PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

ESCUELA DE POSGRADO

SIGNIFICADOS DE LAS MEDIDAS DE TENDENCIA CENTRAL. UN ESTUDIO CON ALUMNOS UNIVERSITARIOS DE CARRERAS DE HUMANIDADES.

Tesis para optar el grado académico de magister en Enseñanza de las Matemáticas que presenta:

JAVIER SAYRITUPAC GUTIERREZ

ASESOR

Dr. ULDARICO MALASPINA JURADO

JURADO

Dra. NORMA RUBIO GOYCOCHEA

Mg. AUGUSTA OSORIO GONZALES

Diciembre del 2013

RESUMEN

En el presente trabajo de investigación se analiza los significados personales e institucionales de las medidas de tendencia central en un estudio con alumnos de primeros ciclos de las carreras de humanidades de la Pontificia Universidad Católica del Perú.

Para realizar este análisis se consideró como marco teórico el Enfoque Ontosemiótico de la Cognición e Instrucción Matemática (EOS), el cual nos brindó las herramientas necesarias para el análisis de nuestro objeto de estudio medidas de tendencia central, a través de sus elementos de significado: lenguaje, situaciones, definiciones, procedimientos, proposiciones y argumentos. Se analizaron tres libros de texto usados para la preparación y/o desarrollo del curso. Posterior a esto se elaboró un cuestionario que se aplicó a los estudiantes para luego analizar sus respuestas y evidenciar los conflictos que se presentaron. Se realizó también una entrevista a los profesores del curso. Se hicieron configuraciones cognitivas para analizar las respuestas de los estudiantes.

La metodología empleada fue básicamente de tipo cualitativo e interpretativo, y se complementó con alguna información de carácter cuantitativo, especialmente al presentar los resúmenes de los resultados.

El objetivo general de este trabajo fue analizar los significados personales e institucionales en torno a las medidas de tendencia central: media, mediana y moda, considerando un curso básico de estadística para estudiantes de humanidades de la Pontificia Universidad Católica del Perú.

A manera de resumen de las conclusiones obtenidas, podemos manifestar que los significados de referencia reflejados en los textos analizados, por una parte son restringidos a considerar las medidas de tendencia central como medidas de resumen, sin dar una perspectiva de la media como un estimador del parámetro μ

de la población; y por otra parte, enfatizan los aspectos algorítmicos y de cálculo, y no la comprensión conceptual de estas medidas; sin embargo, esto está presente en los significados pretendidos, como se refleja en las entrevistas realizadas a los docentes. A pesar de ello, no se encuentra entre los significados institucionales implementados y como consecuencia, tampoco en los significados personales logrados de los estudiantes.

AGRADECIMIENTOS

A Dios, el autor de la vida, quien por su gracia, me permite alcanzar este logro.

A un gran maestro, el Dr. Uldarico Malaspina Jurado. Quien además de su acertada asesoría, dio muestras de ser lo que es, una excelente persona. Su paciencia, estímulo y corrección oportuna, han permitido la culminación de mi tesis.

A la profesora Cecilia Gaita, por brindarme todo el apoyo necesario para llevar a cabo la aplicación del cuestionario de evaluación y su preocupación permanente para la finalización de mi tesis.

Al profesor Jorge Bazán Guzmán, por sus aportes, sugerencias y observaciones en los aspectos teóricos de las medidas de tendencia central.

A los profesores del curso Matemáticas por su apertura y colaboración en las diversas etapas de la tesis.

A mis amados padres y mis hermanos, por haberme apoyado de manera incondicional en los momentos que más lo necesité.

A mis queridos amigos, jóvenes y adolescentes, con quienes compartimos una gran amistad y una misma fe.

DEDICATORIA

A mis padres Esteban y Maura, ejemplos de vida y amor.

A mi hermana Oriele, tu esfuerzo no fue en vano.

Lista de figuras

FIGURA 3. Media como centro de gravedad o punto de equilibrio. 36

FIGURA 4. Posiciones relativas de la media, mediana y moda.................................48

FIGURA 5. Registro del movimiento de los astros realizadas por los astrónomos babilonios. *Fuente: http://www.astroyciencia.com/2007/06/19/tablillas-babilonicas/* 70

FIGURA 6. Respuesta del alumno A13 al ítem 2b.3..133

FIGURA 7. Respuesta del alumno A6, al ítem 2.b.3..137

FIGURA 8. Respuesta del alumno A6, al ítem 3b. 139

FIGURA 9. Respuesta del alumno A4, al ítem 3b. 140

FIGURA 10. Respuesta del alumno A47, al ítem 3b..141

FIGURA 11. Respuesta del alumno A36, al ítem 3b..142

Lista de tablas

TABLA 1. Libros analizados para el significado de referencia. 25
TABLA 2. Términos y expresiones verbales encontradas en el texto A. 32
TABLA 3. Términos y expresiones verbales encontradas en el texto B. 32
TABLA 4. Notaciones y símbolos encontrados en el texto A. 33
TABLA 5. Notaciones y símbolos encontrados en el texto B. 34
TABLA 6. Tablas estadísticas y gráficos encontrados en el texto A. 35
TABLA 7. Tablas estadísticas y gráficos encontrados en el texto B. 35
TABLA 8. Resumen comparativo de los textos analizados A y B. 62
TABLA 9. Términos y expresiones verbales encontrados análisis del texto Matemáticas para no matemáticos. 78
TABLA 10. Notaciones y símbolos encontrados en el análisis del texto Matemáticas para no matemáticos. 79
TABLA 11. Tablas estadísticas encontradas en el análisis del texto Matemáticas para no matemáticos. 80
TABLA 12. Resumen de las situaciones-problema encontrados en el análisis del texto Matemáticas para no matemáticos. 84
TABLA 13. Definiciones encontradas en el análisis del texto Matemáticas para no matemáticos. 86
TABLA 14. Propiedades encontradas en el análisis del texto Matemáticas para no matemáticos. 87
TABLA 15. Algoritmos encontrados en el análisis del texto Matemáticas para no matemáticos 88
TABLA 16. Argumentos encontrados en el análisis del texto Matemáticas para no matemáticos. 90
TABLA 17. Resumen sobre las respuestas de los docentes a la entrevista. 91
TABLA 18. Descripción la apertura de la clase relativa a las medidas de tendencia central. 93
TABLA 19. Descripción del desarrollo de la clase relativa a las medidas de tendencia central. 95
TABLA 20. Descripción del cierre de la clase relativa a las medidas de tendencia central. 97
TABLA 21. Resumen de los elementos de significado implementado. 98
TABLA 22. Configuración de la solución experta al ítem1. 103
TABLA 23. Configuración de la solución experta al ítem2. 106
TABLA 24. Configuración de la solución experta al ítem3. 109
TABLA 25. Configuración de la solución experta al ítem 4. 111
TABLA 26. Configuración de la solución experta al ítem 5. 114
TABLA 27. Configuración de la solución experta al ítem6. 116
TABLA 28. Configuración de la solución experta al ítem7. 118

TABLA 29. Resumen de los elementos de significado presentes en el cuestionario. 119
TABLA 30. Frecuencia de tipologías de significados personales declarados en el ítem 1. 122
TABLA 31. Frecuencia de tipologías de significados personales declarados en el ítem 2a. 124
TABLA 32. Frecuencia de tipologías de significados personales declarados en el ítem 2b.1....124
TABLA 33. Frecuencia de tipologías de significados personales declarados en el ítem 2b.2....125
TABLA 34. Frecuencia de tipologías de significados personales declarados en el ítem 2b.3....125
TABLA 35. Frecuencia de tipologías de significados personales declarados en el ítem 3a. 126
TABLA 36. Frecuencia de tipologías de significados personales declarados en el ítem 3b. 126
TABLA 37. Frecuencia de tipologías de significados personales declarados en el ítem 4. 127
TABLA 38. Frecuencia de tipologías de significados personales declarados en el ítem 5a. 128
TABLA 39. Frecuencia de tipologías de significados personales declarados en el ítem 5b.1....128
TABLA 40. Frecuencia de tipologías de significados personales declarados en el ítem 5b.2....129
TABLA 41. Frecuencia de tipologías de significados personales declarados en el ítem 6. 129
TABLA 42. Frecuencia de tipologías de significados personales declarados en el ítem 7a. 130
TABLA 43. Frecuencia de tipologías de significados personales declarados en el ítem 7b. 130
TABLA 44. Resumen cuantitativo de las respuestas a los ítems propuestos........131
TABLA 45. Configuración cognitiva del alumno A13, respecto al ítem 2.b.3. 133
TABLA 46. Configuración cognitiva del alumno A25, respecto al ítem 2.b.3.........134
TABLA 47. Configuración cognitiva del alumno A14, respecto al ítem 2.b.3. 136
TABLA 48. Configuración cognitiva del alumno A6, respecto al ítem 2.b.3. 137
TABLA 49. Configuración cognitiva del alumno A6, respecto al ítem 3b. 139
TABLA 50. Configuración cognitiva del alumno A36, respecto al ítem 3b. 142
TABLA 51.Configuración cognitiva del alumno A13, respecto al ítem 5a. 145

INDICE GENERAL

Resumen……………..………………………………………………………………...….I

Agradecimientos…………………………………………………………………..….III

Lista de Figuras………………………………………………………………….….V

Lista de Tablas………………………………………………………………….……VI

INTRODUCCIÓN…………………………………………………………………..XI

1. DEFINICIÓN DEL PROBLEMA.. 1
1.1 Planteamiento y justificación del tema.. 1
1.2 Antecedentes.. 3
2. MARCO TEÓRICO 10
2.1. Enfoque ontosemiótico de la cognición e instrucción matemática (EOS)... 10
2.1.1. Supuestos en que se basa.. 11
2.2 Panorama general sobre cómo pueden ser vistas las medidas de tendencia central.. 18
3. OBJETIVOS Y METODOLOGÍA 19
3.1. Objetivos.. 19
3.2. Metodología... 20
3.2.1. Población de estudio. Su contexto. 20
3.2.2. Etapas... 21
4. SIGNIFICADO DE REFERENCIA.. 24
4.1 Análisis de los libros de texto seleccionados... 24
4.1.1 Descripción general de los elementos de significado en los libros de texto. 25
4.1.2 Descripción detallada de los elementos de significado y ejemplos específicos encontrados. 31
4.1.3 Tabla de resumen comparativo.. 62
4.1.4 Conclusiones sobre el análisis de los libros de texto. 65

4.2 Significado institucional de referencia 66
4.3. Comentarios sobre el significado de referencia en relación con los fundamentos teóricos de las medidas de posición central. 76
5. SIGNIFICADO INSTITUCIONAL PRETENDIDO 77
5.1 Análisis del texto matemáticas para no matemáticos 77
5.1.1 Descripción de los elementos de significado encontrados en el libro de texto 77
5.2 Descripción de las respuestas de los docentes a la entrevista.................... 90
5.3 Conclusiones sobre el significado pretendido... 92
6. SIGNIFICADO INSTITUCIONAL IMPLEMENTADO 93
6.1 Descripción de la secuencia de clase relativo al tema de medidas de tendencia central. 93
6.2 Elementos del significado implementado.. 98
7. SIGNIFICADO PERSONAL 101
7.1. Descripción y análisis de contenido del cuestionario............................... 101
7.1.1. Soluciones expertas a los ítems del cuestionario y elementos de significado intervinientes. 101
7.1.2 Resumen de los elementos de significado que intervienen en la solución de los ítems del cuestionario.. 119
7.2. Análisis de las respuestas a los ítems propuestos. 121
7.2.1 Estudio cuantitativo de las respuestas. 121
7.2.2 Estudio cualitativo de las respuestas (configuraciones cognitivas).. ... 132
8. CONCLUSIONES Y RECOMENDACIONES 148
8.1 Conclusiones 148
8.2 Resumen de las relaciones entre los diversos significados de las medidas de tendencia central. 150
8.2.1 Fundamentos teóricos y los significados institucionales de referencia………. 150
8.2.2 Significado institucional de referencia y el pretendido.. 151
8.2.3 Significado institucional pretendido y el implementado. 151
8.2.4 Significado institucional implementado y el significado personal logrado.. 151
8.2.5 Un breve comentario que relaciona los diversos significados. 152
8.3 RECOMENDACIONES.. 152

REFERENCIAS.. 154
APÉNDICE.. 157
ANEXOS 164

INTRODUCCIÓN

La importancia que ha adquirido la Estadística en las diversas áreas de estudio como la medicina, biología, sociología, economía, etc., y aún en el quehacer cotidiano de las personas hace necesario que se empiece a educar a los futuros ciudadanos con una formación general en estadística. Dentro de esta formación general para la buena cultura estadística del ciudadano, las medidas de tendencia central, media, mediana y moda se presentan como conceptos que deben formar parte necesaria de dicha formación, y más aún en los profesionales de todas las especialidades, por ser conceptos estadísticos básicos y de gran aplicación práctica en la vida cotidiana y profesional.

Según Batanero, Godino, Holmes y Vallecillos (1994) en el proceso de enseñanza y aprendizaje de los conceptos de media, mediana y moda suelen presentarse diversos errores y dificultades. Esto se debe, entre otros aspectos, a la complejidad de estos conceptos (Mayén, S. 2009). El aprendizaje de la media, mediana y moda, no pasa por aprender los algoritmos de cálculo, sino más bien por comprender sus conceptos, características, su uso pertinente frente a ciertas situaciones y su interpretación en contextos específicos.

De lo anterior, y motivado por las experiencias docentes del investigador, en su calidad de asistente de docencia de cursos en los que se impartían temas de estadística, se realizó el presente trabajo de investigación, que analiza los significados institucionales y personales de las medidas de tendencia central en un estudio con alumnos de primeros ciclos de las carreras de humanidades de la Pontificia Universidad Católica del Perú (PUCP). Se usó para nuestro análisis, el marco teórico del Enfoque Ontosemiótico de Instrucción y Cognición Matemática.

Esta investigación consta de ocho capítulos, que se describen a continuación:

En el capítulo 1, se aborda el problema de investigación, mostrando la importancia de nuestro objeto de estudio, las medidas de tendencia central, su enseñanza y aprendizaje en alumnos de carreras de humanidades. Se presenta además las diversas investigaciones sobre las medidas de tendencia central, previas a nuestro trabajo. Finalmente, se formula las preguntas de investigación.

En el capítulo 2, se presenta primero el marco teórico sobre el cual se fundamenta nuestra investigación, el Enfoque Ontosemiótico de la Cognición e Instrucción Matemática. Finalmente, se presenta un panorama general, que muestra las distintas formas como pueden ser vistas las medidas de tendencia central (como medida de resumen o como un estimador del parámetro μ de la población).

En el capítulo 3, se presenta el objetivo general de la investigación, así como los cinco objetivos específicos que se desprenden del mismo. Por otro lado, se presenta la población sobre la cual se hace el estudio y se explica la metodología usada mostrando las etapas de la investigación.

En el capítulo 4, sección 4.1, se describe el significado institucional de referencia a través del análisis a dos libros de texto recomendados a los estudiantes que llevaron el curso Matemáticas. Para esto, en la sección 4.1.1 se brinda una descripción general de los elementos de significado: Lenguaje, situaciones, procedimientos, conceptos, propiedades y argumentos, analizados. Posteriormente, en la sección 4.1.2 se muestra una descripción detallada de los elementos de significado, mostrando algunos ejemplos específicos encontrados. En la sección 4.1.3 se muestra un cuadro de resumen que compara los resultados de los dos libros de texto analizados. Finalmente, en la sección 4.2, se describe el significado institucional de referencia mostrando algunos datos de carácter histórico.

En el capítulo 5, se describe el significado institucional pretendido, para esto, en la sección 5.1 se muestra el análisis al texto guía usado en el curso “matemáticas para no matemáticos”, identificando los diversos elementos de significado. Por otro lado, en la sección 5.2, se describe las respuestas de los docentes a la entrevista que se diseñó con el fin d conocer los logros que pretenden alcanzar con sus alumnos, así como sus percepciones y concepciones sobre la enseñanza y aprendizaje sobre la media, mediana y moda.

En el capítulo 6, se describe la secuencia de clases relativos al tema de medidas de tendencia central, mostrando elementos importantes encontrados en el desarrollo del mismo dentro del aula.

En el capítulo 7, sección 7.1, se analiza el contenido del cuestionario aplicado a los alumnos, describiendo los elementos de significado intervinientes en la solución a los problemas propuestos. Se resumen, en un cuadro, dichos elementos de significado. En la sección 7.2 se realiza primero un estudio cuantitativo de las respuestas de los alumnos a los problemas propuestos y, en segundo lugar, se realiza un estudio cualitativo, mediante configuraciones cognitivas a algunas de las respuestas de los alumnos (aquellas donde se presentaron muy bajos porcentajes de respuestas correctas).

En el capítulo 8, se establecen las conclusiones considerando nuestros objetivos planteados, se resumen las relaciones entre los diversos significados institucionales y el personal, y se brindan algunas recomendaciones para la enseñanza y el aprendizaje de las medidas de tendencia central.

CAPITULO I

DEFINICIÓN DEL PROBLEMA

1.1 PLANTEAMIENTO Y JUSTIFICACIÓN DEL TEMA

En las últimas décadas se ha hecho evidente los grandes cambios y avances vertiginosos en las ciencias y la importancia que va adquiriendo la Estadística dentro de ésta, debido a su amplia aplicación en diversas áreas de estudio como la medicina, biología, sociología, economía, etc., y aún en el quehacer cotidiano de las personas. Por ello, resulta muy importante educar a los nuevos ciudadanos con una formación general en estadística, que permita, entre muchas otras cosas, saber interpretar de manera correcta las informaciones que se presentan a diario en los diferentes medios, para formar opiniones más fundamentadas y tomar decisiones con criterios más objetivos.

Dentro de esta formación general para la buena cultura estadística del ciudadano, las medidas de tendencia central, media, mediana y moda se presentan como conceptos que deben formar parte necesaria de dicha formación, y más aún en los profesionales de todas las especialidades, por ser conceptos estadísticos básicos y de gran aplicación práctica en la vida cotidiana y profesional, como por ejemplo, al estudiar la esperanza de vida, la tasa de natalidad, el índice de precios, etc.

Del estudio de Mayen, S. (2009), consideramos las siguientes tres razones generales sobre la importancia de hacer un estudio sobre la media, mediana y la moda: a) la complejidad de estos conceptos (a pesar de ser conceptos básicos); b) su importancia para la construcción posterior de otros conceptos estadísticos (resulta útil para los alumnos que más adelante llevarán cursos de estadística como parte de su formación profesional) y c) su uso frecuente en la vida diaria, como ya lo habíamos mencionado.

Ahora bien, es en la etapa escolar donde se inician los estudios de los conceptos básicos de estadística y desde allí se presentan una diversidad de deficiencias (limitaciones) y errores, en su enseñanza y aprendizaje. Esto se debe, entre otras cosas, a la falta de preparación (de tipo conceptual y didáctico) de los profesores en esta área; al tiempo dedicado para la enseñanza de los temas y, como menciona Bazán (2006), a la ausencia de textos adecuados para la enseñanza de temas sobre estadística y probabilidad en este nivel, ya que los que se usan presentan errores de tipo conceptual o didáctico. Todo ello hace que los alumnos lleguen a la universidad con pocos conocimientos, concepciones erróneas y poco interés por aprender estos conceptos estadísticos.

Se constituye entonces, en todo un reto para la universidad a través de sus docentes universitarios de primeros ciclos, lograr que sus alumnos adquieran de manera correcta los conocimientos básicos en estadística y encuentren el sentido práctico de su uso y aplicación en su vida cotidiana y profesional. Esto resulta más relevante en el caso de estudiantes de las carreras de humanidades, pues, por lo general, presentan mayores dificultades para asimilar estos conceptos de manera correcta.

Interesados en esta problemática, realizamos un estudio sobre las medidas de tendencia central, y usamos para ello como nuestro marco teórico de referencia al Enfoque Ontosemiótico de la Cognición e Instrucción Matemática (al que llamaremos de ahora en adelante EOS), que se ha venido desarrollando por Godino y colaboradores como: Font, Batanero, entre otros. Esta teoría en educación matemática, resulta beneficiosa para nuestros fines, pues brinda herramientas valiosas para analizar los factores que condicionan los procesos de enseñanza y aprendizaje de las matemáticas; en nuestro caso particular, las medidas de tendencia central. Una de estas herramientas es la teoría de los significados, que tiene como uno de sus objetivos conocer y analizar los sistemas de prácticas que realizan una persona o institución (significados personales e institucionales) al resolver una determinada situación problema (en nuestro caso, sobre media, mediana o moda) y ver qué tan “cercanos” (acoplados) se encuentran estos significados.

En virtud a ello, en la presente investigación hacemos un análisis de los significados personal e institucional del objeto matemático “*Medidas de Tendencia Central”*; primero, por las razones mencionadas en los párrafos anteriores; segundo, porque estos conceptos suelen presentarse con errores y dificultades en el proceso de enseñanza y aprendizaje (Batanero, Godino, Holmes y Vallecillos, 1994), no solo en la educación básica, sino también, en la educación universitaria, y en particular a los estudiantes de humanidades, que es la población a la cual se dirige nuestra investigación; y finalmente, porque dichos significados nos permitirán evaluar los conocimientos y dificultades de los estudiantes sobre los conceptos de media, mediana y moda, y contribuirán a tener herramientas valiosas para quienes se interesen en diseñar nuevas situaciones didácticas, elaborar nuevos textos o realizar nuevas investigaciones que involucren estos conceptos; todo ello con la finalidad de buscar mejorar la comprensión de estos conceptos por el estudiante.

No hemos encontrado investigaciones que aborden estos conceptos y su problemática en el contexto peruano, por ello en esta investigación pretendemos, en la realidad del Perú y concretamente en Estudios Generales Letras de la Pontificia Universidad Católica del Perú, responder a las siguientes preguntas: ¿Cuáles son los significados institucionales de las medidas de tendencia central en su enseñanza a estudiantes de carreras de humanidades en las que no se usa intensivamente las matemáticas? ¿Cuáles son los significados personales de los estudiantes de estas carreras sobre los conceptos de medidas de tendencia central?

