COLECCIÓN INVESTIGACIÓN

La **Estadística** **NO PARAMÉTRICA** *en los* **Libros de texto universitarios**

Situaciones problema para promover la investigación científica en el aula

Víctor Miguel Ángel Burbano Pantoja
Margoth Adriana Valdivieso Miranda
Ángela Saray Burbano Valdivieso

EDITORIAL UPTC

LA ESTADÍSTICA NO PARAMÉTRICA EN LOS LIBROS DE TEXTO UNIVERSITARIOS

Situaciones problema para promover la investigación científica desde el aula

Víctor Miguel Ángel Burbano Pantoja
Margoth Adriana Valdivieso Miranda
Ángela Saray Burbano Valdivieso

EDITORIAL
U P T C

2023

La estadística no paramétrica en los libros de texto universitarios: situaciones problema para promover la investigación científica desde el aula / Non-Parametric Statistics in University Textbooks: Problem Situations to Promote Scientific Research from the Classroom / Burbano Pantoja, Víctor Miguel Ángel; Valdivieso Miranda, Margoth Adriana; Burbano Valdivieso, Ángela Saray. Tunja: Editorial UPTC, 2023. 184 p.

ISBN (ePub) 978-958-660-730-8

Incluye referencias bibliográficas

1. Estadística. 2. Investigación científica. 3. Libros de texto universitarios. 4. Métodos no paramétricos.
5. Prueba de hipótesis. 6. Situaciones problema.

(Dewey 519.5/21) (Thema 375831-PBT- Probabilidad y Estadística)

Primera edición, 2023
Versión digital
La estadística no paramétrica en los libros de texto universitarios: situaciones problema para promover la investigación científica desde el aula
Non-Parametric Statistics in University Textbooks: Problem Situations to Promote Scientific Research from the Classroom

ISBN (ePub) 978-958-660-730-8
Colección Investigación UPTC N.º 268
Proceso de arbitraje doble ciego
Recepción: agosto de 2022
Aprobación: enero de 2023

Editorial UPTC
Edificio Administrativo, piso 4
La Colina, bloque 7, casa 5
Avenida Central del Norte 39-115, Tunja, Boyacá
comite.editorial@uptc.edu.co
www.uptc.edu.co

Libro financiado por la Vicerrectoría de Investigación y Extensión - Dirección de Investigaciones de la UPTC, Convocatoria 01 de 2022 "Investigar da más" estímulo económico a grupos de investigación por productividad resultado de proyectos de investigación.

Impreso y hecho en Colombia - Printed and made in Colombia

Libro resultado de investigación del proyecto "El pensamiento aleatorio y los sistemas de datos a través de medios virtuales" con SGI 2605.
Citar este libro / Cite this book
Burbano Pantoja, V., Valdivieso Miranda, M. & Burbano Valdivieso, Á. (2023). *La estadística no paramétrica en los libros de texto universitarios: situaciones problema para promover la investigación científica desde el aula.* Editorial UPTC.

doi: https://doi.org/10.19053/9789586607308

RESUMEN

La inferencia estadística es un elemento esencial dentro del método científico experimental. Los procesos de inferencia pueden desarrollarse por medio de la aplicación de métodos cuantitativos paramétricos o de modelos estadísticos no paramétricos. Los primeros se utilizan cuando se cumplen supuestos estrictos sobre la forma de la distribución de probabilidad de la población de la cual se ha de seleccionar la muestra aleatoria que sirve de base para estimar uno o más parámetros desconocidos de la población, construir intervalos de confianza para tales parámetros o determinar regiones críticas para rechazar la hipótesis nula. Los modelos estadísticos no paramétricos son técnicas estadísticas que requieren supuestos más flexibles o usan estadísticas de prueba libres de distribución, las cuales generalmente se definen por medio de los rangos vinculados a los datos de la variable o variables por investigar, ya sea que se trate de un problema de inferencia relacionado con una muestra aleatoria, dos, tres o más muestras.

Este trabajo investigativo indaga acerca de la manera como algunos libros de texto universitarios examinan los tópicos de estadística no paramétrica referentes a la prueba de hipótesis. Para este propósito, se utiliza una metodología basada en un enfoque mixto, es decir, en métodos cuantitativos y cualitativos de investigación (análisis de contenido e investigación documental). Además, contempla diversas situaciones problema concretas, que permiten promover la investigación científica desde los claustros universitarios. Esta obra está dirigida a estudiantes, profesores, profesionales e investigadores interesados en aplicar modelos estadísticos no paramétricos clásicos en la prueba de hipótesis en sus trabajos de investigación o en incrementar sus saberes sobre tales modelos. Para abordarla se necesitan conocimientos básicos de estadística descriptiva e inferencial desde una mirada paramétrica.

Palabras clave: estadística, investigación científica, libros de texto universitarios, métodos no paramétricos, prueba de hipótesis, situaciones problema.

ABSTRACT

Statistical inference is an essential element within the experimental scientific method. Inference processes can be developed through the application of parametric quantitative methods or using non-parametric statistical models. The former is used when strict assumptions are met about the shape of the probability distribution of the population from which the random sample is to be selected as the basis for estimating one or more unknown population parameters, constructing confidence intervals for such parameters or determining critical regions to reject the null hypothesis. Non-parametric statistical models are statistical techniques that require more flexible assumptions or use distribution-free test statistics, which are generally defined by the ranges associated with the data of the variable or variables to be investigated, whether it is an inference problem associated with one random sample, two samples, or three or more samples.

This research work focuses on the way in which some university textbooks address the topics of non-parametric statistics associated with hypothesis testing. For this purpose, a methodology based on a mixed approach is used, since it is supported by both quantitative and qualitative research methods (content analysis and documentary research). In addition, it considers several concrete problem situations, susceptible to promote scientific research from university cloisters. This book is aimed at students, professors, professionals and researchers interested in applying classical non-parametric statistical models in hypothesis testing in their research work or in increasing their knowledge about such models. To approach it, basic knowledge of descriptive and inferential statistics from a parametric perspective is needed.

Keywords: statistics, scientific research, university textbooks, non-parametric methods, hypothesis testing, problem situations.

TABLA DE CONTENIDO

Introducción 7

1. Aspectos teóricos y metodología utilizada 11
1.1 Aspectos teóricos 11
1.2 El problema por investigar 23
1.3 Metodología y resultados 33

2. La estadística no paramétrica en los libros de texto universitarios 45
2.1 Reportes de la literatura en torno a la estadística no paramétrica 46
2.2 La estadística no paramétrica en los libros de texto del nivel universitario 54
2.3 La estadística no paramétrica en el currículo universitario 64

3. Situaciones problema con una muestra para promover la investigación en el aula 67
3.1 Situaciones alusivas al trabajo con datos provenientes de una variable de tipo cualitativo población-muestra 67
3.2 Prueba de hipótesis basada en la distribución binomial 76
3.3 Prueba chi-cuadrado para una muestra 87
3.4 Prueba K-S de Kolmogorov-Smirnov para una muestra 93

4. Situaciones problema con dos muestras para la prueba de hipótesis en investigación formativa 103
4.1 Situaciones referidas a la prueba de McNemar 103
4.2 Prueba del signo 108
4.3 Prueba de la mediana 117
4.4 Prueba del rango signado de Wilcoxon 120
4.5 Prueba U de Mann-Whitney 128

5. Situaciones problema no paramétricas relacionadas con más de dos muestras 137
5.1 Prueba de Kruskal-Wallis 137
5.2 Prueba de Friedman con k muestras 146
5.3 Coeficiente de correlación de Spearman 156
5.4 Prueba de independencia para tablas de contingencia 164

6. Conclusiones 169
6.1 Hallazgos y discusión de resultados 169
6.2 Principales conclusiones 173

Referencias 176

Introducción

Desde tiempos inmemoriales, las personas han desarrollado actividades diversas relacionadas con los modos de producción de bienes y servicios, las cuales también han incluido procesos técnicos para recolectar, organizar, representar e interpretar diferentes tipos de datos relativos a la observación reiterada de los fenómenos presentes tanto en la naturaleza como en la sociedad. Con el paso del tiempo, tales técnicas fueron sistematizadas en un campo de conocimiento llamado estadística, inicialmente asumida por el Estado (Burbano et al., 2021); así, por medio de técnicas estadísticas el Estado cobraba los impuestos, controlaba la natalidad, administraba la riqueza y la ejecución de diferentes procesos productivos, entre otras actividades (Rioboó et al., 1997).

En los albores del siglo XX, con la formulación axiomática de la probabilidad emergieron y se consolidaron los procesos de inferencia estadística con los trabajos de Pearson, Neyman y Fisher (citados por García y Ríos, 1998); así, la estadística descriptiva, que había prevalecido como una técnica del método científico, empezó a fortalecerse apoyándose en las distribuciones de probabilidad para generar inferencias sobre los parámetros de una población y las pruebas de hipótesis que los incluyeron (Hurtado y Silvente, 2012).

Tomando en cuenta que, actualmente, la inferencia estadística paramétrica se ha consolidado como una herramienta fundamental para el método científico experimental, los científicos y académicos han recomendado incluir su estudio en los planes curriculares de diferentes carreras universitarias y, por supuesto, se han escrito libros de texto de estadística para el nivel universitario con el propósito de socializar este tipo de saber requerido en la formación universitaria y en el ejercicio profesional (Vera et al., 2021).

Formalmente, la estadística inferencial paramétrica se ocupa de elaborar procesos para estimar los parámetros de una población especificada y de efectuar pruebas de hipótesis a partir de los datos de una muestra aleatoria proveniente de una población con una distribución de probabilidad F que ha de cumplir determinados supuestos en cuanto a su forma (Gutiérrez y De la Vara, 2008); sin embargo, no siempre se pueden tener tales supuestos y resulta conveniente utilizar modelos estadísticos no paramétricos, los cuales asumen supuestos más flexibles y estadísticas basadas en rangos como una alternativa para efectuar la prueba de hipótesis (Corzo, 2005) inherentes a un proceso de investigación.

En estas circunstancias, en el ámbito universitario resulta conveniente examinar contenidos de estadística no paramétrica tendientes a promover la investigación en el aula con el apoyo del profesor y los libros adecuados. Aunque, al parecer, los libros de texto universitarios de estadística analizan principalmente tópicos de estadística descriptiva, probabilidad, distribuciones de probabilidad, inferencia estadística paramétrica, modelos de regresión y elementos de diseño experimental, entre otros, y otorgan menos oportunidad al estudio de contenidos referidos a los modelos estadísticos no paramétricos, considerados necesarios para desarrollar investigación en ciencias humanas, sociales, naturales y aplicadas (Ríos y Peña, 2020) desde perspectivas alternativas a la visión paramétrica.

Con la finalidad de mitigar esta problemática, desde el Grupo GICI (Grupo Interdisciplinario en Ciencias) avalado por la Universidad Pedagógica y Tecnológica de Colombia y reconocido en la categoría B por Minciencias, se ha decidido indagar sobre la manera como los libros de texto universitarios plantean los métodos estadísticos no paramétricos y, en ese sentido, proponer algunas situaciones problema que incluyan tales métodos con la intención de promover la investigación formativa y científica en el estudiantado. Además, esta obra puede aportar elementos conceptuales y didácticos que potencien los procesos de inferencia centrados en la prueba de hipótesis con estos métodos.

La metodología usada para este trabajo recoge elementos de la investigación cualitativa y cuantitativa, que se conjugan en un enfoque de corte mixto (Hernández y Mendoza, 2018). La parte cualitativa se centró en el uso de técnicas de investigación documental (Uribe, 2005) y del "análisis textual de contenido" (Bardín, 1986), cuyos resultados generaron cuatro categorías de conocimiento. La parte cuantitativa recoge las situaciones problema planteadas y tratadas a través de un algoritmo adoptado de Burbano y Valdivieso (2016), para llevar a cabo la prueba de hipótesis en consonancia con los métodos no paramétricos. Tales situaciones forman parte de los resultados del trabajo.

Esta obra fue estructurada en seis capítulos. En el primero se exponen los aspectos teóricos y la metodología empleada en el trabajo investigativo. En el segundo se explicita el primer resultado de investigación (primera categoría) emergente del análisis textual, que fue codificado como *la estadística no paramétrica en los libros de texto universitarios*. En el tercero se indica la segunda categoría (resultado 2), titulada *situaciones problema con una muestra para promover la investigación en el aula*. En el cuarto se describe la tercera categoría (resultado 3), llamada *situaciones problema con dos muestras para la prueba de hipótesis en investigación formativa*. En el quinto se expone la cuarta categoría (resultado 4), codificada como *situaciones problema no paramétricas asociadas a más de dos muestras*. Y en el sexto se presentan los principales hallazgos y las conclusiones.

Esta obra está dirigida a las personas motivadas por iniciarse en los modelos estadísticos no paramétricos clásicos o en usarlos para probar hipótesis en sus trabajos de investigación; en particular, puede resultar de gran ayuda para los estudiantes y profesores universitarios interesados en realizar pruebas de hipótesis con técnicas estadísticas alternativas a los métodos paramétricos; así mismo, para diversos investigadores y profesionales que deseen incrementar sus conocimientos sobre la estadística no paramétrica.

Los autores de la obra son docentes-investigadores del proyecto con SGI 2605, dirigido a investigar sobre el pensamiento aleatorio, los sistemas de datos y la aplicación de la estadística. Tal proyecto fue desarrollado según la modalidad de convocatoria abierta por la Vicerrectoría de Investigaciones (VIE) de la UPTC y, por consiguiente, tuvo el aval respectivo de su Comité de Ética.

1

Aspectos teóricos y metodología utilizada

En el presente capítulo se examinan tres subcategorías —los aspectos teóricos, el problema que guía la investigación y la metodología asumida—, que han proporcionado los elementos teóricos para la consolidación del proceso investigativo de este trabajo. En la primera se exponen aspectos conceptuales de la estadística tanto descriptiva como inferencial y se indaga sobre el tratamiento de la estadística no paramétrica en los libros de texto universitarios. En la segunda se indican diferentes aspectos de tipo contextual que han contribuido a identificar y justificar el problema de investigación. Y en la tercera subcategoría se describe la metodología empleada, que contiene una trayectoria para desarrollar el proceso investigativo.

1.1 Aspectos teóricos

Desde hace muchos años, el ser humano ha realizado actividades relacionadas con la recolección, organización, representación e interpretación de la información proveniente de la observación sistemática de los fenómenos naturales y sociales que ocurren en su entorno inmediato. Con el transcurso del tiempo, estas actividades fueron enmarcadas en un campo particular del conocimiento denominado estadística; al principio, este campo fue asumido como una actividad administrativa propia del Estado (Burbano et al., 2021); así, mediante la estadística, el Estado podía cobrar impuestos, controlar la natalidad, administrar la riqueza y los procesos productivos focalizados en la fabricación y comercialización de artículos de consumo, el conteo del ganado y el almacenamiento de los productos agrícolas, entre otras actividades (Rioboó et al., 1997).

A lo largo de varios siglos, la estadística se fue consolidando como una técnica para analizar los datos provenientes de diversas fuentes. Paulatinamente, se constituyó en una herramienta del método científico experimental y en la actualidad se ha establecido como la “ciencia de los datos” (Gunderson y Aliaga, 2005). Por otro lado, desde hace miles de años el hombre ha tratado de comprender y cuantificar los fenómenos no deterministas (aleatorios) en un intento por asimilar el azar y la idea de lo probable, presentes tanto en la naturaleza como en la sociedad. En estas circunstancias se empezó a gestar el concepto de probabilidad, el cual estuvo relacionado inicialmente con diversas creencias, aspectos morales y filosóficos, pero también fue mutando hacia procesos aritméticos relacionados con el reparto de apuestas en los juegos de azar y suerte (Burbano y Valdivieso, 2020).

Desde un punto de vista matemático, la probabilidad empieza a surgir con la búsqueda de soluciones al problema de reparto de apuestas provenientes de la práctica reiterada de los juegos de azar. A finales de la Edad Media, con los trabajos de Tartaglia, Pascal y Fermat, entre otros, se inicia el cálculo de probabilidades. Posteriormente, matemáticos como Bernoulli, Bayes, De Moivre, por mencionar solo algunos, contribuyeron con sus trabajos investigativos para que Laplace en 1812 formulara el concepto de probabilidad clásica restringido a espacios muestrales finitos y al concepto de equiprobabilidad. En el siglo XX, la probabilidad alcanza finalmente el estatus de rama de las matemáticas con la definición axiomática debida a Kolomogorov (1933); en este contexto, la probabilidad es entendida como una medida numérica de la posibilidad de que ocurra un evento (Martínez y Caballero, 2016).

Durante más de dos mil años predominó la estadística descriptiva destinada a realizar procesos de indagación centrados en la recolección de información y el procesamiento de esta a través de la organización de los datos, la representación mediante diversos tipos de gráficos, el análisis y la interpretación de resultados, a fin de generar algunas conclusiones fundamentadas en la exploración de los datos (Burbano y Valdivieso, 2016); aunque, solamente en la primera mitad del siglo XX fue posible establecer un puente entre la probabilidad y la estadística descriptiva. En efecto, gracias a los trabajos de Pearson, Neymar y Fisher se generaron procesos de tipo estadístico que dieron origen a una teoría sólida para desarrollar procesos inferenciales a través del testeo de hipótesis. Para ello, se asume una determinada distribución de probabilidad para los datos de una variable de interés en una muestra aleatoria tomada de una determinada población (García y Ríos, 1998).

De forma intuitiva, la estadística inferencial se usa para obtener conclusiones generalizables y válidas para una población delimitada como objeto de estudio, con base en una muestra aleatoria seleccionada de forma pertinente de esa población (Hurtado y Silvente, 2012). Formalmente, esta estadística se dirige a realizar procesos para estimar los parámetros (valores desconocidos) de una población especificada y a efectuar pruebas de hipótesis (docimasia) que contengan dichos parámetros (Gutiérrez y De la Vara, 2008). Los parámetros son objetos matemáticos que corresponden a ciertos valores, los cuales se pueden considerar verdaderos, únicos y válidos para toda la población objeto de investigación; se calculan con los datos de la variable de interés en los individuos de una población específica. En inferencia estadística, por lo general, los parámetros son desconocidos y se estiman siguiendo diversos métodos (Burbano y Valdivieso, 2016).

De acuerdo con Corzo (2005), un proceso de inferencia estadística tiene que ver con la extracción de una muestra aleatoria X de una población determinada, la cual presenta una distribución de probabilidad F (distribución muestreada) con uno o más parámetros que se desconocen. Sobre tales parámetros se pueden efectuar tres clases de inferencia estadística: i) estimación puntual, ii) estimación por intervalos de confianza, y iii) prueba de hipótesis, las cuales contienen tales parámetros.

Generalmente, los métodos para realizar los procesos de inferencia requieren que se cumplan ciertos supuestos referidos a la forma de la distribución F (simetría, apuntamiento y centralidad, entre otros) con el propósito de construir regiones críticas o de rechazo de la hipótesis nula. Cada vez que la hipótesis nula sea rechazada, resulta razonable aceptar la hipótesis alternativa o de investigación.

Según Burbano y Valdivieso (2016), una hipótesis estadística es una afirmación o conjetura que el investigador hace sobre los parámetros de la población por estudiar, con el fin de establecer si estos han cambiado o se mantienen en un valor fijo. Un proceso de inferencia estadística suele incluir dos clases de hipótesis: la nula, que se acostumbra a simbolizar con H_0, y la hipótesis alternativa, que se puede denotar con H_1. La hipótesis nula se formula en términos en que el parámetro sea al menos igual a un valor; esta hipótesis es la que se ha de aceptar o rechazar (Krueger, 2001). La hipótesis alternativa se plantea en términos en que el parámetro ha disminuido, ha aumentado o ha resultado diferente al valor consignado en la hipótesis nula. Este proceso conduce a tener una prueba unilateral izquierda, unilateral derecha o bilateral.

Krueger (2001) afirma que en los procesos inferenciales pueden cometerse dos tipos de errores: el error tipo I se comete cuando se rechaza la hipótesis nula, dado que esta era verdadera; la probabilidad de cometer este tipo de error se denota con α e indica la probabilidad de rechazar H_0 siendo cierta; se considera razonable que α sea menor o igual que 0.05; por lo tanto $1 - \alpha$ corresponde a la probabilidad de aceptar la hipótesis nula, dado que es cierta. Por otra parte, el error tipo II suele cometerse cuando se acepta la hipótesis nula, dado que aquella es falsa; la probabilidad de cometer este otro tipo de error suele denotarse con β; en consecuencia $1 - \beta$ corresponde a la probabilidad de rechazar H_0 dado que es falsa (Krueger, 2001). Por consiguiente, en un proceso de prueba de hipótesis las conclusiones obtenidas pueden estar afectadas por estos tipos de errores.

En el ámbito de la estadística inferencial suele realizarse el proceso de prueba de hipótesis desde una mirada paramétrica o no paramétrica. La estadística inferencial paramétrica incluye el cumplimiento estricto de algunos supuestos sobre la forma de la distribución de probabilidad muestral, entre ellos, la simetría, la normalidad de los datos, el apuntamiento, la centralidad. De acuerdo con Delicado (2008), los modelos estadísticos para desarrollar procesos de inferencia estadística desde un enfoque paramétrico comprenden los siguientes elementos: i) una variable aleatoria real X con función de distribución de probabilidad F, ii) la distribución F ha de pertenecer a una familia de distribuciones de probabilidad, la cual debe poderse indexar a través del parámetro Θ con dimensión finita; este hecho se denota de la siguiente manera:

$$X \sim F \;\text{ con }\; F \in \Im_{\Theta} = \left\{ F_{\theta} : \theta \in \Theta \subseteq \mathbb{R}^{k} \right\}$$

En este contexto, iii) $\Im_{\Theta}$ suele denominarse el modelo estadístico paramétrico.

Por otro lado, la estadística no paramétrica incluye un conjunto de métodos para procesos de inferencia estadística focalizados en métodos de prueba de hipótesis anclados en modelos estadísticos no paramétricos; en este contexto, una variable aleatoria X está regida por un modelo no paramétrico si sobre su distribución de probabilidad F se hacen muy pocos supuestos sobre la forma de F (Corzo, 2005) o se debilitan los supuestos sobre F para hacer la inferencia sobre los parámetros. Por ejemplo, se puede exigir que F sea una función de distribución absolutamente continua o que F sea simétrica en torno a su mediana, simplemente que la función de densidad proveniente de F tenga dos derivadas continuas (Delicado, 2008). En esta situación, en diversos casos se utilizan estadísticas (variables aleatorias) cuya distribución muestral no depende de F; es decir, las pruebas o los contrastes de hipótesis son libres de la distribución de los datos, también son conocidas como *distribution-free tests*.

Los modelos estadísticos no paramétricos aparecieron durante la segunda mitad del siglo XX. Después de 1950 surgieron los métodos paramétricos clásicos, varios de los cuales se fundamentan en estadísticas de rangos y se constituyen en una alternativa para testear hipótesis sobre la media, la mediana, diferencia de medias, análisis de varianza cuando los datos están en una escala ordinal; además, se han generado pruebas específicas en aquellos casos en los que los datos se encuentran en una escala nominal (Gibbons, 1997). Entre los modelos no paramétricos clásicos están los basados en la distribución binomial, la prueba del signo, la del rango signado de Wilcoxon, la prueba de Mann-Whitney, la de Kruskal-Wallis, la de Friedman, el coeficiente de correlación de Spearman y la prueba K-S de Kolmogorov-Smirnov, por mencionar solo algunas (Lehmann y D'Abrera, 1975).

Durante el último tercio del siglo XX surgieron modelos no paramétricos más complejos, destinados a la estimación no paramétrica de funciones de densidad (curvas), a la valoración de funciones de regresión, verosimilitud local, estimación por splines, modelos semiparamétricos, entre otros (Wasserman, 2006). La investigación con este tipo de modelos se ha fortalecido por cuanto actualmente es posible apoyarse en *software* estadístico especializado y programas computacionales para realizar el respectivo testeo de las hipótesis. Los métodos no paramétricos clásicos se han constituido en una herramienta poderosa para la investigación formal desde una mirada cuantitativa (positivista) en áreas como las ciencias del comportamiento, las ciencias sociales y humanas (Siegel, 1970), la administración de empresas (Lind et al., 2015), las ingenierías (Olaya, 2012; Canavos, 1988), entre otras.

Enseguida, se abordan algunos aspectos conceptuales acerca del desarrollo de procesos de inferencia estadística: variable aleatoria, función de distribución de probabilidad, población, individuo, variable, dato, muestra, muestra aleatoria, rangos en una muestra aleatoria, estimadores.

Según Burbano y Valdivieso (2020), dado un espacio de probabilidad $(\Omega, \Im, P)$ donde Ω es un conjunto no vacío que contiene todos los resultados de un experimento aleatorio denominado espacio muestral, $\Im$ es una familia de subconjuntos del espacio muestral con una estructura de sigma álgebra y P una medida de probabilidad definida desde el sigma álgebra hacia el intervalo desde cero hasta uno. Entonces, una variable aleatoria real X corresponde a una función medible definida desde el espacio muestral Ω hacia el conjunto de los números reales, de tal modo que, para todo evento A en el sigma álgebra de Borel se ha de cumplir que la imagen inversa de A con la función X corresponde a un evento en el sigma álgebra $\Im$; es decir, la variable aleatoria se constituye en un mecanismo para transformar los resultados del espacio muestral en números reales.

Según Blanco (2004), una función de distribución F para la variable aleatoria X es una función definida desde los números reales hacia el intervalo [0,1] que debe cumplir las siguientes condiciones: i) $F(x)$ tiene valores desde cero hasta uno, para todo número real x, ii) $F(x)$ es monótona no decreciente, iii) *F(x)* es una función real continua por la derecha, iv) $F(\infty) = \underset{x \to \infty}{Lim} F(x) = 1$, v) $F(-\infty) = \underset{x \to -\infty}{Lim} F(x) = 0$.

La población se constituye por un conjunto de individuos, sobre el cual recae un proceso investigativo en torno a una o varias características de acuerdo con un objetivo específico; estadísticamente, tal población se considera finita (de tamaño *N*) siempre y cuando se puedan listar todos los individuos que la conforman, en otro caso, se dice que esa población es infinita (Burbano et al., 2021). Cada individuo es un elemento metodológico para ejecutar el proceso de indagación; estos individuos pueden corresponder a objetos, personas, animales, entre otros, con características comunes por ser investigadas en concordancia con un objetivo (Valdivieso, 2011); tal objetivo se utiliza para delimitar el objeto de estudio en el tiempo y en el espacio.

Las variables son cada una de las características que interesa analizar o investigar en los individuos que constituyen una población o conforman una muestra representativa de aquella. Cuando se trabaja con una muestra y los datos de una variable específica, se calculan determinados valores denominados estimadores, los cuales, al cambiar de una muestra a otra, se constituyen en variables aleatorias susceptibles de tener una función de distribución de probabilidad, llamada distribución muestral (Lindgren, 1993). Una muestra aleatoria, llamada también muestra probabilística, se conforma por "*n*" individuos (tamaño *n*) seleccionados a través de un diseño muestral que incluye un muestreo probabilístico; por ejemplo, un muestreo aleatorio simple, estratificado o por conglomerados (Lind et al., 2015).

Los datos corresponden a cada uno de los valores admitidos para una variable de interés, los cuales, con frecuencia, se denotan con letras minúsculas acompañados de subíndices; por ejemplo, la secuencia $x_1, x_2, x_3..., x_n$ representa los datos de una variable *X* especificada.

Generalmente, las variables suelen clasificarse en cuantitativas y cualitativas. Una variable se define como cualitativa o categórica cuando sus datos permiten clasificar a los individuos en grupos mutuamente disjuntos (excluyentes) llamados categorías; para este tipo de variable, los datos se miden en una escala nominal. Una variable cuantitativa mide cantidades o intensidades de la característica por estudiar. Las variables cuantitativas suelen clasificarse en discretas y continuas. Una variable *X* es discreta cuando admite valores en un conjunto finito o en un conjunto contable de números reales; en particular, toma valores en el conjunto de números enteros; por ejemplo, puntajes de cierta variable (Siegel, 1970). Este tipo de variable se puede medir en una escala ordinal, lo cual implica que sus datos son susceptibles de ser ordenados tanto de menor a mayor como de mayor a menor.

Una variable *X* se denomina continua cuando admite cualquier valor dentro de un intervalo de números reales. Este tipo de variable puede medirse en escala de intervalo cuando el cero es relativo, lo cual quiere decir que el cero no denota ausencia de la característica que se esté midiendo; por ejemplo, al considerar los datos de la variable temperatura, el dato cero grados centígrados no significa ausencia de temperatura. Además, una variable continua está en escala de razón cuando el cero es absoluto, lo que muestra que el cero expresa ausencia de la cantidad de la característica que se esté midiendo; por ejemplo, si *X* representa las utilidades de una empresa en millones de pesos por mes, el valor cero implica que no se obtuvieron utilidades en ese mes (Burbano et al., 2021).

Una muestra aleatoria corresponde a un conjunto de "*n*" variables aleatorias $X_1, X_2, X_3..., X_n$, las cuales están idénticamente distribuidas; esto significa que tales variables son independientes y tienen la misma función de distribución *F* (Mayorga, 2003; Lindgren, 1993). Ahora, si $x_1, x_2, x_3..., x_n$ es una realización de tal muestra aleatoria (conjunto de valores de la variable *X*), entonces se ha de cumplir que la función de probabilidad conjunta se puede expresar con el producto de sus correspondientes marginales, o sea:

$$f(x_1, x_2, x_3, .., x_n) = f(x_1).f(x_2).f(x_3)...f(x_n)$$

Por ejemplo, si *X* es una variable aleatoria con distribución normal con media 90 y desviación estándar de 3, su función de distribución de probabilidad está dada por:

$$F(x) = P(X \le x) = \int_{-\infty}^{x} \frac{1}{3\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{t-90}{3}\right)^2\right] dt$$

La función de densidad de probabilidad para esta variable es la siguiente:

$$f(t) = \frac{1}{3\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{t-90}{3}\right)^2\right] \text{ con } t \in \mathbb{R}$$

Una muestra aleatoria $X_1, X_2, X_3..., X_n$ para estudiar una variable aleatoria *X* con la mencionada distribución de probabilidad para una realización $x_1, x_2, x_3..., x_n$ ha de satisfacer las siguientes igualdades (Burbano y Valdivieso, 2016):

$$f(x_1) = \frac{1}{3\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x_1 - 90}{3}\right)^2\right] \text{ con } x_1 \in \mathbb{R}$$

$$f(x_2) = \frac{1}{3\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x_2 - 90}{3}\right)^2\right] \text{ con } x_2 \in \mathbb{R}$$

De similar modo se prosigue hasta obtener:

$$f(x_n) = \frac{1}{3\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x_n - 90}{3}\right)^2\right] \text{ con } x_n \in \mathbb{R}$$

En los casos en los cuales haya necesidad de hacer un análisis que involucre dos o más variables con sus respectivas muestras aleatorias con el propósito de indagar sobre su efecto simultáneo, es conveniente efectuar un análisis multivariado de datos o el testeo de sistemas de hipótesis que pueden incluir dos o más parámetros (Hair y Taham, 2008), ya sea considerando modelos paramétricos o no paramétricos. En una muestra aleatoria $X_1, X_2, X_3..., X_n$ para estudiar una variable aleatoria X, se denota con R_i el rango del dato x_i dentro de la secuencia $x_1, x_2, x_3..., x_n$ y se define como el puesto que ocupa x_i en la sucesión ordenada $X_{(1)} < X_{(2)} < X_{(3)} < ... < X_{(n)}$ (Hettmansperger, 1984; Corzo, 2005).

Finalmente, un estimador corresponde a una variable aleatoria que se construye como una función de las variables que conforman una muestra aleatoria (Mayorga, 2003). El estimador construido no debe depender de parámetro alguno referido a la distribución de probabilidad F que representa a la población de donde se ha obtenido la muestra aleatoria (Lindgren, 1993); es decir, un estimador T se denota del siguiente modo:

$$T = f(X_1, X_2, X_3 ..., X_n)$$

Algunos estimadores usuales en inferencia estadística —tanto paramétrica como no paramétrica— son el promedio, la varianza, el porcentaje o proporción, el coeficiente de variación muestral, entre otros; los cuales se denotan con $\bar{X}$, S^2, $\hat{p}$, CV y definen como sigue (Burbano et al., 2021):

Promedio muestral: $$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Varianza muestral: $$S^2 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n}$$

El coeficiente de variación muestral: $$CV = \frac{S}{\bar{X}}$$

El porcentaje muestral: $$\hat{p} = \frac{x}{n}$$

donde x representa la cantidad de individuos con la característica X por estudiar en la muestra de tamaño n.

1.2 El problema por investigar

En esta sección se presentan diferentes aspectos de tipo contextual que han contribuido a identificar y justificar el problema de investigación.

Entre estos se pueden mencionar la necesidad de abordar algunos tópicos de estadística no paramétrica en el currículo universitario, el planteamiento de interrogantes que harían posible el desarrollo del trabajo investigativo y aspectos relativos al tratamiento del tema de modelos estadísticos no paramétricos en la literatura respectiva.

