inferencial. Esta división permite abordar con precisión la organización de los datos y la obtención de conclusiones generalizables a partir de muestras (Daniel y Cross, 2018; Pagano et al., 2022; Figura 2).

Figura 2

Clasificación de la estadística



Fuente: elaboración propia (2025).

Estadística descriptiva: se encarga de la recopilación, organización, presentación y análisis de un conjunto de datos con el propósito de describir sus características principales, sin inferir más allá de la información observada (Kwak y Kim, 2017; Nørskov et al., 2021). Entre sus herramientas más comunes se encuentran las tablas de frecuencia, los gráficos y las medidas de tendencia central y dispersión.

Estadística inferencial: permite, a partir del estudio de una muestra representativa, estimar, generalizar y probar hipótesis sobre una población. Esta rama se fundamenta en el cálculo de probabilidades y en el uso de modelos matemáticos para evaluar la incertidumbre inherente a cualquier proceso de inferencia (Agresti, 2018; Adam, 2020).

1.1.3. Utilidad e importancia de la estadística

La estadística es fundamental en la recolección, organización, análisis, interpretación y presentación de datos cuantitativos y cualitativos, lo que permite comprender de manera objetiva diversos fenómenos en múltiples contextos. Su utilidad trasciende la descripción numérica al facilitar la toma de decisiones y la formulación de inferencias a partir de la evidencia empírica (Daniel y Cross, 2018; Pagano et al., 2022).

La estadística revela patrones, tendencias y variaciones en conjuntos de datos; además, posibilita el seguimiento de la evolución de un fenómeno. Este seguimiento es esencial en la salud pública, la economía, la educación, la industria y las ciencias sociales. Por ejemplo, el análisis de series temporales permite identificar cambios significativos en variables clave y anticipar comportamientos futuros mediante modelos predictivos (Adam, 2020).

En México, el Instituto Nacional de Estadística y Geografía (INEGI) desempeña un papel central al encargarse de la recopilación, procesamiento, análisis y difusión de información estadística y geográfica del país. Los datos publicados por este organismo ofrecen insumos confiables que permiten conocer la evolución histórica de aspectos demográficos, económicos, sociales, ambientales y territoriales, con el fin de elaborar políticas públicas y evaluar programas gubernamentales (INEGI, 2023).

La estadística tiene aplicación transversal en todas las disciplinas del conocimiento (humanidades; ciencias exactas, naturales y sociales). En la investigación científica y en entornos laborales, permite abordar problemas específicos mediante la formulación de hipótesis, la evaluación de intervenciones y la generación de conclusiones con evidencia (WHO, 2021).

El problema se define como la discrepancia entre una situación real y una deseada. La estadística proporciona una descripción objetiva de la realidad, facilitando al investigador o al tomador de decisiones el análisis de sus objetivos, la identificación de brechas y la selección de estrategias para su resolución. Así, la estadística no solo describe el estado actual de los fenómenos, sino que transforma el entorno mediante la generación de conocimiento útil y verificable (OECD, 2020).

1.1.4. Uso de la estadística en el campo de las ciencias de la salud

La estadística proporciona los fundamentos metodológicos para la recolección, análisis, interpretación y presentación de datos relacionados con la salud y la enfermedad. Su aplicación se extiende desde la investigación biomédica básica hasta la salud pública; en la evaluación de intervenciones, la formulación de políticas sanitarias y la mejora de la atención médica basada en evidencia (Gordis, 2020; WHO, 2021).

La estadística también tiene aplicación en los ensayos clínicos, que representan el estándar de referencia para la evaluación rigurosa de la eficacia y seguridad de nuevos tratamientos, medicamentos o intervenciones sanitarias. En estos estudios, el diseño experimental, la determinación del tamaño de la muestra, la asignación aleatoria y el análisis de los resultados se sustentan en técnicas estadísticas que garantizan la validez interna y externa de los hallazgos. El cálculo del tamaño muestral se fundamenta en parámetros como la variabilidad esperada de los datos y el tamaño del efecto clínicamente relevante, con el fin de asegurar la potencia estadística del estudio. Asimismo, las pruebas de significancia y los intervalos de confianza permiten establecer si las diferencias observadas entre grupos de intervención y control pueden atribuirse al tratamiento evaluado y no al azar (Ioannidis et al., 2014; Chow y Zhang, 2020; Pagano et al., 2022).

En el campo de la epidemiología, la estadística se emplea para estudiar la distribución, frecuencia y determinantes de las enfermedades en poblaciones humanas. Mediante el uso de técnicas estadísticas, es posible identificar factores de riesgo, estimar tasas de incidencia y prevalencia, así como evaluar el impacto de intervenciones preventivas y terapéuticas. Durante emergencias sanitarias, como la pandemia por la COVID-19, los modelos estadísticos epidemiológicos predijeron el curso de la enfermedad, estimaron la carga de morbimortalidad y guiaron decisiones estratégicas relacionadas con confinamientos, distribución de recursos y campañas de vacunación masiva. La precisión y robustez de estos modelos contribuyeron a la mitigación del impacto social y sanitario de dicha contingencia (Adam, 2020; Jewell et al., 2020).

La bioestadística es una disciplina que aplica los principios estadísticos al estudio de fenómenos biológicos y médicos, en investigaciones genéticas y genómicas, en las que se analizan grandes volúmenes de datos derivados del estudio del ADN, ARN o proteínas. Por ejemplo, los estudios de asociación del genoma completo (GWAS) utilizan modelos estadísticos complejos para identificar correlaciones entre variaciones genéticas y enfermedades como el cáncer, la diabetes o trastornos neurodegenerativos.

Estas investigaciones permiten avanzar hacia una medicina de precisión, en la que los tratamientos se diseñan y adaptan a las características genéticas y clínicas individuales (Ioannidis, 2019; Nørskov et al., 2021).

La estadística permite monitorear y evaluar el estado de salud de las poblaciones a través del análisis de encuestas nacionales de salud, registros administrativos y sistemas de vigilancia epidemiológica. La información obtenida mediante estos instrumentos revela prioridades sanitarias para asignar recursos y diseñar programas. Por ejemplo, los datos sobre la prevalencia de enfermedades crónicas no transmisibles, como la hipertensión o la obesidad, han fundamentado campañas preventivas, políticas fiscales y programas de intervención dirigidos a poblaciones vulnerables (OECD, 2020; WHO, 2021).

En la práctica clínica cotidiana, la estadística se incorpora mediante la medicina basada en evidencia. Las guías clínicas y protocolos de tratamiento se elaboran a partir de revisiones sistemáticas y metaanálisis, que combinan los resultados de múltiples estudios mediante técnicas estadísticas, lo que permite obtener conclusiones más robustas y generalizables. Asimismo, los modelos de predicción clínica, construidos con regresiones logísticas, curvas ROC u otros métodos estadísticos, permiten estimar la probabilidad de que un paciente desarrolle una enfermedad o experimente un desenlace adverso, contribuyendo a una atención médica más personalizada y eficiente (Brozek et al., 2021).

La estadística es un componente transversal y estratégico en las ciencias de la salud, ya que su adecuada aplicación permite garantizar la calidad metodológica de las investigaciones, optimizar los recursos disponibles, anticipar riesgos sanitarios y ofrecer una atención médica basada en evidencia científica sólida, lo que repercute de manera positiva en la salud individual y colectiva (Adam, 2020; WHO, 2021).

1.1.5. Fuentes de información estadística

Las fuentes de información deben contener datos e indicadores estadísticos confiables. Su adecuada selección y utilización permiten realizar análisis válidos que orienten la toma de decisiones, el diseño de políticas públicas y la investigación científica. Las fuentes estadísticas se clasifican en dos categorías: directas e indirectas, de acuerdo con la proximidad del dato al fenómeno observado (OECD, 2020; WHO, 2021).

Directas

Proviene del lugar y del momento en el que ocurre el fenómeno de interés. Estas fuentes permiten obtener información de primera mano, sin mediación ni transformación previa. Entre las más representativas se encuentran:

Historias clínicas: registros detallados de la atención médica brindada a los pacientes, que incluyen diagnósticos, tratamientos, resultados de estudios auxiliares y evolución clínica. Se utilizan en estudios de cohortes, evaluación de efectividad terapéutica y seguimiento de enfermedades crónicas (Gliklich et al., 2014).

Encuestas de salud: instrumentos estructurados que recopilan información autoinformada sobre condiciones de salud, estilos de vida y factores de riesgo en individuos o poblaciones. Se emplean en investigaciones epidemiológicas, estudios de prevalencia y determinantes sociales de la salud.

Registros hospitalarios: documentos administrativos que recopilan datos sobre admisiones, egresos, intervenciones quirúrgicas, diagnósticos y procedimientos médicos realizados en instituciones hospitalarias. Son útiles para evaluar la utilización y calidad de los servicios de salud (OECD, 2020).

Pruebas de laboratorio: reportes derivados de estudios clínicos como análisis de sangre, biopsias y estudios de imagen. Su análisis permite el monitoreo diagnóstico, la validación de pruebas y la evaluación de nuevas tecnologías.

Datos de vacunación: registros individuales o poblacionales sobre la administración de vacunas, incluyendo fechas, tipos y dosis, esenciales para estudios de cobertura vacunal y efectividad inmunológica (WHO, 2021).

Indirectas

Se componen de información que no proviene directamente del evento observado, sino de registros administrativos o bases secundarias. Entre ellas destacan:

Registros de seguros médicos: reclamaciones, coberturas y costos asociados a la atención médica. Son útiles en estudios de economía de la salud y equidad en el acceso a servicios.

Informes de estadísticas vitales: información gubernamental sobre nacimientos, defunciones, matrimonios y divorcios. Constituyen un insumo clave para el análisis demográfico, la mortalidad y la esperanza de vida (United Nations, 2022).

Bases de datos administrativas: conjuntos de datos recolectados sistemáticamente por instituciones de salud que permiten planear recursos y evaluar políticas.

Registros de farmacia: información sobre la dispensación de medicamentos, utilizada en estudios de adherencia terapéutica y farmacovigilancia.

Encuestas nacionales de salud: investigaciones poblacionales de gran escala realizadas por organismos oficiales, que permiten conocer el estado de salud general de la población y las tendencias en salud pública (INEGI, 2023).

En general, el acceso a las fuentes directas puede estar limitado por barreras éticas, logísticas o institucionales. En tales casos, las fuentes indirectas representan una alternativa viable y se emplean de forma complementaria. La literatura especializada distingue otras formas de clasificación de las fuentes de información:

- **Primarias:** información recolectada directamente por el investigador, como encuestas, entrevistas o mediciones clínicas.
- **Secundarias:** datos previamente recolectados y publicados por otras entidades, utilizados con fines distintos a los originalmente previstos.
- **Internas:** información generada dentro de una organización.
- **Externas:** información obtenida fuera de la organización emisora del análisis.

Indicadores nacionales

Uno de los desafíos contemporáneos en estadística aplicada es la generación de información relevante para tomar decisiones. Para ello, los indicadores estadísticos permiten cuantificar fenómenos sociales, económicos y sanitarios. Algunos ejemplos de indicadores nacionales de la salud son: A) Población total y por grupos etarios, B) Tasa de natalidad, C) Tasa de mortalidad general y específica, D) Derechohabencia en instituciones de salud, E) Prevalencia de discapacidad y F) Estadísticas sobre salud mental. Estos indicadores se encuentran disponibles en fuentes oficiales como el INEGI, cuya plataforma (<http://www.inegi.gob.mx>) constituye un repositorio actualizado de datos de interés nacional.

Indicadores internacionales

A nivel global, diversos organismos multilaterales y centros de investigación han desarrollado indicadores que permiten comparar países y evaluar tendencias transnacionales. Estos indicadores reflejan la creciente complejidad de los fenómenos contemporáneos y promueven una visión integral del desarrollo humano y la gobernanza. Algunos ejemplos relevantes son: A) Índice de Progreso Real (IPR), B) Índice de Competitividad Mundial, C) Índice de Desarrollo Humano (IDH), D) Índice de Percepción de la Corrupción, E) Índice de Confianza en las Instituciones y F) Índice de Calidad de Vida de la OCDE. El uso sistemático y riguroso de estos indicadores nacionales e internacionales enriquece el análisis estadístico para formular políticas que mejoren los sistemas de salud y bienestar social.

1.2. Conceptos de estadística

La estadística constituye una disciplina fundamental para el desarrollo y consolidación del conocimiento en las ciencias de la salud. Su aplicación es amplia y transversal, abarcando desde la investigación biomédica básica hasta la práctica clínica, la salud pública y la gestión sanitaria. La estadística proporciona las herramientas necesarias para recolectar, organizar, analizar e interpretar datos con objetividad y precisión, lo cual permite formular hipótesis, evaluar intervenciones, establecer patrones, identificar relaciones causales y generar conclusiones fundamentadas en evidencia empírica (Frieden, 2017; Daniel y Cross, 2018; Ioannidis, 2019). El uso adecuado de la estadística facilita una evaluación crítica de la literatura científica, al promover la comprensión e interpretación de los resultados de estudios observacionales, ensayos clínicos, revisiones sistemáticas y metaanálisis (Abelsson et al., 2021; WHO, 2021).

Por ello, el dominio de los conceptos estadísticos fundamentales resulta imprescindible para los profesionales y estudiantes del área de la salud, quienes deben ser capaces de interpretar datos, aplicar criterios de significancia y confiabilidad, además de distinguir entre correlación y causalidad en sus contextos específicos de actuación (Wasserstein y Lazar, 2016; Figura 3).

Figura 3

Conceptos de estadística



Fuente: elaboración propia (2025).

1.2.1. Definiciones básicas

Estos términos constituyen la base de los métodos estadísticos, por lo que su comprensión es esencial en todas las etapas de un estudio: planeación, recolección de datos, análisis e interpretación (Daniel y Cross, 2018; Pagano et al., 2022). A continuación, se describen las principales definiciones:

<p style="text-align: center;">Censo</p> 	<p>Es una investigación exhaustiva que recopila información sobre todos los elementos de una población. Su objetivo es obtener datos precisos, completos y actualizados sobre las características observables de todos los individuos. Un ejemplo es el Censo Nacional de Población que recaba información demográfica, económica y social de cada habitante del país (INEGI, 2020; United Nations, 2022).</p>
<p style="text-align: center;">Parámetro</p> 	<p>El parámetro es una característica numérica que describe una propiedad específica de una población. En la práctica, los parámetros suelen ser desconocidos y deben estimarse a partir de los datos de una muestra. Ejemplos comunes de parámetros incluyen el promedio de edad, la desviación estándar del ingreso o la proporción de individuos con una condición de salud en una población (Agresti, 2018; Pagano et al., 2022).</p>
<p style="text-align: center;">Estadístico</p> 	<p>El estadístico es un valor numérico que resume una característica de una muestra y se utiliza como estimación del parámetro poblacional correspondiente. Entre los estadísticos más empleados se encuentran la media muestral, la mediana, la proporción y la varianza de la muestra (Agresti, 2018; Pagano et al., 2022).</p>
<p style="text-align: center;">Población</p> 	<p>La población se define como el conjunto completo de elementos, personas, objetos o unidades que comparten una o varias características medibles y constituyen el objeto de estudio de una investigación estadística. Por ejemplo, una población puede estar integrada por todos los empleados de una empresa, los estudiantes de una universidad, los hogares de una ciudad o los productos de una línea de producción. El análisis estadístico tiene como finalidad describir, comprender o predecir fenómenos en la población (Agresti, 2018; Triola, 2020).</p>

Muestra



La muestra es un subconjunto representativo extraído de una población que se utiliza cuando resulta inviable o innecesario estudiar la totalidad de sus elementos. El análisis muestral debe reflejar las características de la población de origen. El tamaño de la muestra, aunque considerablemente menor que el de la población, debe garantizar la validez de las inferencias. Ejemplos de aplicación incluyen estudios sobre los alumnos de una facultad, los habitantes de una comunidad o los clientes de una empresa (Triola, 2020; Pagano et al., 2022).

Individuo



El individuo, también denominado unidad estadística, es un elemento que integra una población o una muestra. Un individuo puede corresponder a una persona, un objeto, una institución o un grupo. Por ejemplo, en un censo económico, cada establecimiento comercial es un individuo, aunque esté conformado por múltiples personas (Agresti, 2018; Triola, 2020).

Muestreo



El muestreo es el proceso mediante el cual se selecciona una muestra representativa de una población, mediante métodos probabilísticos o no probabilísticos, con el fin de obtener datos adecuados para realizar inferencias válidas sobre la población de origen (Levine et al., 2020; Lohr, 2022).

Dato



El dato se define como cada uno de los valores obtenidos durante la ejecución de un estudio estadístico. Los datos constituyen la materia prima de la estadística y pueden adoptar distintas formas, como numérica, categórica u ordinal. Por ejemplo, al lanzar una moneda cinco veces, se obtienen cinco datos que podrían ser: sol, sol, águila, sol, águila (Triola, 2020; Pagano et al., 2022).

Variable



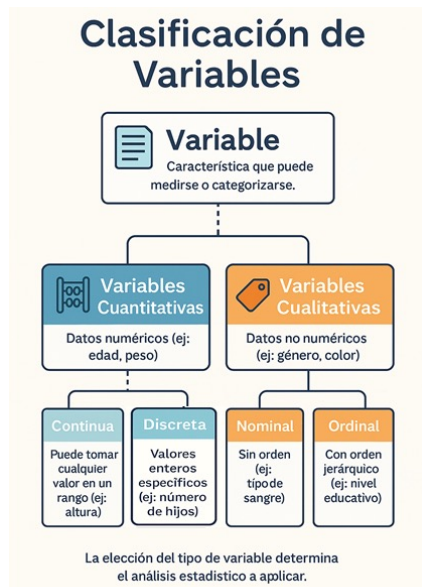
La variable es una característica medible u observable de los individuos que conforman una población o una muestra, cuyo valor puede diferir. Las variables se clasifican en cuantitativas y cualitativas (Agresti, 2018; Triola, 2020).

1.2.2. Tipos de variables

En estadística, una variable es toda característica observable y medible que puede adquirir diferentes valores entre los elementos de una población o muestra. Su correcta identificación y clasificación determinan el tipo de análisis estadístico aplicable, así como la forma en que los datos serán recolectados, organizados y presentados (Agresti, 2018; Triola, 2020). Las variables se dividen en dos categorías: cuantitativas y cualitativas, cada una con subclasificaciones (Pagano et al., 2022; Figura 4).

Figura 4

Tipos de variables



Fuente: elaboración propia (2025).

1.2.2.1. Variables cuantitativas

Las variables cuantitativas pueden expresarse mediante valores numéricos: representan cantidades o magnitudes, permiten realizar operaciones aritméticas con sentido matemático, se subdividen en continuas y discretas (Triola, 2020; Pagano et al., 2022):

- **Cuantitativas continuas:** las variables cuantitativas continuas pueden tomar un número infinito de valores dentro de un intervalo determinado. Son susceptibles de fraccionamiento, lo que implica que entre dos valores

cualesquiera pueden existir múltiples valores intermedios. Ejemplos de este tipo de variables incluyen el peso corporal (por ejemplo, entre 50.0 y 57.9 kg), la estatura (entre 1.70 y 1.76 metros) y el ingreso mensual (entre \$ 15 030.40 y \$ 25 299.90) (Triola, 2020; Pagano et al., 2022).

- **Cuantitativas discretas:** las variables cuantitativas discretas solo pueden asumir valores enteros y no admiten subdivisión dentro de las unidades consideradas. Generalmente, se obtienen mediante conteo. Por ejemplo, el número de hijos en una familia, los estudiantes en un aula o los empleados en una empresa (Triola, 2020; Pagano et al., 2022).

1.2.2.2. Variables cualitativas

Las variables cualitativas, también denominadas categóricas, representan cualidades, atributos o categorías no numéricas, cuya función principal es clasificar a los individuos en grupos con características comunes. Estas variables no permiten realizar operaciones aritméticas, aunque sí pueden contarse o agruparse. Se clasifican en dos tipos: nominales y ordinales (Agresti, 2018; Triola, 2020).

- **Cualitativas nominales:** Las variables cualitativas nominales son aquellas que no presentan un orden lógico o jerárquico entre sus categorías. Sus modalidades son mutuamente excluyentes y no pueden organizarse en función de criterios de magnitud. Ejemplos de este tipo de variables incluyen el estado civil (soltero, casado, viudo), la preferencia por una marca, el sexo biológico o el lugar de residencia (Agresti, 2018; Triola, 2020).
- **Cualitativas ordinales:** las variables cualitativas ordinales son aquellas cuyas categorías pueden organizarse en jerarquías o niveles que expresan un orden. Los valores asignados pueden implicar un mayor o menor grado de una característica, aunque las distancias entre las categorías no son necesariamente iguales ni cuantificables. Ejemplos de este tipo de variables incluyen el nivel de escolaridad (primaria, secundaria, preparatoria, licenciatura), la calidad del servicio (deficiente, regular, buena, excelente), el nivel socioeconómico (bajo, medio, alto) y los días de la semana (Agresti, 2018; Triola, 2020).

La distinción entre los tipos de variables constituye uno de los pilares conceptuales del diseño de estudios científicos, tanto en el ámbito de las ciencias de la salud como en otras áreas del conocimiento (Triola, 2020; Pagano et al., 2022).

1.2.3. Proceso de una investigación estadística

El proceso de una investigación estadística es una secuencia estructurada de etapas orientadas a identificar, analizar y resolver un problema mediante el uso sistemático de datos. Esta metodología permite transformar información empírica en conocimiento útil (Agresti, 2018; Lohr, 2022). A continuación, se describen las etapas fundamentales del proceso de investigación estadística:

El planteamiento del problema

Esta etapa inicial tiene como propósito delimitar con claridad el fenómeno a investigar, así como establecer el contexto, los objetivos y el alcance del estudio. Para ello, es necesario considerar los siguientes elementos:

- El campo temático en el que se desarrolla el fenómeno de interés.
- Las limitaciones de espacio, tiempo y recursos que condicionan el estudio.
- La definición operacional de los conceptos y variables a analizar.
- Los objetivos específicos y metas que se pretende alcanzar mediante la investigación.

El planteamiento adecuado del problema permite orientar todo el proceso estadístico posterior y asegurar la pertinencia metodológica del estudio (Hernández-Sampieri et al., 2021; Lohr, 2022).

La planeación del trabajo

Esta etapa consiste en el diseño estructurado de las fases que integran la investigación, así como en la selección de las herramientas y recursos necesarios para su ejecución. En este momento del proceso se establecen:

- Las fases secuenciales del estudio.
- Las fuentes primarias y secundarias de los datos requeridos.
- La metodología estadística aplicable en cada fase.
- Los plazos estimados para cada etapa.
- Los recursos humanos, materiales, tecnológicos y financieros disponibles.

Una adecuada planeación del trabajo garantiza la optimización del tiempo y los recursos y asegura la calidad metodológica en cada una de las fases de la investigación (Creswell y Creswell, 2018; Hernández-Sampieri et al., 2021).

La recolección de los datos

Esta etapa está destinada a recolectar la información necesaria mediante procedimientos que deben seleccionarse de acuerdo con el tipo de variable, el diseño del estudio y la naturaleza de la población. La recopilación puede realizarse bajo diferentes modalidades:

- Completa (censo): se obtiene información de todos los elementos de la población.
- Parcial (muestreo): se extrae información de una parte representativa de la población.
- Directa: los datos se recogen en el momento y lugar en que ocurre el fenómeno observado.
- Indirecta: los datos se obtienen de fuentes secundarias, como encuestas, registros o bases de datos.
- Permanente: la información se registra de manera continua en el tiempo.
- Periódica: la recolección se realiza en intervalos regulares.
- Ocasional: se efectúa de manera puntual en respuesta a un evento específico.

La adecuada elección de la técnica de recolección de datos asegura la validez y confiabilidad de la información, lo que constituye la base para un análisis estadístico riguroso (Hernández-Sampieri et al., 2021; Lohr, 2022). Las técnicas específicas para cada forma de recolección se sintetizan en el siguiente esquema (Tabla 1).

Tabla 1. Técnicas para la recolección de los datos

Forma de recolección	Descripción	Técnica estadística asociada	Aplicabilidad
Completa	Se obtiene información de todos los elementos de la población.	Censo	Recomendado cuando la población es finita y accesible (por ejemplo, censo de población).
Parcial	Se extrae información de una parte representativa de la población.	Muestreo	Útil en poblaciones grandes o cuando los recursos son limitados.
Directa	Los datos se recogen en el momento y lugar del fenómeno observado.	Diseño de experimentos	Preferible en estudios controlados o clínicos.
Indirecta	Se obtienen datos de fuentes secundarias (encuestas, registros, archivos).	Encuesta/ consulta documental	Ideal cuando el acceso directo a la población no es viable.

Forma de Recolección	Descripción	Técnica Estadística Asociada	Aplicabilidad
Permanente	La información se recoge de forma continua en el tiempo.	Control estadístico de procesos (CEP)	Empleado en vigilancia epidemiológica, producción industrial o seguimiento clínico.
Periódica	La recolección se realiza en intervalos regulares.	Encuestas sistemáticas / registros	Útil para medir tendencias (por ejemplo, encuestas de salud cada 5 años).
Ocasional	Se efectúa de manera puntual por un evento específico.	Encuesta transversal / estudios ad hoc	Aplicable en eventos extraordinarios o situaciones emergentes (por ejemplo, desastres).

Fuente: elaboración propia (2025).

Crítica de los datos

La información recolectada se valida mediante el proceso de crítica de los datos para depurar la información obtenida. En esta etapa se distinguen dos tipos principales:

- Crítica de fondo: orientada a identificar y corregir errores sistemáticos que puedan comprometer la validez del estudio, como inconsistencias conceptuales, sesgos en la medición o incongruencias en los registros.
- Crítica de forma: dirigida a detectar y corregir errores accidentales o aleatorios, tales como omisiones, duplicaciones o errores tipográficos en los datos.

La crítica de los datos fortalece la validez interna y externa de los resultados (Creswell y Creswell, 2018; Hernández-Sampieri et al., 2021).

La elaboración de los datos

Esta etapa consiste en preparar los datos para su análisis mediante una serie de procesos que buscan organizar y estructurar la información de manera coherente y sistemática. Entre estos procesos se incluyen:

- Simplificación: reducción de la complejidad de los datos sin perder su esencia informativa.
- Ordenación: disposición lógica de los datos para facilitar su interpretación.
- Agrupación: clasificación de los datos en categorías homogéneas.
- Concentración: síntesis de la información mediante el uso de medidas estadísticas representativas.
- Presentación: organización visual de los resultados mediante cuadros, tablas o representaciones gráficas.

La adecuada elaboración de los datos permite optimizar la interpretación de la información recolectada, reduce el riesgo de sesgos y garantiza que el análisis estadístico se realice sobre una base estructurada y confiable (Hernández-Sampieri et al., 2021; Lohr, 2022).

La determinación y análisis de los estadígrafos muestrales

Los estadígrafos son medidas resumen que describen las características de una muestra, tales como la media, la mediana, la desviación estándar y la proporción. Su análisis permite identificar relaciones, patrones y tendencias dentro del conjunto de datos estudiados, lo que constituye una etapa fundamental para la caracterización inicial de la información antes de aplicar procedimientos inferenciales.

La correcta determinación y el análisis de los estadígrafos proporcionan una visión estructurada de la información y facilitan la detección de posibles anomalías, distribuciones atípicas o concentraciones de datos, aspectos esenciales para la validez de estudios en las ciencias de la salud y en otras disciplinas aplicadas (Triola, 2020; Pagano et al., 2022).

La inferencia estadística

Esta fase tiene como finalidad extender los resultados obtenidos en una muestra a la población de la que proviene, mediante generalizaciones de carácter probabilístico. La inferencia estadística permite formular conclusiones más allá de los datos observados y establecer el grado de confianza asociado a dichas conclusiones.

Las herramientas más empleadas en esta etapa incluyen:

- Estimaciones puntuales y por intervalos que permiten aproximar parámetros poblacionales a partir de los valores muestrales.
- Pruebas de hipótesis, utilizadas para contrastar supuestos sobre los parámetros poblacionales.
- Modelos de regresión y correlación, que posibilitan el análisis de relaciones entre variables y la predicción de comportamientos.

El uso adecuado de estas técnicas garantiza la validez y confiabilidad de los resultados y constituye un componente esencial en la investigación científica aplicada a las ciencias de la salud, la economía, la psicología y otras disciplinas (Casella y Berger, 2021; Pagano et al., 2022).

La evaluación de las soluciones

La etapa final del proceso consiste en la valoración de las decisiones que podrían adoptarse a partir de los resultados obtenidos. En esta fase se analiza la probabilidad de éxito de las diferentes alternativas de solución, considerando su viabilidad, impacto y pertinencia con respecto al problema inicialmente planteado.

Esta etapa es esencial para garantizar que los hallazgos estadísticos se traduzcan en acciones concretas y aplicables, ya sea en el ámbito de las ciencias de la salud, en la formulación de políticas públicas o en la gestión organizacional. Asimismo, permite retroalimentar el proceso de investigación, ajustando metodologías para estudios futuros y favoreciendo la mejora continua (Creswell y Creswell, 2018; Hernández-Sampieri et al., 2021).

Metodología general para resolver un problema estadístico

De forma complementaria, se presenta un esquema metodológico de solución de problemas estadísticos para organizar el proceso de análisis de manera lógica y estructurada, garantizando la validez y confiabilidad de los resultados obtenidos. Las fases generales comprenden (Tabla 2).

Tabla 2. Esquema metodológico para la solución de problemas estadísticos

a) Planteamiento del problema
Consiste en la definición del fenómeno a estudiar, la delimitación de la población o muestra, la identificación de las variables relevantes y, en caso necesario, la formulación de hipótesis. Asimismo, se determinan el procedimiento y los medios disponibles para la investigación.
b) Elaboración de un modelo
Implica la propuesta de un modelo teórico que represente el comportamiento esperado de las variables. En algunos casos, dicho modelo puede plantearse de manera preliminar y ajustarse posteriormente según los datos recolectados. Entre los modelos más comunes se encuentran la distribución normal, binomial, de Poisson, uniforme, geométrica o binomial negativa.
c) Extracción de la muestra
Se lleva a cabo mediante técnicas de muestreo probabilístico o mediante diseños experimentales, con el propósito de obtener información representativa de la población de estudio.
d) Tratamiento de los datos
Incluye la depuración de la muestra, la eliminación de errores, la tabulación de los resultados y el cálculo de estadígrafos descriptivos (como la media, la mediana o la varianza). Esta fase corresponde a la estadística descriptiva e incorpora la clasificación y presentación de la información mediante tablas y representaciones gráficas.

e) Estimación de los parámetros estadísticos

Se realiza a través de herramientas de estadística inferencial, mediante las cuales se estiman parámetros poblacionales. Este proceso incluye el cálculo de intervalos de confianza y la aplicación de pruebas de hipótesis, que permiten tomar decisiones sustentadas en los datos (Casella y Berger, 2021; Triola, 2020).

Fuente: elaboración propia (2025).

Este esquema metodológico puede adaptarse a distintos campos del conocimiento, favoreciendo la toma de decisiones con evidencia empírica (Agresti, 2018; Casella y Berger, 2021).

1.2.4. Identificación de conceptos básicos en un contexto de realidad

La estadística facilita la recolección, organización, análisis e interpretación de datos, lo que permite presentar la información de manera clara, objetiva y comprensible. Aun cuando los datos sean limitados y los resultados obtenidos representen aproximaciones, el enfoque estadístico posibilita la formulación de conclusiones fundamentadas. Debido a su carácter transversal, la estadística desempeña un papel fundamental en múltiples actividades humanas. Muchas de las decisiones relevantes en ámbitos como la salud, la economía, la educación, la política y la industria se sustentan en el análisis estadístico de datos previamente recopilados (Agresti, 2018; Triola, 2020). Entre las aplicaciones más comunes de la estadística se encuentran:

- Censos de población, empleados para conocer la distribución demográfica y las características socioeconómicas de los habitantes de un país.
- Estudios sobre la canasta básica, que permiten estimar el costo de los bienes y servicios esenciales para una población.
- Cálculo de índices de inflación, utilizado para medir la variación de precios en el tiempo y su impacto en la economía.
- Estimación de tasas de mortalidad, fundamental en estudios epidemiológicos y de salud pública.
- Análisis epidemiológicos, aplicados al estudio de la distribución y los determinantes de enfermedades en poblaciones humanas.
- Determinación de primas de seguros, con base en modelos de riesgo y cálculos actuariales.
- Establecimiento de tarifas de servicios, mediante análisis de costos, demanda y sostenibilidad económica.
- Estudio de preferencias electorales, orientado a predecir comportamientos de voto y diseñar estrategias políticas.

- Medición de audiencias en medios de comunicación, aplicada para evaluar el alcance e impacto de los contenidos difundidos.

La diversidad de estas aplicaciones evidencia la transversalidad de la estadística y su papel estratégico en la planificación, la gestión y la toma de decisiones en múltiples sectores (Triola, 2020; Pagano et al., 2022).

Campos de aplicación

Aunque con frecuencia se asocia la estadística a estudios de carácter demográfico, económico o sociológico, esta disciplina ha adquirido también gran relevancia en áreas científicas como la medicina, la biología, la física y la ingeniería, en los que se integra como una herramienta metodológica esencial para la construcción de modelos explicativos que describen y predicen el comportamiento de fenómenos naturales, sociales o tecnológicos (Agresti, 2018; Pagano et al., 2022). A continuación, se presentan dos ejemplos de aplicación contextualizada:

Ejemplo empresarial

En el ámbito de las microempresas, es frecuente la ausencia de registros sistemáticos de tiempos y movimientos de producción, lo que dificulta la estimación precisa de costos y la estandarización de procesos. La estadística permite establecer tiempos mínimos, máximos y promedio en la elaboración de productos, optimizando así el control de calidad y reduciendo los gastos operativos.

Ejemplo científico

En el campo de la biología experimental, la estadística facilita el análisis cuantitativo de muestras y la interpretación de resultados en condiciones de laboratorio. Mediante el uso de tablas de frecuencia, representaciones gráficas y modelos probabilísticos, es posible identificar patrones de comportamiento molecular, celular o genético, aspectos fundamentales para el avance científico (Triola, 2020).

1.2.5. Aplicaciones cotidianas de la estadística en el área médica

La estadística proporciona herramientas metodológicas necesarias para el diseño, análisis e interpretación de investigaciones clínicas, epidemiológicas, genéticas y sanitarias. Su utilidad abarca desde la formulación de diagnósticos y el desarrollo de pruebas médicas hasta la planificación de servicios de salud y la evaluación de intervenciones terapéuticas, contribuyendo a la medicina basada en evidencia y a la optimización de recursos sanitarios (Gordis, 2020; Pagano et al., 2022; Figura 5).

Figura 5

Aplicaciones de la estadística en el área médica



Fuente: elaboración propia (2025).

Estudios epidemiológicos

La epidemiología se fundamenta en el análisis cuantitativo de la frecuencia y la distribución de enfermedades en una población. En este contexto, la estadística estudia patrones de aparición, evalúa factores de riesgo y diseña estrategias de prevención y control. Entre sus aplicaciones destacan el análisis de brotes y la medición de indicadores epidemiológicos —prevalencia e incidencia— (Gordis, 2020; Pagano et al., 2022).

a. Análisis de brotes de enfermedades

- **Rastreo de brotes:** la estadística permite describir patrones de propagación de enfermedades infecciosas, facilitando la localización de focos de transmisión. A través de la recopilación y el análisis de datos sobre casos confirmados, contactos cercanos, factores ambientales y movilidad poblacional, es posible trazar cadenas de contagio, delimitar zonas de alto riesgo y diseñar intervenciones focalizadas para contener la diseminación.
- **Modelos predictivos:** los modelos estadísticos permiten simular escenarios futuros del comportamiento de un brote o epidemia. Durante la pandemia por la COVID-19, dichos modelos fueron fundamentales para estimar el número esperado de casos, predecir puntos críticos de saturación hospitalaria y evaluar el impacto potencial de medidas como los confinamientos, la vacunación masiva o el distanciamiento social (Kucharski, 2020).

b. Estudios de prevalencia e incidencia

- Prevalencia: es la proporción de individuos que presentan una enfermedad o condición de salud específica en un momento determinado. Este indicador refleja la carga total de enfermedad en una población y se utiliza comúnmente en estudios transversales. Su análisis permite estimar la magnitud del problema de salud y orientar los recursos hacia su atención.
- Incidencia: es el número de casos nuevos de una enfermedad que ocurren en una población durante cierto tiempo. Este indicador resulta útil para comprender la dinámica y velocidad de propagación de enfermedades, en especial en estudios de cohorte. El cálculo de la incidencia permite monitorear tendencias, identificar brotes incipientes y evaluar la eficacia de intervenciones preventivas (Rothman et al., 2021).

Ensayos clínicos

Los ensayos clínicos constituyen el método estándar para evaluar la seguridad, eficacia y tolerabilidad de intervenciones terapéuticas, preventivas o diagnósticas en humanos. La estadística se aplica en todas las fases del proceso, desde el diseño del estudio hasta la interpretación de los resultados. Esto ayuda a que las conclusiones sean válidas, confiables y aplicables a la población objetivo (Friedman et al., 2015; Piantadosi, 2017).

a. Diseño y análisis de ensayos clínicos

- Diseño experimental: el diseño estadístico de un ensayo clínico implica determinar el tamaño óptimo de la muestra, definir los grupos de comparación (tratamiento vs. control) e implementar mecanismos de aleatorización, con el fin de minimizar sesgos, controlar variables de confusión y maximizar la potencia estadística del estudio. La aleatorización, en particular, permite distribuir equitativamente las características de los participantes entre los grupos, favoreciendo la comparabilidad y la validez interna.
- Análisis de datos: los datos se analizan durante y al finalizar el ensayo para establecer si las diferencias observadas entre grupos son atribuibles al efecto del tratamiento o al azar. Entre las técnicas más empleadas se incluyen las pruebas de hipótesis, los intervalos de confianza y las estimaciones de riesgo relativo o diferencias de medias. Estas herramientas permiten determinar la significancia estadística y evaluar la relevancia clínica de los resultados, lo que ayuda en la toma de decisiones terapéuticas y regulatorias (Pagano et al., 2022).

b. Metaanálisis

- **Combinación de resultados:** el metaanálisis es una técnica estadística que permite integrar los resultados de múltiples estudios independientes sobre una misma intervención, con el propósito de obtener una estimación global más precisa y robusta del efecto terapéutico. Esta técnica resulta útil cuando los estudios individuales presentan tamaños muestrales reducidos, resultados inconsistentes o niveles moderados de evidencia. Mediante modelos de efectos fijos o aleatorios, el metaanálisis pondera los resultados en función de su calidad metodológica y precisión, generando conclusiones más sólidas para la práctica clínica basada en evidencia (Higgins et al., 2021).

Diagnóstico médico

En el ámbito clínico, las pruebas diagnósticas desempeñan un papel esencial en la identificación precisa de enfermedades y condiciones de salud. La estadística contribuye de manera decisiva al desarrollo, validación y evaluación de estas pruebas, garantizando su eficacia para distinguir entre individuos enfermos y sanos, y facilitando así una adecuada toma de decisiones médicas (Fletcher et al., 2019; Zhou et al., 2023).

a. Desarrollo y validación de pruebas diagnósticas

- **Sensibilidad:** se define como la proporción de individuos verdaderamente enfermos que son correctamente identificados como positivos por la prueba. Este indicador mide la capacidad del instrumento para minimizar los falsos negativos. Una alta sensibilidad es especialmente deseable en enfermedades graves o contagiosas, en las que no detectar casos puede tener consecuencias clínicas o epidemiológicas significativas.
- **Especificidad:** corresponde a la proporción de individuos sanos que son correctamente identificados como negativos. Este parámetro evalúa la capacidad de la prueba para evitar falsos positivos, reduciendo así diagnósticos erróneos que podrían derivar en ansiedad, tratamientos innecesarios o sobreutilización de recursos.
- **Curvas ROC:** constituyen una herramienta gráfica que permite evaluar el desempeño de una prueba diagnóstica en distintos puntos de corte. Estas curvas relacionan la sensibilidad (tasa de verdaderos positivos) con el complemento de la especificidad (tasa de falsos positivos) para cada valor posible de la prueba. El área bajo la curva (AUC) proporciona una medida global de la capacidad discriminativa de la prueba. Un AUC cercano a 1

indica una prueba altamente precisa, mientras que un AUC de 0.5 refleja un desempeño equivalente al azar (Šimundić, 2009; Zhou et al., 2023).

Gestión sanitaria

La estadística desempeña un papel estratégico en la planificación y evaluación de los servicios de salud, al identificar necesidades, distribuir recursos y valorar el impacto de las intervenciones sanitarias. Esta función resulta esencial para garantizar la equidad en el acceso, la calidad de la atención y la sostenibilidad de los sistemas de salud (Katz, 2019; WHO, 2021).

a. Planificación y evaluación de servicios de salud

- **Análisis de necesidades:** el análisis estadístico posibilita determinar con precisión las necesidades sanitarias de una población específica. A través del estudio de indicadores clave, como tasas de hospitalización, morbilidad y mortalidad, es posible identificar deficiencias en la cobertura de servicios, detectar zonas con alta demanda insatisfecha y anticipar requerimientos futuros de infraestructura, personal o equipamiento. Por ejemplo, un incremento sostenido en las hospitalizaciones por enfermedades respiratorias en una región podría justificar la apertura de nuevas unidades médicas especializadas.
- **Evaluación de programas:** la estadística facilita la valoración objetiva de programas e intervenciones de salud pública mediante el uso de diseños comparativos, análisis longitudinales y estudios de impacto. La comparación de datos previos y posteriores a una intervención —como una campaña de vacunación o un programa de promoción de estilos de vida saludables— permite determinar su efectividad. Estos resultados constituyen evidencia empírica para la toma de decisiones, la reasignación de recursos y el rediseño de estrategias cuando sea necesario (Frieden, 2017).

b. Control de calidad

- **Monitoreo de indicadores de calidad:** los hospitales y clínicas emplean métodos estadísticos para supervisar indicadores de calidad de la atención, como las tasas de infecciones nosocomiales, la frecuencia de errores médicos y los tiempos de espera. Este monitoreo permite identificar áreas de mejora y diseñar intervenciones correctivas, lo que contribuye a incrementar la seguridad del paciente y optimizar la eficiencia de los servicios sanitarios (Klijs et al., 2021).

Investigación genética y genómica

La estadística ha adquirido un papel central en la investigación genética y genómica, disciplinas que exploran la estructura, función y variabilidad del material genético en los organismos vivos. El análisis estadístico en este campo permite identificar asociaciones entre variantes genéticas y enfermedades, así como comprender los mecanismos moleculares implicados en la expresión génica. Los métodos estadísticos avanzados han posibilitado avances trascendentes en medicina personalizada, diagnóstico molecular y terapias dirigidas (Albalwy et al., 2021; Roberts et al., 2024).

a. Estudios de Asociación Genómica (GWAS)

- **Identificación de genes asociados:** los estudios de asociación del genoma completo (Genome-Wide Association Studies, GWAS) emplean herramientas estadísticas para detectar correlaciones entre millones de variantes genéticas, en particular los polimorfismos de un solo nucleótido (SNPs), y la presencia de enfermedades o características fenotípicas. Estos estudios analizan grandes cohortes de individuos, comparando la frecuencia de variantes entre grupos con y sin una determinada condición. Los resultados permiten identificar loci genéticos de interés, que pueden convertirse en candidatos para investigaciones funcionales y futuras aplicaciones clínicas (Tam et al., 2019).
- **Riesgo genético:** una vez identificadas las variantes asociadas a una enfermedad, se aplican modelos estadísticos para calcular el riesgo genético individual. Este riesgo suele estimarse mediante la suma ponderada de múltiples variantes genéticas, lo que posibilita clasificar a los individuos según su predisposición genética a desarrollar enfermedades como cáncer, diabetes tipo 2 o trastornos cardiovasculares. Esta información constituye un elemento clave para la prevención personalizada y el diseño de intervenciones tempranas (Sims et al., 2020).

b. Análisis de Expresión Génica

- **Microarrays y Secuenciación de ARN:** el análisis de la expresión génica busca cuantificar los niveles de transcripción de genes en condiciones fisiológicas o patológicas. Para ello se utilizan tecnologías como los microarreglos de ADN y la secuenciación de ARN (RNA-seq), que generan grandes volúmenes de datos. El empleo de técnicas estadísticas permite comparar perfiles de expresión entre grupos, identificar genes diferencialmente expresados y explorar patrones globales mediante análisis de componentes principales, clustering jerárquico o modelos lineales

generalizados. Estos métodos facilitan la comprensión de la regulación génica y su implicación en procesos biológicos como la inflamación, proliferación celular o apoptosis (Conesa et al., 2016; Satam et al., 2023).