1.2 ANTECEDENTES

Las medidas de tendencia central han sido objeto de diversas investigaciones en los distintos niveles educativos, como la de Batanero, Godino, y Navas. (1997), en el que un estudio con profesores en formación para el nivel primario encontraron errores de tipo conceptual y dificultad de aplicación práctica del concepto de promedio. Otra investigación realizada por Carvallo (citado en Batanero, C., 2000), arroja como resultados algunos errores frecuentes, como por ejemplo, tomar la mayor frecuencia como valor de la moda, hallar la mediana sin antes ordenar los datos, no tomar en cuenta la frecuencia absoluta en el cálculo de la media como también la no distinción y conveniencia en el uso de estas tres medidas de posición, etc. Otra investigación es la desarrollada por Cobo, B. (2003) quien realiza un estudio sistemático sobre el significado de las medidas de posición central en estudiantes españoles de secundaria y destaca algunos resultados importantes como el hecho de que los libros de texto prestan más importancia a la definición de la media que al estudio de sus propiedades, la escases de problemas que busquen estimar una medida a partir de diversas mediciones en presencia de errores (resulta importante considerar este resultado ya que históricamente fue precisamente en la búsqueda de soluciones a estos tipos de problemas de donde surgió la idea de media), entre otras cosas. Así también, en Mayen, Cobo, Batanero, y Balderas. (2007) se encontró que los alumnos mexicanos que finalizan la educación secundaria tienen dificultades en la comprensión de los conceptos de media, mediana y moda.

Un trabajo más reciente sobre los significados de la media, mediana y moda, lo realizó Mayen, S. (2009). Dicha investigación continúa a la realizada Cobo, B. (2003), pero en el contexto mexicano, con alumnos de secundaria y bachillerato que oscilan entre los 13 y 19 años de edad. En ella se hace un estudio especial sobre la comprensión de estos conceptos y se realiza, además, un análisis semiótico de las respuestas de los alumnos a los ítems que involucran el concepto de la mediana. Se encontraron resultados importantes como el hecho de que persisten en los estudiantes ciertas dificultades, como por ejemplo, confundir las definiciones de media y mediana, siendo el cálculo y la interpretación de la mediana más complicado que el de la media. También no identifican de manera correcta los problemas que involucran el uso de la media ponderada, causando conflictos y errores en su solución, repitiéndose así el resultado encontrado en Pollateseks y Cols. (Citado en Batanero, C. ,1994) donde describe el mismo error consistente en emplear la fórmula $\frac{180+120}{2} = 150$ para resolver la situación siguiente: *"Hay 10 personas en un ascensor, 4 mujeres y 6 hombres. El peso medio de las mujeres es* *de 120 libras y el peso medio de los hombres 180. ¿Cuál es el peso medio de las personas en el ascensor?*

Otro error encontrado es que no consideran el efecto que causa los valores atípicos sobre el cálculo de la media, ignorando la representatividad del mismo. Aquí, es preciso recordar que las medidas de tendencia central se aplican a todo el conjunto de datos (distribución) más bien que a individuos. Otros conflictos que se encontraron están relacionados al cálculo de media, mediana y moda de un conjunto de datos agrupados y presentados por medio de tablas de frecuencias, así como también la estimación de valores e interpretación de datos a partir de gráficos estadísticos.

Nuestro trabajo seguirá las pautas generales de Cobo, B. (2003) y Mayen, S. (2009).

Las medidas de tendencia central, definiciones y algunas consideraciones.

En los trabajos estadísticos, cuando se tiene cierta información válida sobre una determinada variable de estudio, interesa hacer un análisis de los datos obtenidos que permitan describirlos, resumirlos y presentarlos de manera que sean fáciles de interpretar. Una tabla de distribución de frecuencias nos da un primer resumen de la información original, sin embargo, si quisiéramos conocer algo más acerca de cómo es que se distribuyen estos datos o buscar un valor que represente al conjunto de datos, tendríamos que calcular una serie de valores llamados *estadísticos*.

Dentro de estos estadísticos se encuentran las *medidas de tendencia central* (llamados así por su característica principal de acumularse al centro de la distribución de datos), siendo las más comunes, la media, la mediana y la moda, que son nuestro objeto de estudio. Estas medidas describen un valor típico o representativo del conjunto de datos y señalan las tendencias o características del mismo.

Para un mayor conocimiento y precisión sobre nuestro objeto de estudio, pasamos a mostrar cómo se define usualmente la media, la mediana y la moda en los textos de estadística básica.

La media (media aritmética) de un conjunto de datos, viene dado por la suma de todos los datos dividido por el número de ellos.

La mediana es el valor que ocupa la posición central de un conjunto de datos previamente ordenados.

La moda es el dato que se repite con mayor frecuencia.

Se enriquecen estas definiciones haciendo uso de un lenguaje simbólico formal.

Si $x_1, x_2, \ldots, x_n$ son los n valores de una variable X, la media aritmética, o simplemente media, de X puede ser escrita como

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Si existieran valores repetidos que se agrupan (x_i con frecuencia f_i), entonces

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i f_i}{n}$$

Considerando los n valores de una variable ordinal X ordenadas en forma creciente (puede ser también decreciente), se denota al menor valor por $x_{(1)}$, al segundo por $x_{(2)}$, y así sucesivamente:

$$x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n-1)} \leq x_{(n)}$$

Con esta notación, la mediana de la variable X queda determinada por

$$Me(X) = \begin{cases} x_{\left(\frac{n+1}{2}\right)}, & si\ n\ es\ impar \\ \dfrac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2}, & si\ n\ es\ par \end{cases}$$

Es preciso notar que cuando se tiene un número par de datos, puede ocurrir que la mediana no necesariamente sea uno de esos datos.

Finalmente, la moda se define como el valor de la variable X que se presenta con mayor frecuencia.

A continuación precisamos algunos aspectos importantes que serán necesarios comprender para cuando nos enfrentemos a diversas situaciones problemas relacionadas al concepto de media, mediana o moda.

Cuando se quiera elegir un buen representante del conjunto de datos, es necesario tomar en cuenta la variable sobre la cual se está haciendo el estudio y la forma como están distribuidos estos datos. En este sentido, Campbell (1981) observa que si bien se suele situar a la media en el centro del recorrido de la distribución, debido a su reconocimiento como valor "típico" o "representativo", esta propiedad sólo es válida cuando la distribución es simétrica, no así, cuando la distribución es asimétrica (presenta valores atípicos y por tanto la media se desplaza a uno de los extremos). En este caso, la mediana se presenta como un valor más representativo del conjunto de datos.

Como un ejemplo de esta situación podríamos considerar los siguientes valores (ingresos mensuales, en miles de soles, de cinco trabajadores de cierta empresa): 1; 3; 5; 7 y 9, al calcular la media y la mediana sus valores coinciden y es igual a 5.

$$media = \frac{1+3+5+7+9}{5} = 5, \qquad mediana = 5.$$

Ahora, si tuviéramos una distribución similar a la anterior, excepto el valor 9, que sería 34, tendríamos: 1; 3; 5; 7 y 34, y al hallar la media y la mediana observamos que, mientras la mediana sigue siendo igual a 5, la media se incrementó a 10.

$$media = \frac{1+3+5+7+34}{5} = 10, \qquad mediana = 5$$

Surge entonces la pregunta, ¿cuál de las dos medidas representa mejor al conjunto de datos? En la primera distribución resulta indistinta la elección de cualquiera de las dos medidas, pues los resultados son los mismos. Sin embargo, para la segunda distribución es preciso notar que, mientras la mediana se mantiene próxima a la mayoría de los datos (1; 3; 5; 7), no viéndose alterada por el valor atípico 34, la media se ve totalmente influida por este número, resultando un valor (10) que está alejado de los valores de la distribución, no siendo así un valor representativo del conjunto de datos.

Otra situación que ejemplifica la "buena elección" de un representante de un conjunto de datos sería la siguiente: *en la última encuesta realizada por la Pontificia Universidad Católica del Perú a la comunidad universitaria sobre qué tipo de oferta alimenticia preferían para el patio de comidas, se obtuvo los siguientes resultados:*

Tipo de oferta alimenticia	**Número de votos**
Pizzas y pastas	*4 617*
Hamburguesas (hamburguesas de carne, parrillas varias)	*4 264*
Helados y postres divertidos	*3 582*

Ensaladas y comida dietética	*3 406*
Total de votos	***15 869***

Ante esta situación podríamos preguntarnos:

¿Cuál de las tres medidas de tendencia central es el adecuado para representar al conjunto de datos? ¿Cuál es el valor de dicha medida?

Para responder a la pregunta, primero tendríamos que reconocer el tipo de variable que está en estudio, en este caso se trata de una variable cualitativa nominal (tipo de oferta alimenticia que se prefiere) y, por lo tanto, la única medida que adquiere sentido en este caso sería la moda, pues si quisiéramos hallar la media de los datos, tendríamos que operar aritméticamente los tipos de oferta alimenticia, y eso, carece de sentido. De manera similar, para hallar la mediana tendríamos que ordenar los datos, y una vez más, tal arreglo, no es posible.

Ya conocido que la moda es la única medida que puede representar a esta variable cualitativa nominal, queda determinar el valor de dicha moda. Antes de hallar dicho valor, resulta propicio mencionar un error frecuente al que suelen recurrir los estudiantes al calcular dicho valor, y es confundir a la moda con la frecuencia, es decir, para nuestro caso, dar como respuesta de la moda el valor 4 617. En buena cuenta, para este problema, la moda sería “pizzas y pastas” por ser el dato que más votos obtuvo (4 617).

Queda claro que la moda se usa principalmente cuando la variable en estudio es de tipo cualitativo nominal y, por tanto, es imposible calcular la media o cuando lo que interesa conocer es el valor dominante del conjunto. Es preciso notar que, dado que la mediana hace referencia al valor que ocupa la posición central o del medio, para que haya tal centro o medio por lo menos tiene que haber un orden. Por esta razón, no es posible calcular la mediana de datos cualitativos nominales.

Resulta importante considerar el aporte de Russel y Mokros (Citado en Batanero, C., 1994) quienes consideran tres tipos de capacidades que implica la comprensión de la idea de valor “típico” o “representativo”:

- Dado el conjunto de datos, comprender la necesidad de emplear un valor central y elegir el más adecuado.

- Construir un conjunto de datos que tenga un promedio dado.
- Comprender el efecto que, sobre los promedios (media, mediana y moda), tiene un cambio en todos los datos o parte de ellos.

Estas y otras cuestiones que estudiaremos más adelante, resultan útiles para el diseño de los cuestionarios de nuestra investigación y posteriores análisis.

CAPITULO II

MARCO TEÓRICO

En esta sección, presentamos el marco teórico en el cual se fundamenta nuestra investigación. Para ello, explicamos algunos elementos de la Teoría del Enfoque Ontosemiótico de la Cognición e Instrucción Matemática que fueron necesarias para nuestra investigación. Así también, presentamos un resumen, que muestra un panorama general de la forma cómo pueden ser vistas las medidas de tendencia central.

2.1. ENFOQUE ONTOSEMIÓTICO DE LA COGNICIÓN E INSTRUCCIÓN MATEMÁTICA (EOS).

Nuestra investigación queda fundamentada en el Enfoque Ontosemiótico de la Cognición e Instrucción Matemática. Dicha teoría propone las siguientes tres dimensiones en el análisis de la enseñanza y el aprendizaje de las matemáticas: la epistémica, la cognitiva y la instruccional. Cada una de ellas se aborda con herramientas agrupadas en tres modelos teóricos: teoría de los significados institucionales y personales de los objetos matemáticos, teoría de las funciones semióticas y la teoría de las configuraciones didácticas. En nuestro estudio abordaremos las dimensiones epistémica y cognitiva. En la primera de ellas describimos los significados institucionales de las medidas de tendencia central en su enseñanza a estudiantes de humanidades de la PUCP, y en la segunda realizamos un estudio sobre las respuestas de los estudiantes a las preguntas formuladas en nuestro cuestionario, obteniendo así los significados personales.

Basados en Godino y Batanero, (1994) y Godino, Batanero y Font, (2008), hacemos una síntesis de los principales elementos que caracterizan al EOS.

2.1.1. Supuestos en que se basa.

Siguiendo la presentación cronológica presentada por Malaspina, U. (2009) podemos mencionar que el EOS es una teoría que ha surgido en el seno de la Didáctica de las Matemáticas y viene desarrollándose en España por casi tres décadas (1984), con el fin de articular y unificar las diversas nociones teóricas sobre el conocimiento matemático, su enseñanza y aprendizaje. Busca así estudiar los factores que condicionan los procesos de enseñanza y aprendizaje de las matemáticas. Con este propósito, el EOS ha venido desarrollando diversas herramientas teóricas las cuales se han ido complementando y refinando de manera progresiva.

El EOS toma como punto de partida la formulación de una ontología de objetos matemáticos que tiene en cuenta el triple aspecto de la matemática:

- *Como actividad de resolución de problemas, socialmente compartida:* las matemáticas constituyen un quehacer humano, dirigido a dar respuesta a ciertas situaciones problemáticas internas o externas a la propia matemática. Los objetos matemáticos (conceptos, procedimientos, propiedades, etc.) surgen de esta actividad y evolucionan progresivamente.

  Cuando una clase de situaciones-problema comparten soluciones y procesos de resolución, se la considera como un campo de problemas. Para nuestra investigación nos centraremos en los campos de problemas y las actividades de las que emergen los objetos matemáticos “media”, “mediana” y “moda”, conocidas como *medidas de tendencia central.*
- *Como lenguaje simbólico:* las matemáticas pueden ser vistas como un lenguaje simbólico en el que se expresan las situaciones problema y sus soluciones.
- *Como sistema conceptual lógicamente organizado:* la organización lógica de los conceptos, los teoremas y las propiedades también explica el gran número de problemas implicados en el aprendizaje de las matemáticas.

Con el fin de mostrar, por un lado, el triple carácter de la matemática al que nos hemos referido, y por otro, la génesis personal e institucional del conocimiento matemático, y la mutua interdependencia entre estos conocimientos, en el EOS se toma como noción primitiva la de *situación-problema*. A partir de esta idea, se definen conceptos teóricos como práctica, objeto y significado (considerando la “dualidad” personal-institucional), los que resumimos a continuación.

2.1.1.1. Prácticas e instituciones.

Práctica matemática: Es toda actuación o manifestación (verbal, gráfica, etc.) realizada por alguien para resolver problemas matemáticos, comunicar a otros la solución obtenida, validarla y generalizarla a otros contextos y problemas. (Godino, J.D. y Batanero, C. 1994). Por ejemplo, para resolver un problema de estimar un valor desconocido de un objeto a partir de las medidas del mismo, una práctica matemática puede ser sumar todos los valores obtenidos y dividir por el número de ellos.

Estas prácticas pueden ser específicas de una persona o compartidas en el seno de una institución. Es bueno aclarar que en el estudio de las matemáticas, más que una práctica particular ante un problema concreto, interesa considerar los *sistemas de prácticas de las personas* en su actuación ante tipo de situaciones problemáticas.

Institución: Está constituida por las personas involucradas en una misma clase de situaciones problemáticas. El compromiso mutuo con la misma problemática conlleva la realización de unas prácticas sociales compartidas que suelen tener rasgos particulares, generalmente condicionadas por los instrumentos disponibles en la misma, sus reglas y modos de funcionamiento. Así por ejemplo, mientras que en una institución científica, el cálculo de la media mediante el uso de computadoras sería aceptado como lo ideal, en una escuela (que es otra institución) no necesariamente, pues en este caso se pretende asegurar que los alumnos hayan comprendido bien el algoritmo de cálculo. Algunas veces esto obliga a los profesores a forzar a que los alumnos realicen dicho cálculo a mano.

Sistema de prácticas personales: Constituida por las prácticas prototípicas que una persona realiza en su intento de resolver un campo de problemas.

Sistema de prácticas institucionales: Constituida por las prácticas consideradas como significativas[1] para resolver un campo de problemas y compartirlas en el seno de una institución.

[1] Para el EOS una práctica es significativa (o tiene sentido) para una persona o institución, si cumple una función

2.1.1.2. Objetos institucionales y personales.

Objeto institucional: Es un emergente del sistema de prácticas asociadas a un campo de problemas y que son compartidas en una institución. Dicha emergencia es progresiva a lo largo del tiempo.

Objeto personal: Es un emergente del sistema de prácticas personales asociadas a un campo de problemas. La emergencia del objeto es progresiva a lo largo de la historia del sujeto, como consecuencia de la experiencia y el aprendizaje.

2.1.1.3. Significado personal e institucional de un objeto

Cuando se considera una institución escolar, como la educación universitaria, puede ocurrir que el significado construido por un cierto estudiante sobre el objeto media, mediana o moda, en un momento del proceso de aprendizaje, puede no corresponder al significado que se da del mismo objeto en la institución dada, por consiguiente, es necesario hacer una distinción entre el significado personal e institucional de un objeto para luego analizar si existe un acoplamiento entre dichos significados.

Significado[2] institucional de un objeto: Es el sistema de prácticas institucionales asociadas al campo de problemas de las que emergen los objetos institucionales en un momento dado.

Significado personal de un objeto: Es el sistema de prácticas personales asociadas al campo de problemas de las que emergen los objetos personales en un momento dado.

Así, ante la pregunta ¿Qué es el objeto matemático media aritmética? La respuesta que se propone en el EOS es, “el sistema de prácticas que realiza una persona (significado personal) o compartido en el seno de una institución (significado institucional) con el propósito de resolver un tipo de situaciones-problema en los cuales requiere encontrar un representante de un conjunto de datos”.

2.1.1.4. Tipos de significados institucionales y personales

Dado que los significados dependen de los contextos sociales y de los sujetos, su carácter es relativo. En consecuencia, su utilización en el análisis didáctico lleva a

[2] El significado de un objeto matemático debe entenderse en términos de lo que se puede hacer con dicho objeto. Godino y Font, (2007, p.4)

introducir la tipología básica de significados que se resume en la figura 1 (Godino, 2008, p.6).

Así, en el EOS, respecto a los significados institucionales se proponen los siguientes tipos:

Significado de referencia (qué significado se considera en una enseñanza o investigación): Sistema de prácticas que se usa como referencia o patrón para elaborar el significado pretendido del objeto matemático media (luego, mediana y moda). Según Wilhelmi, Godino y Lacasta (2007), en una institución de enseñanza específica, este significado forma una parte del significado holístico del objeto en estudio. Para determinar dicho significado global se hace necesario un estudio histórico y epistemológico sobre el origen y evolución del objeto matemático.

Significado pretendido: sistema de prácticas incluidas en la planificación del proceso de estudio, es decir, en la clase de matemáticas (lo que se pretende enseñar de la media, mediana o moda).

Significado implementado: en un proceso de estudio específico es el sistema de prácticas efectivamente implementadas por el docente (lo que realmente se logra enseñar).

Significado evaluado: el subsistema de prácticas que usa el docente para evaluar los aprendizajes (lo que se evalúa).

Respecto a los significados personales, el EOS, propone los siguientes tipos:

Significado Global: corresponde a la totalidad del sistema de prácticas personales que es capaz de manifestar potencialmente el alumno, relativas al objeto matemático media, mediana o moda (todo lo que el alumno sabe sobre estos objetos).

Este tipo de significado puede incorporar prácticas no previstas en el significado pretendido.

Significado evaluado: lo que podemos evaluar a priori del conocimiento del alumno.

Significado declarado: corresponde a las prácticas efectivamente expresadas, en respuesta a las pruebas de evaluación propuestas, incluyendo tanto las correctas como las incorrectas (desde el punto de vista institucional).

Significado logrado: corresponde a las prácticas manifestadas que son conformes con las prácticas institucionales establecidas.

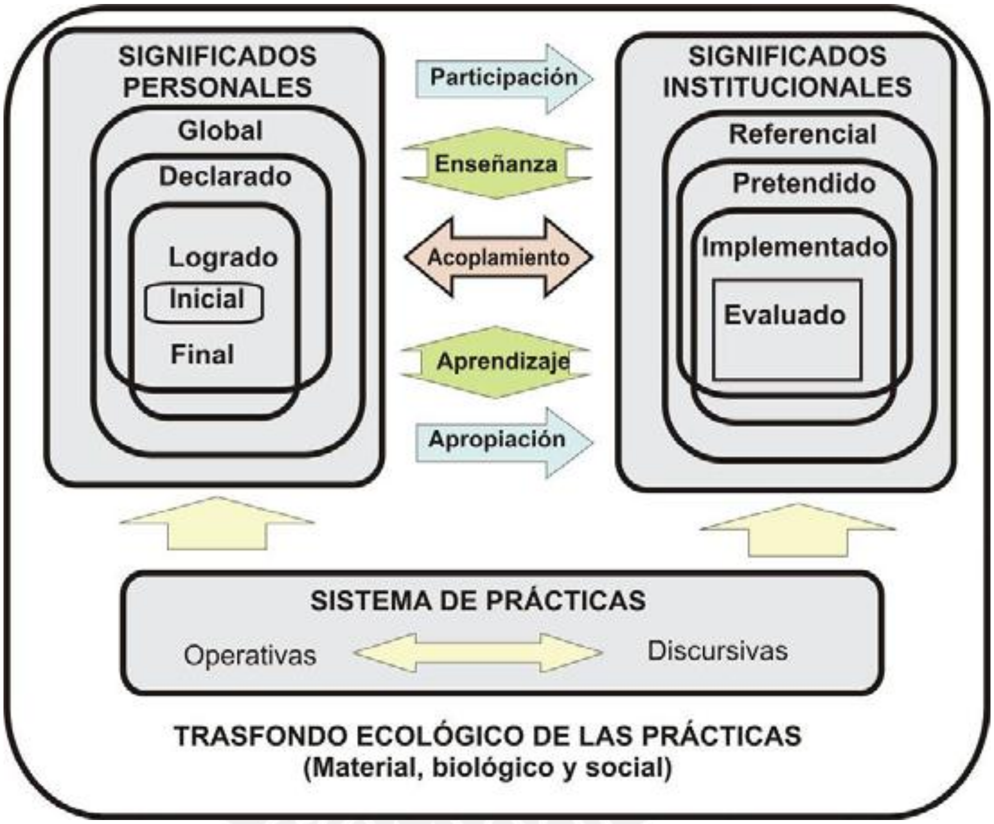


FIGURA 1. Tipos de significados institucionales y personales. Fuente: Godino, et al. (2008)

En la parte central de la figura 2 se indican las relaciones entre la enseñanza y aprendizaje, que supone el acoplamiento progresivo entre los significados personales e institucionales de los objetos matemáticos que emergen del sistema de prácticas. Asimismo, la enseñanza implica la participación del estudiante en la comunidad de prácticas que soporta los significados institucionales y el aprendizaje, en última instancia, supone la apropiación por el estudiante de dichos significados, a esto, el EOS lo llama comprensión personal de un objeto.