1.2.1 Razones para incluir tópicos de estadística no paramétrica en el currículo universitario

Los niveles de formación logrados por el profesor universitario en aspectos disciplinares y didácticos son elementos indispensables para promover el aprendizaje de la estadística y su aplicación a fin de fortalecer la investigación formativa desde el aula (Burbano, 2017). Desde este punto de vista, el sistema educativo universitario ha de generar escenarios adecuados para que sus profesores puedan acceder a procesos de cualificación continuada y contribuir de manera proactiva en la formación profesional del estudiantado y en su desempeño eficiente en la sociedad, la cual espera que los profesionales sean individuos competentes en el contexto real. Por esto se requiere que los futuros profesionales sean formados en las aulas no solo en aspectos académicos, sino tecnológicos y científicos, y que se propicien también estrategias de aprendizaje independiente para que estén en capacidad de continuar aprendiendo a lo largo de su vida laboral (Kant, 2003).

En particular, la estadística es un campo de conocimiento que, gracias a sus métodos cuantitativos específicos, significa un elemento valioso para que el futuro profesional pueda aportar soluciones a problemas reales mediante el procesamiento adecuado de la información inherente a la situación problemática. De allí que desde los salones de clase sea conveniente potenciar el pensamiento y el razonamiento estadístico, de modo que contribuyan a promover la investigación científica con un principio de autonomía (Ministerio de Educación Nacional [MEN], 2006; congreso de la República, 1992). Tanto la probabilidad como la estadística, ya sea descriptiva o inferencial, son herramientas conceptuales que permiten potenciar en el estudiantado su pensamiento aleatorio (estocástico) y el manejo de diversos sistemas de datos —entre ellos, las pruebas de hipótesis— requeridos para tomar mejores decisiones en una variedad de situaciones en las cuales haya presencia de azar, riesgo o incertidumbre (MEN, 2006).

El interés por impartir contenidos de estadística y probabilidad en la educación preuniversitaria y superior se gestó en el contexto internacional y paulatinamente se ha ido adoptando en el escenario nacional. La educación estadística (estocástica) recoge un conjunto de saberes que crecieron con los aportes de la misma estadística ante la necesidad de procesar y analizar datos, y por las estrategias provenientes de las matemáticas y su didáctica (Garfield y Ben Zvi, 2008). Actualmente, la educación estocástica reúne un amplio campo de trabajo investigativo en concomitancia con otras ramas del saber, entre ellas, las ingenierías, las ciencias naturales, las ciencias sociales y humanas, las matemáticas, la física, la biología, la psicología, la economía, la administración de empresas, las telecomunicaciones; además, su enseñanza ha sido auspiciada por organismos internacionales como la Asociación Americana de Estadística, el Instituto Internacional de

Estadística y la Asociación Internacional de Educación Estadística (Mason et al., 1990).

La educación estadística, además de ocuparse de los contenidos que han de ser enseñados en el aula, incluye procesos de investigación que abordan desde diversas perspectivas los procesos de enseñanza y aprendizaje (E–A) de la estadística y la probabilidad; además, se interesa por avanzar en el uso de métodos alternativos para la prueba de hipótesis tales como los modelos estadísticos no paramétricos, así como por intensificar la utilización del *software* estadístico —ya sea libre o con licencia— para implementar procesos investigativos apoyados en métodos cuantitativos de estadística. Por lo tanto, posibilita potenciar y mejorar la formación profesional del estudiantado universitario en diversas titulaciones (Pérez, 2005) con el fin de acrecentar sus competencias académicas, investigativas y laborales.

En el contexto internacional se han propuesto planes curriculares, materiales de apoyo didáctico y estrategias para promover el aprendizaje de la estadística desde la educación primaria hasta la universitaria; empero, también se ha identificado la necesidad de mejorar la formación profesional del profesorado en su conocimiento disciplinar y didáctico acerca de las distintas ramas de la estadística (Barnett, 1982). Estos aspectos deben ser impulsados por las universidades para mitigar procesos de E–A centrados en la mera instrucción, la transmisión mecánica de contenidos, el aprendizaje memorístico de fórmulas (Pinto, 2010) y abrirle paso al desarrollo de actividades dinámicas orientadas al aprendizaje significativo y contextualizado para el estudiante, de modo que paulatinamente lo conecten con los procesos de investigación científica (Burbano, 2017).

Por otro lado, a pesar de que las materias de estadística y probabilidad se han constituido en herramientas que acrecientan el razonamiento de las personas y posibilitan la solución de una variedad de problemas de corte cuantitativo que involucren el manejo de datos, en Colombia tanto el plan curricular en la educación secundaria como en la universitaria incluyen solamente unos contenidos básicos de probabilidad y estadística (Gómez, 2011), y no permiten la explotación de todo su potencial en esta sociedad informatizada, compleja y global, en la cual de forma cotidiana el futuro profesional requerirá tomar decisiones en contextos laborales marcados frecuentemente por la incertidumbre y soportados por gigantescas cantidades de información que deberá procesar y utilizar junto con las tecnologías de la información y de las comunicaciones (TIC) a fin de generar soluciones a diversos problemas (Burbano y Valdivieso, 2020).

Como era de esperarse, la inferencia estadística con sus métodos paramétricos y no paramétricos se aplica en distintos contextos, entre ellos, en el ejercicio profesional en la vida real, en el pronóstico de consecuencias en situaciones azarosas, en el control de calidad de los productos y servicios, su confiabilidad y durabilidad (Burbano et al., 2019). Los procesos de testeo de hipótesis efectuados a partir de muestras aleatorias para una variable de interés, ciertos procesos de simulación soportados en el uso de números aleatorios y el análisis multivariado de datos requieren un conocimiento más profundo de la estadística; en este contexto, un profesor universitario con un alto nivel de dominio de esta materia es un elemento fundamental para planear los contenidos y apoyar el aprendizaje estadístico con variados recursos didácticos como los libros de texto universitarios y los programas de computador pertinentes.

Adicionalmente, el profesor ha de seleccionar adecuadamente los libros de texto, de modo que se conviertan en un elemento de apoyo para el aprendizaje estudiantil; aun cuando, los libros presentan diferentes niveles de complejidad en el tratamiento de los distintos temas; unos se focalizan más en la estadística descriptiva, la probabilidad, la estadística inferencial, y otros posiblemente incluyan tópicos alusivos al manejo de los modelos estadísticos no paramétricos. Por consiguiente, el profesor será el encargado de dar una graduación pertinente a las temáticas por abordar; por ejemplo, para los cursos universitarios introductorios se recomienda trabajar elementos de estadística descriptiva e inferencia estadística básica, fundamentadas en los modelos paramétricos, y para un curso con mayor profundidad se pueden examinar los tópicos referidos al análisis multivariado de datos mediante técnicas paramétricas y no paramétricas (Burbano et al., 2021).

Con frecuencia, ciertos profesores universitarios consideran que con los elementos básicos referidos a probabilidad y estadística, los estudiantes podrán cursar con éxito las demás asignaturas que requieren el uso de estas materias en su plan curricular. Otros argumentan que la inferencia estadística que incluya los métodos no paramétricos para probar hipótesis o la definición de medidas de probabilidad desde el enfoque axiomático resultan demasiado teóricas y pueden ser estudiadas en los cursos de maestría en los cuales se requieran (Pérez, 2005; Burbano y Valdivieso, 2020). Además, una gran cantidad de textos de estadística del nivel universitario plantean de forma incompleta o sesgada las pruebas de hipótesis (Krueger, 2001), abordan la probabilidad desde la mirada clásica de Laplace (Sánchez y Monroy, 2013) o no consideran a cabalidad la forma de la distribución de probabilidad *F* para el testeo de hipótesis o el uso de métodos estadísticos no paramétricos (Corzo, 2005).

Finalmente, en relación con el proceso E–A de la estadística y la probabilidad, un sinnúmero de profesores universitarios explicitan sus creencias de que un conocimiento estadístico-probabilístico profundo será suficiente para enseñar las asignaturas de este tipo en la universidad; además, asumen metodologías de enseñanza similares a las utilizadas para enseñar el álgebra, la geometría, la física, la ingenierías, sin reflexionar que tales metodologías son de tipo determinista y excluyen el tratamiento impredecible de los fenómenos aleatorios o la persistencia de la variabilidad en los datos estadísticos. Estas concepciones requieren ser revaluadas, por cuanto en los últimos tres lustros la investigación educativa ha mostrado que la problemática del aprendizaje estudiantil está relacionada con aspectos de tipo cognitivo, epistemológico, social, didáctico y de aplicación del conocimiento en diversos contextos (Salinas y Alanís, 2009). Este tipo de creencias se perpetúan aún más en la eventualidad de que los libros de texto universitarios con los cuales los profesores preparan sus clases puedan presentar inconsistencias, sesgos o interpretaciones inadecuadas ya sea dentro de la estadística descriptiva (Mayen, 2009) o de la estadística inferencial, y particularmente en lo referente al uso de modelos estadísticos no paramétricos (Delicado, 2008).

1.2.2 Cuestionamientos iniciales para orientar el trabajo investigativo

La ley 30 de 1992 es el marco regulatorio de la autonomía universitaria, y en consonancia con ella el profesor universitario tiene libertad para desarrollar sus cátedras, tomar decisiones sobre las actividades académicas e investigativas, a fin de que el estudiantado pueda desarrollar sus competencias generales y específicas de acuerdo con las asignaturas que conforman el plan de estudios de una determinada carrera universitaria. En este contexto, en el currículo de diversas titulaciones se incluyen elementos de estadística y probabilidad atendiendo las necesidades de formación en correspondencia con el perfil profesional; asimismo, el docente puede seleccionar los libros de texto universitarios donde al estudiantado le sea posible complementar su proceso de aprendizaje y acrecentarlo con sus dinámicas de aprendizaje independiente; en este sentido, un adecuado libro de texto también puede facilitar el aprendizaje de temas puntuales sobre inferencia estadística, como lo es el testeo de hipótesis. En estas circunstancias, un adecuado libro de texto significa un recurso didáctico tanto para el estudiante como para el profesor.

En este orden de ideas, los primeros interrogantes que se plantearon en este trabajo fueron: ¿cuáles ejes temáticos de inferencia estadística son frecuentes en los libros de texto del nivel universitario? En los textos universitarios ¿cuáles temáticas de estadística descriptiva, probabilidad o inferencia estadística son tratadas? En los libros clásicos de estadística universitarios ¿se examinan con suficiente profundidad las medidas de probabilidad y las funciones de distribución?

Por otra parte, refiriéndose a los libros de uso común en el aula, Occeli y Valeiras (2013) señalan que los libros pueden constituirse en instrumentos de mediación en el proceso E–A, los cuales posibilitan la traducción de significados y la concreción de los objetivos establecidos en el plan curricular. Desde esta perspectiva, los libros de texto universitarios permiten presentar unos contenidos específicos con una visión teórico-didáctica singular, destinada a promover la adquisición de conocimientos y el desarrollo de competencias en el estudiantado en concomitancia con lo establecido en el perfil profesional. Ante esto, otra pregunta que surgió fue la siguiente: ¿cuáles aspectos didácticos son observables en los libros de estadística del nivel universitario?

Adicionalmente, y siguiendo las recomendaciones de Wild y Pfannkuch (1999) de que en la enseñanza de la estadística se ha de utilizar un conjunto de pasos para que el estudiantado los aplique en contextos reales y potencie su procesos de investigación formativa, se plantearon las siguientes preguntas: ¿en los libros universitarios de estadística se observa el desarrollo de actividades propias del ciclo investigativo constituido por: recolección de información, organización, formulación de hipótesis, prueba y obtención de conclusiones? ¿De qué manera son expuestos los problemas reales en los libros universitarios de estadística? ¿Se da prelación a los ejercicios académicos frente a la formulación y solución de problemas reales en los cuales se ha de usar la inferencia estadística? ¿Se incluyen en los libros clásicos suficientes tópicos relacionados con el uso de los modelos estadísticos no paramétricos?

Como era de esperarse, las respuestas a las preguntas formuladas no son inmediatas; por lo tanto, una exploración minuciosa de diferentes libros universitarios de estadística puede dar luces para responder a tales interrogantes; en especial, a la siguiente pregunta general que orientó el trabajo investigativo: ¿cuáles textos universitarios contemplan modelos estadísticos no paramétricos como alternativa a los modelos paramétricos clásicos? Para resolver esta pregunta fue necesario inicialmente implementar algunas actividades de investigación documental consistentes en: i) observar las características propias de los libros de texto universitarios utilizados como recursos didácticos para apoyar el aprendizaje de la inferencia estadística, ii) generar reflexión en los profesores universitarios sobre la necesidad de emplear secuencias didácticas en el aula de modo que potencien el uso de la estadística tanto desde el punto de vista académico como investigativo, y iii) elaborar textos didácticos donde se haga un tratamiento especializado de, al menos, las pruebas clásicas no paramétricas.

Por otro lado, se sugiere que en un próximo trabajo investigativo se explore sobre el conocimiento estadístico del profesor universitario en torno a los modelos no paramétricos y a su uso en proyectos de investigación en diversas áreas del saber. Además, se plantea la necesidad de elaborar y publicar libros de estadística como resultado de proyectos de investigación que permitan paulatinamente hacer una transformación en las maneras tradicionales de abordar la inferencia estadística. Para esto pueden utilizarse los modelos y marcos teóricos planteados en Burbano *et al*. (2017), entre otros.

1.3 Metodología y resultados

En la elaboración de los trabajos de investigación es posible adoptar distintas metodologías, las cuales deben ser concordantes con la pregunta planteada para orientar la ejecución del proceso investigativo. De ahí que se requiera utilizar un marco teórico que permita asumir un diseño metodológico constituido por un conjunto de procedimientos y actividades para obtener los resultados esperados en el proceso investigativo (Burbano et al., 2019). Para el caso de esta obra, el proceso inició con la revisión de literatura acerca del tratamiento de tópicos de estadística inferencial y de estadística no paramétrica en los libros de texto universitarios; en particular, lo referente a las pruebas de hipótesis mediante modelos estadísticos no paramétricos.

Para ello, fue conveniente adoptar un método mixto de investigación (Hernández y Mendoza, 2018), el cual conjugó elementos de investigación cualitativa relativos al tratamiento de datos textuales provenientes de fuentes diversas y la emergencia de categoría en el análisis de los libros de texto, así como métodos estadísticos para implementar procesos de inferencia estadística focalizados en socializar un conjunto de situaciones problema en el uso de modelos estadísticos no paramétricos, los cuales pueden promover la investigación científica desde las aulas universitarias. La investigación cualitativa se centró en un "análisis de contenido" (Bardin, 1986) a diversos textos de estadística de potencial utilización en la educación superior; asimismo, se utilizaron técnicas provenientes del ámbito de la investigación documental (Uribe, 2005). Algunos de los resultados que se presentan en las situaciones problema sobre prueba de hipótesis fueron obtenidos al aplicar elementos del círculo hermenéutico (Burbano, 2017) acompañados de la experiencia docente y la reflexión continua de los investigadores.

Para los procesos de inferencia estadística asociados a las situaciones problema, se aplicaron los métodos no paramétricos en el testeo de hipótesis, tanto para una muestra como para dos, tres o más muestras. En este contexto, se asumió un algoritmo semejante al expuesto por Burbano y Valdivieso (2016), constituido por los siguientes pasos:

1) Plantear el sistema de hipótesis por comprobar

H_o: $\Theta = \Theta_o$
H_1: $\Theta > \Theta_o$
H_o: $\Theta = \Theta_o$
H_1: $\Theta < \Theta_o$
H_o: $\Theta = \Theta_o$
H_1: Θ diferente de Θ_o

Donde Θ representa al parámetro de interés y Θ_o un valor de dicho parámetro.

2) Fijar el nivel de significancia para hacer la prueba de hipótesis. Tal nivel debe corresponder a un valor α menor o igual que 0.05.

3) Identificar la dirección de la prueba, a fin de establecer la región crítica o de rechazo de la hipótesis nula. Aquí, el primer par de hipótesis corresponde a una prueba unilateral derecha; el segundo par, a una prueba unilateral izquierda, y el tercer par, a una prueba bilateral.

4) Seleccionar una estadística de prueba en correspondencia con los datos provenientes de la muestra.

5) Tomar una decisión: si el estadístico de prueba cae en la región crítica, entonces se rechaza la hipótesis nula; de forma equivalente, si el P-valor asociado a la estadística de prueba es menor que el nivel de significancia α, entonces se rechaza la hipótesis nula. Así, cada vez que se rechace la hipótesis nula, se ha de aceptar la hipótesis alternativa o de investigación, ya que estadísticamente no hay suficiente evidencia para rechazar la hipótesis alterna. También se puede estimar el respectivo intervalo de confianza para probar las hipótesis.

6) Obtener una conclusión.

Es conveniente mencionar que antes de aplicar el anterior algoritmo, el investigador debe interpretar la situación problema por resolver y establecer si la prueba de hipótesis se debe hacer a través de métodos paramétricos o por modelos estadísticos no paramétricos

1.3.1 Aspectos asociados a una línea de investigación

Este trabajo fue elaborado en el contexto de acción investigativa del Grupo de Investigación en Ciencias (GICI), una de cuyas líneas de investigación tiene que ver con la educación estadística; además, se relaciona con otra línea denominada estadística aplicada a las ciencias. El mencionado grupo ha sido clasificado en la categoría B por Minciencias de Colombia y pertenece a la UPTC. Con esta línea de indagación y otras asumidas por el Grupo GICI se busca promover los procesos de investigación científica desde las aulas universitarias. Este propósito permite tanto a los profesores como a los estudiantes enfrentar diversos retos, ya sean académicos o investigativos, generados en el seno de la sociedad actual, caracterizada por la globalidad, la complejidad y el uso frecuente de las TIC en los ámbitos universitario y laboral.

1.3.2 De los métodos a la acción con procedimientos y resultados

Para este trabajo investigativo se decidió utilizar métodos cuantitativos y cualitativos, a fin de evidenciar los hallazgos y soportar la obtención de resultados; por consiguiente, los procedimientos guardan relación con estos métodos, los cuales se enmarcan en un enfoque mixto de investigación (Hernández y Mendoza, 2018). Estos aspectos propios de la metodología son concordantes con los interrogantes planteados. Por otra parte, la pregunta global formulada que apunta al objeto de estudio y el conjunto de procedimientos permitieron generar acciones concretas de investigación, para indagar y documentar sobre las diferentes maneras en que los autores de los libros de texto universitarios abordan tanto la cantidad de contenidos como las estrategias propias para exponer las temáticas referidas a los métodos estadísticos no paramétricos. Asimismo, posibilitaron hacer aportes para el planteamiento y la solución de una cantidad considerable de situaciones problema desde una mirada cuantitativa y como resultado de una cuidadosa revisión de literatura y del análisis de contenido sobre distintas fuentes.

De acuerdo con Bardín (1986), el análisis de contenido es una técnica cualitativa dirigida a generar categorías y subcategorías al dar sentido e interpretar los datos textuales o información documental acerca de una temática de estudio; en este contexto, el investigador realiza actividades de identificación y clasificación de conceptos implícitos en fragmentos de textos, en partes de un libro o en libros completos, artículos indexados en revistas tanto científicas como de socialización del conocimiento, entre otros. Tales técnicas se centran en el conteo de frases, segmentos de texto acordes a una pregunta de investigación y número de palabras frecuentes dentro de un documento textual específico, un discurso o una estrategia que algún autor haya deseado compartir con los lectores a través de un escrito. En este trabajo dicho análisis se llevó a cabo de manera manual, aunque también es posible hacerlo mediante *software* especializado para tal propósito.

Un algoritmo implícito en un análisis de contenido puede incluir los siguientes pasos: i) codificar la información textual; esta actividad puede incluir una codificación selectiva, abierta o axial de los datos textuales; ii) conformar categorías y subcategorías, lo cual permite dar una mejor estructura a los hallazgos; iii) hacer inferencias de tipo textual; para esto el investigador interpreta cuidadosamente las piezas de información que guarden relación con el objeto de estudio, aquí puede elaborarse un ciclo hermenéutico (Burbano, 2017); iv) ejecutar acciones vinculadas con la investigación documental (ID), y v) dar un tratamiento informático a los documentos de estudio; cuando esto sea posible, por ejemplo, se puede emplear *software* como el Atlas.ti, XLStat, entre otros.

Según Bardín (1986), la acción de codificar es un procedimiento por medio del cual los datos brutos (sin ningún análisis previo) son transformados usando métodos sistemáticos en unidades que hacen posible una descripción detallada y precisa de las principales características que distinguen al contenido o conjunto de segmentos textuales por analizar. Las acciones relacionadas con la categorización implican la elaboración de un aislamiento de los principales elementos que caracterizan al texto que se esté analizando; además, permiten generar agrupaciones de elementos textuales en “categorías” de conocimiento, las cuales proporcionan mensajes de tipo cualitativo variables que se pueden sistematizar. Estos mensajes son el insumo principal para generar inferencias que, a su vez, se van a relacionar con otras categorías o subcategorías que puedan emerger.

Por otro lado, la investigación documental (ID) corresponde a un conjunto de técnicas para analizar e interpretar distintas fuentes ya sean documentales o textuales que puedan proporcionar diversas informaciones. Entre tales fuentes están los artículos, los libros, los textos universitarios y del nivel escolar, conferencias y videos. De acuerdo con Uribe (2005), las características principales de la ID son las siguientes: i) hace posible la revisión de las publicaciones que circulan por diferentes medios en lo referente a un tópico específico o a diversas temáticas relacionadas con un objeto de estudio, ii) la ID es un proceso con alto nivel de rigurosidad y sistematización, centrado en elaborar un análisis pertinente y crítico sobre la información disponible, y iii) es una actividad de tipo científico que asume procesos deductivos e inductivos tendientes a recopilar, sistematizar e interpretar datos cualitativos y cuantitativos con el propósito de producir constructos nuevos o reorganizar los ya existentes; aquí, se genera una interacción entre elementos de corte epistemológico y metodológico en referencia al objeto de estudio; en este sentido, el marco teórico brinda elementos conceptuales para establecer la metodología por seguir para obtener los resultados esperados en el proceso investigativo.

Además, según Guevara (2015), la ID también hace posible la selección de diversas fuentes de información para implementar procesos de clasificación con el propósito de fundamentar la indagación preliminar del tema central y luego proseguir con el uso de conceptos y teorías que permitan dilucidar aún más el tema de interés. En este mismo sentido, Ávila (2015) agrega que la ID permite estructurar y consolidar documentos nuevos escritos por el investigador, en los cuales se consignan nuevas descripciones, se hacen análisis complementarios, se elaboran comparaciones o se critica un tema determinado por medio del análisis minucioso de diversas fuentes adicionales a la información inicial. Además, según Peña y Pirela (2007), el análisis documental es una tarea compleja, ya que puede incluir aspectos de tipo epistemológico, idiomático y lingüístico, cognitivo, psicológico, formativo y socializador, entre otros.

En el presente trabajo investigativo, tanto la ID y el análisis de contenido como los métodos cuantitativos basados en procedimientos estadísticos para la prueba de hipótesis mediante modelos no paramétricos han sido herramientas indispensables para ofrecer respuestas a las inquietudes que paulatinamente les surgían a los investigadores; además, los motivaban constantemente para desentrañar la estructura implícita que no se exhibía a primera vista en los documentos y escritos consultados, de manera que poco a poco permitieron responder a los principales interrogantes que se plantearon al empezar el trabajo. Asimismo, estos métodos facilitaron la ubicación de algoritmos, vocablos, símbolos o segmentos de información relevante con los contenidos y estrategias de solución a las situaciones problema planteadas y resueltas por los mismos investigadores (López, 2002); por ejemplo, fue conveniente rastrear vocablos como hipótesis estadística, “nivel de significancia”, los rangos para los datos de una determinada muestra, región de rechazo, prueba no paramétrica, entre otros; igualmente, fue posible establecer códigos para visualizar la manera como los autores estaban abordando los temas de estadística no paramétrica en los libros de texto universitarios.

Por otra parte, en el presente trabajo también se recurrió a la aplicación del método inductivo-deductivo (de lo particular a lo general, de lo empírico hacia la obtención de conclusiones generales), en lo referente a la construcción de varias secuencias didácticas para abordar la prueba de hipótesis en una muestra, dos muestras y tres o más muestras, a partir de situaciones problema particulares; tales resultados se presentan en su orden en los capítulos tres, cuatro y cinco.

El razonamiento inductivo está presente en este trabajo, pues en cada prueba de hipótesis se inicia con la selección de una muestra aleatoria, se prosigue con el planteamiento de la hipótesis nula y alternativa, se selecciona la estadística de prueba pertinente a fin de aceptar o rechazar la hipótesis nula en correspondencia con una determinada región crítica, para finalmente obtener una conclusión generalizable a toda la población, en el sentido expuesto por Krueger (2001), sin descartar el uso de trayectorias deductivo-inductivas (Newman, 2006).

1.3.2 Los asuntos éticos en la elaboración del trabajo investigativo

Por la naturaleza del presente trabajo, el cual ha centrado su atención en indagar la forma como en los textos del nivel universitario se están usando los modelos estadísticos no paramétricos, este se enmarca en el campo de la educación estadística. Aunque, se aclara, sin desligarse de la idea de que la inferencia estadística no paramétrica debería formar parte de los cursos de estadística y probabilidad que se incluyen dentro del plan curricular de las diversas carreras (titulaciones) universitarias, considerando que estas materias contienen elementos conceptuales que aportan a la aplicación del método científico experimental a las ciencias exactas y naturales, y a las ciencias sociales y humanas, que trascienden los claustros universitarios.

En la elaboración del presente documento se asumieron las normas éticas socializadas en las diferentes convocatorias realizadas por la VIE y la Editorial de la UPTC, requeridas para los procesos de publicación de obras de investigación vinculadas con algún proyecto de investigación ejecutado. Se tuvieron presentes las normas sobre derechos de autor, tales como las consagradas en la Ley 23 de1982 (Burbano et al., 2021); se citaron autores y documentos según las normas de estilo APA (7.ª versión). Por otra parte, las personas que colaboraron en las diversas actividades del proceso investigativo —referidas a la ubicación y manipulación de diferentes fuentes documentales y a la participación en el análisis textual— siempre estuvieron conscientes del bajo riesgo al que podría estar expuesta su salud física, emocional y mental.

1.3.4 Procedimientos de actuación adicionales

Es pertinente indicar que la información recolectada (textual y cuantitativa) provino de fuentes de carácter documental tanto físicas (libros, boletines, revistas, entre otros) como electrónicas (consultas en internet), y los documentos recopilados fueron sometidos a procesos sistemáticos de análisis textual y documental. A los colaboradores de la parte operativa de este trabajo se les informó con anticipación que si se presentaba alguna eventualidad que conllevara fenómenos inesperados que representaran peligro o riesgo para su salud (física y mental), de forma inmediata debían suspender la actividad que se encontraran realizando con las mencionadas fuentes de información y acudir a la entidad prestadora de salud a la cual estuviesen afiliados.

Para alcanzar las metas propuestas del trabajo investigativo que originó la presente obra, se ejecutaron diferentes procedimientos adicionales, entre ellos, i) se socializaron los hallazgos y resultados parciales, ii) se estructuró un primer borrador del libro de investigación con capítulos y secciones donde se explicitaron los resultados del trabajo investigativo en forma de categorías y subcategorías emergentes del análisis de contenido efectuado, y iii) se hicieron las correcciones y puesta a punto de los contenidos que se exponen en esta obra. En cada etapa del proceso se tuvo presente la ejecución de cada actividad consignada en un cronograma que también guardaba relación con las estipuladas en el proyecto con código SGI 2805, denominado “El pensamiento aleatorio y los sistemas de datos”, adscrito a la VIE de la UPTC. Adicionalmente, se utilizó un diagrama de Gantt (ver Tabla 1) para planificar y monitorear las actividades macro organizadas así: la parte **E1** contiene aspectos relativos al proyecto con SGI 2605, y la parte **E2** explicita actividades relativas a la elaboración del presente libro de investigación.

Tabla 1. Diagrama de Gantt con actividades macro

Tipo de actividad	**Periodo de tiempo**
E1	2021-2022
Revisión de fuentes primarias, literatura especializada referida a los libros de texto escolar y del nivel universitario.	Abril de 2019 a noviembre de 2020
Lectura comprensiva de diferentes fuentes documentales tanto escritas como provenientes de Internet, para generar ideas fuerza sobre los modelos estadísticos no paramétricos.	Marzo-julio de 2021
Realización de los primeros análisis de contenido y aplicación de la ID para estructuras las primeras categorías y subcategorías.	Agosto-octubre de 2021
E2	
Tópicos de pruebas de hipótesis con métodos no paramétricos.	Noviembre–diciembre de 2021
Diversas pruebas de hipótesis mediante estadística no paramétrica para una, dos y más muestras.	Enero-febrero de 2022
Recopilación de los capítulos ya escritos y estructuración de un primer borrador de la obra.	Marzo–abril de 2022
Versión ajustada y corregida del libro de investigación.	Enero de 2023

Fuente: elaborado por los autores.

2

La estadística no paramétrica en los libros de texto universitarios

Los métodos estadísticos para la prueba de hipótesis que comprende uno o varios parámetros desconocidos se seleccionan teniendo en cuenta los supuestos sobre la distribución de probabilidad F, la cual contiene el parámetro o parámetros desconocidos de interés. Tales métodos suelen clasificarse en paramétricos y no paramétricos. En los métodos paramétricos se hacen supuestos estrictos sobre la forma de F para generar una región de rechazo para la hipótesis nula. Los procedimientos para el testeo de hipótesis mediante pruebas no paramétricas utilizan menos supuestos o ninguno sobre F; la literatura reporta algunas ventajas de usar unos u otros métodos.

Este capítulo se divide en tres secciones. En la primera se presenta la categoría que sobre reportes de la literatura acerca de la inferencia estadística y el uso de los métodos no paramétricos en la prueba de hipótesis; en la segunda se indican los principales hallazgos en torno a la estadística no paramétrica en los libros de texto; y en la tercera se describen aspectos curriculares relacionados con la inclusión de la estadística no paramétrica en los planes curriculares universitarios. Estas tres categorías son parte de los resultados obtenidos al analizar el contenido de diferentes fuentes documentales como libros y artículos, entre otros. Los procedimientos basados en la ID permitieron generar búsquedas sistemáticas tendientes a sistematizar y estructurar los datos textuales.

2.1 Reportes de la literatura en torno a la estadística no paramétrica

La estadística no paramétrica ha sido abordada desde diferentes enfoques en los libros de texto destinados a atender las necesidades de aprendizaje en los estudiantes del nivel universitario o los requerimientos de las personas que emplean métodos no paramétricos en sus proyectos de investigación. Según Delicado (2008), existen dos familias de modelos estadísticos no paramétricos; aquellos que surgieron entre los años 40 y 50 del siglo XX, que suelen denominarse como métodos no paramétricos clásicos, y los que surgieron a partir de los años 70 del mismo siglo, los cuales son técnicas más avanzadas para la estimación no paramétrica de curvas; por ejemplo, para determinar la función de densidad de probabilidad, la función de confiabilidad y tasa de fallas, la función de regresión y para abordar la regresión no paramétrica y manejo de *splines*.

Entre los métodos no paramétricos clásicos están: i) la prueba del rango signado de Wilcoxon para una muestra, que es una prueba alternativa a la prueba t-student para una muestra con observaciones por parejas cuando las diferencias antes-después no cumplen el supuesto de normalidad, ii) la prueba del signo, la prueba de McNemar o de la significancia de los cambios, la prueba de Mann-Whitney, prueba que se constituye en una alternativa a la prueba t-student para dos pruebas independientes cuando no se cumplen los supuestos allí exigidos, iii) la prueba de Kruskal-Wallis corresponde a una prueba alternativa al análisis de varianza en una vía cuando no se cumplen los supuestos y se tienen tres o más muestras (Lehmann y D'Abrera).

De acuerdo con Corzo (2005), si en un problema de testeo de hipótesis se cumplen todos los supuestos requeridos por los métodos paramétricos empleados en ese tipo de pruebas, ya no es conveniente usar los modelos estadísticos no paramétricos; sin embargo, los métodos no paramétricos presentan algunas ventajas como las siguientes: i) son aplicables cuando la información recolectada (datos) se encuentran en una escala nominal o en una de tipo ordinal, ya que no existen pruebas paramétricas para ello; según Siegel (1970), pruebas de esta naturaleza son aplicables para resolver situaciones problema en el contexto de las ciencias de la conducta y de las ciencias humanas, ii) cuando se presenten datos extremos en una o más muestras, los cuales pueden generar un coeficiente de variación grande; en ese caso puede recurrirse al uso de la prueba de la mediana para el testeo de las hipótesis (Burbano y Valdivieso, 2016), iii) para muestras de tamaño grande; las estadísticas de prueba no paramétricas se comportan de forma normal asintóticamente.

Desde una mirada didáctica, algunas investigaciones se han dedicado a analizar la cantidad de contenidos y las estrategias didácticas utilizadas en la exposición de elementos de estadística no paramétrica. Por ejemplo, Palop y García (2017), centrados en el marco conceptual referido a la teoría de la comunicación, valoraron libros de texto universitarios de estadística teniendo en cuenta variables asociadas al nivel tecnológico, a la funcionalidad del libro desde una mirada didáctica, a la lecturabilidad referida al contenido, a la estructuración del mensaje, la simbolización y el tipo de lenguaje utilizado, y encontraron que los contenidos sobre estadística no paramétrica eran relativamente bajos en comparación con los tópicos de estadística descriptiva, y que se prefería la estadística paramétrica, incluidos los modelos de regresión.

