

Valor estadístico del poder de discriminación de pruebas pedagógicas de ejecución máxima por métodos correlacionales o inferenciales

Statistical value of the discrimination power of pedagogical tests of maximum performance by correlational or inferential methods

Recepción: 20 de octubre de 2024

Aceptación: 27 de noviembre de 2024

<https://doi.org/10.26490/uncp.investigacionyeducacion.2024.5.1.2358>

Cárdenas Ayala, Edgar Aníbal

Universidad Nacional del Centro del Perú, Huancayo, Perú

ecardenas@uncp.edu.pe y anibalito.c@gmail.com.

<https://orcid.org/0000-0002-8888-2401>

Resumen

En el presente trabajo de investigación de tipo descriptivo correlacional, se estudió el estadígrafo psicométrico; poder de discriminación de los ítems de las pruebas pedagógicas de ejecución máxima de tipo dicotómica, que se aplican a los estudiantes de las instituciones educativas de la Región Junín; calculado por métodos correlacionales —correlación biserial-puntual— y por métodos inferenciales —a través de una prueba t de student para datos relacionados— y analizar estadísticamente e interpretar la relación existente entre los resultados obtenidos por ambos métodos, fue su objetivo principal. Tuvo como hipótesis: Existe una relación significativa entre los resultados del estadígrafo psicométrico (promedio), de poder de discriminación calculados, por métodos correlacionales —correlación biserial-puntual—, y por métodos inferenciales —prueba t de student para datos relacionados—; de las pruebas pedagógicas de ejecución máxima de tipo dicotómica, que se aplican a los estudiantes de las instituciones educativas de la Región Junín. Arribando a la conclusión, entre otras, que teniendo en cuenta, el valor de la correlación de Pearson $r = 0,975$; correlación alta y fuerte, con significatividad bilateral promedio de 0,01; y, el valor del coeficiente de determinación 95%; es indistinto utilizar cualesquiera de los métodos, el correlacional o el inferencial para el cálculo del poder de discriminación de las pruebas de ejecución máxima de tipo dicotómica, que se aplican a los estudiantes de las instituciones educativas de la Región Junín.

Keywords: Edumetric, Statistics and educational investigation, Psychometric.

Abstract

STATISTICAL VALUE OF DISCRIMINATION POWER OF PEDAGOGICAL TESTS OF MAXIMUM EXECUTION BY CORRELATIONAL OR INFERNENTIAL METHODS

Edgar Aníbal Cárdenas Ayala

Nationality: Peruvian.

Affiliation: National University of Central Peru. Researcher and professor of Natural Sciences, Quantitative Methods for Scientific Research and Languages (English and Portuguese). Sciences Education Doctor and Master of University Didactics; both by the National University of Central Peru.

Emails: ecardenas@uncp.edu.pe y anibalito.c@gmail.com

ORCID: <https://orcid.org/0000-0002-8888-2401>

In the present descriptive correlational research work, the psychometric statistician was studied; discrimination power of the items of the pedagogical tests of maximum execution of dichotomous type, which are applied to the students of the educational institutions of Junín Region; calculated by correlational methods —biserial-point correlation— and by inferential methods —through a t student test for related data— and to statistically analyze and interpret the relationship between the results obtained by both methods, was its main objective. Its hypothesis was: There is a significant relationship between the results of the psychometric statistic (average), of discrimination power calculated by correlational methods — biserial-point correlation— and by inferential methods —t student test for related data—; of pedagogical tests of execution maximum of dichotomous type, which are applied to the students of educational institutions of Junín Region. Arriving at the conclusion, among others, that taking into account the value of the Pearson correlation $r = 0,975$; high and strong correlation, with an average bilateral significance of 0,01; and the value of the coefficient of determination 95%; it is indistinct to use any of the methods, the correlational or the inferential one, for the calculation of the discrimination power of the maximum dichotomous type tests, which are applied to the students of the educational institutions of Junín Region.

Keywords: Edumetric, Statistics and educational investigation, Psychometric.

Introducción

Desde hace más de una década y de manera casi ininterrumpida se vienen realizando de manera sistemática investigaciones centradas en averiguar las relaciones que existen entre los diferentes estadígrafos psicométricos de los ítems de las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de la Región Junín; y en la presente investigación se realiza el estudio del valor estadístico del estadígrafo psicométrico, poder de discriminación, de las pruebas pedagógicas de ejecución máxima por métodos correlacionales o inferenciales.

Por otro lado, el análisis cuantitativo y cualitativo de datos provenientes de la aplicación de las pruebas pedagógicas, implica, entre otros temas, el cálculo del estadígrafo psicométrico de poder de discriminación; quedando por determinar cuál método —correlacional o inferencial— es el más indicado y pertinente para su cálculo o ambos lo son. Así también, es necesario recalcar que, para la presente investigación, como espacio muestral, sólo se tomaron en consideración a las pruebas pedagógicas de ejecución máxima de tipo dicotómica, por adecuarse más y mejor al cálculo por los métodos correlacionales.

