



UNIVERSIDAD AUTÓNOMA
DE AGUASCALIENTES

CENTRO DE CIENCIAS SOCIALES Y HUMANIDADES
DEPARTAMENTO DE EDUCACIÓN

TESIS

LA COMPRENSIÓN ESTADÍSTICA EN LOS ESTUDIANTES UNIVERSITARIOS:
UN ESTUDIO ENTRE ALUMNOS DEL CENTRO DE CIENCIAS SOCIALES Y
HUMANIDADES DE LA UAA

PRESENTA

María de Jesús Vanessa Mendoza Rivera

PARA OBTENER EL GRADO DE MAESTRA EN INVESTIGACIÓN EDUCATIVA

TUTOR

Dr. Daniel Eudave Muñoz

CO TUTOR

Dr. Jesús Pintos Sosa

COMITÉ TUTORAL:

Dra. Guadalupe Ruiz Cuellar

Aguascalientes, Ags., 15 de noviembre 2013



UNIVERSIDAD AUTÓNOMA
DE AGUASCALIENTES



ANIVERSARIO
UAA

ASUNTO: **CONCLUSIÓN DE TESIS**
DEC. CCS y H. OF. N° 1615/2013

DR. FERNANDO JARAMILLO JUÁREZ,
DIRECTOR GENERAL DE INVESTIGACIÓN Y POSGRADOS,
P R E S E N T E

Por medio del presente me permito comunicarle a usted que el documento final de la tesis titulado **"LA COMPRENSIÓN ESTADÍSTICA DE LOS ESTUDIANTES UNIVERSITARIOS. UN ESTUDIO ENTRE ALUMNOS DEL CENTRO DE CIENCIAS SOCIALES Y HUMANIDADES DE LA UAA"**, de la **C. MARÍA DE JESÚS VANESSA MENDOZA RIVERA**, egresada de la **MAESTRÍA EN INVESTIGACIÓN EDUCATIVA**, respeta las normas y lineamientos establecidos institucionalmente para su elaboración y su autor cuenta con el voto aprobatorio de su tutor y comité tutorial.

Sin más por el momento aprovecho la ocasión para enviarle un cordial saludo..

A T E N T A M E N T E
Aguascalientes, Ags., 12 de Noviembre de 2013
"SE LUMEN PROFERRE"

DR. DANIEL EUDAVE MUÑOZ
DECANO DEL CENTRO DE CIENCIAS SOCIALES Y HUMANIDADES

c.c.p.- DRA. GUADALUPE RUIZ CUELLAR.- Secretaria Técnico del Doctorado en Ciencias Sociales y Humanidades
c.c.p.- DR. LUCIANO RAMÍREZ HURTADO.- Secretario de Investigación y Posgrado del CCSyH
c.c.p.- C. MARÍA DE JESÚS VANESSA MENDOZA RIVERA.- Egresada de la Maestría en Investigación Educativa
c.c.p.- Archivo

ggl



UNIVERSIDAD AUTÓNOMA
DE AGUASCALIENTES



ANIVERSARIO
UAA

DR. DANIEL EUDAVE MUÑOZ
DECANO DEL CENTRO DE
CIENCIAS SOCIALES Y HUMANIDADES
PRESENTE

Por medio de la presente hacemos de su conocimiento que **María de Jesús Vanessa Mendoza Rivera**, egresado de la MAESTRÍA EN INVESTIGACIÓN EDUCATIVA ha presentado el documento final de su tesis de maestría titulado "La comprensión estadística de los estudiantes universitarios. Un estudio entre alumnos del Centro de Ciencias Sociales y Humanidades de la UAA".

La tesis incorpora los elementos teóricos y metodológicos que le permiten ser defendida en el examen de grado reglamentario, por ello se solicita que se proceda a los trámites correspondientes para la presentación del examen de grado de maestría.

ATENTAMENTE
"SE LUMEN PROFERRE"

Aguascalientes, Ags., a 7 de noviembre de 2013

DR DANIEL EUDAVE MUÑOZ
TUTOR

DR. JESUS ENRIQUE PINTO SOSA
CÓ TUTOR

DR. GUADALUPE RUIZ CUELLAR
INTEGRANTE COMITÉ

c.c.p. Archivo Maestría en Investigación Educativa
c.c.p. Interesado

Agradecimientos



Agradezco el apoyo brindado por el Consejo Nacional de Ciencia y Tecnología (CONACYT) para la realización de mis estudios de maestría que concluyeron con esta tesis, como producto final de la Maestría en Investigación Educativa de la Universidad Autónoma de Aguascalientes y por haberme otorgado la beca No. 230408/257210 durante el período de agosto de 2011 a julio de 2013.

Dedicatoria



A mis padres y mis hermanas por el apoyo incondicional.

A mis asesores de tesis que gracias a sus consejos
y experiencia influyeron para concluir este trabajo.

Índice General

Índice General.....	1
Índice de Tablas.....	3
Índice de Gráficas.....	4
Índice de figuras.....	4
Resumen.....	5
Abstrac.....	6
INTRODUCCIÓN.....	7
CAPÍTULO I.....	8
PROBLEMA DE ESTUDIO.....	8
1.1 La importancia de la estadística en la sociedad.....	8
1.1.1 La educación estadística en el contexto internacional.....	8
1.1.2 La educación estadística en el contexto mexicano.....	10
1.1.3 La comprensión estadística como objeto de estudio.....	13
1.2 Preguntas de investigación.....	17
1.3 Justificación de la investigación.....	18
1.4 Objetivos.....	19
1.4.1 El objetivo general.....	19
1.4.2 Objetivos específicos.....	19
CAPÍTULO II.....	20
REVISIÓN DE LA LITERATURA.....	20
2.1 Comprensión de la estadística.....	20
2.1.1 Importancia de la comprensión estadística.....	20
2.1.2 Caracterización de la comprensión estadística.....	22
2.1.3 Niveles de comprensión estadística.....	24
2.2 Educación estadística.....	28
2.2.1 Alfabetización, razonamiento y pensamiento estadísticos.....	29
2.2.2 Interacción de la alfabetización, razonamiento y pensamiento estadístico...	34
2.3 Pruebas de logro.....	35
2.3.1 Pruebas de logro en estadística.....	35
2.3.2 Pruebas de logro en México.....	36
2.3.3 Desarrollo de CAOS- 4.....	38

2.3.4 Estructura de CAOS-4	40
CAPÍTULO III	43
DISEÑO METODOLÓGICO	43
3.1 Enfoque metodológico	43
3.2 Variables	44
3.2.1 Variable dependiente	44
3.2.2 Variables independientes.....	45
3.3 Instrumentos de medición.....	47
3.3.1 CAOS-4	47
3.3.2 Cuestionario y escala de opinión	49
3.5 Pilotaje de los instrumentos.....	50
3.6 Muestra	51
3.7 Reporte de trabajo de campo	53
3.8 Plan de análisis	56
CAPÍTULO IV.....	60
RESULTADOS.....	60
4.1 Descripción de los estudiantes.....	60
4.2 Confiabilidad y validez	63
4.3 La comprensión estadística de los estudiantes	71
CAPITULO V.....	94
DISCUSIÓN.....	94
5.1 Discusión.....	94
Conclusiones	101
Glosario.....	104
Bibliografía	105
Instituciones y Centros	109
Anexos	110

Índice de Tablas

TABLA 1 TAXONOMÍA DE HABILIDADES.....	25
TABLA 2 PERFIL DE LA COMPRESIÓN ESTADÍSTICA	26
TABLA 3 ESTRUCTURA CAOS-4	41
TABLA 4 DEFINICIÓN OPERACIONAL DE COMPRESIÓN ESTADÍSTICA PARA CAOS-4.....	45
TABLA 5 COMENTARIOS DE LOS EXPERTOS QUE REVISARON LA TRADUCCIÓN DE CAOS-4 .	48
TABLA 6 LICENCIATURAS DEL CENTRO DE CIENCIAS SOCIALES Y HUMANIDADES.....	52
TABLA 7 NÚMERO DE INSTRUMENTOS APLICADOS POR CARRERA	53
TABLA 8 NIVEL DE DIFICULTAD Y EVALUACIÓN DEL REACTIVO	58
TABLA 9 PODER DE DISCRIMINACIÓN DE LOS REACTIVOS SEGÚN SU VALOR D.....	59
TABLA 10 FRECUENCIA DE RESPUESTAS DEL CUESTIONARIO DE OPINIÓN	62
TABLA 11 ÍNDICE DE DIFICULTAD DE LA PRUEBA	64
TABLA 12 DIFICULTAD DE LOS ÍTEMS SEGÚN LAS DIMENSIONES DE LA PRUEBA	67
TABLA 13 PODER DE DISCRIMINACIÓN.....	68
TABLA 14 PROMEDIO DE ÍNDICE DE DIFICULTAD Y PODER DE DISCRIMINACIÓN	70
TABLA 15 PUNTAJE TOTAL DE RESPUESTAS CORRECTAS	73
TABLA 16 FRECUENCIAS POR DIMENSIÓN DE CAOS-4	74
TABLA 17 VARIABLE SEXO Y SU RELACIÓN CON EL NIVEL DE COMPRESIÓN ESTADÍSTICA...	91
TABLA 18 VARIABLE CARRERA Y SU RELACIÓN CON EL NIVEL DE COMPRESIÓN ESTADÍSTICA	92
TABLA 19 VARIABLE NUMERO DE CURSOS Y SU RELACIÓN CON EL NIVEL DE COMPRESIÓN ESTADÍSTICA	93
TABLA 20 EJEMPLO DE RECOLECCIÓN DE DATOS.....	97

Índice de Gráficas

GRÁFICA 1 DISTRIBUCIÓN DE CARRERAS Y CURSOS.....	61
GRÁFICA 2 CONCEPTOS ESTADÍSTICOS NO COMPRENDIDOS.....	71
GRÁFICA 3 RESULTADOS POR DIMENSIÓN DE CAOS-4	79
GRÁFICA 4 RECOLECCIÓN DE DATOS	80
GRÁFICA 5 . ESTADÍSTICA DESCRIPTIVA	81
GRÁFICA 6 REPRESENTACIÓN GRÁFICA	82
GRÁFICA 7 DIAGRAMA DE CAJAS.....	83
GRÁFICA 8 DISTRIBUCIÓN NORMAL.....	84
GRÁFICA 9 DATOS BIVARIADOS	85
GRÁFICA 10 PROBABILIDAD	86
GRÁFICA 11 VARIABILIDAD DE LA MUESTRA.....	87
GRÁFICA 12 INTERVALOS DE CONFIANZA.....	88
GRÁFICA 13 PRUEBAS DE SIGNIFICANCIA.....	89

Índice de figuras

FIGURA 1. INTEGRACIÓN DE LA ALFABETIZACIÓN, RAZONAMIENTO Y PENSAMIENTO ESTADÍSTICO	34
FIGURA 2. CICLO DE LA EDUCACIÓN ESTADÍSTICA.....	35
FIGURA 3. CARACTERÍSTICAS DE LOS SUJETOS.....	56
FIGURA 4. DESCRIPCIÓN DE CAOS-4	57

Resumen

Esta investigación se realizó en la Universidad Autónoma de Aguascalientes. Tuvo como objetivo general analizar la comprensión que manifiestan los estudiantes de educación superior que han realizado al menos un curso de estadística en temas relacionados con estadística descriptiva e inferencial. Se recurrió a la metodología cuantitativa, al tipo de estudio no experimental, transversal, descriptivo y correlacional. El muestreo fue no probabilístico formado por 206 estudiantes del Centro de Ciencias Sociales y Humanidades a quienes se les administró una prueba de opción múltiple y una escala de opinión. Cada estudiante contestó 40 reactivos construidos a partir de diez dimensiones. La confiabilidad de la prueba aplicada fue muy baja ($\alpha = 0.13$); ello obliga a tener cautela con la interpretación de los resultados obtenidos en esta investigación y a la vez, abre interrogantes y nuevas vías de indagación para estudios posteriores.

Para dar respuesta a los objetivos planteados en este estudio, en una primera etapa se describieron las variables independientes, es decir, la cantidad de hombres y mujeres así como el semestre y las carreras cursadas por los participantes con la finalidad de conocer las características de la muestra. En una segunda etapa del estudio se sugirió revisar los reactivos para mejorar la confiabilidad y validez, además de aplicarlo en diferentes contextos universitarios ya que ninguna dimensión logró ser contestada con más del 50% de los ítems de forma correcta en el contexto de las carreras de Ciencias Sociales y Humanidades que pertenecen a la Universidad Autónoma de Aguascalientes. Finalmente se clasificó, a los estudiantes que participaron la aplicación de la evaluación, en cuatro niveles: 1. Nivel muy bajo, 2. Nivel bajo, 3. Nivel regular y 4. Nivel alto; se encontró que el 94.2% de los estudiantes cae en la clasificación de bajo, una minoría cae en la clasificación de muy bajo y regular y finalmente ninguno de los estudiantes alcanzó la categoría de alto, esta clasificación fue realizada de acuerdo a los puntajes obtenidos en la prueba de forma individual.

Abstrac

This investigation was realized in the University of Aguascalientes. The objective was analyze the comprehension that university students testify, when they had done one course of statistics in topics like descriptive e inferencial statistic. The methodology of this investigation was quantitative, no experimental, cross, descriptive and correlational. The sampling was no probabilistic and it was constituted by 206 students of Social Science and Humanity Centre. They answer the test with option multiple responses and one opinion scale. Each student answered 40 items constructed by ten dimensions. Inside the results the test had a confidence of $\alpha = 0.13$ this was lower and according to literature the confidence coefficients for one test can be modest but no lower than $\alpha = 0.56$, it is important to have be careful with the results interpretation.

According to goals in this investigation, in one phase are describe the independent variables, how many men and women, semester and bachelor, the purpose is know the characteristics of the sample. In the second phase the study recommended review the items because is important to improve the confidence and validity, in addition is important to apply the test in different university context because no body achieved answered more than 50% of items in correct form in the context of the bachelor in the University of Aguascalientes. Finally the students were classified in three levels according to the correct answers; the levels were very low, low, moderate and high and found that 94.2% of students get only level low and nobody get the level high, this classification was realized according to the correct answers that students got in the test.

INTRODUCCIÓN

De manera inicial esta tesis surge del interés de conocer que es lo que los estudiantes han aprendido de estadística a lo largo de sus cursos académicos, ya que enfrentarse a la tarea de interpretar grupos de datos numéricos en ocasiones no es una actividad sencilla debido a la falta de habilidades ocasionadas por una limitada preparación en el tema de la estadística.

Me ha interesado conocer acerca del tema y profundizar en él, puesto que en México hay pocos instrumentos de evaluación que determinan los conocimientos adquiridos por los estudiantes. Este tipo de instrumentos son básicos para detectar áreas de oportunidad en docentes y/o estudiantes o bien para mejorar planes y programas de estudio.

El trabajo centra la atención en la comprensión estadística de los estudiantes universitarios los cuales son evaluados por medio de la prueba CAOS-4 (*Comprehensive Assessment of Outcomes in Statistics*) la cual aborda los siguientes temas estadísticos: recolección de datos, estadística descriptiva, representación gráfica, diagrama de cajas, distribución normal, datos bivariados, probabilidad, variabilidad de la muestra, intervalos de confianza y pruebas de significancia.

Un aspecto determinante para elegir la prueba utilizada en este trabajo tuvo que ver con el interés de evaluar los conocimientos estadísticos, más allá de la memorización de conceptos y fórmulas. CAOS-4 demanda utilizar de forma reflexiva conocimientos previamente adquiridos para poder responder los cuestionamientos planteados.

La investigación ofrece un panorama de cómo los estudiantes de la Universidad Autónoma de Aguascalientes, específicamente del Centro de Ciencias Sociales y Humanidades, resuelven problemas estadísticos, lo que permite describir los conocimientos que han sido afianzados durante su instrucción y los que no se han logrado dominar.

La tesis está organizada en cinco capítulos. El primero presenta el problema de investigación; aquí se describe puntualmente las preguntas y los objetivos del trabajo. Enseguida se encuentra el capítulo del marco de teórico en el que se desarrolla la teoría que fundamenta la tesis; un tercer capítulo tiene por objetivo describir la metodología de investigación; el cuarto capítulo ofrece los resultados del análisis de los datos. Finalmente, se cierra con un apartado de discusión que surge a partir de los resultados obtenidos.

CAPÍTULO I

PROBLEMA DE ESTUDIO

1.1 La importancia de la estadística en la sociedad

La estadística está cobrando relevancia para muchas áreas de la ciencia, de tal forma que es indiscutible según Batanero (2001) que el siglo XX ha sido importante para la estadística, puesto que ya se considera una de las ciencias metodológicas fundamentales y base importante del método científico experimental.

Tener en mente que prácticamente no existe actividad en la que la estadística no puede estar presente, es decir, aplicamos la estadística en la vida diaria aun sin estar conscientes de ello, en la investigación científica, también utilizamos la estadística pues gracias a la llamada “inferencia estadística” podemos considerar que las características de una población estudiada se pueden establecer conociendo las características de una fracción de esa población. En la industria también se utiliza la estadística por ejemplo, en el diseño de un producto, en estudios de mercado. En otras áreas también se utiliza como en las tecnologías de la información y en la administración pública (ver: <http://www.estadistica2013cimat.mx/impacto>).

La necesidad de crear equipos sólidos de trabajo relacionados con la educación estadística han llevado a esta disciplina a consolidarse en el ámbito nacional e internacional, es por ello que en los siguientes apartados se abordan estas temáticas.

1.1.1 La educación estadística en el contexto internacional.

El Instituto Internacional de Estadística (*ISI, por sus siglas en inglés*) desde su fundación en 1885 se ha centrado en promover la estadística. En 1948 el ISI se consolida como el principal impulsor del tema estadístico y al mismo tiempo se establece el Comité de Educación encargado de promover la formación estadística, de esta forma colabora con la UNESCO y otros organismos internacionales (Vere-Jones, 1995).

Según Batanero (2001) subcomités especiales del ISI se han dedicado a promover la estadística en las instituciones educativas además, de promover conferencias sobre el tema dando origen a los ICOTS (*International Conferences on Teaching Statistics*) que comienzan formalmente en 1982 en la Universidad de Sheffield y se celebran cada cuatro años a partir de entonces.

El Centro de Educación Estadística de la Universidad de Sheffield dio origen a la creación de la revista *Teaching Statistics* dirigida a profesores, el éxito que ha tenido a lo largo de 24 años demuestra el interés que tienen los profesores acerca de los temas

didácticos y de la necesidad de compartir y discutir problemas educativos (Batanero, 2001).

En 1991 el ISI promovió la creación del IASE (*International Association for Statistical Education*), la cual participa activamente en la promoción y desarrollo de la educación estadística en el ámbito internacional (Batanero, 2001).

En el año 2000 el IASE lanzó un boletín, llamado Hipótesis Alternativa, el cual tiene como principal objetivo divulgar información profesional y documentos de interés para la comunidad de educación estadística en Venezuela. En el 2002 se unieron a este proyecto México y España y a partir del 2006 el boletín ya es para toda América Latina. Este es un ejemplo claro de cómo se ha ido incrementando el interés por difundir entre los investigadores y profesionistas de la estadística la enseñanza de ésta, ya que es importante para el desarrollo de la ciencia.

Así como Venezuela, Argentina, es uno de los países en América Latina, que se ha interesado en impulsar el desarrollo y el estudio de la estadística, pues a través de la SAE (Sociedad Argentina de Estadística) que fue creada en junio de 1952 se desarrollan coloquios referentes al tema de forma anual. En 1991 Chile se integró a estos trabajos y actualmente se desarrollan de forma bianual en los Congresos Latinoamericanos de Sociedades de Estadística (ver <http://www.s-a-e.org.ar/sae/>) en los que participan estadísticos latinoamericanos en su mayoría.

Hoy día a nivel mundial, se está desarrollando un evento que impulsa el conocimiento de la estadística tanto de investigadores como de estudiantes y público en general; el evento se ha denominado Statistics2013 (ver <http://www.estadistica2013cimat.mx/boletin01>) y hace referencia al año internacional de esta disciplina. Este evento surge por la iniciativa de cinco sociedades internacionales. Statistics2013 establece tres objetivos principales:

1. Aumentar la conciencia pública sobre el poder y el impacto de las estadísticas relativo a todos los aspectos de la sociedad.
2. Nutrir a la estadística como profesión, especialmente entre los más jóvenes.
3. Fomentar la creatividad y el desarrollo de las ciencias de la probabilidad y estadística.

México participa en este evento fomentando la importancia de la disciplina a través de diversas actividades como seminarios y congresos dirigidos a estudiantes de diferentes niveles académicos así como a empresas gubernamentales y privadas.

Todas las iniciativas que se presentan van muy de la mano, ya que comparten la disposición para promover y nutrir a la estadística tanto a nivel internacional como en América Latina. En el siguiente apartado se describe lo que sucede en el contexto mexicano, ya que también se está integrando a las propuestas impulsadas por los organismos internacionales.

1.1.2 La educación estadística en el contexto mexicano

La Red de Centros de Investigación en Matemática Educativa A.C., tiene como principal objetivo impulsar el desarrollo de la Matemática Educativa a nivel nacional mediante la creación de grupos de investigación así como la promoción y difusión de diferentes actividades de carácter sociocultural que permitan acercar la Matemática a la sociedad a través de la Escuela de Invierno en Matemática Educativa fundada en 1998 y en donde a partir de 2011 se han generado grupos de trabajo y discusión sobre educación estadística.

México, en los últimos años ha dado pie al desarrollo de varios grupos de trabajo los cuales se reportan a continuación en virtud de haber sido descritos en la memoria de la XV Escuela de Invierno en Matemática Educativa del 2012 (ver http://www.red-cimates.org.mx/images/pdf/EIMES/Memorias/memoria_eime_xv.pdf).

En primer lugar se encuentran las redes de trabajo que se han consolidado en el país y en donde se encuentran grupos como el de enseñanza y aprendizaje de la probabilidad y la estadística que pertenece al departamento de matemática educativa del CINVESTAV-IPN y el cual ya tiene varias líneas de investigación entre las que destacan la variación en el aprendizaje de la estadística y la noción de distribución y probabilidad.

También está el Departamento de Ciencias Básicas del Instituto Tecnológico de Chihuahua II, que tiene las líneas de investigación de *Educación en probabilidad, estadística y análisis de datos*, así como la de *Formación del profesorado*.

El Cuerpo Académico *Metodología y aplicaciones de las técnicas y modelos estadísticos*, de la Universidad Veracruzana, tiene las líneas de investigación de modelación estadística, metodología estadística, consultoría estadística y enseñanza de la misma.

Está también el Grupo de Educación Estadística de la Universidad de Sonora quienes trabajan líneas de investigación relacionadas a la educación estadística.

En el Departamento de Matemáticas de la Universidad Autónoma Metropolitana-Iztapalapa, se tiene el Área de probabilidad y estadística, cuyos proyectos principales se

TESIS TESIS TESIS TESIS TESIS

enfocan directamente en la validación de instrumentos de evaluación para desarrollar habilidades matemáticas, metodología estadística, técnicas estadísticas longitudinales, operadores en espacios probabilísticos, optimización de procesos estocásticos y estadística asintótica.

También se encuentra el grupo de investigación de la Benemérita Universidad Autónoma de Puebla con una antigüedad de apenas dos años y con líneas de investigación relacionadas con probabilidad y estadística.

El grupo de Investigación e Innovación en Didáctica de la Estadística del Instituto Tecnológico de Estudios Superiores de Monterrey en su campus Monterrey, quien trabaja sobre temas relacionados con el desarrollo del pensamiento estadístico basado en la variabilidad y aleatoriedad, distribución de probabilidad, intervalos de confianza y pruebas de hipótesis.

Hay un grupo interesado por el mejoramiento del proceso de enseñanza-aprendizaje de la estadística conformado por investigadores de la Universidad Veracruzana, Universidad Autónoma Metropolitana Unidad Iztapalapa y del Centro de Investigación en Matemáticas (CIMAT), Unidad Aguascalientes, el cual tiene por objetivo formar, actualizar e innovar las competencias didácticas de académicos e investigadores que se encargan de enseñar probabilidad y estadística.

Por su parte, el Grupo de investigación sobre didáctica de la probabilidad y estadística de la Universidad Autónoma de Sinaloa, trabaja con las líneas de investigación relacionadas con la tecnología computacional aplicada a la enseñanza de la probabilidad y la estadística así como la formación de profesores en didáctica de la probabilidad y estadística.

El Grupo sobre la enseñanza de la estadística y la probabilidad, de la Universidad Autónoma de Querétaro y del Tecnológico Regional de Querétaro, desarrolla dos líneas de investigación, una que va enfocada a la enseñanza y aprendizaje de los conceptos de la probabilidad y la estadística y la segunda relacionada con el uso de las nuevas tecnologías para la enseñanza de la probabilidad y la estadística.

Finalmente se encuentra la Red de investigación e innovación en educación estadística y matemática educativa del IPN, la cual se caracteriza por las líneas de investigación relacionadas con la profesionalización docente y con la mejora de la actividad docente a través del diseño de actividades de aprendizaje para la enseñanza de la probabilidad y estadística.

Además de estos grupos existen instituciones dedicadas a la formación de capital intelectual en estadística, prueba de ello, es el Centro de Investigación en Matemáticas (CIMAT) que a partir de 1980, se ha dedicado a difundir la importancia de la estadística en la industria, en la vida profesional, en la actividad científica, incluso en la investigación y creación artística así como en la vida cotidiana, de tal forma que trabaja con instituciones de orden gubernamental y privado con la finalidad de orientarles en la interpretación de grandes cantidades de datos, para hacer diagnósticos y extraer conclusiones válidas y pertinentes.

Así mismo la Asociación Mexicana de Estadística (AME) fundada en 1978 con el propósito de contribuir al desarrollo de esta disciplina dentro del país ha apoyado a la creación en 2012 de la Red de Investigación sobre Enseñanza de la Probabilidad y Estadística (RIEPE), la cual manifiesta que aún falta mucho por trabajar. El objetivo principal es lograr conformar una comunidad global debido a que: 1. Los grupos de trabajo colegiado trabajan relativamente aislados; 2. No existe consenso respecto de los problemas que aquejan la educación estocástica en todos los niveles educativos; y 3. Es importante crear nuevos espacios de discusión para la comunidad científica en esta disciplina.

México es un país que se ha ido involucrando en desarrollar instituciones y grupos que ofrecen información relevante al tema, si bien países de América Latina como Venezuela y Chile, por poner un ejemplo, han iniciado más temprano el proceso de difusión y de investigación, nuestro país también está trabajando en aportar información a la comunidad educativa y un ejemplo claro de ello es la reciente creación de la *RIEPE* y de la Escuela de Invierno en Matemática quienes pretenden difundir y fortalecer grupos de trabajo colegiado entre investigadores.

En síntesis, mientras en países como Estados Unidos la enseñanza de la estadística ya ocupa un lugar primordial desde años atrás, en los últimos años México se ha interesado en desarrollar sociedades e instituciones propias para discutir los temas relativos a la enseñanza y el aprendizaje de la probabilidad y la estadística.

1.1.3 La comprensión estadística como objeto de estudio

La estadística se está expandiendo rápidamente en todos los niveles educativos y profesionales. Responder a las necesidades de la sociedad del conocimiento es una tarea inaplazable, por tal motivo, se debe incluir y reforzar en los programas de estudio temas relevantes que ayuden al estudiante a adquirir una adecuada comprensión. Tal como lo ha dicho May (2009): “Las escuelas hoy día, deberían considerar una reforma en la enseñanza de la estadística, con la finalidad de formar a los estudiantes no sólo en la aplicación de fórmulas, sino también en su interpretación y utilidad” (p. 8).

El estudio que realizaron Kaplan, Fisher y Rogness (2010) mostró que el uso de un lenguaje especializado en un campo específico puede hacer que un tema parezca más difícil de lo que realmente es, y entonces, los estudiantes adquieren concepciones erróneas de lo que significa estar estadísticamente educado; esta situación se manifiesta como problemática, ya que los alumnos adquieren conocimientos erróneos o deficientes desde los niveles básicos de instrucción y en un futuro muy próximo estas carencias se manifiestan al tener mayores dificultades en la toma de decisiones cuando se presentan situaciones de incertidumbre o al interpretar datos numéricos y por consiguiente se adquiere una deficiente comprensión estadística.

En la XIII Conferencia Interamericana de Educación Matemática llevada a cabo en Brasil, Inzunza y Jiménez (2011) abordaron el tema de razonamiento y pensamiento estadístico en estudiantes universitarios con una preparación avanzada. Los resultados fueron parecidos a los que reporta Eudave (2007), quien investigó a doce estudiantes de carreras no matemáticas con el propósito de exponer la manera en que los estudiantes conceptualizan las nociones de estadística descriptiva según la carrera. Ambos estudios afirman que aunque los alumnos hayan tomado cursos formales de estadística, estos no necesariamente poseen un razonamiento y pensamiento estadístico avanzado.

Por otro lado, Pinto (2011) investigó a través de un estudio de caso cualitativo, los errores y dificultades de los estudiantes de una universidad mexicana, desde la perspectiva de los profesores. Acotó su indagación a la carrera de Educación, específicamente en el aprendizaje de la representación gráfica en estadística. El autor identifica y agrupa los errores como sigue: a) errores por ignorancia o desconocimiento de conceptos previos, b) errores de reconocimiento y comprensión de la información del gráfico, c) errores en la comprensión de conceptos y d) errores relacionados con el contexto; de esta forma llega a la conclusión de que los errores y las dificultades de los

TESIS TESIS TESIS TESIS TESIS

estudiantes son fundamentales para la formación de los futuros profesores con la finalidad de que se aprenda a diagnosticar para implementar estrategias que faciliten el proceso de aprendizaje de los estudiantes.

Los estudios realizados en cuanto al tema de comprensión estadística se han ido desarrollando a través del tiempo, de tal forma, que hoy en día se encuentran líneas de investigación que dan cuenta de la importancia del tema; instituciones como la *International Commission on Mathematical Instruction* (ICMI, por sus siglas en inglés) y la *International Association for Statistical Education* (IASE, por sus siglas en inglés) han advertido que a pesar de las recomendaciones y los esfuerzos realizados para incrementar la enseñanza de la estadística en las escuelas, los estudiantes adquieren en forma deficiente la comprensión estadística necesaria para manejarse en una sociedad basada en la información (May, 2009).

En 2009, Canché, desarrolló una prueba que determina la medida, en la cual los estudiantes de primaria comprenden e interpretan gráficas estadísticas a través de un estudio cuantitativo. La muestra estuvo conformada por 206 estudiantes de sexto año de primaria y los resultados mostraron que la mayor dificultad se presentaba en los pictogramas y gráficas lineales. Este tipo de evaluación sirve a los docentes para que puedan preparar a los estudiantes al presentar la prueba ENLACE ya que los contenidos de los ítems de la prueba diseñada por Canché, son muy similares a los de la prueba que aplica la SEP cada año a los alumnos de nivel básico.

Salazar (2008) investigó sobre la comprensión de la información gráfica que tienen los estudiantes de educación superior; tomando una muestra de 276 estudiantes de la Universidad Autónoma de Aguascalientes inscritos en los últimos semestres de doce diferentes licenciaturas y reportó como parte de los resultados que “los estudiantes tienen mejor comprensión del concepto de media, que comprensión de la información estadística mostrada en gráficas” (p. 121). Considera que el número de materias de estadística que cursan los estudiantes es vital para mejorar los puntajes que se obtienen en las pruebas estandarizadas; además afirma que el dominio de conceptos estadísticos influye en la comprensión de la información. Estudios como el de Salazar, permiten dar cuenta de la comprensión estadística lograda por los estudiantes universitarios.

Según Schaun y Mattern (1997) de la investigación y de la experiencia de la enseñanza se sabe que son pocas las personas que emplean de forma efectiva la comprensión estadística aun después de la instrucción, es decir, pueden dar solución

TESIS TESIS TESIS TESIS TESIS

técnica a un problema, pero carecen de las habilidades para interpretar los resultados obtenidos o bien utilizan de forma inadecuada los conceptos estadísticos, en otras palabras, existe una falta de conexión entre el aprendizaje de fórmulas estadísticas y la interpretación de los resultados; De acuerdo con Kaplan, Fisher y Rogness (2010), en ocasiones se adquieren conocimientos estadísticos que son poco comprendidos tanto por los estudiantes como por los profesionistas es decir, si los contenidos estadísticos en los diferentes niveles académicos son aplicados en forma deficiente la importancia de la aplicación adecuada de la estadística también será limitada.

Así como se han desarrollado diversos estudios para abordar temas relacionados con la importancia de la comprensión estadística, también se han desarrollado pruebas de logro con la finalidad de evaluar aspectos específicos de estadística.

En estudios como los presentados por Garfiel, delMas, Chance, Poly y Ooms (2005) se demuestra que una problemática para determinar los logros académicos de los alumnos a través de la comprensión que han adquirido al término de los cursos estadísticos, es la carencia de instrumentos válidos que evalúen el razonamiento y el pensamiento estadístico. Para enfrentar esta situación, los investigadores desarrollaron CAOS-4 con la finalidad de medir la comprensión estadística entre sus estudiantes. Cabe mencionar que se han desarrollado varias versiones de este instrumento desde su construcción en 2004 y se ha ido perfeccionando a través del tiempo, a fin de obtener la validez y confiabilidad necesarias en todo instrumento de evaluación e investigación.

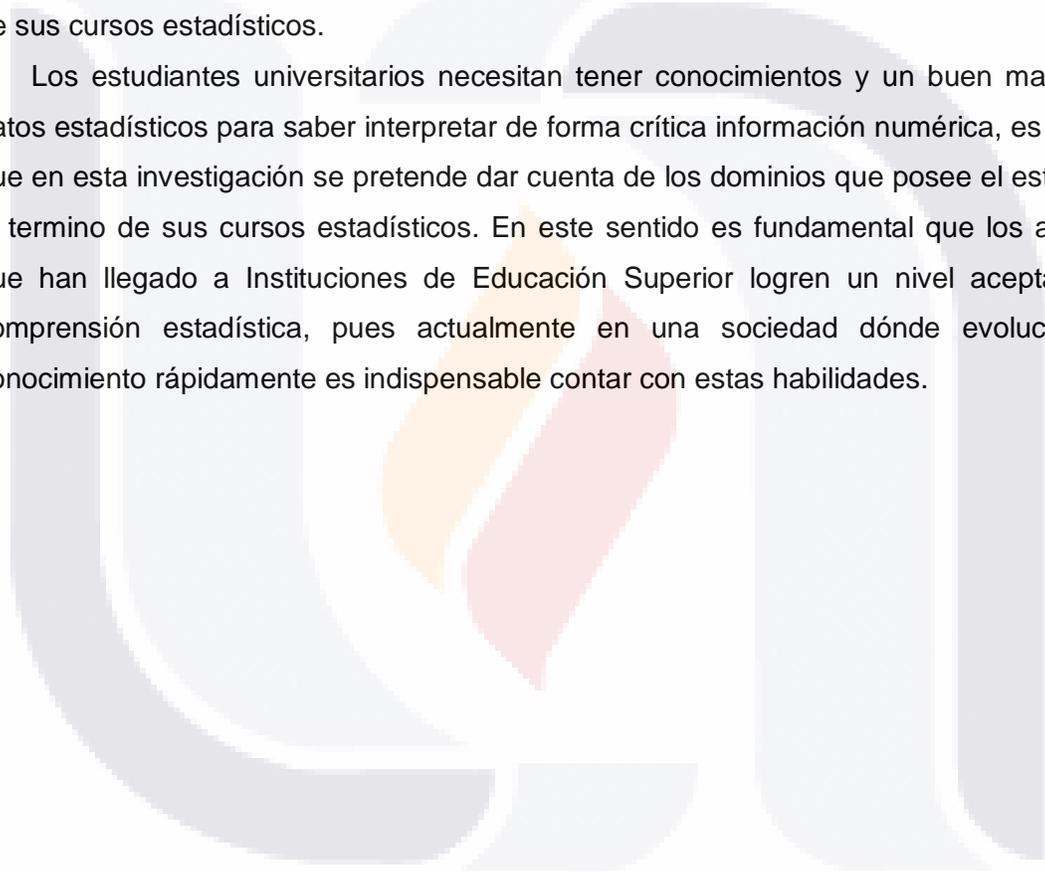
Los esfuerzos desarrollados por impulsar la comprensión estadística se han ido consolidando en México a través de investigaciones y de instituciones interesadas en el tema. Un ejemplo de ello, es el Centro Nacional de Evaluación para la Educación Superior (CENEVAL), quien desde 2008 desarrolló el Examen Transversal para el Nivel de Licenciatura enfocado a la Estadística (EXTRA-ES) con el propósito de identificar los conocimientos que adquieren los estudiantes de licenciatura durante su formación profesional con respecto a las habilidades estadísticas. Con este examen se pretende promover la calidad de la formación estadística de los estudiantes de las diferentes Instituciones de Educación Superior dentro del país.

Los contenidos del examen (ver <http://www.ceneval.edu.mx/ceneval-web/content.do?page=3377#exam04>) incluyen cuatro módulos; uno básico y tres complementarios. El primer módulo se espera que sea respondido en un periodo de tiempo de tres horas por cualquier estudiante cuya formación profesional incluya la

estadística, independientemente de la carrera que curse. Los módulos complementarios se refieren a áreas de especialidad como lo son: 1) métodos estadísticos, 2) muestreo y 3) estadística experimental.

Esta es una prueba de logro que ha sido validada por los expertos en la materia y ha sido diseñada con el rigor metodológico que caracteriza a las evaluaciones de alto impacto y se considera un primer paso en el que se definen los lineamientos de desempeño de los estudiantes y que pretende indicar si el alumno ha alcanzado los conocimientos necesarios para desempeñarse en la sociedad del conocimiento al término de sus cursos estadísticos.

