Horizonte de la Ciencia
ISSN: 2304-4330
ISSN: 2413-936X
horizontedelaciencia@gmail.com
Universidad Nacional del Centro del Perú
Perú

HORIZONTE DE LA CIENCIA
DOI: https://doi.org/10.26490/uncp.horizonteciencia.2020.19.597

Estadística inferencial. Elección de una prueba estadística no paramétrica en investigación científica

Mayay ćhalkachi. Ćhalkachi tapukuy aklay mana tupuna yaćhayp ashinakaaćhu

Kantëgotiro antagaisati. Jikotagantsi monkaratagantsi ogotagantsijenga kengagagntsipage

Okantakoyetari abetsikayetero kamantayeteri, aisati antimakoyete ayoyeteri

Recibido: 06 Septiembre 2019 Corregido: 06 Octubre 2019 Aprobado: 22 Diciembre 2019

Alejandro Ramírez Ríos
Nacionalidad: Peruana, Filiación: Universidad César Vallejo, Perú
Correo: ramirez.estudios@gmail.com; ORCID: https://orcid.org/0000-0003-0976-4974

Ana María Polack Peña
Nacionalidad: Peruana, Filiación: Universidad Nacional Enrique Guzmán y Valle
Correo: anapolack87@gmail.com; ORCID: https://orcid.org/0000-0003-4691-1929

Resumen
El propósito del artículo, fue identificar las pruebas no paramétricas no sujetos a una distribución de probabilidad normalizada para el análisis inferencial adecuado de datos provenientes de muestras pequeñas. Mediante la teoría fundamentada se describió su fundamento y uso: 1 muestra (Binomial, Chi-cuadrado, Kolmogorov-Smirnov, de rachas), 2 muestras independientes (Moses, Kolmogorov-Smirnov, rachas de Wald-Holfowitz, U-Mann Whitney), 2 muestras pareadas (De signo, McNemar, Wilcoxon), m muestras no pareadas (Mediana, Kruskal-Wallis, Jonckeere-Terpstra) y m muestras pareadas (Fridman, Q-Cochran, W-Kendall). Se concluye que estas pruebas son valiosas y robustas, la elección está sujeto al diseño, número y escala de medición de las variables.

Palabras clave:
Estadística no paramétrica; investigación científica.

Lisichiku limaykuna:
mana yupana ćhalkachi, allipyaćhay ashina.

Nibarintsipage Katingaro:
Kantëgotiro ora ogotagantsipage kengagantsipage.

Ñantsipe ayoyeteri:
Monkarayetirori, nesankirori yotantsipe.

Inferential Statistics. Choice of a Non Parametric Statistical Test in Scientific Research

Abstract

The purpose of the article was to identify non-parametric tests not subject to a normalized probability distribution for adequate inferential analysis of data from small samples. Based on the grounded theory, its rationale and use were described: 1 sample (Binomial, Chi-square, Kolmogorov-Smirnov, of streaks), 2 independent samples (Moses, Kolmogorov-Smirnov, streaks of Wald-Holfowitz, U-Mann Whitney), 2 paired samples (Sign, McNemar, Wilcoxon), m unpaired samples (Medium, Kruskal-Wallis, Jonckeere-Terpstra) and m paired samples (Fridman, Q-Cochran, W-Kendall). It is concluded that these tests are valuable and robust, the choice is subject to the design, number and scale of measurement of the variables.

Keywords

Non-parametric statistics; scientific investigation.

Estatística inferencial. Seleção de um ensaio estatístico não paramétrico em pesquisa científica

Resumo

O objetivo do artigo foi identificar os testes não paramétricos não sujeitos a uma distribuição de probabilidade normalizada para análise inferencial adequada de dados de pequenas amostras. Com base na teoria fundamentada, descreveu-se seu fundamento e uso: 1 amostra (Binomial, qui-quadrado, Kolmogorov-Smirnov, de rachas); 2 amostras independentes (Moses, Kolmogorov-Smirnov, rachas de Wald-Holfowitz, U-Mann Whitney); 2 amostras pareadas (dos sinais, McNemar, Wilcoxon), “m” amostras não pareadas (Médio, Kruskal-Wallis, Jonckeere-Terpstra) e “m” amostras pareadas (Fridman, Q-Cochran, W-Kendall). Conclui-se que esses testes são valiosos e robustos, a escolha está sujeita ao design, número e escala de medição das variáveis.

Palavras-chave:

Estatística não paramétrica; investigação científica.

Datos de los autores

Alejandro Ramírez Ríos: Investigador y docente de matemática y estadística. Doctor en Ciencias de la Educación por la Universidad Nacional Enrique Guzmán y Valle, Perú. Magister en Ciencias de la Educación con Mención en Educación Matemática por la Universidad Nacional Enrique Guzmán y Valle, Perú.

Ana María Polack Peña;

Investigador y docente de lengua y literatura e inglés. Doctor en Ciencias de la Educación por la Universidad Nacional Enrique Guzmán y Valle, Perú. Magister en Ciencias de la Educación con Mención en Didáctica de la Comunicación por la Universidad Nacional Enrique Guzmán y Valle, Perú.

Introducción

Uno de los problemas que se observa en muchas investigaciones es el uso inadecuado de los estadísticos de prueba; esto se debe a varios factores: a) desconocimiento de la estadística tanto descriptiva como inferencial, b) poco dominio de la metodología de investigación, c) falta de docentes investigadores y d) desconocimiento del manejo de softwares estadísticos (Excel, Minitab, Stata, Sas, R, Geogebra, etc.). Estos cuatro factores inciden directamente en la elaboración de una investigación, sin el dominio y puesto en práctica alguno de ellos, dicha investigación adolecería de muchas falencias, y un aspecto importante en todo esto es ¿cómo probar las hipótesis en una investigación? ¿Cómo y cuál estadístico de prueba usar adecuadamente?