Salud pública y prevención

La estadística en el ámbito de la salud pública permite vigilar, analizar y evaluar el estado de salud de las poblaciones, así como diseñar e implementar estrategias de prevención y control de enfermedades. Mediante el análisis cuantitativo se identifica la aparición de riesgos sanitarios, se evalúa el impacto de las intervenciones implementadas y se formulan políticas públicas (Frieden, 2017; Gordis, 2020).

a. Vigilancia de la salud

- **Monitoreo de enfermedades:** los sistemas de vigilancia sanitaria recopilan de manera sistemática datos sobre morbilidad, mortalidad y otros eventos de salud, con el propósito de identificar brotes, detectar tendencias epidemiológicas y orientar la respuesta sanitaria. La estadística permite reconocer patrones, establecer umbrales de alerta y activar mecanismos de respuesta ante emergencias sanitarias, como epidemias o desastres naturales (Thacker y Qualters, 2018).
- **Evaluación de intervenciones:** los métodos estadísticos son fundamentales para valorar la efectividad de las intervenciones de salud pública, por medio de diseños pretest-postest, análisis de series temporales y estudios cuasiexperimentales; es posible medir los resultados de campañas de prevención, por ejemplo, programas de cesación del tabaquismo, educación sobre VIH/SIDA o estrategias de vacunación. La comparación de indicadores antes y después de la intervención proporciona evidencia cuantificable de su impacto (Rothman et al., 2021).

b. Estudios de cohorte y casos y controles

- **Estudios de cohorte:** consisten en el seguimiento longitudinal de un grupo de individuos con características comunes, con el objetivo de observar la relación entre determinadas exposiciones (dieta, actividad física o exposición a sustancias) y la aparición de enfermedades a lo largo del tiempo. Los métodos estadísticos permiten estimar riesgos relativos, calcular tasas de incidencia y controlar variables confusoras mediante análisis multivariados. Estos estudios otorgan una elevada capacidad para establecer relaciones causa-efecto temporales (Gordis, 2020).

- Estudios de casos y controles: se basan en la comparación retrospectiva entre un grupo de individuos con una enfermedad (casos) y otro grupo sin la enfermedad (controles), con el fin de identificar factores de riesgo asociados. La estadística es esencial tanto en el diseño del estudio —en la selección de controles adecuados— como en el análisis, mediante la estimación de razones de momios (odds ratios). Esta estrategia permite inferir asociaciones relevantes para la prevención y el diagnóstico temprano de enfermedades (Rothman et al., 2021).

Medicina personalizada

La medicina personalizada, también denominada medicina de precisión, constituye un enfoque emergente que integra características individuales, como el perfil genético, epigenético, molecular, ambiental y clínico para diseñar estrategias de prevención, diagnóstico y tratamiento adaptadas a cada paciente. En este paradigma, la estadística desempeña un papel central al posibilitar el desarrollo y validación de modelos predictivos capaces de anticipar respuestas terapéuticas y pronósticos clínicos a partir del análisis de grandes volúmenes de datos heterogéneos (Collins y Varmus, 2015; Hasin et al., 2017).

a. Modelos predictivos

- Predicción de respuesta al tratamiento: mediante la regresión logística, modelos de supervivencia, redes neuronales y algoritmos de aprendizaje automático (machine learning), es posible construir modelos que predicen cómo responderá un paciente a un tratamiento determinado. Estos modelos se elaboran a partir de bases de datos que integran información clínica, genética, metabólica y de biomarcadores, estimando la probabilidad de eficacia terapéutica, anticipando la aparición de efectos adversos y seleccionando la opción más segura y efectiva para cada individuo (Kourou et al., 2015).
- Terapias personalizadas: los datos estadísticos derivados de investigaciones clínicas y estudios poblacionales permiten establecer patrones terapéuticos aplicables a subgrupos de pacientes con características compartidas. Con base en esta evidencia, es posible personalizar los tratamientos para optimizar sus beneficios clínicos, reducir la toxicidad, evitar el sobretratamiento y mejorar la adherencia terapéutica. Este enfoque favorece una atención centrada en el paciente, orientada por los principios de eficacia individual y sostenibilidad del sistema sanitario (Weissler et al., 2021).

La estadística, por su capacidad para procesar, integrar e interpretar datos complejos provenientes de múltiples fuentes, es indispensable en la medicina moderna. Su aplicación mejora la calidad del cuidado clínico y representa un avance significativo hacia un modelo de salud más preciso, preventivo y participativo, con un impacto positivo en la salud individual y colectiva.

Cierre de la unidad

A lo largo de la presente unidad, se abordó la importancia de la estadística y su aplicabilidad en diversos contextos del ejercicio profesional; se reconoció su utilidad como herramienta fundamental para la recolección, organización, análisis e interpretación de datos con fines diagnósticos, preventivos y de investigación. En la siguiente unidad, se profundizará en el uso de tablas de frecuencias y en la elaboración de representaciones gráficas, con el propósito de visualizar el comportamiento de los datos mediante el análisis de un caso específico.

Unidad

Representación numérica
y gráfica de datos

Presentación de la unidad

En la unidad anterior, se establecieron los fundamentos teóricos de la estadística, concebida como un conjunto de técnicas orientadas a describir datos y facilitar la toma de decisiones en contextos caracterizados por la incertidumbre o la información incompleta. Para el cumplimiento de dichas funciones, la estadística se clasifica en descriptiva e inferencial. La presente unidad se centra en el estudio de la estadística descriptiva, con el propósito de desarrollar habilidades para organizar, presentar e interpretar datos muestrales. Antes de abordar los procedimientos específicos, se introducen los métodos y fuentes para la obtención de los datos que serán objeto de análisis.

Objetivos

Al finalizar el estudio de esta unidad, el estudiante será capaz de:

- Identificar los conceptos fundamentales de la estadística descriptiva.
- Analizar datos cualitativos y cuantitativos representados en tablas y figuras.

Competencia específica

A partir de la integración de los conocimientos declarativos, procedimentales, actitudinales y contextuales abordados en esta unidad, se espera que el estudiante desarrolle la capacidad de:

- Analizar bases de datos para identificar fenómenos e interpretar tablas o figuras.

2.1. Recolección de datos

La recolección de datos constituye una de las etapas fundamentales en todo proceso de investigación científica y estadística. Su propósito es obtener información precisa, válida y confiable que permita describir, analizar o inferir características de una población o fenómeno de interés. La calidad de los datos recolectados influye en la validez de los resultados y en la solidez de las conclusiones derivadas del análisis estadístico (Creswell y Creswell, 2018; Lohr, 2022; Figura 6).

Figura 6

Técnicas para recolección de datos



Fuente: elaboración propia (2025).

Para recolectar los datos, resulta indispensable definir las variables de interés, establecer la población objetivo y seleccionar el método más adecuado en función del tipo de estudio, el diseño metodológico y los recursos disponibles. En términos generales, los datos pueden obtenerse mediante diversas técnicas, por ejemplo:

Censo

El censo recolecta datos de todos los elementos de la población o universo de estudio. Esta metodología se emplea cuando se requiere información exhaustiva y precisa de cada unidad que integra la población. Los estudios demográficos nacionales son un ejemplo de ello. Los censos permiten determinar con exactitud la cantidad de habitantes, su distribución geográfica y sus características socioeconómicas y demográficas (Clynes et al., 2020). Para que un censo sea metodológicamente válido, debe cumplir con dos condiciones:

- **Universalidad:** todos los elementos de la población deben estar incluidos en el proceso de recolección de datos, evitando omisiones o duplicaciones.
- **Simultaneidad:** la obtención de los datos debe realizarse en un periodo específico y breve para minimizar sesgos derivados de variaciones temporales que alteren los resultados.

Encuesta

La encuesta es una técnica de recolección de datos aplicada a una muestra representativa de la población para inferir características del universo de estudio sin necesidad de evaluar a todos sus elementos. Las encuestas se utilizan en estudios sociales, de mercado, de salud pública, educativos, entre otros (Fowler, 2014). Este método se basa en un conjunto estructurado de preguntas, que pueden clasificarse en:

- Preguntas cerradas: muestran opciones de respuesta predefinidas que facilitan la tabulación y el análisis estadístico.
- Preguntas abiertas: permiten respuestas libres que aportan mayor riqueza cualitativa, aunque requieren técnicas más complejas para su análisis.

La calidad de una encuesta depende del diseño del instrumento, la claridad de las preguntas, la capacitación del personal encargado de aplicarla y el rigor en la selección de la muestra. Una encuesta bien diseñada y aplicada puede proporcionar información valiosa para la toma de decisiones, siempre que se respete el principio de representatividad muestral (Groves et al., 2009).

Experimento

El experimento es una técnica de recolección de datos que se fundamenta en la manipulación deliberada de una o más variables independientes, con el propósito de observar su efecto sobre variables dependientes bajo condiciones controladas. A diferencia del censo y la encuesta, el experimento permite establecer relaciones causales entre fenómenos (Shadish et al., 2002). Existen dos enfoques comunes en la experimentación:

- Exploratorio: se aplica cuando no se dispone de una hipótesis definida. Se observa un fenómeno con el fin de identificar patrones o comportamientos sin predicciones. Ejemplo: evaluar la percepción de sabor en una muestra de alimentos sin hipótesis previa sobre cuál será considerado más salado.
- Confirmatorio: se utiliza cuando existe una hipótesis previa sustentada en evidencia teórica o empírica. Ejemplo: un experimento diseñado para medir la velocidad de la luz parte de la hipótesis de que esta es de 300 000 km/s, con el fin de confirmar o refutar dicha afirmación bajo condiciones controladas.

La validez de un experimento depende de la manipulación correcta de las variables, el control de factores externos y la replicabilidad de los resultados. En el ámbito científico, los experimentos constituyen un pilar metodológico para la validación de teorías y la generación de conocimiento objetivo.

2.2. Organización de datos y distribución de frecuencias

La estadística descriptiva, como disciplina orientada a la síntesis y análisis de datos, tiene como finalidad estructurar, representar y resumir la información proveniente de una población o muestra, sin recurrir a procesos de inferencia. A través de este enfoque, es posible transformar conjuntos extensos o desordenados de datos en formatos comprensibles, lo que facilita su análisis, interpretación y aplicación en la toma de decisiones (Triola, 2020).

Los datos pueden organizarse en distribuciones de frecuencia, que permiten observar patrones, tendencias y concentraciones. Estas distribuciones pueden presentarse mediante tablas o figuras, para datos no agrupados y agrupados, en función del tamaño del conjunto y su naturaleza (Levine et al., 2020). Las distribuciones de frecuencia mejoran la visualización de la información antes de aplicar técnicas estadísticas como las medidas de tendencia central, dispersión y forma (Figura 7).

Figura 7

Organización de datos y distribución de frecuencias



Fuente: elaboración propia (2025).

2.2.1. Datos no agrupados

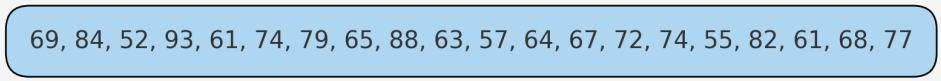
Los datos no agrupados son observaciones que, por su escasa cantidad o por la claridad con que se presentan, no requieren ser organizadas en intervalos de clase. De manera general, corresponden a conjuntos con un número de observaciones menor a 30, aunque este umbral puede variar según el contexto analítico (Triola, 2020). En estos casos, los datos pueden organizarse de forma ascendente o descendente para análisis preliminares de dispersión, valores atípicos y distribución.

Los datos no agrupados suelen representarse en una gráfica de tallo y hojas (stem-and-leaf plot) que permite visualizar la distribución de los datos en orden y conservar los valores individuales de cada observación. Esto es útil en análisis exploratorios (Levine et al., 2020).

Ejemplo: considérese un conjunto de 20 calificaciones obtenidas por participantes en una prueba de coordinación física, posterior al consumo de alcohol en una proporción equivalente al 0.1 % de su peso corporal. Los resultados fueron los siguientes (Figura 8):

Figura 8

Conjunto de calificaciones

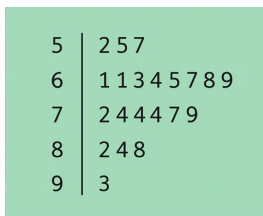


Fuente: elaboración propia (2025).

Para organizar estos datos mediante la gráfica de tallo y hojas, se separan los valores por decenas (tallos) y unidades (hojas). Por ejemplo, el número 65 se descompone como 6 | 5, donde el número “6” representa el tallo (decena) y el “5” la hoja (unidad). La gráfica resultante se muestra a continuación (Figura 9):

Figura 9

Gráfica de tallo y hojas



Fuente: Elaboración propia (2025).

En esta representación, cada fila comienza con un tallo seguido por una barra vertical. A la derecha se disponen las hojas correspondientes a los valores individuales. Este método permite identificar la concentración y dispersión de las observaciones.

2.2.2. Datos agrupados

Cuando el volumen de datos recogidos en un estudio estadístico es considerable, su organización individual resulta ineficiente e incluso confusa para el análisis. En tales casos, se recurre a la técnica de agrupación de datos, que clasifica las observaciones en clases o intervalos de clase. Esta modalidad de presentación permite resumir la información, visualizar su distribución general y facilitar la interpretación de patrones o tendencias, sin perder la representatividad del conjunto original (Levine et al., 2020).

Los datos agrupados resultan útiles cuando el número de observaciones supera los 30 registros o cuando se pretende mostrar el comportamiento global de los datos, con el objetivo de identificar concentraciones, dispersión o irregularidades.

Mediante la construcción de tablas de frecuencia y figuras (histogramas, polígonos de frecuencia u ojivas), los datos agrupados se convierten en un recurso esencial en la estadística descriptiva (Triola, 2020). Para estructurar una tabla de frecuencias con datos agrupados es necesario comprender los siguientes conceptos:

Frecuencia: la frecuencia se define como el número de observaciones que pertenecen a cada intervalo de clase (el conteo directo de los datos que se encuentran dentro de los límites establecidos para una clase específica). Este concepto permite describir la distribución de los datos e identificar tendencias, concentraciones y dispersiones en los valores observados (Triola, 2020). La organización de los datos en una tabla de frecuencias proporciona una representación sistemática de la información, lo que facilita su análisis y la elaboración de representaciones gráficas en las que se visualice el comportamiento de los datos (Levine et al., 2020; Figura 10).

Figura 10

Frecuencia

Frecuencia
Número de veces que aparece un dato en un conjunto.

Dato	Frecuencia
2	2 (porque se repite 2 veces)
3	3 (porque se repite 2 veces)
5	2 (porque se repite 2 veces)
7	4 (porque se repite 4 veces)
Total de datos (N)	11

Fuente: elaboración propia (2025).

Frecuencia acumulada: es una medida estadística que representa la suma progresiva de las frecuencias absolutas en una distribución y su cálculo es secuencial. En consecuencia, en la primera clase, la frecuencia acumulada coincide con la frecuencia absoluta, dado que no existen registros previos (Levine et al., 2020; Triola, 2020).

La frecuencia acumulada permite analizar la distribución de los datos, facilitando la identificación de umbrales, percentiles y medidas de posición. Además, constituye la base para la frecuencia relativa acumulada y para gráficas como la ojiva, utilizada para visualizar la proporción de datos que no exceden cierto valor (Figura 11).

Figura 11

Frecuencia acumulada



Fuente: elaboración propia (2025).

Frecuencia relativa: es una medida estadística que expresa la proporción de veces que ocurre un valor o clase dentro de un conjunto de datos. Se obtiene al dividir la frecuencia absoluta de un valor o intervalo, es decir, el número de observaciones entre el total de datos de la muestra o población. Este indicador puede representarse en forma decimal, fraccionaria o porcentual; ofrece una visión comparativa del peso de cada valor en el conjunto de observaciones. Su uso es relevante en tablas de frecuencias y gráficas (Levine et al., 2020; Triola, 2020; Figura 12).

Figura 12
Frecuencia relativa



Fuente: elaboración propia (2025).

Frecuencia relativa acumulada: es una medida estadística que representa la suma progresiva de las frecuencias relativas de un conjunto de datos, organizados de menor a mayor, para identificar de forma acumulativa la proporción de observaciones que no exceden un determinado valor o intervalo de clase. El último valor de la columna de frecuencias relativas acumuladas siempre es igual a 1 (o al 100 % cuando se expresa en porcentaje), porque se ha considerado la totalidad de las observaciones (Levine et al., 2020; Triola, 2020). Esta medida es útil en el cálculo de percentiles, cuartiles, medidas de posición y gráficas acumulativas (Figura 13).

Figura 13
Frecuencia relativa acumulada



Fuente: elaboración propia (2025).

Distribución de frecuencias

En el análisis estadístico de grandes conjuntos de datos, la agrupación de observaciones en intervalos simplifica la información sin sacrificar su valor analítico. Este procedimiento da lugar a la distribución de frecuencias, que facilita la organización y representación de los datos, permitiendo identificar patrones generales, tendencias y comportamientos dentro de una muestra o población (Triola, 2020).

La distribución de frecuencias consiste en organizar las observaciones en clases o intervalos de clase, a los que se asigna su respectiva frecuencia, entendida como la cantidad de veces que los datos se ubican en cada intervalo. Esta estructura es el punto de partida para el cálculo de medidas estadísticas más complejas como las de tendencia central (media, mediana, moda) y de dispersión (rango, varianza, desviación estándar) (Levine et al., 2020).

Importancia de la distribución de frecuencias

El uso de una distribución de frecuencias proporciona múltiples ventajas metodológicas en el análisis estadístico:

- Síntesis de datos: permite resumir conjuntos extensos de información en una estructura concisa y comprensible, evitando la dispersión de valores individuales.
- Visualización gráfica: facilita la representación del comportamiento de las variables mediante recursos como histogramas, polígonos de frecuencia y ojivas, que permiten identificar rápidamente tendencias o irregularidades.
- Interpretación estadística: mejora la comprensión de los datos al evidenciar patrones de concentración, dispersión o simetría, lo que contribuye a un análisis más preciso y objetivo.
- Base para análisis avanzados: constituye el punto de partida para estudios posteriores, tales como análisis de correlación, regresión e inferencia estadística, en los que se requiere un conocimiento preliminar de la distribución de los datos (Levine et al., 2020; Triola, 2020).

En este sentido, la distribución de frecuencias organiza la información y también sustenta procesos analíticos de mayor complejidad.

Componentes fundamentales sobre intervalos en una distribución de frecuencias

La organización de datos en una tabla de distribución de frecuencias permite estructurar los datos de manera coherente, homogénea y metodológicamente válida (Levine et al., 2020; Triola, 2020). A continuación, se describen los componentes que integran una tabla de distribución de frecuencias para datos agrupados:

Intervalo o rango: el intervalo, también denominado rango de clase, corresponde al conjunto de valores numéricos comprendidos entre dos límites específicos: inferior y superior. Dicho conjunto representa un tramo continuo de la escala de medición dentro del cual se agrupan observaciones con características cuantitativas similares. Desde una perspectiva general, cada intervalo contiene una parte del total de datos observados. Los intervalos se complementan para abarcar todo el rango de variación de la variable analizada.

Intervalo de clase: el intervalo de clase designa el rango utilizado para agrupar datos en una distribución de frecuencias. Cada uno está delimitado por dos valores extremos:

- El límite inferior, que representa el valor mínimo del intervalo.
- El límite superior, que señala el valor máximo que el intervalo puede contener (según las reglas de inclusión, habitualmente cerrado por la izquierda y abierto por la derecha).

Los intervalos de clase deben ser mutuamente excluyentes y colectivamente exhaustivos para que no se traslapen y representen la totalidad del conjunto de datos.

Amplitud del intervalo: la amplitud del intervalo, también denominada longitud o ancho de clase, es la magnitud del tramo numérico que abarca el intervalo de clase. Se calcula como la diferencia entre el límite superior y el inferior del intervalo. La elección de una amplitud constante favorece la uniformidad en la presentación y el análisis de los datos. No obstante, en situaciones específicas puede justificarse el uso de amplitudes desiguales, siempre que se mantenga la consistencia metodológica.

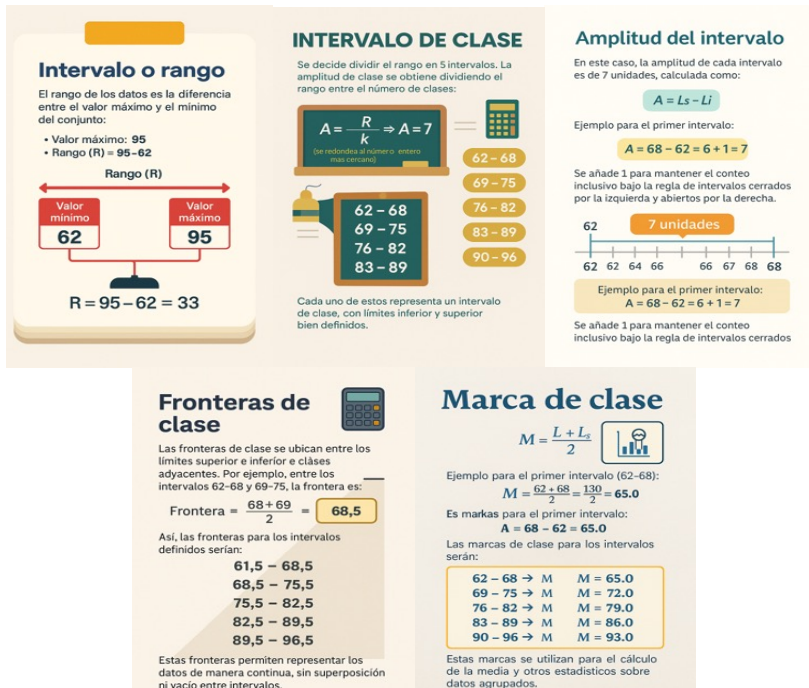
Fronteras de clase: son valores que se ubican entre los límites de intervalos contiguos y se utilizan cuando se presentan datos continuos. Estas fronteras dividen el espacio numérico entre un intervalo y otro, eliminando discontinuidades entre los límites superiores e inferiores de clases consecutivas. Se calculan como el promedio entre el límite superior de un intervalo y el inferior del siguiente; su aplicación es fundamental en la construcción de histogramas y en el análisis gráfico de distribuciones continuas.

Marca de clase: la marca de clase, también conocida como punto medio del intervalo, corresponde al valor central de un intervalo de clase; se obtiene como el promedio aritmético entre el límite inferior y el superior del intervalo. Este valor es un representante numérico de las observaciones en el cálculo de medidas de tendencia central como la media aritmética para datos agrupados y en diversos procedimientos estadísticos que requieren un valor puntual por intervalo. Por tanto, sintetiza los datos distribuidos en rangos.

Ejemplo: considérese el siguiente conjunto de datos que representa las calificaciones obtenidas por 30 estudiantes en una evaluación de conocimientos generales: 62, 68, 70, 75, 64, 81, 85, 87, 91, 93, 95, 72, 74, 77, 79, 83, 66, 69, 71, 76, 80, 88, 84, 86, 89, 90, 92, 94, 73, 78. Para organizar estos datos y facilitar su análisis, se construye una tabla de distribución de frecuencias agrupando las observaciones en intervalos. A continuación, se explican paso a paso sus componentes (Figura 14).

Figura 14

Componentes fundamentales sobre intervalos en una distribución de frecuencias



Fuente: elaboración propia (2025).

Elaboración de una distribución de frecuencias por intervalos

La organización de datos cuantitativos en una distribución de frecuencias por intervalos resume conjuntos extensos de observaciones mediante su agrupación en rangos homogéneos denominados intervalos de clase. Este procedimiento facilita el análisis, la comparación y la representación gráfica de datos de variables continuas o series de datos discretos de gran magnitud (Levine et al., 2020; Triola, 2020).

Para construir una distribución de frecuencias por intervalos, se requiere seguir una secuencia metodológica que asegure la coherencia en la organización de los datos y la validez de las interpretaciones derivadas. Los pasos principales incluyen:

Determinación del rango de los datos

El primer paso consiste en calcular el rango (R), el cual representa la extensión total de los valores que adopta la variable en estudio. Se obtiene como la diferencia entre el valor máximo y el valor mínimo observados:

$$R = \text{Valor máximo} - \text{valor mínimo}$$

Este valor proporciona una base cuantitativa para determinar el número de intervalos necesarios y la amplitud de cada uno (Triola, 2020).

Estimación del número de intervalos o clases (k)

El siguiente paso consiste en definir el número de intervalos de clase (k) en los que se dividirá el rango total. Este puede establecerse de forma empírica usualmente entre 5 y 20 intervalos o calcularse mediante una fórmula teórica, como la regla de Sturges:

$$K = 1 + 3.322\log_{10}(n)$$

En este caso, n corresponde al número total de observaciones. El resultado se redondea al número entero más cercano. Este criterio busca un equilibrio entre el detalle analítico y la simplicidad interpretativa, evitando tanto la pérdida de información como la sobrecarga visual (Levine et al., 2020).

Cálculo de la amplitud del intervalo (A)

La amplitud de cada intervalo (A) se obtiene dividiendo el rango total entre el número de intervalos definidos:

$$A = R / k$$

Por lo general, el valor se redondea al número entero inmediato superior para que todos los datos queden comprendidos en los intervalos. La amplitud debe mantenerse constante en toda la distribución, salvo que se justifique el uso de intervalos abiertos o de amplitud variable.

Definición de los intervalos de clase

Los intervalos se construyen a partir del valor mínimo de la variable y se incrementan en función de la amplitud establecida. Los intervalos deben ser mutuamente excluyentes, de modo que cada observación pertenezca únicamente a un intervalo. Además, ser colectivamente exhaustivos, garantizando que todos los datos queden incluidos. La convención más utilizada es considerar los intervalos cerrados por la izquierda y abiertos por la derecha, es decir, incluyen el límite inferior, pero excluyen el límite superior. Ejemplo: $[60 - 69)$, $[69 - 78)$, $[78 - 87)$, etcétera.

Recopilación de frecuencias absolutas

Una vez definidos los intervalos, se contabiliza el número de observaciones en cada uno para obtener la frecuencia absoluta (f). Este procedimiento puede realizarse manualmente, mediante hojas de cálculo o a través de *software* estadístico. De estas frecuencias absolutas derivan las relativas y acumuladas.

Cálculo de frecuencias acumuladas y relativas

Una vez obtenidas las frecuencias absolutas (f), se obtienen:

- Frecuencia acumulada (F): es la suma progresiva de las frecuencias absolutas desde el primer intervalo hasta el actual. Permite identificar cuántas observaciones se encuentran por debajo de un valor específico.
- Frecuencia relativa (fr): expresa la proporción de observaciones de un intervalo respecto al total de datos. Se calcula mediante la fórmula: $fr = f/n$, donde f es la frecuencia absoluta del intervalo y n el número total de observaciones.
- Frecuencia relativa acumulada (Fr): representa la suma progresiva de las frecuencias relativas. Su valor final debe ser igual a 1 (o 100 % si se expresa en porcentaje).

Cálculo de marcas de clase

Para cada intervalo, se calcula la marca de clase (M) o punto medio que representa todas las observaciones comprendidas en dicho intervalo. Se obtiene mediante el promedio aritmético entre el límite inferior (Li) y el límite superior (Ls) de cada clase:

$$M = Li + Ls / 2$$

Las marcas de clase resultan fundamentales para el cálculo de medidas de tendencia central (como la media aritmética para datos agrupados) y de dispersión (como la varianza y la desviación estándar).

Presentación tabular de los resultados

Una vez obtenidos todos los componentes, se construye la tabla de distribución de frecuencias que integra la información de forma ordenada y sistemática. Los elementos básicos que deben incluirse son (Tabla 3):

Tabla 3. Presentación tabular de resultados

Intervalo de clase	f	F	fr	Fr	Marca de clase (M)
.....

Fuente: elaboración propia (2025).

La presentación tabular constituye el paso final de la organización de los datos y permite la elaboración de histogramas, polígonos de frecuencia y ojivas (Levine et al., 2020; Triola, 2020). Esta estructura permite una lectura clara y sistemática del comportamiento de la variable.

Representación gráfica complementaria (opcional)

Con el propósito de facilitar la comprensión visual de los datos, la estadística descriptiva recurre a diversas herramientas gráficas que permiten identificar patrones y estructuras en un conjunto de observaciones. Entre las más utilizadas se encuentran:

- **Histograma:** gráfico de barras contiguas cuya altura es proporcional a la frecuencia absoluta de cada intervalo de clase. Facilita la visualización de la distribución de los datos y la identificación de su forma (simetría, asimetría, sesgo).

- **Polígono de frecuencia:** representación mediante una línea quebrada que une las marcas de clase con sus respectivas frecuencias. Resulta útil para comparar distribuciones y detectar tendencias en los datos.
- **Ojiva:** curva acumulativa que refleja la frecuencia acumulada o relativa acumulada, permitiendo analizar la concentración de los datos y determinar percentiles o cuartiles de manera gráfica.

Estas representaciones gráficas cumplen un papel metodológico esencial al complementar las tablas de frecuencias, ya que facilitan la detección de características como la simetría, la dispersión, la presencia de valores atípicos y la forma general de la distribución (Triola, 2020; Grainge, 2023). Este procedimiento transforma un conjunto complejo de observaciones en una estructura organizada que permite su interpretación, comparación y análisis.

Ejemplo: se ha aplicado una prueba de razonamiento lógico a un grupo de 30 estudiantes universitarios. Las calificaciones obtenidas (en una escala de 0 a 100) fueron las siguientes: 56, 60, 65, 68, 70, 71, 72, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 88, 89, 90, 91, 92, 93, 94, 95, 96, 98. Se requiere elaborar una distribución de frecuencias por intervalos para organizar la información descriptiva.

Paso 1: determinar el rango de los datos

$$R = \text{Valor máximo} - \text{Valor mínimo} = 98 - 56 = 42$$

Paso 2: estimar el número de intervalos (k)

Se opta por 6 intervalos de clase

Paso 3: calcular la amplitud del intervalo (A)

$$A = R / k = 42 / 6 = 7$$

Para asegurar que todos los valores queden incluidos y se simplifique la agrupación, se redondea a ocho unidades de amplitud.

Paso 4: definir los intervalos de clase

Los intervalos se definen comenzando desde el valor mínimo (56) y sumando la amplitud (8) en cada intervalo:

$$56 - 63, 64 - 71, 72 - 79, 80 - 87, 88 - 95, 96 - 103$$

Nota: aunque la escala es de 0 a 100, el último intervalo se extiende a 103 para incluir el valor 98 con suficiente margen.

Paso 5: contar la frecuencia absoluta (f)

Una vez definidos los intervalos de clase, se contabiliza el número de observaciones en cada uno. Este conteo se denomina frecuencia absoluta (f) y constituye la base de la tabla de distribución de frecuencias (Levine et al., 2020; Triola, 2020; Tabla 4).

Tabla 4. Frecuencia absoluta

Intervalo de clase	f
56 – 63	2
64 – 71	4
72 – 79	6
80 – 87	8
88 – 95	8
96 – 103	2
Total (n)	30

Fuente: elaboración propia (2025).

Paso 6: Calcular frecuencias acumuladas y relativas

Una vez obtenidas las frecuencias absolutas (f), es necesario complementar la tabla de distribución de frecuencias con otras medidas, como:

- Frecuencia acumulada (F): suma progresiva de f
- Frecuencia relativa (fr): f / n
- Frecuencia relativa acumulada (Fr): suma progresiva de fr
- Marca de clase (M): promedio entre los límites inferior y superior de cada intervalo

Paso 7: construir la tabla de distribución de frecuencias

Una vez obtenidos los valores de la frecuencia absoluta (f), la frecuencia acumulada (F), la frecuencia relativa (fr), la frecuencia relativa acumulada (Fr) y la marca de clase (M), se integra la información en una tabla de distribución de frecuencias. Este formato tabular constituye la síntesis estructurada de los datos agrupados, permitiendo su análisis descriptivo, la construcción de representaciones gráficas (como histogramas, polígonos y ojivas) y el cálculo de medidas de tendencia central y dispersión (Tabla 5).

Tabla 5. Tabla de distribución de frecuencias

Intervalo de clase	f	F	fr	Fr	Marca de clase (M)
56–63	2	2	0.067	0.067	59.5
64–71	4	6	0.133	0.200	67.5
72–79	6	12	0.200	0.400	75.5
80–87	8	20	0.267	0.667	83.5
88–95	8	28	0.267	0.934	91.5
96–103	2	30	0.067	1.000	99.5

Fuente: elaboración propia (2025).

Paso 8: interpretación inicial

La interpretación preliminar de la tabla de distribución de frecuencias permite identificar patrones generales en los datos antes de proceder a cálculos más avanzados de estadística descriptiva e inferencial. A partir del análisis anterior, destaca lo siguiente:

- La mayor proporción de observaciones se ubica en los intervalos [80–89] y [90–99], lo que sugiere que la mayoría de los participantes alcanzó calificaciones altas en la prueba aplicada. Este hallazgo permite inferir un buen desempeño global del grupo analizado.
- El análisis de la columna de frecuencias acumuladas muestra que aproximadamente el 85 % de los estudiantes alcanzó una calificación igual o inferior a 89, lo que refleja una distribución sesgada hacia valores altos. Asimismo, el 55 % de los estudiantes obtuvo puntajes iguales o menores a 79, lo que indica que más de la mitad del grupo obtuvo un desempeño intermedio.
- La frecuencia relativa acumulada confirma la progresión porcentual de los datos, evidenciando que el 100 % de las observaciones se encuentra dentro de los cuatro intervalos establecidos, lo que valida la exhaustividad de la clasificación.
- La inclusión de la marca de clase proporciona un valor representativo de cada intervalo para el cálculo de medidas de tendencia central y para la construcción de modelos estadísticos que requieren un valor puntual de cada rango.

Paso 9 (opcional): representación gráfica

La información contenida en la tabla de distribución de frecuencias puede complementarse mediante representaciones gráficas para identificar las características clave de la distribución, tales como simetría, sesgo, concentración de observaciones o presencia de valores atípicos (Triola, 2020; Grainge, 2023).

Este ejemplo ilustra el procedimiento completo para construir una distribución de frecuencias por intervalos, desde el cálculo del rango hasta la interpretación de los resultados. Esta herramienta permite sintetizar datos cuantitativos extensos y facilitar su análisis comparativo.

Tablas

En el campo de la estadística, las tablas constituyen una herramienta esencial para la organización, presentación y análisis de datos tanto cuantitativos como cualitativos. Su empleo permite sintetizar grandes volúmenes de información en un formato estructurado y ordenado, lo que facilita la comprensión, la identificación de patrones y la interpretación de los datos (Grainge, 2023). El uso de tablas posibilita:

- Resumir la información de forma clara y accesible.
- Comparar variables o categorías con mayor precisión.
- Detectar tendencias, concentraciones o irregularidades en los datos.
- Servir de base para la elaboración de representaciones gráficas o el cálculo de medidas estadísticas.

Existen diversos tipos de tablas, cada una con estructura y función específicas; la selección del tipo adecuado dependerá de la naturaleza de los datos, del número de variables involucradas y de los objetivos del análisis. A continuación, se describen las principales modalidades de tablas empleadas en el tratamiento de datos estadísticos:

Tabla de datos

La tabla de datos o tabla simple proporciona una presentación directa y ordenada de los valores obtenidos en un proceso de observación, medición o registro. Este tipo de tabla no implica procesamiento estadístico alguno, ya que solo lista los datos tal como fueron recolectados, organizados en filas o columnas (Triola, 2020).

Su función principal es organizar y visualizar preliminarmente los datos, lo que facilita su revisión inicial y constituye el punto de partida para el desarrollo de representaciones estadísticas más elaboradas (Grainge, 2023). Las tablas de datos son de especial utilidad en la fase inicial de una investigación, ya que permiten:

- Revisar la completitud y consistencia de los datos recolectados.
- Detectar errores de registro o valores atípicos.
- Disponer la información en un formato estructurado que facilite su análisis posterior.

Ejemplo: se han registrado las temperaturas ambientales (en grados Celsius) en siete momentos distintos del día (Tabla 6).

Tabla 6. Tabla de datos de temperaturas ambientales

Hora	Temperatura (°C)
08:00	20.4
10:00	21.2
12:00	22.1
14:00	23.9
16:00	25.3
18:00	26.9
20:00	27.7

Fuente: elaboración propia (2025).

Esta tabla permite observar la evolución de la temperatura a lo largo del día.

Tabla de frecuencias

Esta constituye una forma estructurada y analítica para presentar datos, al contabilizar cuántas veces aparece cada valor observado. Este recurso resulta especialmente útil al trabajar con variables cualitativas o discretas, ya que permite visualizar la distribución de los datos e identificar con facilidad los valores más frecuentes (Triola, 2020).

En su forma más sencilla, la tabla de frecuencias incluye la frecuencia absoluta (f), que corresponde al número de veces que ocurre un valor en el conjunto. Sin embargo, para un análisis más profundo, también pueden incorporarse otras medidas complementarias, tales como:

- Frecuencia relativa (fr): proporción de observaciones que corresponden a un valor específico respecto al total de datos.
- Frecuencia acumulada (F): suma progresiva de las frecuencias absolutas hasta un punto.

- Frecuencia relativa acumulada (Fr): proporción acumulada de observaciones hasta un valor o categoría.

La incorporación de estas medidas permite sintetizar grandes volúmenes de información e identificar patrones, concentraciones y tendencias en los datos (Grainge, 2023). La tabla de frecuencias proporciona un puente entre los datos crudos y representaciones más complejas, como histogramas, polígonos de frecuencia u ojivas acumuladas.

Ejemplo: considerando el siguiente conjunto de observaciones obtenidas en un experimento: 1, 2, 2, 2, 1, 1, 5, 4, 3, 2, 2, 1, 3, 4, 5, 6, 2, 3, 4, 5, 5, 4, 3, 3, 2; se construye la siguiente tabla de frecuencias (Tabla 7):

Tabla 7. Tabla de frecuencias

Valor	Frecuencia absoluta (f)
1	3
2	6
3	5
4	4
5	4
6	1

Fuente: elaboración propia (2025).

Esta tabla permite observar, por ejemplo, que el valor 2 es el más frecuente con 6 apariciones, mientras que el valor 6 solo aparece una vez.

Tabla por intervalos de clase

Cuando los datos cuantitativos son numerosos o presentan una amplia dispersión, resulta conveniente organizarlos en intervalos de clase, especialmente en el caso de variables continuas. Este tipo de tabla permite sintetizar la información y observar la distribución general de los datos en rangos, lo que facilita el análisis comparativo y la identificación de patrones (Triola, 2020). En una tabla por intervalos de clase:

- Cada fila representa un intervalo de clase, definido por un límite inferior y uno superior.
- Se registra la frecuencia absoluta (f), es decir, el número de observaciones dentro de ese intervalo.
- La frecuencia relativa (fr) expresa la proporción de observaciones en relación con el total de datos.

- Asimismo, se suelen agregar la frecuencia acumulada (F) y la relativa acumulada (Fr), que permiten interpretar de manera progresiva la concentración de los valores.
- En algunos casos, se incorpora la marca de clase (M), que corresponde al punto medio del intervalo y es el valor representativo para cálculos posteriores, como la media o la varianza de datos agrupados.

Este tipo de tabla constituye una herramienta metodológica fundamental para el análisis de datos cuantitativos extensos, ya que transforma un conjunto complejo de observaciones en una estructura clara y organizada (Levine et al., 2020).

Ejemplo: en una encuesta sobre desempleo en el área metropolitana de la Ciudad de México, se agrupan los desempleados por rangos de edad (Tabla 8).

Tabla 8. Tabla por intervalos de clase

Rango de edad (años)	Número de desempleados
15 – 24	18
25 – 34	22
35 – 44	15
45 – 54	10
55 - 64	7

Fuente: elaboración propia (2025).

De esta forma, el grupo de 25 a 34 años concentra el mayor número de personas desempleadas, lo que puede ayudar a trazar políticas públicas para ese segmento poblacional.

Tabla de doble entrada

Las tablas de doble entrada, también denominadas tablas de contingencia permiten el análisis conjunto de dos variables categóricas o discretas, organizando los valores de una variable en filas y los de la otra en columnas. Este formato de presentación posibilita:

- La comparación simultánea de categorías.
- El análisis de la distribución conjunta de dos variables.
- La identificación de patrones, asociaciones o independencias estadísticas entre las variables estudiadas.

De forma complementaria, las tablas de contingencia pueden incluir:

- Frecuencias absolutas: número de casos observados en cada cruce de categorías.
- Frecuencias relativas: proporciones respecto al total de casos para dimensionar la importancia relativa de cada celda.
- Totales marginales: sumatorias por filas y columnas, útiles para el análisis de las distribuciones marginales.

Las tablas de doble entrada se emplean en campos como la epidemiología (por ejemplo, para estudiar la relación entre la exposición a un factor de riesgo y la presencia de enfermedad), la educación (para comparar niveles de rendimiento académico según el género) y las ciencias sociales (para explorar asociaciones entre variables sociodemográficas). Su análisis puede enriquecerse mediante la prueba de chi-cuadrado de independencia, que determina si existe una relación significativa entre las variables (Agresti, 2018).

Ejemplo: se ha registrado el número de cirugías realizadas según grupos de edad en una muestra de 100 personas (Tabla 9).

Tabla 9. Tabla de doble entrada

Grupo de edad	Cirugías 0	Cirugías 1	Cirugías 2 o más	Total
18 - 29	10	8	2	20
30 - 49	5	10	5	20
50 - 69	3	12	5	20
70 o más	2	10	8	20
Total	20	40	20	100

Fuente: elaboración propia (2025).

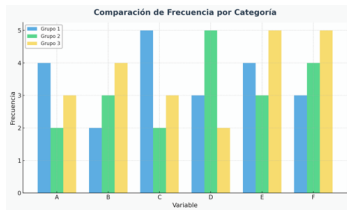
Esta tabla permite observar, por ejemplo, que en el grupo de edad de 70 o más, la mayor proporción de personas ha tenido dos o más cirugías.

2.3. Representación gráfica de datos

La representación gráfica de datos permite visualizar de manera clara, estructurada e inmediata la distribución, el comportamiento y las características de un conjunto de observaciones. Su uso favorece la interpretación de resultados estadísticos y la comunicación efectiva de la información, incluso para públicos no especializados (Montgomery y Runger, 2020). Cada tipo de gráfico se diseña para un tipo de variable (cuantitativa o cualitativa). A continuación, se describen los gráficos más

utilizados en presentaciones de datos estadísticos, junto con sus características esenciales y aplicaciones más frecuentes.

Histograma



El histograma es una representación gráfica de variables cuantitativas continuas. Su construcción se realiza en un sistema de coordenadas rectangulares en el que:

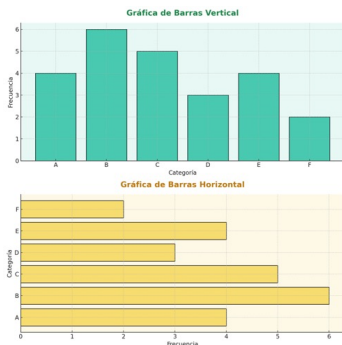
- El eje horizontal (x) representa los límites o fronteras de clase de los intervalos en los que se agrupan los datos.
- El eje vertical (y) indica la frecuencia absoluta, relativa o porcentual de cada intervalo.

Cada intervalo se representa mediante una barra cuya altura es proporcional a la frecuencia observada. Cuando todos los intervalos tienen la misma amplitud, las diferencias en altura reflejan directamente la variación en las frecuencias. En el caso de que los intervalos tengan amplitudes distintas, se recomienda utilizar la densidad de frecuencia como criterio de proporcionalidad para asegurar una interpretación adecuada (Levine et al., 2020). El histograma permite visualizar de manera inmediata aspectos clave de la distribución de los datos, tales como:

Histograma

- La concentración de observaciones en determinados rangos.
- La simetría o asimetría de la distribución.
- La existencia de sesgo positivo o negativo.
- La presencia de valores atípicos.
- La identificación de una o varias modas en la distribución.

Gráfica de barras

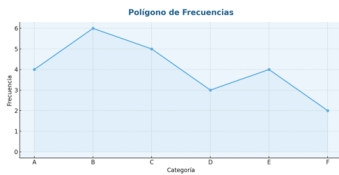


La gráfica de barras es una representación visual que muestra la distribución de variables cualitativas (nominales u ordinales) o cuantitativas discretas. Su principal utilidad consiste en facilitar la comparación de frecuencias absolutas, relativas o porcentuales entre diferentes categorías de una variable (Pagano et al., 2022).

Este recurso gráfico permite identificar de forma rápida los valores más y menos frecuentes, así como la magnitud relativa entre categorías.

A diferencia del histograma, la gráfica de barras se caracteriza por presentar barras separadas que evidencian si las categorías representadas son discretas o no continuas (Triola, 2020). Este detalle gráfico —la presencia de espacios— resalta que no existen valores intermedios entre las categorías y permite diferenciar de manera explícita las observaciones que corresponden a grupos mutuamente excluyentes (Levine et al., 2020).

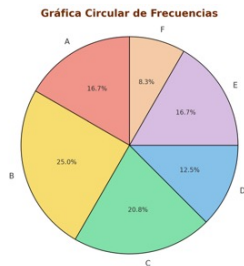
Polígono de frecuencias



El polígono de frecuencias es una representación gráfica para visualizar la distribución de datos agrupados en intervalos de clase. Constituye una alternativa al histograma y se prefiere por su capacidad para mostrar las tendencias generales de la distribución, facilitar la comparación entre diferentes series de datos y ofrecer una visualización continua de la variación de frecuencias (Pagano et al., 2022).

A diferencia del histograma, que utiliza barras para representar la frecuencia de cada intervalo, el polígono de frecuencias se construye mediante una línea poligonal que conecta una secuencia de puntos, donde cada uno representa la frecuencia correspondiente a un intervalo de clase y se ubica en el plano cartesiano mediante un par ordenado definido por la marca de clase (eje X) y la frecuencia (eje Y) (Triola, 2020). El uso de polígonos de frecuencia resulta útil en contextos comparativos, ya que permite superponer varias distribuciones en una misma gráfica, facilitando la identificación de semejanzas, diferencias y desplazamientos. Además, al proporcionar una curva continua, este recurso favorece la interpretación visual de la forma de la distribución, incluyendo aspectos como la simetría, la concentración de datos y la presencia de sesgo (Levine et al., 2020).