2.1.1.5. Objetos intervinientes y emergentes de los sistemas de prácticas.

En el modelo del EOS se asume que en las prácticas matemáticas *intervienen* objetos matemáticos que evocamos al hacer matemáticas y que son representados en forma textual, oral, gráfica o incluso gestual. Por ejemplo, al resolver el problema de estimar una medida a partir de varias mediciones de la misma, intervienen ciertos objetos como “adición”, “división”, “error de medida”, etc. De los sistemas de prácticas matemáticas *emergen* nuevos objetos que dan cuenta de su organización y estructura, así, siguiendo el ejemplo anterior, la media es un emergente que resulta de sumar todas las medidas y dividirla por el número de sumandos. Estos objetos emergentes pueden ser institucionales o personales según lo explicado en el punto **2.3.1.2.**

Para un análisis más profundo y completo de las prácticas matemáticas, el EOS realiza una descomposición de los objetos matemáticos emergentes de dichas prácticas, introduciendo las siguientes entidades (objetos) primarias como componentes del significado de cada objeto:

Lenguaje: se refiere a los términos, expresiones, notaciones o gráficos asociadas a las medidas de tendencia central (en sus diversos registros, visual, oral, gestual, etc.) Por ejemplo, las palabras media o media aritmética, mediana y moda, así como sus representaciones ($\bar{x}, M_e, M_o$), entre otras cosas.

Situaciones problema: se refieren a los problemas o ejercicios o cualquier otra situación problemática, en cuya resolución emergen los conceptos de media, mediana o moda. Por ejemplo la idea de media podría surgir al presentar al alumno la siguiente situación: El peso de un objeto pequeño es medido por 5 estudiantes de una clase, obteniéndose los siguientes valores en milígramos: 3.2; 3.5; 3.3 y 3.4 ¿Cuál sería la mejor estimación al peso real del objeto?

Procedimientos: se refiere, entre otras cosas, a los algoritmos, operaciones que realiza un alumno al enfrentarse a las situaciones problemas, muchos de estos procedimientos pueden ser algo rutinarios. Por ejemplo, para calcular el promedio de cierto número de datos, se suman todos los datos y se divide entre el total de sumandos.

Conceptos-definición: introducidos mediante definiciones o descripciones. Por ejemplo, la definición de mediana como valor central o como valor que divide una distribución ordenada de datos en dos partes iguales.

Proposiciones: enunciados sobre conceptos; es decir, propiedades características y su relación con otros conceptos. Por ejemplo, la media de un conjunto de datos no siempre es igual a uno de los valores de los datos.

Argumentos: enunciados que se emplean para validar o explicar las proposiciones y procedimientos.

2.1.1.6. Comprensión.

Según Godino (2008), básicamente existen dos maneras de entender la comprensión: como un *proceso mental* o como una *competencia*, pero debido a la inclinación pragmatista del EOS, se entiende a la comprensión como una competencia y no tanto como un proceso mental, pues según este marco, un sujeto comprende un determinado objeto matemático cuando es capaz de usarlo de manera competente en diferentes prácticas.

Godino (citado en Mayén, S., 2009) aclara que no es posible observar de manera directa la comprensión personal, mientras que si podemos observar las prácticas personales (significado). En este sentido, para evaluar la comprensión de un objeto matemático según el EOS, será necesario un estudio de la correspondencia entre los significados personales e institucionales de dicho objeto.

2.2 PANORAMA GENERAL SOBRE CÓMO PUEDEN SER VISTAS LAS MEDIDAS DE TENDENCIA CENTRAL.

Las medidas de tendencia central pueden verse desde las siguientes perspectivas:

1. Medida de resumen. Es decir, funciona como indicador de un conjunto de datos. Por ejemplo, el promedio de notas de un grupo de alumnos, la edad promedio de los profesores de alguna universidad, etc.
   En situaciones como éstas, las medidas de tendencia central (MTC) se presentan como un fin, y lo que se busca es conocer el valor que representa mejor al conjunto de datos.
2. Estimador de parámetros. Cuando se va a estudiar el comportamiento de una población definido por una variable aleatoria X y con ciertos parámetros μ y σ , y se desea estimar el valor de μ, se selecciona, bajo ciertas condiciones, una muestra aleatoria y, en este conjunto, se calcula la media muestral (uno de los principales indicadores de tendencia central) como un estimador del parámetro μ de la población. A este proceso se le llama inferencia estadística. Es evidente que, en este caso, las medidas de tendencia central son un medio y no el fin del estudio (Ver Apéndice).

Como se podrá notar, los conceptos de media, mediana y moda presentados en el Apéndice son estudiados bajo el concepto de variable aleatoria ya sea discreta o continua. Estas definiciones resultan imprescindibles para estudiar la estadística inferencial, que es donde se centra la mayor riqueza de toda la estadística. Veremos más adelante, que los libros de texto analizados enfocan las medidas de tendencia central como medidas de resumen.

CAPITULO III

OBJETIVOS Y METODOLOGÍA

Una vez presentado el marco teórico, estamos en condiciones de especificar los objetivos de nuestra investigación y la forma en que tratamos de abordar dichos objetivos**.**

3.1. OBJETIVOS.

OBJETIVO GENERAL: Analizar los significados personales e institucionales en torno a las medidas de tendencia central: media, mediana y moda, considerando un curso básico de estadística para estudiantes de humanidades de la Pontificia Universidad Católica del Perú.

OBJETIVOS ESPECÍFICOS:

Los objetivos específicos que se desprenden de lo anterior son:

1. Describir los significados institucionales referenciales correspondientes a las medidas de tendencia central, mediante el análisis de dos textos de estadística.
2. Describir los significados institucionales pretendidos correspondientes a las medidas de tendencia central, mediante el análisis del texto guía del curso de Matemáticas en Estudios Generales Letras (EEGGLL) y de una entrevista semi estructurada realizada a los profesores del curso.
3. Describir los significados institucionales implementados correspondientes a las medidas de tendencia central, mediante la observación del proceso seguido en el aula, al enseñar las medidas de tendencia central en el curso de Matemáticas en EEGGLL.
4. Describir los significados personales declarados de los estudiantes al finalizar el proceso de estudio, mediante un cuestionario diseñado para tal fin.

5. Proponer recomendaciones para la enseñanza de las medidas de tendencia central a partir de la comparación de los significados personales logrados por los estudiantes, con los significados institucionales implementados.

3.2. METODOLOGÍA

La metodología empleada fue básicamente de tipo cualitativo e interpretativo, y se complementó con alguna información de carácter cuantitativo, especialmente al presentar los resúmenes de los resultados.

3.2.1. Población de estudio. Su contexto.

Población de estudio

Nuestra población de estudio fue de 49 alumnos universitarios de primeros ciclos de EEGGLL de la Pontificia Universidad católica del Perú. Todos ellos matriculados, por primera vez, en el curso “Matemáticas (MAT128)”. La parte experimental de nuestra investigación se llevó a cabo en el semestre 2009-2 y las edades de los alumnos oscilaron entre los 16 y 19 años de edad aproximadamente.

Contexto

La Pontificia Universidad Católica del Perú, a través de la Unidad Académica Estudios Generales Letras, ofrece, a partir del semestre 2007-2 un curso denominado *“Matemáticas”,* dirigido exclusivamente para los alumnos de los primeros ciclos que aspiran a carreras de humanidades, donde las matemáticas no son usadas de manera intensiva, como derecho, periodismo, historia, literatura, entre otras. Este curso surge con el fin de lograr que dichos alumnos adquieran de manera correcta los conceptos básicos de matemática y encuentren sentido al uso práctico y la aplicación en su vida cotidiana y profesional. Para tal fin, se hizo algunas consideraciones importantes, no solo en los contenidos, sino también en la forma de impartirlos, fundamentándose en una metodología activa basada en problemas con contextos cercanos a la realidad de los estudiantes o su futura actividad profesional.

Ya para el 2009-2, fruto de muchas reflexiones y experiencias de los docentes, se publica el primer libro de texto “*Matemáticas para no matemáticos”*, que desde entonces es usado como texto guía en dicho curso, no solo para los alumnos, sino también para el docente, quién desarrolla su clase según el orden temático establecido en el libro, pues allí se están todos los temas a desarrollar durante el curso, expuestos de una manera que los alumnos comprendan.

3.2.2. Etapas

El desarrollo de nuestra investigación se organizó siguiendo las pautas presentadas por Alvarado, H. (2007). En nuestro caso, consideramos cinco etapas, cada una de las cuales nos permitió atender a los objetivos planteados, con el fin de describir los significados institucionales y personales de las medidas de tendencia central en una institución particular (PUCP), más específicamente, en EEGGLL.

Etapa1. *Presentación del fundamento teórico (Cap.2).* En ella se presentó los fundamentos teóricos de las medidas de tendencia central, así como también el marco teórico que sustenta nuestra investigación.

Etapa2. *Análisis de los significados institucionales de referencia y pretendido (Cap.4). E*l propósito de esta fase fue fijar el significado institucional de referencia y el significado pretendido, para luego (en el Cap.7) compararlo con el significado logrado por los estudiantes en el estudio de las medidas de tendencia central.

Para describir el significado de referencia en nuestra investigación, se realizó lo siguiente:

- *Se consideró algunos resultados de las tesis doctorales desarrolladas por Cobo, B. (2003) y Mayén, S. (2009) relativos al significado de referencia de las medidas de tendencia central*.
- *Se seleccionaron los libros de texto a analizar:* "Estadística elemental, lo esencial" *(Johnson y Kuby, 2004)* y "Estadística" *(Murray R. Spiegel y Larry J. Stephens, 2009).* La elección de estos libros de texto se justifican en el capítulo 4.
- *Se analizaron los contenidos respecto a las medidas de tendencia central en los libros de texto seleccionados.* Para el análisis de los libros de texto señalados, se realizaron los siguientes pasos:
    - Selección de aquellos capítulos donde se aborda el estudio de las medidas de tendencia central.

❖ Descripción de los elementos de significado encontrados en los textos (campo de problemas, lenguaje, procedimientos, conceptos, propiedades y argumentos)

Para describir el significado pretendido, se realizó lo siguiente:

- *Se analizaron los contenidos respecto a las medidas de tendencia central en el libro de texto "Matemáticas para no matemáticos"*, el cual es el libro usado por alumnos y profesores en el desarrollo de la clase (esto se describe con mayor precisión en el capítulo 5). Para su análisis, se consideró los pasos descritos anteriormente.
- *Se elaboró una entrevista semi estructurada (anexo 4)* con el objetivo de conocer las creencias y concepciones de los profesores del curso respecto a las Medidas de tendencia central, así como también, qué es lo que esperan logren sus alumnos luego de estudiar estos objetos matemáticos.
- *Entrevistas a los profesores.* Dichas entrevistas se desarrollaron personalmente y algunas fueron grabadas en audio previo consentimiento del profesor.

Etapa 3. *Determinación del significado institucional implementado (Cap.6).* Por medio de observaciones de las sesiones de clase del curso "Matemáticas" en relación al tema "Medidas de tendencia central", se describió la enseñanza tal como fue llevada a cabo. Asimismo, se observaron y anotaron algunos conflictos semióticos que los alumnos pusieron de manifiesto en el transcurso de la enseñanza.

La técnica que se usó para este fin es la "Observación participante", y toda la información se recogió en un *diario de observación* (anexo3) diseñado para tal fin.

Etapa 4. *Análisis del significado personal declarado de los estudiantes.* Para el logro de este objetivo, se realizaron los siguientes pasos:

- *Elaboración de cuestionarios.* Se construyeron los siguientes dos cuestionarios:
  - ❖ *Cuestionario inicial*, de cinco preguntas, que fue tomado antes de llevarse a cabo el desarrollo del tema Medidas de tendencia central y que sirvió para conocer los conocimientos previos de los alumnos.
  - ❖ *Cuestionario final,* compuesto por siete preguntas y que fue tomado luego del desarrollo del tema Medidas de tendencia central. Seis preguntas fueron tomadas del trabajo de Cobo, 2003, pero adaptadas y llevadas a

un contexto más cercano del alumno. Una pregunta (ítem 7) fue creado por el investigador y asesor.

- *Aplicación de los cuestionarios.* Los cuestionarios fueron aplicados poco antes de finalizar el ciclo 2009-II. El cuestionario inicial, como habíamos mencionado, se tomó antes del inicio del tema Medidas de tendencia central y se dio un tiempo de 30 minutos para su solución. La aplicación del cuestionario final, se dio después del desarrollo del tema en cuestión y las preguntas se incluyeron en una prueba de 1 hora con 50 minutos de duración, válida para la evaluación del curso.
- *Análisis de significados personales declarados.* El análisis de las respuestas al cuestionario final permitió caracterizar los elementos de significado puestos en juego por los alumnos. Dicha caracterización sirvió para describir el significado personal logrado y evaluar su comprensión sobre el objeto medidas de tendencia central.

Etapa 5. *Conclusiones.*

Se establecieron algunas conclusiones a partir de las comparaciones entre el significado institucional implementado y el significado personal logrado; asimismo, se propusieron algunas sugerencias de carácter didáctico.

CAPITULO IV

SIGNIFICADO DE REFERENCIA

El significado de referencia es el sistema de prácticas que se usa como referencia o patrón para elaborar o analizar el significado pretendido de un objeto matemático. Según Wilhelmi, et al. (2007), en una institución específica – en nuestro caso la Pontificia Universidad Católica del Perú – este significado forma parte de del significado holístico del objeto de estudio. Así, el objetivo de este capítulo consiste en describir y fijar el significado institucional de referencia sobre las medidas de tendencia central para nuestra investigación y, dado que, para determinar dicho significado se hace necesario, entre otros aspectos, un estudio histórico y epistemológico sobre el origen y evolución del objeto media, mediana y moda, consideramos las siguientes dos etapas:
Primero (4.1), analizamos dos textos que hemos seleccionado para este fin, centrándonos en los capítulos que tratan el tema en estudio, clasificándolo y describiendo los elementos de significado de la media, mediana y moda. Se añaden, además algunos ejemplos que permiten visualizar mejor esta clasificación.
Segundo (4.2), consideramos los diversos aportes encontrados en Cobo, B. (2003) y Mayen, S. (2009) y, en conjunto con 4.1 fijamos el significado de referencia.

El significado de referencia que fijamos resulta útil pues brinda los elementos necesarios para la construcción de los instrumentos de evaluación y la interpretación de las respuestas de los alumnos.

4.1 ANÁLISIS DE LOS LIBROS DE TEXTO SELECCIONADOS.

Para esta parte nos centramos en el análisis de los libros de texto que aparecen el siguiente cuadro:

Texto	Titulo	Autores/Año	Editorial
A	“Estadística elemental, lo esencial”	*(Johnson y Kuby, 2004)*	Thomson
B	“Estadística”	*(Murray R. Spiegel y Larry J. Stephens, 2009)*	McGraw-Hill

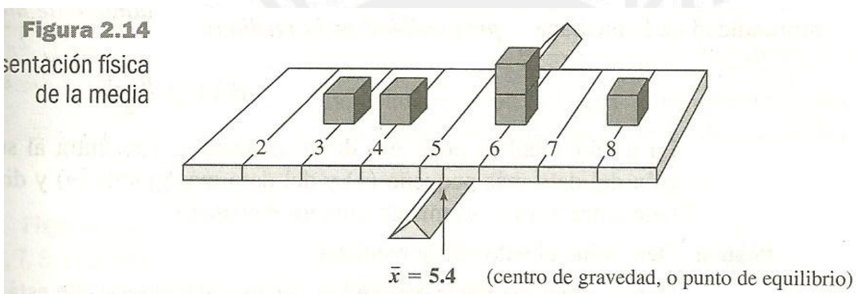
TABLA 1. Libros analizados para el significado de referencia.

El primero se consideró por ser un texto que aparece como bibliografía básica del curso “*Matemáticas*” y el segundo por ser un libro que tiene el propósito de presentar los principios generales de estadística, para cuyo aprendizaje - según los autores - se requiere sólo de conocimientos básicos de álgebra y aritmética elemental, haciéndose viable y de fácil lectura para estudiantes que se dirigen a carreras de humanidades.

En el caso del libro de texto A, el tema de “Medidas de tendencia central” es tratado dentro del capítulo 2 junto con otros temas relativos al análisis descriptivo e interpretación de datos. En tanto, el libro de texto B aborda este tema en todo el capítulo 3.

4.1.1 DESCRIPCIÓN GENERAL DE LOS ELEMENTOS DE SIGNIFICADO EN LOS LIBROS DE TEXTO.

En esta sección se brinda una información general sobre los elementos de significado que se encontraron en los dos libros de texto analizados relativos al tema “Medidas de tendencia central”. Ya en la sección 4.1.2 se desarrolla detenidamente y ejemplifica a cada uno de estos elementos.

TEXTO A *(Johnson y Kuby, 2004)*	
LENGUAJE	Existe una variedad bastante amplia de términos y expresiones que son usadas en el estudio del objeto medidas de tendencia central. Algunos de ellos de uso poco común, como por ejemplo la expresión verbal *"profundidad"* que hace referencia a la posición de la mediana $(\tilde{x})$ y simbolizado por $\boldsymbol{d}(\tilde{x})$. Asocia el término promedio con todas las medidas de tendencia central y no a uno en particular. Se presentan algunos gráficos y figuras variados, de los cuales destaca la siguiente figura (p.57) que brinda una interpretación física de la media: 
SITUACIONES PROBLEMA	Los ejemplos son escasos pero se explican de manera bien detallada. Predominan los ejercicios propuestos, lo cuales, en su mayoría, son situaciones contextualizadas. Se presentan, además, estudios de caso basados en situaciones de interés y en los que usan datos reales.
CONCEPTOS	Las definiciones se presentan de forma directa, evitando el uso de fórmulas complejas y usando más bien - lo que ellos llaman - un *"lenguaje algebraico coloquial".* Se define también otra medida de tendencia central llamada *rango medio*.

TEXTO A	
PROPIEDADES	Son muy pocos los casos donde se presentan las propiedades de manera explícita, ya que el alumno tiene que inducirlas a partir de los ejercicios que se proponen.
ALGORITMOS Y PROCEDIMIENTOS	Los ejemplos que presentan para clarificar los conceptos son bien detallados, paso a paso. Sin embargo, son muy pocos los algoritmos y procedimientos que se explican, pues la mayoría de los problemas son propuestos. Se hace uso de los siguientes software: calculadora, MINITAB y el EXCEL.
ARGUMENTOS	No hay demostraciones y, respecto a la mayoría de las propiedades, se espera que los alumnos puedan inducirlas luego de resolver los problemas propuestos. Se hace uso del lenguaje verbal para justificar ciertos resultados.

TEXTO B *(Murray R. et al, 2009)*	
LENGUAJE	Presenta también una gran variedad de términos y expresiones verbales involucradas en el estudio de las medidas de tendencia central. A estas medidas las llaman también “promedios”, sin embargo, no usan el término promedio como sinónimo de la media aritmética, pues mencionan que, en un sentido estricto, esto no es correcto, pues además de la media hay otros promedios. Así también, se presentan diversos símbolos y notaciones, de las cuales destacan algunos muy particulares para la mediana y la moda ($\tilde{\boldsymbol{X}}\ y\ \hat{\boldsymbol{X}}\ respectivamente$) , así como la aparición de una nueva “forma” para hallar la media (presentada en la siguiente sección), que involucra el uso de las desviaciones de los datos respecto a un cierto valor que es supuesto como media. Los gráficos y diagramas son escasos, ya que encontramos solo algunas tablas básicas de frecuencia, diagramas de tallo y hojas y unas curvas de frecuencia que son usadas para dar una idea geométrica de las medidas de tendencia central.
SITUACIONES- PROBLEMA	Se presentan diversos ejemplos y una gran variedad de problemas resueltos y que además, en gran parte, son problemas contextualizados, los cuales son tomados fundamentalmente del periódico *USA today*. Se encuentran también varios problemas propuestos. Los ejemplos se presentan después de la definición de los conceptos; luego se aborda una diversidad de problemas resueltos que sirven para afianzar los conceptos, así como también, para reflexionar sobre algunas propiedades y características importantes de las medidas de tendencia central.

TEXTO B	
CONCEPTOS	Las definiciones se dan de manera directa, sin presentación previa de algunas situaciones que involucren estos conceptos. Respecto a la media, hemos encontrado diversas formas de definir el mismo concepto, aunque en el fondo enfatizan los diferentes aspectos del significado de los conceptos o sencillamente remiten a diferentes formas de cálculo. En cuanto a la mediana y a la moda las definiciones (para datos no agrupados) son puramente verbales, evitando cualquier tipo de notaciones y fórmulas. Finalmente, se da una idea geométrica para la definición de la mediana, usando el histograma para interpretar el significado de la misma. Se definen también otras medidas de tendencia central como la media geométrica y la media armónica.
PROPIEDADES	Las propiedades son presentadas a través de los diversos problemas resueltos; incluso, para la media aritmética, se presenta una lista aparte de ejercicios donde se hacen explícitas diversas propiedades. En varios de los problemas se pide demostrar dichas propiedades, en otras (en menor número) se espera que el alumno, al resolver ejercicios prácticos, pueda inducirlas.
ALGORITMOS Y PROCEDIMIENTOS	Se explican los diversos algoritmos útiles para el cálculo de las medidas de tendencia central ya sea cuando se presentan distribuciones con datos discretos o agrupados. Para este fin, se valen de una gran variedad de ejercicios resueltos que se presentan en el libro. Es importante notar que en muchos de estos ejercicios, se brinda un espacio para clarificar algunas ideas importantes de estos conceptos, ya sea en el lenguaje, en las notaciones y/o propiedades. El uso de la tecnología se hace evidente, tanto así que uno de los sub temas es dedicado exclusivamente a la explicación en el uso de cinco programas usados en estadística (EXCEL, MINITAB, SPSS, SAS y STATISTIX).