Entonces, para determinar cuál de los dos métodos es el más adecuado, se calculó la correlación r de Pearson, entre los valores promedio de los coeficientes de poder de discriminación calculados por el método correlacional y los valores de la t de student (a los que se les consideró como valores del poder de discriminación por el método inferencial); de las diez pruebas pedagógicas dicotómicas siendo el valor de la correlación $r = 0,975$; correlación alta y fuerte, con significatividad bilateral promedio de 0,01; así también se calculó, el coeficiente de determinación $r^2 = 0,95$; el que se interpreta —mediante el coeficiente de variabilidad explicada ($r^2 \cdot 100 = 95\%$)— esto quiere decir que existe hasta el 95% de coincidencia entre el valor del poder de discriminación calculado por métodos correlacionales con el realizado por métodos inferenciales, y sólo habría un 5% de no coincidencia, en

las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de la Región Junín, lo que estaría poniendo de manifiesto que, ambos métodos son igual de pertinentes para el cálculo del poder de discriminación.

Aproximaciones teóricas y conceptuales

Pruebas de ejecución máxima

Son aquellas que miden las características de la persona de acuerdo a una calificación de correcto o errado y pueden ser de rendimiento, aptitud y habilidad.

a.- De rendimiento

Miden el aprendizaje que se ha producido como resultado de las experiencias en un programa de estudios o entrenamiento. No tienen un tiempo fijo establecido, puede referirse a lo aprendido en una semana, un mes, un año, etc. Se utilizan para tomar decisiones en relación al avance de una estudiante. Pueden ser de tipo dicotómico o policotómico.

b.- De aptitud

Señalan capacidades para adquirir ciertas conductas o habilidades, dadas las oportunidades adecuadas. Miden los resultados de experiencias de aprendizaje generales e incidentales y su marco de referencia se enfoca hacia el futuro. Tienen por finalidad predecir lo que puede aprender en el futuro.

c.- De habilidad

Miden las experiencias de aprendizaje más amplias y tienen su marco de referencia en el presente e indican el poder para realizar una tarea en el aquí y el ahora.

Estadígrafos psicométricos de las pruebas de ejecución máxima

En las investigaciones con enfoque cuantitativo, la metodología estadística interviene en el proceso de validación de los instrumentos de recolección de datos, siendo uno de ellos las pruebas pedagógicas de ejecución máxima y como tal debe reunir los siguientes requisitos esenciales: confiabilidad, validez, poder de discriminación, nivel de dificultad-facilidad y sensibilidad; el presente trabajo de investigación pretende hallar el valor estadístico del estadígrafo psicométrico, poder de discriminación de las pruebas pedagógicas

de ejecución máxima de rendimiento; por métodos correlacionales e inferenciales de manera comparativa, y así evaluar cuál de los dos métodos es el más adecuado; determinar la equivalencia de ambos métodos y recomendar la eficacia. Entonces, las aproximaciones teóricas están enfocadas para el estadígrafo psicométrico, poder de discriminación.

Poder de discriminación

Mide el grado con el cual el ítem es capaz de establecer diferencias entre las personas con niveles altos y bajos de una habilidad, aptitud o conocimiento que está siendo evaluado.

El poder discriminativo de las pruebas pedagógicas representa el grado de diferenciación, separación, distanciamiento o dispersión de los resultados observados mediante las respuestas que el estudiante manifieste ante cada pregunta planteada. Desde la perspectiva psicométrica, se espera que, mayoritariamente, las respuestas acertadas, correctas o buenas provengan de aquellos estudiantes que también muestran un rendimiento o desempeño general adecuado en el total de la prueba, y que las respuestas erradas u omitidas provengan del grupo de estudiantes que presente un rendimiento general también deficiente en la prueba. Es decir, se espera que cada ítem sea capaz de permitir una diferenciación lo más nítida posible entre los estudiantes de alto y bajo rendimiento, lo que, cuantitativa e idealmente, equivaldría a 100%, o en proporciones a 1,00.

También existe un método descriptivo, que al igual que para el método inferencial se separan los resultados de las pruebas considerando el grupo superior (27%) y el grupo inferior (27%), luego se obtienen separadamente para cada ítem el porcentaje de participantes que responden correctamente, ambos datos se restan y el resultado final es la discriminación que tiene el ítem. La fórmula es la siguiente:

$$\text{Dis} = \text{GS} - \text{GI}$$

Dónde: GS: % del grupo superior que contestó correctamente el ítem.

GI: % del grupo inferior que contestó correctamente el ítem.

Para aceptar el ítem, por su valor del coeficiente de poder de discriminación, debe ser positivo e igual o mayor a 0,30.