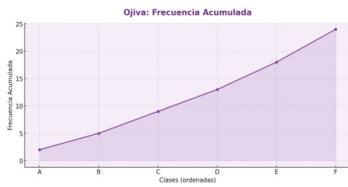
Gráfica circular o de pastel



La gráfica circular, también denominada gráfico de pastel o gráfico de sectores, es una representación visual utilizada para variables cualitativas o cuantitativas discretas, en particular cuando se busca expresar la composición porcentual de un total o la participación relativa de cada categoría dentro de un conjunto de observaciones (Peck et al., 2021).

Este tipo de gráfico se construye mediante la división de un círculo en sectores angulares, donde cada sector corresponde a una categoría de la variable y su tamaño es proporcional a la frecuencia relativa o al porcentaje que dicha categoría representa respecto al total de datos. De esta manera, la gráfica circular proporciona una visión intuitiva e inmediata de la distribución de los datos entre las distintas categorías, lo que facilita la comparación visual de proporciones en informes descriptivos, encuestas y estudios poblacionales. No obstante, su uso debe considerarse con cautela cuando se incluyen demasiadas categorías, ya que el exceso de sectores puede dificultar la interpretación visual y reducir la claridad del mensaje estadístico (Montgomery y Runger, 2020).

Ojiva



La ojiva es una representación gráfica utilizada en estadística descriptiva para mostrar la frecuencia acumulada de un conjunto de datos agrupados en intervalos de clase. Este tipo de gráfica permite observar de manera progresiva cómo se acumulan las observaciones a lo largo de la escala de la variable. Lo anterior ayuda a analizar distribuciones acumuladas e identificar medidas de posición como percentiles, cuartiles o la mediana (Triola, 2020).

Formalmente, la ojiva se construye a partir de los datos contenidos en una tabla de distribución de frecuencias acumuladas, ya sea en su forma absoluta (F) o relativa acumulada (Fr). El gráfico se desarrolla en un sistema de coordenadas cartesianas en el que:

- El eje horizontal (x) representa los límites superiores de cada intervalo de clase.
- El eje vertical (y) representa las frecuencias acumuladas, absolutas o relativas, estas últimas expresadas como fracciones decimales o porcentajes.

La conexión secuencial de los puntos genera una curva ascendente que refleja la tendencia de acumulación de los datos. Dado que su construcción es acumulativa, la curva nunca decrece. El último intervalo es el total de observaciones (si se expresa en F) o el valor $1/100\%$ (si se expresa en Fr). La ojiva permite identificar de manera inmediata la distribución acumulada de los datos y facilita la localización de medidas de posición que orientan la interpretación de fenómenos en contextos clínicos, educativos, sociales o económicos (Levine et al., 2020).

2.4. Análisis de datos

El análisis de datos permite interpretar la información contenida en una distribución de frecuencias, esté agrupada o no. El objetivo de la construcción de dichas distribuciones es facilitar la comprensión del comportamiento de los datos, mediante el examen de sus frecuencias y su representación visual, con el propósito de extraer conclusiones preliminares que orienten la toma de decisiones (Montgomery y Runger, 2020).

El procedimiento analítico suele comenzar con la observación de las frecuencias absolutas (f), es decir, la cantidad de veces que ocurre cada valor o intervalo en el conjunto de datos. A partir de ellas, se derivan otras medidas esenciales, como la relativa (fr), la acumulada (F) y la relativa acumulada (Fr), también conocida como porcentaje acumulado. Estas transformaciones permiten interpretar la distribución de los datos en términos proporcionales o acumulativos (Peck et al., 2021).

La frecuencia relativa se obtiene dividiendo la frecuencia absoluta de cada clase entre el número total de observaciones. Cuando este resultado se multiplica por 100, se obtiene una distribución porcentual, cuya suma equivale al 100% . Este

formato resulta especialmente útil para la construcción de gráficas circulares, ya que muestra de manera clara la proporción que representa cada categoría respecto al total.

Por su parte, la frecuencia acumulada indica cuántas observaciones se encuentran por debajo (o hasta) de una determinada clase. Este valor se obtiene mediante la suma progresiva de las frecuencias absolutas de cada clase con las anteriores. Si dicha suma se expresa en porcentaje, se obtiene la frecuencia relativa acumulada o porcentaje acumulado (útil para determinar medidas de posición, como percentiles, cuartiles y deciles).

Una de las representaciones más utilizadas de estas frecuencias acumuladas es la ojiva, también conocida como gráfico de porcentajes acumulados. Este tipo de gráfica muestra, por ejemplo, el porcentaje de individuos cuya estatura es inferior o superior a un valor específico. La construcción de una ojiva se realiza de manera similar a la de un polígono de frecuencias, aunque en la primera el eje horizontal representa los límites superiores de clase y el eje vertical las frecuencias acumuladas absolutas o relativas. Esta herramienta es útil para establecer rangos de referencia o evaluar la posición relativa de un individuo dentro de un grupo (Levine et al., 2020).

Las distribuciones de frecuencias pueden adoptar diversas formas según la naturaleza de los datos y el fenómeno analizado. No obstante, en la práctica estadística la mayoría de las distribuciones suelen describirse mediante algunos patrones típicos (Montgomery y Runger, 2020; Triola, 2020):

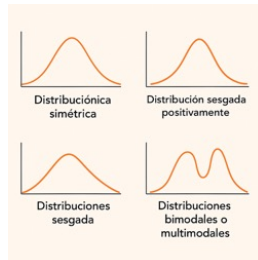
- La distribución simétrica, también conocida como distribución normal o campana de Gauss, se caracteriza por que la moda, la mediana y la media coinciden y los datos se distribuyen a ambos lados del centro. Esta distribución constituye la base de múltiples pruebas inferenciales y modelos probabilísticos.
- La distribución sesgada positivamente se presenta cuando la mayor parte de las observaciones se concentra en los valores bajos de la variable, extendiéndose una cola hacia la derecha o hacia los valores altos. Es común en fenómenos económicos o sociales, como la distribución de ingresos.
- La distribución sesgada negativamente se caracteriza por una concentración de observaciones en los valores altos, con una cola hacia la izquierda o hacia los valores bajos. Este tipo de distribución suele encontrarse, por ejemplo, en variables como la edad de jubilación o los puntajes de pruebas en las que una gran mayoría obtiene calificaciones elevadas.

- Las distribuciones bimodales o multimodales presentan dos o más picos o modas que suelen indicar la existencia de subgrupos dentro de la población analizada. Su identificación es relevante en estudios clínicos, educativos o sociales, ya que puede reflejar heterogeneidad en la muestra, diferencias por género, edad u otras variables estratificadoras.

El reconocimiento de la forma de una distribución describe un conjunto de datos y orienta la selección de técnicas estadísticas para su análisis, dado que muchos procedimientos requieren normalidad o simetría en los datos (Figura 15).

Figura 15

Patrones típicos de distribuciones de frecuencia



Fuente: elaboración propia (2025).

El análisis de datos permite transformar registros numéricos en información estructurada y significativa. Las frecuencias relativas, acumuladas y porcentuales, junto con su representación gráfica, posibilitan una descripción objetiva de la distribución, facilitan la comparación entre categorías y contribuyen a la identificación de patrones, anomalías y puntos clave dentro del conjunto de observaciones. Estas herramientas son la base de la inferencia, la predicción y la interpretación de los fenómenos estudiados (Montgomery y Runger, 2020; Triola, 2020).

Cierre de la unidad

En esta unidad se estudió la identificación de un problema, la recolección y organización de datos mediante distribuciones de frecuencias, así como su representación gráfica. Estas herramientas permiten una primera aproximación al comportamiento de los datos en función del contexto, facilitando su análisis e interpretación inicial.

Unidad

Medidas de tendencia central
y de dispersión

Presentación de la unidad

Para lograr una comprensión integral de cualquier conjunto de datos, es fundamental contar con una descripción sintética que refleje sus principales características. Esta visión general no requiere examinar cada observación individual, sino resumir el comportamiento global de la variable en estudio. Con este propósito, la estadística utiliza dos grandes tipos de indicadores: las medidas de tendencia central, que indican hacia qué valores tienden a agruparse los datos, y las medidas de dispersión que muestran el grado en que estos se separan entre sí; ambas son fundamentales para el análisis e interpretación de datos en contextos reales.

Objetivos

El estudio de esta unidad tiene como finalidad que el estudiante desarrolle las siguientes capacidades:

- Identificar problemas que involucren datos tanto desagrupados como agrupados.
- Seleccionar y obtener una muestra representativa de estos tipos de datos.
- Calcular e interpretar medidas de tendencia central y de dispersión.
- Aplicar la estadística descriptiva para emitir conclusiones basadas en datos.

Competencia específica

Al integrar los conocimientos teóricos (declarativos), los procedimientos técnicos (procedimentales), las actitudes analíticas (actitudinales) y los contextos de aplicación (contextuales), el estudiante será capaz de:

Analizar muestras extraídas de una población, explicando los fenómenos observados a través de la interpretación y representación adecuadas de las medidas de tendencia central y de dispersión.

3.1. Medidas de tendencia central

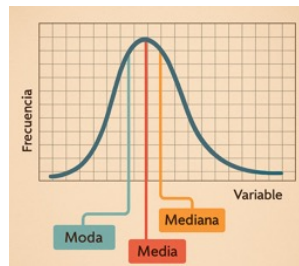
Las medidas de tendencia central son indicadores estadísticos cuyo propósito es representar, mediante un solo valor, el comportamiento general de un conjunto de datos. Su función principal es señalar un valor representativo alrededor del cual tienden a concentrarse las observaciones, ofreciendo una visión resumida del conjunto de datos sin necesidad de detallar su distribución completa (Peck et al., 2021).

Se denominan medidas de tendencia central porque, en la mayoría de las distribuciones, los valores se ubican próximos a la zona central. Estas son útiles en el análisis descriptivo porque permiten sintetizar información compleja, identificar patrones y comparar diferentes poblaciones, muestras o periodos de tiempo (Triola, 2020).

En estadística descriptiva, las tres medidas de tendencia central más utilizadas son: media aritmética, mediana y moda; cada una presenta características específicas, ventajas y limitaciones, que condicionan su aplicabilidad según la naturaleza de los datos y el objetivo del análisis (Figura 16).

Figura 16

Medidas de tendencia central



Fuente: elaboración propia (2025).

Media aritmética: también conocida como promedio, se calcula sumando todos los valores del conjunto y dividiendo el resultado entre el número total de observaciones. Es una medida muy sensible a valores extremos.

Mediana: se define como el valor central de un conjunto de datos ordenados en forma ascendente o descendente. Cuando el número de observaciones es impar, la mediana corresponde al dato que ocupa la posición media; en cambio, cuando el número de observaciones es par, se obtiene calculando el promedio de los dos valores centrales.

Moda: es el valor que ocurre con mayor frecuencia en el conjunto de datos. Puede existir más de una moda (bimodal o multimodal) o ninguna en caso de que todos los valores tengan la misma frecuencia.

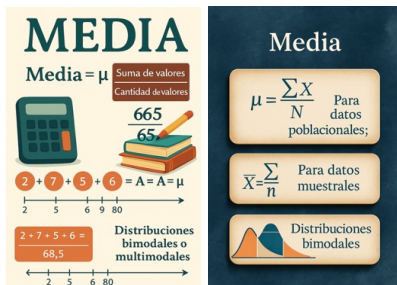
La aplicación de estas medidas depende tanto del tipo de variable como del comportamiento de los datos y del objetivo del análisis. La media se recomienda cuando los datos son cuantitativos y no presentan valores extremos; la mediana es preferible en distribuciones sesgadas o con *outliers*; mientras que la moda aporta información clave en distribuciones categóricas o cuando se desea identificar el valor más frecuente. En conjunto, estas medidas permiten describir la localización central de los datos y constituyen una base esencial para el análisis estadístico descriptivo (Montgomery y Runger, 2020; Triola, 2020; Peck et al., 2021).

3.1.1. Media aritmética

La media aritmética, media o promedio, es una de las medidas de tendencia central más utilizadas debido a su simplicidad de cálculo y su capacidad para representar el valor típico de una variable cuantitativa. Este indicador permite estimar el valor central esperado en un conjunto de datos (Shreffler y Huecker, 2023).

Se denota habitualmente como \bar{x} cuando se calcula sobre una muestra o mediante la letra griega μ cuando se refiere a una población. Su cálculo consiste en dividir la suma total de las observaciones entre el número de elementos del conjunto analizado, lo que proporciona un valor representativo de la distribución (Kocaoglu, 2024). No obstante, una de sus principales limitaciones es su sensibilidad a los valores extremos o atípicos que pueden distorsionar el resultado en distribuciones sesgadas. Por esta razón, debe utilizarse con precaución cuando se presentan valores atípicos (Kocaoglu, 2024; Figura 17).

Figura 17
Media aritmética



Fuente: elaboración propia (2025).

Ejemplo: se dispone de una muestra de cinco observaciones, cuyos valores son 10, 54, 21, 33 y 53. Para calcular la media aritmética muestral se sustituyen los datos como se indica a continuación (Figura 18):

Figura 18

Cálculo de media aritmética

$$\bar{x} = \frac{10 + 54 + 21 + 33 + 53}{5}$$

$$\bar{x} = \frac{171}{5}$$

$$\bar{x} = 34.20$$

Fuente: elaboración propia (2025).

3.1.2. Mediana

La mediana es una medida de tendencia central que representa el valor intermedio de un conjunto de datos ordenados de menor a mayor. En contraste con la media aritmética, la mediana no se ve afectada por valores extremos o atípicos, lo que la convierte en una opción preferente en distribuciones sesgadas o con datos anómalos (Khorana et al., 2023).

Este valor divide el conjunto de observaciones en dos partes iguales: el 50 % de los datos se sitúa por debajo y el otro 50 % por encima. Su cálculo sigue una lógica clara y sistemática: primero se ordenan los datos de forma ascendente y luego se determinan las observaciones (n). Si n es impar, la mediana corresponde al valor que ocupa la posición central; si es par, se calcula como el promedio de los dos valores centrales.

La relevancia de la mediana en estadística descriptiva se basa en su robustez: posee un punto de ruptura (breakdown point) del 50 %, es decir, la presencia de la mitad de los datos como valores extremos no altera su valor central. Esta propiedad la vuelve valiosa en análisis en los que los datos pueden estar contaminados o contener outliers, como ocurre en estudios clínicos o demográficos (Figura 19).

Figura 19
Mediana



Fuente: elaboración propia (2025).

3.1.3. Moda

La moda se define como el valor que aparece con mayor frecuencia en un conjunto de datos. Se trata de la categoría o dato que presenta la frecuencia absoluta más elevada, por lo que refleja el valor más común de la distribución analizada. A diferencia de la media y la mediana, la moda puede aplicarse tanto a variables cuantitativas como a cualitativas y no requiere que los datos estén ordenados para su identificación (Adeniran et al., 2024; Khorana et al., 2023). Resulta útil para determinar la categoría predominante en una distribución, dependiendo del patrón de frecuencia, una distribución puede clasificarse de varias maneras:

- Unimodal: existe una sola moda.
- Bimodal: hay dos valores que comparten la frecuencia máxima.
- Multimodal: tres o más valores presentan frecuencia máxima.
- Amodal: no hay moda definida, es decir, todos los valores tienen la misma frecuencia.

La moda representa la categoría predominante en distribuciones estadísticas y resulta valiosa en el análisis de variables categóricas, donde otras medidas de tendencia central no son aplicables. En estudios de ciencias sociales, biomédicas y de salud pública, la moda se utiliza para describir patrones de respuesta o características en poblaciones (Daniel y Cross, 2018; Pagano et al., 2022; Figura 20).

Figura 20*Moda*

Fuente: elaboración propia (2025).

3.2. Medidas de dispersión

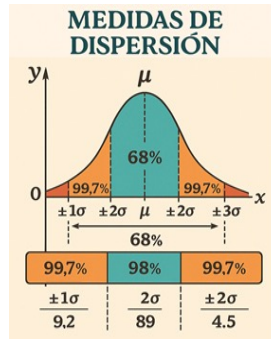
Las medidas de dispersión permiten evaluar el grado de variabilidad o alejamiento de los datos con respecto de una medida de tendencia central, ya sea la media, la mediana o la moda. A diferencia de las medidas de tendencia central, que se enfocan en identificar un valor representativo del conjunto, las medidas de dispersión aportan información acerca de la homogeneidad o heterogeneidad en el grupo de datos (Adeniran et al., 2024).

El análisis de dispersión determina cuán adecuada o representativa es la medida de posición seleccionada. En términos generales, una menor dispersión indica una mayor homogeneidad entre las observaciones, lo que refuerza la representatividad del valor central. Por el contrario, una mayor dispersión sugiere que los datos se encuentran más alejados entre sí, reduciendo la capacidad de síntesis de un único valor central (Daniel y Cross, 2018; Pagano et al., 2022).

Estas medidas permiten identificar patrones de variabilidad relevantes en investigaciones aplicadas a la biomedicina, las ciencias sociales y la salud pública, en donde el grado de dispersión puede modificar la interpretación de los resultados y la validez de las conclusiones (Peck et al., 2021; Figura 21).

Figura 21

Medidas de dispersión



Fuente: elaboración propia (2025).

3.2.1. Recorrido

El recorrido o rango (Re) constituye una de las medidas de dispersión más simples. Esta medida indica la amplitud total de los datos, es decir, la diferencia entre el valor más alto y el más bajo. Su cálculo ofrece una estimación inicial del grado de variabilidad en los datos, aunque su simplicidad conlleva limitaciones analíticas relevantes (Adeniran et al., 2024). Desde el punto de vista matemático, el recorrido se expresa (Figura 22).

Figura 22

Recorrido

$$Re = x_{\max} - x_{\min}$$

Fuente: elaboración propia (2025).

Donde x_{\max} es el valor máximo observado y x_{\min} es el mínimo observado. El recorrido no considera la distribución interna de los datos ni la frecuencia con la que se repiten los valores, por lo que resulta sensible a la presencia de valores atípicos o extremos. A pesar de esta limitación, su cálculo rápido lo convierte en una herramienta útil para obtener una primera aproximación a la dispersión en fases exploratorias del análisis de datos (Daniel y Cross, 2018; Pagano et al., 2022).

Ejemplo: considérese el siguiente conjunto de datos que representa las calificaciones obtenidas por un grupo de estudiantes: 68, 68, 52, 57, 69, 71, 78, 52, 74, 74, 69, 52, 76. Para calcular el recorrido se identifica el valor máximo $x_{\text{máx}} = 78$ y el valor mínimo $x_{\text{mín}} = 52$, aplicando la fórmula (Figura 23):

Figura 23

Cálculo de recorrido

$$Re = 78 - 52 = 26$$

Fuente: elaboración propia (2025).

Por lo tanto, el recorrido de esta distribución es 26, lo que indica una separación de 26 unidades entre el menor y el mayor valor del conjunto de datos.

3.2.2. Varianza

La varianza es una medida de dispersión que cuantifica la variabilidad de los datos respecto a la media aritmética. Se define como el promedio de los cuadrados de las desviaciones de cada valor con respecto de la media del conjunto. De esta manera, refleja el grado en que los valores se distribuyen alrededor del centro de la distribución (Motulsky, 2022). La varianza incorpora todas las observaciones del conjunto, lo que la convierte en una medida más robusta y representativa de la dispersión. Su valor indica qué tan alejados están, en conjunto, los datos respecto al promedio, siendo una herramienta para evaluar la homogeneidad o heterogeneidad de un grupo de observaciones (Daniel y Cross, 2018; Pagano et al., 2022).

El valor de esta medida siempre es mayor o igual que cero, resultado del cuadrado de las diferencias entre los valores y la media. Por lo tanto, nunca puede adoptar valores negativos, aunque presenta la limitación de que su unidad de medida corresponde al cuadrado de la unidad original de los datos, lo que puede dificultar su interpretación directa (Peck et al., 2021; Figura 24).

Figura 24
Varianza

CÁLCULO DE LA VARIANZA

$$\sigma^2 = \frac{\sum(x_i - \mu)^2}{N}$$

Para calcularla en una población.
Donde $\sum(x_i - \mu)^2$ es la suma de los cuadrados de los valores de X menos la media poblacional
N es el número total de datos.

Para calcularla en una muestra	Para calcularla en una muestra
$s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$	$s^2 = \frac{\sum(x_i - \bar{x})^2}{n-1}$
Donde $\sum(x_i - \bar{x})^2$ es la suma de los cuadrados de los valores de X menos la media poblacional	Donde $\sum(x_i - \bar{x})^2$ es la suma de los cuadrados de los valores de X menos la muestral

Fuente: elaboración propia (2025).

Ejemplo: se obtuvo una muestra aleatoria simple de 10 individuos de una población. Las edades en años registradas fueron las siguientes: 42, 28, 28, 61, 31, 23, 50, 34, 32 y 37. Calcule la varianza de las edades de los individuos estudiados; se aplica la fórmula de la media (Figura 25):

Figura 25
Cálculo de la media

$$\bar{x} = \frac{42 + 28 + 28 + 61 + 23}{23 + 50 + 34 + 32}$$

$$\bar{x} = \frac{366/10}{10}$$

$$\bar{x} = 36.6$$

Fuente: elaboración propia (2025).

Posteriormente, se sustituyen los valores en la fórmula correspondiente a la varianza (Figura 26):

Figura 26*Cálculo de la varianza*

$$s^2 = \frac{(42 - 36.6)^2 + (28 - 36.6)^2 + \dots + (37 - 36.6)^2}{9}$$

$$s^2 = \frac{1196.39}{9}$$

$$s^2 = 132.93$$

Fuente: elaboración propia (2025).

3.2.3. Desviación estándar

La desviación estándar o típica es una medida de dispersión que indica cuánto se alejan, en promedio, los datos de la media aritmética. A diferencia de la varianza, cuyos resultados se expresan en unidades cuadradas, la desviación estándar se expresa como lo hacen los datos originales (Motulsky, 2022).

Esta medida proporciona una evaluación más intuitiva de la consistencia o variabilidad interna de un conjunto de datos. En términos generales, una desviación estándar baja indica que la mayoría de los valores se concentran en torno a la media, mientras que una desviación estándar alta refleja una mayor dispersión de los datos (Daniel y Cross, 2018; Pagano et al., 2022).

Se representa mediante los símbolos:

- s , cuando se calcula a partir de una muestra.
- σ , cuando se calcula sobre una población completa.

La desviación estándar se usa en la construcción de intervalos de confianza y en pruebas de hipótesis; permite comparar la variabilidad entre diferentes conjuntos de datos de forma más accesible que la varianza. Además, es fundamental en los modelos de probabilidad y distribuciones estadísticas (Szklo y Nieto, 2023; Figura 27).

Figura 27

Desviación estándar



Fuente: elaboración propia (2025).

Ejemplo: retomando los datos anteriores, en los cuales se obtuvo una muestra aleatoria simple de 10 individuos, donde la media muestral calculada previamente fue 28 y la varianza muestral fue 14.67, se procede a calcular la desviación estándar obteniendo la raíz cuadrada de la varianza (Figura 28):

Figura 28

Cálculo de la desviación estándar

$$s = \sqrt{132.93}$$

$$s = 11.53$$

Fuente: elaboración propia (2025).

Cierre de la unidad

Esta unidad mostró la importancia de las medidas de tendencia central y de dispersión para describir el comportamiento de un conjunto de datos. Mientras las primeras indican los valores representativos, las segundas cuantifican su variabilidad. El uso conjunto de ambas facilita una interpretación más precisa y fundamentada del problema analizado.

Unidad

Principios de muestreo
y estimación

Presentación de la unidad

Esta unidad introduce los fundamentos del muestreo estadístico como herramienta para obtener información sobre una población a partir de una muestra representativa. Se abordan los conceptos básicos, los tipos de muestreo probabilístico, las técnicas de estimación en el muestreo aleatorio simple y el teorema del límite central como base para la inferencia estadística.

Objetivos

El estudio de esta unidad tiene como finalidad que el estudiante pueda:

- Identificar los conceptos fundamentales del muestreo estadístico.
- Reconocer los tipos de muestreo probabilístico y los criterios para su aplicación en distintos contextos.
- Comprender el papel principal del teorema del límite central como base de la inferencia estadística.

Competencia específica

Al integrar los conocimientos teóricos, procedimentales y contextuales desarrollados en esta unidad, el estudiante será capaz de:

- Seleccionar el tipo de muestreo más adecuado y aplicar correctamente técnicas de estimación para el análisis e interpretación de datos poblacionales.

4.1 Conceptos básicos de muestreo

En la práctica estadística, los estudios rara vez se aplican a la totalidad de la población de interés, ya que este procedimiento puede resultar costoso, lento o logísticamente inviable. Como alternativa, se recurre al análisis de una muestra, entendida como un subconjunto representativo de la población, que posibilita inferir propiedades generales sin necesidad de observar a todos los elementos (Daniel y Cross, 2018; Lohr, 2022).

El muestreo constituye, por tanto, una estrategia metodológica esencial para la investigación científica, pues permite ahorrar recursos, acelerar la obtención de resultados y mantener la precisión estadística. No obstante, la validez de las inferencias obtenidas dependerá directamente de la calidad del diseño muestral, de la técnica de selección y del tamaño de la muestra (Lohr, 2022; Figura 29).

Figura 29

Población y muestra



Fuente: elaboración propia (2025).

Para que los resultados obtenidos a partir de una muestra sean válidos y confiables, es indispensable que esta cumpla con criterios relacionados con la determinación del tamaño de la muestra, la representatividad de las características de la población y la técnica de selección de los elementos muestrales (Lohr, 2022). Una muestra correctamente diseñada y seleccionada permite minimizar el sesgo y estimar parámetros poblacionales con un nivel de precisión aceptable (Daniel y Cross, 2018; Levy y Lemeshow, 2020). Solo mediante una muestra correctamente diseñada es posible garantizar la validez estadística de las inferencias realizadas sobre una población (Figura 30).

Figura 30

Análisis muestral



Fuente: elaboración propia (2025).

Los métodos de selección de una muestra se dividen en dos grandes categorías: muestreo probabilístico y muestreo no probabilístico. La elección entre ambos depende de los objetivos del estudio, el diseño metodológico y las condiciones prácticas de acceso a la población (Daniel y Cross, 2018; Lohr, 2022).

La adecuada aplicación del muestreo contribuye a optimizar los recursos disponibles y a mejorar la validez interna y externa de los estudios, aspectos fundamentales en investigaciones aplicadas en los campos de la salud, las ciencias sociales y la economía (Levy y Lemeshow, 2020; Daniel y Cross, 2018; Lohr, 2022).

4.2 Muestreo probabilístico

En el muestreo probabilístico, todos los elementos de una población poseen una probabilidad conocida y distinta de cero de ser seleccionados para integrar la muestra. Esta característica lo diferencia de los métodos no probabilísticos y constituye su principal fortaleza, ya que permite garantizar la representatividad de la muestra respecto a la población de origen, reducir el sesgo de selección y realizar inferencias estadísticas válidas a partir de los resultados (Lohr, 2022).

Debido a su rigor metodológico, el muestreo probabilístico es utilizado en investigaciones científicas, estudios epidemiológicos, encuestas poblacionales y análisis sociales. Este enfoque asegura que las estimaciones de parámetros poblacionales se realicen con un nivel cuantificable de error muestral, lo que otorga solidez y confiabilidad a los hallazgos (Daniel y Cross, 2018; Levy y Lemeshow, 2020). Entre sus modalidades más comunes se incluyen el muestreo aleatorio simple, el sistemático, el estratificado y por conglomerados, cada uno con aplicaciones específicas que dependen de la población, el objetivo del estudio y los recursos disponibles.

4.2.1. Muestreo aleatorio simple

En el muestreo aleatorio simple, cada unidad de la población tiene la misma probabilidad de ser seleccionada de manera aleatoria. Para su aplicación es indispensable contar con un marco muestral completo, es decir, una lista actualizada y exhaustiva de todos los elementos que integran la población de interés. Este requisito asegura que cada unidad pueda ser identificada y seleccionada con igual oportunidad, lo cual minimiza el sesgo (Lohr, 2022).

El muestreo aleatorio simple (Tabla 10) se caracteriza por su sencillez, transparencia y robustez metodológica, siendo empleado en encuestas poblacionales, estudios epidemiológicos y ensayos clínicos. No obstante, presenta limitaciones prácticas cuando la población es muy grande o dispersa geográficamente, debido a la dificultad de construir y mantener un marco muestral confiable (Daniel y Cross, 2018).

Tabla 10. Muestreo aleatorio simple

Muestreo aleatorio simple	
Características	<ul style="list-style-type: none"> • Cada individuo de la población tiene una probabilidad igual e independiente de ser incluido en la muestra • La selección se realiza sin sesgos ni juicios subjetivos • Es posible efectuar la selección con reemplazo (el mismo individuo puede ser seleccionado más de una vez) o sin reemplazo (una vez seleccionado, no puede ser elegido nuevamente) • Requiere la identificación clara del universo de estudio y el acceso a un listado completo de sus miembros (marco muestral)
Procedimiento	<ul style="list-style-type: none"> • Definir la población objetivo y establecer el marco muestral • Determinar el tamaño de la muestra con base en los objetivos del estudio, el nivel de confianza deseado y el error muestral tolerado • Seleccionar los elementos de la muestra de forma aleatoria, utilizando herramientas como: generadores de números aleatorios, tablas de números aleatorios, sorteos manuales (por ejemplo, extracción de papeles o bolas)
Ventajas	<ul style="list-style-type: none"> • Es un método fácil de comprender y aplicar; especialmente en poblaciones de tamaño moderado • Si el tamaño de la muestra es suficiente, garantiza una representación adecuada de la población • Permite aplicar fórmulas estadísticas estándar para estimar parámetros y calcular errores

Muestreo aleatorio simple

Limitaciones	<ul style="list-style-type: none"> • Puede resultar difícil de implementar cuando la población es muy extensa o se encuentra geográficamente dispersa • Requiere disponer de un marco muestral completo y actualizado, lo cual no siempre es posible en la práctica
--------------	---

Fuente: elaboración propia (2025).

Ejemplo

Supóngase que un equipo de investigación desea estimar la prevalencia de hipertensión arterial en la población adulta de una ciudad; se cuenta con un registro actualizado de todos los residentes mayores de 18 años y, a partir de este listado, se seleccionan aleatoriamente 500 individuos para formar parte del estudio. En este procedimiento, cada persona tiene la misma probabilidad de ser incluida, lo que reduce el riesgo de sesgo y permite realizar inferencias válidas sobre la proporción de personas con hipertensión en la población total.

4.2.2 Muestreo sistemático

El muestreo sistemático es una técnica de selección muestral en la cual los elementos de la población se eligen siguiendo un patrón regular, mediante la aplicación de un intervalo de selección constante. Este método constituye una alternativa eficiente al muestreo aleatorio simple, particularmente cuando se dispone de una lista ordenada de todos los elementos que integran la población (Lohr, 2022).

La lógica del procedimiento consiste en seleccionar un primer elemento de manera aleatoria dentro de los primeros casos y, posteriormente, continuar la selección con intervalos fijos determinados a partir de la razón entre el tamaño de la población y el tamaño de la muestra requerida. Este mecanismo asegura que la muestra se distribuya uniformemente a lo largo de toda la población, lo cual favorece la representatividad y reduce posibles sesgos de selección (Daniel y Cross, 2018).

El muestreo sistemático (Tabla 11) se implementa con frecuencia en estudios de gran escala, encuestas institucionales, investigaciones sociales y auditorías, debido a su simplicidad operativa, rapidez y menor costo en comparación con otros métodos probabilísticos. No obstante, puede presentar limitaciones si la población presenta periodicidad en el ordenamiento de los elementos, lo que generaría sesgos sistemáticos en la muestra seleccionada (Levy y Lemeshow, 2020).

Tabla 11. Muestreo sistemático

Muestreo sistemático	
Características	<ul style="list-style-type: none"> • Requiere una lista ordenada de los elementos de la población (marco muestral) • A partir de ese punto, se selecciona cada k-ésimo elemento, donde k representa el intervalo de selección constante • La selección comienza con un punto inicial aleatorio dentro del primer intervalo
Procedimiento	<ul style="list-style-type: none"> • Organizar la población en una lista ordenada y numerada • Calcular el intervalo de selección k mediante la fórmula: $k = N / n$, donde N es el tamaño total de la población y n es el tamaño deseado de la muestra • Elegir un número aleatorio entre 1 y k como punto de inicio • Seleccionar sucesivamente cada k-ésimo elemento a partir del número inicial hasta completar la muestra
Ventajas	<ul style="list-style-type: none"> • Es más rápido, sencillo y práctico que el muestreo aleatorio simple, especialmente en poblaciones grandes • Facilita la implementación cuando se cuenta con una lista ordenada de elementos • Asegura una distribución regular de la muestra a lo largo de la población
Limitaciones	<ul style="list-style-type: none"> • Puede introducir sesgo sistemático si existe algún patrón cíclico en la población que coincida con el intervalo de selección k • Exige contar con un marco muestral ordenado y accesible, lo cual puede no ser viable en algunos contextos

Fuente: elaboración propia (2025).

Ejemplo

Un hospital desea evaluar el nivel de satisfacción de los pacientes atendidos en su servicio de urgencias. Para ello, se dispone de una lista cronológica de los 1 000 pacientes registrados durante un mes. Se pretende seleccionar una muestra de 100 pacientes.

1. Se calcula el intervalo de selección: $k = 1\,000 / 100 = 10$.
2. Se elige un número aleatorio entre 1 y 10, por ejemplo, el número 5.
3. A partir de ahí, se seleccionan los pacientes en las posiciones: 5, 15, 25, 35, ..., 995.

Este procedimiento garantiza una selección rápida, ordenada y uniforme de la muestra, siempre que no exista un patrón periódico en la lista que interfiera con el intervalo.

4.2.3 Muestreo estratificado

El muestreo estratificado es un procedimiento de muestreo probabilístico que consiste en dividir la población en subgrupos homogéneos, denominados estratos, antes de la selección aleatoria de los elementos que integrarán la muestra. Este método resulta especialmente útil cuando se conoce la heterogeneidad estructural de la población y se busca garantizar la representación proporcional o equitativa de todos los subgrupos relevantes para el estudio (Lohr, 2022).

La lógica de esta técnica radica en que, al segmentar la población en estratos internamente homogéneos pero heterogéneos entre sí, se mejora la precisión de las estimaciones y se reduce el error muestral, en comparación con los métodos que no consideran la estratificación (Daniel y Cross, 2018). Existen diferentes modalidades de asignación muestral dentro del muestreo estratificado, entre ellas:

- **Asignación proporcional:** la muestra se distribuye en cada estrato de acuerdo con el tamaño relativo de dicho estrato en la población.
- **Asignación equitativa:** cada estrato aporta el mismo número de elementos, independientemente de su tamaño.
- **Asignación óptima:** la distribución se realiza considerando tanto el tamaño del estrato como la variabilidad interna de sus datos, maximizando la eficiencia estadística (Levy y Lemeshow, 2020).

El muestreo estratificado (Tabla 12) se aplica con frecuencia en estudios poblacionales, investigaciones epidemiológicas y encuestas nacionales, donde resulta indispensable garantizar la inclusión de subgrupos como sexo, edad, nivel socioeconómico o regiones geográficas, con el fin de obtener resultados generalizables y válidos.

Tabla 12. Muestreo estratificado

Muestreo estratificado	
Características	<ul style="list-style-type: none"> • Los estratos deben ser mutuamente excluyentes, es decir, un individuo solo puede pertenecer a un único estrato. • A su vez, los estratos deben ser colectivamente exhaustivos, de manera que todos los individuos de la población estén contemplados en al menos un estrato. • La selección de elementos dentro de cada estrato se realiza de forma aleatoria e independiente.
Procedimiento	<ul style="list-style-type: none"> • División de la población en estratos con base en una característica significativa para el estudio (por ejemplo, edad, sexo, nivel educativo, ingreso económico, etcétera). • Determinación del tamaño de muestra por estrato, a través de dos enfoques principales:

Muestreo estratificado	
Procedimiento	<ul style="list-style-type: none"> - Muestreo proporcional: se asigna un tamaño muestral a cada estrato proporcional a su tamaño respecto al total poblacional. - Muestreo no proporcional: se asignan tamaños muestrales que no dependen del tamaño del estrato, sino de otros criterios como importancia analítica, variabilidad interna o costo. • Selección aleatoria de los individuos dentro de cada estrato.
Ventajas	<ul style="list-style-type: none"> • Aumenta la representatividad de la muestra respecto a subgrupos relevantes de la población. • Disminuye la variabilidad interna de los estratos, lo cual reduce el error de estimación. • Facilita el análisis comparativo entre distintos subgrupos poblacionales.
Limitaciones	<ul style="list-style-type: none"> • Requiere información previa detallada sobre la población para identificar y segmentar adecuadamente los estratos. • Puede resultar más complejo, demandante y costoso que otras técnicas de muestreo simples.

Fuente: elaboración propia (2025).

Ejemplo

Se desea realizar un estudio nacional para determinar la prevalencia de diabetes mellitus tipo 2 por grupo etario, la población se divide en tres estratos:

- Estrato 1: personas menores de 40 años
- Estrato 2: personas entre 40 y 60 años
- Estrato 3: personas mayores de 60 años

Se requiere una muestra de 300 personas. Si se desea aplicar un muestreo estratificado proporcional y se conoce que el 40 % de la población pertenece al estrato 1, el 35 % al estrato 2 y el 25 % al estrato 3, entonces la distribución muestral será la siguiente:

- Estrato 1: 40 % de 300 = 120 personas
- Estrato 2: 35 % de 300 = 105 personas
- Estrato 3: 25 % de 300 = 75 personas

Una vez definidos los tamaños muestrales, se seleccionan aleatoriamente los individuos de cada estrato. Este enfoque asegura que todos los grupos de edad estén representados de forma proporcional para obtener conclusiones comparables entre subgrupos etarios.

4.2.4. Muestreo por conglomerados

El muestreo por conglomerados es una técnica de muestreo probabilístico que consiste en dividir la población en grupos naturales o preexistentes, denominados conglomerados, y seleccionar aleatoriamente algunos de estos grupos. A diferencia del muestreo estratificado, en el cual los estratos son homogéneos en su interior y heterogéneos entre sí, los conglomerados suelen ser internamente heterogéneos, pero se asume que reflejan en conjunto la diversidad de la población (Lohr, 2022). Una vez seleccionados los conglomerados, se puede optar por dos procedimientos:

- Muestreo de una etapa: se estudia la totalidad de los elementos dentro de cada conglomerado elegido.
- Muestreo de dos etapas: tras seleccionar los conglomerados, se aplica un segundo nivel de muestreo en su interior, eligiendo aleatoriamente a una parte de los individuos que los componen.

Este método resulta útil cuando la población es muy amplia o geográficamente dispersa, ya que permite reducir costos y facilitar el acceso a las unidades muestrales.

No obstante, presenta como limitación una menor precisión en comparación con el muestreo estratificado, debido a la posible homogeneidad parcial dentro de los conglomerados seleccionados (Daniel y Cross, 2018; Levy y Lemeshow, 2020).

El muestreo por conglomerados (Tabla 13) se aplica en encuestas nacionales de salud, estudios epidemiológicos comunitarios y evaluaciones educativas, en los que resulta más factible seleccionar grupos como escuelas, hospitales o comunidades en lugar de individuos de forma aislada.

Tabla 13. Muestreo por conglomerados

Muestreo por conglomerados	
Características	<ul style="list-style-type: none"> • Los conglomerados pueden estar definidos por ubicación geográfica (ciudades, colonias, escuelas, hospitales, comunidades rurales, etcétera) o por agrupaciones institucionales o sociales. • Se presume que cada conglomerado es representativo de la población general, lo cual permite realizar inferencias válidas si el diseño está correctamente estructurado. • Es particularmente útil en estudios de gran escala donde no es factible enumerar a todos los individuos de la población.
Procedimiento	<ul style="list-style-type: none"> • División de la población en conglomerados, que constituyen unidades primarias de muestreo. • Selección aleatoria de uno o varios conglomerados del total disponible.

Muestreo por conglomerados	
Procedimiento	<ul style="list-style-type: none"> • Aplicación de una de las dos variantes: <ul style="list-style-type: none"> - Muestreo de una etapa: se incluyen todos los elementos del conglomerado seleccionado. - Muestreo de dos etapas: se selecciona una submuestra aleatoria de individuos dentro de cada conglomerado elegido.
Ventajas	<ul style="list-style-type: none"> • Disminuye los costos logísticos y operativos, especialmente cuando la población se encuentra geográficamente dispersa. • Simplifica la planificación y ejecución del trabajo de campo, ya que no requiere contar con un marco muestral individual, sino solo de conglomerados. • Es eficiente cuando los conglomerados están bien definidos y son fácilmente accesibles.
Limitaciones	<ul style="list-style-type: none"> • Puede generar un mayor error muestral si los conglomerados seleccionados no son suficientemente representativos de la población total. • La heterogeneidad interna de los conglomerados puede afectar la precisión estadística, en especial cuando existe variabilidad significativa entre grupos. • Requiere especial cuidado en el análisis, utilizando técnicas adecuadas que consideren la estructura del muestreo.

Fuente: elaboración propia (2025).

Ejemplo

Supóngase que un programa nacional de salud pública desea evaluar la cobertura de vacunación contra el sarampión en zonas rurales de una región extensa, dado que no es factible realizar un muestreo aleatorio individual en toda la población debido a limitaciones geográficas y de recursos, se decide implementar un muestreo por conglomerados. La región se divide en conglomerados naturales, correspondientes a cada comunidad rural. De una lista de 50 comunidades, se seleccionan aleatoriamente cinco conglomerados. Posteriormente, se decide aplicar un muestreo de una etapa, evaluando a todos los niños menores de 5 años en esas cinco comunidades. Esta estrategia optimiza los recursos y facilita la logística, sin comprometer la representatividad si los conglomerados son seleccionados.

4.3 Estimación en el muestreo aleatorio simple

El muestreo aleatorio simple (MAS) sirve para realizar inferencias sobre una población a partir de una muestra. La estimación se entiende como el conjunto de técnicas empleadas para inferir parámetros poblacionales tales como la media, la proporción o la varianza a partir de los estadísticos de la muestra (Daniel y Cross, 2018).

El propósito esencial de la estimación en el MAS radica en que los resultados derivados del análisis muestral sean representativos de la población de origen. Para ello, el procedimiento de selección de la muestra debe ser aleatorio para que cada elemento de la población tenga la misma probabilidad de ser elegido. Esta característica disminuye la posibilidad de sesgo sistemático y fortalece la validez estadística de las inferencias (Lohr, 2022).

El fundamento teórico que respalda este procedimiento es el Teorema del Límite Central (TLC) en el que, a medida que el tamaño de la muestra aumenta, la distribución muestral de la media u otros estadísticos tiende a aproximarse a una distribución normal, independientemente de la forma original de la distribución poblacional (Wasserman, 2021). Esta propiedad constituye la base para la aplicación de herramientas inferenciales sólidas, tales como la construcción de intervalos de confianza y la realización de pruebas de hipótesis mediante métodos paramétricos sustentados en el supuesto de normalidad.

La estimación puede clasificarse en dos categorías principales. La primera corresponde a la estimación puntual, que consiste en proporcionar un único valor como aproximación del parámetro poblacional. Un ejemplo de este tipo es el uso de la media muestral como estimador puntual de la media poblacional μ , lo cual permite obtener una medida representativa directa, aunque sin expresar el grado de incertidumbre asociado. La segunda categoría es la estimación por intervalo, cuyo propósito es determinar un rango de valores dentro del cual es probable que se encuentre el parámetro poblacional con un nivel de confianza previamente establecido, generalmente del 95 %. Este tipo de estimación incorpora explícitamente la variabilidad muestral y otorga una medida de precisión más robusta que la estimación puntual, ya que delimita el grado de certeza con el cual se puede afirmar que el parámetro se encuentra dentro de dicho rango (Wasserman, 2021; Lohr, 2022).

En el MAS, la precisión de la estimación está determinada por diversos factores, entre los que destacan el tamaño de la muestra, la variabilidad de los datos y el nivel de confianza seleccionado. A mayor tamaño muestral y menor variabilidad en los datos, la estimación obtenida será más precisa. El nivel de confianza, por su parte, define la probabilidad de que el parámetro poblacional se encuentre dentro del intervalo estimado (Lohr, 2022).

La estimación derivada del MAS es clave para la generalización de resultados. Su aplicación permite transformar información muestral en conocimiento inferencial con un grado controlado de incertidumbre, garantizando así la validez metodológica y la utilidad práctica de los hallazgos (Wasserman, 2021; Daniel y Cross, 2018).

4.3.1. Metodología del muestreo aleatorio simple

El MAS constituye un procedimiento fundamental en la estadística inferencial, mediante el cual se obtiene una muestra que representa de forma equitativa a toda una población. Cada elemento de la población posee la misma probabilidad conocida y distinta de cero de ser seleccionado, lo que asegura la aleatoriedad del proceso y permite realizar inferencias válidas sobre los parámetros poblacionales (Lohr, 2022). La metodología para llevar a cabo un muestreo aleatorio simple se organiza en varias etapas sistemáticas:

1. *Definición de la población y del parámetro de interés:* el primer paso consiste en delimitar con precisión el universo de estudio, entendido como la totalidad de elementos (personas, objetos, instituciones o eventos) sobre los cuales se desea obtener información. Asimismo, debe identificarse el parámetro poblacional que se pretende estimar, como la media, la proporción o la varianza, entre otros (Lohr, 2022).

