

Prueba Kolmogorov-Smirnov para dos muestras grandes con dos colas
Kolmogorov-Smirnov test for two large samples with two-tailed samples

Arabel Moraguez Iglesias¹, Roberto Johan Sierra Pérez.², María Eugenia Torres Santander³

DOI: [10.61236/ciya.v9i2.1097](https://doi.org/10.61236/ciya.v9i2.1097)

RESUMEN:

El problema investigado surge a partir de la identificación de deficiencias persistentes en tesis doctorales, maestrías y trabajos de diploma relacionados con la validación de hipótesis. El objetivo de este trabajo estuvo determinado en mostrar cómo validar una hipótesis mediante el empleo de la prueba estadística no paramétrica Kolmogorov Smirnov para muestras grandes y dos colas. Se tomó como referente la tesis de maestría del autor principal de este trabajo a través de la cual, se validaron los criterios de 140 profesores, tomados al azar, de la Educación Técnica y Profesional de la provincia de Holguín y 35 expertos de todo el país, lo que permitió determinar si hubo concordancia en relación a las evaluaciones emitidas por éstos, relacionados con las dimensiones e indicadores para evaluar el impacto educacional de las Escuelas Politécnicas Industriales de dicha provincia (propuestos por el autor). Como resultado se constató, con un 95% de confianza, que los criterios evaluativos de cada indicador emitidos por los expertos fueron coincidentes con los criterios opinados por los profesores de la Educación Técnica y Profesional de la provincia de Holguín, lo que permitió validar, con un mayor grado de confianza, la pertinencia de la hipótesis planteada en la tesis. Se puede concluir que este trabajo muestra de forma práctica cómo aplicar este valioso instrumento que posibilita validar las hipótesis de una investigación.

Palabras claves: estadística, hipótesis, Kolmogorov, prueba de hipótesis, prueba. Smirnov, tesis doctorales, tesis maestrías.

Recibido 12 de diciembre de 2019; revisión aceptada 26 de abril de 2020

ABSTRACT:

The problem investigated arose from the identification of persistent deficiencies in doctoral theses, master's degrees, and diploma theses related to hypothesis validation. The objective of this work was to show how to validate a hypothesis using the nonparametric Kolmogorov-

¹ ORCID: <https://orcid.org/0009-0008-8820-1654>. Email: arabel69@gmail.com

² ORCID: <https://orcid.org/0009-0009-8454-5932>. Email: robertojoansierraperez59@gmail.com

³ ORCID: <https://orcid.org/0009-0007-3167-1324>. Email: epciencia@citmahol.gob.cu

Smirnov test for large, two-tailed samples. The master's thesis of the main author of this work was used as a reference, through which the criteria of 140 randomly selected teachers from the Technical and Vocational Education Department in the province of Holguín and 35 experts from across the country were validated. This made it possible to determine whether there was concordance in relation to the evaluations issued by these teachers, related to the dimensions and indicators for assessing the educational impact of the Industrial Polytechnic Schools of said province (proposed by the author). As a result, it was established, with 95% confidence, that the evaluative criteria for each indicator issued by the experts were consistent with the criteria expressed by the teachers of the Technical and Vocational Education Department in the province of Holguín, which made it possible to validate, with a greater degree of confidence, the relevance of the hypothesis posed in the thesis. It can be concluded that this work demonstrates in a practical way how to apply this valuable tool that enables the validation of research hypotheses.

Keywords: *statistic, hypothesis, Kolmogorov, hypothesis test, proves, Smirnov, doctoral thesis, thesis masters.*

Recibido 12 de diciembre de 2024; revisión aceptada 26 de abril de 2025

1. INTRODUCCIÓN

En las investigaciones, la utilización estadística de las pruebas de hipótesis permiten demostrar, acorde al grado de significación asumido o nivel de confianza, la factibilidad de las hipótesis planteadas en una investigación; sin embargo, por lo general, en la práctica no son explotadas lo suficiente en las tesis doctorales y de maestrías, así como en diversas investigaciones, quizás por el desconocimiento que aún subsiste del potencial que brinda la aplicación consecuente de las pruebas de hipótesis, tanto métricas como no paramétricas.

La mayoría de las tesis consultadas utilizan la estadística descriptiva, mediante el empleo de los distintos gráficos: barra, pastel, poligonal, entre otras, y el manejo de las medidas de tendencia central como: promedio, mediana, moda..., en dichos trabajos.

Por otra parte, se pudo constatar en un buen número de tesis consultadas, que no han sido explotadas las pruebas de hipótesis que viabilizan validar la pertinencia de las hipótesis planteadas, fundamentalmente la aplicación de la prueba no paramétrica Kolmogorov Smirnov, que posibilita comparar, con un determinado nivel de confianza asumido por el investigador: el estado final de una misma muestra o dos muestras independientes con su estado inicial, luego de aplicada la propuesta investigativa.

CIYA. Ciencias de la Ingeniería y Aplicadas, vol. 9 n° 2, julio-diciembre de 2025, pp. 80-97

Para ello los autores han tomado como referente la tesis de maestría de [1] defendida en el Instituto Pedagógico Latinoamericano y Caribeño (IPLAC), en La Habana, debido a que en ésta se puede apreciar cómo se aplicó esta prueba estadística objeto de estudio.

