

CE 02 1203

[♪ Música de presentación ♪ Revista de Educación, un tiempo ahora para Pedagogía Experimental II, otra asignatura de segundo curso. Vamos a hablar de la opción no paramétrica en investigaciones educativas, un tema complejo del que nos va a hablar el profesor Dionisio del Rey Osadornil. Buenas noches, profesor.

Buenas noches. Si me parece, comenzamos con una introducción al tema, profesor, porque estas cuestiones son un poco difíciles para los alumnos. En una primera aproximación al tema de la opción no paramétrica en investigaciones educativas, quisiera esbozar dos ideas preliminares.

La primera está referida a la medida de los fenómenos educativos. Y aquí quisiera aclarar el concepto de medida, que en educación es más amplio y, por tanto, menos riguroso que el utilizado en ciertas magnitudes físicas en las que se reduce a comparar la cantidad con la unidad, siendo el número las veces que esta, la unidad, se encuentra contenida en aquella, la cantidad. En temas educativos, debemos medir variables difícilmente accesibles, es decir, no observables directamente.

Por ejemplo, la inteligencia y todas aquellas variables relacionadas con las operaciones internas de la persona. En consecuencia, al ser muy difícil elaborar una unidad de medida de la misma naturaleza y constante en su valor, no tenemos más remedio que identificarnos con un concepto más amplio de medida, o, en verdad, no podríamos medir fenómenos educativos. Lo que nos lleva a concluir que los niveles de medición más frecuentes en ciencias de la educación son el nominal, concretado en el recuento de frecuencias por clases de objetos, individuos, etc., y el ordinal, que indica el puesto rango que un sujeto o la cuantía de un dato ocupa en un grupo respecto de una característica, distinción cuantitativa.

Nos encontraremos con no pocas dificultades para alcanzar una escala de intervalos satisfactoria, la que, por tener unidad constante de medida, permite intervalos iguales y donde, además, el cero no es absoluto. Y para acabar esta idea de los niveles de medida, podemos afirmar que hay muy pocas variables de importancia en educación que den pie para hacer medidas según escala de razón, proporción o cociente, que añade a la escala de intervalo el cero absoluto, es decir, carencia total del atributo medido. La otra idea preliminar era la referida al principio básico del uso de las pruebas estadísticas de significación, comparación de los resultados obtenidos con los que cabría esperar por la sola intervención del azar, es decir, con los que, en teoría, se esperaban, porque, no lo olvidemos, todo buen estadístico es un escéptico y, hasta donde cabe, un probabilista inveterado que supone que los resultados empíricos se deben al azar hasta que prueba lo contrario, aplicando pruebas estadísticas.

Pues bien, en este contraste con pruebas estadísticas es donde surge la polémica y hasta las posturas encontradas al referirse al nivel de medida logrado como condicionante de la utilización de pruebas paramétricas o no paramétricas. En este punto, si le parece, profesor del

Río, yo creo que es importante que, antes de seguir adelante, nos aclare brevemente estas expresiones, pruebas paramétricas y pruebas no paramétricas, porque, seguramente, están relacionadas con la construcción de modelos o de opciones de investigación empírica o experimental, y es importante que tengamos, desde el principio, la idea clara de qué es una prueba paramétrica y qué es una prueba no paramétrica, profesor del Río. A las dos partes de su pregunta puedo responder conjuntamente.

En ciencias sociales y, por tanto, en educación, hoy se habla de dos paradigmas. El cualitativo, o idealista, naturalista, en el que aquí no vamos a entrar, y el cuantitativo o realista, también denominado positivista, caracterizado por la búsqueda de causas de los fenómenos sociales, pero con escasa atención a los estados subjetivos de los individuos. Aporta, por tanto, explicaciones científicas de naturaleza causal al subordinar casos particulares a leyes generales.

Pues bien, dentro de este paradigma cuantitativo, que es hipotético-deductivo-experimental, y que utiliza una medición penetrante y controlada de los fenómenos, caben dos formas de construir modelos. Primera, formulando supuestos sobre la naturaleza de la población de la que se ha tomado la muestra, y, en concreto, sobre las características o parámetros de la población. Estamos, pues, en este caso, ante los métodos inferenciales paramétricos.

