

# ANÁLISIS ESTADÍSTICO DE REACTIVOS Y CALIDAD DE LAS PRUEBAS OBJETIVAS. APLICACIÓN AL EXAMEN DE ADMISIÓN EN LA ESPOCH

Dr. Rubén Pazmiño Maji

Facultad de Ciencias. InAEM. ESPOCH  
rpazmino@epoch.edu.ec

## RESUMEN

Este trabajo tiene como objetivo el análisis estadístico de los reactivos<sup>2</sup> utilizados en el examen de admisión universitaria para el periodo marzo-agosto 2012 en las cinco carreras de la Facultad de Ciencias de la ESPOCH. La calidad es analizada luego de la ejecución de los exámenes comunes de admisión, donde el cuestionario está formado por 30 preguntas diferentes para uno de los 546 aspirantes, pero todas tomadas de un banco común de 615 reactivos. Los parámetros medidos son el índice de dificultad, la desviación típica y el índice de discriminación considerados desde el punto de vista de la teoría clásica de los ítems. Los resultados de éste trabajo permitirán la autoevaluación del proceso de admisión con el objetivo de la mejora continua de la calidad, además ilustrará su aplicación y optimización en procesos similares de evaluación basada en cuestionarios objetivos. Los resultados obtenidos responden a las características propias del grupo de aspirantes y son un referente para otra población similar. Es importante trabajar en un proceso de estandarización cuyo instrumento no dependa tan fuertemente de la población objetivo.

**Palabras clave:** cuestionario, examen de admisión, índice de dificultad, índice de discriminación.

## ABSTRACT

This work aims to statistical analysis of the reagents used in the university entrance examination for the period March to August 2012 in the five races of the Faculty of Sciences ESPOCH. The quality is analyzed after the execution of the common tests for admission where the questionnaire consists of 30 different questions for one of the 546 applicants, but all taken from a common bank of 615 reagents. The measured parameters are the index of difficulty, the standard deviation and the discrimination index considered from the standpoint of the classical theory of Items. The results of this work will allow the self-assessment of the admission process with the aim of continuous improvement of Quality, also illustrate the application and optimization in similar processes of evaluation based on objective questionnaires. The results correspond to the characteristics of the applicant pool and are a reference for other similar population. It is important to work in a standardization process for which the instrument does not depend so heavily on the target population.

**Keywords:** questionnaire, entrance examination, index of difficulty, discrimination index.



## INTRODUCCIÓN

Considero que la calidad es el estado óptimo deseado y consensuado de determinado proceso. Esta percepción personal debe tener un nivel de objetividad, que puede estar dado mediante ciertas mediciones relacionadas con el proceso. El trabajo propone la mejora de la calidad del instrumento de medición, que son los reactivos. Estos reactivos forman parte del banco de reactivos de la Facultad y servirán para generar los cuestionarios personales para los postulantes a la Facultad de Ciencias de la ESPOCH.

La calidad de los reactivos se medirá desde tres estadísticos básicos:

1. El Índice de Dificultad (ID): Que determina el grado de dificultad de los reactivos. Es deseable cuestionarios con reactivos que varíen en dificultad, pero que no sean ni fáciles ni difíciles. Se han considerado como reactivos difíciles aquellos que tienen un índice de dificultad de 15% o menos y reactivos fáciles aquellos con un índice de dificultad de 85% o más (Matos, 2007). Formalmente, el Sistema Moodle calcula este índice de la siguiente manera:

$$ID = (X_{promedio}) / X_{máx}$$

donde  $X_{promedio}$  es la media obtenida por todos los estudiantes que contestan el ítem, y  $X_{máx}$  es la puntuación máxima obtenida en el ítem. Si los reactivos pueden distribuirse dicotómicamente en categorías correcto/incorrecto, este parámetro coincide con el porcentaje de usuarios que responden correctamente a la pregunta. (Moodle)

2. La desviación típica (DT): Determina cuáles reactivos se respondieron en forma idéntica. Es muy poco probable que un mismo reactivo sea contestado de la misma manera por un grupo diferente. Formalmente, para el sistema Moodle, DT se calcula como la desviación estándar para la muestra de puntuaciones fraccionadas (correctas/máxima) para cada pregunta particular (moodle).

3. El Índice de discriminación (IDISC): Está dado por la capacidad de que el reactivo discrimine al grupo superior de las mejores notas y al grupo inferior con peores notas. Un índice de cero indicaría que el reactivo no discrimina, y cualquiera lo puede responder. Es deseable un valor de 0,4 (Matos, 2007). Para éste trabajo se ha considerado un valor de

0,2. El parámetro se calcula dividiendo en primer lugar a los respondientes en tercios, basándose en la puntuación global obtenida en el cuestionario. A continuación se calcula la media obtenida en el ítem analizado por los grupos superior e inferior, y se sustrae la me-

dia. La expresión matemática es

$$IDISC = \frac{X_{superior} - X_{inferior}}{N}$$

La ESPOCH desde hace algunos años ha llevado a cabo las pruebas de admisión, basada sobre todo en conocimientos, con componentes comunes de cultura

general y reactivos de razonamiento Lógico. Cada facultad y escuela ha definido su propio banco de reactivos, que los ha implementado en el sistema de cuestionarios de Moodle.