Además, se toman como referentes teóricos los distintos clásicos de la Estadística, en especial los trabajos de: [2], [3], [4], [5], [6], [7] y [8], entre otros, para explicar cómo aplicar el estadístico para la prueba de hipótesis Kolmogorov-Smirnov, para dos muestras grandes para dos colas, que permite demostrar que los resultados entre dos grupos independientes investigados fueron diferentes o similares o cuando se desea comparar a un estado inicial con otro final.

También se emplea esta prueba cuando se desea indicar que los resultados de un grupo experimental fueron diferentes a los del grupo o grupos de controles.

Los autores de este trabajo consideran, al igual que: [5], [6], [8] y [9], entre otros, que el empleo de la estadística tiene una gran significación en la investigación científica, ya que la misma requiere de algún tipo de análisis estadístico que posibilite evaluar sus resultados.

Al decir de Montgomery and Runge:

La estadística inferencial es el proceso de uso de los resultados derivados de las muestras para obtener conclusiones acerca de las características de una población. La estadística inferencial nos permite estimar características desconocidas como la medida de la población o la proporción de la población. [6, p. 238]

En algunos casos para resolver un problema de carácter empírico, es preciso llevar a cabo un análisis bastante complejo; otras veces, basta con efectuar un análisis muy simple y directo. La elección de uno u otro tipo de análisis estadístico depende del problema que se plantee en el estudio, así como de la naturaleza de los datos.

En la estadística se pueden estimar parámetros a partir de datos muestrales; sin embargo, con frecuencia, el objetivo de una investigación no es estimar un parámetro sino determinar cuál de dos hipótesis contradictorias acerca de un parámetro es la correcta.

Al igual que: [3], [8] y [10] se considera que los métodos para lograr esto comprenden la parte de la inferencia estadística que recibe el nombre de pruebas de hipótesis.

Se considera que:

Una hipótesis estadística o simplemente hipótesis es una pretensión o aseveración sobre el valor de un solo parámetro (característica de una población o característica de una

CIYA. Ciencias de la Ingeniería y Aplicadas, vol. 9 n° 2, julio-diciembre de 2025, pp. 80-97

distribución de probabilidad), sobre los valores de varios parámetros o sobre la forma de una distribución de probabilidad completa. [3, p. 285]

Por consiguiente, una hipótesis estadística o hipótesis, es una expresión acerca del valor de una sola característica de población o acerca de los valores de varias características de población, que en el ámbito educacional, por citar un ejemplo, podría estar dada por la asunción de que los resultados de un grupo experimental, luego de aplicada determinada metodología de aprendizaje, resultaron superiores a otros grupos (grupo de control) que recibieron la enseñanza de forma tradicional: esto es precisamente lo que se quiere mostrar en este artículo.

Como antecedente de este trabajo, ya hay publicado otro del autor principal [10] en la Revista “Luz” de la Universidad de Holguín, Cuba, similar a éste pero con una cola el cual que puede ser consultado.

Este material sirve para comprobar, mediante la aplicación de la prueba de hipótesis Kolmogorov Smirnov para dos muestras independientes grandes, cómo se puede validar la hipótesis propuesta de que: los resultados emitidos por un grupo de profesores respecto a una serie de indicadores para evaluar la calidad de la educación, fueron similares a los emitidos por el grupo de expertos seleccionados.

Esta metodología está dirigida a los investigadores como instrumento en su trabajo científico que les posibilite validar hipótesis, de aquí su relevancia social.

De los materiales consultados se ha podido constatar que, por lo general, se emplea la prueba Kolmogorov Smirnov cuando se desea realizar una prueba de bondad de ajuste o lo que es lo mismo, para determinar si las distribuciones de los datos de una muestra se comportan como una distribución normal o no.

También puede emplearse para comparar, en un mismo grupo, un estado antes y otro después; es decir, por ejemplo, si los resultados de un estado después han sido superiores a los de un estado antes para un mismo grupo.

Otra variante que se emplea mucho este tipo de prueba de hipótesis cuando se desea comparar los resultados de una propuesta científica entre un grupo experimento con uno o varios grupos de control.

La tercera variante, que es el caso que compete en este trabajo, consiste en comparar los estados de opiniones, criterios o evaluaciones emitidas entre dos grupos independientes para determinar el grado de coincidencia que tienen o no los criterios obtenidos de ambos grupos.

Problemas con cierta similitud a esta propuesta se puede apreciar en el libro de Siegel & Castellán [8]; sin embargo, no se ha observado, con mucha frecuencia, la aplicación de esta metodología aplicada a tesis doctorales y de maestrías

2. METODOLOGÍA

¿En qué consiste una Prueba de Hipótesis?

