

Poder discriminativo y nivel de exigencia de los exámenes de selección

Mg. Aníbal Cárdenas Ayala, UNCP

Resumen

En el presente trabajo se determinó, los estadígrafos edumétricos, poder discriminativo y nivel de dificultad, de los exámenes de admisión que se administraron la UNCP en el 2008. Tuvo como objetivos específicos, medir el grado en que los ítems de los exámenes de admisión a la UNCP son capaces de establecer diferencias entre postulantes con niveles altos o bajos en habilidades, capacidades, aptitudes o conocimientos; clasificar por nivel de dificultad cada uno de los ítems de los exámenes de admisión a la UNCP; así como, determinar el poder predictivo, el que se estará realizando en el año 2013, fecha en la que los ingresantes en el 2008 egresarán. La investigación es tecnológico-aplicada con un diseño descriptivo comparativo. Se llevó a cabo en la Universidad Nacional del Centro del Perú, a lo largo del año 2008. Llegando a la conclusión de que las pruebas de admisión que se administran en la UNCP tienen poder discriminativo aceptable y alto nivel de dificultad.

Palabras clave: Psicometría, Estadística aplicada a la investigación educativa

Abstract

In the present research work was determined, edumetric statistical, discriminative power and difficulty level, of entrance examinations to UNCP. It had like specific objectives, measuring the degree of items of entrance examinations to UNCP are able to establish differences between postulants with high or low levels in abilities, capacities, qualifications or knowledge; and, classifying by level of difficulty every one of items of entrance examinations to UNCP. Verifying the hypothesis: the analysis and interpretation of statistical, discriminative power and difficulty level, will improve the performance of the entrance examinations to UNCP. Investigation is technological-applied with a descriptive comparative design. It made at the National University of the Center of Peru -UNCP-, in 2008. The conclusion was that entrance examinations to UNCP have acceptable discriminative power and have high level of difficulty.

Key words: Psychometric, Statistic applied to the educational investigation

Introducción

Entre los cinco últimos años de la década pasada, se realizaron investigaciones centradas en la validez, confiabilidad, dificultad y sensibilidad de los diferentes tipos de instrumentos de recolección de datos. Por otro lado, el dictado de la asignatura de Estadística en la Unidad de Post Grado de la Facultad de Pedagogía y Humanidades, implicó, entre otros temas, la enseñanza de los estadígrafos edumétricos, poder discriminativo y el nivel de dificultad de los instrumentos de recolección de datos; y como éstos, deberían de ser más reales que imaginarios, se planteó que el análisis estadístico e interpretación de los mencionados estadígrafos, fuera realizado en los exámenes de selección o de admisión a las universidades, específicamente en los exámenes de admisión a la UNCP.

Método

Población y muestra de la investigación

Población

La población de la investigación estuvo conformada por los exámenes de admisión tomados en la UNCP durante los Años Académicos 2007 y 2008, para los postulantes al pregrado.

Muestra

Para la elección de la muestra se utilizó un diseño muestral no aleatorio dirigido de tal manera que la muestra estuvo conformada por los tres exámenes de admisión, para los postulantes al pregrado, tomados en la UNCP durante el Año Académico 2008: 2008-I, marzo; 2008-II, agosto; y, Primera selección, diciembre.

Método

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 tecnológico-aplicada.

El método empleado fue el descriptivo.

El diseño que se empleó en esta investigación fue el descriptivo, que fue útil para someter a análisis la técnica descrita. Dentro de los diseños descriptivos se encuentra el diseño descriptivo comparativo que consiste en recolectar información relevante en varias muestras con respecto a un mismo fenómeno o aspecto de interés y luego caracterizar este fenómeno sobre la base de la comparación de datos recogidos, pudiendo hacerse esta comparación en los datos generales o en una categoría de ellos. En el caso concreto de la presente investigación, se recolectaron los exámenes de admisión tomados en la UNCP con sus resultados de calificación, luego con estos, se calcularon los estadígrafos de poder discriminativo, e índice de dificultad los que fueron analizados e interpretados (Sánchez, 1996, p.78).

