

Análisis cuantitativo de las preguntas de una prueba estandarizada para observar las dificultades del cálculo de probabilidades en estudiantes universitarios

Quantitative analysis of a standardized test aimed to identify university students' difficulties in computing probabilities

Liliana García Barco¹ y Néstor Fernando Méndez Hincapié²

¹Universidad Sergio Arboleda. ²Universidad Pedagógica Nacional, Colombia

Resumen

Se analizaron tres preguntas de una prueba estandarizada para identificar los errores y concepciones sobre la interpretación de tablas y gráficos para el cálculo de probabilidades, de los estudiantes de una universidad colombiana de las carreras de Marketing, Economía e Ingenierías; después de que estos ven un primer curso introductorio de estadística. Este estudio se basa en los elementos lingüísticos (tablas o gráficos) para abordar los conceptos de probabilidad en la enseñanza de la estadística y de los niveles de comprensión gráfica que los estudiantes pueden alcanzar, identificados por Pinto (2010) y Batanero (2001 y 2005). Se encontró que la mayoría de los estudiantes aún no alcanzan el segundo nivel de comprensión gráfica y de tablas una vez finalizado este curso.

Palabras clave: Probabilidad, elementos lingüísticos, pruebas estandarizadas, niveles de comprensión gráfica.

Abstract

Three questions of a standardized test were analyzed to identify the errors and conceptions about the interpretation of tables and graphs for the calculation of probabilities, of the students of a Colombian university of the disciplines of Marketing, Economy and Engineering; after they see a first introductory course in statistics. This study is based on the linguistic elements (tables or graphs) to address the concepts of probability in the teaching of statistics, and the levels of graphic comprehension that students can achieve, identified by Pinto (2010) and Batanero (2001 and 2005). It was found that most of the students still do not reach the second level of graphic comprehension and tables once this course is finished.

Key words: Probability, linguistic elements, standardized tests, levels of graphic comprehension.

1. Introducción

En la Universidad Sergio Arboleda – USA (Colombia), durante los períodos semestrales de 2014-2 hasta el periodo 2016-2, se diseñaron exámenes finales con preguntas tipo test, con cuatro opciones de respuesta y sólo una correcta, para el curso introductorio de Estadística que se ofrece a varias carreras de las escuelas de ingeniería, marketing y economía de esta universidad. Estas pruebas y sus preguntas fueron validadas bajo la teoría clásica de los test (TCT) y por la teoría de respuesta al ítem (TRI).

Las temáticas abordadas en los exámenes finales fueron: estadística descriptiva, fundamentos de probabilidad, variables aleatorias, distribuciones discretas de probabilidad y conceptos básicos de distribuciones continuas de probabilidad. Por otro lado, en las preguntas, los distractores (opciones de respuesta incorrectas) fueron diseñados de tal forma que se incluyesen los errores más comunes que tienen los

García Barco, L. y Méndez Hincapié, N. (2019). Análisis cuantitativo de las preguntas de una prueba estandarizada para observar las dificultades del cálculo de probabilidades en estudiantes universitarios. En J. M. Contreras, M. M. Gea, M. M. López-Martín y E. Molina-Portillo (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística*. Disponible en www.ugr.es/local/fqm126/civeest.html

estudiantes al enfrentarse a problemas del cálculo de probabilidades y de la estadística. Estos a su vez fueron concertados por el grupo de profesores de acuerdo con sus experiencias anteriores.

La Universidad Sergio Arboleda tiene una tradición investigativa en temas relacionados con las actitudes y aptitudes de sus estudiantes en las áreas de matemáticas y estadística (Pérez, 2008; Pérez, Niño y Páez, 2010; citados en Pérez, Aparicio, Bazán y Abdounur, 2015) y por ello ha venido aplicando instrumentos psicométricos, tanto de actitudes propuesta por Aparicio y Bazán (1997) citados por Pérez, *et al* (2015) como de rendimiento óptimo (García, 2014).