<sup>1</sup> Escuela Superior Politécnica de Chimborazo

<sup>2</sup> En este trabajo se utiliza como sinónimos: Reactivos, Items o Preguntas objetivas

En particular en la Facultad de Ciencias existe un Banco común formado por 615 reactivos repartidos de la siguiente forma:

BIOLOGIA	QUIMICA	MATEMATICA	FISICA	CULTURA GENERAL	RAZONAMIENTO LOGICO
114	105	122	124	75	75

Las pruebas se rindieron durante el 23 y 24 de Febrero del 2012 (no se consideraron las pruebas de desempate). Los aspirantes a estudiantes de la Facultad de Ciencias estuvieron distribuidos de la siguiente forma en las 5 Escuelas de la Facultad de Ciencias:

FACULTAD DE CIENCIAS (546)				
ESTADISTICA INFORMÁTICA	BIOFÍSICA	INGENIERÍA QUÍMICA	BIOTECNOLOGÍA AMBIENTAL	BIOQUIMICA Y FARMACIA
214	7	102	159	217

En éste trabajo se analizan las 2725 mediciones y los 615 Reactivos en forma global y por área de conocimiento: Biología, Química, Matemática, Física, cultura general y Razonamiento lógico.

**MATERIALES Y MÉTODOS**

El colectivo de estudio está dado por 615 reactivos que conforman el total, considerado las pruebas generadas en el proceso de admisión en la Facultad de Ciencias. La población está conformada por 2725 mediciones, generadas por los 546 estudiantes aspirantes a las 5 carreras de la Facultad de Ciencias de la ESPOCH. La recolección de la información se realizó mediante los informes generados por el sistema gestor de contenidos Moodle. La depuración, organización y filtrado de las bases de datos fue realizada en el software Libre OfficeCalc, versión 3.4, y el análisis estadístico se realizó en SPSS, versión 19. La metodología utilizada en el análisis de

reactivos es la dada por la Teoría Clásica de los Ítems y se analizaron los Índices de Dificultad, la Desviación Típica y el Índice de Discriminación.

**RESULTADOS Y DISCUSIÓN**

Es importante notar que el banco de 615 reactivos considerados en la prueba de admisión, generaron una base de datos de 2725 mediciones debido a que la prueba fue tomada con 30 reactivos en forma aleatoria, esto provocó que una misma pregunta se repitiese en algunos cuestionarios y en varias escuelas, teniendo elaborar un índice de dificultad único por reactivos partiendo de los diferentes informes dados por el sistema Moodle.

El índice de dificultad, en su forma más simple e intuitiva, se puede describir como el porcentaje de personas que han respondido un determinado reactivo. Si el porcentaje es pequeño indicará que la pregunta es difícil y si el porcentaje es grande significa que la pregunta es fácil. Para nuestro estudio estamos considerando un valor del índice de dificultad menor o igual que 15 como difícil y un valor del índice de dificultad mayor o igual que 85 como fácil. Los reactivos cuyo  $ID \leq 15$  o  $ID \geq 85$ , deben considerarse ser reformulados o eliminados del banco común luego de un estudio más minucioso.

De los 615 reactivos que conforman la prueba de admisión aproximadamente una cuarta parte de ellos, el 23,9 %, deben estudiarse, para decidir si ameritan reformularse o eliminarse por ser difíciles o fáciles. Las tres cuartas partes de reactivos a nivel de toda la facultad están en un índice de dificultad aceptable. Las siguientes ilustraciones muestran lo analizado.

A_ID	Recuento	% del N de la columna
Pregunta Difícil	35	5,7%
Aceptable	468	76,1%
Pregunta Fácil	112	18,2%

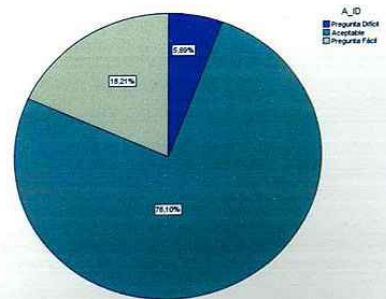


Ilustración 1: Tabla y Gráfico circular sobre la frecuencia de Reactivos mediante el Índice de Dificultad

Se puede notar en el gráfico de cajas y alambre que la media de dificultad de los ítems es de 56,86, bastante cercana a lo ideal, que es 50. Además, como en el coeficiente de asimetría indica que es levemente asimétrica negativa y también es platicúrtica, esto corrobora el resultado de las pruebas de normalidad de que los índices de dificultad de los diferentes ítems no se distribuyen normalmente como lo deseáramos. El histograma así lo muestra gráficamente.

