EDITORIAL NAVEGANTE

LA INVESTIGACIÓN CIENTÍFICA: MÉTODOS ESTADÍSTICOS en las CIENCIAS SOCIALES

Ramiro Valdivia Rodríguez
Karen Meddellit Mendoza Gutiérrez
Nery Rafael Bernabé
Marleni Mary Ayma Jiménez



Ramiro Valdivia Rodríguez
Correo: ramirovaldiviarodriguez@gmail.com
Perfil: Abogado especialista en Gestión Pública, Sub Gerente del Estudio Jurídico Urquizo &Valdivia Asociados SRL, magister en Gestión Pública, egresado del Doctorado en Ciencias Ambientales, egresado de la maestría en Ciencias de la Educación con mención en Docencia Superior e Investigación, Docente Universitario Especialista en Metodología de la Investigación Jurídica
ORCID: https://orcid.org/0000-0001-9098-2711
Afiliación: Universidad Nacional Jorge Basadre Grohmann.

Karen Meddellit Mendoza Gutiérrez
Correo: meddellit.28534@gmail.com
Perfil: Doctora en Gestión Pública y Gobernabilidad, Magister en Gestión Pú-blica, Obstetra especialista en monitoreo fetal y ecografía bostétrica, Licenciada en Obstetricia, Docente universitaria.
ORCID: https://orcid.org/0000-0003-0447-4628
Afiliación: Universidad Nacional Jorge Basadre Grohmann.

Nery Rafael Bernabé
Correo: neryrabe@gmail.com
Perfil: Obstetra Asistencial, Doctora en Salud Pública, Magíster en Salud Pública con mención en Gestión de Servicios de Salud, egresada de la Maestría en Docencia Universitaria.
ORCID: https://orcid.org/0009-0002-4110-1791
Afiliación: Universidad Nacional Jorge Basadre Grohmann.

Marleni Mary Ayma Jiménez
Correo: marfido302005@gmail.com
Perfil: Obstetra Asistencial, Maestro en Ciencias con mención en Salud Pública, Especialista en monitoreo Fetal y Diagnóstico por Imágenes en Obstetricia. Docente Universitario.
ORCID: https://orcid.org/0000-0002-2050-4519
Afiliación: Universidad Nacional Jorge Basadre Grohmann.

LA INVESTIGACIÓN CIENTÍFICA:

Método estadísticos en las Ciencias Sociales

Ramiro Valdivia Rodríguez

Karen Meddellit Mendoza Gutiérrez

Nery Rafael Bernabé

Marleni Mary Ayma Jiménez

EDITORIAL
NAVEGANTE

Todas nuestras publicaciones son sometidas a revisión doble-ciego de pares académicos (*Peer Review Double Blinded*).

Esta publicación cuenta con licencia Creative Commons Reconocimiento - NoComercial - SinObraDerivada 3.0 Unported License.



ISBN: 978-628-7736-39-9

© Ramiro Valdivia Rodríguez
© Karen Meddellit Mendoza Gutiérrez
© Nery Rafael Bernabé
© Marleni Mary Ayma Jiménez

2025
DOI: https://doi.org/10.5281/zenodo.14871680
© Editorial Navegante
www.editorialnavegante.com

Queda prohibida la reproducción bajo cualquier modalidad de toda o una parte de esta obra sin autorización expresa del titular de los derechos.

Edición electrónica: Editorial Navegante

Editado en Colombia/ *Published in Colombia*

LA INVESTIGACIÓN CIENTÍFICA:

Método estadísticos en las Ciencias Sociales

Ramiro Valdivia Rodríguez

Karen Meddellit Mendoza Gutiérrez

Nery Rafael Bernabé

Marleni Mary Ayma Jiménez

EDITORIAL
NAVEGANTE

PRÓLOGO

La investigación en las ciencias sociales se enfrenta a un desafío constante: transformar la complejidad del comportamiento humano y las dinámicas sociales en datos cuantificables y significativos. En este contexto, la estadística se convierte en una herramienta fundamental, proporcionando métodos rigurosos para analizar, interpretar y fundamentar los hallazgos científicos.

Este libro, *La Investigación Científica: Métodos Estadísticos en las Ciencias Sociales*, surge con el propósito de servir como una guía clara y accesible para quienes desean comprender y aplicar los métodos estadísticos en sus estudios. Lejos de ser una mera recopilación de fórmulas y procedimientos, la obra busca demostrar cómo la estadística no solo organiza información, sino que permite dar sentido a los fenómenos sociales, ofreciendo respuestas basadas en evidencia y evitando interpretaciones erróneas.

Los autores han estructurado este trabajo de manera que el lector pueda recorrer desde los principios básicos de la estadística hasta su aplicación en el análisis de datos. Se abordan tanto la estadística descriptiva—clave para resumir y visualizar información—como la inferencial, que posibilita extrapolar conclusiones a poblaciones más amplias. Además, se incluyen ejemplos prácticos y estudios de caso que ilustran cómo estas herramientas pueden aplicarse a problemáticas reales en disciplinas como la sociología, la psicología, la economía y la educación.

El enfoque de este libro no solo es teórico, sino también aplicado. En un mundo donde las decisiones deben sustentarse en datos confiables, la correcta interpretación de los resultados estadísticos es esencial. Así, esta obra no solo enseña a calcular, sino también a razonar, cuestionar y comunicar los hallazgos de manera efectiva.

Los conocimientos aquí reunidos permitirán al lector fortalecer su capacidad investigativa, mejorar la calidad de sus estudios y desarrollar una visión crítica frente a la información que consume y produce. La estadística, lejos de ser un mero recurso técnico, se presenta como un pilar del pensamiento científico y una aliada indispensable en la construcción del conocimiento en las ciencias sociales.

ÍNDICE

PRÓLOGO 6
ÍNDICE DE FIGURAS 11
ÍNDICE DE TABLAS 12
INTRODUCCIÓN 13

CAPÍTULO I

APORTES DE LA ESTADÍSTICA 15

1.1. Puntuaciones individuales 15
1.2. Descripción de grupos para tomar decisiones 19
1.3. Aplicación de la estadística en modelos 22
1.4. Estadística desde la percepción científica 26

CAPÍTULO II

APROXIMACIONES TEÓRICAS DE LA ESTADÍSTICA 29

2.1. Definición y objetivos 29
2.2. Estadística deductiva o descriptiva 33
2.3. Estadística inductiva o inferencial 37

CAPÍTULO III

MÉTODOS ESTADÍSTICOS 41

3.1 Historia 41

3.2 Clasificación según sus fines 44

3.2.1. Exploratorios 44

3.2.2. Confirmatorios 46

3.3. Clasificación según sus presupuestos teóricos 49

3.3.1. Paramétricos 49

3.3.2. No paramétricos 51

3.4. Clasificación según la clase de variable utilizada 54

3.4.1. Univariados 54

3.4.2. Multivariados 57

CAPÍTULO IV

MÉTODOS EN LA ESTADÍSTICA DESCRIPTIVA
E INFERENCIAL 64

4.1. Métodos en la estadística descriptiva 64

4.1.1. Métodos que permiten la búsqueda y la obtención de la información 64

4.1.2. Métodos para sistematizar y organizar la información 70

4.2. Métodos en la estadística inferencial 76

4.2.1. Métodos para calcular el componente de la tendencia secular en las series cronológicas 76

4.2.2. Métodos para determinar el espacio muestral en la teoría de probabilidades..... 78

4.2.3. Métodos utilizados en la teoría de la estimación..... 80

4.2.4. Método de la prueba de hipótesis o teoría de decisión 83

REFLEXIONES FINALES..... 86

REFERENCIAS BIBLIOGRÁFICAS..... 88

ÍNDICE DE FIGURAS

Figura 1
Gráfica radial de demanda global por parte de las instituciones educativas 35

Figura 2
Medidas o parámetros de la estadística descriptiva 37

Figura 3
Enfoques de inferencia en el enfoque clásico estadístico 39

Figura 4
Esquema del paradigma Análisis Exploratorio de Datos (AED) 45

Figura 5
Pruebas paramétricas y no paramétricas 55

ÍNDICE DE TABLAS

Tabla 1
Cuadro comparativo de métodos exploratorio y confirmatorios 48

Tabla 2
Comparativo entre los métodos univariados y multivariados 61

INTRODUCCIÓN

La investigación científica en las ciencias sociales se ha enriquecido considerablemente gracias a la aplicación de herramientas estadísticas. Este libro, *La Investigación Científica: Métodos Estadísticos en las Ciencias Sociales*, se concibe como una guía esencial que explora en profundidad los aportes de la estadística de manera general a los diferentes ámbitos de las ciencias sociales, contribuyendo con un compendio de conocimiento sobre los métodos estadísticos fundamentales de la estadísticas descriptiva e inferencial y sus diferentes aplicaciones.

La obra aborda la estadística como un sólido apoyo en el proceso de investigación, proporcionando a los lectores una comprensión clara de cómo los métodos estadísticos permiten analizar datos, extraer conclusiones significativas y tomar decisiones informadas en el ámbito de las ciencias sociales. Se exploran los métodos estadísticos descriptivos para resumir y presentar datos de manera efectiva, así como los métodos inferenciales para generalizar hallazgos de muestras a poblaciones más amplias.

El libro se distingue por su enfoque práctico y aplicado. A través de ejemplos de estudios de casos ficticios, que demuestran cómo los métodos estadísticos pueden utilizarse para abordar problemas de investigación concretos en diversas disciplinas de las ciencias sociales. Se presta especial atención a la interpretación de resultados y a la comunicación efectiva de hallazgos estadísticos.

La obra tiene como objetivo principal: Comprender los fundamentos de la estadística, para que el lector se forme una base sólida en los conceptos y principios de la estadística, que incluye entre otros conocimientos la comprensión de variables, escalas de medición, distribuciones de probabilidad y conceptos clave de la estadística descriptiva e inferencial y como las mismas son aplicables a los ámbitos de las ciencias sociales.

CAPÍTULO I

APORTES DE LA ESTADÍSTICA

1.1. Puntuaciones individuales

La capacidad humana para realizar juicios intuitivos sobre atributos cuantitativos como la talla o el peso se basa en referentes contextuales y experiencias previas, pero la interpretación de variables psicológicas como: la autoestima, la ansiedad, la estabilidad emocional o la inteligencia, presenta mayores desafíos debido a su naturaleza abstracta y a la dificultad de establecer unidades de medida directas. Es en este punto, donde la estadística ofrece un marco conceptual -factores, constructos claves, relaciones de probabilidad- y herramientas precisas para abordar tales problemática, comprenderlas y darles respuesta (Pérez-Juste, 2012).

Barreto-Villanueva (2012) y Pérez-Juste (2012), refiere que mediante la descripción de distribuciones de datos (media, desviación estándar) y el cálculo de puntuaciones individuales o tipificadas (z-scores, cuartiles), que permiten situar a un individuo en relación a un grupo de referencia, cuantificando así su posición relativa, la estadística facilita la comparación

entre diferentes variables y la identificación de patrones y relaciones, lo que resulta elemental para la construcción y validación de teorías en las ciencias sociales.

Las puntuaciones individuales, como refiere Pérez-Juste (2012), representan un sistema de medición y comprobación en el análisis estadístico, ya que permiten transformar datos crudos en información comparable y comprensible. Al asignar un valor numérico específico a cada observación- la altura de una persona o el puntaje obtenido en un examen- se facilita la comparación entre individuos y la identificación de patrones o tendencias dentro de un conjunto de datos.

Para entender la relevancia y el aporte de la aplicación de las puntuaciones individuales en las ciencias sociales, se puede considerar, por ejemplo, un estudio sociológico sobre el nivel de satisfacción laboral de los empleados de una empresa: Cada empleado podría asignar una puntuación del 1 al 10 para expresar su grado de satisfacción, una puntuación individual de 8 indica el nivel de satisfacción de un empleado en particular. Sin embargo, el verdadero significado de esta puntuación se revela al compararse con la puntuación promedio de todos los empleados de la empresa.

Mediante el cálculo de la media y la desviación estándar de las puntuaciones de satisfacción, se determina si el empleado en cuestión se encuentra por encima, por debajo o en línea con la satisfacción general de sus colegas. Esta información resulta invaluable para identificar áreas de mejoras, en este caso, en las condiciones laborales y diseñar estrategias para aumentar la satisfacción de los empleados (Pérez-Juste, 2012).

Corroborando lo anterior, Calderón y Ledesma (2012), manifiestan que para la interpretación de variables complejas como las psicológicas o sociodemográficas, que incluyen por ejemplo: inteligencia, facilidad relacional o la misma satisfacción laboral, las puntuaciones individuales se vuelven útiles y necesarias, ya que en estos casos la medición son menos directa y las escalas de medición pueden variar, pero a pesar de estas limi-

taciones, la estadística ofrece una serie de herramientas para cuantificar y comparar este tipo de variables, que podrían ser muy abstractas en ciertos estudios, dando resultados veraces.

Las puntuaciones típicas (o z-scores), por ejemplo, permiten expresar una puntuación individual en términos de cuántas desviaciones estándar se encuentra por encima o por debajo de la media del grupo, facilitando la comparación entre diferentes variables, incluso si se miden en escalas distintas. Otros índices estadísticos, como los cuartiles o los percentiles, proporcionan información adicional sobre la posición relativa de un individuo dentro de una distribución (Calderón y Ledesma, 2012).

Para Daza (2015), Las puntuaciones individuales son instrumentos esenciales en la transformación de datos, en información útil, ya que al permitir situar un individuo en relación con un grupo de referencia, permiten realizar comparaciones significativas y extraer conclusiones sobre características individuales y colectivas socialmente hablando, teniendo en cuenta que existen mediciones individuales específicas para cada caso, debido a las limitaciones inherentes que conlleva a la medición de ciertas variables.

Por estas razones, son variadas las puntuaciones individuales que aportan veracidad científica a las ciencias sociales, según Daza (2015), las más relevantes en el área social podrían ser:

a) Puntuaciones directas: Representan los valores originales obtenidos al aplicar un instrumento de medición a un conjunto de individuos o unidades de análisis. Proporcionan valores crudos, sin ningún tipo de transformación y reflejan la magnitud de la variable en su escala de medición original. En el contexto de las ciencias sociales, las puntuaciones directas son el punto de partida para una amplia gama de análisis estadísticos; su interpretación puede resultar un tanto limitada, debido a que las escalas de medición utilizadas en las ciencias sociales suelen ser

ordinales o de intervalo, lo que restringe las operaciones matemáticas permitidas, dando origen a las puntuaciones transformadas.

b) Puntuaciones transformadas, como se ha mencionado con anterioridad, se derivan a partir de las puntuaciones directas mediante operaciones matemáticas, como la estandarización o la normalización. Estas transformaciones permiten comparar puntuaciones obtenidas en diferentes escalas o distribuciones, facilitando así la interpretación y la realización de inferencias estadísticas en las ciencias sociales.

c) Puntuaciones proporcionales y porcentuales: A diferencia de las ciencias naturales, donde a menudo se buscan respuestas correctas o incorrectas, en las ciencias sociales las variables suelen ser de naturaleza cualitativa o cuantitativa ordinal, lo que dificulta la aplicación directa de puntuaciones normativas. En este contexto, este tipo de puntuaciones adquieren relevancia, ya que, las proporcionales como su nombre lo indican, expresan la proporción de individuos que presentan una determinada categoría o valor de una variable en relación al total de la muestra; mientras que las segundas representan este mismo valor en términos de porcentaje. Ambas puntuaciones permiten describir la distribución de las respuestas en un cuestionario, identificar tendencias y comparar grupos, sin requerir de un referente normativo externo; sin embargo, es importante destacar que estas medidas no ofrecen información sobre el rendimiento relativo de los individuos dentro de un grupo, ni permiten establecer juicios de valor sobre las respuestas.

d) Puntuaciones típicas: Se cuentan entre ella las puntuaciones z o estandarizadas, las mismas, constituyen una transformación lineal de los datos originales que permite estandarizar las distribuciones y facilitar la comparación entre diferentes variables o grupos. En el contexto de las ciencias sociales, estas puntuaciones expresan la posición relativa de un individuo dentro de una distribución en términos de desviaciones estándar respecto a la media, por ejemplo: un valor positivo de z indica que la puntuación está por encima de la media, mientras que un valor

negativo señala una puntuación por debajo. La utilidad de las puntuaciones típicas radica en su capacidad para transformar distribuciones con diferentes medias y desviaciones estándar en una distribución estándar con media 0 y desviación estándar 1, lo que facilita la interpretación y comparación de los resultados. Además, permiten identificar valores atípicos y evaluar la homogeneidad de los grupos.

e) Puntuaciones cuantiles: Son medidas de posición que dividen un conjunto de datos en partes iguales, proporcionando una descripción detallada de la distribución de una variable. En las ciencias sociales, los cuantiles son una herramienta esencial para analizar la variabilidad, identificar valores atípicos y comparar grupos, contribuyendo a una mejor comprensión de los fenómenos sociales.

f) Puntuaciones percentiles: Ofrecen una medida de posición que indica el porcentaje de datos que se encuentran por debajo de un determinado valor. En las ciencias sociales, los percentiles son ampliamente utilizados para describir y comparar distribuciones, especialmente cuando se trabaja con variables continuas. Por ejemplo, el percentil 50 corresponde a la mediana, que divide la distribución en dos partes iguales. Los percentiles permiten interpretar los resultados de manera más intuitiva, ya que indican el rango en el que se ubica un individuo en relación con el resto del grupo. Sin embargo, es importante tener en cuenta que los percentiles son sensibles a la forma de la distribución y pueden no ser directamente comparables entre diferentes conjuntos de datos.