Gamarra *et al*. (2006) manifiestan que en diversos trabajos de investigación con frecuencia se utilizan los métodos estadísticos paramétricos para realizar la prueba de hipótesis sobre varianzas poblacionales, cuando los datos de una variable en la muestra provienen de una población con distribución normal o con una gran aproximación normal; con todo, cuando tales supuestos no se cumplen, hay necesidad de recurrir al uso de modelos estadísticos no paramétricos basados en estadísticas de rangos. Estos autores elaboraron un trabajo investigativo centrado en la estadística no paramétrica, cuyo objetivo consistió en aplicar un paquete estadístico denominado NCSS a determinadas situaciones de prueba de hipótesis factibles de testearse mediante métodos no paramétricos y comparar sus resultados con los obtenidos al aplicar métodos paramétricos cuando fuera posible. En este contexto, este tipo de estrategias didácticas favorecen la aplicación de procesos inferenciales en distintos campos del conocimiento como las ingenierías, la biología y la administración de empresas, entre otros.

También Gamarra *et al*. (2006) señalan que en Colombia la literatura reporta muy pocos trabajos referidos al manejo y la aplicación de los métodos no paramétricos, mucho menos cuando se trata de implementarlos con la ayuda de paquetes computacionales. Estos mismos autores destacan el trabajo de Corzo (2005) consignado en el libro titulado *Estadística no paramétrica*, publicado por la Universidad Nacional de Colombia. Se trata de un libro de texto para el nivel universitario dirigido principalmente a estudiantes de estadística, matemáticas y ciencias naturales sin descartar su uso en otras ramas del saber. Allí, entre otros, se han examinado de manera detallada y comprensible un alto número de tópicos recogidos del libro de Hettmansperger (1984) sobre los principales modelos estadísticos no paramétricos clásicos. En la presente obra se proporcionan estrategias para facilitar la comprensión estudiantil sobre las estadísticas no paramétricas basadas en rangos sin dejar de lado el tópico de rachas.

Un tratamiento más puntual y profundo sobre aspectos teóricos y procedimentales del tratamiento de datos mediante técnicas propias de la estadística no paramétrica se puede encontrar en libros como los siguientes (citados por Delicado, 2008): Lehmann y D'Abrera (1975), Pratt y Gibbons (1981), Gibbons (1997), Leach (1982), Hettmansperger (1984), Gibbons y Chakraborti (1992), Hollander y Wolfe (1999), Hájek, Sidak y Sen (1999), Lehmann y Casella (2001). Además, resulta de gran ayuda el libro de texto universitario de Siegel (1970), en el cual se expone ampliamente el uso de los métodos no paramétricos en el campo de las ciencias de la conducta, aplicables también en las ciencias sociales y humanas.

Considerando que en los procesos de inferencia estadística se utilizan las distribuciones de probabilidad, es recomendable revisar con antelación algunos textos que tratan el concepto de probabilidad desde una mirada fundamental y didáctica (Burbano et al., 2021), tales como el libro de David (1998), en el que se ha elaborado un recorrido de tipo fenomenológico sobre probabilidad atendiendo aspectos históricos desde tiempos inmemoriales hasta el siglo XX; el libro de Hacking (1975), en el cual se estudia detalladamente la evolución que ha tenido el concepto de lo probable; y el de Kendall (1978), donde se describen los progresos de la probabilidad y su interacción con los procesos evolutivos de la estadística hasta que esta última se consolidó con la formulación de los métodos paramétricos en el siglo pasado.

Sobre el uso de la estadística no paramétrica, Rivera (2003), en México, desarrolla en el ámbito de la ingeniería industrial una prueba de contrastes ortogonales referida a diseños experimentales mediante el uso de la estadística no paramétrica con métodos intensivos de cómputo; la autora menciona que en diseño experimental es fundamental identificar factores que poseen un efecto significativo en un determinado experimento, en el cual también interesan las interacciones que se produzcan; lo anterior con el propósito de mejorar un proceso o todo un sistema productivo. Asimismo, indica que este tipo de situación problema frecuentemente se trata mediante diseños experimentales de tipo paramétrico; sin embargo, una alternativa flexible la constituyen los métodos no paramétricos que posibilitan una buena estimación de los parámetros incluidos cuando se utilizan métodos robustos e intensivos de cómputo.

Por otra parte, Vera *et al*. (2021) mencionan que en las últimas décadas la inclusión de los contenidos referidos a la estadística impartida en los claustros universitarios se ha incrementado de manera significativa. En la práctica, la mayoría de las carreras profesionales en la universidad incluyen por lo menos una asignatura de estadística en su plan de estudios. En Chile, como en otros países de Latinoamérica, se han aumentado los contenidos de estadística en los cursos universitarios, aunque también se ha empezado hacerlo desde la educación primaria, secundaria y fuera de las aulas, para incrementar el uso cotidiano de la estadística en los procesos productivos y en la cultura de las personas al hacer diversos razonamientos e inferencias de tipo estadístico para tomar decisiones en presencia de incertidumbre e información procesada de forma pertinente (Vera et al., 2021).

A pesar de eso, en los cursos de la estadística universitaria es frecuente cerrar con tópicos alusivos a la inferencia estadística focalizados en la prueba de hipótesis que incorpora distribuciones de probabilidad para elaborar inferencias en torno a uno o varios parámetros desconocidos de la población de interés desde una mirada paramétrica y con poca utilización de los métodos no paramétricos. Además, el proceso E–A resulta complejo en cuanto subyacen diversas relaciones entre las dimensiones epistemológica, didáctica, cognitiva, histórica, social y cultural entre el contenido por enseñar, el profesor universitario y el estudiantado (Soto y Cantoral, 2014). Un elemento esencial en el proceso de inferencia estadística es el de parámetro; no obstante, en diversos libros de texto universitarios poco se enfatiza en hacer una diferenciación fuerte entre parámetros y estimadores, lo cual provoca confusiones en el estudiantado al momento de efectuar las pruebas de hipótesis (Vera et al., 2021).

Por otra parte, Solís (2014), en su trabajo referido a las escalas de medición de los datos de las variables vinculadas a procesos de investigación psicológica, menciona que la mayoría de las variables intervinientes en los procesos de tipo psicológico no se miden en una escala de intervalo, sino mediante puntajes relacionados con una escala ordinal; pese a lo cual, se aplican métodos paramétricos para su análisis de forma legítima, aunque debería recurrirse en varios casos a modelos estadísticos no paramétricos. Así, tanto en la psicología como en otras disciplinas algunos autores persisten en el uso de estadística paramétrica para analizar datos de variables en escala ordinal, aunque, según Siegel (1970), sería más apropiado tratar este tipo de situaciones a través de estadísticas basadas en rangos y métodos no paramétricos.

En cuanto a las escalas de medición de las variables y su relación con las pruebas no paramétricas específicas, Ríos y Peña (2020) indican que es pertinente elaborar primero una identificación de los diversos tipos de pruebas no paramétricas, las cuales generalmente no están sujetas a seguir los supuestos de una distribución de “probabilidad normalizada” (son libres de distribución), son robustas y su uso está ligado al diseño de la investigación, el número de variables y la escala de medición de los datos de esas variables consideradas en un determinado estudio. Las pruebas sugeridas para una muestra, cuando los datos están medidos en una escala nominal (variables categóricas), son la prueba binomial y la chi-cuadrado (Pearson); además, la prueba K-S de Kolmogorov-Smirnov y la prueba de rachas se utilizan cuando los datos están al menos en una escala ordinal o de intervalo.

Para el caso de dos muestras independientes se sugiere utilizar la prueba de Mann-Whitney cuando los datos están por lo menos en una escala ordinal y preferentemente cuando no se haya cumplido el supuesto de normalidad y de homocedasticidad. La prueba de Wald-Wolfowitz (rachas) con variables dicotómicas se usa para probar que dos muestras "autónomas" han sido seleccionadas de una misma población (Ríos y Peña, 2020). Para datos pareados (muestra antes-después) se emplea la prueba del rango de Wilcoxon cuando la variable de interés es continua. Si se pretende determinar la igualdad de porcentajes (datos en escala nominal) se utiliza la estadística de McNemar, y se recurre a la prueba del signo para testear si dos medianas son iguales en la población o igualdad de medias cuando la distribución F resulte simétrica. Los datos deben estar, al menos, en una escala ordinal (Siegel, 1970).

Para k muestras, se suele emplear la prueba de Kruskal-Wallis a fin de probar la igualdad de medias con similares requerimientos que la prueba U de Mann-Whitney. Se trata de una prueba alternativa al análisis paramétrico de varianza en una vía cuando no se cumplen los supuestos al considerar tres o más muestras de una variable continua. Para k muestras relacionadas, se sugiere usar la prueba de Friedman, en este escenario, los k tratamientos se relacionan en algún sentido, además cada muestra ha de presentar n elementos que constituyen n bloques; en este contexto, a las puntuaciones de cada bloque se les asigna un rango desde 1 hasta k; se prueba que las medias son iguales, es decir, que las medias de los tratamientos provienen de una misma población (Lehmann y D'Abrera, 1975). En algunos otros casos se pueden utilizar pruebas de concordancia o de correlación basadas en estadísticas de rangos.

2.2 La estadística no paramétrica en los libros de texto universitarios

Desde una mirada conceptual, de acuerdo con Alzate *et al*. (1999), en general los libros de texto universitarios son obras que han de incluir aspectos didácticos tendientes a promover el proceso E–A de contenidos temáticos específicos en el estudiantado; puede tratarse de materiales impresos o digitales destinados tanto al profesorado como a los estudiantes que cursan un conjunto determinado de materias en una carrera profesional. En particular, los libros de estadística universitarios son elementos de tipo didáctico que corresponden a materiales impresos que incluyen conceptos (texto), símbolos, expresiones matemáticas referidas al tratamiento de las variables, estimadores, parámetros, procedimientos y representaciones gráficas, entre otros, con uno o más contenidos acordes con los programas de asignatura prescritos como parte de un determinado plan curricular. Tales contenidos son secuenciados en unidades didácticas que generalmente conforman los capítulos del libro de texto universitario.

Adicionalmente, Occelli y Valeiras (2013) señalan que los libros de texto para cualquier nivel educativo, incluido el universitario, son herramientas mediadoras entre un saber específico, el estudiantado y el docente. Tales libros traducen y concretan los significados y contenidos sugeridos en el microcurrículo prescrito por la autoridad educativa correspondiente, que en el aula de clase se han de reflejar en acciones específicas para promover el aprendizaje de tópicos particulares a través de actividades planificadas por el profesor. Aunque —hay que mencionarlo— en distintas ocasiones se han constituido en receptáculos de saberes parcelados con estrategias didácticas implícitas, destinados a regular la actividad de docencia que se desarrolla en un lapso específico de tiempo.

En cuanto a los libros de estadística y de matemáticas en general, Van Dormolen (1986) menciona que existen distintos tipos de libros de texto, entre los cuales están aquellos que exponen contenidos (conceptos y definiciones), prosiguen con la presentación de ejemplos, el desarrollo de ejercicios típicos y plantean problemas para ser resueltos con la teoría expuesta. En otros casos, los libros de texto mezclan en cada capítulo los aspectos teóricos con ejemplos, situaciones problema para promover el aprendizaje y ejercicios con respuesta incluida. aunque, en algunos libros de texto subyace una teoría de corte pedagógico que facilita la asimilación de los conceptos y su aplicación en contextos reales o de investigación ya sea formativa o de corte científico; en otros, se propicia la construcción del conocimiento, el aprendizaje significativo o la transposición didáctica, que consiste en suscitar una transformación del saber sabio (científico) expuesto a través de contenidos, en saber para ser enseñado a los estudiantes incluyendo estrategias de enseñanza (Chevallard, 1991), pero en correspondencia con el desarrollo de competencias básicas y profesionales cuando se trata del entorno universitario.

En un trabajo elaborado por Díaz (2014) se analizan diversos libros de texto españoles que incluyen tópicos de estadística. En tales libros se da prelación al uso de gráficos para analizar información, se continúa con la estadística descriptiva y de manera somera se recurre a los procesos de inferencia. Por su parte, en el libro de texto universitario de la autora

Valdivieso (2011) hay un tratamiento de la “estadística descriptiva” centrado en el trabajo independiente del estudiante. Primero, se ofrecen conceptos estadísticos; luego, se exponen situaciones de aprendizaje con ejemplos, y finalmente se involucra al estudiantado en actividades con retroalimentación continua. A lo largo de cada capítulo se hace especial énfasis en el ciclo de Wild y Pfannkuch (1999), además de actividades orientadas a la recolección, organización, representación e interpretación de la información recabada de contextos reales, destacando la distinción entre estimadores y parámetros, aspectos necesarios en los cursos de inferencia estadística.

Por su parte, Burbano y Valdivieso (2016) han elaborado un libro de texto universitario orientado a procesos de inferencia estadística y desarrollado a través de métodos paramétricos. En el primer capítulo se abordan algunos tipos de muestreo probabilístico, los conceptos de muestra aleatoria, estimadores y parámetros; en el segundo se trabajan las distribuciones de algunos estimadores y se determina el tamaño de ciertas muestras; en el tercer capítulo se hace la estimación puntual de parámetros y la construcción de intervalos de confianza; en el cuarto se efectúan diversos testeos de hipótesis estadísticas que incluyen parámetros como la media, la proporción, la varianza, la diferencia de medias, la diferencia de proporciones y el cociente de varianzas, desde una mirada netamente paramétrica. Al finalizar cada capítulo se plantean actividades para fortalecer el estudio independiente del estudiantado. Y en el capítulo quinto se ofrece la retroalimentación para las actividades propuestas.

En el libro de texto de Lind *et al.* (2012) se examina ampliamente la estadística descriptiva con aplicaciones a situaciones problema relacionadas con la administración de empresas y la economía; también se trabajan tópicos alusivos a los modelos de regresión lineal y al control estadístico de la calidad. La parte concerniente a la inferencia estadística concentra más tópicos referidos al testeo de hipótesis mediante métodos paramétricos; no obstante, dedica un capítulo al estudio de algunos de los modelos estadísticos no paramétricos orientados hacia la prueba de hipótesis. Ahora bien, su estructura metodológica recoge la secuencia: exposición de miniproyectos investigativos, presentación de aspectos conceptuales, ejemplos y ejercicios de aplicación relativos a tales miniproyectos. Cabe anotar que se requiere la aplicación de diversos preconceptos relacionados con la administración, la estadística y las matemáticas fundamentales.

En el libro de texto universitario de Canavos (1998) se examinan varios temas relacionados con la estadística descriptiva, las variables aleatorias y su distribución de probabilidad, las distribuciones muestrales clásicas paramétricas, la estimación paramétrica puntual y por intervalo, el testeo de hipótesis para la media, la proporción, la varianza, la diferencia de medias, la diferencia de proporciones y el cociente de varianzas por medio de métodos paramétricos, diseños de experimentos, modelo de regresión y control de calidad, además incluye los métodos no paramétricos; entre ellos, la prueba del signo, la prueba chi-cuadrado, la de Wilcoxon, la de Mann-Whitney, la prueba de Kruskal-Wallis, la prueba de Friedman, la prueba de Kolmogorov-Smirnov y otras pruebas para la bondad de ajuste.

En el libro de texto de Lindgren (1993), escrito en inglés, sugerido para cursos universitarios y de posgrado, se exponen tópicos relacionados con probabilidad, variables aleatorias, distribuciones paramétricas, muestreo aleatorio, estimación puntual y por intervalo, prueba de hipótesis, análisis de datos categóricos, distribuciones multivariadas, modelos lineales y de regresión, y elementos de la teoría de la decisión; además, se dedica un capítulo al estudio de los métodos no paramétricos clásicos. Aquí, los contenidos son presentados con un alto nivel de complejidad desde la estadística teórica y los ejercicios presentan un alto grado de dificultad.

El libro de texto universitario de Lind *et al*. (2015) está orientado a trabajar la estadística en el ámbito de la economía y los negocios. En este texto se aplica el conocimiento teórico a situaciones problema relacionadas con los negocios y a aspectos económicos trabajados desde una metodología de proyectos; igualmente, se aborda la estadística descriptiva, la inferencia estadística paramétrica y otros tópicos. En uno de sus veinte capítulos, se examinan los principales métodos no paramétricos clásicos que hacen hincapié en la prueba de hipótesis y se aplican a situaciones del contexto real en una determinada empresa productora de bienes y servicios. Esta obra presenta una estructura didáctica agradable para el estudiantado, la cual incluye conceptos referidos a los negocios, la economía y la estadística desde un nivel básico.

Otro libro de texto para el nivel universitario es *Estadística matemática con aplicaciones*, de Freund y Miller (2000), el cual está dirigido principalmente a estudiantes de matemáticas e ingenierías por su alto nivel de abstracción; no obstante, puede utilizarse en las carreras de economía, administración y psicología en algunas temáticas. El texto inicia con probabilidad y las principales distribuciones de probabilidad paramétricas, distribuciones muestrales, prueba de hipótesis (con teoría y aplicaciones), regresión y correlación, análisis de varianza, y dedica uno de sus 16 capítulos a las pruebas no paramétricas con ejemplos de aplicación a diversas disciplinas; entre ellos, la prueba del signo, la prueba chi-cuadrado, la de Wilcoxon, la de Mann-Whitney, la prueba de Kruskal-Wallis, la de Friedman, la de Kolmogorov-Smirnov y el coeficiente de correlación de Spearman basado en la teoría de rangos.

El libro de Pérez (2005), titulado *Técnicas estadísticas con SPSS 12: aplicaciones al análisis de datos*, trata ampliamente los principales tópicos de estadística descriptiva e inferencial (univariada y multivariada) principalmente desde un enfoque paramétrico, mediante el uso directo del *software* SPSS; empero, en el capítulo nueve, de los veinte que conforman el libro, se exponen ampliamente los métodos no paramétricos y su utilización a través del mencionado *software*; para cada tópico estadístico se describe el procedimiento y se analizan los resultados obtenidos a través de dicho paquete estadístico.

En el libro de Chamorro y Revelo (2008), *Simulación, un primer contacto*, los autores utilizan diversas distribuciones de probabilidad (discretas y continuas) para generar números aleatorios y simular valores de variables aleatorias, las cuales a su vez se utilizan para modelar y analizar sistemas reales desde un punto de vista investigativo. En este contexto, se utilizan algunas pruebas de hipótesis (paramétricas y no paramétricas) para el ajuste de los datos reales a determinados modelos de simulación; entre ellas, la prueba de normalidad de Kolmogorov-Smirnov y la prueba de Friedman. Este libro resulta de primordial importancia para los estudiantes de ingeniería de sistemas, industrial, electrónica y otras, donde con frecuencia se usa la inferencia estadística computacional.

En el texto de Montgomery y Runger (2012), al principio se analizan elementos de estadística descriptiva, probabilidad y variables aleatorias, inferencia estadística paramétrica, regresión, diseños experimentales, control de calidad; además hay un capítulo centrado en la estadística no paramétrica. Con una similar estructura, en los claustros universitarios se utiliza frecuentemente el libro de texto de Anderson *et al*. (2012), pero este da una mayor prelación a la estimación por intervalo y a la prueba de hipótesis. Otro libro con estructura semejante utilizado en diferentes carreras universitarias es el de Newbold *et al*. (2013), que destaca los modelos no paramétricos y presenta algunos ejemplos de aplicación. En otros textos, como los de Walpole *et al*. (2012) y Devore (2012), también se observa que hacen poco énfasis en el uso de los métodos no paramétricos porque solamente dedican un capítulo a ejemplificar tales métodos.

En cuanto a libros de texto especializados para el nivel universitario, se puede mencionar el de Siegel (1970), el cual se dedica completamente a los métodos no paramétricos aplicados a las ciencias de la conducta, la psicología y las ciencias humanas. El texto inicia describiendo las fortalezas y debilidades de utilizar los modelos estadísticos no paramétricos, las escalas de medición, la forma de seleccionar los métodos en correspondencia con las escalas de medición de las variables, y después se centra en las principales pruebas de hipótesis que contienen datos provenientes de una variable aleatoria en una muestra, dos muestras independientes, muestras pareadas y k muestras; varios de los conjuntos de datos allí utilizados son tomados de diversos trabajos de investigación.

En el libro de texto en inglés de Lehmann y D'Abrera (1975) se sugiere que este sea utilizado en cursos universitarios tanto de pregrado en estadística como de otras disciplinas y de posgrado. Allí se exponen únicamente tópicos de estadística no paramétrica para una, dos y tres o más muestras, diseños experimentales cuyas pruebas de hipótesis se sustentan en la teoría de rangos, testeo de aleatoriedad e independencia estadística. Los contenidos también promueven situaciones de investigación en diferentes ramas del conocimiento, pero se requieren sólidos conocimientos sobre estadística teórica para muestras pequeñas y grandes, y sobre la convergencia asintótica de las estadísticas de rangos hacia la distribución normal, la chi-cuadrado u otras.

En el libro de Randles y Wolfe (1979) se hace una introducción a la teoría estadística no paramétrica. Esta obra inicia con las estadísticas basadas en rangos, prosigue con las pruebas libres de distribución, las estadísticas del tipo U clásicas y de comportamiento asintótico. Allí se describe la función de potencia y se estudian sus propiedades, se abordan de manera fundamental los problemas de estimación y testeo de hipótesis mediante estadísticas basadas en rangos, se asumen los problemas de localización en una, dos y más muestras, y se analizan situaciones de correlación, independencia y de diversos diseños experimentales.

En el libro de Hettmansperger (1984) se trata exhaustivamente el problema de localización para una muestra, de la prueba del signo, la estimación puntual, la estabilidad, las pruebas más potentes, el problema de las muestras pareadas y la prueba del rango signado de Wilcoxon, la eficiencia relativa, las distribuciones para las estadísticas de puntajes, pruebas para estadísticas para puntajes winzorizados, el problema de localización en dos y más muestras, la prueba de Mann-Whitney, la prueba de Friedman, el coeficiente de correlación de Spearman y de Kendall, diseños experimentales no paramétricos, el modelo de regresión lineal no paramétricos, entre otros. Se manejan altos niveles de abstracción.

En el libro de Manoukian (1986), *Mathematical Nonparametric Statistics*, se hace una aproximación matemática a los diferentes modelos estadísticos no paramétricos. Se inicia con los modos de convergencia de variables aleatorias, el teorema de Slutsky y de Helly-Bray, Levy-Cramer, teorema central del límite y el de Liapunov; después se examina la teoría de las estadísticas de orden y su distribución de probabilidad, se hace un desarrollo axiomático de las pruebas no paramétricas para una, dos, tres y k muestras, sin descartar el uso de estadísticas de puntajes y diseños experimentales, el uso de la estadística de Klotz y el problema tanto de localización como de escala.

En el libro de Hájek *et al*. (1999), *Theory of Ranks Tests*, también se efectúa un tratamiento matemático de los principales modelos estadísticos no paramétricos, organizados a lo largo de diez capítulos;

Igualmente, hay un desarrollo axiomático de las pruebas no paramétricas para una, dos, tres y *k* muestras basadas en rangos, el uso de estadísticas de puntajes, el problema tanto de localización como de escala, y el testeo de hipótesis en modelos de regresión con base en estadísticas de rangos.

El libro de Lehmann y Casella (2001), *Theory of Point Estimation*, presenta un tratamiento teórico de la estimación puntual en los modelos estadísticos no paramétricos. Este texto está organizado en seis capítulos, comienza con elementos de la teoría de la medida, teoría de la probabilidad, la familia exponencial, convergencia en probabilidad y estadísticas suficientes; prosigue con familias no paramétricas, estimadores UMVU, el problema de localización y escala, modelos lineales, riesgo promedio y optimización, estimación simultánea, admisibilidad, optimización asintótica, entre otros. En este libro se manejan altos niveles de axiomatización.

En el libro de Wasserman (2006), titulado *All of Nonparametric Statistics*, hay una aproximación teórica a los modelos estadísticos no paramétricos que trascienden a los modelos clásicos. Esta obra se divide en diez capítulos; empieza con una contextualización de la inferencia estadística no paramétrica; después se refiere a la estimación de la función de distribución y de funcionales estadísticos, la distribución empírica, intervalos de confianza mediante *bootstrap*, técnicas de suavizado y *kernels*, regresión no paramétrica, verosimilitud local, estimación de densidades mediante *kernel*, teoría minimax, inferencia no paramétrica usando funciones ortogonales, *wavelets* y métodos adaptativos e inferencia semiparamétrica, entre otros temas.

Finalmente, en el libro de Marshal y Olkin (2007), *Life Distributions*, se expone un considerable número de modelos estadísticos con estructura no paramétrica, semiparamétrica y los que incluyen a familias paramétricas; además, se presentan familias de distribuciones mixtas en las cuales son de mucha importancia los parámetros de localización y de escala. Por supuesto, existe muchísima más literatura sobre los modelos no paramétricos que va más allá de los alcances del presente libro de investigación, pero que dejan abierta la posibilidad para emprender futuros trabajos investigativos focalizados en otras técnicas no paramétricas para la estimación de densidades de probabilidad o el uso de la teoría de Wavelets.

2.3 La estadística no paramétrica en el currículo universitario

La estadística no paramétrica es un tópico que al parecer resulta poco frecuente en el currículo universitario; su baja presencia puede atribuirse a que en los libros de texto universitarios se incluye someramente un capítulo del tema de métodos no paramétricos o no se lo incluye. En estas circunstancias, el profesor universitario tampoco lo encuentra atractivo, se abstiene de leerlo y frecuentemente no lo utiliza en sus clases. Al respecto, Gattuso y Pannone (2002) indican que se requiere una mejor preparación del profesorado que no ha logrado la suficiente formación en educación estocástica tanto desde la parte disciplinar como didáctica; de acuerdo con Mendonça *et al*. (2006), un buen número de profesores solamente han abordado un curso de estadística durante su formación profesional, con poco énfasis en las aplicaciones de la estadística en el mundo real y mucho menos en métodos no paramétricos.

En este contexto, resulta conveniente aumentar los tópicos de estadística universitaria en el currículo de la formación del profesorado y de las carreras profesionales donde esta se imparte; asimismo, resultaría de gran utilidad incluir el uso de los modelos estadísticos no paramétricos. Algunas pruebas internacionales y nacionales han informado que los resultados obtenidos por el estudiantado en estadística son generalmente bajos (Bos et al., 2014). Para mitigar esta situación, en alguna medida los profesores han de acudir a los libros de texto actualizados, para incrementar los contenidos de estadística en el currículo y estudiarla desde una mirada paramétrica y no paramétrica.

En Colombia, bien sea en la formación del profesorado de matemáticas o en otras profesiones, los tópicos de estadística expuestos se organizan en uno o dos cursos que generalmente se dedican a trabajar elementos de estadística descriptiva, cálculo de probabilidades e inferencia estadística paramétrica (Gómez, 2011). Además, en consonancia con lo establecido en la Ley 30 de 1992 acerca de la autonomía, cada universidad fija las asignaturas y los contenidos requeridos en el proceso de formación del estudiantado a nivel universitario, lo que dificulta incorporar más contenidos de estadística y mucho menos si se trata de métodos no paramétricos.

En general, los planes curriculares incluyen pocas asignaturas de carácter optativo donde se podrían contener elementos de estadística no paramétrica orientada hacia la investigación científica desde el aula. En la formación básica de las carreras universitarias (primeros tres semestres) se ha observado la incorporación de asignaturas como cálculos, físicas, y una materia o dos de estadística en las que se tratan temáticas muy básicas apoyadas con libros de texto destinados a estudiantes de matemáticas, ciencias e ingenierías, donde se utiliza esporádicamente el *software* estadístico (Pérez, 2005).

En cuanto a la probabilidad y la estadística, la Universidad Pedagógica Nacional ha aumentado sus contenidos y ofrece un curso completo sobre este tópico, la Universidad de Cundinamarca se ha propuesto dictar la materia denominada pensamiento aleatorio, la Universidad del Tolima ha incursionado con la historia de las matemáticas y de la estadística (Gómez, 2011), en la UPTC se ha detectado que se ofrecen dos cursos de estadística en las principales carreras universitarias que oferta cada semestre. A pesar de esto, cabe destacar que algunas universidades como la Nacional, la del Valle, la de Córdoba, Santo Tomás, Sergio Arboleda, El Bosque, entre otras, ofrecen la carrera profesional de estadística, donde su plan curricular sí contiene diversos tipos de estadística, incluida la “no paramétrica”.

3

Situaciones problema con una muestra para promover la investigación en el aula

En este capítulo se presentan variadas situaciones problema relacionadas con el uso de tópicos de estadística no paramétrica, para promover la investigación científica en el salón de clases universitario. En la primera sección se hace referencia al manejo de datos pertenecientes a una variable cualitativa, los cuales están clasificados en dos o más categorías o modalidades y considerados ya sea en una población o en una muestra. En la segunda se explicita la prueba de hipótesis basada en la distribución binomial y en una sola muestra aleatoria. En la tercera se indica la prueba chi-cuadrado con una muestra aleatoria donde la variable cualitativa presenta más de dos categorías. En la cuarta se aborda la prueba K-S de Kolmogorov-Smirnov, para analizar si los datos de una variable continua provienen de una población con una distribución específica. En este marco, se aplica un conjunto de pasos destinados a la ejecución del proceso de inferencia estadística, referido a la prueba de hipótesis.

3.1 Situaciones alusivas al trabajo con datos provenientes de una variable de tipo cualitativo en una población o en una muestra

Para este tipo de situaciones hay que recordar que una variable es cualitativa cuando corresponde a una característica que puede estudiarse en los individuos de una población o de una muestra, pero que sus datos son factibles de clasificarse en categorías o grupos mutuamente excluyentes; es decir, que un individuo no ha de estar en dos o más grupos a la vez. Si los individuos pertenecen a la población objeto de estudio, será suficiente con realizar estadística descriptiva sobre los datos que conforman las categorías de la variable; en cambio, si los datos de la variable de interés corresponden a una muestra, entonces es posible hacer inferencia estadística y comprobar las hipótesis a que haya lugar.

Por otra parte, los datos de la variable X pueden conformar solamente dos categorías o también pueden clasificarse en tres o más grupos. En seguida, se ilustra una situación que puede implicar el uso de inferencia estadística o solamente elementos de estadística descriptiva.

Una empresa registrada en la Cámara de Comercio con el nombre de PAW Productores de artículos W, en el mes de abril de 2022 presentaba un total 50 trabajadores, quienes laboraban en 4 áreas específicas en la ciudad de Cali, codificadas de la siguiente manera:

G: Gestión, P: Producción, C: Comercialización, T: Talento humano
F: Finanzas

G, T, P, P, P, P, P, C, F, C, C, T, G, T, T, P, P, P, P, C, F, F, T, P, G,
G, T, P, P, P, P, P, C, F, C, C, T, G, T, T, P, P, P, P, C, F, F, T, P, G.

Los datos se recolectaron mediante un formulario llamado censo. En este caso, se trata de 50 individuos que pertenecen a una población (Burbano et al., 2021). Por lo tanto, sobre estos datos solamente se harán procesos relativos a la estadística descriptiva. En primera instancia, se identificará la variable y luego se organizarán sus datos en 5 grupos mutuamente excluyentes.

La variable de interés se denota y se define de la siguiente forma:

X: área de trabajo donde se desempeña cada uno de los trabajadores de la empresa PAW Productores de artículos W, localizada en la ciudad de Cali, en el mes de abril de 2022.

Recuérdese que los datos corresponden a cada uno de los valores que admite la variable X; la frecuencia absoluta es la cantidad de veces que se repite cierto dato, y la frecuencia relativa corresponde a un porcentaje alusivo a cada frecuencia relativa.

Después de organizar los datos, los 5 grupos o categorías mutuamente excluyentes son los siguientes:

G, G, G, G, G, G, P, P, P, P, P, P, P, P, P, P, P, P, P, P, P, P, P, P, P, P,
C, C, C, C, C, C, C, C, T, T, T, T, T, T, T, T, T, T, F, F, F, F, F, F

La primera categoría está constituida por los trabajadores que laboran en el área de gestión (G), la segunda categoría la conforman quienes trabajan en el área de producción (P), la tercera corresponde a quienes trabajan en la de comercialización (C), la cuarta es la de talento humano (T) y la quinta categoría es la de finanzas (F).

La frecuencia absoluta para la categoría G es de 6, para P es de 20, para la categoría C es de 8, para T es de 10, y para la categoría F es de 6, para un total de N = 50 individuos que constituyen la población objeto de estudio. Ahora, se puede proseguir a organizar los datos en una tabla de frecuencias absolutas, como se presentan en la Tabla 3.1. Además, los valores que se calculen con estos datos se suelen llamar parámetros, porque se están refiriendo a la población de trabajadores de toda la empresa PAW.