Cálculo del poder de discriminación por los métodos correlacional e inferencial

Por el método correlacional

Esencialmente, la correlación es la medida de la relación entre dos variables, sin implicar necesariamente que una de las variables sea la causa de la otra, en algunos casos puede serlo, en otros casos las variables están relacionadas porque ambas lo están a una tercera o son causa de ella. Existen relaciones lineales y no lineales (curvilíneas). Las necesarias para el presente trabajo de investigación son las lineales.

El valor de los coeficientes de correlación varía entre -1 hasta +1, pasando por cero. La mayor parte de los coeficientes de correlación indican: la cuantía de la relación; el signo proporciona información sobre el sentido de la relación, si es positivo quiere decir que, al aumentar el valor de una de las variables la otra también aumenta —las variables están en relación directa—, en cambio si el signo es negativo quiere decir que, si aumenta el valor de una de las variables la otra disminuye, y viceversa —están en relación inversa—. Si el valor de la correlación es cero, indica una ausencia total de relación entre las dos variables, es decir independencia de las variables.

Existen diferentes coeficientes de correlación. De acuerdo a las escalas de medida de las variables: para las ordinales, coeficiente de correlación de Spearman; para las de intervalo, coeficiente de correlación de Pearson—. También, existen muchos otros coeficientes de correlación, con características, fórmulas propias y formas de ser calculados, así, por ejemplo: correlación biserial, correlación punto biserial (biserial puntual), correlación cuádruple, correlación tetracórica, correlación de Kendall. Finalmente, existen correlaciones especiales: correlación parcial y correlación múltiple.

Para la consecución, así como para los efectos de la presente investigación, el método correlacional aplicado es el punto biserial, porque en la validación de las pruebas pedagógicas, una de las variables es continua y la otra es dicotómica (la respuesta es correcta o incorrecta). La magnitud de esta capacidad diferenciadora de los ítems se expresa por índices de discriminación, cuyo cálculo se realiza mediante el método correlacional, coeficiente de correlación punto biserial ($r_{p,bis}$).

$$P.D. = r_{p.bis} = (p_i - X_t)/s_t \cdot (p/q)^{1/2}$$

o

$$P.D. = \frac{p_i - X_t}{s_t} \sqrt{\frac{p}{q}}$$

Dónde:

P.D.: Poder de discriminación del ítem.

p_i : promedio de las puntuaciones totales de los estudiantes que respondieron bien el ítem analizado.

X_t : promedio de las puntuaciones totales de todos los estudiantes Considerados en el análisis.

s_t : desvío de las puntuaciones totales de todos los estudiantes considerados en el análisis.

p y q : proporciones de los estudiantes que, respondieron bien, mal u omitieron el ítem analizado.

Tabla 1

Clasificador del poder de discriminación del ítem

TIPO DE DISCRIMINACIÓN	VALOR
Discriminación muy buena	de 0,40 a 0,99
Discriminación aceptable	de 0,30 a 0,39
Discriminación intermedia	de 0,20 a 0,29
Discriminación inaceptable	de 0,05 a 0,19

Fuente: (Santibáñez, 2001, p.187)

Algunos autores —entre otros, Santibáñez— ponen de manifiesto que el valor obtenido debe ser positivo y para aceptar el ítem debe ser igual o mayor a 0,30; es decir, poseer un poder de discriminación aceptable, para los efectos de la presente investigación éste será el referente. Pero, también hay otros autores que manifiestan que el valor del poder de discriminación del ítem podría ser igual o mayor a 0,20.

Mediante el método correlacional, se obtiene un valor del coeficiente de poder de discriminación para cada ítem de la prueba,

Por el método inferencial

El tercer estadio en que el que interviene la metodología estadística, en las investigaciones con enfoque cuantitativo, es la inferencia estadística, ésta ayuda al investigador a encontrar significatividad en sus resultados al relacionar o comparar dos o más muestras.

En la inferencia estadística se comparan dos o más grupos de datos para poder determinar si las posibles diferencias a encontrarse entre ellos son diferencias reales o son debidas solamente al azar.

Así, por ejemplo, si un investigador está comparando a un grupo experimental con uno de control con respecto al nivel de aprendizaje alcanzado en la enseñanza de nuevos contenidos con dos metodologías diferentes. El investigador puede encontrar para cada grupo sus medidas de tendencia central (media y desviación estándar, por ejemplo), y desea saber si uno de los grupos es superior al otro o si son iguales. Puede ocurrir que a simple vista los puntajes pueden aparecer como diferentes, pero para saber si realmente son diferentes debe de recurrir a la comparación de las medias en función de su variabilidad para determinar la probabilidad en que las diferencias obtenidas entre las medias son diferencias reales y no diferencias debidas al azar o casualidad.

Existen dos tipos de pruebas de inferencia estadística, es decir, maneras de llevar a cabo el análisis de los datos, dependiendo cada uno del cumplimiento de supuestos, así como también del tipo de escala de medición empleada. Estas dos maneras de procesar los datos reciben los nombres de pruebas estadísticas paramétricas y pruebas estadísticas no paramétricas.