El diseño puede diagramarse de la siguiente manera:

M ₁	O ₁
M ₂	O ₂
M ₃	O ₃

Donde:

- M₁, M₂, M₃ : Exámenes de admisión administrados en la UNCP.
O₁, O₂, O₃ : Estadígrafos edumétricos, poder discriminativo e índice de dificultad.

Variables de la investigación

Exámenes de admisión a la UNCP.

Análisis e interpretación de los estadígrafos edumétricos, poder discriminativo e índice de dificultad.

Técnicas e instrumentos de recopilación de datos

Los instrumentos fueron los exámenes de admisión, para postulantes al pregrado, tomados en la UNCP durante el Año Académico 2008.

Procedimientos

Se analizaron los exámenes de admisión a la UNCP, calculándoles sus estadígrafos edumétricos, poder discriminativo e índice de dificultad. Por último, los datos cuantitativos fueron analizados e interpretados.

Técnicas y procedimientos de procesamiento y análisis de datos

Siendo la investigación cuantitativa, se aplicaron procedimientos estadísticos, calculando los estadígrafos de poder discriminativo e índice de dificultad; mediante las fórmulas respectivas y los paquetes informáticos estadísticos (SPSS y MINITAB).

Resultados

Análisis estadístico de los exámenes de selección

La población objetivo estuvo conformada por los exámenes de admisión tomados en la UNCP durante los Años Académicos 2007 Y 2008.

La población accesible y muestra fueron los exámenes de admisión tomados en la UNCP durante el Año Académico 2008 en sus tres procesos de selección de postulantes:

Examen de admisión 2008-I, marzo; con cinco tipos de prueba.

Examen de admisión 2008-II, agosto; con dos tipos de prueba.

Examen de primera selección, diciembre; con dos tipos de prueba.

Clasificación de las pruebas de admisión aplicadas en la UNCP por áreas y grupos

En el examen de admisión 2008-I se aplicaron cinco tipos de pruebas, las que respondieron a cinco áreas y las áreas agruparon a las diferentes facultades por afinidad de conocimientos a ser evaluados en las respectivas pruebas:

ÁREA	FACULTADES
I	Medicina Humana Enfermería Pedagogía: Biología y Química
II	Agronomía Ciencias Forestales y del Ambiente Zootecnia
III	Arquitectura Ingeniería Eléctrica y Electrónica Ingeniería Mecánica Ingeniería Metalúrgica y de Materiales Ingeniería de Minas Ingeniería Química Ingeniería en Industrias Alimentarias Ingeniería Civil Pedagogía: Matemática y Física
IV	Administración de Empresas Contabilidad Economía
V	Antropología. Sociología. Trabajo Social. Ciencias de la Comunicación. Pedagogía: Educación Física. Español y Literatura. CC. SS. E Historia. Educación Primaria. Educación Inicial.

En el examen de admisión 2008-II y para el examen de primera selección se aplicaron dos tipos de pruebas, las que respondieron a dos grupos y cinco áreas; los grupos y las áreas agruparon a las diferentes facultades por afinidad de conocimientos a ser evaluados en las respectivas pruebas:

GRUPO	ÁREA	FACULTADES
I	I	Medicina Humana Enfermería Pedagogía: Biología y Química
	III	Arquitectura Ingeniería Eléctrica y Electrónica Ingeniería Mecánica Ingeniería Metalúrgica y de Materiales Ingeniería de Minas Ingeniería Química Ingeniería en Industrias Alimentarias Ingeniería Civil Pedagogía: Matemática y Física
II	II	Agronomía Ciencias Forestales y del Ambiente Zootecnia
	IV	Administración de Empresas Contabilidad Economía
	V	Antropología Sociología Trabajo Social Ciencias de la Comunicación Pedagogía: Educación Física Español y Literatura CC. SS. E Historia Educación Primaria Educación Inicial