Tabla 3.1. Organización de los datos de *X*

X: Área de trabajo	Frecuencia absoluta
G	6
P	20
C	8
T	10
F	6
Total	N=50

Fuente: los autores

En la Tabla 3.1 se observa que el área de la empresa con más trabajadores es la de producción (P) con 20 trabajadores; es decir, P es la que presenta la mayor frecuencia absoluta. También, con base en la misma tabla, se determina que las categorías G y F son las que muestran la frecuencia absoluta menor o más baja. En esta población también es posible calcular la medida descriptiva correspondiente al parámetro llamado la moda de X, la cual se suele denotar con $M_o(X)$. Entonces, la moda corresponde al dato que presenta la mayor frecuencia absoluta; en esta situación la moda es el dato P (trabajadores en el área de producción). Por lo tanto, se puede denotar y escribir así:

$$\mathrm{M_o}(X) = \mathrm{P}$$

Una actividad subsiguiente propia de la estadística descriptiva corresponde al proceso de interpretación de la información. En este caso: ¿cómo se interpreta la moda? Recuérdese que la moda es un indicativo de la tendencia de los datos; por lo tanto, aquí indica que en las áreas de la empresa PAW hay una tendencia a que sus trabajadores se ubiquen en el área de producción (P).

Otra actividad inherente a la estadística descriptiva es la representación de la información; para elaborar una representación de los datos pertenecientes a una variable cualitativa es recomendable utilizar un diagrama de pastel o torta, basado en las frecuencias absolutas o en las frecuencias relativas.

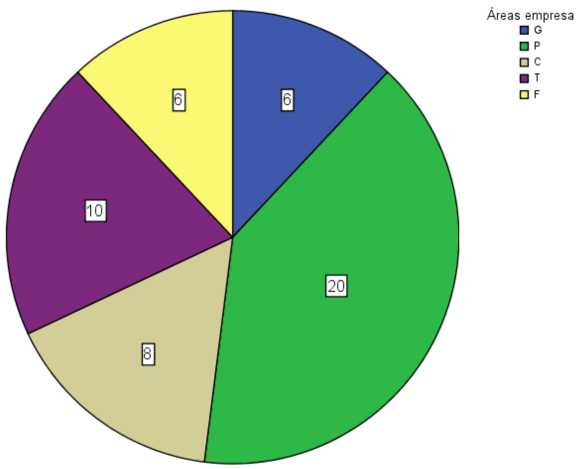

Figura 3.1. Diagrama de torta para la variable X
Fuente: los autores

A continuación, se presenta una tabla de frecuencias relativas asociadas a las 5 categorías en las cuales se han organizado los datos de la variable X (Tabla 3.2). De acuerdo con la Tabla 3.2, se puede interpretar que 20 de los 50 trabajadores de la empresa PAW laboran en el área de producción (P); es decir, 20/50 = 0.4 = 0.40 = 40 %; este porcentaje indica que el 44 % de los trabajadores de la mencionada empresa trabajan en el área de producción (P).

Tabla 3.2. Tabla de frecuencias relativas y absolutas

X: Área de trabajo	Frecuencia absoluta	Frecuencia relativa
G	6	0.12 = 12 %
P	20	0.40 = 40 %
C	8	0.16 = 16 %
T	10	0.20 = 20 %
F	6	0.12 = 12 %
Total	N=50	1=100 %

Fuente: los autores

Esta afirmación también equivale a que 40 de cada 100 trabajadores trabajan en el área P, o que 20 de cada 50 trabajan en dicha área. De forma semejante, 10 de los 50 trabajadores trabajan en el área de talento humano (T), lo cual representa al 20 % de los trabajadores quienes trabajan en el área T de la empresa PAW. Así se puede continuar con la interpretación de las frecuencias para las demás categorías.

Algunas veces se acostumbra a representar las frecuencias relativas correspondientes a las categorías de una variable cualitativa mediante un diagrama de barras separadas; en el eje horizontal se ubican las categorías y en el eje vertical se escriben los porcentajes respectivos, como se presentan en la Figura 3.2.

Finalmente, se podría obtener una conclusión como sigue: los trabajadores de la empresa PAW están organizados en 5 áreas de trabajo, de las cuales el área de producción es la que contiene más trabajadores y el área de gestión junto al área de finanzas son las que poseen menos trabajadores. En dicha empresa hay una tendencia a que los trabajadores se ubiquen en el área de producción más que en las otras áreas.

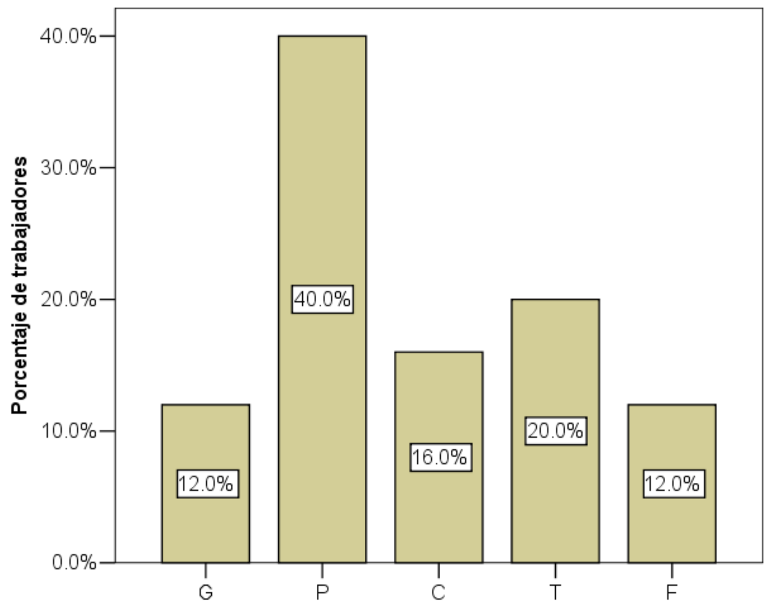

Figura 3.2. Porcentaje de trabajadores por área de trabajo
Fuente: los autores

A continuación, se presenta otra situación hipotética en la cual se puede requerir el uso de estadística descriptiva, inferencia estadística o ambas.

En un estudio sobre control de calidad para el producto A en la empresa T, se toma una muestra aleatoria de 20 unidades, codificadas de la siguiente forma:

B: Buena calidad, D: Defectuoso

B, B, B, D, B, B, B, B, B, B, B, B, B, B, D, B, B, B, B, B.

Los datos fueron colectados a través de un formulario llamado encuesta. En esta situación hipotética se trata de 20 unidades del artículo A (20 individuos), las cuales corresponden a una muestra (Burbano et al., 2021). Por consiguiente, en referencia a estos datos es posible efectuar procesos sobre la estadística descriptiva y la estadística inferencial. En primer lugar, se identifican las variables y luego se organizan en dos categorías mutuamente excluyentes; estos dos grupos también reciben el nombre de modalidades.

Ahora, la variable de interés, por tener dos categorías, también suele llamarse dicotómica, la cual se mide en una escala nominal y se denota y se define así:

X: estado de calidad de las unidades del producto A seleccionadas aleatoriamente de un proceso productivo en la empresa T.

Al organizar los datos en dos categorías, se obtiene el siguiente arreglo:

B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, D, D.

La primera categoría está conformada por las unidades de buena calidad (B), la segunda categoría la constituyen las unidades defectuosas (D). La frecuencia absoluta para la categoría B es 18 y para la categoría D es 2. Se prosigue a organizar los datos en una tabla de frecuencias absolutas, como se indica en la Tabla 3.3. Adicionalmente, los valores que se calculan con estos datos reciben el nombre de estimadores, porque están vinculados a la muestra aleatoria.

Tabla 3.3. Organización de los datos de X

X: Estado de calidad	Frecuencia absoluta
B	18
D	2
Total	n =20

Fuente: los autores

Con base en la Tabla 3.3 se puede determinar que la categoría B es la que presenta la mayor frecuencia absoluta con 18 unidades de buena calidad en la muestra; en contraste, D es la que presenta la menor cantidad de datos, en la muestra de tamaño n = 20. En este caso, la moda para la variable X es B, la cual se denota y escribe de la siguiente manera:

$$m_o(X) = B$$

Para esta situación hipotética, la moda indica que hay una tendencia de que las unidades del artículo A en la empresa T que fueron escogidas aleatoriamente resulten de buena calidad. Una actividad adicional relacionada con la estadística descriptiva consiste en representar la información por medio de un diagrama de torta o pastel, como se puede observar en la Figura 3.3.

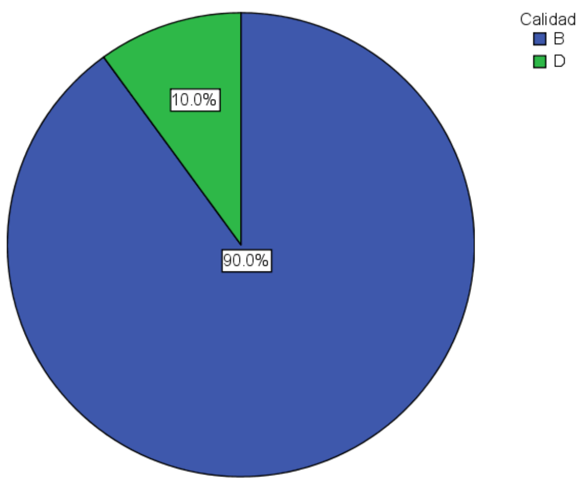

Figura 3.3. Estado de calidad de las unidades del producto A
Fuente: los autores

En seguida, en la Tabla 3.4 se pueden observar las frecuencias relativas a la variable dicotómica X.

En correspondencia con la Tabla 3.4, se interpreta que 18 de los 20 artículos de la muestra aleatoria resultaron de buena calidad, lo cual representa el 90 % de la muestra; en contraste, 2 de los 20 artículos de la mencionada muestra resultaron defectuosos y corresponden al 10 % de esta muestra.

Tabla 3.4. Frecuencias absolutas y relativas para la variable X

X: Estado de calidad	Frecuencia Absoluta	Frecuencia relativa
B	18	0.9 = 90 %
D	2	0.1 = 10 %
Total	n=20	1=100 %

Fuente: los autores

En la Figura 3.4 se pueden observar las respectivas frecuencias relativas.
Finalmente, se puede obtener la siguiente conclusión acerca de los datos de la muestra. Los datos están organizados en dos categorías, de estas, la categoría B es la que más unidades presenta del producto A. En este contexto hay una tendencia a que las unidades de la muestra resulten de buena calidad. Probablemente, al tomar otra muestra aleatoria se obtengan resultados semejantes o quizá diferentes; en consecuencia, a partir de una muestra aleatoria, también será factible realizar procesos de inferencia estadística.

3.2 Prueba de hipótesis basada en la distribución binomial

En diversas circunstancias de investigación es posible que se tenga que trabajar con una población conformada solo por dos categorías; por ejemplo, productos con dos resultados posibles: bueno o defectuoso; también en poblaciones cuyos resultados posibles sean: soltero, casado, masculino, femenino, entre otros. En una población conformada por dos categorías, saber que la proporción o porcentaje de individuos en la primera clase es *p*, implica que la proporción en la otra categoría es *1 – p* = *q*.

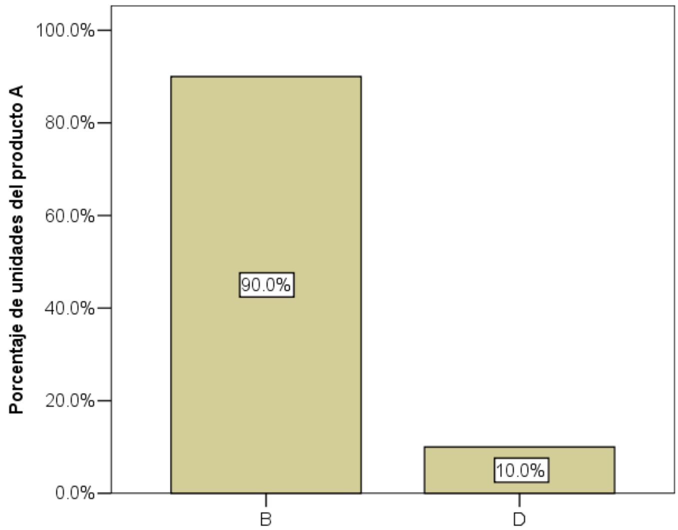

Figura 3.4. Estado de calidad del producto A en la muestra
Fuente: los autores

A pesar de que la proporción *p* puede cambiar de una población a otra, este valor *p* puede considerarse fijo en una población específica; no obstante, aun si se conoce este valor en una determinada población, no se puede esperar que una muestra aleatoria proveniente de tal población contenga exactamente la proporción *p* en la primera categoría y la proporción *q* en la segunda categoría. En este contexto, a los efectos aleatorios en el proceso de muestreo suele atribuírseles el hecho de que la muestra no reproduzca con exactitud los valores de *p* y *q* en esa población.

En estas circunstancias, la distribución muestral correspondiente a la proporción (porcentaje) *p* observada en una muestra aleatoria tomada en una población de dos categorías es una distribución binomial. Entonces, H_o es la hipótesis nula, la cual informa que p corresponde a un valor específico en la población; esto indica cuán razonables es que la proporción muestral (estimador) provenga de una población de parámetro *p*, o si por el contrario hay diferencias significativas.

En esta situación, para una muestra de tamaño *n*, la probabilidad de que la variable aleatoria *X* sea igual a la cantidad *x* de individuos en la primera categoría y *n*-*x* en la segunda, se puede calcular a través de la distribución binomial (Burbano y Valdivieso, 2015):

$$f(x) = P(X = x) = \begin{cases} \binom{n}{x} p^x q^{n-x} & si\ x = 0,1,2,...,n \\ 0 & \text{en otro caso} \end{cases}$$

Al expresar el combinatorio en términos de números factoriales, la distribución binomial también se puede escribir de la siguiente manera:

$$f(x) = P(X = x) = \begin{cases} \dfrac{n!}{x!(n-x)!} p^x q^{n-x} & si\ x = 0,1,2,...,n \\ 0 & \text{en otro caso} \end{cases}$$

Donde *n*! = 1*2*3*…**n*; por ejemplo: 4! = 1*2*3*4 = 24; 0! = 1; 1! = 1 y

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

Por ejemplo (Burbano et al., 2021):

$$\binom{4}{2} = \frac{4!}{2!(4-2)!} = \frac{24}{2(2)} = \frac{24}{4} = 6$$

Supóngase que con base en la muestra aleatoria:

B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, D, D

se quiere probar la hipótesis de que no hay diferencias significativas entre la proporción de unidades del producto A que resultan de buena calidad y la proporción de las que resultan defectuosas con un nivel de significancia α del 5 % (α = 0.05).

En concordancia con lo expuesto en Burbano y Valdivieso (2016), un conjunto de pasos para probar esta hipótesis es el siguiente:

1) Planteamiento del sistema de hipótesis

Como se trata de probar que no hay diferencias significativas, la hipótesis nula puede incluir la afirmación de que p = 0.5 y la hipótesis alternativa de que p es mayor que 0.5; por lo tanto, el sistema de hipótesis por comprobar se podría escribir así:

H_o: p=0.5
H_1: $p > 0.5$

2) Se fija el nivel de significancia en el valor α = 0.05.

3) La dirección de la prueba irá hacia la derecha; es decir, corresponde a una prueba unilateral derecha.

4) La estadística de prueba se soporta en la distribución binomial.

5) Decisión: si el p-*valor* es menor que α = 0.05, entonces se rechaza la hipótesis nula; en caso contrario, no habrá evidencia suficiente para rechazarla y será aceptada.

Aquí, el *p*-valor será igual a sumar las probabilidades para X=18, 19 y 20 calculadas sobre el modelo binomial, es decir:

P-valor = $P(X \geq 18) = P(X = 18) + P(X = 19) + P(X = 20)$

$$P(X = 18) = \binom{20}{18}(0.5)^{18}(0.5)^{20-18} = 190(0.5)^{20} = 0.00018$$

$$P(X = 19) = \binom{20}{19}(0.5)^{19}(0.5)^{20-19} = 20(0.5)^{20} = 0.000019$$

$$P(X = 20) = \binom{20}{20}(0.5)^{20}(0.5)^{20-20} = 1(0.5)^{20} = 0.000009$$

En consecuencia,

p-valor = $P(X \geq 18) =$ 0.00018 + 0.000019 + 0.000009 = 0.000208

Como el *p*-valor resultó inferior al nivel de significancia α = 0.05, entonces se rechaza la hipótesis nula y se acepta la hipótesis alternativa.

6) Con un nivel de significancia del 5 %, se concluye que sí hay diferencias significativas en la proporción poblacional *p* referida a la cantidad de unidades del producto A con respecto a la proporción de unidades defectuosas.

Otra forma de resolver esta misma situación problema consiste en utilizar una aproximación de la distribución binomial por una distribución normal estándar cuando el tamaño de la muestra n sea lo suficientemente grande, así:

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

Donde $\hat{p}$ corresponde a la proporción muestral y p_0 al valor de la proporción poblacional que aparece en la hipótesis nula, $q_0 = 1 - p_0$; es decir:

$$\hat{p} = \frac{x}{n}$$

Cuando se trabaja con la cantidad x de la variable aleatoria X, una expresión equivalente para la estandarización mencionada es la siguiente:

$$Z = \frac{x - np_0}{\sqrt{np_0 q_0}}$$

Para muestras pequeñas se usa el factor de corrección por continuidad. Esta corrección consiste en disminuir en $0.5/n$ el valor observado en la proporción muestral $\hat{p}$ y el valor esperado en la proporción p_0 poblacional. Cuando $\hat{p} < p_0$ se hace un incremento de $0.5/n$; en cambio, cuando $\hat{p} > p_0$ se hace una disminución de $0.5/n$; en consecuencia,

si $\hat{p} < p_0$ entonces se utiliza $Z = \dfrac{(\hat{p} + 0.5/n) - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$

si $\hat{p} > p_0$ entonces se emplea $Z = \dfrac{(\hat{p} - 0.5/n) - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$

Para cuando se trabaje con la cantidad x de la variable aleatoria X, solamente se agrega o disminuye 0.5 a la x; por lo tanto, se utilizarán las siguientes expresiones:

Si $x < np_0$ entonces se utiliza $Z = \frac{(x+0.5)-np_0}{\sqrt{np_0q_0}}$

Si $x > np_0$ entonces se emplea $Z = \frac{(x-0.5)-np_0}{\sqrt{np_0q_0}}$

Para solucionar la situación problema objeto de estudio a partir de la muestra aleatoria:

B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, D, D.

se procedería de la siguiente forma:

1) Planteamiento del sistema de hipótesis

H_o: p=0.5
H_1: $p > 0.5$

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$.

3) La prueba es unilateral derecha.

4) La estadística de prueba se soporta ahora en una distribución normal estándar.

5) Decisión: si el p-valor es menor que $\alpha = 0.05$ entonces se rechaza la hipótesis nula; en caso contrario, se la acepta.

Como

$\hat{p} = \frac{x}{n} = \frac{18}{20} = 0.9$ resultó mayor que p_0 = 0.5 entonces se emplea

$$Z = \frac{(\hat{p} - 0.5/n) - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{(0.9 - 0.5/20) - 0.5}{\sqrt{\frac{(0.5)(0.5)}{20}}} = \frac{(0.9 - 0.025 - 0.5}{\sqrt{\frac{0.25}{20}}}$$

$$Z = \frac{0.375}{0.1118} = 3.35$$

En esta situación, el *p*-valor = $P(Z \geq 3.35) = 1 - 0.9996 = 0.0004$

Como el *p*-valor =0.0004 resultó inferior al nivel de significancia α = 0.05, entonces se rechaza la hipótesis nula y se acepta la hipótesis alternativa.

6) Con un nivel de significancia del 5 %, se concluye que *p* es mayor que 0.5; por lo tanto, sí existen diferencias significativas en la proporción poblacional *p* referida a la cantidad de unidades del producto A que resultan de buena calidad con respecto a la proporción de unidades defectuosas. Se ha comprobado que el porcentaje de unidades de buena calidad en la población es superior al 50 %.

En esta misma situación, un inspector de la empresa T, basado en su experiencia de supervisar el proceso de producción, afirma que la proporción (porcentaje) en la población de unidades del producto A es superior al 91 %; usar la muestra obtenida para aceptar o rechazar la afirmación del inspector con un nivel de significancia del 5 %.

Aquí se dispone de la muestra aleatoria de tamaño *n* = 20, dada por:

B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, B, D, D.

Se procedería del siguiente modo:

1) Planteamiento del sistema de hipótesis

H_o: p=0.91
H_1: $p > 0.91$

2) se fija el nivel de significancia en el valor $\alpha = 0.05$.

3) La prueba es unilateral derecha.

4) La estadística de prueba se soporta ahora en una distribución normal estándar.

5) Decisión: si el p-valor es menor que $\alpha = 0.05$ entonces se rechaza la hipótesis nula; en caso contrario, se la acepta.

Como

$\hat{p} = \frac{x}{n} = \frac{18}{20} = 0.90$ resultó menor que $p_0 = 0.91$ entonces se emplea

$$Z = \frac{(\hat{p} + 0.5/n) - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{(0.9 + 0.5/20) - 0.91}{\sqrt{\frac{(0.91)(0.09)}{20}}} = \frac{(0.9 + 0.025 - 0.91}{\sqrt{\frac{0.0819}{20}}}$$

$$Z = \frac{0.015}{0.06399} = 0.2344$$

En esta situación, el p-valor = $P(Z \geq 0.23) = 1 - 0.591 = 0.409$

Como el P-valor =0.409 resultó mayor que el nivel de significancia α = 0.05, entonces se acepta la hipótesis nula H_o: p=0.91.

6) Con un nivel de significancia del 5 %, se concluye que *p* es igual a 0.91; por lo tanto, el inspector no tiene la razón. En este contexto, la proporción poblacional p relacionada con la cantidad de unidades del producto A que resultan de buena calidad es del 91 %.

La siguiente situación problema ha sido adaptada de Burbano y Valdivieso (2016). Un importador de frutas chilenas afirma que el 96 % de las unidades de un lote grande que le acaba de llegar a su bodega están en buenas condiciones y el resto de la fruta se perderá porque ha arribado en malas condiciones. El vendedor chileno difiere de esa afirmación y le solicita que se tome una muestra aleatoria de 200 unidades de las mencionadas frutas; se encontró que 194 unidades de la muestra sí estaban en buenas condiciones; con la información de esa muestra se desea probar la afirmación del importador con un nivel de significancia del 5 %.

En esta situación, el tamaño de la muestra n = 200 puede considerarse grande. Los datos de la muestra aleatoria también presentan dos categorías, la primera conformada por las frutas que llegan en buenas condiciones a la bodega y la segunda constituida por las frutas arruinadas; en consecuencia, se puede utilizar la aproximación de una distribución binomial por una distribución normal estándar. El proceso de prueba de hipótesis para esta situación puede contener los siguientes pasos:

1) Planteamiento del sistema de hipótesis

H_o: p=0.96
H_1: p diferente de 0.96

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$.

3) Ahora la prueba es bilateral.

4) La estadística de prueba se soporta ahora en una distribución normal estándar.

5) Decisión: si el *p*-valor es menor que $\alpha/2 = 0.025$ entonces se rechaza la hipótesis nula; en caso contrario, dicha hipótesis será aceptada.

Aquí resulta que $1- \alpha = 0.95$; $Z_{\alpha/2} = Z_{0.025} = -1.96$; $Z_{1-\alpha/2} = Z_{0.975} = 1.96$ como se ilustra en la Figura 3.5

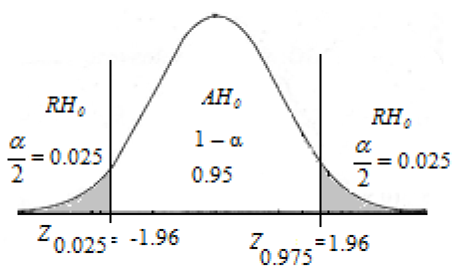

Figura 3.5. Prueba bilateral para la proporción
Fuente: los autores.

Como
$\hat{p} = \frac{x}{n} = \frac{194}{200} = 0.97$, ahora se emplea la siguiente estadística de prueba

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{0.97 - 0.96}{\sqrt{\frac{(0.96)(0.04)}{200}}} = \frac{0.01}{0.013856} = 0.7217$$

En esta situación, el *p*-valor = $P(Z \geq 0.7217) = 1 - 0.7246 = 0.2758$

Como el p-valor =0.2758 resultó mayor que el nivel de significancia $\alpha/2$ = 0.025, entonces se acepta la hipótesis nula H_o: p=0.96.

En este contexto, también se observa que el valor de Z=0.7217 cae en la región AH_0 comprendida entre los valores -1.96 y 1.96 (ver Figura 3.5), por lo tanto, se acepta H_o: p=0.96.
6) Con un nivel de significancia del 4 %, se concluye que p es igual a 0.96; por lo tanto, el importador tiene la razón. En este contexto, la proporción poblacional p asociada a la cantidad de unidades de fruta que llega en buenas condiciones a la bodega es del 96 %.

3.3 Prueba chi-cuadrado para una muestra

Con frecuencia, cuando un investigador emprende sus tareas, su interés puede focalizarse en el número de individuos cuyos datos pueden clasificarse en una categoría de varias que recogen el total de la información. Por ejemplo, la opinión de un grupo de individuos puede clasificarse en una de cinco categorías: muy de acuerdo, de acuerdo, indiferente, en desacuerdo o muy en desacuerdo. En estas circunstancias, la prueba chi-cuadrado es utilizada para probar si existen diferencias significativas entre el número observado de individuos (datos) que se clasifican en cada categoría y el número esperado de datos en cada categoría, con la hipótesis nula.

La hipótesis establece la proporción de individuos que se clasifican en cada una de las categorías de una población de la que se presume provengan los datos; en otros términos, mediante la prueba de hipótesis nula se puede determinar cuáles son las frecuencias esperadas. La estadística de prueba corresponde a una chi-cuadrado, la cual permite precisar si las frecuencias observadas están próximas a las frecuencias esperadas, en cuyo caso ha de aceptarse la hipótesis nula; de lo contrario, si tales diferencias son muy grandes, se ha de rechazar esta hipótesis.

La estadística de prueba puede expresarse de la siguiente manera:

$$Chi-cudrado=\sum_{i=1}^{k}\frac{\left(O_i-e_i\right)^2}{e_i}$$

Donde O_i corresponde al número observado de individuos clasificados en la categoría *i*, e_i hace referencia al número esperado de individuos en la categoría *i* en concordancia con la hipótesis nula; en este contexto, e_i se obtiene del cociente entre el total de datos *n* (tamaño de la muestra) dividido entre el número de categorías (*k*); el símbolo de la sumatoria indica que se han de sumar todas las *k* categorías. La frecuencia relativa esperada puede estimarse así: $f_i = e_i/n = 1/k$. Si se produce un número alto de concordancias entre las frecuencias observadas y esperadas, es posible que dicha suma sea pequeña, en cuyo caso el valor de chi-cuadrado resultará pequeño y por lo tanto se acepta la hipótesis nula; en caso contrario, si chi-cuadrado resulta lo suficientemente grande, entonces se rechaza esta hipótesis.

Como la estadística (estimador) chi-cuadrado con la hipótesis nula sigue una distribución teórica chi-cuadrado con *k-1* grados de libertad, si *p*-valor es menor que el nivel de significancia α, entonces se rechaza la hipótesis nula; de lo contrario, será aceptada. Es necesario mencionar que cuando *k* sea igual a 2, cada frecuencia esperada debe ser mayor o igual que 5; si *k* es mayor que 5, la chi-cuadrado no se utiliza cuando más del 20 % de las frecuencias esperadas sean menores que 5 (Cochran, 1954).

Algunas veces, las frecuencias esperadas aumentan al combinar categorías adyacentes cuando tal combinación tenga sentido (Burbano et al., 2019).

La siguiente situación problema ha sido planteada de forma semejante a la expuesta por Canavos (1988): el gerente de la empresa WA quiere saber si el número de trabajadores de su empresa que asisten a la oficina de talento humano para recibir consejería laboral se encuentra distribuido de manera equitativa durante los 5 días hábiles de cada semana. De una muestra aleatoria tomada durante 4 semanas completas de labor activa de sus trabajadores, se obtuvieron los siguientes datos sobre X: número de trabajadores que asistieron a la oficina de talento humano para recibir consejería laboral.

Lunes: 50
Martes: 36
Miércoles: 33
Jueves: 40
Viernes: 46

Con un nivel de significancia del 5 %, probar la hipótesis de que el número de trabajadores que asisten a la oficina de talento humano de la empresa WA se distribuyen de manera equitativa durante los 5 días de la semana.

El proceso de prueba de hipótesis para esta situación puede seguir los siguientes pasos:

1) Planteamiento del sistema de hipótesis

H_o: f_i = 0.2 para i=1, 2, 3, 4, 5
H_1: f_i difiere de 0.2 para algunos i=1, 2, 3, 4, 5

De forma equivalente, las hipótesis anteriores se pueden plantear de la siguiente forma:

H_o: e_i = 41 para i=1, 2, 3, 4, 5
H_1: e_i difiere de 41 para algunos i=1, 2, 3, 4, 5

2) Se fija el nivel de significancia en el valor α = 0.05.

3) Se trata de una prueba unilateral derecha sobre una distribución teórica chi-cuadrado con *k-1* = 5-1 = 4 grados de libertad

$$\chi^2_{0.95,4} = 9.4877$$

4) La estadística de prueba se soporta en los datos de la muestra, con *n* = 205, *k = 5*; e_i = 205/5 = 41.

$$Chi = \sum_{i=1}^{k} \frac{(O_i - e_i)^2}{e_i} = \frac{(50-41)^2}{41} + \frac{(36-41)^2}{41} + \frac{(33-41)^2}{41} + \frac{(40-41)^2}{41} + \frac{(46-41)^2}{41}$$

$$Chi = \sum_{i=1}^{k} \frac{(O_i - e_i)^2}{e_i} = \frac{(9)^2}{41} + \frac{(-5)^2}{41} + \frac{(-8)^2}{41} + \frac{(-1)^2}{41} + \frac{(5)^2}{41}$$

$$Chi - cuadrado = \sum_{i=1}^{k} \frac{(O_i - e_i)^2}{e_i} = \frac{81}{41} + \frac{25}{41} + \frac{64}{41} + \frac{1}{41} + \frac{25}{41} = \frac{196}{41} = 4.78$$

5) Decisión: si el *p*-valor es menor que α = 0.05 entonces se rechaza la hipótesis nula; en caso contrario, dicha hipótesis será aceptada.

El *p*-valor = P(Chi >4.78) =0.3106

Como el *p*-valor resultó mayor que 0.05 entonces se acepta la hipótesis nula H_o: $e_i = 41$.

Aquí también resulta que 1- α = 0.95; $\chi^2_{0.95,4} = 9.4877$

Como chi-cuadrado = 4.78 resulta menor que 9.4877, entonces se decide aceptar la hipótesis nula. En esta situación problema, por cualesquiera de las dos formas se acepta la hipótesis nula.

6) Con un nivel de significancia del 5 %, se concluye que no existe ninguna razón para considerar que el número de trabajadores que asisten a la oficina de talento humano de la empresa WA no se distribuyen de manera equitativa durante los 5 días de la semana; es decir, que los mencionados trabajadores asisten de manera uniforme con valor esperado de 41 trabajadores por semana.

La siguiente situación problema ha sido adaptada de Siegel (1970): los individuos aficionados a las carreras equinas en torno a una pista circular piensan que hay ventajas para los equinos que ocupan determinadas posiciones en la posta. En una pista para 8 equinos, la posición 1 es aquella que se encuentra más próxima a la baranda interior de esa pista; en cambio, la posición 8 es la que queda en el lado más alejado de esa primera baranda (la más exterior). Los siguientes datos corresponden al número de triunfos logrados por los equinos en cada una de las 8 pistas.

Tabla 3.5. Número de triunfos de los equinos por posición en la pista

Pista	1	2	3	4	5	6	7	8
Número de triunfos	30	20	19	26	18	11	16	12

Fuente: los autores

Con un nivel de significancia del 5 %, se quiere probar si existe algún efecto en correspondencia con la posición en la cual compitieron los equinos durante un periodo determinado de carreras.

El proceso de prueba de hipótesis para esta situación puede contener los siguientes pasos:

1) Planteamiento del sistema de hipótesis

H_o: f_i = 0.125 para i=1, 2, 3, …, 8
H_1: f_i difiere de 0.125 para algunos i=1, 2, 3, …, 8

De forma equivalente, las hipótesis anteriores se pueden plantear de la siguiente forma:

H_o: e_i = 19 para i=1, 2, 3, …, 8
H_1: e_i difiere de 19 para algunos i=1, 2, 3, …, 8

2) Se fija el nivel de significancia en el valor α = 0.01.

3) Se trata de una prueba unilateral derecha sobre una distribución teórica chi-cuadrado con k-1 = 8-1 = 7 grados de libertad

$$\chi^2_{0.95,7} = 14.0671$$

4) La estadística de prueba se soporta en los datos de la muestra, con n = 152, $k = 8$; e_i = 152/8 = 19.

$$Chi = \sum_{i=1}^{8} \frac{(O_i - e_i)^2}{e_i} = \frac{(30-19)^2}{19} + \frac{(20-19)^2}{19} + \frac{(19-19)^2}{19} + \frac{(26-19)^2}{19}$$
$$+ \frac{(18-19)^2}{19} + \frac{(11-19)^2}{19} + \frac{(16-19)^2}{19} \frac{(12-19)^2}{19}$$

$$Chi=\sum_{i=1}^{8}\frac{\left(O_i-e_i\right)^2}{e_i}=\frac{\left(11\right)^2}{19}+\frac{\left(1\right)^2}{19}+\frac{\left(0\right)^2}{19}+\frac{\left(7\right)^2}{19}+\frac{\left(-1\right)^2}{19}+\frac{\left(-8\right)^2}{19}+\frac{\left(3\right)^2}{19}\frac{\left(-7\right)^2}{19}$$

$$Chi-cuadrado=\frac{121}{19}+\frac{1}{19}+\frac{0}{19}+\frac{49}{19}+\frac{1}{19}+\frac{64}{19}+\frac{9}{19}+\frac{49}{19}=\frac{294}{19}=15.4737$$

5) Decisión: si el *p*-valor es menor que α = 0.05, entonces se rechaza la hipótesis nula; en caso contrario, dicha hipótesis será aceptada.