<table>
<tr><th colspan="2">TEXTO B</th></tr>
<tr><td>ALGORITMOS Y PROCEDIMIENTOS</td><td>
Encontramos, además, una nueva forma de calcular la media (AM8), la cual resulta interesante, ya que si el valor escogido para A coincide con la media de la distribución de datos, la relación dada en (5) o (6) nos dice que la suma de las desviaciones respecto a la media es cero.

4. Si se cree o se supone que un número A (que puede ser cualquier número) es la media aritmética y si $d_j = X_j - A$ son las desviaciones de X_j de A, entonces las ecuaciones (1) y (2) se convierten, respectivamente, en
$$\bar{X} = A + \frac{\sum_{j=1}^{N} d_j}{N} = A + \frac{\sum d}{N} \qquad (5)$$
$$\bar{X} = A + \frac{\sum_{j=1}^{K} f_j d_j}{\sum_{j=1}^{K} f_j} = A + \frac{\sum fd}{N} \qquad (6)$$
donde $N = \sum_{j=1}^{N} f_j = \sum f$. Obsérvese que las fórmulas (5) y (6) se resumen en la ecuación $\bar{X} = A + \bar{d}$ (ver problema 3.18).
</td></tr>
<tr><td>ARGUMENTOS</td><td>
Los algoritmos se explican a través de diversos ejemplos y problemas resueltos. Algunas propiedades se deducen a partir de los ejercicios y otros son demostrados usando un razonamiento algebraico. En ambos casos, se hace uso del razonamiento verbal para justificar los resultados y/o demostraciones.
</td></tr>
</table>

Comentario: El análisis muestra que los libros de referencia son inadecuados en el sentido que enfatizan en el algoritmo de cálculo y no en el concepto y sus propiedades.

4.1.2 DESCRIPCIÓN DETALLADA DE LOS ELEMENTOS DE SIGNIFICADO Y EJEMPLOS ESPECÍFICOS ENCONTRADOS.

En esta parte se describe, de forma detallada, a cada uno de los diferentes lenguajes, situaciones-problema, lenguajes, definiciones, propiedades, algoritmos y argumentos que se encontraron en el análisis de los libros de texto, relativo a la media, mediana y moda. Se presentan, además, en casa caso, algunos ejemplos de los mismos libros para que ayuden a clarificar las ideas que describen a cada uno de los elementos.

LENGUAJE

Presentamos las palabras, notaciones y representaciones que se han encontrado del objeto abstracto *medidas de tendencia central,* usadas para representar a dichos conceptos y sus propiedades, así como para describir los problemas y sus datos. Para la clasificación nos basamos en la distinción que hace Cobo, B. (2003).

Términos y expresiones verbales

<table>
<tr><th>Libro de texto A</th><th>Específicas de la matemática</th><th>Aparecen en matemáticas y el lenguaje cotidiano pero no siempre con el mismo significado</th><th>Significados iguales o próximos en ambos contextos</th></tr>
<tr><td colspan="2">Medidas de tendencia central, estadístico, distribución de frecuencias, distribución de frecuencias agrupadas y no agrupadas, frecuencia absoluta y relativa, escala (vertical y horizontal), histograma, diagrama de barras, diagrama de pastel, diagrama de caja y de brazos, diagrama de tallo y hojas, diagrama de pareto, gráficos de puntos, datos</td><td>Media, mediana, moda, rango medio, dispersión, datos, tendencia, población, mínimo, máximo, ordenados, solución, valor, normal, campana, simétrica, desviación (respecto a la media), media aritmética, estimar, promedio, distribución, profundidad o posición, ocurrencias.</td><td>Central, centro, mitad, variación, creciente, decreciente, ascendente, descendente, menor, mayor, punto de equilibrio, punto de gravedad, observaciones, tamaño, amplitud, longitud, ancho, frecuente, típico, representativo, fórmula.</td></tr>
</table>

cuantitativos, datos cualitativos, par, impar, números de clase, intervalos de clase, límite de clase, extremos, marca de clase o punto medio de la clase, clase modal, , media muestral, media poblacional, mediana muestral, mediana poblacional.		

TABLA 2. Términos y expresiones verbales encontradas en el texto A.

Libro de texto B	**Específicas de la matemática**	**Aparecen en matemáticas y el lenguaje cotidiano pero no siempre con el mismo significado**	**Significados iguales o próximos en ambos contextos**
Medidas de tendencia central, distribución de frecuencias, variable, media aritmética ponderada, factores de ponderación o pesos, intervalo de clase, marca de clase o punto medio de clase, clase mediana, clase modal, curvas de frecuencia, sesgadas o asimétricas, simétricas, uniforme, campana, unimodal, bimodal, histograma, polígono de frecuencias,		*Media, mediana, moda, promedio, orden, ordenación, magnitud, media aritmética, clase, datos agrupados, desviación (respecto a la media), punto de equilibrio, valor central.*	*Central, centro, menor, mayor, representativo, típico, frecuencia, fórmula, amplitud.*

TABLA 3. Términos y expresiones verbales encontradas en el texto B.

Notaciones y símbolos

Las notaciones y símbolos son usadas para representar a los conceptos y realizar las operaciones entre los mismos. Nos permiten denotar de manera abreviada y precisa a los conceptos y proposiciones. En la siguiente tabla presentamos las notaciones y símbolos encontrados en nuestro análisis del texto.

Libro de texto A			
	Relacionadas a la media	**Relacionadas a la mediana**	**Relacionadas a la moda**
Se representa por	$\bar{x}$, μ,	M , $\tilde{x}$, Q_2 , P_{50} Me, D_5	Moda Mo
Símbolos que usan para el estudio	$\sum x$, $\sum f$, $\sum xf$	$d(\tilde{x}) = \frac{n+1}{2}$	
Fórmulas establecidas	$\bar{x} = \frac{\sum x}{n}$, $\bar{x} = \frac{\sum xf}{\sum f}$		

TABLA 4. Notaciones y símbolos encontrados en el texto A.

Libro de texto B			
	Relacionadas a la media	**Relacionadas a la mediana**	**Relacionadas a la moda**
Se representa por	$\bar{X}$	$\tilde{X}$, P_{50} Q_2 , D_5	$\hat{X}$

Símbolos que usan para el estudio	$f_j,\ X_j,\ d_j = X_j - A$ $(A\ es\ cualquier\ número\ supuesto\ como\ media\ aritmética)$ $d_j = cu_j$ $(c\ es\ la\ amplitud\ de\ los\ intervalos\ de\ clase)$ $\sum_{j=1}^{k} X_j,\ \sum X,$ $\sum_{j=1}^{k} f_j = \sum f = N$ $\sum_{j=1}^{k} (f_j X_j),\ \sum fX$	**Posición que ocupa el dato central:** **si n es impar:** $\frac{n+1}{2}$ **si n es par:** entre $\frac{n}{2}\ y\ \frac{n}{2} + 1$	
Fórmulas establecidas	$\bar{X} = \frac{X_1 + X_2 + \cdots X_n}{N},$ $\bar{X} = \frac{\sum_{j=1}^{k} X_j}{N},$ $\bar{X} = \frac{\sum_{j=1}^{k} f_j X_j}{\sum_{j=1}^{k} f_j}$ $\bar{X} = \frac{\sum fX}{\sum f}$		

TABLA 5. Notaciones y símbolos encontrados en el texto B.

Tablas estadísticas y gráficos

En esta sección presentamos las tablas y gráficos (diagramas) que hemos encontrado en nuestro análisis. Dichas tablas y gráficos aparecen, bien como parte del tema medidas de tendencia central o en otros que involucran a estos conceptos, por ejemplo el tema medidas de dispersión.

Libro de texto A	
Tablas	**Gráficos (diagramas)**
❖ Listado de datos: ▪ Ordenados/no ordenados ❖ Tablas de frecuencia: ▪ Datos agrupados/no agrupados ▪ Intervalos iguales ▪ Extremos del intervalo coincidentes ▪ Intervalos semi-cerrados ▪ Un intervalo cerrado.	❖ Diagrama de barras ❖ Diagrama de pareto ❖ Diagrama de caja y brazos. ❖ Diagrama de tallo y hojas ❖ Gráfico de puntos ❖ Diagrama de pastel ❖ Histogramas

TABLA 6. Tablas estadísticas y gráficos encontrados en el texto A.

Libro de texto B	
Tablas	**Gráficos**
❖ Listado de datos: ▪ Ordenados/no ordenados ❖ Tablas de frecuencia: ▪ Datos agrupados ▪ Intervalos iguales ▪ Intervalos semi-cerrados ▪ Un intervalo cerrado.	❖ Histograma ❖ Curvas de frecuencia simétricas y asimétricas.

TABLA 7. Tablas estadísticas y gráficos encontrados en el texto B.

Gráficas (recta numérica balanceada en un punto de apoyo)

La media puede ser vista como el punto de equilibrio de una balanza, indicando así su caracterización como centro de gravedad de la distribución. Hemos encontrado esta representación física de la media a través del siguiente gráfico (p.57, texto A)

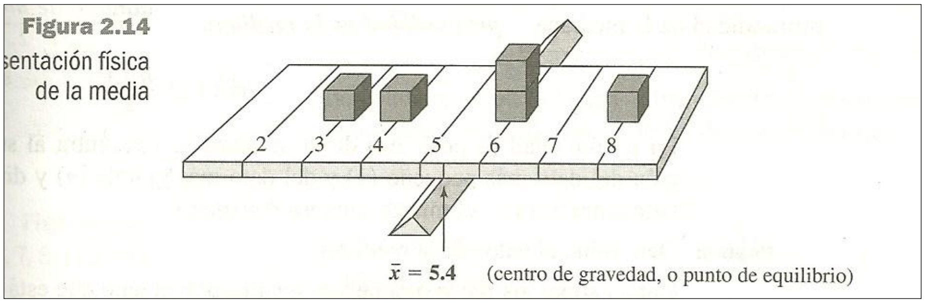


FIGURA 2. Media como centro de gravedad o punto de equilibrio.

SITUACIONES-PROBLEMA

En esta parte, presentamos las diversas situaciones-problema encontradas en nuestro análisis y, en cuya solución, emergen los objetos matemáticos, media, mediana y moda.

Situaciones-problema asociadas a la media

PM1. Estimar una medida a partir de diversas mediciones realizadas en presencia de errores.

En muchas situaciones necesitamos medir una cantidad X desconocida de una cierta magnitud; por ejemplo, la longitud de una pequeña varilla en milímetros, el peso de cierto instrumento en gramos, etc. Ahora bien, debido a la imperfección de los instrumentos que usamos para realizar estas mediciones, hace que, en mediciones sucesivas, obtengamos distintos números como medidas de X, y al no tener ninguna razón para pensar que el verdadero valor esté más cercano a uno u otro de los datos obtenidos, nos lleva a buscar el valor más representativo del conjunto de datos. Dicho valor, se consigue calculando la suma total de las medidas y dividiendo por el número de datos. Tal medida es conocida como la media. Se toma a la media como solución por sus propiedades como estimador, tales como ser insesgado o tener mínima varianza.

Este tipo de situaciones se encontraron en ambos libros de texto ya sea como ejemplo o como ejercicio propuesto.

Texto A (p.97 “ejercicio propuesto”)

2.126. Se supone que la gasolina bombeada de un gasoducto de suministro tiene un octanaje nominal de 87.5. Durante 13 días consecutivos se tomó una muestra que fue analizada y se encontraron los siguientes resultados:

88.6	86.4	87.2	88.4	87.2	87.6	86.8
86.1	87.4	87.3	86.4	86.6	87.1	

a. **Encuentre la media de la muestra.**
b. Encuentre la desviación estándar de la muestra
c. ¿Considera usted que estos porcentajes parecen promediar 87.5? Explique su respuesta. (retenga estas soluciones para resolver el ejercicio 9.42.)

Texto B (p.70 “ejemplo”)

3.7 Un científico mide diez veces el diámetro de un cilindro y obtiene los valores 3.88, 4.09, 3.92, 3.97, 4.02, 3.95, 4.03, 3.92, 3.98 y 4.06 centrímetros (cm). Hallar la media aritmética de estas mediciones.

SOLUCIÓN

$$\bar{X} = \frac{\sum X}{N} = \frac{3.88 + 4.09 + 3.92 + 3.97 + 4.02 + 3.95 + 4.03 + 3.92 + 3.98 + 4.06}{10} = \frac{39.82}{10} = 3.98 \text{ cm}$$

PM3. Obtener un elemento representativo de un conjunto de valores dados cuya distribución es aproximadamente simétrica.

Otra aplicación típica de la media consiste en servir de elemento representativo de un conjunto de datos x_i , $i = \overline{1,n}$ cuya distribución es aproximadamente simétrica. Esto se deriva por la propiedad de localización central que posee la media (la cual veremos más adelante). Si la distribución es asimétrica, convendrá considerar al valor más frecuente (moda) o el valor central del conjunto de datos ordenados (mediana) como el valor más representativo.

Este tipo de situaciones se encontraron en ambos libros de texto

Texto A (p.70 “ejercicio propuesto”)

Ejercicio 2.65. A los reclutas de una academia de policía se les solicitó presentar un examen que mide su capacidad para hacer ejercicio. Esta capacidad (medida en minutos) se obtuvo para cada uno de los 20 reclutas:

25	27	30	33	30	32	30	34	30	27
26	25	29	31	31	32	34	32	33	30

a. Trace una gráfica de puntos de los datos

b. Encuentre la media

c. Encuentre el rango

d. Encuentre la varianza

e. Encuentre la desviación estándar

Texto B (p.70 “ejemplo”)

3.6 Las calificaciones de un estudiante en seis exámenes fueron 84, 91, 72, 68, 87 y 78. Hallar la media aritmética de estas calificaciones.

SOLUCIÓN

$$\bar{X} = \frac{\sum X}{N} = \frac{84 + 91 + 72 + 68 + 87 + 78}{6} = \frac{480}{6} = 80$$

El término *promedio* suele emplearse como sinónimo de *media aritmética*. Sin embargo, estrictamente hablando, esto no es correcto, ya que además de la media hay otros promedios.

Situaciones-problema asociadas a la mediana

PME1. Encontrar un resumen estadístico de posición central, en situaciones en las que la media no es suficientemente representativa.

Existen situaciones donde la media de una colección de datos no aporta una información suficiente o adecuada. Esto ocurre cuando la distribución es asimétrica (la media está muy alejada del centro) o tiene valores atípicos. Podríamos pensar en la moda como un valor representativo, pero si dicha colección presenta varias modas, lo más conveniente será dividir al conjunto de datos, previamente ordenado, en dos partes, con el mismo número de datos, buscando un valor central llamado mediana. El siguiente ejemplo (p.97) muestra una situación de este campo de problemas. Al calcular la media resulta $\bar{x} = 20{,}03$ y este valor no es representativo del conjunto de datos, pues ha sido afectado por la dispersión de los mismos, alejándose así del centro. La mediana $M_e = 17$ se presenta como el mejor representante.

(Texto A, p. 97)

2.127 El siguiente conjunto de datos proporciona las edades de 118 conocidos delincuentes que cometieron el robo de un automóvil el año pasado en Garden City, Michigan.

11	14	15	15	16	16	17	18	19	21	25	36
12	14	15	15	16	16	17	18	19	21	25	39
13	14	15	15	16	17	17	18	20	22	26	43
13	14	15	15	16	17	17	18	20	22	26	46
13	14	15	16	16	17	17	18	20	22	27	50
13	14	15	16	16	17	17	19	20	23	27	54
13	14	15	16	16	17	18	19	20	23	29	59
13	15	15	16	16	17	18	19	20	23	30	67
14	15	15	16	16	17	18	19	21	24	31	
14	15	15	16	16	17	18	19	21	24	34	

a. Encuentre la media. **b.** Encuentre la mediana. **c.** Encuentre la moda.
d. Encuentre Q_1 y Q_3. **e.** Encuentre P_{10} y P_{95}.

(Texto B, p.77)

3.25 En los resultados de MINITAB, a continuación, se presenta el tiempo, por semana, que 30 usuarios de Internet pasaron haciendo búsquedas, así como la mediana de estos 30 tiempos. Verificar la mediana. ¿Se considera que este promedio es típico (representativo) de estos 30 tiempos? Compárense los resultados con los hallados en el problema 3.8.

```
MTB > print cl
```

Muestra de datos

```
tiempo
3   4   4   5   5   5   5   5   5   6
6   6   6   7   7   7   7   7   8   8
9  10  10  10  10  10  10  12  55  60

MTB > median cl
```

Mediana de columna

```
Median of time = 7.0000
```

SOLUCIÓN

Obsérvese que los dos valores de en medio son 7 y que la media de estos dos valores de en medio es 7. En el problema 3.8 se encontró que la media es 10.4 horas. La mediana es más típica (representativa) de estos tiempos que la media.

Comentario: es evidente que la media (10,4 horas) está muy alejado del centro, esto, debido a que se vio afectada por los valores atípicos 55 y 60 horas.

Situaciones-problema asociadas a la moda

PMO1. Obtener como valor representativo de un conjunto de datos, el más frecuente de ellos, en situaciones en las que lo que interesa es el valor dominante del conjunto.

En situaciones en las que solo se desea conocer el dato que más se repite dentro de un conjunto de datos, se hace uso de la moda, la cual no es difícil de calcular, pues solo se necesita observar el número de veces que se repite cada dato (frecuencia) y proceder a clasificarlos; así, la moda es el dato que se presenta con mayor frecuencia. Cobo (2003) aclara que la moda depende de todas las observaciones sólo por su frecuencia absoluta y no por su valor, y además, el mayor inconveniente que presenta es que si la variable es continua no está bien definida puesto que la clase modal dependerá de la amplitud de los intervalos de clase. Otra cuestión es que algunas distribuciones pueden presentar más de una moda. Aún así, su uso es recomendable, en situaciones en la cual no se requiere una gran precisión. Ambos libros de texto presentan este tipo de situaciones.

El siguiente es un ejemplo de este campo.

Texto A (p.60).

> **Ejercicio 2.42** El número de automóviles por departamento que pertenecen a una muestra de propietarios en un gran complejo es 1, 2, 2, 2, 1, 2, 3, 2. ¿Cuál es la moda?

DEFINICIONES

Definiciones de la media

DM1. Definición de media como la suma ponderada de cada uno de los valores de la variable multiplicada por su peso o frecuencia. En nuestro análisis, encontramos esta definición cuando se les asignan a los datos ciertos factores de ponderación (pesos) o cuando son presentados en una tabla de frecuencias.

Texto B (p.62)

MEDIA ARITMÉTICA PONDERADA

Algunas veces, a los números $X_1, X_2, \ldots, X_K$ se les asignan ciertos *factores de ponderación* (o *pesos*) $w_1, w_2, \ldots, w_K$ que dependen del significado o importancia que se les asigne a estos números. En este caso, a

$$\bar{X} = \frac{w_1X_1 + w_2X_2 + \cdots + w_KX_k}{w_1 + w_2 + \cdots + w_K} = \frac{\sum wX}{\sum w} \qquad (3$$

se le llama *media aritmética ponderada.* Obsérvese la semejanza con la ecuación (*2*), la cual se puede considerar com una media aritmética ponderada con pesos $f_1, f_2, \ldots, f_K$.

Texto A (p.73).

media de distribución de frecuencias: $\boldsymbol{x\ barra} = \dfrac{\boldsymbol{suma\ de\ x,\ usando\ frecuencias}}{\boldsymbol{número,\ usando\ frecuencias}}$

$$\bar{x} = \frac{\Sigma xf}{\Sigma f} \qquad \textbf{(2.10)}$$

DM2. Definición de media como promedio aritmético de un conjunto de datos. Esta definición hace referencia a la característica de la media como valor central.

Texto A (p.56).

MEDIA (media aritmética)

Promedio que quizá sea el más conocido. Se representa por $\bar{x}$ (que se lee como **"x barra"** o **"media muestral"**). La media se encuentra sumando todos los valores de la variable *x* (la suma de valores *x* se simboliza como $\sum x$) y dividiendo entre el número de estos valores, *n* ("tamaño de muestra"). Lo anterior se expresa con una fórmula como

media muestral: $\boldsymbol{x\ barra} = \dfrac{\boldsymbol{suma\ de\ x}}{\boldsymbol{número\ de\ x}}$

$$\bar{x} = \frac{\Sigma x}{n} \qquad \textbf{(2.1)}$$

Texto B (p.62)

La *media aritmética,* o brevemente la *media,* de un conjunto de N números $X_1, X_2, X_3, \ldots, X_N$ se denota así: $\bar{X}$ (que se lee "X barra") y está definida como

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \cdots + X_N}{N} = \frac{\sum_{j=1}^{N} X_j}{N} = \frac{\sum X}{N} \qquad (1)$$

Definiciones de la mediana

DME1. Mediana como centro de la distribución de datos. Esta definición enfatiza la idea de la mediana como centro de la distribución previamente ordenado. Presentamos las siguientes definiciones:

Texto A (p.58).

MEDIANA
Valor de los datos que ocupa la posición central cuando los datos se ordenan según su tamaño. Se representa por x (se lee como "x tilde" o "mediana muestral").