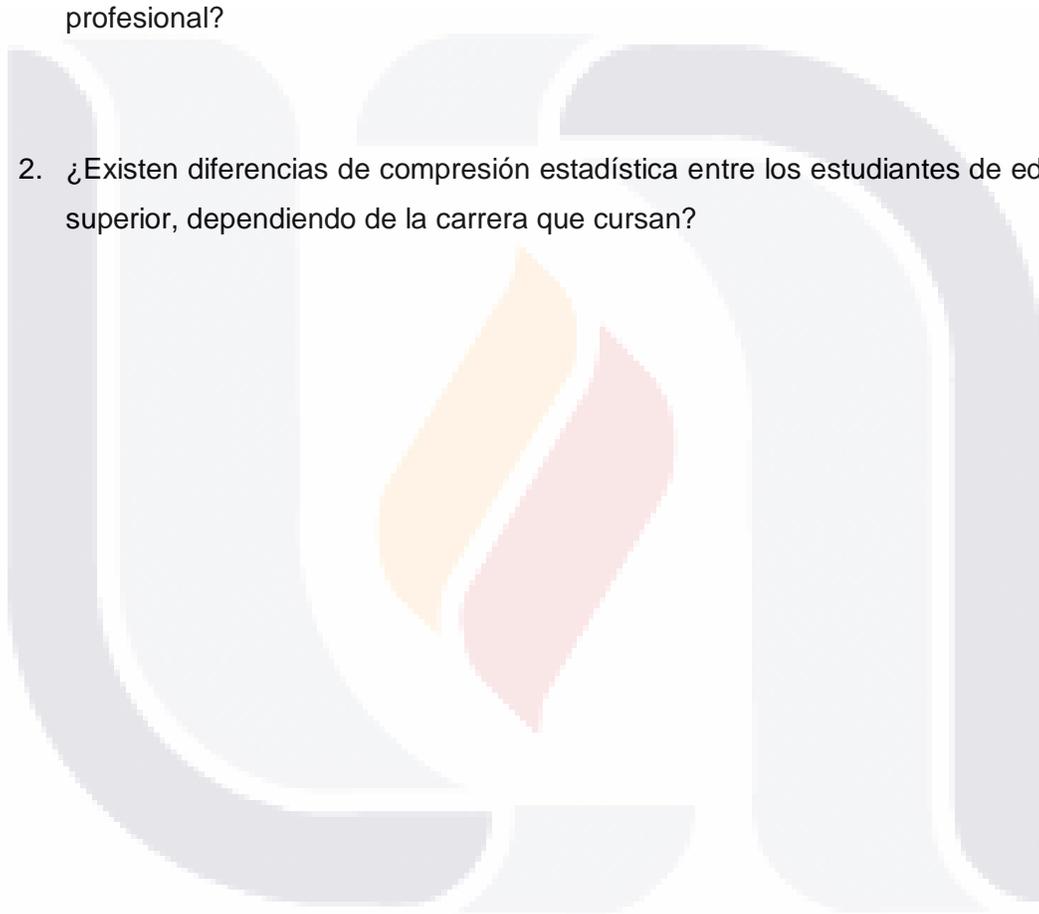
Los estudiantes universitarios necesitan tener conocimientos y un buen manejo de datos estadísticos para saber interpretar de forma crítica información numérica, es por ello que en esta investigación se pretende dar cuenta de los dominios que posee el estudiante al término de sus cursos estadísticos. En este sentido es fundamental que los alumnos que han llegado a Instituciones de Educación Superior logren un nivel aceptable de comprensión estadística, pues actualmente en una sociedad dónde evoluciona el conocimiento rápidamente es indispensable contar con estas habilidades.



1.2 Preguntas de investigación

Con base en los planteamientos que se han venido desarrollando, las preguntas de investigación que rigen este trabajo son las siguientes.

1. ¿Qué grado de comprensión estadística se manifiesta en los estudiantes de educación superior que han realizado cursos de estadística durante su formación profesional?
2. ¿Existen diferencias de comprensión estadística entre los estudiantes de educación superior, dependiendo de la carrera que cursan?



1.3 Justificación de la investigación

Gal (2002) afirmó que la alfabetización estadística es necesaria hoy en día y los adultos deben estar completamente conscientes de las tendencias y fenómenos de importancia social como el crecimiento de la población, la propagación de enfermedades, la producción industrial, los logros educativos o las tendencias de empleo, es decir, la mayoría de los adultos son consumidores de información estadística, la cual debe poderse interpretar de forma crítica. En esta misma línea Shi, He y Tao (2009) reconocieron la importancia de fomentar la comprensión estadística en todos los niveles educativos para que los ciudadanos puedan enfrentarse a tareas estadísticas que les implica la lectura y comprensión numérica.

Es de llamar la atención que a pesar de que se ha reconocido la importancia de la comprensión estadística en diversos ámbitos, autores como Salazar reportan que los alumnos reprueban la materia de estadística o bien se le da poca importancia a la misma por los estudiantes y las instituciones educativas, de tal forma que los conocimientos adquiridos por los estudiantes, en ocasiones suelen ser deficientes, y esto queda demostrado en los resultados de diversas investigaciones, ya que se desconoce, o en el mejor de los casos se conoce de una forma muy limitada, el modo de presentar o interpretar información estadística, cuando ésta es básica para la futura práctica profesional, pues a medida que se logra la comprensión estadística se pueden ofrecer interpretaciones de datos numéricos de manera más crítica y así proponer alternativas de acción oportunas ante un planteamiento estadístico determinado.

Es por ello que el desarrollo de esta tesis busca generar conocimiento con relación a la comprensión estadística, con la finalidad de dar cuenta del grado de asimilación que se manifiesta en los estudiantes de educación superior al responder a planteamientos que requieren de habilidades específicas adquiridas previamente desde la formación básica, y que concluyen en la carrera de elección del estudiante.

Con el propósito de ofrecer una alternativa de medición referente a la competencia de la comprensión estadística, se obtuvo el instrumento de evaluación CAOS-4, construido por Garfield, delMas, Chance, Poly y Ooms en 2005 con la finalidad de que sea utilizado de forma íntegra o tomando sólo algún aspecto, para ser retomado en futuras investigaciones, de tal forma que sea útil para la evaluación óptima de la comprensión estadística en los estudiantes universitarios ya que aún es un área muy poco explorada.

1.4 Objetivos

De acuerdo con la problemática planteada y con base en las preguntas de investigación, se formularon los siguientes objetivos que guiaron la investigación.

1.4.1 El objetivo general

Analizar la comprensión que manifiestan los estudiantes de educación superior al enfrentarse a temas de estadística, específicamente aquellos que ya han realizado al menos un curso de estadística durante la carrera.

1.4.2 Objetivos específicos

- 1) Describir los temas estadísticos que los estudiantes universitarios comprenden mejor.
- 2) Indicar si la comprensión estadística tiene relación con el género, la carrera y el cuestionario de opinión.

CAPÍTULO II

REVISIÓN DE LA LITERATURA

2.1 Comprensión de la estadística

A lo largo de este apartado se hace una revisión de la importancia de la comprensión estadística, así como las dificultades epistemológicas a las que se enfrenta el constructo, finalmente se describe cómo, a través del tiempo, se han desarrollado diversas pruebas que han servido para evaluar la comprensión estadística de forma válida y confiable.

2.1.1 Importancia de la comprensión estadística

La comprensión estadística se ha convertido en una necesidad a lo largo del tiempo en las sociedades actuales. La industria, la medicina, la economía, la política, y prácticamente todos los sectores de la sociedad están confiando cada vez más en los datos numéricos para tomar decisiones con impacto social y económico, y entonces la formación estadística debe convertirse en una parte integral de la formación académica de un estudiante.

En el mundo en el que vivimos, es primordial que el ser humano sea capaz de orientarse en la red de la información, el reto es que el estudiante logre las habilidades necesarias para enfrentar los cambios del día a día. Es decir, los individuos deben de conocer o por lo menos tener el mínimo de conocimiento de cómo los datos numéricos presentados en los medios de comunicación masiva están organizados, analizados e interpretados con la finalidad de emitir juicios críticos basados en el análisis y síntesis de la información cuantitativa (Ottaviani, 1998).

Autores como Schau y Mattern (1997) sostienen que cuando se logra la comprensión estadística también se alcanzan otros dominios como son, el razonamiento estadístico y por ende la solución de problemas, es decir, el pensamiento estadístico. En el mismo sentido Ben – Zvi y Garfield (2004) hacen mención de la importancia de adquirir dominios específicos que ayuden a explicar el qué y el cómo de los procesos con la finalidad de obtener interpretaciones de datos estadísticos válidas y confiables para la toma de decisiones, y para lograrlo es necesaria la comprensión de un lenguaje especializado en estadística.

Desafortunadamente, entre los estudiantes la estadística es frecuentemente vista como un tema desalentador ya que utiliza un lenguaje y una forma de pensar diferente que muy pocas veces es comprendida (Reading, 2002).

Con la finalidad de impulsar el conocimiento de esta disciplina, a través del colegio de profesionistas, *Guidelines for Assessment and Instruction in Statistical Education (GAISE)*, se recomienda que en los cursos introductorios de estadística se ponga énfasis en el entendimiento de conceptos y símbolos como parte de la función de conocer, con la finalidad de dejar de lado la memorización (Goldman y McKenzie, 2006), con esta propuesta se pretende involucrar al estudiante en proyectos reales del mundo en el que vive y de esta forma evaluar sus avances.

Godino (1996, citado en Batanero, 2001 p. 68), se adentra epistemológicamente al significado de comprensión afirmando que está íntimamente ligado a cómo se concibe el conocimiento y describió el término a través de diversas preguntas, lo que permite reflexionar acerca de lo que se entiende por comprensión estadística, puesto que en ocasiones tiende a ser difuso y difícil llegar a una definición concreta. “Se requiere responder a preguntas tales como: ¿Cuál es la estructura del objeto a comprender?, ¿Qué formas o modos posibles de comprensión existen para cada concepto?, ¿Qué aspectos o componentes de los conceptos estadísticos es posible y deseable que aprendan los estudiantes en un momento y circunstancias dadas?, ¿Cómo se desarrollan estos componentes?”

Friel, Curcio y Bright (2001), aunque habla de la comprensión gráfica, definen el término comprensión como el entendimiento de la información en forma escrita o simbólica; otros autores hablan de una conexión de estructuras cognitivas o redes mentales que desarrolla el sujeto y que le permiten acceder a formas más elevadas de conocimiento, siendo las deseables el razonamiento y la comprensión estadística (Shaun y Mattern, 1997). Es decir, la comprensión estadística se centra en la interpretación óptima que se le da a la información estadística ofrecida en un contexto determinado.

Cada uno de los autores intenta incursionar de forma muy particular en la definición del término de comprensión, de tal forma que se encuentra una diversidad de enfoques. Para fines de este trabajo se retomará la postura de Friel, Curcio y Bright que fue abordada en el 2001.

Hasta aquí se ha descrito el tema de la importancia de la comprensión estadística, la cual se considera básica para un sujeto estadísticamente educado, sin embargo, es necesario conocer las características que conforman este constructo.

2.1.2 Caracterización de la comprensión estadística

Es importante destacar que la comprensión estadística se debe ir desarrollando a lo largo de toda la instrucción que reciben los alumnos, desde los niveles básicos hasta la universidad, pasando por diferentes etapas; algunos autores pueden llamarle niveles, otros pueden incluso realizar un perfil del estudiante, pero el fin último es alcanzar una comprensión estadística al término de los cursos.

Diversos autores han abordado el tema desde su propio marco de referencia. Por ejemplo, Chau y Mattern (1997) observaron una fragilidad importante en los estudiantes, quienes al parecer carecen de una adecuada conexión entre el razonamiento conceptual y la solución de problemas, en este sentido, pareciera ser que los conceptos son un ente aislado de la comprensión estadística; por ejemplo pueden calcular la desviación estándar, sin embargo la interpretación que se hace de ésta en ocasiones es incorrecta. Sin una adecuada comprensión, los alumnos difícilmente logran un efectivo y eficiente razonamiento estadístico y en consecuencia tampoco se llega adecuadamente a la solución de problemas.

En la actualidad la educación estadística orienta a los profesores para hacer los temas estadísticos más interesantes y desafiantes y de esta forma lograr que los conceptos sean integrados en redes o estructuras cognitivas para la adecuada solución de problemas, pero según Broers (2008), repetidamente se encuentra que sólo una pequeña minoría de estudiantes tienen éxito en la construcción de estos esquemas o estructuras, es por ello que hay que familiarizar a los estudiantes con la estadística, demostrando la utilidad de ésta.

Para una adecuada comprensión estadística, las redes mentales o estructuras cognitivas son básicas; en muchas teorías de redes mentales el concepto de esquema se vuelve importante, ya que es considerado un mecanismo de almacenamiento mental que es estructurado como una red de conocimiento según Marshall (1995, citado en Shaun y Mattern, 1997).

Los esquemas se van formando cuando el sujeto está expuesto repetidamente a la solución de problemas. El alumno puede crear un esquema de tres formas diferentes, la primera se da abstrayendo las características más relevantes de una tarea; la segunda forma, es a través de asimilar estas características en los esquemas ya existentes; y finalmente la tercera, consiste en construir algunos nuevos esquemas. Con base en este planteamiento, se dice que una propiedad fundamental que define un esquema es la

TESIS TESIS TESIS TESIS TESIS

presencia de conexiones o redes mentales interconectados entre si (Shaun y Mattern 1997).

Los estudiantes de estadística pueden intentar formar modelos mentales. Un primer paso es tratar de abstraer los componentes importantes del problema; si los estudiantes no han adquirido los esquemas pertinentes y adecuados o no los evocaron oportunamente, una opción alternativa es intentar formar modelos mentales a través de recordar fórmulas, sin embargo este proceso se vuelve cada vez más difícil puesto que al utilizar la memorización se corre el riesgo de que en corto tiempo se olvide el proceso. Inicialmente puede ser más fácil memorizar una fórmula para resolver un problema particular que formar una red mental que permita asimilar la información dada en un determinado contexto; sin embargo la variedad en el número de fórmulas y problemas diferentes llega a ser abrumadora para los estudiantes, ya que existen muchos problemas que no pueden ser resueltos al utilizar una simple fórmula y encontrar la respuesta, sino que se requiere de un proceso más complejo (Skemp, 1987, citado en Shaun y Mattern, 1997).

De acuerdo con el modelo que tiene bases de la teoría del desarrollo cognitivo, Shaun y Mattern en 1997 afirman que los estudiantes aprenden estadística a partir de las siguientes operaciones: 1) asimilar (conectar) nueva información dentro de sus redes cognitivas, 2) formar nuevas conexiones entre el conocimiento que ya existe en sus redes, 3) reorganizar (acomodar) sus esquemas conectados para que coincidan con la información entrante, y por último, 4) eliminar conceptos incorrectos y conexiones. Esta forma de llegar a la comprensión pretende formar redes cognitivas duraderas que difícilmente se olvidan.

La interpretación de la información estadística depende de la comprensión conceptual que posee el sujeto; de esta forma los estudiantes pueden explicar los procedimientos apropiados o identificar las condiciones que son válidas para el uso de un procedimiento estadístico, es decir, utilizan el razonamiento y el pensamiento estadístico. Por ejemplo, si un profesor quiere desarrollar la comprensión del estudiante con relación al término de mediana dentro del contexto de la estadística, se diseñan actividades para orientar a los estudiantes a descubrir por qué la mediana es una medida de promedio, además de contrastar otras medidas de tendencia central, para que el alumno sepa identificar cuándo y dónde se debe usar.

En síntesis, la comprensión estadística hace referencia al desarrollo de estructuras cognitivas o redes mentales que le permiten al estudiante acceder a formas más elevadas de conocimiento. El objetivo es preparar a ciudadanos con niveles de comprensión estadística adecuados y por ende lograr que los ciudadanos adquieran niveles óptimos de alfabetización, razonamiento y pensamiento estadístico.

En el siguiente apartado se describirán los niveles de comprensión estadística que diversos autores han desarrollado para abordar el tema en cuestión y los cuales servirán de base para hacer interpretaciones sólidas en el capítulo IV que corresponde a los resultados de la investigación.

2.1.3 Niveles de comprensión estadística

Aprender es mucho más que recolectar información, ya que involucra sintetizar y estructurar las nuevas ideas con las ya existentes, para lograr una mejor comprensión. En 1995, Shamos enunció tres niveles de comprensión denominados: vocabulario básico, comprensión de procesos científicos y comprensión del impacto en la sociedad (citado en Budé, 2006, p. 3).

Watson por su parte en 1997 identificó: la comprensión de la terminología básica, la incorporación del lenguaje estadístico en un contexto amplio y el desarrollo de una actitud crítica hacia conceptos más sofisticados. Las tres habilidades descritas según Watson se van desarrollando a lo largo de los años de instrucción estadística basada en métodos de enseñanza innovadores; por ejemplo el uso de *software*, puede tener efectos positivos en la comprensión de los conceptos básicos, pero aun así, se encuentran alumnos que siguen manifestando dificultades conceptuales.

Ambos autores, es decir, Shamos y Watson, coinciden en que el nivel inicial es el que se refiere a la incorporación del vocabulario básico para después pasar a niveles más avanzados de conocimiento. Sin embargo, para Huberty, Dresden y Bak (1993, citado en Broers, 2008) el proceso se da de diferente manera, pues en primer lugar surge una habilidad computacional, luego el conocimiento proposicional y hasta el final se menciona la comprensión conceptual.

También Skemp (1979, citado en Kelly, Sloane y Whittaker, 1997) describe dos diferentes tipos de aprendizaje: 1) el aprendizaje basado en la comprensión instrumental, el cual consiste en el reconocimiento de una tarea para la cual se conoce una regla particular, 2) la comprensión relacional, que tienen que ver con el esforzarse para aprender con significado, es decir, se relaciona una tarea al esquema adecuado, no se

aplican fórmulas o reglas sin comprender el motivo del uso de esa fórmula. Es importante que el profesor se encuentre alerta de cuando el alumno está pasando del conocimiento instrumental al relacional para poder decir que el objetivo de la instrucción se cumplió, pero también se debe tener en cuenta que no es una tarea fácil.

delMas en 2002, distingue los objetivos de la educación estadística en alfabetización, razonamiento y pensamiento; así mismo, puntualiza una lista de términos que orientan los requerimientos para demostrar la comprensión de un dominio que va, desde los conocimientos básicos hasta llegar a analizar y generalizar resultados de un problema estadístico específico. En la tabla 1 se observa la taxonomía que el autor propone para distinguir cada una de las habilidades adquiridas en un estudiante de estadística.

Tabla 1
Taxonomía de habilidades.

Alfabetización Básica	Razonamiento	Pensamiento
Identificar	¿Por qué? ¿Cómo? Explicar (el proceso)	Aplicar
Describir		Criticar
Traducir		Evaluar
Interpretar		Generalizar
Lectura		

Nota: Adaptado de delMas, 2002

Cabe precisar que esta taxonomía no es exhaustiva de los dominios descritos. La alfabetización estadística básica puede tomar en cuenta otro tipo de habilidades como lo es comunicar y organizar datos así como lo señaló Rumsey (2002); por su parte Garfield (2002) aportó una taxonomía para el razonamiento estadístico y Wild y Pfankuch (1999) hicieron lo propio con el pensamiento estadístico, todos estos autores abordan los conceptos con la finalidad de delimitar cada uno de ellos y así determinar cuando un ciudadano esta educado estadísticamente y en consecuencia ha adquirido una comprensión de la disciplina.

Reconocer cuál es el nivel de comprensión estadística que poseen los estudiantes es, por un lado, ayudar a los educadores a identificar y a reflexionar sobre los procesos de enseñanza y, por otro, tener una medida general de lo que significa tener comprensión estadística. De esta manera, los profesores podrían utilizar un perfil estadístico como una forma de registrar el progreso de los estudiantes en el nivel de comprensión durante su escolarización. En la tabla 2, se presenta el modelo propuesto por Reading (2002), que generó al evaluar a 180 estudiantes entre 13 y 18 años en Australia.

Tabla 2
Perfil de la comprensión estadística.

Dominios	General	Recolección de datos	de Representación	Reducción de datos	de Interpretación e inferencia
Elemental	No ven los valores de los datos en un argumento, no interactúa con los datos.	Sugiere el uso de un resultado previo en un lugar de recoger los datos.	Sólo menciona el título o las variables involucradas.	Reduce los datos, pero en un valor que no está directamente relacionado a los datos.	Describe patrones, pero imposible el pronóstico.
Ciclo 1:	Centrarse sobre datos individuales, la variación no se ve como global.	Interés por los aspectos físicos de la recolección de datos como el tiempo y el costo.	Interactúa con datos en términos estadísticos, describe puntos individuales de datos.	Reducir los datos a una forma más utilizable basada en ítems de datos individuales.	Describe patrones pero predice usando la experiencia personal y no los datos.
Ciclo 2:	Centrarse sobre características de los datos, la variación se ve como global.	Interés por la calidad o exactitud de los datos recolectados identificando variables relevantes para su consideración.	Interactúa con los datos en términos estadísticos, describe características o comportamiento de los datos.	Reduce los datos a una forma más utilizable basada en las características de los datos.	Predice mediante descripciones de modelos, hace uso de los datos para justificar.

Nota: Adaptado de Reading, 2002

Reading (2002), reportó que Holmes en 1980, identificó cinco áreas estadísticas básicas: recolección de datos, representación de datos, reducción de datos, probabilidad¹, y la interpretación e inferencia. Estas cinco áreas forman parte importante del perfil; la otra parte consiste en los diversos niveles de comprensión estadística, la cual contiene la forma elemental, el ciclo uno y el dos. Con estas descripciones, se intenta cubrir todas las áreas básicas de estadística de tal manera, que en las especificaciones se puedan identificar las etapas en el desarrollo de la comprensión estadística.

Según el autor, este perfil aún se encuentra en desarrollo, sin embargo los aspectos son útiles para evaluar la comprensión de los estudiantes. La finalidad con la que Reading realizó este perfil es para dar una descripción de lo que se espera de un estudiante promedio en un grupo específico de estudiantes con respecto a la comprensión estadística.

La finalidad de definir un perfil es simplemente impulsar la comprensión dejando de lado el aprendizaje memorístico de fórmulas que con el paso del tiempo el sujeto tiende a olvidar, pues sólo se logra un modo elemental del aprendizaje en los años de instrucción y en el futuro se tendrán problemas al aplicar los conocimientos adquiridos a situaciones reales fuera del salón de clase. Entonces, sería factible involucrar al estudiante en la resolución de problemas reales y de la vida cotidiana, para que de esta forma encuentre el sentido estadístico en el mundo en el que vive.

Tradicionalmente los maestros han prestado más atención al desarrollo del plan de estudios sobre el contenido de la disciplina y se espera que los estudiantes lleguen a comprender por medio de la repetición conceptos que necesitan ser enseñados de formas diferentes, reforzando así la memorización en lugar de la comprensión.

Reading (2002) a través de Shaughnessy (1997), ha identificado la falta de comprensión como una causa importante de incapacidad de los estudiantes para razonar y pensar estadísticamente. Según Budé (2006), para evaluar la comprensión básica se solicita por ejemplo que se respondan preguntas de opción múltiple que hacen referencia a procesos básicos como lo son definiciones o bien fórmulas. En el caso de niveles más profundos de comprensión se hace más complejo el proceso ya que es esencial que las preguntas sean las apropiadas para el dominio que se está evaluando, de tal forma que se hagan planteamientos de la vida real en contextos específicos.

¹ El área de probabilidad no ha sido incluido en el perfil de la comprensión estadística por que es un tema ya bastante estudiado por otros autores, sin embargo será incluida en el futuro según los autores.

Los aspectos mencionados por delMas en la tabla 2 y el perfil descrito por Reading en la tabla 3 son útiles para evaluar la comprensión de los alumnos y saber qué esperar de un estudiante promedio con respecto a la comprensión que debe tener en cierto periodo de tiempo.

A lo largo de este apartado se ha mencionado en repetidas ocasiones los términos de razonamiento y pensamiento estadístico, los cuales deben ser definidos de forma muy específica; en los siguientes párrafos se describen estos términos.

2.2 Educación estadística

Un primer paso es definir los términos de alfabetización, razonamiento y pensamiento estadísticos, que son los dominios fundamentales para delimitar el campo de estudio.

Shamos (1995) sugiere que para abordar los conceptos previamente mencionados hay que formar un vocabulario básico que sea común para toda la comunidad científica, de tal forma que sea comprensible con la finalidad de lograr un impacto importante en la ciencia y la tecnología en la sociedad. Sin el conocimiento y comprensión del vocabulario común poco se puede hacer con la definición de los conceptos.

Siguiendo con la postura de Shamos, y centrando la atención en la educación estadística, autores como delMas, Garfield, Rumsy y Chance, entre otros, han realizado un esfuerzo para definir los términos y el enfoque teórico con el cual se aborda la educación estadística. Ben-Zvi y Garfield (2004) afirman que no existe una distinción entre los términos utilizados por parte de los investigadores y estadísticos, lo que trae como consecuencia que los objetivos que se pretende alcanzar al educar estadísticamente a una población aún son poco claros y por lo tanto, no es fácil evaluarlos.

Alfabetización, razonamiento y pensamiento estadístico se presentan como resultados deseables de aprendizaje en educación estadística; es por ello que durante el desarrollo de este apartado se hace una distinción de éstos con base en las definiciones desarrolladas en las diferentes investigaciones reportadas.

2.2.1 Alfabetización, razonamiento y pensamiento estadísticos

La delimitación de los dominios que se abordan en la educación estadística, no ha sido una tarea fácil de resolver y a través de las investigaciones se ha logrado una distinción y un acercamiento conceptual el cual es expuesto en las siguientes líneas.

Alfabetización estadística

Gal (2002) realizó un esfuerzo por definir el término como la habilidad para interpretar, evaluar críticamente y comunicar información, es decir, implica más que sólo un conocimiento mínimo de la materia, engloba actitudes, creencias y una perspectiva crítica. Se advierte que esta definición va encaminada a los sujetos que consumen los datos y no aquellos que los producen, de tal forma que quedan fuera de la definición.

Garfield (1999, citada en Rumsey, 2002) define la alfabetización estadística como la comprensión del lenguaje estadístico, es decir, palabras, símbolos y términos, así como ser capaz de leer y entender la estadística en los medios de comunicación.

Finalmente, Ben – Zvi y Garfield (2004, p. 7), la definen como habilidades que deben incluir la capacidad de organizar datos, construir tablas y trabajar con diferentes representaciones de datos. La alfabetización estadística también incluye la comprensión de conceptos, vocabulario y símbolos, al igual que una comprensión de la probabilidad como medida de incertidumbre.

En síntesis, según los niveles de comprensión descritos previamente, la alfabetización se encuentra en el nivel básico en donde se tienen que adquirir los conocimientos de los conceptos y del lenguaje estadístico, así como el uso de símbolos, también implica reconocer y organizar los datos. Es la habilidad básica que debe adquirir un estudiante alfabetizado.

La alfabetización es un objetivo clave para lograr los siguientes dominios referentes al razonamiento y pensamiento estadístico, incluso según Rumsey (2002), los maestros de estadística están de acuerdo que en cualquier curso introductorio de estadística se deberían incrementar los conocimientos de los estudiantes y prepararlos para una carrera en la actual era de la información.

La alfabetización estadística aún es un concepto demasiado amplio según ha declarado Rumsey (2002) y en su lugar menciona la competencia estadística ya que resulta más acotado; los aspectos importantes que considera son cinco que se describen a continuación.

En primer lugar, están aquellos que tienen que ver con la conciencia de datos, que son parte de la vida cotidiana y son aspectos del mundo del trabajo; también tiene que ver con el conocimiento de que frecuentemente los datos se utilizan de forma inadecuada; finalmente según la autora, es importante tener en cuenta que las decisiones que se toman sobre la base de los datos puede tener un fuerte impacto en nuestras vidas.

En segundo lugar, tienen que ver con uso de términos, que es un prerrequisito para interpretar mensajes estadísticos, se trata de conocer los conceptos de estadística y probabilidad; Scheaffer, Watkins y Landwehr (1998, citado en Gal, 2002) revisaron varios currículos y con base en esto describieron numerosas áreas esenciales para incluirlas en el estudio de temas estadísticos en la preparatoria, algunas de ellas tienen que ver con:

- El sentido numérico y la interpretación de tablas y gráficas.
- Aspectos para planear una encuesta tal como constituir una buena muestra o un método de recolección de datos.
- El proceso de análisis de datos para determinar la frecuencia de los mismos.
- La relación entre probabilidad y estadística tal como determinar características de muestras aleatorias y conocer la significancia estadística.
- Finalmente otra área esencial es el razonamiento inferencial, que tiene que ver con intervalos de confianza o pruebas de hipótesis.

En tercer lugar, Rumsey (2002) aborda el conocimiento de los conceptos básicos de la recopilación de datos, así como generar estadística descriptiva. En cuarto lugar la interpretación en un nivel básico, lo que implica obtener conclusiones adecuadas a partir de la información estadística. Finalmente, para Rumsey las habilidades básicas para la comunicación son muy importantes, ya que si bien, la interpretación es clave para demostrar el entendimiento, la comunicación permite transmitir la información a otras personas, es una habilidad que se tiene que desarrollar a lo largo de la formación.

Razonamiento estadístico

El razonamiento estadístico también juega un papel muy importante para educar estadísticamente a una población, por lo que Garfield (2002), lo define como la forma en que las personas razonan con ideas estadísticas y dan sentido a la información. Rumsey (2002), en una forma breve explica que el razonamiento es la capacidad de cuestionar, comparar y explicar.

Autores como Ben – Zvi y Garfield (2004) afirman que el razonamiento estadístico involucra hacer interpretaciones de conjuntos, representación, o bien, resúmenes estadísticos de datos. También involucra conectar un concepto con otro, o incluso combinar ideas acerca de los datos y el azar. Para Ben – Zvi y Garfield, razonamiento significa entender, ser capaz de explicar procesos estadísticos y de interpretar completamente los resultados.

Incluso Garfield (2002) afirma que algunos investigadores de la educación se han centrado en seis tipos de razonamiento: 1.razonamiento sobre los datos; 2. acerca de las representaciones de los datos; 3. referente a las medidas estadísticas; 4.sobre la incertidumbre; 5. referente a la muestra; y finalmente, 6. la asociación.

Un modelo que guía a los diferentes tipos de razonamiento previamente mencionados es el desarrollado por Garfield (2002), que incluye una serie de pasos que tienen que ocurrir para llegar a desarrollar el razonamiento estadístico en un estudiante. El modelo incluye los siguientes niveles:

Nivel 1. Razonamiento idiosincrásico: El estudiante conoce algunas palabras y símbolos estadísticos, los usa sin plena comprensión de ellos, a menudo incorrectamente. Por ejemplo, los estudiantes han aprendido términos como media, mediana y desviación estándar como medidas de resumen, pero los usan indiscriminadamente.

Nivel 2. Razonamiento verbal: El estudiante tiene una comprensión verbal de algunos conceptos; puede dar una definición correcta, pero no comprenderlos completamente.

Nivel 3. Razonamiento de transición: El estudiante es capaz de identificar correctamente una o dos dimensiones de un proceso estadístico sin integrar completamente estas dimensiones, por ejemplo que una muestra grande conduce a un intervalo de confianza más ajustado.

Nivel 4. Razonamiento de procedimiento: El estudiante es capaz de identificar correctamente las dimensiones de conceptos estadísticos o procesos pero no los integra o comprende completamente.

TESIS TESIS TESIS TESIS TESIS

Nivel 5. Proceso de razonamiento integrado. El estudiante tiene una comprensión completa de un proceso estadístico. El estudiante puede explicar el proceso en sus propias palabras con plena confianza.

Estos niveles delimitan un proceso específico que se va construyendo durante la formación del razonamiento estadístico y donde previamente se alfabetizó estadísticamente al estudiante.

Pensamiento estadístico

El proceso de educación estadística, continúa con un tercer dominio que delimita lo que de aquí en adelante se entenderá como pensamiento estadístico.

El término que se intenta definir, ha pasado por una serie de transformaciones, incluso no es un concepto aislado ya que retoma partes de la alfabetización y el razonamiento estadístico y con la finalidad de construir un concepto que implique los dos dominios previos, la Asociación Americana de Estadística (ASA) y la Asociación Matemática de América (MAA) (Cobb, 1992 citado en Chance, 2002) involucran aspectos como la necesidad de datos, la importancia de la producción de datos, la omnipresencia de la variabilidad y la medición de la variabilidad.

El pensamiento estadístico ha sido definido por Snee (1990, citado en Chance, 2002) como el proceso de pensamiento que reconoce que la variación está a nuestro alrededor y presente en todo lo que hacemos, todo el trabajo es una serie de procesos interconectados, e identifica, caracteriza, cuantifica, y reducen la variación que proporciona oportunidades de mejora.

Wild y Pfannkuch (1999), tratan de identificar elementos importantes del pensamiento estadístico y abordan un ciclo de cuatro dimensiones. En la primera dimensión llamada ciclo de investigación se encuentran particularidades como el definir un problema, planificar, recolectar datos, planear un análisis, generar hipótesis y realizar interpretaciones y nuevas ideas. La dimensión dos está caracterizada por buscar explicaciones, seguir procedimientos, reconocer la necesidad de los datos, transmitir mensajes en datos, considerar la importancia de la variación, considerar el razonamiento con modelos estadísticos, además de integrar la estadística y el contexto. La tercera dimensión está representada por generar explicaciones, buscar información interna y externa, interpretar información, es decir, compara y conecta y por último abarca aspectos como criticar y realizar juicios. Finalmente en la cuarta dimensión llamada de disposición, se aborda la imaginación y la curiosidad del individuo y está abierto a ideas que cambian

TESIS TESIS TESIS TESIS TESIS

las preconcepciones. De esta forma los autores tratan de dar explicación al pensamiento estadístico y realizan una discusión a profundidad de cada una de las dimensiones.

Por otro lado Chance (2002, p. 5) sigue el enfoque de Wild y Pfannkuch al afirmar que el pensamiento estadístico es el proceso que implica el resumen de datos, la resolución de problemas en particular, el razonamiento por medio de procedimientos y la explicación de conclusiones, es decir, lo que es único en el pensamiento estadístico, según el autor, es la habilidad para ver el proceso como un todo para entender la relación y significado de la variación en este proceso; tener la capacidad de explorar datos más allá de lo que está escrito en los textos para de esta forma generar nuevas preguntas.

En 2004, Ben – Zvi y Garfield hablan del pensamiento estadístico y puntualizan que adquirir este dominio implica la comprensión de la naturaleza de la muestra, además de entender por qué se necesitan modelos para simular fenómenos aleatorios, cómo los datos se producen para estimar probabilidades. Para los autores, el pensamiento estadístico incluye la capacidad de comprender y utilizar el contexto de un problema para formar investigaciones, elaborar conclusiones y reconocer y entender todo el proceso. En resumen, los pensadores estadísticos son capaces de criticar y evaluar los resultados de un problema resuelto o de un estudio estadístico.

En esta etapa se involucra tanto al proceso de alfabetización como al de razonamiento, no pueden ser elementos independientes uno del otro. En conclusión, la alfabetización está estrechamente relacionada con los métodos y conceptos aprendidos en un curso; el razonamiento involucra hacer resúmenes de datos y ser capaz de explicar procesos estadísticos y además de interpretar resultados; finalmente el pensador estadístico es capaz de explicar la relación de la variación en un conjunto de datos, además elabora sus propias conclusiones a partir de la comprensión del proceso estadístico

Ya se han definido los términos de alfabetización, razonamiento y pensamiento estadístico, pero en el siguiente apartado se encuentran las características que permiten integrar estos conceptos, es decir, necesariamente debe existir uno para que se pueda dar el otro.

2.2.2 Interacción de la alfabetización, razonamiento y pensamiento estadístico

Autores como delMas (2002, p.5), han señalado que las definiciones de alfabetización, razonamiento y pensamiento estadísticos están muy relacionadas entre sí, pero a pesar de ello hace un esfuerzo por distinguirlas y a la vez integrarlas. Este autor afirma que la alfabetización está vinculada con habilidades y conocimientos que son necesarios para desarrollar los otros dos dominios y propone un diagrama de Venn para representar esta idea. Sosteniendo que cada dominio tiene contenidos que son independientes de los otros dos y donde hay cierta superposición entre ellos, la figura 1 ejemplifica la idea propuesta por el autor. En el diagrama de Venn que a continuación se expone, la alfabetización estadística está en la base y es el pilar para que de aquí surja el razonamiento y el pensamiento como formas de conocimiento más avanzadas.

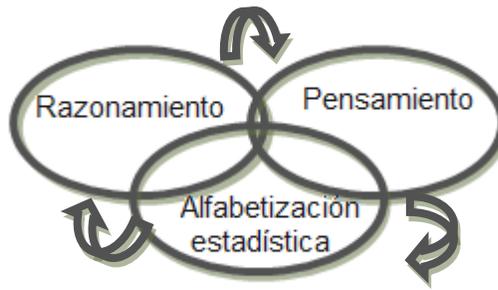
Figura 1. Integración de la alfabetización, razonamiento y pensamiento estadístico



Fuente: Adaptado de delMas, 2002

La instrucción en estadística transforma el razonamiento y en consecuencia el pensamiento de los individuos; si el objetivo de la alfabetización es realizar un análisis crítico, la meta del razonamiento es justificar por qué una respuesta es o no adecuada y el fin último del pensamiento estadístico es identificar la variabilidad; la propuesta a partir de esta distinción genera necesariamente un ciclo, siendo la alfabetización lo que da lugar al razonamiento y al pensamiento y si falta alguno difícilmente se accede al siguiente dominio, de esta forma se logra educar estadísticamente a los ciudadanos. Ver Figura 2.

Figura 2. Ciclo de la educación estadística.



Fuente: Adaptado de delMas, 2002

Las propuestas desarrolladas anteriormente, muestran cómo están relacionados los conceptos sin necesidad de que estos sean tratados de una forma aislada, pues como se comentó, se encuentran interconectados, de tal forma que se logra un trabajo cíclico entre cada dominio, cada uno con objetivos específicos para que puedan ser evaluados al término de los cursos de estadística.

Una alternativa para evaluar los dominios previamente descritos, es a través de las pruebas de logro que son diseñadas con una metodología estricta, de tal manera que los resultados de las evaluaciones al ser aplicadas a los estudiantes puedan aportar información valiosa con relación a la comprensión que se ha adquirido a lo largo de la instrucción estadística, es por ello que en el siguiente apartado, se describen algunas de las evaluaciones diseñadas para dominios específicos.

2.3 Pruebas de logro

Las pruebas de ejecución máxima o también llamadas de logro están diseñadas para medir los logros de los estudiantes, como resultado de la exposición a una experiencia de aprendizaje definida. Este tipo de pruebas pueden estandarizarse a nivel nacional, regional o bien local (Cohen y Swerdlik, 2002, p. 305)

2.3.1 Pruebas de logro en estadística

En 2000, Garfield y Chance (citado en delMas, Garfield, Ooms y Chance, 2007), han expuesto que muchos instrumentos de evaluación son los exámenes finales que diseñan los maestros y que frecuentemente son poco apropiados si estos se centran en los procedimientos y definiciones más que en la comprensión conceptual. En virtud de que es importante contar con elementos válidos y confiables de evaluación, en la actualidad ya hay intentos para evaluar el uso de la estadística, y algunos de los ejemplos se mencionan a continuación.

Para evaluar el razonamiento estadístico, se desarrolló la evaluación The Statistical Reasoning Assessment (SRA, por sus siglas en inglés) la cual se centra en gran medida en la probabilidad, pero esta prueba carece de ítems relacionados con la recolección de datos y la inferencia estadística (Garfield, 2003), lo que ha obligado a desarrollar otras evaluaciones más completas. Otro intento fue el Inventario de Conceptos Estadísticos (SCI, por sus siglas en inglés) desarrollado para evaluar la comprensión estadística pero éste fue escrito para una audiencia específica de estudiantes de ingeniería en estadística (delMas, Garfield, Ooms y Chance, 2007).

Garfield, delMas, Chance, Poly y Ooms, (2005) propusieron desarrollar un instrumento de evaluación que tuviera una cobertura más amplia que incluyera el contenido estadístico normalmente cubierto en el primer curso de estadística y poder aplicarlo a la mayor parte de los estudiantes de estos cursos.

Estos ejemplos nos muestran que si bien hay intentos por evaluar el empleo de la estadística, aún hay mucho por hacer en esta área, pues se han estandarizado pruebas en contextos específicos, pero es importante que otros países comiencen a retomar estos instrumentos para adecuarlos a las necesidades de la población y de esta forma consolidar pruebas que permitan dar cuenta de aspectos específicos a los que hace referencia la educación estadística.

En el siguiente apartado se mencionan las pruebas de logro desarrolladas en México con la finalidad de conocer específicamente el trabajo en evaluación estadística de instituciones dedicadas a la evaluación.