Las pruebas estadísticas paramétricas fueron creadas para datos provenientes, principalmente de escalas de intervalo. Su aplicación en la prueba de una hipótesis de investigación requiere del cumplimiento de los siguientes supuestos:

- Distribución normal, la utilización es más válida con datos que provienen de una población cuya distribución es normal;

es decir, una distribución de los datos en forma simétrica alrededor de la media.

- Homogeneidad de la varianza, esto, hace referencia al grado de dispersión que pueden tener cada uno de los grupos de datos que se compara. Se hace más fácil la comparación de grupos y la interpretación de los resultados, cuando las varianzas son iguales u homogéneas.

Las pruebas estadísticas no paramétricas, para su aplicación, mayormente no exigen el cumplimiento de supuestos, como es el caso de las paramétricas, no requieren una distribución normal ni varianza homogéneas de los grupos. Son aplicadas principalmente a datos provenientes de escalas nominales y ordinales, aunque es también posible su aplicación a muestras grandes donde las suposiciones paramétricas no son satisfechas o su cumplimiento es dudoso. De preferencia se aplican a muestras chicas (menos de 30 casos).

En la metodología inferencial, siempre es necesario saber cómo elegir la prueba estadística apropiada, aquí se dan algunas recomendaciones para este menester:

- 1) Primero se debe determinar el número de variables dependientes e independientes en el estudio realizado. Para los propósitos estadísticos muchas veces se consideran a las variables interviniéntes como variables independientes.
- 2) Luego se identifica el tipo de escala en que han sido mensuradas las variables: de intervalo, nominal u ordinal.
- 3) Si ambas variables (independiente y dependiente) están medidas en una escala de intervalo, puede emplearse técnicas de correlación.
- 4) Cuando las variables independientes y dependientes son de intervalo, las pruebas como la t de student y el análisis de varianza son empleadas.
- 5) Las mediciones de tipo ordinal generalmente requiere la aplicación de técnicas No paramétricas, y se recomienda el análisis de Chi Cuadrado para una combinación de variable independiente y variable dependiente nominal.

- 6) Para el tratamiento o análisis de más de una variable independiente y dependientes simultáneamente, es necesario el empleo de técnicas estadísticas bastante complejas, como: Análisis de varianza, correlación múltiple, análisis factorial, etc.

Finalmente, para lograr realizar el cálculo del poder de discriminación por el método inferencial, se separan los resultados de las pruebas considerando el grupo superior (27%) y el grupo inferior (27%), Se hace una prueba de hipótesis t de student para comparar los dos grupos y así probar el poder de discriminación de todo el instrumento; es decir, mediante el método inferencial, se halla el poder de discriminación de toda la prueba.

Metodología de la investigación

Población

Se tomó como población de la investigación a todas las pruebas pedagógicas objetivas de ejecución máxima de tipo dicotómica, que aplicaron a sus alumnos, en las diferentes instituciones educativas del nivel primario, secundario y tecnológico de la Región Junín, los estudiantes de la Maestría y Doctorado en Educación, en los años académicos 2023 al 2024, en la asignatura de Análisis de datos cuantitativos y cualitativos; las mismas que se consignan en los anexos del presente trabajo de investigación.

Muestra

Para la elección de la muestra se utilizó un diseño muestral no aleatorio dirigido, de tal manera que, la muestra estuvo conformada por diez pruebas pedagógicas de ejecución máxima de tipo dicotómicas, que aplicaron —los estudiantes de la Maestría en Educación— a sus alumnos en las instituciones educativas de la Región Junín en los dos últimos Años Académicos Escolares: 2023 al 2024.

Tipo de investigación

Dentro del campo de la investigación educativa y de acuerdo a las características de la hipótesis y los objetivos se enmarcó dentro del tipo de investigación descriptivo correlacional.

Método básico de la investigación

El método utilizado en la investigación fue el descriptivo.

Diseño de la investigación

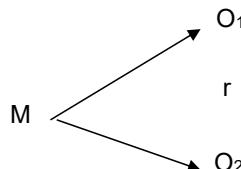
El diseño a emplearse en esta investigación fue el descriptivo, porque resulta imposible el control experimental riguroso de las variables. Además, este diseño se aadecua perfectamente a la presente investigación.

Dentro de los diseños descriptivos se encuentra el diseño descriptivo correlacional, y es el que se aplicó para determinar la relación entre los resultados del estadígrafo psicométrico, poder de discriminación, por los dos métodos (correlacional e inferencial) de las pruebas pedagógicas de ejecución máxima que se aplican a los estudiantes de las instituciones educativas de la Región Junín.

En el caso concreto de la presente investigación, se aplicaron las pruebas pedagógicas, luego con los datos recolectados, se calcularon los estadígrafos psicométricos de poder de discriminación, por los dos métodos puestos de manifiesto en el párrafo anterior, los que fueron analizados. Sánchez (2006:105).