Los autores consideran conveniente que en lugar de comenzar a dar definiciones matemáticas, que para el lector que no conozca sobre el asunto le resulte incomprensible, explicar cómo emplear, en general, cualquier prueba de hipótesis que establece la estadística; es por ello que se deben seguir los siguientes 8 pasos [6] :

2.1 Metodología para la aplicación de una prueba de hipótesis

- 1) Por el contexto del problema, identificar el parámetro de interés.
- 2) Establecer la hipótesis nula, H_0 .
- 3) Establecer la hipótesis alternativa apropiada, H_1 .
- 4) Elegir un nivel de significación α .
- 5) Establecer un estadístico de prueba apropiado.
- 6) Establecer la región de rechazo del estadístico.
- 7) Calcular las cantidades muestrales necesarias, sustituirlas en la ecuación del estadístico de la prueba y calcular ese valor.
- 8) Decidir si deberá rechazarse o no, H_0 y contextualizar la decisión en el problema. [6]

Ahora se explicará detalladamente cómo contextualizar estos pasos:

1) Identificar los parámetros de interés

Es decir, establecer cuáles son los parámetros que serán objeto de evaluación y análisis estadístico a considerar.

2) Establecer la hipótesis nula, H_0

Siempre la hipótesis de nulidad se refiere a establecer una igualdad entre los criterios obtenidos entre ambos grupos: puede ser mediante la igualdad de promedios, medianas, entre otros [5].

3) Establecer la hipótesis alternativa H_a o H_1

La hipótesis alternativa siempre será contraria o contradecirá a la hipótesis nula, por lo general, es lo que se quiere demostrar; en este caso es considerar que los criterios evaluativos obtenidos por ambos grupos objeto de análisis fueron desiguales.

4) Seleccionar o establecer el grado de significación α o nivel de confianza a asumir

El grado de significación estadística o como algunos autores les llama también nivel de significancia, entre otros, no es más que considerar la probabilidad de cometer un error: cuando se niega una hipótesis cierta, se comete un error del Tipo I o Alfa (α); por el contrario, si se acepta una hipótesis falsa, entonces se incurre en un error del Tipo II o Beta (β) (figura 1) [2], [6].

En la estadística se trabaja preferiblemente las pruebas de hipótesis asumiendo cometer un error del Tipo I, es por ello que el grado de significación Alfa se da en términos probabilísticos en 0,05 (cuando se trabaja con el 95 % de confianza) o 0,01 (cuando se trabaja con un 99 % de confianza), lo que no impide que se puedan trabajar con otros.

En la figura 1 se representan las dos zonas de probabilidades en que se trabajan las dos hipótesis: no perder de vista que el área de probabilidad asignada a la hipótesis de nulidad es mucho mayor (puede ser una probabilidad igual a 0,95).

Sin embargo, la probabilidad **Alfa (α)**, que se le asigna a la zona de rechazo y la que corresponde a la hipótesis alternativa (lo que se desea demostrar), es mucho menor (zona extrema derecha e izquierda de las colas).

- Si la hipótesis alternativa no establece una dirección, en este caso, cuando los valores o criterios entre ambos grupos difieren: se trabaja con dos colas.

Al trabajar con dos colas el valor de alfa asumido se debe dividir por 2 para asignarle a cada cola el valor de $\frac{\alpha}{2}$.

- Cuando en la hipótesis alternativa se establece una dirección:
 - a) Si se considera que los resultados de una muestra fueron superiores al estado inicial de la misma; o si los resultados de un grupo experimento fueron superiores a los del grupo de control: se trabaja con la cola de la derecha.

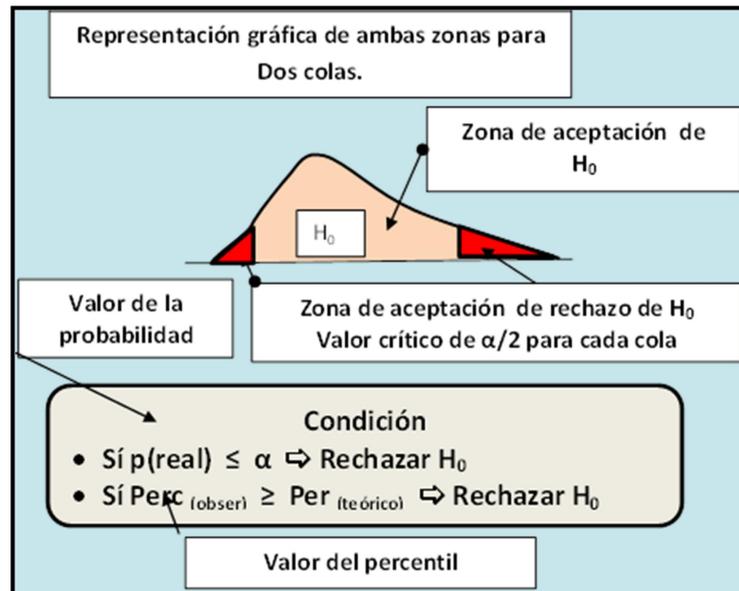


Figura 1. Representación gráfica de las zonas que plantea cada hipótesis y sus correspondientes colas

- b) Cuando los resultados de una muestra fueron inferiores al estado inicial o si los resultados de un grupo experimento fueron inferiores a los del grupo de control: se trabaja con la cola de la izquierda.

En la estadística se trabaja preferiblemente las pruebas de hipótesis asumiendo cometer un error del Tipo I, es por ello que el grado de significación Alfa se da en términos probabilísticos en 0,05 (cuando se trabaja con el 95 % de confianza) o 0,01 (cuando se trabaja

con un 99 % de confianza), lo que no impide que se puedan trabajar con otros niveles de confianza.

Es por ello que en las pruebas de dos colas a cada valor de los extremos de la curva se le asigna el valor de $\alpha/2$.