De ahí la denominación de paramétricas, que será a las pruebas utilizadas en este modelo u opción. Esos supuestos sobre las poblaciones y sus parámetros o características se resumen en a. Garantía de independencia de cada observación o dato respecto de los demás. b. Distribución normal de las poblaciones de las que se obtuvieron las muestras.

c. Tener varianzas iguales las poblaciones origen de las muestras, es decir, igualdad de dispersión, lo que técnicamente se llama homoscedasticidad. y d. Nivel de medición de intervalo o de razón. La segunda forma o modelo sería la utilización de analogías entre los conjuntos de datos y las fórmulas matemáticas existentes, como, por ejemplo, el desarrollo del binomio o las leyes de probabilidad.

Este otro conjunto de métodos inferenciales no parte de ningún supuesto relativo a las poblaciones, como estábamos en la forma o modelo anterior, y no considera que los estadísticos o valores representativos de las muestras reproduzcan o representen a los parámetros, o sea, a los valores de la población. Por ello, se denominan métodos inferenciales no paramétricos. Un nombre mejor para la estadística no paramétrica es, o sería, estadística sin distribución, porque los métodos no paramétricos se aplican a todas las distribuciones en vez de adaptarse solo a las de cierta forma particular.

Otros hablan de métodos de distribución libre, puesto que no importa la forma de la distribución del universo o población del que procede la muestra o las muestras sometidas a la prueba estadística. Antes de esta aclaración que nos ha hecho sobre el significado de las opciones paramétrica y no paramétrica, nos hablaba de que existe una polémica con posturas enfrentadas en relación con el nivel de medida logrado como condicionante de la utilización de pruebas paramétricas y no paramétricas. ¿Podría explicarnos estas dos posturas, profesor del

Río? Sí, existen dos posiciones extremas en cuanto a la repercusión de las cuatro escalas de medida en el tipo de pruebas estadísticas a utilizar, que, recordémoslo, eran cuatro escalas de medida, nominal, ordinal, de intervalo y de razón.

En primer lugar, tenemos la postura dogmática, defendida por Stevens, Senders y Siebel, dicen estos autores. Resumo, las mediciones nominales y ordinales son las realizadas más comúnmente en las ciencias de la conducta, por tanto, en educación. Los datos medidos por escalas nominales y ordinales deben analizarse por métodos no paramétricos, ya que no exigen medidas más precisas que las mencionadas.

Los datos medidos con escalas de intervalo o de proporción, o también llamada a este nivel de medida razón o cociente, deben analizarse por métodos paramétricos, si los supuestos del modelo estadístico paramétrico son sostenibles. Esta posición dogmática advierte, no obstante, de que, al ser problemático en ciencias de la conducta el logro de una escala de intervalo satisfactoria y sumamente difícil la consecución de una escala de razón o proporción, las pruebas paramétricas difícilmente podrán aplicarse para probar hipótesis. Vamos con la segunda postura, la pragmática, defendida por Anderson, Kaiser, Lohr, Glass-Stanley, Kerlinger y Amon.

Parten los defensores de esta postura pragmática o flexible del supuesto de que cualquier conjunto de números es susceptible de procesarse estadísticamente mediante cualquier técnica, con algunas limitaciones, evidentemente, pero, dicen sus defensores, no conocemos ninguna técnica que no pueda emplearse en razón de que los números que se le han escrito no son los idóneos, no son los indicados. Desde esta posición flexible y abierta se adoptaría un compromiso razonable, usar aquellas técnicas estadísticas que creamos más apropiadas en cada caso y que nos puedan llevar a conclusiones pedagógicas razonables. La experiencia parece confirmar que podemos llegar a resultados numéricos interpretables pedagógicamente, aunque las técnicas estadísticas utilizadas no hayan sido las más oportunas según el nivel de medición requerido por los datos.

Como el profesor Dionisio del Río Sabernil, estamos compartiendo un espacio de Pedagogía Experimental II dedicado a la opción no paramétrica en investigaciones educativas. Hasta ahora hemos hecho una introducción para que ustedes se sitúen y, a partir de ahora, ya nos vamos a centrar en nuestro tema. La opción no paramétrica en investigaciones educativas.