El diseño puede diagramarse de la siguiente manera:

Figura 1



Dónde:

M: Pruebas pedagógicas de tipo dicotómica, que se aplican a los estudiantes de las instituciones educativas de la Región Junín.

O₁: Estadígrafo psicométrico (promedio) de poder de discriminación calculado, por métodos correlacionales —correlación biserial-puntual—.

O₂: Estadígrafo psicométrico de poder de discriminación calculado, por métodos inferenciales —prueba t de student para datos relacionados—
r : Coeficiente de correlación.

Variables de la investigación

Variable 01: (V-01)

Estadígrafo psicométrico (promedio) de poder de discriminación calculado, por métodos correlacionales —correlación biserial-puntual— a los ítems de las pruebas pedagógicas de ejecución

máxima de tipo dicotómica, que se aplican a los estudiantes de las instituciones educativas de Región Junín.

Variable 02: (V-02)

Estadígrafo psicométrico de poder de discriminación calculado, por métodos inferenciales —prueba t de student para datos relacionados— a las pruebas pedagógicas de ejecución máxima de tipo dicotómica, que se aplican a los estudiantes de las instituciones educativas de Región Junín.

Técnicas e instrumentos de recopilación de datos

Procedimientos de recopilación de datos

Los docentes —mencionados en el anterior acápite— fueron capacitados (oportunamente) en el diseño, elaboración y aplicación de pruebas; así como, en el análisis estadístico y el cálculo del estadígrafo psicométrico de poder de discriminación por los métodos correlacionales e inferenciales. Los diez instrumentos de recolección de datos, fueron aplicados por los docentes a sus estudiantes en sus respectivas instituciones educativas.

Técnicas y procedimientos de procesamiento y análisis de datos

- Siendo la investigación de enfoque cuantitativo, para el cálculo se utilizó el paquete estadístico IBM SPSS Statistics 25, con los menús correspondientes: estadísticos descriptivos —media, suma y desvío—, estadísticos inferenciales —comparación de medias—. El cálculo de los estadígrafos se realizó para cada uno de los ítems y para el puntaje total obtenido en cada una de las diez pruebas. Los datos, pueden observarse en los Anexos respectivos.
- El poder de discriminación, mediante métodos correlacionales se hizo mediante el cálculo de medias, desvíos y la aplicación de la fórmula respectiva, y para la interpretación de los valores de los coeficientes de poder de discriminación de cada uno de los ítems de las pruebas, se tuvo en cuenta el rango de valores: superiores a 0,30 —de 0,30 a 0,39 cualitativamente indican poder de discriminación aceptable y de 0,40 a 0,99 cualitativamente indican poder de discriminación muy bueno—; (Santibáñez, 2001, p. 187).
- El poder de discriminación por métodos inferenciales se hizo mediante la prueba t de student para muestras relacionadas o emparejadas.

Resultados

Tabla 2

Resumen de valores del poder de discriminación por el método correlacional e inferencial

PRUEBA PEDAGÓGICA	Valor promedio del coeficiente de poder de discriminación por el método correlacional	Valor de la t de student (poder de discriminación por el método inferencial)
Nº01	0,489	-12,649
Nº02	0,472	-14,749
Nº03	0,442	-10,729
Nº04	0,352	-10,759
Nº05	0,569	-11,357
Nº06	0,426	-9,437
Nº07	0,469	-10,537
Nº08	0,587	-11,624
Nº09	0,412	-12,648
Nº10	0,402	-11,981

Tabla 3

Correlación entre el poder de discriminación correlacional (valor promedio) y el poder de discriminación inferencial (valor de la t de student)

	PODER DE DISCRIMINACIÓN CORRELACIONAL	PODER DE DISCRIMINACIÓN INFERENCIAL
PODER DE DISCRIMINACIÓN CORRELACIONAL	Correlación de Pearson (r)	1
	Sig. (bilateral)	0,000
	N	10
PODER DE DISCRIMINACIÓN INFERENCIAL	Correlación de Pearson (r)	0,975**
	Sig. (bilateral)	0,000
	N	10