En el caso particular del área de estadística, se aplicó en el primer semestre de 2013 una prueba para medir las actitudes hacia la estadística a 545 estudiantes de primer semestre de la USA que cursaron carreras de economía, administración e ingenierías, con edades entre 17 y 25 años. Esta investigación consideró tres escalas: la de Estrada (2002), la de Cazorla y otros (1999) y una escala conjunta basada en las dos (Pérez, *et al*, 2015). Los resultados de esta investigación en relación con las actitudes de los estudiantes hacia la Estadística mostraron que ellos reconocen que esta es de utilidad, que puede ser usada en diferentes disciplinas y que se requiere que la universidad contribuya con su preparación en este campo. Sin embargo, se mostró también en esta investigación el poco uso, habilidad y gusto por la clase de estadística (Pérez, *et al*, 2015).

Como resultado de esa investigación, se planteó la necesidad de elaborar “talleres y programas encaminados a cambiar las actitudes negativas y a reforzar los conocimientos más débiles.” (Pérez, *et al*, 2015, p. 144). Con esta idea aquí planteada, y con lo desarrollado dentro de esta institución sobre las pruebas psicométricas, se diseñaron evaluaciones finales tipo test de rendimiento óptimo a los estudiantes que finalizaban la asignatura Estadística I en los períodos 2014-2 a 2016-2 inclusive, y cuya particularidad en el desarrollo de esta materia era mostrar problemas de aplicación que le dieran sentido al estudiante para reforzar la habilidad y uso de la estadística.

Ahora bien, para reforzar dicha habilidad, nos acogemos a lo expuesto por Pinto (2010), sobre la importancia de que los estudiantes alcancen “habilidades para leer y comprender gráficos y tablas estadísticas” (p. 99) ya que “es la primera categoría de alfabetización cuantitativa (...) a desarrollar en los estudiantes” (p. 101), dado que la sociedad actual se mueve en un mundo de la información (periódicos, revistas, redes sociales, TV, entre otros) donde estos elementos lingüísticos aparecen y se requiere de la habilidad lectora para su comprensión. Por eso en este trabajo nos centramos en analizar preguntas concernientes a la interpretación de gráficos y tablas.

Por lo anterior, y continuando con el diagnóstico a los estudiantes de la USA, en la presente investigación se describe el análisis psicométrico de la prueba realizada en el semestre 2014-2 y posteriormente el análisis de algunas preguntas de este examen para identificar los errores más comunes que siguen presentando los estudiantes relacionados con la interpretación de tablas y gráficos y su razonamiento en el cálculo de probabilidades, después de haber finalizado el curso, con la idea de entregar insumos que permitan mejorar la enseñanza de esta materia.

En el apartado 2 de este artículo se hace un breve resumen del surgimiento de la probabilidad y su enseñanza, enfatizando en los elementos lingüísticos necesarios para su representación, análisis e interpretación con fines didácticos. Posteriormente se hace una corta explicación del análisis psicométrico bajo la TCT y la TRI utilizadas como modelos para analizar el examen realizado en el semestre 2014-2 para validar la prueba

en conjunto y sus preguntas. Se finaliza explicando el método de investigación utilizado en este trabajo y la descripción y discusión de los resultados alcanzados realizando el análisis detallado a tres preguntas relacionadas con el tema de probabilidad y su interpretación y razonamiento a través de tablas y gráficos. En el numeral 3 se presentan las conclusiones.

2. Marco teórico

A pesar de la formalización matemática de la probabilidad, su desarrollo histórico muestra que los diferentes significados que se produjeron en esta surgieron de ideas intuitivas que fueron evolucionando a lo largo de la historia para resolver problemas vinculados al azar. Esto mismo sucede con el desarrollo del aprendizaje que tienen los estudiantes: ideas intuitivas que deben evolucionar para construir concepciones adecuadas del cálculo de probabilidades, los cuales deben ser entendidos y asumidos por los profesores frente al proceso de enseñanza (Batanero, 2005).