2.3.2 Pruebas de logro en México

Uno de los referentes importantes para mencionar las pruebas de logro estandarizadas en México es el Instituto Nacional para la Evaluación de la Educación (INEE), ya que esta institución se encarga de realizar evaluaciones estandarizadas a estudiantes de educación obligatoria. Una de las características del instituto, es ofrecer información con relación al rendimiento académico de los estudiantes a través de la prueba EXCALE que ellos mismos han estandarizado. Entre las áreas que evalúa la prueba se encuentra la de matemáticas y dentro de ésta se incluye la parte estadística en temas de probabilidad e interpretación de gráficas, aunque las pruebas están dirigidas a la educación básica (Backhoff, Andrade, Sánchez y Peon, 2008).

Otra institución que se dedica a la evaluación dentro del país y es importante mencionar es el Centro Nacional de Evaluación para la Educación Superior, A.C.

(CENEVAL) y su principal actividad educativa se centra en el diseño y aplicación de instrumentos de evaluación de conocimientos, habilidades y competencias a personas de diferentes niveles educativos, centrandose gran parte de su trabajo evaluativo a la educación media y superior.

A través del Examen Nacional de Ingreso (EXANI) y el Examen General de Egreso de Licenciatura (EGEL), entre otros exámenes que se han desarrollado con el fin de conocer los aprendizajes adquiridos, se realiza una breve evaluación de la competencia de estadística, es decir, lo que delMas y colaboradores denominan alfabetización estadística.

Hasta 2008, en México no se tenían reportes de evaluaciones estandarizadas que midieran específicamente la competencia estadística, en este sentido, el CENEVAL² se dio a la tarea de desarrollar el examen EXTRA-ES³ que previamente ya se había descrito. Este examen se caracteriza por que se evalúan los conocimientos y habilidades estadísticas que han adquirido los estudiantes de licenciatura; en el área básica se evalúan temas como fundamentos del pensamiento estadístico, estructura y generación de datos; descripción, organización e interpretación de los datos; y finalmente nociones de inferencia. En el segundo módulo denominado métodos estadísticos se evalúan dos áreas, la primera hace referencia a la inferencia paramétrica y no paramétrica y la segunda se refiere al análisis de varianza y regresión. En el tercer módulo, que se refiere al muestreo, se evalúan las áreas relacionadas con fundamentos de muestreo y descripción, organización e inferencia a partir de los datos en muestreo. El módulo de estadística experimental está estructurado por dos áreas principales, la primera se refiere a los fundamentos de estudios estadísticos y diseños de experimentos y la segunda hace referencia al análisis de diseños experimentales.

Cabe destacar que en algunas preguntas que componen esta prueba, el estudiante utiliza un formulario estadístico para poder obtener la respuesta correcta de una serie de respuestas de opción múltiple que contiene el ítem; en otros reactivos el estudiante debe recordar conceptos o términos estadísticos para seleccionar las respuestas correctas. En conclusión el EXTRA-ES es una evaluación importante para determinar si un estudiante ha logrado una alfabetización estadística adecuada, es decir, como afirma Budé (2006) para evaluar la comprensión básica se diseñan preguntas con relación a definiciones o fórmulas, sin embargo los dominios de razonamiento y pensamiento estadísticos deben

² Ver el capítulo del planteamiento del problema.

³ Ver anexo C referente a lo que evalúa el examen de estadística diseñado por el CENEVAL.

ser fortalecidos con planteamientos de la vida real, es por ello que en el marco de esta investigación se ha retomado el instrumento de evaluación CAOS-4 que centra su interés en el razonamiento y pensamiento estadístico de los estudiantes de nivel superior, evaluación que se ha mencionado en párrafos previos y será retomado más ampliamente en el siguiente apartado.

2.3.3 Desarrollo de CAOS- 4

En respuesta a las dificultades que los estudiantes tienen con el aprendizaje y comprensión estadística, en 1900 se inició en Estados Unidos un movimiento de reforma con la finalidad de transformar la enseñanza de los niveles introductorios de estadística (Cobb, 1992 citado en Ben-Zvi y Garfield, 2004). Moore (1997, citado en Ben-Zvi, 2004) describió el movimiento de reforma como cambios en el contenido, la pedagogía y la tecnología.

Como parte de este movimiento surgieron varios institutos, entre ellos *The National Science Foundation* (NSF) que fundó el proyecto *Assessment Resource Tools for Improving Statistical Thinking* (ARTIST) para direccionar los cambios de evaluación en educación estadística representados por Garfield y Gal, quienes destacaron la necesidad de desarrollar elementos de evaluación fiables, válidos y prácticos. La página de ARTIST (ver <https://app.gen.umn.edu/artist/>), en la actualidad ofrece una extensa variedad de recursos de evaluación para evaluar la alfabetización, razonamiento y pensamiento estadístico de los estudiantes (delMas, Garfield, Ooms y Chance, 2007).

Una propiedad importante del proyecto ARTIST fue la de proponer la evaluación CAOS-4 con la intención de que sea una evaluación fiable compuesta de un conjunto de ítems que los estudiantes de cualquier curso introductorio de estadística pudieran comprender. Teniendo en cuenta que se puede realizar una evaluación fiable, se incluye una segunda meta que se refiere a identificar las áreas donde los estudiantes logran o no avances significativos en su razonamiento y comprensión estadística.

La prueba de CAOS-4 se desarrolló en un periodo de tres años al conseguir ítems de maestros expertos; también se trabajó para escribir preguntas que incluyera las áreas no cubiertas por los ítems iniciales que se formularon para construir la prueba,, además de que se revisaron a profundidad los ítems y así mismo se logró obtener retroalimentación de asesores y evaluadores expertos.

TESIS TESIS TESIS TESIS TESIS

Con relación a la validez según Aiken (1996, p. 95) se define como el grado en el cual una prueba mide aquello para lo que se diseñó. Es decir, qué tan bien mide lo que pretende medir en un determinado contexto (Cohen y Swerdlik, 2002, p. 157).

Para determinar la validez de contenido, los autores de CAOS-4 llevaron a cabo una serie de etapas que fueron perfeccionando el instrumento para el contexto en el que fue aplicado originalmente y que se describe en los siguientes párrafos.

La primera versión de CAOS, contaba con 34 ítems de opción múltiple; esta versión fue usada en un estudio piloto con estudiantes universitarios de estadística introductoria durante el otoño de 2004. Los datos del estudio piloto fueron usados para hacer una revisión adicional de CAOS, resultando una segunda versión que contenía 37 ítems de opción múltiple.

La segunda versión llamada CAOS-2, estaba lista para lanzarse como una evaluación en línea en enero de 2005. La administración de la prueba en línea requería del registro de los maestros, además era un medio para que los estudiantes accedieran de forma segura a la prueba. Los resultados de los análisis de datos de la prueba de CAOS-2 fueron usados para hacer cambios adicionales, lo que produjo la tercera versión denominada CAOS-3.

Aunque las puntuaciones obtenidas por los estudiantes a partir de la evaluación CAOS-3 indicaba que la prueba medía lo que debía medir, los profesores y los evaluadores expertos, realizaron sugerencias de cambios. Esta retroalimentación fue usada para sumar y eliminar reactivos de la prueba, además de realizar revisiones extensas para producir la versión final de la prueba, llamada CAOS-4, que actualmente consta de 40 ítems de opción múltiple.

En Marzo de 2006, el análisis final de validez de contenido de CAOS-4 se realizó por los investigadores de la Universidad de Minnesota. La evaluación se aplicó a un total de 1470 estudiantes y con esta muestra se realizó un análisis de consistencia interna de los 40 ítems en el que se generó un alfa de Cronbach de 0.82 y con estos resultados, la consistencia interna se consideró aceptable.

Esta propuesta que hace el grupo de delMas, Garfield, Ooms y Chance. (2007), ofrece información valiosa sobre lo que los estudiantes parecen aprender y comprender después de completar un primer curso de estadística en la universidad.

Cabe señalar, según delMas, Garfield, Ooms y Chance (2007), que todos los elementos en la prueba CAOS-4 fueron pensados para exigir a los estudiantes a razonar

y pensar, no para calcular, usar fórmulas o definiciones de recuerdo, a diferencia de muchos exámenes diseñados por un docente.

Cuando los profesores conocieron los hallazgos, muchos de ellos informaron que se sorprendieron al ver los resultados obtenidos por los estudiantes y causó reflexión sobre su propia enseñanza. Ese es uno de los propósitos más importantes de la prueba CAOS-4, proporcionar información útil a los profesores de enseñanza de las estadísticas, con el fin de que puedan ver si sus estudiantes están aprendiendo a razonar y pensar estadísticamente y de esta forma, promover cambios en la instrucción si es necesario.

2.3.4 Estructura de CAOS-4

Para conocer la estructura de la prueba CAOS-4⁴, se presenta la tabla 3 que explica la distribución de las dimensiones evaluadas por el instrumento. Para saber qué mide cada reactivo del instrumento, se recomienda consultar el anexo D.

⁴ Revisar el anexo D, para encontrar más detalle de la descripción de cada reactivo que compone la prueba.

Tabla 3
Estructura CAOS-4

Dimensión	Definición Operacional	No. Pregunta	Total
Recolección de datos	Seleccionar una muestra representativa de la población así como lograr hacer una distribución adecuada del muestreo para poder generalizar los datos de la población.	32, 34, 35 y 38	4
Estadística descriptiva	Comparar grupos mediante las diferencias en promedios, así como estimar desviación estándar y compararlas.	12, 14 y 15,	3
Representación gráfica	Describir e interpretar con un histograma la distribución de una variable.	1, 3, 4 y 5	4
Diagrama de cajas	Capacidad para determinar el diagrama de cajas que ofrece estimaciones precisas de desviación estándar y mediana	2, 8, 9 y 10	4
Distribución normal	Describir una distribución a partir de un histograma identificando los datos. Determinar que sucede si la mediana es más grande que la media.	6, 11 y 33	3
Datos bivariados	Describir una relación bivariada mostrada en un diagrama de dispersión cuando hay y no hay un valor atípico.	20, 21 y 22,	3
Probabilidad	Simular datos para encontrar la probabilidad de un valor observado, así como identificar el sentido de la aleatorización.	7, 24, 36 y 37	4
Variabilidad de la muestra	Identificar que en dos grupos no se requieren muestras iguales, especialmente si ambos conjuntos de datos son grandes, así mismo señalar que existe más variabilidad en muestra pequeñas que en las grandes.	13, 16, 17 y 18	4
Intervalos de Confianza	Interpretación de los límites de confianza.	28, 29, 30 y 31	4
Pruebas de significancia	Conocer que los bajos valores de p son deseables en estudios de investigación.	19, 23, 25, 26, 27, 39 y 40	7

Nota: Adaptado de delMas, J. Garfield y A. Ooms, 2007

A través de la revisión de la literatura hecha en este capítulo se percibe que la comprensión estadística es un tema relevante para diversos autores en el ámbito internacional, no obstante, en México este tema está cobrando relevancia a través de los grupos de trabajo y de las investigaciones realizadas hasta hoy, de tal forma que actualmente se habla de niveles de comprensión estadística que han surgido de los esfuerzos por identificar a un estudiante que ha logrado los conocimientos mínimos indispensables de la alfabetización estadística.

Al referirse al proceso de alfabetización estadística necesariamente se hace mención de los términos de razonamiento y pensamiento estadístico ya que van muy de la mano y evaluar estos dominios no es un proceso sencillo de seguir.



CAPÍTULO III

DISEÑO METODOLÓGICO

3.1 Enfoque metodológico

Esta sección se centra en describir el diseño metodológico que se empleó en el proceso de investigación, con la finalidad de delinear las estrategias de obtención de información y de análisis respecto a la comprensión estadística de los estudiantes universitarios; se abordan entonces apartados clave que orienten hacia el análisis de resultados.

En el capítulo 2 correspondiente al marco teórico se realizó un breve recorrido del estudio de la educación estadística y de la comprensión de la misma, en la que se aprecian diferentes metodologías de investigación

El diseño metodológico que se eligió para el desarrollo de este estudio fue de corte cuantitativo, a través de un diseño no experimental, transversal, descriptivo y correlacional. Según lo expuesto por Hernández, Fernández y Baptista (2006) en primer lugar se denominó cuantitativo porque se pretendió explicar el fenómeno de la comprensión estadística al medir diferentes variables y encontrando su posible relación; en segundo lugar, un indicador según estos autores, de que es un estudio basado en un diseño no experimental, es porque no se realizó experimento alguno sino que se midieron factores ya acontecidos en los sujetos de estudio que fueron seleccionados mediante un criterio establecido; también se consideró transversal, porque la recopilación de datos se dio en un momento único en el tiempo, “es como tomar una fotografía de algo que sucede” (p. 208). Finalmente se denominó descriptivo y correlacional porque se indagó la incidencia de una o más variables en una población y porque posteriormente se pretendió establecer la relación parcial entre las variables tipo $X_1 \Theta Y_1$ en los subgrupos de estudio.

Las variables que se tomaron en cuenta para este estudio son presentadas en los siguientes párrafos, en donde se describen cada una de ellas, así como su definición operacional.

3.2 Variables

En este apartado se presentan las variables dependientes e independientes que fueron consideradas para el proceso de investigación. La operacionalización de la variable dependiente y de las variables independientes se realizó con base en lo ya planteado por Hernández, Fernández y Baptista (2006).

3.2.1 Variable dependiente

La variable comprensión estadística es una construcción que se apoya en lo que diversos autores como Garfield y delMas han aportado a la literatura con la finalidad de tener una visión más clara de lo que significa el concepto que de por sí es complejo de observar. A partir de lo que se ha discutido a lo largo de esta tesis, se definirá comprensión estadística como el entendimiento de la información en forma escrita o simbólica, para responder a planteamientos que requieren interpretación de datos estadísticos. de manera crítica con la finalidad de explicar el proceso estadístico.

La comprensión estadística, como ya se ha venido discutiendo, se evaluará a través de la prueba CAOS-4⁵ y en este entendido en la tabla 4, se describen los dominios e indicadores de cada una de las dimensiones que componen la prueba.

⁵ Revisar anexo D para más detalles de los reactivos.

Tabla 4
Definición operacional de comprensión estadística para CAOS-4

Dimensión	No. Pregunta	Dominios	Indicador
Recolección de datos	32, 34, 35 y 38	Pensamiento	Evaluar Aplicar Generalizar
Estadística descriptiva	12, 14 y 15,	Pensamiento Razonamiento	Critica Explica el proceso
Representación gráfica	1, 3, 4 y 5	Alfabetización	Describir Interpretar
Diagrama de cajas	2, 8, 9 y 10	Alfabetización	Identificar Interpretar
Distribución normal	6, 11 y 33	Alfabetización Pensamiento	Descripción Interpretar Evaluar
Datos bivariados	20, 21 y 22,	Alfabetización	Interpretar Traducir
Probabilidad	7, 24, 36 y 37	Razonamiento Pensamiento	Explica Criticar Evaluar Aplicar
Variabilidad de la muestra	13, 16, 17 y 18	Pensamiento	Evaluar Aplicar
Intervalos de confianza	28, 29, 30 y 31	Pensamiento	Evaluar Criticar
Pruebas de significancia	19, 23, 25, 26, 27, 39 y 40	Pensamiento	Evaluar Aplicar

3.2.2 Variables independientes

El aprendizaje de la estadística se desarrolla en contextos concretos y es preciso conocerlos para identificar aquellas variaciones en los resultados de pruebas específicas como la que aquí se presenta, es decir, las variables dependientes pueden o no afectar a la variable dependiente.

Los factores relevantes que afectan el proceso de comprensión estadística pueden ser de diferente índole; según Durón y Oropeza (1999, citados en Izar, Ynzunza y López, 2011) se pueden encontrar factores fisiológicos que se sabe que afectan, pero es difícil

precisar en que medida afectan pues interactúan con otros factores; algunos ejemplos de este tipo de factores están referidos a la deficiencia de órganos de los sentidos, desnutrición y problemas de salud. Otro tipo de factores son los pedagógicos, que tienen que ver con la calidad de la enseñanza; Entre los más destacados, se encuentra el número de alumnos por maestro, los métodos y materiales didácticos utilizados, la motivación de los estudiantes y profesores y finalmente el tiempo que los profesores dedican a preparar sus clases. Los factores psicológicos que tienen que ver con procesos de memoria, con la percepción; y finalmente se enumeran los factores sociológicos que tiene que ver con el ambiente que rodea al estudiante, con aspectos familiares y socioeconómicos

Algunas de las variables consideradas en este estudio son personales es decir: sexo, edad, carrera y semestre.

También se incluyen las variables:

- Opiniones respecto a la estadística

Hace referencia al grado de acuerdo o desacuerdo que manifiesta el alumno tener hacia la estadística que se enseña como parte del currículo, dentro de su formación universitaria.

- Número de materias de estadística cursadas en la universidad

Incluye el número total de las materias de estadística que ha cursado el estudiante durante su formación universitaria.

De acuerdo con la delimitación de las variables expuestas en este apartado, a continuación se describirán los instrumentos que ayudaron a la recolección de información y los cuales fueron clave para exponer los resultados en el siguiente capítulo.

3.3 Instrumentos de medición

Para obtener la información de las variables previamente descritas, se utilizaron dos instrumentos, el primero fue la prueba CAOS-4 que ya fue descrita en apartados anteriores⁶ y el segundo fue un cuestionario y una escala de opinión.

3.3.1 CAOS-4

Para esta investigación, se hizo una traducción al español⁷ de la prueba CAOS-4 diseñada por los investigadores Joan Garfield, Bob delMas, Ann Ooms, Beth Chance y Cal Poly en 2005. La prueba CAOS-4, como ya se mencionó, está estructurada para evaluar las siguientes áreas temáticas: recolección de datos, estadística descriptiva, representación gráfica, diagrama de caja, distribución normal, datos bivariados, probabilidad, variabilidad de la muestra y pruebas de significación.

La prueba CAOS-4 se obtuvo de la página de web ARTIST diseñada por el equipo de Garfield y delMas de la Universidad de Minnesota, desarrollada con la finalidad de proponer un instrumento de evaluación que pudiera tener una cobertura amplia del contenido estadístico abordado en un primer curso básico de estadística.

El Dr. Robert delMas, encargado de administrar la página ARTIST, en febrero de 2012, dio la autorización vía correo electrónico para realizar la traducción del instrumento del idioma inglés a español con el propósito de aplicarlo a estudiantes de nivel superior en la Universidad Autónoma de Aguascalientes.

El trabajo de traducción del instrumento lo realizó la responsable de la investigación, en tres momentos: una primera etapa consistió propiamente en la traducción del documento; posteriormente se sometió a un proceso de revisión que incluía verificar la redacción y las instrucciones y estuvo a cargo de dos estudiantes de la maestría en investigación educativa; en tercer lugar el documento fue revisado por el Mtro. Piry Herrera, profesor de inglés del Departamento de Idiomas de la Universidad Autónoma de Aguascalientes, el cual validó que la traducción fuera fiel y que las preguntas tuvieran el sentido del documento original y ésta última versión fue enviada al Dr. delMas⁸ para que validara el trabajo que se había realizado. En todas las revisiones hubo observaciones de redacción y de traducción las cuales se se atendieron para mejorar la calidad del

⁶ Revisar capítulo 2 apartado de desarrollo de CAOS-4.

⁷ Tener en cuenta que no se anexa la prueba en español, tan sólo se encontraran algunos reactivos de la prueba.

⁸ Mensaje en una lista de correos electrónicos del 1 de octubre de 2012, Re: Studentes CAOS

instrumento en tanto, suponen dar mayor confiabilidad al trabajo realizado. En la tabla 5 se reportan las principales observaciones recabadas durante el proceso.

Tabla 5
Comentarios de los expertos que revisaron la traducción de CAOS-4

Apartado	Comentarios
Traducción	Es un instrumento en idioma inglés, por lo que la traducción de la prueba CAOS-4 debe ser específica, se encontraron cuestiones de precisión en la redacción de los ítems que fueron marcados para su posterior corrección.
Estructura de las preguntas	La redacción de las preguntas debe ser clara, sin preámbulos innecesarios; en el diseño de gráficas y tablas, se necesita obtener un tamaño adecuado de estas, de tal forma que la figura se pueda observar sin problemas para su interpretación.
Extensión de la prueba	Es un instrumento extenso que puede tornarse tedioso para los alumnos y esto puede influir en los resultados de la evaluación.
Observaciones generales	Se tuvieron entrevistas breves con cada uno de los jueces que evaluaron la prueba, para conocer la pertinencia del instrumento y se determinó que éste se aplicaría en su totalidad a la muestra universitaria propuesta ya que al ser aplicada ofrecerá información relevante para dar respuesta a los planteamientos planteados en el capítulo de planteamiento del problema.

Las sugerencias y comentarios realizados por cada uno de los expertos se integraron en la versión final del instrumento y se hicieron las correcciones pertinentes para cada caso cuidando la fidelidad del instrumento así como la extensión del mismo.

La versión en español estuvo lista a finales del mes de octubre de 2012 y el documento impreso quedó listo para aplicación la primera semana del mes de noviembre de ese mismo año.

Es fundamental resaltar que CAOS-4 no se había aplicado en América Latina, por lo que en este trabajo sólo se presentan algunos de los reactivos de la traducción del instrumento. La versión en inglés de la evaluación CAOS-4 ha sido usada en otros países como Holanda, Mónaco y Reino Unido, además, un estudiante de maestría en Turquía, modificó algunos de los ítems del instrumento para crear una evaluación de comprensión estadística utilizando contextos que pudieran ser familiares para los estudiantes de aquella región (delMas⁹, 2012)

La prueba CAOS-4 se integró con 40 reactivos de opción múltiple distribuidos de la siguiente forma: 3 reactivos de recolección de datos, 3 de estadística descriptiva, 4 de representación gráfica, 4 de diagrama de cajas, 3 de distribución normal, 3 de datos bivariados, 4 de probabilidad, 5 de variabilidad de la muestra, 4 de intervalo de confianza y finalmente 7 ítems de pruebas de significancia. (Ver Tabla 4)

Los reactivos de CAOS-4 procuran evaluar el manejo de la información y no la memorización del conocimiento, de tal forma, las opciones de respuesta de cada ítem permiten la reflexión de los estudiantes para encontrar las posibles soluciones a los planteamientos presentados. Este tipo de reactivos, en dónde se evita el uso de la memorización de fórmulas y conceptos, solicita un nivel más complejo de conocimiento por parte de los alumnos, o lo que es lo mismo, el razonamiento y el pensamiento estadístico, como ya se explico anteriormente.

3.3.2 Cuestionario y escala de opinión

Uno de los objetivos de la investigación fue conocer la opinión de los estudiantes respecto a la estadística y para ello se desarrolló un cuestionario de 8 preguntas en forma de escala; además, se logró obtener información acerca de sexo, edad y carrera como parte de los datos generales.(Ver Anexo A del cuestionario de opinión)

Para la construcción de la escala de opinión se revisó y se tomó como base la escala de Salazar (2008) referente a las actitudes de los estudiantes hacia la estadística, adicionalmente, se revisó una serie de cuestionarios que se consideraron pertinentes, con la finalidad de adaptar una estructura al cuestionario final que se utilizó en esta investigación.

Posteriormente, el instrumento se validó con la opinión de jueces expertos; y finalmente se obtuvo la primera versión del cuestionario el cual contenía 10 preguntas que se aplicó en una prueba piloto junto con la evaluación CAOS-4.

⁹ Mensaje en una lista de correos electrónicos del 27 de febrero de 2012. Re: use the CAOS in Latin America

En una primera etapa, el cuestionario de opinión comprendió opciones de respuesta de opción múltiple y dicotómicas; después del primer pilotaje, el instrumento fue transformado en una escala; posteriormente la escala fue validada nuevamente por expertos en el tema añadiendo algunas preguntas que se consideraron fundamentales; finalmente se aplicó el cuestionario a 30 estudiantes de mercadotecnia para verificar la precisión y claridad del instrumento.

Los ítems que se integraron en el cuestionario final tenían el propósito de describir las opiniones de los estudiantes por medio de una escala tipo likert, además de recabar información de orden general como sexo, edad, licenciatura a la que pertenecían y semestre.

3.5 Pilotaje de los instrumentos

En este apartado se describe el proceso del pilotaje el cual se realizó en dos etapas con la finalidad de verificar que la traducción realizada de la evaluación CAOS-4 fuera comprensible para los sustentantes a los que se dirigió y así mismo estimar el tiempo aproximado de respuesta en formato de lápiz y papel, además fue importante comprobar que el cuestionario diseñado para recolectar información respecto a las opiniones de los estudiantes fuera claro y conciso.

Los participantes del pilotaje fueron escogidos con características similares a la muestra original, empleando las mismas condiciones de aplicación de la llevada a cabo con la muestra definitiva, durante el mes de noviembre de 2012.

El criterio de selección de los estudiantes de la primera aplicación piloto fue bajo el criterio de un muestreo no probabilístico intencional; el proceso de selección de los estudiantes se llevó acabo de la siguiente manera:

Se convocó a los estudiantes del Centro de Ciencias Sociales y Humanidades para participar en este proceso de piloteo teniendo una respuesta favorable y en la que participaron en una primera etapa 68 alumnos de sexto semestre de las licenciaturas de Psicología y Asesoría Psicopedagógica.

En esta primera etapa se verificó la duración de la aplicación que oscilaba entre 40 y 50 minutos. Según Aiken (1996), en este tipo de pruebas que se aplican a nivel medio superior y superior, una regla general es dar un minuto para cada reactivo de opción múltiple. De tal manera que, una prueba con 50 reactivos de opción múltiple por lo regular es apropiada para un periodo de tiempo de 50 minutos en el nivel de educación superior,

esto quiere decir que la evaluación contaba con la extensión y el tiempo necesario según la cantidad de reactivos.

Durante esta primera aplicación se detectó la necesidad de incorporar un apartado en el que se solicitó al estudiante circular en el cuadernillo de preguntas o bien, escribir en la hoja de respuestas, cualquier término o enunciado que no se entendiera al contestar el ítem. La finalidad de esta instrucción era validar que el estudiante comprendiera la estructura de la pregunta en su totalidad, de tal forma que el contexto fuera familiar para poder responder los ítems. Se esperaba encontrar respuestas tales como: degeneración macular o pez lobina, sin embargo los estudiantes registraron palabras tales como: desviación estándar, valor p , intervalo de confianza ecuación de regresión entre otros, por lo que se tomó la decisión de mantener la instrucción para futuras aplicaciones. Los resultados de esta instrucción, sirvieron para orientar lo que los estudiantes desconocen no solo del contexto de los ítems empleados en la evaluación, sino de temas específicos referentes a la estadística.

La primera aplicación de los instrumentos orientó la estructura del cuestionario de opinión que contenía 10 preguntas, entre ellas las de datos generales, y para la aplicación final sólo se consideraron 8 de ellas.

En una segunda etapa del pilotaje se incorporó la carrera de mercadotecnia con una participación de 30 estudiantes de quinto semestre, para ese entonces ya se habían realizado correcciones a la escala de opinión, así mismo, ya se tenía incorporado en la hoja de respuestas de CAOS-4 el apartado donde los participantes registrarían las palabras u oraciones que no eran comprendidas en el momento de contestar el ítem.

Cuando se logró tener el instrumento listo a partir del último pilotaje, se comenzó a realizar las gestiones para la aplicación final, en el siguiente bloque se encuentra la descripción de la muestra que se tomó.

3.6 Muestra

Después de haber definido como objeto de estudio la comprensión estadística en estudiantes universitarios, se definió la muestra a partir de la población seleccionada, que fue la población estudiantil del Centro de Ciencias Sociales y Humanidades, con la finalidad de determinar cuáles eran los logros alcanzados por los estudiantes en materia de estadística, ya que la mayoría de las carreras de este centro involucra al menos un curso de estadística durante su formación universitaria. (ver tabla 6)

Tabla 6
Licenciaturas del Centro de Ciencias Sociales y Humanidades

Licenciatura	Materia	Semestre que imparte	Horas por semana
Asesoría Psicopedagógica	Análisis de información educativa	2	5
Ciencias Políticas y Administración Pública	Estadística I	2	5
	Estadística II	3	5
Comunicación e Información	Estadística	5	5
Licenciatura en Comunicación Organizacional	Estadística I	2	5
	Estadística II	3	5
Enseñanza del Idioma Inglés	*Como parte del perfil se encuentra el manejo de software en estadística.	6 (investigación educativa I)	4
		7 (Investigación educativa II)	6
Historia	Estadística I	4	5
Psicología	Estadística I	6	5
	Estadística II	7	5
Sociología	Estadística I	1	5
	Estadística II	2	6
Trabajo Social	Estadística I	2	5

La selección de la muestra en este estudio fue definida por un muestreo no probabilístico intencional, y precisamente por la falta de aleatoriedad el estudio no se considera válido para generalizar los resultados.

La matrícula del Centro de Ciencias Sociales y Humanidades para el ciclo escolar 2012-2013 fue de 2,491 estudiantes repartidos en las 12 carreras que se ofertan. Se consideraron a los alumnos del quinto semestre en adelante, por ser quienes ya habían cursado la materia de estadística. El total de la matrícula de quinto, séptimo y noveno semestre fue de 957 estudiantes, sin considerar las carreras que fueron excluidas del estudio por no tener en su curricula la materia de estadística, como Derecho, Filosofía y Docencia del Francés y del Español como Lenguas Extranjeras. De los 957 estudiantes

matriculados, sólo se tomaron en cuenta 269 alumnos, en la tabla 7, se indica la distribución de los participantes según la profesión.

Tabla 7
Número de instrumentos aplicados por carrera

<i>Licenciatura</i>	<i>Semestre participante</i>	<i>Aplicados</i>	<i>Inscritos</i>
Comunicación e información	7	20	29
Comunicación organizacional	5	15	31
Enseñanza del idioma inglés	7	35	46
Ciencias Políticas	7	18	24
Historia	7	16	18
Sociología	7	4	6
Asesoría Psicopedagógica	5	37	44
Psicología	9	38	47
Trabajo Social	7	23	24

De las nueve carreras participantes, las licenciaturas de comunicación y Psicología, habían terminado los cursos correspondientes a su formación dos semestres antes del momento de la aplicación de CAOS-4; para el caso de Asesoría Psicopedagógica e Historia había ocurrido tres semestres atrás; en la carrera de Ciencias Políticas los cursos habían terminado cuatro semestres antes; para las carreras de Sociología y Trabajo Social los alumnos finalizaron la asignatura cinco semestres atrás antes de la aplicación.

Licenciatura en Enseñanza del Inglés fue contemplada dentro de la muestra participante, ya que en el perfil de egreso del programa de estudio puntualiza el manejo de software especializados en estadística, logrando este cometido con las asignaturas de Investigación Educativa I y II cursadas en los semestres seis y siete respectivamente, aunque no llevan ningún curso de estadística como tal.

La muestra estuvo delimitada por aquellos estudiantes que en general ya habían concluido los cursos de estadística propuestos en el currículo que formaban parte de los estudios universitarios y que estaban adscritos al Centro de Ciencias Sociales y Humanidades.

3.7 Reporte de trabajo de campo

Todas las aplicaciones de CAOS-4 se realizaron con los estudiantes del Centro de Ciencias Sociales y Humanidades, de la siguiente forma: en primer lugar, se contactó a los jefes de departamento responsables de la coordinación de cada carrera, quienes apoyaron el programa de fechas y horarios de aplicación; a cada uno de los jefes de departamento se solicitó que el horario de la aplicación no fuera la última hora de clase de los estudiantes y que preferentemente, un maestro estuviera en el aula durante todo el

proceso de recolección de datos. Con base en la calendarización obtenida previamente, se acudió a las aulas correspondientes para llevar a cabo el trabajo de campo.

Esta actividad tuvo lugar durante las primeras tres semanas del mes de noviembre de 2012 y previendo que quedara listo el proceso durante estas fechas, ya que en la última semana de este mes los alumnos concluían las clases del semestre agosto – diciembre 2012.

La aplicación tuvo lugar en una sola visita al salón de clases y sólo para el caso de la Licenciatura en Asesoría Psicopedagógica y Trabajo Social se reprogramó una segunda visita con la finalidad de conseguir la participación de la mayor cantidad de alumnos, pues en la primera sesión de trabajo que se realizó, se encontraba menos de la mitad del grupo.

Al inicio de cada aplicación el docente presentaba al aplicador con la finalidad de dar a conocer el motivo de la visita. La secuencia de aplicación que se llevó a cabo fue en el siguiente orden: primero se repartía una hoja de respuestas y el cuestionario de opinión, posteriormente el cuadernillo que incluía 40 reactivos de CAOS-4; una vez que todos los estudiantes tenían el material de trabajo se pedía la atención de los participantes y se solicitaba que se situaran en la primera página del cuadernillo, la cual contenía los objetivos de la aplicación así como el tiempo de inversión requerido para la actividad; se leía textualmente el escrito y acto seguido se explicaba la forma de responder la hoja de respuestas y el cuestionario de opinión, finalmente se solicitaba a los estudiantes que la entrega del cuadernillo fuera de forma individual.

Al finalizar la lectura de las instrucciones del llenado de los instrumentos se registraron una diversidad de comentarios como los siguientes:

- Los estudiantes del grupo de la Lic. en Enseñanza del Inglés, declaraban que en su carrera jamás llevaron materias de estadística y habían elegido la carrera porque no llevaban números y de estadística sabían muy poco.
- Las estudiantes de Licenciatura en Asesoría Psicopedagógica exteriorizaban que ya habían terminado sus materias de estadística y pensaron que jamás la volverían a utilizar, además se lograba apreciar expresiones de desaprobación ante la tarea a ejecutar y comentarios tales como “está muy difícil y no me acuerdo de nada”, ello no significó que al paso de los minutos las estudiantes no se hayan concentrado en la tarea.

- En la Licenciatura de Historia los alumnos hicieron comentarios como los siguientes: ¡No creí que volvería a ver estadística!, ¿La van a quitar del plan de estudios?, incluso hubo estudiantes que expresaron ¡Ya reprobé!;
- Las estudiantes de la Licenciatura de Trabajo Social, cooperaron con la tarea sin problema en una primera aplicación, sin embargo en una segunda aplicación, hubo comentarios del tipo ¡no me acuerdo de esto!, incluso una alumna contestó al azar las preguntas del instrumento terminando en no más de 15 minutos de tal forma que se identificó para posteriormente ser eliminado del resto de las aplicaciones.
- En la Licenciatura de Comunicación Organizacional los alumnos cooperaron con la actividad pero hicieron comentarios del tipo ¡está muy difícil!, ¡hace mucho que vimos estadística y ya no me acuerdo!
- Finalmente los estudiantes de las carreras de Ciencias Políticas, Psicología, Sociología y Comunicación e Información, trabajaron en la tarea de contestar la evaluación sin hacer comentario alguno respecto a la actividad.

Las hojas de respuesta fueron foliadas manualmente por el aplicador al momento de la entrega de las mismas, acto seguido se revisaba que el llenado de la información fuera la adecuada. Los folios estuvieron asignados por el nombre de la carrera y un número consecutivo asignado en el momento en que el estudiante entregaba el material. Este código sirvió para integrar una base de datos utilizando el paquete SPSS y así poder identificar a los estudiantes que contestaron el instrumento, además de que permitiría hacer los análisis estadísticos correspondientes.

La distribución de los estudiantes que contestaron el cuestionario de datos generales y de opinión, así como la prueba CAOS-4, se muestran en la Tabla 7. Cabe mencionar que en el periodo agosto – diciembre 2012 algunos estudiantes de las diferentes carreras seleccionadas para participar en el estudio se encontraban de intercambio estudiantil, tal era el caso de 6 alumnos de la Licenciatura de Asesoría Psicopedagógica y 3 más de la Licenciatura de la Enseñanza del Inglés; algunos otros ya no asistieron a terminar el semestre como es el caso de un estudiante de la carrera de Comunicación e información y 6 más de la Lic. en Enseñanza del Inglés. Bajo este panorama se obtuvieron los datos reportados.

Según los resultados mostrados en la tabla anterior, el total de la muestra era de 269 estudiantes y sólo se aplicaron 206 instrumentos lo que corresponde al 76.5% del total de

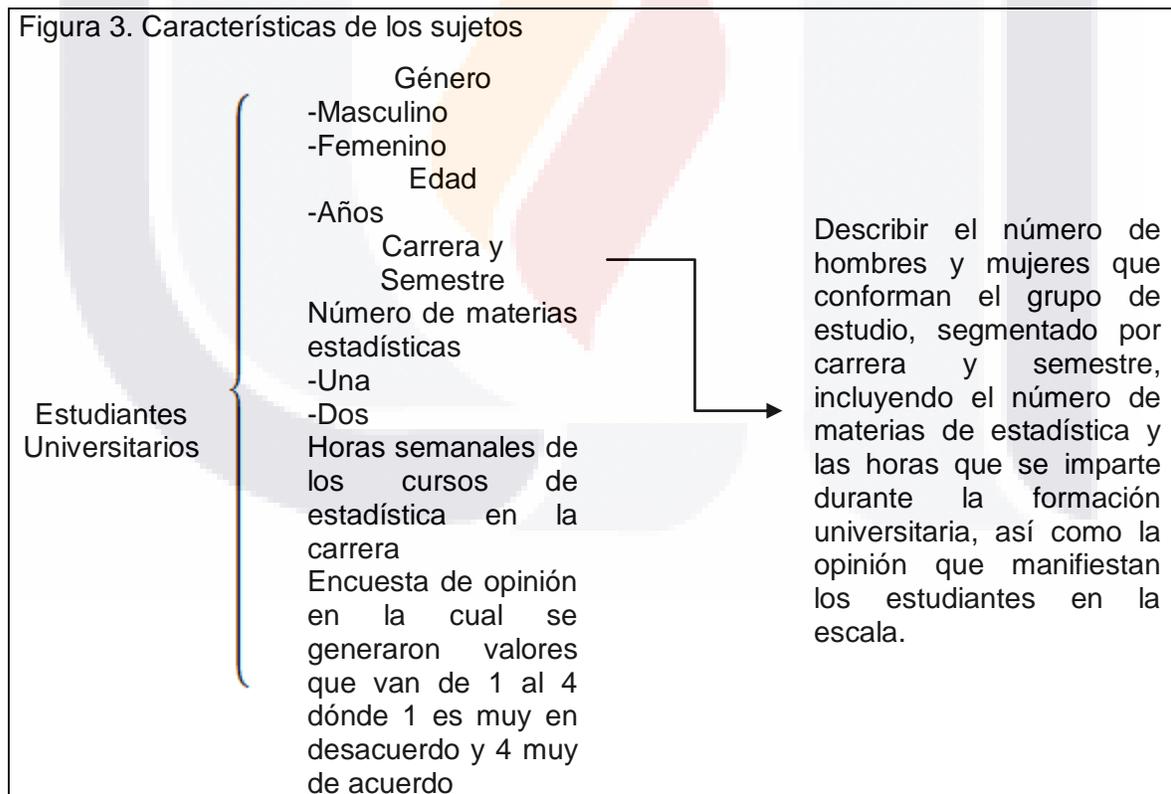
alumnos inscritos y considerados inicialmente como la muestra total contemplada para el estudio.

Con base en estos resultados en el siguiente apartado se plantea el plan de análisis que guiará el reporte del capítulo IV.