Análisis del nivel de exigencia de las pruebas de admisión aplicadas en la UNCP

El análisis del nivel de exigencia de las pruebas de admisión se hizo en términos de la dificultad, la que se relaciona con la proporción de postulantes que respondieron adecuadamente a un ítem o a toda la prueba:

$$D = \text{N}^\circ \text{ de postulantes que aprobaron} / \text{N}^\circ \text{ de postulantes evaluados}$$

Se midió para cada uno de los 80 ítems que conformaron cada una de las pruebas y para el calificación total, obtenidos por los postulantes, para el cálculo fueron de gran ayuda el paquete estadístico SPSS.15 y el MINITAB 6,0. La magnitud del nivel de exigencia se expresa en proporciones que fluctúan entre 0,00 y 1,00. Los resultados para cada examen y por tipo de prueba y área respectiva se presentan en el siguiente cuadro:

Examen de admisión 2008-I

Tipo de prueba por área	Nivel de exigencia (dificultad)
ÁREA I	0,26
ÁREA II	0,38
ÁREA III	0,41
ÁREA IV	0,36
ÁREA V	0,44

Los cinco tipos de prueba del examen tuvieron un nivel de exigencia DIFÍCIL, para fines psicométricos, esto se obtuvo comparando con el Cuadro de Clasificación de Nivel de Dificultad (Cortada, 1999, p.182).

Examen de admisión 2008-II

Tipo de prueba por grupo	Nivel de exigencia (dificultad)
GRUPO I	0,31
GRUPO II	0,41

Los dos tipos de prueba del examen tuvieron un nivel de exigencia DIFÍCIL para fines psicométricos, esto se obtuvo comparando con el Cuadro de Clasificación de Nivel de Dificultad (Cortada, 1999, p.182).

Examen de admisión de primera selección 2009

Tipo de prueba por grupo	Nivel de exigencia (dificultad)
GRUPO I	0,29
GRUPO II	0,37

Los dos tipos de prueba del examen tuvieron un nivel de exigencia DIFÍCIL para fines psicométricos, esto se obtuvo comparando con el Cuadro de Clasificación de Nivel de Dificultad (Cortada, 1999, p. 182).

Análisis del poder discriminativo de las pruebas de admisión aplicadas en la UNCP

El poder discriminativo de las pruebas de admisión representa el grado de diferenciación, separación, distanciamiento o dispersión de los resultados observados mediante las respuestas que el postulante manifieste ante cada pregunta planteada. Desde la perspectiva psicométrica, se espera que, mayoritariamente, las respuestas acertadas, correctas o buenas provengan de aquellos postulantes 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 postulantes 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 postulantes de alto y bajo rendimiento, lo que, cuantitativa e idealmente, equivaldría a 100% o, en proporciones, a 1,00. La magnitud de esta capacidad diferenciadora de los ítems se expresa por índices de discriminación, cuyo cálculo se realizó mediante el método correlacional, coeficiente de correlación punto serial ($r_{p,bis}$)

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

Donde:

- p_i : promedio de las puntuaciones totales del postulante que respondió bien el ítem analizado.
- X_t : promedio de las puntuaciones totales de todos los postulantes considerados en el análisis.
- s_t : desvío de las puntuaciones totales de todos los postulantes considerados en el análisis.
- p y q : proporciones de los postulantes que, respondieron bien, mal u omitieron el ítem analizado.

Se calculó para cada uno de los 80 ítems que conformaron cada una de las pruebas y para el calificativo total, obtenidos por los postulantes, para el cálculo fueron de gran ayuda el paquete estadístico SPSS.15 y el MINITAB 6,0. La magnitud del poder de discriminación se expresa en proporciones que fluctúan entre 0,00 y 1,00. Los resultados para cada examen y por tipo de prueba y área respectiva se presentan en el siguiente cuadro:

Examen de admisión 2008-I

Tipo de prueba por área	Poder de discriminación
ÁREA I	0,31
ÁREA II	0,20
ÁREA III	0,29
ÁREA IV	0,39
ÁREA V	0,39

Los promedios del poder de discriminación de los 80 ítems, para los cinco tipos de prueba del examen tuvie-

ron un poder de discriminación entre intermedio y aceptable, lo que se obtuvo comparando con el Cuadro de Clasificación de la Discriminación (Cortada, 1999, p.192) pero, para fines psicométricos, es necesario que el poder discriminativo sea de 0,30 como mínimo. Entonces sólo serían aceptadas como pruebas con poder de discriminación a las pruebas de las áreas I, IV y V.