Siguiendo a Batanero (2005) sobre los significados de probabilidad, en el curso de Estadística I de la USA se abordan las definiciones clásica, frecuencial y subjetiva de la probabilidad, y en el desarrollo de estas definiciones se enfatiza en los elementos lingüísticos (elementos que le dan significado a un objeto matemático “como símbolos, palabras o gráficos para representar los datos y soluciones, al igual que las operaciones y conceptos usados” (Batanero, 2005; p 250); por ello en este trabajo se muestran los resultados de algunas preguntas bajo los elementos lingüísticos (gráficos y tablas) que hacen referencia al cálculo de probabilidades compuestas como es el caso del principio aditivo y el principio multiplicativo.

Pinto (2010) identificó y agrupó seis aspectos diferentes en que se han hecho investigaciones sobre representaciones gráficas en la enseñanza de la estadística. Dos de ellos son: “investigaciones sobre el estudio de las concepciones, errores y dificultades de los estudiantes”, y el otro aspecto es el “desarrollo de instrumentos para medir la comprensión gráfica” (p. 103). Aquí, con los instrumentos ya elaborados, que son pruebas estandarizadas validadas, se intenta exponer los errores y dificultades más comunes de los estudiantes de la USA. Además, las preguntas propuestas las utilizamos como instrumento para medir la comprensión gráfica y de tablas de los estudiantes.

Por otro lado, Batanero (2001), citando a Curcio (1989), expone que existen cuatro niveles de comprensión de gráficos o tablas que son:

- Primer nivel: *leer los datos*, lectura literal del gráfico o tabla sin interpretación de estos.
- Segundo nivel: *leer dentro de los datos*, requiere interpretación de datos, comparación de cantidades y uso de otros conceptos.
- Tercer nivel: *leer más allá de los datos*, asociado a predicciones e inferencias que no están expuestas directamente en los gráficos o tablas.
- Cuarto nivel: *leer detrás de los datos*, relacionado con la confiabilidad de los datos.

Esta investigación se relaciona con el cálculo de probabilidades asociado a las representaciones gráfica para observar el nivel en que se encuentran los estudiantes de la USA.

2.1. Psicometría

La psicometría recoge el conjunto de los diferentes modelos que hacen posible la medición de variables psicológicas por medio de test, y los procesos necesarios para que se realice de forma adecuada (Martínez, 1995). Se asume que es posible cuantificar un rasgo psicológico o conducta de los individuos, que en nuestro caso es el conocimiento que tienen los estudiantes de la USA sobre probabilidad a partir del examen final de la asignatura Estadística I.

Hay dos modelos de mayor importancia: la teoría clásica de los test (TCT) y la teoría de respuesta al ítem (TRI) (Martínez, 1995). La TCT es un modelo clásico que se basa en que la *puntuación observada* (resultado de una prueba o test) está compuesta por la suma de la *puntuación verdadera* (o puntuación real de la persona en el rasgo medido) y del *error de medida*. La TRI se basa en el supuesto clave de que hay una relación funcional entre la competencia de los sujetos, con la probabilidad de responder los ítems correctamente, denominada curva característica del ítem CCI. Aquí se trabajará el modelo de Rasch o modelo de un parámetro porque se tiene una prueba de 111 estudiantes (Chávez y Saade, 2010).

Según Chávez y Saade (2010) hay al menos dos indicadores básicos de las características de un ítem o reactivo: Primero, la *dificultad*, es decir, un índice que indica qué tan fácil o qué tan difícil resulta un ítem o reactivo para una población. Segundo, la discriminación, que indica qué tan eficiente es un ítem para diferenciar entre una persona que sabe de otra que no.

Desde la TCT la dificultad se mide a través del índice de dificultad definido como n_A/n_r donde n_A es el número de sujetos que respondieron acertadamente al ítem y n_r es el número de sujetos que respondieron el mismo. La homogeneidad se calcula como la correlación de Pearson ítem-test. La Tabla 1 muestra las posibles clasificaciones de estos índices tomados de Bazán (2000) y Chávez y Saade (2010).