El *p*-valor = P(Chi >15.4737) =0.0304

Como el *p*-valor resultó menor que 0.05, entonces se rechaza la hipótesis nula H_o: $e_i = 19$.

Aquí también resulta que 1- α = 0.95; $\chi^2{}_{0.95,7}=14.0671$

Como chi-cuadrado = 15.4737 resulta mayor que 14.0671, entonces se decide rechazar la hipótesis nula. En esta situación problema, por cualquiera de las dos maneras se rechaza la hipótesis nula.

6) Con un nivel de significancia del 5 %, se concluye que sí existe un efecto significativo de la posición en la cual compiten los equinos en la pista circular analizada; es decir, que los mencionados equinos se ven afectados en el número de triunfos en correspondencia con la posición de donde inicien su competencia en la posta; así entonces, el valor esperado difiere de 19 triunfos para cada posición.

3.4 Prueba K-S de Kolmogorov-Smirnov para una muestra

Esta prueba se utiliza para analizar si los datos de una variable *X* continua se ajustan a una distribución de probabilidad específica. Por ejemplo, en diversas ocasiones se requiere analizar si un conjunto de datos cuantitativos con n pequeño y en escala de razón o al menos en una escala de intervalo, provienen de una distribución normal con media μ conocida y desviación estándar σ también conocida. Así mismo, con esta prueba se puede determinar si los datos de una muestra aleatoria provienen de una distribución $F_0(x)$ de interés como la distribución exponencial, lognormal, Weibull, gamma, entre otras.

Esta prueba K-S consiste en comparar los valores obtenidos mediante la función empírica $\hat{F}_n(x)$ (función de distribución acumulativa). Las observaciones (datos) se ordenan de menor a mayor de la siguiente manera: $x_{(1)}, x_{(2)}, x_{(3)}, ..., x_{(n)}$, la función empírica se denota y define como sigue:

$$\hat{F}_n(x) = \begin{cases} 0 & \text{si } x < x_{(1)} \\ \dfrac{k}{n} & \text{si } x_{(k)} \le x < x_{(k+1)} \\ 1 & \text{si } x \ge x_{(n)} \end{cases}$$

El proceso de prueba de hipótesis para esta situación puede seguir los siguientes pasos:

1) Planteamiento del sistema de hipótesis

H_o: $F(x) = F_0(x)$
H_1: $F(x)$ difiere de $F_0(x)$

2) Se fija el nivel de significancia en un valor α menor o igual que 0.05.

3) Se trata de una prueba bilateral.

4) La estadística de prueba se soporta en calcular el siguiente valor asociado a la prueba K-S, así (Burbano et al., 2019):

$$D = \underset{1 \le i \le n}{Sup} \left| \hat{F}_n(x_i) - F_0(x_i) \right|$$

En este contexto, el valor D representa la "diferencia absoluta más grande" que se pueda obtener entre la función acumulativa y la probabilidad calculada en el valor x_i sobre la distribución de probabilidad propuesta.

De acuerdo con Burbano *et al.* (2019), en distintos casos es posible utilizar la distribución asintótica para *D*, la cual se puede escribir para n lo suficientemente grande, en los siguientes términos:

$$\underset{n \to \infty}{Lim} P\left(D \le \frac{\lambda}{\sqrt{n}} = D_\alpha \right) = 1 - \exp\left(-2\lambda^2\right) = 1 - \alpha$$

En este caso, D_α se ha de seleccionar de tal modo que:

$P(\text{Re}\,chazar\ \text{Ho/Ho es cierta})=\text{P}(D > D_\alpha\ /\ \text{Los datos siguen la distribución propuesta})=\alpha$

Por ejemplo, para un nivel de significancia α = 0.05, el valor de λ se puede obtener de la siguiente ecuación:

1- *exp(-2λ²)* = 1- α/2

De la solución de esta ecuación resulta un valor λ = 1.36 aprox., en consecuencia, se tiene (Canavos, 1988):

$$D_\alpha = \frac{\lambda}{\sqrt{n}} = \frac{1.36}{\sqrt{n}}$$

5) Decisión: la región crítica o de rechazo de la hipótesis nula ha de satisfacer lo siguiente:

$$P\left(D > \frac{\lambda}{\sqrt{n}} = D_{\alpha}\right) = \alpha$$

En estas circunstancias, si D es mayor que D_{α} entonces se rechaza la hipótesis nula; en caso contrario, se aceptará esta hipótesis.

6) Obtención de una conclusión.

La siguiente situación problema se ha formulado de forma similar a la presentada en Lehmann y D'Abrera (1975): la variable X representa la cantidad de combustible en litros que en una muestra aleatoria de tamaño n =10 los vehículos de la marca A consumen por un lapso de tiempo L; los datos son los siguientes: 13.8, 13.9, 14, 14.2, 12.5, 12.8, 13, 13.4, 13.5, 13.6. Con un nivel de significancia del 5 % se quiere probar la hipótesis de que estos datos provienen de una distribución normal de parámetros $\mu = 13$ y $\sigma = 1$.

El proceso de prueba de hipótesis para esta situación puede incluir los siguientes pasos:

1) Planteamiento del sistema de hipótesis

H_o: $F(x) = F_0(x)$
H_1: $F(x)$ difiere de $F_0(x)$

Donde $F_0(x)$ corresponde a una distribución normal con $\mu = 13$ y $\sigma = 1$.

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$.

3) Se trata de una prueba bilateral.

4) La estadística de prueba se soporta en calcular el siguiente valor vinculado a la prueba K-S, así:

$$D = \underset{1 \le i \le n}{Sup} \left| \hat{F}_n(x_i) - F_0(x_i) \right|$$

En primera instancia se ordenan los datos de menor a mayor: 12.5, 12.8, 13, 13.4, 13.5, 13.6, 13.8, 13.9, 14, 14.2, los cuales corresponden a la serie ordenada (por el rango o posición de cada dato) $x_{(1)}, x_{(2)}, x_{(3)}, \ldots, x_{(10)}$.

Para cada dato se utiliza la estandarización

$$Z = \frac{x - \mu}{\sigma}$$

Para cada valor de Z obtenido, se determina su probabilidad en una distribución normal estándar; por ejemplo, para x = 12.5, su estandarización es

$$Z = \frac{12.5 - 13}{1} = -0.5$$

Su probabilidad a través de la distribución normal estándar es $F_0(x_{(1)})$ = $F_0(12.5)$ = 0.3085.

Para x= 12.8, su estandarización es

$$Z = \frac{12.8 - 13}{1} = -0.2$$

Su probabilidad a través de la distribución normal estándar es $F_0(x_{(2)})$ = 0.4207. Para los demás valores de x se procederá de similar manera. Por medio de la función acumulativa o distribución empírica se obtienen los valores para $\hat{F}_n(x)$, por ejemplo, $\hat{F}_n(x_{(1)}) = \hat{F}_n(12.5) = \frac{1}{10} = 0.1$,

$\hat{F}_n(x_{(2)}) = \hat{F}_n(12.8) = \frac{2}{10} = 0.2$; así sucesivamente.

Es conveniente mencionar que, en caso de presentarse empates, se divide el número de empates entre el valor de n. Con estos resultados se conforma la Tabla 3.6.

Como D representa la "diferencia absoluta más grande", entonces para este caso D =0.2254.

En esta situación, $D_\alpha = 0.410$

5) Decisión: como D resultó menor que D_α entonces no se rechaza la hipótesis nula, se acepta que los datos de la variable X se ajustan a una distribución normal.

Tabla 3.6. Cálculo de la estadística D, prueba de Kolmogorov-Smirnov

$X_{(i)}$	$\hat{F}_n(x_{(i)})$	$F_0(x_{(i)})$	D
12.5	0.1	0.3085	0.2085
12.8	0.2	0.4207	0.2207
13	0.3	0.5	0.2
13.4	0.4	0.6554	0.2554
13.5	0.5	0.6915	0.1915
13.6	0.6	0.7257	0.1257
13.8	0.7	0.7881	0.0881
13.9	0.8	0.8159	0.0159
14	0.9	0.8413	0.0587
14.2	1.0	0.8849	0.1151

6) Con un nivel de significancia del 5 % se concluye que los datos de la variable X provienen de una distribución normal con $\mu = 13$ y $\sigma = 1$.

La situación problema que se expone a continuación guarda alguna similitud con la indicada por Canavos (1988): la variable X representa el número de respuestas correctas obtenidas en una muestra de 16 estudiantes en una prueba interna para el ingreso a la universidad A: 397, 400, 405, 408, 413, 425, 438, 453, 242, 265, 300, 323, 347, 353, 371, 388. Con un nivel de significancia del 5 % se quiere probar la hipótesis de que estos datos provienen de una distribución normal de parámetros $\mu = 375$ y $\sigma = 50$.

El proceso de prueba de hipótesis para esta situación puede comprender los siguientes pasos:

1) Planteamiento del sistema de hipótesis

H_o: $F(x) = F_0(x)$
H_1: $F(x)$ difiere de $F_0(x)$

Donde F_0(x) corresponde a una distribución normal con $\mu = 375$ y $\sigma = 50$.

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$.

3) Se trata de una prueba bilateral.

4) La estadística de prueba se soporta en calcular el siguiente valor asociado a la prueba K-S, así:

$$D = \underset{1 \le i \le n}{Sup} \left| \hat{F}_n(x_i) - F_0(x_i) \right|$$

Ahora se ordenan los datos de menor a mayor: 242, 265, 300, 323, 347, 353, 371, 388, 397, 400, 405, 408, 413, 425, 438, 453, los cuales corresponden a la serie ordenada teniendo en cuenta el rango de cada dato, así: $x_{(1)}, x_{(2)}, x_{(3)}, ..., x_{(16)}$

Para cada dato se utiliza la estandarización

$$Z = \frac{x - \mu}{\sigma}$$

Para cada valor de Z obtenido se determina su probabilidad en una distribución normal estándar; por ejemplo, para x=242, su estandarización es

$$Z = \frac{242 - 375}{50} = -2.66$$

Su probabilidad a través de la distribución normal estándar es $F_0(x_{(1)}) = F_0(242) = 0.0039$.

Para x= 265, su estandarización es

$$Z = \frac{265 - 375}{50} = -2.2$$

Su probabilidad a través de la distribución normal estándar es $F_0(x_{(2)}) = F_0(265) = 0.0139$. Para los demás valores de x se procederá de similar manera. Por medio de la función acumulativa o distribución empírica se obtienen los valores para $\hat{F}_n(x)$, por ejemplo,

$\hat{F}_n(x_{(1)}) = \hat{F}_n(242) = \frac{1}{16} = 0.0625$,

$\hat{F}_n(x_{(2)}) = \hat{F}_n(265) = \frac{2}{16} = 0.125$; así sucesivamente.

Con estos resultados se conforma la Tabla 3.7.

Como D representa la "diferencia absoluta más grande", entonces para este caso D =0.1207.

En esta situación, D_α = 0.328

5) Decisión: como D resultó menor que D_α entonces no se rechaza la hipótesis nula, se acepta que los datos de la variable X se ajustan a una distribución normal.

6) Con un nivel de significancia del 5 %, se concluye que los datos de la variable X provienen de una distribución normal con con $\mu = 375$ y $\sigma = 50$.

Tabla 3.7. Cálculo de la estadística *D*, prueba de Kolmogorov-Smirnov

$X_{(i)}$	$\hat{F}_n(x_{(i)})$	$F_0(x_{(i)})$	D
242	0.0625	0.0039	0.0586
256	0.125	0.0139	0.1111
300	0.1875	0.0668	0.1207
323	0.25	0.1492	0.1008
347	0.3125	0.2877	0.0248
353	0.375	0.33	0.045
371	0.4375	0.4681	0.0306
388	0.5	0.6026	0.1026
397	0.5625	0.67	0.1075
400	0.625	0.6915	0.0665
405	0.6875	0.7257	0.0382
408	0.75	0.7454	0.0046
413	0.8125	0.7764	0.0361
425	0.875	0.8413	0.0337
438	0.9375	0.8962	0.0413
453	1	0.9406	0.0594

Es conveniente mencionar que este tipo de prueba no paramétrica también se puede realizar de forma directa mediante diversos paquetes estadísticos, entre ellos, el *software* libre R, el *software* con licencia SPSS, Minitab, economapas, SAS, o el lenguaje de programación Python, entre otros.

4

Situaciones problema con dos muestras para la prueba de hipótesis en investigación formativa

En este capítulo se exponen diversas situaciones problema relacionadas con la utilización de algunos tópicos de estadística no paramétrica centrados en la prueba de hipótesis con base en dos muestras aleatorias. Este tipo de pruebas se utilizan cuando el interés del investigador recae en determinar si existen diferencias significativas entre dos tratamientos o si un tratamiento puede ser considerado mejor que otros desde un punto de vista estadístico.

La primera sección se ocupa del manejo de la prueba de McNemar para la significación de los cambios. En la segunda se examina la prueba del signo. En la tercera se expone la prueba de la medina. En la cuarta se presenta la prueba del rango signado de Wilcoxon o prueba de rangos con signo de Wilcoxon. En la quinta se estudia la prueba U de Mann-Whitney. En el contexto de dos muestras también se aplica un conjunto de pasos para elaborar un proceso de inferencia estadística ligado a la prueba de hipótesis.

4.1 Situaciones referidas a la prueba de McNemar

En un diseño de dos muestras relacionadas, pareadas o por parejas (antes-después) es usual utilizar una prueba paramétrica denominada t-student, la cual se basa en el supuesto de que las diferencias de la variable cuantitativa de interés se han de distribuir normalmente, con independencia de la población de donde provienen los datos de esa muestra; además, los datos deben estar en una escala de intervalo o de razón. La prueba de McNemar es una prueba no paramétrica destinada a establecer la significancia de los cambios con base en dos muestras relacionadas (antes-después), cuyos datos se pueden colocar en una escala ordinal o en una nominal. Aquí cada individuo puede utilizarse como su propio control.

Para determinar la significancia de cualquier cambio observado con esta prueba se elabora una tabla de contingencia de tamaño 2x2, que presenta cuatro entradas a, b, c, d, las cuales corresponden a frecuencias absolutas que representan los dos conjuntos de respuestas referentes al antes y al después de los individuos que conforman la muestra; se utilizan los signos – y + para representar respuestas diferentes. En la Tabla 4.1 las celdas A y D recogen los casos para los cuales existen cambios entre la primera y la segunda respuesta; un individuo se clasifica en la celda A cuando ha cambiado de + hacia -; en cambio, se ubica en la celda D si ha pasado de – hacia +. Si no se ha observado ningún cambio antes = +, después = + entonces el individuo se clasifica en la celda B; si se observa el resultado antes = -, después = - se ubica en C.

Tabla 4.1. Tabla de contingencia para probar la significación de los cambios

	Después	Después
Antes	Menos (-)	Más (+)
Más (+)	A	B
Menos (-)	C	D

Fuente: los autores

En este contexto, la suma A+D representa el número de individuos que generaron cambios en las parejas (antes, después) de la variable de estudio X. Se espera que la mitad de (A+D) cambie en una dirección y la otra mitad en la otra dirección, en concordancia con la hipótesis nula. La situación problema entonces se reduce a utilizar la estadística de prueba chi-cuadrado con un solo grado de libertad para probar la hipótesis nula de igualdad de frecuencias esperadas. Así, para las frecuencias observadas A y D resulta que:

$$Chi-cudrado=\sum_{i=1}^{2}\frac{\left(O_i-e_i\right)^2}{e_i}=\frac{\left(A-0.5(A+D)\right)^2}{0.5(A+D)}+\frac{\left(D-0.5(A+D)\right)^2}{0.5(A+D)}$$

Después de efectuar las operaciones correspondientes, la anterior expresión de la chi-cuadrado con un grado de libertad, se reduce a:

$$Chi-cudrado=\sum_{i=1}^{2}\frac{\left(O_i-e_i\right)^2}{e_i}=\frac{\left(A-D\right)^2}{(A+D)}$$

Según Yates (1934), la corrección se efectúa por continuidad, ya que la distribución chi-cuadrado es continua y se está aproximando una distribución discreta por una continua, por lo tanto, se utiliza la estadística de prueba dada por:

$$Chi-cudrado=\frac{\left(\left|A-D\right|-1\right)^2}{(A+D)}$$

La siguiente situación problema ha sido adaptada de Siegel (1970): un psicólogo se ha interesado en el estudio de la iniciación de los estudiantes del nivel de preescolar en los contactos sociales. Este psicólogo ha observado que estos estudiantes que llegan a la institución educativa T suelen iniciar sus contactos sociales con las personas adultas más que con otros compañeros de su edad; además, afirma que a medida que aumente su familiaridad, experiencia e interacción con los individuos de la institución educativa T, los estudiantes de este nivel comenzarán a tener más contactos sociales con sus homólogos. Para probar esta hipótesis decide observar a 25 estudiantes de dicho nivel en cuanto a su iniciación en los contactos sociales en su primer día en la institución. Hace la clasificación correspondiente y luego los observa después de 20 días laborables y los clasifica nuevamente. Los resultados se presentan en la Tabla 4.2.

Tabla 4.2. Tabla de contingencia para probar la significación de los cambios en la iniciación de los estudiantes

	Después	Después
Antes	Homólogo	Adulto
Adulto	15	3
Homólogo	3	4

Fuente: los autores

El proceso de prueba de hipótesis para esta situación puede incluir los siguientes pasos:

1) Planteamiento del sistema de hipótesis

H_o: f_i = 0.5 para i=1, 2
H_1: f_1 es mayor que f_2

De forma equivalente, las hipótesis anteriores se pueden plantear de la siguiente forma:

H_o: $P_A = P_D = 0.5$
H_1: $P_A > P_D$

Donde P_A representa la probabilidad de que los estudiantes cambien su objeto de iniciación de adulto a homólogo, P_D representa la probabilidad de que los estudiantes cambien su objeto de iniciación de homólogo a adulto.

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$. $1 - \alpha = 0.95$

3) Se trata de una prueba unilateral derecha sobre una distribución teórica chi-cuadrado con *k-1* = 2-1 = 1 grado de libertad

$$\chi^2_{0.95,1} = 3.8415$$

4) La estadística de prueba se soporta en los datos de la muestra, con n = 25, *k =2*.

A = 15 es el número de infantes que cambiaron su objeto de iniciación pasando de adulto a infante (homólogo); D = 4 corresponde al número de infantes cuyo objeto de iniciación fue infante y luego pasaron a adulto; B = 3 y C = 3 representan a los infantes que se mantuvieron sin cambios en sus objetos de iniciación.

$$Chi-cudrado = \frac{\left(|A-D|-1\right)^2}{(A+D)} = \frac{\left(|15-4|-1\right)^2}{(15+4)} = \frac{(10)^2}{19} = \frac{100}{19} = 5.2631$$

5) Decisión: si el *p*-valor es menor que $\alpha = 0.05$ entonces se rechaza la hipótesis nula; en caso contrario, dicha hipótesis será aceptada.

El *p*-valor = P(Chi >5.2631) =1-0.9782=0.025 aprox.
Al usar un paquete estadístico se establece que este *p*-valor es de 0.0218

Como el *p*-valor resultó menor que 0.05 entonces se rechaza la hipótesis nula H_o: $P_A = P_D = 0.5$ y se acepta la hipótesis alternativa H_1: $P_A > P_D$.

Aquí también resulta que 1- α = 0.95; $\chi^2_{0.95,1} = 3.8415$

Como chi-cuadrado = 5.2631 resulta mayor que 3.8415, entonces se decide rechazar la hipótesis nula. En esta situación problema, por cualesquiera de las dos formas se rechaza la hipótesis nula. En consecuencia, se acepta la hipótesis alternativa planteada por el psicólogo.

6) Con un nivel de significancia del 5 %, se concluye que los infantes del nivel de preescolar adoptan una tendencia significativa a cambiar su objeto de iniciación al pasar de su interacción social con adultos hacia los infantes (homólogos) transcurridos 20 días laborables en la institución educativa T. Es decir, hay evidencias suficientes para aceptar la afirmación hecha por el psicólogo.

4.2 Prueba del signo

Con frecuencia esta prueba se utiliza como una alternativa a la prueba t-student de una muestra, donde la hipótesis nula se plantea en términos de que la media poblacional μ es igual a un valor μ_o y la hipótesis alternativa consiste en uno de tres casos: a) $\mu > \mu_o$, b) $\mu < \mu_o$, c) μ difiere de μ_o. Para aplicar la prueba del signo solamente se supone que la variable de interés *X* en la población muestreada es continua y simétrica. Si en vez de la media poblacional se trabaja con la mediana poblacional, la cual en este documento será denotada por μ_e entonces solamente se requiere que la variable X en tal población sea continua; en estas circunstancias, la prueba del signo que incluye a la mediana resulta adecuada para cuando datos de la muestra aleatoria resulten heterogéneos y el promedio no los represente adecuadamente.

Para operacionalizar la variable X, en esta prueba se sustituye cada valor de la muestra que supere al valor de μ_o con un signo + y cada dato inferior a μ_o se reemplaza por un signo menos (-); cuando un dato de la muestra resulta igual a μ_o solo se descarta. En este contexto, para la cantidad n constituida por signos positivos y negativos se asocia una prueba binomial donde la hipótesis nula se plantea en términos de que parámetro es $p = 0.5$ la cual significa que la proporción de signos positivos es igual a la de signos negativos; la hipótesis alternativa podría plantearse así: a) $p > 0.5$, b) $p < 0.5$, c) p difiere de 0.5. En el caso de que n resulte lo suficientemente grande, se utiliza la aproximación de una binomial por una normal con $p = q = 0.5$, así:

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

Donde x corresponde a la cantidad de signos positivos.

Es conveniente mencionar que la prueba del signo también se puede aplicar para muestras por parejas, cuando el investigador tiene interés en determinar que entre ambas condiciones (antes, después) se generan diferencias significativas. Aquí, cada individuo puede constituirse en su propio control. En este contexto, la hipótesis nula debe incluir la condición de que la probabilidad del evento $P(X_D > X_A) = P(X_D < X_A) = 0.5$. Aquí se maneja el siguiente criterio, si la diferencia entre X_D y X_A

resulta positiva, entonces se coloca un signo +, si la diferencia resulta negativa, se pone un signo menos y si la diferencia es cero, la pareja se descarta. En este caso, para la cantidad n constituida por signos positivos y negativos, si x representa la cantidad de signos positivos, entonces se puede usar una prueba binomial de parámetros n y p donde la hipótesis nula se plantea en términos de que parámetro es $p = 0.5$ lo cual significa que la proporción de signos positivos es igual a la de signos negativos; la hipótesis alternativa podría plantearse así: a) $p > 0.5$, b) $p < 0.5$, c) p difiere de 0.5. En el caso de que n resulte lo suficientemente grande, nuevamente se utiliza la aproximación de una binomial por una normal con $p = q = 0.5$.

La siguiente situación problema se ha adaptado de Freund y Miller (2000): un ingeniero industrial hizo 20 mediciones en cuanto a la resistencia de la rotura medida en libras, referidas a la clase de cinta de algodón TW de dos pulgadas, y obtuvo los siguientes datos:

164, 166, 161, 190, 162, 172, 159, 152, 170, 163, 164, 140, 173, 166, 149, 167, 173, 164, 188, 174

El ingeniero afirma que la media de esta clase de cintas es superior a 161. Usar un nivel de significancia del 5 % para probar la hipótesis planteada por el ingeniero.

Para solucionar la situación problema objeto de estudio a partir de la muestra aleatoria:

164, 166, 161, 190, 162, 172, 159, 152, 170, 163, 164, 140, 173, 166, 149, 167, 173, 164, 188, 174 se procedería de la siguiente forma:

1) Planteamiento del sistema de hipótesis

H_o: μ = 161
H_1: μ > 161

De forma equivalente se pueden plantear las siguientes hipótesis:

H_o: p=0.5
H_1: p > 0.5

2) Se fija el nivel de significancia en el valor α = 0.05.

3) La prueba es unilateral derecha.

4) La estadística de prueba se soporta ahora en una distribución binomial.

Para cada dato de la muestra que exceda a 161 se le asigna un signo más, para cada valor de la muestra que resulte inferior a 161 se le coloca un signo menos y se descarta el único valor de 161 en la muestra que es igual a la media (en la hipótesis nula); en esta situación, se obtiene la siguiente secuencia de signos:

+, +, descarte, +, +, +, -, -, +, +, +, -, +, +, -, +, +, +, +, +

Así, n= 19

Se define la variable X: número de signos positivos en la muestra de tamaño 19; x=0,1,2,…,15

P-valor =
$$P(X \geq 15) = P(X = 15) + P(X = 16) + P(X = 17) + P(X = 18) + P(X = 19)$$

$$P(X = 15) = \binom{19}{15}(0.5)^{15}(0.5)^{19-15} = 3876(0.5)^{19} = 0.00739$$

$$P(X=16)=\binom{19}{16}(0.5)^{16}(0.5)^{19-16}=969(0.5)^{19}=0.00184$$

$$P(X=17)=\binom{19}{17}(0.5)^{17}(0.5)^{19-17}=171(0.5)^{19}=0.00033$$

$$P(X=18)=\binom{19}{18}(0.5)^{18}(0.5)^{19-18}=19(0.5)^{19}=0.00004$$

$$P(X=19)=\binom{19}{19}(0.5)^{19}(0.5)^{19-19}=1(0.5)^{19}=0.000002$$

Por lo tanto,

P-valor = $P(X\geq 15)=0.0096$ para cuando p = 0.5

5) Como el p-valor resultó inferior al nivel de significancia α = 0.05, entonces se rechaza la hipótesis nula y se acepta la hipótesis alternativa.

6) Con un nivel de significancia del 5 %, se concluye que el promedio de la resistencia a la rotura de la clase de cinta de algodón TW de 2 pulgadas excede las 161 libras; en este caso, se acepta como cierta la afirmación hecha por el ingeniero industrial.

La situación problema que se aborda a continuación se formula de manera similar a la expuesta por Freund y Miller (2000): un ingeniero químico ha tomado 40 mediciones referidas a la cantidad de toneladas de óxido de azufre, las cuales han sido emitidas por una planta industrial perteneciente a la empresa NW, y ha obtenido los siguientes datos:

16, 14, 19, 28, 18, 17, 21, 24, 26, 8, 23, 19, 16, 5, 23, 13, 14, 22, 23, 25, 18, 22, 27, 18, 15, 21, 23, 16, 19, 12, 18, 9, 22, 17, 30, 12, 19, 16, 23, 13.

El ingeniero afirma que la emisión promedio de óxido de azufre en la empresa NW es inferior a 20.3 toneladas; utilizar un nivel de significancia del 2 % para probar la hipótesis planteada por el ingeniero.

Para solucionar esta situación problema mediante la prueba del signo, primero se obtiene el conjunto conformado por los signos más y los signos menos; al comparar cada dato con la media de 20.3, si el dato está por arriba se pondrá un signo más, si está por debajo se colocará un signo menos y se descartarán los empates (cuya diferencia resulte cero).

-, -, -, +, -, -, +, +, +, -, +, -, -, -, +, -, -, +, +, +, -, +, +, -, -, +, +, -, -, -, -, -, +, -, +, -, -, -, +, -.

Se procedería de la siguiente forma:

1) Planteamiento del sistema de hipótesis

H_o: $\mu = 20.3$
H_1: $\mu < 20.3$

De forma equivalente se pueden plantear las siguientes hipótesis:

H_o: p=0.5
H_1: $p < 0.5$

2) Se fija el nivel de significancia en el valor $\alpha = 0.02$.

3) La prueba es unilateral izquierda.

4) La estadística de prueba se soporta ahora en una distribución normal.

Al observar el conjunto conformado por los signos más y menos, resulta que,

n= 40

Se define la variable X: número de signos positivos en la muestra de tamaño 40; x=0,1, 2,…,40 con p = 0.5 = q se hace la aproximación de una distribución binomial por una normal, para x =16 signos positivos, así:

$$Z = \frac{x - 0.5n}{\sqrt{n(0.5)(0.5)}} = \frac{16 - 0.5(40)}{\sqrt{40(0.5)(0.5)}} = \frac{16 - 20}{\sqrt{10}} = -1.26$$

P-valor =

$$P(X \leq 16) = P\left(\frac{X - 0.5n}{\sqrt{n(0.5)(0.5)}} \leq \frac{16 - 0.5n}{\sqrt{n(0.5)(0.5)}}\right) = P(Z \leq -1.26) = 0.1038$$

Por lo tanto,

P-valor = $P(X \leq 16) = 0.1038$ para cuando p = 0.5

5) Como el p-valor resultó superior al nivel de significancia α = 0.02, entonces se acepta la hipótesis nula H_o: μ = 20.3.

De otra forma, como el valor de la estadística de prueba Z = -1.26 es mayor que el valor de la distribución teórica normal estándar $Z_\alpha = Z_{0.02}$ = -2.33, entonces se acepta la hipótesis nula.

6) Con un nivel de significancia del 2 %, se concluye que el promedio de emisión de óxido de azufre en la empresa NWW es de 20.3 toneladas; en este contexto, no hay evidencia suficiente para aceptar la afirmación hecha por el ingeniero químico, en esta situación el ingeniero no tiene la razón.

La siguiente situación problema se ha formulado de manera semejante a la expuesta en Lind *et al*. (2015): el director de sistemas de la empresa

MW recomendó la implementación de un curso de capacitación para los subgerentes; el objetivo era incrementar las competencias en el manejo de bases de datos; algunos subgerentes opinaron que valdría la pena tomar el curso, otros lo consideraron sin ningún valor; en esta situación se determinó que las sesiones del curso empezarían el primer día del siguiente mes. Se seleccionó una muestra aleatoria de 15 subgerentes; un grupo de expertos estableció el nivel de competencia de cada subgerente en referencia al manejo de tales bases, codificado así: Excelente =5, Sobresaliente = 4, Bueno = 3, Aceptable = 2 y Deficiente =1. Una vez los subgerentes de la muestra terminaron el curso, el mismo grupo de expertos calificó nuevamente a los subgerentes. Se obtuvieron las siguientes parejas de resultados (antes, después):

(4,5); (2,5); (5,4); (1,4); (5,5); (3,4); (1,2); (4,5); (3,1); (1,3); (3,4); (2,5); (3,2); (3,4); (1,3).
El director afirma que las competencias sobre el manejo de bases de datos se ha de incrementar una vez termine el curso; usar un nivel de significancia del 5 % para probar la hipótesis planteada por el director.

Para solucionar la situación problema objeto de estudio mediante la prueba del signo, primero se obtiene el conjunto conformado por los signos más y los signos menos; al comparar cada diferencia (Después – Antes), si tal diferencia está por arriba de cero se asignará un signo más, si está por debajo se colocará un signo menos y se descartarán los empates (cuya diferencia resulte cero).

+, +, -, +, descarte, +, +, +, -, +, +, +, -, +, +

Se procedería de la siguiente forma:

1) Planteamiento del sistema de hipótesis

H_o: $\mu_D = \mu_A$
H_1: $\mu_D > \mu_A$

Donde μ_D es la media poblacional de las puntuaciones después del curso y μ_A es la media antes de tomar el curso. La hipótesis alternativa refleja la afirmación hecha por el director.

De forma equivalente se pueden plantear las siguientes hipótesis:

H_o: p=0.5
H_1: p > 0.5

En términos de la mediana poblacional (μ_e) de las diferencias, las hipótesis equivalentes serían:

H_o: $\mu_e = 0$
H_1: $\mu_e > 0$

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$.

3) La prueba es unilateral derecha.

4) La estadística de prueba se soporta ahora en una distribución binomial.

Como de la muestra se obtuvo:

+, +, -, +, descarte, +, +, +, -, +, +, +, -, +, +

Así n= 14

Se define la variable X: número de signos positivos en la muestra de tamaño 14; x=0,1,2,…,14; como la cantidad de signos positivos es x=11,

P-valor = $P(X \geq 11) = P(X = 11) + P(X = 12) + P(X = 13) + P(X = 14)$

$$P(X=11)=\binom{14}{11}(0.5)^{11}(0.5)^{14-11}=364(0.5)^{14}=0.0222$$

$$P(X=12)=\binom{14}{12}(0.5)^{12}(0.5)^{14-12}=91(0.5)^{14}=0.0055$$

$$P(X=13)=\binom{14}{13}(0.5)^{13}(0.5)^{14-13}=14(0.5)^{14}=0.00085$$

$$P(X=14)=\binom{14}{14}(0.5)^{14}(0.5)^{14-14}=1(0.5)^{14}=0.00006$$

Por lo tanto,

P-valor = $P(X \geq 11) = 0.02861$ para cuando p = 0.5

5) Como el *p*-valor resultó inferior al nivel de significancia α = 0.05, entonces se rechaza la hipótesis nula y se acepta la hipótesis alternativa.