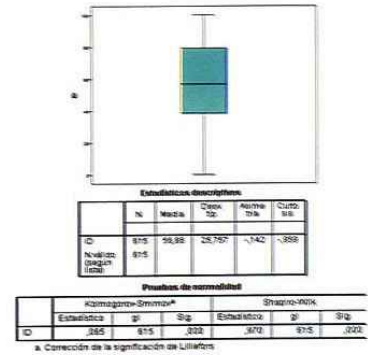
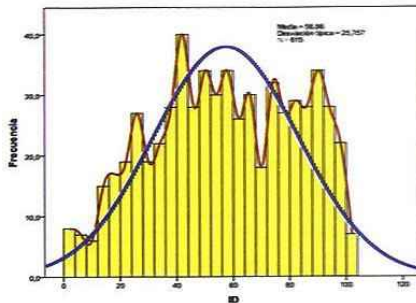


Ilustración 1: Tabla y Gráfico circular sobre la frecuencia de reactivos mediante el índice de Dificultad

Si realizamos el estudio del índice de dificultad por área podemos ver que las áreas de cultura general y razonamiento lógico tienen demasiados reactivos fáciles que deberían ser reformuladas o eliminadas previo su estudio. En química llama la atención el número de reactivos fáciles, mientras en matemática, y sobre todo en física, el número de reactivos difíciles. El siguiente cuadro resume lo dicho:

AREA	A_DESVT	
	Inaceptable	Aceptable
BIOLOGIA	1	113
FÍSICA	2	122
MATEMÁTICA	6	116
QUÍMICA	2	103
Total Especialidad	11	454
CULTURA GENERAL	4	71
RAZONAMIENTO LÓGICO	2	73
Total Generales	6	144
Total Facultad de Ciencias	17	598

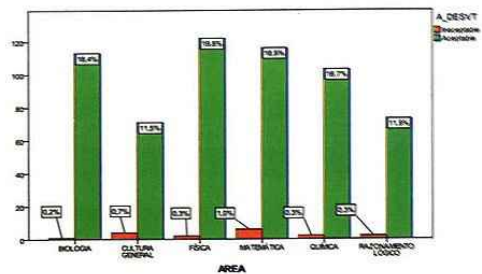


Ilustración 3: Frecuencias de preguntas fáciles, difíciles y aceptables por área

Sobre el rendimiento medio por área, se puede analizar que tanto biología como matemática mantienen una dificultad media de aproximadamente 50, que es lo deseado. Los reactivos de química se alejan superiormente de 50 y son más difíciles que los de biología, física y matemática. Los reactivos de física son los más difíciles, poco menos de la mitad los pueden contestar (el 40%). Los reactivos de cultura general y razonamiento lógico en general son fáciles, llaman la atención los reactivos de los casos 591 (correspondiente al reactivo 44), 587 (correspondiente al reactivo 47) y 591 (correspondiente al reactivo 51) que son difíciles en su grupo. El siguiente gráfico muestra lo analizado:

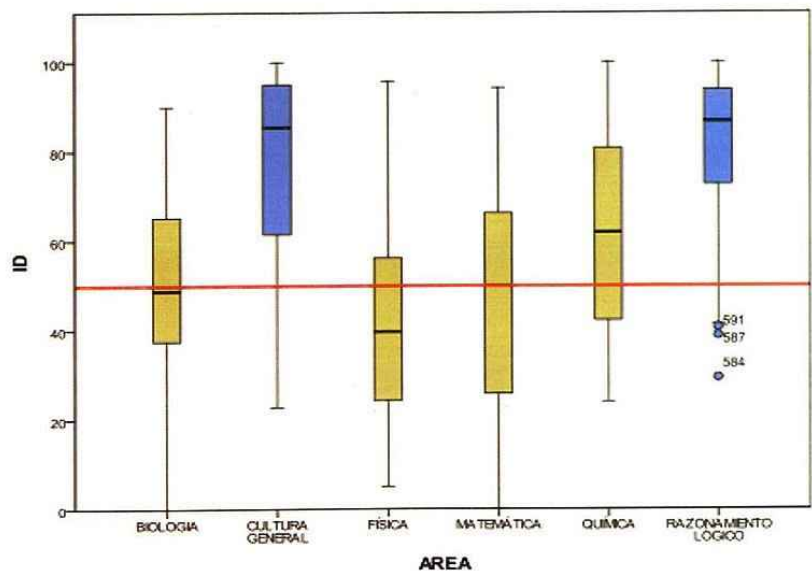


Ilustración 4: Gráfico de cajas y alambres comparativo de la dificultad media entre áreas