3.8 Plan de análisis

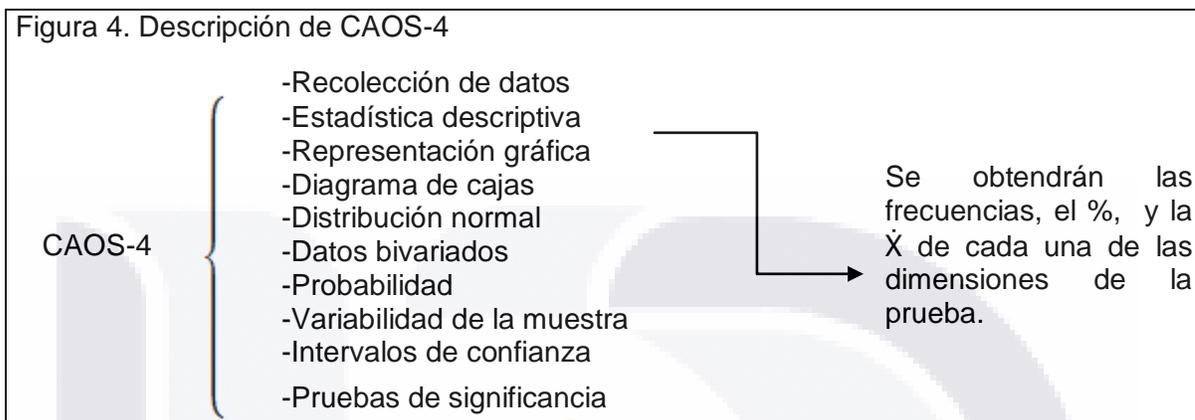
El plan de análisis tiene la finalidad de utilizar los datos obtenidos en el trabajo de campo con el propósito de comprender la relación que se va a establecer entre las variables determinadas previamente. En este proceso se diseñó un modelo de descripción y relación entre las variables que permitió encaminar el análisis de la información así como la parte de la discusión de los resultados.

En la primera etapa del análisis se trató individualmente a las variables independientes, ya que la intención era describir detalladamente a los estudiantes universitarios que formaron la muestra para este estudio. En la figura 3 se detallan las características de los sujetos basadas en las variables descritas previamente.



Posteriormente, se realizó un proceso similar al anterior, pero ahora con los resultados de la prueba CAOS-4. En la figura 4, se observa que se llegó al análisis a

través de la estadística descriptiva, es decir, valores como la media, desviación estándar, máximos y mínimos, y se calculó una medida de resumen en términos de porcentajes de respuesta correctas.



Como parte de un primer análisis se acordó¹⁰ considerar que la comprensión estadística estaría presente si al menos el 70% de los ítems del total de la prueba fueron contestados de manera correcta en la evaluación, lo que corresponde a 28 reactivos de los 40 ítems que componen la prueba.

También se realizó un análisis por reactivos, es decir, se consideró que cada ítem contestado por los estudiantes debería tener al menos 70% de respuestas correctas para determinar que los alumnos conocían acerca del tema que se estaba evaluando.

De esta forma se pretende evaluar la comprensión estadística del estudiante universitario, pero para ello también se tomaron en cuenta las variables independientes que formaron parte del estudio.

Como parte del proceso de validación de la prueba, también se pretendió obtener el índice de dificultad para cada uno de los ítems de CAOS-4 a través de la fórmula 1

$$P_i = \frac{A_i}{N_i}$$

¹⁰ Consultando con el Dr. Marín, se consideró la pertinencia de tener un punto de corte provisional que sirviera de referencia para identificar los niveles de comprensión, fue entonces que se tomó de manera tentativa el 70% por representar 2/3 partes de los contenidos de la prueba.

Donde:

P_i = Al índice de dificultad del reactivo.

A_i = Número de aciertos en el reactivo

N_i = Número de aciertos más el numero en el errores del reactivo.

Así como determinar el poder de discriminación a través de la siguiente fórmula 2:

$$D_i = \frac{GA_{aciertos} - GB_{aciertos}}{N_{grupo\ mayor}}$$

Donde:

D_i = Índice de discriminación del reactivo

$GA_{aciertos}$ = Número de aciertos en el reactivo del 25 % de personas con las puntuaciones más altas en el test.

$GB_{aciertos}$ = Número de aciertos en el reactivo del 25 % de personas con las puntuaciones más bajas en el test.

$N_{grupo\ mayor}$ = Número de personas en el grupo más numeroso (GA o GB)

En resumen, “entre más alto es el índice de discriminación, el reactivo diferenciará mejor a las personas con altas y bajas calificaciones” (Backhoff, Larrazolo y Rosas 2000).

La interpretación de los datos generados a partir de la aplicación de estas fórmulas se integra en las tablas 8 y 9 lo que da cuenta acerca de la calidad de cada uno de los reactivos (Pérez, Acuña, Arratia, 2008).

En la tabla 8 se muestran los valores de P que es el resultado de aplicar la fórmula de nivel de dificultad; de esta forma se determinó una evaluación para los reactivos que va de altamente fácil a altamente difícil.

Tabla 8
Nivel de dificultad y evaluación del reactivo

P =	Evaluación del reactivo
>0.86	Altamente fáciles
0.74 – 0.86	Medianamente fáciles
0.53-0.73	Dificultad media
0.33-0.52	Medianamente difíciles
<-0.32	Altamente difíciles

Nota: Adaptado de J. Pérez, N. Acuña, y E. Arratia, 2008

En la tabla 9 se muestran los valores de D que son el resultado de aplicar la fórmula de poder de discriminación, de esta forma se determinó en la evaluación la calidad de cada uno de los 40 reactivos y su posible recomendación para mejorarlos, además de aplicar la prueba $t - student$ para identificar los reactivos que están discriminando de forma efectiva o deficiente en la evaluación.

Tabla 9
Poder de discriminación de los reactivos según su valor D

D=	Calidad	Recomendaciones
>0.39	Excelente	Conservar
0.30-0.39	Buena	Posibilidad de mejorar
0.20-0.29	Regular	Necesidad de revisar
0.00-0.20	Pobre	Descartar o revisar a profundidad
<-0.01	Pésima	Descartar definitivamente

Nota: Adaptado de J. Pérez, N. Acuña, y E. Arratia, 2008

En la segunda etapa del análisis el interés se centra en establecer las relaciones entre las variables independientes con la variable dependiente a través de la prueba Chi cuadrada según se muestra en la figura 5, de esta forma se pretende dar respuesta a los objetivos de la investigación, cabe aclarar que se descarta la posibilidad de establecer relaciones de causalidad entre las variables relacionadas.

Figura 5. Relación de variables.

Género			
-Masculino	Θ		
-Femenino		Comprensión	estadística
Carrera		determinada	a través de
Semestre	Θ	CAOS-4	
Numero de materias			
Encuesta de opinión	Θ		

Finalmente se utilizó el Alfa de Chronbach con el propósito de validar la consistencia interna de la evaluación adaptada al idioma español.

Con base en este plan de análisis se presentan en el capítulo de resultados los hallazgos del estudio.

CAPÍTULO IV

RESULTADOS

En este capítulo se pretende dar respuesta principalmente a los objetivos de investigación y a las preguntas que guiaron este estudio, es decir, ¿qué grado de comprensión se manifiesta en los estudiantes de educación superior que han realizado cursos de estadística? ¿Existen diferencias de comprensión estadística de los estudiantes de educación superior dependiendo de la carrera y el semestre que cursan?

Los resultados se presentan en tres apartados denominados: descripción de los estudiantes; confiabilidad y validez de los instrumentos; y comprensión estadística de los universitarios a través de CAOS-4.

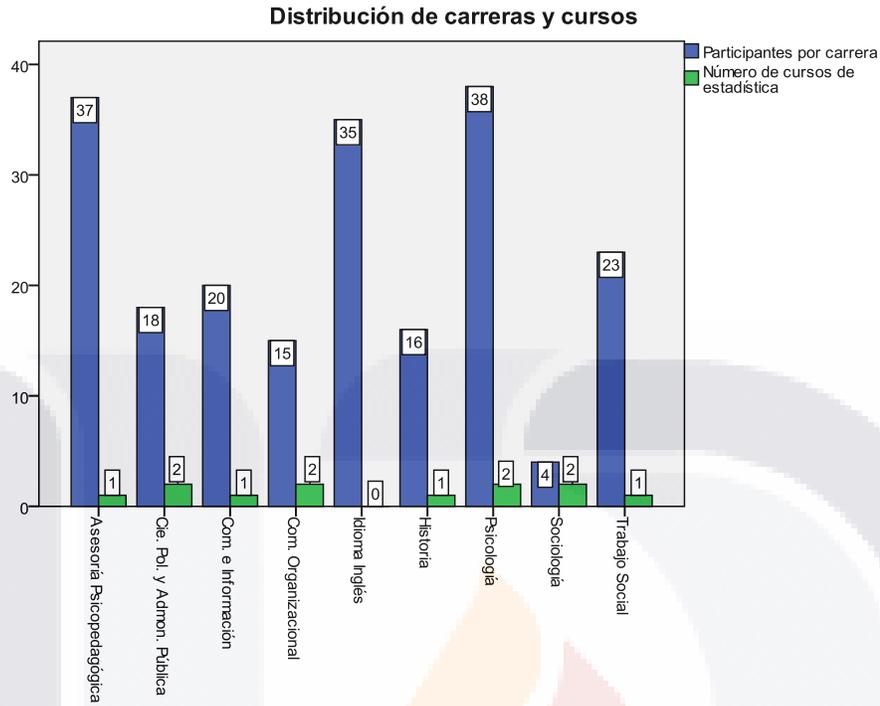
En la primera etapa se describe a los estudiantes de acuerdo a las variables independientes obtenidas como parte de la recolección de la información, con la finalidad de conocer a detalle las características de la muestra.

En el apartado de la confiabilidad y validez se abordan los procesos para determinar los niveles de dificultad de cada reactivo y de la prueba en su conjunto, así como sus niveles de confiabilidad según la prueba Alfa de Cronbach.

Finalmente, en la parte de la comprensión estadística se describen los logros obtenidos por los estudiantes en la aplicación de CAOS-4, así como los resultados de la prueba X^2 al utilizarla con las variables independientes consideradas en la investigación.

4.1 Descripción de los estudiantes

De los 206 estudiantes que participaron en el estudio, el 28.6% eran hombres y 71.4% eran mujeres, todos ellos con una edad que va desde los 19 hasta los 40 años con una media de edad de 22 años. En la gráfica 1 se observa la participación de 37 estudiantes del cuarto semestre de la licenciatura en asesoría psicopedagógica, 38 participantes del noveno semestre de psicología, y 131 estudiantes que pertenecen al séptimo semestre del resto de las carreras. Del total de la muestra, sólo 35 (17%) estudiantes no habían tomado ningún curso de estadística, 96 (46.1%) de ellos sólo tomaron un curso y 75 (36.4%) de los alumnos tomaron dos cursos de estadística como parte de su formación profesional.



Gráfica 1 Distribución de carreras y cursos.

Siguiendo con la descripción de los resultados, en la tabla 10 se presentan las frecuencias y porcentajes de cada una de las respuestas que dieron los estudiantes con respecto al cuestionario de opinión aplicado como parte de este trabajo de investigación.

Tabla 10
 Frecuencia de respuestas del cuestionario de opinión

N = 206	Muy en desacuerdo	En desacuerdo	De acuerdo	Muy de acuerdo	No contestó
Los cursos de estadística que he recibido en la universidad me han gustado.	32	64	84	18	8
	15.50%	31.10%	40.80%	8.70%	3.90%
Me interesa aprender estadística	16	77	94	17	2
	7.80%	37.40%	45.60%	8.30%	1.00%
Los cursos ayudan a mi formación	7	24	134	38	3
	3.40%	11.70%	65.00%	18.40%	1.50%
La estadística es difícil de aprender	17	89	64	31	5
	8.30%	43.20%	31.10%	15%	2.40%
Se mucho de estadística	70	118	16	0	2
	34.00%	57.30%	7.80%	0%	1.00%
Recibir más cursos de estadística	46	70	77	11	2
	22.00%	34.00%	37.40%	5.50%	1.00%
Suficientes conocimientos para ejercer mi profesión	23	91	85	4	3
	11.20%	44.20%	41.30%	1.90%	1.50%
Los cursos de estadística estaban enfocados a la carrera	62	65	55	16	8
	30.10%	31.60%	26.70%	7.80%	3.90%

Como puede apreciarse en la Tabla 10, del total de la población encuestada, al 40% de la muestra les agradó en forma moderada los cursos de estadística que recibieron durante su formación profesional, mientras que el 31.1% estuvo en desacuerdo en que los cursos de estadística les gustaron..

Según la opinión de los estudiantes el 57.3% de ellos consideró que conoce limitadamente de estadística y ninguno consideró saber mucho de este tema. En esta misma línea el 44.2% de la muestra opinó que tampoco tiene suficientes conocimientos estadísticos para ejercer su profesión, sólo el 1.9% considera tener los suficientes conocimientos estadísticos para ejercer la profesión

Por otro lado, solamente al 45.6% de la población le interesaría aprender más acerca de la estadística y el 37% de ellos estuvieron en desacuerdo con aprender más acerca del tema; el 43.2% opina que la estadística no es tan difícil de aprender y sólo el 15% piensa que sí lo es; el 37.4% están de acuerdo con que deben tomar cursos adicionales de estadística; el 31.6% de los estudiantes consideran que sus cursos no estaban totalmente enfocados directamente a su que hacer profesional y el 30% está en total desacuerdo que sus cursos estadísticos se ofrecían con un enfoque dirigido a su profesión.

4.2 Confiabilidad y validez

Determinar la confiabilidad de la prueba CAOS-4 en el contexto mexicano y específicamente a partir de su aplicación entre estudiantes del Centro de Ciencias Sociales y Humanidades de la Universidad Autónoma de Aguascalientes es fundamental para dar respuesta a los objetivos de investigación planteados.

Según Aiken (1996) antes de utilizar una prueba con cierta seguridad como instrumento de medición preciso, se debe obtener la suficiente información con respecto de su confiabilidad y validez (p. 77).

Por confiabilidad se entiende que el instrumento mide una característica relativamente estable, es decir, se refiere al atributo de consistencia (Cohen y Swerdlik, 2002). Para el caso de la comprensión estadística, en este estudio se utilizó el alfa de Cronbach. Este cálculo se realizó con la elaboración de una base de datos capturada en el programa SPSS, que contenía los resultados de la prueba CAOS-4; las opciones de respuesta de la prueba eran de opción múltiple las cuales fueron capturadas con valores de 1 a 4 y posteriormente recodificadas en respuestas dicotómicas siendo 1 la respuesta correcta y 0 la respuesta incorrecta.

Los valores recodificados sirvieron para obtener la confiabilidad de la prueba, teniendo un valor de $\alpha = 0.13$, con un total de 206 sujetos evaluados. Con este resultado se observó que el nivel de confiabilidad obtenido es muy bajo, pues de acuerdo con Aiken (1996), los coeficientes de confiabilidad para una prueba de aprovechamiento pueden ser bastante modestos pero no inferiores a $\alpha = 0.56$.

Los factores que posiblemente influyeron en el proceso de confiabilidad del instrumento se refieren al tamaño de la muestra ya que fue muy pequeña y cabe aclarar que existe más variabilidad en muestras pequeñas que en las grandes, también es importante considerar la escasa correspondencia entre lo que mide la prueba y lo que realmente conocen los estudiantes, otro factor que influyó en el nivel de confiabilidad de la prueba fue la dificultad de los ítems y el poder de discriminación, que se reportan en la tabla 11 y 13 respectivamente, y que fueron interpretados con base en la categorización realizada por Pérez, Acuña, Arratia en el 2008, con la finalidad de analizar aquellos ítems que realmente aportaron información relevante en la prueba CAOS-4.

Para calcular el índice de dificultad de los reactivos, en primer lugar se elaboró una tabla de frecuencias y posteriormente se aplicó la fórmula 1, que consiste en dividir el

número de aciertos entre el número de aciertos más el número de errores. Los resultados de cada ítem se reportan en la tabla 11.

Tabla 11
Índice de dificultad de la prueba

Categoría de dificultad	N/ 206	Incorrecto	Correcto	Índice de dificultad
Dificultad media	Ítem 1	96 46.6%	110 53.4%	.53
	Ítem 8	92 44.7%	114 55.3%	.55
	Ítem 11	64 31.1%	142 68.9	.69
	Ítem 12	65 31.6%	141 68.4%	.68
	Ítem 18	83 40.3	123 59.7%	.60
	Ítem 20	63 30.6%	143 69.4%	.69
	Ítem 23	94 45.6%	112 54.4%	.54
	Ítem 31	92 44.7%	114 55.3%	.55
	Ítem 34	93 45.1%	113 54.9%	.55

Tabla 11 (Continuación)

Categoría de dificultad	N/ 206	Incorrecto	Correcto	Índice de dificultad
Medianamente difícil	Ítem 2	116 56.3%	90 43.7%	.44
	Ítem 13	116 56.3%	90 43.7%	.44
	Ítem 15	127 61.7%	79 38.3%	.38
	Ítem 21	111 53.9%	95 46.1%	.46
	Ítem 22	99 48.1%	107 51.9%	.52
	Ítem 24	122 59.2%	84 40.8%	.41
	Ítem 25	107 51.9%	99 48.1%	.48
	Ítem 26	125 60.7%	81 39.3%	.39
	Ítem 27	119 57.8%	87 42.2%	.42
	Ítem 28	126 61.2%	80 38.8%	.39
	Ítem 29	132 64.1%	74 35.9%	.36
	Ítem 30	138 67.0%	68 33.0%	.33
	Ítem 32	125 60.7%	81 39.3%	.39
	Ítem 33	143 69.4%	63 30.6%	.31
	Ítem 35	121 58.7%	85 41.3%	.41
	Ítem 36	133 64.6%	73 35.4%	.35

Tabla 11 (Continuación)

Categoría de dificultad	N/ 206	Incorrecto	Correcto	Índice de dificultad
Altamente difícil	Ítem 3	151 73.3%	55 26.7%	.27
	Ítem 4	143 69.4%	63 30.6%	.31
	Ítem 5	178 86.4%	28 13.6%	.14
	Ítem 6	191 92.7%	15 7.3%	.07
	Ítem 7	170 82.5%	36 17.5%	.17
	Ítem 9	170 82.5%	36 17.5%	.17
	Ítem 10	174 84.5%	32 15.5%	.16
	Ítem 14	178 86.4%	28 13.6%	.14
	Ítem 16	194 94.2%	12 5.8%	.06
	Ítem 17	149 72.3%	57 27.7%	.28
	Ítem 19	140 68.0%	66 32.0%	.32
	Ítem 37	162 78.6%	44 21.4%	.21
	Ítem 38	166 80.6%	40 19.4%	.19
	Ítem 39	141 68.4%	65 31.6%	.32
	Ítem 40	149 72.3%	57 27.7%	.28

Se observa en esta tabla que el índice de dificultad promedio que se obtiene de la suma de todos los índices de dificultad de los ítems entre la cantidad de ítems es de 0.37, es decir, según la interpretación de la tabla 8, los reactivos caen en la categoría de medianamente difícil para los 206 estudiantes que participaron en la aplicación de la evaluación CAOS-4.

Siguiendo con el análisis del índice de dificultad, en la tabla 12 se muestra la distribución que tuvieron los ítems en la prueba, de tal forma que se observa la distribución de estos según la dimensión a la que corresponde cada uno de ellos.

Tabla 12
 Dificultad de los ítems según las dimensiones de la prueba

Dimensiones de la prueba	N/206	Dificultad de los ítems
Recolección de datos	Ítem 32	Medianamente difícil
	Ítem 34	Dificultad media
	Ítem 35	Medianamente difícil
	Ítem 38	Altamente difícil
Estadística descriptiva	Ítem 12	Dificultad media
	Ítem 14	Altamente difícil
	Ítem 15	Medianamente difícil
Representación gráfica	Ítem 1	Dificultad media
	Ítem 3	Altamente difícil
	Ítem 4	Altamente difícil
	Ítem 5	Altamente difícil
Diagrama de cajas	Ítem 2	Medianamente difícil
	Ítem 8	Dificultad media
	Ítem 9	Altamente difícil
	Ítem 10	Altamente difícil
Distribución normal	Ítem 6	Altamente difícil
	Ítem 11	Dificultad media
	Ítem 33	Medianamente difícil
Datos bivariados	Ítem 20	Dificultad media
	Ítem 21	Medianamente difícil
	Ítem 22	Medianamente difícil
Probabilidad	Ítem 7	Altamente difícil
	Ítem 24	Medianamente difícil
	Ítem 36	Medianamente difícil
	Ítem 37	Altamente difícil
Variabilidad de la muestra	Ítem 13	Medianamente difícil
	Ítem 16	Altamente difícil
	Ítem 17	Altamente difícil
	Ítem 18	Dificultad media
Intervalos de confianza	Ítem 28	Medianamente difíciles
	Ítem 29	Medianamente difíciles
	Ítem 30	Medianamente difíciles
	Ítem 31	Dificultad media
Prueba de significancia	Ítem 19	Altamente difícil
	Ítem 23	Dificultad media
	Ítem 25	Medianamente difícil
	Ítem 26	Medianamente difícil
	Ítem 27	Medianamente difícil
	Ítem 30	Altamente difícil
	Ítem 40	Altamente difícil

Como se puede observar en la tabla, en su mayoría los reactivos se clasifican en medianamente difíciles, solo unos pocos se encuentran en la categoría de dificultad media y ningún reactivo se clasificó dentro de la categoría de fácil.

Con la finalidad de determinar la discriminación de los ítems, se ejecutó la prueba *t* de student, cuyos resultados indicaron que los reactivos 1, 4, 6, 7, 9, 10, 15, 16, 20, 23, 26, 28, 30, 31 y 40 no discriminaron adecuadamente, con un valor $p >.05$. En este mismo sentido, como una estrategia alternativa, para obtener el poder de discriminación se utilizó la base de datos capturada en el programa SPSS con el fin de obtener el cuartil más bajo y el cuartil más alto de la base de datos, es decir, se calcula la diferencia entre el 25 por ciento de los estudiantes que respondieron correctamente el reactivo y tuvieron las más altas calificaciones, de esta forma se seleccionaron los casos con las puntuaciones más altas; también se tomo en cuenta al 25 por ciento de los estudiantes que respondieron correctamente el reactivo y tuvieron las más bajas calificaciones, los casos con las puntuaciones más bajas, por lo tanto, se realizó el análisis con 100 casos del total de la población participante. Los datos se recodificaron en una variable distinta y se generó un nuevo índice con el que se aplicó la fórmula para obtener el poder de discriminación y el cual se reporta en la tabla 13.

Tabla 13
Poder de discriminación

	Poder de discriminación	Calidad	Recomendación
Ítem 1	0.07	Pésima	Descartar definitivamente
Ítem 2	0.17	Pobre	Revisar a profundidad
Ítem 3	0.16	Pobre	Revisar a profundidad
Ítem 4	0.10	Pobre	Revisar a profundidad
Ítem 5	0.20	Regular	Necesidad de revisar
Ítem 6	-0.01	Pésima	Descartar definitivamente
Ítem 7	0.07	Pésima	Descartar definitivamente
Ítem 8	0.19	Pobre	Revisar a profundidad
Ítem 9	0.07	Pésima	Descartar definitivamente
Ítem 10	0.04	Pésima	Descartar definitivamente
Ítem 11	0.29	Regular	Necesidad de revisar
Ítem 12	0.14	Pobre	Revisar a profundidad
Ítem 13	0.29	Regular	Necesidad de revisar
Ítem 14	0.16	Pobre	Revisar a profundidad
Ítem 15	0.01	Pésima	Descartar definitivamente
Ítem 16	0.05	Pésima	Descartar definitivamente
Ítem 17	0.22	Regular	Necesidad de revisar
Ítem 18	0.17	Pobre	Revisar a profundidad
Ítem 19	0.13	Pobre	Revisar a profundidad
Ítem 20	0.06	Pésima	Descartar definitivamente

Tabla 13 (continuación)

	Poder de discriminación	Calidad	Recomendación
Ítem 21	0.30	Buena	Posibilidad de mejorar
Ítem 22	0.20	Pobre	Revisar a profundidad
Ítem 23	0.05	Pésima	Descartar definitivamente
Ítem 24	0.16	Pobre	Revisar a profundidad
Ítem 25	0.14	Pobre	Revisar a profundidad
Ítem 26	0.07	Pésima	Descartar definitivamente
Ítem 27	0.20	Pobre	Revisar a profundidad
Ítem 28	-0.08	Pésima	Descartar definitivamente
Ítem 29	0.13	Pobre	Revisar a profundidad
Ítem 30	0.04	Pésima	Descartar definitivamente
Ítem 31	0.07	Pésima	Descartar definitivamente
Ítem 32	0.14	Pobre	Revisar a profundidad
Ítem 33	0.11	Pobre	Revisar a profundidad
Ítem 34	0.17	Pobre	Revisar a profundidad
Ítem 35	0.17	Pobre	Revisar a profundidad
Ítem 36	0.22	Regular	Necesidad de revisar
Ítem 37	0.19	Pobre	Revisar a profundidad
Ítem 38	0.11	Pobre	Revisar a profundidad
Ítem 39	0.13	Pobre	Revisar a profundidad
Ítem 40	0.10	Pobre	Revisar a profundidad

Esta tabla indica que, en principio, 13 de los reactivos que forman parte de CAOS-4 deben ser descartados definitivamente de la prueba porque no discriminan entre aquellos estudiantes que puntúan alto en la evaluación y los que puntúan bajo, los otros 27 ítems deben ser revisados para mejorarlos y se pueda lograr un poder de discriminación de los reactivos óptimo, con la finalidad de que se conserven en la evaluación. Sin embargo, convendría probar el CAOS-4 con una muestra más grande y más heterogénea, que tome incluso estudiantes de otras Instituciones de Educación Superior, pues al parecer, los contenidos estadísticos contemplados en los cursos de la UAA no son del todo afines a los considerados por el CAOS-4.

Finalmente, en la tabla 14, se muestra un resumen general del comportamiento de los ítems según las dimensiones que contiene la prueba aplicada.

Tabla 14

Promedio de índice de dificultad y poder de discriminación de la prueba

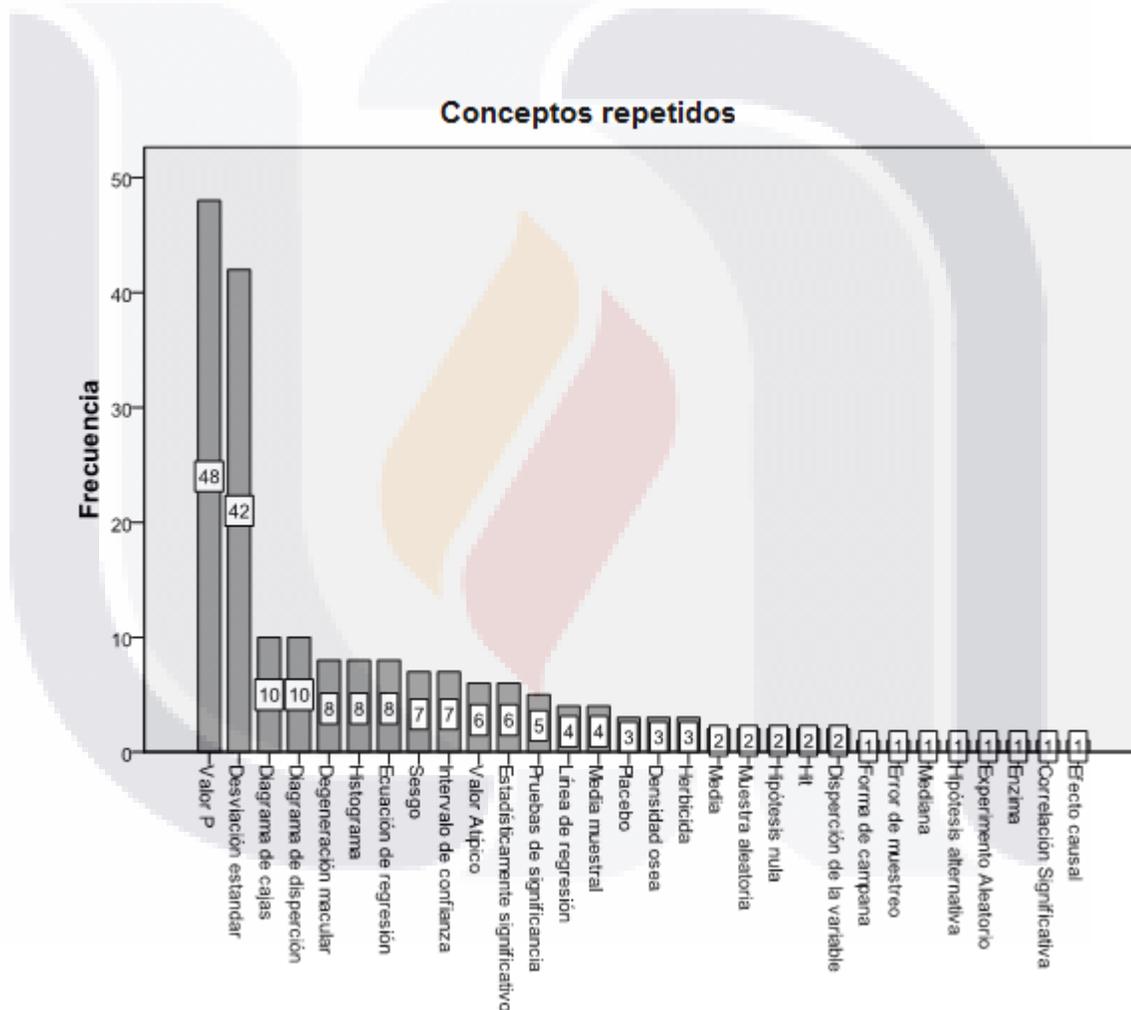
Dimensión	Poder de discriminación	Índice de dificultad
Recolección de datos	0.14 Pobre, revisar a profundidad	0.38 (medianamente difícil)
Estadística descriptiva	0.10 Pobre, revisar a profundidad	0.4 (mediana mente difícil)
Representación gráfica	0.13 Pobre, revisar a profundidad	0.31 (altamente difícil)
Diagrama de cajas	0.11 Pobre, revisar a profundidad	0.33 (altamente difícil)
Distribución normal	0.13 Pobre, revisar a profundidad	0.35 (medianamente difícil)
Datos bivariados	0.18 Pobre, revisar a profundidad	0.55 (dificultad media)
Probabilidad	0.16 Pobre, revisar a profundidad	0.28 (altamente difícil)
Variabilidad de la muestra	0.18 Pobre, revisar a profundidad	0.34 (medianamente difícil)
Intervalos de Confianza	0.04 Pésima, descartar	0.40 (medianamente difícil)
Pruebas de significancia	0.11 Pobre, revisar a profundidad	0.39 (medianamente difícil)
Total 40 ítems	0.13 Pobre, revisar a profundidad	0.37 medianamente difícil

Los resultados indican que la prueba, se clasifica en la categoría de medianamente difícil y se tienen que revisar la mayoría de los reactivos con la finalidad de mejorar la calidad del instrumento y poderlo utilizar en posteriores investigaciones.

4.3 La comprensión estadística de los estudiantes

Para analizar los conceptos estadísticos que los estudiantes universitarios comprenden, se llevaron a cabo tres tipos de análisis que se describen en los siguientes párrafos.

Un dato importante que se obtuvo en esta investigación, se refirió a la información que ofrecieron los estudiantes al escribir o circular las palabras que no fueron comprendidas al contestar los ítems de la prueba, lo que dio por resultado la gráfica 2, en donde se representa la frecuencia de dichas palabras.



Gráfica 2 Conceptos estadísticos no comprendidos.

La última instrucción añadida al instrumento CAOS-4 ofrece información importante ya que el estudiante determinó libremente aquellos conceptos o palabras que el mismo consideró que entendió de forma deficiente para poder responder los ítems

correctamente, de tal manera que las palabras más repetidas para esta instrucción fueron en primer lugar Valor p y en segundo lugar, desviación estándar, con un porcentaje de 23.3% (48) y 20.4% (42) respectivamente. Las variables con frecuencia de repetición uno fueron ocho, y corresponden a efecto causal, correlación significativa, enzima, experimento aleatorio, hipótesis alternativa, mediana, error de muestreo y forma de campana. Estos datos dan a conocer aquellos términos estadísticos que el estudiante debe conocer y por consiguiente estar familiarizado para dar respuesta a planteamientos estadísticos.

Siguiendo con un análisis descriptivo, se retomaron las siguientes medidas de tendencia central y de dispersión generadas con la prueba CAOS-4: una media de 34.8% (13.92) una mediana de 35% (14), la desviación estándar alcanzada fue de 8% (3.20), un valor mínimo de 12.5% (5) con un valor máximo de 57.5% (23) reactivos, lo cual dista mucho de acercarse al reporte de aplicación de CAOS-4 presentado en 2006 por Garfield, delMas, Chance, Poly y Ooms, donde los datos que se reportaron en términos de porcentajes fueron los siguientes: una media de 55.77%, una mediana de 53.75%, la desviación estándar de 16.13%, con un valor mínimo de 18% y un máximo de 100%.

El primer análisis planteado como ya se había mencionado se determinó por el puntaje total del número de respuestas correctas, es decir, la comprensión estadística estaría presente si al menos el 70% de los ítems fueron contestados de manera correcta en la evaluación, lo que corresponde a 28 reactivos de los 40 ítems que componen la prueba.

En la tabla 15 se ofrece información del número mínimo y máximo de aciertos obtenidos en la prueba, la frecuencia con la que estos se presentaron y el porcentaje que le corresponde a cada fila respectivamente.

Tabla 15
Puntaje total de respuestas correctas

Aciertos	Frecuencia	Porcentaje
5	1	0.5%
7	4	1.9%
8	4	1.9%
9	5	2.4%
10	16	7.8%
11	12	5.8%
12	26	12.6%
13	29	14.1%
14	23	11.2%
15	25	12.1%
16	18	8.7%
17	18	8.7%
18	6	2.9%
19	12	5.8%
20	1	0.5%
21	3	1.5%
22	1	0.5%
23	2	1.0%
Total	206	100%

Bajo esta forma de análisis se detecta que ningún estudiante logró un puntaje de 28 reactivos y tan sólo uno de cada 100 estudiantes logran obtener 23 reactivos correctos, es decir, apenas el 57.5% del total de la prueba y tan sólo un 0.5% de los estudiantes obtuvo 22 reactivos correctos que corresponde al 55% del total de la evaluación CAOS-4; finalmente cabe destacar que 13 es el promedio de respuestas correctas en toda la evaluación.

En segundo lugar, para dar respuesta al objetivo de investigación planteado, se realizó un análisis por reactivos, es decir, si el 70% de las respuestas fueron contestadas correctamente, se acordó considerar que los estudiantes conocían acerca del tema y para ello se generó la tabla 16, en la que se observa la frecuencia de cada uno de los reactivos contestados correctamente, los incorrectos, e incluso también los que no fueron contestados, así como los porcentajes de los mismos.

Tabla 16
Frecuencias por dimensión de CAOS-4

Dimensión		Correcto	Incorrecto	No contestó
Recolección de datos	Pregunta 32	75	125	6
		36.40%	60.70%	2.90%
	Pregunta 34	106	93	7
		51.50%	45.10%	3.40%
	Pregunta 36	63	133	10
30.60%		64.60%	4.90%	
Pregunta 38	40	166	0	
		19.40%	80.60%	0%
Estadística descriptiva	Pregunta 12	140	65	1
		68.00%	31.60%	0.50%
	Pregunta 14	28	170	8
		13.60%	82.50%	3.90%
Pregunta 15	71	127	8	
	34.50%	61.70%	3.90%	
Representación Gráfica	Pregunta 1	110	96	0
		53.40%	46.60%	0%
	Pregunta 3	50	151	5
		24.30%	73.30%	2.40%
	Pregunta 4	59	143	4
28.60%		69.40%	1.90%	
Pregunta 5	28	170	8	
		13.60%	82.50%	3.90%
Diagrama de cajas	Pregunta 2	88	116	2
		42.70%	56.30%	1.00%
	Pregunta 8	75	92	39
		36.40%	44.70%	18.90%
	Pregunta 9	36	166	4
17.50%		80.60%	1.90%	
Pregunta 10	32	172	2	
		15.50%	83.50%	1.00%

Tabla 16 (continuación)

Dimensión		Correcto	Incorrecto	No contestó
Distribución normal	Pregunta 6	15	178	13
		7.30%	86.40%	6.30%
	Pregunta 11	141	64	1
		68.40%	31.10%	0.50%
	Pregunta 33	54	143	9
		26.20%	69.40%	4.40%
Datos bivariados	Pregunta 20	141	63	2
		68.40%	30.60%	1.00%
	Pregunta 21	90	111	5
		43.70%	53.90%	2.40%
	Pregunta 22	103	99	4
		50.00%	48.10%	1.90%
Probabilidad	Pregunta 7	36	166	4
		17.50%	80.60%	1.90%
	Pregunta 24	78	122	6
		37.90%	59.20%	2.90%
	Pregunta 36	63	133	10
		30.60%	64.60%	4.90%
Pregunta 37	42	162	2	
	20.40%	78.60%	1.00%	
Varibilidad de la muestra	Pregunta 13	87	116	3
		42.20%	56.30%	1.50%
	Pregunta 16	12	193	1
		5.80%	93.70%	0.50%
	Pregunta 17	53	149	4
		25.70%	72.30%	1.90%
Pregunta 18	120	83	3	
	58.30%	40.30%	1.50%	
Intervalos de confianza	Pregunta 28	77	126	3
		37.40%	61.20%	1.50%
	Pregunta 29	74	132	0
		35.90%	64.10%	0%
	Pregunta 30	68	138	0
		33.00%	67.00%	0%
Pregunta 31	113	92	1	
	54.90%	44.70%	0.50%	

Tabla 16 (continuación)

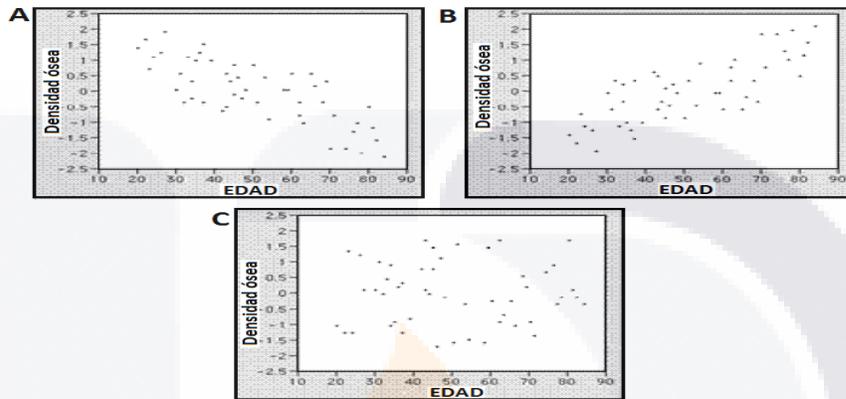
Dimensión		Correcto	Incorrecto	No contestó
Pruebas de significancia	Pregunta 19	49	140	17
		23.8%	68.0%	8.3%
	Pregunta 23	107	94	5
		51.9%	45.6%	2.4%
	Pregunta 25	86	107	13
		41.7%	51.9%	6.3%
	Pregunta 26	68	125	13
		33.0%	60.7%	6.3%
	Pregunta 27	75	119	12
		36.4%	57.8%	5.8%
Pregunta 39	56	141	9	
	27.2%	68.4%	4.4%	
Pregunta 40	50	149	7	
	24.3%	72.3%	3.4%	

Bajo esta segunda forma de análisis se observa que tan sólo los reactivos 11, 12 y 20, lograron obtener un 68% de respuestas correctas, es decir, de los 206 estudiantes a los que se les aplicó la prueba, 141 contestaron correctamente esos reactivos. Los contenidos¹¹ evaluados en cada uno de estos ítems se refieren a: la capacidad para comparar dos grupos considerando donde está la mayoría de los datos y centrándose en la distribución como entidades individuales; la capacidad para comparar grupos mediante la comparación de las diferencias en promedios; y finalmente, la capacidad para coincidir con un gráfico de dispersión a una descripción verbal de una relación bivalente.