Acotando lo anterior, se ha establecido dos direcciones al considerar que los criterios de los expertos y profesores difieren.

5) Establecer un estadístico de prueba apropiado

Este es un paso muy importante porque aquí es donde el investigador debe seleccionar o escoger el tipo de prueba de hipótesis aplicar en dependencia de:

- La cantidad de grupos a considerar: un mismo grupo, dos grupos o varios grupos.
- La escala de la variable en la que se está trabajando: dicotómica, dicotómica nominal, ordinal, de intervalo o en la escala de razón.
- El tipo de muestra: intencional (grupo de expertos), aleatoria (la que se sugiere siempre considerar), muestras pequeñas (hasta 30 elementos) y muestras grandes (más de 30 elementos) [2].
- Para el caso que compete se aplicará la prueba de Hipótesis Kolmogorov-Smirnov para dos muestras grandes para dos colas, por las razones que posteriormente serán expuestas:

6) Establecer la región de rechazo para el estadístico

La condición de rechazo se muestra en la ecuación 1, queda establecida en función al tipo de prueba que se aplica, por ejemplo, en el caso que concierne se establece la siguiente [11]:

$$\text{Sí} \quad D_{T_{\text{máx}(\text{observado})}} \geq D_{(\text{tabla calculado})} \Rightarrow \text{Rechazar } H_0 \text{ en favor de } H_1$$

(1).

Donde:

$D_{T_{\text{máx}(\text{observado})}}$: Diferencia máxima obtenida entre los dos grupos comparados de los resultados de sus probabilidades acumuladas.

$D_{(\text{tabla calculada})}$: Diferencia de la tabla del estadígrafo Kolmogorov-Smirnov en función del tamaño de la muestra (grande o pequeña), grado de significación Alfa asumido y número de elementos considerados en las muestras.

7) Calcular las cantidades muestrales necesarias, sustituirlas en la ecuación del estadístico de la prueba y calcular ese valor

Es importante acotar que en dependencia del tipo de prueba de hipótesis lógicamente será correspondido con una fórmula o estadístico relacionado con el mismo, en este caso la expresión a trabajar será la que se observa en la ecuación 2, [8]:

$$D_{(tabla\ calculado)} = 1.36 \cdot \sqrt{\frac{(n_1 + n_2)}{(n_1 \cdot n_2)}} (I) \quad (2)$$

8) Decidir si deberá rechazarse o no H_0 y contextualizar la decisión en el problema

En este paso se determina si se rechaza o no la hipótesis de nulidad en favor de la alternativa, por lo que el investigador establece el grado de pertinencia y validez de la hipótesis de su tesis.

3. ANÁLISIS DE RESULTADOS

Moraguez [1] en su tesis de maestría elaboró una propuesta de dimensiones e indicadores que posibilitan evaluar el impacto de las escuelas politécnicas en la provincia de Holguín, Cuba, establecidos en 5 dimensiones y 24 indicadores, con su correspondiente operacionalización, los que fueron sometidos al criterio o consenso de 35 expertos de todo el país, mediante la aplicación del Método Delphi: [12], [13], [14].

Resulta necesario hacer un paréntesis y dar una breve explicación en qué consiste el Método Delphi, que no es más que obtener los criterios u opiniones informadas de un grupo de especialistas que son elegidos selectivamente en función del grado de conocimiento que tienen acerca de un determinado objeto de estudio a los cuales se les solicitan opiniones sobre determinado asunto.

Es imprescindible aclarar que para la selección de los posibles expertos resulta necesario elaborar un listado de los posibles expertos, algunos autores les llaman “bolsa de posibles expertos” [13], [15], [16] relacionados con el objeto de estudio investigado.

Estos posibles expertos son preseleccionados teniendo en cuenta los años de experiencia y los conocimientos adquiridos acerca de la temática objeto de estudio a los cuales se les envía vía electrónica un instrumento a través del cual se le informa de la investigación que se llevará a cabo y se les conmina a participar en ella: es importante que el posible experto tenga disposición de cooperar con la investigación.

Además, en este documento se les pide que deben de completar una serie de preguntas a fin de que ellos mismos se autoevalúen: el grado de conocimiento que tienen sobre el tema y cómo fue que adquirieron dichos conocimientos: estudio de autores, elaboración de trabajos científicos acerca del tema, entre otros, se recomienda consultar el artículo [13], [14].

Una vez recogidas las informaciones emitidas por los expertos acerca del grado de conocimientos que éstos consideran tienen sobre el tema, se procede a contabilizar las

CIYA. Ciencias de la Ingeniería y Aplicadas, vol. 9 n° 2, julio-diciembre de 2025, pp. 80-97

puntuaciones obtenidas y todos aquellos que obtuvieron una puntuación mayor o igual de 0,75 puntos (el máximo es 1) se considerará como experto.

A partir de este momento se conforma el grupo de expertos los cuales serán consultados en varias oportunidades en aspectos relacionados con el objeto de estudio del investigador.

En el caso que compete se elaboró la propuesta teniendo en cuenta los criterios de los expertos en dos vueltas.