Solamente 11 reactivos de los 615 son respondidos por los estudiantes exactamente igual, la mayoría de ellos son en matemática seguidos por los de cultura general. Esto se debe a que quizás son muy evidentes sus respuestas debido a que los distractores no fueron seleccionados adecuadamente u otras razones externas a los reactivos. El siguiente cuadro muestra detalladamente lo analizado:

	A_ID			
	Pregunta Difícil	Pregunta Fácil	Inaceptable	Aceptable
BIOLÓGICA	8	5	11	103
CULTURA GENERAL	0	38	38	37
FÍSICA	15	3	18	108
MATEMÁTICA	14	10	24	98
QUÍMICA	0	15	15	90
RAZONAMIENTO LÓGICO	0	41	41	34

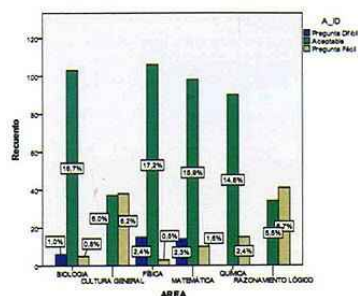


Ilustración 5: Frecuencia de los Reactivos que fueron contestados idénticamente por área

Respecto a los reactivos que discriminan a los aspirantes con buenas calificaciones de aquellos con malas calificaciones se puede ver que hay solo un 14,1% de reactivos que no discriminan adecuadamente y que se debería analizar su reformulación o eliminación. Aproximadamente el 86% de reactivos son buenos discriminadores. Es importante notar que se trabajó con un nivel de 0,2 como límite inferior para ser un buen discriminador, aunque lo deseable es de 0,4.

ÁREA	BIOLÓGICA	A_IDISC			
		Aceptable	Óptimo	Aceptables+Óptimos	Inaceptable
		Recuento	Recuento	Recuento	Recuento
		82	34	98	18
	FÍSICA	72	34	108	18
	MATEMÁTICA	73	38	109	13
	QUÍMICA	83	33	98	9
	Total Especialidades	270	137	407	68
	CULTURA GENERAL	48	13	81	14
	RAZONAMIENTO LÓGICO	45	15	80	15
	Total Generales	93	28	121	29
	Total Facultades de Ciencias	353	165	528	87

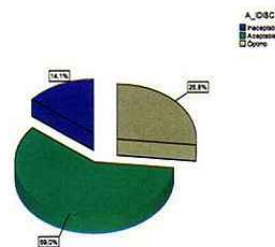


Ilustración 6: Tabla y gráfico de frecuencias de los aceptables, óptimos e inaceptables desde el punto de vista del índice de discriminación por área y globales.

El análisis realizado sólo explora los rasgos básicos de la teoría del análisis de reactivos. El análisis completo debe realizarse en el contexto en el que se realizó la prueba, y luego del contraste conjunto determinar cuáles reactivos deben adecuarse o eliminarse. Se tiene además la información de los tres ámbitos analizados (Índice de dificultad, Desviación Típica e Índice de Discriminación) por cada uno de los 615 Reacti-

vos que permitirá un estudio minucioso por cada uno de ellos.

Esta investigación desea mostrar cómo analizar objetivamente los reactivos para poder aplicarlos luego como un instrumento de medición válido, eficiente, de calidad y consensado para que cumpla los objetivos implícitos en él.

### AGRADECIMIENTOS

Este trabajo está dedicado a Carmita, Cris y Gaby por su paciencia, comprensión y por soportar tantas horas de trabajo. Un agradecimiento a las autoridades de la Facultad de Ciencias, a mis estudiantes y a todos aquellos que deseen aportar con la mejora de éste trabajo.

### REFERENCIAS BIBLIOGRÁFICAS

1. DUSSAILLANT, FRANCISCA. 2003. TÉCNICAS DE MEDICIÓN EN PRUEBAS DE ADMISIÓN A LAS UNIVERSIDADES. SANTIAGO DE CHILE : S.N., 2003.
2. López, Dr. Agustín Tristán. 2003. Modelo para el análisis de reactivos objetivos por computadora. www.ieesa-kalt.com. [En línea] 2003. [Citado el: 20 de Marzo de 2011.] www.ieesa-kalt.com.
3. Matos, Menigno Hidalgo. 2007. Nueva Cultura Evaluativa. Nueva Cultura Evaluativa. Perú : Instituto de Desarrollo Humano, 2007.
4. Moodle, wiki. WIKI COLABORATIVA MOODLE. [En línea] WIKI MOODLE. [Citado el: 26 de Marzo de 2012.] http://docs.moodle.org/all/es/Informe\_del\_cuestionario.
5. Stage, Christina. 2003. Teoría clásica de la medición o teoría de Respuesta al ítem. Umea-Suecia : s.n., 2003.