¿Qué parece, profesor? Así es. Al respecto, habría que comenzar recordando los criterios de una buena elección de prueba, sea esta paramétrica o no paramétrica. Las exigencias de una prueba paramétrica se resumen en ciertos supuestos sobre las poblaciones y sus parámetros, supuestos que no son exigibles en las pruebas no paramétricas.

Estos son los supuestos más corrientes, y repito porque creo haberlo dicho anteriormente, pero, como es importante, no me importa repetirlo. Primero, selección aleatoria, es decir, al azar, de la muestra que garantice la independencia de cada observación o dato respecto de los demás. Segundo, distribución normal de las poblaciones de las que se obtuvieron las muestras.

Tercero, tener varianzas iguales las poblaciones origen de las muestras. Decíamos que esto recibe el nombre técnico de homoscedasticidad o igual dispersión. Cuarto, nivel de medición de intervalo o de razón.

Entonces, profesor, si se dieran estos supuestos, ¿se preferiría una prueba paramétrica a otra no paramétrica? Por supuesto, nos decidiríamos por una prueba paramétrica, ya que cualquier contraste paramétrico tiene mayor potencia o eficiencia estadística, es decir, probabilidad mayor de rechazar hipótesis de nulidad falsas, $1 - \text{menor beta}$, o a igual potencia, la prueba paramétrica permite el ahorro de sujetos en las muestras. Asimismo, su desarrollo estadístico es superior, lo que permite proseguir los análisis más allá que en el caso de las pruebas no paramétricas. Finalmente, las pruebas paramétricas permiten hacer estimaciones de los parámetros o valores o características de la población.

Ahora bien, en las numerosísimas investigaciones educativas en que no se dan los supuestos descritos exigidos por las pruebas paramétricas, no nos quedará otra opción que las pruebas no paramétricas. No obstante, la gran aplicabilidad de estas pruebas no paramétricas la pagan con una menor potencia respecto de las pruebas paramétricas homólogas, para rechazar hipótesis de nulidad falsa. Y ya de lleno en la opción no paramétrica, parece oportuno que nos indique las ventajas de las pruebas no paramétricas.

Las ventajas más destacables de las pruebas no paramétricas se refieren a, en primer lugar, un más fácil y rápido cálculo, por regla general. En segundo lugar, unos mínimos supuestos de sus respectivos modelos, lo que conlleva aplicabilidad con niveles de medida ordinal y nominal. Y también aplicabilidad con muestras procedentes de distintas poblaciones, al no requerir los supuestos poblacionales de las pruebas paramétricas.

Por tanto, son útiles a un nivel de significación determinado, cualquiera que sea la forma de la distribución de la cual procedan las muestras. Y, además, son adecuadas para muestras pequeñas o muy pequeñas, incluso con datos o número de sujetos inferiores a 10, en las que suelen ser tan potentes como las pruebas paramétricas, sin olvidar que, en ocasiones, son la única alternativa. Bueno, pues ya que hemos visto las ventajas, ¿y le parece que vamos a ver los inconvenientes de las pruebas no paramétricas? Efectivamente, tienen sus inconvenientes, como también los tienen las pruebas paramétricas, y a los que ya aludíamos resumiéndose, en las pruebas paramétricas, digo, en la desventaja de su menor aplicabilidad al exigir el cumplimiento de una serie de supuestos referentes a las poblaciones de las que se toman las muestras.

Pero, bien, vamos con los inconvenientes de las pruebas no paramétricas, que se sintetizan en su menor potencia o eficiencia estadística, lo que conlleva precisar mayor número de datos o de sujetos que las pruebas paramétricas para alcanzar la potencia de éstas. Otro inconveniente es su desarrollo estadístico menor. Profesor del Río, si le parece, nos podría dar algunos criterios para la elección entre ambos tipos de pruebas paramétricas o no paramétricas, unos criterios que sintetizasen, o más bien que nos den una idea clara de todo lo que hemos dicho

hasta ahora, que nos sirva de resumen, profesor del Río.

Sí, los criterios de elección entre ambos tipos de contraste se refieren a tres aspectos fundamentales, y que voy a explicar por separado. Potencia, modelo estadístico y nivel de medida. Vamos con el primer aspecto, criterio de potencia.