**. La correlación es significativa en el nivel 0,01 (bilateral).

A las diez pruebas pedagógicas de tipo dicotómica, se les hizo la correlación de Pearson entre los valores promedio de los coeficientes de poder de discriminación calculados por el método correlacional y los valores de la t de student, a los que se les consideró como valores del poder de discriminación por el método inferencial; siendo el valor de la correlación de Pearson $r = 0,975$; correlación alta y fuerte, con significatividad bilateral promedio de 0,01; por ser positiva, indica que existe una correlación directa entre los dos métodos de cálculo del poder de discriminación; es decir, se podría utilizar indistintamente cualesquiera de los dos métodos para calcular el poder de discriminación, por cierto que, el valor del poder discriminación por métodos correlacionales es para cada uno de los ítems de la prueba pedagógica y el poder de discriminación de toda la prueba pedagógica es el promedio de los valores de poder de discriminación; en cambio, el poder de discriminación por métodos inferenciales es en términos de aceptación de la hipótesis alterna en la docimasia de hipótesis, la que manifiesta que existe diferencia significativa entre los resultados obtenidos por el 27% superior de los estudiantes y el 27% inferior de los estudiantes, por lo tanto, la prueba pedagógica a la que fueron sometidos tiene poder de discriminación. Por otro lado, el coeficiente de determinación es $r^2 = 0,95$; el que se interpreta —mediante el coeficiente de variabilidad explicada (r^2) $100 = 95\%$ — esto quiere decir que existe hasta el 95% de coincidencia entre el valor del poder de discriminación calculado por métodos correlacionales con el realizado por métodos inferenciales, y sólo habría un 5% de no coincidencia.

Docimasia de hipótesis para la correlación del poder de discriminación por métodos correlacionales e inferenciales.

- Hipótesis nula (H_0): La correlación entre los valores promedio de los coeficientes de poder de discriminación calculados por el método correlacional y los valores de la t de student, a los que se les consideró como valores del poder de discriminación por el método inferencial de las pruebas pedagógicas de ejecución máxima de tipo dicotómica, no es significativa.

$$H_0: p = q$$

- Hipótesis alterna (H_1). La correlación entre los valores promedio de los coeficientes de poder de discriminación calculados por el método correlacional y los valores de la t de student, a los que se les consideró como valores del poder de discriminación por el método inferencial de las pruebas pedagógicas de ejecución máxima de tipo dicotómica, es significativa. (Bilateral)

$$H_1: p \neq q$$

- Distribución de la muestra: la distribución de la muestra es la distribución t de student con grados de libertad $gl = N-2$; teniendo en cuenta que son diez pruebas pedagógicas de ejecución máxima consideradas en el espacio muestral, será:

$$gl = 10 - 2 = 8$$

- Nivel de confianza: 95%; nivel de significación $\alpha = 0,05$; y un error de 5%.
- Prueba estadística: se eligió la prueba t de student bilateral, en función de la $r = 0,975$ de Pearson, para las diez pruebas pedagógicas de ejecución máxima.

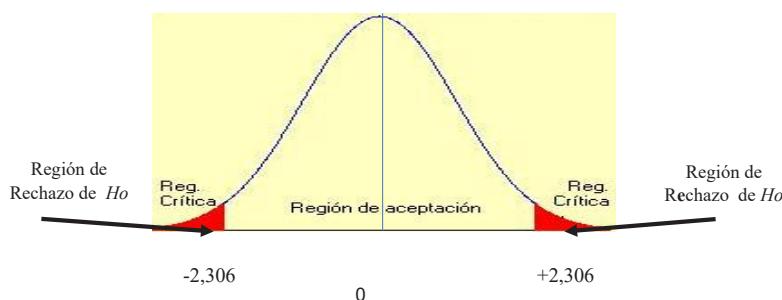
$$t = \frac{r}{\sqrt{\frac{1-r^2}{N-2}}}$$

$$t = \frac{0,975}{\sqrt{\frac{1-0,95}{8}}} = \frac{0,975}{0,079} = 12,342$$

t calculada = + 12,342

- Región crítica: $t_{tabulada} = +/- 2,306$. Puesto que H_1 es bilateral, la región crítica abarca todos los valores de $t \geq +2,306$ o $t \leq -2,306$.

Figura 2



- **Decisión:** Puesto que la $t_{calculada}$ está ubicada en la región de rechazo del H_0 ; es decir, la $t_{calculada} > t_{tabulada}$, se acepta la H_1 .
- **Conclusión estadística**
La correlación entre los valores promedio de los coeficientes de poder de discriminación calculados por el método correlacional y los valores de la t de student, a los que se les consideró como valores del poder de discriminación por el método inferencial de las pruebas pedagógicas de ejecución máxima, es significativa de tipo dicotómica

Discusión final de resultados

Para tratar de dar solución al problema de la presente investigación: ¿Qué relación existe entre los resultados del estadígrafo psicométrico de poder de discriminación calculados, por métodos correlacionales —correlación biserial-puntual—, y por métodos inferenciales —prueba t de student para datos relacionados—; de las pruebas pedagógicas de ejecución máxima de tipo dicotómica, que se aplican a los estudiantes de las instituciones educativas de la Región Junín?

En primer lugar, el poder de discriminación, está explicado mediante el estadígrafo de relación, correlación punto biserial; para cada uno de los ítems de las pruebas, en los que se tuvo en cuenta el rango de valores: superiores a 0,30 —de 0,30 a 0,39 cualitativamente indican poder de discriminación aceptable y de 0,40 a 0,99 cualitativamente indican poder de discriminación muy bueno—; tal como lo prescribe (Santibáñez, 2001, p. 187).