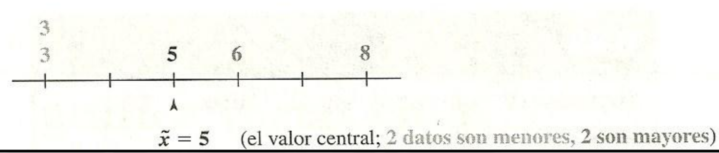
Texto B (p.64)

La *mediana* de un conjunto de números acomodados en orden de magnitud (es decir, en una ordenación) es el valor central o la media de los dos valores centrales.

DME2. Mediana como valor de la variable estadística que divide en dos subconjuntos de igual tamaño. Esta definición resalta la idea de la mediana como el valor del conjunto que divide al mismo en dos partes iguales. En nuestro análisis no hemos encontrado una definición de este tipo, sin embargo, en el texto A, cuando se estudia el procedimiento para encontrar el valor de la mediana, se llama dos veces la atención al lector respecto a esta característica:

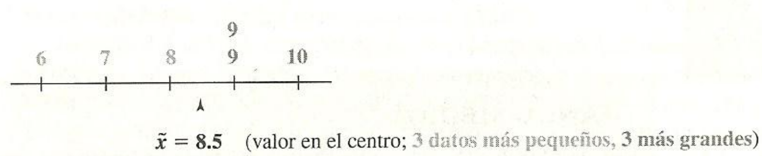
"Observe que la mediana esencialmente separa al conjunto de datos ordenados en dos subconjuntos de igual tamaño" (p.58).

Figura 2.15
{3, 3, 5, 6, 8}



"Observe que de nuevo la mediana separa al conjunto de datos ordenados en dos subconjuntos de igual tamaño" (p.59).

Figura 2.16
6, 7, 8, 9, 9, 10}



DME3. Mediana como valor que define a la recta vertical que divide al histograma en dos regiones de áreas iguales. Esto resulta de una interpretación geométrica de la

mediana. No aparecen en otras investigaciones. Esta definición sólo la encontramos en el texto B (p.64)

> Geométricamente, la mediana es el valor de X (abscisa) que corresponde a una recta vertical que divide al histograma en dos partes que tienen la misma área. A este valor de X se le suele denotar $\tilde{X}$.

Definiciones de la moda

DMO1. La moda es el valor más frecuente de la variable estadística. Esta definición nos dice que la moda es el valor que más se repite. Es una definición sencilla y fácil de entender. Encontramos las siguientes definiciones:

Texto A (p.60)

> **MODA**
> Es el valor de *x* que ocurre más frecuentemente.

Texto B (p.64)

> La *moda* de un conjunto de números es el valor que se presenta con más frecuencia; es decir, es el valor más frecuente. Puede no haber moda y cuando la hay, puede no ser única.

PROPIEDADES

La media, mediana y moda, como representantes de un conjunto de datos, nos brindan cierta información de cómo se comportan los datos en una distribución; sin embargo, resulta necesario saber cuándo una de estas medidas resulta ser más representativa o adecuada frente a otra, pues cada una de ellas posee ciertas características particulares. Para esto, será necesario conocer las propiedades de estas medidas, que nos darán una mejor información para el buen manejo las mismas. Siguiendo el modelo de Cobo (2003), las propiedades se han clasificado como: numéricas, algebraicas y estadísticas.

Propiedades Numéricas

N1. La media, mediana y moda de un conjunto de datos son siempre valores pertenecientes al rango de la variable. La media, mediana y moda nunca resultan mayores al máximo de los datos, ni menores al mínimo de los mismos. Esta propiedad sólo aparece de manera implícita en los diversos ejemplos que se muestran sobre las medidas de tendencia central.

N2. Un representante de un conjunto de datos puede ser uno de ellos (caso de la moda o de la mediana, si hubiera un número impar de datos) o algún valor, que no necesariamente coincide con uno de los datos, obtenido por un promedio adecuado (la media aritmética o la mediana, cuando se tiene un número par de datos).

Los ejemplos que se presentan en *AM1 y AM2 (p.53)* muestran claramente que la media y la mediana no necesariamente coincide con uno de los valores de los datos ni tampoco pertenecen al mismo conjunto numérico, pues en ellos se ve que, mientras los datos son números enteros, la media y mediana son números decimales. No ocurre lo mismo con la moda que, por definición, tiene que ser uno de los valores de la distribución.

No hemos encontrado una declaración explícita de la siguiente propiedad.

N3. En el cálculo de la media se tienen en cuenta todos los valores de los datos, pero no en la mediana y en la moda. Encontramos el siguiente ejercicio propuesto (texto A, p.96) donde, implícitamente, aparece esta propiedad. Originalmente, para la muestra A, tendríamos: $\bar{x} = 5.17,\ M_e = 5\ y\ M_o = 5.$ Ahora, si cambiáramos el 8 por el 9, el único promedio que cambiaría sería la media. En situaciones cuando los datos sean homogéneos, resulta conveniente escoger a la media como representante, pues ésta utiliza todos los datos. En comparación, la mediana solo considera el orden de los datos, pero no su magnitud.

2.123 En la siguiente tabla se presentan dos muestras A y B. Observe que las dos muestras son iguales, excepto que el 8 en A ha sido sustituido por un 9 en B.

A:	2	4	5	5	7	8
B:	2	4	5	5	7	9

¿Cuál es el efecto de cambiar el 8 por e 9 sobre cada una de las siguientes estadísticas?

a. **Media**
b. **Mediana**
c. **Moda**
d. Rango medio
e. Rango
f. Varianza
g. Desviación estándar

N4. El valor numérico de la media cambia cuando se cambia cualquier dato, o cuando se añade un dato (incluso el cero) al conjunto y todos los otros se mantienen igual, salvo que el dato añadido sea igual al de la media anterior.

En nuestro análisis, encontramos una sección en el libro que habla sobre *"El arte del engaño estadístico"* (texto A, p.94) y en ella se analiza la siguiente situación:

> *"Suponga que una pequeña empresa de negocios emplea a ocho personas que ganan entre $300 y $350 a la semana. El dueño de la empresa se remunera así mismo con $1250 a la semana, y reporta al público que el salario medio que reciben los empleados de su empresa es de $430 a la semana…"*

Paralelo a esto, en un pequeño cuadro, el autor formula al estudiante la siguiente pregunta: *¿es posible que ocho empleados ganen entre $300 y $350 y que uno gane $1250 a la semana y que la media sea $430? Compruebe su respuesta.*

Se espera que el alumno, haciendo uso del algoritmo *AM7 (p.56)* responda de manera afirmativa a esta pregunta.

El autor señala que el reporte dado por el dueño de la empresa es un ejemplo de *mala estadística* (entendemos como mal uso de la estadística), pues el público pensaría que los empleados ganan alrededor de $430 a la semana.

De todo esto debemos notar que, mientras hay ocho personas que ganan entre $300 y $350, hay uno que gana $1250, valor que está bastante alejado de los demás (valor atípico) y que afecta significativamente a la media debido a la propiedad que estamos estudiando. Por lo tanto, en distribuciones asimétricas o que presenten datos atípicos la media no resulta ser un buen representante.

Propiedades Algebraicas

A5. Las medidas de tendencia central conservan los cambios de origen y escala. Si sumamos o multiplicamos a cada uno de los datos por cierto valor; la media, mediana y moda son también sumados o multiplicados por ese mismo valor. Encontramos el siguiente ejercicio propuesto (texto A, p.107) donde se espera que el alumno deduzca esta propiedad, además de analizar si es también válida para la desviación estándar.

2.11 Las hojas de respuesta de un examen fueron calificadas por computadora. Después se descubrió que a cada calificación debían sumársele dos puntos. El estudiante A pensó que a la calificación promedio también debía incrementársele en dos puntos. El estudiante B agregó que a la desviación estándar también debía incrementársele dos puntos. ¿Quién tiene razón? Justifique su respuesta.

A6. La media de la suma de dos o más variables, es igual a la suma de las medias de dichas variables. Esta propiedad sólo se cumple cuando las variables tienen el mismo número de elementos. La encontramos solo en el texto B (p.73)

Si $Z_1 = X_1 + Y_1$, $Z_2 = X_2 + Y_2, \ldots, Z_N = X_N + Y_N$, probar que $\bar{Z} = \bar{X} + \bar{Y}$.

SOLUCIÓN

Por definición

$$\bar{X} = \frac{\sum X}{N} \qquad \bar{Y} = \frac{\sum Y}{N} \qquad \bar{Z} = \frac{\sum Z}{N}$$

Por lo tanto $\bar{Z} = \frac{\sum Z}{N} = \frac{\sum (X+Y)}{N} = \frac{\sum X + \sum Y}{N} = \frac{\sum X}{N} + \frac{\sum Y}{N} = \bar{X} + \bar{Y}$

en donde los subíndices de X, Y y Z se han omitido y donde $\sum$ significa $\sum_{j=1}^{N}$.

A7. La moda puede no existir o, si existe, no ser única. La media y la mediana siempre existen en datos numéricos. Para el texto A (p.60) hemos encontrado explícitamente lo siguiente: *"Si dos o más valores de una muestra están empatados en cuanto a mayor frecuencia (número de ocurrencias), se dice que no hay moda".* Esto nos dice que la moda puede no existir, y si existe, entonces debe ser única. Por el contrario, en el texto B, encontramos lo siguiente: *"La moda de un conjunto de números es el valor que se presenta con más frecuencia; es decir, es el valor más frecuente. Puede no haber moda y cuando la hay, puede no ser única".* Además, especifica que, si la distribución posee sólo una moda, se llamará *unimodal* y si hubiere dos modas *bimodal*.

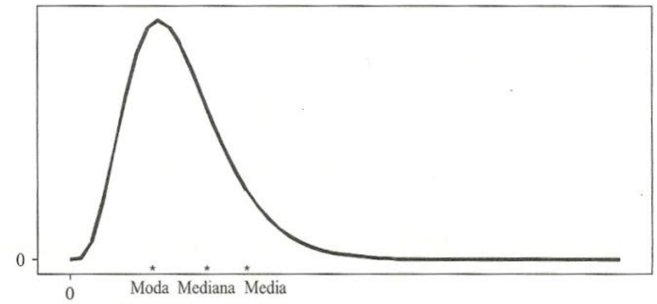
Las siguientes propiedades no aparecen en investigaciones anteriores y la hemos encontrado en el texto B.

A8. Relación empírica entre la media, mediana y moda. En las curvas de frecuencias unimodales que son ligeramente sesgadas (asimétricas), se tiene la siguiente relación empírica (p.64):

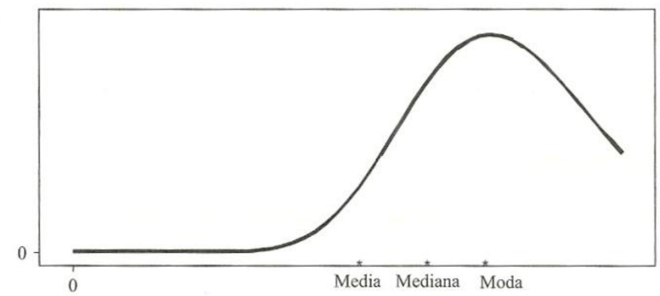
$$\boxed{Media - moda = 3(media - mediana)}$$

Se brinda, además, la siguiente idea geométrica (p.65).

las figuras 3-1 y 3-2 se muestran las posiciones relativas de la media, la mediana y la moda en curvas de fre-
as sesgadas a la derecha o a la izquierda, respectivamente. En las curvas simétricas, la media, la mediana y la
coinciden.



ura 3-1 **Posiciones relativas de la media, la mediana y la moda en curvas de frecuencias sesgadas a la derecha.**



ra. 3-2 **Posiciones relativas de la media, la mediana y la moda en curvas de frecuencias sesgadas a la izquierda.**

FIGURA 3. Posiciones relativas de la media, mediana y moda.

A9. Relación entre la media aritmética, geométrica y armónica. Se encontró lo siguiente: La media geométrica de un conjunto de números positivos $X_1,\ X_2, \ldots, X_N$ es menor o igual que su media aritmética, pero mayor o igual que su media armónica. En símbolos,

$$H \leq G \leq \bar{X}$$

La igualdad es válida sólo cuando todos los números $X_1,\ X_2, \ldots, X_N$ son idénticos.

Propiedades Estadísticas

E1. La media, mediana y moda son representantes de un colectivo. Estas medidas nos brindan información de todo el conjunto y no de uno de ellos en particular. Hemos encontrado expresamente lo siguiente:

"Las medidas de tendencia central son valores numéricos que localizan, de alguna manera, el centro de un conjunto de datos" (texto A, p.56).

“Los promedios son valores representativos de un conjunto de datos. Dichos promedios tienden a encontrarse en el centro de dicho conjunto. Se pueden definir varios tipos de promedios; los más usados son la media aritmética, la mediana, y la moda” (texto B, p.62)

E2. La media coincide con el centro de un conjunto de datos, semejante al centro de la gravedad. En nuestro análisis esta propiedad se encontró solo en el texto A, especificada así:

> *“Una representación física de la media puede obtenerse al pensar en una recta numérica balanceada en un punto de apoyo […] La media es el valor que equilibra los pesos sobre la recta numérica.” (p.57)*

En este sentido la media está en el centro del conjunto de datos y por ello puede identificarse como el centro de gravedad o punto de equilibrio.

E3. En distribuciones simétricas, la media coincide con la mediana y la moda (en distribuciones unimodales). En una distribución de frecuencias unimodales cuyos valores se distribuyen más o menos simétricamente, las tres medidas de tendencia central resultan ser representativas al conjunto de datos observados de dicha distribución. Si la distribución fuera perfecta, estas medidas coinciden. Esta propiedad sólo aparece en el texto B y la expresión que hace referencia a la misma aparece en la grafica mostrada en la propiedad algebraica *A8* donde se menciona que: *“En las curvas simétricas, la media, la mediana y la moda coinciden”.*

E4. La media es un estadístico poco resistente, muy sensible a la variación de los datos, especialmente en los valores atípicos.

Ya habíamos estudiado en *N4* que el valor de la media cambia significativamente cuando se presentan valores atípicos, esto nos muestra que esta medida es poco resistente y muy sensible a la variación de los datos. En estos casos, la mediana se presenta como un mejor representante pues solo considera el o los valores centrales de la distribución. Otra buena opción sería la moda, pues tanto la mediana como la moda son estadísticos más resistentes.

Esta propiedad aparece también en el texto B, expresada así: *"Una gran desventaja de la media es que es fuertemente afectada por los valores atípicos (o valores extremos)".*

E5. La suma de las desviaciones de los datos de un conjunto respecto de su media es cero. La diferencia entre cada dato y la media se llama *desviación* respecto de la media. La suma de las desviaciones de todos los datos es cero. Esta propiedad se desprende directamente del concepto de media como punto de equilibrio. En ambos textos se hacen explícitas esta propiedad.

Texto A (p.66).

Ejercicio 2.57

La suma $\sum(x - \bar{x})$ siempre es igual a cero. ¿Por qué? Consulte la definición de media (p. 56) y considere si puede justificar esta afirmación.

También encontramos una nota donde se confirma esta propiedad, así como también se hace referencia a la posibilidad que la suma de las desviaciones no sea exactamente cero.

Texto A (p.67).

NOTAS

1. Para encontrar $\bar{x}$ se usa la suma de todas las x.
2. En el supuesto de que se use el valor exacto de $\bar{x}$, la suma de las desviaciones, $\sum(x - \bar{x})$ siempre es cero. Aplique este hecho para comprobar sus cálculos, como se hizo en la tabla 2.8.
3. Si se usa un valor redondeado de $\bar{x}$, entonces $\sum(x - \bar{x})$ no siempre es exactamente cero. No obstante, estará razonablemente próxima a cero.
4. La sumas de las desviaciones al cuadrado se encuentra elevando al cuadrado cada desviación y luego sumando los valores obtenidos.

El conjunto de datos en el ejercicio 2.58 está más disperso que el conjunto de la tabla 2.8, por lo que su varianza es mayor. En la figura 2.22 se muestra una comparación de estas dos muestras.

Texto B (p.73)

3.16 Probar que la suma de las desviaciones de $X_1, X_2, \ldots, X_N$ respecto a su media $\bar{X}$ es igual a cero.

SOLUCIÓN

Sean $d_1 = X_1 - \bar{X}$, $d_2 = X_2 - \bar{X}, \ldots, d_N = X_N - \bar{X}$ las desviaciones de $X_1, X_2, \ldots, X_N$ de su media, $\bar{X}$. Entonces

$$\text{La suma de las desviaciones} = \sum d_j = \sum (X_j - \bar{X}) = \sum X_j - N\bar{X}$$
$$= \sum X_j - N\left(\frac{\sum X_j}{N}\right) = \sum X_j - \sum X_j = 0$$

donde se usa $\sum$ en vez de $\sum_{j=1}^{N}$. Si se desea, también se puede omitir el subíndice j de X_j siempre que éste se *sobreentienda*.

E6. Respecto a la media, la suma de los cuadrados de las desviaciones es mínima. La suma de los cuadrados de las desviaciones de los datos respecto a la media es menor que la de las desviaciones respecto a cualquier otro valor. Esta propiedad sólo aparece en el texto B (p.63)

En un conjunto de números X_j, la suma de los cuadrados de sus desviaciones respecto a un número a es un mínimo si y sólo si $a = \bar{X}$ (ver el problema 4.27).

ALGORITMOS Y PROCEDIMIENTOS

Existen diversas técnicas de cálculo para resolver los problemas relacionados con los promedios. Dichas técnicas van de acorde a la medida de tendencia central involucrada en el problema, considerando además, los distintos tipos de variables, como también la forma en que los datos son presentados. Presentamos de manera detallada las técnicas encontradas en nuestro análisis.

Cálculo de la media

AM1. Cálculo de la media de una variable discreta con datos aislados (sin agrupar). El cálculo de la media resulta bastante simple, pues basta con sumar todos los datos y dividirlo por el número total de estos. Resulta ser una aplicación directa de la definición de media. En ambos textos aparece este algoritmo.

A modo de ejemplo, presentamos la siguiente situación en la que pide hallar la media.

(texto A, p.57).

Un conjunto de datos consta de cinco valores: 6, 3, 8, 6 y 4. Encuentre la media.

Solución

Al aplicar la fórmula (2.1) se encuentra

$$\bar{x} = \frac{\Sigma x}{n} = \frac{6+3+8+6+4}{5} = \frac{27}{5} = 5.4$$

En consecuencia, la media de esta muestra es **5.4.**

AM2. Cálculo de la media de una variable discreta con los datos presentados en una tabla de frecuencias. La media se calcula siguiendo los siguientes pasos:

1. Se multiplica cada valor de la variable (dato) por su frecuencia absoluta respectiva.
2. Se suman todos los productos obtenidos
3. Se divide la suma obtenida en el paso 2 por el número total de datos (suma de las frecuencias absolutas)

Este procedimiento requiere de una correcta interpretación y uso de la tabla por parte del alumno, así como comprender la idea de ponderación. En ambos textos aparece esta propiedad. A modo de ejemplo, presentamos la siguiente situación donde se explica este algoritmo.

Texto A (p.73)

x	f	xf
1	5	5
2	9	18
3	8	24
4	6	24
Total	28	71

Media de distribución de frecuencias: $\boldsymbol{x\ barra} = \frac{\boldsymbol{suma\ de\ x,\ \ usando\ frecuencias}}{\boldsymbol{n\acute{u}mero,\ \ usando\ frecuencias}}$

$$\bar{x} = \frac{\sum xf}{\sum f}$$

El valor medio de x para a la distribución de frecuencias de esta tabla se encuentra aplicando la fórmula

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{71}{28} = 2{,}536$$

AM3. Variable continua o discreta con datos agrupados en clases. Aquí no se conocen los valores específicos de los datos, sino sólo la clase a la que pertenecen y el número de datos en cada una de ellas. Encontramos las siguientes dos formas de calcular la media:

Texto A: usa el algoritmo *AM2* considerando el valor x como la marca de clase (llamado también punto medio), que es el valor numérico que está exactamente a la mitad de la clase correspondiente, y se obtiene al tomar los extremos de la clase y dividirlo entre dos.

Este algoritmo se presenta a través del siguiente ejemplo (p.72).

Número de clase	**Límites de clase**	**Marca de clase (x)**	f	xf	x^2f
1	[35;45[	40	2	80	3 200
2	[45;55[	50	2	100	5 000
3	[55;65[	60	7	420	25 200
4	[65;75[	70	13	910	63 700
5	[75;85[	80	11	880	70 400
6	[85;95[	90	11	990	89 100
7	[95;105[	100	4	400	40 000
			$\sum f=$ 50	$\sum xf=$ 3780	$\sum x^2f=$ 296 600

El valor medio de x para la distribución de frecuencias de esta tabla es

$$\bar{x} = \frac{\sum xf}{\sum f} = \frac{3780}{50} = 75{,}6$$

Texto B: presenta el siguiente algoritmo para el cálculo de la media aritmética cuando se tiene datos agrupados.

CÁLCULO DE LA MEDIA ARITMÉTICA PARA DATOS AGRUPADOS

Cuando se presentan los datos en una distribución de frecuencias, se considera que todos los datos que caen en un intervalo de clase dado coinciden con la marca o punto medio del intervalo. Para datos agrupados, interpretando a las X_j como las marcas de clase, a las f_j como las correspondientes frecuencias de clase, a A como cualquier marca de clase supuesta y $d_j = X_j - A$ como la desviación de X_j respecto de A, las fórmulas (*2*) y (*6*) son válidas.

A los cálculos empleando las fórmulas (*2*) y (*6*) se les suele conocer como *método largo* y *método abreviado*, respectivamente (ver los problemas 3.15 y 3.20).

Si todos los intervalos de clase son de una misma amplitud c, las desviaciones $d_j = X_j - A$ se pueden expresar como cu_j, donde u_j puede tener valores enteros positivos o negativos o cero (es decir, 0, ±1, ±2, $\pm3, \ldots$) con lo que la fórmula (*6*) se convierte en

$$\bar{X} = A + \left(\frac{\sum_{j=1}^{K} f_j u_j}{N}\right) = A + \left(\frac{\sum fu}{N}\right)c \qquad (7)$$

lo que es equivalente a la ecuación $\bar{X} = A + c\bar{u}$ (ver problema 3.21). A esta ecuación se le conoce como *método codificado* para calcular la media. Es un método muy breve recomendado para datos agrupados cuando los intervalos de clase tienen todos la misma amplitud (ver problemas 3.22 y 3.23). Obsérvese que en el método codificado los valores de la variable *X se transforman* en valores de la variable *u* de acuerdo con $X = A + cu$.