6) Con un nivel de significancia del 5 %, se concluye que las competencias de los subgerentes sobre el manejo de bases de datos en la empresa MW se incrementaron una vez ellos terminaron el curso. En esta situación, se acepta como cierta la afirmación hecha por el director.

4. 3 Prueba de la mediana

La prueba de la mediana es una técnica no paramétrica que se aplica cuando los datos de la variable de interés no cumplan el supuesto de normalidad o cuando los datos sean bastante heterogéneos (allí el promedio muestral puede no ser un buen estimador de la media poblacional, ya que no resulta robusto y se deja afectar por los datos extremos). El procedimiento para la prueba de hipótesis de la mediana es similar a la prueba del signo; primero se obtiene el conjunto conformado por los signos más relacionados con los datos cuyo valor sea superior a la mediana, se coloca un signo menos para cada dato que resulte inferior a la mediana y se descartan aquellos que coincidan con la mediana poblacional.

La situación problema que se desarrolla concuerda con lo expuesto por Lind *et al.* (2015): el departamento de investigaciones de la empresa "Psicología del consumidor" afirma que el gasto semanal mediano para satisfacer las necesidades básicas de los matrimonios jóvenes en la ciudad de Bogotá a marzo de 2022 es de 210.000 pesos. Un sociólogo investigador afirma que es menor, y para comprobarlo ha encuestado a 35 matrimonios sobre el mencionado gasto semanal. Obtuvo los siguientes datos:

580, 181, 172, 211, 240, 267, 95, 234, 190, 165, 50, 230, 139, 140, 223, 237, 256, 180, 221, 379, 180, 150, 212, 234, 160, 190, 121, 181, 90, 229, 175, 450, 121, 190, 162.

Utilizar un nivel de significancia del 5 % para probar la hipótesis planteada por el sociólogo investigador.

+, -, -, +, +, +, -, +, -, -, -, +, -, -, +, +, +, -, +, +, -, -, +, +, -, -, -, -, -, +, -, +, -, -, -.

Para resolver la situación problema se procedería de la siguiente forma:

1) Planteamiento del sistema de hipótesis

H_o: $\mu_e = 210$
H_1: $\mu_e < 210$

De forma equivalente se pueden plantear las siguientes hipótesis:

H_o: p=0.5
H_1: $p < 0.5$

2) Se fija el nivel de significancia en el valor α = 0.05.

3) La prueba es unilateral izquierda.

4) La estadística de prueba se soporta ahora en una distribución normal.

Al observar el conjunto conformado por los signos más y menos, resulta que,

n= 35

Se define la variable X: número de signos positivos en la muestra de tamaño 35; x=0,1,2,…,35 con $p = 0.5 = q$ se hace la aproximación de una distribución binomial por una normal, para x =15 signos positivos, así:

$$Z = \frac{x-0.5n}{\sqrt{n(0.5)(0.5)}} = \frac{15-0.5(35)}{\sqrt{35(0.5)(0.5)}} = \frac{15-17.5}{\sqrt{8.75}} = \frac{-2.5}{2.958} - 0.8452$$

P-valor =
$$P(X \leq 15) = P\left(\frac{X-0.5n}{\sqrt{n(0.5)(0.5)}} \leq \frac{15-0.5n}{\sqrt{n(0.5)(0.5)}}\right) = P(Z \leq -0.8452) = 0.1977$$

Por lo tanto,

P-valor = $P(X \leq 15) = 0.1977$ para cuando $p = 0.5$

5) Como el P-valor resultó superior al nivel de significancia α = 0.05, entonces se acepta la hipótesis nula H_o: $\mu_e = 210$.

De otra forma, como el valor de la estadística de prueba Z=-0.8452 es mayor que el valor de la distribución teórica normal estándar $Z_\alpha = Z_{0.05} = -1.645$ entonces se acepta la hipótesis nula.

6) Con un nivel de significancia del 5 %, se concluye que el gasto semanal mediano destinado a satisfacer las necesidades básicas de los matrimonios jóvenes en la ciudad de Bogotá a marzo de 2022 es de 210.000 pesos, en esta situación el sociólogo investigador no tiene la razón.

4.4 Prueba del rango signado de Wilcoxon

Esta prueba tiene en cuenta la magnitud de las diferencias cuando se trata de muestras por parejas. En esta situación ya no se pierde tanta información como en la prueba del signo. En la prueba del rango signado de Wilcoxon se ordenan las diferencias sin considerar su signo, se procede a asignar el rango 1 a la menor diferencia en valor absoluto, el rango 2 a la diferencia en valor absoluto que ocupa la segunda posición, y así sucesivamente se asigna el rango n a la diferencia que en valor absoluto resulte más grande. Si se presentan diferencias iguales a cero se descartan. Además, si dos o más diferencias absolutas resultan iguales, a cada una se le asigna la media de los rangos comprendidos. En esta prueba se halla el valor de la estadística T^+ al sumar los rangos correspondientes a las diferencias positivas, el valor de la estadística T^- al sumar los rangos correspondientes a las diferencias negativas y el valor de la estadística T se obtiene como el mínimo entre T^+ y T^-; así entonces $T = \min (T^+, T^-)$.

Como $T^+ + T^- = n(n+1)/2$ entonces todas las pruebas basada en estas estadísticas resultan equivalentes. Además, de acuerdo con Lehmann y D'Abrera (1975), las estadísticas T^+ y T^- asumen valores de una variable aleatoria, cuyos valores están en el intervalo desde cero hasta $n(n+1)/2$, cada una con distribución simétrica respecto a su valor esperado que equivale $n(n+1)/4$.

En la prueba de hipótesis de la forma

H_o: $\mu = \mu_o$
H_1: μ diferente de μ_o

La hipótesis nula se rechaza cuanto T sea menor o igual que T_α

En la prueba de hipótesis de la forma para muestras
H_o: $\mu = \mu_o$
H_1: $\mu > \mu_o$

La hipótesis nula se rechaza cuanto T^- sea menor o igual que $T_{2\alpha}$

En la prueba de hipótesis de la forma
H_o: $\mu = \mu_o$
H_1: $\mu < \mu_o$

La hipótesis nula se rechaza cuanto T^+ sea menor o igual que $T_{2\alpha}$

Ahora, para la prueba de hipótesis basada en muestras por parejas, el planteamiento y el criterio de rechazo de la hipótesis nula es el siguiente:

H_o: $\mu_1 = \mu_2$
H_1: μ_1 difiere de μ_2

La hipótesis nula se rechaza cuanto T sea menor o igual que T_α

H_o: $\mu_1 = \mu_2$
H_1: $\mu_1 > \mu_2$

La hipótesis nula se rechaza cuanto T^- sea menor o igual que $T_{2\alpha}$

En la prueba de hipótesis de la forma
H_o: $\mu_1 = \mu_2$
H_1: $\mu_1 < \mu_2$

La hipótesis nula se rechaza cuanto T^+ sea menor o igual que $T_{2\alpha}$

Es conveniente mencionar que para n mayor o igual que 15, la distribución de probabilidad de la estadística T^+ se puede aproximar por una normal estándar, así:

$$Z = \frac{T^+ - E(T^+)}{\sqrt{Var(T^+)}}$$

Donde:

$$E(T^+) = \frac{n(n+1)}{4}, \quad Var(T^+) = \frac{n(n+1)(2n+1)}{24}$$

La siguiente situación problema ha sido planteada siguiendo lo expresado por Freund y Miller (2000): un ingeniero de petróleos ha tomado 15 mediciones en relación con el octanaje de cierta clase de combustible producido en la refinería WX, y ha registrado los siguientes datos: 97.4, 95.1, 97.2, 95.9, 96.7, 100.2, 97.3, 95.2, 93.1, 99, 96, 97.5, 98.1, 98.4, 94.8.

El ingeniero afirma que el promedio del octanaje de esa clase de combustible es de 98.4 en la refinería WX; un usuario difiere de esa afirmación y espera probarlo con la misma muestra obtenida por el ingeniero; utilizar un nivel de significancia del 5 % para probar la hipótesis planteada por el ingeniero.

Para solucionar la situación problema mediante la prueba del rango signado de Wilcoxon, en primera instancia se obtienen las diferencias con respecto a la media poblacional de μ = 98.4, luego se asignan los rangos a sus correspondientes valores absolutos, se determinan los valores de los rangos referidos a T^+ y de T^-, a partir de los cuales se establece el valor del estadístico T. Los resultados se pueden observar en la Tabla 4.3. Se descartará cuando se presenten empates (cuya diferencia resulte cero).

De los resultados obtenidos en la Tabla 4.3, se determina que:

Tabla 4.3. Valores de las estadísticas T^+, T^-

Mediciones	Diferencias	Rango absoluto	Rangos para T^+	Rangos para T^-
97.4	-1	4		4
95.1	-3.3	12		12
97.2	-1.2	6		6
95.9	-2.5	10		10
96.7	-1.7	7		7
100.2	1.8	8	8	
97.3	-1.1	5		5
95.2	-3.2	11		11
93.1	-5.3	14		14
99	0.6	2	2	
96	-2.4	9		9
97.5	-0.9	3		3
98.1	-0.3	1		1
98.4	0.0	Descarte		
94.8	-3.6	13		13

Fuente: los autores

n =14 al considerar el descarte

$T^+ = 2 + 8 = 10$
$T^- = 1+3+4+5+6+7+9+10+11+12+13+14+15 = 95$

$T = \min (T^+, T^-) = \min (10, 95) = 10$

El conjunto de pasos para probar las hipótesis son los siguientes:

1) Planteamiento del sistema de hipótesis

H_o: μ = 98.4
H_1: μ difiere de 98.4

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$.

3) La prueba es bilateral.

4) La estadística de prueba corresponde a la estadística T

n= 14

T = 10

Para $\alpha = 0.05$ con n = 14 resulta que $T_\alpha = T_{0.05} = 21$ (Freund y Miller, 2000, p. 590)

5) Como el valor de la estadística T =10 es menor que $T_{0.05} = 21$, entonces se rechaza la hipótesis nula H_o: μ = 98.4 y se acepta la hipótesis alternativa.

6) Con un nivel de significancia del 5 %, se concluye que el promedio del octanaje de esa clase de combustible difiere de 98.4 en la refinería WX; por consiguiente, la afirmación hecha por el ingeniero de petróleos no es cierta, en esta situación el usuario tiene la razón.

La siguiente situación problema sigue los lineamientos sugeridos por Lind *et al*. (2015): en un estudio de mercado el investigador ha seleccionado una muestra aleatoria de 15 clientes. A cada cliente se le entrega una pieza del producto A comestible que se acompaña con la salsa tradicional y se le solicita que califique su sabor con una escala desde 1 hasta 10 puntos; un valor cercano a 10 será indicativo de que al cliente le ha gustado el sabor del producto A con la salsa tradicional, y un valor próximo a cero será indicativo de que no le gustó. Después de un tiempo, a los mismos 15 clientes se les entrega una pieza del producto A pero acompañada con una salsa de ciruelas, y nuevamente se les solicita que califiquen su sabor en una escala desde 1 hasta 10 puntos. Los resultados se pueden observar en la Tabla 4.4.

El investigador afirma que en general el promedio de los puntajes referidos al sabor de la pieza del producto A acompañado con la salsa de ciruelas es mayor que cuando se acompaña el producto A con la salsa tradicional; usar un nivel de significancia del 5 % para probar la hipótesis planteada por el investigador.

Para solucionar la situación problema objeto de estudio mediante la prueba del rango signado de Wilcoxon, primero se obtienen las diferencias entre las puntuaciones obtenidas de saborear el producto A acompañado de la salsa de ciruelas y las reportadas al saborear este producto acompañado de la salsa tradicional; luego se asignan los rangos a sus correspondientes valores absolutos, se determinan los valores de los rangos referidos a T^+ y de T^- , a partir de estos se establece el valor del estadístico T de ser necesario; de lo contrario, se puede recurrir al uso de la aproximación del estadístico T^+ por medio de una distribución normal.

Los resultados se pueden observar en la Tabla 4.4. Se descartará cuando se presenten empates (cuya diferencia resulte cero).

De los resultados presentados en la Tabla 4.3, se deduce que:

n =14

$T^+ = 1+3+13+6+9+9+6+2+14+12 = 75$

$T^- = 6+11+9+4 = 30$

$T = \min (T^+, T^-) = \min (75, 30) = 30$

El conjunto de pasos para probar las hipótesis son los siguientes:

Tabla 4.4. Valores de las estadísticas T^+, T^-

Puntajes Salsa tradicional	Puntajes Salsa de ciruelas	Diferencias	Rango absoluto	Rangos para T^+	Rangos para T^-
6	7	1	1	1	
8	4	-4	6		6
1	3	2	3	3	
2	9	7	13	13	
6	10	4	6	6	
8	8	0	Descarte		
2.5	7	4.5	9	9	
8	3	-5	11		11
5	9.5	4.5	9	9	
5	9	4	6	6	
6.5	8	1.5	2	2	
1	9	8	14	14	
6.5	2	-4.5	9		9
7	3.5	-3.5	4		4
2	8	6	12	12	

Fuente: los autores

1) Planteamiento del sistema de hipótesis

H_o: $\mu_1 = \mu_2$
H_1: $\mu_1 < \mu_2$

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$.

3) La prueba es unilateral izquierda.

4) La estadística de prueba corresponde a la estadística T^+

$n = 14$

Usando la aproximación normal para T^+, se obtiene:

$$E(T^+) = \frac{n(n+1)}{4} = \frac{14(14+1)}{4} = 52.5,$$

$$Var(T^+) = \frac{14(14+1)(2(14)+1)}{24} = \frac{14(15)(29)}{24} = 253.75$$

$$Z = \frac{T^+ - E(T^+)}{\sqrt{Var(T^+)}} = \frac{75-52.5}{\sqrt{253.75}} = \frac{22.5}{15.9295} = 1.4124$$

Para $\alpha = 0.05$ resulta que $Z_\alpha = Z_{0.05} = -1.645$

5) Como $Z = 1.4124$ es mayor que $Z_{0.05} = -1.645$ entonces se acepta la hipótesis nula; por consiguiente, se acepta que H_o: $\mu_1 = \mu_2$

De otra forma, para $\alpha = 0.05$ con $n = 14$ resulta que $T_{2\alpha} = T_{0.10} = 26$ (Freund y Miller, 2000, p. 590), como $T^+ = 75$ no es menor o igual que $T_{2\alpha} = T_{0.10} = 26$ entonces no se rechaza la hipótesis nula; luego no existen diferencias significativas en las calificaciones otorgadas por los sabores del producto A cuando se acompañan de las salsas de ciruelas o de la salsa tradicional.

Por otra parte, el valor de la estadística $T^{-} = 30$ tampoco es menor o igual que $T_{2\alpha} = T_{0.10} = 26$; lo cual también respalda la decisión tomada.

6) Con un nivel de significancia del 5 %, se concluye que, en general, el promedio de los puntajes referidos al sabor de la pieza del producto A acompañado con la salsa de ciruelas no es mayor que cuando se acompaña el producto A con la salsa tradicional; por consiguiente, la afirmación hecha por el investigador no es estadísticamente cierta.

4.5 Prueba U de Mann-Whitney

La prueba U es no paramétrica y se constituye en una alternativa para la prueba t-student para dos muestras independientes, cuando no se cumple el supuesto de normalidad para las dos poblaciones de las cuales se muestrea, pero los datos han de corresponder a una variable continua. La hipótesis nula se plantea en términos de que se muestrea de poblaciones idénticas (sus medias poblacionales son iguales) y la hipótesis alterna indica que tales poblaciones son distintas (sus medias son desiguales).

Ahora, si n_1 es el tamaño de la primera muestra cuyos datos son $x_1, x_2, x_3, ..., x_{n_1}$ y n_2 es el tamaño de la segunda muestra cuyos datos son $x_1, x_2, x_3, ..., x_{n_2}$ en primera instancia se conforma una muestra combinada, la cual estará constituida por $n_1 + n_2$ datos escrito de la siguiente forma: $x_1, x_2, x_3, ..., x_{n_1}, x_{n_1+1}, x_{n_1+2}, x_{n_1+3}, ..., x_{n_1+n_2}$, se prosigue a ordenar de menor a mayor los $n_1 + n_2$ datos y se determina el rango (posición que ocupa cada dato en la secuencia ordenada) para cada dato de la muestra combinada; se calcula la estadística W_1 como la suma de los rangos de los datos correspondientes a la primera muestra y W_2 como la suma de los rangos para los datos de la segunda muestra; en este caso, no interesa si se trabaja con W_1 o con W_2, puesto que $W_1 + W_2$ será siempre igual a la suma de los primeros $n_1 + n_2$ números enteros positivos, es decir:

$$W_1 + W_2 = \frac{(n_1 + n_2)(n_1 + n_2 + 1)}{2}$$

Posteriormente, se obtienen las siguientes estadísticas relacionadas con W_1 y W_2, así:

$$U_1 = W_1 - \frac{n_1(n_1 + 1)}{2}$$

$$U_2 = W_2 - \frac{n_2(n_2 + 1)}{2}$$

$$U = \min(W_1, W_2)$$

En esta situación, las pruebas que se basan en U, U_1 o U_2 son equivalentes a las pruebas que se basan en W_1 o W_2 (Lehmann y D'Abrera, 1975). Ahora, $U_1 + U_2 = n_1 * n_2$; además, cada una de las variables U_1, U_2 tienen distribuciones idénticas, también cada una es simétrica con respecto al valor $(n_1 * n_2)/2$.

Para la prueba de hipótesis de la forma:

H_o: $\mu_1 = \mu_2$
H_1: μ_1 difiere de μ_2

La hipótesis nula se rechaza cuanto U sea menor o igual que U_α

H_o: $\mu_1 = \mu_2$
H_1: $\mu_1 > \mu_2$

La hipótesis nula se rechaza cuanto U_2 sea menor o igual que $U_{2\alpha}$

En la prueba de hipótesis de la forma

H_o: $\mu_1 = \mu_2$
H_1: $\mu_1 < \mu_2$

La hipótesis nula se rechaza cuanto U_1 sea menor o igual que $U_{2\alpha}$

Es pertinente señalar que, para cuando n_1 y n_2 corresponden a tamaños mayores que 8, la distribución de probabilidad de las variables U_1 y U_2 pueden aproximarse por una normal estándar, así:

$$Z = \frac{U_1 - E(U_1)}{\sqrt{Var(U_1)}}$$

Donde:

$$E(U_1) = \frac{n_1 n_2}{2}, \quad Var(U_1) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

De similar modo:

$$Z = \frac{U_2 - E(U_2)}{\sqrt{Var(U_2)}}$$

Donde:

$$E(U_2) = \frac{n_1 n_2}{2}, \quad Var(U_2) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

Además,

$$Z = \frac{U - E(U)}{\sqrt{Var(U)}}$$

Con:

$$E(U) = \frac{n_1 n_2}{2}, \quad Var(U) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

La siguiente situación problema ha sido adaptada de Freund y Miller (2000): un ingeniero eléctrico desea comparar dos clases de señales luminosas, para eso ha seleccionado una muestra aleatoria de tamaño n_1 = 9 de la marca A y ha medido sus tiempos de iluminación en minutos; además ha seleccionado otra muestra aleatoria de tamaño n_2 = 10 de la marca B y ha medido sus correspondientes tiempos de iluminación en minutos, así:

Marca A: 15.9, 12.3, 14.2, 17.6, 18.0, 15.1, 16.4, 14.0, 17.9
Marca B: 16.2, 20.8, 15.7, 19.3, 17.2, 22.2, 19.9, 13.2, 16.3, 20.4

El ingeniero afirma que el promedio de los tiempos de iluminación de las señales luminosas de la marca A es inferior al promedio para la marca B; utilizar un nivel de significancia del 5 % para probar la hipótesis formulada por el ingeniero.

Para solucionar la situación problema mediante la prueba *U* de Mann-Whitney, en primera instancia se obtiene la muestra combinada y después se asignan los rangos desde 1 hasta 19; luego se establece que los rangos correspondientes a los datos de la primera muestra son: 1, 3, 4, 5, 7, 10, 12, 13, 14; en consecuencia, los rangos para los datos de la segunda muestra son: 2, 6, 8, 9, 11, 15, 16, 17, 18, 19. Con estos resultados se deduce que,

$W_1 = 1 + 3 + 4 + 5 + 7 + 10 + 12 + 13 + 14 = 69$

$W_2 = 2 + 6 + 8 + 9 + 11 + 15 + 16 + 17 + 18 + 19 = 121$

$$U_1 = W_1 - \frac{n_1(n_1+1)}{2} = 69 - \frac{9(9+1)}{2} = 69 - 45 = 24$$

$$U_2 = W_2 - \frac{n_2(n_2+1)}{2} = 121 - \frac{10(10+1)}{2} = 121 - 55 = 66$$

$$U = \min(W_1, W_2) = \min(24, 66) = 24$$

El conjunto de pasos para probar las hipótesis respectivas son los siguientes:

1) Planteamiento del sistema de hipótesis

H_o: $\mu_1 = \mu_2$
H_1: $\mu_1 < \mu_2$

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$.

3) La prueba es unilateral izquierda.

4) La estadística de prueba corresponde a la estadística U_1 que se ha de comparar con $U_{2\alpha}$ para tomar una decisión sobre la hipótesis nula.

Para $\alpha = 0.05$ el valor $U_{2\alpha} = U_{2(0.05)} = U_{0.10} = 24$ al considerar $n_1 = 9$ y $n_2 = 10$ (Freund y Miller, 2000, p. 591)

5) Como $U_1 = 24$ es igual a $U_{0.10} = 24$, se cumple que $U_1 \leq U_{2\alpha}$ en consecuencia se rechaza la hipótesis nula y se acepta la hipótesis alternativa H_1: $\mu_1 < \mu_2$

6) Con un nivel de significancia del 5 %, se concluye que el promedio de los tiempos de iluminación de las señales luminosas de la marca A es inferior al promedio para la marca B; por consiguiente, la afirmación hecha por el ingeniero eléctrico es cierta.

En la siguiente situación problema hipotética de investigación se consideran dos muestras aleatorias, cuyos datos reflejan el aumento de peso de dos muestras aleatorias de pavos, los cuales han sido alimentados mediante dos dietas distintas, pero manteniendo un ambiente de vida diaria similar. La primera muestra aleatoria resultó de tamaño $n_1 = 16$; además, la segunda muestra aleatoria también tiene tamaño $n_2 = 16$, sus correspondientes valores de su ganancia de peso son los siguientes:

Dieta A: 15.3, 9.1, 9.7, 12.5, 13, 12.2, 17.4, 13.5, 14.1, 22.6, 13.7, 11, 13.9, 10.8, 13.3, 9.2

Dieta B: 22.8, 20.3, 14.4, 18.6, 18.2, 17.8, 12.9, 11, 15.2, 14.8, 14.3, 19.1, 20.1, 19.7, 13.8, 17.9

El dueño de la granja donde se crían los pavos afirma que la dieta B, en promedio, genera una mayor ganancia de peso que la dieta A. Usar un nivel de significancia del 5 % para probar la hipótesis formulada por el dueño de la granja.

Para resolver esta situación problema se propone utilizar la prueba U de Mann-Whitney; para esto, en primer lugar se obtiene la muestra combinada y después se asignan los rangos desde 1 hasta 32; luego se establece que los rangos correspondientes a los datos de los pavos alimentados con la dieta A son: 1, 2, 3, 4, 5.5, 7, 8, 10, 11, 12, 13, 15, 16, 21, 22, 31; ahora los rangos de los datos para los pavos alimentados con la dieta B son: 5.5, 9, 14, 17, 18, 19, 20, 23, 24, 25, 26, 27, 28, 29, 30, 32. Con estos resultados se deduce que,

$$W_1 = 1+2+3+5.5+7+8+10+11+12+13+15+16+21+22+31 = 181.5$$

W_2= 5.5+9+14+17+18+19+20+23+24+25+26+27+28+29+30+32= 346.5

$$U_1 = W_1 - \frac{n_1(n_1+1)}{2} = 181.5 - \frac{16(16+1)}{2} = 181.5 - 136 = 45.5$$

$$U_2 = W_2 - \frac{n_2(n_2+1)}{2} = 346.5 - \frac{16(16+1)}{2} = 346.5 - 136 = 210.5$$

$$U = \min(W_1, W_2) = \min(45.5, 210.5) = 45.5$$

El conjunto de pasos para probar las hipótesis respectivas son los siguientes:

1) Planteamiento del sistema de hipótesis

H_o: $\mu_1 = \mu_2$
H_1: $\mu_1 < \mu_2$

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$.

3) La prueba es unilateral izquierda.

4) La estadística de prueba corresponde a la estadística U_1 que se debe comparar con $U_{2\alpha}$ para tomar una decisión sobre la hipótesis nula; aunque, para $n_1 = 16$ y $n_2 = 16$ no suele encontrarse el valor de U en los libros de texto universitarios. Por consiguiente, se recurre a su aproximación por medio de la distribución normal estándar. En este contexto

$$E(U_1) = \frac{n_1 n_2}{2} = \frac{(16)(16)}{2} = \frac{256}{2} = 128$$

$$Var(U_1) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12} = \frac{(16)(16)(16+16+1)}{12} = \frac{256(33)}{12} = 704$$

$$Z = \frac{45.5 - 128}{\sqrt{704}} = \frac{-82.5}{26.533} = -3.1093$$

Por otra parte, para $\alpha = 0.05$ en una prueba unilateral izquierda, el valor teórico en la distribución normal estándar es $Z_\alpha = Z_{0.05} = -1.645$

5) Como el valor Z = -3.1093 es menor que $Z_\alpha = Z_{0.05} = -1.645$ entonces se rechaza la hipótesis nula y se acepta la hipótesis alternativa H_1: $\mu_1 < \mu_2$

6) Con un nivel de significancia del 5 %, se concluye que la dieta B destinada a alimentar los pavos en promedio genera una mayor ganancia de peso en los pavos que la dieta A; por consiguiente, la afirmación hecha por el dueño de la granja es estadísticamente cierta.

5

Situaciones problema no paramétricas relacionadas con más de dos muestras

En este capítulo se analizan distintas situaciones problema relativas al uso frecuente de pruebas no paramétricas relacionadas con más de dos muestras aleatorias. Esta clase de pruebas estadísticas se emplean cuando el interés del investigador recae en determinar si existen diferencias significativas entre tres o más tratamientos o si un tratamiento trae mejores ventajas que otros desde una mirada estadística. En la primera sección se aborda la prueba de Kruskal-Wallis, que corresponde a una generalización de la prueba basada en el estadístico U para dos muestras. En la segunda se hace referencia a la prueba de Friedman, la cual pertenece a una prueba basada en rangos para más de dos muestras relacionadas que incluye el análisis de varianza de dos clasificaciones por medio del uso de la teoría de rangos. En la tercera se indica el manejo del coeficiente de correlación por rangos de Spearman. En la cuarta se trabaja la prueba de independencia para tablas de contingencia. En el contexto de tres o más muestras también es factible aplicar un conjunto ordenado de pasos para elaborar un proceso de inferencia estadística vinculado a la prueba de hipótesis.

5.1 La prueba de Kruskal-Wallis

La prueba de Kruskal-Wallis también se conoce con el nombre de prueba no paramétrica *H*. Esta prueba se puede considerar como una generalización de la prueba *U* de Mann Whitney, ya que puede incluir tres o más muestras aleatorias independientes provenientes de poblaciones en las cuales resulta de interés el estudio de una determinada variable continua. La prueba *H* puede utilizarse cuando en el análisis de varianza en una sola vía no se cumpla el supuesto de normalidad para los datos de las k muestras.

Ahora, si se tienen *k* muestras donde n_1 es el tamaño de la primera muestra cuyos datos son $x_1, x_2, x_3, ..., x_{n_1}$, n_2 es el tamaño de la segunda muestra cuyos datos son $x_1, x_2, x_3, ..., x_{n_2}$, así sucesivamente n_k corresponde al tamaño de la *k*-ésima muestra cuyos datos son $x_1, x_2, x_3, ..., x_{n_k}$. En primer lugar, se conforma una muestra combinada, la cual estará constituida por $n_1 + n_2 +\ldots+ n_k$ datos; se prosigue a ordenar de menor a mayor los $n_1 + n_2 +\ldots+ n_k$ datos y se determina el rango (posición que ocupa cada dato en la secuencia ordenada) para cada dato de la muestra combinada, se calcula R_i como la suma de los rangos de los datos correspondientes a la *i*-ésima muestra, luego se determina la estadística *H*, de la siguiente manera:

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

En este contexto, $n = n_1 + n_2 +\ldots+ n_k$. Para cuando los tamaños de las *k* muestras son relativamente grandes, la distribución de probabilidad de la estadística *H* se puede aproximar por medio de una distribución chi-cuadrado con *k-1* grados de libertad. Aunque en la práctica el valor del estadístico *H* se compara con el valor crítico de una distribución chi-cuadrado.

Un procedimiento para la prueba de hipótesis puede incluir los siguientes pasos:

1) Planteamiento del sistema de hipótesis

H_o: $\mu_1 = \mu_2 = \ldots = \mu_k$
H_1: las μ_i no son todas iguales

O de forma equivalente:

H_o: $\mu_1 = \mu_2 = \ldots = \mu_k$
H_1: algunas de las μ_i son diferentes

En algunas ocasiones también las hipótesis se suelen plantear como sigue:

H_o: $f_1(x) = f_2(x) = \ldots = f_k(x)$
H_1: las $f_i(x)$ no son todas idénticas

En este contexto, las $f_i(x)$ son las funciones de densidad de probabilidad de las poblaciones de donde provienen las k muestras.

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$.

3) La prueba es unilateral derecha.

4) La estadística de prueba corresponde a la estadística H que se ha de comparar con una chi-cuadrado con k-1 grados de libertad $\chi^2_{1-\alpha,\, k-1}$

5) Para tomar una decisión sobre la hipótesis nula se aplica el siguiente criterio: si H resulta mayor o igual que el valor teórico $\chi^2_{1-\alpha,\, k-1}$ entonces se rechaza la hipótesis nula; en caso contrario, se aceptará.

6) Conclusión

La siguiente situación problema puede considerarse como un problema hipotético de investigación, el cual puede asemejarse al expuesto por Canavos (1988). Un investigador ha tomado 4 muestras aleatorias independientes de casas, las cuales han sido vendidas en los dos últimos meses y están ubicadas en 4 zonas de la ciudad de Medellín en Colombia. El investigador quiere establecer si existen diferencias sobre el valor de estas casas en las 4 zonas. Tal valor tiene en cuenta la relación entre el precio de venta de cada casa y el precio registrado por el inmueble en el recibo de impuesto predial. La información procesada para el valor de las casas en cada muestra es la siguiente:

Zona 1: 2.19, 2.05, 2.14, 2.25, 2.29

Zona 2: 2.08, 2.23, 2.26, 2.1, 2.18, 2.14

Zona 3: 1.98, 2.19, 2.08, 1.93, 2.23, 2.18

Zona 4: 2.12, 2.14, 2.31, 2,12, 2.19

En este contexto, n_1 = 5 casas en la zona 1, n_2 = 6 casas ubicadas en la zona 2, n_3 = 6 para la zona 4 y n_4 = 5 de la zona 4.

El investigador afirma que el promedio del valor de las casas es el mismo, ya que no existen diferencias al relacionar su precio de venta y el precio que aparece en el recibo del impuesto predial. Usar un nivel de significancia del 5 % para probar la hipótesis formulada por el investigador.

Para solucionar esta situación problema se puede utilizar la prueba H de Kruskal-Wallis. Por consiguiente, primero se obtiene la muestra combinada y después se asignan los rangos desde 1 hasta 22; luego se establece que los rangos correspondientes a los datos de la primera muestra son: 15, 3, 10, 19, 21; los rangos para los datos de la segunda muestra son: 4.5, 17.5, 20, 6, 12.5, 10; los rangos para los datos de la tercera muestra son: 2, 15, 4.5, 1, 17.5, 12.5 y los rangos para los datos de la cuarta muestra son: 7.5, 10, 22, 7.5, 15. Se observa que en la muestra combinada se presentaron varios empates, para los cuales se tomó el promedio de sus respectivos rangos. Con estos valores de los rangos se determina que:

$R_1 = 15 + 3 + 10 + 19 + 21 = 68$

$R_2 = 4.5 + 17.5 + 20 + 6 + 12.5 + 10 = 70.5$

$R_3 = 2 + 15 + 4.5 + 1 + 17.5 + 12.5 = 52.5$

$R_4 = 7.5 + 10 + 22 + 7.5 + 15 = 62$

El tamaño de la muestra combinada es $n = n_1 + n_2 + \ldots + n_k$, entonces $n=5+6+6+5=22$.

