

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



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

© Víctor Miguel Ángel Burbano Pantoja, 2023

© Margoth Adriana Valdivieso Miranda, 2023

© Ángela Saray Burbano Valdivieso, 2023

© Universidad Pedagógica y Tecnológica de Colombia, 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

Rector, UPTC

Enrique Vera López

Comité Editorial

Dr. Carlos Mauricio Moreno Téllez

Dr. Jorge Andrés Sarmiento Rojas

Dra. Yolima Bolívar Suárez

Mg. Pilar Jovanna Holguín Tovar

Dra. Nelsy Rocío González Gutiérrez

Dra. Ruth Maribel Forero Castro

Dr. Óscar Pulido Cortés

Mg. Edgar Nelson López López

Editor en jefe

Ph. D. Witton Becerra Mayorga

Coordinadora editorial

Mg. Andrea María Numpaque Acosta

Corrección de estilo

Claudia Helena Amarillo Forero

Diseño

Grupo de investigación GICI

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. Se permite la reproducción parcial o total, con la autorización expresa de los titulares del derecho de autor. Este libro es registrado en Depósito Legal, según lo establecido en la Ley 44 de 1993, el Decreto 460 de 16 de marzo de 1995, el Decreto 2150 de 1995 y el Decreto 358 de 2000.

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.