En el campo de la investigación cada institución propone una estructura para realizar investigaciones científicas. Es importante resaltar que, en el enfoque cuantitativo, es esencial la formulación de hipótesis. El dilema de la mayoría de los investigadores es como probar dichas hipótesis o cual estadístico utilizar para el efecto; porque existen las llamadas pruebas paramétricas y no paramétricas. El objetivo del artículo es brindar información respecto al uso de las estadísticas no paramétricas en la prueba de hipótesis que requiere la investigación.

La importancia de las pruebas no paramétricas o de dispersión libre es que no están expuestas a requisitos previos específicos como las pruebas paramétricas. Valiosos para tamaños pequeños de muestras menores a 30 casos, en situaciones donde la variable que nos interesa es una escala ordinal. En ciencias sociales y la pedagogía, el uso de pruebas no paramétricas para el análisis inferencial es normal ya que hay numerosas variables que no persiguen las condiciones paramétricas. Es necesario mencionar que el análisis inferencial forma parte de los protocolos de investigación con el fin de obtener conclusiones inferidas de la muestra y generalizada a la población.

Método

El artículo se centra netamente en la técnica de investigación documentada, el cual se apoyó en fuentes documentales de estudio como artículos científicos y libros de texto de los mismos autores y creadores de las pruebas no paramétricas. Para entender con claridad el tema propuesto se optó por el método descriptivo. Es decir, describir las cualidades básicas de las pruebas no paramétricas mediante el diseño de teoría fundamentada y las teorías sustantivas. Esto permite al lector tener conocimiento adecuado de la estructura y comportamiento de las variables al hacer uso estas pruebas. Para los criterios de análisis, organización, estructura y finalmente el desarrollo del tema, se tuvo en consideración algunos aspectos, que se detallan:

Como primer aspecto, tomar en cuenta las cualidades de las pruebas no paramétricas: a) sencillas de aplicar; b) los datos presentan jerarquía, c) cuando dos observaciones son de diferente población, d) única opción para muestras pequeñas (menor o igual a 30), e) para calcular se ordena por rangos o frecuencia y f) las hipótesis son formuladas sobre medianas y rangos.

Como segundo aspecto, considerar tres requisitos importantes para esta elección: a) el diseño de la investigación (experimental: experimental puro, pre-experimental y cuasi experimental, y no experimental: transversal o longitudinal), b) el número de mediciones

(mediciones antes y después, varios tratamientos de una variable) y c) escala de medición de las variables (cualitativa: nominales u ordinales, cuantitativas: discretas y continuas).

Como tercer aspecto, la existencia de diversidad de pruebas no paramétricas en la estadística utilizadas en la investigación, existiendo diversas clasificaciones, dependiendo de los requisitos y/o de los autores e investigadores. Ferrán (2012) clasifica en pruebas para una muestra y los demás en no paramétricas. Visauta (2007) clasifica en no paramétricas todas las pruebas. Se clasifica por variable, tipo y número de muestras; otros autores consideran el tipo de muestra para esta clasificación y considera diferente terminología y número de pruebas no paramétricas para identificarlos; mientras que Pérez, García, Gil y Galán (2009) a la clasificación anterior adiciona las pruebas de libre elección correlacionales. Esta clasificación en el presente artículo, toma en cuenta el número de muestras (una, dos o más), tipo de muestras (independientes o dependientes) y el tipo de variables (cualitativa o cuantitativa), más no las correlacionales.

Desarrollo del contenido

Estadísticos no paramétricos para 1 muestra

En la tabla 1 se presenta un resumen de las estadísticas no paramétricas para 1 muestra, tiene el objetivo evaluar el nivel de ajuste a una determinada distribución de probabilidad los datos. La característica que los identifica es que las variables son categóricas cuya variable dependiente es nominal, a excepción de la prueba K-S cuya variable dependiente es ordinal o de intervalo.

Tabla1
Resumen de las pruebas no paramétricas sugeridas para una muestra

N° de muestras	Prueba sugerida	Variables
1 muestra	Prueba Chi-cuadrado de Pearson (*)	VD: Nominal
	Prueba Binomial	
	Prueba de Rachas	
	Prueba de Kolmogorov-Smirnov	VD: Ordinal/Intervalo

*Se puede utilizar la prueba exacta de Fisher cuando el grupo tiene como valor menor que 5

Prueba de Chi-cuadrado de Pearson: Se usa para verificar si se acomoda o no a una distribución teórica, los datos obtenidos de una variable categórica. Este ajuste puede ser distribución uniformemente o binomialmente o multinomialmente. Esta prueba denotada por χ^2 fue creada por Pearson (1900) con la finalidad de establecer comparaciones entre las frecuencias observadas y esperadas de una muestra y que se debe encontrar si se sigue una distribución teórica. A través de las tablas de contingencia el Chi-cuadrado verifica Ho de la independencia de dos variables, cada una con sus respectivas categorías (dos o más categorías).