Ahora se procede a efectuar el cálculo de la estadística *H*, así:

$$H = \frac{12}{n(n+1)}\left(\sum_{i=1}^{k=4}\frac{R_i^2}{n_i}\right) - 3(n+1)$$

$$H = \frac{12}{n(n+1)}\left(\frac{R_1^2}{n_1}+\frac{R_2^2}{n_2}+\frac{R_3^2}{n_3}+\frac{R_4^2}{n_4}\right) - 3(n+1)$$

$$H = \frac{12}{22(22+1)}\left(\frac{68^2}{5}+\frac{70.5^2}{6}+\frac{52.5^2}{6}+\frac{62^2}{5}\right) - 3(22+1)$$

$$H = \frac{12}{506}\left(\frac{4624}{5}+\frac{4970.25}{6}+\frac{2756.25}{6}+\frac{3844}{5}\right) - 69 = 1.7$$

El procedimiento para la prueba de hipótesis es:
1) Planteamiento del sistema de hipótesis

H_o: $\mu_1 = \mu_2 = \mu_3 = \mu_4$
H_1: algunas de las μ_i son diferentes para i = 1,2,3,4

O de forma equivalente:

H_o: $f_1(x) = f_2(x) = f_3(x) = f_4(x)$
H_1: las $f_i(x)$ no son todas idénticas

En este contexto, las $f_i(x)$ son las funciones de densidad de probabilidad de las 4 zonas de donde provienen las 4 muestras.

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$.

3) La prueba es unilateral derecha.

4) La estadística de prueba es $H = 1.7$

Ahora, el valor teórico de la distribución chi-cuadrado $\chi^2_{1-\alpha,\, k-1}$ con k-1 = 4-1 = 3 grados de libertad para $\alpha = 0.05$ es

$$\chi^2_{1-\alpha,\, k-1} = \chi^2_{1-0.05,\, 4-1} = \chi^2_{0.95,\, 3} = 7.8147$$

5) Como $H = 1.7$ no es mayor o igual que $\chi^2_{1-\alpha,\, k-1} = 7.8147$ entonces no se rechaza la hipótesis nula; por consiguiente, se acepta la hipótesis nula H_o: $\mu_1 = \mu_2 = \mu_3 = \mu_4$

6) Con un nivel de significancia del 5 % se concluye que el promedio del valor de las casas es el mismo, ya que no existen diferencias al relacionar su precio de venta y el precio que aparece en el recibo del impuesto predial; por consiguiente, la afirmación hecha por el investigador es cierta. Es decir, no hay evidencias suficientes para demostrar que existan diferencias significativas en el valor de las casas vendidas en las 4 zonas.
La siguiente situación problema puede considerarse como una situación hipotética de investigación, susceptible de replicarse en el salón de clases universitario. Un investigador en ciencias de la educación ha decidido tomar 3 muestras aleatorias independientes referidas a las calificaciones

obtenidas por los estudiantes de la asignatura de probabilidad y estadística ofrecida en 3 cursos diferentes de ingeniería (civil, electrónica e industrial) en una universidad de la ciudad de Bogotá en Colombia. El investigador quiere establecer si existen diferencias sobre las calificaciones obtenidas por los estudiantes de esta asignatura en esos cursos de ingeniería. La información recolectada varía en el rango desde 0.0 hasta 5.0 y fue la siguiente:

Muestra en ingeniería civil: 4.7, 4.4, 4.6, 3.7, 4.3, 4.9

Muestra en ingeniería electrónica: 4.2, 4.1, 3.8, 4.0, 3.1, 3.5, 3.9

Muestra en ingeniería industrial: 4.5, 3.4, 3.5, 3.6, 3.4

El investigador afirma que en promedio no hay diferencias significativas en las calificaciones de la asignatura de probabilidad y estadística en los 3 cursos. Utilizar un nivel de significancia del 5 % para probar la hipótesis formulada por el investigador en ciencias de la educación.

Para solucionar la situación problema objeto de investigación se puede utilizar la prueba H de Kruskal-Wallis, por lo tanto, en primera instancia se obtiene la muestra combinada y después se asignan los rangos desde 1 hasta 18. En esta situación, n_1 = 6 estudiantes, n_2 = 7, n_3 = 5. El tamaño de la muestra combinada es n = 6+7+5 = 18; en la muestra combinada se ha colocado el rango entre paréntesis a la derecha de cada uno de los datos, así:

4.7 (17), 4.4 (14), 4.6 (16), 3.7 (7), 4.3 (13), 4.9 (18), 4.2 (12), 4.1 (11), 3.8 (8), 4.0 (10), 3.1 (1), 3.5 (4.5), 3.9 (9), 4.5 (15), 3.4 (3), 3.5 (4.5), 3.6 (6), 3.3 (2)

Así, los rangos correspondientes a los datos de la muestra en ingeniería civil son: 17, 14, 16, 7, 13, 18. Los rangos correspondientes a los datos de la muestra en ingeniería electrónica son: 12, 11, 8, 10, 1, 4.5, 9. Los rangos correspondientes a los datos de la muestra en ingeniería industrial son: 15, 3, 4.5, 6, 2.

Se puede observar que en la muestra combinada se presentaron dos empates, para los cuales se tomó el promedio de sus respectivos rangos y se obtuvo 4.5; con estos rangos se deduce que:

$R_1 = 17 + 14 + 16 + 7 + 13 + 18 = 85$

$R_2 = 12 + 11 + 8 + 10 + 1 + 4.5 + 9 = 55.5$

$R_3 = 15 + 3 + 4.5 + 6 + 2 = 30.5$

Ahora, se procede a efectuar el cálculo de la estadística *H*, del siguiente modo:

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

$$H = \frac{12}{n(n+1)}\left(\frac{R_1^2}{n_1}+\frac{R_2^2}{n_2}+\frac{R_3^2}{n_3}\right) - 3(n+1)$$

$$H = \frac{12}{18(18+1)}\left(\frac{85^2}{6}+\frac{55.5^2}{7}+\frac{30.5^2}{5}\right) - 3(18+1)$$

$$H = \frac{12}{18(19)}\left(\frac{7225}{6}+\frac{3080.25}{7}+\frac{930.25}{5}\right) - 3(19)$$

$$H = \frac{12}{342}\left(1204.1666 + 440.0357 + 186.05\right) - 57 = 0.0351(1830.2523) - 57$$

$$H = 64.2418 - 57 = 7.2418$$

El procedimiento para la prueba de hipótesis es:

1) Planteamiento del sistema de hipótesis

H_o: $\mu_1 = \mu_2 = \mu_3$
H_1: algunas de μ_i son diferentes para i = 1,2,3

O de forma equivalente:

H_o: $f_1(x) = f_2(x) = f_3(x)$
H_1: las $f_i(x)$ no son todas idénticas

En este contexto, las $f_i(x)$ son las funciones de densidad de probabilidad de los 3 cursos de ingeniería de donde provienen las 3 muestras.

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$.

3) La prueba es unilateral derecha.

4) La estadística de prueba es $H = 7.2418$
Ahora, el valor teórico de la distribución chi-cuadrado $\chi^2_{1-\alpha,\, k-1}$ con $k-1$ = 3-1 = 2 grados de libertad para $\alpha = 0.05$ es,

$$\chi^2_{1-\alpha,\, k-1} = \chi^2_{0.95,\, 2} = 5.9915$$

5) Como $H = 7.2418$ es mayor que $\chi^2_{1-\alpha,\, k-1} = 5.9915$ entonces se rechaza la hipótesis nula; por consiguiente, se acepta la hipótesis alternativa así,

H_1: algunas de μ_i son diferentes para i = 1,2,3

6) Con un nivel de significancia del 5 %, se concluye que en promedio sí hay diferencias significativas en las calificaciones de la asignatura de

probabilidad y estadística en los 3 cursos en la universidad ubicada en la ciudad de Bogotá, Colombia; por lo tanto, la afirmación hecha por el investigador en ciencias de la educación no es cierta. Es decir, hay evidencias suficientes de que estadísticamente existen diferencias significativas en las calificaciones promedio en los 3 cursos. También se puede afirmar que las poblaciones de donde provienen los datos de las muestras no son todas idénticas, algunas son diferentes.

5.2 Prueba de Friedman con *k* muestras

Esta prueba puede considerarse como una extensión de la prueba de Wilcoxon para muestras por parejas (dos tratamientos). Ahora se tienen tres o más tratamientos en un diseño de bloques completamente al azar y los tamaños de muestras igualados con n datos en cada tratamiento. La prueba de Friedman resulta pertinente para estudiar *k* mayor que dos tratamientos con un solo factor donde puede estar presente un factor externo y se tienen mediciones de una variable de interés en, por lo menos, una escala ordinal. Esta prueba resulta útil para establecer si los efectos atribuibles a los tratamientos son los mismos (Lehmann y D'Abrera, 1975; Siegel, 1970).

Para aplicar esta prueba se conforma un bloque para cada una de las n condiciones referidas a los factores externos, de tal forma que cada bloque contiene un dato (observación) perteneciente a cada uno de los k tratamientos (Canavos, 1988). Adicionalmente, se trabaja en el supuesto de que los tratamientos se asignan de manera aleatoria y que no hay interacción entre los bloques y los tratamientos. En total hay nk observaciones, las cuales permiten conformar un arreglo matricial donde las columnas son los tratamientos (1, 2, …, j,…*k*) y las filas son los bloques (1, 2, ..., n), como se puede observar en la Tabla 5.1.

Tabla 5.1 Arreglo matricial para aplicar la prueba de Friedman

		Tratamientos			
		1	2 …	j …	k
Bloques	1	x_{11}	x_{12}	x_{1j}	x_{1k}
	2	x_{21}	x_{22}	x_{2j}	x_{2k}
	:	:	:	:	:
	n	x_{n1}	x_{n2}	x_{nj}	x_{nk}

Fuente: los autores

Para aplicar la prueba de Friedman se comienza conformando la matriz con los tratamientos y los bloques. Luego, para cada bloque (fila) se asigna un rango desde 1 hasta *k* (posición que ocupa cada dato en ese bloque); en este caso, para cualquier bloque los rangos deben corresponder a una permutación aleatoria de los números enteros desde 1 hasta *k*. Después, se calcula R_j como la suma de los rangos del tratamiento j con j =1,2,…,*k*. Se prosigue a determinar la estadística *Q* de Friedman, de la siguiente manera (Lehmann y D'Abrera, 1975):

$$Q = \frac{12}{nk(k+1)}\left(\sum_{j=1}^{k} R_j^2\right) - 3n(k+1)$$

Para cuando los tamaños de las k muestras son relativamente grandes, la distribución de probabilidad de la estadística *Q* se puede aproximar por medio de una distribución chi-cuadrado con *k-1* grados de libertad. Aunque en la práctica el valor del estadístico *Q* se compara con el valor crítico de una distribución chi-cuadrado (Friedman, 1937).

En la prueba de Friedman, la hipótesis nula indica que los efectos atribuibles a los tratamientos son los mismos; es decir, las poblaciones de interés tienen distribuciones idénticas. Además, las medias de los rangos de las columnas (tratamientos) son casi iguales.

Un procedimiento para la prueba de hipótesis puede incluir los siguientes pasos:

1) Planteamiento del sistema de hipótesis

H_o: $f_1(x) = f_2(x) = ... = f_k(x)$
H_1: las $f_i(x)$ no son todas idénticas

En este contexto, las $f_i(x)$ son las funciones de densidad de probabilidad de las poblaciones de donde provienen las k muestras o tratamientos.

O de forma equivalente:

H_o: $\mu_1 = \mu_2 = \ldots = \mu_k$
H_1: algunas de las μ_i son diferentes

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$.

3) La prueba es unilateral derecha.

4) La estadística de prueba corresponde a la estadística Q que se ha de comparar con una chi-cuadrado con $k-1$ grados de libertad $\chi^2_{1-\alpha,\, k-1}$

5) Para tomar una decisión sobre la hipótesis nula se aplica el siguiente criterio: si Q resulta mayor o igual que el valor teórico $\chi^2_{1-\alpha,\, k-1}$ entonces se rechaza la hipótesis nula; en caso contrario, se aceptará.

6) Conclusión

La siguiente situación problema se formuló a partir de Canavos (1988): 4 jueces expertos han sido encargados de calificar el desempeño de 5 finalistas en una competencia de aptitud verbal, las calificaciones van de 2.0 hasta 10.0 como se indican en la Tabla 5.2.

Tabla 5.2. Arreglo matricial para aplicar la prueba de Friedman con $k=4$ y $n=5$

		Tratamientos			
		1	2	3	4
Bloques	1	8.6	8.7	8.3	8.5
	2	9.9	9.8	9.5	9.7
	3	8.0	8.2	7.6	8.3
	4	9.8	9.9	9.7	9.7
	5	9.0	9.3	8.2	8.8

Fuente: los autores, con base en Canavos (1988)

El investigador quiere establecer si existen diferencias discernibles en las calificaciones que otorgan los jueces expertos. Usar un nivel de significancia del 5 % para probar la hipótesis del investigador.

Para solucionar esta situación problema se puede recurrir al uso de la prueba Q de Friedman; por consiguiente, en primer lugar, se asignan los rangos desde 1 hasta $k=4$ para cada uno de los 5 bloques (filas) $n=5$; el rango se ha colocado entre paréntesis en el lado derecho en cada una de las observaciones, como se indica en la Tabla 5.3.

Así, la suma de los rangos del tratamiento 1 consiste en sumar los rangos de los datos en la columna 1, de la siguiente forma:

$R_1 = 3+4+2+3+3 = 15$

La suma de los rangos del tratamiento 2 consiste en sumar los rangos de los datos en la columna 2, así:

$R_2 = 4+3+3+4+4 = 18$

Al proceder de similar forma, resulta que:

$R_3 = 5.5$

$R_4 = 11.5$

Tabla 5.3. Arreglo matricial para aplicar la prueba de Friedman con rangos *k=4* y *n=5*

		Tratamientos			
		1	2	3	4
Bloques	1	8.6(3)	8.7(4)	8.3(1)	8.5(2)
	2	9.9(4)	9.8(3)	9.5(1)	9.7(2)
	3	8.0(2)	8.2(3)	7.6(1)	8.3(4)
	4	9.8(3)	9.9(4)	9.7(1.5)	9.7(1.5)
	5	9.0(3)	9.3(4)	8.2(1)	8.8(2)

Fuente: los autores, con base en Canavos (1988)

Ahora, se procede a efectuar el cálculo de la estadística Q, del siguiente modo:

$$Q = \frac{12}{nk(k+1)}\left(\sum_{j=1}^{k=4} R_j^2\right) - 3n(k+1)$$

$$Q = \frac{12}{nk(k+1)}\left(R_1^2 + R_2^2 + R_3^2 + R_4^2\right) - 3n(k+1)$$

$$Q = \frac{12}{(5)(4)(4+1)}\left(15^2 + 18^2 + 5.5^2 + 11.5^2\right) - 3(5)(4+1)$$

$$Q = \frac{12}{(5)(4)(5)}\left(225 + 324 + 30.25 + 132.25\right) - 3(5)(5)$$

$$Q = \frac{12}{100}\left(711.5\right) - 75 = 85.38 - 75 = 10.38$$

El procedimiento para la prueba de hipótesis es:

1) Planteamiento del sistema de hipótesis

H_o: $f_1(x) = f_2(x) = f_3(x) = f_4(x)$
H_1: las $f_i(x)$ no son todas idénticas

En este contexto, las $f_i(x)$ son las funciones de densidad de probabilidad de las poblaciones de donde provienen las 4 muestras o tratamientos.

O de forma equivalente:

H_o: $\mu_1 = \mu_2 = \mu_3 = \mu_4$
H_1: algunas de μ_i son diferentes para i = 1,2,3,4

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$.
3) La prueba es unilateral derecha.

4) La estadística de prueba es $Q = 10.38$

Ahora, el valor teórico de la distribución chi-cuadrado $\chi^2_{1-\alpha,\, k-1}$ con $k-1 = 4-1 = 3$ grados de libertad para $\alpha = 0.05$ es,

$$\chi^2_{1-\alpha,\, k-1} = \chi^2_{0.95,\, 3} = 7.8147$$

5) Como $Q = 10.38$ es mayor que $\chi^2_{1-\alpha,\, k-1} = 7.8147$ entonces se rechaza la hipótesis nula; por consiguiente, se acepta la hipótesis alternativa así,

H_1: algunas de μ_i son diferentes para $i = 1,2,3,4$; lo cual indica que los efectos de los tratamientos no son los mismos.

6) Con un nivel de significancia del 5 %, se concluye que existen diferencias significativas entre las calificaciones que otorgan los jueces expertos en correspondencia con el desempeño de los finalistas en la competencia de aptitud verbal. Por lo tanto, sí existen diferencias discernibles en las calificaciones que otorgan los jueces expertos.

La siguiente situación problema hipotética se soporta en Grosslight y Radlow (1956) y en la teoría del reforzamiento, para realizar actividades de entrenamiento con los operarios de la empresa TT. Se requiere indagar sobre el efecto de 3 patrones distintos de reforzamiento en el aprendizaje discriminatorio de un conjunto específico de rutinas que van a implementar los operarios; 30 operarios fueron asignados completamente al azar a cada uno de los 3 tratamientos (patrones de reforzamiento) conformando 3 muestras aleatorias de tamaño 10 (bloques de igual tamaño); en el primer tratamiento se trabaja con total reforzamiento (100 % de los ensayos tenían reforzamiento), el segundo con reforzamiento parcial (la mitad de los ensayos incluyeron reforzamiento) y el tercero con un solo ensayo de reforzamiento.

Tabla 5.3. Arreglo matricial para aplicar la prueba de Friedman con *k=3* y *n=10*

		Tratamientos		
		1	2	3
Bloques	1	4.0	4.5	4.3
	2	3.9	4.4	3.5
	3	4.5	4.2	4.0
	4	4.1	4.5	4.2
	5	4.3	4.4	4.5
	6	4.0	3.5	4.4
	7	4.8	4.5	4.6
	8	4.1	4.6	4.2
	9	4.7	4.5	4.2
	10	3.8	3.5	4.0

Fuente: los autores, con base en Grosslight y Radlow (1956)

Una vez terminado el adiestramiento, el grado de aprendizaje se midió a través de la rapidez con la cual los distintos operarios lograron aprender las rutinas; para ello se asignó una valoración desde 1.0 hasta 5.0. Los datos se pueden observar en la Tabla 5.3.

El investigador afirma que la utilización de los diversos patrones de reforzamiento genera un aprendizaje diferencial, el cual se exhibe por la capacidad de transferencia que muestra el operario. Utilizar un nivel de significancia del 5 % para probar la hipótesis del investigador.

Para solucionar la situación problema se recurre al empleo de la prueba Q de Friedman; por lo tanto, al principio se asignan los rangos desde 1 hasta $k=3$ para cada uno de los 10 bloques (filas) n = 10; el rango se ha colocado entre paréntesis en el lado derecho en cada una de las observaciones, como se indica en la Tabla 5.4.

Tabla 5.4. Arreglo matricial para aplicar la prueba de Friedman con rangos $k=3$ y $n=10$

		Tratamientos		
		1	2	3
Bloques	1	4.0 (1)	4.5 (3)	4.3 (2)
	2	3.9 (2)	4.4 (3)	3.5 (1)
	3	4.5 (3)	4.2 (2)	4.0 (1)
	4	4.1 (1)	4.5 (3)	4.2 (2)
	5	4.3 (1)	4.4 (2)	4.5 (3)
	6	4.0 (2)	3.5 (1)	4.4 (3)
	7	4.8 (3)	4.5 (1)	4.6 (2)
	8	4.1 (1)	4.6 (3)	4.2 (2)
	9	4.7 (3)	4.5 (2)	4.2 (1)
	10	3.8 (2)	3.5 (1)	4.0 (3)

Fuente: los autores, con base en Grosslight y Radlow (1956)

Tomando en cuenta los rangos de la Tabla 5.4, se establece que la suma de los rangos del tratamiento 1 consiste en sumar los rangos de los datos en la columna 1. Se obtiene:

$R_1 = 1+2+3+1+1+2+3+1+3+2 = 19$

Al proceder de manera similar y sumar los rangos de la columna 2, resulta que:

$R_2 = 21$

$R_3 = 20$

Ahora, se procede a efectuar el cálculo de la estadística Q, de la siguiente manera:

$$Q = \frac{12}{nk(k+1)}\left(\sum_{j=1}^{k=3} R_j^2\right) - 3n(k+1)$$

$$Q = \frac{12}{nk(k+1)}\left(R_1^2 + R_2^2 + R_3^2\right) - 3n(k+1)$$

$$Q = \frac{12}{(10)(3)(3+1)}\left(19^2 + 21^2 + 20^2\right) - 3(10)(3+1)$$

$$Q = \frac{12}{(10)(3)(4)}\left(361 + 441 + 400\right) - 3(10)(4)$$

$$Q = \frac{12}{120}\left(1202\right) - 120 = 0.1(1202) - 120 = 120.2 - 120 = 0.2$$

El procedimiento para la prueba de hipótesis es:

1) Planteamiento del sistema de hipótesis

H_o: los distintos patrones de reforzamiento no tienen efecto diferencial
H_1: los distintos patrones de reforzamiento sí tienen efecto diferencial

H_o: $f_1(x) = f_2(x) = f_3(x)$
H_1: las $f_i(x)$ no son todas idénticas

En estas circunstancias, las $f_i(x)$ son las funciones de densidad de probabilidad de las poblaciones de donde provienen los 3 tratamientos.

También de manera equivalente, se puede escribir:

H_o: $\mu_1 = \mu_2 = \mu_3$
H_1: algunas de μ_i son diferentes para i = 1,2,3

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$.

3) La prueba es unilateral derecha.

4) La estadística de prueba es $Q = 0.2$

Por lo tanto, el valor teórico de la distribución chi-cuadrado $\chi^2_{1-\alpha,\, k-1}$ con $k-1 = 3-1 = 2$ grados de libertad para $\alpha = 0.05$ es,

$$\chi^2_{1-\alpha,\, k-1} = \chi^2_{0.95,\, 2} = 5.9915$$

5) Como $Q = 0.2$ no es mayor que $\chi^2_{1-\alpha,\, k-1} = 5.9915$ entonces no se rechaza la hipótesis nula; por consiguiente, se acepta la hipótesis nula, así,

H_o: los distintos patrones de reforzamiento no tienen efecto diferencial

6) Con un nivel de significancia del 5 %, se concluye que los distintos patrones de reforzamiento no tienen efecto diferencial; esto es, no existen diferencias significativas en el efecto generado por los patrones de reforzamiento; las valoraciones obtenidas por los operarios sobre la transferencia del aprendizaje no dependen de los patrones de reforzamiento. Luego, la afirmación hecha por el investigador no es cierta estadísticamente.

5.3 Coeficiente de correlación de Spearman

Este coeficiente corresponde a una medida de asociación para los datos de dos variables cuantitativas en escala ordinal, en las cuales sus respectivos datos pueden ordenarse y a estos se les puede asignar un determinado rango. Por lo tanto, el coeficiente de correlación de Spearman es una medida de asociación perteneciente a los métodos estadísticos no paramétricos. Para calcular este coeficiente, en primer lugar se asignan los rangos desde 1 hasta n para los datos x_i de la variable X, luego se asignan los rangos desde 1 hasta n para los datos y_i de la variable cuantitativa Y, los cuales han conformado parejas de la firma (x_i , y_i) para $i = 1,2,\ldots,n$. Se prosigue a calcular las diferencias d_i entre el rango de cada x_i y su correspondiente rango para y_i. Finalmente, se aplica la siguiente expresión, la cual permite obtener el mencionado coeficiente, el cual se denota y se define de la siguiente manera:

$$r_s = 1 - \frac{6\left(\sum_{i=1}^{n} d_i^2\right)}{n(n^2-1)}$$

Si se presentan empates en los rangos para algunas observaciones (datos), se asigna el promedio como rango para las observaciones empatadas.

Cuando la cantidad n de parejas (x_i, y_i) para $i = 1,2,\ldots, n$ es lo suficientemente grande, la variable aleatoria R_s cuyos valores son los r_s. Según la hipótesis nula de que no hay correlación entre X e Y, se puede aproximar por medio de una distribución normal estándar, así (Gibbons, 1997):

$$Z = \frac{R_s - E(R_s)}{\sqrt{Var(R_s)}}$$

Donde:

$$E(R_s) = 0,$$

$$Var(R_s) = \frac{1}{n-1}$$

En el proceso de inferencia estadística, para un nivel de significancia α, se plantean las siguientes hipótesis:

H_o: no hay correlación entre *X* e *Y*
H_1: sí hay correlación entre *X* e *Y*

En esta situación se procede como una prueba bilateral.

Cuando las hipótesis se plantean de la forma:

H_o: la correlación entre *X* e *Y* en la población es cero
H_1: hay una asociación positiva entre los rangos

Se trata de una prueba unilateral derecha; de acuerdo con Lind *et al*. (2015), cuando la muestra es de tamaño n mayor o igual que 10 para analizar la significancia de la prueba, se utiliza una distribución teórica t-student con *n-2* grados de libertad para determinar la región crítica y se obtiene el valor t siguiente a partir de la distribución de los rangos:

$$t = r_s \sqrt{\frac{n-2}{1-r_s^{\,2}}}$$

Se aplica la siguiente regla de decisión: si t es mayor que $t_{1-\alpha,\, n-2}$ entonces se rechaza la hipótesis nula.

En contraste, cuando las hipótesis se plantean así:
H_o: la correlación entre *X* e *Y* en la población es cero
H_1: hay una asociación negativa entre los rangos

Se aplica la siguiente regla de decisión: si t es menor que $t_{\alpha,\, n-2}$ entonces se rechaza la hipótesis nula.

La siguiente situación problema es un caso hipotético de investigación. Se requiere indagar sobre la asociación entre el número de horas que usa un trabajador y la cantidad de artículos a los que les ha hecho un proceso de terminado (pintura y empacado) en la empresa TK. Para esto se ha tomado una muestra aleatoria de 11 trabajadores a quienes se les han medido las variables X: número de horas laboradas por el trabajador, Y: número de artículos a los que el trabajador le ha hecho el proceso de terminado. Los datos se pueden observar en la Tabla 5.5.

Un investigador del área de procesos en la empresa TK afirma que sí existe una asociación significativa entre las variables *X* e *Y*. Utilizar un nivel de significancia del 5 % para probar la hipótesis del investigador.

Tabla 5.5. Información de horas laboradas y cantidad de artículos terminados en la empresa TK

Trabajador	*X*	*Y*
1	21	100
2	9	57
3	6	45
4	12	80
5	14	73
6	11	71
7	6	55
8	19	95
9	16	86
10	3	34
11	9	66

Fuente: los autores

Para solucionar esta situación problema se utiliza el coeficiente de correlación de Spearman, ya que los datos están en una escala ordinal (los datos son factibles de ser ordenados). Inicialmente, se asignan los rangos desde 1 hasta n=11 para los datos tanto de la variable X como de la variable Y, poniendo atención si se presentan empates. El rango se ha colocado entre paréntesis en el lado derecho en cada una de las observaciones, como se indica en la Tabla 5.6. Además, se ha determinado la diferencia de rangos y el cuadrado de tal diferencia.

Tabla 5.6. Rangos para los datos de las variables X e Y

Trabaj.	X	Y	d_i	d_i^2
1	21 (11)	100 (11)	0	0
2	9 (4.5)	57 (4)	0.5	0.25
3	6 (2.5)	45 (2)	0.5	0.25
4	12 (7)	80 (8)	-1	1
5	14 (8)	73 (7)	1	1
6	11 (6)	71 (6)	0	0
7	6 (2.5)	55 (3)	-0.5	0.25
8	19 (10)	95 (10)	0	0
9	16 (9)	86 (9)	0	0
10	3 (1)	34 (1)	0	0
11	9 (4.5)	66 (5)	-0.5	0.25

Fuente: los autores

Con base en los rangos y las diferencias al cuadrado de esos rangos presentados en la Tabla 5.6, se determina que la suma de las diferencias al cuadrado es igual a 3.

Ahora, se procede a efectuar el cálculo del coeficiente de correlación de Spearman, así:

$$r_s = 1 - \frac{6\left(\sum_{i=1}^{n} d_i^2\right)}{n(n^2-1)} = 1 - \frac{6(3)}{11(11^2-1)} = 1 - \frac{18}{1320} = 1 - 0.0136 = 0.9864$$

Este valor indica que existe una asociación fuerte y positiva entre el número de horas laboradas por cada trabajador y el número de artículos a los que el trabajador les ha hecho el proceso de terminado en la empresa TK.

El procedimiento para la prueba de hipótesis es:

1) Planteamiento del sistema de hipótesis

H_o: no hay correlación entre X e Y en la población
H_1: sí hay correlación entre X e Y

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$.

3) La prueba es bilateral.

4) La estadística de prueba es una normal estándar donde para el valor específico $r_s = 0.9864$ de la variable aleatoria R_s resulta:

$$E(R_s) = 0,$$
$$Var(R_s) = \frac{1}{n-1} = \frac{1}{10} = 0.1$$

$$Z = \frac{R_s - E(R_s)}{\sqrt{Var(R_s)}} = \frac{0.9864 - 0}{\sqrt{0.1}} = 0.9864\sqrt{10} = 3.1192$$

Ahora, el valor teórico de la distribución normal estándar con $\alpha = 0.05$, $\alpha/2 = 0.025$ es el siguiente,

$Z_{\alpha/2} = Z_{0.025} = -1.96$; $Z_{1-\alpha/2} = Z_{0.975} = 1.96$;

5) Como Z = 3.1192 es mayor que $Z_{1-\alpha/2} = Z_{0.975} = 1.96$, entonces se rechaza la hipótesis nula; por consiguiente, se acepta la hipótesis alterna, así,

H_1: sí hay correlación entre *X* e *Y* en la población

6) Con un nivel de significancia del 5 %, se concluye que existe una asociación significativa entre el número de horas laboradas por cada trabajador y el número de artículos a los que el trabajador les ha hecho el proceso de terminado en la empresa TK. Además, se ha evidenciado que dicha relación es fuerte y positiva.

La siguiente situación problema se basa en lo expuesto por Lind *et al*. (2015). La empresa de ingeniería Ingenius admite que estudiantes de ingeniería de distintas universidades hagan su pasantía (práctica de grado) en esa empresa, siempre y cuando los candidatos a pasantes superen dos tipos de pruebas con la tutela de un instructor. El instructor otorga una calificación a cada candidato durante una entrevista en el campus universitario, la cual está puntuada desde 1 hasta 16. La segunda prueba se refiere a la habilidad que debe demostrar el candidato en el manejo de herramientas computacionales; en esta, el instructor otorga una puntuación desde 0 hasta 12. Para esto se ha tomado una muestra aleatoria de 13 candidatos. Los datos se pueden observar en la Tabla 5.7.

X: puntuaciones obtenidas en la prueba 1, *Y*: puntuaciones obtenidas en la prueba 2.

El gerente de la empresa Ingenius afirma que sí existe una asociación positiva entre las variables *X* e *Y*. Usar un nivel de significancia del 5 %

para probar la hipótesis del investigador.

Tabla 5.7. Información de las puntuaciones obtenidas en las dos pruebas

Candidato	*X*	*Y*
1	9	5
2	11	5
3	10	5
4	5	4
5	13	7
6	12	10
7	12	10
8	8	7
9	9	7
10	14	10
11	11	6
12	13	10
13	15	12

Fuente: los autores

Para hallar una solución a esta situación problema, nuevamente se emplea el coeficiente de correlación de Spearman, ya que los datos están en una escala ordinal. Al comienzo se asignan los rangos desde 1 hasta n=13 para los datos tanto de la variable X como de la variable Y, poniendo atención si se presentan empates. El rango se ha colocado entre paréntesis en el lado derecho de cada uno de los datos de cada variable, como se indica en la Tabla 5.8. Además, se ha determinado la diferencia de rangos y el cuadrado de esa diferencia.

Con base en los rangos y las diferencias al cuadrado de esos rangos presentados en la Tabla 5.8, se deduce que la suma de las diferencias al cuadrado es igual a 78.5.

Tabla 5.8. Rangos para los datos de las variables *X* e *Y*

Candidato	*X*	*Y*	d_i	d_i^2
1	15 (13)	12 (13)	0	0
2	11 (6.5)	5 (3)	3.5	12.5
3	10 (5)	5 (3)	2	4
4	5 (1)	4 (1)	0	0
5	13 (10.5)	7 (7)	3.5	12.5
6	12 (8.5)	10 (10.5)	-2	4
7	12 (8.5)	10 (10.5)	-2	4
8	8 (2)	7 (7)	-5	25
9	9 (3.5)	7 (7)	-3.5	12.5
10	14 (12)	10 (10.5)	1.5	2.25
11	11 (6.5)	6 (5)	1.5	2.25
12	13 (10.5)	10 (10.5)	0	0
13	9 (3.5)	5 (3)	0.5	0.25

Fuente: los autores

Ahora, se procede a efectuar el cálculo del coeficiente de correlación de Spearman, así:

$$r_s = 1 - \frac{6\left(\sum_{i=1}^{n} d_i^2\right)}{n(n^2-1)} = 1 - \frac{6(78.5)}{13(13^2-1)} = 1 - \frac{471}{2184} = 1 - 0.2156 = 0.7844$$

Este valor del coeficiente de correlación de Spearman indica que existe una asociación fuerte y positiva entre las variables *X*: puntuaciones obtenidas en la prueba 1 e *Y*: puntuaciones obtenidas en la prueba 2.

El procedimiento para la prueba de hipótesis es:
1) Planteamiento del sistema de hipótesis

H_o: la correlación entre *X* e *Y* en la población es cero
H_1: hay una asociación positiva entre los rangos

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$.

3) La prueba es unilateral derecha.