¹¹ Revisar anexo B referente a CAOS-4 test y anexo D ítems de CAOS-4

Ejemplo de reactivo:

20. La densidad ósea se mide típicamente como un puntaje estandarizado con una media de 0 y una desviación estándar de 1. Las puntuaciones más bajas corresponden a una densidad ósea menor. ¿Cuál de las siguientes gráficas muestran que a medida que envejecen las mujeres tienden a tener menor densidad ósea?



- a. Gráfica A
- b. Gráfica B
- c. Gráfica C

Luego le sigue el reactivo 18, que evalúa la comprensión del significado de variabilidad en el contexto de medidas repetidas y en un contexto donde se desea la pequeña variabilidad. y que fue contestado por 120 estudiantes, lo que corresponde a tan sólo el 58.3%. Bajo esta forma de análisis se observa que ningún reactivo alcanzó un porcentaje del 70%.

Ejemplo de reactivo:

1. Juan vive a unos 10 kilómetros de la universidad donde planea asistir por 10 semanas a un curso de verano. Hay dos rutas principales que pueden llevarlo a la escuela, una a través de la ciudad y una por el pueblo. La ruta de la ciudad es más corta en kilómetros, pero tiene más semáforos. La ruta del campo, es más larga en kilómetros y tiene pocos semáforos. Juan establece un experimento aleatorio en el que cada día lanza una moneda para decidir qué ruta tomar ese día. Se registran los siguientes datos para 5 días de viaje en cada ruta.

Ruta del Pueblo – 17min, 15min, 17min, 16min, 18min

Ruta de la ciudad – 18min, 13min, 20min, 10min, 16min

TESIS TESIS TESIS TESIS TESIS

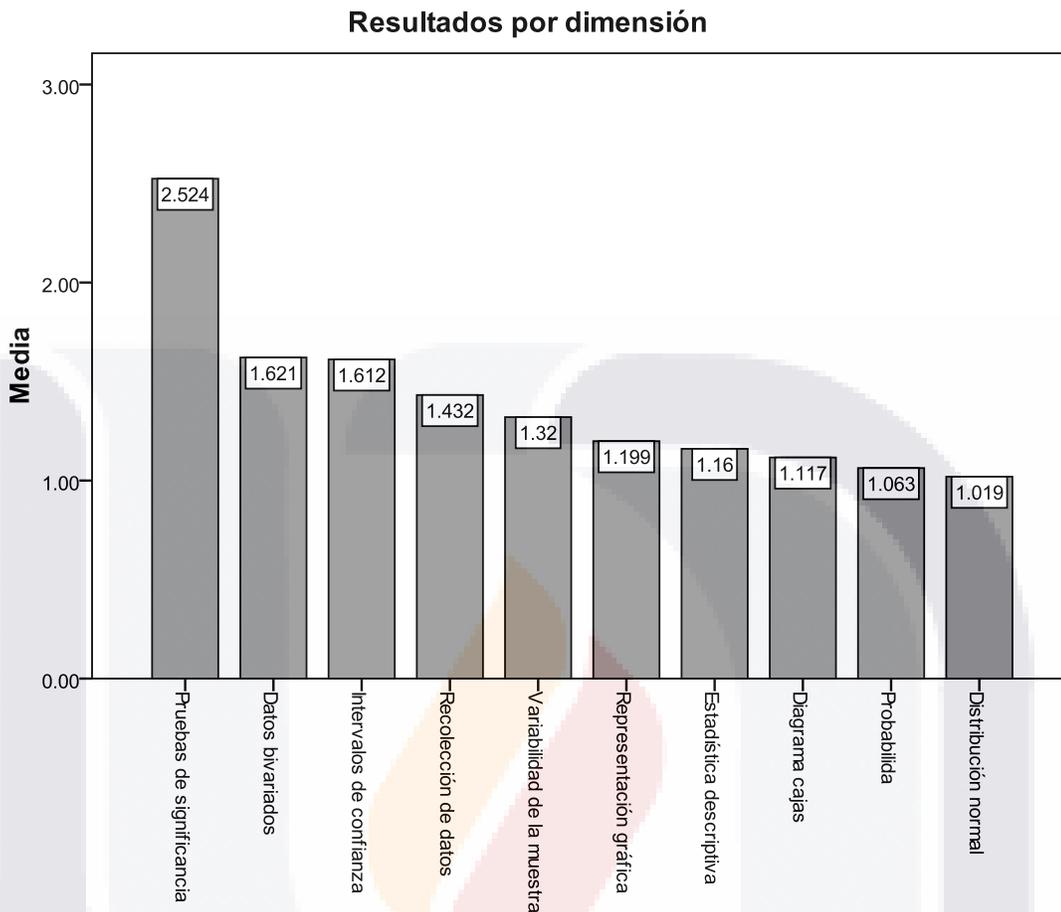
Es importante para Juan llegar a tiempo a sus clases, pero no quiere llegar demasiado temprano porque incrementaría la tarifa del estacionamiento. Basados en los datos registrados, ¿Cuál ruta podría aconsejarle que tome?

- a. La ruta del pueblo, porque los tiempos son consistentes entre 15 y 18 minutos.
- b. La ruta de la ciudad, porque puede llegar en 10 minutos en un buen día y el promedio de tiempo es menor que la ruta del pueblo.
- c. Debido a que el tiempo en las dos rutas tienen mucha coincidencia, ninguna ruta es mejor que la otra. Podría seguir lanzando la moneda.

Una tercera forma de análisis que se propone es tomar en cuenta las dimensiones de CAOS-4, la cual se reporta en la gráfica 3.

En primer lugar cabe mencionar que 9 de las diez dimensiones propuestas para la prueba están compuestas de tres a cuatro reactivos cada una; tan solo la dimensión de pruebas de significancia está compuesta por 7 ítems. De esta forma se observa en la gráfica que la media de respuesta obtenida por la muestra es de una pregunta correcta por dimensión, de tal forma que, en una aplicación similar a la reportada en este estudio con sujetos que tengan características similares a los aquí descritos, con base en los resultados obtenidos y con una muestra representativa, se puede hacer una afirmación predictiva, de tal forma que se espera que en una dimensión que consta de tres ítems sean contestadas correctamente el 33.3% de las preguntas lo que corresponde a una pregunta por dimensión. En el caso de dimensiones que son evaluadas con cuatro reactivos se espera que sean contestadas correctamente tan solo el 25% de estas, es decir, se espera que se responda tan solo una pregunta del total de la dimensión. El caso del grupo de preguntas que componen la dimensión pruebas de significancia se esperaría que 3 de los 7 reactivos sean contestados correctamente, es decir, el 42.8% del total de la dimensión. En resumen, ninguna dimensión¹² logró ser contestada con más del 50% de los ítems de forma correcta en el contexto de las carreras del área de las ciencias sociales en la Universidad Autónoma de Aguascalientes.

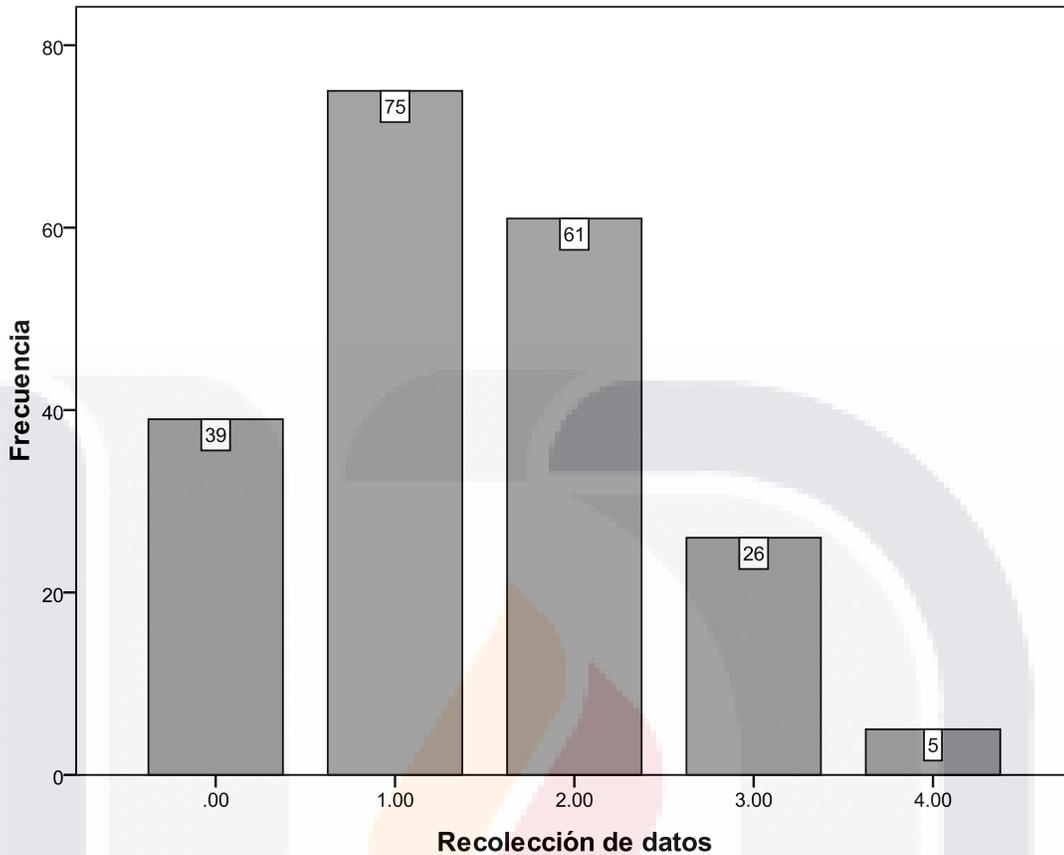
¹² Revisar el Anexo E para una descripción detallada de las respuestas otorgadas por los estudiantes al contestar la prueba CAOS-4



Gráfica 3 Resultados por dimensión de CAOS-4

Siguiendo esta forma de análisis a continuación se muestran las gráficas a detalle de las diez dimensiones. Para estos reactivos se utilizaron valores entre cero y uno; donde cero fue codificado como respuesta incorrecta y uno como respuesta correcta.

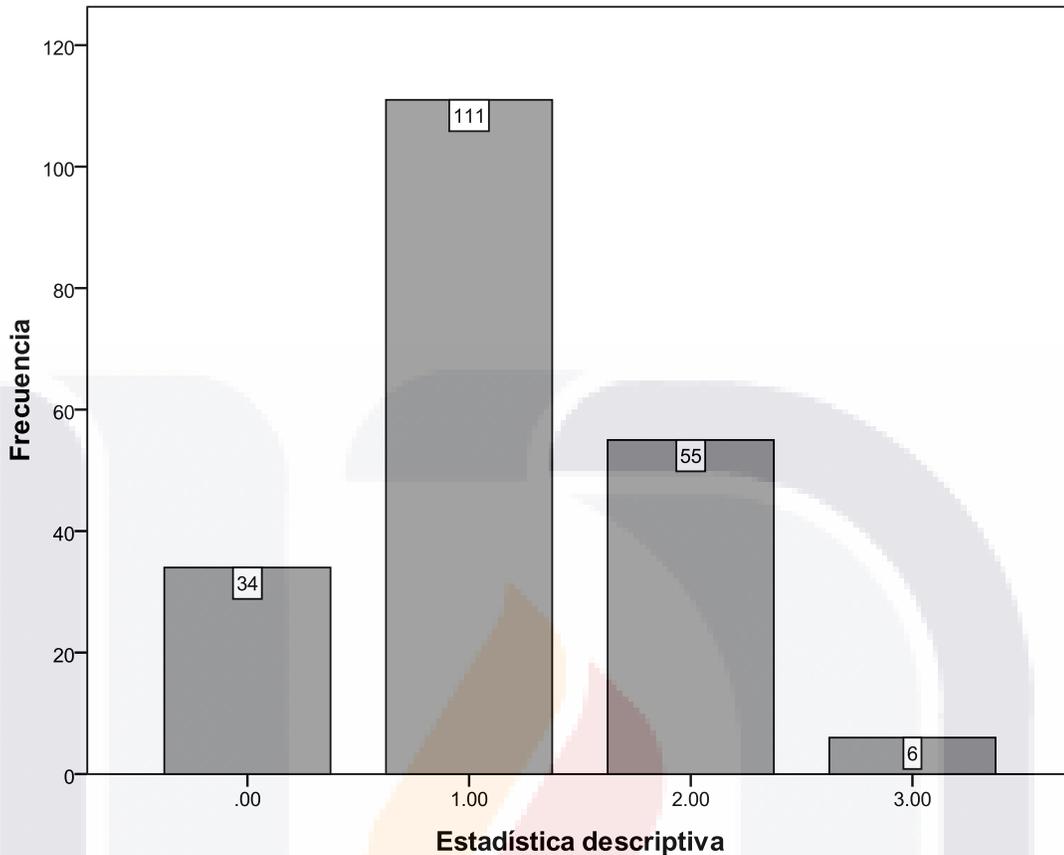
La primera dimensión que se analiza es la que se denomina *recolección de datos* que se mide con cuatro ítems, las preguntas están enfocadas a que el estudiante seleccione muestras representativas de una población. La gráfica 4 muestra la frecuencia de las respuestas obtenidas.



Gráfica 4 Recolección de datos

Se observa que de los 206 estudiantes 75 (36.4%) tuvieron tan sólo una respuesta correcta, solo 5 (2.4%) de ellos obtuvieron el 100% de los reactivos correctos y 39 (18.9%) no tuvieron ninguna respuesta correcta. La media de esta dimensión es de 1.4 y la desviación típica 1.01.

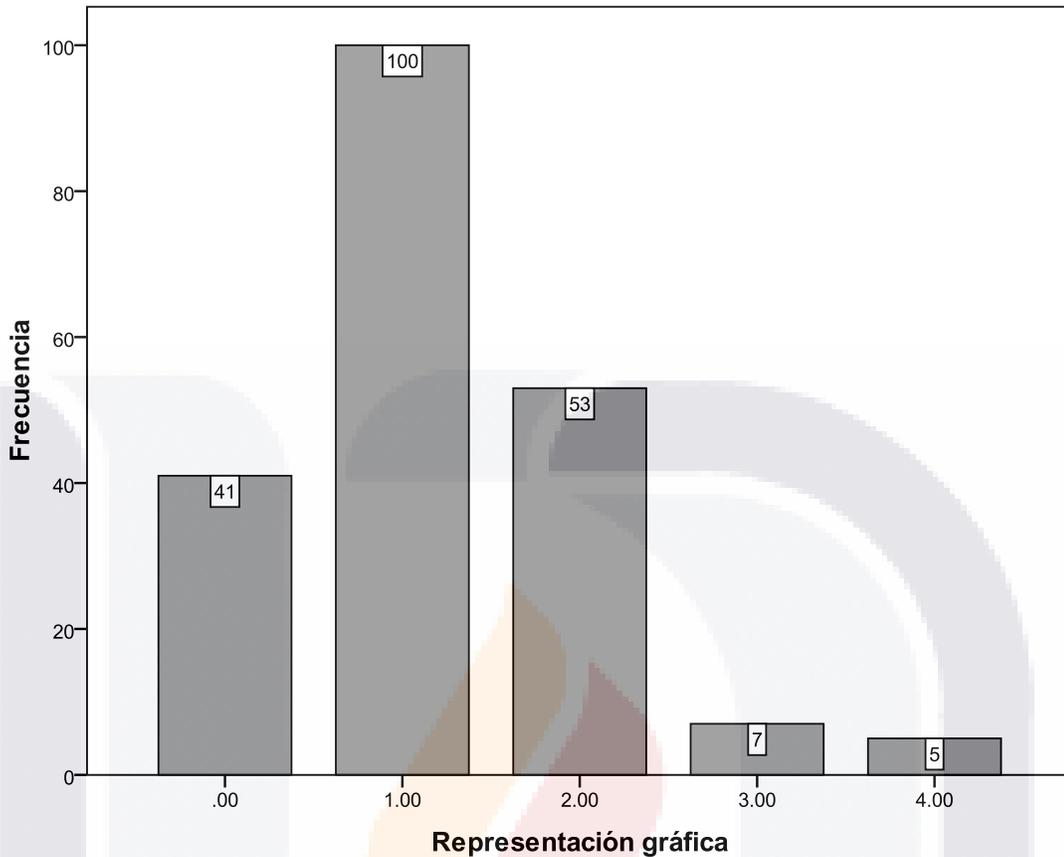
La segunda dimensión que compone la prueba es la referente a la *estadística descriptiva*, compuesta por 3 reactivos con valores dicotómicos. En esta dimensión las preguntas van enfocadas a estimar desviación estándar y de esta forma hacer comparaciones. La gráfica 5 muestra la frecuencia de las respuestas.



Gráfica 5. Estadística descriptiva

De los 206 estudiantes 111 (56.8%) tuvieron tan sólo una respuesta correcta, sólo 6 (2.9%) de ellos obtuvieron el 100% de los reactivos correctos y 34 (16.5%) no tuvieron ninguna respuesta correcta. Cabe aclarar –como se mostró en la gráfica 1- que los estudiantes reportan repetidamente, como uno de los conceptos difíciles de entender para poder contestar los ítems, la desviación estándar. La media de esta dimensión es de 1.1 y la desviación típica es de 0.72.

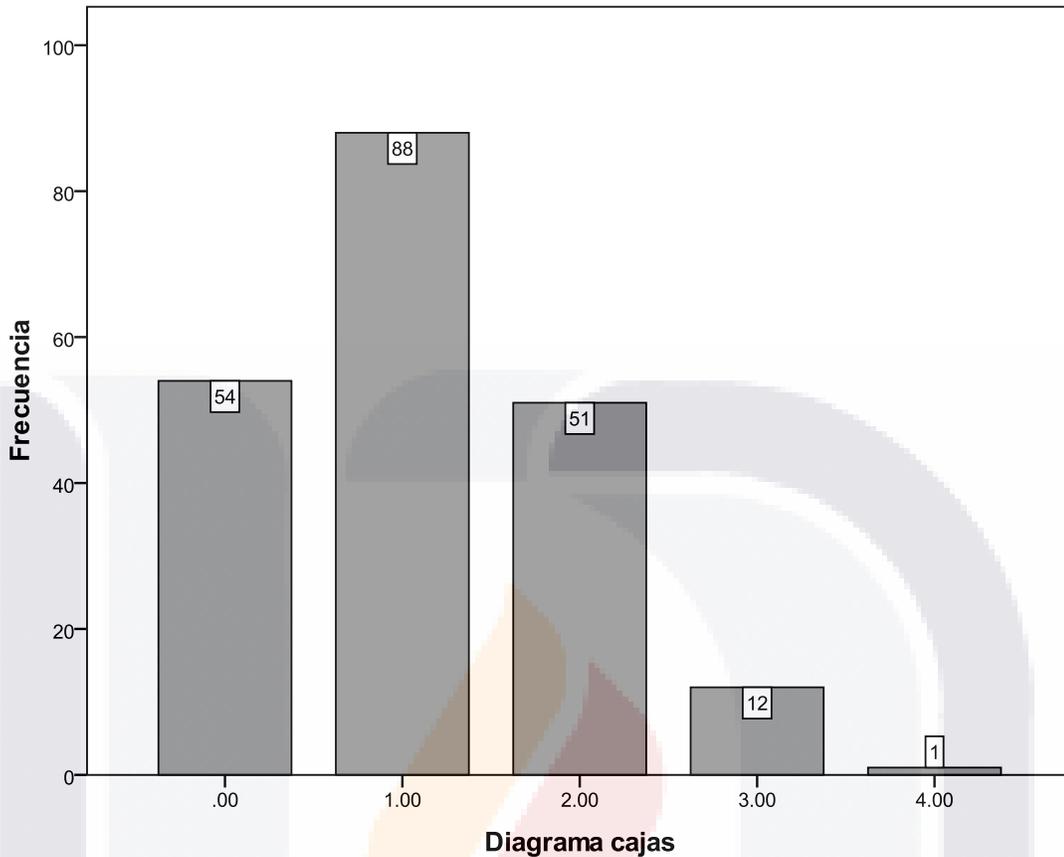
La tercera dimensión se refiere a la *representación gráfica* y está compuesta por 6 reactivos. En esta dimensión las preguntas van enfocadas a interpretar histogramas. La gráfica 5 muestra la frecuencia de las respuestas obtenidas.



Gráfica 6 Representación gráfica

De los 206 estudiantes 100 (48.5%) tuvieron tan sólo una respuesta correcta, solo 5 (2.4%) de ellos obtuvieron el 100% de los reactivos correctos y 41 (19.9%) no tuvieron ninguna respuesta correcta. La media de esta dimensión es de 1.19 y la desviación típica es de 0.88.

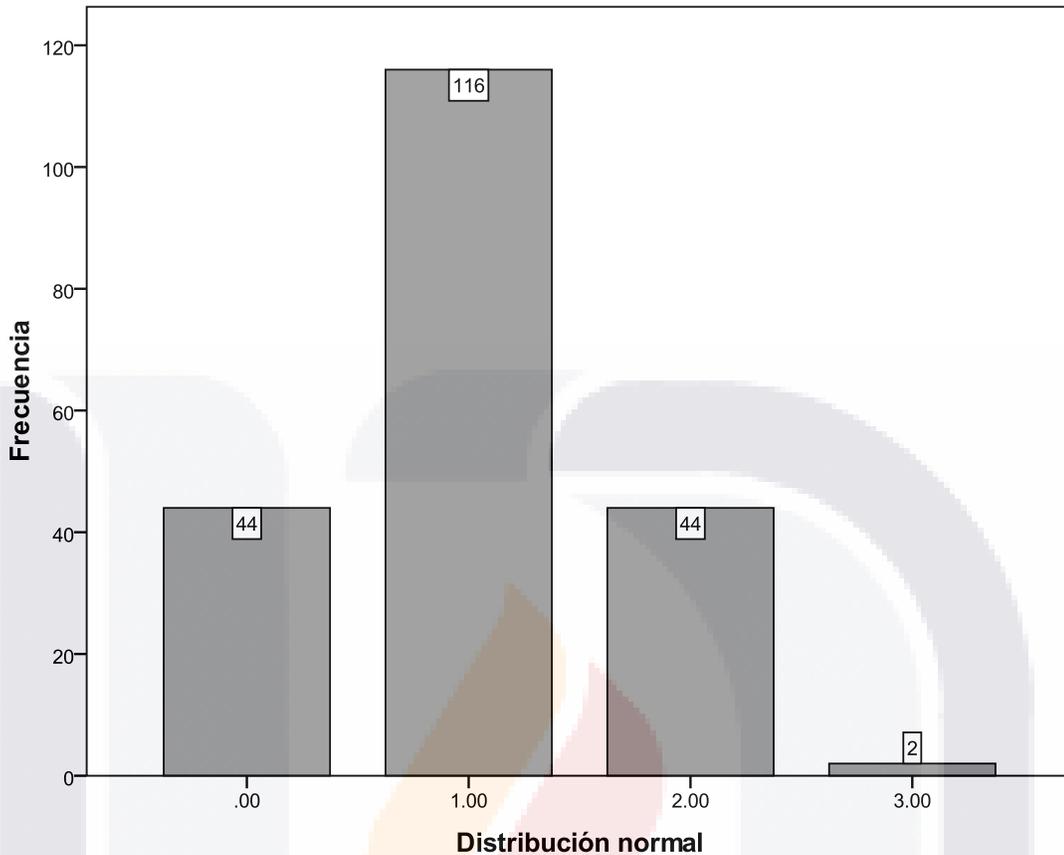
La cuarta dimensión es de *diagrama de cajas*, compuesto por 4 reactivos. Las preguntas están enfocadas a elegir el diagrama de cajas correcto. La gráfica 7 muestra la frecuencia de las respuestas obtenidas.



Gráfica 7 Diagrama de cajas

De los 206 estudiantes 88 (42.7%) tuvieron tan sólo una respuesta correcta, sólo uno (0.5%) de ellos obtuvo el 100% de los reactivos correctos y 54 (26.2%) no tuvieron ninguna respuesta correcta. La media de esta dimensión es de 1.11 y la desviación típica es de 0.88.

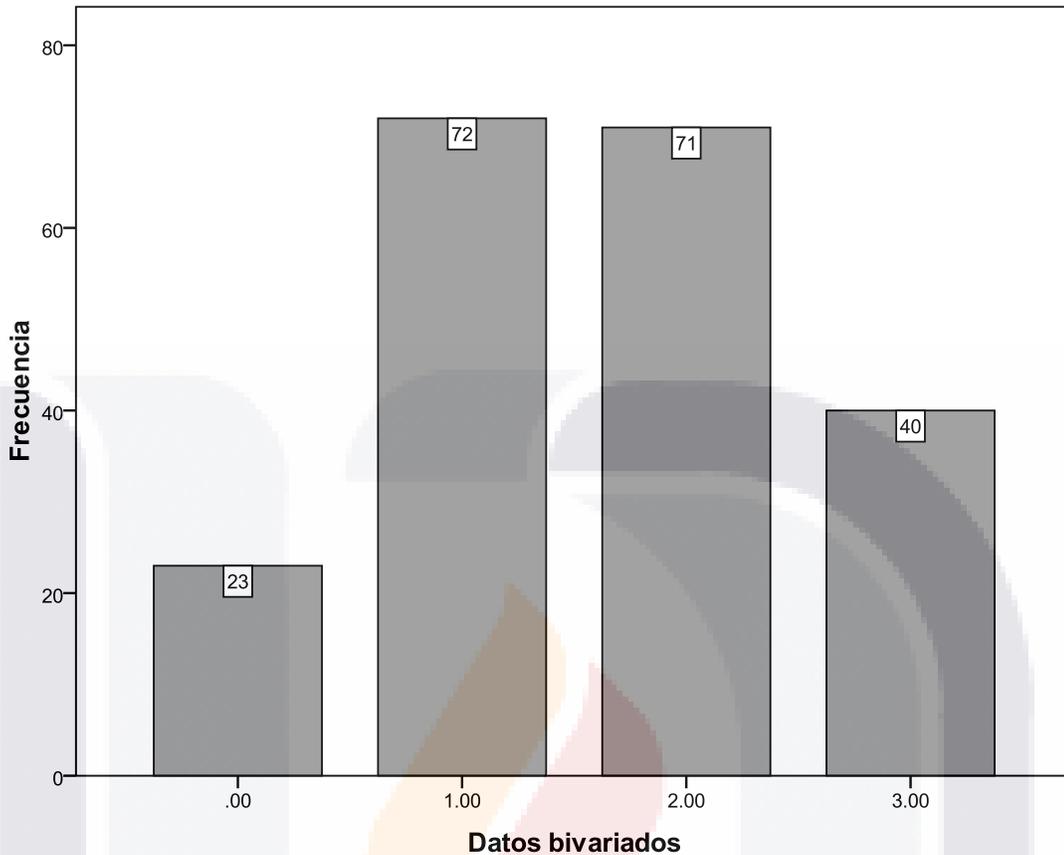
La quinta dimensión es la *distribución normal* compuesta por tres reactivos. Las preguntas están enfocadas a determinar qué sucede si la mediana es más grande que la media. La gráfica 8 muestra la frecuencia de las respuestas obtenidas.



Gráfica 8 Distribución normal.

De los 206 estudiantes 116 (56.3%) tuvieron tan sólo una respuesta correcta, sólo dos (1.0%) de ellos obtuvo el 100% de los reactivos correctos y 44 (21.4%) no tuvieron ninguna respuesta correcta. La media de esta dimensión es de 1.01 y la desviación típica es de 0.68.

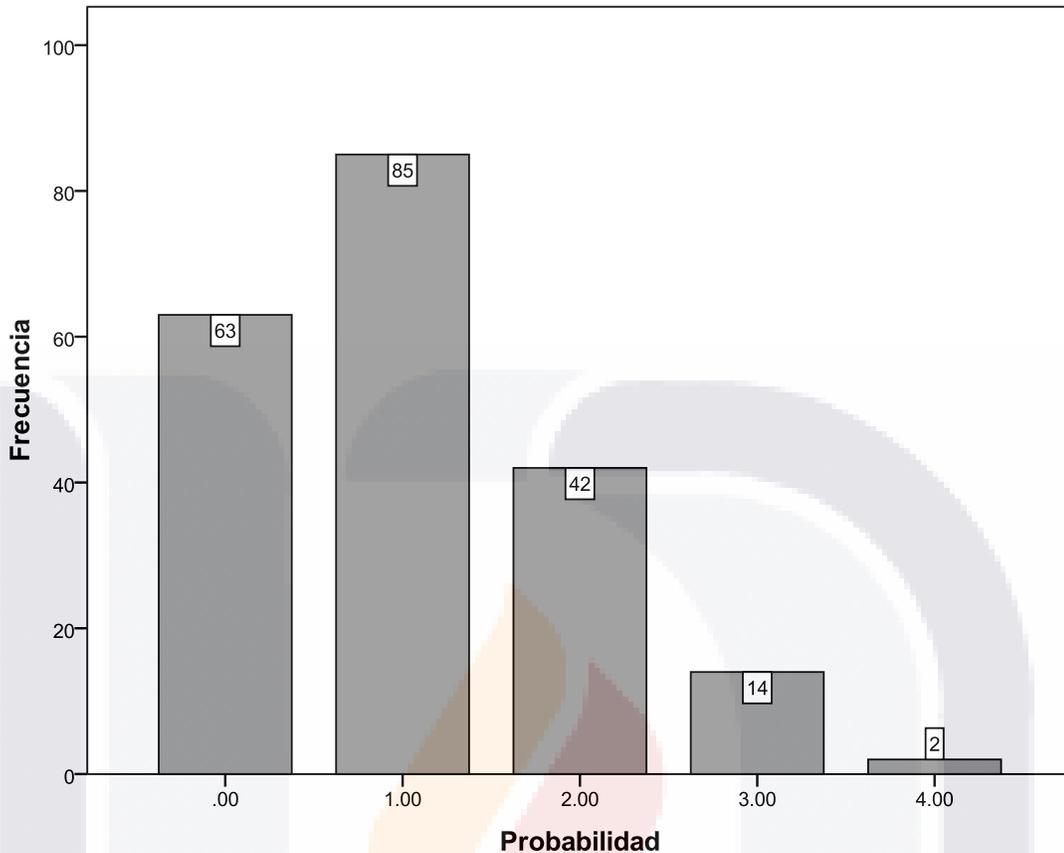
La sexta dimensión está compuesta por tres reactivos que representan a los *datos bivariados*, los cuales evalúan una relación bivariada mostrada en un diagrama de dispersión. La gráfica 9 muestra la frecuencia de las respuestas obtenidas.



Gráfica 9 Datos bivariados

De los 206 estudiantes 72 (35%) tuvieron tan sólo una respuesta correcta, 40 (19.4%) de ellos obtuvo el 100% de los reactivos correctos y 23 (11.2%) no tuvieron ninguna respuesta correcta. La media de esta dimensión es de 1.62 y la desviación típica es de 0.92. En conclusión, esta dimensión es la que reporta mejores puntajes en los resultados obtenidos en la evaluación.

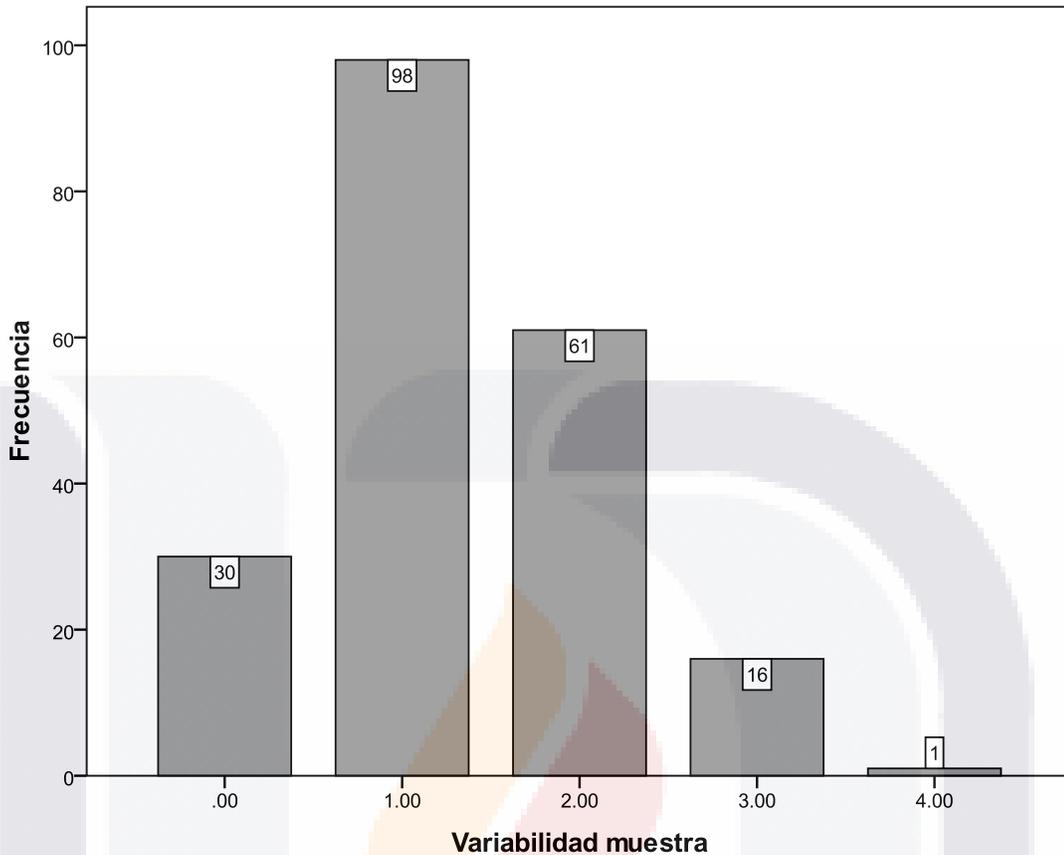
La séptima dimensión está compuesta por 4 reactivos y evalúa la *probabilidad*, es decir, trata de simular datos para encontrar la probabilidad de un valor observado. La gráfica 10 muestra la frecuencia de las respuestas obtenidas en esta dimensión.



Gráfica 10 Probabilidad

De los 206 estudiantes 85 (41.3%) tuvieron tan sólo una respuesta correcta, tan sólo 2 (1.0%) de ellos obtuvo el 100% de los reactivos correctos y 63 (30.6%) no tuvieron ninguna respuesta correcta. La media de esta dimensión es de 1.62 y la desviación típica es de 0.92. Cabe resaltar que en esta dimensión hubo el mayor porcentaje de alumnos que no tuvieron ninguna respuesta correcta.

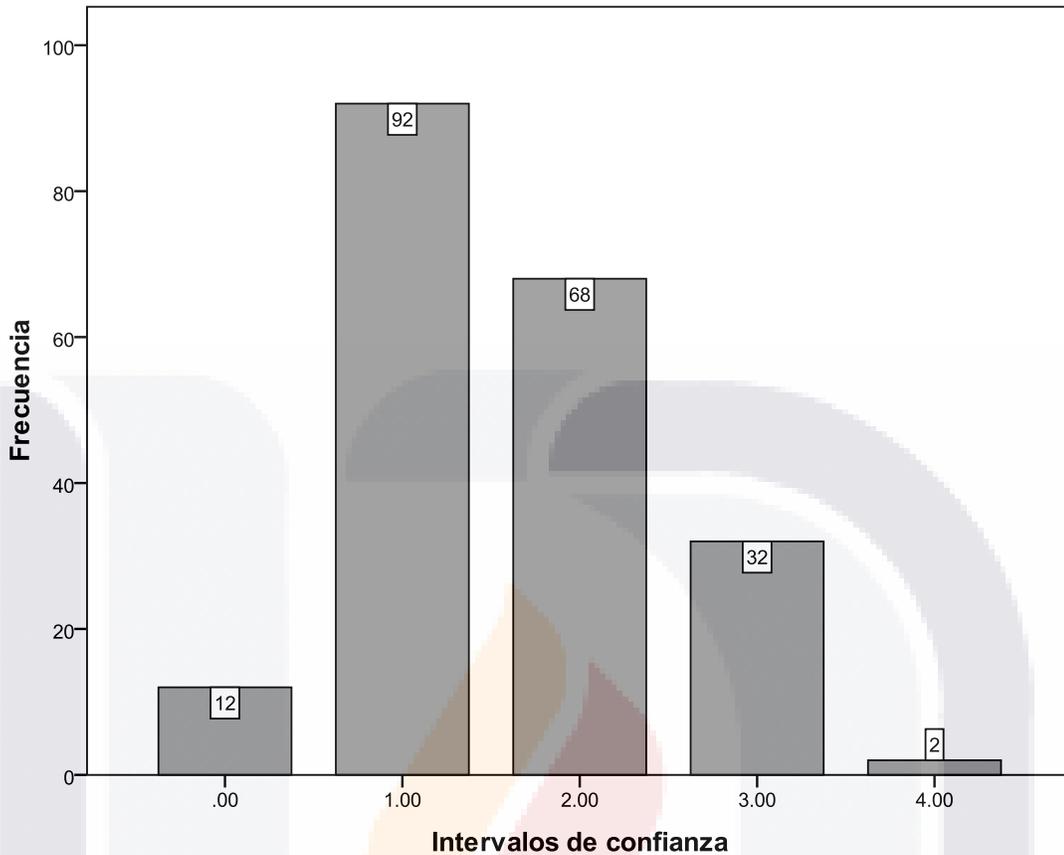
La octava dimensión es la que se refiere al tema de la *variabilidad de la muestra* la cual está enfocada a que el estudiante indique que existe más variabilidad en muestra pequeñas que en las grandes. Está compuesta por cuatro reactivos. La gráfica 11 muestra la frecuencia de las respuestas obtenidas en esta dimensión.



Gráfica 11 Variabilidad de la muestra.

De los 206 estudiantes 98 (47.6%) tuvieron tan sólo una respuesta correcta, tan sólo uno (0.5%) de ellos obtuvo el 100% de los reactivos correctos y 30 (14.6%) no tuvieron ninguna respuesta correcta. La media de esta dimensión es de 1.32 y la desviación típica es de 0.83.