2. *Enumeración y codificación de la población:* una vez delimitada la población, resulta indispensable enumerar todas las unidades de análisis que la integran, asignando un número de identificación único. Esta lista exhaustiva constituye el marco muestral para garantizar la aleatoriedad en la selección de la muestra y evitar sesgos de cobertura (Daniel y Cross, 2018).

3. *Determinación preliminar del tamaño de muestra:* en esta etapa se establecen los criterios estadísticos fundamentales para la estimación, entre los que destacan:

- Nivel de confianza deseado (comúnmente 95 % o 99 %).
- Margen de error o error muestral tolerable (por ejemplo, ± 5 %).
- Variabilidad esperada de la característica de interés (expresada en desviación estándar o proporción).

A partir de estos elementos es posible calcular un tamaño muestral preliminar, empleando fórmulas específicas para el MAS (Wasserman, 2021).

4. *Cálculo del tamaño muestral óptimo:* Una vez definidos los elementos fundamentales del diseño muestral, que incluyen la población de estudio, el nivel de confianza, el margen de error aceptable y la estimación de la variabilidad esperada, resulta indispensable proceder al cálculo del tamaño muestral óptimo (Berner y Amrhein, 2022; Lohr, 2022; Figura 31). El cálculo del tamaño de muestra puede abordarse desde dos escenarios principales:

- Cuando se conoce el tamaño de la población, en estudios donde el tamaño poblacional total (N) es conocido y finito, el cálculo del tamaño de la muestra (n) debe considerar la corrección para poblaciones finitas.

- Cuando se desconoce el tamaño de la población, en situaciones donde el tamaño total de la población no es conocido, o se considera muy grande (por ejemplo, en estudios con poblaciones infinitas o indeterminadas), se utiliza una fórmula simplificada que no aplica la corrección para población finita.

Figura 31

Fórmulas para cálculo de la muestra

Desconocimiento del tamaño de la población	Conocimiento del tamaño de la población
$n = \frac{Z^2 pq}{E^2}$	$n = \frac{Z^2 pqN}{NE^2 + Z^2 pq}$

Fuente: elaboración propia (2025).

Donde:

n = tamaño de la muestra

N = tamaño de la población

Z = valor crítico de la distribución normal estándar correspondiente al nivel de confianza deseado (por ejemplo, 1.96 para un 95 % de confianza)

p = proporción esperada de éxito (en ausencia de datos previos se utiliza 0.5)

$q = 1 - p$

e = margen de error permitido

5. *Selección aleatoria de los sujetos*: el último paso corresponde a la selección aleatoria de los sujetos que conformarán la muestra definitiva, mediante tablas de números aleatorios, la aplicación de generadores computarizados de números pseudoaleatorios o sorteos manuales. Resulta indispensable que la selección se ejecute con rigor metodológico, de modo que se asegure la ausencia de sesgo en la elección de los participantes. Asimismo, debe garantizarse que todas las unidades de la población conserven la misma probabilidad de ser incluidas en la muestra (Lohr, 2022; Daniel y Cross, 2018).

Ejemplo

Considérese una investigación cuyo propósito es analizar las relaciones de género en el noviazgo, enfocándose en la presencia de manifestaciones de violencia física y psicológica.

Paso 1: Definir la población de estudio y las variables a estudiar

En este caso, el objeto de estudio está constituido por las conductas violentas que se producen en relaciones afectivas entre jóvenes.

La población objetivo está conformada por los estudiantes que cursan el último año de la carrera de medicina, específicamente aquellos que mantienen una relación de noviazgo al momento del estudio. A partir del censo académico y los criterios de inclusión, se identificó que el número total de individuos que cumplen con estas condiciones asciende a 386 personas. Así, se establece que:

- Población: los 386 estudiantes del último año de medicina que tienen pareja.
- Variables de estudio: violencia física y violencia psicológica en el contexto del noviazgo.

Retomando el ejemplo del estudio sobre relaciones de género en el noviazgo entre estudiantes de último año de la carrera de medicina, previamente se identificó una población de 386 estudiantes que cumplían con el criterio de mantener una relación afectiva al momento del levantamiento de datos. Para proceder con el muestreo, es indispensable asignar un número a cada estudiante, numerándolos del 1 al 386.

Paso 2: Numerar todas las unidades de análisis que integran la población

Esta numeración no implica ningún tipo de jerarquización o clasificación, sino que cumple una función organizativa y técnica; la lista numerada servirá como marco muestral a partir del cual se seleccionarán los sujetos que integrarán la muestra final, utilizando procedimientos de aleatorización que garanticen igualdad de oportunidades para todos los integrantes de la población.

Paso 3: Definir la población, determinar el porcentaje de error y porcentaje de confianza

Una vez definida la población y enumerados sus integrantes, es necesario calcular el tamaño de la muestra. Para ello, deben considerarse tres factores estadísticos esenciales:

Nivel de confianza deseado: representa la probabilidad con la que se espera que los resultados muestrales incluyan el verdadero valor poblacional. En estudios sociales y de salud, se emplea comúnmente un nivel de confianza del 95 %, correspondiente a un valor crítico de $Z = 1.96$ en la distribución normal estándar.

Margen de error aceptable: se refiere a la diferencia máxima tolerada entre el parámetro poblacional real y su estimación muestral. En este caso, se establece un margen de error del 5 % (0.05) aceptable para estudios exploratorios. La suma de la variabilidad positiva (p) y la negativa (q) siempre debe ser igual a 1, es decir, $p + q = 1$. Cuando no se tienen antecedentes sobre la investigación, los valores de variabilidad son $p = q = 0.5$.

Variabilidad esperada en la población (p): cuando no se cuenta con estudios previos que proporcionen una estimación aproximada de la proporción esperada, se recomienda usar $p = 0.5$, ya que maximiza la variabilidad y garantiza una muestra suficientemente grande.

Paso 4: Determinar el tamaño de la muestra óptimo para el estudio

Una vez definidos la población, el porcentaje de confianza, el porcentaje de error y el nivel de variabilidad esperada, se procede al cálculo del tamaño de la muestra utilizando una fórmula adecuada. En el ejemplo presentado, la población está compuesta por 386 estudiantes de último año de medicina que mantienen una relación de noviazgo y se desea analizar las manifestaciones de violencia física y psicológica en estas relaciones.

Para este caso, al tratarse de una población finita y conocida, se aplica la fórmula de MAS con corrección para población finita (Figura 32):

Figura 32

Cálculo del tamaño de muestra óptimo

$$n = \frac{(1.96)^2 \cdot 0.5 \cdot 0.5 \cdot 386}{(0.05)^2 \cdot (386) + (1.96)^2 \cdot 0.5 \cdot 0.5}$$

$$n = \frac{3.8416 \cdot 0.25 \cdot 386}{0.0025 \cdot 386 + 3.8416 \cdot 0.25}$$

$$n = \frac{370.024}{0.9625 + 0.9604} = \frac{370.024}{1.9229} \approx 192.4$$

Fuente: elaboración propia (2025).

Por lo tanto, el tamaño óptimo de la muestra es de 193 estudiantes (redondeando al entero superior), quienes deberán ser seleccionados aleatoriamente del total de la población, este procedimiento garantiza que la muestra sea representativa del total de la población estudiantil con las características establecidas, permitiendo realizar inferencias válidas con un 95 % de confianza y un margen de error del 5 %.

Paso 5: Seleccionar los sujetos de la muestra usando números aleatorios

Con base en el ejemplo aplicado sobre las relaciones de género en el noviazgo entre estudiantes del último año de la carrera de medicina, donde se identificó una población total de 386 individuos con pareja y se determinó que el tamaño de la muestra debía ser de 193 estudiantes, el siguiente paso consiste en seleccionar aleatoriamente a los participantes que formarán parte de dicha muestra. El procedimiento por seguir es el siguiente:

- Numerar a los estudiantes: se asigna un número consecutivo a cada uno de los 386 estudiantes identificados como población de estudio, comenzando con el número 1 y finalizando con el 386. Esta numeración permite identificar de forma única a cada estudiante.
- Generar números aleatorios: a continuación, se generan 193 números aleatorios, ya que ese es el tamaño de la muestra requerida. Para ello, puede utilizarse una hoja de cálculo, como Microsoft Excel o Google Sheets, empleando la función =ALEATORIO.ENTRE(1,386), la cual permitirá obtener valores dentro del rango poblacional. También es posible utilizar generadores digitales de números aleatorios o técnicas manuales, como sorteos con papeletas numeradas.
- Seleccionar los individuos correspondientes: finalmente, se revisan los números generados y se identifican los estudiantes cuyos números de identificación coincidan con ellos. Estos estudiantes formarán parte de la muestra en el estudio sobre manifestaciones de violencia física y psicológica en las relaciones de noviazgo.

Este proceso asegura que la selección sea equitativa, al otorgar a cada estudiante de la población una probabilidad igual de ser elegido.

4.3.2 Teorema del Límite Central

El Teorema del Límite Central (TLC) constituye uno de los fundamentos teóricos esenciales de la estadística inferencial. Este teorema establece que, cuando se extraen múltiples muestras aleatorias de tamaño suficientemente grande, $n \geq 30$, a partir de una población con cualquier tipo de distribución, ya sea normal o no,

la distribución de las medias muestrales tenderá a aproximarse a una distribución normal a medida que aumenta el número de muestras, siempre que la varianza de la población sea finita.

En términos formales, si una variable aleatoria X , con media μ y desviación estándar σ , no sigue inicialmente una distribución normal, las medias muestrales obtenidas a partir de muestras aleatorias de tamaño n tienden a distribuirse aproximadamente de manera normal conforme el tamaño muestral aumenta. Esta propiedad, derivada del TLC, resulta válida siempre que la varianza poblacional sea finita y constituye el fundamento que justifica la utilización de procedimientos inferenciales basados en la normalidad en contextos donde la distribución original de los datos no la presenta (Wasserman, 2021; Rice, 2023).

Importancia del Teorema del Límite Central

La importancia del TLC radica en que constituye el fundamento teórico de gran parte de la estadística inferencial, incluso cuando se trabaja con poblaciones cuya distribución no es normal. Al extraer muestras aleatorias de tamaño suficientemente grande, la distribución de las medias muestrales tiende a aproximarse a una distribución normal, independientemente de la forma original de la población.

En consecuencia, aun si la variable de interés presenta una distribución sesgada, asimétrica o desconocida, es posible realizar inferencias estadísticas válidas sobre la media poblacional mediante herramientas basadas en la distribución normal, siempre que se utilicen muestras de tamaño adecuado (Wasserman, 2021; Rice, 2023).

En la práctica, el TLC permite aplicar pruebas de hipótesis, construir intervalos de confianza y estimar parámetros poblacionales a partir de la media muestral. Su utilidad es relevante en contextos donde resulta difícil o imposible conocer con precisión la forma de la distribución poblacional. Gracias a este teorema, la estadística adquiere un carácter robusto y aplicable a diversos escenarios científicos y profesionales, lo que posibilita la toma de decisiones confiables bajo condiciones de incertidumbre. Por estas razones, el TLC se considera una de las piedras angulares de la estadística moderna (Rice, 2023).

Ejemplo

Un equipo de investigación en salud se propone estimar el nivel promedio de presión arterial sistólica (PAS) en adultos mayores de una ciudad. Se sabe que la población presenta una gran variabilidad: algunas personas muestran valores elevados de presión arterial, otras tienen valores normales y algunas registran niveles bajos.

Esta heterogeneidad en los datos da lugar a una distribución no simétrica, es decir, que no sigue una forma de campana o una distribución normal. Esta situación plantea el siguiente cuestionamiento metodológico: ¿cómo es posible estimar con precisión el valor promedio de la PAS si la distribución poblacional no es normal?

Para resolver este problema, se aplica el TCL, que infiere que la distribución de las medias muestrales tiende a normalizarse conforme incrementa el número de muestras de tamaño suficiente.

Aplicación del Teorema del Límite Central

- Selección de muestras aleatorias: el equipo médico toma múltiples muestras aleatorias de adultos mayores, cada una con un tamaño de $n = 30$ pacientes. En cada muestra, se mide la presión arterial sistólica de los pacientes y se calcula la media correspondiente.
- Ejemplo de resultados en tres muestras distintas: muestra 1 (30 pacientes), media de PAS = 130 mmHg; muestra 2 (30 pacientes), media de PAS = 125 mmHg; Muestra 3 (30 pacientes), media de PAS = 132 mmHg.
- Repetición del procedimiento: al repetir este proceso un número considerable de veces con distintas muestras, se obtiene una colección de medias muestrales. Aunque los datos originales no siguen una distribución normal, la distribución de esas medias muestrales tiende a formar una curva normal o de campana conforme a lo establecido por el TLC.
- Aplicación estadística posterior: esta propiedad permite al equipo investigador utilizar la distribución normal como base para el análisis inferencial, lo que facilita el cálculo de intervalos de confianza para estimar la media poblacional de PAS. Además, se pueden realizar pruebas de hipótesis para evaluar, por ejemplo, la efectividad de un tratamiento antihipertensivo, aun si la población total no presenta una distribución normal de los valores de presión arterial.

Cierre de la unidad

Esta unidad ha consolidado el conocimiento del muestreo y la estimación estadística en el ámbito de la salud. Se abordaron los principales métodos de muestreo probabilístico aleatorio simple, sistemático, estratificado y por conglomerados, destacando su relevancia para obtener muestras representativas que posibiliten inferencias válidas sobre una población. Asimismo, se analizó la importancia del Teorema del Límite Central, que justifica el uso de la distribución normal para el análisis inferencial aun cuando la población original no presente dicha forma. Estos conceptos son clave para sustentar la medicina basada en evidencia, al permitir una toma de decisiones clínica y en salud pública más precisa, confiable y orientada al bienestar poblacional.

Unidad

Inferencia estadística

Presentación de la unidad

La unidad introduce los fundamentos de la inferencia estadística para extraer conclusiones sobre una población a partir del análisis de una muestra. Se abordan conceptos clave como parámetros, estimadores, intervalos de confianza y pruebas de hipótesis. Además, se enfatiza la interpretación del valor p , la significancia estadística y la identificación de errores tipo I y II, elementos esenciales para una toma de decisiones informada en el análisis de datos.

Objetivos

Mediante el estudio de esta unidad se pretende que el estudiante sea capaz de:

- Distinguir entre poblacionales y estimadores muestrales.
- Aplicar intervalos de confianza como herramienta para estimar parámetros y evaluar su precisión.
- Ejecutar pruebas de hipótesis estadísticas, interpretando correctamente su significado y el valor de significancia estadística.

Competencia específica

Al integrar los aspectos declarativos, procedimentales, actitudinales y contextuales de esta unidad, el estudiante será capaz de:

- Comprender y diferenciar los conceptos de parámetros y estimadores; aplicar intervalos de confianza para evaluar la incertidumbre de las estimaciones; y llevar a cabo pruebas de hipótesis estadísticas.
- Interpretar los resultados obtenidos, valorando su significancia estadística mediante el análisis del valor p , identificando la probabilidad de cometer errores tipo I y II, y seleccionando entre enfoques paramétricos o no paramétricos según la naturaleza de los datos analizados.

5.1. Parámetros y estimadores

En el ámbito de la estadística inferencial, los conceptos de parámetro y estimador constituyen pilares para el análisis riguroso de poblaciones a partir de observaciones muestrales. La comprensión adecuada de estos términos resulta indispensable para establecer inferencias válidas y confiables, especialmente cuando la totalidad de una población es impracticable por limitaciones logísticas, económicas o éticas. Esta situación adquiere particular relevancia en las ciencias de la salud, donde la toma de decisiones basada en evidencia depende en gran medida del uso apropiado de estos conceptos (Wasserman, 2021; Lohr, 2022).

Parámetros poblacionales

Un parámetro poblacional se define como un valor numérico fijo que resume o describe alguna característica de una población completa. Dicho valor es constante pero generalmente desconocido, ya que el acceso a la totalidad de la población resulta inalcanzable en la práctica. Entre los parámetros más empleados se encuentran:

- La media poblacional (μ), que representa el valor promedio de una variable cuantitativa en toda la población.
- La proporción poblacional (p), que indica la fracción de individuos que poseen una determinada característica dentro del total poblacional.
- La desviación estándar poblacional (σ) mide la dispersión de los valores respecto a la media poblacional.

En el contexto de la salud pública, por ejemplo, un parámetro puede corresponder a la media de la presión arterial sistólica en adultos mayores de un país, o bien a la proporción de individuos con diagnóstico confirmado de diabetes *mellitus* tipo 2 en una región específica. Sin embargo, debido a la imposibilidad de acceder a todos los individuos de la población, estos valores rara vez se conocen con certeza, lo que justifica recurrir a estimadores muestrales para aproximarlos de manera válida (Daniel y Cross, 2018; Haller y Krause, 2021).

Estimadores muestrales

Un estimador se define como una función estadística calculada a partir de los datos de una muestra extraída de la población, cuyo propósito es aproximar o inferir el valor del parámetro correspondiente. A diferencia del parámetro, que es constante, aunque generalmente desconocido, el estimador es una variable aleatoria; su valor depende de la muestra particular seleccionada (Wasserman, 2021). Entre los estimadores más utilizados en la práctica se encuentran:

- La media muestral (\bar{x}), utilizada como estimador de la media poblacional (μ).
- La proporción muestral (p^\wedge), empleada para estimar la proporción poblacional (p).
- La desviación estándar muestral (s) es utilizada para aproximar la desviación estándar de la población (σ).

Un ejemplo aplicado en el campo de la salud se presenta cuando se desea estimar la media de glucosa en sangre en pacientes con diabetes de una ciudad. En este caso, se selecciona una muestra aleatoria de pacientes, se miden sus niveles de glucosa y se calcula la media muestral. Este valor constituye un estimador del parámetro poblacional correspondiente, lo que permite realizar inferencias válidas sobre la población sin necesidad de evaluar a todos los individuos (Lohr, 2022).

Propiedades deseables de los estimadores

Para que un estimador sea adecuado, debe cumplir con ciertas propiedades fundamentales que garantizan la validez y precisión de las inferencias realizadas.

- Insesgadez (o falta de sesgo): se refiere a la capacidad del estimador para producir, en promedio, un valor igual al del parámetro poblacional que se pretende estimar. Un estimador es insesgado si su valor esperado coincide con el verdadero valor del parámetro, lo que implica que no tiende sistemáticamente a sobreestimar ni a subestimar.
- Eficiencia: entendida como la comparación entre distintos estimadores insesgados de un mismo parámetro. El estimador más eficiente es aquel que presenta la menor varianza, lo que significa que sus posibles valores muestrales se encuentran menos dispersos alrededor del valor real del parámetro.
- Consistencia: alude a la capacidad del estimador para aproximarse al valor verdadero del parámetro conforme aumenta el tamaño de la muestra. En otras palabras, un estimador es consistente si, al incrementarse el número de observaciones, la probabilidad de que la estimación se acerque al valor real se incrementa progresivamente (Wasserman, 2021; Rice, 2023).

En el ámbito de la medicina basada en evidencia, estas propiedades resultan esenciales, ya que permiten inferir características de grandes poblaciones a partir de muestras representativas. Por ejemplo, no resulta factible estudiar a todos los pacientes hipertensos de un país para conocer la media de la presión arterial sistólica. En su lugar, se selecciona una muestra adecuada, se calcula la media muestral y, mediante su intervalo de confianza, se establecen inferencias que respaldan

decisiones clínicas y de salud pública, tales como la modificación de esquemas terapéuticos o la definición de prioridades de intervención (Daniel y Cross, 2018).

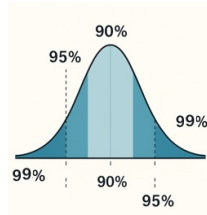
El estudio de los parámetros y estimadores no solo permite formalizar la representación de fenómenos poblacionales, sino que también proporciona las bases técnicas necesarias para realizar inferencias con validez científica. Esta labor resulta indispensable en la toma de decisiones fundamentadas en datos, especialmente en disciplinas donde el rigor metodológico es esencial para garantizar la calidad del conocimiento generado. En este sentido, el dominio de estos conceptos constituye una competencia fundamental para todo profesional que emplee la estadística como herramienta de análisis en contextos académicos, clínicos o de investigación, ya que posibilita la interpretación crítica de la información, la reducción del sesgo y la generación de conclusiones reproducibles y generalizables (Daniel y Cross, 2018; Wasserman, 2021; Rice, 2023).

5.1.1. Intervalos de confianza

El intervalo de confianza (IC) permite estimar un parámetro poblacional a partir de datos muestrales y expresar, mediante un rango, el nivel de incertidumbre asociado al proceso de estimación. Cuando se obtiene una muestra aleatoria de una población y se calcula una estadística muestral, como la media o la proporción, se obtiene un estimador puntual del parámetro correspondiente. No obstante, este valor puntual no refleja la variabilidad inherente al muestreo, por lo que resulta insuficiente para comunicar la precisión de la estimación (Navidi, 2021).

Para solventar esta limitación, se construyen intervalos de confianza que proporcionan un rango de valores plausibles dentro del cual, con una determinada probabilidad, se encuentra el verdadero valor del parámetro poblacional. El nivel de confianza —habitualmente establecido en 90 %, 95 % o 99 %— indica la proporción de veces que, al repetir el proceso de muestreo bajo las mismas condiciones, los intervalos calculados contendrían efectivamente el valor real del parámetro (Wasserman, 2021; Figura 33).

Un ejemplo ilustrativo se observa en el ámbito de la salud pública: si se calcula un intervalo de confianza del 95 % para la media de la presión arterial sistólica en adultos mayores de una comunidad, ello implica que, si se repitiera el procedimiento de muestreo múltiples veces, aproximadamente el 95 % de los intervalos construidos incluirían el valor poblacional verdadero de la media. En otras palabras, se puede afirmar con un 95 % de confianza que el intervalo obtenido a partir de la muestra contiene el parámetro poblacional real (Daniel y Cross, 2018).

Figura 33*Nivel de confianza*

Fuente: elaboración propia (2025).

Elementos que componen un intervalo de confianza

Los intervalos de confianza se construyen a partir de distintos componentes que permiten garantizar su precisión y validez en el contexto de la inferencia estadística:

- **Estimador puntual:** es el valor calculado a partir de la muestra que constituye la base del intervalo. Puede corresponder a la media muestral (\bar{x}) o a la proporción muestral (p^{\wedge}). Dicho estimador resume la información de la muestra y aproxima el parámetro poblacional de interés (Pereira et al., 2020).
- **Nivel de confianza ($1-\alpha$):** indica la probabilidad con la que se espera que el intervalo contenga el valor verdadero del parámetro poblacional. Los niveles más empleados en la práctica estadística son 90 %, 95 % y 99 %, siendo el 95 % el estándar en la mayoría de las investigaciones biomédicas y sociales (Wasserstein et al., 2019).
- **Error estándar (SE):** representa la variabilidad del estimador puntual, que depende tanto del tamaño de la muestra como de la dispersión de los datos. A menor error estándar, mayor precisión del intervalo de confianza. Su cálculo se fundamenta en principios de teoría del muestreo y distribución de probabilidades, lo que permite ajustar la estimación a la variabilidad inherente al proceso de selección muestral (Pagano et al., 2022; Motulsky, 2022).
- **Distribución estadística de referencia:** la selección de la distribución estadística de referencia constituye un aspecto esencial en la construcción de intervalos de confianza, ya que determina la forma en que se ajusta la estimación a la variabilidad muestral y al tamaño de la muestra analizada. La elección depende de la naturaleza del estimador y de las condiciones del estudio:

- * Para muestras grandes ($n \geq 30$), se emplea la distribución normal estándar (Z), en virtud del Teorema del Límite Central (TLC) —conforme aumenta el tamaño de la muestra, la distribución de los estimadores tiende a aproximarse a una distribución normal, independientemente de la forma de la población original— (Wasserman, 2021).
- * Para muestras pequeñas ($n < 30$) y cuando se desea estimar una media poblacional con varianza desconocida, se utiliza la distribución t de Student. Esta distribución presenta colas más amplias que la normal, con lo que corrige la mayor incertidumbre introducida por el tamaño reducido de la muestra y la necesidad de estimar la varianza a partir de los datos disponibles (Mishra et al., 2019; Peck et al., 2021).

Fórmulas generales para el cálculo del intervalo de confianza

Las fórmulas generales para la construcción de intervalos de confianza permiten estimar, con un grado de certeza, el rango dentro del que se encuentra un parámetro poblacional a partir de la información obtenida en una muestra. La aplicación de estas expresiones depende de la naturaleza del estimador, el tamaño de la muestra y el conocimiento de la varianza poblacional (Chow y Liu, 2021).

Para una media poblacional

El intervalo de confianza para una media poblacional constituye una de las aplicaciones más frecuentes en estadística inferencial. Su propósito es estimar, a partir de una muestra representativa, el rango de valores dentro del cual se espera que se ubique la media verdadera de la población con un nivel de confianza preestablecido. Esta metodología permite cuantificar la incertidumbre inherente al proceso de muestreo (Figura 34).

Figura 34

Intervalo de confianza para una media poblacional

$$IC = \bar{x} \pm z \cdot \frac{\sigma}{\sqrt{n}}$$

Fuente: elaboración propia (2025).

Donde:

\bar{x} : media muestral

z : valor z correspondiente al nivel de confianza (por ejemplo, 1.96 para 95 %)

σ : desviación estándar poblacional

n : tamaño de la muestra

IC: intervalo de confianza

Para una proporción poblacional

El intervalo de confianza para una proporción poblacional se usa para estimar, a partir de los datos de una muestra, el rango dentro del cual se encuentra la proporción verdadera de individuos que poseen una determinada característica en la población. Esta técnica resulta útil en áreas como la epidemiología, la investigación social y la salud pública, donde con frecuencia se requiere estimar prevalencias, tasas o proporciones asociadas a fenómenos de interés. El uso de intervalos de confianza en este contexto permite no solo expresar la estimación puntual de la proporción, sino también cuantificar la incertidumbre asociada al proceso de muestreo, garantizando mayor solidez en la interpretación de los resultados (Woodward, 2023; Figura 35).

Figura 35

Intervalo de confianza para una proporción poblacional

$$IC = \hat{p} \pm z \cdot \sqrt{\frac{z^2(1-\hat{p})}{n}}$$

Fuente: elaboración propia (2025).

Donde:

\hat{p} : proporción muestral

z : valor z para el nivel de confianza

n : tamaño de la muestra

IC: intervalo de confianza para la proporción poblacional

A continuación, se presentan dos ejemplos aplicados que ilustran el uso de los intervalos de confianza en estimaciones de la media y de la proporción poblacional, empleando información muestral representativa y los fundamentos estadísticos previamente descritos.

Ejemplo 1: estimación de la media poblacional

En un hospital, se seleccionó una muestra aleatoria de 50 pacientes para medir sus niveles de presión arterial sistólica; la media obtenida fue de 130 mmHg, con una desviación estándar muestral de 15 mmHg. Se desea calcular un intervalo de confianza del 95 % para estimar el valor medio de la presión arterial en la población. Aplicando la fórmula del intervalo de confianza para la media (Figura 36) y considerando una distribución normal (dado que $n \geq 30$):

Figura 36

Cálculo de intervalo de confianza para una media poblacional

$$IC = \bar{x} \pm z_{\alpha/2} \left(\frac{s}{\sqrt{n}} \right)$$

$$IC = 130 \pm 1.96 \left(\frac{15}{\sqrt{50}} \right) = 130 \pm 4.15$$

$$\Rightarrow IC = (125.85, 134.15)$$

Fuente: elaboración propia (2025).

Este intervalo indica que, con un 95 % de confianza, la media poblacional de presión arterial sistólica se encuentra entre 125.85 mmHg y 134.15 mmHg.

Ejemplo 2. Estimación de una proporción poblacional

En un estudio clínico, se evaluaron 100 pacientes y se identificó que 40 de ellos presentan niveles elevados de colesterol; la proporción muestral (\hat{p}) es 0.40. Se busca estimar, con un 95 % de confianza, el porcentaje verdadero de pacientes con colesterol elevado en la población total. Se emplea la fórmula del intervalo de confianza para una proporción (Figura 37):

Figura 37

Cálculo de intervalo de confianza para una proporción poblacional

$$IC = \hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

$$IC = 0.40 \pm 1.96 \cdot \sqrt{\frac{0.40(0.60)}{100}} = 0.40 \pm 0.096$$

$$\Rightarrow IC = (0.304, 0.496)$$

Fuente: elaboración propia (2025).

Este resultado sugiere que, con un 95 % de confianza, la proporción de pacientes con niveles elevados de colesterol en la población se encuentra entre el 30.4 % y el 49.6 %. Ambos ejemplos muestran cómo los intervalos de confianza permiten tomar decisiones informadas a partir de datos muestrales, proporcionando un marco cuantitativo para expresar la incertidumbre inherente a la estimación de parámetros poblacionales.

Los intervalos de confianza se emplean de manera extensiva en la investigación médica, epidemiológica, social y económica, debido a que permiten estimar parámetros poblacionales con un margen de error controlado y evaluar la precisión de las estimaciones realizadas. Por ejemplo, en la comparación de la eficacia de dos tratamientos, los intervalos de confianza aplicados a las tasas de recuperación posibilitan determinar si existe una diferencia estadísticamente significativa entre ambos grupos (Woodward, 2023).

En el contexto de la medicina basada en evidencia, los intervalos de confianza constituyen una herramienta estadística de relevancia fundamental, ya que cuantifican la incertidumbre inherente a las estimaciones derivadas de muestras. Su utilidad trasciende el ámbito meramente numérico y se proyecta directamente en la calidad de las decisiones clínicas y en la solidez metodológica de la investigación científica. La importancia de los intervalos de confianza puede analizarse en tres dimensiones principales (Woodward, 2023):

- Toma de decisiones clínicas fundamentadas: los intervalos de confianza permiten a los profesionales de la salud, epidemiólogos y responsables de políticas sanitarias evaluar con objetividad la precisión de las estimaciones obtenidas.
- Evaluación comparativa de intervenciones médicas: estos intervalos resultan esenciales para valorar la efectividad de distintos tratamientos o intervenciones sanitarias. Un ejemplo se observa al analizar la diferencia de medias en los niveles de presión arterial entre dos grupos o al comparar proporciones de respuesta a un fármaco. Los intervalos de confianza ofrecen un marco sólido para determinar si una diferencia observada es estadísticamente significativa y clínicamente relevante.
- Validez y transparencia en la investigación científica: en la comunicación de resultados de estudios clínicos, los intervalos de confianza proporcionan una representación clara y precisa de la variabilidad de los datos y de la fiabilidad de las estimaciones.

El uso de intervalos de confianza en el ámbito de la salud constituye una base cuantitativa para evaluar la fiabilidad de los resultados. Su aplicación sistemática los consolida como un recurso indispensable en la práctica clínica, la investigación biomédica y la salud pública contemporánea.

5.2. Pruebas de hipótesis

Las pruebas de hipótesis evalúan la validez de afirmaciones o suposiciones a partir del análisis de datos de muestras representativas. Su propósito principal es determinar, dentro de un marco probabilístico, si una observación empírica puede atribuirse al azar o si, por el contrario, constituye evidencia suficiente para inferir un efecto real o una diferencia estadísticamente significativa (Petrie y Sabin, 2023).

En las ciencias de la salud, el uso de pruebas de hipótesis adquiere una relevancia particular, ya que permite valorar la eficacia de tratamientos, establecer asociaciones entre factores de riesgo y enfermedades, así como validar diferencias observadas entre grupos poblacionales. Por ejemplo, estas pruebas son aplicadas en la comparación de terapias farmacológicas, en la evaluación de intervenciones quirúrgicas o en el análisis de la efectividad de programas preventivos en salud pública. El procedimiento se inicia con la formulación de una hipótesis nula (H_0), que se asume cierta de manera provisional, y de una hipótesis alternativa (H_a), que plantea un efecto o diferencia respecto a la condición establecida en H_0 . Mediante el análisis de los datos muestrales y pruebas estadísticas específicas, se estima la probabilidad de que los resultados observados se deban únicamente al azar. Así, se decide si se rechaza o se mantiene la hipótesis nula. Las pruebas de hipótesis garantizan la solidez científica de las conclusiones y fortalecen la confianza en la aplicación práctica de los resultados en contextos clínicos y epidemiológicos (Woodward, 2023).

5.2.1. Hipótesis estadísticas

Las hipótesis estadísticas son proposiciones cuantificables que pueden ser sometidas a verificación empírica mediante el análisis de datos. Constituyen el punto de partida de toda prueba estadística y permiten contrastar afirmaciones sobre parámetros poblacionales como medias, proporciones o varianzas (Hernández-Sampieri et al., 2021). Existen dos tipos de hipótesis que se formulan de manera complementaria:

Hipótesis nula (H_0): corresponde a la afirmación que se considera válida en ausencia de evidencia contraria. Representa la ausencia de efecto, diferencia o relación entre las variables en estudio. Se mantiene como válida hasta que los datos proporcionen evidencia suficiente para rechazarla.

Ejemplos: “el nuevo tratamiento no presenta diferencias con respecto al tratamiento estándar” o “no existe asociación entre el hábito tabáquico y la incidencia de cáncer pulmonar”.

Hipótesis alternativa (H_a): se plantea como la contraposición a H_0 y representa la existencia de un efecto, diferencia o asociación. Es la hipótesis que se desea comprobar. Ejemplos: “el nuevo tratamiento es más eficaz que el convencional” o “el consumo de tabaco se asocia significativamente con el riesgo de desarrollar cáncer pulmonar”.

El proceso de decisión en la prueba de hipótesis implica determinar si los resultados muestrales son lo suficientemente incompatibles con H_0 como para justificar su rechazo en favor de H_a . Las pruebas de hipótesis no demuestran la veracidad o falsedad absoluta de H_0 ; en cambio, evalúan la coherencia de los datos empíricos con dicha afirmación bajo un nivel de significancia preestablecido (Pagano et al., 2022).

5.2.2. Significancia alfa y significancia estadística (valor p)

El proceso de contrastación de hipótesis requiere establecer reglas de decisión que permitan discernir si los resultados son atribuibles al azar o si representan evidencia estadísticamente significativa. Para ello, se emplean dos herramientas fundamentales: el nivel de significancia (α) y el valor p.

Nivel de significancia (α): este parámetro se define como el umbral de tolerancia al error tipo I; es la probabilidad de rechazar la hipótesis nula cuando en realidad es verdadera. Representa el criterio de corte para determinar la significancia estadística de los resultados observados. En la práctica, el valor más utilizado es $\alpha = 0.05$, lo que implica aceptar un 5 % de probabilidad de un falso positivo. Sin embargo, la elección de α debe ajustarse al contexto de la investigación, considerando las implicaciones clínicas, éticas y económicas de los errores estadísticos (Motulsky, 2022).

Valor p (probabilidad observada): el valor p corresponde a la probabilidad de obtener, bajo el supuesto de que H_0 es verdadera, un valor del estadístico de prueba igual o más extremo que el observado en los datos muestrales. En términos prácticos, cuanto menor es el valor p, mayor es la evidencia empírica contra la hipótesis nula, lo que incrementa la plausibilidad de aceptar la hipótesis alternativa (Wasserstein et al., 2019).

Criterios de decisión:

- Si $p < \alpha$, se rechaza H_0 , lo que indica que los resultados son estadísticamente significativos y respaldan la hipótesis alternativa.
- Si $p \geq \alpha$, no se rechaza H_0 : lo que implica que no se dispone de evidencia suficiente para afirmar la existencia de un efecto o diferencia (Amrhein et al., 2019).

Ejemplo: supóngase que se desea evaluar la eficacia de un nuevo fármaco para la reducción de colesterol en comparación con un placebo. Tras realizar un estudio clínico aleatorizado, se obtiene un valor p de 0.03 y se ha establecido un nivel de significancia de $\alpha = 0.05$. Dado que $p < \alpha$, se rechaza la hipótesis nula y se concluye que existe evidencia estadísticamente significativa de que el nuevo fármaco reduce los niveles de colesterol más que el placebo. Las pruebas de hipótesis proporcionan una base formal para tomar decisiones racionales y replicables en contextos de incertidumbre, contribuyendo al rigor metodológico de la investigación biomédica.

5.2.3. Errores tipo I y II

En el análisis estadístico mediante pruebas de hipótesis, la toma de decisiones conlleva una incertidumbre inherente. Esto implica que, al rechazar o no rechazar una hipótesis nula (H_0), pueden cometerse errores. Estos se clasifican en dos categorías: tipo I y II.

Error Tipo I (α): rechazo incorrecto de la hipótesis nula

El error tipo I ocurre cuando se rechaza la hipótesis nula cuando es verdadera. En otras palabras, se concluye que existe un efecto, diferencia o relación estadísticamente significativa cuando en realidad no la hay (como cuando se considera que un tratamiento es eficaz cuando en realidad no lo es). La magnitud de este error se controla mediante el nivel de significancia (α), que representa la probabilidad máxima tolerada de incurrir en un falso positivo. Por convención, los valores más utilizados son $\alpha = 0.05$, 0.01 o 0.10, dependiendo del rigor del estudio. Por ejemplo, un valor de $\alpha = 0.05$ implica aceptar hasta un 5 % de probabilidad de rechazar H_0 de manera incorrecta (Kim, 2015; Andrade, 2019; Amrhein et al., 2019).

Ejemplo: un estudio concluye que un nuevo medicamento reduce significativamente la presión arterial ($p < 0.05$), cuando en realidad no existe tal efecto. Esta afirmación corresponde a un error tipo I.

Error tipo II (β): aceptación incorrecta de la hipótesis nula

El error tipo II ocurre cuando no se rechaza la hipótesis nula, siendo falsa; en otras palabras, se concluye que no existe efecto o diferencia significativa cuando sí la hay. En la práctica clínica, esto puede significar ignorar un tratamiento realmente eficaz. La probabilidad de incurrir en este error se denota con la letra griega β y su complemento ($1 - \beta$) representa la potencia estadística del estudio, es decir, la capacidad del análisis para detectar diferencias reales. Un diseño adecuado de estudio aspira a mantener $\beta \leq 0.20$, lo que equivale a una potencia del 80 % o superior.

Ejemplo: un ensayo clínico no encuentra diferencias significativas entre un nuevo antibiótico y el estándar ($p \geq 0.05$), cuando en realidad el nuevo antibiótico es más eficaz. Esta situación constituye un error tipo II.

Relación entre α y β : el dilema del equilibrio

Existe una relación inversa entre la probabilidad de cometer errores de tipo I y II. Disminuir el nivel de significancia (α), por ejemplo, de 0.05 a 0.01, reduce la probabilidad de incurrir en un error tipo I; sin embargo, esta acción puede incrementar la probabilidad de cometer un error tipo II (β), especialmente cuando el tamaño de la muestra es reducido o la variabilidad de los datos es elevada (Lakens, 2022; Condon et al., 2023).

Por esta razón, en el diseño de estudios resulta indispensable establecer un equilibrio adecuado entre α y β , tomando en cuenta el contexto clínico, la magnitud esperada del efecto y las implicaciones prácticas de cada tipo de error. En términos generales, reducir α ofrece protección frente a falsos positivos, pero incrementa el riesgo de falsos negativos. En cambio, aumentar el tamaño de la muestra permite disminuir simultáneamente α y β , lo cual mejora la precisión y potencia del estudio (Ioannidis, 2019; Wasserstein et al., 2019).

5.2.4. Estadística paramétrica y no paramétrica

En el análisis estadístico inferencial, la selección del método depende de las características intrínsecas de los datos y de los supuestos sobre la población de origen. En este marco, se distinguen dos enfoques principales: la estadística paramétrica y la estadística no paramétrica. Cada uno posee aplicaciones específicas, ventajas y limitaciones, por lo que su correcta utilización resulta esencial para garantizar la validez de los resultados científicos (Dehaene et al., 2021; Kallogjeri y Piccirillo, 2023).

Estadística paramétrica

La estadística paramétrica supone que los datos siguen una distribución conocida, generalmente la distribución normal (Thams et al., 2022). Este enfoque utiliza parámetros poblacionales como la media (μ) y la desviación estándar (σ), lo que permite obtener estimaciones más eficientes y con mayor potencia estadística cuando los supuestos se cumplen (Sullivan, 2018; Norman y Streiner, 2021).

Características principales:

- Requiere que los datos sean cuantitativos, continuos y con distribución aproximadamente normal.
- Supone homogeneidad de varianzas y selección aleatoria de las muestras.
- Los resultados son más precisos y potentes si los supuestos estadísticos se cumplen.

Pruebas paramétricas comunes:

- Prueba t de Student: empleada para comparar medias entre dos grupos, ya sean independientes o pareados.
- ANOVA (análisis de la varianza): utilizado para comparar medias entre tres o más grupos.
- Regresión lineal: aplicada al estudio de la relación entre una variable dependiente y una o más variables independientes.

Estadística no paramétrica

La estadística no paramétrica es un enfoque alternativo a la estadística paramétrica porque no requiere supuestos estrictos sobre la distribución de los datos. Este tipo de pruebas resulta útil cuando los datos no cumplen con el requisito de normalidad, presentan asimetrías relevantes o se encuentran en escalas ordinales, como rangos o categorías ordenadas. Aunque presentan menor potencia estadística que sus equivalentes paramétricos, su flexibilidad y menor sensibilidad a los valores atípicos las convierten en herramientas fundamentales en numerosos escenarios de investigación biomédica, social y epidemiológica (Mishra et al., 2019; Amaral y Line, 2021).

Características principales:

- Son adecuadas para datos no normales o de escala ordinal.
- No requieren homogeneidad de varianzas entre grupos.
- Poseen menor sensibilidad frente a valores extremos (outliers), lo que otorga mayor robustez en contextos donde la variabilidad es elevada.

Pruebas no paramétricas comunes:

- Prueba de Mann-Whitney U: es una alternativa a la prueba t para comparar dos grupos independientes.
- Prueba de Wilcoxon: compara dos muestras relacionadas o pareadas.
- Prueba de Kruskal-Wallis: equivalente no paramétrica del ANOVA con tres o más grupos.

La elección de pruebas no paramétricas permite mantener la validez de las conclusiones cuando los supuestos estadísticos de los métodos paramétricos no se cumplen, garantizando así la pertinencia metodológica del análisis (Conover, 1999; Dehaene et al., 2021).

Ejemplo: un equipo de investigación en el área de medicina interna está evaluando la eficacia de un nuevo tratamiento farmacológico diseñado para reducir los niveles de colesterol LDL (lipoproteína de baja densidad) en pacientes adultos con dislipidemia. El objetivo es determinar si dicho tratamiento tiene un efecto significativamente mayor que el tratamiento estándar utilizado hasta ahora.

Formulación de la pregunta de investigación: ¿el nuevo tratamiento reduce más los niveles de colesterol LDL que el tratamiento estándar actualmente prescrito?

Formulación de las hipótesis estadísticas:

- Hipótesis nula (H_0): no hay diferencia en la eficacia entre el nuevo tratamiento y el tratamiento estándar. Es decir, la media de reducción del colesterol es igual en ambos grupos ($\mu_{\text{nuevo}} = \mu_{\text{estándar}}$).
- Hipótesis alternativa (H_a): el nuevo tratamiento es más eficaz que el tratamiento estándar; por lo tanto, la media de reducción del colesterol es mayor en el grupo que recibió el nuevo fármaco ($\mu_{\text{nuevo}} > \mu_{\text{estándar}}$).

Datos recolectados

Se seleccionaron dos grupos de pacientes adultos con características clínicas similares (Tabla 14):

Tabla 14. Grupos de pacientes adultos

Grupo	Número de pacientes	Reducción promedio en colesterol LDL	Desviación estándar
Nuevo tratamiento	50	30 mg/dL	10 mg/dL
Tratamiento estándar	50	25 mg/dL	10 mg/dL

Fuente: elaboración propia (2025).

Ambos grupos fueron tratados durante el mismo periodo y bajo condiciones controladas. Los datos presentaron una distribución normal y que las varianzas eran homogéneas, por lo que se usa una prueba t para muestras independientes.

Prueba estadística aplicada

Se realizó una prueba t de Student para dos muestras independientes a fin de comparar las medias de dos grupos cuando se cumplen los supuestos de normalidad y varianza homogénea.

- Nivel de significancia adoptado (α): 0.05
- Resultado obtenido: valor $p = 0.02$

Interpretación

El valor p de 0.02 indica que existe solo un 2 % de probabilidad de que la diferencia observada entre los dos tratamientos sea producto del azar si en realidad no hubiera diferencia en la población. Como este valor p es menor al nivel de significancia establecido ($\alpha = 0.05$), se procede a rechazar la hipótesis nula.

Conclusión

Con un nivel de confianza del 95 %, se concluye que el nuevo tratamiento presenta una reducción significativamente mayor en los niveles de colesterol LDL en comparación con el tratamiento estándar. Este resultado respalda la hipótesis de que el nuevo fármaco es más efectivo, al menos en la muestra estudiada.

En términos clínicos, estos hallazgos podrían justificar futuras investigaciones con muestras más amplias y la implementación del nuevo tratamiento en guías clínicas, en tanto se corroboren la eficacia y seguridad del medicamento en estudios más extensos y con seguimiento a largo plazo.