4) La estadística de prueba es una t-student donde para el valor específico $r_s = 0.7844$ de la variable aleatoria R_s resulta:

$$t = r_s\sqrt{\frac{n-2}{1-r_s^2}} = 0.7844\sqrt{\frac{13-2}{1-0.7844^2}} = 0.7844\sqrt{\frac{11}{1-0.6153}}$$

$$t = 0.7844\sqrt{\frac{11}{0.3847}} = 0.7844(5.3473) = 4.1944$$

Ahora, el valor teórico de la distribución t-student con $\alpha = 0.05$, y n-2 = 13 -2 =11 grados de libertad es,

$t_{1-\alpha\,,\, n-2} = t_{0.95,\,11} = 1.796$;

5) Como t = 4.1944 es mayor que $t_{1-\alpha,\, n-2} = Z_{0.95\,,\,11} = 1.796$ entonces se rechaza la hipótesis nula; por consiguiente, se acepta la hipótesis alterna, así,

H_1: hay una asociación positiva entre los rangos.

6) Con un nivel de significancia del 5 %, se concluye que existe una asociación fuerte y positiva entre las puntuaciones obtenidas en la prueba 1 y las puntuaciones obtenidas en la prueba 2 por los candidatos a realizar la pasantía en la empresa Ingenius.

5.4 Prueba de independencia para tablas de contingencia

Una tabla de contingencia es un arreglo matricial para clasificar a los individuos de acuerdo con dos variables cualitativas (nominales), cada una de las cuales presenta dos o más categorías (modalidades). En este tipo de tabla, cada fila corresponde a una categoría de la primera variable y cada columna a una modalidad de la segunda variable. Cuando se tiene una tabla con m filas y n columnas, generalmente interesa indagar si las filas y las columnas son independientes (Peña y Romo, 1997), lo cual también es un indicativo de que las variables son independientes. En este contexto, la hipótesis nula se formula en esos términos; en cambio, la hipótesis alternativa indicará que no hay independencia, lo cual significará que las variables están asociadas o relacionadas.

La estadística de prueba corresponde a la expresión chi-cuadrado siguiente, la cual permite determinar si no existen discrepancias entre las frecuencias observadas y las frecuencias esperadas, en cuyo caso ha de aceptarse la hipótesis nula; de lo contrario, si tales discrepancias son muy grandes, se rechaza esta hipótesis.

La estadística de prueba puede expresarse de la siguiente manera:

$$Chi - cudrado = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left(O_{ij} - e_{ij}\right)^2}{e_{ij}}$$

Donde O_{ij} corresponde al número observado de individuos clasificados en la categoría *i* de la primera variable y en la modalidad *j* de la segunda variable, e_{ij} hace referencia al número esperado de individuos en la categoría *i* en correspondencia con la modalidad *j*. En este contexto, e_{ij} se obtiene al multiplicar la marginal *i* por la marginal *j* y dividir entre el total de individuos. Si se produce un número alto de concordancias entre las frecuencias observadas y esperadas, es posible que dicha suma de cuadrados sea pequeña, en cuyo caso el valor de chi-cuadrado resultará pequeño y por lo tanto se ha de aceptar la hipótesis nula; en caso contrario, se ha de rechazar esta hipótesis. Como la estadística (estimador) chi-cuadrado con la hipótesis nula sigue una distribución teórica chi-cuadrado con (m-1)(n-1) grados de libertad, si P-valor es menor que el nivel de significancia α, entonces se rechaza la hipótesis nula; de lo contrario, será aceptada.

La siguiente situación problema se soporta en lo expuesto en Peña y Romo (1997): un investigador quiere indagar si el número de reclamaciones recibidas en la oficina de atención al cliente de la empresa NPN clasificadas de acuerdo con el tipo de producto en las categorías E, F y G y la edad del cliente quien hace el reclamo son independientes. Por edad, los clientes son susceptibles de ser clasificados en el grupo de 25 a 35 años o en el de clientes mayores que 35 años. Usar un nivel de significancia del 5 % para la prueba de hipótesis. En la Tabla 5.9 se presentan los datos referidos a una muestra de tamaño 150 clientes

Tabla 5.9. Tabla de contingencia para tipo de producto y edad del cliente

Producto	25-35 años	Más de 35	Marginal
E	10	30	40
F	15	45	60
G	18	32	50
Marginal	43	107	150

Fuente: los autores

En la Tabla 5.10 se presentan las frecuencias esperadas correspondientes a cada una de las celdas de la Tabla 5.9. Por ejemplo, la frecuencia esperada e_{11} = 40(43)/150 = 11.47, e_{12} = 40(107)/150 = 28.53; e_{32} = 50(107)/150 = 35.67; de manera similar se obtienen las demás frecuencias esperadas.

El proceso de prueba de hipótesis para esta situación puede seguir los siguientes pasos:

Tabla 5.10. Tabla de contingencia con frecuencias esperadas edad del cliente

Producto	25-35 años	Más de 35	Marginal
E	11.47	28.53	40
F	17.2	42.8	60
G	14.33	35.67	50
Marginal	43	107	n = 150

Fuente: los autores

1) Planteamiento del sistema de hipótesis

H_o: las variables son independientes
H_1: las variables no son independientes

De forma equivalente, las hipótesis anteriores se pueden plantear de la siguiente forma:

H_o: no existe asociación entre las variables
H_1: sí existe asociación entre las variables

2) Se fija el nivel de significancia en el valor $\alpha = 0.05$. $1 - \alpha = 0.95$

3) Se trata de una prueba unilateral derecha sobre una distribución teórica chi-cuadrado con (m-1)(n-1) = (3-1)(2-1) =3 * 1 = 3 grados de libertad

$$\chi^2_{0.95,3} = 7.8147$$

4) La estadística de prueba se soporta en los datos de la muestra y las respectivas tablas de contingencia, así:

$$Chi-cudrado = \sum_{i=1}^{m}\sum_{j=1}^{n}\frac{\left(O_{ij} - e_{ij}\right)^2}{e_{ij}}$$

$$Chi - cudrado = \frac{(10 - 11.47)^2}{11.47} + \frac{(30 - 28.53)^2}{28.53} + \frac{(15 - 17.2)^2}{17.2} + \frac{(45 - 42.8)^2}{42.8} + \frac{(18 - 14.33)^2}{14.33} + \frac{(32 - 35.67)^2}{35.67}$$

$$Chi - cudrado = 0.1884 + 0.0757 + 0.2814 + 0.113 + 0.9399 + 0.3775 = 1.9759$$

5) Decisión: si el P-valor es menor que $\alpha = 0.05$ entonces se rechaza la hipótesis nula; en caso contrario, dicha hipótesis será aceptada. De forma equivalente, si chi-cuadrado es mayor que $\chi^2_{0.95,3}$ entonces se rechaza la hipótesis nula.

Como chi-cuadrado = 1.9759 resultó menor que $\chi^2_{0.95,3} = 7.8147$ entonces se decide aceptar la hipótesis nula.

6) Con un nivel de significancia del 5 %, se concluye que el número de reclamaciones recibidas en la oficina de atención al cliente de la empresa NPN, clasificadas de acuerdo con el tipo de producto en las categorías E, F y G y la edad del cliente quien hace el reclamo, son independientes. Estos resultados indican que entre el número de reclamaciones y la edad del cliente no existe una asociación estadísticamente significativa.

6

Conclusiones

En este capítulo se dan a conocer los hallazgos principales del análisis de los resultados obtenidos en el trabajo de investigación. Dichos hallazgos están explicitados en las categorías y subcategorías provenientes del análisis documental y de contenido sobre los datos tanto textuales como cuantitativos, y de aplicar procesos de ID a un gran número de fuentes relativas al tema de la estadística no paramétrica en los libros de texto universitarios. Además, los procesos de indagación posibilitaron la formulación de situaciones problema para potenciar la investigación científica en el aula. También se exponen las conclusiones y se indican algunas recomendaciones.

6.1 Hallazgos y discusión de resultados

Los procesos de inferencia estadística pueden ser desarrollados mediante la aplicación de métodos cuantitativos paramétricos o de modelos estadísticos no paramétricos. Los primeros se utilizan cuando se han cumplido determinados supuestos sobre la forma de la distribución de probabilidad de la población de la cual se ha obtenido la muestra aleatoria que sirve de base para estimar uno o más parámetros desconocidos de la población, construir intervalos de confianza para tales parámetros o generar regiones críticas para rechazar la hipótesis nula. En diversos casos, se debe cumplir el supuesto de normalidad, entre otros supuestos.

Los modelos estadísticos no paramétricos son técnicas estadísticas que requieren supuestos más flexibles (Corzo, 2005) o que utilizan estadísticas de prueba libres de distribución (Delicado, 2008), las cuales generalmente se definen a través de los rangos relativos a los datos de la variable o variables por investigar, ya sea que se trate de un problema de inferencia relativo a una muestra aleatoria, dos, tres o más muestras. Este trabajo investigativo ha generado resultados centrados en la aplicación de procesos relacionados con la prueba de hipótesis mediante métodos no paramétricos clásicos (Ríos y Peña, 2020), en situaciones concretas susceptibles de promover la investigación científica desde los claustros universitarios.

Con base en el análisis documental y de contenido se ha determinado que, con frecuencia, los libros de texto universitarios dedican varios de sus capítulos a abordar tópicos acerca del análisis exploratorio de datos, la inferencia estadística paramétrica, y una gran variedad de modelos estadísticos soportados en el uso de variables aleatorias, entre otros, pero solamente destinan un capítulo a presentar algunos métodos no paramétricos acompañados de ejemplos de tipo académico con poco énfasis en la investigación formativa y sin el suficiente soporte desde la teoría estadística; por consiguiente, sería deseable incluir elementos teóricos semejantes a los expuestos en Hettmansperger (1984), Hájek *et al*. (1999), Corzo (2005), Wasserman (2006) o Delicado (2008).

Aunque los modelos estadísticos no paramétricos son herramientas estadísticas robustas con un gran potencial para desarrollar procesos de investigación científica, entre otros, en las ciencias naturales, en las ciencias sociales y de la conducta, en los libros de textos del nivel universitario no se les ha otorgado el sitio que se merecen, a pesar de que tales modelos son versátiles y pueden utilizarse en diversas carreras universitarias como la psicología, la sociología, la administración de empresas, la economía, la biología, la química, las ingenierías, la estadística misma y la simulación, entre otras (Vera et al., 2021). Además, en algunos textos universitarios se ha evidenciado que dichos modelos posibilitan efectuar una moderada transposición didáctica de las pruebas de hipótesis en la forma como lo expone Chevallard (1991).

En los libros de texto universitarios que incluyen un solo capítulo referido al análisis de algunos métodos no paramétricos se ha demostrado que utilizan distintos enfoques didácticos que van desde el tratamiento algorítmico de las técnicas estadística con la aplicación directa de fórmulas para testear hipótesis (Canavos, 1988; Freund y Miller, 2000) hasta aquellos que, con la intervención apropiada del profesor universitario, pueden promover el trabajo investigativo en el estudiantado (Siegel, 1970; Lehmann y D'Abrera, 1975).

En este trabajo se han aportado diversas situaciones problema para promover la investigación científica desde el aula, las cuales se soportan en un conjunto estructurado de pasos para comprobar las hipótesis de manera semejante a como se exponen en Burbano y Valdivieso (2016). Este tipo de procesos pueden ser complementados por el profesor con la ayuda del modelo expuesto por Wild y Pfannkuch (1999).

La revisión de literatura ha permitido establecer que la escogencia del modelo estadístico no paramétrico está relacionada con el tipo de escala con la cual se han medido los datos de la variable de interés (Vera at al., 2021). Además, se ha determinado que cuando los datos pertenecen a una variable en escala nominal o categórica no existen pruebas paramétricas para testear las hipótesis; en ese caso ha de recurrirse a una prueba no paramétrica basada en la distribución binomial. Varias de las pruebas no paramétricas se fundamentan en estadísticas definidas a través de la teoría de rangos, de modo que los datos han de medirse al menos en una escala ordinal o corresponder a puntajes (Randles y Wolfe, 1979).

Los resultados también han demostrado un bajo porcentaje de contenidos referidos a las pruebas no paramétricas en los libros de texto; para mitigar este tipo de deficiencia, en esta obra de investigación se han expuesto diversas situaciones problema que, a criterio de los autores, pueden promover el acrecentamiento de procesos investigativos en el estudiantado universitario, los profesores y los profesionales desde una mirada cuantitativa. Tales situaciones están relacionadas con la prueba basada en la distribución binomial, la chi-cuadrado y la prueba de Kolmogorov-Smirnov para una muestra; las pruebas de MacNemar, del signo, la del rango signado de Wilcoxon y la Mann-Whitney para dos muestras; las pruebas de Kruskal-Wallis, Friedman para k muestras; esto no excluye el uso de pruebas no paramétricas adicionales reportadas en textos especializados sobre el tema.

Los hallazgos también permitieron determinar que existen libros especializados que tratan diversos tópicos acerca de los modelos estadísticos no paramétricos, tanto desde un punto de vista clásico como contemporáneo, los cuales manejan diferentes niveles de profundización y tratamiento formal. Entre tales obras se pueden mencionar: el libro de Lehmann y D’Abrera (1975), Randles y Wolfe (1979), Hettmansperger

(1984), Manoukian (1986), Hájek *et al*. (1999), Lehmann y Casella (2001), Corzo (2005), Wasserman (2006) y Marshal y Olkim (2007).

Para terminar es conveniente mencionar que para el diseño e implementación de obras como el presente libro de investigación es pertinente asumir un marco teórico que sirva de referencia y posibilite elaborar una metodología que trace un camino para seguir durante el proceso de indagación; además, que permita realizar un tratamiento sistemático de la información obtenida de diversas fuentes documentales, tales como libros de texto universitarios, artículos, libros especializados referidos a los modelos estadísticos no paramétricos, ponencias, trabajos de pregrado, maestría o doctorado. En este campo, según Burbano *et al*. (2021), cuando el investigador y sus colaboradores tienen acceso a un conjunto amplio de fuentes (primarias y secundarias), es conveniente usar técnicas de investigación documental, de análisis textual y de tratamiento de contenido (Bardín, 1986) con la finalidad de enriquecer paulatinamente la obra que se está elaborando. Para ello, también es pertinente acudir al análisis interpretativo a través de la hermenéutica (Guevara, 2015; Burbano, 2017), con el propósito de construir un manuscrito con altos niveles de originalidad en concomitancia con la ID (Uribe, 2005) y con metodologías que propicien la promoción de la investigación científica y formativa desde el aula.

6.2 Principales conclusiones

Con fundamento en el marco conceptual y metodológico se concluye que los modelos estadísticos no paramétricos se han de utilizar solamente cuando en la prueba de hipótesis resulte improcedente la aplicación de los métodos paramétricos de inferencia estadística; esto implica que es conveniente comenzar por verificar los supuestos sobre la forma de la distribución de probabilidad que subyace tras la población de donde se ha seleccionado la muestra, y si hay incumplimiento de tales supuestos, entonces se emplean tales modelos.

Los hallazgos demuestran que los libros de texto universitarios de estadística presentan bajos porcentajes de contenidos referidos al uso de los métodos no paramétricos y se focalizan principalmente en tópicos de estadística descriptiva, probabilidad, distribuciones de probabilidad, inferencia estadística paramétrica, modelos de regresión y elementos de diseño experimental, entre otros; de allí que sea necesario consultar libros especializados en estadística no paramétrica a fin de abordar con suficiente profundidad los temas sobre este tipo de estadística, para aplicarlos en procesos de investigación científica a través de la prueba de hipótesis.

Los planes curriculares de diferentes carreras (titulaciones) incluyen uno o dos cursos de estadística, dentro de los cuales se da poca prelación al tratamiento y uso de los métodos no paramétricos. Con el presente trabajo se contribuye a la dinamización de la investigación científica con enfoque experimental y se promueven los procesos de investigación desde el aula, porque al aplicar los métodos no paramétricos también se fortalece el uso del método científico que comprende la observación (recogida de datos), la formulación de hipótesis, la prueba de estas (mediante estadísticas basadas en rangos) y la obtención de conclusiones.

En este libro se va más allá del uso cotidiano de la estadística descriptiva y la inferencia estadística paramétrica, pues se hace un aporte epistemológico, didáctico y metodológico, que fortalece el proceso E–A de ciertos tópicos de la prueba de hipótesis mediante las situaciones problema planteadas y resueltas que incluyen el uso de métodos no paramétricos clásicos. Además, la utilización de técnicas relativas al análisis de contenido y de investigación documental ha posibilitado crear espacios de reflexión académica que confirman la necesidad de promover la elaboración de textos universitarios que examinen más ampliamente los aspectos fundamentales y de aplicación de los modelos no paramétricos.

En estas circunstancias, la indagación realizada también genera elementos de juicio para recomendar a los profesores universitarios que utilizan la estadística, que elaboren un análisis crítico desde su actividad académica que les permita seleccionar mejor los libros de texto que emplean para orientar a sus estudiantes, planifiquen e implementen situaciones problema semejantes a las expuestas en esta obra, tendientes a promover la investigación científica desde los claustros universitarios con metodologías alternativas. Desde luego, esta obra puede presentar limitaciones de distinta índole, pero también puede constituirse en un aporte para que otros investigadores produzcan diversos avances en esta dirección.

Referencias

Alzate, M., Gómez, M. y Romero, F. (1999). *Textos escolares y representaciones sociales de familia, definiciones, dimensiones y campos de investigación*. Universidad Tecnológica de Pereira/Cargraphics, S.A.

Anderson, D., Sweeney. D. y Williams, T. (2012). *Estadística para negocios y economía*. Cengage Learning.

Ávila, H. (2015). *Introducción a la metodología de la investigación*. Eumed.net. https://www.eumed.net/libros-gratis/2006c/203/

Bardín, L. (1986). *El análisis de contenido*. Editorial Akal.

Barnett, V. (1982). *Teaching Statistics in Schools Throughout the World*. International Statistical Institute.

Blanco, L. (2004). *Probabilidad*. Editorial Unibiblos/Universidad Nacional de Colombia.

Bos, M. et al. (2014). *Brief 12: Colombia Pisa 2012. Logros y desafíos pendientes.* Scholar Harvard. Hppt://scholar.harvard.edu/files/Alejandro_ganimian/files/brief_12.pdf

Burbano, V. M. A. (2017). *Un acercamiento a la didáctica de la probabilidad desde el PCK para nivel preuniversitario*. Editorial de la UPTC.

Burbano, V. M. A. y Valdivieso, M. (2015). *Elementos de probabilidad: apoyo al estudio independiente.* Editorial de la UPTC.

Burbano, V. M. A. y Valdivieso, M. (2016). *Inferencia estadística básica: apoyo al estudio independiente*. Editorial de la UPTC.

Burbano, V. M. A. y Valdivieso, M. A. (2020). *Una mirada histórica de las medidas de probabilidad reales desde la investigación documental*. Editorial de UPTC.

Burbano, V., Valdivieso, M. y Aldana, E. (2017). Conocimiento base para la enseñanza: un marco aplicable a la didáctica de la probabilidad. *Revista de Investigación, Desarrollo e Innovación*, *7*(2), 269-285.

Burbano, V., Valdivieso, M. y Burbano, A. (2019). *Confiabilidad: un enfoque paramétrico en la práctica investigativa.* Editorial de la UPTC.

Burbano, V., Valdivieso, M. y Valdivieso, A. (2021). *El pensamiento aleatorio y los sistemas de datos en los libros de texto: situaciones alternativas de investigación formativa.* Editorial de la UPTC.

Canavos, G. (1988). *Probabilidad y estadística. Aplicaciones y métodos*. McGraw Hill.

Corzo, J. A. (2005). *Estadística no paramétrica: métodos basados en rangos.* Editorial Universidad Nacional de Colombia.

Chamorro, L. V. y Revelo, O. (2008). *Simulación, un primer contacto*. Editorial Universitaria Universidad de Nariño.

Cochran, W. G. (1954). Some Methods for Strengthening the Common chi-cuadrado test. *Biometrics*, *10*, 417-451.

Congreso de la República de Colombia. (1982). *Ley 23, sobre los derechos de autor.* https://propiedadintelectual.unal.edu.co/fileadmin/recursos/innovación/docs/normatividad_pi/ley23_1982.pdf.

Congreso de la República de Colombia. (1992). *Ley 30, por la cual se organiza el servicio público de la Educación Superior.* Imprenta Nacional.

Chevallard, Y. (1991). *La transposición didáctica. Del saber sabio al saber enseñado.* Aique.

David, F. N. (1998). *Games, Gods, and Gambling: A History of Probability and Statistical Ideas*. Courier Corporation.

Delicado, P. (2008). *Curso de modelos no paramétricos*. Departament d'Estadística i Investigació Operativa, Universitat Politécnica de Catalunya.

Devore, J. (2012). *Probabilidad y estadística para ingeniería y ciencias*. (7.ª ed.). Cengage Learning.

Díaz-Levicoy, D. (2014). *Un enfoque empírico de las gráficas estadísticas en los libros de texto de educación primaria española*. [Tesis de fin de máster]. Universidad de Granada, España. https://www.ugr.es/~batanero/pages/ARTICULOS/TFMDanilo.pdf

Freund, J. y Miller, I. (2000). *Estadística matemática con aplicaciones*. Prentice Hall

Friedman, M. (1937). The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. *J. Amer, statist, Ass, 32*, 675-701

Gamarra, C. H., Pérez, A. A. y Quiseno, R. A. (2006). *Estadística no paramétrica*. [Trabajo de pregrado]. Universidad de Sucre.

García, S. R. y Ríos, I. S. (1998). La teoría de la decisión, de Pascal a Von Newmann. *Historia de la Matemática*, 11-42. https://eudml.org/doc/42701

Garfield, J. y Ben-Zvi, D. (2008). *Developing Students' Statistical Reasoning. Connecting Research and Teaching Practice*. Springer.

Gatusso, L. y Pannone, M. (2002). Teacher's Training in a Statistic Teaching Experimentation. In B. Phillips (Ed.), *Proceeding of the Sixth International Conference on Teaching Statistics*, (pp. 685-692). Cope Town. International Association for Statistical Education e International Statistical Institute.

Gibbons, J. D., y Chakraborti, S. (1992). *Nonparametric Statistical Inference* (Third ed.). Marcewl Dekker.

Gibbons, J. D. (1997). *Nonparametric Methods for Quantitative Analysis* (Third ed.). American Sciences Press.

Grosslight, J. H. y Radlow, R. (1953). Studies in Partial Reinforcement: I. Pattering Effect of the Nonreinforcement Sequence in a Discrimination Situation. *J, Comp Physiol, Psychol*, en prensa.

Gutiérrez, H. y De la Vara, R. (2008). *Análisis y diseño de experimentos*. McGraw Hill.

Gómez, E. (2011). *Bases para la definición sistemática del conocimiento matemático para enseñar probabilidad en futuros profesores*. [Tesis de fin de máster]. Universitat de Granada.

Guevara, R. (2015). *El estado del arte en investigación. ¿Análisis de los conocimientos acumulados o indagación por nuevos sentidos?* Editorial Universidad Pedagógica Nacional de Colombia.

Gunderson, B. y Aliaga, M. (2005). *Interactive Statistics*. Prentice Hall.

Hacking, I. (1995). *El surgimiento de la probabilidad*. Gedisa.

Hernández, S. y. Mendoza, C. (2018). *Metodología de la investigación. Las rutas cualitativa, cuantitativa y mixta*. McGraw Hill Education.

Hair, A. y Taham, B. (2008). *Análisis multivariante*. (5.ª ed.). Prentice Hall.

Hájek, J., Sidak, Z. y Sen, P. (1999). *Theory of Ranks Tests*. Academic Press.

Hettmansperger, T. P. (1984). *Statistical Inference Based on Ranks*. John Willey & Sons.

Hollander, M. y Wolfe, D. A. (1999). *Nonparametric Statistical Methods* (Second ed.). Wiley & Sons.

Hurtado, M. y Silvente, V. (2012). Cómo aplicar las pruebas paramétricas bivariadas t-student y ANOVA en SPSS. Caso práctico. *Reire*, *5*(2), 83-100.

Kant, I. (2003). *Pedagogía* (vol. 85). Ediciones Akal.

Kendall, M. (1978). The Beginnings of the Probability Calculus. In Pearson and Kendall (Eds.), *Studies in the History of Statistics and Probability* (vol. I). Charles Griffin.

Kolmogorov, A. N. (1956). *Foundation of the Theory of Probability*. Chelsea.

Krueger, J. (2001). Null Hypothesis Significance Testing: On the Survival of a Flawed Method. *American Psychologist*, *56*(1), 16.

Leach, C. (1982). *Fundamentos de estadística. Enfoque no paramétrico para ciencias sociales*. Limusa.

Lehmann, E. L. y D'Abrera, H. J. (1975). *Nonparametrics: Statistical Methods Based on Ranks*. Holden-Day Publisher.

Lehmann, E.L. y Casella, G. (20021). *Theory of Point of Estimation*. Springer.

Lind, D. A., Marchal, W. G., Wathen, S. A., Obón, D. P. y León, J. (2012). *Estadística aplicada a los negocios y la economía*. McGraw-Hill/Interamericana Editores.

Lind, D. A., Marchal, W. G. y Wathen, S. A. (2015). *Statistical Techniques in Business & Economics: Econ 209*. McGraw-Hill Education.

Lindgren, B. (1993). *Statistical Theory*. Chapman-Hall.

López, F. (2002). El análisis de contenido como método de investigación. *Revista de Educación*, *4*(1), 167-179.

Manoukian, E. B. (1986). *Mathematical Nonparametric Statistics*. Gordon and Breach Science Publishers.

Martínez, A. y Caballero, J. (2016). Borel y Steingaus: dos piezas claves para comprender el surgimiento de la probabilidad moderna. *Lecturas Matemáticas, 37*(1), 37-61.

Marshal, A. W. y Olkin, I. (2007). *Life Distributions*. Springer.

Mason, R, McKenzie, J. y Ruberg, S. (1990). A Brief History of America Statistical Association, 1839-1989. *American Statistician, 44*(2), 68 – 73.

Mayen, S. (2009). *Comprensión de las medidas de tendencia central en estudiantes mexicanos de educación secundaria y bachillerato*. [Tesis doctoral no publicada]. Universidad de Granada, España.

Mayorga, H. (2003). *Inferencia estadística*. Editorial Unibiblos/ Universidad Nacional.

Mendonça, T., Cautinho, C. y Almouloud, S. (2006). Mathematics Education and Statistics Education: Meeting Points and Perspectives. In A. Rossman y B. Chance (Eds.). *Proceedings of the Seventh International Conference on Teaching Statistics*. [CD- ROM]. Salvador (Bahia), Brazil. International Association for Statistical Education and International Statistical Institute

Ministerio de Educación Nacional de Colombia –MEN- (2006). *Estándares básicos de competencias en matemáticas, lenguaje, ciencias y ciudadanía*. MEN.

Montgomery, D. y Runger, G. (2012). *Probabilidad y estadística aplicadas a la ingeniería*. México.

Newbold, P., Carlson, W. y Thorne, B. (2013) *Estadística para administración y economía*. Pearson/Prentice Hall.

Newman, G. (2006). El razonamiento inductivo y deductivo dentro del proceso de investigación en ciencias experimentales y sociales. *Laurus, 12*(1), 180-205.

Olaya, O, J. (2012). *Métodos de regresión no paramétrica*. Programa Editorial Univalle.

Occeli, M. y Valeiras, N. (2013). Los libros de texto en ciencias como objeto de investigación: una revisión bibliográfica. *Enseñanza de las Ciencias, 31*(2), 133-153.

Peña, T. y Pirela, J. (2007). La complejidad del análisis documental. *Información, Cultura y Sociedad*, *16*(1), 55-81.

Peña, D. y Romo, J. (1997). *Introducción a la estadística para las ciencias sociales.* McGraw-Hill.

Pérez, C. (2005). *Técnicas estadísticas con SPSS 12: aplicaciones al análisis de datos*. Pearson/Prentice Hall.

Pinto, J. E. (2010). *Conocimiento didáctico del contenido sobre la representación de datos estadísticos: un estudio de casos con profesores de estadística de las carreras de psicología y educación*. [Tesis doctoral no publicada]. Universidad de Salamanca, España.

Palop, M. P. y García, P. A. (2017). El libro de texto como objeto de estudio y recurso didáctico para el aprendizaje: fortalezas y debilidades. *Revista Electrónica Interuniversitaria de Formación del Profesorado, 20*(1), 201-217.

Pratt, J. W. y Gibbons, J. D. (1981). *Concepts of Nonparametric Theory*. Springer-Verlag.

Randles, R. y Wolfe, D. (1979). *Introduction to the Theory of Nonparametric Statistics*. John Wiley & Sons.

Rioboó, J., González, P. y Tato, M. (1997). Resumen histórico de la evolución de la estadística. *Estudios de Economía aplicada, 8*, 141-162.

Ríos, A. R. y Peña, A. M. P. (2020). Estadística inferencial. Elección de una prueba estadística no paramétrica en investigación científica. *Horizonte de la Ciencia*, *10*(19), 191-208.

Rivera, A. F. (2003). *Una prueba de contrastes ortogonales antes de un experimento en estadística no paramétrica utilizando métodos* *intensivos de computo (MIC)*. [Tesis de Licenciatura en Ingeniería Industrial], Universidad de las Américas, México.

Salinas, P. y Alanís, J. (2009). Hacia un nuevo paradigma en la enseñanza del cálculo. *Revista Latinoamericana de Investigación en Matemática Educativa, 12*(3), 355-382.

Sánchez, E. y Monroy, J. (2013). La cuantificación del azar, una articulación de las definiciones, subjetiva, frecuencial y clásica de la probabilidad y la probabilidad condicionada. *Revista de la Didáctica de la Estadística, 2*(1), 39-46.

Siegel, S. (1970). *Diseño experimental no paramétrico aplicado a las ciencias de la conducta* (N.° BF39 S5e). Editorial F. Trillas.

Solís, M. S. (2014). ¿Por qué algunos aun prohíben utilizar estadística paramétrica para analizar datos ordinales? *Enseñanza e Investigación en Psicología, 19*(2), 1-14.

Soto, D. y Cantoral, R. (2014). Discurso matemático escolar y exclusión. *Bolet. Educ. Matemát, 28*(50), 1525-1544.

Uribe, J. (2005). *La investigación documental y el estado del arte como estrategia de intervención en ciencias sociales*. Editorial Universidad Piloto de Colombia

Valdivieso, M. (2011). *Estadística descriptiva, apoyo al estudio independiente. Contiene comando de R*. Editorial UPTC.

Van Dormolen, J. (1986). Textual Analysis. In B. Christiansen, A. G. Howson and M. Otte (Eds.), *Perspectives on Mathematics education* (pp. 141-171). Reidel.

Vera, S. A., Hernández, B. R. R., Ortiz, L. V., Huerta, J. A. A. y Pontoni, Á. C. (2021). El problema de la transparencia didáctica del parámetro en los textos de estadística. *Interciencia, 46*(11), 416-422.

Wasserman, L. (2006). *All of Nonparametric Statistics*. Springer.

Walpole, R., Myers, R. y Myers, S. (2012). *Probabilidad y estadística para ingeniería y ciencias*. (9.ª ed.). Pearson Educación.

Wild, C. y Pfannkuch, M. (1999). Statistical Thinking in Empirical Enquiry. *International Statistical Review, 67*(3), 223-265.

Colección de Investigación UPTC N.° 268

La inferencia estadística es una herramienta esencial dentro del método científico experimental. Los procesos inferenciales pueden desarrollarse mediante la aplicación de métodos cuantitativos paramétricos o de modelos estadísticos no paramétricos. Los primeros se utilizan cuando se cumplen supuestos estrictos sobre la forma de la distribución de probabilidad de la población de la cual se ha de seleccionar la muestra aleatoria para estimar uno o más parámetros desconocidos de la población, construir intervalos de confianza para tales parámetros o determinar regiones críticas de rechazo de la hipótesis nula. Los modelos estadísticos no paramétricos son técnicas estadísticas que requieren supuestos más flexibles o usan estadísticas de prueba libres de distribución, las cuales generalmente se definen por medio de los rangos vinculados a los datos de la variable o variables por investigar, ya sea que se trate de un problema de inferencia relacionado con una muestra aleatoria, dos, tres o más muestras.

Este trabajo investigativo indaga acerca de la manera como algunos libros de texto universitarios examinan los tópicos de estadística no paramétrica referentes a la prueba de hipótesis. Para este propósito, se utiliza una metodología basada en un enfoque mixto de investigación e involucra diversas situaciones problema concretas, que permiten promover la investigación científica desde los claustros universitarios. Esta obra está dirigida a estudiantes, profesores, profesionales e investigadores interesados en aplicar modelos estadísticos no paramétricos clásicos en la prueba de hipótesis en sus trabajos de investigación o en incrementar sus saberes sobre tales modelos. Para abordarla se necesitan algunos conocimientos básicos de estadística descriptiva e inferencial desde una mirada paramétrica.

Uptc Universidad Pedagógica y Tecnológica de Colombia
VIGILADA MINEDUCACIÓN
ACREDITACIÓN INSTITUCIONAL DE ALTA CALIDAD MULTICAMPUS
RESOLUCIÓN 023655 DE 2021 MEN / 6 AÑOS
Vie Vicerrectoría de Investigación y Extensión
EDITORIAL UPTC