Observación: En la expresión de en medio se comete un error, ya que el paréntesis debe estar multiplicado por la constante *"c" (amplitud del intervalo).*

*AM4. Cálculo gráfico. Solo en el texto B (p.79), s*e encontró esta situación en la que se pide determinar la media a partir de la lectura de gráficos (diagrama de tallo y hojas).

En la figura 3-3 se muestra una representación de tallo y hoja que proporciona el número de muertes en accidentes de tránsito en 2005 relacionados con el alcohol en los 50 estados y Washington, D.C.

Representación de tallo y hoja: Muertes

Representación de tallo y hoja: Muertes N = 51
Leaf Unit = 10

14	0	22334556667889
23	1	122255778
(7)	2	0334689
21	3	124679
15	4	22669
10	5	012448
4	6	3
3	7	
3	8	
3	9	
3	10	
3	11	
3	12	
3	13	
3	14	7
2	15	6
1	16	
1	17	1

Figura 3-3 MINITAB, representación de tallo y hoja de las muertes en accidentes de tránsito relacionados con el alcohol.

Encontrar la media, la mediana y la moda de las muertes relacionadas con el alcohol dadas en la figura 3-3.

AM5.Cálculo mediante calculadora u ordenador. En general, se ingresan todos los valores de los datos junto con su frecuencia correspondiente, a partir de ahí, se puede calcular la media, además de otras medidas. El uso es simple, en el sentido de que el alumno no necesita entender el algoritmo, sino solo ingresar los datos de manera correcta. Hemos encontrado el uso de calculadora, MINITAB, EXCEL, STATISTIX, SPSS y el SAS, en especial, en el texto B, se usan todos estos programas sin excepción. Presentamos los siguientes ejemplos (texto A, p.57).

s siguientes comandos se calcula la media.

ITAB

oducir los datos en C1; luego,
inuar con:

oger:	**Calc > Column Statistics**
leccionar:	**Mean**
roducir:	Variable de entrada: **C1**

MICROSOFT® EXCEL

Introducir los datos en la columna A y activar una celda para la respuesta; luego, continuar con:

Escoger:	**Paste function, f_x* > Statistical > AVERAGE > OK**
Introducir:	Número 1: **(A2:A6 or select cells)** [Empezar en A1 si no se utiliza renglón de encabezado (título de la columna)]

*Método alternativo: Escoger: Insert > Function. . .

AM7. Dada una media, construir una distribución. Cuando el valor de la media es conocida y lo que se requiere es construir una distribución cuya media sea ese valor, será necesario conocer bien el uso del algoritmo *AM1*, ya que ello nos llevará a determinar los posibles valores de la distribución. La distribución más sencilla que se puede formar es considerando todos sus valores igual a la media.

Un problema que involucra este algoritmo solo se encontró en el texto A (p.64).

Ejercicio 2.54 Empezando con los valores 70 y 100, sume tres valores a la muestra, de modo que ésta tenga lo siguiente (justifique su respuesta en cada caso):

- **a. Una media de 100**
- b. Una mediana de 70
- c. Una moda de 87
- d. Un rango medio de 70
- e. Una media de 100 y una mediana de 70
- f. Una media de 100 y una moda de 87
- g. Una media de 100 y un rango medio de 70
- h. Una media de 100, una mediana de 70 y una moda de 87

El siguiente algoritmo no se encontró en investigaciones anteriores:

AM8. Media aritmética como suma de una "media supuesta A" y el promedio de las desviaciones de los datos respecto a dicha media. Se encontró este algoritmo en el libro de texto B (p.63). A partir de este algoritmo de cálculo se puede deducir la propiedad E5 (la suma de las desviaciones respecto a la media es cero), esto, cuando la media supuesta A coincide con la media $\bar{X}$ de la distribución. A continuación presentamos el algoritmo:

4. Si se *cree* o *se supone* que un número A (que puede ser cualquier número) es la *media aritmética* y si $d_j = X_j - A$ son las desviaciones de X_j de A, entonces las ecuaciones (*1*) y (*2*) se convierten, respectivamente, en

$$\bar{X} = A + \frac{\sum_{j=1}^{N} d_j}{N} = A + \frac{\sum d}{N} \tag{5}$$

$$\bar{X} = A + \frac{\sum_{j=1}^{K} f_j d_j}{\sum_{j=1}^{K} f_j} = A + \frac{\sum fd}{N} \tag{6}$$

donde $N = \sum_{j=1}^{N} f_j = \sum f$. Obsérvese que las fórmulas (*5*) y (*6*) se resumen en la ecuación $\bar{X} = A + \bar{d}$ (ver problema 3.18).

Cálculo de la mediana

AME1. Datos no agrupados en clases, número de datos impar. La mediana es el valor del dato que ocupa la posición central, respecto al total de datos, supuestos los datos ordenados según su tamaño. Este algoritmo es común en ambos libros de texto. En el libro se ejemplifica este algoritmo con la siguiente ilustración (texto A, p.58).

Encuentre la mediana del conjunto de datos {6, 3, 8, 5, 3}.

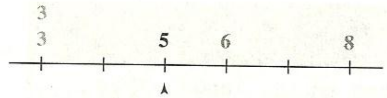
Solución

PASO 1 Los datos, ordenados de manera creciente, son 3, 3, 5, 6 y 8.

PASO 2 Profundidad de la mediana: $d(\tilde{x}) = \frac{n+1}{2} = \frac{5+1}{2} = 3$ (la "3a" posición).

PASO 3 Es decir, la mediana es el tercer número desde cualquier extremo en los datos ordenados, o bien, $\tilde{x} = 5$. Observe que la mediana esencialmente separa el conjunto de datos ordenado en dos subconjuntos de igual tamaño (vea la figura 2.15).

Figura 2.15
{3, 3, 5, 6, 8}



$\tilde{x} = 5$ (el valor central; 2 datos son menores, 2 son mayores)

Además, en el texto se hace la aclaración que $d(\tilde{x})$ es la profundidad (posición) de la mediana y no el valor de la misma. Así también, al ser el número de datos impar, el valor de la mediana siempre será uno de los datos.

AME2. Datos no agrupados en clases, número de datos par. La mediana es la semi suma de los dos valores que ocupan las posiciones centrales. Ambos textos presentan esta propiedad. El procedimiento se explica en base al siguiente ejemplo (texto A, p.59).

Encontrar la mediana de la muestra 9, 6, 7, 9, 10, 8.

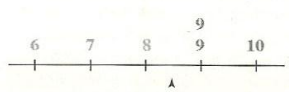
Solución

PASO 1 Los datos, ordenados de manera creciente, son 6, 7, 8, 9, 9 y 10.

PASO 2 Profundidad de la mediana: $d(\tilde{x}) = \frac{n+1}{2} = \frac{6+1}{2} = 3.5$ (la "3.5ava" posición)

PASO 3 Es decir, la mediana está a la mitad entre las porciones de datos tercera y cuarta. Para encontrar el número situado a la mitad de dos valores cualesquiera, se suman los dos valores y el resultado se divide entre 2. En este caso, se suman el tercer valor (8) y el cuarto valor (9); luego se divide entre 2. La mediana es $\tilde{x} = \frac{8+9}{2} = 8.5$, número que está a medio camino entre los dos números (vea la figura 2.16). Observe que de nuevo la mediana separa el conjunto de datos ordenados en dos subconjuntos del mismo tamaño.

Figura 2.16
7, 8, 9, 9, 10}



$\tilde{x} = 8.5$ (valor en el centro; 3 datos más pequeños, 3 más grandes)

AME3, AME4. Datos sin presentados en tablas de frecuencias; casos de un número par o impar de datos. Sólo hemos encontrado ejercicios, al final del capítulo del texto A, que involucran este algoritmo. No se explica cómo realizar el cálculo de la mediana, pero se espera el alumno sea capaz de hacerlo con las herramientas teóricas dadas hasta el momento. Incluso se le pide que muestre dicho procedimiento.

Tenemos el siguiente caso, cuando el número de datos es par (p.107).

2.7 Una muestra de las compras de varios clientes de la Corner Convenience Store dio por resultado la siguiente muestra de datos:

x = número de artículos comprados por cliente

x	f
1	6
2	10
3	9
4	8
5	7

a. ¿Qué representa el "2"?
b. ¿Qué representa el "9"?
c. ¿Cuántos clientes fueron necesarios para obtener esta muestra?
d. ¿Cuántos artículos fueron comprados por los clientes de esta muestra?
e. ¿Cuál es el mayor número de artículos comprados por un cliente?

Encuentre cada una de las siguientes estadísticas (muestre las fórmulas y el procedimiento):
f. moda **g.** mediana **h.** rango medio
i. media **j.** varianza **k.** desviación estándar

El siguiente caso se presenta cuando el número de datos es impar (p.78)

2.84 La cantidad de dinero que los adultos dicen que gastarán en regalos navideños fue descrita en "What 'Santa' Will Spend" del *USA Today* del 23 de noviembre de 1994.

Cantidad	**Nada**	**\$1-\$300**	**\$301-\$600**	**\$601-\$1 000**	**Más de \$1 000**	**No sabe**
Porcentaje que respondió	1%	24%	30%	20%	14%	11%
			Promedio: \$734			

Sea la siguiente distribución, que representa la parte de la muestra que sabía:

Cantidad, x	0	150	450	800	1500
Frecuencia	1	24	30	20	14

a. Encuentre la media de la distribución de frecuencias.
b. ¿Considera que el promedio reportado hubiese podido ser la media? Explique su respuesta.
c. Encuentre la mediana de la distribución de frecuencias.
d. ¿Considera que el promedio reportado hubiese podido ser la mediana? Explique su respuesta.
e. Encuentre la moda de la distribución de frecuencias.
f. ¿Considera que el promedio reportado hubiese podido ser la moda? Explique su respuesta.
g. ¿El promedio reportado hubiese podido ser el rango medio? En caso afirmativo, ¿cuál fue la mayor cantidad de dinero reportada?

AME5. Datos agrupados en clases. Esta propiedad se encontró en ambos textos, sin embargo, en el texto A, sólo aparece como ejercicio propuesto. Por el contrario, en el texto B, se presenta un algoritmo de cálculo, además de diversos ejemplos de aplicación.

Texto B (p.64)

En datos agrupados, la mediana se obtiene por interpolación, como se expresa por la fórmula

$$\text{Mediana} = L_1 + \left(\frac{\frac{N}{2} - (\sum f)_1}{f_{\text{mediana}}} \right) c$$

donde L_1 = frontera inferior de la clase mediana (es decir, de la clase que contiene la mediana)
N = número de datos (es decir, la frecuencia total)
$(\sum f)_1$ = suma de las frecuencias de todas las clases anteriores a la clase mediana
f_{mediana} = frecuencia de la clase mediana
c = amplitud del intervalo de la clase mediana

AME6. Datos presentados en un gráfico, diagrama de tallo y hojas. Este algoritmo se encontró en ambos libros de texto, aunque muy escasamente. *El ejemplo sería el mismo de AM4.*

Cálculo de la moda

AMO1. Variable discreta con datos aislados (sin agrupar). Se determina que valor o valores ocurre con más frecuencia. En el texto A se menciona que, si hubiera dos o más valores que presentan una igualdad respecto a la mayor frecuencia, se dice que no hay moda (esto es una convención del autor). Por su parte, en el texto B se sostiene que, si hay dos valores que presentan mayor frecuencia, la distribución es bimodal. Presentamos el siguiente ejemplo (texto A, p.60).

- En la muestra 6, 7, 8, 9, 9, 10, la moda es 9.
- En la muestra 3, 3, 4, 5, 5, 7 tanto el 3 como el 5 aparecen un número igual de veces. No hay ningún valor que aparezca con mayor frecuencia; así, esta muestra no tiene moda.

AMO2. Variable discreta con datos aislados, presentados en una tabla de frecuencias. La moda será aquél dato que presente la mayor frecuencia. Se encontraron diversos ejercicios que involucran este algoritmo. Los ejemplos pueden ser los mismos que presentamos para el algoritmo de cálculo de la mediana *AME3*. El texto B no presenta estos algoritmos.

AMO3. Datos agrupados en clases. Esta propiedad se encontró en ambos textos, sin embargo, en el texto A, sólo aparece como ejercicio propuesto. Por el contrario, en el texto B, se presenta un algoritmo de cálculo, además de algunos ejemplos de aplicación.

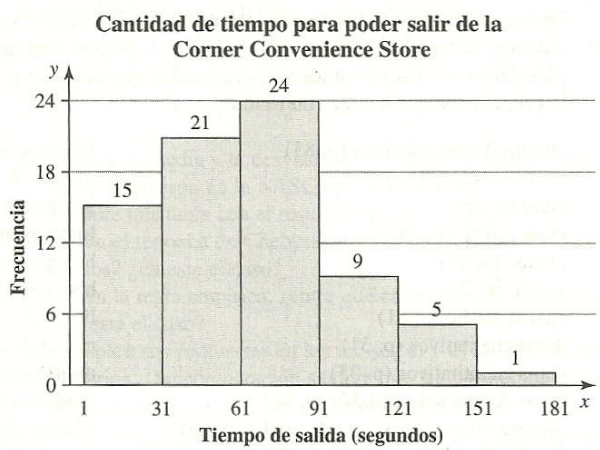
Texto B (p.64)

> En el caso de datos agrupados, para los que se ha construido una curva de frecuencia que se ajuste a los datos, la moda es el valor (o los valores) de X que corresponden al punto (o puntos) máximos de la curva. A este valor de X se le suele denotar $\hat{X}$.

AMO4. Cálculo a partir de un diagrama de barras. Dicho cálculo dependerá del tipo de variable en estudio. Para el caso del texto A, no se presenta el procedimiento que se debe de seguir. En tanto, en el texto B, se da un algoritmo de cálculo.

Texto A (p.106).

2.6 El siguiente histograma representa los resultados de un estudio sobre consumidore realizado por la Corner Convenience Store. Encuentre la respuesta para cada una d preguntas siguientes.



a. ¿Cuál es el ancho de clase?
b. ¿Cuál es la marca de clase para 31-61?
c. ¿Cuál es el límite superior para la clase 61-91?
d. ¿Cuál es la frecuencia de la clase 1-31?
e. ¿Cuál es la frecuencia de la clase que contiene el mayor valor observado de x?
f. ¿Cuál es el límite inferior de la clase con la mayor frecuencia?
g. ¿Cuántos datos muestra este histograma?
h. ¿Cuál es el valor de la moda?
i. ¿Cuál es el valor del rango medio?
j. Estime el valor del 90avo percentil, P_{90}.

Texto B (p.64)

En una distribución de frecuencia o en un histograma la moda se puede obtener mediante la fórmula siguiente:

$$\text{Moda} = L_1 + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right)c \qquad (9)$$

donde L_1 = frontera inferior de la clase modal (es decir, de la clase que contiene la moda)
Δ_1 = exceso de frecuencia modal sobre la frecuencia en la clase inferior inmediata
Δ_2 = exceso de frecuencia modal sobre la frecuencia en la clase superior inmediata
c = amplitud del intervalo de la clase modal

ARGUMENTOS

"Todos los enunciados, propiedades, problemas y algoritmos se ligan entre sí mediante argumentos o razonamientos que se usan para comprobar las soluciones de los problemas o demostrar las propiedades y relaciones" *(Alvarado, 2007).* A continuación presentamos los argumentos encontrados en nuestro análisis, respecto al estudio de las medidas de tendencia central.

ARG1. Justificación con ejemplos o contraejemplos. Las definiciones, propiedades y demás elementos de las medidas de tendencia central pueden validarse o hacerse más comprensibles por medio de ejemplos o contraejemplos. Estos ejemplos, por lo general, vienen acompañados de algún argumento que justifica su resultado.

ARG3. Razonamientos algebraicos deductivos. Las demostraciones solo aparecen en el texto B. un ejemplo sería el siguiente (p.73)

3.16 Probar que la suma de las desviaciones de $X_1, X_2, \ldots, X_N$ respecto a su media $\bar{X}$ es igual a cero.

SOLUCIÓN

Sean $d_1 = X_1 - \bar{X}$, $d_2 = X_2 - \bar{X}, \ldots, d_N = X_N - \bar{X}$ las desviaciones de $X_1, X_2, \ldots, X_N$ de su media, $\bar{X}$. Entonces

$$\text{La suma de las desviaciones} = \sum d_j = \sum (X_j - \bar{X}) = \sum X_j - N\bar{X}$$
$$= \sum X_j - N\left(\frac{\sum X_j}{N}\right) = \sum X_j - \sum X_j = 0$$

donde se usa $\sum$ en vez de $\sum_{j=1}^{N}$. Si se desea, también se puede omitir el subíndice j de X_j siempre que éste se *sobreentienda*.

ARG4. Razonamientos verbales deductivos. Según Mayén (2009) este tipo de argumentos generalmente se utilizan para interpretar los resultados obtenidos de algún cálculo realizado. En el ejemplo tomado para el algoritmo *AME2* se explica el cálculo de la mediana y concluye haciendo énfasis a la característica principal de la mediana, expresándolo así: *“Observe que de nuevo la mediana separa al conjunto de datos ordenados en dos subconjuntos del mismo tamaño.”*

4.1.3 TABLA DE RESUMEN COMPARATIVO.

A continuación presentamos, en diversas tablas, los elementos que aparecen en cada uno de los libros de texto analizados, de manera que se logre tener una visión más clara de los elementos que son tratados en uno y otro texto.

TABLA 8. Resumen comparativo de los textos analizados A y B.

SITUACIONES-PROBLEMA	Johnson y Kuby	Murray y Larry
PM1. Estimar una medida a partir de diversas mediciones en presencia de errores.	x	x
PM3. Obtener un elemento representativo de una distribución aproximadamente simétrica.	x	x
PME1. Encontrar un resumen estadístico en situaciones en las que la media nos es suficientemente representativa.	x	x
PMO1. Obtener como valor representativo de un conjunto de datos, el más frecuente de ellos, en situaciones en las que lo que interesa es el valor dominante del conjunto.	x	x

DEFINICIONES	Johnson y Kuby	Murray y Larry
DM1. Media como suma ponderada.	x	x
DM2. Media como promedio aritmético.	x	x
DME1. Mediana como valor central.	x	x
DME2. Mediana como valor que divide en dos subconjuntos de igual tamaño.	x	
DME3. Mediana como valor que define a la recta vertical que divide al histograma en dos regiones de áreas iguales.		x
DMO1. Moda como valor más frecuente.	x	x

PROPIEDADES	Johnson y Kuby	Murray y Larry
NUMERICAS		
N1. Pertenecen al rango de la variable.	x	x
N2. La moda coincide con un valor de los datos, la media y mediana no siempre.	x	x
N3. La media toma en cuenta todos los valores del conjunto, no así la mediana ni la moda.	x	x
N4. El valor numérico de la media cambia cuando se cambia cualquier dato, o cuando se añade un dato (incluso el cero) al conjunto y todos los otros se mantienen igual, salvo que el dato añadido sea igual al de la media anterior.	x	x
ALGEBRAICAS		
A5. Conservan cambios de origen y escala.	x	x
A6. La media de la suma de dos o más variables es la suma de las medias de dichas variables.		x
A7. La moda puede no existir o, si existe, no ser única.	x	x
A8. Relación empírica entre la media, mediana y moda.		x
A9. Relación entre la media aritmética, geométrica y armónica.		x
ESTADÍSTICAS		
E1. Representan a un colectivo.	x	x
E2. La media coincide con el centro del conjunto de datos.	x	x
E3. Las tres medidas coinciden en distribuciones simétricas.		x
E4. La media es bastante sensible a la variación de los datos del conjunto.	x	x
E5. La suma de las desviaciones de los datos de un conjunto respecto a la media es cero	x	x
E6. Respecto a la media, la suma de los cuadrados de las desviaciones es mínima		x

ALGORITMOS Y PROCEDIMIENTOS	Johnson y Kuby	Murray y Larry
AM1. Media para datos sin agrupar.	x	x
AM2. Media para datos presentados en tabla de frecuencias. (ponderada)	x	x
AM3. Media para datos agrupados en clases.	x	x
AM4. Cálculo de la media a partir de gráficos.		x
AM5. Cálculo de la media con calculadora u ordenador.	x	x
AM7. Dada la media, construir una distribución.	x	
AM8. Cálculo de la media a partir de una media supuesta.		x
AME1. Mediana para datos sin agrupar (n impar).	x	x
AME2. Mediana para datos sin agrupar (n par).	x	x
AME3. Mediana para datos presentados en tabla de frecuencias (n impar)	x	
AME4. Mediana para datos presentados en tabla de frecuencias (n par).	x	
AME5. Mediana para datos agrupados en clases.	x	x
AME6. Cálculo de la mediana a partir de un gráfico.	x	x
AMO1. Moda para datos sin agrupar.	x	x
AMO2. Moda para datos sin agrupar presentados en tabla de frecuencias.	x	x
AMO3 .Moda para datos agrupados en clases.	x	x
AMO4. Cálculo de la moda a partir de un gráfico.	x	x

ARGUMENTOS	Johnson y Kuby	Murray y Larry
ARG1. Justificación de ejemplos	x	x
ARG3. Razonamiento algebraico deductivo		x
ARG4. Razonamiento verbal deductivo	x	x

4.1.4 CONCLUSIONES SOBRE EL ANÁLISIS DE LOS LIBROS DE TEXTO.

1. En los libros de texto que analizamos, las medidas de tendencia central son estudiadas solo como medidas de resumen de un conjunto de datos. No se encontró situaciones que busque relacionar este concepto con la estadística inferencial. Resultaría importante, que se pueda insertar algunas situaciones y aspectos básicos de la misma, para que el alumno vaya comprendiendo que la importancia de las medidas de tendencia central no sólo consiste en calcularlas, sino más bien en usarlas para tomar decisiones acertadas frente a ciertos fenómenos.