Ejemplo:
Título de la investigación: actitud y enseñanza del curso de estadística aplicada.
Ho: la actitud es independiente de la enseñanza del curso de estadística aplicada
H1: la actitud es dependiente de la enseñanza del curso de estadística aplicada

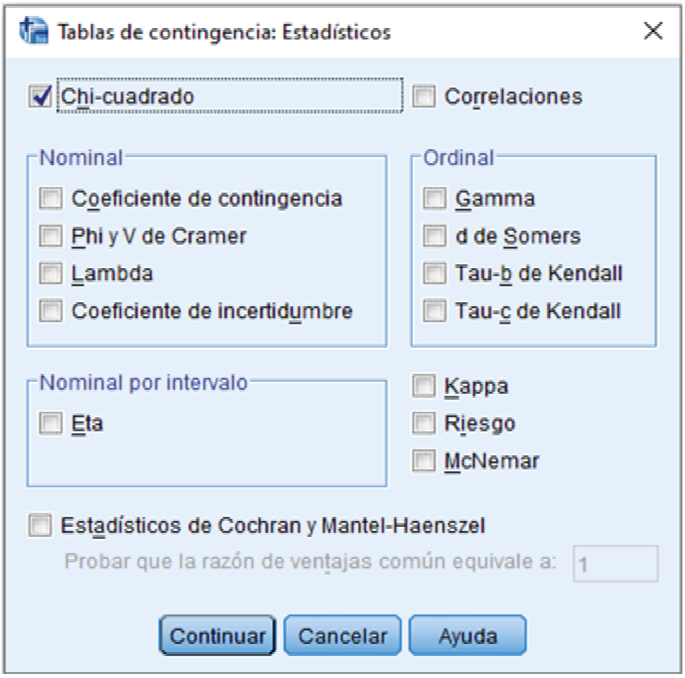


Figura 1: cuadro de secuencia del estadístico Chi-cuadrado

La tabla 2 de contingencia muestra el recuento y porcentajes de la relación entre la actitud y la enseñanza del curso de estadística aplicada

Tabla 2
Tabla de contingencia entre actitud y enseñanza del curso de estadística aplicada

			Enseñanza del curso de Estadística Aplicada				Total
			Sección A	Sección B	Sección C	Sección D	
Actitud	En contra	Frecuencia	3	1	0	0	4
		% del total	25,0%	8,3%	0,0%	0,0%	33,3%
	Indiferente	Frecuencia	0	2	2	0	4
		% del total	0,0%	16,7%	16,7%	0,0%	33,3%
	A favor	Frecuencia	0	0	1	3	4
		% del total	0,0%	0,0%	8,3%	25,0%	33,3%
Total		Frecuencia	3	3	3	3	12
		% del total	25,0%	25,0%	25,0%	25,0%	100,0%

La tabla 3, nos presenta la información adecuada para determinar una decisión: el estadístico Chi-cuadrado resultó 16,000, con 6 gl. y P-valúe 0,014 (0,014 < 0,05). Como el nivel de significancia es 0,014 < 0,05, es rechazada Ho; en consecuencia, la actitud es dependiente de la enseñanza del curso de estadística aplicada

Tabla 3
Pruebas de Chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	16,000[a]	6	,014
Razón de verosimilitudes	18,729	6	,005
Asociación lineal por lineal	9,167	1	,002
N de casos válidos	12		

Prueba binomial: La prueba binomial, llamada prueba de bondad de ajuste, cuyo objetivo es verificar la suposición de que las frecuencias observadas de aciertos de una variable dicotómica se ajustan a una proporción teórica binomial; se realiza el contraste de hipótesis sobre proporciones y cuantiles. Cuando se ha seleccionado una variable que no es dicotómica se debe dicotomizar indicando el punto de corte para los dos grupos, este proceso es importante cuando se contrasta la hipótesis sobre la mediana o sobre otro cuantil.

Prueba de Rachas: Es una prueba de independencia, sirve para verificar si los datos de una muestra se distribuyen aleatoriamente por debajo y por encima de un determinado punto de corte, como el promedio, mediana, la moda o un punto elegido por el investigador. Además, permite averiguar si la cantidad de rachas observado en alguna muestra es tan pequeño o tan grande, para el rechazo de la hipótesis de independencia entre las rachas observadas.

Prueba de Kolmogorv-Smirnov (K-S): La prueba K-S permite contrastar la Ho si la distribución de los datos sigue una probabilidad teórica ya sea de tendencia normal, uniforme, de Poisson o exponencial. La diferencia con las tres pruebas anteriores, la prueba K-S es una prueba de ajuste de variables cuantitativas con escala de medida ordinal o de intervalo.

Estadísticos no paramétricos para 2 muestras autónomas

Presentamos un resumen (tabla 4) de los estadísticos no paramétricos para dos muestras autónomas. Tiene como objetivo comparar proporciones y medianas u otra característica equivalente al procedimiento de comparar promedios.

Tabla 4
Resumen de los estadísticos no paramétricos sugeridas para 2 muestras autónomas

N° de muestras	Tipo de muestras	Prueba sugerida	Variables
Dos muestras	Muestras independientes	U de Mann-Whitney (**) Prueba de Rachas de Wald-Wolfowitz Z de Kolmogorv-Smirnov Reacciones extremas de Moses	VI: Dicotómica VD: Ordinal/Intervalo

** Prueba no paramétrica del equivalente paramétrico t de Student

En el software SPSS se encuentra siguiendo la ruta: Analizar ››› Pruebas no paramétricas ››› Cuadro de diálogos antiguos ››› dos muestras independientes. Aparece la siguiente ventana:

Horizonte de la Ciencia 10 (19) jul-dic 2020 FE/UNCP. ISSN (Impreso): 2304-4330/ ISSN (En Línea): 2413-936X
Ramírez, A. y Polack, A. Estadística inferencial. Elección de una prueba estadística no paramétrica... 191-208