Tabla 2. Clasificación de los niveles de dificultad y homogeneidad de las preguntas según Bazán (2000) y Chávez y Saade (2010)

Índice de dificultad	Valores	Clasificación
Muy difícil	[0,00 – 0,25)	MD
Difícil	[0,25 – 0,45)	D
Normal	[0,45 – 0,55)	N
Fácil	[0,55 – 0,75)	F
Muy fácil	[0,75 – 1,00]	MF

Tabla 2. Clasificación de los niveles de dificultad y homogeneidad de las preguntas según Chávez y Saade (2010)

Índice de homogeneidad	Valores	Clasificación
Pésima descartar	[-1,00 0,20)	PD
Regular revisar	[0,20 – 0,29)	RR
Buena mejorar	[0,29 – 0,39)	BM
Conservar	[0,39 – 1,00]	C

Desde la TRI, el modelo de Rasch establece que la CCI es una función logística y sólo tiene en cuenta el parámetro de dificultad b_i y asume que el índice de discriminación es 1 e igual para todos los ítems en este modelo.

2.2. Método de investigación.

El trabajo desarrollado consistió en una etapa de diagnóstico de los problemas que presentan los estudiantes de la Universidad Sergio Arboleda que han cursado la asignatura Estadística I. De acuerdo con lo visto por los propios docentes en el desarrollo de esta, se elaboró un examen final tipo test de rendimiento óptimo de 20 preguntas cada una con cuatro opciones de respuesta, una de ellas correcta. Las otras opciones o distractores se escogieron según los errores comunes que los propios docentes a lo largo de este curso y de su experiencia previa, han detectado que cometen los estudiantes con más frecuencia. De esta manera se tiene información sobre un grupo mayoritario de estudiantes de esta Institución que cursan esta asignatura en la jornada diurna.

Una vez realizadas las preguntas, se implementaron en todos los cursos a manera de examen final que representa parte de la nota final del curso. Para este examen posteriormente se realizó un análisis psicométrico bajo las teorías Clásica de los Tests (TCT) y la Teoría de Respuesta al Ítem (TRI) para comprobar la buena calidad de los ítems, reactivos o preguntas.

Finalmente se analizan las respuestas a los ítems para determinar las dificultades que presentan los estudiantes, según las opciones escogidas, que servirán de insumo para diseñar estrategias metodológicas de enseñanza en cursos posteriores.

2.3. Descripción y discusión de resultados

Se aplicó un test de rendimiento óptimo como examen final de la asignatura estadística I a 111 estudiantes de carreras de economía, ingeniería y administración en el período 2014-2. Esta prueba consta de 20 preguntas de elección múltiple con cuatro opciones de respuesta. La prueba abarcó temáticas como: análisis de gráficos y tablas, probabilidad, probabilidad condicional, teorema de Bayes y algunas distribuciones discretas de probabilidad.

Los resultados del análisis del índice de dificultad e índice de homogeneidad bajo la TCT se muestran en la Tabla 3. Se utilizó el programa BILOG-MG para Windows versión 3.0.2327.2 para realizar el análisis psicométrico que se muestran hasta la pregunta 9 en la Figura 1.

Tabla 3. Resultados de los índices de dificultad y homogeneidad obtenidos en Excel.

Pregunta	1	2	3	4	5	6	7	8	9	10
Índice de dificultad	0,94 MF	0,39 D	0,34 D	0,92 MF	0,43 D	0,52 N	0,56 F	0,50 N	0,55 N	0,61 F
Índice de homogeneidad	0,12 PD	0,32 BM	0,42 C	0,20 RR	0,36 BM	0,45 C	0,34 BM	0,36 BM	0,43 C	0,46 C
Pregunta	11	12	13	14	15	16	17	18	19	20
Índice de dificultad	0,48 N	0,41 D	0,34 D	0,72 F	0,48 N	0,58 F	0,40 D	0,39 D	0,44 D	0,53 N
Índice de homogeneidad	0,33 BM	0,38 BM	0,44 C	0,14 PD	0,14 PD	0,19 PD	0,43 C	0,30 BM	0,41 C	0,34 BM

La Figura 1 muestra el ajuste bajo la TRI de un parámetro o modelo de Rasch dado que sólo se tienen 111 sujetos (Chávez y Saade, 2010). De acuerdo con este análisis, la dificultad, parámetro b , tuvo valores entre -2.688 y 0.901 lo que indica que en general la

prueba no fue muy difícil y por el contrario dos preguntas resultaron bastante fáciles, los ítems 1 y 4.