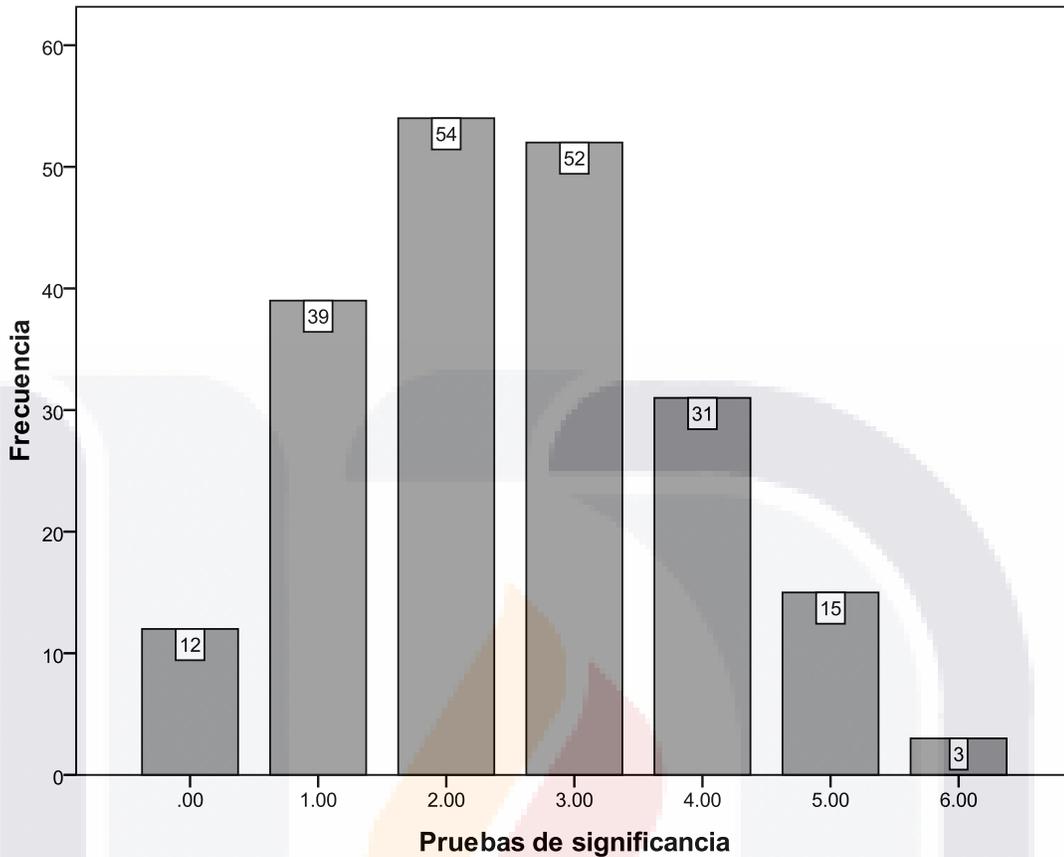
La novena dimensión se refiere a los *intervalos de confianza*, los cuales son evaluados con cuatro reactivos en donde se tienen que interpretar los límites de confianza. En la gráfica 12 se muestra la frecuencia de respuestas correctas de cada una de las preguntas que componen la dimensión



Gráfica 12 Intervalos de confianza

De los 206 estudiantes 92 (44.7%) tuvieron tan solo una respuesta correcta, tan sólo dos (1.0%) de ellos obtuvo el 100% de los reactivos correctos y 12 (5.8%) estudiantes más no tuvieron ninguna respuesta correcta. La media de esta dimensión es de 1.61 y la desviación típica es de 0.85.

La décima y última dimensión se refieren a las *pruebas de significancia*, que fueron evaluadas con 7 ítems. Esta dimensión se enfoca a conocer el valor *p*. En la gráfica 13 se observan las frecuencias de respuesta por cada uno de los ítems que componen la dimensión.



Gráfica 13 Pruebas de significancia.

De los 206 estudiantes tan sólo tres (1.5%) de ellos obtuvo el 100% de los reactivos correctos y 12 (5.8%) estudiantes no tuvieron ninguna respuesta correcta. La media de esta dimensión es de 2.52 y la desviación típica es de 1.37.

Después de haber analizado y dado respuesta al primer objetivo de investigación en el siguiente apartado se presentan los resultados para dar cuenta del segundo objetivo de investigación.

El segundo objetivo que dio pie al desarrollo de esta investigación, fue indicar si existe relación de comprensión estadística con el género, la formación académica (carrera) y el total de materias de estadística cursadas.

Para dar respuesta a este objetivo, lo primero que se realizó fue un análisis de la curva normal y así definir 4 categorías clasificadas de la siguiente forma: muy bajo, bajo, regular y alto, con la finalidad de categorizar una variable numérica –el puntaje total de los estudiantes- de tal forma de poder hacer un análisis de asociación.

La media teórica se determinó con un valor de 20 puesto que es el 50% de los reactivos de los 40 que componen la evaluación, siguiendo con esta línea, se hicieron varios puntos de corte donde se indica una relación espacial entre la media y la desviación estándar con respecto al área que se encuentra por debajo de la curva.

Según Pagano (1999), “cuando un conjunto de datos está distribuido de forma normal, 34.13% del área que se encuentra por debajo de la curva está contenida entre la (μ) y un dato igual a $\mu + 1\sigma$, 13.59% del área esta contenida entre un dato igual a $\mu + 1\sigma$ y un dato de $\mu + 2\sigma$; 2.15% del área está contenida entre los datos de $\mu + 2\sigma$ y $\mu + 3\sigma$ y 0.13 % del área que está más allá de $\mu + 3\sigma$. Esto representa el 50% del área” (p. 82).

Siguiendo lo propuesto por Pagano, los puntos de corte para la evaluación CAOS-4 se formaron con los datos del puntaje total que obtuvieron los estudiantes, es decir, entre 7 y 23 reactivos correctos.

El primer punto de corte, corresponde al 32.5% (13 reactivos) del área que se encuentra por debajo de la curva, es decir, esto corresponde al 65% justo al centro de la curva.

El segundo punto de corte es del 17.5% (7 reactivos), es decir, al sumarlo con el primer punto de corte se obtiene el 50% del total de los reactivos, de tal forma que a ambos lados de la campana se tendría la misma cantidad de datos.

Con la ayuda de estos puntos de corte, se realizó una clasificación de cuatro grupos, los cuales fueron denominados en muy bajo que se agrupó en sujetos con puntuación que van de 0-7 ítems correctos; el grupo de bajo con puntuaciones de 8 a 19 ítems correctos; uno más que corresponde a regular con puntuaciones de 20-32 ítems correctos; y finalmente el grupo de alto con puntuaciones que van de 33-40 ítems correctos.

Según la clasificación realizada a través del análisis propuesto por Pagano, el 2.4% (5) de la población que contestó la evaluación cae en la categoría de muy bajo, el 94.2% (194) de los estudiantes caen en la clasificación de bajo, sólo el 3.4% (7) se considera

estudiante regular en el tema de la comprensión estadística y ninguno de los estudiantes entró en el grupo que comprendía alumnos con un buen puntaje en la prueba. En resumen, los valores de los reactivos contestados en la prueba varían de 7 hasta 23, es decir, la forma de distribución de los datos esta más cargada hacia la izquierda, esto significa que los alumnos que presentaron la prueba tienden a estar en la categoría de bajo y solo unos pocos en la categoría de muy bajo.

Como parte del trabajo de tesis se tenía contemplado hacer un análisis de correlación, sin embargo al presentarse poca variabilidad de los datos, solo se realizó el ejercicio siendo esperable los valores poco significativos.

La relación entre las variables propuestas en este objetivo de investigación se comprobó utilizando la prueba no paramétrica X^2 ya que los datos no mostraron una distribución normal, la población era mayor a 30 casos y las variables eran de tipo nominal y ordinal.

La primera hipótesis planteada para este tipo de análisis fue:

Las personas del sexo masculino logran una comprensión estadística más elevada con relación a las personas del sexo femenino.

En la tabla 17 se observa que los datos mostraron suficiente evidencia para afirmar que las variables comprensión estadística y sexo no se encuentran relacionadas (X^2 , $p = 0.196$).

Tabla 17
Variable sexo y su relación con el nivel de comprensión estadística (NT = 206)

Sexo	N	Puntuación en el índice		
		comprensión muy baja	comprensión baja	comprensión regular
		%	%	%
Masculino	59	3.4%	89.8%	6.8%
Femenino	147	2.0%	95.9%	2.0%
Total	206	2.4%	94.2%	3.4%

Pearson Chi cuadrada: $p < .05$

Se concluye que tanto mujeres como hombres tuvieron el mismo aprovechamiento al contestar la prueba sin encontrar diferencias significativas entre géneros.

Una expectativa que se tenía, estaba enfocada a encontrar diferencias significativas entre el tipo de carrera que los alumnos cursaban y el nivel de comprensión adquirido, es decir, los que cursan la carrera de enseñanza del idioma inglés tienen menor nivel de

comprensión estadística que el resto del grupo ya que ellos no han cursado la materia como tal. La tabla 18 muestra los datos obtenidos al ejecutar la prueba de X^2 .

Tabla 18
Variable Carrera y su relación con el nivel de comprensión estadística (NT =206)

Carrera	N	Puntuación en el índice		
		comprensión muy baja	comprensión baja	comprensión regular
		%	%	%
Asesoría Psicopedagógica	37	0%	94.6%	5.4%
Ciencias políticas y Administración Pública	18	11.1%	88.9%	0%
Comunicación e información	20	5.0%	90.0%	5.0%
Comunicación organizacional	15	0%	100%	0%
Enseñanza del idioma Inglés	35	2.9%	94.3%	2.9%
Historia	16	0%	100%	0%
Psicología	38	0%	94.7%	5.3%
Sociología	4	0%	100%	0%
Trabajo Social	23	4.3%	91.3%	4.3%
<i>Total</i>	<i>206</i>	<i>2.4%</i>	<i>94.2%</i>	<i>3.4%</i>

Pearson Chi cuadrada: $p < .05$

La prueba de hipótesis indica que los datos mostraron suficiente evidencia para afirmar que las variables comprensión estadística y carrera no se encuentran relacionadas ($X^2, p = 0.724$).

Una tercera hipótesis estaba enfocada a encontrar diferencias significativas entre aquellos estudiantes que durante su formación profesional cursaron más de una materia de estadística, es decir, tienen mayor nivel de comprensión que el resto del grupo. La tabla 19 muestra los datos obtenidos al ejecutar la prueba de X^2 .

Tabla 19
Variable número de cursos y su relación con el nivel de comprensión estadística (NT= 206)

Número de cursos	N	Puntuación en el índice		
		comprensión muy baja	comprensión baja	comprensión regular
		%	%	%
Ningún curso	35	2.9%	94.3%	2.9%
Un curso	96	2.1%	93.8%	4.2%
Dos cursos	75	2.7%	94.7%	2.7%
<i>Total</i>	<i>206</i>	<i>2.4%</i>	<i>94.2%</i>	<i>3.4%</i>

Pearson Chi cuadrada: $p < .05$

La prueba de hipótesis indica que los datos mostraron suficiente evidencia para afirmar que las variables número de cursos de estadística y comprensión estadística no se encuentran relacionadas ($X^2, p = 0.982$).

La cuarta hipótesis tiene que ver con la encuesta de opinión en donde se esperaba encontrar diferencias significativas entre las respuestas obtenidas y la comprensión estadística. Para cada una de las ocho preguntas que contenía la escala la prueba de X^2 no fue significativa obteniendo valores de $p = 0.202$ hasta $p = 0.898$ lo cual indica que no se encontró ninguna relación.

Dando por terminado el análisis de los datos, en el siguiente capítulo se desarrolla la discusión de los resultados.

CAPITULO V

DISCUSIÓN

Este capítulo presenta la discusión de los resultados, así como un apartado en dónde se describen las futuras investigaciones que se generan a partir de la aquí desarrollada; finalmente se encuentran las recomendaciones que se ofrecen derivadas del estudio.

5.1 Discusión

La evaluación de CAOS -4 tuvo una serie de revisiones que buscaban determinar que la traducción realizada fuera la adecuada para el contexto mexicano, por lo que la primera aplicación formal de esta evaluación fue fiel a la prueba original. Al someter la prueba a un proceso de confiabilidad interna a través de alfa de Cronbach los resultados no fueron alentadores ya que estos estuvieron muy por debajo de lo que Aiken (1996) ha reportado como el valor óptimo para las pruebas de aprovechamiento. Sin embargo, esto no quiere decir que la prueba deba descartarse, como ha referido Morales (2007) un coeficiente de confiabilidad bajo no indica necesariamente que el instrumento sea malo y que no es posible utilizarlo, si no más bien, que debe mejorar en cuanto a su calidad de confiabilidad y validez, de tal forma que se realice una revisión exhaustiva de cada una de las preguntas por dimensión o ítem por ítem para mejorar la calidad del instrumento, ya que son pocas las evaluaciones existentes en México que intentan conocer el manejo de la estadística en los estudiantes a diferencia de algunos otros países que ya han comenzado a estudiar el razonamiento y el pensamiento como constructos más complejos.

En México el constructo de comprensión estadística no se ha abordado como tal, existen evaluaciones tales como Extra-Es que evalúan los conocimientos estadísticos de los estudiantes universitarios, pero su finalidad es totalmente diferente ya que se pretende que el alumno resuelva problemas estadísticos con ayuda de formularios y a través de la memorización, abordando de forma limitada el razonamiento y pensamiento estadístico. Como lo ha mencionado Budé (2006) para evaluar la comprensión básica se solicita que se respondan preguntas de opción múltiple que hacen referencia a procesos básicos como lo son definiciones o fórmulas, para evaluar niveles más profundos de comprensión se hace más complejo el proceso ya que es esencial que las preguntas sean las apropiadas para el dominio que se está evaluando, de tal forma que se hagan planteamientos de la vida real en contextos específicos.

Por otro, lado estudiar la comprensión estadística en un contexto totalmente diferente y a través de la aplicación de la prueba CAOS-4 la cual fue diseñada para una población específica, se considera que es un factor de influencia importante para los resultados presentados en este trabajo de tesis, ya que las características personales, sociológicas, psicológicas, pedagógicas entre otras de los estudiantes son totalmente diferentes. Es por ello que es importante seguir en esta línea de investigación para poder aportar mayor literatura en el contexto mexicano.

Un aspecto importante que hay que tomar en cuenta es el currículo, que en el contexto estadounidense es sólo una base para que el profesor desarrolle sus propios contenidos educacionales, lo que permite que sea flexible; en este mismo contexto hay una tendencia a seguir los estándares educativos nacionales que incluyen niveles educativos desde preescolar hasta nivel medio (k-12), incluso también nivel superior, de tal manera que los investigadores educativos también se interesan en evaluar lo que los estudiantes han comprendido específicamente en estadística y así la prueba CAOS-4 fue desarrollada en virtud de que se carecían de evaluaciones para determinar la comprensión estadística de los estudiantes de nivel superior y los resultados de esas evaluaciones reportan que los alumnos lograron un desempeño adecuado en la ejecución de la prueba con un mínimo de 18% (7) de respuestas correctas y un máximo del 100% (40) de preguntas contestadas correctamente.

En México, el currículo único es propio de la educación básica, bajo el cual se tiene que desarrollar un curso y de esta forma se pretende que el estudiante logre adquirir los conocimientos propuestos, pero prevalecen los enfoques tradicionales centrados en la aplicación de fórmulas y procedimientos, sin abordar el desarrollo del razonamiento estadístico, por lo que no es de extrañar que en la prueba aplicada, los participantes según los resultados obtenidos, se encuentran ubicados en la categoría de bajo desempeño de tal forma que el porcentaje mínimo de respuestas correctas fue del 12.5%, es decir, 5.5% menos que los estudiantes de la Universidad de Minnesota, con un porcentaje máximo de 57.5% es decir, 42.5% menos. Con la debida cautela debido a la baja confiabilidad del instrumento, se puede llegar a concluir que los resultados obtenidos reafirman lo que en la investigación y en la experiencia de la enseñanza se ha encontrado, es decir, sabe que son pocas las personas que emplean de forma efectiva la comprensión estadística aun después de la instrucción; pueden dar solución técnica a un

problema, pero carecen de las habilidades para interpretar los resultados obtenidos o bien utilizan de forma inadecuada los conceptos estadísticos (Schaun y Mattern, 1997).

Los contenidos de los cursos de estadística en la Universidad Autónoma de Aguascalientes, suelen ser muy parecidos de una carrera a otra y ahí existe quizá una parte del problema, ya que hay una falta de contextualización de la materia dados los usos que puede tener en un área profesional u otra; los estudiantes universitarios llevan la materia de estadística, en la cual se tiene previsto revisar temas como estadística descriptiva, probabilidad y estadística inferencial, además, la materia tiene como objetivo principal la introducción al análisis estadístico que contempla actividades de aprendizaje, referidas a situaciones reales por lo que el estudiante logrará obtener conocimientos suficientes para entender el comportamiento de algunos fenómenos relacionados con su área de estudio y en consecuencia desarrollará una habilidad para aplicar una metodología estadística que tiene que ver con el análisis de datos además de evaluar y comunicar los resultados obtenidos.

El objetivo planteado para los cursos de estadística en el contexto de la Universidad Autónoma de Aguascalientes es muy ambicioso ya que para lograr un análisis estadístico como lo refiere el objetivo general, el estudiante debe conocer los aspectos básicos de alfabetización, como describir, traducir e interpretar grupos de datos y saberlos representar a través de símbolos o gráficas, para posteriormente pasar a la siguiente etapa según lo descrito en la literatura¹³, es decir, el razonamiento, donde el estudiante explica el proceso estadístico en la resolución de planteamientos y finalmente adquiere procesos más avanzados como el pensamiento estadístico. En la tabla 20 se muestra un ejemplo del contenido a revisar en el tema de recolección de datos en la materia de estadística que forma parte del programa¹⁴ de la Universidad Autónoma de Aguascalientes para las carreras del Centro de Ciencias Sociales y Humanidades así mismo se encuentra información relacionada con lo que evalúa CAOS-4 y EXTRA-EX.

¹³ Revisar tabla 1. Taxonomía de habilidades

¹⁴ Revisar anexo F relacionado con el plan de estudios de la materia de estadística de UAA

Tabla 20
Ejemplo recolección de datos

Contenido plan de estudios	Objetivo plan de estudios	CAOS-4	EXTRA-EX
Poblaciones y muestras Censos y encuestas El muestreo aleatorio simple El muestreo sistemático Toma de datos Mediciones observadas y esperadas	Comprenderá la utilidad del muestreo, sus bondades y limitaciones, el proceso de obtener información de interés.	Seleccionar una muestra representativa de la población así como lograr hacer una distribución adecuada del muestreo para poder generalizar los datos de la población.	Fundamentos de muestreo incluye: Conceptos, definiciones y características Origen de los datos incluye el cálculo de la muestra y la selección del método de muestreo adecuado.

Como se aprecia en la tabla 19, el objetivo del plan de estudios para el tema de recolección de datos está enfocado a la comprensión, lo cual supone que abarca la alfabetización, el razonamiento y el pensamiento estadístico y no dista mucho de lo que evalúa CAOS-4, ya que todas las preguntas de la prueba que incluyen esta dimensión están enfocadas al pensamiento estadístico, donde se utilizan habilidades de crítica y evaluación de datos, es decir, se da por sentado que el estudiante ya maneja los dominios previos; Extra-Ex se centra en un conocimiento básico ya que los módulos evaluados en este ejemplo se enfocan al cálculo y a la memorización.

A pesar de que el objetivo del plan de estudios revisado se refiere a la comprensión del tema de recolección de datos, en la prueba aplicada en esta investigación, 36.4% de los estudiantes lograron contestar 1 pregunta de las 4 que componen la dimensión y tan solo el 2.4% de la población contestó correctamente el total de las preguntas; aquí cabe la pregunta, ¿Que influye para que un estudiante conteste correctamente una pregunta y no las cuatro planteadas?, si es de suponerse que el objetivo de la prueba es similar al del plan de estudios que ellos cursaron, o bien habría que preguntarse qué aspectos hay que modificar de la dimensión de la prueba para mejorar la calidad de ésta.

Éste es un ejemplo claro de que la educación estadística en México aún tiene que fortalecerse en temas básicos y consecutivamente en los temas más complejos, de tal forma que el estudiante aprenda en el plano conceptual, pero también llegue a la interpretación de los datos, por que como afirma Shau y Mattern (1997), pareciera ser que los conceptos son un ente aislado de la comprensión estadística; por ejemplo los estudiantes pueden calcular la desviación estándar, sin embargo la interpretación que se hace de ésta en ocasiones es incorrecta. El análisis realizado es sencillo pero tiene que llevarse a cabo detalladamente con los contenidos estadísticos implementados en el contexto de la educación superior para adaptar la prueba al contexto mexicano¹⁵.

Al seleccionar una prueba de logro cuando ésta ya ha sido estandarizada es importante verificar que tenga la dificultad apropiada y un contenido que vaya de acuerdo con los objetivos de enseñanza de la escuela. Esto significa que el nivel de conocimiento o capacidad de los sujetos y el contenido y objetivos del plan de estudios deben determinarse antes de decidir qué pruebas aplicar (Aiken, 1996). Es por eso que CAOS-4 debe ser revisada a la luz de los planes de estudio para adecuar cada uno de los ítems que no han logrado un índice de discriminación y de dificultad óptimos de tal forma que la prueba sea estandarizada para evaluar no solo la alfabetización, si no también el razonamiento y la comprensión estadística.

Al realizar el análisis de discriminación de los ítems se obtiene como resultado que, mientras no se pruebe con muestras más grandes y heterogéneas, la prueba tiene que ser usada con todos sus reactivos, para posteriormente mejorar la calidad de la prueba, es decir, revisar y/o modificar los planteamientos de cada una de las preguntas, así como de las opciones de respuesta, puesto que muchas de ellas necesitan mejorarse, según los datos obtenidos en esta investigación. Tomar en cuenta las diez dimensiones de CAOS-4 para reformular las preguntas supondría llevar un nuevo proceso para certificar la confiabilidad y validez de cada uno de los componentes de la prueba y de esta forma preparar un prueba con mayor precisión en cuanto al número de reactivos que deben ser fáciles y difíciles en el contexto mexicano.

Así el objetivo general de esta investigación, fue analizar la comprensión estadística que manejaban los estudiantes de educación superior que han realizado al menos un

¹⁵ Previamente habría que hacer una revisión de los planes y programas de otras universidades del estado de Aguascalientes y del país, pues puede darse el caso de que el problema del aparente no fomento del razonamiento estadístico sea propio de la Universidad Autónoma de Aguascalientes, o quizás sólo de las carreras del área de ciencias sociales.

TESIS TESIS TESIS TESIS TESIS

curso de estadística a través de la prueba CAOS-4; para dar respuesta a este objetivo se realizó un análisis descriptivo de los datos en donde se logra percibir que la mayoría de los estudiantes caen en un nivel de comprensión estadística bajo y ninguno de ellos logra un nivel de comprensión óptimo en donde todos los reactivos hayan sido contestados de forma correcta. Las carreras que destacan con sujetos en la categoría denominada como puntaje regular corresponden a Psicología y Asesoría Psicopedagógica con tan solo dos estudiantes cada una, de tal manera, que habría que investigar que factores influyeron para que los sujetos de estas carreras hayan logrado puntajes regulares. Por otro lado la carrera de Ciencias Políticas y Administración Pública cae dentro de los puntajes más bajos que obtuvieron los estudiantes al contestar la prueba, esto es de llamar la atención porque de las tres carreras mencionadas, únicamente asesoría psicopedagógica lleva un solo curso de estadística como parte de su plan de estudios y el resto lleva dos.

Aunque la validez y confiabilidad de la prueba fue limitada, ofrece información relevante para futuros trabajos de investigación ya que en 2008, Salazar encontró que es importante el número de materias de estadística cursadas para mejorar los puntajes que se obtienen en las pruebas estandarizadas y en esta investigación aún no es concluyente esta afirmación.

Los resultados, llevan a concluir que en el contexto mexicano, específicamente el de la Universidad Autónoma de Aguascalientes, los estudiantes del Centro de Ciencias Sociales y Humanidades en la primera aplicación formal de CAOS-4, no lograron obtener un puntaje óptimo según los criterios establecidos para este estudio, es decir, más de 23 reactivos correctos, sin embargo no hay suficiente evidencia que afirme que los resultados muestran que existe una limitada comprensión estadística que le permita al alumno utilizarla como parte de su práctica profesional, de tal manera que aún se tiene que perfeccionar el instrumento para aplicarlo en diferentes contextos universitarios para llegar a concluir esta aseveración.

Estos resultados llevan a discutir que en la actualidad, países de primer mundo están adaptando nuevas tendencias para que sus estudiantes sean capaces de leer datos numéricos. México, un país que está teniendo grandes cambios educativos, está haciendo también lo propio, sin embargo aún queda mucho por hacer; según Pinto (2010), en la literatura se detectan por lo menos seis focos-problemas y entre los cuales resaltan tres: 1. el uso y manejo incorrecto de la estadística; 2. la falta de contextualización y aplicación de la estadística dirigida a profesionales de áreas diferentes a las matemáticas; 3. la

diversidad en los perfiles de la formación inicial y continua de los profesores de estadística. Estos temas son clave para mejorar la enseñanza de la disciplina en cuestión.

El primer objetivo específico que se planteó en la investigación fue describir las dimensiones estadísticas que los estudiantes universitarios comprenden, de tal forma que las pruebas de significancia, datos bivariados e intervalos de confianza son las que mejor se comportan en la prueba al obtener mayores tasas de respuesta, es decir, los estudiantes conocen más en relación a estos temas. En cuanto a las dimensiones que tienen menor tasa de respuesta se reporta el tema de probabilidad, y distribución normal que son temas que se son revisados en el curso de estadística que se ofrece en el Centro de Ciencias Sociales y Humanidades para las carreras evaluadas. En conclusión, los programas de estudio y los cursos de estadística deberían enfocarse a reforzar la comprensión de los estudiantes desde la alfabetización, hasta lograr un pensamiento y en consecuencia una comprensión estadística, de tal manera que los alumnos logren integrar conceptos y solución de problemas, suponiendo que efectivamente hubiera un interés por fomentar el razonamiento y el pensamiento estadístico.

El segundo objetivo específico planteado al inicio de este trabajo era el indicar si existe relación con los puntajes de comprensión estadística y un cuestionario de opinión, género y formación académica. Los resultados no mostraron diferencias significativa, los resultados de la evaluación concluyen que la mayoría de los estudiantes lograron obtener la clasificación de bajo, es decir no hay diferencia entre hombres ni mujeres, tampoco entre carreras y mucho menos entre las opiniones que manifiestan los estudiantes; la prueba no logra discriminar diferencias entre las variables, esto significa que todos están en igualdad de circunstancias.

Por último, el cuestionario de opinión aplicado en esta investigación refleja que tan sólo el 46.5% de la población total se interesa en aprender más acerca de estadística es decir están conscientes de que aún falta desarrollar habilidades para ejercer la disciplina de forma adecuada; el resto opina que la estadística es poco útil y aburrida; es decir, los resultados de la prueba refleja que los propios alumnos detectan la necesidad de desarrollar habilidades estadísticas óptimas.

Conclusiones

El trabajo de tesis desarrollado ha permitido reflexionar lo que aún falta por hacer para seguir fortaleciendo la cultura estadística en el contexto mexicano, ya que según Araujo (2006) “el público en general tiene una percepción generalizada con relación a que la estadística miente, es fría o bien es poco confiable”.

Siguiendo las propuestas del mismo autor, es importante reconocer que la estadística es parte de los conocimientos básicos que todo profesionista necesita para poder resolver problemas estadísticos, pues en muchas ocasiones esos conocimientos se ven limitados a fórmulas sin sentido, y entonces, a pesar de cubrir una serie de cursos estadísticos los estudiantes responden no sentirse preparados para ejercer la disciplina en el ámbito profesional cuando son cuestionados respecto al tema.

En consecuencia, es primordial poder seguir realizando investigación para fortalecer las habilidades estadísticas del estudiante y a través de pruebas objetivas y otros instrumentos de medición, tales como entrevistas y observaciones, reconocer las áreas de oportunidad tanto pedagógicas y académicas durante la formación profesional. Esto no quiere decir que haya que formar profesionales especializados en estadística, si no más bien, prepararlos de forma adecuada para enfrentar los retos en diversos ambientes como por ejemplo el laboral y el social.

Por lo tanto, cabe la pregunta ¿los resultados de los estudiantes de la Universidad Autónoma de Aguascalientes que pertenecen al Centro de Ciencias Sociales y Humanidades son similares a los estudiantes que pertenecen a otras universidades? Este tipo de preguntas refuerzan la importancia de tener un instrumento adecuado para medir la comprensión estadística dentro del país.

En México aún queda mucho camino por recorrer, ya que apenas se están consolidando los grupos de trabajo que se centran en impulsar este tipo de investigaciones en todos los niveles educativos.

Limitaciones del trabajo

Las limitaciones de esta investigación se enuncian en los siguientes párrafos.

El estudio sólo se llevó a cabo con estudiantes de semestres avanzados del Centro de Ciencias Sociales y Humanidades de la Universidad Autónoma de Aguascalientes, pues se suponía que ellos habían adquirido el bagaje necesario para contestar la prueba al haber cursado una o dos materias de estadística, dejando de lado a otros centros como son, el de Ciencias Agropecuarias, Arte y cultura, Ciencias Básicas, Ciencias Económico

Administrativas, Ciencias Empresariales, Ciencias de la Ingeniería y Ciencias de la Salud, los que convendría incluir en una muestra mucho más amplia.

El análisis de este estudio se vio limitado a una evaluación de lápiz y papel las futuras investigaciones deberían incluir una entrevista con aquellos estudiantes que obtuvieron los puntajes más altos y los puntajes más bajos para conocer a detalle los procesos de construcción de conocimiento del estudiante.

Finalmente, otra limitación considerable es la confiabilidad de la prueba además de que se pudieron considerar otras variables diferentes a la carrera o el sexo, tales como las pedagógicas o psicológicas.

Futuras investigaciones

Este trabajo ha permitido avanzar en el conocimiento de la comprensión estadística en el ambiente universitario, sin embargo, es importante indicar algunas posibles investigaciones que surgen a partir de los resultados de este estudio.

Para ampliar el conocimiento de la comprensión se plantean al menos cinco futuras investigaciones que pretendan continuar con una línea de investigación respecto al tema.

- a. En cuanto a la prueba CAOS -4 se refiere, se propone analizar los resultados de la prueba en otros contextos universitarios para contrastar los datos con los que se han obtenido en esta investigación. Un primer punto es administrar la prueba a muestras mucho más grandes en universidades de diferentes tipos en México. Un segundo punto implica que esta muestra sea para estudiantes de áreas matemáticas y no matemáticas con la finalidad de conocer la variación en cuanto a las respuestas.
- b. Modificar, adaptar y/o rediseñar la prueba por cada dimensión con la finalidad de volver a realizar el proceso de validez y confiabilidad que supondrían dar mayor consistencia al instrumento. Entre otros aspectos, es necesario: a) elegir todos los reactivos que deben ser eliminados y contrastarlos con los programas de estudio para saber si realmente el estudiante logró la comprensión estadística necesaria o bien son temas en la prueba que no se contemplan en aquellos; b) revisar las opciones de respuesta de aquellos reactivos en los que el estudiante no están discriminando para mejorar la prueba; c) analizar los contenidos de los objetivos estadísticos que se tienen a nivel nacional en el caso que los hubiera y que llevan las carreras de ciencias sociales y humanidades específicamente con la finalidad de adaptar los niveles cognitivos de la prueba y posteriormente ampliarla a otras

disciplinas como ciencias exactas, ciencias biológicas, entre otras; y finalmente d) con base en el análisis de los reactivos es importante eliminar o modificar aquellos reactivos que no son compatibles en el contexto mexicano.

- c. A la luz de los resultados de está y otras investigaciones, analizar los planes y programas estadísticos que se llevan desde nivel básico hasta nivel superior para determinar aquellos temas que deben ser fortalecidos por los profesores durante los cursos de estadística.
- d. Otro tema de investigación se refiere a las dificultades que tienen los profesores para la enseñanza de la estadística, específicamente en el contexto universitario y de esta forma elaborar cursos para profesores que atiendan las debilidades y refuercen y mejoren las fortalezas que ya se tienen con la finalidad de influir en los procesos de aprendizaje de los estudiantes. Por ejemplo, enfocarse específicamente a elaborar un perfil profesional del profesor de estadística para conocer quiénes son los maestros que entrenan a los estudiantes en estos temas; otra investigación está enfocada a conocer los métodos de enseñanza que utilizan los profesores de estadística; y finalmente conocer las actitudes del profesor ante esta disciplina ya que mucho influye en la transmisión de conocimientos.

Glosario

Alfabetización Estadística: La habilidad para interpretar, evaluar críticamente y comunicar información, es decir, involucra palabras, símbolos y términos, así como ser capaz de leer y entender la estadística en los medios de comunicación.

CAOS-4: Comprehensive Assessment of Outcomes in Statistics

Comprensión estadística: Conexión de estructuras cognitivas o redes mentales que desarrolla el sujeto y que le permiten acceder a formas más elevadas de conocimiento, siendo las deseables el razonamiento y la comprensión estadística

Confiabilidad: El instrumento mide una característica relativamente estable, es decir, se refiere al atributo de consistencia

Distribución de frecuencias: Agrupación de datos que indican el número de observaciones en cada categoría.

Ji-cuadrada: Calcula si los resultados estadísticos de un experimento se alejan significativamente o no de los resultados esperados del modelo teórico.

Muestra: Subconjunto de casos o individuos de una población estadística.

Pensamiento Estadístico: Habilidad para ver el proceso como un todo para entender la relación y significado de la variación en este proceso; tener la capacidad de explorar datos más allá de lo que está escrito en los textos

Pruebas de logro: Las pruebas de ejecución máxima o también llamadas de logro están diseñadas para medir los logros de los estudiantes, como resultado de la exposición a una experiencia de aprendizaje definida.

Razonamiento Estadístico: Razonamiento: entender, ser capaz de explicar procesos estadísticos y ser capaz de interpretar completamente los resultados

Validez: El grado en el cual una prueba mide aquello para lo que se diseñó; que tan bien mide lo que pretende medir en un determinado contexto

Bibliografía

- Aiken, L.R. (1996). *Tests psicológicos y evaluación*. México: Prentice Hall. Hispanoamericana.
- Araujo, C., (2006). Statistical Illiteracy in Latin America: A consequence of the different visions about the meaning of statistics. *International Conferences on Teaching Statistics 7*, pp. 1-6.
- Batanero, C. (2001). *Didáctica de la Estadística* (Versión DX Reader). Recuperado de http://www.pucrs.br/famat/viali/graduacao/matematica/material/referencias/didactic_aestadistica.pdf
- Backhoff, E., Andrade, E., Sánchez, A. y Peon, M. (2008). *Estudio comparative del aprendizaje en sexto de primaria en México 2005-2007: Español y Matemáticas*. México: INNE. Recuperado de: <http://www.inee.edu.mx/archivosbuscador/2007/01/INEE-20070164-comparativo08.pdf>
- Backhoff, E., Larrazolo, N. y Rosas, M. (2000). Nivel de dificultad y poder de discriminación del Examen de Habilidades y Conocimientos Básicos (EXHCOBA). *Revista Electrónica de Investigación Educativa*, 2(1). Consultado el 21 de abril de 2013 en: <http://redie.uabc.mx/vol2no1/contenido-backhoff.html>
- Ben-Zvi, D. y Garfield, J. (2004). *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*. Dordrecht: Kluwer Academic Publishers.
- Broers, J. (2008). Helping students to build a conceptual understanding of elementary statistics. *The American statistician*, 62, (2), pp. 161-166. doi: 10.1198/000313008X302091
- Budé, L. (2006). Assessing student's understanding of statistics. *International Conferences on Teaching Statistics*, 7, pp. 1-6.
- Canché, L., (2009). *La comprensión gráfica de los alumnos del nivel primaria* (tesis de maestría). Universidad Autónoma de Yucatán. Mérida de Yucatán.
- Chance, B. L. (2002). Components of statistical thinking and implications for instruction and assessment. *Journal of Statistics Education*, 10(3).
- Cohen, R., y Swerdlik, E. (2002). *Pruebas y Evaluación Psicológicas. Introducción a las pruebas y a la medición*. México: McGrawHill
- delMas, R. (2002). Statistical literacy, reasoning and learning: A commentary. *Journal of Statistics Education*, 10(3).

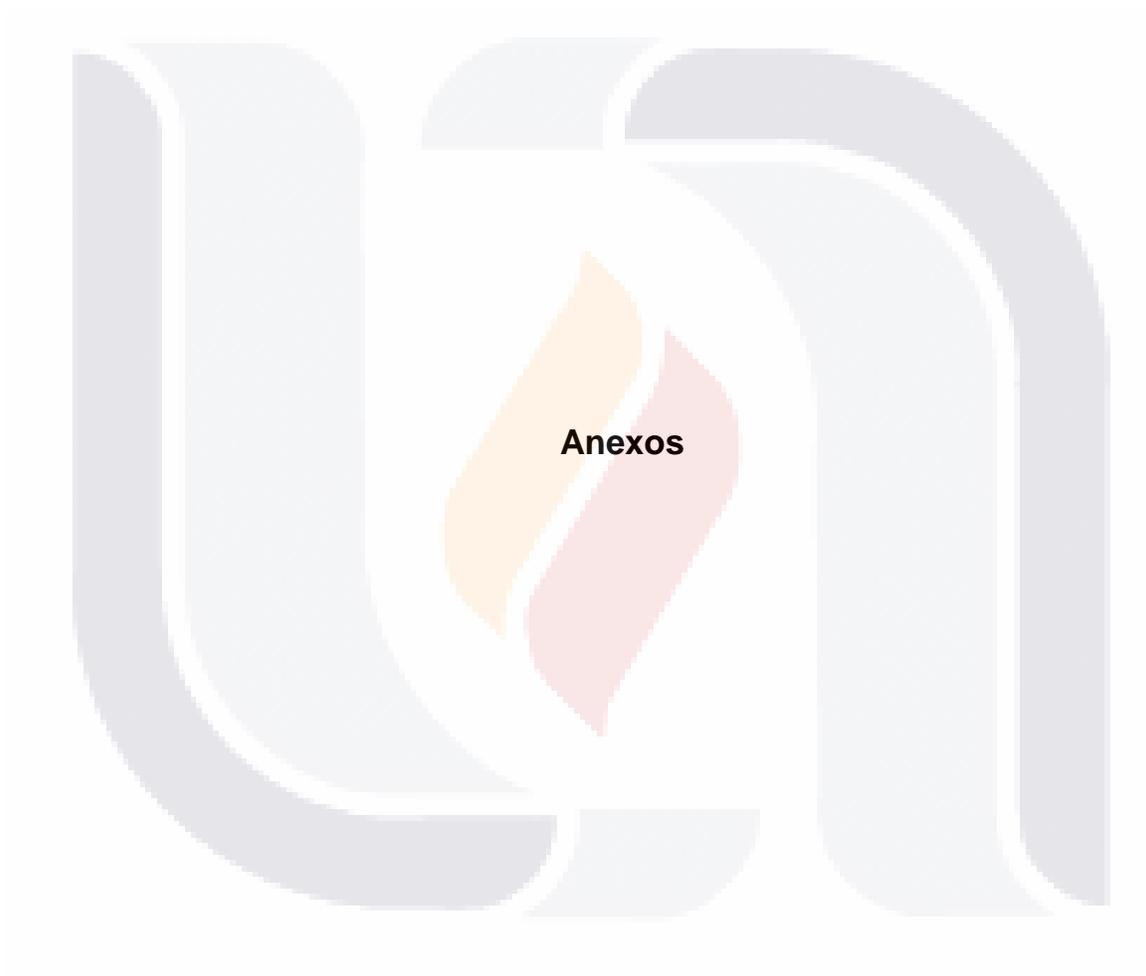
- delMas, R., Garfield, J. Ooms, A., y Chance, B. (2007) Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal*, 6(2), pp. 28-58
- Eudave, D. (2007). El aprendizaje de la estadística en estudiantes universitarios de profesiones no matemáticas. *Educación Matemática*, 19 (2), pp. 41-66
- Friel, S. N., Curcio, F. R. and Bright, G. W. (2001). Making sense of graphs: critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, Vol. 32, No. 2, pp. 124-158.
- Gal, I. (2002). Adults' Statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, 70, 1-51.
- Garfield, J. (2002). The Challenge of Developing Statistical Reasoning. *Journal of Statistics Education*. 10 (3). Recuperado de www.amstat.org/publications/jse/v10n3/garfield.html
- Garfield, J. (2003). Assessing statistical reasoning. *Statistics Education Research Journal*, 2(1), 22-38.
- Garfield, J., delMas B., Chance, B., Poly, C. y Ooms, A. (2005). Comprehensive Assessment of Outcomes for a first course in Statistics (CAOS). *Web ARTIST Project*. Consultado el 3 de octubre de 2011 en: <http://app.gen.umn.edu/artist/>
- Garfield, J., delMas B., Chance, B., Poly, C. y Ooms, A. (2006). Summary Statistics for a National Sample of Undergraduates. *Web ARTIST Project*. Consultado el 3 de octubre de 2011 en: <http://app.gen.umn.edu/artist/>
- Goldman, R. y McKenzie, J. (2006). Questions to assess the understanding of statistical concepts. *International Association of Statistical Education*. Recuperado el 26 de noviembre de 2011, en la base de datos cots. Net
- Hernandez, R., Fernández, C. y Baptista P. (2006) Metodología de la investigación. México: The McGraw-Hill.
- Inzunza, S. y Jiménez, J., (2011). Caracterización del razonamiento y pensamiento Estadístico en Estudiantes universitarios: el caso de pruebas de hipótesis. *Trabajo presentado en XIV edición de la Escuela de Invierno en Matemática Educativa*. Recuperado de http://webcache.googleusercontent.com/search?q=cache:http://matematicas.redua.mx/home/eventos/2011/esc_invierno/Programa_Completo.pdf

- Izar, J., Ynzunza, C y López, H., (2011). Factores que afectan el desempeño académico de los estudiantes de nivel superior en Rioverde, San Luis Potosí, México. *Revista de Investigación Educativa*, 12. Consultado el 17 de septiembre de 2013 en: <http://www.uv.mx/cpue/num12/opinion/completos/izar-desempeno%20academico.html>
- Kaplan, J., Fisher, D. y Rogness, T. (2010). Lexical Ambiguity in Statistics: How students use and define the words: association, average, confidence, random and spread. *Journal of Statistics Education*, 18, (2), pp. 1-22.
- Kelly, A., Sloane, F y Whittaker, A. (1997). Simple Approaches to Assessing Underlying Understanding of Statistical Concepts. En Gal, I. y Garfield, J. (Eds). *The Assessment Challenge in Statistics Education* (pp. 85-90). Recuperado de <http://www.stat.auckland.ac.nz/~iase/publications/assessbkref>.
- May, R. (2009). *La representación gráfica en estadística a nivel superior: un análisis de libros de texto en Psicología y educación* (Tesis maestría). Universidad Autónoma de Yucatán. Mérida de Yucatán.
- Morales, P. (2007). La fiabilidad de los test y escalas. *Universidad Pontificia Comillas* (1-37). Recuperado de <http://www.upcomillas.es/personal/peter/estadisticabasica/Fiabilidad.pdf>
- Ottaviani, M. G. (1998). Developments and perspectives in statistical education. Ponencia invitada en la Sesión: Statistics for Monitoring Educational Systems. Publicado en los Proceedings of the Joint IASS/IAOS Conference. Statistics for Economic and Social Development. Aguascalientes, México
- Pagano, R. (1999). *Estadística para las ciencias del comportamiento*. México: International Thomson Editores, S.A. de C.V.
- Pérez, J., Acuña, N. y Arratia, E. (2008). Nivel de dificultad y poder de discriminación del tercer y quinto examen parcial de la cátedra de cito-histología 2007 de la carrera de medicina de la UMSA. *Revista-Cuadernos* 53 (2). Consultado el 1 de febrero de 2013 en: <http://www.revistasbolivianas.org.bo/pdf/chc/v53n2/v53n2a03.pdf>
- 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, Salamanca, España.