1.2. Descripción de grupos para tomar decisiones

La estadística se convirtió en una herramienta importante y de gran utilidad en las ciencias sociales, permitiendo caracterizar, describir y comparar grupos de la manera más precisa posible, a través de técnicas como la estadística descriptiva (medidas de tendencia central y dispersión) y la inferencial (pruebas de hipótesis y estimación). A partir de estas técnicas, los investigadores pueden identificar patrones, tendencias y dife-

rencias significativas entre grupos, considerando variables demográficas, socioeconómicas y psicológicas, siendo crucial garantizar la representatividad de las muestras para que las conclusiones obtenidas sean generalizables a la población de interés (Pérez-Tejada, 2008).

Entendiendo que la representatividad de una muestra, refiere la capacidad de un subgrupo de reflejar de manera precisa las características de un conjunto más grande, conocido como población. Para López y Diez (2017), la representatividad de la muestra captura la diversidad y heterogeneidad en proporciones adecuadas, de allí radica la importancia de la misma y la adecuada caracterización que se realice del subgrupo, ya que una correcta representación de una muestra permite confiar en que las conclusiones a las que se llega a partir de la misma, son válidas para el grupo más amplio que se pretende estudiar. Por ejemplo, si se busca conocer las opiniones de los estudiantes universitarios sobre un nuevo programa académico, es fundamental que la muestra incluya a estudiantes de diferentes carreras, años y procedencias geográficas, para así asegurar que los resultados sean generalizables a toda la población estudiantil.

Pérez-Tejada (2008), reseña que las técnicas estadística que garantizan una precisa representación de cada elemento de la muestra de la población en estudio, tiene una probabilidad alta de ser seleccionado en la ciencias sociales, ya que reduce el sesgo y aumenta la confiabilidad de los resultados en cuanto a una buena caracterización del grupo mayor, y dicha caracterización no solo suministra una descripción detallada, sino que también sienta las bases para establecer relaciones causales entre variables y realizar inferencias más amplias. Al modelar matemáticamente los fenómenos sociales, la estadística proporciona un marco riguroso para probar hipótesis y tomar decisiones informadas, contribuyendo así al desarrollo de teorías, políticas públicas y medidas de mejoras sociales sólidas y con más precisión .

Como señala Levin y Rubin (2019), la identificación de esquemas distintivos en los datos y la cuantificación de las diferencias entre grupos son

aspectos clave en la investigación en ciencias sociales. Esta capacidad es indispensable para comprender los fenómenos sociales y generar conocimiento científico, por lo tanto, la caracterización de grupos es esencial para ubicar a los individuos en un contexto más amplio y comprender sus atributos en relación con un grupo de referencia, ya que al comparar las características de un individuo con las de un grupo, no solo se puede identificar patrones y tendencias, sino otorgar posibles explicaciones de ciertos fenómenos tanto generales como específicos, socialmente complejos como el suicidio en determinadas poblaciones.

Por su parte, Pérez-Juste (2012), refiere que la estadística proporciona la caracterización rigurosa y objetiva de los grupos apoyándose en medidas como la media, mediana, moda, varianza y desviación estándar, mediantes las mismas, se describe la distribución de los datos, se detectan diferencias significativas entre grupos, convirtiendo la estadística en una disciplina que encuentra aplicación en diversas áreas y permite a partir de la caracterización de los grupos tomar decisiones más acertadas y optimizar en sí mismo dicho proceso. Por ejemplo, Pérez-Juste (2012), manifiesta que a partir de una adecuada caracterización grupal:

a) Los educadores, pueden decidir segmentar a los estudiantes en función de sus características sociodemográficas, estilos de aprendizaje y rendimiento académico. Esta segmentación, a su vez, permite diseñar estrategias pedagógicas personalizadas y realizar intervenciones educativas focalizadas en aquellos estudiantes que presentan mayores dificultades.

b) Los psicólogos, llegan a identificar patrones de comportamiento, que enriquecen tanto los diagnósticos de trastornos mentales, como la evaluación de la eficacia de diferentes intervenciones terapéuticas, con lo cuales, pueden construir perfiles psicológicos detallados que facilitan la selección de tratamientos adecuados y eficaces

c) Los sociólogos, luego de analizar las dinámicas sociales, pueden identificar las desigualdades y predecir tendencias, en fin, comprender las

causas subyacentes de los fenómenos sociales y diseñar políticas públicas más efectivas con base a la data obtenida de la representación de una muestra poblacional.

d) Los economistas, lograr segmentar mercados, evaluar la rentabilidad de inversiones y predecir la demanda de productos y servicios, esto contribuye en la toma de decisiones empresariales más oportunas y acertadas, igualmente en la formulación de políticas económicas más precisas y exitosas.

Pérez-Tejada (2008); Levin y Rubin (2019), refieren que la estadística, al facilitar la identificación y caracterización de grupos homogéneos y heterogéneos dentro de una población, proporciona una base empírica sólida para la toma de decisiones informadas en diversos campos, esta capacidad de segmentación no solo es esencial para la práctica profesional, la investigación aplicada sino que también es básica para generar y avanzar en el conocimiento científico en el ámbito de las ciencias sociales.

1.3. Aplicación de la estadística en modelos

Pérez-Tejada (2008), refiere que la estadística inferencial constituye una herramienta elemental en la investigación en ciencias sociales, permitiendo trascender la mera descripción de los datos empíricos- como se ha mencionado con anterioridad- al analizar información proveniente de muestras representativa, obtenidas a partir de instrumentos como test, encuestas y observaciones, los investigadores buscan establecer generalizaciones más amplias o realizar inferencias sólidas sobre características desconocidas de poblaciones más grandes. En este contexto la estadística inferencial, al proporcionar un marco riguroso para el análisis de datos, sea hace fundamental en la construcción y evaluación mediantes modelos en las ciencias sociales, ya que contribuye a determinar la generalidad con más precisión en la caracterización de una población completa.

En esta misma línea de investigación, Pérez-Juste (2012), menciona que los datos obtenidos a partir de una muestra representativa de una población más amplia, representan la materia prima para la construcción de modelos estadísticos. Los estadísticos (características descriptoras), calculados a partir de los datos muestrales, son utilizados para estimar los parámetros poblacionales (valores inferidos a una población en un explícito atributo), y esta estimación, inherentemente sujeta a un error de muestreo, es un proceso fundamental en la inferencia estadística. Por tal razon, la elección del modelo estadístico más adecuado para analizar los datos es crucial, ya que permite realizar inferencias válidas y mas solidas sobre la población de estudio.

La decisión de aplicar un modelo estadístico particular a un conjunto de datos no es arbitraria, dicha decisión se basa en una serie de supuestos teóricos y en la evaluación de la adecuación del modelo a los datos empíricos. Para Pérez-Tejada (2008); González y Pérez de Vargas (2009), Levin y Rubin (2019), los modelos que se adecuan a los datos empíricos, cada uno llega a estar diseñado para abordar diferentes tipos de problemas. Por tal razón, algunos de los modelos más comunes incluyen:

a) Los modelos de regresión, permiten en ciencias sociales analizar de manera cuantitativa la relación entre variables. Al modelar cómo una variable dependiente (aquella que se busca explicar) varía en función de una o más variables independientes (factores que se consideran influyentes), estos modelos permiten no solo describir patrones de asociación, sino también realizar inferencias causales bajo ciertas condiciones. Por ejemplo, en sociología, un modelo de regresión podría utilizarse para evaluar el impacto del nivel educativo y el género en los ingresos salariales, permitiendo identificar si existen desigualdades salariales y cuáles son los factores que las explican. En términos predictivos, los modelos de regresión ofrecen la posibilidad de estimar valores futuros de la variable dependiente a partir de valores conocidos de las variables independientes, lo que resulta de gran utilidad en diversas áreas como la economía, la psicología y la ciencia política.

b) Los modelos de análisis de varianza (ANOVA), se utilizan especialmente cuando se busca comparar las medias de diferentes grupos o categorías, permitiendo identificar de manera rigurosa si existen diferencias significativas entre los grupos estudiados, lo que a su vez facilita la inferencia sobre las relaciones causales subyacentes. Por ejemplo, en un estudio sobre el rendimiento académico, un ANOVA podría emplearse para comparar las calificaciones promedio de estudiantes provenientes de diferentes niveles socioeconómicos; de esta forma, se podría determinar si existe un efecto significativo del nivel socioeconómico sobre el rendimiento, más allá de lo que podría atribuirse al azar. En términos más generales, los ANOVA permiten predecir la probabilidad de que las diferencias observadas entre grupos sean genuinas y no simplemente el resultado de la variabilidad aleatoria, esta capacidad predictiva es esencial para construir teorías sólidas y generalizables en las ciencias sociales.

c) Los modelos lineales generalizados (MLG), útiles cuando se busca modelar relaciones entre variables donde la variable dependiente no sigue una distribución normal, estos modelos permiten abordar una amplia gama de fenómenos sociales, como la probabilidad de que un individuo adopte un determinado comportamiento (variable binaria), la frecuencia con la que ocurre un evento (variable de conteo), o la intensidad de una actitud (variable ordinal). Al flexibilizar la función de enlace, los MLG adaptan la relación entre las variables predictoras y la variable de respuesta a diferentes tipos de distribuciones, lo que los hace especialmente útiles para analizar datos complejos y heterogéneos propios de las ciencias sociales. Por ejemplo, un MLG de regresión logística podría utilizarse para predecir la probabilidad de que un individuo vote por un determinado candidato político en función de variables sociodemográficas y de actitudes políticas; de esta manera, los MLG no solo permiten describir patrones en los datos, sino también realizar inferencias causales y generar predicciones sobre fenómenos sociales de interés.

d) Los modelos de ecuaciones estructurales, permiten analizar de manera sofisticada relaciones causales complejas entre múltiples variables, tanto observadas como latentes. Estos modelos no solo describen las asociaciones entre variables, sino que además permiten inferir sobre la dirección y magnitud de los efectos causales, facilitando la construcción de teorías más robustas y la comprensión profunda de los fenómenos sociales. Un ejemplo ilustrativo es el estudio de la relación entre el capital social, la confianza interpersonal y el bienestar subjetivo, donde los modelos de ecuaciones estructurales pueden desentrañar la mediación de la confianza interpersonal en el vínculo entre el capital social y el bienestar, proporcionando evidencia empírica sobre mecanismos causales subyacentes. Estos modelos, permiten predecir no solo los valores de variables observables, sino también los de variables latentes, lo que a su vez facilita la formulación de hipótesis más precisas y la evaluación de la efectividad de intervenciones sociales.

e) Los modelos multivariados, permiten analizar de manera simultánea la relación entre múltiples variables dependientes y un conjunto de variables independientes. Esta capacidad de modelar relaciones complejas entre diversos factores resulta especialmente valiosa para comprender fenómenos sociales multidimensionales. Por ejemplo, en el campo de la sociología, un modelo multivariado podría utilizarse para analizar cómo factores como el nivel educativo, el ingreso familiar y la discriminación afectan simultáneamente las oportunidades laborales de un individuo. De esta manera, los modelos multivariados no solo describen la realidad social, sino que también permiten realizar inferencias causales y predecir tendencias futuras, contribuyendo a la formulación de políticas públicas más informadas.

La elección del modelo adecuado depende de: La naturaleza de las variables (Si son continuas, discretas, ordinales o nominales); La distribución de los datos, (Si siguen una distribución normal o alguna otra distribución conocida); La relación entre las variables, (Si es lineal o no lineal); y los objetivos del estudio, (Si se busca describir, explicar o predecir fenómenos).

Además, la elección del modelo estadístico adecuado es un paso crucial en el proceso de investigación y requiere una comprensión profunda de los fundamentos teóricos y de los datos a analizar (Levin y Rubin, 2019).

1.4. Estadística desde la percepción científica

Para Pérez-Tejada (2008), la investigación científica se constituye como una estructurada y sólida empresa intelectual destinada a comprender los fenómenos naturales y sociales, con el objetivo primordial de desarrollar teorías firmes que proporcionen explicaciones coherentes y comprensivas de realidades generales y puntuales de dichos fenómenos. Siendo esto posible, mediante el cumplimiento de un riguroso método científico, que permite a los investigadores la construcción de modelos teóricos que no solo describen los eventos observados, sino que también consienten realizar predicciones precisas y, en muchos casos, intervenir en algunos procesos naturales y sociales. Este ciclo de observación, hipótesis, experimentación y teorización es lo que caracteriza a la ciencia como una actividad productiva y vivaz, en constante evolución, cuyo fin último es ampliar el conocimiento humano y mejorar la calidad de vida.

Calduch (s.f); Calderón y Ledesma (2012), en la misma línea de opinión, refieren que en la investigación científica la operación del paradigma de contrastación empírica constante, conlleva a las teorías a competir por el ofrecimiento de explicaciones más profundas, densas, pero comprehensivas de los fenómenos naturales y sociales sometidos a exploración. Siendo el papel de la Estadística como disciplina científica y de sostén, la protagonista en la incorporación de datos informativos que sostienen las diferentes experimentaciones y observaciones rigurosas de los fenómenos en estudio; esta integración de nuevos datos permite la evaluación y corroboración de la firmeza de la coherencia interna del cuerpo teórico.

Dicha dinámica competitiva, fomenta el progreso científico, ya que las teorías que no resisten el escrutinio comprobativo de la data estadística son gradualmente descartadas en favor de aquellas que ofrecen un mayor

poder explicativo basado en datos formales cuantitativos que precisen no solo el ¿Qué? y el ¿Por qué? de los fenómenos en estudio, sino que por igual, aperturen nuevas vías de investigación, estimulen la exploración de ámbitos aún desconocidos del conocimiento y generen nuevas preguntas e hipótesis como indicadores de la fecundidad científica (Calduch, s.f.; Calderón y Ledesma, 2012).

Pérez-Tejada (2008), reseña que los procedimientos y pasos sistemáticos que se siguen en el estudio de la investigación científica experimental, como método riguroso para establecer relaciones causales, encuentra en la estadística un aliado indispensable, ya que esta disciplina proporcionar un marco teórico sólido para el diseño experimental y el análisis de datos, porque permite cuantificar la incertidumbre y el error inherentes a cualquier medición, facilitando así la inferencia causal.

Barreto Villanueva (2012); Pérez-Juste (2012), declaran que a través de técnicas estadísticas, los investigadores pueden generalizar los resultados obtenidos a partir de muestras específicas, controlando el sesgo y maximizando la validez de sus conclusiones. Esta sinergia entre la estadística y la experimentación científica es fundamental en el ámbito de las ciencias sociales, donde la complejidad de los fenómenos estudiados demanda un análisis de datos sofisticado. De esta manera, la estadística no solo es una herramienta auxiliar, sino un componente esencial en el proceso de construcción del conocimiento científico social, desde la concepción de un estudio hasta la interpretación de sus resultados.

Calderón y Ledesma (2012); Barreto-Villanueva (2012), igualmente reseñan que desde la perspectiva científica, la evolución de las Ciencias Sociales y las Ciencias Humanas a la categoría científica, sobre todo en las últimas décadas ha estado ligada a la estadística, debido a la necesidad de cuantificar y analizar fenómenos sociales, cuyas variables requieren el establecimiento de relaciones causales; lo que ha impulsado la interdisciplinariedad en la dicha perspectiva científica, permitiendo a investigadores de diversas áreas emplear herramientas estadísticas para medir la

magnitud, frecuencia e influencia de dichas variables sociales. Desde la sociología hasta la psicología, pasando por la economía, la antropología y la ciencia política, la estadística se ha convertido en un lenguaje común para describir y modelar la realidad social y generar conocimientos sólidos, confiables y veraces.

En resumen Pérez-Tejada (2008), Pérez-Juste (2012); Barreto Villanueva (2012), para ratificar la importancia de la estadística como factor determinante de la cientificidad de las disciplina, aseguran que desde la perspectiva científica el valor de una teoría se evalúa en función de su capacidad explicativa y predictiva, siendo las teorías más consistentes aquellas que explican un mayor número de fenómenos, tienen menos excepciones y, sobre todo, generan predicciones precisas y novedosas que pueden ser sometidas, en repetidas ocasiones a pruebas empíricas rigurosas; y en cuanto a la cuantificación de las variables y el uso de modelos estadísticos, permiten a diversas ciencias, formalizar predicciones más precisas y evaluar el grado de incertidumbre asociado a esas predicciones, posibilitando la determinación de las discrepancias entre las predicciones teóricas y los datos observados, y si dichas discrepancias se deben a errores de medición o a limitaciones de la propia teoría; y toda esta dinámica la sostiene y robustece el andamio estadístico.

CAPÍTULO II

APROXIMACIONES TEÓRICAS DE LA ESTADÍSTICA

2.1. Definición y objetivos

Para Carrasco (s.f), la investigación científica, entendida como un proceso sistemático y herramienta de exploración, descripción, comprensión, explicación y hasta transformación de los de los fenómenos sociales, mediante la formulación de objetivos claros y precisos, delimitaciones claras del ámbito de estudio y diseño de un plan de indagación, puede llegar a dar respuesta a interrogantes específicos. En el caso de la realidad social, es común que los problemas de investigación no se presenten de manera aislada, por lo que el investigador debe a través de una observación cuidadosa y una revisión exhaustiva de la literatura existente, identificar, delimitar y reformular los problemas de investigación, de manera tal, que sean susceptibles de ser abordados empíricamente. Una vez definido el problema de investigación, el investigador selecciona, en función del marco teórico y de los objetivos planteados, la metodología más

apropiada y las técnicas de recolección de datos que le permitan obtener información relevante y confiable.