ITEM	Parámetro a		Parámetro b		Parámetro c		Valor p de la prueba Ji	
	INTERCEPT S.E.	SLOPE S.E.	THRESHOLD S.E.	LOADING S.E.	ASYMPTOTE S.E.	CHISQ (PROB)	DF	
ITEM0001	2.688 0.393*	1.000 0.067*	-2.688 0.681*	0.707 0.048*	0.000 0.000*	0.1 (0.9429)	2.0	
ITEM0002	-0.648 0.204*	1.000 0.067*	0.648 0.354*	0.707 0.048*	0.000 0.000*	7.5 (0.3780)	7.0	
ITEM0003	-0.901 0.215*	1.000 0.067*	0.901 0.373*	0.707 0.048*	0.000 0.000*	3.7 (0.7182)	6.0	
ITEM0004	2.084 0.314*	1.000 0.067*	-2.084 0.543*	0.707 0.048*	0.000 0.000*	6.0 (0.1127)	3.0	
ITEM0005	-0.487 0.203*	1.000 0.067*	0.487 0.352*	0.707 0.048*	0.000 0.000*	9.6 (0.2141)	7.0	
ITEM0006	-0.095 0.205*	1.000 0.067*	0.095 0.356*	0.707 0.048*	0.000 0.000*	4.9 (0.5629)	6.0	
ITEM0007	0.061 0.201*	1.000 0.067*	-0.061 0.348*	0.707 0.048*	0.000 0.000*	1.6 (0.9545)	6.0	
ITEM0008	-0.173 0.201*	1.000 0.067*	0.173 0.347*	0.707 0.048*	0.000 0.000*	1.5 (0.9592)	6.0	
ITEM0009	0.022 0.205*	1.000 0.067*	-0.022 0.355*	0.707 0.048*	0.000 0.000*	3.3 (0.7732)	6.0	

Figura 1. Resultado de los parámetros del modelo Rasch bajo la TRI obtenidos en BILOG para los nueve primeros ítems.

Por otro lado, el valor p de la prueba ji muestra (valor entre paréntesis de la columna 7 de la Figura 1) que todos superan el valor 0.05 lo que significa que se ajustan bien al modelo de Rasch (Chávez y Saade, 2010).

El anterior análisis nos permite concluir que bajo la TRI la prueba diseñada para este período se ajustó muy bien al modelo de un parámetro. Sin embargo, desde la TCT no se logró el objetivo de discriminación necesario para la mitad de la prueba. Aun así, el objetivo de esta prueba no es discriminar ya que lo que se busca es determinar los posibles errores o concepciones que tienen los estudiantes de la Universidad Sergio Arboleda después de su primer curso de Estadística.

En este sentido, y para los propósitos del presente trabajo, se analizaron las preguntas 2, 3 y 9 que corresponde a ítems de temas relacionados con la comprensión a primer y segundo nivel de las tablas y gráficas, donde se pide calcular probabilidades. Las preguntas 2 y 3 resultaron difíciles (D) y la 9 normal (N) de acuerdo con la escala de clasificación de la Tabla 1 y los resultados de la Tabla 2. Se quiere mirar el comportamiento de los distractores para dar luces sobre los errores o concepciones que cometen los estudiantes sobre la interpretación de tablas y gráficas en el cálculo de probabilidades o porcentajes.

2.4. Análisis

Ya realizado el análisis psicométrico de la prueba y el conjunto de preguntas, se escogieron las preguntas 2, 3 y 9 para este análisis porque están relacionadas con el

tema de gráficos o tablas, y además presentaron índices de dificultad para los estudiantes Difícil o Normal, y sus índices de homogeneidad fueron superiores a 0.3, Buena Mejorar o Conservar, que indica que las preguntas son adecuadas.