Cierre de la unidad

Al concluir esta unidad, se habrá consolidado una comprensión integral sobre las pruebas de hipótesis, esenciales para la toma de decisiones fundamentadas en contextos clínicos y de investigación. La integración de los conceptos de hipótesis estadísticas, niveles de significancia, valor p , errores tipo I y II, así como la diferenciación entre métodos paramétricos y no paramétricos, proporciona al estudiante las herramientas analíticas necesarias para interpretar datos, evaluar intervenciones terapéuticas.

Unidad

Inferencia sobre una muestra

Presentación de la unidad

La presente unidad aborda la inferencia estadística aplicada al análisis de una sola muestra. Se estudian métodos paramétricos (pruebas z y t) y no paramétricos (Kolmogórov-Smirnov, chi-cuadrada y binomial), esenciales para interpretar resultados en variables cuantitativas y cualitativas. Estas herramientas son clave en la toma de decisiones basadas en evidencia, especialmente en áreas como la medicina y las ciencias sociales.

Objetivos

El estudio de esta unidad tiene como propósito que el estudiante adquiera el conocimiento para:

- Comprender los principios fundamentales y aplicaciones de las pruebas z y t en el análisis de una muestra con datos cuantitativos.
- Identificar los supuestos estadísticos que sustentan dichas pruebas y evalúe su cumplimiento en datos reales.
- Analizar datos cualitativos mediante pruebas no paramétricas como Kolmogórov-Smirnov, chi-cuadrada y binomial, reconociendo sus fundamentos y aplicaciones.
- Aplicar de manera crítica e informada los procedimientos estadísticos aprendidos para interpretar adecuadamente los resultados de investigaciones empíricas.
- Desarrollar habilidades para comunicar hallazgos estadísticos con precisión y solidez conceptual.

Competencia específica

Al integrar los aspectos declarativos, procedimentales, actitudinales y contextuales de esta unidad, será capaz de:

Seleccionar de manera adecuada las pruebas estadísticas pertinentes al análisis de datos provenientes de una sola muestra, distinguiendo entre métodos paramétricos y no paramétricos según las características de los datos, dominar herramientas como las pruebas z, t, Kolmogórov-Smirnov, chi-cuadrada y binomial, así como la interpretación crítica de los resultados obtenidos y formular conclusiones sólidas, sustentadas en evidencia empírica y fundamentadas en principios estadísticos, fortaleciendo su capacidad para tomar decisiones informadas en contextos académicos, clínicos o de investigación.

6.1. Datos cuantitativos

La inferencia estadística aplicada a datos cuantitativos permite estimar parámetros poblacionales y contrastar hipótesis con variables numéricas, tales como la media, la varianza o la desviación estándar. Estas pruebas constituyen herramientas esenciales en la investigación médica y clínica, al facilitar la evaluación objetiva de tratamientos, procedimientos o condiciones fisiológicas mediante el análisis de datos derivados de muestras representativas.

Por ejemplo, mediante estas técnicas es posible determinar si un fármaco reduce de manera significativa los niveles de glucosa en sangre, si un nuevo abordaje quirúrgico disminuye el tiempo promedio de hospitalización, o si la presión arterial observada en un grupo de pacientes difiere de los valores esperados en la población general (Montgomery et al., 2021; Norman y Streiner, 2021).

Las principales pruebas empleadas para este tipo de análisis son la prueba z y la prueba t , cuya selección depende de factores como el tamaño muestral y el conocimiento previo sobre la desviación estándar de la población. La prueba z es apropiada cuando se dispone de muestras grandes ($n \geq 30$) y se conoce la desviación estándar poblacional, mientras que la prueba t se utiliza cuando el tamaño muestral es pequeño ($n < 30$) y/o la desviación estándar poblacional es desconocida (Kim, 2015).

En las siguientes secciones se detallan las características formales de cada una de estas pruebas, su formulación matemática, los supuestos que las sustentan y su aplicación en contextos clínicos reales. Esta comprensión permite fundamentar con rigor los análisis y decisiones estadísticas sobre variables cuantitativas.

6.1.1. Prueba z y t para una muestra

En estadística inferencial, las pruebas z y t para una muestra constituyen herramientas fundamentales que permiten contrastar si la media observada en una muestra difiere significativamente de una media poblacional conocida o hipotética. El objetivo de estas pruebas es establecer, con base en la evidencia empírica, si una diferencia observada es atribuible al azar o si refleja un efecto real en la población (Kwak y Kim, 2017).

Hipótesis estadísticas

El procedimiento inicia con la formulación de dos hipótesis complementarias:

- **Hipótesis nula (H_0):** afirma que no existe diferencia entre la media muestral y la media poblacional
- **Hipótesis alternativa (H_a):** sostiene que existe una diferencia significativa, en concordancia con el planteamiento del estudio.

Estas hipótesis constituyen el marco lógico de interpretación del análisis estadístico. A través del cálculo del estadístico de prueba y su comparación con valores críticos de referencia o con el valor p, se determina si debe rechazarse o no la hipótesis nula, permitiendo sustentar inferencias basadas en datos cuantitativos.

Prueba z para una muestra

La prueba z se aplica en condiciones específicas:

- Cuando se conoce la desviación estándar de la población (σ).
- Cuando el tamaño de la muestra es mayor o igual a 30 ($n \geq 30$).
- En muestras menores a 30, si se asume que la población sigue una distribución normal.

Esta prueba resulta útil en estudios donde se dispone de información poblacional sólida, derivada de investigaciones previas o grandes bases de datos, lo que permite evaluar con precisión si la media muestral difiere significativamente de un valor poblacional teórico o esperado (Montgomery et al., 2021).

Fórmula de la prueba z

La prueba z se utiliza para determinar si la media de una muestra difiere de manera significativa de una media poblacional conocida o hipotética. Su formulación matemática se expresa de la siguiente manera (Figura 38):

Figura 38

Fórmula prueba z

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

Fuente: elaboración propia (2025).

\bar{x} : Media de la muestra.

μ : media poblacional (hipotética o conocida).

σ : desviación estándar de la población.

n: Tamaño de la muestra.

Este valor z se interpreta en función de la distribución normal estándar, cuyo valor esperado es 0 y cuya desviación estándar es 1. Si el valor calculado excede el umbral del nivel de significancia (α), se considera que la diferencia observada es estadísticamente significativa. En la práctica, para un nivel de confianza del 95 % ($\alpha = 0.05$), el valor crítico de z es ± 1.96 . Esto implica que, si el valor calculado se encuentra fuera de este rango, la hipótesis nula es rechazada en favor de la alternativa, sugiriendo que la media muestral difiere significativamente de la media poblacional.

Ejemplo: un médico desea evaluar la efectividad de un nuevo medicamento antihipertensivo. Estudios previos establecen que la presión arterial sistólica (PAS) promedio de la población general es de 120 mmHg. Se desea saber si, tras administrar el nuevo medicamento, la PAS promedio de los pacientes disminuye significativamente respecto a este valor.

Tamaño de la muestra (n): 25 pacientes.

Media muestral (\bar{x}): 116 mmHg

Desviación estándar poblacional (σ): 10 mmHg

Nivel de significancia (α): 0.05 (95 % de confianza)

Hipótesis:

$H_0: \mu = 120$ mmHg (el medicamento no reduce la presión arterial).

$H_a: \mu < 120$ mmHg (el medicamento reduce la presión arterial).

Figura 39

Cálculo de estadístico z

$$z = \frac{116 - 120}{\frac{10}{\sqrt{25}}} = \frac{-4}{2} = -2.0$$

Fuente: elaboración propia (2025).

Interpretación

El valor crítico de z para una prueba unilateral con $\alpha = 0.05$ es -1.645. Dado que el estadístico calculado $z = -2.0$ es menor que el valor crítico, se rechaza la hipótesis nula, lo que indica que existe evidencia estadística para afirmar que el medicamento produce una disminución significativa en la presión arterial sistólica en la muestra evaluada. La prueba z es esencial en la investigación cuantitativa para evaluar si las

diferencias observadas en una muestra reflejan cambios reales en la población o si son atribuibles a la variabilidad aleatoria inherente al proceso de muestreo.

Prueba t para una muestra

La prueba t para una muestra constituye una herramienta fundamental en la estadística inferencial, utilizada para contrastar si la media observada en una muestra difiere significativamente de un valor de referencia poblacional. Este procedimiento es particularmente pertinente cuando se desconoce la desviación estándar poblacional y el tamaño de la muestra es relativamente pequeño ($n < 30$).

A diferencia de la prueba z, que requiere el conocimiento de la desviación estándar de la población, la prueba t se basa en la desviación estándar de la muestra como estimador de la variabilidad. Este aspecto confiere a la distribución de probabilidad utilizada una mayor dispersión y dependencia del número de grados de libertad ($n-1$), modelándose mediante la distribución t de Student, que ofrece un marco robusto para el análisis en contextos clínicos, sociales y experimentales donde no se dispone de parámetros poblacionales conocidos y se trabaja con muestras reducidas (Navidi, 2021; Lakens, 2022; Pagano et al., 2022). La fórmula general de la prueba t es (figura 40):

Figura 40

Fórmula prueba t

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

Fuente: elaboración propia (2025).

Donde:

\bar{x} : es la media muestral

μ : es la media poblacional hipotética

s: es la desviación estándar de la muestra

n: es el tamaño de la muestra

Ejemplo: un médico investigador evalúa la eficacia de un nuevo medicamento hipoglucemiante destinado al tratamiento de la diabetes *mellitus* tipo 2, con base en datos históricos y literatura científica. Se establece que el nivel promedio de glucosa en sangre en ayuno de pacientes diabéticos sin tratamiento es de $\mu = 140$ mg/dL.

Para evaluar el efecto del nuevo fármaco, se administra el tratamiento a un grupo de 16 pacientes durante un mes y posteriormente se registran los niveles de glucosa en sangre en ayuno. El análisis muestral arroja una media de 130 mg/dL, con una desviación estándar de 12 mg/dL. Se plantea la siguiente hipótesis estadística:

- H_0 : $\mu = 140$ mg/dL (el medicamento no reduce los niveles de glucosa).
- H_a : $\mu < 140$ mg/dL (el medicamento reduce los niveles de glucosa).
- Nivel de significancia: $\alpha = 0.05$ (confianza del 95 %).

Aplicando la fórmula (figura 41):

Figura 41

Cálculo de estadístico t

$$t = \frac{130 - 140}{12/\sqrt{16}} = \frac{-10}{3} = -3.33$$

Fuente: elaboración propia (2025).

Interpretación

Se obtiene un valor t de -3.33 , que debe compararse con el valor crítico de la distribución t de Student con 15 grados de libertad ($n-1$). Para $\alpha = 0.05$ en una prueba unilateral, el valor crítico es -1.753 . Dado que $-3.33 < -1.753$, se rechaza la hipótesis nula. Así, existe evidencia de que el nuevo medicamento reduce los niveles de glucosa en sangre en pacientes con diabetes tipo 2, al menos en la muestra analizada. No obstante, la interpretación clínica requiere considerar otros factores como la magnitud del efecto, la relevancia clínica del descenso observado, la seguridad del medicamento y su aplicabilidad en poblaciones más amplias.

La prueba t para una muestra permite evaluar, con fundamentos estadísticos robustos, si una intervención clínica tiene un efecto medible sobre una variable cuantitativa cuando la información poblacional completa no está disponible.

6.2. Datos cualitativos

El análisis estadístico de datos cualitativos constituye una herramienta esencial en disciplinas como la medicina, las ciencias sociales, la psicología, la epidemiología y la biología (Woodward, 2023). A diferencia de los datos cuantitativos, que se expresan en valores numéricos y permiten operaciones aritméticas, los datos cualitativos representan atributos, categorías o cualidades no numéricas, cuya función principal es clasificar a los sujetos u objetos de estudio en grupos con características comunes.

Estas variables pueden ser de tipo nominal (sin un orden intrínseco, como el grupo sanguíneo o el tipo de enfermedad) u ordinal (con un orden implícito, como el grado de satisfacción o la severidad de los síntomas).

La interpretación adecuada de los datos cualitativos reviste una importancia particular en contextos clínicos y poblacionales, dado que múltiples decisiones diagnósticas, terapéuticas y estratégicas se fundamentan en la distribución y la frecuencia de estas categorías. Ejemplos relevantes incluyen la evaluación de diferencias en la prevalencia de enfermedades según el sexo, la aceptación de un tratamiento en distintos grupos etarios o la asociación entre la clase social y el acceso a servicios médicos (Agresti, 2019; Whatley, 2022). Para el análisis estadístico de variables cualitativas existen métodos específicos que permiten evaluar hipótesis, determinar asociaciones, contrastar distribuciones empíricas y estimar probabilidades categóricas. Entre las pruebas más empleadas se destacan:

- Kolmogórov-Smirnov (K-S), utilizada para comparar distribuciones empíricas con una distribución teórica y con datos ordinales.
- Chi-cuadrada (χ^2), ampliamente utilizada para analizar la independencia entre variables categóricas y evaluar la bondad de ajuste de una distribución observada respecto de una distribución esperada.
- Distribución binomial, empleada para modelar situaciones dicotómicas (éxito/fracaso), permitiendo calcular probabilidades acumuladas en eventos de naturaleza binaria (Sheskin, 2020; Bewick et al., 2005).

La selección de la técnica estadística adecuada depende de la naturaleza de la variable cualitativa y la estructura del conjunto de datos. Estas herramientas permiten determinar la significancia estadística de los hallazgos y estimar con precisión el grado de incertidumbre asociado a las decisiones inferenciales (Woodward, 2023).

6.2.1. Kolmogórov-Smirnov

La prueba de Kolmogórov-Smirnov (K-S) constituye una técnica no paramétrica empleada para contrastar la congruencia entre una distribución empírica y una distribución teórica previamente definida, o bien para evaluar si dos muestras independientes provienen de la misma distribución poblacional. Aunque su uso se ha consolidado en el análisis de variables cuantitativas continuas, su aplicación puede extenderse a datos ordinales o categóricos siempre que estos se transformen en proporciones acumuladas o frecuencias relativas estandarizadas (Motulsky, 2022; Pagano et al., 2022).

La prueba compara funciones de distribución acumulada, ya sea empírica contra teórica (prueba de una muestra) o empírica contra empírica (prueba de dos muestras). El estadístico de Kolmogórov-Smirnov corresponde a la máxima diferencia absoluta entre dichas distribuciones acumuladas. Este valor se contrasta con puntos críticos tabulados o con valores p derivados de simulaciones, lo que permite decidir, bajo un nivel de significancia preestablecido, si se acepta o se rechaza la hipótesis nula. En el caso de la prueba de una muestra, la hipótesis nula plantea que los datos provienen de la distribución teórica especificada; en la prueba de dos muestras, sostiene que ambas derivan de la misma distribución poblacional.

Destaca su carácter no paramétrico, lo que la hace independiente de supuestos de normalidad o de homogeneidad de varianzas, a diferencia de pruebas paramétricas más restrictivas. Sin embargo, su aplicación a datos cualitativos es menos común y suele restringirse a contextos en los que las categorías se expresan en escalas ordinales o en representaciones acumulativas.

Aplicaciones clínicas y metodológicas

En la investigación clínica y en la gestión hospitalaria, la prueba de Kolmogórov-Smirnov verifica la adecuación de los datos a distribuciones teóricas como la normal, la exponencial o la uniforme. Ejemplos de aplicación incluyen la evaluación de tiempos de espera en servicios de urgencias, la duración de estancias hospitalarias, las puntuaciones en escalas de calidad percibida y la distribución de indicadores funcionales en rehabilitación. La validez de los análisis posteriores depende de la comprobación de la forma de la distribución subyacente (Sheskin, 2020; Konopatov et al., 2024).

Ejemplo: la gerencia de un hospital desea mejorar su servicio de consulta externa en relación con los tiempos de espera de los pacientes, quienes han reportado que son un factor determinante en la percepción de la calidad del servicio. Dado que muchas técnicas estadísticas para el modelado y optimización de procesos asumen una distribución normal de los datos, es indispensable verificar si esta suposición se cumple en los tiempos de espera observados.

Descripción del estudio

Se recopila una muestra aleatoria de los tiempos de espera registrados en 50 pacientes que acudieron a consulta externa durante una semana; los valores obtenidos oscilan entre 12 y 95 minutos. A partir de estos datos, se construye la función de distribución empírica de los tiempos de espera y se compara con una distribución normal con media y desviación estándar calculadas a partir de la muestra.

Hipótesis de prueba

- **Hipótesis nula (H_0):** los tiempos de espera en la consulta externa siguen una distribución normal.
- **Hipótesis alternativa (H_a):** los tiempos de espera en la consulta externa no siguen una distribución normal.

Datos de los tiempos de espera (en minutos):

12	15	17	18	20	22	24	25	27	30
31	32	34	35	37	38	40	41	43	45
46	48	50	51	53	55	57	58	60	62
63	65	67	68	70	71	73	75	77	78
80	82	83	85	87	88	90	92	94	95

Interpretación

Mediante la prueba de Kolmogórov-Smirnov, se calcula la diferencia máxima entre la función de distribución empírica de los tiempos de espera observados y la distribución normal teórica correspondiente. Si esta diferencia excede el valor crítico determinado para el tamaño muestral y el nivel de significancia ($\alpha = 0.05$), se rechaza la hipótesis nula y se concluye que los datos no se distribuyen normalmente. La correcta aplicación de esta prueba permite establecer con rigor si los tiempos de espera se ajustan a los supuestos de normalidad, lo cual condiciona el uso posterior de técnicas paramétricas.

6.2.2. Chi-cuadrada

La prueba chi-cuadrada (χ^2) constituye una de las herramientas estadísticas más empleadas en el análisis de datos cualitativos, al permitir contrastar la congruencia entre las frecuencias observadas en una muestra y las frecuencias que se esperarían bajo un modelo teórico o una hipótesis estadística. Su carácter versátil y su aplicabilidad en múltiples disciplinas, particularmente en las ciencias médicas, epidemiológicas y sociales, la convierten en una prueba fundamental de la estadística inferencial (Pilamunga et al., 2024).

En términos conceptuales, la prueba se basa en la comparación entre las frecuencias observadas (f_o) y las frecuencias esperadas (f_e), con el fin de evaluar si las discrepancias son atribuibles al azar o si constituyen evidencia estadísticamente significativa. La hipótesis nula (H_0) plantea la ausencia de relación o diferencia significativa entre las variables en estudio, mientras que la hipótesis alternativa (H_a) sostiene la existencia de una asociación o ajuste deficiente al modelo esperado.

Tipos de prueba chi-cuadrada

Existen principalmente dos aplicaciones de esta prueba:

- Prueba de bondad de ajuste: valúa si la distribución de una variable categórica se ajusta a una distribución teórica o histórica. Es útil en escenarios donde se desea verificar la homogeneidad o uniformidad en la distribución de una característica, como la prevalencia esperada de un fenotipo genético en una población.
- Prueba de independencia (o asociación): analiza si dos variables categóricas están relacionadas a través de una tabla de contingencia que cruza sus categorías. Se trata de la forma más utilizada en medicina y ciencias sociales, ya que permite determinar asociaciones significativas entre factores de riesgo y desenlaces clínicos, así como entre intervenciones terapéuticas y resultados sanitarios (Llauce et al., 2024).

Aplicaciones metodológicas y clínicas

En el ámbito de la salud, la prueba chi-cuadrada se emplea con frecuencia en estudios observacionales, encuestas poblacionales y análisis de registros clínicos. Ejemplos de aplicación incluyen la exploración de la asociación entre el sexo y la prevalencia de una enfermedad, el vínculo entre una intervención terapéutica y la respuesta clínica, o la relación entre la satisfacción de los pacientes y la calidad de los servicios hospitalarios.

Ejemplo: la administración de un hospital busca determinar si el nivel de satisfacción de los pacientes está asociado con el tipo de servicio médico que reciben; para ello, se diseña un estudio transversal con una muestra de 150 pacientes que, al finalizar su atención, completan un cuestionario de satisfacción. Las categorías consideradas son:

- Tipo de servicio médico: consulta externa, urgencias, hospitalización.
- Nivel de satisfacción: satisfecho, neutral, insatisfecho.

Tabla de contingencia (hipotética)

Tabla 15. Tabla de contingencia satisfacción de los pacientes

Nivel de satisfacción	Consulta externa	Urgencias	Hospitalización	Total
Satisfecho	30	20	15	65
Neutral	5	10	15	50
Insatisfecho	15	10	10	35
Total	70	40	40	150

Fuente: elaboración propia (2025).

Formulación de hipótesis

- Hipótesis nula (H₀): no existe relación entre el tipo de servicio médico recibido y el nivel de satisfacción del paciente.
- Hipótesis alternativa (H_a): existe una relación significativa entre el tipo de servicio médico recibido y el nivel de satisfacción del paciente.

Procedimiento

Se calculan las frecuencias esperadas para cada celda de la tabla bajo la hipótesis de independencia, mediante la fórmula (Figura 42):

Figura 42

Fórmula de frecuencias esperadas

$$f_e = \frac{(\text{total fila}) \times (\text{total columna})}{\text{total general}}$$

Fuente: elaboración propia (2025).

Posteriormente, se aplica la fórmula del estadístico de prueba chi-cuadrada (Figura 43):

Figura 43

Fórmula estadístico chi-cuadrada

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

Fuente: elaboración propia (2025)

Donde:

f_o es la frecuencia observada

f_e es la frecuencia esperada

El valor resultante se compara con el valor crítico de la distribución chi-cuadrada con grados de libertad (Figura 44):

Figura 44

Fórmula grados de libertad distribución chi cuadrada

$$gl = (r - 1)(c - 1)$$

Fuente: elaboración propia (2025).

Si el valor calculado supera el valor crítico (o si el valor p es menor al nivel de significancia $\alpha = 0.05$), se rechaza la hipótesis nula.

Interpretación

Si se concluye que existe una asociación significativa, esto puede indicar que la experiencia del paciente, medida por su nivel de satisfacción, está influida por el tipo de atención médica recibida. Esta información puede ser utilizada para diseñar estrategias diferenciadas de mejora en los servicios hospitalarios.

6.2.3. Binomial

La distribución binomial constituye uno de los modelos fundamentales de la estadística inferencial para el análisis de variables aleatorias discretas, aplicables a experimentos donde cada observación puede tener dos resultados mutuamente excluyentes: éxito o fracaso. Esta característica la convierte en una herramienta de gran utilidad en contextos biomédicos y clínicos, donde es frecuente evaluar fenómenos dicotómicos, como la presencia o ausencia de una enfermedad, la efectividad o ineficacia de un tratamiento, o la ocurrencia o no de un evento adverso (Rosner, 2020).

En términos formales, la distribución binomial permite calcular la probabilidad de observar un número específico de éxitos en una secuencia de ensayos independientes, cada uno con la misma probabilidad de éxito. Su formulación se sustenta en cinco características esenciales (Ríos y Peña, 2020):

- Resultados dicotómicos (éxito/fracaso): cada ensayo da lugar a uno de dos posibles resultados. El “éxito” se refiere a la ocurrencia del evento de interés (por ejemplo, la respuesta favorable a un fármaco), mientras que el “fracaso” indica su ausencia.
- Número fijo de ensayos (n): el total de repeticiones del experimento está previamente determinado y permanece constante. En el ámbito clínico, este valor puede corresponder al número total de pacientes incluidos en un estudio.
- Independencia entre ensayos: se asume que el resultado de un ensayo no influye en los resultados de los demás, lo que garantiza la validez del modelo probabilístico.
- Probabilidad constante de éxito (p): la probabilidad de éxito es la misma en cada ensayo. En investigación médica, esto se traduce en suponer que un tratamiento mantiene la eficacia en todos los pacientes del estudio bajo condiciones controladas.
- Variable aleatoria discreta (X): se define como el número total de éxitos observados en los n ensayos. Su función de probabilidad se expresa mediante la fórmula (Figura 45):

Figura 45

Distribución de probabilidad variable aleatoria discreta

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Fuente: elaboración propia (2025).

Donde:

n: es el número total de ensayos

k: es el número de éxitos deseados

p: es la probabilidad de éxito

$\binom{n}{k}$: representa el número de combinaciones posibles para obtener k éxitos en n ensayos

Aplicaciones clínicas y metodológicas

La distribución binomial posee un amplio espectro de aplicaciones en medicina y ciencias de la salud, donde numerosos fenómenos pueden representarse como eventos dicotómicos (éxito/fracaso, presencia/ausencia, positivo/negativo). Su

utilidad es destacada en estudios observacionales, ensayos clínicos controlados y validación de pruebas diagnósticas, dado que permite cuantificar probabilidades y evaluar la consistencia de resultados bajo condiciones de variabilidad controlada (Rosner, 2020). Entre los usos clínicos más relevantes se incluyen:

- Evaluación de tratamientos: la distribución binomial permite calcular la probabilidad de que un número específico de pacientes se recupere después de recibir una nueva terapia, considerando la eficacia esperada del tratamiento.
- Diseño de estudios clínicos: estima el tamaño de muestra necesario para investigar eventos dicotómicos, como curación frente a no curación o éxito terapéutico frente a fracaso (Althubaiti, 2023).
- Análisis de pruebas diagnósticas: facilita la estimación de probabilidades asociadas a la obtención de resultados positivos o negativos al aplicar una prueba diagnóstica con sensibilidad y especificidad conocidas (Ríos y Peña, 2020).

Propiedades de la distribución binomial

Además de su estructura probabilística sustentada en ensayos dicotómicos independientes con probabilidad constante de éxito, la distribución binomial presenta propiedades matemáticas que favorecen su aplicación práctica en contextos clínicos y epidemiológicos (Wasserman, 2021).

Esperanza matemática $E(X)$: la esperanza matemática o media de la distribución binomial se expresa como (Figura 46):

Figura 46

Expresión esperanza matemática

$$E(X) = n \cdot p$$

Fuente: elaboración propia (2025).

Donde:

n: representa el número total de ensayos o pacientes

p: es la probabilidad de éxito en cada ensayo

Esta expresión representa el número promedio de éxitos esperados en un número fijo de ensayos. Por ejemplo, si un tratamiento tiene una eficacia del 70 % y se administra a 20 pacientes, se espera en promedio que 14 de ellos respondan positivamente al tratamiento.

Varianza $\text{Var}(X)$: la varianza, que mide la dispersión de los datos respecto a la media, se calcula mediante la fórmula (Figura 47):

Figura 47

Fórmula varianza $\text{var}(x)$

$$\text{Var}(X) = n \cdot p \cdot (1 - p)$$

Fuente: elaboración propia (2025).

Una varianza mayor indica mayor variabilidad en los posibles resultados observados. En medicina, esta propiedad permite cuantificar la incertidumbre asociada a la respuesta terapéutica esperada en un conjunto de pacientes.

Simetría y sesgo: la forma de la distribución binomial está determinada por el valor de p .

- Si $p = 0.5$, la distribución es perfectamente simétrica.
- Si $p < 0.5$, la distribución presenta sesgo hacia la derecha (asimetría negativa).
- Si $p > 0.5$, la distribución presenta sesgo hacia la izquierda (asimetría positiva).

Esta característica permite interpretar los resultados, ya que las distribuciones sesgadas pueden alterar la percepción de la probabilidad acumulada y los valores extremos.

Aplicaciones prácticas en medicina

La distribución binomial constituye una herramienta estadística de gran utilidad en el campo de la salud, ya que permite modelar fenómenos clínicos dicotómicos, tales como la presencia o ausencia de una enfermedad, la respuesta positiva o negativa a un tratamiento o el resultado de una prueba diagnóstica. Su aplicación práctica abarca múltiples áreas de la medicina basada en evidencia, proporcionando un marco probabilístico sólido para la interpretación de resultados (Ríos y Peña, 2020; Rosner, 2020). Entre las aplicaciones más frecuentes destacan:

- Evaluación de tratamientos médicos: la distribución binomial permite estimar la probabilidad de que un número específico de pacientes mejore tras recibir un tratamiento, lo que resulta esencial para validar la eficacia y seguridad de nuevas intervenciones clínicas y terapéuticas (Althubaiti, 2023).
- Interpretación de pruebas diagnósticas: facilita el cálculo de probabilidades de obtener resultados positivos o negativos, en función de la sensibilidad y especificidad establecidas para una prueba diagnóstica.
- Diseño de estudios epidemiológicos: se emplea para estimar la frecuencia esperada de aparición de un evento, como una infección o recaída, dentro de una cohorte bajo el supuesto de riesgo constante, lo cual permite estructurar diseños metodológicos robustos.

Ventajas del uso de la distribución binomial en medicina

El empleo de la distribución binomial en contextos médicos presenta múltiples ventajas metodológicas y aplicadas, que fortalecen la validez de los análisis estadísticos clínicos y epidemiológicos. Una de sus principales ventajas radica en la capacidad de generar estimaciones precisas incluso en escenarios con datos limitados, que resulta de gran utilidad en estudios piloto, investigaciones preliminares o ensayos con poblaciones reducidas, donde métodos estadísticos más complejos podrían resultar inaplicables (Wasserman, 2021).

Asimismo, la distribución binomial proporciona un marco predictivo para realizar estimaciones probabilísticas sobre la ocurrencia de eventos clínicos dicotómicos. Este aspecto incrementa la objetividad de las decisiones médicas al reducir la dependencia de juicios subjetivos, otorgando mayor solidez a las conclusiones.

Otra ventaja relevante se relaciona con el diseño de estudios clínicos, ya que la distribución binomial facilita el cálculo del tamaño de muestra necesario para alcanzar una potencia estadística adecuada, considerando tanto la proporción esperada de éxito como el nivel de significancia. Esta característica asegura que los estudios se encuentren correctamente dimensionados (Althubaiti, 2023).

Finalmente, su estructura matemática sencilla y la claridad de su interpretación permiten su uso no solo por especialistas en estadística, sino también por profesionales de diversas disciplinas vinculadas a la salud. Esto la convierte en una herramienta versátil para la docencia, la investigación biomédica y la práctica clínica.

Ejemplo: un médico desea estimar la probabilidad de que 10 a 15 pacientes tratados con un nuevo antihipertensivo respondan favorablemente; se conoce que la eficacia del fármaco es del 60 % ($p = 0.6$).

Datos

$n=15$ (número de pacientes tratados),

$p=0.6$ (probabilidad de respuesta positiva)

$k=10$ (número de éxitos deseado)

Aplicando la fórmula binomial (figura 48)

Figura 48

Fórmula distribución binomial

$$P(X = 10) = \binom{15}{10} \cdot (0.6)^{10} \cdot (0.4)^5$$

$$P(X = 10) = 3003 \cdot 0.00605 \cdot 0.01024 \approx 0.186$$

Fuente: elaboración propia (2025).

Interpretación

Existe aproximadamente un 18.6 % de probabilidad de que exactamente 10 de los 15 pacientes tratados con el nuevo medicamento experimenten una reducción significativa de la presión arterial. Este cálculo permite al médico tomar decisiones informadas sobre la eficacia del tratamiento en grupos de tamaño similar y proporciona una base para proyecciones en estudios clínicos más amplios.

Cierre de la unidad

Al concluir esta unidad, se consolida una comprensión sólida sobre las técnicas de inferencia estadística aplicadas a una muestra, abarcando datos cuantitativos y cualitativos. Esta base metodológica permite interpretar resultados con rigor científico, evaluar hipótesis con fundamentos estadísticos y tomar decisiones informadas en contextos clínicos, académicos y de investigación.

Unidad

Inferencia sobre muestras
independientes

Presentación de la unidad

Esta unidad aborda el análisis comparativo entre dos o más muestras independientes mediante pruebas estadísticas paramétricas (como t de Student y ANOVA) y no paramétricas (como U de Mann-Whitney y Kruskal-Wallis). Su aplicación permite evaluar diferencias significativas entre grupos en investigaciones clínicas y científicas, asegurando decisiones con evidencia estadística confiable.

Objetivos

Mediante el estudio de esta unidad se pretende alcanzar los siguientes propósitos fundamentales:

- Comparar datos entre muestras independientes utilizando técnicas estadísticas adecuadas, tanto paramétricas como no paramétricas, con base en la naturaleza y distribución de los datos.
- Aplicar ANOVA (análisis de varianza) para identificar diferencias significativas en contextos donde intervienen más de dos grupos o muestras independientes.
- Evaluar la pertinencia metodológica de cada prueba, seleccionando entre enfoques paramétricos o no paramétricos según el tamaño de muestra, distribución de los datos y nivel de medición de las variables involucradas.

Competencia específica

Al integrar los componentes declarativos, procedimentales, actitudinales y contextuales, el estudiante será capaz de:

Comparar datos de dos o más muestras independientes mediante pruebas estadísticas paramétricas y no paramétricas, seleccionando de manera crítica la técnica más apropiada en función de las características de los datos y del diseño del estudio. Aplicar correctamente pruebas como la t de Student, U de Mann-Whitney, ANOVA de un factor y de k factores, así como Kruskal-Wallis. Desarrollar la capacidad de interpretar los resultados y evaluar la significancia estadística, considerando el contexto del problema de investigación y fortaleciendo la toma de decisiones fundamentada en evidencia dentro de escenarios profesionales y de investigación.

7.1. Dos muestras

En el análisis estadístico, la comparación de dos muestras independientes constituye una estrategia clave para evaluar diferencias entre grupos expuestos a condiciones distintas. Esta comparación es especialmente pertinente en el ámbito biomédico, donde interesa determinar si un tratamiento produce un efecto distinto frente a otro, si existen diferencias entre grupos de pacientes con perfiles clínicos heterogéneos o si determinadas exposiciones modifican el riesgo de enfermedad (Aguinis et al., 2021; Higgins et al., 2024).

La elección de la prueba depende de las propiedades de los datos. Cuando las variables aproximan una distribución normal y las varianzas son homogéneas, es apropiado utilizar pruebas paramétricas como la *t* de Student para muestras independientes; si se sospecha heterocedasticidad, se recomienda la *t* de Welch en lugar de la *t* clásica. En escenarios donde los supuestos paramétricos no se cumplen—por ejemplo, distribuciones marcadamente asimétricas, presencia de valores atípicos influyentes o escalas de medición ordinales— se prefiere un enfoque no paramétrico como la *U* de Mann–Whitney (Bangdiwala, 2021; West et al., 2021; Zhou et al., 2023).

Debe considerarse, además, que los modelos gaussianos y la *t* de Student mantienen una robustez razonable ante desviaciones moderadas de la normalidad, de modo que el incumplimiento leve de este supuesto no obliga automáticamente a abandonar los métodos paramétricos; la decisión debe sostenerse en la magnitud de la asimetría, el tamaño muestral, la presencia de *outliers* y la igualdad de varianzas (Knief y Forstmeier, 2021; Zhou et al., 2023).

La *U* de Mann–Whitney no es, estrictamente, una “prueba de medianas” en todos los contextos: contrasta diferencias en la distribución (estocástica) entre grupos y solo puede interpretarse como una diferencia de medianas cuando las distribuciones comparten forma (misma asimetría y dispersión). Asimismo, el tratamiento de los empates (*ties*) y la posible desigualdad de varianzas pueden afectar su interpretación (Neuhäuser y Ruxton, 2024; Bangdiwala, 2021).

Tanto la *t* de Student/Welch como la *U* de Mann–Whitney proporcionan un valor *p* para inferir si la diferencia observada es estadísticamente significativa; no obstante, su aplicación rigurosa exige alinear el método con el contexto clínico, reportar estimaciones de efecto con intervalos de confianza y documentar supuestos, decisiones y resultados conforme a guías de reporte vigentes (Higgins et al., 2024; Aguinis et al., 2021). A continuación, se examinan ambas pruebas con criterios de aplicación, ventajas, limitaciones y ejemplos prácticos en contextos biomédicos.

7.1.1. t de Student

La prueba t de Student es una herramienta estadística paramétrica de amplia aplicación en la investigación cuantitativa, cuyo propósito es determinar si existen diferencias estadísticamente significativas entre las medias de dos grupos. Esta prueba ha sido fundamental en el desarrollo del análisis comparativo en campos como la medicina, la psicología, la biología y las ciencias sociales, al permitir contrastar el efecto de tratamientos, intervenciones o condiciones distintas entre grupos independientes o relacionados (Ruxton, 2006; West et al., 2021).

Su fundamento teórico se basa en la distribución t de Student, propuesta por William Gosset bajo el seudónimo “Student”, cuando el tamaño de la muestra es pequeño y la desviación estándar poblacional es desconocida. El procedimiento parte de una hipótesis nula (H_0) que plantea la igualdad entre medias poblacionales y una hipótesis alternativa (H_a) que postula la existencia de una diferencia significativa (Wasserman et al., 2024).

Tipos de prueba t de Student

Existen dos modalidades principales, seleccionadas en función del diseño experimental y de la naturaleza de las muestras comparadas:

1. Prueba t para muestras independientes: se emplea cuando se comparan dos grupos distintos entre sí, es decir, cuando no existe relación directa entre las observaciones de un grupo y las del otro. Es frecuente su uso en ensayos clínicos en los que se evalúan los efectos de dos tratamientos administrados a dos grupos de pacientes.

Hipótesis estadística

- Hipótesis nula (H_0): las medias de ambos grupos son iguales ($\mu_1 = \mu_2$).
- Hipótesis alternativa (H_a): las medias de los dos grupos difieren ($\mu_1 \neq \mu_2$).

Supuestos necesarios para su aplicación

- Normalidad de la distribución de los datos, verificada mediante pruebas de Shapiro-Wilk o Kolmogórov-Smirnov (Razali y Wah, 2011).
- Homogeneidad de varianzas, evaluada con las pruebas de Levene o Bartlett (Zhou et al., 2023).
- Independencia entre los grupos.

2. Prueba t para muestras relacionadas: se utiliza cuando las mediciones de un grupo están emparejadas con las del otro, como ocurre en estudios de antes y después en los mismos sujetos. Es común en medicina evaluar el impacto de una intervención específica en una misma cohorte de pacientes.

Hipótesis estadística

- Hipótesis nula (H_0): la media de las diferencias entre las mediciones emparejadas es cero ($\mu_d = 0$).
- Hipótesis alternativa (H_a): la media de las diferencias difiere de cero ($\mu_d \neq 0$).

Procedimiento para aplicar la prueba t de Student

- Formular las hipótesis: establecer claramente la hipótesis nula y la alternativa según el objetivo del análisis.
- Cálculo de la estadística t: depende del tipo de prueba; se utilizan fórmulas específicas que involucran las medias, las varianzas y los tamaños de las muestras.

Para muestras independientes (Figura 49):

Figura 49

Fórmula de la prueba t de Student para muestras independientes

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

Fuente: elaboración propia (2025).

Donde S_p^2 es la varianza combinada.

Para muestras pareadas (Figura 50):

Figura 50

Fórmula de la prueba t de Student para muestras pareadas

$$t = \frac{\bar{d}}{s_d / \sqrt{n}}$$

Fuente: elaboración propia (2025).

Donde \bar{d} es la media de las diferencias y S_d la desviación estándar de dichas diferencias.

- Determinación del valor crítico: con base en la distribución t para un nivel de significancia α (comúnmente 0.05) y los grados de libertad correspondientes.
- Toma de decisión estadística: si el valor calculado de t excede en valor absoluto al valor crítico, o si el valor p es menor que α , se rechaza la hipótesis nula.

Ventajas

- Constituye una prueba estadística robusta y de fácil interpretación (Ruxton, 2006).
- Es adecuada para variables continuas que se aproximen a una distribución normal (Razali y Wah, 2011).
- Permite detectar diferencias significativas incluso con tamaños de muestra relativamente pequeños, por ejemplo, en contextos clínicos con participantes limitados (West et al., 2021).

Limitaciones

- Es sensible a violaciones del supuesto de normalidad, especialmente en muestras pequeñas (Knief y Forstmeier, 2021).
- Requiere homogeneidad de varianzas en el caso de muestras independientes; de no cumplirse este supuesto, se recomienda emplear la corrección de Welch (West et al., 2021; Zhou et al., 2023).
- En presencia de distribuciones no normales o datos ordinales, resulta más adecuado aplicar pruebas no paramétricas como la U de Mann-Whitney (Bangdiwala, 2021).

Ejemplo: en un estudio de intervención clínica se desea evaluar la eficacia de un nuevo medicamento antidiabético. Los 60 pacientes se dividen aleatoriamente en dos grupos: uno recibe el fármaco activo y el otro un placebo.

- Grupo de tratamiento ($n = 30$): media de glucosa = 125 mg/dL, desviación estándar = 12 mg/dL.
- Grupo placebo ($n = 30$): media de glucosa = 135 mg/dL, desviación estándar = 15 mg/dL.
- Nivel de significancia: $\alpha = 0.05$.

Datos del estudio (Tabla 16)

Tabla 16. Datos de pacientes diabéticos

Grupo	Glucosa
Tratamiento	110
Tratamiento	115
Tratamiento	108
Placebo	125
Placebo	130
Placebo	128

Fuente: elaboración propia (2025).

Se verifica la normalidad en ambos grupos mediante la prueba de Shapiro-Wilk y la homogeneidad de varianzas mediante la prueba de Levene. Ambos supuestos se cumplen. Se aplica una prueba t para muestras independientes, obteniéndose un valor $p = 0.007$.

Interpretación

Dado que el valor p es menor que 0.05, se rechaza la hipótesis nula. Se concluye que existe una diferencia estadísticamente significativa en los niveles de glucosa entre los grupos, lo que sugiere un efecto clínico del nuevo tratamiento. La prueba t de Student es un instrumento estadístico esencial para comparar dos grupos y emitir inferencias válidas sobre diferencias en medias. Su correcta aplicación, bajo el cumplimiento de los supuestos requeridos, proporciona un marco riguroso para la evaluación de hipótesis en investigaciones científicas, contribuyendo a la consolidación de conclusiones basadas en evidencia empírica.

7.1.2 U de Mann-Whitney

La prueba U de Mann-Whitney, también denominada suma de rangos de Wilcoxon para muestras independientes, constituye una alternativa no paramétrica a la prueba t de Student para comparar dos grupos independientes. Esta técnica se utiliza cuando los datos no cumplen con los supuestos de normalidad o cuando las variables son de tipo ordinal, es decir, aquellas que representan categorías ordenadas sin que la diferencia entre niveles sea cuantificable (Bangdiwala, 2021; Nachar, 2008).

Su aplicabilidad se extiende a estudios clínicos y biomédicos en los que las condiciones de los datos no permiten el uso de pruebas paramétricas. A diferencia de las pruebas basadas en medias, la U de Mann–Whitney evalúa si las distribuciones de los dos grupos difieren significativamente entre sí, considerando la posición relativa de los valores en una escala común de rangos (Neuhäuser y Ruxton, 2024).

Aplicaciones clínicas y metodológicas

- Comparación de dos grupos cuando los datos son ordinales (por ejemplo, niveles de dolor en una escala del 1 al 10).
- Análisis de muestras pequeñas en las que no se puede asumir la normalidad.
- Evaluación de diferencias entre distribuciones, no solo entre medias (Nachar, 2008).

Procedimiento estadístico

El procedimiento para aplicar la prueba U de Mann–Whitney consta de los siguientes pasos:

- Recolección de datos: se obtienen los valores de los dos grupos independientes comparados.
- Asignación de rangos: se combinan todos los valores de ambas muestras en un conjunto ordenado en el que a cada observación se le asigna un rango (de menor a mayor).
- Cálculo de la suma de rangos por grupo: se determina la suma total de los rangos de cada grupo.
- Cálculo de la estadística U: se utiliza la suma de rangos y los tamaños de las muestras para determinar el valor de U mediante la fórmula correspondiente (Figura 51).

Figura 51

Fórmula de la estadística U

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 \quad \text{y} \quad U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2$$

Fuente: elaboración propia (2025).

Donde n_1 y n_2 son los tamaños de las muestras y R_1 y R_2 son las sumas de rangos de cada grupo. Se utiliza el valor más pequeño entre U_1 y U_2 para la comparación estadística.

- Determinación del valor crítico: se compara el valor calculado de U con los valores críticos de la distribución de Mann-Whitney o se obtiene el valor p asociado.
- Toma de decisión estadística: si el valor p es menor que el nivel de significancia (generalmente $\alpha = 0.05$), se rechaza la hipótesis nula.

Hipótesis estadística

- Hipótesis nula (H_0): las distribuciones de los dos grupos son iguales.
- Hipótesis alternativa (H_a): las distribuciones difieren entre los grupos (en forma, posición o ambas).

Ventajas

- No requiere que los datos cumplan los supuestos de normalidad ni de homogeneidad de varianzas (Nachar, 2008).
- Puede aplicarse con datos ordinales o transformados en rangos (Bangdiwala, 2021).
- Presenta mayor tolerancia frente a valores extremos o distribuciones asimétricas (Neuhäuser y Ruxton, 2024).
- Es útil en contextos clínicos reales en los que las condiciones ideales de los modelos paramétricos no se cumplen.

Limitaciones

- No estima la magnitud de la diferencia entre grupos, lo que limita su interpretación en términos de tamaño del efecto.
- Posee menor potencia estadística que la prueba t de Student cuando los supuestos paramétricos sí se cumplen (Ruxton, 2006).
- No permite incorporar ajustes en las covariables, a diferencia de técnicas como el análisis de covarianza (ANCOVA).