Examen de admisión 2008-II

Tipo de prueba por grupo	Poder de discriminación
GRUPO I	0,28
GRUPO II	0,32

Los promedios del poder de discriminación de los 80 ítems, para los dos tipos de prueba del examen tuvieron un poder de discriminación entre intermedio y aceptable, lo que se obtuvo comparando con el Cuadro de Clasificación de la Discriminación (Cortada, 1999, p. 192) pero, para fines psicométricos, es necesario que el poder discriminativo sea de 0,30 como mínimo. Entonces sólo serían aceptadas como pruebas con poder de discriminación a las pruebas del grupo II.

Examen de admisión de primera selección 2009

Tipo de prueba por grupo	Poder de discriminación
GRUPO I	0,21
GRUPO II	0,34

Los promedios del poder de discriminación de los 80 ítems, para los dos tipos de prueba del examen tuvieron un poder de discriminación entre intermedio y aceptable, lo que se obtuvo comparando con el Cuadro de Clasificación de la Discriminación (Cortada, 1999, p. 192) pero, para fines psicométricos, es necesario que el poder discriminativo sea de 0,30 como mínimo. Entonces sólo serían aceptadas como pruebas con poder de discriminación a las pruebas del grupo II.

Prueba de hipótesis “t Student” para el poder discriminativo de las pruebas de admisión aplicadas en la UNCP, la prueba de hipótesis se realizó a un grupo muestral tomado aleatoriamente para cada examen de admisión: En el examen de admisión 2008-I, se tomó aleatoriamente a la prueba del Área II, cuyo número total de postulantes fueron 368, se tomó al 5% de ellos como conformantes de la muestra (18 postulantes), el 27% eran aprobados y 27% desaprobados. Intencionalmente se tomaron los resultados de los postulantes a la Facultad de Ciencias Forestales y del Ambiente por ser la de mayor cantidad de postulantes.

En el examen de admisión 2008-II, se tomó aleatoriamente al Grupo I cuyo número total de postulantes fueron 2043, se tomó al 1% de ellos como conformantes de la muestra (20 postulantes), el 27% eran aprobados y 27% desaprobados. Intencionalmente se tomaron los resultados de los postulantes a la Facultad de Medicina Humana por ser la más representativa y apetecida por los postulantes.

En el examen de admisión de Primera Selección, se tomó aleatoriamente al Grupo II cuyo número total de postulantes fueron 1851, se tomó al 1% de ellos como conformantes de la muestra (18 postulantes), el 27% eran aprobados y 27% desaprobados. Intencionalmente se tomaron los resultados de los postulantes a la Facultad de Contabilidad por ser la de mayor cantidad de postulantes.

Para realizar las pruebas de hipótesis, con los datos provenientes de los resultados de los tres exámenes de admisión, se procedió a enunciar en términos estadísticos formales, los resultados se obtuvieron mediante la aplicación del paquete SPSS (VER ANEXO), de la siguiente manera:

Hipótesis nula (H₀): No existe diferencia significativa entre los resultados obtenidos por el 27% de los postulantes aprobados y el 27% de los postulantes desaprobados, por lo tanto las pruebas de admisión a las que fueron sometidos no tienen poder discriminativo.