Se quiere observar y analizar la distribución de las opciones de respuesta que dieron los estudiantes a las preguntas 2, 3 y 9 como se muestra en la Figura 2.

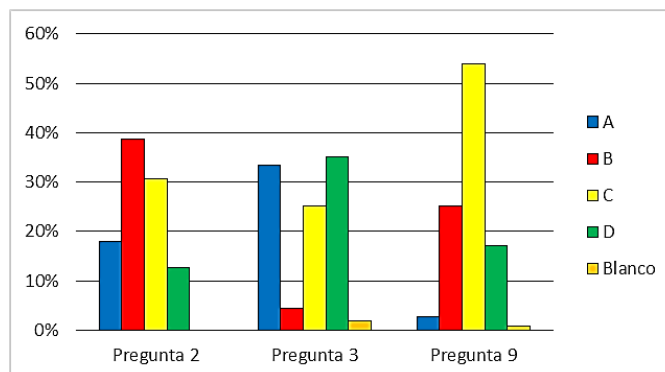


Figura 2. Distribución de las respuestas de los estudiantes a las preguntas 2, 3 y 9

Análisis de la pregunta 2

Pregunta 2. Una empresa tiene un inventario de máquinas que se clasifican en N: nuevas, U: usadas, A: automáticas y M: manuales, las cantidades de máquinas con estas características está dada en la siguiente tabla.

	N: Nuevas	U: Usadas
A: Automáticas	200	60
M: Manuales	40	100

Si se selecciona aleatoriamente una máquina, la probabilidad de que sea manual o usada es:

A. 0.25 B. 0.50 C. 0.75 D. 0.40 (Opción correcta B)

Esta pregunta pide encontrar la probabilidad de la unión entre M (manual) y U (usada). La opción B es la correcta. Fue respondida por 43 estudiantes, casi el 39%, que corresponde a la suma de las máquinas manuales y usadas (100), las máquinas manuales nuevas (40) y las máquinas usadas automáticas (60). Esto indica que más de la mitad de los estudiantes no llegan al segundo nivel de comprensión de tablas, *leer dentro de los datos*, porque no se alcanza la destreza en el cálculo de probabilidades.

El distractor que tuvo mayor opción, después de la opción correcta, fue C, como se muestra en la Figura 2, de aproximadamente el 31%, y que corresponde a la suma de todas las máquinas manuales (140) con todas las máquinas usadas (160), que son 300. Es decir, sumaron el 100 (manuales y usadas) dos veces, y por ello llegaron al resultado 0.75.

La opción A fue respondida por el 18% y corresponde a 0.25, es decir 100 de 400 máquinas, por lo que escogieron máquinas que son simultáneamente manuales y usadas. En este caso los estudiantes confunden el operador lógico de disyunción con el de la conjunción. La opción D fue respondida por 14 estudiantes (13% aproximadamente) en la que sólo contaron las máquinas usadas, 160, y no contaron las 40 máquinas manuales nuevas.

El anterior análisis nos permite concluir que dentro de los distractores se exponen principios aditivos como la opción C que indica que este grupo de estudiantes alcanzan

a identificar que la pregunta requiere del proceso aditivo, pero no distinguen que se trata de conjuntos no disyuntos, es decir, que tienen elementos en común. Por eso contaron 140 máquinas manuales y 160 máquinas usadas sin reconocer que 100 de estas máquinas pertenecen simultáneamente a estos dos conjuntos.

Análisis de la pregunta 3

Pregunta 3. Los empleados de la compañía *ABC Ltda* están separados en tres sectores: administración, operación de planta y ventas. A continuación, se muestra la información clasificada:

	Mujer	Hombre	Total
Administración	20	30	50
O. de Planta	60	140	200
Ventas	100	50	150
Totales	180	220	400

Si se escoge una de estas personas al azar, la probabilidad de que sea mujer si trabaja en la división de ventas es:

- A. $100/150$ B. $10/150$ C. $100/400$ D. $100/180$ (Opción correcta A)

Aquí se quiere preguntar sobre una probabilidad condicional. La opción A, que es la correcta, fue escogida por 37 estudiantes, un 33.3% (según lo muestra la Figura 2) y es la segunda opción más escogida. Se debe reconocer que la frase “si trabaja en la división de ventas” es porque este es el grupo de referencia 150 personas, de las cuales 100 son mujeres.