Por otra parte fueron seleccionados al azar a un grupo de 140 profesores de la Educación Técnica y Profesional de la provincia de Holguín, Cuba, a los cuales se les aplicó el mismo instrumento que les fue aplicado a los expertos, a fin de comprobar si existía o no diferencias de criterios acerca de los indicadores y dimensiones propuesta en la investigación.

Luego se compararon ambos resultados de los 140 profesores, seleccionados al azar, de la Educación Técnica y Profesional (ETP en lo adelante) que laboraban en las distintas escuelas politécnicas de la provincia antes citada y el grupo de 35 expertos de todo el país, a fin de analizar el grado de coincidencia que tenían los criterios emitidos por ambos grupos relacionados con las dimensiones e indicadores propuesto por el investigador.

Las dimensiones e indicadores fueron categorizados en escala ordinal en las categorías de: Indispensable para medir la variable (I), Muy útil para medir la variable (MU), Útil... (U), Quizás sirve... (Q) y No sirve... (NS).

Es importante acotar que las cinco categorías ordinales establecidas son las preferidas para la aplicación del método Delphi o Método de Consulta de Expertos [10], [12], [14].

Las categorías se establecieron tomando en cuenta la operacionalización que se hicieron de las dimensiones e indicadores (variables), las cuales fueron entregadas a todos los encuestados para que tuvieran una mayor claridad a la hora de emitir criterios acerca de éstos [1, p. Anexo 4 hasta 5.5].

Los criterios fueron recogidos mediante una encuesta a ambos grupos y procesados estadísticamente a través del empleo o aplicación de la prueba de hipótesis: la Prueba Kolmogorov-Smirnov para dos muestras grandes con dos colas, que permitió demostrar que los criterios emitidos por ambos grupos no resultaron diferentes, lo que denota la pertinencia de la propuesta.

Este tipo de prueba es recomendada por Siegel & Castellan [8] en el libro Estadística no paramétrica aplicada a las Ciencias de la Conducta.

Se debe tener en cuenta el tipo de escala a que obedecen estos parámetros para con ello determinar el tipo de prueba de hipótesis a aplicar.

4. RESULTADOS Y DISCUSIÓN

4.1 Aplicación del estadígrafo

Se explica a continuación cómo aplicar esta prueba de hipótesis a través de un ejemplo tomado de la tesis de maestría de Moraguez [1].

El investigador aplicó el método de consulta a expertos a 35 de ellos tomados de todo el país y con este mismo cuestionario encuestó a 140 profesores del territorio, seleccionados al azar con más de 15 años de experiencia promedio, solicitándoles el criterio que tenían acerca de un conjunto de dimensiones (5) e indicadores (25) (propuesto por el autor) que permiten evaluar el impacto educacional, como indicador de calidad educacional.

De todos ellos se tomó como referente en este trabajo la dimensión I. Efecto Académico con el indicador 1.1.2. Nivel de actualidad técnica y tecnológica sobre la base de evaluarlos en las categorías de Indispensable (I), Muy Útil (MU), Quizás Sirve (Q) y No Sirve (N) para medir la variable.

1) Por el contexto del problema, identificar el parámetro de interés

Primeramente, se elaboró una tabla de frecuencia absoluta, pero teniendo en cuenta ordenar la escala en orden creciente (se recomienda siempre hacerlo así).

Tabla 1. Frecuencia absoluta de los criterios de los expertos y profesores encuestados

Frecuencia absoluta						
Indicador	N	Q	U	MU	I	Total
Profesores	0	2	8	40	90	140
Expertos	0	0	0	6	29	35

Nota: Fuente: Moraguez [1, Anexos 13, 28]

Observe que lo primero que se hace es colocar los indicadores evaluativos en orden creciente, en este caso desde No Sirve (N) hasta Indispensable (I) como se muestra en la tabla 1.

Se aclara que, aunque en este problema las muestras tomadas son desiguales, esto no siempre puede ser así, pues se pueden tomar de tamaños iguales; en el caso que una muestra sea

grande y la otra pequeña, que el caso que compete, a la hora de aplicar el método se debe considerar como prueba para muestras grandes.

2) Establecer las hipótesis de nula, H_0

H_0 : Hipótesis de nulidad: Los criterios entre los expertos y profesores del territorio en relación con los indicadores propuestos no difieren.

3) Establecer la hipótesis alternativa apropiada, H_1

Recordar que la hipótesis alternativa es aquella hipótesis que se quiere demostrar y se denota por H_1 , otros autores la denotan por H_a . A ésta se le asigna el grado de significación Alfa (error del Tipo I) y además determina el tipo de cola a trabajar.

Si por ejemplo, se quiere demostrar que los resultados de un grupo experimental fueron superiores a los de un grupo de control (o varios grupos de controles), se considera que se está trabajando con una dirección: los resultados, por ejemplo, de la media μ_2 del grupo experimental fueron superiores a la media del grupo de control (μ_1); luego se discurre en esta hipótesis que: $\mu_2 > \mu_1$ y ello implica trabajar con una cola.

Por otra parte, si se creyera que los resultados del grupo de experimento son diferentes a los del grupo de control; ello infiere que los valores de la media son diferentes ($\mu_2 \neq \mu_1$) y por tanto pudiera ser que: $\mu_2 > \mu_1$ o $\mu_2 < \mu_1$, lo que indica trabajar con dos colas, (aspecto que a tratar en este trabajo).