El análisis de los datos obtenidos permitirá al investigador explorar, comprender e interpretar los fenómenos estudiados, generando así nuevo conocimiento sobre la realidad social. En el caso de las ciencias exactas o naturales las técnicas o metodologías estadísticas, en su papel de herramientas matemáticas para el análisis causal, han experimentado una transformación significativa a lo largo del tiempo, su desarrollo histórico revela cómo estas herramientas se han adaptado y sofisticado para responder a las demandas de investigaciones cada vez más complejas. Comprender esta evolución es esencial para utilizar estas técnicas de manera efectiva y rigurosa, y en ambos campos de estudios el papel de las aproximaciones teóricas de las estadística juegan y accionan sus mejores contribuciones (Sagaró del Campo y Zamora, 2019).

López y Diez (2017), refieren que la estadística como andamio en el análisis de datos en diversas disciplinas científicas, se sustenta en una variedad de aproximaciones teóricas que guían su aplicación e interpretación, adecuándose a los datos y a los problemas en cuestión, es decir, las aproximaciones teóricas lejos de ser estáticas, evolucionan constantemente en respuesta a los desafíos y avances metodológicos.

Box (1976); Fisher (1935) como se citaron en Lizasoain (2024), definen a las aproximaciones teóricas en estadística como un conjunto de herramientas y marcos conceptuales que proporcionar métodos de análisis de datos, influyen en la conceptualización de los fenómenos estudiados y median en el diseño de las investigaciones, cuya selección particular está condicionada por la naturaleza del problema de investigación, las características de los datos disponibles y los objetivos del análisis. En este sentido, las aproximaciones teóricas actúan como lentes a través de las cuales los investigadores observan la realidad y construyen conocimiento, orientando así el proceso de investigación desde su concepción hasta su interpretación final.

Phillips y Pugh (2011), relatan que en el ámbito de la investigación estadística, la construcción de modelos suele priorizar la explicación de las relaciones observadas entre variables cuantificables, esta orientación pragmática se justifica por la necesidad de obtener resultados aplicables en contextos prácticos, donde la medición directa es más accesible. En consecuencia, los modelos estadísticos tienden a centrarse en la dimensión empírica de los fenómenos, dejando en segundo plano una exploración exhaustiva de sus fundamentos teóricos y filosóficos. Para estos autores, las aproximaciones teóricas en estadística son enfoques que priorizan la construcción de modelos matemáticos y estadísticos diseñados para describir y explicar las relaciones entre variables observables, en lugar de buscar una representación completa y filosóficamente rigurosa de la realidad. Estas aproximaciones se caracterizan por su énfasis en la utilidad práctica, la cuantificación y la capacidad de generar predicciones basadas en datos empíricos, muy apropiadas con este énfasis para las ciencias exactas y la economía.

Por su parte, Lizasoain (2024), manifiesta que las aproximaciones teóricas en estadística se refieren al grado en que un modelo estadístico busca explicar los fenómenos subyacentes a los datos, más allá de simplemente describir las relaciones observadas entre variables. Mientras que las aproximaciones empíricas se centran en la cuantificación y modelización de los datos disponibles, las aproximaciones teóricas aspiran a construir modelos que reflejen una comprensión más profunda de los procesos generadores de datos, basándose en teorías y principios científicos. En otras palabras, las aproximaciones teóricas buscan establecer un puente entre los datos observados y los mecanismos causales que subyacen a ellos, muy apropiado este lente para las ciencias sociales.

Las aproximaciones teóricas, tienen en ambas campos de investigación, variados objetivo o finalidades, que según Kunh (1962); Bunge (1983); Becker (1998); como fueron referidos por Phillips y Pugh (2011), son los siguientes:

1) Modelizar fenómenos sociales, ya que permite representar de manera simplificada y cuantitativa la complejidad de los fenómenos sociales. A través de la identificación y selección de variables relevantes, los investigadores construyen modelos teóricos que buscan capturar los patrones, relaciones y tendencias subyacentes en los datos. Es decisivo reconocer que estos modelos son representaciones idealizadas de la realidad, y como tales, están sujetos a limitaciones inherentes.

2) Comparar el modelo y realidad, una vez construido el modelo, se procede a compararlo con los datos empíricos. Las discrepancias entre el modelo y la realidad permiten identificar las variables o factores que no han sido considerados en la construcción del modelo, lo que a su vez sugiere áreas para futuras investigaciones y refinamientos o formulación de modelos más precisos.

3) Realizar inferencias, La estadística inferencial tiene como objetivo primordial generalizar los hallazgos obtenidos a partir del análisis de una muestra representativa, con el fin de inferir características y tendencias en una población completa. A través de diversas técnicas estadísticas, los investigadores buscan estimar parámetros poblacionales con un determinado grado de confianza, lo que permite realizar afirmaciones sólidas acerca de fenómenos más amplios.

4) Contrastar hipótesis, el contraste de hipótesis proporciona un marco riguroso para la evaluación de teorías, al someter las hipótesis a contrastación empírica, los investigadores buscan reducir la incertidumbre y aumentar la confianza en las conclusiones obtenidas. Este proceso iterativo de formulación de hipótesis, diseño de experimentos y análisis de datos es fundamental para el avance del conocimiento científico.

5) Identificar las variables, relevantes o significativas para el fenómeno en estudio, mediante la observación cuidadosa, con el fin de reducir la complejidad del análisis y facilitar la construcción del modelo.

6) Otras varias: Descubrir patrones ocultos y relaciones significativas entre variables, cuantificando la incertidumbre inherente a los procesos

de medición y observación. Al evaluar el grado de confianza en los resultados obtenidos, estas aproximaciones permiten realizar inferencias sólidas y objetivas sobre la población de estudio. Además, al construir modelos estadísticos que capturan las tendencias y relaciones subyacentes en los datos, las aproximaciones teóricas facilitan la predicción de futuros eventos o valores, lo que resulta invaluable en diversos campos como la economía, la ciencia y la ingeniería.

Es importante dejar claro, lo que refieren Pooper (1962): Merton (1942), como se citaron en Phillips y Pugh (2011), estos autores sintetizan que el objetivo primordial de las aproximaciones teóricos de la estadísticas en las ciencias sociales es facilitar la toma de decisiones fundamentadas, ya que al cuantificar riesgos y beneficios, así como evaluar rigurosamente la evidencia empírica, estos modelos ofrecen una base sólida para la elección de alternativas. Los modelos en ciencias sociales, a diferencia de sus contrapartes en física y economía suelen adoptar la forma de teorías de alcance medio, que establecen conclusiones a través de una deducción lógica rigurosa.

Estas conclusiones son sometidas a una comparación exhaustiva con otros enunciados teóricos, con el fin de identificar las relaciones lógicas subyacentes, tal como lo plantea Popper. En este sentido, los modelos permiten extraer una mayor cantidad de información de los datos disponibles en comparación con los métodos estadísticos tradicionales, al posibilitar la identificación de patrones y relaciones más sutiles que escapan a los análisis comparativos simples entre grupos experimentales y de control (Popper, 1962, como se citó en Phillips y Pugh, 2011).

2.2. Estadística deductiva o descriptiva

La palabra deducir, detona un razonamiento inferido a partir de un principio general que conlleva a sacar una conclusión sobre algo, dicha conclusión representar o detalla los aspectos importantes o relevantes por medio del lenguaje. La estadística descriptiva es una rama de la disciplina

que se encarga de recolectar, organizar, analizar, clasificar y presentar datos de interés de una manera clara y concisa, en una muestra, en este caso poblacional, con el objetivo de describir las características a partir de un conjunto de datos, sin hacer generalizaciones más allá de esos datos. En otras palabras, es la realización de una deducción (razonamiento) a partir de una data informativa (principio general), que conlleva la descripción de un fenómeno o hecho solo con base a dicha data (González y Pérez de Vargas, 2009).

Calderón y Ledesma (2012), indican que la estadística descriptiva abarca cualquier tratamiento de datos numéricos que involucre generalizaciones, esta disciplina agrupa diversas técnicas destinadas al manejo y procesamiento de un compendio de datos, con la finalidad de caracterizar múltiples datos numéricos. La estadística descriptiva busca resaltar las propiedades de estos conjuntos, lo cual puede lograrse a través de métodos gráficos o analíticos.

La connotación estadística deductiva es menos común y, a menudo, se utiliza como sinónimo de estadística descriptiva. Sin embargo, dicha connotación a veces es empleada para referir un enfoque de la estadística descriptiva más amplio, es decir que incluye no solo la descripción de los datos, sino también la identificación de patrones y tendencias, pero en líneas generales es la rama de la estadística constituida por el conjunto de instrumentos y temas correspondidos con la descripción de selecciones de observaciones estadísticas, referidos bien a una muestra poblacional determinada o a la totalidad de la misma (González y Pérez de Vargas, 2009) (Ver Figura 1)

Phillips y Pugh (2011), refieren que las metodologías de la estadística descriptiva se han consolidado en todas las áreas de conocimientos, tanto sociales como naturales, como una instrumentación apreciable que aportan grandes ventajas en cuanto a los niveles de confianza y eficacia en la investigación.

Figura 1

Gráfica radial de demanda global por parte de las instituciones educativas



Nota: Tomado de Buján (2017).

En el ámbito social, su aplicación es especialmente beneficiosa para la descripción de poblaciones, el análisis de encuestas, la comparación de grupos y la exploración de relaciones entre variables; por ejemplo, un sociólogo podría emplear estas técnicas para caracterizar la distribución de edades en una ciudad, comparar los ingresos de distintos grupos sociales, analizar las actitudes hacia cuestiones políticas o fenómenos psicosociales como la ansiedad generalizada y la depresión.

En el caso de las ciencias naturales y exactas, la estadística descriptiva es básica para resumir datos experimentales, describir propiedades de materiales y analizar resultados de experimentos; Por ejemplo, un biólogo podría utilizarla para describir la distribución de tamaños de una población de insectos, mientras que un físico podría usarla para analizar los resultados de una medición de temperatura (Pérez-Tejada, 2008).

En su obra "Estadística Aplicada: Una Visión Instrumental", María Teresa González Manteiga y Alberto Pérez de Vargas (2009), presentan dos enfoques claramente identificables de esta categoría de la estadística, los mis-

mos son: La estadística descriptiva unidimensional y bidimensional, resaltando sus respectivas finalidades. La estadística unidimensional se centra en el análisis de un único conjunto de datos, permitiendo una caracterización detallada de variables individuales a través de medidas de tendencia central, dispersión y distribución y su principal objetivo es ofrecer una visión integral del comportamiento de una variable específica, facilitando la identificación de patrones y características fundamentales dentro de la población estudiada.

Un estudio de la estadística descriptiva unidimensional en el área social, sería una encuesta para evaluar la satisfacción de los estudiantes en una universidad sobre la comunicación asertiva de los profesores, la metodología aplicada seria la recopilación de datos de una sola variable, como la satisfacción general en una escala del 1 al 5. Se calcula la media, la mediana y la moda de las respuestas, así como dispersión (desviación estándar) para entender el nivel de satisfacción general de los estudiantes (González y Pérez de Vargas, 2009).

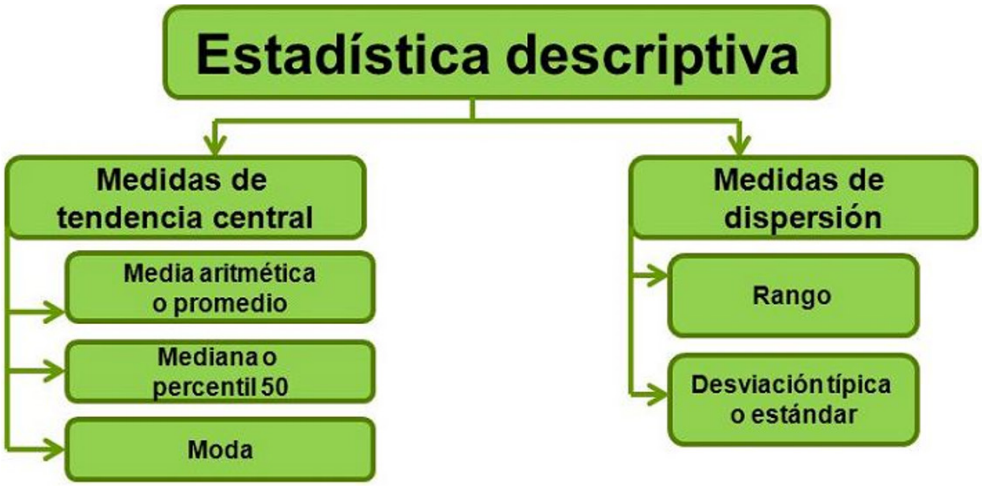
Por otro lado, la estadística descriptiva bidimensional aborda el análisis simultáneo de dos variables, permitiendo explorar las relaciones y correlaciones existentes entre ellas, este enfoque es de gran utilidad para comprender fenómenos complejos donde múltiples factores interactúan, siendo su objetivo principal develar patrones de asociación, proporcionando así un entendimiento más profundo de la dinámica entre las variables analizadas (González y Pérez de Vargas, 2009).

Un ejemplo en el ámbito social de esta metodología de investigación, sería el estudio de la relación entre los ingresos de una población y su nivel educativo en una determinada región. Se valorarían dos variables: “nivel educativo” (dividido en categorías como secundaria, pregrado, posgrado) e “ingresos” (en rangos). Para dicha evaluación se pueden crear tablas cruzadas para mostrar cómo varían los ingresos promedio según el nivel educativo; gráficos de dispersión o diagramas de caja pueden ser utiliza-

dos para visualizar la relación entre ambas variables (González y Pérez de Vargas, 2009) (Ver Figura 2).

Figura 2

Medidas o parámetros de la estadística descriptiva



Nota: Tomado de Buján (2017).

Ambos ejemplos ilustran cómo se pueden utilizar técnicas descriptivas unidimensionales y bidimensionales, para obtener información útil en contextos sociales, facilitando la toma de decisiones basada en datos, enriqueciendo el análisis estadístico y refrendando la cientificidad de las ciencias sociales.

2.3. Estadística inductiva o inferencial

La inferencia, un proceso cognitivo fundamental, consiste en la derivación de conclusiones a partir de premisas o hipótesis establecidas. Este mecanismo, objeto de estudio de la lógica, permite establecer relaciones causales y establecer conocimiento nuevo a partir de información existente. La validez de una inferencia no depende del contenido especí-

fico de las premisas, sino de la estructura lógica subyacente. Es decir, una inferencia es válida cuando la conclusión se sigue necesariamente de las premisas, independientemente del tema al que se refiera. La lógica, como disciplina, se encarga de analizar y formalizar estas estructuras, permitiendo identificar inferencias válidas y descartar aquellas que no lo son. De este modo, la inferencia se convierte en una herramienta esencial para la construcción de argumentos sólidos y la evaluación crítica de la información (Levin y Rubin, 2019).

Partiendo de lo anterior, Calderón y Ledesma (2012), sostienen que la estadística inferencial es el marco conceptual que sustenta la toma de decisiones en contextos de incertidumbre. Al generalizar, predecir o estimar a partir de datos, las técnicas estadísticas empleadas permiten extraer conclusiones sólidas y objetivas sobre el fenómeno en estudio, proporcionando una base sólida para la acción, ya que estas técnicas proporcionan las herramientas necesarias para ir más allá de la simple descripción de los datos, permitiendo comprender patrones y tendencias subyacentes basadas en la evidencia empírica.

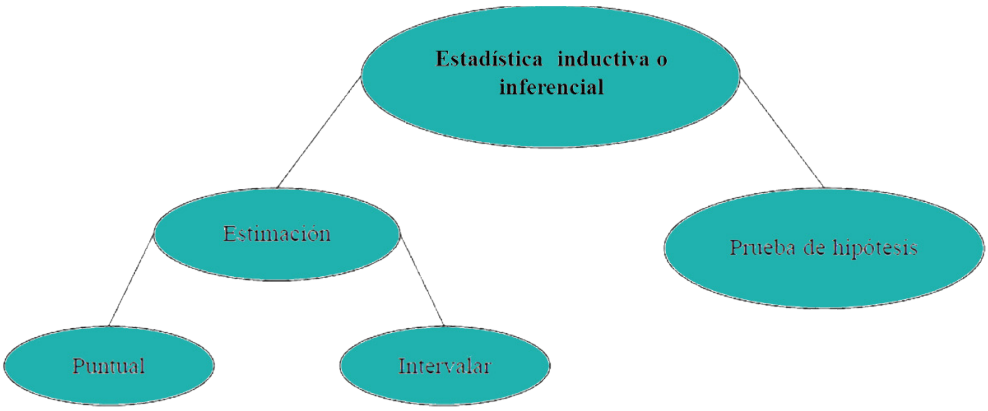
Calduch (s.f.), de acuerdo con Calderón y Ledesma (2012), enfatiza que la estadística inductiva o inferencial, es una herramienta que juega un papel crucial en la generalización de resultados, incluso ante la imposibilidad de estudiar a todos los individuos de una población, es decir a partir de muestras, ya que al emplear la teoría de probabilidades, esta disciplina permite realizar inferencias sólidas sobre poblaciones, siendo el análisis de varianza una técnica destacada para comparar múltiples grupos y determinar si existen diferencias significativas entre ellos.