- Pinto, J. (2011). Conocimiento del profesor sobre los errores y dificultades de los estudiantes en el aprendizaje de la representación gráfica en Estadística. *Memoria Electrónica del XI Congreso Nacional de Investigación Educativa*, Noviembre, México.
- Reading, C., (2002) Profile for statistical understanding. *International Association of Statistical Education*, 6, pp. 1-6
- Rumsey, J. (2002). Statistical Literacy as a Goal for Introductory Statistics Courses. *Journal of Statistics Education* 10 (3). Recuperado de www.amstat.org/publications/jse/v10n3/rumsey2.html
- Salazar, B. (2008). *Comprensión de la información estadística en estudiantes universitarios*, (Tesis de la Maestría) Universidad Autónoma de Aguascalientes, Aguascalientes, México
- Shamos, M.H. (1995). *The myth of scientific literacy*. New Brunswick, NJ: Rutgers University Press.
- Shi. N., He, X. y Tao. J. (2009). Understanding Statistics And Statistics Education: A Chinese Perspective. *Journal Of Statistics Education*, 17, (3).
- Schaun y Mattern (1997). Assessing Students' Connected Understanding of Statistical Relationships. En Gal, I. y Garfield, J. (eds.) *The Assessment Challenge in Statistics Education* (pp. 91-104). Recuperado de <http://www.stat.auckland.ac.nz/~iase/publications/assessbkref>.
- Vere-Jones, D. (1995). The coming age of statistical education. *International Statistical Review*, 63(1), pp. 3-23.
- Watson. J., (1997). Assessing Statistical Thinking Using the Media. En I. Gal. y J. Garfield (editors). *The Assessment Challenge in Statistics Education*. (pp. 107-121). International Statistical Institute.
- Wild, J., y Pfannkuch, M., (1999). Statistical Thinking in Empirical Enquiry. *International Statistical Review*, 67, (3), pp. 233-265

Instituciones y Centros

- Asociación Mexicana de Estadística (2013). *AME*. Recuperado el 1 de mayo de 2013 de:
<http://www.amestad.mx/index.php>
- CENEVAL. (2008) *Examen Transversal por Campo de Conocimiento para el nivel Licenciatura-Estadística (ExTra-Es)* Recuperado el 29 de abril de 2013 de:
<http://www.ceneval.edu.mx/ceneval-web/content.do?page=3377#exam04>
- Red CIMATES A.C. (2012). *Red de Centros de Investigación en Matemática Educativa A. C.* Recuperado el 29 de abril de 2013 de: <http://www.red-cimates.org.mx/>
- Red de Investigación y Educación en Probabilidad y Estadística. (2012) *RIEPE*. Recuperado el 29 de abril de 2013 de: <http://red.amestad.mx/origen.php>.
- Sociedad Argentina de estadística (2013). *SAE*. Recuperado el 1 de mayo de 2013 de:
http://www.s-a-e.org.ar/sae/index.php?option=com_content&view=article&id=87&Itemid=157
- Statistics2013. (2013). *Año Internacional de la Estadística*. Recuperado el 25 de noviembre de 2012 de: <http://www.estadistica2013cimat.mx/boletin01>
- University of Minnessota (2013). *AIMS Project Adapting and Implementing Innovative Material in Statistics*. Recuperado el 14 de marzo de 2013 de
<http://www.tc.umn.edu/~aims/aimsttopics.htm>



CAOS -4

FOLIO

Anexo A. Cuestionario de opinión

Cuestionario Estudiante

Por favor, contesta con tranquilidad todas las preguntas de este cuestionario. Debes responder siempre de acuerdo a tu experiencia o manera de pensar. Tus respuestas serán secretas. Lee atentamente cada pregunta y marca con una equis (X) en el cuadro que corresponda a tu respuesta.

1. ¿Eres hombre o mujer? *(Marca con una equis (X) una sola alternativa).*

 Hombre

 Mujer

2. ¿Cuántos años tienes? *(Escribe el número de años, por ejemplo, 20).*

 Años

3. ¿Qué licenciatura estudias? *(Marca sólo UNA casilla)*

1. Lic. En Asesoría Psicopedagógica	
2. Lic. En ciencias Políticas y Administración Pública.	
3. Lic. En Comunicación e Información.	
4. Lic. En Comunicación Organizacional.	
5. Lic. En el idioma inglés.	
6. Lic. En Historia.	
7. Lic. En Psicología.	
8. Lic. En Sociología.	
9. Lic. En Trabajo Social.	

4. Indica que semestre cursas actualmente: *(Escribe el número de semestre, por ejemplo, 8)*

Semestre

	Muy en desacuerdo	En desacuerdo	De acuerdo	Muy de acuerdo
1. Los cursos de estadística que he recibido en la universidad me han gustado.				
2. Me interesa mucho aprender estadística.				
3. Los cursos de estadística ayudan a mi formación profesional.				
4. Pienso que la estadística es difícil de aprender.				
5. Considero que se mucho de estadística.				
6. Me gustaría recibir más cursos de estadística durante mi formación.				
7. Tengo los suficientes conocimientos estadísticos requeridos para ejercer mi profesión.				
8. Los cursos de estadística que recibí estaban enfocados a mi carrera.				

¡Muchas gracias por tu ayuda al contestar este cuestionario!

Pensando en tus cursos de estadística, ¿qué tan de acuerdo o en desacuerdo estás con cada una de las siguientes afirmaciones? *(Marca con una equis (X) UNA sola alternativa para cada afirmación).*

CAOS -4

FOLIO

HOJA DE RESPUESTAS

1	a	b	C	d	
2	a	b	C		
3	a	b	C	d	
4	a	b	C	d	
5	a	b	C	d	
6	a	b	C	d	
7	a	b	C	d	e
8	a	b	C	d	
9	a	b	C	d	
10	a	b	C	d	
11	a	b			
12	a	b			
13	a	b			
14	a	b	C	d	e
15	a	b	C	d	e
16	a	b	C	d	e
17	a	b	C	d	
18	a	b	C		
19	a	b	C		
20	a	b	C		

21	A	b	c	
22	A	b	c	
23	A	b	c	
24	A	b	c	d
25	A	b		
26	A	b		
27	A	b		
28	A	b		
29	A	b		
30	A	b		
31	A	b		
32	A	b	c	
33	A	b	c	
34	A	b	c	
35	A	b	c	
36	A	b	c	
37	A	b	c	d
38	A	b	c	d
39	A	b	c	d
40	A	b	c	d

Escribe los conceptos que no entiendes	
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	

Anexo B. Prueba CAOS-4

**Comprehensive Assessment of Outcomes
for a first course in Statistics (CAOS)**

CAOS 4

**Developed by the Web ARTIST Project
<https://app.gen.umn.edu/artist/>**

**Funded by a grant from the National Science Foundation
NSF CCLI ASA- 0206571**

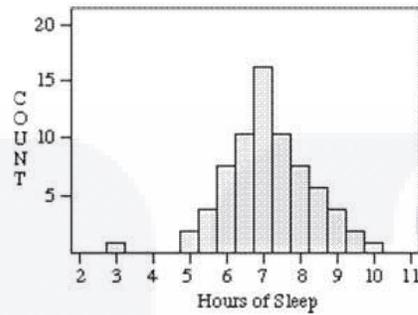
**Principal Investigators:
Joan Garfield and Bob delMas, University of Minnesota
Beth Chance, Cal Poly – San Luis Obispo
Post-doctoral Research Assistant:
Ann Ooms, University of Minnesota**

Version 31

September 8, 2005

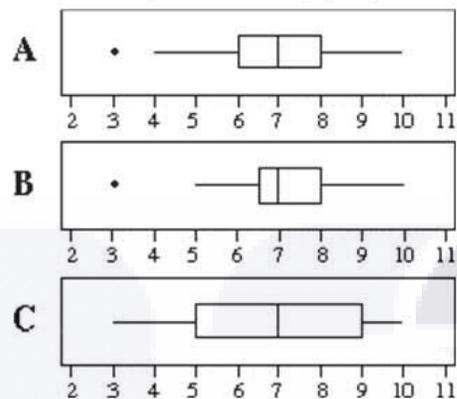
ARTIST CAOS 4 POSTTEST

The following graph shows a distribution of hours slept last night by a group of college students.



1. Select the statement below that gives the most complete description of the graph in a way that demonstrates an understanding of how to statistically describe and interpret the distribution of a variable.
 - a. The bars go from 3 to 10, increasing in height to 7, then decreasing to 10. The tallest bar is at 7. There is a gap between three and five.
 - b. The distribution is normal, with a mean of about 7 and a standard deviation of about 1.
 - c. Most students seem to be getting enough sleep at night, but some students slept more and some slept less. However, one student must have stayed up very late and got very few hours of sleep.
 - d. The distribution of hours of sleep is somewhat symmetric and bell-shaped, with an outlier at 3. The typical amount of sleep is about 7 hours and overall range is 7 hours.

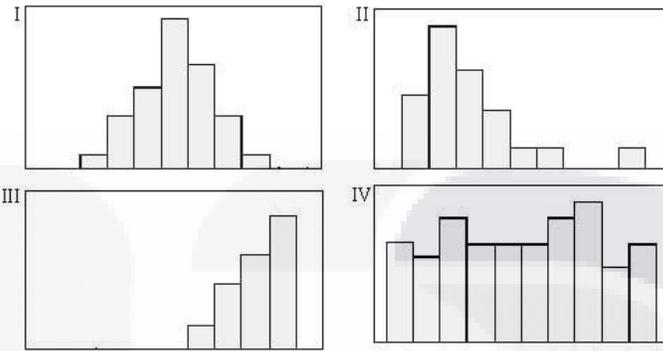
2. Which box plot seems to be graphing the same data as the histogram in question 1?



- a. Boxplot A.
- b. Boxplot B.
- c. Boxplot C.

Items 3 to 5 refer to the following situation:

Four histograms are displayed below. For each item, match the description to the appropriate histogram.



3. A distribution for a set of quiz scores where the quiz was very easy is represented by:
 - a. Histogram I.
 - b. Histogram II.
 - c. Histogram III.
 - d. Histogram IV.

4. A distribution for a set of wrist circumferences (measured in centimeters) taken from the right wrist of a random sample of newborn female infants is represented by:
 - a. Histogram I.
 - b. Histogram II.
 - c. Histogram III.
 - d. Histogram IV.

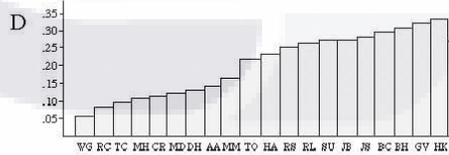
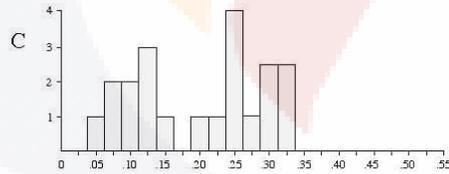
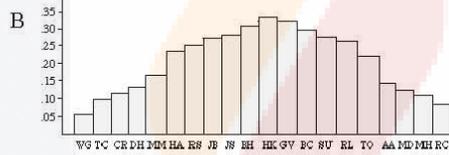
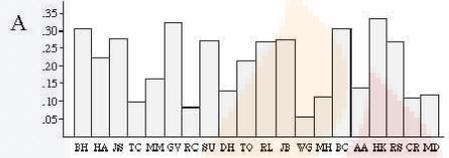
5. A distribution for the last digit of phone numbers sampled from a phone book (i.e., for the phone number 968-9667, the last digit, 7, would be selected) is represented by:
 - a. Histogram I.
 - b. Histogram II.
 - c. Histogram III.
 - d. Histogram IV.

6. A baseball fan likes to keep track of statistics for the local high school baseball team. One of the statistics she recorded is the proportion of hits obtained by each player based on the number of times at bat as shown in the table below. Which of the following graphs gives the best display of the distribution of proportion of hits in that it allows the baseball fan to describe the shape, center and spread of the variable, proportion of hits?

Player	Proportion of hits
BH	0.305
HA	0.229
JS	0.281
TC	0.097
MM	0.167
GV	0.333
RC	0.085

Player	Proportion of hits
SU	0.270
DH	0.136
TO	0.218
RL	0.267
JB	0.270
WG	0.054
MH	0.108

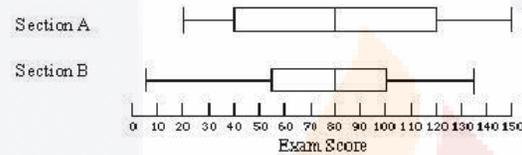
Player	Proportion of hits
BC	0.301
AA	0.143
HK	0.341
RS	0.261
CR	0.115
MD	0.125



7. A recent research study randomly divided participants into groups who were given different levels of Vitamin E to take daily. One group received only a placebo pill. The research study followed the participants for eight years to see how many developed a particular type of cancer during that time period. Which of the following responses gives the best explanation as to the purpose of randomization in this study?
- To increase the accuracy of the research results.
 - To ensure that all potential cancer patients had an equal chance of being selected for the study.
 - To reduce the amount of sampling error.
 - To produce treatment groups with similar characteristics.
 - To prevent skewness in the results.

Items 8 to 10 refer to the following situation:

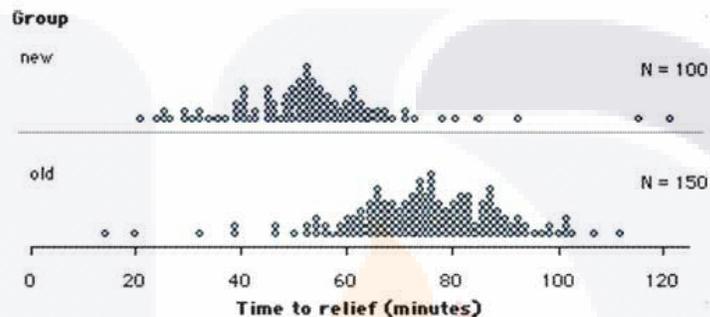
The two boxplots below display final exam scores for all students in two different sections of the same course.



8. Which section would you expect to have a greater standard deviation in exam scores?
- Section A.
 - Section B.
 - Both sections are about equal.
 - It is impossible to tell.
9. Which data set has a greater percentage of students with scores at or below 30?
- Section A.
 - Section B.
 - Both sections are about equal.
 - It is impossible to tell.
10. Which section has a greater percentage of students with scores at or above 80?
- Section A.
 - Section B.
 - Both sections are about equal.

Items 11 to 13 refer to the following situation:

A drug company developed a new formula for their headache medication. To test the effectiveness of this new formula, 250 people were randomly selected from a larger population of patients with headaches. 100 of these people were randomly assigned to receive the new formula medication when they had a headache, and the other 150 people received the old formula medication. The time it took, in minutes, for each patient to no longer have a headache was recorded. The results from both of these clinical trials are shown below. Items 11, 12, and 13 present statements made by three different statistics students. For each statement, indicate whether you think the student's conclusion is valid.



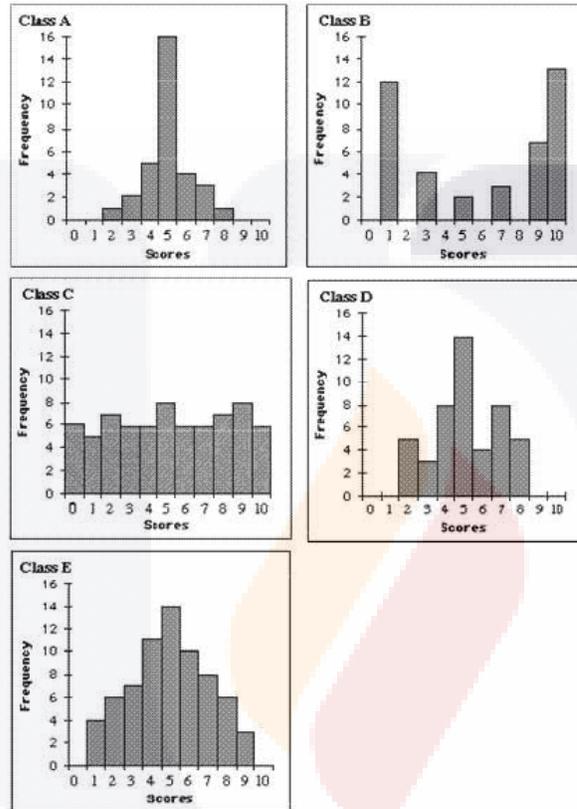
11. The old formula works better. Two people who took the old formula felt relief in less than 20 minutes, compared to none who took the new formula. Also, the worst result - near 120 minutes - was with the new formula.
 - a. Valid.
 - b. Not valid.

12. The average time for the new formula to relieve a headache is lower than the average time for the old formula. I would conclude that people taking the new formula will tend to feel relief about 20 minutes sooner than those taking the old formula.
 - a. Valid.
 - b. Not valid.

13. I would not conclude anything from these data. The number of patients in the two groups is not the same so there is no fair way to compare the two formulas.
 - a. Valid.
 - b. Not valid.

Items 14 and 15 refer to the following situation:

Five histograms are presented below. Each histogram displays test scores on a scale of 0 to 10 for one of five different statistics classes.



14. Which of the classes would you expect to have the lowest standard deviation, and why?
- Class A, because it has the most values close to the mean.
 - Class B, because it has the smallest number of distinct scores.
 - Class C, because there is no change in scores.
 - Class A and Class D, because they both have the smallest range.
 - Class E, because it looks the most normal.

15. Which of the classes would you expect to have the highest standard deviation, and why?
 - a. Class A, because it has the largest difference between the heights of the bars.
 - b. Class B, because more of its scores are far from the mean.
 - c. Class C, because it has the largest number of different scores.
 - d. Class D, because the distribution is very bumpy and irregular.
 - e. Class E, because it has a large range and looks normal.

16. A certain manufacturer claims that they produce 50% brown candies. Sam plans to buy a large family size bag of these candies and Kerry plans to buy a small fun size bag. Which bag is more likely to have more than 70% brown candies?
 - a. Sam, because there are more candies, so his bag can have more brown candies.
 - b. Sam, because there is more variability in the proportion of browns among larger samples.
 - c. Kerry, because there is more variability in the proportion of browns among smaller samples.
 - d. Kerry, because most small bags will have more than 50% brown candies.
 - e. Both have the same chance because they are both random samples.

17. Imagine you have a barrel that contains thousands of candies with several different colors. We know that the manufacturer produces 35% yellow candies. Five students each take a random sample of 20 candies, one at a time, and record the percentage of yellow candies in their sample. Which sequence below is the most plausible for the percent of yellow candies obtained in these five samples?
 - a. 30%, 35%, 15%, 40%, 50%.
 - b. 35%, 35%, 35%, 35%, 35%.
 - c. 5%, 60%, 10%, 50%, 95%.
 - d. Any of the above.

18. Jean lives about 10 miles from the college where she plans to attend a 10-week summer class. There are two main routes she can take to the school, one through the city and one through the countryside. The city route is shorter in miles, but has more stoplights. The country route is longer in miles, but has only a few stop signs and stoplights. Jean sets up a randomized experiment where each day she tosses a coin to decide which route to take that day. She records the following data for 5 days of travel on each route.

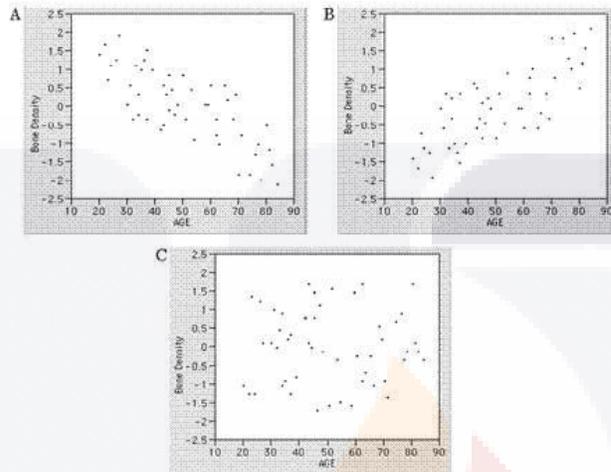
Country Route - 17, 15, 17, 16, 18

City Route - 18, 13, 20, 10, 16

It is important to Jean to arrive on time for her classes, but she does not want to arrive too early because that would increase her parking fees. Based on the data gathered, which route would you advise her to choose?

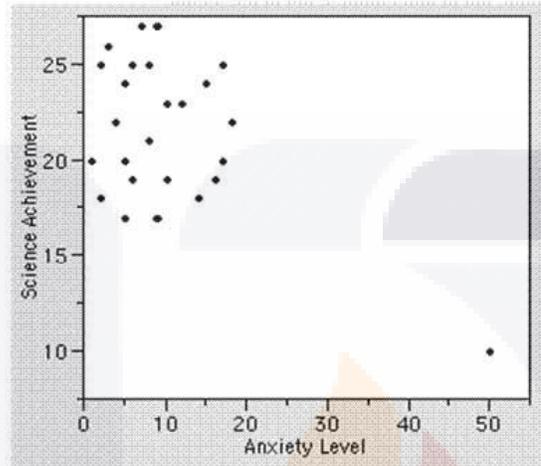
- a. The Country Route, because the times are consistently between 15 and 18 minutes.
 - b. The City Route, because she can get there in 10 minutes on a good day and the average time is less than for the Country Route.
 - c. Because the times on the two routes have so much overlap, neither route is better than the other. She might as well flip a coin.
19. A graduate student is designing a research study. She is hoping to show that the results of an experiment are statistically significant. What type of p -value would she want to obtain?
- a. A large p -value.
 - b. A small p -value.
 - c. The magnitude of a p -value has no impact on statistical significance.

20. Bone density is typically measured as a standardized score with a mean of 0 and a standard deviation of 1. Lower scores correspond to lower bone density. Which of the following graphs shows that as women grow older they tend to have lower bone density?



- a. Graph A.
- b. Graph B.
- c. Graph C.

21. The following scatterplot shows the relationship between scores on an anxiety scale and an achievement test for science. Choose the best interpretation of the relationship between anxiety level and science achievement based on the scatterplot.



- a. This graph shows a strong negative linear relationship between anxiety and achievement in science.
 - b. This graph shows a moderate linear relationship between anxiety and achievement in science.
 - c. This graph shows very little, if any, linear relationship between anxiety and achievement in science.
22. Researchers surveyed 1,000 randomly selected adults in the U.S. A statistically significant, strong positive correlation was found between income level and the number of containers of recycling they typically collect in a week. Please select the best interpretation of this result.
- a. We can not conclude whether earning more money causes more recycling among U.S. adults because this type of design does not allow us to infer causation.
 - b. This sample is too small to draw any conclusions about the relationship between income level and amount of recycling for adults in the U.S.
 - c. This result indicates that earning more money influences people to recycle more than people who earn less money.

Items 23 and 24 refer to the following situation:

A researcher in environmental science is conducting a study to investigate the impact of a particular herbicide on fish. He has 60 healthy fish and randomly assigns each fish to either a treatment or a control group. The fish in the treatment group showed higher levels of the indicator enzyme.

23. Suppose a test of significance was correctly conducted and showed no statistically significant difference in average enzyme level between the fish that were exposed to the herbicide and those that were not. What conclusion can the graduate student draw from these results?
 - a. The researcher must not be interpreting the results correctly; there should be a significant difference.
 - b. The sample size may be too small to detect a statistically significant difference.
 - c. It must be true that the herbicide does not cause higher levels of the enzyme.

24. Suppose a test of significance was correctly conducted and showed a statistically significant difference in average enzyme level between the fish that were exposed to the herbicide and those that were not. What conclusion can the graduate student draw from these results?
 - a. There is evidence of association, but no causal effect of herbicide on enzyme levels.
 - b. The sample size is too small to draw a valid conclusion.
 - c. He has proven that the herbicide causes higher levels of the enzyme.
 - d. There is evidence that the herbicide causes higher levels of the enzyme for these fish.

Items 25 to 27 refer to the following situation:

A research article reports the results of a new drug test. The drug is to be used to decrease vision loss in people with Macular Degeneration. The article gives a p -value of .04 in the analysis section. Items 25, 26, and 27 present three different interpretations of this p -value. Indicate if each interpretation is valid or invalid.

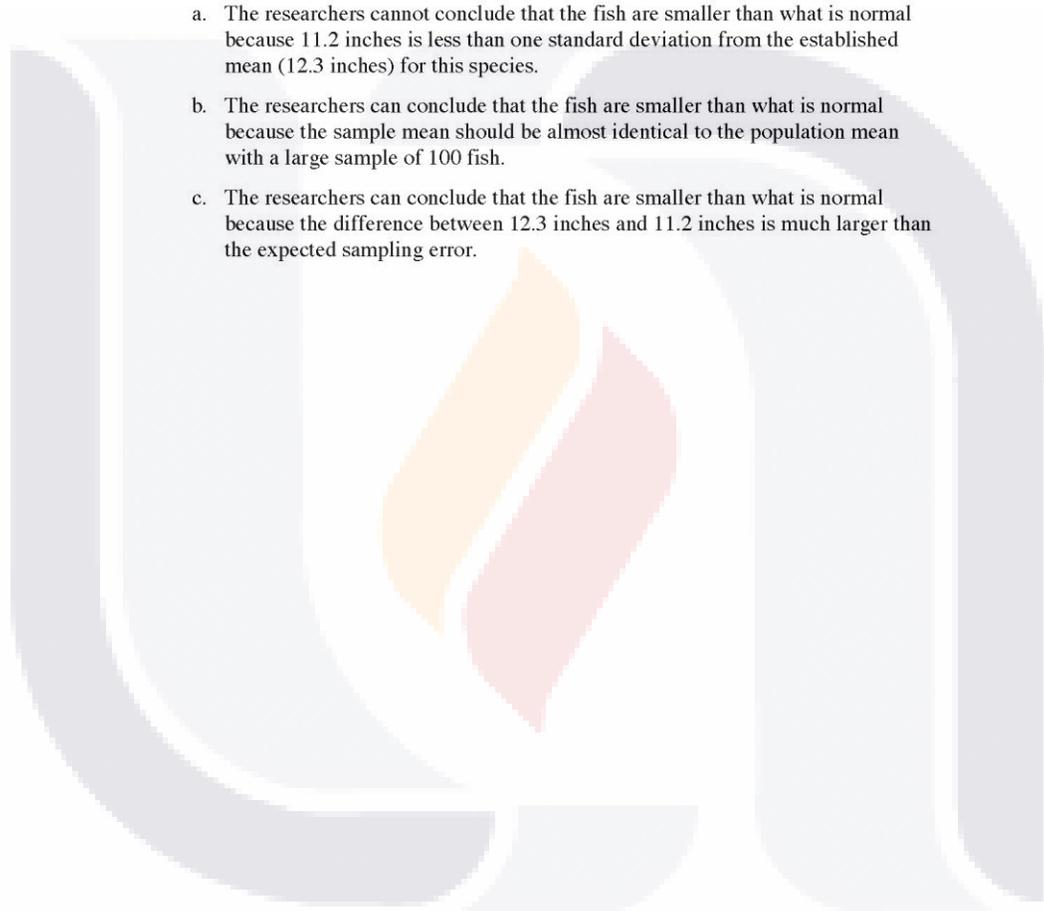
25. The probability of getting results as extreme as or more extreme than the ones in this study if the drug is actually not effective.
 - a. Valid.
 - b. Invalid.
26. The probability that the drug is not effective.
 - a. Valid.
 - b. Invalid.
27. The probability that the drug is effective.
 - a. Valid.
 - b. Invalid.

Items 28 to 31 refer to the following situation:

A high school statistics class wants to estimate the average number of chocolate chips in a generic brand of chocolate chip cookies. They collect a random sample of cookies, count the chips in each cookie, and calculate a 95% confidence interval for the average number of chips per cookie (18.6 to 21.3). Items 28, 29, and 30 present four different interpretations of these results. Indicate if each interpretation is valid or invalid.

28. We are 95% certain that each cookie for this brand has approximately 18.6 to 21.3 chocolate chips.
 - a. Valid.
 - b. Invalid.
29. We expect 95% of the cookies to have between 18.6 and 21.3 chocolate chips.
 - a. Valid.
 - b. Invalid.
30. We would expect about 95% of all possible sample means from this population to be between 18.6 and 21.3 chocolate chips.
 - a. Valid.
 - b. Invalid.
31. We are 95% certain that the confidence interval of 18.6 to 21.3 includes the true average number of chocolate chips per cookie.
 - a. Valid.
 - b. Invalid.

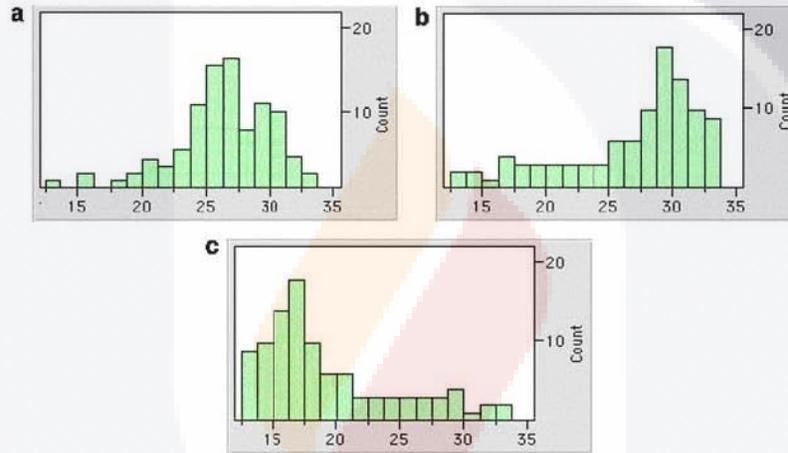
32. It has been established that under normal environmental conditions, adult largemouth bass in Silver Lake have an average length of 12.3 inches with a standard deviation of 3 inches. People who have been fishing Silver Lake for some time claim that this year they are catching smaller than usual largemouth bass. A research group from the Department of Natural Resources took a random sample of 100 adult largemouth bass from Silver Lake and found the mean of this sample to be 11.2 inches. Which of the following is the most appropriate statistical conclusion?
- The researchers cannot conclude that the fish are smaller than what is normal because 11.2 inches is less than one standard deviation from the established mean (12.3 inches) for this species.
 - The researchers can conclude that the fish are smaller than what is normal because the sample mean should be almost identical to the population mean with a large sample of 100 fish.
 - The researchers can conclude that the fish are smaller than what is normal because the difference between 12.3 inches and 11.2 inches is much larger than the expected sampling error.



A study examined the length of a certain species of fish from one lake. The plan was to take a random sample of 100 fish and examine the results. Numerical summaries on lengths of the fish measured in this study are given.

Mean	26.8mm
Median	29.4mm
Standard Deviation	5.0 mm
Minimum	12.m mm
Maximum	33.4mm

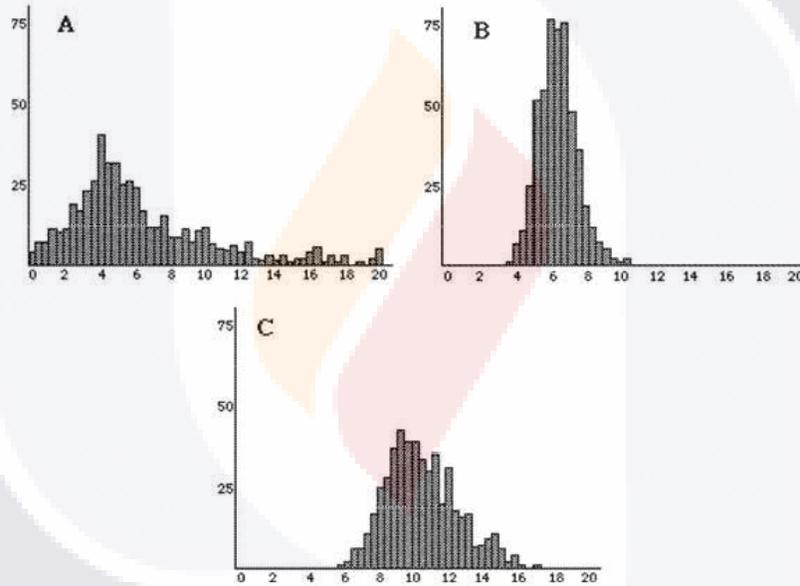
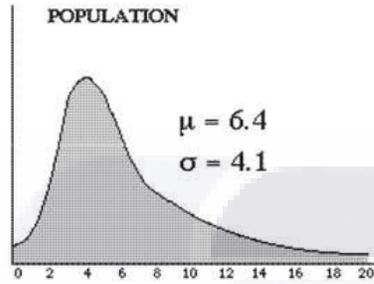
33. Which of the following histograms is most likely to be the one for these data?



- a. Histogram a.
- b. Histogram b.
- c. Histogram c.

Items 34 and 35 refer to the following situation:

Four graphs are presented below. The graph at the top is a distribution for a population of test scores. The mean score is 6.4 and the standard deviation is 4.1.



34. Which graph (A, B, or C) do you think represents a single random sample of 500 values from this population?
- Graph A
 - Graph B
 - Graph C
35. Which graph (A, B, or C) do you think represents a distribution of 500 sample means from random samples each of size 9?
- Graph A
 - Graph B
 - Graph C
36. This table is based on records of accidents compiled by a State Highway Safety and Motor Vehicles Office. The Office wants to decide if people are less likely to have a fatal accident if they are wearing a seatbelt. Which of the following comparisons is most appropriate for supporting this conclusion?