El distractor D fue el más elegido (más del 35%), es decir, escogieron como grupo de referencia a todas las mujeres, y de ellas las que trabajen en ventas. Es decir, se comprendió al revés de lo que se quería. La opción C fue la tercera más escogida con más del 25%, se entendió sobre la probabilidad de que al escoger una persona al azar esta sea mujer y trabaje en ventas, del grupo total de empleados de la compañía.

Finalmente, menos del 5% escogió la opción B. Reconocen los estudiantes que se debe elegir entre 150 que es el número de personas que hay en la división de ventas, pero no verificaron que el numerador debe ser 100 y no 10. De todas maneras, fue un distractor de muy baja escogencia en comparación con los otros.

Esto indica que los estudiantes deben tener claro que la probabilidad no se puede calcular sobre el total de integrantes de la muestra, sino de un grupo particular al cual se hace referencia. Entonces los estudiantes no pudieron entender la pregunta más que la propia lectura de la tabla. Es decir, no se alcanza el segundo nivel de comprensión de tablas de *leer dentro de los datos*.

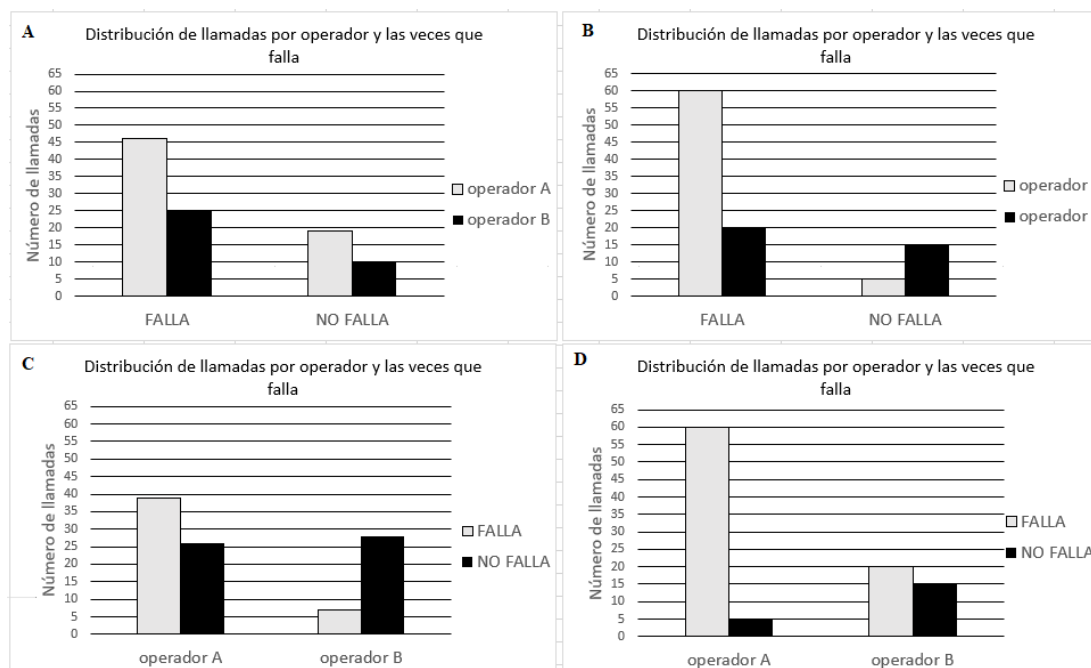
Análisis de la pregunta 9

En esta pregunta se busca que el estudiante logre una la traducción o correspondencia entre el enunciado y el gráfico, pues la traducción requiere un cambio en la forma de una comunicación (Friel, Curcio y Bright, 2001), y se observa que fue respondida correctamente, opción C, (Figura 2) por la mayoría de los estudiantes, más del 54%.