La estadística inductiva o inferencial según Pérez-Tejada (2008), ofrece dos enfoques principales para lograr generalizar datos relacionados con los parámetros de una población, estos son: la estimación y la contrastación de hipótesis. La estimación, por su parte, se divide en dos tipos: puntual e intervalar:

- La estimación puntual, consiste en asignar un único valor a un parámetro poblacional desconocido, basándose en el valor correspondiente de un estadístico calculado a partir de la muestra. Sin embargo, es importante reconocer que toda estimación puntual conlleva un margen de error inherente, dado que es prácticamente imposible obtener un valor exacto para el parámetro poblacional.
- La estimación intervalar, por otro lado, proporciona un rango de valores dentro del cual se espera que se encuentre el verdadero valor del parámetro poblacional con una determinada probabilidad. Este rango, conocido como intervalo de confianza, permite cuantificar la incertidumbre asociada a la estimación puntual. La necesidad de recurrir a la estimación se fundamenta en la imposibilidad práctica de estudiar a toda la población, puesto que al trabajar con una muestra, se introduce un error de muestreo que refleja la diferencia entre el valor del parámetro poblacional y el valor del estadístico utilizado como estimador. Afortunadamente, la teoría estadística proporciona herramientas para calcular la probabilidad de cometer este tipo de error, lo que permite evaluar la precisión de las estimaciones obtenidas (Pérez-Tejada, 2008).

Figura 3

Enfoques de inferencia en el enfoque clásico estadístico



Nota: Tomado de Pérez-Tejada (2008).

El otro método riguroso para evaluar afirmaciones sobre características poblacionales son las pruebas de hipótesis, este procedimiento comienza con la formulación de una hipótesis nula, que representa la situación que se desea refutar. A continuación, se establece una hipótesis alternativa, que expresa la afirmación que se desea respaldar y con base en la muestra, se calcula un estadístico de prueba, el cual se compara con un valor crítico. Si el estadístico cae en la región de rechazo, se concluye que existe evidencia suficiente para rechazar la hipótesis nula en favor de la alternativa (Pérez-Tejada, 2008) (Ver Figura 3).

En la estadística inductiva o inferencial, los estadísticos son funciones calculadas a partir de los datos muestrales y sirven como estimadores de parámetros poblacionales, dichos parámetros representan características numéricas de la población completa, que tanto en los contrastes de hipótesis como en la estimación por intervalos, la distribución del estadístico muestral es fundamental para realizar inferencias sobre la población. La principal diferencia radica en su objetivo: mientras que los contrastes de hipótesis evalúan una afirmación específica sobre el parámetro (hipótesis nula) y la aceptan o rechazan en función del valor del estadístico, la estimación por intervalo proporciona un rango de valores plausibles para el parámetro con una determinada confianza, que contrarrestar la incertidumbre (Pérez-Tejada, 2008).

CAPÍTULO III

MÉTODOS ESTADÍSTICOS

3.1 Historia

Para Carrasco (s.f.), de manera general los métodos estadísticos y la Estadística como disciplina que se ocupa de la recolección, análisis e interpretación de datos, ha evolucionado significativamente desde sus orígenes. Inicialmente los métodos estadísticos fueron concebidos como un conjunto de herramientas, que permitían describir y cuantificar fenómenos sociales, políticos y económicos propios de los Estados y las sociedades humanas, que debido a la incorporación de la teoría de la probabilidad experimentaron una transformación profunda; es decir, gracias a esta confluencia, se originaron las dos grandes etapas en su desarrollo: La estadística, cuyos métodos estadísticos se caracterizaban por descriptivos, centrada en la organización y resumen de datos, y la estadística cuyos métodos estadísticos se determinaron inferenciales, por permitir realizar generalizaciones y tomar decisiones a partir de muestras.

Barreto-Villanueva (2012), refiere que los orígenes de la aplicación de los métodos estadísticos se remontan a las antiguas civilizaciones, donde se realizaban censos, conteos, registros y gestiones con fines administrativos. Sin embargo, la consolidación de esta metodología y posterior asignación como "Estadística" y disciplina científica se produjo a partir del siglo XVI, con la publicación de Liber de Ludo Aleae de Girolamo Cardano, obra pionera en el estudio de los juegos de azar y la probabilidad. Este hecho conllevó a la metodología estadística a ser un compendio de temas concernientes al Estado, la Economía y el tratamiento de asuntos de la administración pública de los países de la época.

A lo largo de los siglos XVII y XVIII, figuras como Gottfried Achenwall y Abraham de Moivre, contribuyeron a establecer los fundamentos de la estadística, definiendo su objeto de estudio y desarrollando métodos analíticos que estudiaban fenómenos poblacionales, económicos y políticos, donde la Escuela de Göttingen, fundada por Achenwall, y la Escuela de Aritmética Política de Londres desempeñaron un papel crucial en la difusión y consolidación de dichos métodos estadísticos. Paralelamente, en Inglaterra, por una parte, la Escuela de Aritmética-Política se dedicaba a estudios demográficos y actuariales, sentando las bases para el desarrollo de los métodos estadísticos aplicados; mientras que, por otro lado, Abraham de Moivre, matemático francés exiliado en Inglaterra, realizó contribuciones fundamentales a la estadística demográfica y a la teoría de la probabilidad, estableciendo los cimientos para los avances posteriores en esta área (Barreto-Villanueva, 2012).

La irrupción de la Revolución Industrial, según Carrasco (s.f.), demandó un nuevo paradigma en la gestión de la producción, por la necesidad de controlar la calidad y optimizar los procesos productivos, impulsando la sofisticación de las herramientas o metodologías estadísticas. En este contexto, los métodos estadísticos dejaron de ser una mera herramienta descriptiva para convertirse en un sostén primordial en la toma de decisiones empresariales, al contar con información segura y constante tanto productiva como comercial.

Figuras como Karl y Egon Pearson, Gossett, Neyman y, especialmente, Ronald Fisher, fueron pioneros en el desarrollo de métodos estadísticos más rigurosos y sofisticados. Fisher, en particular, realizó contribuciones seminales en diversos campos, desde la genética hasta la agronomía, consolidando su posición como uno de los padres de la estadística moderna, sus trabajos sentaron las bases para el desarrollo de una disciplina que trascendió los límites de la industria, encontrando aplicaciones en prácticamente todas las áreas del conocimiento. La obra de Kolmogorov, publicada en 1933, aportó una sólida base teórica probabilística y matemática, enriqueciendo aún más el arsenal metodológico de la estadística. En conjunto, estos avances consolidaron la aplicación de métodos estadísticos, como mecanismos científicos, rigurosos y versátiles, capaces de abordar problemas complejos y proporcionar soluciones basadas en datos (Barreto-Villanueva, 2012).

Para Barreto-Villanueva (2012); Carrasco (s.f.); Zagaró del Campo y Zamora (2019), la comprensión y complejidad del mundo, provocaron la evolución de los métodos estadísticos, y técnicas como: la teoría de juegos, el análisis de series temporales y el diseño de muestras, surgidas de la necesidad de tomar decisiones informadas en contextos cada vez más heterogéneos y cambiantes, han sido relevantes en este desarrollo. El siglo XX marcó un hito en la historia de la metodología estadística, gracias a avances como el cálculo de probabilidades y las pruebas de hipótesis, que permitieron cuantificar la incertidumbre y validar conclusiones científicas, la segunda mitad del siglo, con el advenimiento de las computadoras, aceleró exponencialmente la capacidad de análisis de datos, dando lugar a nuevos métodos como el análisis de varianza (ANOVA) y el control estadístico de procesos.

En la actualidad, la estadística se encuentra en la convergencia de diversas disciplinas, como la informática y la matemática, dando origen a campos como la minería de datos y el aprendizaje automático. Las redes sociales, con su inmenso volumen de información, son un claro ejemplo de cómo la metodología estadística se aplica para ajustar comportamientos, perso-

nalizar experiencias y optimizar procesos; La investigación operativa, con sus raíces en la estadística, juega un papel crucial en la optimización de sistemas complejos, desde la logística hasta la gestión de recursos (Barreto-Villanueva, 2012; Meneses, 2019).

En síntesis, la historia de la metodología estadística es una historia de adaptación y crecimiento, impulsada por la necesidad de comprender y dar sentido a un mundo cada vez más complejo. Desde sus orígenes como herramienta para la toma de decisiones en negocios y gobierno, la estadística se ha extendido a prácticamente todas las áreas del conocimiento, convirtiéndose en una disciplina esencial para el desarrollo científico y tecnológico (Zagaró del Campo y Zamora, 2019).

3.2 Clasificación según sus fines

3.2.1. Exploratorios

Para Calderón y Ledesma (2012) y Salgado-Ugarte (2013), el Análisis Exploratorio de Datos (AED, por sus siglas en inglés), es un método estadístico que permite a los investigadores adentrarse en un conjunto de datos de manera similar a un detective, sus raíces se encuentran en la estadística descriptiva de mediados del siglo XX, revitalizada por el movimiento del Data Analysis en los años 80, pero desarrollada principalmente por John W. Tukey, la cual revolucionó la manera en que se abordan los datos en la investigación científica, gracias a su enfoque inductivo, que prioriza la exploración de los datos sin imponer restricciones previas.

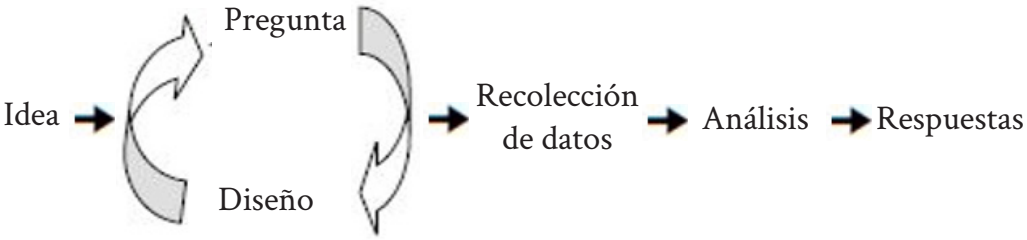
Esta metodología contrasta con los métodos estadísticos confirmatorios más tradicionales, ya que, mediante una variedad de técnicas gráficas y numéricas, el AED permite a los investigadores: detectar y corregir errores en la recopilación de datos, identificar patrones y tendencias subyacentes, evaluar la adecuación de las distribuciones a modelos teóricos (como la normalidad y la homogeneidad de varianzas) y reconocer valores atípicos. Al proporcionar una visión integral y flexible de los datos, el AED facilita

la selección de modelos estadísticos apropiados y la generación de hipótesis más sólidas (Salgado-Ugarte, 2013).

Parra (2002), es una metodología que emplea técnicas estadísticas descriptivas y visualizaciones gráficas, que como enfoque de datos conlleva a identificar valores atípicos, distribuciones no normales, y relaciones entre variables que podrían pasar desapercibidas en un análisis numérico. Además el proceso de investigación es iterativo, es decir, permite generar hipótesis y preguntas de investigación que serán exploradas en profundidad en etapas posteriores del análisis. El AED no solo sirve para limpiar y preparar los datos, sino también construye una comprensión intuitiva de los mismos (Salgado y Ugarte, 2013) (Ver Figura 4).

Figura 4

Esquema del paradigma Análisis Exploratorio de Datos (AED)



Nota: Adaptado de Parra (2002).

Entre las técnicas del AED más comunes se encuentran los diagramas de caja y bigotes, que ofrecen una representación visual concisa de la distribución de una variable, destacando valores atípicos; los diagramas de tallo y hojas, por su parte, proporcionan una vista detallada de la distribución, combinando elementos de un histograma y una tabla de frecuencias; y finalmente, los diagramas de dispersión son fundamentales para visualizar la relación entre dos variables, permitiendo identificar patrones de asociación, correlación y tendencias (Anderson et al., 1999 como se citaron en Salgado y Ugarte, 2013).

Un ejemplo de la aplicación de esta metodología estadística en las ciencias sociales, podría ser: La realización de un estudio sobre el impacto de las redes sociales en la participación política, el investigador podría utilizar un diagrama de barras para comparar la frecuencia de uso de diferentes plataformas de redes sociales entre grupos de edad. Asimismo, podría emplear un diagrama de dispersión para explorar la relación entre el tiempo dedicado a las redes sociales y el nivel de participación en actividades políticas; y a través de estas visualizaciones, el investigador podría identificar patrones interesantes, como una mayor propensión a utilizar ciertas plataformas entre los jóvenes y una correlación positiva entre el uso intensivo de las redes sociales y la participación en debates en línea (Salgado y Ugarte, 2013).

3.2.2. Confirmatorios

Parra (2002), refiere que el análisis confirmatorio de datos es un método estadístico ampliamente utilizado en las ciencias sociales, basado en la comprobación de hipótesis preestablecidas y caracterizado por el empleo de instrumentación de ambas tipologías de la estadística descriptiva (como la media y la varianza) e inferenciales (pruebas de hipótesis, correlación y regresión) para evaluar la validez de un modelo teórico. Debido a su naturaleza estructurada y a la disponibilidad de software estadístico especializado (como SPSS, SAS, SPAD y STATGRAPHICS), el análisis confirmatorio se ha convertido en el método de análisis de datos más comúnmente enseñado y aplicado en esta disciplina.

Este enfoque ha sido de mucha popularidad y utilidad, a pesar de sus limitantes y críticas, siendo una de las principales su tendencia a reducir los datos a un conjunto limitado de indicadores estadísticos, lo que puede llevar a una pérdida de información valiosa. Además, este método al centrarse en la confirmación de hipótesis preexistentes, puede descuidar la exploración de patrones y relaciones inesperadas que podrían surgir de los datos. Por otro lado, la facilidad con la que se pueden realizar análisis confirmatorios utilizando software estadístico ha contribuido a una cierta

mecanización del proceso, lo que puede llevar a una aplicación poco reflexiva de las técnicas estadísticas (Parra, 2002).

Hoaglin et al. (1983), como fueron citado por Salgado y Ugarte (2013), mencionan que los métodos estadísticos tradicionales, son métodos confirmatorios ya que son medios diseñados para confirmar hipótesis establecidas con anterioridad, que suelen basarse en supuestos estrictos sobre la distribución de los datos, pero en la práctica, los datos reales rara vez cumplen con estas condiciones ideales. La incapacidad de los métodos clásicos para adaptarse a datos no normales o con valores atípicos limita su aplicabilidad en muchas situaciones reales, y esta limitación ha impulsado el desarrollo de métodos estadísticos robustos y resistentes, capaces de producir resultados confiables incluso cuando los supuestos tradicionales no se cumplen.

Salgado y Ugarte (2013), refiere que Los métodos confirmatorios parten de un modelo teórico específico (a menudo lineal) que establece una relación entre variables; A través de pruebas de significancia estadística, se evalúa si los datos observados respaldan dicha relación, descartando la posibilidad de que los resultados se deban al azar. Esta rigidez metodológica presenta restricciones debido a que su enfoque es deductivo y no exploratorio, limitando así la capacidad de descubrir hallazgos inesperados que podrían enriquecer el conocimiento sobre el fenómeno en estudio.

Pérez-Tejada (2008); Levin y Rubin (2019), muestran que el enfoque confirmatorio en estadística, utilizado por las ciencias sociales, se sustenta en un conjunto de técnicas analíticas diseñadas para probar hipótesis específicas y modelar relaciones entre variables. Entre las más comunes destacan:

a) La regresión lineal, que permite cuantificar la asociación entre una variable dependiente y una o más independientes, como por ejemplo: estudiar cómo afecta el nivel educativo de las mujeres a su participación en la fuerza laboral;

b) El análisis de varianza (ANOVA), empleado para comparar medias entre diferentes grupos, tal como evaluar si existen diferencias significativas en la satisfacción laboral entre mujeres que trabajan a tiempo completo, medio tiempo y desde casa;

c) Los modelos estructurales de ecuaciones, que exploran relaciones causales complejas, como determinar el impacto de la violencia de género en la salud mental de las mujeres; y

d) Las pruebas de hipótesis, que permiten determinar si los resultados obtenidos son estadísticamente significativos o si se deben al azar, como al contrastar si la proporción de votantes por un determinado partido político difiere significativamente entre hombres y mujeres. Estas técnicas, al ser aplicadas en las ciencias sociales, contribuyen a generar conocimiento riguroso y a validar teorías sobre fenómenos sociales complicados.

Tabla 1

Cuadro comparativo de métodos exploratorio y confirmatorios

Características	Método	
	Exploratorio	Confirmativo
Objetivo	Descubrir patrones	Confirmar hipótesis
Hipótesis	No hay hipótesis previas	Parte de una hipótesis
Flexibilidad	Alta flexibilidad	Menor flexibilidad
Inferencia	Inductiva	Deductiva

NOTA: Tomado de Salgado y Ugarte (2013).

Para realizar un análisis confirmatorio de datos eficaz, que logre un descubrimiento de nuevos conocimientos y comprenda fenómenos sociales, es indefensable complementarlo con otras técnicas de análisis de datos, como el análisis exploratorio, que permiten una indagación más abierta y flexible de los datos (Parra, 2002) (Ver Tabla 1).

3.3. Clasificación según sus presupuestos teóricos

3.3.1. Paramétricos

La Federación Científico-Europea de Osteópatas (SEFO) (2019), define la estadística paramétrica como una rama primordial de la inferencia estadística, que centra su funcionalidad en procedimientos analíticos que presuponen una distribución teórica conocida para los datos; esta distribución, generalmente representada por una función matemática, se caracteriza por un conjunto finito de parámetros numéricos. Por ejemplo, si se asume que las alturas de una población siguen una distribución normal, los parámetros desconocidos, como la media y la desviación estándar, son los objetivos principales de la inferencia paramétrica y la elección de utilizar este enfoque requiere un conocimiento previo o una fuerte evidencia de que los datos se ajustan a una distribución específica. En caso contrario, se recurre a métodos no paramétricos, los cuales no hacen suposiciones tan restrictivas sobre la distribución de los datos.