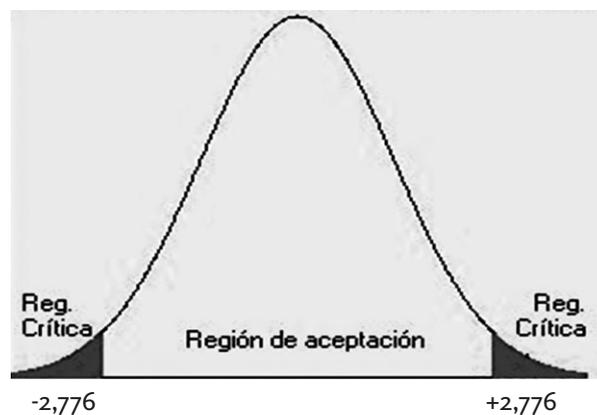
Hipótesis alterna (H_1): Existe diferencia significativa entre los resultados obtenidos por el 27% de los postulantes aprobados y el 27% de los postulantes desaprobados, por lo tanto las pruebas de admisión a las que fueron sometidos tienen poder discriminativo.

Prueba estadística: Se elige la prueba t de Student para muestras relacionadas.

Nivel de confianza: 95%.

Distribución de la muestra: la distribución de la muestra es la distribución t de Student con grados de libertad (gl) = 4.

Región crítica: $t_{\text{tabulada}} = \pm 2,776$. Puesto que H_1 es bilateral, la región crítica abarca todos los valores de $t \geq +2,776$ o $t \leq -2,776$.



Resultados de la prueba t de Student

Examen de admisión	t de Student calculada
2008-I	52,561
2008-II	166,854
PRIMERA SELECCIÓN	40,474

Decisión: Puesto que las t calculadas están ubicadas dentro de la región crítica, es decir, las $t_{\text{calculadas}} > t_{\text{tabuladas}}$, para los tres exámenes de admisión, se acepta la H_1

INTERPRETACIÓN:

Existe diferencia significativa entre los resultados obtenidos por el 27% de los postulantes aprobados y el 27% de los postulantes desaprobados, por lo tanto las pruebas de admisión a las que fueron sometidos tienen poder discriminativo.

Discusión

La investigación aplicada y descriptiva comparativa que realizó el análisis e interpretación de los estadígrafos de poder discriminativo y dificultad de los exámenes de admisión a la UNCP; se corrobora por los resultados obtenidos.

Todos los tipos de prueba de los exámenes de admisión administrados durante el año 2008 --tomado como espacio temporal-- tuvieron un nivel de exigencia alto, tal como lo manifiesta Cortada (1999, p.182) para fines psicométricos en el Cuadro de Clasificación de Nivel de Exigencia; este nivel, medido en términos de índice de dificultad se considera difícil, ya que todos los índices de dificultad obtenidos estuvieron comprendidos

entre 0,26 y 0,44 (Nunnally y Berstein, 1995) es decir, inferiores a 0,5 que vendría a ser el índice de dificultad intermedio, adecuado.

Respecto al poder de discriminación de todos los tipos de prueba de los exámenes de admisión aplicados, en el espacio temporal ya mencionado; presentaron un nivel discriminativo entre intermedio y aceptable, lo que se dedujo comparando con el Cuadro de Clasificación de la Discriminación (Cortada, 1999, p. 192) pero, para fines psicométricos, es necesario que el poder discriminativo sea de 0,30 como mínimo (Nunnally y Berstein, 1995) Entonces sólo serían aceptadas como pruebas con poder de discriminación 5 de las 9 pruebas aplicadas; pero, realizada la prueba de hipótesis t de student y tomando los resultados de ésta como criterio válido (Visauta y Martori, 2003) se acepta que todas los tipos de prueba de los exámenes de admisión, discriminan contundentemente, por su grado de conocimientos a los postulantes, ya que aprueban e ingresan sólo los postulantes con alto grado de conocimientos y los que aún no lo poseen se siguen preparando o en su defecto toman otras opciones académicas.