En el ejemplo tomado se tiene que la hipótesis alternativa se asumió como:

H_1 : Hipótesis alternativa, (contradice a la hipótesis de nulidad): Los criterios entre los expertos y profesores del territorio en relación con los indicadores propuestos difieren o son desiguales.

Observe que, en este caso, la hipótesis alternativa no establece una dirección (los criterios difieren, es por ello que se trabaja con dos colas).

4) Elegir un nivel de significación α

Moraguez [1], consideró en su tesis trabajar, con un 95 % de confianza (0,95), por consiguiente, el grado de significación Alfa (α) es igual a 0,05; que es la probabilidad de cometer un error del Tipo I (rechazar una hipótesis cierta).

Alfa (α) = 0,05 (Cometer un error del Tipo I: Negar una hipótesis cierta en un 5%), por lo tanto le corresponde a cada cola el valor de $\alpha/2 = 0,25$

5) Establecer un estadístico de prueba apropiado

Se aplicó la prueba de Hipótesis Kolmogorov-Smirnov para dos muestras grandes para dos colas, por las siguientes razones:

- I. Esta es una prueba de bondad de ajuste; esto es, se interesa en el grado de acuerdo entre la distribución de un conjunto de valores de la muestra (puntajes observados, que pueden estar en escala ordinal, de intervalo o de razón) y alguna distribución específica. Determina si razonablemente pueden pensarse que los puntajes en la muestra provengan de una población que tenga distribución teórica.
- II. Es una prueba muy poderosa [3], [5], [8], entre otros, recomendada cuando se desea comparar un estado inicial (antes) con un estado final (después).
- III. Es una prueba no paramétrica muy recomendada para utilizar con escalas ordinales o de intervalo, en este caso la escala es ordinal por las categorías: N, Q, U, MU, e I.
- IV. Se utiliza para muestras grandes cuando el tamaño de la misma excede de 40 elementos, Moraguez [1] tomó una de 140 profesores y 35 expertos (tabla 1). Se establece que para muestras grandes ambas pueden tener el mismo o diferentes tamaños, como es el caso que compete.
- V. Cuando se quiere demostrar que los resultados de un grupo experimental fueron superiores a los de un grupo de control o varios grupos de controles se trabaja con una cola o también comparando, para un mismo grupo, un estado inicial con otro final; otro caso de empleo puede ser al comparar los resultados entre dos grupos.

6) Establecer la región de rechazo para el estadístico

La condición de rechazo que establece esta prueba es la siguiente:

$$D_{T_{\text{máx}}(\text{observado})} \geq D_{(\text{tabla calculado})} \Rightarrow \text{Rechazar } H_0 \text{ en favor de } H_1 \quad (3)$$

7) Calcular las cantidades muestrales necesarias, sustituirlas en la ecuación del estadístico de la prueba y calcular ese valor

Siegel & Castellan [8] plantean que el estadístico $D_{(\text{tabla calculado})}$ se determinará mediante la siguiente expresión:

$$D_{(\text{tabla calculado})} = 1,36 \cdot \sqrt{\frac{m+n}{mn}} \quad (4)$$

Donde:

$D_{(\text{tabla calculada})}$: Es el valor calculado por la expresión I) tomada de la tabla 2 que está en función del valor Alfa asumido para dos colas.

m: Tamaño de la muestra 1.

n: Tamaño de la muestra 2.

Si ambos tamaños son iguales entonces la expresión se puede expresar en términos de m o n, según considere el investigador. Fuente: Siegel and Castellan [8].

Tabla 2. Valores críticos de $D_{(m, n, \alpha)}$ para dos muestras grandes

Tabla L_{III}. Valores críticos de $D_{m,n}$ para la prueba de Kolmogorov-Smirnov para dos muestras (muestras grandes, bidireccional).*

Nivel de significación	Valor de $D_{m,n}$ que indica el rechazo de H_0 en el nivel de significación indicado, donde $D_{m,n} = \text{máximo } S_m(X) - S_n(X) $
0.10	$1.22 \sqrt{\frac{m+n}{mn}}$
0.05	$1.36 \sqrt{\frac{m+n}{mn}}$
0.025	$1.48 \sqrt{\frac{m+n}{mn}}$
0.01	$1.63 \sqrt{\frac{m+n}{mn}}$
0.005	$1.73 \sqrt{\frac{m+n}{mn}}$
0.001	$1.95 \sqrt{\frac{m+n}{mn}}$

* Adaptada de Smirnov, N., "Tables for estimating the goodness of fit of empirical distributions", en *Annals of Mathematical Statistics*, núm. 19, 1948, págs. 280-281, por cortesía del editor.

Los pasos a seguir para determinar este estadígrafo son los siguientes:

- Elaborar la tabla de frecuencias absoluta de los parámetros de interés, observe la tabla 1, página 9.
- Calcular la matriz de frecuencia relativa acumulada (Tabla 3), tomando como referente la tabla 1 anterior:

Tabla 3. Frecuencia relativa acumulada.