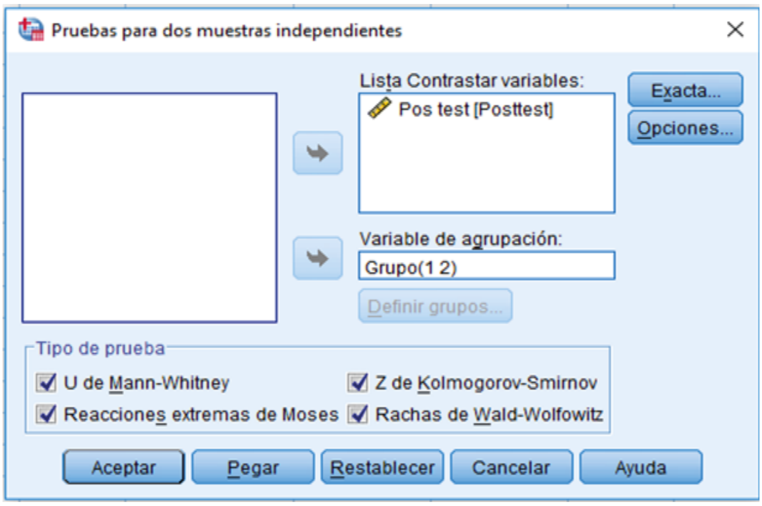


Figura 2: cuadro de secuencia de estadísticos no paramétricas para 2 muestras autónomas

Prueba U de Mann-Whitney: Introducido por Mann y Whitney (1947), sirve para verificar la Ho de que 2 muestras aleatorias autónomas provienen de dos poblaciones iguales o de una misma población, cuando no se cumple el supuesto de normalidad y homocedasticidad, medidos mínimo en escala ordinal. Se puede afirmar que es equivalente a la prueba H de Kruskal-Wallis para 2 grupos y a la prueba de suma de rangos de Wilcoxon. Esta prueba es la alternativa no paramétrica de la prueba paramétrica t de Student que compara promedios, mientras que la U de Mann –Whitney compara las diferencias entre dos medianas, por lo que se basa en rangos.

Prueba de Kolmogorov-Smirnov: Se usa para probar la Ho de que dos muestras autónomas se han cogido de una misma población. En este caso, compara de dos grupos la función de distribución acumulada. Esta prueba es frágil a cualquier tipo de comparación entre dos distribuciones descriptivas. Propuesto por Kolmogorov (1933) para obtener diferencia de rangos y posteriormente complementada por Smirnov (1939) para diferencia de rangos en valor absoluto, de allí el nombre de Z de K-S, que usando SPSS se puede obtener la diferencia extrema en valor absoluto.

Prueba de Rachas de Wald-Wolfowitz: La Ho prueba que las 2 muestras autónomas se han cogido de una sola población. Es similar a la prueba K-S, porque es frágil a cualquier tipo de diferencias entre las medidas descriptivas y otras, propuesto por Wald y Wolfowitz (1940).

Prueba de reacciones extremas de Moses: Método no paramétrico diseñado por Moses (1952) para variables ordinales, se usa para ver el grado de variabilidad o dispersión de dos grupos o distribuciones. Sirve para saber cuántos valores extremos del grupo experimental influyen en el grupo control.

Ejemplo:
Título de la investigación: Aplicación del MProy para el aprendizaje de la estadística en alumnos universitarios.

Ho: la media del GC es igual a la media del GE $(Ho: \mu_{GC} = \mu_{GE})$
H1: la media del GC es diferente a la media del GE $(Ho: \mu_{GC} \neq \mu_{GE})$

La tabla 5 muestra los rangos promedios de cada grupo y la suma de los mismos y la tabla 6, muestra que el sig. = 0,000 < 0,05, rechazamos Ho y aceptamos H1. Es decir, la media del GC es diferente a la media del GE *(Ho:* $\mu_{GC} \neq \mu_{GE}$*)*

Tabla 5
Rangos. Post test

Post test	N	Rango promedio	Suma de rangos
Grupo control	28	14,61	409,00
Grupo experimental	28	42,39	1187,00
Total	56		

Tabla 6.
Prueba U de Mann-Whitney

	Pos test
U de Mann-Whitney	3,000
W de Wilcoxon	409,000
Z	-6,422
Sig. asintót. (bilateral)	,000

La tabla 7 muestra el estadístico de reacciones extremas de Moses. Vemos que el nivel de significancia tanto en la amplitud observada como en la recortada del upo control es menor a 0,05, podemos decir existen reacciones extremas; la última fila indica el número de casos eliminados por arriba y por abajo para obtener la amplitud recortada.

Tabla 7
Estadístico de Moses

		Pos test
Amplitud observada del grupo control	N	30
	Sig. (unilateral)	,000
Amplitud recortada del grupo control	N	23
	Sig. (unilateral)	,000
Valores atípicos recortados de cada extremo		1

La prueba de Kolmigorov-Smirnov de la tabla 8 indica que el valor crítico es menor 0,05, podemos rechazar Ho aceptar H1 y concluir que el promedio del GC es distinto del promedio del GE . *(Ho:* $\mu_{GC} \neq \mu_{GE}$*)* El estadístico de Rachas de Wald-Wolfowitz de la tabla 9 indica que el sig.= 0,000 es menor a 0,05, rechazamos Ho y aceptamos H1, concluyendo que la media del GC es diferente a la media del GE.