Ejemplo: en un ensayo clínico controlado, se comparan dos medicamentos analgésicos para determinar cuál reduce de manera más eficaz el dolor postoperatorio; se reclutan 30 pacientes, asignados aleatoriamente a dos grupos (15 por tratamiento). Cada paciente califica su dolor en una escala analógica de 1 a 10.

Dado que la variable es ordinal y los datos no cumplen con el supuesto de normalidad (verificado mediante la prueba de Shapiro-Wilk), se utiliza la prueba U de Mann-Whitney para contrastar los grupos.

Datos del estudio (Tabla 17)

Tabla 17. Datos de pacientes postoperados

Grupo	Dolor
Medicamento 1	4
Medicamento 1	3
Medicamento 1	5
Medicamento 2	6
Medicamento 2	7
Medicamento 2	5

Fuente: elaboración propia (2025).

La prueba produce un valor p de 0.121, por lo que no hay evidencia para concluir que uno de los medicamentos es más eficiente que el otro para reducir el dolor postoperatorio.

7.2. Tres o más muestras

En el ámbito de la estadística aplicada a las ciencias médicas, resulta frecuente la necesidad de comparar más de dos grupos independientes con el propósito de evaluar la eficacia de distintos tratamientos, establecer diferencias clínicas entre cohortes de pacientes o analizar la influencia de múltiples condiciones experimentales. Esta necesidad metodológica excede las capacidades de las pruebas diseñadas para dos grupos, como la t de Student, que no son adecuadas en estos escenarios, ya que su uso repetido incrementa la probabilidad de error tipo I, es decir, la detección de diferencias significativas donde en realidad no existen (Armstrong, 2014).

Para solventar esta limitación, se emplean procedimientos estadísticos diseñados para manejar comparaciones múltiples de forma simultánea. Entre las herramientas más reconocidas se encuentran el ANOVA (análisis de varianza) de un factor y de múltiples factores, que pertenecen al enfoque paramétrico, así como la prueba de Kruskal–Wallis, representativa del enfoque no paramétrico. Estas técnicas permiten identificar diferencias significativas entre grupos preservando la integridad estadística del análisis, controlando los niveles de error y asegurando inferencias válidas (McDonald, 2014; Mishra et al., 2019).

7.2.1. Anova de un factor

El análisis de varianza de un factor (ANOVA de un factor) es una técnica estadística paramétrica cuyo propósito es comparar las medias de tres o más grupos independientes, con el fin de identificar si al menos uno de ellos difiere significativamente de los demás. Esta herramienta se emplea en estudios médicos, tanto experimentales como observacionales, para investigar múltiples tratamientos, dosis o modalidades terapéuticas aplicadas a distintos grupos de pacientes (McDonald, 2014).

Supuestos del ANOVA de un factor

- Independencia de las observaciones: los datos recolectados en cada grupo deben ser independientes, es decir, que las mediciones de un grupo no influyan en las de otro.
- Normalidad de los datos: dentro de cada grupo, se espera que los datos sigan una distribución normal. Este supuesto puede verificarse mediante pruebas como Shapiro–Wilk o Kolmogorov–Smirnov (Razali y Wah, 2011).
- Homogeneidad de varianzas: las varianzas entre los grupos deben ser aproximadamente iguales, lo cual puede comprobarse mediante la prueba de Levene o de Bartlett (Zhou et al., 2023).

Fundamento metodológico

El ANOVA de un factor divide la variación total observada en los datos en dos componentes fundamentales:

- Variabilidad entre grupos: corresponde a las diferencias atribuibles al efecto del tratamiento o de la condición experimental. Se representa como la variación entre las medias de los grupos.
- Variabilidad dentro de los grupos: representa la variabilidad individual inherente a cada grupo, atribuida a factores no controlados o aleatorios.

El estadístico resultante del ANOVA, conocido como F de Fisher, se calcula como la razón entre la varianza entre grupos y la varianza dentro de los grupos. Un valor elevado del estadístico F indica que la variabilidad explicada por el modelo (entre grupos) es significativamente mayor que la variabilidad residual, lo que constituye evidencia suficiente para rechazar la hipótesis nula (Field, 2018).

Hipótesis de ANOVA

- Hipótesis nula (H_0): todas las medias poblacionales son iguales ($\mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$).
- Hipótesis alternativa (H_a): al menos una de las medias poblacionales difiere.

Cálculo del estadístico F y toma de decisiones

Una vez verificados los supuestos, el procedimiento analítico del ANOVA de un factor se desarrolla en varias etapas metodológicas:

- Cálculo del estadístico F: se determina a partir de la relación entre la varianza entre grupos y la varianza dentro de los grupos; el estadístico F se interpreta como una medida de cuán grande es la variabilidad explicada por los tratamientos respecto de la variabilidad aleatoria o residual. Su fórmula es (Figura 52):

Figura 52

Fórmula del estadístico F

$$F = \frac{MS_{\text{entre grupos}}}{MS_{\text{dentro de los grupos}}} = \frac{SS_{\text{entre}}/df_{\text{entre}}}{SS_{\text{dentro}}/df_{\text{dentro}}}$$

Fuente: elaboración propia (2025).

Donde:

MS: cuadrado medio

SS: suma de cuadrados

df: grados de libertad correspondientes

- Obtención del valor p: este valor se calcula con base en la distribución F para los grados de libertad correspondientes; el valor p indica la probabilidad de obtener un valor de F igual o extremo, asumiendo que la hipótesis nula es verdadera.
- Comparación con el nivel de significancia: tradicionalmente se utiliza un nivel de significancia de $\alpha = 0.05$. Si el valor p es menor que α , se concluye que existe evidencia estadísticamente significativa para rechazar la hipótesis nula.
- Pruebas post hoc: en caso de que el ANOVA indique diferencias significativas globales, es necesario realizar comparaciones múltiples entre los grupos para identificar específicamente cuáles difieren entre sí. Las pruebas post hoc más utilizadas son:

- Tukey: adecuada cuando hay igualdad de varianzas y tamaño de muestra similar.
- Bonferroni: más conservadora, útil cuando se desea reducir al mínimo el riesgo de error tipo I.
- Scheffé: flexible, aunque menos potente, útil cuando se comparan contrastes no planificados.

Aplicaciones clínicas y metodológicas

El ANOVA de un factor ha demostrado gran utilidad en estudios clínicos y biomédicos que requieren comparar el efecto de múltiples intervenciones o condiciones experimentales sobre una variable continua de interés. Algunos ejemplos son:

- Comparación de tratamientos farmacológicos: evaluar si tres medicamentos distintos producen efectos diferentes sobre la presión arterial sistólica en pacientes hipertensos (Field, 2018).
- Estudios dietéticos: determinar si existen diferencias estadísticamente significativas en los niveles de colesterol LDL entre pacientes que siguen una dieta baja en grasa, dieta mediterránea o dieta cetogénica (McDonald, 2014).
- Evaluación de la recuperación postoperatoria: comparar el tiempo de recuperación (en días) entre pacientes sometidos a tres protocolos distintos de rehabilitación.

Ventajas

- Reducción del error tipo I: permite realizar múltiples comparaciones sin inflar artificialmente el riesgo de falsos positivos, como ocurriría al aplicar pruebas t múltiples (Armstrong, 2014).
- Eficiencia analítica: integra la información de todos los grupos en un solo modelo estadístico, optimizando la interpretación de los resultados (Field, 2018).
- Versatilidad: puede aplicarse a una amplia gama de diseños experimentales en contextos clínicos, farmacológicos, epidemiológicos y sociales.

Limitaciones

- Sensibilidad a los supuestos: el incumplimiento de la normalidad o de la homogeneidad de las varianzas puede comprometer la validez del análisis, generando resultados espurios (Mishra et al., 2019; Zhou et al., 2023).

- Necesidad de pruebas complementarias: aunque el ANOVA permite identificar diferencias globales, no indica qué grupos específicos difieren entre sí sin recurrir a pruebas post hoc como Tukey, Bonferroni o Scheffé (Field, 2018).

El ANOVA de un factor compara tres o más grupos independientes. En contextos que evalúen la eficacia relativa de distintas intervenciones terapéuticas, protocolos clínicos o condiciones experimentales, esta prueba ofrece un marco metodológico robusto, siempre que se cumplan sus supuestos y se complemente con análisis posteriores.

Ejemplo: en un estudio clínico aleatorizado, un grupo de investigadores busca comparar la eficacia de tres medicamentos antihipertensivos diferentes (A, B y C) en la reducción de la presión arterial sistólica. Para ello, se reclutaron 90 pacientes con diagnóstico de hipertensión arterial, asignados aleatoriamente en tres grupos iguales de 30 participantes; cada grupo recibió uno de los medicamentos durante ocho semanas. Al finalizar el tratamiento, se midió la presión arterial sistólica en milímetros de mercurio (mmHg) en cada paciente para evaluar si existían diferencias estadísticamente significativas entre las medias de presión arterial sistólica de los tres grupos.

Hipótesis

- H_0 : las medias de presión arterial sistólica son iguales para los tres grupos.
- H_a : al menos una media es diferente.

Datos recolectados (Tabla 18)

Tabla 18. Datos de pacientes hipertensos

Medicamento	Presión arterial sistólica (mmHg)	Medicamento	Presión arterial sistólica (mmHg)	Medicamento	Presión arterial sistólica (mmHg)
A	142	B	127	C	133
A	139	B	139	C	134
A	143	B	130	C	129
A	148	B	125	C	129
A	139	B	134	C	139
A	139	B	124	C	142

Medicamento	Presión arterial		Medicamento	Presión arterial		Medicamento	Presión arterial	
	sistólica (mmHg)			sistólica (mmHg)			sistólica (mmHg)	
A	148		B	131		C	135	
A	144		B	120		C	140	
A	138		B	123		C	137	
A	143		B	131		C	132	
A	138		B	134		C	137	
A	138		B	131		C	143	
A	141		B	129		C	135	
A	130		B	128		C	143	
A	131		B	123		C	122	
A	137		B	126		C	139	
A	135		B	128		C	135	
A	142		B	135		C	134	
A	135		B	132		C	135	
A	133		B	121		C	125	
A	147		B	132		C	134	
A	139		B	128		C	137	
A	140		B	127		C	142	
A	133		B	133		C	132	
A	137		B	135		C	131	
A	141		B	135		C	132	
A	134		B	126		C	140	
A	142		B	128		C	137	
A	137		B	132		C	132	
A	139		B	135		C	138	

Fuente: elaboración propia (2025).

Supuestos del ANOVA evaluados

- Independencia: garantizada mediante asignación aleatoria.
- Normalidad: verificada mediante la prueba de Shapiro-Wilk, sin evidencia de desviaciones significativas.
- Homogeneidad de varianzas: evaluada mediante la prueba de Levene, que no mostró diferencias significativas entre las varianzas de los grupos.

Resultados estadísticos

Estadístico F: 31.22

Valor p: 6.01×10^{-11}

El valor p es mucho menor que el nivel de significancia habitual ($\alpha = 0.05$). Por tanto, se rechaza la hipótesis nula, dado que existen diferencias estadísticamente significativas entre al menos dos de los medicamentos evaluados. De esta forma, no todos los tratamientos producen el mismo efecto en la presión arterial sistólica; sin embargo, el ANOVA no especifica cuáles difieren entre sí. Por ello, se recomienda realizar un análisis *post hoc* (como la prueba de Tukey) para identificar los pares de medicamentos con diferencias significativas. Este análisis permite seleccionar el tratamiento más eficaz entre varias alternativas basadas en evidencia cuantitativa, contribuyendo a la optimización terapéutica y a una mejor atención al paciente con hipertensión arterial.

7.2.2. Anova de k factores

El análisis de varianza de k factores (ANOVA factorial) constituye una extensión del modelo de ANOVA de un solo factor, orientado a evaluar de forma simultánea el efecto de dos o más variables independientes, denominadas factores, sobre una variable dependiente cuantitativa. A diferencia del ANOVA de un factor, que únicamente examina la influencia de una sola condición experimental, el ANOVA factorial permite analizar tanto los efectos principales de cada factor por separado como las interacciones entre ellos. Esta característica lo convierte en una herramienta eficaz para estudios experimentales complejos en los que los factores no actúan de manera aislada, sino que pueden influir mutuamente (Field, 2018).

Supuestos fundamentales del ANOVA de k factores

Para garantizar la validez de los resultados obtenidos mediante este modelo, deben cumplirse los siguientes supuestos estadísticos:

- Independencia de las observaciones: las mediciones dentro de cada combinación de niveles de los factores deben ser independientes, de modo que la ocurrencia de un resultado no afecte a los demás.
- Normalidad: la variable dependiente debe aproximarse a una distribución normal en cada celda factorial (es decir, en cada combinación de niveles de los factores). Este supuesto puede verificarse mediante las pruebas de Kolmogorov–Smirnov o Shapiro–Wilk (Razali y Wah, 2011).

- Homogeneidad de varianzas: las varianzas deben ser comparables entre los distintos grupos mediante las pruebas de Levene o Bartlett (Zhou et al., 2023).

Ventajas

Este enfoque presenta diversas ventajas para analizar datos con múltiples variables independientes.

- Evaluación simultánea de múltiples factores: permite valorar en una sola prueba el efecto de varias variables independientes, reduciendo la necesidad de realizar múltiples análisis separados y, con ello, el riesgo de error tipo I (Armstrong, 2014).
- Identificación de interacciones: es capaz de detectar si el efecto de un factor depende del nivel de otro —útil en investigaciones médicas, farmacológicas y de ciencias del comportamiento— (Field, 2018).
- Incremento en el poder estadístico: al incorporar múltiples fuentes de variabilidad explicada, se reduce el error residual, lo que fortalece la capacidad del modelo para detectar efectos reales.

Limitaciones

A pesar de sus fortalezas, el ANOVA factorial también presenta ciertas restricciones:

- Tamaño muestral: requiere contar con un tamaño de muestra suficiente en cada celda factorial, lo que puede constituir un desafío logístico y ético en estudios clínicos con múltiples combinaciones de tratamientos (McDonald, 2014).
- Complejidad interpretativa: las interacciones de orden superior (tres o más factores) pueden dificultar la interpretación, especialmente si los efectos no son lineales o presentan patrones cruzados.

Aplicación en contextos médicos

El ANOVA de k factores tiene aplicaciones clínicas y biomédicas en contextos con múltiples condiciones que interactúan. Un ejemplo ilustrativo es un estudio diseñado para evaluar el efecto combinado de dos tipos de medicamentos hipoglucemiantes (Factor A) y tres niveles de ejercicio físico (Factor B) sobre los niveles de glucosa en sangre en pacientes con diabetes tipo 2. El modelo factorial permite responder a preguntas clave:

- ¿Existe un efecto significativo del medicamento en la reducción de la glucosa?
- ¿Influye el nivel de actividad física en el resultado?
- ¿El efecto de un medicamento depende del ejercicio realizado?

La inclusión de la interacción $A \times B$ en el modelo ofrece una perspectiva más completa al detectar patrones clínicamente relevantes que no serían visibles mediante análisis univariados o modelos simplificados. De este modo, se optimizan la eficacia terapéutica y la personalización del tratamiento (Maxwell et al., 2017).

El ANOVA de k factores ofrece una aproximación integral y rigurosa para el análisis de fenómenos multifactoriales complejos. Su correcta implementación enriquece la interpretación clínica de los hallazgos y contribuye a la mejora continua en la práctica médica basada en evidencia.

Ejemplo: un estudio analiza el efecto combinado de dos factores sobre el tiempo de recuperación postoperatoria en pacientes sometidos a cirugía ortopédica: el tipo de fisioterapia (Factor A) y el uso de analgésicos (Factor B). Se reclutaron 60 pacientes y se asignaron aleatoriamente a una de las combinaciones posibles entre los tres niveles del factor fisioterapia (sin fisioterapia, fisioterapia estándar y fisioterapia intensiva) y los dos niveles del factor analgésico (con o sin analgésico). El desenlace medido fue el número de días necesarios para alcanzar una movilidad funcional postquirúrgica estable.

Hipótesis

- H_0 : no existen diferencias estadísticamente significativas en el tiempo de recuperación debido a los efectos de los factores y su interacción.
- H_a : existen diferencias significativas en el tiempo de recuperación atribuibles al tipo de fisioterapia, al uso de analgésicos o a la interacción entre ambos factores.

Datos recolectados (Tabla 19)

Tabla 19. Datos de pacientes en recuperación postoperatoria de cirugía ortopédica

Tipo de fisioterapia	Uso de analgésico	Tiempo de recuperación	Tipo de fisioterapia	Uso de analgésico	Tiempo de recuperación
Sin fisioterapia	Sin analgésico	32.4835	Fisioterapia estándar	Con analgésico	18.0833
Sin fisioterapia	Sin analgésico	29.3086	Fisioterapia estándar	Con analgésico	19.2573
Sin fisioterapia	Sin analgésico	33.2384	Fisioterapia estándar	Con analgésico	15.5746
Sin fisioterapia	Sin analgésico	37.6151	Fisioterapia estándar	Con analgésico	15.2151
Sin fisioterapia	Sin analgésico	28.8292	Fisioterapia estándar	Con analgésico	23.2501
Sin fisioterapia	Sin analgésico	28.8293	Fisioterapia estándar	Con analgésico	25.4249
Sin fisioterapia	Sin analgésico	37.8960	Fisioterapia estándar	Con analgésico	19.7119
Sin fisioterapia	Sin analgésico	33.8371	Fisioterapia estándar	Con analgésico	24.0141
Sin fisioterapia	Sin analgésico	27.6526	Fisioterapia estándar	Con analgésico	21.4465
Sin fisioterapia	Sin analgésico	32.7128	Fisioterapia estándar	Con analgésico	17.4195
Sin fisioterapia	Con analgésico	37.3282	Fisioterapia intensiva	Sin analgésico	17.3409
Sin fisioterapia	Con analgésico	28.871	Fisioterapia intensiva	Sin analgésico	19.0713
Sin fisioterapia	Con analgésico	30.3376	Fisioterapia intensiva	Sin analgésico	22.4336
Sin fisioterapia	Con analgésico	22.8762	Fisioterapia intensiva	Sin analgésico	16.4451
Sin fisioterapia	Con analgésico	27.2780	Fisioterapia intensiva	Sin analgésico	15.5745
Sin fisioterapia	Con analgésico	30.5546	Fisioterapia intensiva	Sin analgésico	16.4947
Sin fisioterapia	Con analgésico	24.2450	Fisioterapia intensiva	Sin analgésico	20.7462
Sin fisioterapia	Con analgésico	31.8784	Fisioterapia intensiva	Sin analgésico	18.9862

Tipo de fisioterapia	Uso de analgésico	Tiempo de recuperación	Tipo de fisioterapia	Uso de analgésico	Tiempo de recuperación
Sin fisioterapia	Con analgésico	26.9968	Fisioterapia intensiva	Sin analgésico	16.4107
Sin fisioterapia	Con analgésico	28.5415	Fisioterapia intensiva	Sin analgésico	19.5398
Fisioterapia estándar	Sin analgésico	27.9538	Fisioterapia intensiva	Con analgésico	10.7538
Fisioterapia estándar	Sin analgésico	25.6854	Fisioterapia intensiva	Con analgésico	13.7380
Fisioterapia estándar	Sin analgésico	24.5374	Fisioterapia intensiva	Con analgésico	13.9718
Fisioterapia estándar	Sin analgésico	23.7955	Fisioterapia intensiva	Con analgésico	12.5931
Fisioterapia estándar	Sin analgésico	19.0859	Fisioterapia intensiva	Con analgésico	14.5161
Fisioterapia estándar	Sin analgésico	22.1206	Fisioterapia intensiva	Con analgésico	16.2121
Fisioterapia estándar	Sin analgésico	23.1574	Fisioterapia intensiva	Con analgésico	20.6585
Fisioterapia estándar	Sin analgésico	29.2284	Fisioterapia intensiva	Con analgésico	15.5237
Fisioterapia estándar	Sin analgésico	26.3744	Fisioterapia intensiva	Con analgésico	15.7726
Fisioterapia estándar	Sin analgésico	17.9478	Fisioterapia intensiva	Con analgésico	14.7766

Fuente: elaboración propia (2024).

El resultado del ANOVA de k factores mostró lo siguiente:

- El tipo de fisioterapia tiene un efecto altamente significativo sobre el tiempo de recuperación ($p = 0.027$). Al menos uno de los niveles de tratamiento se asocia con una duración significativamente distinta de la recuperación.
- El uso del analgésico también mostró un efecto significativo por sí mismo ($p = 0.001$), lo que sugiere que su administración acelera la recuperación.

El ANOVA de k factores permite concluir que tanto el tipo de fisioterapia como la administración de analgésicos impactan de manera independiente en la duración de la recuperación; la ausencia de interacción sugiere que los beneficios de un programa intensivo de fisioterapia se mantienen independientemente del uso concomitante de analgésicos y viceversa. Estos hallazgos pueden guiar decisiones

clínicas sobre el diseño de protocolos postoperatorios al optimizar los recursos terapéuticos para mejorar la recuperación funcional.

7.2.3. Kruskal-Wallis

En la investigación médica y en otras disciplinas de las ciencias de la salud, es frecuente trabajar con datos que no cumplen los supuestos necesarios para las técnicas estadísticas paramétricas como el análisis de varianza (ANOVA). Ante esta limitación, la prueba de Kruskal–Wallis constituye una alternativa robusta y eficiente para comparar tres o más grupos independientes, sin requerir que los datos sigan una distribución normal ni que las varianzas sean homogéneas (Mishra et al., 2019).

También denominada ANOVA de una vía por rangos, la prueba de Kruskal–Wallis es una extensión directa de la prueba de Mann–Whitney U para más de dos grupos. Se utiliza cuando la variable de interés es al menos ordinal y las observaciones de los distintos grupos son independientes entre sí. Es especialmente útil en estudios con escalas tipo Likert, evaluaciones subjetivas o mediciones que presentan sesgos de distribución o valores atípicos extremos (Nachar, 2008).

Fundamento metodológico

La lógica de la prueba se basa en reemplazar los valores observados por sus respectivos rangos dentro del conjunto de datos. Posteriormente, se calcula la suma de rangos para cada grupo y se obtiene el estadístico H de Kruskal–Wallis. Dicho estadístico se distribuye aproximadamente como un chi-cuadrado, con grados de libertad equivalentes al número de grupos menos uno, siempre que el tamaño muestral por grupo sea suficientemente grande (McDonald, 2014).

Supuestos metodológicos

Para que los resultados de la prueba de Kruskal-Wallis sean válidos, deben cumplirse los siguientes supuestos:

- Independencia de las observaciones: las mediciones realizadas en los distintos grupos no deben influenciarse entre sí.
- Escala de medición ordinal o superior: la variable dependiente debe poseer un orden lógico que permita asignar rangos significativos (Bangdiwala, 2021).

Aplicaciones clínicas y metodológicas

La prueba de Kruskal–Wallis posee un amplio rango de aplicaciones en investigación biomédica –estudios observacionales, ensayos clínicos preliminares o investigaciones exploratorias–. Ejemplos representativos incluyen:

- Comparación del nivel de dolor reportado entre pacientes que recibieron tres tipos distintos de tratamiento analgésico, cuando el dolor se cuantifica en escalas ordinales (por ejemplo, de 1 a 10).
- Evaluación de la satisfacción del paciente en diferentes hospitales o clínicas utilizando escalas cualitativas.
- Análisis de diferencias en puntuaciones clínicas no normalizadas –por ejemplo, cuestionarios sobre calidad de vida o escalas funcionales– entre grupos sometidos a diferentes intervenciones (Neuhäuser y Ruxton, 2024).

Ventajas

- Independencia de los supuestos paramétricos: no exige normalidad ni homogeneidad de las varianzas, lo que amplía su aplicabilidad.
- Robustez ante valores atípicos: su construcción por rangos lo hace menos sensible a observaciones extremas que sesgan las medias.
- Aplicabilidad en muestras pequeñas: su rendimiento es adecuado incluso con tamaños de muestra reducidos.

Limitaciones

- No identifica diferencias específicas entre grupos: si se detecta una diferencia global significativa, se requiere la aplicación de pruebas post hoc, como las de Dunn o Conover, para determinar entre qué grupos se presentan las diferencias (Armstrong, 2014).
- Menor potencia estadística: en comparación con ANOVA, cuando se cumplen los supuestos paramétricos, puede resultar menos sensible para detectar diferencias reales (McDonald, 2014).

La prueba de Kruskal–Wallis constituye una herramienta estadística esencial para comparar tres o más grupos cuando no se cumplen los supuestos de las pruebas paramétricas; permite realizar inferencias válidas a partir de datos ordinales, no normales y heterocedásticos. La correcta interpretación de sus resultados permite a los profesionales tomar decisiones informadas y basadas en evidencia, incluso en contextos metodológicamente complejos.

Ejemplo: en el ámbito de la salud, la satisfacción del paciente constituye un indicador clave de calidad en la atención médica, a fin de evaluar las diferencias en la percepción del servicio entre diversas instituciones, un grupo de investigadores llevó a cabo un estudio transversal comparativo entre tres hospitales públicos (Hospital A, B y C), utilizando una escala ordinal del 1 al 5 para medir la satisfacción de los usuarios (1 = muy insatisfecho, 5 = muy satisfecho). Se recolectaron datos de 50 pacientes por hospital, seleccionados mediante muestreo aleatorio simple durante el mismo tiempo, para garantizar la comparabilidad.

Hipótesis

- H_0 : las distribuciones de satisfacción son iguales en los tres hospitales.
- H_a : al menos una de las distribuciones de satisfacción difiere significativamente entre los hospitales.

Datos recolectados (Tabla 20):

Tabla 20. Datos recolectados sobre la satisfacción de pacientes sobre calidad de atención médica

Hospital	Puntuación satisfacción	Hospital	Puntuación satisfacción	Hospital	Puntuación satisfacción
A	2	B	5	C	2
A	4	B	4	C	4
A	3	B	5	C	4
A	2	B	4	C	4
A	1	B	3	C	5
A	1	B	5	C	3
A	1	B	1	C	4
A	3	B	2	C	5
A	2	B	1	C	3
A	3	B	3	C	2
A	1	B	3	C	3
A	4	B	2	C	3
A	3	B	4	C	5
A	1	B	3	C	5
A	1	B	2	C	4
A	1	B	3	C	5
A	2	B	2	C	5

Hospital	Puntuación satisfacción	Hospital	Puntuación satisfacción	Hospital	Puntuación satisfacción
A	2	B	4	C	3
A	2	B	1	C	5
A	1	B	5	C	4
A	2	B	4	C	5
A	1	B	2	C	5
A	1	B	1	C	4
A	2	B	4	C	3
A	2	B	4	C	3
A	3	B	4	C	4
A	1	B	4	C	5
A	2	B	1	C	5
A	2	B	3	C	2
A	1	B	2	C	4
A	2	B	4	C	4
A	1	B	4	C	3
A	1	B	3	C	3
A	4	B	1	C	4
A	4	B	3	C	5
A	3	B	3	C	4
A	2	B	4	C	4
A	1	B	4	C	5
A	2	B	4	C	4
A	2	B	3	C	5
A	1	B	2	C	5
A	2	B	4	C	3
A	1	B	4	C	4
A	4	B	3	C	4
A	1	B	4	C	3
A	2	B	3	C	2
A	2	B	3	C	4
A	2	B	3	C	4
A	2	B	1	C	2
A	1	B	2	C	3

Fuente: elaboración propia (2025).

La variable de estudio es ordinal y los datos no cumplen con los supuestos de normalidad ni de homocedasticidad requeridos por un análisis paramétrico (como ANOVA). Por ello, se utiliza la prueba no paramétrica de Kruskal-Wallis para comparar tres o más muestras independientes sin requerir una distribución específica.

Resultados

Estadístico H = 23.01

Valor p = 0.00001068

Dado que el valor p es significativamente menor que el umbral común de significancia ($\alpha=0.05$), se rechaza la hipótesis nula; esto indica que existe una diferencia estadísticamente significativa en los niveles de satisfacción reportados por los pacientes entre al menos dos de los hospitales evaluados.

La prueba de Kruskal-Wallis permitió detectar diferencias en la percepción del servicio hospitalario desde la perspectiva de los usuarios. No obstante, se recomienda aplicar pruebas post hoc no paramétricas (como Dunn-Bonferroni) para identificar las comparaciones con significancia.

Cierre de la unidad

Al finalizar esta unidad, habrá aprendido sobre los métodos estadísticos aplicables a muestras independientes de dos o más grupos. Estos métodos permiten a los estudiantes analizar datos de manera crítica, interpretar resultados con precisión y aplicar estas herramientas para la toma de decisiones fundamentadas en la práctica médica y científica, fortaleciendo su capacidad para desarrollar investigaciones y evaluar tratamientos basados en evidencia.

Unidad

A decorative graphic consisting of two overlapping circles. The top circle is a dark purple, and the bottom circle is a lighter purple. They overlap in the center, creating a darker purple area. The text is centered within this overlapping area.

Inferencia sobre muestras
relacionadas

Presentación de la unidad

Esta unidad aborda el análisis estadístico de datos obtenidos a partir de mediciones repetidas en las mismas unidades de estudio. Se estudian pruebas como la *t* de Student para muestras relacionadas, Wilcoxon, McNemar, Friedman y *Q* de Cochran, útiles para evaluar cambios o diferencias en contextos clínicos o experimentales. Estas pruebas permiten interpretar con precisión la significancia estadística dentro de un mismo grupo.

Objetivos

A través del desarrollo de los contenidos teóricos y prácticos de esta unidad, se pretende que el estudiante logre:

- Aplicar correctamente pruebas estadísticas diseñadas para el análisis de datos relacionados, con base en la naturaleza de la variable y el diseño del estudio.
- Distinguir entre las pruebas utilizadas para comparar dos condiciones dependientes y tres o más medidas relacionadas en el mismo grupo de sujetos.
- Interpretar los resultados obtenidos en estas pruebas, considerando los supuestos estadísticos, el tipo de variable, el contexto del problema y la pertinencia del análisis.

Competencia específica

Al integrar los aspectos declarativos, procedimentales, actitudinales y contextuales de esta unidad, el estudiante será capaz de:

Analizar muestras relacionadas de una población para explicar fenómenos mediante la interpretación estadística e inferencial, con medidas de tendencia central, dispersión y pruebas adecuadas al diseño dependiente, con un enfoque orientado a la solución de problemas en contextos reales de salud e investigación científica.

8.1. Dos muestras

El análisis estadístico de dos muestras relacionadas constituye un componente esencial en diseños experimentales y observacionales con mediciones repetidas en las mismas unidades de análisis. Este enfoque permite evaluar cambios, efectos o diferencias que se producen en un mismo sujeto, sistema o entidad antes y después de una intervención, o bajo dos condiciones distintas. Su principal ventaja radica en ofrecer una mayor sensibilidad para detectar efectos sutiles al controlar la variabilidad interindividual (Field, 2018).

En este contexto, se consideran tres pruebas estadísticas fundamentales: la *t* de Student para muestras relacionadas, la prueba no paramétrica de Wilcoxon y la prueba de McNemar. Cada una de estas técnicas se ha desarrollado para abordar distintas estructuras de datos y tipos de variables: continuas, ordinales o dicotómicas. La prueba *t* de Student para muestras relacionadas constituye una herramienta paramétrica empleada cuando los datos son continuos y se asume normalidad en la distribución de las diferencias. Esta prueba compara las medias de dos mediciones emparejadas y determina si existe una diferencia estadísticamente significativa entre ellas (McDonald, 2014).

Por su parte, la prueba de Wilcoxon para rangos con signo ofrece una alternativa no paramétrica que no requiere normalidad cuando se trabajan datos ordinales o cuando las diferencias no se distribuyen normalmente. Esta prueba evalúa tanto la magnitud como la dirección de los cambios entre dos condiciones relacionadas, proporcionando una solución robusta frente a violaciones de los supuestos paramétricos (Nachar, 2008). Finalmente, la prueba de McNemar se aplica a datos dicotómicos emparejados en diseños antes–después, donde el interés radica en detectar cambios en las proporciones de una misma muestra, como la mejoría o el empeoramiento de un estado clínico (Fagerland et al., 2013).

La elección adecuada entre estas pruebas depende del tipo de variable (numérica, ordinal o categórica), de la distribución de los datos y de los supuestos estadísticos que puedan verificarse. Una selección apropiada garantiza la validez del análisis estadístico y la precisión de las inferencias.

8.1.1. *t* de Student

La prueba *t* de Student para muestras relacionadas, también denominada *t* pareada o *t* dependiente, es una técnica paramétrica que compara las medias de dos mediciones obtenidas de manera emparejada sobre la misma unidad experimental. Resulta particularmente útil en diseños pretest–postest, ensayos clínicos longitudinales o estudios con dos condiciones distintas para un mismo grupo de sujetos (Ruxton, 2006).

Características principales

Esta prueba tiene como objetivo determinar si existe una diferencia estadísticamente significativa entre las medias de dos condiciones relacionadas. La naturaleza del emparejamiento radica en que cada valor de una muestra se vincula con otro, lo que permite controlar la variabilidad interindividual y aumentar la precisión del análisis. Entre sus principales características destacan:

- Es aplicable a datos cuantitativos continuos.
- Requiere que los datos provengan de mediciones repetidas en las mismas unidades (por ejemplo, presión arterial antes y después de un tratamiento).
- Supone que las diferencias entre los pares de observaciones siguen una distribución normal (Razali y Wah, 2011).

Hipótesis estadísticas

- Hipótesis nula (H_0): la media de las diferencias entre las dos mediciones es igual a cero ($\mu_d=0$), lo que implica que no hay efecto del tratamiento o de la intervención.
- Hipótesis alternativa (H_a): la media de las diferencias es distinta de cero ($\mu_d \neq 0$), lo cual sugiere un cambio significativo entre las condiciones evaluadas.

Estadístico de prueba (Figura 53):

Figura 53

Estadístico de prueba t Student

$$t = \frac{\bar{d}}{s_d / \sqrt{n}}$$

Fuente: elaboración propia (2025).

Donde:

\bar{d} es la media de las diferencias entre pares.

S_d es la desviación estándar de las diferencias.

n es el número de pares de observaciones.

Supuestos

Para que la prueba t para muestras relacionadas sea válida, deben cumplirse los siguientes supuestos:

- Normalidad de las diferencias: Las diferencias entre las observaciones emparejadas deben seguir una distribución aproximadamente normal. Este supuesto puede verificarse mediante las pruebas de Shapiro-Wilk y de Kolmogórov-Smirnov aplicadas a la distribución de las diferencias.
- Independencia: Aunque las mediciones están relacionadas dentro de cada par, los pares deben ser independientes entre sí.

Ventajas

- Ofrece una mayor eficiencia estadística al reducir la variabilidad atribuible a las diferencias individuales, al tratarse de un diseño intrasujeto.
- Es de fácil aplicación e interpretación cuando se cumplen los supuestos.
- Tiene mayor poder estadístico que la prueba para muestras independientes en contextos adecuados, lo que facilita la detección de efectos sutiles.

Limitaciones

- Es sensible a la presencia de valores atípicos que pueden distorsionar la media y la desviación estándar.
- No es adecuada cuando las diferencias entre pares no siguen una distribución normal; se recomienda emplear pruebas no paramétricas como la de Wilcoxon.

La prueba t de Student para muestras relacionadas es una herramienta aceptada en el análisis de datos dependientes, siempre y cuando se verifique el cumplimiento de los supuestos estadísticos requeridos. Su correcta aplicación contribuye a la generación de evidencia confiable en investigaciones clínicas, experimentales y observacionales.

Ejemplo: un investigador clínico busca determinar si un programa de ejercicio aeróbico de ocho semanas produce una reducción significativa en la presión arterial sistólica de pacientes con diagnóstico de hipertensión. Para este fin, se reclutan 20 pacientes adultos con presión arterial elevada y se registra su presión arterial sistólica antes y después de completar el programa de intervención física.

Diseño del estudio

- Variable dependiente: presión arterial sistólica (mmHg).
- Variable independiente: tiempo (antes vs. después del programa de ejercicio).
- Tipo de diseño: comparación de dos muestras relacionadas (mediciones repetidas en los mismos individuos).

Hipótesis estadística

- Hipótesis nula (H_0): no existe diferencia en la presión arterial sistólica antes y después del programa.
- Hipótesis alternativa (H_a): existe una diferencia en la presión arterial sistólica antes y después del programa.

Tabla 21. Datos recolectados de pacientes con hipertensión arterial

Paciente	Antes del programa	Después del programa	Diferencia (antes – después)
1	150	140	10
2	145	135	10
3	148	138	10
4	152	142	10
5	149	139	10
6	147	137	10
7	146	136	10
8	151	141	10
9	150	140	10
10	148	138	10
11	147	137	10
12	149	139	10
13	150	140	10
14	151	141	10
15	148	138	10
16	149	139	10
17	147	137	10
18	150	140	10
19	148	138	10
20	146	136	10

Fuente: elaboración propia (2025).

En este caso, todas las diferencias son de 10 mmHg; por lo tanto, $\mu = 10$ y $Sd = 0$. El resultado indica que todos los pacientes mejoraron en la misma magnitud, algo altamente inusual en contextos clínicos reales. Dado que el valor de $p = 0.000$ es menor al nivel de significancia de $p = 0.05$, se rechaza la hipótesis nula, que establece que no hay diferencia en la presión arterial sistólica antes y después del programa de ejercicio. Esto sugiere que existe una diferencia estadísticamente significativa en la presión arterial de los pacientes hipertensos, lo que implica que el programa de ejercicio aeróbico tiene un efecto positivo en la reducción de la presión arterial sistólica.

8.1.2 Wilcoxon

La prueba de Wilcoxon para rangos con signo (prueba de Wilcoxon) es una técnica no paramétrica utilizada para evaluar si existen diferencias significativas entre dos mediciones relacionadas o apareadas cuando los datos no cumplen con los supuestos de normalidad requeridos por la t de Student para muestras relacionadas. Esta prueba se fundamenta en el análisis de los rangos de las diferencias entre pares de observaciones, ignorando la magnitud exacta y enfocándose en la dirección y la intensidad relativa de dichas diferencias (Conover, 1999; McDonald, 2014).

Características principales

- **Naturaleza no paramétrica:** no exige que la variable dependiente siga una distribución normal, lo que la hace adecuada para datos ordinales o continuos con asimetrías, sesgos o valores extremos.
- **Diseño con datos apareados:** se aplica cuando las mediciones se realizan sobre las mismas unidades de análisis en dos momentos distintos (por ejemplo, antes y después de una intervención) o en dos condiciones diferentes.
- **Sensibilidad a la dirección del cambio:** considera tanto la dirección como la magnitud relativa de los cambios, asignando rangos a las diferencias absolutas y luego sumando los signos positivos y negativos.

Hipótesis estadística

- **Hipótesis nula (H_0):** las distribuciones de las dos mediciones relacionadas son iguales; es decir, no hay diferencias sistemáticas entre las condiciones comparadas.
- **Hipótesis alternativa (H_a):** las distribuciones de las dos mediciones relacionadas son diferentes; hay evidencia de un cambio o un efecto sistemático entre ambas condiciones.

Procedimiento estadístico

- Cálculo de las diferencias: se obtiene la diferencia entre cada par de observaciones.
- Eliminación de ceros: se excluyen los pares cuya diferencia sea cero, ya que no aportan información sobre la dirección del cambio.
- Rango de las diferencias absolutas: se ordenan de menor a mayor y se asignan rangos numéricos.
- Asignación de signo: se conservan los signos originales (positivo o negativo) de las diferencias en los rangos.
- Cálculo de las sumas de rangos: se obtiene la suma de los rangos positivos (W^+) y la de los negativos (W^-).
- Estadístico de prueba (W): el valor de W corresponde al menor de las dos sumas de rangos ($W = \min(W^+, W^-)$).
- Comparación con valores críticos o p-valor: el estadístico W se compara con una tabla crítica específica o se interpreta su valor p para determinar la significancia estadística.

Supuestos

- Las observaciones deben ser independientes dentro de los pares.
- Las diferencias entre pares deben tener una escala al menos ordinal.
- Es adecuado para muestras pequeñas y medianas, aunque puede aplicarse en muestras grandes mediante una aproximación normal (Gibbons y Chakraborti, 2011).

Ventajas

- Flexibilidad ante violaciones de la normalidad: es útil cuando los datos presentan sesgos, asimetrías o valores atípicos.
- Aplicación a datos ordinales: puede emplearse en variables medidas en escalas ordinales, como escalas de dolor, satisfacción o preferencia.
- Mayor robustez: resiste mejor la influencia de valores extremos, evitando distorsiones en los resultados.

Limitaciones

- Menor potencia estadística: cuando los datos cumplen la normalidad, es menos eficiente que la prueba t pareada (Ruxton, 2006).
- No estima la magnitud del efecto: solo indica si existen diferencias significativas, sin cuantificar el tamaño del cambio.

Aplicaciones clínicas y metodológicas

La prueba de Wilcoxon ha demostrado gran utilidad para evaluar el efecto de una intervención en un mismo grupo de pacientes con mediciones repetidas. Su aplicación es pertinente cuando las variables de interés se expresan en escalas ordinales, como las de dolor, satisfacción o calidad de vida, o cuando las diferencias entre las mediciones no se distribuyen con normalidad.

En la práctica médica, esta prueba se emplea, por ejemplo, para analizar la reducción del dolor al aplicar diferentes métodos fisioterapéuticos, valorar los cambios en los niveles de ansiedad antes y después de una intervención psicológica, o comparar las variaciones en las puntuaciones de cuestionarios de calidad de vida en estudios longitudinales. De este modo, la prueba de Wilcoxon permite detectar cambios significativos cuando los supuestos paramétricos no se cumplen.

Ejemplo: un médico investigador plantea determinar si un nuevo tratamiento farmacológico produce una reducción significativa en los niveles séricos de colesterol total en pacientes con diagnóstico de hipercolesterolemia. Se recluta una muestra de 15 pacientes, a quienes se les mide el nivel de colesterol antes de iniciar el tratamiento y nuevamente al completar un régimen terapéutico de 12 semanas. Dado que se trata de datos relacionados (antes vs. después en los mismos individuos) y no se puede asegurar que las diferencias sigan una distribución normal, se opta por aplicar la prueba no paramétrica de Wilcoxon para rangos con signo.

Hipótesis estadísticas

- Hipótesis nula (H_0): no hay diferencia en los niveles de colesterol antes y después del tratamiento (las distribuciones son iguales).
- Hipótesis alternativa (H_a): existen diferencias significativas en los niveles de colesterol antes y después del tratamiento (las distribuciones son distintas).

Datos recolectados (Tabla 22):

Tabla 22. Datos recolectados de pacientes con diagnóstico de hipercolesterolemia

Paciente	Antes	Después
1	265	240
2	278	252
3	250	236
4	270	258
5	260	245

Paciente	Antes	Después
6	275	248
7	255	232
8	262	243
9	268	240
10	280	250
11	272	249
12	258	246
13	261	244
14	276	260
15	269	251

Fuente: elaboración propia (2025).

Aplicación de la prueba de Wilcoxon

- Cálculo de las diferencias entre cada par (antes – después).
- Eliminación de ceros: no hay diferencias iguales a cero.
- Cálculo de los valores absolutos y asignación de rangos.
- Asignación de signos a los rangos según el signo de la diferencia.
- Suma de los rangos positivos (W^+) y negativos (W^-).
- Determinación del estadístico de prueba (W): el menor entre W^+ y W^- .
- Comparación con el valor crítico de la tabla de Wilcoxon para $n = 15$ o cálculo del p-valor correspondiente.

En este contexto, si el valor de W observado es menor que el valor crítico de Wilcoxon para un nivel de significancia $\alpha = 0.05$ (aproximadamente $W \leq 25$ para $n = 15$) o si el p-valor asociado es menor a 0.05, se rechaza la hipótesis nula, concluyendo que el tratamiento tiene un efecto estadísticamente significativo sobre la reducción de los niveles de colesterol.

Dado que el valor de $p = 0.001$, el análisis mediante la prueba de Wilcoxon muestra que el tratamiento farmacológico de 12 semanas produce una reducción significativa en los niveles de colesterol total en los pacientes evaluados, lo que respalda su eficacia clínica en el manejo de la hipercolesterolemia.

8.1.3. McNemar

La prueba de McNemar es una herramienta estadística no paramétrica diseñada específicamente para evaluar cambios en variables categóricas dicotómicas en estudios de diseño pareado o repetido. Su uso es especialmente pertinente cuando se pretende analizar si una intervención ha producido una modificación significativa

en la proporción de sujetos que presentan determinada característica antes y después del tratamiento. Se trata de una de las pruebas más empleadas en el análisis de frecuencias emparejadas, como ocurre en estudios clínicos con evaluaciones “antes-después”, pruebas diagnósticas o intervenciones educativas con resultados binarios (Fagerland et al., 2013).

Características

La prueba de McNemar se centra en los pares discordantes, es decir, aquellos sujetos cuya respuesta cambia de una categoría a otra entre las dos mediciones (por ejemplo, de “no mejoró” a “mejoró”). El procedimiento parte de la premisa de que los pares concordantes (los que permanecen en la misma categoría en ambas mediciones) no aportan información sobre cambios atribuibles a la intervención, por lo que no contribuyen al contraste estadístico.

Hipótesis estadísticas

- Hipótesis nula (H_0): las proporciones de respuesta en ambas condiciones (antes y después) son estadísticamente iguales.
- Hipótesis alternativa (H_a): las proporciones difieren significativamente entre ambas condiciones.