Según el criterio de potencia, en igualdad de condiciones, el investigador debe elegir la prueba más potente, es decir, aquella con la que pueda rechazar más fácilmente hipótesis nulas, hipótesis de nulidad falsas. Estoy repitiendo la denominación de hipótesis nulas, que quiere decir hipótesis de no diferencia, que son las hipótesis contrarias a las de investigación. Pues bien, las pruebas paramétricas resultan ser más potentes que las correspondientes no paramétricas, esto en cuanto al criterio de potencia.

B, según el modelo estadístico, las condiciones aludidas en el criterio de potencia son las exigidas por cada modelo estadístico. En el caso de las pruebas paramétricas, ya vimos que las condiciones para su utilización eran mucho más exigentes y restrictivas que en el caso de las no paramétricas. En tercer lugar, C, según el nivel de medición, las pruebas paramétricas exigen una medición a nivel de intervalo como mínimo, o a nivel de intervalos como mínimo, mientras a las no paramétricas les bastan las mediciones ordinales o incluso nominales.

Bien, yo creo que con esta sistematización podemos tener más claro el tema hasta aquí y podemos seguir avanzando, porque probablemente, profesor del Río, quepa la discusión ahora en torno a estos tres criterios que nos ha comentado para elegir la prueba más adecuada. Por supuesto, pero vayamos por partes. Como ya dijimos, deben preferirse las pruebas paramétricas por su mayor potencia siempre que se cumplan las condiciones o suposiciones de los respectivos modelos y del nivel de medición.

No hacerlo así, es decir, no aplicar pruebas paramétricas pudiendo, supondría una notable pérdida de información que pudiera derivar en hipótesis nulas, falsas, no rechazadas. Pero si no se cumplen tales requisitos, lo que hacemos es conceder a los datos una información suplementaria que no poseen, lo que igualmente puede falsear los resultados. Sin embargo, esta discusión conviene enfocarla en forma no generalizada, sino de modo analítico, considerando aspecto por aspecto.

Pues comencemos, si le parece, por el argumento de potencia, referido sobre todo a las pruebas no paramétricas. En cuanto a potencia o eficiencia estadística, las pruebas no paramétricas pueden alcanzar idéntica potencia que las paramétricas aumentando el valor de n , es decir, del número de observaciones, de datos, de sujetos. Y aunque ahorrar sujetos u observaciones es importante, al menos interesante, este argumento de ahorro debe ser secundario en la investigación.

Las pruebas no paramétricas suelen informarnos de su potencia-eficiencia, esto es, del número de observaciones, de datos, de sujetos, que deben incluir para alcanzar la misma potencia que su correspondiente prueba paramétrica. Y en cuanto al criterio de modelo estadístico, ¿qué

cabría decir, profesor del Río? Distingamos por apartados o condiciones. A, independencia de observaciones.

Es una exigencia de ambos tipos de pruebas, de las paramétricas y de las no paramétricas. B, supuesto de normalidad, exigible en pruebas paramétricas, parece ser, según estudios experimentales, por ejemplo, de Pearson, de Boggs, de Anderson, que los efectos de su incumplimiento, digo, del supuesto de normalidad, de incumplimiento del supuesto de normalidad, son prácticamente irrelevantes en cuanto al valor alfa, es decir, en cuanto a la probabilidad de error tipo 1, que consiste, este error tipo 1, en rechazar hipótesis de nublidad, siendo estas correctas realmente. C, homogeneidad de varianza.

Exigible, por supuesto, en las pruebas paramétricas, y que cuando este supuesto no se cumple, los resultados, si se utilizan pruebas paramétricas, pueden llegar a ser graves. Los efectos más graves se dan cuando los N, es decir, el número de datos, de observaciones o sujetos, de los grupos, son desiguales. En tales casos, si lo que desea el investigador es, evitar errores de tipo 1, repito, que era rechazar hipótesis nulas que son correctas, debe procurar que las varianzas mayores se correspondan con los grupos de mayor número de datos o sujetos.