Al respecto, Bautista-Díaz et al., 2020, refiere con más precisión que los métodos paramétricos requieren que la variable de interés se mida en una escala de intervalo o de razón, lo que permite realizar cálculos matemáticos como la media y la desviación estándar. Un ejemplo de escala de intervalo es la edad, donde las diferencias entre los valores tienen un significado numérico (por ejemplo, la diferencia entre 20 y 30 años es la misma que entre 30 y 40 años).

Otro ejemplo de la aplicabilidad de esta metodología en el área social, sería la ejecución de un estudio que busca determinar si existe una relación

entre la edad laboral de los empleados y su nivel de satisfacción laboral en una determinada empresa. Para llevar a cabo este estudio, se podría recopilar datos sobre la edad y el nivel de satisfacción de una muestra representativa de empleados, aplicando la escala medición, tanto la edad como la satisfacción laboral (medida, por ejemplo, en una escala de Likert de 1 a 5) son variables cuantitativas.

La edad se mide en una escala de razón, lo que permite realizar operaciones matemáticas como calcular la media y la desviación estándar, y en cuanto a la satisfacción laboral, aunque ordinal, puede tratarse como una variable cuasi-continua si se cumplen ciertas condiciones, y tener resultados satisfactorios de la aplicabilidad del método. Más sin embargo, si se buscara analizar si existe una diferencia en la satisfacción laboral entre hombres y mujeres, el género sería una variable categórica nominal, que no tiene un orden inherente y no podemos realizar operaciones matemáticas como calcular una media, en este caso analizar este tipo de variable, requiere la utilización de pruebas estadísticas no paramétricas como la prueba de Mann-Whitney U (González y Pérez de Vargas, 2009).

Las técnicas estadísticas paramétricas más comunes aplicadas en las ciencias sociales, según González y Pérez de Vargas (2009); Bautista-Díaz et al. (2020) y la SEFO (2019), son:

- Análisis de regresión: Esta técnica permite modelar la relación entre una variable dependiente y una o más variables independientes, siendo ideal para predecir resultados y evaluar la influencia de diferentes factores en un fenómeno social.
- ANOVA (Análisis de la Varianza): Se emplea para comparar las medias de tres o más grupos y es útil para determinar si existen diferencias significativas entre diferentes categorías o tratamientos.
- T de Student: Esta prueba se utiliza para comparar las medias de dos grupos, es una de las pruebas más básicas y ampliamente utilizadas en la investigación social.

- Correlación: Mide la fuerza y dirección de la relación lineal entre dos variables continuas, permitiendo identificar si existe una asociación entre ellas.
- Análisis factorial: Reduce un gran número de variables a un conjunto más pequeño de factores subyacentes, con gran utilidad para identificar patrones y estructuras latentes en los datos.

En definitiva, la SEFO (2019), refiere que la aplicabilidad de los métodos paramétricos está limitada por los tipos de datos que pueden analizar, ya que esta metodología presupone que los datos se ajustan a una distribución de probabilidad conocida, como la normal, y que la variable de interés se mide en una escala al menos de intervalo, variables categóricas nominales, como: género; color de ojos, adultez, juventud no cumplen con estos requisitos, ya que los métodos paramétricos son adecuados a datos que tengan una estructura numérica específica.

3.3.2. No paramétricos

Para Levin y Rubin (2019), la aplicación de pruebas estadísticas paramétricas, que presuponen una distribución normal de los datos, se encuentra limitada en situaciones donde esta suposición no se cumple. Aunque las pruebas de bondad de ajuste puedan sugerir una distribución aproximadamente normal, la incertidumbre inherente a estas pruebas hace que esta conclusión sea siempre provisional.

En estos casos, donde las muestras pueden ser pequeñas o provenir de poblaciones con distribuciones desconocidas, las pruebas no paramétricas ofrecen una alternativa. A diferencia de las pruebas paramétricas, las no paramétricas no hacen suposiciones restrictivas sobre la forma de la distribución de la población, lo que las convierte en herramientas valiosas para un amplio rango de aplicaciones, puesto que en lugar de centrarse en parámetros específicos de la población, las hipótesis de las pruebas no paramétricas se formulan en términos de características más generales de la distribución, como la mediana o la forma general de la curva.

Por su parte, para la SEFO (2019), las pruebas no paramétricas se caracterizan por su versatilidad y robustez, ya que cuando las distribuciones normales o paramétricas tradicionales no son aplicables, y existen situaciones en las que es posible realizar inferencias estadísticas sólidas, esta metodología basadas en rangos o frecuencias, es más robusta ante violaciones de los supuestos de normalidad y homogeneidad de varianzas; y dicha robustez las convierte en una opción preferible cuando se trabaja con muestras pequeñas, datos ordinales o distribuciones no normales, permitiendo obtener resultados confiables incluso en condiciones menos ideales.

Levin y Rubin (2019), realizan una caracterización de las ventajas y desventajas de los métodos no paramétricos, siendo las ventajas más resaltantes, las siguientes:

- En primer lugar, estos métodos son distribucionalmente libres, porque como se ha mencionado, exigen supuestos sobre la distribución de la población, como la normalidad;
- En segundo lugar, su aplicación suele ser más sencilla y intuitiva, ya que estos métodos, generalmente, evitan cálculos complejos como el cálculo de desviaciones estándar, optando por transformaciones de los datos en rangos o clasificaciones, simplificando notablemente los análisis estadísticos; y Finalmente, los métodos no paramétricos son especialmente útiles cuando los datos son ordinales o cuando la escala de medición es imprecisa.

En cuanto a sus desventajas, las pruebas no paramétricas, a pesar de su utilidad, presentan dos limitaciones principales, según Levin y Rubin (2019), las mismas son:

- En primer lugar, las pruebas no paramétricas, a pesar de su utilidad, implican una pérdida de información, puesto que al transformar los datos originales en rangos, se desechan las magnitudes exactas de las observaciones. Por ejemplo, los valores 13.33 y 1,189.42, aunque muy diferentes en escala, podrían representarse ambos por el rango 5 si am-

bos ocupan la última posición en un conjunto de datos ordenados, esta pérdida de información puede limitar la capacidad de detectar diferencias sutiles entre grupos o de realizar inferencias precisas.

- En segundo lugar, las pruebas no paramétricas suelen ser menos eficientes que sus contrapartes paramétricas cuando los supuestos de normalidad se cumplen. Esto significa que, para alcanzar un nivel de confianza determinado, los intervalos de confianza obtenidos mediante pruebas no paramétricas tienden a ser más amplios. En otras palabras, se requiere una muestra más grande o una diferencia más marcada entre los grupos para rechazar la hipótesis nula; esta pérdida de eficiencia se compensa con la mayor robustez de las pruebas no paramétricas, ya que no dependen de supuestos estrictos sobre la distribución de los datos.

Perez-Tejada (2008); Levin y Rubin (2019); la SEFO (2019); Bautista-Díaz et al. (2020), menciona en el conjunto de herramientas e instrumentos utilizados, por esta metodología, los más utilizados son:

- Prueba de chi-cuadrado: Se emplea para evaluar si existe una diferencia significativa entre las frecuencias observadas y las esperadas en una o más categorías, siendo utilizada para analizar datos nominales o ordinales.
- U de Mann-Whitney: Diseñada para comparar las ubicaciones de dos muestras independientes, es una alternativa no paramétrica a la prueba t de Student.
- T de Wilcoxon: Se utiliza para comparar las ubicaciones de dos muestras relacionadas y es la contraparte no paramétrica de la prueba t para muestras pareadas.
- H de Kruskal-Wallis: Permite comparar las ubicaciones de más de dos muestras independientes; es una extensión de la U de Mann-Whitney para múltiples grupos.

- Prueba de Friedman: Se emplea para comparar las ubicaciones de múltiples muestras relacionadas, y la misma, es la extensión de la T de Wilcoxon para diseños de bloques.

Estas diversas herramientas del compendio de los métodos no paramétricos encuentran una amplia aplicación en las ciencias sociales, ya que los estudios suelen ser ordinales o nominales, y no siempre cumplen con los supuestos de normalidad requeridos para las pruebas paramétricas. Un ejemplo interesante podría ser el estudio de la comunicación asertiva y su impacto en el clima organizacional, se podrían utilizar pruebas no paramétricas como la prueba de Kruskal-Wallis para comparar los niveles de satisfacción laboral entre empleados que han participado en un programa de entrenamiento en comunicación asertiva y aquellos que no. Asimismo, al analizar las tasas de divorcio en parejas adultas post-pandemia, se podrían emplear pruebas como la prueba de Mann-Whitney U para determinar si existen diferencias significativas en las tasas de divorcio entre parejas que experimentaron diferentes niveles de estrés durante la pandemia.

Para Bautista-Díaz et al. (2020), la selección de una prueba estadística, ya sea paramétrica o no paramétrica, es un paso crucial en el análisis de datos, los métodos no paramétricos, al ser menos restrictivos, permiten analizar una amplia gama de datos. Sin embargo, cuando los supuestos paramétricos se cumplen, estos últimos suelen proporcionar resultados más precisos; la Figura 5 ofrece una guía práctica para tomar una decisión informada, considerando las características específicas de los datos a analizar y los objetivos de la investigación.

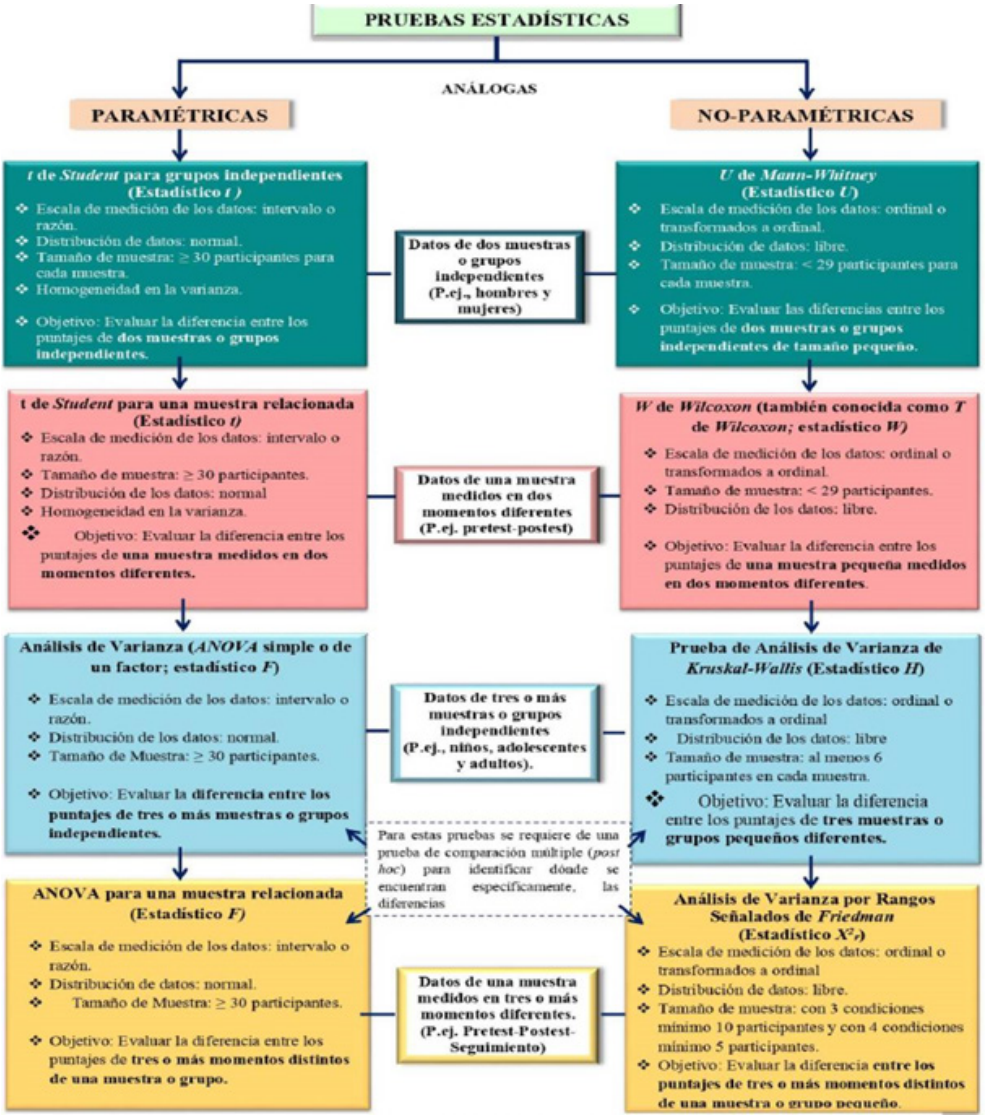
3.4. Clasificación según la clase de variable utilizada

3.4.1. Univariados

Para Sagaró del Campo y Zamora (2019), las raíces de los métodos univariados se hunden en los albores de la civilización, encontrando sus primeras aplicaciones en los censos, estos antiguos recuentos pobla-

Figura 5

Pruebas paramétricas y no paramétricas



Nota: Tomado de Bautista-Díaz et al. (2020). Diagrama- guía de las pruebas paramétricas y no paramétricas, sus objetivos, características y funciones

cionales buscaban describir y profundizar en características básicas de la sociedad, como edad, sexo y ocupación, sentando las bases para el análisis estadístico univariado. Los métodos univariados son dentro del arsenal estadístico, los que permiten a los investigadores explorar y comprender en profundidad una única variable a la vez. Esta rama de la estadística se centra en describir, analizar y sintetizar la información contenida en un conjunto de datos provenientes de una sola característica o atributo (Calderón y Ledesma, 2012).

González y Pérez de Vargas (2009); Trejos (2014), declaran que específicamente en el campo investigativo, los métodos univariados cumplen varias funciones y utilidades, las más importantes son:

a) Exploración de datos, lo que permite al investigador familiarizarse con los datos y detectar posibles problemas o inconsistencias;

b) Resumen de datos, ya que proporcionan un resumen conciso y comprensible de la información contenida en una variable;

c) Realizar inferencia estadística, ya que al profundizar minuciosamente en una variable e identificar patrones, tendencias y valores atípicos se pueden realizar inferencias sobre una población a partir de una muestra.

Los métodos univariados se caracterizan por tener foco en una sola variable, siendo la característica distintiva de estos métodos. Por otra parte la media, la mediana y la moda son las medidas más comunes utilizadas para describir el centro de una distribución de datos, mientras que la varianza, la desviación estándar y el rango describen la variabilidad de los datos alrededor de la medida de tendencia central. Además la visualización de datos, como gráficos, histogramas, diagramas de caja y curvas de densidad, son herramientas esenciales para visualizar la distribución de los datos (González y Pérez de Vargas, 2009; Trejos, 2014).

A partir de lo anterior, por ejemplo, un investigador en ciencias sociales podría emplear métodos univariados para analizar la tasa de suicidios en una población específica a lo largo del tiempo. Al calcular medidas de tendencia central (media, mediana, moda) y de dispersión (desviación estándar, rango), podría identificar patrones y tendencias en los datos. Por otro lado, para estudiar el impacto de la persistencia del COVID-19 en la vida social de los pacientes, el investigador podría utilizar frecuencias y porcentajes para describir la distribución de variables como la soledad percibida, la ansiedad social o la dificultad para mantener relaciones interpersonales.

En el caso de variables cuantitativas, como la edad y el ingreso, los investigadores suelen emplear histogramas para visualizar la frecuencia con la que se presentan diferentes valores. Estos gráficos permiten identificar la forma de la distribución (normal, sesgada, etc.), la presencia de valores atípicos y el rango de variabilidad. Por ejemplo, un histograma de la edad de una población puede revelar si la mayoría de los individuos se concentran en un rango de edad específico o si existe una distribución más uniforme. Para variables cualitativas, como el género, la ocupación o el nivel educativo, los gráficos de barras, de sectores o circulares son los más utilizados, ya que estos gráficos facilitan la comparación de las frecuencias de cada categoría. Un gráfico de barras, por ejemplo, podría mostrar la proporción de personas que se identifican con cada género en una muestra determinada (Trejo, 2014).

Estos análisis univariados permitirían obtener una primera aproximación a la magnitud y naturaleza de variados problemas medibles cuantitativa y cualitativamente, sentando las bases para investigaciones más profundas que consideran la relación entre múltiples variables.

3.4.2. Multivariados

Con respecto a los enfoques metodológicos estadísticos denominados multivariados, Sagaró del Campo y Zamora (2019), relatan que la

segunda mitad del siglo XX marcó un hito en la evolución de la estadística, caracterizado por la adopción de enfoques más sofisticados y la expansión de sus aplicaciones; el surgimiento de los computadores fue un catalizador crucial, permitiendo el análisis de grandes volúmenes de datos y el desarrollo de técnicas multivariadas. Estos métodos, capaces de analizar simultáneamente múltiples variables, revolucionaron disciplinas como la medicina, donde se volvieron indispensables para el estudio de factores de riesgo y la evaluación de tratamientos. Aparte la transición desde modelos deterministas a modelos probabilísticos fue fundamental, ya que la incertidumbre inherente a las relaciones entre variables llevó a la adopción de conceptos como el riesgo relativo y el riesgo atribuible, reemplazando la noción de causalidad directa por la de asociaciones causales.

Por otro lado, la regresión logística, una técnica clave en este contexto, permitió cuantificar el impacto de múltiples factores sobre la probabilidad de un evento, como desarrollar una enfermedad; sus orígenes se remontan a mediados del siglo XX, pero fue con el advenimiento de los computadores que su aplicación se popularizó, gracias a la capacidad de calcular de manera eficiente los estimadores necesarios. La historia de los métodos multivariados, refleja una evolución hacia un enfoque más probabilístico y computacional en la estadística, permitiendo a los investigadores abordar preguntas de investigación cada vez más sofisticadas (Sagaró del Campo y Zamora, 2019).