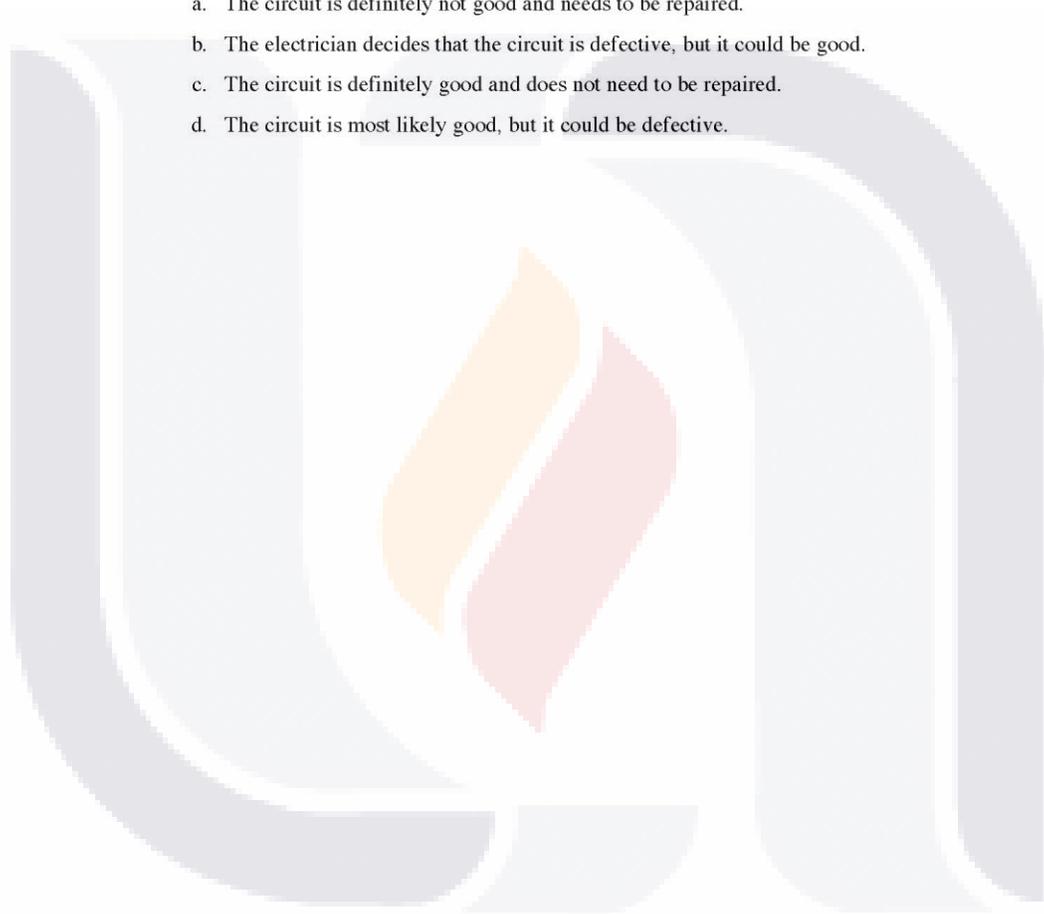
Safety Equipment in Use	Injury		ROW TOTAL
	Nonfatal	Fatal	
Seat Belt	412,368	510	412,878
No Seat Belt	162,527	1,601	164,128
COLUMN TOTAL	574,895	2,111	577,006

- Compare the ratios $510/412,878$ and $1,601/164,128$
- Compare the ratios $510/577,006$ and $1,601/577,006$
- Compare the numbers 510 and 1,601

37. A student participates in a Coke versus Pepsi taste test. She correctly identifies which soda is which four times out of six tries. She claims that this proves that she can reliably tell the difference between the two soft drinks. You have studied statistics and you want to determine the probability of anyone getting at least four right out of six tries just by chance alone. Which of the following would provide an accurate estimate of that probability?
- Have the student repeat this experiment many times and calculate the percentage time she correctly distinguishes between the brands.
 - Simulate this on the computer with a 50% chance of guessing the correct soft drink on each try, and calculate the percent of times there are four or more correct guesses out of six trials.
 - Repeat this experiment with a very large sample of people and calculate the percentage of people who make four correct guesses out of six tries.
 - All of the methods listed above would provide an accurate estimate of the probability.
38. A college official conducted a survey to estimate the proportion of students currently living in dormitories about their preference for single rooms, double rooms, or multiple (more than two people) rooms in the dormitories on campus. Which of the following does NOT affect the college official's ability to generalize the survey results to all dormitory students?
- Five thousand students live in dormitories on campus. A random sample of only 500 were sent the survey.
 - The survey was sent to only first-year students.
 - Of the 500 students who were sent the survey, only 160 responded.
 - All of the above present a problem for generalizing the results.
39. The number of people living on American farms has declined steadily during the last century. Data gathered on the U.S. farm population (millions of people) from 1910 to 2000 were used to generate the following regression equation: Predicted Farm Population = $1167 - .59(\text{YEAR})$. Which method is best to use to predict the number of people living on farms in 2050?
- Substitute the value of 2050 for YEAR in the regression equation, and compute the predicted farm population.
 - Plot the regression line on a scatterplot, locate 2050 on the horizontal axis, and read off the corresponding value of population on the vertical axis.
 - Neither method is appropriate for making a prediction for the year 2050 based on these data.
 - Both methods are appropriate for making a prediction for the year 2050 based on these data.

40. The following situation models the logic of a hypothesis test. An electrician uses an instrument to test whether or not an electrical circuit is defective. The instrument sometimes fails to detect that a circuit is good and working. The null hypothesis is that the circuit is good (not defective). The alternative hypothesis is that the circuit is not good (defective). If the electrician rejects the null hypothesis, which of the following statements is true?

- a. The circuit is definitely not good and needs to be repaired.
- b. The electrician decides that the circuit is defective, but it could be good.
- c. The circuit is definitely good and does not need to be repaired.
- d. The circuit is most likely good, but it could be defective.



CAOS 4 ANSWER KEY

1. D	11. B	21. C	31. A
2. B	12. A	22. A	32. C
3. C	13. B	23. B	33. B
4. A	14. A	24. D	34. A
5. D	15. B	25. A	35. B
6. C	16. C	26. B	36. A
7. D	17. A	27. B	37. B
8. A	18. A	28. B	38. A
9. D	19. B	29. B	39. C
10. C	20. A	30. B	40. B

Anexo C. Extra-Es

AREA BÁSICA		
I. Fundamentos del pensamiento estadístico y estructura y generación de datos.	II. Descripción, organización e interpretación de datos	III. Nociones de inferencia
<p>1.1. Propósitos y usos de la estadística</p> <ul style="list-style-type: none"> • Principios de la estadística descriptiva • Noción de inferencia estadística • Estadística para la toma de decisiones • Síntesis de datos para la generación de conocimiento <p>1.2. Conceptos y definiciones elementales</p> <ul style="list-style-type: none"> • Definición de población, muestra y unidad de estudio • Tipos de variables y escalas de medición • Frecuencias, proporciones y porcentajes, en un contexto estadístico • Medidas descriptivas de tendencia central, de dispersión y percentiles <p>1.3. Origen de los datos</p> <ul style="list-style-type: none"> • Estudio observacional vs. estudio experimental • Población objetivo (unidad de muestreo, variables respuesta y parámetros) • Muestreo aleatorio y no aleatorio 	<p>2.1. Procedimientos según tipo de variable</p> <ul style="list-style-type: none"> • Estadísticos, tablas y diagramas <p>2.2. Una variable categórica: estadísticos, tablas y diagramas</p> <ul style="list-style-type: none"> • Cálculo de frecuencias, proporciones y porcentajes para variables categóricas • Interpretación de tablas de frecuencias y porcentajes (simples o acumulados) para variables categóricas ordinales • Interpretación de gráficas de barras, circulares y pictogramas <p>2.3. Una variable cuantitativa: estadísticos, tablas y diagramas</p> <ul style="list-style-type: none"> • Cálculo de medidas de tendencia central, dispersión y percentiles • Interpretación de medidas de tendencia central, dispersión y percentiles • Interpretación de histogramas y diagramas de cajas <p>2.4. Dos variables categóricas: tablas y gráficas</p>	<p>3.1. Procedimientos de estimación</p> <ul style="list-style-type: none"> • Estimación puntual y por intervalo • Estimación de parámetros de un modelo de regresión lineal simple <p>3.2. Conceptos fundamentales de prueba de hipótesis</p> <ul style="list-style-type: none"> • Tipos de hipótesis • Errores tipo I y tipo II, y la probabilidad de error tipo I • Definición del valor p <p>3.3. Pruebas de hipótesis con una muestra</p> <ul style="list-style-type: none"> • Uso del valor del estadístico para una prueba de hipótesis sobre una media de una variable con distribución normal e interpretación del resultado • Uso del valor del estadístico para una prueba de hipótesis sobre una proporción de una población con

	<ul style="list-style-type: none"> • Interpretación de tablas de contingencia • Interpretación de diagramas de barras de dos variables categóricas 2.5. Dos variables cuantitativas: estadísticos y diagramas • Interpretación de diagramas de dispersión • Cálculo e interpretación del coeficiente de correlación de Pearson 2.6. Una variable categórica y una variable cuantitativa: estadísticos y diagramas • Interpretación de tablas de medias y desviaciones estándar por categoría • Interpretación de diagramas de barras y de cajas por categoría 2.7. Datos atípicos • Identificación de datos atípicos • Medidas descriptivas adecuadas a datos atípicos 	<p>distribución binomial e interpretación del resultado</p> <p>3.4. Pruebas de hipótesis con dos muestras independientes</p> <ul style="list-style-type: none"> • Obtención del valor del estadístico para una prueba de hipótesis sobre la diferencia de medias con dos muestras independientes con varianzas iguales.
--	---	--

MÉTODOS ESTADÍSTICOS	
I. Inferencia paramétrica y no paramétrica	II. Análisis de varianza y de regresión
<p>1.1. Estimación</p> <ul style="list-style-type: none"> • Estimación por intervalo de medias y varianzas. y proporciones • Estimadores insesgados <p>1.2. Pruebas de hipótesis con una muestra</p> <ul style="list-style-type: none"> • Pruebas de hipótesis sobre varianzas y bondad de ajuste para normalidad <p>1.3. Pruebas de hipótesis con dos muestras independientes y pareadas</p> <ul style="list-style-type: none"> • Obtención del valor del estadístico para una prueba de hipótesis sobre la diferencia de medias con datos pareados <p>1.4. Estadística no paramétrica</p> <ul style="list-style-type: none"> • Correlación • Identificación del método no paramétrico alternativo a pruebas paramétricas • Comparación de dos poblaciones con variable ordinal • Interpretación de los resultados de una prueba de independencia para variables categóricas • Interpretación de los resultados de una prueba de independencia para variables categóricas. 	<p>2.1. Pruebas de hipótesis con más de dos muestras independientes</p> <ul style="list-style-type: none"> • Prueba de hipótesis sobre la igualdad de las medias para más de dos poblaciones mediante el Análisis de Varianza <p>2.2. Regresión</p> <ul style="list-style-type: none"> • Identificación de curvatura • Regresión lineal simple y múltiple

MUESTREO	
I. Fundamentos del muestreo	II. Descripción, organización e inferencia a partir de los datos en muestreo
<p>1.1. Conceptos, definiciones y características</p> <ul style="list-style-type: none"> • Marco muestral • Ventajas y limitantes de los muestreos • Preguntas para un cuestionarios de una encuesta • Base de datos para un muestreo • Existencia de estratos y conglomerados • Muestreo sistemático (con inicio aleatorio) <p>1.2. Origen de los datos</p> <ul style="list-style-type: none"> • Reconocimiento de fuentes potenciales de sesgo en muestreos aleatorios y no aleatorios • Reconocimiento de las características de las poblaciones finitas (contable) e infinitas (no contable) • Selección del método de muestreo de acuerdo con la estructura de la población • Cálculo del tamaño de una muestra aleatoria • Población objetivo, tipos de muestreo (aleatorio simple, estratificado con muestreo aleatorio simple dentro de los estratos, por conglomerados con muestreo aleatorio simple en ellos, y sistemático) 	<p>2.1. Estimación de parámetros en muestreo</p> <ul style="list-style-type: none"> • Estimación del total, del promedio o de la proporción poblacional y sus varianzas en un muestreo aleatorio simple de poblaciones finitas • Estimación del total, del promedio o de la proporción poblacional y sus varianzas en un muestreo estratificado de poblaciones finitas • Estimación del total, del promedio o de la proporción poblacional y sus varianzas en un muestreo por conglomerados de poblaciones finitas • Estimación del total, del promedio o de la proporción poblacional en un muestreo sistemático de poblaciones finitas <p>2.2. Interpretación de la estimación por intervalo para proporciones, promedios y totales en muestreo</p> <ul style="list-style-type: none"> • Interpretación de las estimaciones por intervalo de proporciones en muestreo • Interpretación de las estimaciones por intervalo de un promedio en diferentes tipos de muestreo • Interpretación de las estimaciones por intervalo de un total poblacional en diferentes tipos de muestreo

ESTADÍSTICA EXPERIMENTAL	
I. Fundamentos de estudios estadísticos y diseños de experimentos	II. Análisis de diseños experimentales
<p>1.1. Estudios experimentales y no experimentales</p> <ul style="list-style-type: none"> • Identificar factores que afectan a la validez interna o externa • Identificar fuentes o factores de confusión y su control <p>1.2. Unidad experimental</p> <ul style="list-style-type: none"> • Identificar y valorar la representatividad de la unidad experimental <p>1.3. Estudios experimentales</p> <ul style="list-style-type: none"> • Identificar fuentes de variación en un sistema experimental (factores y su naturaleza, nivel, tratamiento, efecto principal, efecto de interacción, unidad experimental, variables de respuesta) • Determinar la formación de bloques para control de factores de confusión en la investigación comparativa • Identificar los elementos que determinan el número de repeticiones (réplicas) 	<p>2.1. Análisis de diseños experimentales factoriales</p> <ul style="list-style-type: none"> • Identificar y analizar diseños experimentales unifactoriales • Identificar y analizar diseños experimentales factoriales con o sin interacción <p>2.2. Análisis de diseños experimentales: bloques al azar</p> <ul style="list-style-type: none"> • Identificar y analizar diseños experimentales: bloques al azar <p>2.3. Análisis de diseños experimentales: cuadro latino</p> <ul style="list-style-type: none"> • Identificar y analizar diseños experimentales: cuadro latino

Anexo D. Descripción de los ítems de CAOS-4

ITEM	RESULTADO DE APRENDIZAJE	DOMINIO
1	Capacidad para describir e interpretar el conjunto de distribuciones de una variable como se muestra en un histograma, incluyendo la referencia en el contexto de los datos.	Alfabetización
2	Capacidad para reconocer dos representaciones gráficas diferentes de los mismos datos (histograma y gráfica de cajas)	Alfabetización
3	Capacidad para visualizar un histograma para una descripción de una variable (distribución sesgada negativamente para las puntuaciones en una prueba fácil)	Alfabetización
4	Capacidad de visualizar y de acuerdo con un histograma una descripción de una variable (Una distribución de campana para circunferencias de muñecas de la mano en niñas recién nacidas.	Alfabetización
5	Capacidad de visualizar y de acuerdo con un histograma una descripción de una variable (distribución uniforme para el último dígito de los números de teléfono incluidos en la muestra a partir de un directorio telefónico.	Alfabetización
6	Comprender que para describir adecuadamente la distribución (la forma, el centro y la dispersión) de una variable cuantitativa, un gráfico como el histograma es necesario.	Alfabetización
7	Comprender el propósito de la aleatorización en un experimento.	Razonamiento
8	Capacidad para determinar cual de los dos diagramas de cajas representa la más grande desviación estándar.	Alfabetización
9	Comprender que el diagrama de cajas no ofrece estimaciones precisas a los porcentajes de datos por encima o por debajo de los valores excepto por los	Alfabetización

	cuartiles.	
10	Comprender la interpretación de mediana en el contexto de diagrama de cajas.	Alfabetización
11	Capacidad para comparar dos grupos considerando donde esta la mayoría de los datos y centrándose en la distribución como entidades individuales.	Pensamiento
12	Capacidad para comparar grupos mediante la comparación de las diferencias en promedios.	Pensamiento
13	Entender que comparar dos grupos no requiere muestras iguales en cada grupo, especialmente si ambos conjuntos de datos son grandes.	Pensamiento
14	Capacidad para estimar correctamente y comparar las desviaciones estándar para diferentes histogramas. Comprender que la más baja desviación estándar podría ser para un gráfico con la menor dispersión (típicamente).	Razonamiento
15	Capacidad para estimar correctamente las desviaciones estándar para diferentes histogramas. Comprender que la más alta desviación estándar puede ser para un gráfico con mayor dispersión (típicamente).	Razonamiento
16	Comprender que la estadística de pequeñas muestras varía más que las estadísticas de muestras grandes.	Pensamiento
17	Comprender los patrones esperados en la variabilidad de la muestra.	Pensamiento
18	La comprensión del significado de variabilidad en el contexto de medidas repetidas y en un contexto donde se desea la pequeña variabilidad.	Pensamiento
19	Comprender que los bajos valores de p son deseables en estudios de investigación.	Pensamiento
20	Capacidad para coincidir con un gráfico de dispersión a una descripción verbal de una relación bivariante.	Alfabetización

21	Capacidad para describir correctamente una relación bivariada mostrada en un diagrama de dispersión cuando hay un valor atípico.	Alfabetización
22	Comprender que la correlación no implica causación.	Alfabetización
23	Comprender que la significancia estadística no es garantía de que no hay ningún efecto.	Pensamiento
24	Comprender que un diseño experimental con asignación aleatoria admite la inferencia causal.	Pensamiento
25	Capacidad para reconocer una interpretación de un valor $-p$	Pensamiento
26	Capacidad para reconocer una interpretación incorrecta de un valor $-p$ (probabilidad de que el tratamiento no es efectivo)	Pensamiento
27	Capacidad de reconocer una interpretación incorrecta del valor $-p$ (problema del tratamiento efectivo)	Pensamiento
28	Capacidad para detectar errores de interpretación de un nivel de confianza (el porcentaje de datos en una muestra entre el limite de confianza)	Pensamiento
29	Capacidad para detectar una mala interpretación de un nivel de confianza (porcentaje de los datos de una población entre los límites de confianza.	Pensamiento
30	Capacidad para detectar errores de interpretación de un nivel de confianza (porcentaje del total del posible significado de muestra entre los límites de confianza.	Pensamiento
31	Habilidad para interpretar correctamente un intervalo de confianza.	Pensamiento
32	Comprender como el error de muestra es usado para hacer una inferencia informal acerca de una media muestral.	Pensamiento
33	Entender que una distribución con la mediana más grande que la media es más probable que se desvíe hacia la izquierda.	Alfabetización

34	La ley de comprensión de grandes cantidades para una muestra amplia mediante la selección de una muestra representativa de una población dada el tamaño de la muestra.	Pensamiento
35	Capacidad para seleccionar una distribución de muestreo apropiada para una población y el tamaño de la muestra.	Pensamiento
36	Comprender como calcular los coeficientes apropiadamente para encontrar la probabilidad condicional utilizando una tabla de datos.	Pensamiento
37	Comprender como simular datos para encontrar la probabilidad de un valor observado.	Pensamiento
38	Comprensión de los factores que permiten que una muestra de datos sea generalizada a la población.	Pensamiento
39	Comprender que no es conveniente extrapolar usando un modelo de regresión.	Pensamiento
40	Comprender la lógica de una prueba de significancia cuando la hipótesis nula es rechazada	Pensamiento

Anexo E. Alternativas de respuesta

*La respuesta correcta se indica con un **

Preguntas	Rendimiento	A	B	C	D	E
Ítem 1	Alto	2	15	7	*37	
	Bajo	8	17	11	*32	
Ítem 2	Alto	15	*36	10		
	Bajo	16	*24	28		
Ítem 3	Alto	6	3	*22	29	
	Bajo	7	6	*11	42	
Ítem 4	Alto	*20	7	11	22	
	Bajo	*13	19	15	19	
Ítem 5	Alto	9	25	8	*17	
	Bajo	22	25	25	*3	
Ítem 6	Alto	12	27	*4	14	
	Bajo	22	17	*5	20	
Ítem 7	Alto	8	13	16	*15	8
	Bajo	9	24	16	*10	8
Ítem 8	Alto	*32	17	4	7	
	Bajo	*19	25	10	11	
Ítem 9	Alto	17	24	5	*15	
	Bajo	21	31	4	*10	
Ítem 10	Alto	35	9	*12	5	
	Bajo	44	10	*9	4	
Ítem 11	Alto	8	*53			
	Bajo	34	*33			
Ítem 12	Alto	*50	11			
	Bajo	*40	27			
Ítem 13	Alto	23	*38			
	Bajo	48	*18			
Ítem 14	Alto	*16	4	25	8	6
	Bajo	*5	10	34	11	4
Ítem 15	Alto	9	*23	12	9	5
	Bajo	16	*22	14	5	9
Ítem 16	Alto	0	5	*6	2	48
	Bajo	7	12	*2	6	41
Ítem 17	Alto	*25	14	2	19	
	Bajo	*10	19	8	28	
Ítem 18	Alto	*46	11	4		
	Bajo	*34	16	17		
Ítem 19	Alto	16	*21	18		
	Bajo	22	*12	26		
Ítem 20	Alto	*44	9	8		
	Bajo	*38	20	8		

Continuación Anexo E

Preguntas	Rendimiento	A	B	C	D	E
Ítem 21	Alto	11	12	*38		
	Bajo	22	25	*17		
Ítem 22	Alto	*38	12	10		
	Bajo	*24	22	21		
Ítem 23	Alto	5	*34	21		
	Bajo	19	*30	15		
Ítem 24	Alto	7	8	15	*30	
	Bajo	10	17	18	*19	
Ítem 25	Alto	*29	28			
	Bajo	*19	42			
Ítem 26	Alto	33	*23			
	Bajo	43	*18			
Ítem 27	Alto	26	*31			
	Bajo	44	*17			
Ítem 28	Alto	41	*20			
	Bajo	40	*26			
Ítem 29	Alto	34	*27			
	Bajo	50	*18			
Ítem 30	Alto	37	*24			
	Bajo	47	*21			
Ítem 31	Alto	*39	22			
	Bajo	*34	33			
Ítem 32	Alto	25	11	*25		
	Bajo	19	29	*15		
Ítem 33	Alto	29	*22	8		
	Bajo	37	*14	11		
Ítem 34	Alto	*38	14	8		
	Bajo	*26	25	12		
Ítem 35	Alto	13	*26	21		
	Bajo	14	*14	31		
Ítem 36	Alto	*29	11	18		
	Bajo	*14	22	26		
Ítem 37	Alto	4	*20	22	14	
	Bajo	15	*7	23	23	
Ítem 38	Alto	*15	16	15	15	
	Bajo	*7	12	16	33	
Ítem 39	Alto	11	14	*22	12	
	Bajo	8	26	*13	16	
Ítem 40	Alto	*15	19	11	15	
	Bajo	*17	12	12	22	

Del total de muestra que fue de 206 estudiantes, para este anexo se ha escogido el 25% de los estudiantes con las puntuaciones más altas y el 25% de los estudiantes con las puntuaciones más bajas, es decir, tan sólo 100 estudiantes.

En la tabla se puede observar como se han distribuido las respuestas entre las opciones de cada pregunta. En la fila superior se encuentran los estudiantes que tienen las puntuaciones más altas, en la fila inferior están los alumnos con las puntuaciones más bajas en el desempeño de la prueba; la respuesta correcta en todos los casos se señala con un asterisco.

Análisis Ítem por Ítem

- Ítem 1. La alternativa correcta (inciso D) la han escogido el 60.7% de los estudiantes que tuvieron alto rendimiento en la ejecución de la prueba. Sin embargo la respuesta se puede confundir con el inciso B, es decir, es una buena alternativa de respuesta incorrecta que atrae al que no sabe o bien no entiende del tema; sin embargo cabe destacar que la discriminación del ítem fue pésima y lo recomendado es descartar el reactivo.
- Ítem 2. La alternativa correcta (inciso B) ha sido escogida por el 59% de los estudiantes del grupo superior, es decir, ellos son los que conocen del tema, los incisos A y C son buenos distractores y son buenas respuestas incorrectas porque el grupo inferior se inclina por escogerlos, sin embargo hay que revisar a profundidad el ítem.
- Ítem 3. La alternativa correcta es el inciso C. Los del grupo superior distribuyen las respuestas entre dos alternativas, la C y la D, puede ser que ambas respuestas sean correctas o bien la pregunta ambigua. Los incisos A y B son buenos distractores. Sin embargo aun hay que revisar el ítem para mejorar la calidad de éste.
- Ítem 4. La alternativa correcta es el inciso A, ha sido escogido por el 32.8% de la población del grupo superior, pero el inciso D debe ser revisado porque también puede ser la respuesta correcta. Los distractores B y C son buenos.
- Ítem 5. La respuesta correcta es el inciso D, pero aun los que saben del tema se inclinan por la opción de respuesta B, por lo que debe revisarse para verificar que no sea otra respuesta correcta o bien la pregunta sea ambigua.
- Ítem 6. La respuesta correcta es el inciso C, la cual es poco discriminada por los estudiantes, por lo que convendría revisar las opciones de respuesta del reactivo o

bien descartar definitivamente ya que la alternativa B y A parecen ser posibles opciones de respuesta.

- Ítem 7. Los estudiantes del 25% del grupo que más saben se distribuyen entre tres alternativas B, C y D (respuesta correcta), al menos es una pregunta que conviene examinar a profundidad o bien descartar definitivamente.
- Ítem 8. La respuesta correcta es el inciso A, la han escogido el 52.5% del grupo superior, el resto se distribuye en los inciso B, C, y D, lo que indica que son buenos distractores, pero aún hay que revisar a profundidad el reactivo.
- Ítem 9. Es un reactivo que debe revisarse a profundidad ya que el grupo superior se inclina por la opción de respuesta A y B que se consideran distractores y el inciso D (correcto) es considerado como respuesta correcta en menor medida.
- Ítem 10. Pocos son los estudiantes del grupo superior que no logran discriminar entre la respuesta correcta (opción C) y los distractores (opción A), por lo que convendría revisar las opciones de respuesta para mejorar la calidad del reactivo.
- Ítem 11. Es una pregunta con opciones de respuesta dicotómicas, en donde el grupo superior identifica la respuesta correcta (B), es decir se trata de una pregunta que discriminan bien los de este grupo, diferencia claramente a los que saben de los que no saben. Los del grupo inferior optaron por la respuesta A, lo que indica que es una buena alternativa de respuesta incorrecta que atrae al que no sabe o bien no entiende.
- Ítem 12. Se trata de una respuesta dicotómica donde el grupo superior e inferior identifican adecuadamente la respuesta correcta (A), lo que quiere decir que conocen del tema aunque la opción B es un buen distractor podría mejorarse la pregunta para poder distinguir mejor entre aquellos que saben y no saben.
- Ítem 13. Es una pregunta con opciones de respuesta dicotómica, donde el grupo superior identifica la respuesta correcta (B), sin embargo podría mejorarse la pregunta.
- Ítem 14. Es un reactivo que no discrimina adecuadamente, ya que el distractor C es confundido con la respuesta correcta (A), por lo que tiene que revisarse la pregunta para verificar que no sea ambigua o bien las opciones de respuesta que sean las adecuadas para poder identificar claramente entre la respuesta correcta e incorrecta.

- TESIS TESIS TESIS TESIS TESIS
- Ítem 15. Es un reactivo en el que tanto el grupo superior e inferior identifican la respuesta correcta (B) en la misma frecuencia, los distractores (A, C, D y E) tienen frecuencias de respuestas similares por lo que no se identifica claramente el grupo que sabe del que no sabe, por lo que el ítem debe de revisarse a profundidad el ítem o bien descartarse.
 - Ítem 16. Es un reactivo que no discrimina adecuadamente, ya que el inciso E (incorrecto) tiene la mayor tasa de respuesta tanto en el grupo superior como en el inferior y el inciso C (correcta) tiene menos frecuencia de respuestas. Por lo que es importante eliminar o mejorar el inciso E o bien revisar a profundidad la pregunta.
 - Ítem 17. Es un ítem donde el grupo superior discrimina la respuesta correcta de los distractores; no sucede lo mismo con el grupo inferior. Aquí se logra distinguir claramente el grupo superior del inferior.
 - Ítem 18. El grupo superior obtuvo una tasa de respuestas correctas de 75.4% (46) y el grupo de bajo rendimiento obtuvo un 50% (34) el resto se distribuyó en los incisos b y c que eran los distractores. Se considera un reactivo que debe ser revisado para mejorar su calidad.
 - Ítem 19. El grupo superior obtuvo una tasa de respuestas correctas de 34.4% (21) y el grupo de bajo rendimiento obtuvo un 17.6% (12). Se considera que es un reactivo que debe revisarse a profundidad ya que los distractores tienen una frecuencia muy similar a la respuesta correcta.
 - Ítem 20. Es un ítem donde el grupo de alto rendimiento logra discriminar muy bien los distractores de la respuesta correcta, no sucede lo mismo con el grupo de bajo rendimiento. Es un ítem que se debe descartar de la prueba, según el poder de discriminación.
 - Ítem 21. Es un reactivo que tiene tres opciones de respuesta; el grupo superior logra obtener el 62.3% (38) de respuestas correctas, el grupo de bajo rendimiento apenas alcanza el 25% (17). Es un ítem que es bien discriminado por el grupo de alto rendimiento.
 - Ítem 22. Es un reactivo que en la tabla de poder de discriminación se califica como pobre, el cual debe revisarse a profundidad. La respuesta correcta es el inciso A, en donde el grupo de alto rendimiento alcanzó puntajes de 62.3% (38) y el grupo

de bajo rendimiento sólo obtuvo un puntaje de 35.3% (24). El grupo de bajo rendimiento no discriminó entre la respuesta correcta de los distractores.

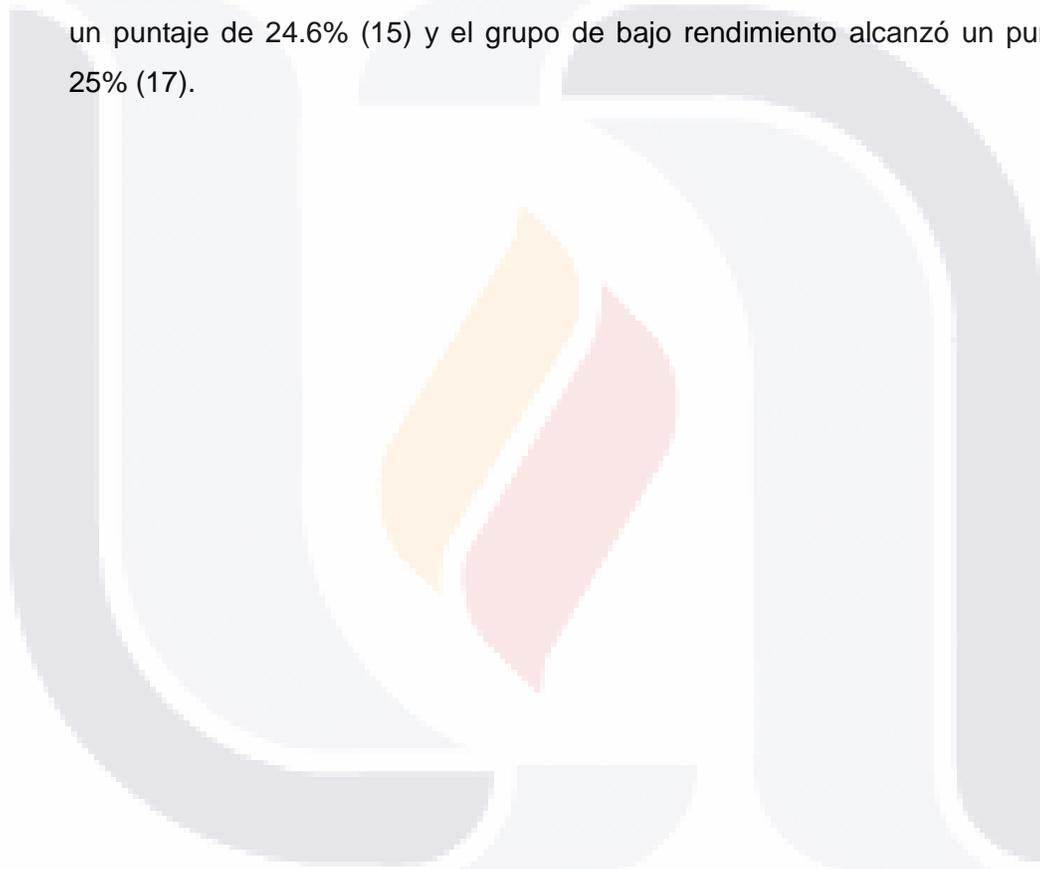
- Ítem 23. Es un reactivo que no discrimina adecuadamente y debe descartarse ya que los estudiantes de ambos grupos (alto y bajo rendimiento) no discriminaron la respuesta correcta de los distractores.
- Ítem 24. Es un reactivo que es necesario revisar a profundidad según la tabla de poder de discriminación. La opción correcta es el inciso D, en donde el grupo de alto rendimiento obtuvo el 49.2% (30) de respuestas correctas y el de bajo rendimiento tan sólo obtuvo el 27.9% (19).
- Ítem 25. Es un reactivo dicotómico que tiene que revisarse a profundidad; se lograron puntajes del grupo superior de 47.5% (29) y el grupo inferior tuvo un puntaje de 27.9% (19). El inciso distractor, tuvo puntajes incluso más altos que la respuesta correcta.
- Ítem 26. Es una pregunta dicotómica la cual debe ser descartada de la prueba según los resultados de la tabla de poder de discriminación ya que las respuestas correctas obtenidas en ambos grupos (alto y bajo rendimiento) son menores que los distractores. El 37.7% (23) del grupo de alto rendimiento logró contestar acertadamente el reactivo, el resto lo contestó de forma equivocada. Para el grupo de bajo rendimiento el porcentaje obtenido fue de 26.5% (18).
- Ítem 27. Es una pregunta dicotómica, la cual debe ser revisada para mejorar la calidad del ítem. El puntaje obtenido en el grupo de alto rendimiento fue de 50.8% (31) y en el de bajo rendimiento fue de 25% (17).
- Ítem 28. Es una pregunta dicotómica que debe ser descartada de la prueba ya que las respuestas incorrectas prevalecen, pues tan solo, el 32.8% (20) del grupo superior y el 38.2% del grupo de bajo rendimiento contestaron acertadamente la pregunta.
- Ítem 29. Es una pregunta dicotómica que debe ser revisada a profundidad. El grupo de alto rendimiento logró obtener un puntaje de 44.3% (27) respuestas correctas y el grupo de bajo rendimiento obtuvo 26.5% (18).
- Ítem 30. Es un reactivo dicotómico que debe ser descartado de la prueba que tanto el grupo de alto rendimiento como el de bajo rendimiento obtuvieron puntajes correctos muy por debajo de aquellos que fueron incorrectos. Tan solo el 39.3% (24) obtuvo la respuesta correcta dentro del grupo de alto rendimiento y el 30.9%

(21) fue del grupo de bajo rendimiento, es decir, menos de la mitad de ambos grupos.

- Ítem 31. Es un reactivo dicotómico que debe ser descartado de la prueba según la tabla de poder de discriminación. El grupo de alto rendimiento obtuvo un porcentaje de respuestas correctas de 63.9% (39) y el grupo de bajo rendimiento obtuvo un porcentaje de 50% (34).
- Ítem 32. Es un reactivo que necesita ser revisado a profundidad, ya que el inciso A respuesta distractora y el inciso C respuesta correcta, tienden a confundirse, ya que los estudiantes eligieron ambos inciso en frecuencias similares. El grupo de alto rendimiento obtuvo un 41% (25) de respuestas correctas y el grupo de bajo rendimiento tan solo obtuvo el 22.1% (15).
- Ítem 33. Es un reactivo que necesita ser revisado a profundidad ya que las respuestas incorrectas prevalecen. El 36.1% (22) del grupo de alto rendimiento obtuvo puntajes acertados al igual que el 20.6% (14) del grupo de bajo rendimiento
- Ítem 34. Es un ítem que debe ser revisado a profundidad ya que el inciso A y el B suelen confundirse, pues fueron elegidos en la misma frecuencia por el grupo de bajo rendimiento, es decir, el distractor se confunde con la respuesta correcta. El grupo de alto rendimiento obtuvo un porcentaje de 62.3% (38) y el grupo de bajo rendimiento obtuvo el 38.2% (26) de respuestas correctas.
- Ítem 35. Es un reactivo que debe ser revisado a profundidad. El grupo de alto rendimiento obtuvo un puntaje de 42.6% (26) y el grupo de bajo rendimiento obtuvo un puntaje de 29.6% (14). Los incisos A y C suelen confundirse con la respuesta correcta.
- Ítem 36. Es un reactivo regular, el cual debe revisarse. Los estudiantes de alto rendimiento obtuvieron un puntaje de 47.5% (29) y los de bajo rendimiento obtuvieron un puntaje de 20.6% (14).
- Ítem 37. Es un ítem que debe revisarse a profundidad. Los puntajes obtenidos son los siguientes: 32.8% (20) para el grupo de alto rendimiento y de 10.3% (7) para el grupo de bajo rendimiento. Los incisos C y D que funcionan como distractores suelen confundirse con la respuesta correcta.
- Ítem 38. Es un reactivo que debe revisarse a profundidad ya que los incisos distractores (B, C y D) suelen comportarse de igual forma que la respuesta

correcta. El grupo de alto rendimiento obtuvo un puntaje de 24.5% (15) y grupo de bajo rendimiento obtuvo un puntaje de 10.3% (7).

- Ítem 39. Es un reactivo que debe ser revisado a profundidad. El puntaje obtenido en el grupo de alto rendimiento fue de 36.1% (22) y en el grupo de bajo rendimiento fue de 19.1% (13). Los incisos distractores (A, B y D) se pueden confundir con la respuesta correcta.
- Ítem 40. Es un ítem que debe ser descartado de la prueba ya que prevalecen los incisos distractores sobre la respuesta correcta. El grupo de alto rendimiento logró un puntaje de 24.6% (15) y el grupo de bajo rendimiento alcanzó un puntaje de 25% (17).



Anexo F. Programa de la materia de estadística

CONTENIDOS	OBJETIVOS
ESTADÍSTICA DESCRIPTIVA	
Fenómenos Aleatorios Experimentos aleatorios Definición de estadística	El alumno valorará la presencia de la variabilidad en los fenómenos aleatorios, para que justifique el uso y la utilidad de la estadística como método de análisis.
Poblaciones y muestras Censos y encuestas El muestreo aleatorio simple El muestreo sistemático Toma de datos Mediciones observadas y esperadas	Comprenderá la utilidad del muestreo, sus bondades y limitaciones, el proceso de obtener información de interés.
Variables cualitativas y cuantitativas Variables discretas y continuas	Clasificará adecuadamente variables, tanto por su tipo como por su densidad.
Tablas de clasificación Proporciones y porcentajes Gráfica de barras Gráfica de sector	Organizará adecuadamente la información contenida en variables categóricas
Medidas de tendencia central. Medidas de dispersión Cuantiles Gráfico de puntos	Obtendrá las diferentes medidas de resumen para variables métricas.
Tablas de frecuencia Histograma Ojiva	Organizará adecuadamente la información contenida en variables métricas mediante técnicas de agrupamiento en clases.
Esquema de tallos y hojas Resumen de valores literales Diagrama de caja y valores aberrantes	Organizará adecuadamente la información contenida en variables métricas mediante el enfoque exploratorio.

ENFOQUES DE PROBABILIDAD

<p>Enfoques de probabilidad. Espacio muestral Cálculo de probabilidades frecuentistas Probabilidad condicional e independencia</p>	<p>Identificará a la probabilidad como una medida de la ocurrencia de un evento aleatorio y verificará sus propiedades en ejemplos concretos.</p>
<p>Variables aleatorias Distribución de probabilidad Valor esperado, varianza y desviación estándar.</p>	<p>Construirá la distribución de probabilidad de una variable aleatoria discreta.</p>
<p>Distribución Bernoulli Función de probabilidad binomial Efectos de los parámetros.</p>	<p>Caracterizará el modelo binomial y construirá la distribución de probabilidad de la variable asociada, aplicada a ejemplos concretos.</p>
<p>Características de la curva normal. Los parámetros de la distribución. La distribución normal estándar. Aplicaciones.</p>	<p>Utilizará la distribución normal como modelo para el estudio de una variable aleatoria continua y su aplicación en situaciones concretas</p>

INFERENCIA ESTADÍSTICA

<p>Concepto de estadístico. La distribución de la media muestral La distribución de la proporción muestral. El error estándar. El Teorema del Límite Central.</p>	<p>Construirá la distribución de estadísticos muestrales.</p>
<p>El proceso de Inferencia estadística. Las características de un estimador. Estimación de la media poblacional. Estimación de la proporción poblacional Estructura de un intervalo de confianza. Intervalo de Confianza para la media. Intervalo de confianza para la proporción. Cálculo del tamaño de la muestra</p>	<p>Estimará parámetros poblacionales mediante la información contenida en una muestra, utilizando un proceso de inferencia estadística.</p>
<p>Concepto de hipótesis. Las hipótesis nula y alternativa. Errores tipo I y tipo II. El estadístico de prueba. El nivel de significancia. La significancia observada (valor p). Prueba para una media Prueba para una proporción. Prueba para dos medias con muestras independientes. Prueba para comparar las varianzas de dos poblaciones (Homocedasticidad)</p>	<p>Utilizará el proceso de Inferencia para probar hipótesis relativas a los parámetros de una y dos poblaciones determinadas, en el contexto de situaciones concretas.</p>