Procedimiento

La prueba se basa en una tabla de contingencia 2 x 2 construida a partir de los datos pareados. Esta tabla contiene las frecuencias siguientes (Tabla 23):

Tabla 23. Tabla de contingencia de datos pareados

	Después = sí	Después = no
Antes = sí	a	b
Antes = no	c	d

Fuente: elaboración propia (2025).

- a: Casos que fueron positivos antes y después (concordancia positiva)
- b: Casos que fueron positivos antes y negativos después (discordancia)
- c: Casos que fueron negativos antes y positivos después (discordancia)
- d: Casos que fueron negativos antes y después (concordancia negativa)

La prueba de McNemar se enfoca en los valores b y c, es decir, en los sujetos que cambiaron de categoría. El estadístico de prueba se calcula como (Figura 54):

Figura 54

Fórmula estadística de prueba de McNemar

$$\chi^2 = \frac{(b - c)^2}{b + c}$$

Fuente: elaboración propia (2025).

Este valor se compara con una distribución chi-cuadrada con un grado de libertad. En contextos de pequeños tamaños muestrales (por ejemplo, si $b + c < 25$), se recomienda aplicar la corrección de continuidad de Yates o utilizar la prueba exacta de McNemar.

Supuestos

- Las observaciones deben ser pares dependientes (es decir, provienen del mismo sujeto en dos momentos o condiciones).
- La variable debe ser dicotómica, codificada típicamente como 0/1, sí/no, presente/ausente.
- La muestra debe estar compuesta por una cantidad suficiente de pares discordantes para garantizar una inferencia robusta.

Ventajas

- Se trata de una prueba simétrica, sencilla de aplicar e interpretar.
- No requiere normalidad en la distribución de los datos ni homogeneidad de varianzas.
- Ideal para el análisis de cambios cualitativos en ensayos clínicos, intervenciones educativas o validación de pruebas diagnósticas.

Limitaciones

- No es aplicable a variables politómicas ni continuas.
- Su utilidad disminuye cuando los valores de b y c son muy pequeños, lo que puede afectar la potencia estadística.
- No ofrece información sobre la magnitud del cambio, solo sobre su significancia estadística.

La prueba de McNemar se aplica en estudios de efectividad de tratamientos donde se evalúa un resultado binario como la presencia o ausencia de una condición clínica; también se utiliza para comparar la concordancia entre dos métodos diagnósticos, o para analizar cambios en actitudes, percepciones o conocimientos tras una intervención educativa, constituye una herramienta estadística robusta para el análisis de proporciones dicotómicas emparejadas, siendo fundamental en investigaciones clínicas que involucran diseños longitudinales o estudios de cambio. Su correcta aplicación permite establecer, con base en la evidencia, si una intervención ha logrado modificar el estado de los individuos en una variable categórica.

Ejemplo: un investigador se propone determinar si una campaña de concienciación sobre la seguridad vial logra modificar significativamente el comportamiento de los conductores respecto al uso del cinturón de seguridad. Para ello, se realiza un estudio observacional con diseño antes-después, en el que se registran las respuestas de 100 conductores antes y después de la intervención.

Planteamiento

- Variable de estudio: uso del cinturón de seguridad (sí/no).
- Diseño: medición emparejada en dos momentos (pre y poscampaña).
- Objetivo: evaluar si existe un cambio estadísticamente significativo en la proporción de conductores que usan el cinturón tras la intervención educativa.

Datos recolectados

Se construye una tabla de contingencia 2 x 2 con base en los comportamientos antes y después de la campaña (Tabla 24):

Tabla 24. Datos recolectados de campaña de concienciación sobre seguridad vial

	Después = Sí	Después = No	Total
Antes = Sí	a = 30	b = 10	40
Antes = No	c = 50	d = 10	60
Total	80	20	100

Fuente: elaboración propia (2025).

Interpretación de la tabla

- a (concordancia positiva): 30 conductores usaban el cinturón antes y después.
- b (discordancia negativa): 10 conductores usaban el cinturón antes, pero dejaron de usarlo después.
- c (discordancia positiva): 50 conductores no usaban el cinturón antes, pero comenzaron a usarlo después.
- d (concordancia negativa): 10 conductores no usaban el cinturón ni antes ni después.

Aplicación de la prueba de McNemar

La prueba se enfoca en los valores discordantes (b y c; Figura 55):

Figura 55

Fórmula para aplicación de la prueba de McNemar

$$\chi^2 = \frac{(b - c)^2}{b + c} = \frac{(10 - 50)^2}{10 + 50} = \frac{1600}{60} \approx 26.67$$

Fuente: elaboración propia (2025).

El valor crítico para χ^2 con 1 grado de libertad al nivel de significancia $\alpha = 0.05$ es 3.841; dado que $26.67 > 3.841$, el resultado es estadísticamente significativo. Además, el valor exacto de $p = 0.000$, lo que permite rechazar la hipótesis nula de igualdad de proporciones.

Interpretación

Los resultados muestran que la campaña de concienciación produjo un cambio significativo en el comportamiento de los conductores respecto al uso del cinturón de seguridad; el número de personas que comenzaron a usar el cinturón después de la intervención (50 casos) fue mayor que el de las que dejaron de usarlo (10 casos), lo que se refleja en una mejora significativa. Por tanto, la campaña logró promover el uso del cinturón para evitar riesgos en la vía pública.

8.2. Tres o más muestras

La comparación de tres o más muestras relacionadas es esencial para evaluar cambios longitudinales en variables dependientes. Esta situación se presenta con frecuencia en diseños experimentales u observacionales en los que los mismos sujetos son evaluados en tres o más momentos en el tiempo, o bien bajo diferentes condiciones o tratamientos. El objetivo central de este análisis es determinar si las diferencias observadas entre las condiciones pueden atribuirse al efecto de las intervenciones, tratamientos o contextos evaluados, más allá de la variabilidad aleatoria (Field, 2018).

Desde el punto de vista técnico, el análisis de múltiples mediciones intrasujeto requiere técnicas estadísticas que consideren la dependencia entre las observaciones. A diferencia de los análisis entre grupos independientes, los estudios con muestras relacionadas implican observaciones no independientes, lo que invalida el uso de pruebas que asumen independencia entre mediciones (Girden, 1992).

En estudios que cumplen con los supuestos de normalidad —es decir, cuando las diferencias entre las condiciones presentan una distribución normal y se verifica esfericidad (una forma específica de homogeneidad de varianzas para datos repetidos)—, el ANOVA de medidas repetidas constituye la herramienta paramétrica de referencia (Keselman et al., 1998). No obstante, en la práctica clínica y social, estos supuestos no siempre se cumplen, particularmente cuando se trabaja con escalas ordinales, datos dicotómicos, tamaños muestrales pequeños, presencia de valores atípicos o distribuciones marcadamente asimétricas.

Ante estas limitaciones, las pruebas no paramétricas surgen como alternativas robustas y metodológicamente sólidas para comparar tres o más mediciones relacionadas. Entre las más empleadas destacan:

- Prueba de Friedman, que permite comparar k tratamientos o condiciones en datos de tipo ordinal o continuo no normales en los mismos sujetos (Gibbons y Chakraborti, 2011).
- Prueba Q de Cochran, utilizada cuando la variable de interés es dicotómica y se busca evaluar la proporción de respuestas (por ejemplo, éxito/fracaso) en varias condiciones dentro del mismo grupo (Conover, 1999).

Ambas pruebas se fundamentan en el análisis de rangos o frecuencias, en lugar de medias o varianzas, lo que las hace menos sensibles a supuestos estrictos y más adecuadas para el tratamiento de datos clínicos, educativos y sociales donde las condiciones ideales de normalidad y homogeneidad rara vez se cumplen. En suma, la comparación de tres o más muestras mediante métodos paramétricos o no paramétricos constituye una estrategia indispensable para el análisis riguroso

de datos dependientes, incluso en condiciones de variabilidad, heterogeneidad y limitaciones muestrales que caracterizan los estudios aplicados en salud y ciencias humanas.

8.2.1. Friedman

La prueba de Friedman es una técnica estadística no paramétrica diseñada para evaluar si existen diferencias significativas entre tres o más tratamientos, condiciones o momentos de medición aplicados a un mismo conjunto de sujetos o unidades experimentales. Se le reconoce como una extensión de la prueba de Wilcoxon para muestras relacionadas. Esta prueba es adecuada cuando los datos no cumplen los supuestos de normalidad exigidos por el ANOVA de medidas repetidas, pero se desea mantener un diseño intrasujeto que controle la variabilidad interindividual (Gibbons y Chakraborti, 2011).

En estudios clínicos, psicológicos, educativos y organizacionales en los que las mediciones se repiten en los mismos individuos bajo diferentes condiciones, la prueba de Friedman permite detectar diferencias significativas entre condiciones sin que la asimetría, la curtosis o los valores extremos comprometan la validez del análisis (Conover, 1999).

Hipótesis estadísticas

Como en todo análisis estadístico de contrastes de hipótesis, la prueba de Friedman se basa en la siguiente formulación:

- Hipótesis nula (H_0): no existen diferencias estadísticamente significativas entre las medianas de las condiciones evaluadas; es decir, las distribuciones de las variables dependientes son homogéneas.
- Hipótesis alternativa (H_a): al menos una de las condiciones presenta diferencias estadísticamente significativas respecto de las demás, lo que sugiere un efecto del tratamiento o de la condición experimental.

Este planteamiento permite investigar si las variaciones observadas pueden atribuirse al efecto de los tratamientos o se explican por la variabilidad aleatoria esperada bajo la hipótesis de igualdad.

Supuestos

- Las observaciones deben ser dependientes dentro de cada sujeto y pares independientes entre sí.
- La variable dependiente debe ser ordinal o continua no normal.

- Cada sujeto debe tener mediciones en todas las condiciones evaluadas (diseño balanceado).

Ventajas

- No requiere supuestos de normalidad o de homogeneidad de varianzas.
- Controla la variabilidad interindividual al tratarse de un diseño intrasujeto.
- Es robusta frente a valores atípicos o distribuciones sesgadas.

Limitaciones

- No permite identificar directamente qué condiciones difieren entre sí; en caso de significancia, se requieren pruebas post hoc (por ejemplo, Nemenyi o Dunn–Bonferroni).
- Presenta menor potencia estadística que el ANOVA de medidas repetidas cuando se cumplen los supuestos paramétricos.
- Su interpretación se limita a diferencias en rangos medianos, sin estimar directamente el tamaño del efecto.

Procedimiento metodológico

El proceso de aplicación de la prueba de Friedman implica los siguientes pasos operativos:

1. Estructuración del conjunto de datos: se construye una matriz de datos en la cual las filas representan a los sujetos o unidades experimentales, mientras que las columnas corresponden a las condiciones, tratamientos o momentos de medición. Este diseño asegura que cada sujeto sea su propio control, lo cual incrementa la sensibilidad del análisis.
2. Asignación de rangos intrasujeto: en cada fila, es decir, por sujeto, se ordenan los valores correspondientes a las diferentes condiciones y se asignan rangos ascendentes del menor al mayor valor; en caso de empates, se asigna el promedio de los rangos que ocuparían los valores empatados. Esta transformación a rangos elimina la influencia de la magnitud absoluta y se enfoca en la posición relativa de las mediciones.
3. Suma de rangos por condición ($\sum R_j$): una vez asignados los rangos dentro de cada sujeto, se suman los rangos correspondientes a cada columna (condición). Estas sumas de rangos permiten cuantificar la posición relativa acumulada de cada tratamiento.
4. Cálculo del estadístico de Friedman: el estadístico de prueba de Friedman se calcula mediante la siguiente fórmula (Figura 56):

Figura 56

Fórmula para cálculo de estadístico de Friedman

$$X_F^2 = \frac{12}{nk(k+1)} \sum R_j^2 - 3n(k+1)$$

Fuente: elaboración propia (2025).

Donde:

n es el número de sujetos

k es el número de condiciones o tratamientos

R_j es la suma de rangos para la j-ésima condición

Este estadístico sigue una distribución aproximada chi-cuadrada con k-1 grados de libertad bajo la hipótesis nula cuando el tamaño de muestra es mayor a 10.

5. Toma de decisión estadística: se compara el valor calculado del estadístico con el valor crítico de la distribución chi-cuadrada para un nivel de significancia establecido (por ejemplo, $\alpha = 0.05$). Alternativamente, se puede calcular el valor p asociado. Si $p < \alpha$, se rechaza la hipótesis nula y se concluye que existen diferencias significativas entre las condiciones evaluadas.

El principal aporte de la prueba de Friedman radica en su capacidad para controlar la variabilidad intersujeto sin asumir normalidad ni igualdad de varianzas, características que la convierten en una técnica robusta y versátil. Además, al basarse en rangos, es menos sensible a valores atípicos y a escalas no métricas, como las ordinales.

No obstante, una limitación importante es que, en caso de detectar una diferencia significativa global, la prueba de Friedman no especifica cuáles condiciones difieren entre sí. Para ello, es necesario realizar comparaciones múltiples post hoc, como las pruebas de Nemenyi o Wilcoxon con ajuste de Bonferroni, que permitan identificar los pares responsables de las diferencias.

Aplicaciones clínicas y metodológicas

En medicina y ciencias del comportamiento, la prueba de Friedman se emplea frecuentemente en situaciones como:

- Evaluar el efecto de un fármaco en diferentes momentos (pretratamiento, postratamiento inmediato y seguimiento).
- Comparar el rendimiento cognitivo de un grupo bajo tres tipos distintos de estímulo.
- Medir la satisfacción de pacientes en diferentes fases del proceso clínico.

La prueba de Friedman constituye un recurso estadístico imprescindible en estudios longitudinales o de diseño cruzado con múltiples condiciones relacionadas, especialmente cuando los supuestos paramétricos no se cumplen. Su adecuada aplicación favorece la interpretación rigurosa de resultados y fortalece la validez de las inferencias en entornos clínicos, educativos y sociales.

Ejemplo: con el objetivo de evaluar la eficacia relativa de tres tratamientos farmacológicos en la reducción de la presión arterial sistólica en pacientes con hipertensión arterial sistémica, un equipo de investigadores diseñó un estudio de tipo intrasujeto, en este se seleccionaron 12 pacientes que recibieron, en distintos periodos y bajo condiciones controladas, tres medicamentos diferentes identificados como medicamento A, B y C. Cada paciente recibió los tres tratamientos en un orden aleatorio; posteriormente se midió su presión arterial sistólica después de cada intervención.

Este diseño permite controlar la variabilidad interindividual, ya que cada paciente actúa como su propio control. Para el análisis estadístico de las diferencias observadas entre los tratamientos, se utilizó la prueba no paramétrica de Friedman debido a que no se asumió la normalidad de los datos y se trataba de comparar tres condiciones relacionadas.

Datos recolectados

La siguiente tabla presenta los valores de presión arterial sistólica (en mmHg) medidos después de la administración de cada uno de los medicamentos a los 12 pacientes (Tabla 25).

Tabla 25. Datos recolectados de pacientes con hipertensión arterial

Paciente	Medicamento A	Medicamento B	Medicamento C
1	140	135	130
2	150	148	145
3	138	136	132
4	145	140	138

Paciente	Medicamento A	Medicamento B	Medicamento C
5	155	150	147
6	142	140	136
7	148	144	142
8	135	132	128
9	152	150	146
10	147	143	140
11	149	145	143
12	153	149	146

Fuente: elaboración propia (2025).

Hipótesis estadísticas

- Hipótesis nula (H_0): no existen diferencias estadísticamente significativas en la presión arterial sistólica entre los tres medicamentos.
- Hipótesis alternativa (H_a): al menos uno de los medicamentos produce un efecto significativamente diferente en la reducción de la presión arterial sistólica.

Aplicación del procedimiento de Friedman

- Asignación de rangos por paciente: a cada paciente se les asignaron rangos a las mediciones de cada tratamiento (1 para la presión más baja, 3 para la más alta). Esta transformación se realizó fila por fila, lo que permitió controlar la variabilidad interindividual.
- Suma de rangos por tratamiento: se sumaron los rangos asignados a cada tratamiento. Este paso es esencial para calcular el estadístico de Friedman.
- Cálculo del estadístico de Friedman ($\chi^2 F$): utilizando la fórmula correspondiente y considerando 12 sujetos y 3 tratamientos, se obtuvo el valor del estadístico de Friedman, que fue comparado con la distribución chi-cuadrada con $k - 1 = 2$ grados de libertad.
- Resultado estadístico: el análisis arrojó un valor de $p = 0.000$, que es significativamente menor al nivel de significancia habitual de $\alpha = 0.05$.

Interpretación

Dado que el valor p fue inferior al umbral de significancia ($p < 0.05$), se rechaza la hipótesis nula; por tanto, se concluye que existen diferencias estadísticamente significativas en la presión arterial sistólica entre al menos uno de los tratamientos administrados. En otras palabras, al menos uno de los medicamentos evaluados muestra una eficacia significativamente distinta en la reducción de la presión

arterial en comparación con los otros dos. Para identificar específicamente qué medicamentos difieren entre sí en su efecto antihipertensivo, se recomienda realizar pruebas post hoc, como la de Nemenyi, con ajustes adecuados para comparaciones múltiples.

Este análisis respalda la utilidad del diseño dentro de sujetos y de la prueba de Friedman para evaluar la eficacia comparativa de tratamientos médicos en condiciones donde la distribución de los datos no garantiza el uso de pruebas paramétricas. La evidencia obtenida constituye una base sólida para futuras decisiones terapéuticas, orientadas a seleccionar el tratamiento más eficaz en la práctica clínica basada en evidencia.

8.2.2. Q de Cochran

El estadístico Q de Cochran constituye una prueba no paramétrica diseñada para analizar datos dicotómicos en contextos de medidas repetidas. Se aplica cuando un mismo grupo de sujetos es evaluado bajo tres o más condiciones distintas y la variable de interés adopta únicamente dos categorías (por ejemplo, éxito/fracaso, presencia/ausencia, sí/no). Esta prueba es una extensión de la prueba de McNemar para contrastar proporciones emparejadas entre diferentes tratamientos o momentos de medición (McDonald, 2014).

El Q de Cochran busca comparar la eficacia de diversas intervenciones sobre una variable dicotómica dependiente en un mismo conjunto de unidades experimentales. Gracias a su diseño, la prueba controla la variabilidad intrasujeto y permite identificar diferencias estadísticamente significativas en la frecuencia de respuestas positivas (o negativas) atribuibles a distintas condiciones (Conover, 1999).

Supuestos

Para la aplicación adecuada de la prueba Q de Cochran deben cumplirse los siguientes criterios metodológicos:

- La variable dependiente debe ser dicotómica.
- Las observaciones deben ser emparejadas, es decir, cada sujeto debe haber sido evaluado en todas las condiciones.
- El número de sujetos debe ser constante en todas las condiciones.

El objetivo de esta prueba es contrastar la hipótesis nula que establece que las proporciones de éxito (o de respuestas positivas) son iguales en todas las condiciones, frente a una hipótesis alternativa que plantea que al menos una condición presenta una proporción significativamente distinta.

Hipótesis estadísticas

- Hipótesis nula (H_0): las proporciones de éxito son iguales en todas las condiciones evaluadas.
- Hipótesis alternativa (H_a): al menos una condición presenta una proporción de éxito diferente a las demás.

Procedimiento metodológico

El procedimiento para llevar a cabo la prueba Q de Cochran consta de los siguientes pasos estructurados:

1. Organización de los datos: se construye una matriz en la que cada fila representa a un sujeto y cada columna a una condición. En cada celda se registra una respuesta dicotómica (por ejemplo, 1 = éxito, 0 = fracaso) obtenida bajo esa condición específica.
2. Cálculo de totales marginales: se computan las sumas por filas (número total de éxitos por sujeto) y por columnas (número de éxitos por condición), lo que permite cuantificar la variabilidad entre sujetos y entre condiciones.
3. Cálculo del estadístico Q: se utiliza la siguiente fórmula (Figura 57).

Figura 57

Cálculo del estadístico Q de Cochran

$$Q = \frac{(k - 1) \left[k \sum_{j=1}^k C_j^2 - T^2 \right]}{k \cdot T - \sum_{i=1}^n R_i^2}$$

Fuente: elaboración propia (2025).

Donde:

k = número de condiciones

n = número de sujetos

C_j^2 = total de respuestas positivas en la condición j

R_i^2 = total de respuestas positivas del sujeto i

T = total general de respuestas positivas

4. Prueba de significancia: el valor calculado del estadístico Q se compara con la distribución chi-cuadrado con k - 1 grados de libertad. Si el valor p asociado es menor al nivel de significancia previamente establecido (habitualmente $\alpha = 0.05$), se rechaza la hipótesis nula.

Ventajas

- Aplicabilidad a datos dicotómicos: es una de las pocas pruebas diseñadas específicamente para datos binarios en diseños de medidas repetidas.
- Control de la variabilidad intrasujeto: al utilizar un diseño emparejado, se mejora la sensibilidad del análisis frente a métodos para muestras independientes.
- No requiere normalidad: ideal para contextos clínicos o sociales donde las respuestas no son continuas ni siguen una distribución paramétrica.

Limitaciones

- Restricción a variables dicotómicas: no puede utilizarse con variables categóricas, politómicas ni ordinales.
- Sensibilidad a estructuras incompletas: requiere que todos los sujetos tengan observaciones completas en todas las condiciones.
- Necesidad de pruebas post hoc: si se detecta una diferencia global significativa, es necesario realizar comparaciones múltiples adicionales (por ejemplo, pruebas de McNemar por pares con corrección de Bonferroni) para identificar específicamente dónde radica la diferencia.

Aplicaciones clínicas y metodológicas

La prueba Q de Cochran es útil para evaluar la eficacia de múltiples intervenciones binarias sobre los mismos individuos, por ejemplo, determinar si hay diferencias en la respuesta inmunológica a diferentes vacunas, en el cumplimiento terapéutico frente a diversas modalidades de seguimiento, o en la precisión de diagnóstico de un mismo grupo de radiólogos evaluando tres técnicas de imagen distintas. La prueba Q de Cochran ofrece una solución estadística robusta para el análisis de proporciones en medidas repetidas con datos dicotómicos, en contextos en los que las alternativas paramétricas no son adecuadas ni válidas.

Ejemplo: un equipo de investigación en psicología clínica desea evaluar la eficacia comparativa de tres enfoques terapéuticos en la reducción de síntomas de ansiedad en pacientes con diagnóstico confirmado de trastorno de ansiedad generalizada (TAG). Para ello, se diseña un estudio con medidas repetidas en el cual un mismo grupo de pacientes participa, de forma controlada y secuenciada, en tres tipos distintos de intervención psicológica: terapia cognitivo-conductual (TCC), terapia de aceptación y compromiso (ACT) y terapia psicodinámica (TP).

Cada paciente es sometido a las tres intervenciones en distintos momentos, siguiendo un esquema de diseño cruzado (crossover), y al término de cada fase se evalúa si presenta mejoría clínica en los niveles de ansiedad, según un criterio

dicotómico (mejoría: 1; sin mejoría: 0) basado en un umbral clínicamente validado en escalas estandarizadas.

Datos recolectados

Los datos recolectados se organizan en la siguiente matriz binaria (Tabla 26):

Tabla 26. Datos recolectados de pacientes con trastorno de ansiedad generalizada

Paciente	Terapia cognitivo-conductual (TCC)	Terapia de aceptación y compromiso (ACT)	Terapia psicodinámica (TP)
1	1	0	1
2	1	1	1
3	0	0	1
4	1	1	0
5	1	1	1
6	0	1	1
7	1	0	0
8	0	1	1
9	1	1	1
10	1	0	1

Fuente: elaboración propia (2025).

Hipótesis planteadas

- Hipótesis nula (H_0): las proporciones de pacientes que presentan mejoría son iguales en las tres intervenciones.
- Hipótesis alternativa (H_a): al menos una de las terapias difiere significativamente en su proporción de eficacia.

Aplicación de la prueba Q de Cochran

Se calcula el número de pacientes que experimentaron mejoría tras cada tipo de intervención: TCC: 7 pacientes, ACT: 6 pacientes y TP: 8 pacientes. Asimismo, se identifican los patrones de respuesta binaria por paciente y se aplica el estadístico Q de Cochran para determinar si existen diferencias estadísticamente significativas entre las proporciones de éxito observadas en cada terapia. En esta prueba, todos los participantes fueron evaluados con los tres tratamientos (diseño completo) y los datos son dicotómicos y emparejados. El análisis se realiza en software estadístico y el resultado arrojado para el valor de p asociado al estadístico Q es $p = 0.210$.

Interpretación

Dado que el valor p es mayor que el nivel de significancia habitual ($\alpha = 0.05$), no se rechaza la hipótesis nula; por tanto, no se encontraron diferencias estadísticamente significativas en la efectividad de las tres intervenciones psicológicas para la reducción de ansiedad en los pacientes evaluados. Este análisis sugiere que, en esta muestra específica de pacientes con trastorno de ansiedad generalizada, las tres terapias psicológicas presentan niveles de eficacia comparables en términos de proporción de individuos que reportan mejoría clínica tras la intervención. Aunque no se observan diferencias estadísticamente significativas, esta conclusión debe considerarse en el contexto del tamaño de muestra limitado y de los posibles efectos cruzados entre intervenciones. Investigaciones con mayor poder estadístico podrían explorar más a fondo los efectos diferenciales entre estos enfoques terapéuticos.

Cierre de la unidad

Esta unidad proporcionó las herramientas estadísticas esenciales para analizar datos dependientes obtenidos mediante mediciones repetidas en los mismos sujetos. Se revisaron pruebas paramétricas y no paramétricas t de Student, Wilcoxon, McNemar, Friedman y Q de Cochran, adecuadas para distintos tipos de variables y diseños. El aprendizaje adquirido permite aplicar e interpretar estas pruebas con rigurosidad metodológica, fortaleciendo la capacidad de tomar decisiones informadas.

Unidad



Correlación y regresión

Presentación de la unidad

Esta unidad aborda el análisis de relaciones entre variables mediante técnicas de correlación bivariada y modelos de regresión, tanto lineal como logística. Estas herramientas permiten identificar asociaciones, construir modelos predictivos y tomar decisiones fundamentadas en el análisis estadístico, siendo aplicables en múltiples disciplinas científicas y profesionales.

Objetivos

A través del estudio sistemático de esta unidad, se pretende que el estudiante sea capaz de:

- Evaluar la asociación entre dos variables mediante coeficientes de correlación bivariada, interpretando su fuerza y su dirección.
- Construir, estimar e interpretar modelos de regresión lineal, comprendiendo sus supuestos, limitaciones y aplicaciones en contextos reales.
- Aplicar modelos de regresión logística para el análisis de variables dicotómicas o multinomiales, interpretando los coeficientes en términos de probabilidades u *odds ratios* y evaluando la calidad del ajuste del modelo.

Competencia específica

Al integrar los contenidos conceptuales, procedimentales, actitudinales y contextuales de esta unidad, el lector desarrollará la competencia para:

Analizar relaciones entre variables utilizando técnicas de correlación y regresión, seleccionando y aplicando modelos lineales y logísticos adecuados para interpretar y predecir comportamientos en diversos contextos. Esta competencia incluye la capacidad de fundamentar decisiones estadísticas en las características de los datos, así como de evaluar la pertinencia de los modelos según los objetivos del estudio.

9.1. Correlación

El análisis estadístico de correlaciones permite no solo detectar asociaciones, sino también caracterizar su fuerza y dirección. Este tipo de análisis se utiliza ampliamente en disciplinas como la medicina, la economía, la psicología, la biología, la ingeniería y las ciencias sociales, donde la identificación de patrones entre variables resulta esencial para interpretar fenómenos complejos y generar conocimiento empírico (Mukaka, 2012).

En términos formales, la correlación cuantifica la intensidad y la dirección de una relación lineal entre dos variables mediante el coeficiente de correlación. Este puede evidenciar una asociación positiva (ambas variables aumentan o disminuyen en conjunto), negativa (una variable aumenta mientras la otra disminuye) o nula (ausencia de asociación lineal). El coeficiente de correlación de Pearson es el más utilizado para variables de nivel de intervalo o de razón con distribución normal, mientras que el coeficiente de Spearman se emplea en variables ordinales o en aquellas que no cumplen los supuestos de normalidad (Schober et al., 2018).

La correlación no implica causalidad. Una asociación estadística significativa entre dos variables no necesariamente indica que una cause a la otra; ambas podrían ser influenciadas por un tercer factor no observado o ser producto del azar. Por esta razón, la interpretación de los coeficientes de correlación debe realizarse con cautela, considerando el diseño del estudio, los supuestos estadísticos y el marco teórico en el que se inscribe la investigación (Murdock, 2022).

Por su parte, el análisis de regresión complementa y amplía el estudio de la correlación al modelar matemáticamente la relación entre variables. La regresión lineal simple permite estimar cómo cambia una variable dependiente en función de una variable independiente, mientras que la regresión múltiple incorpora varios predictores para evaluar su impacto relativo sobre la variable de interés (Montgomery et al., 2021). Estos modelos generan predicciones, identifican relaciones funcionales y evalúan factores que influyen de manera significativa en un fenómeno.

Cuando la variable dependiente es categórica (por ejemplo, éxito/fracaso o presencia/ausencia de una enfermedad), se utilizan modelos de regresión logística, que permiten estimar la probabilidad de ocurrencia de un evento en función de los predictores. Esta técnica es especialmente útil en la investigación clínica, epidemiológica y social, donde los desenlaces suelen expresarse en términos binarios o multinomiales (Hosmer et al., 2013). En conjunto, los análisis de correlación y regresión ofrecen herramientas para comprender, modelar y predecir fenómenos complejos a partir de datos empíricos. Su dominio técnico y conceptual resulta indispensable para la investigación científica rigurosa.

9.2. Correlaciones bivariadas

El análisis de correlaciones bivariadas evalúa la relación entre dos variables cuantitativas. Su propósito es estimar, describir y cuantificar el grado de asociación lineal entre dichas variables, así como determinar si la asociación observada es estadísticamente significativa. La correlación no debe interpretarse como causalidad, ya que el hecho de que dos variables estén asociadas no implica necesariamente que una sea causa de la otra. La correlación identifica patrones de covariación, no mecanismos causales (Murdock, 2022). La interpretación de la correlación se basa en la dirección y la magnitud de la asociación:

- Correlación positiva: se presenta cuando un incremento en una variable se asocia con un incremento en la otra, o bien cuando ambas disminuyen simultáneamente. Un ejemplo clásico se encuentra en la relación entre el tiempo dedicado al estudio y el rendimiento académico: a mayor tiempo de estudio, suele corresponder un mejor desempeño.
- Correlación negativa: se observa cuando un aumento en una variable se acompaña de una disminución en la otra. Un caso empírico frecuente es la relación inversa entre el consumo de tabaco y la capacidad pulmonar: a medida que aumenta el número de cigarrillos consumidos, la función pulmonar tiende a disminuir.
- Correlación nula: indica la ausencia de una relación lineal significativa. En estos casos, los cambios en una variable no permiten predecir de manera consistente el comportamiento de la otra.

Coefficientes de correlación

Para cuantificar el grado de asociación entre dos variables se emplean distintos coeficientes estadísticos, cuya elección depende de la naturaleza de los datos y de los supuestos estadísticos que puedan verificarse:

- Coeficiente de correlación de Pearson (r): es el más utilizado cuando ambas variables son continuas, cuantitativas y distribuidas normalmente. Su valor oscila entre -1 y 1 ; valores cercanos a ± 1 indican una asociación lineal fuerte, mientras que valores próximos a 0 sugieren una asociación débil o inexistente. Este coeficiente requiere como supuestos la linealidad y la homocedasticidad (Schober et al., 2018).
- Coeficiente de correlación de Spearman (ρ): se emplea con datos no normales o con variables ordinales. Transforma los datos en rangos y mide la asociación entre ellos (útil en presencia de asimetría o valores atípicos) (Mukaka, 2012).

- Coeficiente de correlación de Kendall (τ): Constituye una medida alternativa para variables ordinales o para muestras pequeñas. Evalúa la concordancia entre pares de observaciones; es más conservadora y robusta ante la presencia de empates (Gibbons y Chakraborti, 2011).

Aplicaciones

- La utilidad del análisis de correlaciones bivariadas trasciende el ámbito teórico de la estadística y se ha consolidado como una técnica esencial en diversas disciplinas:
- Medicina: para examinar la relación entre variables biomédicas, como el nivel de colesterol y el riesgo de enfermedad cardiovascular.
- Psicología: para estudiar la asociación entre estados emocionales (como ansiedad o depresión) y el rendimiento cognitivo.
- Economía: para investigar la relación entre indicadores macroeconómicos, como el producto interno bruto y la tasa de empleo.
- Educación: para explorar el vínculo entre la asistencia escolar y el desempeño académico de los estudiantes.

En conjunto, el análisis de correlación bivariada proporciona una base sólida para la formulación de hipótesis, la generación de líneas de investigación y la construcción de modelos más complejos, como los de regresión, que permiten estudiar con mayor profundidad las relaciones funcionales entre variables.

Ejemplo: un investigador se propone analizar si existe una relación significativa entre la edad de los pacientes y su presión arterial sistólica (expresada en milímetros de mercurio, mmHg). Para tal propósito, se selecciona una muestra de 15 pacientes adultos, en quienes se registran las siguientes variables: edad (en años) y PAS. Los datos recolectados se detallan en la siguiente tabla: Datos recolectados (Tabla 27)

Tabla 27. Datos recolectados sobre edad y presión arterial sistólica

Paciente	Edad (años)	PAS (mmHg)
1	35	120
2	42	125
3	50	130
4	55	138
5	60	142

Paciente	Edad (años)	PAS (mmHg)
6	63	145
7	68	150
8	70	155
9	75	160
10	78	162
11	80	165
12	82	168
13	85	170
14	87	172
15	90	175

Fuente: elaboración propia (2025).

Análisis estadístico

Para evaluar la fuerza y dirección de la relación entre la edad y la PAS, se calcula el coeficiente de correlación de Pearson (r), una medida estadística que cuantifica la asociación lineal entre dos variables cuantitativas. El coeficiente r puede tomar valores entre -1 y $+1$:

- Valores cercanos a $+1$ indican una correlación positiva fuerte.
- Valores cercanos a -1 indican una correlación negativa fuerte.
- Valores próximos a 0 indican una correlación débil o nula.

Asimismo, se realiza una prueba de significancia estadística para determinar si el valor de r es significativamente distinto de cero, mediante el p -valor.

Resultados

- Coeficiente de correlación (r): el valor de $r = 0.95$ indica una correlación positiva muy fuerte entre la edad y la PAS, es decir, a medida que aumenta la edad, tiende a aumentar la presión arterial sistólica.
- p -valor: como el p -valor es menor que 0.001 , el resultado es estadísticamente significativo al nivel de significancia convencional ($\alpha = 0.05$); por tanto, se rechaza la hipótesis nula de que no existe correlación entre las variables. Se concluye que la asociación observada no se debe al azar.

Interpretación

El análisis realizado demuestra la existencia de una relación positiva y significativa entre la edad de los pacientes y su presión arterial sistólica, este hallazgo es consistente con la evidencia médica acumulada, la cual indica que el envejecimiento conlleva

una progresiva rigidez arterial y un incremento en la resistencia vascular periférica, lo que se traduce en aumentos graduales de la presión arterial sistólica.

Desde el punto de vista clínico, estos resultados subrayan la importancia de una vigilancia estrecha de la presión arterial en la población adulta mayor, así como la necesidad de estrategias preventivas personalizadas para mitigar el riesgo cardiovascular asociado al envejecimiento.

9.3. Modelo de regresión lineal simple y múltiple

La regresión lineal constituye una de las herramientas fundamentales del análisis estadístico para el estudio de las relaciones funcionales entre variables cuantitativas. A diferencia del análisis de correlación, cuyo propósito es describir la intensidad y la dirección de una relación, la regresión se centra en modelar dicha relación mediante una ecuación matemática, lo que permite explicar el comportamiento de una variable dependiente (respuesta) en función de una o más variables independientes (predictoras) (Montgomery et al., 2021; Gelman et al., 2020).

9.3.1. Regresión lineal simple

La regresión lineal simple describe la relación entre una variable dependiente (Y) y una única variable independiente (X) mediante una función lineal. Su forma general se expresa como (Figura 58):

Figura 58

Fórmula general regresión lineal simple

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Fuente: elaboración propia (2025).

Donde:

Y: variable dependiente (respuesta).

X: variable independiente (predictora).

β_0 : Intercepto o constante, que representa el valor esperado de Y cuando X=0.

β_1 : Coeficiente de regresión, que indica el cambio esperado en Y por cada unidad de cambio en X.

ε : término de error aleatorio que representa la variabilidad de Y no explicada por X.

Supuestos

El modelo de regresión lineal simple se basa en los siguientes supuestos:

- **Linealidad:** existe una relación lineal entre X y Y .
- **Independencia:** las observaciones son independientes entre sí.
- **Homocedasticidad:** la varianza de los errores es constante a lo largo de los valores de X .
- **Normalidad del error:** los errores (ϵ) se distribuyen normalmente con media cero.

Cuando estos supuestos se cumplen, el modelo ofrece estimaciones eficientes, insesgadas y mínimas en varianza para los coeficientes β_0 y β_1 .

Interpretación de los coeficientes

- **β_0 (intercepto):** representa el valor estimado de Y cuando $X=0$; en algunos contextos puede carecer de interpretación práctica si $X=0$ no tiene sentido.
- **β_1 (pendiente):** indica el cambio promedio en Y por cada unidad adicional de X ; si $\beta_1 > 0$, la relación es positiva; si $\beta_1 < 0$, es negativa.

Medida de ajuste del modelo

El grado en que el modelo explica la variación en la variable dependiente se cuantifica mediante el coeficiente de determinación (R^2), que varía entre 0 y 1; un R^2 cercano a 1 indica que el modelo explica una gran proporción de la variabilidad de Y .

Ejemplo: un equipo de investigación en ciencias de la salud desea analizar la relación entre la altura (en cm) y el peso corporal (en kg) en una muestra de adultos jóvenes, con el fin de construir un modelo predictivo que permita estimar el peso de un individuo a partir de su estatura. Se parte del supuesto de que existe una relación lineal entre ambas variables.

Datos recolectados

Se selecciona aleatoriamente una muestra de 10 individuos y se registran sus valores de altura y peso (Tabla 28):

Tabla 28. Datos recolectados sobre altura y peso

Paciente	Altura X (cm)	Peso Y (kg)
1	160	55
2	165	58
3	170	62
4	175	65
5	180	70
6	185	72
7	190	76
8	195	79
9	200	83
10	205	85

Fuente: elaboración propia (2025).

Formulación del modelo

Se aplica un análisis de regresión lineal simple, asumiendo un modelo del tipo: $Y = \beta_0 + \beta_1 X + \varepsilon$, donde Y representa el peso y X la altura. Mediante métodos de mínimos cuadrados, se estiman los coeficientes: $\hat{Y} = -49.2 + 0.65X$.

Interpretación de los coeficientes

- Intercepto ($\hat{\beta}_0 = -49.2$): representa el valor estimado del peso cuando la altura es 0, aunque carece de una interpretación práctica directa (pues ninguna persona mide 0 cm); es un parámetro matemático necesario para definir la recta.
- Pendiente ($\hat{\beta}_1 = 0.65$): indica que, por cada incremento de 1 cm en la altura, el peso corporal estimado aumenta en promedio 0.65 kg; esto sugiere una relación positiva entre ambas variables.

Medida de ajuste

El coeficiente de determinación $R^2 = 0.985$ indica que el 98.5 % de la variabilidad observada en el peso corporal puede ser explicada por el modelo a partir de la altura del individuo; esto sugiere un ajuste excelente del modelo a los datos.

Interpretación

El modelo construido permite predecir el peso de una persona a partir de su estatura, asumiendo una relación lineal entre las variables. Esta herramienta puede ser de utilidad en contextos clínicos y nutricionales en los que se requiera una estimación

aproximada del peso corporal con base en mediciones antropométricas básicas. No obstante, se recomienda validar el modelo considerando otros factores como el sexo, la edad o la composición corporal, para mejorar su precisión predictiva.

9.3.2. Regresión lineal múltiple

La regresión lineal múltiple constituye una extensión del modelo de regresión lineal simple que incorpora dos o más variables independientes para medir la influencia simultánea de múltiples factores sobre una variable dependiente. Este enfoque permite modelar fenómenos complejos en los que intervienen diversas dimensiones (Kutner et al., 2005). La ecuación del modelo se representa como (Figura 59):

Figura 59

Ecuación del modelo de regresión lineal múltiple

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon$$

Fuente: elaboración propia (2025).

Donde:

Y: Variable dependiente

X_1, X_2, \dots, X_k : Variables independientes

$\beta_0, \beta_1, \dots, \beta_k$: Coeficientes del modelo

ε : Término de error

Ventajas

- Control de variables de confusión: permite ajustar el efecto de una variable considerando la influencia de otras.
- Mayor poder explicativo: aumenta el R² y mejora la capacidad predictiva del modelo.
- Detección de relaciones más complejas: facilita la identificación de interacciones o efectos no observables en modelos simples.

En regresión múltiple se deben vigilar problemas como la colinealidad entre variables predictoras, que puede distorsionar la interpretación de los coeficientes, así como los residuos atípicos o influyentes, que pueden afectar la estabilidad del modelo. Para ello, se emplean diagnósticos estadísticos como el VIF (Variance Inflation Factor), los residuos estandarizados y los gráficos de dispersión parcial.

Ejemplo: en el contexto clínico, la predicción del nivel de glucosa en sangre es fundamental para la prevención y el manejo de la diabetes. Para este fin, se propone construir un modelo de regresión lineal múltiple que permita estimar los niveles de glucosa (variable dependiente Y) a partir de un conjunto de factores relevantes, tales como:

- X_1 : Calidad de la dieta (valor numérico basado en puntuación nutricional estandarizada)
- X_2 : Nivel de actividad física semanal (minutos de ejercicio moderado a vigoroso)
- X_3 : Edad del paciente (años)

Datos recolectados (Tabla 29)

Tabla 29. Datos recolectados sobre calidad de dieta, nivel de actividad física, glucosa y edad

Paciente	Dieta (X_1)	Actividad física (X_2)	Edad (X_3)	Glucosa (Y)
1	85	180	25	92
2	70	120	32	110
3	60	90	40	130
4	88	240	27	89
5	55	60	45	145
6	75	150	34	108
7	65	80	41	125
8	90	200	30	94
9	50	45	50	155
10	80	160	29	100

Fuente: elaboración propia (2025).

Modelo propuesto

Se ajusta el siguiente modelo de regresión lineal múltiple: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$. A través del método de mínimos cuadrados ordinarios (MCO), se estiman los coeficientes: $\hat{Y} = 170.2 - 0.45X_1 - 0.12X_2 + 0.75X_3$

Interpretación de los coeficientes

- **Intercepto ($\beta^{\wedge}_0=170.2$):** representa el nivel estimado de glucosa cuando los valores de dieta, actividad física y edad son cero. Aunque no es interpretable clínicamente, es necesario para definir el modelo.
- **Coefficiente de dieta ($\beta^{\wedge}_1=-0.45$):** por cada punto adicional en la calidad de la dieta, se espera una reducción promedio de 0.45 mg/dL en la glucosa, manteniendo constantes las demás variables.
- **Coefficiente de actividad física ($\beta^{\wedge}_2=-0.12$):** indica que un incremento de 10 minutos semanales de ejercicio reduce la glucosa estimada en 1.2 mg/dL, en promedio.
- **Coefficiente de edad ($\beta^{\wedge}_3=+0.75$):** por cada año adicional de edad, el nivel estimado de glucosa aumenta en 0.75 mg/dL, controlando por dieta y ejercicio.

Evaluación del modelo

- **$R^2=0.91$:** el modelo explica el 91 % de la variabilidad total en los niveles de glucosa (excelente ajuste).
- **Valor p global < 0.001:** el modelo completo es estadísticamente significativo.
- **Pruebas de individualidad:** los coeficientes de dieta y edad son significativos a un nivel de $\alpha = 0.05$, mientras que la actividad física se aproxima a la significancia marginal.

Interpretación

El modelo permite estimar con elevada precisión los niveles de glucosa en función de la dieta, la actividad física y la edad; la dirección de los coeficientes concuerda con el conocimiento biomédico actual: una mejor alimentación y un mayor ejercicio físico se asocian con niveles más bajos de glucosa, mientras que la edad incrementa el riesgo glucémico. Estos resultados pueden ser útiles para intervenciones preventivas personalizadas en pacientes con riesgo de disglucemia o de diabetes tipo 2.

Aplicaciones de la regresión lineal

La regresión lineal puede modelar relaciones entre variables y realizar predicciones. A continuación, se ilustran algunas de sus principales aplicaciones:

- **Finanzas:** estimación del precio de acciones, bonos o fondos de inversión en función de variables económicas como las tasas de interés, la inflación o el producto interno bruto.

- **Salud pública:** evaluación de la asociación entre hábitos de vida saludables (por ejemplo, nivel de ejercicio físico) y desenlaces en la salud como la esperanza de vida, la incidencia de enfermedades cardiovasculares o los niveles de presión arterial.
- **Marketing:** medición del efecto de las estrategias de publicidad y promoción sobre el volumen de ventas, la participación de mercado o la fidelidad del consumidor.

Ejemplo: con el objetivo de identificar factores asociados al riesgo de diabetes tipo 2, un grupo de investigadores evalúa la relación entre los niveles de glucosa en sangre en ayuno (mg/dL) y dos variables independientes: la edad (años) y el índice de masa corporal (IMC) (kg/m^2) en una muestra de 15 pacientes con factores de riesgo metabólico.