La opción B fue escogida por el 25%, muestra que estos estudiantes no realizan ninguna operación y simplemente identifican fallas de 60 y 20 de las 100 llamadas y no reparan en que si falla son porcentajes de uno u otro operador.

La opción D fue escogida por poco más del 17% y muestra el mismo problema detectado en la opción B, sólo que la gráfica invierte las variables categóricas en su representación gráfica.

Pregunta 9) Juan posee dos operadores de telefonía en su celular, el operador A y el operador B. Normalmente el 65% de las veces usa el operador A y el 35% el operador B. El primer operador falla 60% de las veces y el segundo, 20%. Si Juan hace 100 llamadas. ¿Cuál de los siguientes gráficos representa el enunciado?



Opción correcta C

La opción A fue escogida por menos del 3% y se trata de un distractor que no tiene nada que ver con los datos y se podría decir que carece de lógica dado el enunciado. Esto muestra que los estudiantes sí pretendieron responder la pregunta intentando sacar información de las gráficas para compararla con el enunciado y no respondieron al azar.

También se concluye que los estudiantes, en un alto porcentaje, no *leen dentro de los datos* porque no usan los conceptos de probabilidad para identificar la gráfica correcta.

3. Conclusiones y comentarios finales

Según lo expuesto en este trabajo, se encuentra que la mayoría de los estudiantes que respondieron a esta prueba tienen muchas dificultades alrededor de la comprensión de las tablas y los gráficos al primer y segundo nivel aún finalizado el primer curso introductorio de estadística, lo que requiere un mayor esfuerzo en proponer didácticas que permitan solucionar estos problemas. Particularmente, los estudiantes en su mayoría no comprenden la diferencia lingüística de una disyunción y una conjunción para el cálculo de probabilidades. Por otro lado presentan dificultades con leer dentro de los datos y hacer traducciones de enunciados a través de gráficos.

Este estudio se extendió a otros temas que se ven en esta asignatura, como las variables aleatorias, distribuciones de probabilidad, entre otros, que son temas tratados en estos

exámenes, pero aquí no se abordaron por la extensión. Estos resultados pueden servir como puntos de partida de otras investigaciones más detalladas y particulares que permitan desarrollar nuevas metodologías, necesarias en la investigación en educación estocástica.

4. Referencias

- Batanero, C. (2001). *Didáctica de la estadística*. Granada, España: Universidad de Granada.
- Batanero, C. (2005). Significados de la probabilidad en la educación secundaria. *Revista Latinoamericana de Investigación en Matemática Educativa*, 8(3), 247-263.
- Bazán, J. (2000). *Evaluación psicométrica de las preguntas y pruebas crecer 96*. Lima, Perú: Ministerio de Educación. Disponible en: https://www.ime.usp.br/~jbazan/download/eval_psicometrica_preguntas_pruebas_crecer_96.pdf
- Chávez, C., y Saade, A. (2010). *Procedimientos básicos para el análisis de reactivos. Cuaderno técnico 8*. México: Centro Nacional de Evaluación para la Educación Superior.
- Curcio, F. R. (1989). *Developing graph comprehension*. Reston, VA: N.C.T.M.
- Friel, S. N., Curcio, F. R., y Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, 124-158.
- García, L. (2014). Construcción de un test para medir los conocimientos y aptitudes en matemáticas de estudiantes de primer año de universidad. En C. C. Suárez. (Ed.), *Cuadernos de la maestría en docencia e investigación universitaria. Tomo II* (pp. 359-375). Bogotá: Universidad Sergio Arboleda.
- Martínez, R. (1995). *Psicometría: Teoría de los test psicológicos y educativos*. Madrid: Síntesis.
- Pérez, L., Aparicio, A., Bazán, J. y Abdounur, O. (2015). Actitudes hacia la estadística de los estudiantes universitarios de Colombia. *Educación Matemática*, 27(3), 111-149.
- Pinto, J (2010). *Conocimiento didáctico del contenido sobre la representación de datos estadísticos: estudios de casos con profesores de estadística en carreras de Psicología y educación*. Tesis Doctoral. Universidad de Salamanca.