Indicador	N	Q	U	MU	I
Profesores (S_{x1})	0,0000	0,0143	0,0714	0,3571	1,0000
Expertos (S_{x2})	0,0000	0,0000	0,0000	0,1714	1,0000
$D/(S_{x1}) - (S_{x2})/$	0,0000	0,0143	0,0714	0,1857	0,0000
$D_{(T\text{máx}(\text{observado}))}$	0,19				

Fuente: Moraguez [1],

CIYA. Ciencias de la Ingeniería y Aplicadas, vol. 9 n° 2, julio-diciembre de 2025, pp. 80-97

Se debe acotar que la frecuencia acumulada se obtiene de ir sumando la frecuencia absoluta del nivel precedente e irle sumando las frecuencias absolutas de los niveles que les suceden de manera tal que el último valor acumulado debe coincidir con el tamaño de la muestra.

La frecuencia relativa no es más que el cociente obtenido entre cada una de las frecuencias absoluta entre el total de la muestra; cuando este valor se multiplica por 100 se convierte en valor porcentual.

En la Tabla 3 para calcular el valor de la frecuencia relativa acumulada de la primera celda (0,000), resulta de dividir la primera frecuencia absoluta (Tabla 1) de la evaluación de “N” de los profesores S_{x1} (0), entre el total de la muestra n_1 (140), $0/140 = 0,000$.

De igual forma se procede para determinar la frecuencia relativa de los evaluados de “N” de los expertos S_{x2} (0) (segunda celda debajo del valor 0,000); se obtiene el cociente 0,000 que resulta de dividir los cero casos con “N” de los expertos entre el total de ellos o de la muestra n_2 , que es de 35; luego: $0/35 = 0,000$.

Ahora se obtendrá el siguiente valor de la frecuencia relativa acumulada para los casos evaluados de “Q” emitidas por los profesores (S_{x1}), que en este caso es igual a 2 casos, por lo que resulta que a la frecuencia relativa anterior 0,000 se le debe sumar el cociente $2/140 = 0,0143$; luego $0,000 + 0,0143 = 0,0143$.

De manera análoga se procede a calcular todos los valores de cada una de las celdas y se completa la tabla.

Finalmente para determinar el valor acumulado de la última celda de la Tabla 2 resulta de sumarle la frecuencia acumulada anterior (0,3571) con el cociente que se obtiene de dividir la frecuencia absoluta de los profesores que consideraron de Indispensable este indicador (90) entre el total de profesores de esta muestra (140), resultando la operación: $0,3571 + 90/140 = 0,9999$, (se recuerda que como se trabaja con valores decimales la frecuencia acumulada relativa final casi siempre no llega 1, pero si a un valor muy próximo, por lo tanto se asume igual a 1 por ser la máxima probabilidad de que ocurra un suceso.

Es importante recordar que el último valor de la frecuencia relativa acumulada para ambos casos debe ser igual a la unidad por ser esta la máxima probabilidad de que ocurra un suceso, también conocida como primera ley o propiedad de las probabilidades.

De manera análoga se procede para encontrar el resto de los valores de la tabla 3 para el caso de los criterios emitidos por los expertos según se puede apreciar.

- c) **Determinar el valor absoluto de las diferencias entre los estados “Antes” y “Después” o de comparar el grupo de profesores con el de expertos.**

Una vez calculados los valores de las frecuencias relativas acumuladas de ambos casos para cada variable (N, Q, U, MU, I), se calculan las diferencias absolutas entre los estados “Antes” y “Después”, que en este caso para los criterios emitidos entre los profesores y expertos

En la penúltima fila de la Tabla 3 se puede apreciar las diferencias absolutas entre las variables $S_{(x1)}$ y $S_{(x2)}$.

- d) **Seleccionar el valor máximo de estas diferencias**

Como se aprecia en la Tabla 3 el valor máximo de los valores absolutos de las diferencias observadas es igual a 0,1857, que se aproxima al valor de 0,19, luego este valor se denota por:

$$D_{TM\acute{a}x(observado)} = 0,19$$

Donde: $D_{TM\acute{a}x(observado)}$: Valor de la tabla correspondiente a la diferencia máxima observada.

Ya se tiene el valor de $D_{TM\acute{a}x(observado)}$, así como el de m (140) y n (35), por lo que simplemente hay que sustituir y calcular el valor del estadístico en la expresión (I), luego:

$$\begin{aligned} D_{(tabla\ calculado)} &= 1,36 \cdot \sqrt{\frac{m+n}{mn}} \quad (tabla\ 2) \\ &= 1,36 \cdot \sqrt{\frac{140+35}{140 \cdot 35}} = 0,257 \end{aligned}$$

$$D_{(tabla\ calculado)} = 0,257$$

- 8) Decidir si deberá rechazarse o no, H_0 y contextualizar la decisión en el problema**

Siempre hay que recordar que la condición para rechazar la hipótesis de nulidad en favor de la alternativa es que se cumpla la condición:

$$\text{Sí } D_{Tm\acute{a}x(observado)} \geq D_{(tabla\ calculado)} \Rightarrow \text{Rechazar } H_0 \text{ en favor de } H_1$$

Como $D_{Tm\acute{a}x(observado)} < D_{(tabla\ calculado)}$, debido a que la probabilidad observada 0,19 es menor que la probabilidad de la tabla 0,257 ello implica que no cae en la zona de rechazo H_0 y por consiguiente hay que aceptarla; por lo tanto, se puede presuponer, con un 95 % de confianza, que los criterios emitidos entre los profesores y los expertos acerca de la propuesta del

indicador 1.1.2 Efecto académico coinciden, lo que significa que hubo una valoración coincidente entre los expertos y profesores de experiencia, con relación a las dimensiones e indicadores propuestos en la tesis, que es lo que el investigador quería demostrar.