En segundo lugar, para realizar el cálculo del poder de discriminación por el método inferencial, se separaron los resultados de las pruebas considerando el grupo superior (27%) y el grupo inferior (27%), Se hizo una prueba de hipótesis t de student para comparar los dos grupos y así probar el poder de discriminación de todo el instrumento; es decir, mediante el método inferencial, se halla el poder de discriminación de toda la prueba.

En la presente investigación, se han observado estas consideraciones teóricas y procedimentales; y, los valores calculados del poder de discriminación de los ítems de las pruebas pedagógicas de ejecución máxima de tipo dicotómica, conformantes del espacio muestral; están enmarcados entre los valores requeridos.

Es así como, a los valores promedio de los coeficientes de poder de discriminación calculados por el método correlacional y a los valores de la t de student —a los que se les consideró como valores del poder de discriminación por el método inferencial—; de las diez pruebas pedagógicas de tipo dicotómicas, se les hizo la correlación de Pearson entre; siendo el valor de la correlación de Pearson $r = 0,975$; correlación alta y fuerte, con significatividad bilateral promedio de 0,01; que por ser positiva, indica que existe una correlación directa entre los dos métodos de cálculo del poder de discriminación; es decir, se podría utilizar indistintamente cualesquiera de los dos métodos para calcular el poder de discriminación, por cierto que, el valor del poder discriminación por métodos correlacionales es para cada uno de los ítems de la prueba pedagógica y el poder de discriminación de toda la prueba pedagógica es el promedio de estos valores de poder de discriminación; en cambio, el poder de discriminación por métodos inferenciales es en términos de aceptación de la hipótesis alterna en la docimasia de hipótesis, la que manifiesta que existe diferencia significativa entre los resultados obtenidos por el 27% superior de los estudiantes y el 27% inferior de los estudiantes; por lo tanto, la prueba pedagógica a la que fueron sometidos tiene poder de discriminación. Esta forma de interpretar, es estrictamente matemática, y no estaría revelando relaciones de causa–efecto, como demarca el espíritu de la investigación.

Que se pueden utilizar indistintamente, el método correlacional o el método inferencial para el cálculo del poder de discriminación de las pruebas pedagógicas de tipo dicotómica, es un hecho; por eso se hizo necesario calcular el coeficiente de determinación, cuyo valor es $r^2 = 0,95$; el que se interpreta —mediante el coeficiente de variabilidad explicada ($r^2 \cdot 100 = 95\%$)— esto quiere decir que existe hasta el 95% de coincidencia entre el valor del poder de discriminación calculado por métodos correlacionales con los realizados por métodos inferenciales, y sólo habría un 5% de no coincidencia. Por cierto, que, para emitir cualquier juicio valorativo sobre lo expuesto, habría que analizar no solo diez pruebas pedagógicas, como es el caso de la presente investigación, sino muchas (muchísimas más); no perder de vista que a la presente investigación también se la podría considerar como exploratoria y sería un punto de partida para otras investigaciones que se pudieran realizar.

Observando que, los resultados de los coeficientes de correlación de Pearson —entre el poder de discriminación correlacional (valor promedio) y el poder de discriminación inferencial (valor de la *t* de student) de las pruebas de ejecución máxima utilizadas como espacio muestral del presente estudio— pese a ser alto y fuerte, no se puede aseverar la causalidad entre los cálculos del poder de discriminación por ambos métodos; para todas las pruebas pedagógicas de ejecución máxima que pudieran haber; es más, esta generalización sería riesgosa, considerando que la muestra para la presente investigación no fue tan vasta; tampoco aleatoria.

Así también, contrastando los resultados del presente trabajo, con los obtenidos por Cárdenas (2009, 2010, 2011, 2015, 2020, 2021 y 2023), en sus trabajos de investigación se observa que hay que continuar realizando una mayor cantidad de observaciones respecto al cálculo y a las relaciones que existen entre el poder de discriminación con otros estadígrafos psicométricos, que existen en las pruebas pedagógicas de ejecución máxima de tipo dicotómica, aplicadas a los estudiantes de las instituciones educativas de la Región Junín.

Todo lo expresado líneas arriba, se escapa de los alcances de una buena práctica pedagógica, porque ésta tiene otras características, y los estadígrafos psicométricos investigados no son los únicos, sino que existen otros que también influyen en la elaboración de una prueba pedagógica; tales como los estadígrafos edumétricos. Santibáñez (2001).

Conclusiones

1. El poder de discriminación de gran parte de los ítems de las pruebas pedagógicas de ejecución máxima de tipo dicotómica, que se aplican a los estudiantes de las instituciones educativas de la Región Junín, calculado por el método correlacional punto biserial en términos cualitativos, está en el rango de aceptable y muy bueno.
2. El poder de discriminación de las pruebas pedagógicas de ejecución máxima de tipo dicotómica, que se aplican a los estudiantes de las instituciones educativas de la Región Junín, calculado por el método inferencial es también aceptable.