Para Meneses (2019), el análisis estadístico con base en un enfoque multivariante, es un conjunto sofisticado de herramientas diseñadas para desentrañar las intrincadas relaciones que subyacen entre múltiples variables. Al construir modelos estadísticos multidimensionales, esta disciplina permite aislar la influencia individual de cada variable en un fenómeno dado, superando las limitaciones de los análisis univariados. De este modo, el análisis multivariado se convierte en un instrumento preciso para describir, explicar y, en muchos casos, predecir fenómenos complejos en diversas áreas del conocimiento.

Trejo, (2014), por su parte, refiere que esta metodología proporcionar marcos teóricos solidos sobre la correlación entre múltiples variables estudiadas haciendo posible cuantificar la contribución de cada variable a la variabilidad total de un fenómeno y, por consiguiente, identificar los factores más relevantes para explicar o predecir su ocurrencia. Esta capacidad de modelar relaciones complejas convierte al análisis multivariante en un factor de gran peso y esencial para la toma de decisiones informadas en ámbitos como las ciencias sociales, la economía y las ciencias naturales.

Con respeto a la utilidad de los métodos multivariados, específicamente a nivel de metodologías investigativas, según Barreto-Villanueva (2012); Trejos (2014) y Meneses (2019), estos métodos contribuyen a:

a) Identificar relaciones causales, al controlar el efecto de variables confusoras, los métodos multivariados permiten establecer relaciones causales más sólidas entre las variables de interés;

b) Modelar fenómenos complejos, al considerar múltiples variables interrelacionadas;

c) Hacer predicciones, mediante otorgarle valor a una variable en función de los valores de otras variables; y

d) Reducir la dimensionaldad, mediante técnicas como el análisis de componentes principales, se logra reducir un gran número de variables a un conjunto más pequeño de factores subyacentes, facilitando la interpretación de los resultados.

El análisis multivariado ofrece un conjunto diverso de técnicas estadísticas para explorar relaciones complejas entre múltiples variables. Si bien estas técnicas comparten el objetivo de analizar datos multidimensionales, cada una posee fortalezas particulares:

a) El análisis factorial, por ejemplo, se destaca en la reducción de dimensionalidad, identificando factores latentes que subyacen a un conjunto de variables observadas;

b) El análisis de componentes principales, al igual que el análisis factorial confirmatorios, son esenciales para la construcción y validación de instrumentos de medición, permitiendo resumir información de múltiples ítems en unos pocos componentes o factores;

c) El análisis de regresión logística, en cambio, se orienta a modelar relaciones causales entre variables, prediciendo la probabilidad de ocurrencia de eventos categóricos, como en el análisis de fenómenos políticos; y

d) El análisis de series de tiempo multivariado, se especializa en el estudio de la evolución de múltiples variables a lo largo del tiempo, capturando patrones y tendencias dinámicas, como en el análisis de indicadores económicos (Meneses, 2019).

Los métodos multivariados permiten estudiar fenómenos complejos interrelacionados, ejemplo de ellos en las ciencias sociales podría ser: Un investigador que busque estudiar la influencia de los retos virales en las redes sociales sobre los niveles de ansiedad y depresión en adolescentes, este estudio se puede llevar a cabo mediante técnicas como el análisis de regresión múltiple, ya que se podría modelar la relación entre la frecuencia de exposición a estos retos, el tiempo invertido en las redes sociales y la sintomatología depresiva y ansiosa, controlando por variables sociodemográficas y factores de riesgo psicosocial. Alternativamente, un análisis de conglomerados permitiría identificar grupos de adolescentes con patrones específicos de uso de redes sociales y niveles de bienestar psicológico, lo que a su vez podría informar intervenciones preventivas más focalizadas.

Tabla 2

Comparativo entre los métodos univariados y multivariados

Características	Método	
	Exploratorio	Confirmativo
Enfoque	Análisis individual de cada variable	Estudio de múltiples variables en conjunto para identificar relaciones y patrones
Técnicas	Medidas de tendencia central (media, mediana, moda), medidas de dispersión (varianza, desviación estándar), gráficos (histograma, boxplot), pruebas de hipótesis (t-student, chi-cuadrado)	Análisis de componentes principales, análisis factorial, análisis de conglomerados, análisis discriminante, regresión múltiple, MANOVA
Historia de utilidad	Tradicionalmente utilizados en estadística básica, ampliamente aplicados en diversas disciplinas	Surgieron en el siglo XX con el desarrollo de computadoras y software estadístico, cada vez más utilizados en campos como marketing, finanzas, ciencias sociales, etc
Funciones	Describir la distribución de una variable, identificar valores atípicos, realizar pruebas de hipótesis sobre un solo parámetro	Identificar relaciones entre variables, reducir la dimensionalidad de los datos, clasificar observaciones, predecir valores de una variable en función de otras

Características	Método	
	Exploratorio	Confirmativo
Objetivos	Obtener un resumen conciso de los datos de una sola variable	Explorar la estructura subyacente de los datos, encontrar patrones complejos y relaciones causales
Inductivo o deductivo	Pueden ser ambos, dependiendo del objetivo del análisis	Generalmente inductivos, ya que buscan descubrir patrones a partir de los datos
Aplicaciones	Estudios descriptivos, análisis exploratorio de datos, pruebas de hipótesis simples	Estudios de correlación, análisis de causalidad, segmentación de mercado, evaluación de modelos predictivos
Otros	Más sencillos de interpretar y aplicar.	Requieren conocimientos estadísticos más avanzados y software especializado

NOTA: Tomado de González y Pérez de Vargas (2009) y Meneses (2019).

Es importa aclarar que los métodos univariados y multivariados se complementan de manera esencial en las ciencias sociales, mientras los primeros proporcionan solidez en los estudios individuales de una variables, los segundo permiten exploraciones en múltiples variables, y en conjunto llegan construir modelos realistas de los fenómenos sociales (Meneses, 2019) (Ver Tabla 2).

CAPÍTULO IV

MÉTODOS EN LA ESTADÍSTICA DESCRIPTIVA E INFERENCIAL

4.1. Métodos en la estadística descriptiva

4.1.1. Métodos que permiten la búsqueda y la obtención de la información

La recolección de datos en la investigación científica es un proceso meticuloso que depende en gran medida de los objetivos del estudio y de las características de la población a analizar. La elección de los métodos estadísticos descriptivos adecuados, así como la definición precisa de la muestra y el diseño experimental, son fundamentales para garantizar la obtención de información confiable y representativa. El muestreo, ya sea probabilístico o no probabilístico, permite inferir características de una población a partir del análisis de una porción representativa de la misma; por su parte, el diseño experimental establece las condiciones bajo las cuales se recolectan los datos, asegurando la validez y la replicabilidad de los resultados (Pérez-Tejada, 2008).

a) Muestreo no probabilístico:

Según Pimienta (2000), este método también conocido como muestreo por juicio, es una técnica de selección de muestras en la cual los elementos no son elegidos al azar, sino que se seleccionan de manera intencional basándose en el criterio del investigador.

Esta metodología se fundamenta en la suposición de que la muestra seleccionada es representativa de la población, a pesar de que no se utiliza un procedimiento aleatorio. La elección de los elementos puede estar influenciada por características específicas, conveniencia o juicio experto, limitando la generalización de los resultados a toda la población. Los investigadores recurren al muestreo no probabilístico cuando la representatividad de la muestra no es la prioridad principal o cuando existen limitaciones prácticas para realizar un muestreo aleatorio. Ejemplos comunes incluyen la selección de casos extremos, la elección de voluntarios o la creación de muestras por cuotas (Pimienta, 2000).

El muestreo no probabilístico, ofrece herramientas valiosas en diversas disciplinas. Métodos como el muestreo por conveniencia (o casual), común en estudios piloto o exploratorios de las disciplina de medicina, arqueología, historia, astronomía, física y química, toma sus muestras con sujetos voluntarios o unidades que se consiguen de manera casual, limitando su generalización; Por otro lado, el muestreo por juicio (o experto) delega la selección de la muestra a individuos con conocimiento especializado, asegurando que los casos seleccionados sean informativos para el objetivo de estudio, tomando en cuenta que la subjetividad del experto puede introducir sesgos; El muestreo de captura-recaptura es fundamental en ecología y biología, permitiendo estimar tamaños poblacionales de especies móviles a través de la marcaje y recaptura de individuos; y Finalmente, el muestreo por cuotas es ampliamente utilizado en encuestas de opinión, donde se establecen cuotas numéricas para asegurar que ciertos subgrupos de la población estén representados en proporción a su tama-

ño, aunque la selección dentro de cada cuota suele ser arbitraria (Pimienta, 2000; Perez-Tejada, 2008)

b) Muestreo probabilístico:

Para Levin y Rubin (2019), El muestreo probabilístico es una técnica estadística que busca obtener una muestra representativa de una población, con el objetivo de inferir características de la población completa a partir del análisis de la muestra, su característica distintiva es que cada elemento de la población tiene una probabilidad conocida y no nula de ser seleccionado; dicha probabilidad conocida permite calcular el error de muestreo, es decir, la diferencia entre el valor de un estadístico calculado en la muestra y el verdadero valor del parámetro en la población, que al conocer este error, podemos establecer un nivel de confianza en los resultados obtenidos.

La utilidad del muestreo probabilístico radica en su capacidad para generar inferencias válidas sobre la población, lo cual es relevante cuando se requiere tomar decisiones basadas en datos, ya que permite cuantificar la incertidumbre asociada a dichas decisiones; sin embargo, su implementación puede ser más compleja y costosa que otros métodos de muestreo (Levin y Rubin, 2019).

Existen diversos tipos de muestreo probabilístico, cada uno con sus características y aplicaciones específicas, las cuales están caracterizadas por González y Pérez de Vargas (2009); Levin y Rubin (2019) de la siguiente manera:

- Muestreo aleatorio simple: Es el método más básico, donde cada elemento de la población tiene la misma probabilidad de ser seleccionado, y aunque este método es conceptualmente útil y sencillo, su aplicación práctica puede resultar limitada en poblaciones grandes, donde la obtención de una lista completa de la misma resulta dificultoso. Por ejemplo, para estudiar las experiencias de inmigrantes en una ciudad,

se podría generar una lista aleatoria de números de identificación de extranjeros y seleccionar una muestra aleatoria de ellos.

- Muestreo sistemático: En este método, se selecciona un punto de partida aleatorio y luego se seleccionan elementos a intervalos regulares. Por ejemplo, para investigar las necesidades de desplazados ambientales en un campamento, se podría seleccionar a cada quinta persona que ingrese al campamento; o en un estudio sobre la opinión de los inmigrantes sobre los servicios sociales, se podría seleccionar cada 10° inmigrante de una lista que exponga su opinión.
- Muestreo estratificado: Ideal cuando la población es heterogénea, se divide la población en grupos homogéneos (estratos) según características relevantes (edad, género, origen); luego, se selecciona una muestra aleatoria de cada estrato. Por ejemplo, para estudiar el maltrato hacia la mujer, se podría estratificar la población por edad, nivel educativo y grupo étnico, y luego seleccionar una muestra aleatoria de cada estrato.
- Muestreo por conglomerados: Se utiliza cuando la población está naturalmente agrupada en conglomerados (por ejemplo, escuelas, barrios). Se selecciona una muestra de conglomerados y luego se estudian todos los elementos dentro de los conglomerados seleccionados. Por ejemplo, para estudiar las actitudes hacia la inmigración en un país, se podrían seleccionar aleatoriamente varias ciudades y luego encuestar a todos los habitantes de esas ciudades.

Pimienta (2000), refiere que tanto el muestreo probabilístico como el no probabilístico buscan obtener información a partir de una porción de la población. Sin embargo, difieren en la forma en que se seleccionan las unidades de análisis; el muestreo probabilístico se basa en procedimientos aleatorios, inferencias estadísticas y cuantificación de la incertidumbre de las estimaciones. Por el contrario, el muestreo no probabilístico utiliza criterios subjetivos o de conveniencia para conformar la muestra, la generalización de resultados es limitada, así como la precisión en las evaluaciones de las estimaciones.

c) Los diseños de experimentos

Según Pérez-Tejada (2008), estos métodos son parte fundamental de las herramientas estadísticas que utiliza la investigación científica y la ingeniería, los cuales tienen como propósito principal establecer relaciones de causa y efecto entre variables. Mediante la manipulación sistemática de una o más variables independientes (factores), se busca cuantificar su impacto sobre una variable dependiente (respuesta); esta metodología se sustenta en la premisa de que la variabilidad observada en los datos experimentales es resultado tanto de la influencia de los factores manipulados como de factores aleatorios inherentes al proceso de medición.

La utilidad de los diseños de experimentos radica en su capacidad para optimizar procesos, mejorar productos y tomar decisiones basadas en evidencia empírica, ya que al permitir identificar los factores más relevantes y sus interacciones, se pueden establecer estrategias para maximizar o minimizar la respuesta de interés.

Según Marin (s.f.); Montgomery (2015), en función del número de factores que se manipulan simultáneamente, los diseños experimentales se clasifican en:

- Diseños de un factor, ya que se evalúa el efecto de una única variable independiente sobre la respuesta, estos diseños son útiles para explorar el comportamiento de un proceso en un rango limitado de condiciones. Los diseños experimentales de un factor incluye:

  1) Diseños completamente al azar, Los tratamientos se asignan aleatoriamente a las unidades experimentales;

  2) Diseños en bloques al azar, las unidades experimentales se agrupan en bloques homogéneos y los tratamientos se asignan aleatoriamente dentro de cada bloque; y

3) Diseños de medidas repetidas, cada unidad experimental recibe todos los tratamientos, y se miden las respuestas para cada tratamiento.

- Diseños de múltiples factores, esta metodología estadistica permiten estudiar el efecto de varios factores y sus interacciones de manera simultánea, son más eficientes que los de un factor, ya que permiten obtener más información con un menor número de experimentos. La tipología de los diseños experimentales multifactoriales incluye:

  1) Diseños factoriales completos, que estudian todas las posibles combinaciones de los niveles de los factores;

  2) Diseños factoriales fraccionados, estos estudian solo una fracción de las posibles combinaciones de los niveles de los factores; y

  3) Diseños de superficie de respuesta, estos se utilizan para optimizar la respuesta en función de los niveles de los factores.

Los diseños experimentales no se limitan a las ciencias naturales, sino que también encuentran una amplia aplicación en las ciencias sociales, por la manipulación de las variables y la medición de sus efectos de manera controlada, estos diseños contribuyen a comprender variados fenómenos sociales que antes no tenían respuestas. Por ejemplo, en el ámbito de la educación: Un investigador podría utilizar un diseño experimental para evaluar la eficacia de diferentes métodos de enseñanza en el rendimiento académico de los estudiantes, al asignar aleatoriamente a los estudiantes a distintos grupos y aplicarles métodos de enseñanza variados, se podría determinar cuál de ellos produce mejores resultados (Marin, s.f.).

Otro ejemplo lo encontramos en el campo de la psicología, un psicólogo podría estar interesado en investigar el impacto de diferentes tipos de terapia en la reducción de la ansiedad, al comparar los resultados de un grupo de pacientes que reciben una terapia cognitivo-conductual con otro

grupo que recibe una terapia psicodinámica, se podría evaluar la efectividad relativa de cada enfoque.

La elección del diseño experimental adecuado depende de diversos factores, como el número de factores a estudiar, el rango de valores de cada factor y el nivel de precisión deseado. Los diseños factoriales completos y fraccionarios son de los más utilizados en la práctica, y es de gran importancia los que pueden generar como conocimiento sólido y confiable sobre el comportamiento humano y las dinámicas sociales (Marin, s.f.).

4.1.2. Métodos para sistematizar y organizar la información

En el ámbito de la investigación científica, el proceso de sistematizar y organizar la información que se recopila, es relevante a la hora de construir el conocimiento científico, porque le otorga orden y estructura al cuerpo o data informativa, facilitando su análisis a la hora de clasificar la información relevante de la que no lo es, evitando la acumulación de la misma o errores de omisión, por tal razón dentro de la estadística como disciplina es importante contar con metodología especializada en dicha área (Corona, 2015).

Para Meneses (2019); Levin y Rubin (2019), esta metodología se define como el proceso estructurado de ordenar, clasificar y analizar los datos recopilados durante la investigación, con el objetivo de facilitar su interpretación y extracción de conclusiones relevantes. Su utilidad radica en permitir a los investigadores identificar estándares, predisposiciones y relaciones significativas entre las variables estudiadas, lo que a su vez impulsa la generación de nuevas hipótesis y el avance del campo de estudio. El objetivo metodológico de la sistematización y organización es transformar los datos brutos en información útil y comprensible, que pueda ser utilizada para responder a las preguntas de investigación planteadas.

Para Trejo (s.f); Calderón y Ledesma (2012); Meneses (2019); Levin y Rubin (2019), en el proceso de sistematizar y organizar la información son

cruciales ciertas técnicas que forman parte de esta metodología, entre las cuales se cuenta:

a) Las técnicas univariadas analíticas:

Específicamente las tablas de frecuencia definidas como representaciones tabulares que condensan la distribución de una variable individual, exhiben la frecuencia con la que cada uno de sus valores se manifiesta en un conjunto de datos. Su utilidad no solo es la descripción de una variable sino la identificación de estándares, disposiciones y anomalías dentro de la información recopilada. En el ámbito de las ciencias sociales, y particularmente en la psicología de las emociones, las técnicas univariadas analíticas demuestran su valía en una diversidad de escenarios, por ejemplo, un caso de estudio paradigmático podría ser el análisis de la distribución de respuestas emocionales ante un estímulo específico en un grupo de participantes. Mediante la construcción de una tabla de frecuencia, el investigador podría visualizar de manera concisa la proporción de individuos que experimentaron cada emoción (por ejemplo, alegría, tristeza, miedo) y la intensidad asociada a cada una de ellas (Trejo, s.f; Calderón y Ledesma, 2012; Levin y Rubin, 2019).