Tabla 8
Prueba de Kolmogorov-Smirnov

		Pos test
Diferencias más extremas	Absoluta	,964
	Positiva	,964
	Negativa	,000
Z de Kolmogorov-Smirnov		3,608
Sig. asintót. (bilateral)		,000

Horizonte de la Ciencia 10 (19) jul-dic 2020 FE/UNCP. ISSN (Impreso): 2304-4330/ ISSN (En Línea): 2413-936X
Ramírez, A. y Polack, A. Estadística inferencial. Elección de una prueba estadística no paramétrica... 191-208

Tabla 9
Prueba de Rachas de Wald-Wolfowitz

	Número de rachas	Z	Sig. asintót. (unilateral)
Mínimo posible	2^c	-7,283	,000
Máximo posible	4^c	-6,743	,000

c. Hay 1 empates inter-grupos que implican 7 casos.

Estadísticos no paramétricos para 2 muestras pareadas

Presento un resumen (tabla 10) de los estadísticos no paramétricos para 2 muestras relacionadas. Su objetivo es comparar proporciones y medianas u otra característica equivalente al procedimiento de comparar medias.

Tabla 10
Resumen de estadísticos no paramétricos sugeridas para 2 muestras relacionadas

N° de muestras	Tipo de muestras	Prueba sugerida	Variables
Dos muestras	Muestras dependientes	X^2 de McNemar	VI: Dicotómina VD: Nominal
		Prueba de signos T de Wilcoxon (**)	VI: Dicotómica VD: Ordinal/Intervalo

** Prueba no paramétrica del equivalente paramétrico t de Student

En el software SPSS se encuentra siguiendo la ruta: Analizar >>> Pruebas no paramétricas >>> Cuadro de diálogos antiguos >>> dos muestras independientes. Aparece la ventana:

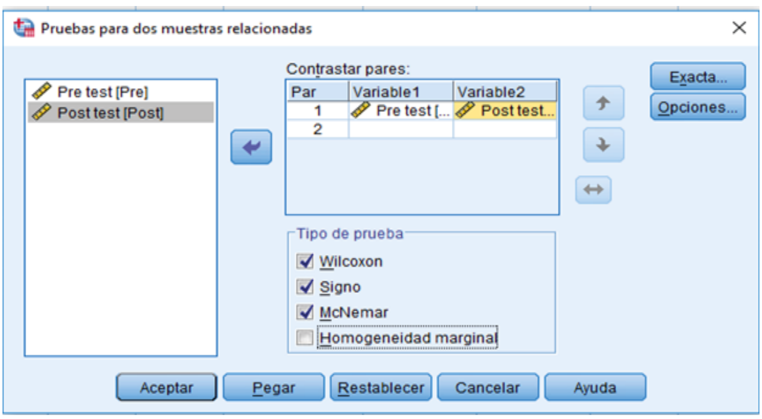


Figura 4: cuadro de secuencia de los estadísticos no paramétricos para 2 muestras relacionadas

Prueba de Wilcoxon: Se usa para verificar la Ho de igualdad entre 2 medianas poblacionales, la variable debe ser continua y observaciones emparejadas; es decir, datos de la misma muestra con medición de pre y post prueba (Wilcoxon, 1945). Es el equivalente no paramétrico de la estadística paramétrica t de Student para 2 muestras emparejadas (Pérez, 2001).

Prueba de X^2 McNemar: Este estadístico diseñado por McNemar (1947), se usa para demostrar la Ho si las proporciones son iguales. Se usa cuando las respuestas de cada persona se obtienen dos veces: antes y después de ocurrido un suceso en particular. En otras palabras, se aplica a diseños de tipo pre y post (diseño experimental) y probar la potencia del tratamiento efectuado entre los dos. Es necesario arreglar los datos en función a una tabla cruzada de 2x2.

Prueba de los signos: Se usa para verificar la Ho si dos medianas poblacionales son iguales, es decir, tener aproximadamente la misma cantidad de signos más (+) y signos menos (-). También se usa para identificar cuál de las variables es mayor que la otra o verificar la tendencia que puede seguir un conjunto de variables ordinales positivas. Tiene afinidad con la prueba binomial.

Ejemplo:
Objetivo: comparar medias. (Influencia de una metodología en aprendizaje de la estadística)
Ho: El promedio del GE del pre test es igual a la media del GE en el post test *(Ho: $\mu_{GC} = \mu_{GE}$)*
H1: El promedio del GE del pre test es diferente a la media del GE en el post test
(Ho: $\mu_{GC} \neq \mu_{GE}$)

Se muestra el rango promedio y la suma de rangos positivos y negativos (tabla 11), como los empates, las indicaciones debajo de la tabla permiten la interpretación adecuada. También se muestra la prueba de Wilcoxon (tabla 12), cuyo sig. = 0,000 es menor a 0,05, luego Ho es rechazada y aceptada H1, concluyendo que el promedio del GE del pre test es diferente a la media del GE en el post test *(Ho: $\mu_{GC} \neq \mu_{GE}$)*

Tabla 11
Rango promedio T de Wilcoxon

Post test - Pre test	N	Rango promedio	Suma de rangos
Rangos negativos	0[a]	,00	,00
Rangos positivos	28[b]	14,50	406,00
Empates	0[c]		
Total	28		

a. Post test < Pre test
b. Post test > Pre test
c. Post test = Pre test

Tabla 12
Prueba de Wilcoxon

	Pretest - Post test
Z	-4,646[b]
Sig. asintót. (bilateral)	,000

a. Prueba de los rangos con signo de Wilcoxon
b. Basado en los rangos negativos.