3. Teniendo en cuenta, el valor de la correlación de Pearson $r = 0,975$; correlación alta y fuerte, con significatividad bilateral promedio de 0,01; y el valor del coeficiente de determinación 95%; es indistinto utilizar cualesquiera de los métodos, el correlacional o el inferencial para el cálculo del poder de discriminación de las pruebas de ejecución máxima de tipo dicotómica, que se aplican a los estudiantes de las instituciones educativas de la Región Junín

Referencias bibliográficas

- Aiken, L. (1995). *Test psicológicos y evaluación*. Prentice Hall Hispanoamericana.
- Ary; Jacobs y Razaviev. (1992). *Introducción a la investigación pedagógica*. McGraw-Hill.
- Calvo, E. (2015). *Estadística: software estadísticos versus formulas estadísticas*. [Trabajo de investigación]. Centro de Producción de Documentos-UPSMP.
- Cárdenas Ayala, A. (2009). *Poder discriminativo, nivel de exigencia y capacidad predictiva de los exámenes de selección* [Investigación no publicada] I.I.FPH-UNCP.
- Cárdenas Ayala, A. (2010). *Relación entre la validez, el poder discriminativo y nivel de dificultad de las pruebas pedagógicas*. Revista Horizonte de la ciencia, 01, 119-126. Huancayo, Perú. <http://revistas.uncp.edu.pe/index.php/horizontedelaciencia>.
- Cárdenas Ayala, A. (2015). *Confiabilidad y poder de discriminación de las pruebas pedagógicas de ejecución máxima*. Revista Prospectiva universitaria, V-13, 01 y 02, 85-90. UNCP. <http://www.uncp.edu.pe/universidad/investigación/index.php>
- Cárdenas Ayala, A. (2020). *Poder de discriminación y validez de pruebas de ejecución máxima*. I.E.I., FE-UNCP, Revista Investigación y educación, 1, 8-21. <https://revistas.uncp.edu.pe/index.php/invest>
- Cárdenas Ayala, A. (2022). *Nivel de exigencia y poder de discriminación de pruebas pedagógicas de ejecución máxima*. Revista Investigación y educación, 1, 79-90 <https://doi.org/10.26490/uncp.investigaciónyeducación.2022.3.1.1668>
- Cárdenas Ayala, A. (2023). *Sensibilidad y poder de discriminación de pruebas pedagógicas de ejecución máxima*. Revista Investigación y educación, VOL. 4, 1, 40-55. <https://doi.org/10.26490/uncp.investigaciónyeducación.2023.4.1.1886>

- Delgado; Escurra y Torres. (2006). *La medición en psicología y educación*. Hozlo.
- Fermín, M. (1995). *La evaluación, los exámenes y las calificaciones*. Kapeluz.
- Gonzalvo, G. (1994). *Diccionario de metodología estadística*. Morata S.A.
- Hernández; Fernández y Baptista. (2003). *Metodología de la investigación*. McGraw-Hill.
- Lozada y Sánchez. (2010). *Estadística inductiva*. Universidad de Holguín “Oscar Lucero Moya”.
- Nunnally y Bernstein. (1995). *Teoría psicométrica*. Mc Graw-Hill.
- Piris y Zacharzewski. (2017). *Sociedad estadística*. Universidad Nacional de Misiones.
- Quezada, N. (2010). *Metodología de la investigación-Estadística aplicada en la investigación*. MACRO.
- Rubio, A. (2018). *Cuidado al calcular las medidas de forma con el software Excel*. [Trabajo de investigación]. UNALM.
- Sánchez y Reyes. (2006). *Metodología y Diseños en la Investigación Científica*. Visión Universitaria.
- Santa, M. (2015). *Glosario Español-Inglés y English-Spanish de terminología estadística*. Universidad Nacional Abierta y a Distancia.
- Schifelbein, E. (1999). *Teoría, técnicas, procesos y casos en el planeamiento de la educación*. Ateneo.
- Santibañez, J. (2001). *Manual para la evaluación del aprendizaje estudiantil (conceptos, procedimientos, análisis e interpretación para el proceso evaluativo)*. México: Trillas.
- Santisteban, C. (2009). *Principios de psicometría*. Síntesis.
- Tavella, N. (1991). *Análisis de los ítems en la construcción de instrumentos Psicométricos*. Trillas.
- Visauta y Martori i Cañas. (2003). *Análisis estadístico con SPSS para Windows*, (Volumen II). McGraw-Hill.



© Los autores. Este artículo es publicado por la revista Investigación y Educación del Instituto Especializado de Investigación la Facultad de Educación de la Universidad Nacional del Centro del Perú. Es un artículo de acceso abierto, distribuido bajo los términos de la Licencia Atribución-No Comercial 4.0 Internacional. (<https://creativecommons.org/licenses/by-nc/4.0/>), que permite el uso no comercial y distribución en cualquier medio, siempre que la obra original sea debidamente citada.