Esta información, susceptible de ser complementada con medidas de tendencia central y dispersión, permitiría establecer un perfil emocional del grupo estudiado y compararlo con otros grupos de referencia. Entendiendo, por una parte, que las medidas de tendencia central, se conciben como parámetros estadísticos que señalan el punto central de un conjunto de datos o muestras, útiles para resumir dicho conjunto en un solo valor representativo, facilitando la comprensión y el análisis de bloques de información, proporciona una visión generalizada y concisa de los datos en cuestión. Y por la otra, que las medidas de dispersión, son valores que se definen como valores numéricos que señalan el grado de variabilidad presente en un conjunto de datos, estas permiten cuantificar cuánto se alejan los datos individuales del valor central, es decir, si estos se encuentran agrupados o dispersos. Las técnicas univariados analíticas, son útiles para

resumir, sistematizar, organizar y analizar datos de manera eficiente en la generación de conocimiento científico solido en diversas disciplinas del área social (Trejo, s.f; Calderón y Ledesma, 2012; Levin y Rubin, 2019).

b) Las técnicas univariados gráficas:

Mientras que el análisis numérico ofrece una precisión detallada de los datos, las técnicas univariadas gráficas complementan esta información al proporcionar una perspectiva visual que facilita la interpretación y comunicación de los resultados ya que los gráficos permiten identificar de manera intuitiva predisposiciones, modelos y distribuciones que pueden ser difíciles de apreciar en tablas de datos, en el ámbito de la investigación como se ha mención con anterioridad, la visualización de gráficos facilitan la formulación de hipótesis, la identificación de anomalías y la generación de nuevas preguntas de investigación (Sulbarán, 2013).

Es importante mencionar que estas técnicas llegan a ser efectivas, ya que se sustentan en principios psicológicos y cognitivos, debido a que la mente humana está diseñada para procesar información visual de manera más rápida y eficiente que la información numérica, y con los gráficos se aprovecha esta capacidad innata para comunicar ideas complejas de manera simple y efectiva. Igualmente la visualización de datos ha sido respaldada por numerosas investigaciones en el campo de la psicología cognitiva y la ciencia de la información, y estudios han demostrado que los gráficos pueden mejorar la comprensión, la retención y la comunicación de información, así como facilitar la toma de decisiones basadas en la data proveniente de ellos (Levin y Rubin, 2019).

Entre las herramientas de la metodología univariados gráficas, que permiten sistematizar y organizar la información, según García-Marín (2022), se cuentan:

1) Gráfico de pastel, es un gráfico circular dividido en porciones, donde cada porción representa una categoría y su tamaño es proporcio-

nal a la frecuencia o porcentaje de dicha categoría en el conjunto de datos, esta herramienta es útil para visualizar la distribución de una variable cualitativa;

2) Gráfico de barras, es un gráfico que utiliza barras rectangulares de longitudes proporcionales a los valores que representan, muy útil para comparar las frecuencias o valores de diferentes categorías de una variable, ya sea cualitativa o cuantitativa discreta;

3) Histograma, es un gráfico similar al de barras, pero utilizado para variables cuantitativas continuas, donde el eje horizontal se divide en intervalos de clase y la altura de cada barra representa la frecuencia de observaciones dentro de ese intervalo; esta herramienta permite visualizar la distribución de frecuencias de la variable; y

4) Gráfico de líneas, es un gráfico que utiliza líneas para conectar puntos de datos, mostrando la evolución de una variable a lo largo del tiempo o en función de otra variable, de gran utilidad para identificar tendencias y patrones en datos secuenciales.

Estas herramientas son de gran utilidad en las ciencias sociales, por ejemplo, en el área de la sociología de masas, son técnicas que ayudan al abordaje de temáticas concerniente a la opinión pública, comportamiento de los consumidores, tendencias y dinámicas sociales. Un caso de estudio podría ser sobre la opinión pública en relación a la aprobación de un nuevo plan gubernamental comunal. Se realiza una encuesta a una muestra representativa de la población y se recogen datos sobre la opinión de cada individuo (favorable, desfavorable, neutral o indiferente), se procede a elaborar un gráfico de pastel, ya que es la herramienta ideal para representar la proporción de individuos que expresan una opinión favorable, desfavorable, neutral o indiferente, ofreciendo una visión global de la opinión pública; Los gráficos de barras, por su parte, permitirían comparar estas opiniones entre diferentes subgrupos demográficos (género, edad, etc.), facilitando la identificación de posibles diferencias significativas, si

se dispone de datos cuantitativos sobre el nivel de conocimiento acerca del plan; un histograma proporcionaría una representación visual de la distribución de este conocimiento en la población, permitiendo identificar, en caso de que existiera un grupo mayoritario con un nivel de conocimiento bajo o alto.

c) Las técnicas multivariadas:

Incluyen métodos para reducir variables y formar grupos, permiten analizar simultáneamente múltiples variables y sus interrelaciones, son métodos estadísticos avanzados, se erigen como herramientas esenciales en la investigación contemporánea que al analizar de manera simultánea múltiples variables y las intrincadas relaciones que se establecen entre ellas, le otorgan la capacidad analítica a los investigadores para abordar fenómenos complejos de las ciencias, donde una sola variable no sería suficiente para comprender la totalidad del panorama (Meneses, 2019). La utilidad de estas técnicas se manifiesta en su habilidad para sintetizar y organizar grandes volúmenes de información, que guía la modelización, esquematización y recurrencias relacional de data de hechos que pueden pasar desapercibidos.

Las técnicas multivariadas según Calderón y Ledesma (2012), resaltan las técnicas que reducen variables y las que forman grupos. Estos dos componentes metodológicos contribuyen a simplificar conjuntos de datos extensos y complejos, facilitando su interpretación y comprensión, puesto que al reducir el número de variables, se pueden identificar los factores más importantes que influyen en un fenómeno determinado. La formación de grupos, por su parte, permite agrupar observaciones similares en función de sus características.

Complementado lo anterior, García-Marín (2022), refiere que los diagramas de barras son útiles para representar la relación entre una variable categórica y una numérica, mientras que los diagramas de dispersión multivariantes permiten visualizar la relación entre dos variables, ya sean

numéricas o categóricas. Estas herramientas de visualización facilitan la identificación de plantillas y preferencias en la organización de los datos, lo que llega ser de gran utilidad en muchos aspectos de la investigación donde se acciona la toma de decisiones.

En el contexto de las ciencias sociales, las técnicas multivariadas se convierten en aliadas valiosas, por ejemplo al estudiar el maltrato de género en México, las mismas pueden ser empleadas para analizar encuestas y registros y armar una data estructurada de incidentes de violencia contra la mujer, identificando factores de riesgo comunes, patrones de comportamiento de los agresores y las consecuencias psicológicas y sociales que sufren las víctimas. Asimismo, pueden ser utilizadas para evaluar la efectividad de programas de prevención y atención a víctimas de violencia de género, comparando los resultados obtenidos en diferentes contextos y poblaciones.

Con respecto a las técnicas que reducen variables y forman grupos, se pueden utilizar para este ejemplo, ya que ayudan en el análisis de componentes principales, reduciendo la dimensionalidad de los datos e identificando las variables más relevantes; y el análisis de conglomerados, contribuye por su parte, a agrupa a los individuos o casos en función de sus similitudes. Estas técnicas les ofrecen a los investigadores un abanico de posibilidades para abordar el complejo fenómeno del maltrato de género desde una perspectiva multidimensional y obtener conclusiones más precisas y significativas (García- Marín, 2022).

La representación gráfica es fundamental en el análisis univariado, ya que los histogramas, diagramas de caja y diagramas de tallo y hoja son herramientas comunes para visualizar la distribución y dispersión de variables cuantitativas, pero para variables cualitativas, los gráficos de barras, de bastones o circulares son útiles para representar las frecuencias de cada categoría de manera proporcional (Meneses, 2019).

4.2. Métodos en la estadística inferencial

4.2.1. Métodos para calcular el componente de la tendencia secular en las series cronológicas

Los métodos para calcular la tendencia secular en series cronológicas son herramientas del análisis de datos, ya que permiten identificar y modelar la dirección general que sigue una variable a lo largo del tiempo, donde el componente de la serie, representa su comportamiento a largo plazo, siendo el mismo esencial para comprender la evolución de un fenómeno y realizar proyecciones futuras (Santiago de la Fuente, 2014).

Para la Universidad Autónoma del Estado de México (UAEMéx.) (2010), en estadística la tendencia secular es el movimiento frecuente y general –creciente, decreciente o estable- de una serie de datos a lo largo de un periodo de tiempo, dichas frecuencias son un componente básico de las series temporales, las cuales se pueden describir a través de funciones matemáticas, medias móviles o alisamiento exponencial. Las tendencias seculares, se pueden aplicar en estudios o ámbitos de variadas disciplinas como: epidemiologia, demografía, finanzas, negocios y economía, ya que son útiles para análisis de fenómenos sociales complicados, donde intervienes factores externos como crisis económicas, conflictos armados y disrupciones tecnológicas pueden llegar a ocultar la tendencia.

Para Mauricio (2007); UAEMéx.(2010), entre los métodos más utilizados para estimar la tendencia secular, se pueden contar:

a) Las funciones matemáticas, se busca o ajusta la función matemática -lineal, polinómica, exponencial, entre otras- a los datos para modelar la tendencia, siendo útil este método cuando la tendencia es clara y se ajusta bien a una forma matemática específica;

b) Los semipromedios, se divide la serie de datos en dos mitades y se calcula el promedio de cada mitad, siendo la línea que une estos dos pro-

medios la representación de la tendencia secular, este método es sencillo, pero puede ser sensible a valores atípicos;

c) Los promedios móviles, se calcula el promedio de un número fijo de puntos de datos consecutivos y se desplaza este promedio a lo largo de la serie, este procedimiento suaviza las fluctuaciones a corto plazo y revela la tendencia subyacente, dicho método es útil para identificar tendencias en datos ruidosos; y

d) Mínimos cuadrados, con este método se ajusta una línea (o curva) a los datos que minimiza la suma de los cuadrados de las diferencias entre los valores observados y los valores predichos por la línea, este método es ampliamente utilizado y proporciona una estimación precisa de la tendencia.

Cada uno de estos métodos tiene sus propias ventajas y limitaciones, y su elección dependerá de las características específicas de la serie y del objetivo del análisis. Sin embargo, en conjunto, estas herramientas proporcionan un marco sólido para comprender y modelar la tendencia secular en series cronológicas, lo que a su vez facilita la toma de decisiones informadas en diversos campos de aplicación (UAEMéx., 2010).

La aplicación de esta serie de métodos para calcular el componente de la tendencia secular en una serie cronológica, en el área socioeconómica podría ser de manera general: Un estudio sobre la evolución de la desigualdad de ingresos en un país latinoamericano (Perú) durante los últimos 30 años. En primer lugar se recopilas datos del coeficiente de Gini (medida de desigualdad) para cada año; luego se aplican los tres métodos (semipromedios, promedios móviles y mínimos cuadrados) a esta serie cronológica; Posteriormente se comparas los resultados para identificar la tendencia secular de la desigualdad (¿aumento, disminución o estabilidad); y Finalmente se analizan los factores sociopolíticos que pueden haber influido en esta tendencia (ejemplo: políticas económicas, programas sociales, cambios demográficos) (Santiago de la Fuente,2014).

4.2.2. Métodos para determinar el espacio muestral en la teoría de probabilidades

Pérez-Tejada (2008); González y Pérez de Vargas (2009), refieren que en la teoría de probabilidades, la determinación del espacio muestral es un elemento clave en los análisis de datos de la estadística inferencial. El espacio muestral, definido como el conjunto de todos los resultados posibles de un experimento aleatorio, es esencial para calcular la probabilidad de eventos específicos y su correcta identificación, dicho análisis ayuda a comprender y modelar fenómenos fortuitos, ya sean discretos (con un número finito o infinito numerable de resultados) o continuos (con un número infinito no numerable de resultados). En otras palabras, la delimitación precisa del espacio muestral es, por lo tanto, un paso crítico en el proceso de inferencia estadística, ya que proporciona el marco necesario para calcular probabilidades con las cuales se toman decisiones orientadas por una data.

Las técnicas de conteo y análisis combinatorio, tienen como objetico principal determinar el número de posibles resultados (espacio muestral) en un experimento aleatorio, posibilitando los cálculos de probabilidades en torno a un fenómeno o acontecimiento en proceso investigativo, tanto en el ámbito natural como en el social, Para la Universidad Nacional Autónoma de México (UNAM) (2015); Acosta (2019), estas técnicas incluyen:

a) El teorema fundamental del conteo, para acciones o eventos, el cual establece que si un evento puede ocurrir de n maneras y otro evento puede ocurrir de m maneras, entonces ambos eventos pueden ocurrir juntos de n × m maneras, es decir, este principio es básico para calcular el número de posibles resultados en experimentos que involucran múltiples pasos o decisiones;

b) Los diagramas de árbol, son representaciones gráficas que muestran todos los posibles resultados de una secuencia de eventos, donde cada

rama del árbol representa una posible opción o resultado, facilitando la visualización y el conteo de todas las combinaciones posibles.

c) Combinaciones: Permiten calcular el número de formas (subconjuntos posibles de datos), esta herramienta es esencial en diversas áreas- estadística, probabilidad y optimización- ya que permite calcular probabilidades de eventos compuestos y diseñar experimentos. Existen dos tipos principales de combinaciones: Combinaciones simples y combinaciones con repetición.

d) Permutaciones: Son herramientas matemáticas útiles para resolver problemas donde el orden de los elementos es importante, y se adaptan a diferentes escenarios según la naturaleza de los elementos y la forma de ordenarlos. Las permutaciones pueden ser: simples, con repeticiones, con grupos de elementos iguales y circulares

e) Variaciones: Similares a las permutaciones, pero solo se seleccionan algunos elementos del conjunto.

Un ejemplo de la aplicación de estas técnicas en el ámbito psicológico, podría ser en un estudio sobre el impacto de tres factores de riesgo en el suicidio de adolescentes con TDAH, donde se especifique;

a) Depresión: (presente o ausente);

b) Acoso escolar: (presente o ausente); y

c) Impulsividad: (alta o baja).

En este caso se podría crear, por ejemplo, un diagrama de árbol para visualizar las 8 combinaciones posibles de estos factores ($2 \times 2 \times 2 = 8$). Luego, se procede a calcular la probabilidad de suicidio para cada combinación utilizando datos de la muestra de adolescentes con TDAH. Al analizar los resultados, la combinaciones podrían arrojar que la combinación de depresión, acoso escolar y alta impulsividad, tiene la probabilidad de

suicidio más alta; esta información podría ser de gran utilidad para diseñar intervenciones específicas para adolescentes con esta combinación de factores de riesgo.

4.2.3. Métodos utilizados en la teoría de la estimación

Para González y Pérez de Vargas (2009), la teoría de la estimación, es una rama importante de la estadística, ya que se enfoca en la inferencia de valores numéricos para parámetros poblacionales desconocidos a partir de datos empíricos. Este proceso de inferencia se basa en la comprensión de que los datos empíricos poseen un componente aleatorio inherente y que los parámetros poblacionales, al describir un entorno profundo, influyen en la distribución de los datos medidos.

La estimación, a menudo llevada a cabo mediante el muestreo, implica la selección de un subconjunto de una población más amplia para proyectar conclusiones sobre la totalidad. En este sentido, la teoría de la estimación proporciona el marco teórico y metodológico para realizar inferencias precisas y confiables sobre poblaciones basándose en información muestral, lo que resulta esencial en diversas áreas del conocimiento (Levin Y Rubin, 2019).

La teoría de la estimación se nutre de diversos métodos estadísticos, cada uno con sus propias fortalezas y aplicaciones. Entre los más importantes y relevantes, según González y Pérez de Vargas (2009); Levin y Rubín (2019), destacan:

a) La estimación puntual:

Busca identificar un valor único que represente la mejor aproximación del parámetro poblacional desconocido, este proceso implica la selección de un estimador puntual, cuya naturaleza varía según el parámetro que se desea estimar y las características intrínsecas de la población en estudio, ya que se busca aquel parámetro que exhiba propiedades desea-

bles como la insesgadez, la eficiencia y la consistencia. Algunos ejemplos comunes de estimadores puntuales incluyen la media muestral, empleada para inferir la media poblacional; la proporción muestral, utilizada para estimar una proporción poblacional; y la varianza muestral, que sirve para aproximar la varianza poblacional. La elección del estimador puntual óptimo recae en un análisis cuidadoso de la población y del parámetro en cuestión, buscando aquel que exhiba propiedades deseables como la insesgadez, la eficiencia y la consistencia.

b) La estimación por intervalos:

Se distingue de la estimación puntual al proporcionar un rango de valores plausibles para un parámetro poblacional desconocido, en lugar de un único valor. Este rango, denominado intervalo de confianza, se construye con un nivel de confianza predefinido, lo que implica que, si el proceso de muestreo se repitiera numerosas veces, el intervalo contendría el verdadero valor del parámetro en un porcentaje de las veces igual al nivel de confianza. Los intervalos de confianza en la práctica estadística, ofrecen una medida tangible de la incertidumbre asociada a las estimaciones y permiten a los investigadores y tomadores de decisiones evaluar la fiabilidad de sus conclusiones.

c) El método de Máxima Verosimilitud (MV):

Esta función, definida como la probabilidad de obtener los datos observados bajo diferentes valores del parámetro desconocido, es el eje central del método MV. La premisa fundamental radica en que el estimador más plausible para un parámetro desconocido es aquel que maximiza la función de verosimilitud. En otras palabras, el método MV busca el valor del parámetro que hace que los datos observados sean lo más probables posible. Este enfoque, respaldado por sólidas propiedades estadísticas como la consistencia y la eficiencia, le confiere relevancia en la estimación de parámetros, proporcionando estimadores óptimos en una amplia gama de situaciones y modelos estadísticos.

d) Método de los Momentos:

Basado en la igualación de los momentos muestrales (calculados a partir de los datos) con los momentos poblacionales (expresados en función de los parámetros desconocidos). Al resolver este sistema de ecuaciones, se obtienen estimadores para los parámetros, dicho método puede ser más sencillo de aplicar que el de Máxima Verosimilitud en algunos casos, no siempre conduce a estimadores óptimos.

e) Estimación Bayesiana:

Este enfoque combina la información muestral con el conocimiento previo sobre los parámetros (expresado a través de una distribución a priori) para obtener una distribución a posteriori que representa la información actualizada sobre los parámetros. Esta metodología permite incorporar información subjetiva o proveniente de estudios previos en el análisis, lo que puede ser útil cuando se cuenta con poca información muestral.