La tabla 13 presenta las diferencias positivas, negativas y los empates, la información debajo de la tabla permiten identificar estas diferencias. Se muestra la prueba de los signos (tabla 14) cuyo estadístico Z es -5,103 (tamaño de la muestra > a 25) y el sig. = 0.000 es menor a 0,05, luego rechazamos Ho y aceptamos H1, es decir que el promedio del GE del pre test es diferente a la media del GE en el post test *(Ho: $\mu_{GC} \neq \mu_{GE}$)*

Horizonte de la Ciencia 10 (19) jul-dic 2020 FE/UNCP. ISSN (Impreso): 2304-4330/ ISSN (En Línea): 2413-936X
Ramírez, A. y Polack, A. Estadística inferencial. Elección de una prueba estadística no paramétrica... 191-208

Tabla 13
Frecuencias

Post test - Pre test	N
Diferencias negativas[a]	0
Diferencias positivas[b]	28
Empates[c]	0
Total	28

a. Post test < Pre test
b. Post test > Pre test
c. Post test = Pre test

Tabla 14
Prueba de los signos

	Post test - Pre test
Z	-5,103
Sig. asintót. (bilateral)	,000

a. Prueba de los signos

Estadísticos no paramétricos para m muestras autónomas

Se presenta el resumen de los estadísticos no paramétricos para m muestras autónomas, están diseñadas para verificar datos que provienen de diseños con variable independiente categórica (más de dos categorías definidas en más de muestras o grupos) y con VD cuantitativa, mínimo con escala ordinal. El objetivo es comparar proporciones y medianas u otra característica equivalente al procedimiento de comparar medias o promedios.

Tabla 15
Resumen de los estadísticos no paramétricos sugeridas para m muestras autónomas

N° de muestras	Tipo de muestras	Prueba sugerida	Variables
Mayor a dos muestras	Muestras independientes	Prueba de la mediana H de Kruskal-Wallis (***) Prueba de Jonkheere-Terpstra	VI: Politómica VD: Ordinal/Intervalo

*** Prueba no paramétrica del equivalente paramétrico ANOVA

En el software SPSS se encuentra siguiendo la ruta: Analizar ››› Pruebas no paramétricas ››› Cuadro de diálogos antiguos ››› k muestras independientes, se muestra la siguiente ventana:

Figura 3: cuadro de secuencia de los estadísticos no paramétricos para m muestras autónomas.

Prueba H de Kruskal-Wallis: Introducido por Kruskal y Wallis (1952), sirve para probar la Ho de que las m muestras autónomas provienen de poblaciones similares o de una misma población, aquí la variable que se estudia tiene una distribución continua, con escala mínimamente ordinal. Es una ampliación del estadístico U de Mann-Whitney, es la prueba equivalente al estadístico paramétrico del ANOVA de un factor completamente aleatorizado

Prueba de la mediana: Parecida al estadístico Chi-cuadrado, sin embargo, en lugar de utilizar dos variables categóricas, ahora son cuantitativas. Sirve para verificar la Ho de que m grupos o muestras surgen de poblaciones que tengan iguales medianos, cuando no cumple el requisito de normalidad o cuando la variable es cuantitativa discreta.

Prueba de Jonckheere-Terpstra: Llama también prueba exacta de J-T, es más robusta que las pruebas anteriores (la mediana y kruskal-Wallis), cuando la distribución posee ordenación a priori (ascendente o descendente) de las m poblaciones de las cuales son extraídas las muestras (Jonckheere, 1954; Terpstra, 1952).

Ejemplo

El título de la investigación: Programa de tratamiento salud para la depresión en estudiantes universitarios.

Ho: El programa de tratamiento salud no influye en la depresión de estudiantes universitarios

H1: El programa de tratamiento salud influye en la depresión de estudiantes universitarios
La tabla 16 muestra el rango promedio de depresión después del tratamiento, precisa tamaño y rangos promedios obtenidos de los seis grupos. La tabla 17 muestra que el nivel de sig. = 0,001 < 0,05. Luego se rechaza Ho. Por lo tanto, el programa de tratamiento salud influye en la depresión de estudiantes universitarios

Horizonte de la Ciencia 10 (19) jul-dic 2020 FE/UNCP. ISSN (Impreso): 2304-4330/ ISSN (En Línea): 2413-936X
Ramírez, A. y Polack, A. Estadística inferencial. Elección de una prueba estadística no paramétrica... 191-208

Tabla 16
Rangos. Depresión final

Tratamiento	N	Rango promedio
Sin tratamiento	10	48,45
Tratamiento AH	12	30,35
Tratamiento YX	10	29,50
Tratamiento YZX	8	17,25
Tratamiento ABC	10	36,05
Tratamiento psicológico	10	21,40
Total	60	

Tabla 17
Estadísticos de prueba H

	Después
H de Kruskal-Wallis	20.452
gl	5
Sig. asintótica	.001

a. Prueba de Kruskal Wallis

b. Variable de agrupación: Tratamiento

Presentamos las frecuencias (tabla 18) de los tratamientos por encima y debajo de la mediana. En el estadístico de prueba (19) el nivel de sig. = 0,003 es menor a 0,05. Luego rechazamos Ho. Por lo tanto, el programa de tratamiento salud influye en la depresión de estudiantes universitarios