5. CONCLUSIONES

Resulta sustancial resaltar la importancia que tiene la utilización de las pruebas de hipótesis estadísticas que permiten validar hipótesis con mayor confianza, significando el rol que desempeña la utilización de esta poderosa prueba de hipótesis no paramétrica, que posibilita comparar, con un determinado nivel de confianza asumido por el investigador: el estado final de una misma muestra o dos muestras independientes con su estado inicial, luego de aplicada la propuesta investigativa.

Se recomienda el empleo de esta prueba no paramétrica cuando se desea comparar dos muestras grandes o una misma muestra en dos momentos: un estado antes con otro después de aplicada la propuesta investigativa.

6. BIBLIOGRAFÍA

- [1] A. Moraguez, «La evaluación de la eficiencia externa en las Escuelas Politécnicas Industriales de la provincia de Holguín.», Tesis en opción al título de Máster en Ciencias de la Maestría en Planeamiento, Administración y Supervisión de Sistemas Educativos, Instituto Latinoamericano y Caribeño IPLAC., La Habana, Cuba, 2001.
- [2] J. L. Devore, Probabilidad y estadística para ingeniería y ciencias. México: THOMSON EDITORES, 2000.
- [3] J. L. Devore, Probabilidad y estadística para ingeniería y ciencias, Séptima Edición. México: Thomson Editores, 2008. [En línea]. Disponible en: <https://intranetua.uantof.cl/facultades/csbasicas/matematicas/academicos/jreyes/DOCENCIA/APUNTES/APUNTES%20PDF/Probabilidad%20y%20Estadistica%20para%20Ingenieria%20y%20Ciencias%20-%20Jay%20Devore%20-%20Septima%20Edicion.pdf>
- [4] J. Freud, Estadística elemental moderna. La Habana. Cuba, 1977.
- [5] C. Montgomery y G. C. Runge, «Probabilidad y Estadística aplicadas a la Ingeniería», Edit. McGraw-Hill, Ed., California (Impreso en México), 1996, p. 995 p.
- [6] D. Montgomery y G. Runge, Probabilidad y Estadística aplicadas a la Ingeniería, Segunda Edición. México: Edit. McGraw-Hill, 2003. [En línea]. Disponible en: https://www.academia.edu/34899097/Montgomery_y_Runger_Probabilidad_y_Estadistica_Aplicada_a_La_Ingenieria
- [7] S. Siegel, Diseño experimental no paramétrico. La Habana: Edit Rev., 1987.

CIYA. Ciencias de la Ingeniería y Aplicadas, vol. 9 n° 2, julio-diciembre de 2025, pp. 80-97

- [8] S. Siegel y J. Castellan, Estadística no paramétrica aplicadas a ciencias de la conducta, Primera re. Angewandte Chemie International Edition, 6(11), 951–952. Eitorial Trillas, 1998. [En línea]. Disponible en: https://pauyecologia.wordpress.com/wp-content/uploads/2016/11/estadisticas_no_parametricas-siegel5b15d-1.pdf
- [9] C. Sánchez, «Pruebas de hipótesis no paramétricas de Kolmogorov-Smirnov», Monografías.com. Accedido: 20 de enero de 2025. [En línea]. Disponible en: <https://www.monografias.com/trabajos11/docima/docima>
- [10] A. Moraguez, M. de P. Espinosa, y L. Morales, «La prueba de hipótesis Kolmogorov-Smirnov para dos muestras grandes con una cola.», Luz, vol. 16, n.o 3, pp. 77-89, 2017.
- [11] S. Siegel, Estadística no paramétrica aplicada a las ciencias de la conducta. México, 1975.
- [12] M. J. J. Jiménez, M. V. A. Hidalgo, y M. A. B. Villarreal, «Método neutrosófico para evaluar los factores de riesgo de obesidad en adolescentes del barrio Vivienda Popular Tulcán 2023», Neutrosophic Comput. Mach. Learn. ISSN 2574-1101, vol. 31, pp. 137-148, ene. 2024.
- [13] A. Moraguez, «El método Delphi.», Luz, vol. 5, n.o 4, 2006, Accedido: 7 de enero de 2025. [En línea]. Disponible en: <https://www.redalyc.org/pdf/5891/589165888006.pdf>
- [14] A. Moraguez y J. R. García, «El método Delphi, un análisis académico al respecto.», Rev. Dimens. Empres. Fac. Cienc. Adm. Económicas Contab. Univ. Autónoma Caribe, vol. 7, n.o 11, p. 15, 2009.
- [15] L. Capistrós y C. Rizo, «Indicadores e investigación educativa.», 23 de agosto de 1998.
- [16] J. Díaz, «Prueba Kolmogorov-Smirnov», Enigmáticamente (JSD). Accedido: 10 de febrero de 2023. [En línea]. Disponible en: file:///C:/Users/Arabel/Downloads/Paquete_estadistico_SPSS_Prueba_Kolmogo.pdf