La aplicabilidad en las ciencias sociales, podría ser en el área de la sociología cultural, con un ejemplo de preferencias musicales utilizando el método de Máxima Verosimilitud (MV), en un estudio sobre preferencias musicales, se encuestó a 100 adolescentes, revelando que 60 de ellos prefieren la música urbana. Asumiendo que la preferencia musical sigue una distribución de Bernoulli con un parámetro p (probabilidad de preferir música urbana), el objetivo es estimar p mediante el método MV. La función de verosimilitud, que representa la probabilidad de observar los datos (60 de 100 preferencias por la música urbana) bajo diferentes valores de p, se maximiza para encontrar el estimador de máxima verosimilitud.

En este caso, el estimador resulta ser la proporción observada de preferencias por la música urbana, es decir:

$$\frac{60}{100} = 0.6$$

Este valor sugiere que el 60 % de los adolescentes encuestados prefieren la música urbana, proporcionando una estimación basada en la maximización de la probabilidad de los datos observados. En muchos casos, se pueden utilizar varios métodos diferentes y comparar sus resultados para obtener una comprensión más completa del fenómeno estudiado (González y Pérez de Vargas, 2009).

4.2.4. Método de la prueba de hipótesis o teoría de decisión

Levin y Rubín (2019), trata la prueba de hipótesis por separado de la teoría de decisiones, ya que, aunque difieren en su enfoque y alcance ambas son herramientas para tomar decisiones. La prueba de hipótesis, un procedimiento estadístico esencial, que se centra en la evaluación de una afirmación o hipótesis sobre una población, utilizando evidencia muestral para determinar si se debe aceptar o rechazar dicha hipótesis, teniendo como objetivo principal el discernir si los resultados observados en una investigación son estadísticamente significativos o si podrían haber ocurrido por azar.

Por otro lado, la teoría de decisiones, es un marco más amplio, abarca un análisis lógico y cuantitativo de todos los factores que pueden influir en el resultado de una decisión, considerando tanto la incertidumbre como las posibles consecuencias. Mientras que la prueba de hipótesis se enfoca en la validez de una afirmación específica, la teoría de decisiones proporciona un enfoque integral para la toma de decisiones en situaciones complejas y con múltiples factores en juego (González y Pérez de Vargas, 2009).

Dentro del ámbito de la inferencia estadística, tanto la prueba de hipótesis como la teoría de decisiones emplean diversos métodos específicos para alcanzar sus objetivos, es decir, hay tanto métodos como propósitos investigativos, pero en este apartado se explicaran de ambas metodologías dos de los más utilizados y relevantes. González y Pérez de Vargas (2009); Levin y Rubín (2019), relatan que en la prueba de hipótesis, dos métodos destacados son las pruebas t de Student y las pruebas de chi-cuadrado:

- Las pruebas t, particularmente útiles cuando se trabaja con muestras pequeñas y se desconoce la desviación estándar de la población, se centran en determinar si existe una diferencia estadísticamente significativa entre las medias de dos grupos; este método evalúa si la diferencia observada entre las medias muestrales es lo suficientemente grande como para descartar la posibilidad de que se deba simplemente al azar.
- Las pruebas de chi-cuadrado, se aplican a datos categóricos, es decir, datos que representan cualidades o atributos, siendo su objetivo principal analizar si existe una asociación significativa entre dos variables categóricas, revelando si las categorías de una variable están relacionadas de alguna manera con las categorías de la otra. En conjunto, las pruebas t y las pruebas de chi-cuadrado socorren en la extracción de conclusiones significativas sobre las poblaciones de interés.

En el campo de la teoría de decisiones, dos métodos notables son el análisis de árbol de decisiones y la teoría de juegos:

- El análisis de árbol de decisiones, una metodología gráfica y analítica, permite a quien toma las decisiones visualizar y evaluar sistemáticamente diferentes cursos de acción, así como sus posibles resultados y probabilidades asociadas, el objetivo principal radica en facilitar la selección de la mejor opción en situaciones complejas donde existe incertidumbre sobre los eventos futuros.
- La teoría de juegos, es un campo interdisciplinario que combina elementos de matemáticas, economía y ciencias de la computación, se centra en el análisis de decisiones en situaciones estratégicas donde interactúan múltiples agentes o jugadores, en este proceso cada agente tiene sus propios objetivos y estrategias, y sus decisiones pueden afectar los resultados de los demás. La teoría de juegos proporciona modelos y herramientas para comprender el comportamiento estratégico en escenarios competitivos o colaborativos, permitiendo a los analistas predecir y explicar las decisiones de los agentes y los resultados resultantes.

Un ejemplo sobre la aplicabilidad del análisis de árbol de decisiones, podría ser: En el contexto del prolongado conflicto entre Hamás e Israel, la liberación de rehenes israelíes en manos de Hamás emerge como un escenario crítico. Israel, buscando la liberación con mínimas concesiones y riesgos futuros, puede emplear el análisis de árbol de decisiones. Este enfoque permite visualizar las alternativas de negociación (negociación directa o indirecta, presión internacional, acción militar limitada) y sus posibles resultados (éxito o fracaso), asignando probabilidades basadas en datos y análisis de expertos. Al calcular el valor esperado de cada alternativa, se ponderan los resultados por sus probabilidades, facilitando la selección de la estrategia de negociación más prometedora para Israel en este complejo escenario.

REFLEXIONES FINALES

Para Sagaró del Campo y Zamora (2019), en el ámbito de las ciencias sociales, la aplicación de métodos estadísticos ha experimentado un auge significativo, impulsado por el desarrollo acelerado de esta disciplina, pero la variedad de técnicas estadísticas disponibles resultan abrumadoras, incluso para expertos en la materia, cayéndose en excesos en sus aplicaciones. Sin embargo, desde mediados del siglo XIX, una técnica en particular ha demostrado estar presente en estudios con enfoque social: el análisis de regresión, esto se debe a que dicha técnica permite examinar las relaciones entre variables y realizar predicciones, lo cual resulta esencial para comprender la complejidad de los fenómenos sociales.

Por tal razón, es crucial reconocer que los métodos estadísticos, si bien llegan a ser herramientas eficaces, pueden ser malinterpretados y utilizados de forma inapropiada por personas sin experiencia. La estadística requiere un manejo cuidadoso y una comprensión profunda de sus fundamentos teóricos y su aplicación en las ciencias sociales exige una combinación de habilidades técnicas y un juicio crítico para evitar conclusiones erróneas o sesgadas (Sagaró del Campo y Zamora, 2019).

Enrique y Peña (2020), menciona que los fenómenos sociales son multifactoriales y cambiantes, influenciados por la conducta humana y múltiples factores que actúan en conjunto, haciéndolos más difíciles de modelar estadísticamente que los fenómenos naturales o exactos, es decir, no todos

los métodos estadísticos son adecuados para cada fenómeno. Pero a pesar de estas limitaciones, ha sido la estadística en la investigación social la que le ha podido dar fundamentos a la variabilidad de la conducta humana, mediante la identificación de tendencias generales a partir del estudio de grupos de personas; siempre que el investigador ha tenido cuidado y criterios en la elección de la herramienta estadística adecuada para cada investigación, resultando conclusiones apropiadas y exitosas.

Por su parte, Calderón y Ledesma (2012), refieren que la estadística, en esencia, es una disciplina que demanda un alto grado de responsabilidad y ética por parte de quienes la utilizan, y en la investigación científica social, la honestidad y la transparencia en el manejo de la información son fundamentales para garantizar la validez y la confiabilidad de los estudios sociales, observándose todos estos valores en la inexistencia en los estudios sociales de cualquier tipo de manipulación de datos para obtener resultados deseados.

Con respeto a la actualidad metodológica de la estadística, Sagaró del Campo y Zamora (2019), mencionan que la convergencia de diversas disciplinas, como la bioestadística, la computación, los sistemas de información, la inteligencia artificial y el análisis en línea, permite abordar el estudio de las ciencias sociales desde una perspectiva más amplia y sofisticada. La disponibilidad de software estadístico especializado, el acceso a grandes volúmenes de datos poblacionales y la capacidad de identificar modelos válidos mediante técnicas de descubrimiento del conocimiento en bases de datos, brindan a los investigadores un abanico de posibilidades sin precedentes. Además, la colaboración internacional y la transdisciplinariedad, donde confluyen matemáticos, filósofos, cibernéticos, biólogos, epidemiólogos, economistas, sociólogos y geógrafos, enriquecen el debate científico y promueven el desarrollo de nuevas metodologías. La aplicación de técnicas estadísticas diversas, adaptadas a las particularidades de cada fenómeno social estudiado, permite a los investigadores obtener resultados más precisos y relevantes.

REFERENCIAS BIBLIOGRÁFICAS

Acosta, G. (2019). *Análisis Combinatorio.* https://departamentodematematicas.itam.mx/sites/default/files/departamentodematematicasitammx/paginas/notasconteoalgsup2_otono2019.pdf

Barreto-Villanueva, A. (2012). El progreso de la Estadística y su utilidad en la evaluación del desarrollo. *Papeles de Población, 18* (73), 1-31. http://www.redalyc.org/articulo.oa?id=11224638010

Bautista-Díaz, M., Victoria-Rodríguez, E., Vargas-Estrellac, L. y Hernández-Chamosad, C. (2020). Pruebas estadísticas paramétricas y no paramétricas: su clasificación, objetivos y características. *Educación y Salud Boletín Científico Instituto de Ciencias de la Salud, 9* (17), 78-81. https://repository.uaeh.edu.mx/revistas/index.php/ICSA/issue/archive.

Bujan, A. (2017). *Estadística descriptiva.* http://www.enciclopediafinanciera.com/definicion-estadistica-descriptiva.html

Calderón, M. y Ledesma, G. (2012). Los métodos estadísticos en las investigaciones sociales. *Pedagogía y Sociedad, 15*, (34), 1-9. https://www.redalyc.org/articulo.oa?id=581777892004

Calduch, R. (s.f.). *Métodos y técnicas de investigación en relaciones internacionales* [Curso de Doctorado, Universidad Complutense de Madrid, España] https://www.ucm.es/data/cont/media/www/pag-55163/2Metodos.pdf

Carrasco, S. (s. f.). *Aproximaciones a las estadísticas de las ciencias sociales.* 1-25. https://www.bibliotecaenlinea.universidaddorados.edu.mx/contaduria-publica/Estadistica/aproximacion%20estadistica.pdf

Corona, R. (2015). Los métodos estadísticos como fuente de mejora de la calidad en las empresas de manufactura. *NovaRua: Revista Universitaria de Administración, 6*(10), 16-25. https://dialnet.unirioja.es/servlet/articulo?codigo=8043232

Daza, M (2015). *Estadística aplicada a la Educación.* https://www.cartagena99.com/recursos/alumnos/apuntes/Estadistica%20Educacion%20Social_T6_p.pdf

Enrique, F. y Peña, M. (2020). Improcedencias al usar la estadística en las investigaciones sociales. *VARONA, 70.* https://www.redalyc.org/journal/3606/360671237004/html/

Federación Científico-Europea de Osteópatas (SEFO) (2019). *Estadística Paramétric*a. https://www.scientific-european-federation-osteopaths.org/wp-content/uploads/2019/01/Estadística-paramétrica.pdf

García-Marín, A. (2022). *Visualización de datos univariantes y multivariantes. Catálogos de técnicas.* [Trabajo de Grado, Universidad Politécnica de Madrid, España]. https://oa.upm.es/71215/1/TFG_ALFONSO_GARCIA_MARIN.pdf

González, M. y Pérez de Vargas, A. (2009). *Estadística aplicada una visión instrumental Teoría y más de 500 problemas resueltos o propuestos con solución.* Ediciones Díaz de Santos. https://www.editdiazdesantos.com/wwwdat/pdf/9788479789138.pdf

Levin, L., y Rubin, D. (2019). *Estadística para administración y economía.* Pearson Educación. https://www.economicas.unsa.edu.ar/afinan/informacion_general/book/ebooks/Estadistica%20para%20administracion%20y%20economia%20Levin%20et%20al.pdf

Lizasoain, L. (2024). El análisis estadístico de datos en la investigación educativa. *Revista electrónica interuniversitaria de formación del profesorado, 2*(27), 217-232. https://produccioncientifica.ucm.es/documentos/66391e-733f575266e3bf7d6c

López, A. y Diez, T. (2017). Aproximación de la estadística a las ciencias sociales: una mirada crítica. *Revista Cubana de Educación Superior, 3*6(2), 148-156. http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0257-43142017000200014&lng=es&tlng=es

Marín, J. (s.f.). *Introducción al Diseño de Experimentos.* https://halweb.uc3m.es/esp/personal/personas/jmmarin/esp/disenno/introde.pdf

Mauricio, J. (2007). *Introducción al análisis de series temporales.* https://www.ucm.es/data/cont/docs/518-2013-11-11-JAM-IAST-Libro.pdf

Meneses, J. (2019). *Introducción al análisis multivariante.* Ediciones de la Universitat Oberta de Catalunya. https://femrecerca.cat/meneses/publication/introduccion-analisis-multivariante/introduccion-analisis-multivariante.pdf

Montgomery, D. (2015). *Diseño de Experimentos. Principios y Metodología*. Editorial Limusa Wiley. https://es.slideshare.net/slideshow/diseo-y-analisis-de-experimentos-montgomery-ocr/55532757

Parra, J. (2002). Análisis exploratorio y análisis confirmatorio de datos. *Espacio Abierto, 11*(1). http://www.redalyc.org/articulo.oa?id=12211106

Pérez- Juste, R. (2012). *Estadísticas aplicadas a las ciencias sociales. Cursos destinados a la preparación de futuros estudiantes de las asignaturas de estadísticas de grados universitarios en C.C. Sociales.* UNED. https://www.intecca.uned.es/upload/noticias/20130614114556u_curso0deestadisticaaplicada.pdf

Pérez-Tejada, H. (2008). Estadística para las ciencias sociales, del comportamiento y de la salud. Cengage Learning Editores, S.A. https://www.uv.mx/rmipe/files/2015/09/Estadistica-para-las-ciencias-sociales-del-comportamiento-y-de-la-salud.pdf

Pimienta, R. (2000). Encuestas probabilísticas vs no probabilísticas. *Política y Cultura, 13*, 263-276. http://www.redalyc.org/articulo.oa?id=26701313

Phillips, E. y Pugh, D. (2011). Modelos y aproximaciones teóricas: Fines doctorales. En Uzcátegui, A. (Comp). *La Tesis Doctoral.* https://expresionesdeintelectuales.blogia.com/2011/030901-modelos-y-aproximaciones-teoricas-fines-doctorales.php

Sagaró del Campo, N. y Zamora, L. (2019). Evolución histórica de las técnicas estadísticas y las metodologías para el estudio de la causalidad en ciencias médicas. *MEDISAN, 23*(3), 534-556. https://www.redalyc.org/journal/3684/368460217013/html/

Salgado-Ugarte, I. (2013). *Métodos estadísticos exploratorios y confirmatorios para análisis de datos. Un Enfoque Biométrico*. National Autonomous University of Mexico.

Santiago de la Fuente, F. (2014). *Series temporales*. https://www.estadistica.net/PAU2/series-temporales.pdf

Sulbarán, D. (2013). *Introducción al análisis univariado de datos*. [Trabajo de ascenso docente, Universidad Central de Venezuela, Caracas-Venezuela]. http://saber.ucv.ve/bitstream/10872/17386/1/Análisis%20Univariado%20de%20Datos.pdf

Trejos, J. (2014). *Análisis Multivariado de datos*. [Trabajo de ascensos, Universidad de Costa Rica]. https://gc.scalahed.com/recursos/files/r161r/w24529w/S10/NotasAMD.pdf

Universidad Autónoma del Estado de México (UAEMéx.) (2010). *Series de Tiempo*. https://gc.scalahed.com/recursos/files/r161r/w24481w/S9_R02.pdf

Universidad Nacional Autónoma de México (UNAM) (2015). *Fundamentos de la teoría de la probabilidad*. https://www.paginaspersonales.unam.mx/files/977/ApuntesTema_2.pdf



EDITORIAL
NAVEGANTE