2. Respecto a los algoritmos y procedimientos podemos notar que aparecen la mayoría de los encontrados en investigaciones anteriores. Sin embargo, son pocos los ejemplos que se presentan para explicitar dichos algoritmos; puesto que la mayoría de situaciones que involucran el uso de estos algoritmos, aparecen solo como ejercicios propuestos.

   Respecto al Texto B, podemos decir que:

3. No se hace mención a los tipos de variables y tampoco se presentan problemas que involucren a las variables cualitativas, dejando de lado un aspecto importante en cuanto al uso adecuado de los promedios según el tipo de variable en estudio.

4. Al ser un texto que presenta, casi en su totalidad, problemas resueltos, puede inducir a los estudiantes a una mera memorización y mecanización, además de no estimular el razonamiento y la creatividad para buscar otro tipo de soluciones.

5. Existe una ausencia de problemas que involucran el cálculo de la mediana y la moda para datos cualitativos, así como el cálculo de las mismas a partir de gráficos.

4.2 SIGNIFICADO INSTITUCIONAL DE REFERENCIA

Con los elementos que hemos encontrado en los textos analizados, los aportes de investigaciones como las de Cobo, B. (2003) y Mayen, S. (2009) y otras indagaciones sobre las medidas de tendencia central, fijamos y presentamos el significado institucional de referencia para nuestra investigación, que considera los elementos necesarios en el contexto de nuestro estudio.

LENGUAJE

En nuestro estudio se hizo evidente el uso de una gran variedad de términos y expresiones algebraicas en el estudio de la media, la mediana y la moda, aunque algunos de uso poco común, tal como lo habíamos mencionado y presentado en 4.1.1 y 4.1.2. Así, para nuestra investigación haremos las siguientes consideraciones necesarias:

- Cuando nos refiramos a las medidas de tendencia central, entenderemos estos como la media, la mediana y la moda, por su uso más frecuente.
- Relacionamos el término promedio con la “media”, aunque sabemos que existen otros promedios.
- Los símbolos más usuales y que usaremos para la media, mediana y moda serán $\bar{x}, M_e\ y\ M_o$, respectivamente.

SITUACIONES-PROBLEMA

Además de las situaciones-problema que se encontraron en el análisis de los textos seleccionados, existen otras más (Cobo, 2003 y Mayén, 2009), de las cuales presentamos aquellas que consideramos necesarias para nuestra investigación. Estas son:

Situaciones-problema asociadas a la media

PM2. Obtener una cantidad equitativa al hacer un reparto para conseguir una distribución uniforme. Presentamos como un resultado de Cobo 2003, p. 105. En el ejemplo que sigue, aparece la idea de distribución uniforme, y la media de los salarios se presenta como la cantidad equitativa si todos los empleados tuvieran igual salario.

El comercio
En un pequeño comercio hay cinco empleados cuyos sueldos mensuales son: 80000, 80000, 80000, 100000 y 400000 pesetas. a) Halla la moda y la mediana e interprétalas. b) ¿Cuánto cobran entre todos los empleados al mes? Si esta cantidad se repartiera por igual, ¿cuánto cobraría cada uno? ¿Alguno de ellos cobra este sueldo medio? c) ¿Qué valor de los tres: moda, mediana o media, crees que representa mejor los sueldos de los empleados de este pequeño comercio? d) ¿Cuál es el mayor sueldo? ¿Y el menor? ¿Cuál es la diferencia entre estos dos valores?
3º Matemáticas. Secundaria. Edelvives, pg.328

PM4. Conocer el valor que se obtendrá con mayor probabilidad al tomar un elemento al azar de una población para una variable con distribución aproximadamente simétrica. Este tipo de problemas se presenta, por ejemplo, en la predicción de la esperanza de vida o el beneficio esperado en una inversión en bolsa. Se toma la media de la variable de la población como predicción, como valor esperado, por sus propiedades muestrales derivadas del teorema central de límite. Presentamos este algoritmo como un resultado de investigaciones anteriores sobre la media. Un ejemplo donde la media se presenta como valor esperado sería el siguiente.

La media en fluidez verbal de una clase de un colegio es de 400. Si extraemos una muestra aleatoria de 5 estudiantes y resulta que la puntuación de los 4 primeros es de 380, 420, 600, 400. ¿Cuál debería ser aproximadamente la puntuación esperada para el quinto estudiante? (Batanero, C. 2000, pág.8)

Situaciones-problema asociadas a la mediana

PME2. Encontrar un resumen estadístico de posición central para variables ordinales. Cuando se estudian variables cualitativas ordinales, las diferencias del valor numérico de la variable no corresponden a las diferencias de la magnitud medida y; por lo tanto, la mediana se presenta como un elemento más representativo y coherente a comparación de la media. Por ejemplo, al tener los calificativos *malo*, *regular*, *bueno*, *muy bueno y excelente*, hablar de la media de estos datos carece de sentido. Aún si cuantificáramos estos datos codificándolos con valores numéricos, la media como representante de los datos sería inapropiada e incoherente.

La solución consiste entonces en ordenar los datos y tomar como representante del conjunto de datos el valor de la variable que ocupa la posición central; es decir, el valor que deja tantos valores inferiores como superiores a él. En este contexto, podríamos preguntarnos si la moda puede ser también un representante del conjunto de datos, y la respuesta es que sí. Sin embargo, la moda solo toma en cuenta la frecuencia con que aparece cierto valor de la variable, pero no su ordenación. Por lo tanto, la mediana proporciona una información más completa que la moda, siempre que sea posible calcularla. Incluimos este campo como un resultado de investigaciones anteriores sobre la mediana. Presentamos un ejemplo encontrado en Cobo (2003). p. 47.

Un profesor califica a sus alumnos del siguiente modo: I=Insuficiente, A=Aprobado, N=Notable, S=Sobresaliente. En la siguiente tabla tenemos las notas que ha puesto a dos grupos de alumnos:

Grupo 1	I A A N N S S I I I A A A N S S I A A S S S S
Grupo 2	S S I I A N A N I I S N A S I N N

¿Qué grupo ha obtenido mejores notas?
¿Cuál sería el promedio más apropiado para representar estos datos?

(Godino, 1999)

Situaciones-problema asociadas a la moda

PMO2. Encontrar un valor representativo en datos cualitativos. En los datos cualitativos nominales, esta es la única medida de posición central que es posible calcular. Incluimos este campo, como un resultado de Cobo (2003), p.47.

Barrera	Número	Porcentaje
Falta de apoyo administrativo	10	43.5
tiempo	5	21.7
dinero	4	17.4
apatía	3	13.1
Falta de compensación	1	4.3
Total	23	100

Los siguientes datos muestran las dificultades encontradas para llevar a cabo un cierto tipo de programas de salud en Florida. ¿Cuál considerarías la "barrera típica" ¿Por qué?
(Johnson, 1992, pg. 45)

Es claro que en este ejemplo la única medida que toma sentido y es posible calcular es la moda, pues aunque se codifiquen numéricamente a las "barreras", las operaciones aritméticas para calcular la media carecerían de sentido e interpretación. Tampoco podríamos establecer un orden entre estas "barreras".

Los campos de problemas para nuestra investigación queda determinado por:

CAMPOS DE PROBLEMAS	
PM1	Media como mejor estimación
PM2	*Media como reparto equitativo*
PM3	Media en distribuciones simétricas
PM4	*Media como valor probable*
PME1	Mediana, si la media no es representativa
PME2	*Representante en datos cualitativos ordinales*
PMO1	Moda como valor más frecuente
PMO2	*Moda para datos cualitativos nominales*

DEFINICIONES

Media, mediana y moda*. Sus orígenes.*

Desde la antigüedad, el hombre realizaba actividades matemáticas como las de medir y contar. Los babilonios tuvieron gran interés en estudiar y explicar el universo y lo hacían a través de las observaciones, acompañadas de una disciplina teórica en la que las matemáticas tuvieron un papel de primer orden. Los astrónomos asentaron en tabillas de arcilla (figura 4) registros sobre los movimientos de los astros y planetas y, según Plackett (citado en Cobo, B. 2003), resolvieron problemas de estimación calculando la suma total de observaciones y dividiendo por el número total de datos. Este mismo autor (citado en Chan, C. 2009) señala que Ptolomeo usó la misma técnica, de manera general, para estimar la cantidad por la cual la duración de un año excede de 365 días. Esta práctica se ha conservado hasta la actualidad, dando origen a lo que hoy conocemos por media aritmética. Así, la media, - entre otras cosas - nos ayuda a resolver situaciones como "estimar una cantidad desconocida, en presencia de errores de medida".

FIGURA 4. Registro del movimiento de los astros realizadas por los astrónomos babilonios. ***Fuente: http://www.astroyciencia.com/2007/06/19/tablillas-babilonicas/***

Este concepto evolucionó, presentándose primero como útil implícito en la solución de problemas extra matemáticos y más tarde como un objeto de estudio en sí mismo. Fue precisamente, el estudio y caracterización de sus propiedades el que llevó progresivamente a la aplicación del concepto “media aritmética” en la solución de otras situaciones problemas. Sin embargo, cuando se añaden más condiciones a las situaciones problemas “surgen” otros conceptos. Así por ejemplo, si la distribución es muy asimétrica, la media no sería un valor representativo, por lo que conviene calcular otra medida, llamada *mediana*. Así también, si los datos no se pudieran ordenar (caso de las variables cualitativas nominales) habría que calcular un valor representativo de la distribución, llamada *moda.* Al respecto, Batanero, C. (2001) refiere: “*De los problemas primitivos planteados y de aquellos con otras condiciones (primero prácticos más tarde teóricos), han llevado a la definición de media, a la identificación de sus propiedades, más tarde a la definición de otras medidas de tendencia central como la mediana o moda, que son preferibles a la media en algunas situaciones concretas”.*

Estos valores reciben el nombre de promedios; más concretamente, medidas de tendencia central, pues hay una tendencia de los datos que se va observando a situarse alrededor de tales medidas. Existen otras medidas de tendencia central, pero en nuestra investigación consideramos a la media, mediana y moda como nuestros objetos de estudio.

Presentamos las definiciones para la media, la mediana y la moda que usaremos en nuestra investigación.

- ***Definición de media.*** Es considerada la principal medida de tendencia central. La media de una muestra se representa por $\overline{\boldsymbol{x}}$ y se calcula mediante la siguiente expresión:

$$\overline{\boldsymbol{x}} = \frac{\sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{f}_i}{\boldsymbol{N}}$$

Donde $\boldsymbol{N}$ es el total de valores observados, $\boldsymbol{x_i}$ cada uno de los valores observados y $\boldsymbol{f_i}$ la frecuencia con que aparece cada $\boldsymbol{x_i}$**.**

Si los datos que se presentan fueran agrupados en intervalos de igual longitud, para hallar la media, debemos considerar a las marcas de clase x'_i en lugar de las x_i (la marca de clase x'_i de un intervalo $[a_i;\ b_i]$ se obtiene al usar la siguiente relación: $x'_i = \frac{a_i+b_i}{2}$). En símbolos, tendríamos:

$$\overline{\boldsymbol{x}} = \frac{\sum_{i=1}^{n} \boldsymbol{x}'_i \boldsymbol{f}_i}{\boldsymbol{N}}$$

En este caso, debemos considerar que el valor de la media es solo una estimación (aproximado).

- ***Definición de mediana.*** La mediana es conocido también como un estadístico de orden, pues nos indica su posición central en el conjunto de datos ordenados***.*** En Batanero, C. Godino, J. D. (2001) se presenta la siguiente definición:

> *"Si suponemos ordenados de menor a mayor todos los valores de una variable estadística, se llama mediana al número (o valor de la variable[3]) tal que existen tantos valores de la variable superiores o iguales como inferiores o iguales a él"*

[3] Texto subrayado, agregado por el investigador.

Es preciso notar que, dado que la mediana hace referencia al valor que ocupa la posición central o del medio, para que haya tal centro o medio por lo menos tiene que haber un orden (por eso no será posible calcular la mediana en variables cualitativas nominales).

Esta medida es representada por $\mathrm{M_e}$ y para su cálculo será preciso distinguir cuando se tiene datos agrupados y sin agrupar, así como también, si el número de datos "n" es par o impar. En efecto, para datos sin agrupar tenemos:

- Si "n" es par, la mediana M_e es el valor que resulta de calcular el promedio de los datos que ocupan las posiciones $\frac{n}{2}$ y $\frac{n+1}{2}$.
- Si "n" fuera impar, la mediana M_e resulta ser el dato que ocupa la posición $\frac{n+1}{2}$.

Es preciso notar que cuando tenemos un número par de datos, puede ocurrir que la mediana no necesariamente sea uno de esos datos.

- ***Definición de Moda.*** Es el valor de la variable que tiene mayor frecuencia (se repite más veces). La moda puede no existir, y si existe, no ser única. Si la distribución presenta una sola moda se llamará *unimodal*, si existen dos *bimodal* y si ninguna variable se repite más veces que los otros; o si hay más de dos que se repiten más veces, diremos que no existe moda.

Una definición que se presenta en Cobo (2003) y que relaciona la moda con su representación gráfica es la siguiente:

DMO2. La moda es el valor que corresponde al máximo del diagrama de barras o histograma.

Consideramos las siguientes definiciones.

DEFINICIONES	
DM1	Media como suma ponderada
DM2	Media como promedio
DME1	Mediana como valor central
DME2	Mediana como valor que divide al conjunto en dos partes iguales
DME3	Mediana, cuya recta divide al histograma en dos áreas iguales
DMO1	Moda, valor más frecuente

PROPIEDADES

Además de las propiedades que se encontraron en el análisis de los textos seleccionados, existen otras más (Cobo, 2003). A continuación presentamos dichas propiedades restantes.

Algebraicas

A4. Al conmutar los valores de una distribución, no afecta el valor de la media y ni de la moda. No siempre ocurre lo mismo con la mediana. Recordemos que para el cálculo de la mediana se precisa de un orden, por eso, si cambiamos el orden de los datos, el valor de la mediana se puede ver fácilmente afectada.

Estadísticas

E7. Para datos agrupados en intervalos con alguno de ellos abierto, también es preferible la mediana a la media.

E8. Existen modas tanto para variables cuantitativas como cualitativas. En variables cualitativas nominales, la moda es el único estadístico calculable y que adquiere sentido. Si la variable es continua sólo se puede hallar un valor aproximado de la moda.

E9. En distribuciones no unimodales, la mediana es mejor representante del conjunto de datos que la media.

Todas las propiedades de la media, la mediana y la moda se presentan en la siguiente tabla.

PROPIEDADES
NUMÉRICAS
N1. Pertenecen al rango de la variable
N2. La moda coincide con un valor de los datos, la media y mediana no siempre.
N3. La media toma en cuenta todos los valores del conjunto, no así la mediana ni la moda
N4. El valor de la media cambia cuando se modifica cualquier dato del conjunto
ALGEBRAICA

A4. Son operaciones conmutativas, excepto la mediana.
A5. Conservan cambios de origen y escala
A6. La media de la suma de dos o más variables es la suma de las medias de dichas variables
A7. La moda puede no existir o, si existe, no ser única
A8. Relación empírica entre la media, la mediana y la moda
A9.Relación entre la media aritmética, geométrica y armónica
ESTADÍSTICA
E1. Representan a un colectivo
E2. La media coincide con el centro del conjunto de datos
E3. Las tres medidas coinciden en distribuciones simétricas
E4. La media es bastante sensible a la variación de los datos del conjunto
E5. La suma de las desviaciones de los datos respecto a la media es cero
E6. Respecto a la media, la suma de los cuadrados de las desviaciones es mínima
E7. Es preferible la mediana a la media en datos agrupados en intervalos
E8. Existe la moda para variables cualitativas y cuantitativas.
E9. La mediana es mejor representante que la media en distribuciones no unimodales

En general, solo cuatro de las propiedades se hacen explícitas en los textos analizados, las demás se tienen que deducir de algunos problemas propuestos o aparecen de manera implícita en algunos ejemplos.

ALGORITMOS

El siguiente algoritmo no se encontró en ninguno de los libros de texto analizados:

AM6. Inversión del algoritmo del cálculo de la media. En situaciones donde lo que se conoce es la media y más bien falta conocer uno de los valores de los datos resulta necesario invertir el algoritmo usado para calcular la media.

Los algoritmos que entran en juego en el estudio de la media, mediana y moda, son:

ALGORITMOS Y PROCEDIMIENTOS
AM1. Media para datos sin agrupar.
AM2. Media para datos presentados en tabla de frecuencias
AM3. Media para datos agrupados
AM4. Cálculo de la media a partir de gráficos
AM5. Cálculo de la media con calculadora u ordenador
AM6. Uso del algoritmo en sentido inverso
AM7. Dada la media, construir una distribución
AME1. Mediana para datos aislados (n impar)
AME2. Mediana para datos aislados (n par)
AME3. Mediana para datos presentados en tabla de frecuencias (n impar)
AME4. Mediana para datos presentados en tabla de frecuencias (n par)
AME5. Mediana para datos agrupados
AME6. Cálculo de la mediana a partir de un gráfico
AMO1. Moda para datos sin agrupar
AMO2. Moda para datos presentados en tabla de frecuencias
AMO3.Moda para datos agrupados
AMO4. Cálculo de la moda a partir de un gráfico.

ARGUMENTOS

Presentamos los argumentos que se encontraron en nuestro análisis:

ARGUMENTOS
ARG1. Justificación de ejemplos
ARG3. Razonamiento algebraico deductivo
ARG4. Razonamiento verbal deductivo

4.3. COMENTARIOS SOBRE EL SIGNIFICADO DE REFERENCIA EN RELACIÓN CON LOS FUNDAMENTOS TEÓRICOS DE LAS MEDIDAS DE POSICIÓN CENTRAL.

- Las revisiones de diversas investigaciones en didáctica sobre el tema de medidas de tendencia central, así como el análisis a los dos libros de texto, muestran un estudio solo a nivel de conjuntos de datos, es decir, el cálculo de las medidas de tendencia central se presenta como un fin y no como un medio que sirva para tomar decisiones acertadas en la solución de un problema, en el marco de la estadística inferencial, como resulta al tratarse según los fundamentos teóricos expuestos en el Apéndice 1.
- Cabe destacar que es muy importante que los textos y profesores que tratan estos temas, asuman una visión más amplia, considerando situaciones que busquen relacionar este concepto con la inferencia estadística, al menos, en su aspecto más básico.
- Teniendo en cuenta la perspectiva presente en los textos al presentar las medidas de tendencia central, como lo hemos manifestado en el primer comentario, nuestro significado de referencia se ha restringido a esta perspectiva.

CAPITULO V

SIGNIFICADO INSTITUCIONAL PRETENDIDO

El objetivo de este capítulo es describir el significado pretendido. Es decir, lo que el docente espera realizar y alcanzar con sus estudiantes en el desarrollo del tema "Medidas de tendencia central". Para esto realizamos lo siguiente:

- El análisis de la sección donde se estudia el tema de "Medidas de tendencia central" en el libro de texto "Matemáticas para no matemáticos".
- Una entrevista semiestructurada a los docentes.

Con esto se espera concluir con el primer objetivo, que es describir los significados institucionales correspondientes a las medidas de tendencia central.

5.1 ANÁLISIS DEL TEXTO MATEMÁTICAS PARA NO MATEMÁTICOS.

Como habíamos mencionado, en el texto "Matemáticas para no matemáticos" se encuentran todos los contenidos a desarrollarse en el curso "Matemáticas" (MAT128). Es decir, se presentan todos los temas que se espera, los alumnos comprendan. Cada profesor y alumno usan el texto para sus fines académicos correspondientes (enseñanza y aprendizaje, respectivamente).

El tema "Medidas de tendencia central" es tratado dentro del capítulo 3 (Análisis de datos) a lo largo de 16 páginas (p.121-p.136).

Los elementos que se encontraron las presentamos a continuación:

5.1.1 Descripción de los elementos de significado encontrados en el libro de texto.

LENGUAJE

Existe una gran variedad de términos involucrados en el estudio de las medidas de tendencia central. Las expresiones verbales media, media aritmética y promedio son usadas de manera indistinta, aunque sabemos que este último tiene un significado más amplio. Respecto a los símbolos y notaciones, éstos son usados en las formalizaciones de las definiciones de los objetos de estudio como también en los procedimientos. Así mismo, encontramos el uso de tablas de frecuencia, histogramas, gráfico de puntos y de sector circular.

A continuación, presentamos los elementos encontrados:

Términos y expresiones verbales

Específicas de la matemática	**Aparecen en matemáticas y el lenguaje cotidiano pero no siempre con el mismo significado**	**Significados iguales o próximos en ambos contextos**
Medidas de tendencia central, distribución de datos, distribución de frecuencias, frecuencia acumulada, frecuencia absoluta y relativa, variable cualitativa nominal y ordinal, variable cuantitativa discreta y continua, par, impar, variable estadística, marca de clase, intervalo, grafico de barras, bimodal.	*Media, mediana, moda, promedio, categoría, orden, promedio aritmético, posición, media aritmética, clase, datos agrupados, datos sin agrupar, desviación (respecto a la media).*	*Central, Muestra, representante, frecuencia, creciente, decreciente, fórmula, amplitud.*

TABLA 9. Términos y expresiones verbales encontrados análisis del texto Matemáticas para no matemáticos.