Conclusiones

- 1.- En el examen de admisión 2008-I, los cinco tipos de prueba tuvieron un nivel de exigencia DIFÍCIL; en el examen de admisión 2008-II, los dos tipos de prueba tuvieron un nivel de exigencia DIFÍCIL; en el examen de admisión de Primera Selección 2009, los dos tipos de prueba tuvieron un nivel de exigencia DIFÍCIL.
- 2.- En el examen de admisión 2008-I, sólo los ítems de las pruebas de las áreas I, IV y V serían aceptados como ítems con poder de discriminación; en el examen de admisión 2008-II, sólo los ítems de las pruebas del grupo II serían aceptados como ítems con poder de discriminación; en el examen de admisión de Primera Selección 2009, sólo los ítems de las pruebas del grupo II serían aceptados como ítems con poder de discriminación.
- 3.- Mediante la prueba de hipótesis "t de student" se determinó que en los tres exámenes de admisión administrados durante el año 2008, existen diferencias significativas entre los resultados obtenidos por el 27% de los postulantes aprobados y el 27% de los postulantes desaprobados, por lo tanto las pruebas de admisión a las que fueron sometidos tienen poder discriminativo.

Referencias bibliográficas:

Aiken, L. (1995). Test psicológicos y evaluación. México D.F., México: Editorial Prentice Hall Hispanoamericana.

Ary; Jacobs y Razavieh. (1992). Introducción a la investigación pedagógica. México D.F., México: Editorial McGraw-Hill.

Cárdenas Ayala, A. (2007). Pertinencia de los Instrumentos de recolección de datos con los métodos estadísticos de confiabilidad. (I.I. FPH-UNCP, trabajo de investigación inédito). Huancayo, Perú.

----- (2008). Análisis estadístico de los instrumentos de recolección de datos a través de los estadígrafos de deformación y apuntamiento. (I.I.FPH-UNCP, trabajo de investigación inédito). Huancayo, Perú.

----- (2010). Relación entre la validez, el poder discriminativo y nivel de dificultad de las pruebas pedagógicas. Horizonte de la ciencia, 01, 109-119. Huancayo, Perú.

----- (2011). Índice de discriminación de los ítems politómicos en las pruebas pedagógicas. Prospectiva universitaria, V- 8, 01, 119-126. Huancayo, Perú.

----- (2012). Análisis estadístico del índice de dificultad de los ítems politómicos en las pruebas pedagógicas. Horizonte de la ciencia, 02, 74-81. Huancayo, Perú.

Delgado; Ecurra y Torres. (2006). La medición en psicología y educación. Lima, Perú: Editorial Hozlo S.R.L.

Fermín, M. (1995). La evaluación, los exámenes y las calificaciones. Buenos Aires, Argentina: Editorial Kapeluz.

Hambleton y Novik. (1989). Toward an integration of theory and method for criterion referenced test. New York, USA: Editorial Measurmet.

Hernández; Fernández y Baptista. (2003). Metodología de la investigación. México: Editorial McGraw-Hill.

Quezada, N. (2010). Metodología de la investigación-Estadística aplicada en la investigación. Lima, Perú: Empresa Editora MACRO.

Sánchez y Reyes. (2006). Metodología y Diseños en la Investigación Científica. Lima, Perú: Editorial Visión Universitaria.

Schiefelbein, E. (1999). Teoría, técnicas, procesos y casos en el planeamiento de la educación. Buenos Aires, Argentina: Editorial Ateneo.

Santisteban, C. (2009). Principios de psicometría. Madrid, España: Editorial Síntesis.

Tembrick, T. (1981). Evaluación. Madrid, España: Editorial Narcea.

Thorndike y Hagen. (1991). Medición y evaluación en psicología y educación. México: Editorial Trillas.

Visauta y Martori i Cañas. (2003). Análisis estadístico con SPSS para windows, (Volumen II). Madrid, España: Editorial McGraw-Hill.

REFERENCIAS DE INTERNET

Quantitative data processing: Churchman / Ackoff / Arnoff

<http://www.yale.university.research.center.org.usa>, septiembre, 2007.