Tabla 18
Frecuencias

	Tratamiento. Después					
	Sin trat.	Trat. AH	Trat. YX	Trat. YZX	Trat. ABC	Trat. Psi.
> Mediana	9	4	3	0	4	3
<= Mediana	1	6	7	10	6	7

Tabla 19
Estadísticos de prueba

	Tratamiento después
N	60
Mediana	13.00
Chi-cuadrado	18.120[b]
gl	5
Sig. asintótica	.003

La prueba exacta de J-T (tabla20) ofrece las diferentes mediciones, la cantidad de tratamientos, el tamaño de la muestra y el sig. Asintótica bilateral. Como sig.=0,004 es menor a 0,05, rechazamos Ho. Por consiguiente, el programa de tratamiento salud influye en la depresión de estudiantes universitarios

Tabla 20
Estadísticos de Jonkheere-Terpstra

	Después
Número de niveles en Tratamiento	6
N	60
Estadístico J-T observado	529.500
Estadístico J-T de media	750.000
Desviación estándar del estadístico J-T	76.429
Desv. Estadística J-T	-2.885
Sig. asintótica(bilateral)	.004

a. Variable de agrupación: Tratamiento

Estadísticos no paramétricos para m muestras relacionadas

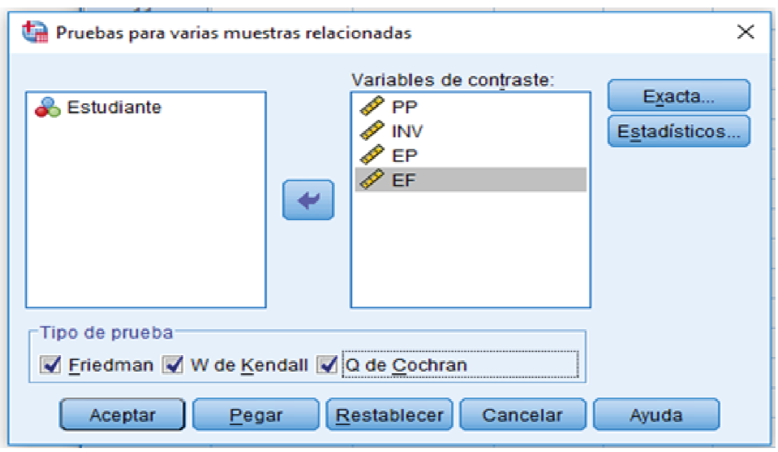
Se presenta el resumen de los estadísticos no paramétricos para m muestras dependientes. Permiten verificar datos que provienen de diseños con medidas repetidas. El objetivo de estas pruebas es comparar proporciones y medianas u otra característica equivalente al procedimiento de comparar medias o promedios.

Tabla 21
Resumen de los estadísticos no paramétricos sugeridas para m muestras dependientes

N° de muestras	Tipo de muestras	Prueba sugerida	Variables
Mayor a dos muestras	Muestras dependientes	F de Friedman (***)	VI: Politómica VD: Ordinal/Intervalo
		W de Kendall	
		Q de Cochran	VI: Dicotómica VD: Nominal

*** Prueba no paramétrica del equivalente paramétrico ANOVA

En el software SPSS se encuentra siguiendo la ruta: Analizar ››› Pruebas no paramétricas ››› Cuadro de diálogos antiguos ››› m muestras relacionadas, se muestra la siguiente ventana:



Horizonte de la Ciencia 10 (19) jul-dic 2020 FE/UNCP. ISSN (Impreso): 2304-4330/ ISSN (En Línea): 2413-936X
Ramírez, A. y Polack, A. Estadística inferencial. Elección de una prueba estadística no paramétrica... 191-208

Prueba F de Friedman: Esta prueba introducida por Friedman (1937) se usa cuando se compara k medias poblacionales en muestras relacionadas. El diseño lo forma m muestras o tratamientos pareados y 1 muestra aleatoria de n elementos o bloques autónomos entre sí y autónomos de los tratamientos. Las puntuaciones deben ser convertidos en rangos asignados independientemente para cada elemento o bloque. En estas condiciones, sirve para probar la Ho de la igualdad de los m tratamientos o equivalente las muestras provienen de la misma población.

Coeficiente de concordancia W de kendall: Kendall y Babington-Smith (1939) fueron los propulsores de este coeficiente en forma independiente. Se usa para estudiar la concordancia entre dos o más conjuntos de rangos. Posee las mismas características que el estadístico de Friedman.

Prueba Q de Cochran: Diseñada por Cochran (1950), para probar la Ho de que varias variables dicotómicas dependientes entre sí, tienen igual promedio; es similar al estadístico de Friedman, la diferencia es que la variable medida o dependiente es una variable dicotómica (puede tomar solo dos valores)

Ejemplo
Título de la investigación: Aplicación del MProy en el aprendizaje de la estadística aplicada
Ho: El promedio del GE del pre test es igual al promedio del GE en el post test *(Ho:* $\mu_{GC} = \mu_{GE}$*)*
H1: El promedio del GE del pre test es diferente al promedio del GE en el post test
(Ho: $\mu_{GC} \neq \mu_{GE}$*)*

Se presenta el rango promedio (tabla 22) del aprendizaje de la estadística aplicada con 4 notas de cada alumno en distintos momentos. Así mismo los estadísticos de Friedman (tabla 23), se aprecia que el sig.=0,000 es menor a 0,05, lo cual se rechaza Ho. Por lo tanto, el promedio del GE del pre test es diferente al promedio del GE en el post test
(Ho: $\mu_{GC} \neq \mu_{GE}$*)*