Estos hechos aconsejan que, aunque puede prescindirse de probar la normalidad de las poblaciones, es conveniente el contraste de homogeneidad de varianzas, para tener seguridad de que las decisiones que se toman, a partir de los resultados de las pruebas estadísticas, lo sean realmente al nivel de probabilidad decidido por el investigador. Sin embargo, y a pesar de todo, algunos autores, como Kerlinger, manifiestan una clara posición favorable a las pruebas paramétricas. Quizá en el nivel de medición las posiciones ya no estén tan acordes, como consecuencia de las dos posturas que nos acaba de describir, a la hora de elegir entre pruebas paramétricas o no paramétricas.

De las dos posturas, la dogmática o radical, recordémoslo, preconizaba el uso de pruebas no paramétricas en ciencias de la conducta, mientras no se pudiese lograr una medida de intervalo, como mínimo, que es la exigencia mínima también de las pruebas paramétricas. La postura flexible y abierta propugnaba usar aquellas técnicas estadísticas que se creyesen más apropiadas en cada caso, las paramétricas, con preferencia las no paramétricas, y que nos pudiesen llevar a conclusiones pedagógicas razonables, aunque el nivel de medición no llegase exactamente al de intervalo. No obstante, si hay acuerdo general en que, primero, salvo casos rarísimos, cabe hablar de pruebas paramétricas para escalas de razones, y no es conveniente o no es rentable utilizar pruebas no paramétricas en tales escalas.

Segundo, asimismo, no deben utilizarse pruebas paramétricas cuando la medición no supere el nivel nominal. El punto de fricción está en las medidas ordinales, en las de intervalo, repetimos, es viable el contraste paramétrico, y en la posibilidad o no de aplicación de técnicas paramétricas a esas mediciones ordinales, que son las más comunes en ciencias de la educación. En este campo, algunos autores, los de la postura flexible, por supuesto, proponen la utilización de técnicas paramétricas en escalas ordinales, si bien precisando la necesidad de

contar con instrumentos de medición lo más finos posible, nivel de cuasi-intervalo, como dice Glass-Stanley.

Además, advierten de la prudencia en la interpretación de resultados, especialmente cuando se encuentran en zonas limítrofes a la significación estadística. Finalmente, y dentro de la opción no paramétrica, profesor del Río, ¿qué criterios habríamos de seguir para elegir la prueba estadística más adecuada? Al optar por el grupo de pruebas no paramétricas y seleccionar la más adecuada, habrá que tener en cuenta los siguientes aspectos o criterios. Primero, nivel de medición, nominal, ordinal, y, en algunos casos, de intervalo.

Segundo, número de grupos del diseño. Tercero, independencia o correlación entre las observaciones o datos. Cuarto, categorías de que consta la variable o en que esta variable aparece clasificada.

Quinto, potencia-eficiencia. Sexto, simetría o no de la población. Séptimo, número de casos de la muestra.

Y octavo, finalidad de la investigación. Ya sea esta finalidad patentizar diferencias en medidas de tendencia central o ya mostrar una diferencia de cualquier tipo. El elenco o lista de pruebas no paramétricas utilizables puede verse, por ejemplo, en la obra estadística no paramétrica aplicada a las ciencias de la conducta, de Sidney Seagal, de la editorial Trillas, donde pueden ver, nuestros oyentes, 25 pruebas desde las de un solo grupo, binomial, G-cuadrado, colmogrófono, semisnocio, rachas, pasando por las de dos grupos hasta las de tres o más grupos, y las relativas a las medidas de correlación.

Bien, pues, hasta aquí ha llegado hoy este tiempo que hemos compartido con el profesor Dionisio del Río Sádornil, dedicado especialmente a los estudiantes de Pedagogía Experimental. Les hemos hablado de la opción no paramétrica en investigaciones educativas, y como nos apuntaba el profesor del Río, tienen ustedes bibliografía en esa guía del curso para poder ampliar estas cuestiones. Muchas gracias, profesor, y buenas noches.

Muchas gracias a ustedes, buenas noches. El próximo jueves, en Revista de Educación, cuestiones relacionadas con primer curso. Muchas gracias por su atención y buenas noches.

Tiempo ahora en la UNED para los programas de enseñanza abierta.