Datos recolectados

Se recolectaron datos de 15 pacientes con riesgo de diabetes (Tabla 30)

Tabla 30. Datos recolectados de pacientes con riesgo de diabetes

Paciente	Edad (años)	IMC (kg/cm^2)	Glucosa en ayuno (mg/dL)
1	35	24.5	98
2	42	26.3	105
3	50	27.8	110
4	55	29.1	120
5	60	30.5	130
6	63	31.8	135
7	68	33.0	140
8	70	34.2	145
9	75	35.0	155
10	78	36.5	160
11	80	37.1	165
12	82	38.4	170
13	85	39.0	175
14	87	40.2	180
15	90	41.0	185

Fuente: elaboración propia (2025).

Análisis estadístico

1. Regresión lineal simple

Se estima un modelo de regresión lineal simple considerando la edad como único predictor de los niveles de glucosa: $Y = \beta_0 + \beta_1 X_1 + \varepsilon$, donde Y: glucosa en ayuno (mg/dL), X_1 : edad (años).

Resultados del modelo

Coefficiente de regresión (β^1) = 1.65

Intercepto (β^0) = 39.4

Coefficiente de correlación de Pearson (r) = 0.97

$R^2 = 0.94$

Interpretación

Existe una correlación positiva y muy fuerte entre la edad y los niveles de glucosa en ayuno; el modelo sugiere que, por cada año adicional de edad, el nivel de glucosa en ayuno aumenta en promedio 1,65 mg/dL. Este hallazgo está en concordancia con la evidencia científica que asocia el envejecimiento con un deterioro progresivo de la sensibilidad a la insulina.

2. Regresión lineal múltiple

Se ajusta un modelo de regresión múltiple incorporando la edad y el IMC como variables independientes: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$, donde X_1 : edad (años), X_2 : IMC (kg/m^2).

Resultados del modelo

$\beta^0 = 42.3$

$\beta^1 = 1.61$ (edad)

$\beta^2 = 0.10$ (IMC)

$R^2 = 0.98$

Interpretación

Al incluir el IMC, el modelo explica el 98 % de la variabilidad en los valores de glucosa en ayuno. No obstante, en este conjunto de datos, la edad continúa siendo el predictor principal, mientras que el coeficiente asociado al IMC no alcanza significación estadística ($p > 0.05$). Esto sugiere que, aunque existe una relación positiva entre el IMC y la glucosa, su efecto no es independiente del impacto de la edad sobre la glucemia.

Los modelos desarrollados evidencian que la edad tiene una fuerte asociación positiva con los niveles de glucosa en sangre en ayuno; el modelo de regresión múltiple mejora el ajuste general, aunque no añade un valor explicativo significativo respecto del IMC en esta muestra. Estos hallazgos son útiles para el monitoreo clínico y para la identificación temprana de sujetos con riesgo de disglucemia en contextos de salud pública y medicina preventiva.

9.4. Modelo de regresión logística

La regresión logística permite modelar la relación entre una o más variables independientes y una variable dependiente categórica. A diferencia de la regresión lineal, cuyo objetivo es predecir valores continuos, la regresión logística está diseñada para estimar la probabilidad de ocurrencia de un evento categórico, generalmente dicotómico (por ejemplo, presencia/ausencia de una enfermedad, éxito/fracaso). Su fundamento matemático se basa en la función logística, también denominada función sigmoide, que transforma una combinación lineal de predictores en una probabilidad entre 0 y 1.

Modelo matemático

La forma funcional de la regresión logística binaria se expresa como (Figura 60):

Figura 60

Modelo matemático del modelo de regresión logística

$$P(Y = 1) = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_n X_n}}$$

Fuente: elaboración propia (2025).

Donde:

$P(Y=1)$: es la probabilidad de que ocurra el evento de interés (por ejemplo, estar enfermo)

β_0 : es el intercepto del modelo

$\beta_1, \beta_2, \dots, \beta_k$: son los coeficientes de regresión asociados a las variables independientes X_1, X_2, \dots, X_k

e : representa el número de Euler, base de los logaritmos naturales

Al aplicar el *logit* (logaritmo de las probabilidades), se obtiene una transformación lineal que facilita la interpretación de los coeficientes del modelo (Figura 61):

Figura 61

Fórmula de logaritmo de las probabilidades

$$\text{logit}(P) = \log\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

Fuente: elaboración propia (2025).

Tipos de regresión logística

- Regresión logística binaria: aplicable cuando la variable dependiente posee únicamente dos categorías mutuamente excluyentes. Ejemplo: sobrevivió/no sobrevivió, positivo/negativo, éxito/fracaso.
- Regresión logística multinomial: utilizada cuando la variable de resultado tiene más de dos categorías sin un orden jerárquico definido, como en un diagnóstico diferencial entre enfermedades.
- Regresión logística ordinal: se emplea cuando las categorías poseen un orden lógico o secuencial, como los niveles de satisfacción (bajo, medio, alto) o grados de severidad clínica.

Aplicaciones de la regresión logística

Dada su versatilidad, la regresión logística tiene múltiples aplicaciones en contextos donde se requiere tomar decisiones con base en predicciones categóricas:

- Medicina: evaluación del riesgo de padecer enfermedades crónicas como diabetes *mellitus* tipo 2, hipertensión o cáncer, a partir de factores clínicos y demográficos (edad, IMC, antecedentes familiares, entre otros).
- Sociología: análisis de variables que influyen en la probabilidad de que un individuo vote por un partido, participe en procesos electorales o adopte comportamientos sociales.
- Marketing y comportamiento del consumidor: predicción de la probabilidad de compra de un producto en función de variables como el historial de compras, la interacción con campañas publicitarias o el perfil demográfico.

La correlación permite cuantificar la fuerza y dirección de una relación entre dos variables. La regresión extiende esta capacidad al ámbito de la predicción y del modelado matemático. La regresión lineal simple y múltiple son ideales para variables continuas, mientras que la regresión logística es la herramienta idónea cuando la variable dependiente es de tipo categórico. El conocimiento riguroso de estas técnicas no solo permite desarrollar modelos explicativos sólidos.

Ejemplo: un grupo de investigadores clínicos interesados en la prevención de enfermedades metabólicas propone un estudio observacional transversal para evaluar la probabilidad de desarrollar diabetes mellitus tipo 2 en pacientes con sobrepeso u obesidad, considerando dos factores de riesgo bien establecidos: el índice de masa corporal (IMC) y la presión arterial sistólica (PAS). La variable dependiente del estudio es dicotómica: presencia de diabetes (1 = sí, 0 = no).

Datos recolectados

Se recolectaron datos de 20 pacientes adultos, distribuidos según los niveles de IMC y PAS, junto con el diagnóstico confirmado de diabetes tipo 2. A continuación, se resumen los datos del estudio (Tabla 31).

Tabla 31. Datos recolectados sobre IMC y PAS de adultos

Paciente	IMC (kg/cm ²)	PAS (mmHg)	Diabetes (1 = Sí, 0 = No)
1	24.5	120	0
2	26.3	130	0
3	27.8	135	0
4	29.1	140	1
5	30.5	145	0
6	31.8	150	1
7	33.0	155	1
8	34.2	160	1
9	35.0	165	1
10	36.5	170	0
11	37.1	175	1
12	38.4	180	1
13	39.0	185	1
14	40.2	190	1
15	41.0	195	1
16	42.5	200	1
17	43.1	205	1
18	44.8	210	1
19	46.0	215	1
20	47.5	220	1

Fuente: elaboración propia (2025).

Desarrollo del modelo

Se aplicó una regresión logística binaria para modelar la probabilidad de diagnóstico de diabetes en función del IMC y la presión arterial sistólica; la forma del modelo estimado fue (Figura 62):

Figura 62

Fórmula del modelo de regresión logística binaria

$$\log \left(\frac{P}{1 - P} \right) = \beta_0 + \beta_1 \cdot \text{IMC} + \beta_2 \cdot \text{PAS}$$

Fuente: elaboración propia (2025).

Donde:

P es la probabilidad de tener diabetes tipo 2.

β_0 representa el intercepto del modelo.

β_1 y β_2 son los coeficientes asociados a los predictores IMC y PAS, respectivamente.

Interpretación

El modelo global mostró ser estadísticamente significativo ($p < 0.05$), lo que indica que, en conjunto, los predictores incluidos explican una proporción significativa de la varianza en la ocurrencia del evento (diagnóstico de diabetes tipo 2). Sin embargo, ninguno de los coeficientes individuales (IMC ni PAS) resultó estadísticamente significativo de manera aislada ($p > 0.05$); a pesar de que ambos factores están asociados con el riesgo de diabetes, en este conjunto de datos específico no se puede concluir que uno sea un predictor robusto del diagnóstico.

Una posible explicación de este hallazgo es la presencia de colinealidad entre IMC y PAS, dado que ambos factores están fisiológicamente correlacionados y podrían estar captando el mismo componente de riesgo. Asimismo, el tamaño limitado de la muestra ($n=20$) reduce el poder estadístico del análisis y la estabilidad de los coeficientes. Las pruebas de bondad de ajuste, como la prueba de Hosmer-Lemeshow y los indicadores pseudo- R^2 , sugirieron que el modelo tiene un ajuste aceptable con una capacidad razonable de discriminación entre casos positivos y negativos.

Este ejemplo ilustra la utilidad de la regresión logística binaria como herramienta predictiva en contextos clínicos, especialmente cuando se trabaja con variables categóricas. A pesar de que el modelo general fue significativo, los resultados deben interpretarse con cautela debido al tamaño muestral reducido y a la posible multicolinealidad entre los predictores. Se recomienda replicar el análisis con una muestra más amplia, explorar modelos con transformaciones de variables y aplicar análisis de colinealidad para mejorar la precisión y la robustez del modelo.

Cierre de la unidad

Al concluir esta unidad, se habrán incorporado fundamentos conceptuales y metodológicos que permiten analizar relaciones entre variables, construir modelos de predicción y fundamentar decisiones basadas en evidencia estadística. Las técnicas abordadas, correlación, regresión lineal y regresión logística constituyen herramientas analíticas clave en contextos donde se requiere comprender la magnitud y dirección de asociaciones, así como estimar probabilidades o valores esperados en función de variables independientes. Estas competencias son aplicables en diversos campos del conocimiento, como la biomedicina, las ciencias sociales, la psicología, la economía y el análisis de datos en general, fortaleciendo la capacidad investigativa y la toma de decisiones basada en datos.

Referencias

- Abelsson, T., Morténus, H., Karlsson, A. K., Bergman, S. y Baigi, A. (2021). Evidence-based practice in primary healthcare from the managerial point of view—a national survey. *BMC Health Services Research*, 21(1), 1014. <https://doi.org/10.1186/s12913-021-07023-w>
- Adam, D. (2020). Special report: The simulations driving the world's response to COVID-19. *Nature*, 580(7802), 316-319. <https://link.gale.com/apps/doc/A619849970/AONE?u=anon~1d59940e&sid=googleScholar&xid=113c4c16>.
- Adeniran, I. A., Efunniyi, C. P., Osundare, O. S. y Abhulimen, A. O. (2024). Data-driven decision-making in healthcare: Improving patient outcomes through predictive modeling. *Engineering Science & Technology Journal*, 5(8), 59-67. <https://doi.org/10.56781/ijsrms.2024.5.1.0040>
- Agresti, A. (2018). *Statistical Methods for the social sciences* (5^a. ed.). Pearson Education.
- Agresti, A. (2019). *An Introduction to Categorical Data Analysis* (3^a. ed.). Wiley.
- Aguinis, H., Vassar, M. y Wayant, C. (2021). On reporting and interpreting statistical significance and p values in medical research. *BMJ Evidence-Based Medicine*, 26(2), 39-42. <https://doi.org/10.1136/bmjebm-2019-111264>
- Albalwy, F., Brass, A. y Davies, A. (2021). A blockchain-based dynamic consent architecture to support clinical genomic data sharing (ConsentChain): Proof-of-concept study. *JMIR Medical Informatics*, 9(11), e27816. <https://doi.org/10.2196/27816>
- Alhubaiti, A. (2023). Sample size determination: A practical guide for health researchers. *Journal of General and Family Medicine*, 24(2), 72-78. <https://doi.org/10.1002/jgf2.600>
- Amaral, E. D. y Line, S. R. (2021). Current use of effect size or confidence interval analyses in clinical and biomedical research. *Scientometrics*, 126(11), 9133-9145. <https://doi.org/10.1007/s11192-021-04150-3>
- Amrhein, V., Greenland, S. y McShane, B. (2019). Scientists rise up against statistical significance. *Nature*, 567(7748), 305-307. <https://www.nature.com/articles/d41586-019-00857-9>
- Andrade, C. (2019). The P value and statistical significance: misunderstandings, explanations, challenges, and alternatives. *Indian Journal of Psychological Medicine*, 41(3), 210-215. https://doi.org/10.4103/IJPSYM.IJPSYM_193_19
- Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalmic and Physiological Optics*, 34(5), 502-508. <https://doi.org/10.1111/opo.12131>
- Bangdiwala, S. I. (2021). Statistical considerations for research. *International Journal of Injury Control and Safety Promotion*, 28(3), 290-300. <https://doi.org/10.1080/17457300.2021.1957583>

- Berner, D. y Amrhein, V. (2022). Why and how we should join the shift from significance testing to estimation. *Journal of Evolutionary Biology*, 35(6), 777-787. <https://doi.org/10.1111/jeb.14009>
- Bewick, V., Cheek, L. y Ball, J. (2005). Statistics review 14: Logistic regression. *Critical Care*, 9(1), 112. <https://link.springer.com/article/10.1186/CC3045>
- Brozek, J. L., Canelo-Aybar, C., Akl, E. A., Bowen, J. M., Bucher, J., Chiu, W. A. et al. (2021). GRADE Guidelines 30: the GRADE approach to assessing the certainty of modeled evidence—an overview in the context of health decision-making. *Journal of Clinical Epidemiology*, 129, 138-150. <https://doi.org/10.1016/j.jclinepi.2020.09.018>
- Casella, G. y Berger, R. L. (2021). *Statistical inference* (3^a. ed.). CRC Press.
- Chow, S. C. y Liu, J. P. (2021). *Design and Analysis of Clinical Trials* (4^a. ed.). CRC Press.
- Chow, S. C. y Zhang, W. (2020). Statistical evaluation of clinical trials under COVID-19 pandemic. *Therapeutic Innovation & Regulatory Science*, 54(6), 1551-1556. <https://doi.org/10.1007/s43441-020-00182-8>
- Clynes, M. A., Harvey, N. C., Curtis, E. M., Fuggle, N. R., Dennison, E. M. y Cooper, C. (2020). The epidemiology of osteoporosis. *British Medical Bulletin*, 133(1), 105-117. <https://doi.org/10.1093/bmb/ldaa005>
- Collins, F. S. y Varmus, H. (2015). A new initiative on precision medicine. *New England Journal of Medicine*, 372(9), 793-795. <https://doi.org/10.1056/NEJMp1500523>
- Condon, D., Drougas, A. y Abrokwah, M. (2023). The mean may not mean what you think it means: The use and misuse of measures of central tendency. *The Journal of Applied Business and Economics*, 25(4), 74-88. <https://www.proquest.com/openview/fc55a44fdac76367f0c2623a0c581638/1?pq-origsite=gscholar&cbl=38282>
- Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A. et al. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology*, 17(1), 13. <https://doi.org/10.1186/s13059-016-0881-8>
- Conover, W. J. (1999). *Practical Nonparametric Statistics* (3^a. ed.). John Wiley & Sons.
- Creswell, J. W. y Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5^a. ed.). SAGE Publications.
- Daniel, W. W. y Cross, C. L. (2018). *Biostatistics: A Foundation for analysis in the health sciences*. John Wiley & Sons.
- Dehaene, H., De Neve, J. y Rosseel, Y. (2021). A Wilcoxon–Mann–Whitney test for latent variables. *Frontiers in Psychology*, 12, 754898. <https://doi.org/10.3389/fpsyg.2021.754898>
- Fagerland, M. W., Lydersen, S. y Laake, P. (2013). The McNemar test for binary matched-pairs data: mid-p and asymptotic are better than exact conditional. *BMC Medical Research Methodology*, 13(1), 91. <https://doi.org/10.1186/1471-2288-13-91>

- Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics* (4^a. ed.). SAGE Publications.
- Fletcher, R. H., Fletcher, S. W. y Fletcher, G. S. (2019). *Clinical epidemiology: The essentials* (6^a. ed.). Wolters Kluwer.
- Fowler, F. J. (2014). *Survey research methods* (5^a. ed.). SAGE Publications.
- Frieden, T. R. (2017). Evidence for health decision making—beyond randomized, controlled trials. *New England Journal of Medicine*, 377(5), 465-475. <https://doi.org/10.1056/NEJMr1614394>
- Friedman, L. M., Furberg, C. D., DeMets, D. L., Reboussin, D. M. y Granger, C. B. (2015). *Fundamentals of Clinical Trials* (5^a. ed.). Springer.
- Gelman, A., Hill, J. y Vehtari, A. (2020). *Regression and Other Stories*. Cambridge University Press.
- Gibbons, J. D. y Chakraborti, S. (2011). *Nonparametric Statistical Inference* (5^a. ed.). CRC Press.
- Girden, E. R. (1992). *ANOVA: Repeated Measures*. Sage Publications.
- Gliklich, R. E., Dreyer, N. A. y Leavy, M. B. (2014). *Registries for evaluating patient outcomes: A user's guide* (4^a. ed.). Agency for Healthcare Research and Quality (AHRQ).
- Gordis, L. (2020). *Epidemiology* (6^a. ed.). Elsevier. <https://www.elsevier.com/books/epidemiology/gordis/9780323552295>
- Grainge, M. J. (2023). Teaching Medical Statistics Within the Context of Evidence Based Medicine. En D. Farnell y R. Medeiros. (Eds.), *Teaching Biostatistics in Medicine and Allied Health Sciences* (pp. 19-29). Springer International Publishing. https://doi.org/10.1007/978-3-031-26010-0_3
- Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E. y Tourangeau, R. (2009). *Survey methodology* (2^a. ed.). Wiley.
- Haller, B. y Krause, F. (2021). *Reproducible Medical Research with R*. CRC Press.
- Hasin, Y., Seldin, M. y Luskis, A. (2017). Multi-omics approaches to disease. *Genome Biology*, 18(1), 83. <https://doi.org/10.1186/s13059-017-1215-1>
- Hernández-Sampieri, R., Mendoza, C. P. y Torres, C. (2021). *Metodología de la investigación* (7^a. ed.). McGraw-Hill Education.
- Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J. et al. (Eds.). (2021). *Cochrane handbook for systematic reviews of interventions* (2^a. ed.). Cochrane.
- _____. (2024). *Cochrane Handbook for Systematic Reviews of Interventions* (version actualizada). Cochrane / Wiley.
- Hosmer, D. W., Lemeshow, S. y Sturdivant, R. X. (2013). *Applied Logistic Regression* (3^a. ed.). Wiley.
- Instituto Nacional de Estadística y Geografía [INEGI]. (2020). *Censo de población y vivienda 2020*. <https://www.inegi.org.mx/programas/ccpv/2020/>
- _____. (2023). *Encuestas nacionales de salud*. <https://www.inegi.org.mx/programas/>

- Ioannidis, J. P. (2019). The importance of predefined rules and prespecified statistical analyses: do not abandon significance. *Jama*, *321*(21), 2067-2068. <https://doi.org/10.1001/jama.2019.4582>
- Ioannidis, J. P., Greenland, S., Hlatky, M. A., Khoury, M. J., Macleod, M. R., Moher, D. et al. (2014). Increasing value and reducing waste in research design, conduct, and analysis. *The Lancet*, *383*(9912), 166-175. [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(13\)62227-8/abstract?_eventId=login=](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(13)62227-8/abstract?_eventId=login=)
- Jewell, N. P., Lewnard, J. A. y Jewell, B. L. (2020). Predictive mathematical models of the COVID-19 pandemic: underlying principles and value of projections. *Jama*, *323*(19), 1893-1894. <https://doi.org/10.1001/jama.2020.6585>
- Kallogjeri, D. y Piccirillo, J. F. (2023). A simple guide to effect size measures. *JAMA Otolaryngology–Head & Neck Surgery*, *149*(5), 447-451 <https://doi.org/10.1001/jamaoto.2023.0159>
- Katz, M. H. (2019). *Study design and statistical analysis: A practical guide for clinicians* (2^a. ed.). Springer.
- Keselman, H. J., Algina, J., Kowalchuk, R. K. y Wolfinger, R. D. (1998). A comparison of two approaches for selecting covariance structures in the analysis of repeated measurements. *Communications in Statistics-Simulation and Computation*, *27*(3), 591-604. <https://doi.org/10.1080/03610919808813497>
- Khorana, A., Pareek, A., Ollivier, M., Madjarova, S. J., Kunze, K. N., Nwachukwu, B. U. et al. (2023). Choosing the appropriate measure of central tendency: mean, median, or mode? *Knee Surgery, Sports Traumatology, Arthroscopy*, *31*(1), 12-15. <https://doi.org/10.1007/s00167-022-07204-y>
- Kim, H. Y. (2015). Statistical notes for clinical researchers: Type I and type II errors in statistical decision. *Restorative Dentistry & Endodontics*, *40*(3), 249-252. <https://doi.org/10.5395/rde.2015.40.3.249>
- Klijs, B., Mitratza, M., Harteloh, P. P., Moll van Charante, E. P., Richard, E., Nielen, M. M. et al. (2021). Estimating the lifetime risk of dementia using nationwide individually linked cause-of-death and health register data. *International Journal of Epidemiology*, *50*(3), 809-816. <https://doi.org/10.1093/ije/dyaa219>
- Knief, U. y Forstmeier, W. (2021). Violating the normality assumption may be the lesser of two evils. *Behavior Research Methods*, *53*(6), 2576-2590. <https://doi.org/10.3758/s13428-021-01587-5>
- Kocaoglu, B. (2024). Data Science. En *Logistics Information Systems: Digital Transformation and Supply Chain Applications in the 4.0 Era* (pp. 235-285). Springer Nature Switzerland.
- Konopatov, A. V., Shidlovsky, Y. V., Shtil, A. A. y Bylino, O. V. (2024). Kolmogórov-Smirnov and Shapiro-Wilk single distribution analysis methods in examining sample structure,

- strength of genetic background, and effects of interventions in lifespan and aging studies. *BioRxiv*, 1-32. <https://doi.org/10.1101/2024.12.12.628153>
- Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V. y Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, 13, 8-17. <https://doi.org/10.1016/j.csbj.2014.11.005>
- Kucharski, A. J. (2020). *The rules of contagion: Why things spread—and why they stop*. Basic Books.
- Kutner, M. H., Nachtsheim, C. J., Neter, J. y Li, W. (2005). *Applied Linear Statistical Models* (5^a. ed.). McGraw-Hill.
- Kwak, S. G. y Kim, J. H. (2017). Central limit theorem: The cornerstone of modern statistics. *Korean Journal of Anesthesiology*, 70(2), 144-156. <https://doi.org/10.4097/kjae.2017.70.2.144>
- Lakens, D. (2022). *Improving Your Statistical Inferences*. Cambridge University Press.
- Levine, D., Stephan, D. F., Szabat, K. A. y Berenson, M. L. (2020). *Statistics for managers using Microsoft Excel* (9^a. ed.). Pearson Education.
- Levy, P. S. y Lemeshow, S. (2020). *Sampling of Populations: Methods and Applications* (5^a. ed.). Wiley.
- Llauce, R. Y., Saldaña, M. Y. y Olivera, M. F. (2024). *Pruebas de Hipótesis Chi-Cuadrado* [Trabajo de ingeniería, Universidad Nacional de Jaén]. Repositorio Institucional Digital. <https://repositorio.unj.edu.pe/items/01662520-8ce3-470a-8ac0-c1f13349999f>
- Lohr, S. L. (2022). *Sampling: Design and analysis* (3^a. ed.). Chapman & Hall/CRC. <https://doi.org/10.1201/9780429296284>
- Maxwell, S. E., Delaney, H. D. y Kelley, K. (2017). *Designing Experiments and Analyzing Data: A Model Comparison Perspective* (3^a. ed.). Routledge.
- McDonald, J. H. (2014). *Handbook of Biological Statistics* (3^a. ed.). Sparky House Publishing.
- Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C. y Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of Cardiac Anaesthesia*, 22(1), 67-72.
- Montgomery, D. C. y Runger, G. C. (2020). *Applied statistics and probability for engineers* (7^a. ed.). John Wiley & Sons.
- Montgomery, D. C., Peck, E. A. y Vining, G. G. (2021). *Introduction to Linear Regression Analysis* (6^a. ed.). Wiley.
- Motulsky, H. (2022). *Intuitive Biostatistics* (4^a. ed.). Oxford University Press.
- Mukaka, M. M. (2012). A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal*, 24(3), 69-71. <https://doi.org/10.4314/mmj.v24i3>
- Murdock, M. (2022). Causation, Relationship, Association, and Correlation: Narrative Review. *EC Orthopaedics*, 13, 1-12. <https://eicon.net/assets/ecor/pdf/ECOR-13-00921.pdf>

- Nachar, N. (2008). The Mann-Whitney U: A test for assessing whether two independent samples come from the same distribution. *Tutorials in Quantitative Methods for Psychology*, 4(1), 13-20. <https://www.tqmp.org/RegularArticles/vol04-1/p013/p013.pdf>
- Navidi, W. (2021). *Statistics for Engineers and Scientists* (5^a. ed.). McGraw-Hill Education.
- Neuhäuser, M. y Ruxton, G. D. (2024). A note on the Wilcoxon-Mann-Whitney test and tied observations. *PlosOne*, 19(8), e0309074.
- Norman, G. R. y Streiner, D. L. (2021). *Biostatistics: The Bare Essentials* (6^a. ed.). PMPH USA.
- Nørskov, A. K., Lange, T., Nielsen, E. E., Gluud, C., Winkel, P., Beyersmann, J. et al. (2021). Assessment of assumptions of statistical analysis methods in randomised clinical trials: The what and how. *BMJ Evidence-Based Medicine*, 26(3), 121-126. <https://doi.org/10.1136/bmjebm-2019-111268>
- OECD [Organisation for Economic Co-operation and Development]. (2020). *Health at a glance 2020: OECD indicators*. OECD Publishing. <https://www.oecd.org/health/health-at-a-glance/>
- Pagano, M., Gauvreau, K. y Mattie, H. (2022). *Principles of biostatistics*. Chapman and Hall/CRC. <https://doi.org/10.1201/9780429340512>
- Peck, R., Olsen, C. y Devore, J. L. (2021). *Introduction to statistics and data analysis* (6^a. ed.). Cengage Learning.
- Petrie, A. y Sabin, C. (2023). *Medical Statistics at a Glance* (4^a. ed.). Wiley.
- Piantadosi, S. (2017). *Clinical trials: A methodologic perspective* (3^a. ed.). Wiley
- Pilamunga, B. O., López, M. B., Monar, K. R. y Sarango, A. F. (2024). Chi Cuadrado y tablas de contingencia aplicado en SPSS. *Código Científico Revista de Investigación*, 5(E3), 499-513. <http://revistacodigocientifico.itslosandes.net/index.php/1/article/view/329>
- Razali, N. M. y Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogórov-Smirnov, lilliefors and anderson-darling tests. *Journal of Statistical Modeling and Analytics*, 2(1), 21-33. https://www.nbi.dk/~petersen/Teaching/Stat2017/Power_Comparisons_of_Shapiro-Wilk_Kolmogorov-Smirnov.pdf
- Rice, J. A. (2023). *Mathematical Statistics and Data Analysis* (4^a. ed.). Cengage Learning.
- Ríos, A. R. y Peña, A. M. (2020). Estadística inferencial. Elección de una prueba estadística no paramétrica en investigación científica. *Horizonte de la Ciencia*, 10(19), 191-208. <https://revistas.uncp.edu.pe/index.php/horizontedelaciencia/article/download/597/840>
- Roberts, M. C., Holt, K. E., Del Fiore, G., Baccarelli, A. A. y Allen, C. G. (2024). Precision public health in the era of genomics and big data. *Nature Medicine*, 30(7), 1865-1873. <https://doi.org/10.1038/s41591-024-03098-0>
- Rosner, B. (2020). *Fundamentals of Biostatistics* (8^a. ed.). Cengage Learning.
- Rothman, K. J., Greenland, S. y Lash, T. L. (2021). *Modern epidemiology* (4^a. ed.). Wolters Kluwer.

- Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to Student's t-test and the Mann-Whitney U test. *Behavioral Ecology*, 17(4), 688-690. <https://doi.org/10.1093/beheco/ark016>
- Satam, H., Joshi, K., Mangrolia, U., Waghoo, S., Zaidi, G., Rawool, S. et al. (2023). Next-generation sequencing technology: current trends and advancements. *Biology*, 12(7), 997. <https://doi.org/10.3390/biology12070997>
- Schober, P., Boer, C. y Schwarte, L. A. (2018). Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, 126(5), 1763-1768. <https://doi.org/10.1213/ANE.0000000000002864>
- Shadish, W. R., Cook, T. D. y Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Sheskin, D. J. (2020). *Handbook of Parametric and Nonparametric Statistical Procedures* (5^a. ed.). CRC Press.
- Shreffler, J. y Huecker, M. R. (2023). *Exploratory data analysis: Frequencies, descriptive statistics, histograms, and boxplots*. StatPearls Publishing. <https://www.ncbi.nlm.nih.gov/books/NBK557570/>
- Sims, R., Hill, M. y Williams, J. (2020). The multiplex model of the genetics of Alzheimer's disease. *Nature Neuroscience*, 23(3), 311-322. <https://doi.org/10.1038/s41593-020-0599-5>
- Šimundić, A. M. (2009). Measures of diagnostic accuracy: basic definitions. *Journal of the International Federation of Clinical Chemistry and Laboratory Medicine*, 19(4), 203-211. <https://pubmed.ncbi.nlm.nih.gov/articles/PMC4975285/>
- Sullivan, L. M. (2018). *Essentials of Biostatistics in Public Health* (3^a. ed.). Jones & Bartlett Learning.
- Szklo, M. y Nieto, F.J. (2023). *Epidemiology: Beyond the Basics* (5^a. ed.). Jones & Bartlett Learning.
- Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G. y Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics*, 20(8), 467-484. <https://doi.org/10.1038/s41576-019-0127-1>
- Thacker, S. B. y Qualters, J. R. (2018). *Public health surveillance* (2^a. ed.). Oxford University Press.
- Thams, N., Oberst, M. y Sontag, D. (2022). Evaluating robustness to dataset shift via parametric robustness sets. *Advances in Neural Information Processing Systems*, 35, 16877-16889. https://proceedings.neurips.cc/paper_files/paper/2022/hash/6b7f9d9c1217a748391800871ff7d17d-Abstract-Conference.html
- Triola, M. F. (2020). *Elementary statistics* (13^a. ed.). Pearson Education.
- United Nations (UN). (2022). *World population prospects 2022: Summary of results*. Department of Economic and Social Affairs, Population Division. <https://www.un.org/development/desa/pd/content/world-population-prospects-2022>

- Wasserman, L. (2021). *All of Statistics: A Concise Course in Statistical Inference* (2^a. ed.). Springer.
- Wasserman, L., Ramdas, A. y Balakrishnan, S. (2024). *Statistical Inference via Data Science*. Cambridge University Press.
- Wasserstein, L., Schirm, A. L. y Lazar, N. A. (2019). Moving to a world beyond “ $p < 0.05$ ”. *The American Statistician*, 73(sup1), 1-19. <https://doi.org/10.1080/00031305.2019.1583913>
- Wasserstein, L. y Lazar, N. A. (2016). The ASA statement on p-values: context, process, and purpose. *The American Statistician*, 70(2), 129-133. <https://doi.org/10.1080/00031305.2016.1154108>
- Weissler, E. H., Naumann, T., Andersson, T., Ranganath, R., Elemento, O., Luo, Y. et al. (2021). The role of machine learning in clinical research: transforming the future of evidence generation. *Trials*, 22(1), 537. <https://doi.org/10.1186/s13063-021-05489-x>
- West, B. T., Welch, K. B. y Galecki, A. T. (2021). *Linear Mixed Models: A Practical Guide Using Statistical Software* (2^a. ed.). CRC Press.
- Whatley, M. (2022). One-Way ANOVA and the chi-square test of independence. En *Introduction to quantitative analysis for international educators* (pp. 57-74). Springer International Publishing. https://doi.org/10.1007/978-3-030-93831-4_5
- WHO [World Health Organization]. (2021). *Together on the road to evidence-informed decision-making for health in the post-pandemic era: a call for action* (No. WHO/SCI/RFH/2021.08). World Health Organization. <https://iris.who.int/bitstream/handle/10665/350883/WHO-SCI-RFH-2021.08-eng.pdf>
- Woodward, M. (2023). *Statistics for Epidemiology* (2^a. ed.). CRC Press.
- Zhou, Y., Zhu, Y. y Wong, W. K. (2023). Statistical tests for homogeneity of variance for clinical trials and recommendations. *Contemporary Clinical Trials Communications*, 33, 101119. <https://doi.org/10.1016/j.conctc.2023.101119>

Glosario de términos

Análisis de varianza (ANOVA)

Técnica estadística que permite comparar las medias de tres o más grupos independientes.

Bias (Sesgo)

Error sistemático que se introduce en el muestreo o en el análisis de datos.

Binomial, distribución

Calcula la probabilidad de obtener un número específico de éxitos en una secuencia de ensayos independientes, todos con la misma probabilidad de éxito.

Bioestadística

Subdisciplina de la estadística aplicada a la biología y la salud.

Bootstrap

Método para estimar errores estándar o intervalos de confianza en el que se selecciona aleatoriamente una pequeña muestra de observaciones de la muestra original, se calculan las estimaciones y la muestra se devuelve a la muestra original. Este proceso continúa muchas veces para producir una distribución en la que basarán las estimaciones.

Chi-cuadrada (χ^2)

Prueba que analiza la independencia entre variables categóricas o la bondad de ajuste de una distribución observada con respecto a una distribución esperada.

Coefficiente de correlación (r)

Medida de la relación lineal entre dos mediciones numéricas realizadas en el mismo conjunto de sujetos. Va de -1 a $+1$, donde 0 indica que no hay relación.

Coefficiente de determinación (R^2)

Medida que indica la proporción de la variabilidad de la variable dependiente explicada por el modelo.

Coefficiente de regresión

La b en la ecuación de regresión simple $Y = a + bX$. A veces se interpreta como la pendiente de la línea de regresión. En regresión múltiple, las b son ponderaciones aplicadas a las variables predictoras.

Coefficiente de variación

Relación entre la desviación estándar y la media, expresada como porcentaje.

Cohorte

Grupo de sujetos que permanecen juntos en el mismo estudio a lo largo del tiempo.

Confianza (nivel de)

Probabilidad de que un intervalo de confianza contenga el valor verdadero del parámetro poblacional.

Correlación

Medida que indica la fuerza y dirección de la relación entre dos variables. No implica causalidad.

Cuartil

Percentil 25 o percentil 75, llamados primero y tercer cuartiles, respectivamente.

Curva ROC

Gráfico que representa la sensibilidad frente a 1-especificidad para diferentes puntos de corte de una prueba diagnóstica.

Desviación estándar

Medida que indica cuánto se alejan, en promedio, los datos respecto a la media.

Diagrama de caja

Gráfico que muestra tanto las frecuencias como la distribución de las observaciones. Es útil para comparar dos distribuciones.

Diagrama de dispersión

Gráfico bidimensional que muestra la relación entre dos características numéricas o variables.

Diagrama de tallo y hojas

Pantalla gráfica para datos numéricos. Es similar tanto a una tabla de frecuencias como a un histograma.

Diagrama o gráfico de barras

Cuadro o gráfico que se utiliza con características nominales para mostrar los números o porcentajes de observaciones con la característica de interés.

Dicotómica, observación

Medida nominal que tiene solo dos resultados.

Distribución binomial

Distribución de probabilidad discreta que describe el número de éxitos en una secuencia de ensayos independientes.

Distribución de frecuencias

En un conjunto de observaciones numéricas, la lista de valores que ocurren junto con la frecuencia de su ocurrencia.

Distribución normal

Distribución de probabilidad continua simétrica en forma de campana.

Ensayo (estudio) clínico

Estudio experimental de un fármaco o procedimiento en el que los sujetos son seres humanos.

Error estándar (SE)

Medida de la variabilidad del estimador puntual. Depende del tamaño de la muestra y de la variabilidad de los datos.

Error tipo I

Error que ocurre al rechazar una hipótesis nula verdadera.

Error tipo II

Error que ocurre al no rechazar una hipótesis nula falsa.

Estadística

Es la disciplina que se encarga de recolectar información cuantitativa relacionada a individuos, grupos, series de hechos, entre otros.

Estadístico

Valor numérico que resume una característica de una muestra.

Estimación

Proceso de utilizar información de una muestra para sacar conclusiones sobre los valores de los parámetros en una población.

Estudio de cohorte

Estudio observacional en el que se sigue a un grupo de individuos a lo largo del tiempo.

Estudio doble ciego

Ensayo clínico en el que ni los sujetos ni los investigadores saben qué tratamiento se ha utilizado.

Estudio experimental

Estudio comparativo que implica una intervención o manipulación. Se denomina ensayo cuando se trata de seres humanos.

Estudio longitudinal

Estudio que se lleva a cabo durante un tiempo prolongado.

Estudio observacional

Estudio que no implica intervención ni manipulación. Se denomina caso Testigo.

Estudio transversal

Estudio observacional que analiza datos de una población en un momento determinado.

Frecuencia absoluta

Número de veces que ocurre un valor en un conjunto de datos.

Frecuencia relativa

Proporción de veces que ocurre un valor respecto al total de observaciones.

Hipótesis alternativa (H_1)

Afirmación que se acepta si se rechaza la hipótesis nula.

Hipótesis nula (H_0)

Afirmación que se considera válida en ausencia de evidencia contraria.

Histograma

Representación gráfica de barras contiguas proporcional a la frecuencia absoluta.

Intervalo de confianza

Rango de valores dentro del cual se espera que se encuentre un parámetro poblacional.

Kolmogórov-Smirnov (K-S)

Prueba no paramétrica, evalúa la congruencia entre una distribución empírica y una distribución teórica previamente definida, o bien, para contrastar si dos muestras independientes provienen de la misma distribución.

McNemar, prueba

Prueba de chi cuadrada para comparar proporciones a partir de dos grupos dependientes o pareados.

Media

Promedio aritmético de un conjunto de datos.

Mediana

Valor que divide un conjunto de datos ordenados en dos partes iguales.

Moda

Valor que ocurre con mayor frecuencia en un conjunto de datos.

Muestreo

Proceso mediante el cual se selecciona una muestra representativa a partir de una población.

Nivel de significancia (α)

Probabilidad máxima tolerada de cometer un error tipo I.

Odds ratio (OR)

Medida de asociación entre una exposición y un resultado.

Ojiva

Representación gráfica que muestra la frecuencia acumulada de un conjunto de datos agrupados en intervalos de clase.

P-valor

Probabilidad de obtener un resultado igual o más extremo que el observado bajo la hipótesis nula.

Parámetro

Característica numérica que describe una propiedad específica de una población.

Población

Conjunto completo de elementos que comparten una o varias características observables.

Prueba de chi-cuadrado

Prueba estadística utilizada para comparar frecuencias observadas con esperadas.

Prueba de hipótesis

Enfoque de la inferencia estadística que da como resultado la decisión de rechazar o no la hipótesis nula.

Prueba t de Student

Prueba estadística para comparar medias entre dos grupos.

Rango (Re)

Medidas de dispersión que indica la diferencia entre el valor más alto y el valor más bajo dentro de un conjunto de datos.

Regresión

Técnica estadística que permite modelar la relación entre una variable dependiente y una o más independientes.

Sesgo

Error relacionado con las formas en que las poblaciones objetivo y muestreadas difieren; también llamado error de medición es una amenaza para la validez de un estudio.

t, prueba

Determina si existen diferencias estadísticamente significativas entre las medias de dos grupos.

Teorema del límite central

Teorema que establece que la distribución de medias es aproximadamente normal, si el tamaño de la muestra es lo suficientemente grande ($n \geq 30$), independientemente de la distribución subyacente de las medidas originales.

U de Man – Whitney – Wilcoxon, prueba

Ver Wilcoxon, prueba.

Valor p

Probabilidad de observar un resultado tan extremo, o más extremo, que el que realmente se observa por casualidad (es decir, si la hipótesis nula es cierta).

Variable

Es una característica de la población o muestra que desea estudiarse.

Variable cualitativa

Representa cualidades, atributos o categorías no numéricas.

Variable cuantitativa

Valores numéricos que representan cantidades o magnitudes y permiten realizar operaciones aritméticas con sentido matemático.

Variable dependiente

Variable cuyos valores son los resultados de un estudio; también llamada variable de respuesta o criterio.

Variable independiente

La variable explicativa o predictora en un estudio.

Varianza

Medida de dispersión que indica la variabilidad total de los datos respecto a la media aritmética.

Wilcoxon, prueba

Prueba no paramétrica que compara dos muestras dependientes con datos ordinales u observaciones numéricas que no tienen distribución normal.

Algunas obras de los coordinadores:

- Tecnologías aplicadas en el área de ciencias de la salud
- Educación en salud
- Temas selectos de medicina
- Manual Grecolatino de Terminología Médica

Consulta estos títulos dentro del catálogo de Libros UAT del Consejo de Publicaciones en el siguiente enlace:



<https://libros.uat.edu.mx>

 <https://publicaciones.uat.edu.mx>

Equipo editorial

Coordinación: Venancio Vanoye Eligio

Gestión y administración: Jessica Abigail Rodríguez Tinajero, María Teresa Maldonado Sada

Revisión y corrección de estilo: José Luis Énder Velarde García, Jorge Alberto Vázquez Herrera

Diseño y maquetación: Erika González Navarro, Wendy Castillo Cruz, Lorena E. Cortez Rodríguez

Bioestadística: principios, métodos y aplicaciones
de Miriam Janet Cervantes López, Jaime Cruz Casados,
Laura Nelly Cruz Casados y Raúl Sadrach Obando Carmona, coordinadores,
publicado por la Universidad Autónoma de Tamaulipas en marzo de 2026.
La revisión y diseño editorial correspondieron al Consejo de Publicaciones UAT.



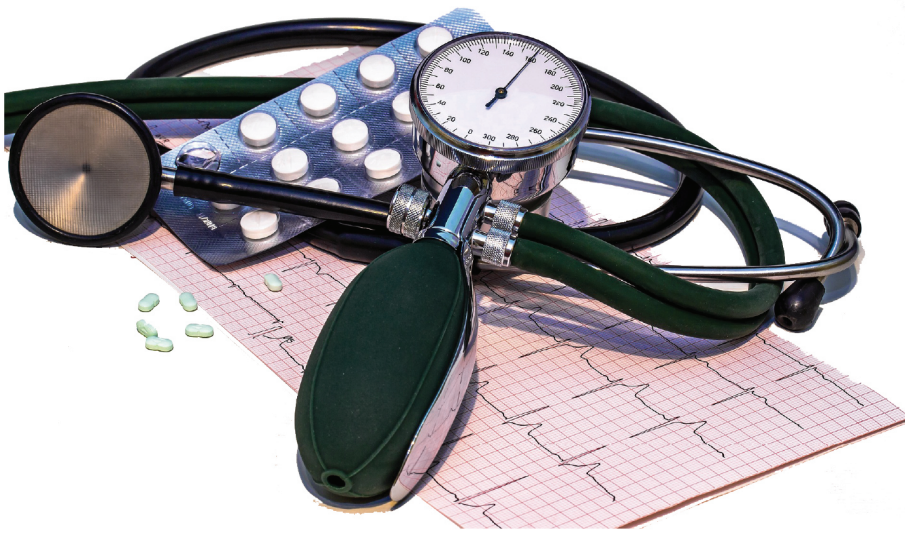
La formación médica exige habilidades clínicas, conocimientos biológicos y un entendimiento profundo de los métodos que permiten transformar las observaciones en información sistematizada. De esta forma, la estadística ocupa un lugar central en el currículo de las ciencias de la salud, pues convierte la experiencia en evidencia y la evidencia en acción sanitaria fundamentada.

Bioestadística: principios, métodos y aplicaciones es una obra dirigida a estudiantes de medicina, con una vocación pedagógica y una rigurosa fundamentación científica. Su enfoque progresivo parte de los conceptos básicos y avanza hacia técnicas de inferencia, regresión y análisis multivariado, con el objetivo de que el lector construya una competencia estadística sólida, crítica y contextualizada para los desafíos reales de la práctica médica.

Este libro articula teoría, método y aplicación clínica. Cada unidad está respaldada por criterios metodológicos contemporáneos y por ejemplos de escenarios médicos reales: desde el análisis de brotes epidémicos hasta el diseño de ensayos clínicos; desde la vigilancia epidemiológica hasta la medicina personalizada.

En tiempos en que la medicina se redefine por la integración de tecnología, datos y decisiones éticas, es prioritario formar médicos capaces de leer críticamente, investigar y participar en políticas sanitarias basadas en la evidencia. Este libro fomenta el compromiso con una medicina informada, justa y técnicamente solvente.

ISBN UAT: 978-607-69439-1-5




Consejo de
Publicaciones