Tabla 22
Rangos promedio

	Rango promedio
Promedio de prácticas	3,90
Investigación	2,75
Examen parcial	1,50
Examen final	1,85

Tabla 23
Estadísticos de Friedman

Estadísticos de contraste[a]	
N	10
Chi-cuadrado	21,989
gl	3
Sig. asintót.	,000

a. Prueba de Friedman

Se puede observar el rango promedio (tabla 24) del aprendizaje de la estadística aplicada con 4 notas de cada alumno en distintos momentos. Así mismo los estadísticos de Friedman (tabla 25), se aprecia que el sig.=0,000 es menor a 0,05, lo cual se rechaza Ho. Por lo tanto, el promedio del GE del pre test es diferente al promedio del GE en el post test *(Ho:* $\mu_{GC} \neq \mu_{GE}$*)*

Tabla 24
Rangos promedio

	Rango promedio
Promedio de prácticas	3,90
Investigación	2,75
Examen parcial	1,50
Examen final	1,85

Tabla 25
Estadístico W de Kendall

Estadísticos de contraste	
N	10
W de Kendall[a]	,733
Chi-cuadrado	21,989
Gl	3
Sig. asintót.	,000

a. Coeficiente de concordancia de Kendall

Conclusiones

Se puede concluir que no existe clasificación estandarizada de las pruebas no paramétricas para el análisis inferencial en las investigaciones científicas.

La literatura respecto a los estadísticos de prueba es variada en cuanto a los autores, es necesario que el investigador adopte una posición investigativa para el buen uso de éstas pruebas, cuyo uso en muestras pequeñas es de vital importancia.

Las interpretaciones correctas del análisis inferencial dependerán del buen manejo por parte del investigador sobre las condiciones que presenta cada uno de los estadísticos no paramétricos, ya que estas pruebas no necesitan supuestos y la mayoría trabaja con la mediana.

En los ejemplos presentados, nos damos cuenta de la importancia del diseño, tipo de variables, tamaño de la muestra, entre otros, para el uso de estos estadísticos inferenciales en la investigación científica.

Horizonte de la Ciencia 10 (19) jul-dic 2020 FE/UNCP. ISSN (Impreso): 2304-4330/ ISSN (En Línea): 2413-936X
Ramírez, A. y Polack, A. Estadística inferencial. Elección de una prueba estadística no paramétrica... 191-208

Referencias bibliográficas

Berlanga, V. y Rubio, M. (2012) Clasificación de pruebas no paramétricas. Cómo aplicarlas en SPSS.Wallis *REIRE, Revista d'Innovació i Recerca en Educació*, Vol. 5, núm. 2, 101-113. Accesible en: http://www.ub.edu/ice/reire.htm

Cochran, W. (1954). Some methods for strengthening the common 2 tests. *Biometrics*. Vol. 10, No. 4 (Dec., 1954), pp. 417-451.

Ferran, M. (2002). *SPSS para Windows. Análisis estadístico*. Madrid: Osborne McGraw-Hill, 2001.

Friedman, M. (1937). *The use of ranks to avoid the assumption of normality implicit in the analysis of variance*. J Am Stat Assoc.

Hernández, R., Fernandez, C. y Baptista; P. (2014). *Metodología de la Investigación*. 6ta. edición. México, DF, México: McGraw-Hill Interamericana Editores.

Jonckheere, A. (1954). A distribution-free k-sample test against ordered alternatives. *Biométrika.* 41: 133–145. *Doi: 10.2307/2333011.*

Kendall, M. (1955). *Rank correlations methods*. Hafner Press, NYC; Edición: Second Edition (1955)

Kendall, M. y Babington-Smith, B. (1939). *Tables of Random Sampling Numbers*. Cambridge, England: Cambridge University Press.

Kolmogorov, A. (1933). *Sulla determinazione empirical di una legge di distribuzione*. Giornale Inst Ital Altuari.

Kruskal, W. y Wallis, W. (1952). *Use of ranks in one-criterion variance analysis*. J Am Stat Assoc.

McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*. pp. 153-157.

Mann, H. y Whitney, D. (1947). *On a test of whether one of two random variables is stochastically larger than the other*. Ann Math Stat.

Moses, L. (1952). *Non-parametric statistics for psychological research*. Psychol Bull.

Pearson, K. (1900). On a criterion that a given system of deviations from the probable in the case of correlated system of variables is Duch that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine*; 50: 157-75

Pérez, C. (2001). *Técnicas estadísticas y SPSS*. Madrid: Prentice Hall.

Pérez, R., García, J., Gil, J. y Galán, A. (2009). *Estadística aplicada a la Educación*. Madrid: UNED - Pearson.

Smirnov, N. (1939). Estimate of deviation between empirical distribution functions in two independent samples. Bull Moscow University.

Terpstra, T. (1952). The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking. *Indagationes Mathematicae. 14: 327–333.*

Visauta, B. (2007). *Análisis estadístico con SPSS para Windows. Estadística básica*. Madrid: McGraw-Hill.

Wald, A.; Wolfowitz, J. (1940). On a Test Whether Two Samples are from the Same Population. *Ann. Math. Statist*. Volume 11, Number 2, 147--162. doi:10.1214/aoms/1177731909. https://projecteuclid.org/euclid.aoms/1177731909

Wallis, W. (1939). *The coefficient of concordance was independently developed, the correlation ratio for ranked data*.

Wilcoxon, F. (1945). Individual comparison by ranking methods. *Biometrics Bulletin*, Vol. 1, No. 6. (Dec., 1945), pp. 80-83.