Notaciones y símbolos

	Relacionadas a la media	Relacionadas a la mediana	Relacionadas a la moda
Se representa por	$\overline{x}$	P_{50} , **Me**	**Mo**
Símbolos que usan para el estudio	$f_i,\ x_i,\ x_i'$, $\sum_{i=1}^{n} x_i,\ \sum_{i=1}^{k}(f_i)$ $\sum_{i=1}^{k}(x_i f_i), \sum_{i=1}^{k}(x_i' f_i)$	**Posición que ocupa el dato central:** **si n es impar:** $\frac{n+1}{2}$ **si n es par:** entre $\frac{n}{2}$ y $\frac{n}{2}+1$	
Fórmulas establecidas	$\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n},$ $\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n},$ $\overline{x} = \frac{\sum_{i=1}^{k}(x_i f_i)}{\sum_{i=1}^{k} f_i}$ $\overline{x} \approx \frac{\sum_{i=1}^{k}(x_i' f_i)}{\sum_{i=1}^{k}(f_i)}$	$M_e = \begin{cases} X_{\frac{n+1}{2}}, & n\ impar \\ X_{\frac{n}{2}} + X_{\frac{n+1}{2}}, & n\ par \end{cases}$	

TABLA 10. Notaciones y símbolos encontrados en el análisis del texto Matemáticas para no matemáticos.

Tablas estadísticas y gráficos

<table>
<tr><th>Tablas</th><th>Gráficos</th></tr>
<tr><td>❖ Listado de datos:
▪ Ordenados/no ordenados
❖ Tablas de frecuencia:
▪ Datos agrupados/no agrupados
▪ Intervalos iguales
▪ Intervalos semi-cerrados
▪ Un intervalo cerrado.</td><td>❖ Gráfico de barras
❖ Gráfico de puntos
❖ Gráfico de sectores circulares</td></tr>
</table>

TABLA 11. Tablas estadísticas encontradas en el análisis del texto Matemáticas para no matemáticos.

SITUACIONES-PROBLEMA

Situaciones-problema asociadas a la media

Situación 18

Jocelyn registra las notas de sus 7 primeras actividades en el curso de Matemáticas:

14; 10; 13; 18; 12; 15; 15

a) Calcule la media aritmética de las notas de las actividades de Jocelyn.

b) Si Jocelyn desea que el promedio aritmético de sus 8 primeras notas de actividades en el curso de Matemáticas sea 14, ¿qué nota deberá obtener en la próxima actividad?

Solución propuesta

a) La media aritmética de las 7 notas está dada por la suma de las 7 notas dividida entre el número de notas.
Es decir, la media aritmética es $\frac{14+10+13+18+12+15+15}{7}$, que equivale a $\frac{97}{7}$, aproximadamente 13,86.

b) Si se llama x a la nota que desea obtener Jocelyn en la octava actividad, entonces la suma de las 8 notas dividida entre 8 debe dar 14. Es decir: $\frac{97+x}{8}=14$. Ello será posible si $97+x=112$. Luego, necesitará obtener $112-97=15$ en su siguiente actividad.

Para abordar el estudio de la media, se presenta la situación 18 *(p.129)* en la cual se pide calcular la media de un conjunto de datos discretos que forman una distribución aproximadamente simétrica *(PM3)*. No se encontraron situaciones donde se pida estimar una medida a partir de mediciones realizadas en presencia de errores *(PM1),* obtener una cantidad equitativa al hacer un reparto para conseguir una distribución uniforme (*PM2)* y conocer el valor que se obtendrá con mayor probabilidad al tomar un elemento al azar de una población para una variable con distribución aproximadamente simétrica *(PM4).*

No considerar la situación *PM1* es obviar un aspecto importante en el estudio de la media, la razón inicial de su existencia, pues fue precisamente en la búsqueda de soluciones a este tipo de problemas donde se originó el concepto de media.

En el texto, después de dar el concepto de media, se hace el siguiente comentario sobre la media: *"Es la medida de tendencia central que posee las mejores propiedades para trabajar en estadística inferencial" (p.130)*. Sin embargo, todo queda allí. Plantear situaciones como *PM4* ayudaría a tener una primera visión acerca del papel que juega la media en la estadística inferencial.

Situaciones-problema asociadas a la mediana.

Situación 15
Camila es una delegada estudiantil que recibe el encargo de analizar conjuntos de notas de un determinado curso. Una de las tareas que debe hacer es encontrar, para cada conjunto de notas, el valor M_e que divide el conjunto en dos conjuntos con la misma cantidad de elementos.

a) Un conjunto A de notas es el siguiente: 08, 10, 14, 15, 15, 17, 19. ¿Qué valor de M_e obtendrá Camila?
b) Otro conjunto B de notas es el siguiente: 18, 14, 16, 15, 12, 13. ¿Puede obtener Camila el valor de M_e como lo hizo en el caso anterior? ¿Qué debería hacer?

La situación que se usa para introducir el concepto de mediana *(situación 15, p.122-123)* hace referencia a la necesidad de "encontrar un valor que divide al conjunto en dos conjuntos con la misma cantidad de elementos". Si bien en ella se muestra tal característica de la mediana, no se hace mención al porqué resulta útil el uso de tal medida y tampoco se muestra frente a qué situaciones se hace necesario tal uso.

No se ha encontrado situación alguna que enfatice e implique el uso necesario de la mediana. Por ejemplo, situaciones donde la distribución posea datos atípicos, que hace que la media no sea representativa *(PME1)*; o también, donde la variable de estudio sea cualitativa ordinal *(PME2)*.

Plantear situaciones de este tipo resulta importante, pues ponen en evidencia a la media como una medida no representativa de la distribución o carente de sentido (cuando las variables o elementos de la distribución presentan las características mencionadas). Ante ello, surge la necesidad de buscar medidas que puedan representar de manera adecuada a la distribución, como es el caso de la mediana.

Situaciones-problema asociadas a la moda.

Situación 14

William es un estudiante norteamericano que visita el Perú por intercambio estudiantil. Es aficionado al fútbol y quiere hacerse hincha del equipo peruano de mayor aceptación. Para ello William realizó una encuesta entre algunos amigos peruanos en la que les preguntó por el equipo de fútbol de su preferencia. Los resultados se muestran a continuación:

Equipo de fútbol favorito	Cantidad de hinchas
Sporting Cristal	5
Cienciano	4
Universitario de Deportes	8
Alianza Lima	6

a) ¿Cuál es la variable de estudio y de qué tipo es?
b) ¿Cuántas personas fueron encuestadas?
c) Si William elige el equipo del que se hará hincha basándose en esta información, ¿qué equipo tendría que elegir?
d) ¿Qué nombre podría llevar en Estadística la respuesta seleccionada en c)?

Se presentan situaciones *(como la situación 14, p.122)* que inducen a usar el valor más frecuente de una distribución de datos como representante de la misma *(PMO1)*, así como también, el uso necesario de la moda cuando se presentan variables cualitativas nominales *(PMO2)*.

Cuando se presenta la solución del problema, se considera como respuesta al ítem c la elección del equipo de Universitario de deportes por el hecho de ser el equipo que tiene mayor cantidad de hinchas (mayor frecuencia).

El problema termina con un comentario que expresa: *“Generalmente, al considerar un conjunto de datos, se busca un valor que lo represente en el sentido de que se parezca a muchos de los datos. Hay una gran variedad de formas de elegir este representante; en la situación 14 se eligió como representante al valor de mayor frecuencia: la moda.”*

Presentan a la moda como **una** solución al problema (*“en la situación 14* ***se eligió****…*”) y más bien no se alerta sobre el hecho que en tal situación la única medida que adquiere sentido y es posible calcular es la moda, por tratarse de una variable cualitativa nominal.

A continuación presentamos los campos de problemas encontrados:

Campos de problemas (Cecilia Gaita, et al., 2009)		
PM1	Media como mejor estimación	
PM2	*Media como reparto equitativo*	
PM3	Media en distribuciones simétricas	x
PM4	*Media como valor probable*	
PME1	Mediana, si la media no es representativa	
PME2	*Representante en datos cualitativos ordinales*	
PMO1	Moda como valor más frecuente	x
PMO2	*Moda para datos cualitativos nominales*	x

TABLA 12. Resumen de las situaciones-problema encontrados en el análisis del texto Matemáticas para no matemáticos.

CONCEPTOS

Definición de media

La definición que se presenta en el libro hace referencia a la existencia de la media para variables cuantitativas. Dicha definición se presenta verbalmente como también simbólicamente.

(p. 129-130)

Media aritmética de un conjunto de datos

La media aritmética de un conjunto de datos de una variable cuantitativa es el número que resulta de sumar todos los datos y dividir dicha suma entre el número de datos.

Así, si se tienen los datos: $x_1; x_2; x_3; \ldots ; x_n$ y se denota la media aritmética de estos datos por $\overline{x}$, entonces:

$$\overline{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

Esta fórmula se suele expresar mediante la notación sigma para sumas, así:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Definición de mediana

La definición que encontramos hace referencia a la mediana como el valor que divide a un conjunto de datos en dos subconjuntos con el mismo número de elementos *(DME2)*;

así como también, al valor que ocupa la posición central de la distribución de datos *(DME1)*. En todo ello, se menciona que la variable debe ser cuantitativa, así como también, que los datos deben ser ordenados para su cálculo. Además, se hace mención que es posible determinar la mediana para datos cualitativos ordinales pero no se aborda en el curso.

(p.124)

La mediana (M_e)

La mediana es un valor que deja por debajo de él la misma cantidad de datos que hay por encima de él.

La mediana de un conjunto de datos de una variable cuantitativa (ordenados previamente de manera creciente o decreciente) será:

- El dato que ocupa la posición central, si el número de datos es impar.
- El promedio de los dos datos centrales, si el número de datos es par.

Es importante notar que si el número de datos n es impar, entonces el dato central ocupa la posición $\frac{n+1}{2}$. En cambio, si n es par hay dos términos centrales que ocupan los lugares $\frac{n}{2}$ y $\frac{n}{2}+1$.

Para el caso de variables cualitativas ordinales también se podría definir la mediana luego de ordenar previamente los datos. Pero en lo que sigue solo se calculará la mediana de variables cuantitativas.

Definición de moda

La definición que se presenta es la más típica y hace referencia al dato que ocurre más frecuentemente, es decir, el que se repite más veces *(DMO1)*. Es importante notar el convenio que se establece, pues esta nos dice que la moda puede no existir, y si existe, pueden ser a lo más dos; en ese caso la distribución recibe el nombre de bimodal.

(p.122)

La moda (M_o)

La moda de un conjunto de datos es el puntaje o categoría que ocurre con la mayor frecuencia, o —equivalentemente— la que se repite más veces.
Se adoptarán las siguientes convenciones:

- Si hay dos valores o categorías que se repiten igual número de veces y dicha cantidad de veces es la más alta, diremos que hay dos modas y que la distribución de datos es bimodal.
- Si ninguna categoría o valor se repite más veces que los otros; o si hay más de dos que se repiten más veces, admitiremos que no hay moda.

Definiciones (Cecilia Gaita, et al., 2009)		
DM2	Media como promedio	x
DME1	Mediana como valor central	
DME2	Mediana como valor que divide al conjunto en dos partes iguales	x
DME3	Mediana, cuya recta divide al histograma en dos áreas iguales	
DMO1	Moda, valor más frecuente	x

TABLA 13. Definiciones encontradas en el análisis del texto Matemáticas para no matemáticos.

PROPIEDADES

Las propiedades que se encontraron ya sea de forma explícita o implícita son las siguientes:

Propiedades de la media, mediana y moda. (Cecilia Gaita, et al., 2009)	
NUMÉRICAS	
N1. Pertenecen al rango de la variable	x
N2. La moda coincide con un valor de los datos, la media y mediana no siempre.	x
N3. La media toma en cuenta todos los valores del conjunto, no así la mediana ni la moda	x
N4. El valor de la media cambia cuando se modifica cualquier dato del conjunto	x
ALGEBRAICAS	
A4. Son operaciones conmutativas, excepto la mediana.	x
A5. Conservan cambios de origen y escala	x
A6. La media de la suma de dos o más variables es la suma de las medias de dichas	

variables	
A7. La moda puede no existir o, si existe, no ser única (a lo más, dos modas)	x
A8. Relación empírica entre la media, la mediana y la moda	
A9.Relación entre la media aritmética, geométrica y armónica	
ESTADISTICAS	
E1. Representan a un colectivo	x
E2. La media coincide con el centro del conjunto de datos	
E3. Las tres medidas coinciden en distribuciones simétricas	x
E4. La media es bastante sensible a la variación de los datos del conjunto	x
E5. La suma de las desviaciones de los datos respecto a la media es cero	
E6. Respecto a la media, la suma de los cuadrados de las desviaciones es mínima	
E7. Es preferible la mediana a la media en datos agrupados en intervalos	
E8. Existe la moda para variables cualitativas y cuantitativas.	x
E9. La mediana es mejor representante que la media en distribuciones no unimodales	

TABLA 14. Propiedades encontradas en el análisis del texto Matemáticas para no matemáticos.

Comentario:

En el texto aparecen todas las propiedades numéricas. *N1* de forma implícita y las restantes *(N2, N3 y N4)* explícitamente. De las propiedades algebraicas existen tres a las que no se hace mención *(A6, A8 y A9)*, aunque tal aparición no resulta trascendental dado los objetivos del curso. Sólo las propiedades *A5 y A7* son presentadas de forma clara, pues *A4* queda implícito en la solución de diversas situaciones. Finalmente, en cuanto a las propiedades estadísticas, no aparecen cinco de ellas *(E2, E5, E6, E7 y E9)*, aunque el uso de dos de éstas *(E6 y E7)* no son necesarias para lo que se pretende en el curso. Por otro lado, no considerar la propiedad *E5* podría considerarse como una desventaja para el alumno en su comprensión sobre la media, pues esta propiedad nos da una idea importante acerca de la misma, al relacionarla con el punto de equilibrio de una distribución de datos; esto, pues el valor de la media es tal que la suma de sus diferencias con los datos que la exceden con las diferencias de los datos que están por debajo de esta, resulta cero.

ALGORITMOS Y PROCEDIMIENTOS

En general, al iniciar el estudio de las medidas de tendencia central, se hace una breve descripción de la importancia de su uso tanto en el quehacer cotidiano como profesional. Luego, para abordar los objetos de estudio (media, mediana y moda) se presenta, para cada uno, una situación problema contextualizada que involucra a estos conceptos. Se espera que los alumnos, evocando a sus conocimientos previos y compartiendo ideas con sus compañeros, intenten resolver dicha situación, pues se pretende que el alumno genere su propio conocimiento. Finalmente, el libro presenta una solución propuesta, que el profesor explica en clase, para luego formalizar el concepto involucrado. En este proceso, encontramos diversos algoritmos que son usados en la resolución de estos problemas, los cuales presentamos en la siguiente tabla.

Algoritmos y procedimientos (Cecilia Gaita, et al., 2009)	
AM1. Media para datos sin agrupar	x
AM2. Media para datos presentados en tablas de frecuencias	x
AM3. Media para datos agrupados	x
AM4. Cálculo de la media a partir de un gráfico	
AM5. Cálculo de la media con calculadora u ordenador	
AM6. Uso del algoritmo en sentido inverso	x
AM7. Dada la media, construir una distribución	x
AME1. Mediana para datos aislados (n impar)	x
AME2. Mediana para datos aislados (n par)	x
AME3. Mediana para datos presentados en tabla de frecuencias (n impar)	x
AME4. Mediana para datos presentados en tabla de frecuencias (n par)	x
AME5. Mediana para datos agrupados	
AME6. Cálculo de la mediana a partir de un gráfico	
AMO1. Moda para datos sin agrupar	x
AMO2. Moda para datos en tabla de frecuencias	x
AMO3. Moda para datos agrupados	
AMO4. Cálculo de la moda a partir de un gráfico.	x

TABLA 15. Algoritmos encontrados en el análisis del texto Matemáticas para no matemáticos

Comentario.

No todos los procedimientos que aparecen se explican en el texto, pues varios de los problemas para cuya solución se necesita del uso de estos algoritmos, son problemas propuestos. Así también, para el cálculo de la media, no se presentan ejemplos ni ejercicios que indiquen el uso de calculadora u otro ordenador (*AM5*). Otros procedimientos que tampoco aparecen son: *AM4, AME5, AME6 y AMO3*. De estos, no considerar problemas que involucren el uso adecuado de algoritmos útiles para una correcta interpretación de gráficos (*AM4 y AME6*) resulta un vacío que debería tratarse dada la importancia de la misma. Respecto a la omisión de los algoritmos *(AME6 y AMO3)* no involucra mayor problema, pues éstos no forman parte de los objetivos de estudio del curso debido a que demandan razonamientos más complejos.

ARGUMENTOS

En general, encontramos diversas situaciones problemas, así como ejemplos que ayudan a entender mejor las definiciones de media, mediana y moda como también los procedimientos necesarios para su cálculo. Varias de estas situaciones o ejemplos presentan razonamientos verbales que interpretan los resultados obtenidos en su solución, un ejemplo sería el siguiente (p. 122):

Así por ejemplo, se tienen los siguientes datos acerca de la escala de pago de 20 alumnos de Estudios Generales Letras:

3 ; 1 ; 2 ; 1 ; 5 ; 5 ; 4 ; 3 ; 4 ; 5

2 ; 4 ; 3 ; 3 ; 2 ; 5 ; 3 ; 2 ; 4 ; 5

Se observa que la escala 3 y las escala 5 ocurren cinco veces cada una y esa es la mayor cantidad de veces que se repite uno de los datos. Es decir, dicho conjunto de datos tiene dos modas: escala 3 y escala 5.

Por otro lado, los argumentos que involucran razonamientos algebraicos deductivos (demostraciones) son muy escasos ya que sólo aparece en una situación problema propuesto (situación que presentamos en *A5, cap.V).* En dicha situación se tendrá que justificar la respuesta y para ello es necesario realizar un razonamiento algebraico, donde el alumno tendrá que recurrir a la definición de media, como también recordar algunas propiedades de la suma.

La siguiente tabla resume los argumentos encontrados:

Argumentos (Cecilia Gaita, et al., 2009)	
ARG1. Justificación con ejemplos	x
ARG2. Uso de gráficos	
ARG3. Razonamiento algebraico deductivo	x
ARG4. Razonamiento verbal deductivo	x

TABLA 16. Argumentos encontrados en el análisis del texto Matemáticas para no matemáticos.

5.2 Descripción de las respuestas de los docentes a la entrevista.

Se entrevistó a tres profesores (de siete) con el fin de extraer información relevante acerca de las concepciones que tienen acerca del tema de medidas tendencia central (en particular media, mediana y moda), como por ejemplo, su importancia, así como otras consideraciones tales como: qué es lo que esperaban lograr con sus alumnos al final del estudio del tema o cuál es su percepción sobre el libro de texto “Matemáticas para no matemáticos”, así como también conocer lo errores frecuentes que han identificado en los alumnos a lo largo de su desempeño como docente. Se usó para ello una entrevista semi-estructurada (ANEXO 4) y algunas de las entrevistas se grabaron previo consentimiento del profesor.

A continuación presentamos un cuadro que resume las respuestas de los docentes:

Pregunta	Profesor A	Profesor B	Profesor C
Importancia	Uso cotidiano	Uso cotidiano/profesional	*U*so académico (trabajos de investigación de los estudiantes, uso en cursos posteriores)
Espera que los alumnos logren	*D*istinguir de forma adecuada los tipos de variables. *U*so adecuado de los conceptos. *R*esolver problemas contextualizados. *S*aber cuándo usar cierto promedio, reconociendo sus ventajas y desventajas.	*I*nterpretar cada una de las MTC. *M*anejo de algoritmos. *S*aber cuándo usar cierto promedio, reconociendo sus ventajas y desventajas.	*D*istinguir de forma adecuada los tipos de variables. *U*so adecuado de los conceptos. Interpretar cada una de las MTC. *U*so de recursos informáticos (Excel y calculadora).
Qué hacer para lograr ese objetivo	*Q*ue los alumnos trabajen con casos reales, concretos y cotidianos.	*D*esarrollar actividades de aprendizaje. *P*articipación activa de los alumnos (que construyan sus propios conocimientos). *U*so de problemas contextualizados. *R*etroalimentación (preguntas reflexivas)	*P*ráctica. Que resuelvan situaciones cotidianas y contextualizadas. *G*uía permanente por parte del profesor y asistentes. *Q*ue el alumno construya su propio conocimiento.
Contribución del texto en este objetivo	Sí, porque el libro está diseñado en el contexto del alumno. Allí se encuentran las situaciones que se usan en clase.	*S*í, porque los alumnos de letras tienen habilidades distintas a los de ingeniería. *R*equiere otro enfoque con problemas contextualizados, donde el alumno encuentre la aplicabilidad de estos conceptos. El texto posee estos requerimientos.	*S*í, porque presentan problemas contextualizados que motivan al alumno a buscar soluciones.
Dificultades más frecuentes en los alumnos	*N*o identifican la variable en estudio.	*L*os alumnos presentan distintos niveles de conocimiento. *A*ctitud de los alumnos hacia la matemática. *E*l alumno entendió y no reforzó después.	*C*álculo a partir de tablas.

TABLA 17. Resumen sobre las respuestas de los docentes a la entrevista.

5.3 CONCLUSIONES SOBRE EL SIGNIFICADO PRETENDIDO.

1. Respecto al texto guía, podemos mencionar que, en general, está diseñado en un contexto cercano al estudiante, lo que hace que su lectura y estudio sea más ágil y motivador. Sin embargo, Existen algunos vacíos en la introducción del concepto de media (no se consideran situaciones donde se pida estimar la medida de un objeto a partir de diversas mediciones realizadas en presencia de errores), el estudio de sus propiedades y características (no se enfatiza sobre la pertinencia y conveniencia del uso de la media, mediana y moda frente a ciertas situaciones), el planteamiento de algunas situaciones problemáticas, así como en el uso de argumentos.
2. Respecto a los profesores entrevistados, todos hacen mención a la importancia de las medidas de tendencia central ya sea en la vida cotidiana, profesional o académica. Esperan que sus alumnos aprendan a resolver problemas contextualizados, saber cuándo usar qué medida de tendencia central, interpretarlas correctamente, usar la calculadora, entre otros logros. Todos coinciden en que el libro es idóneo para desarrollar el curso y brindan para ello diversas razones. Es interesante hacer notar que varios de los logros que desean alcanzar con sus estudiantes no se encuentran en el texto, según lo expuesto en el ítem 1.

CAPITULO VI

SIGNIFICADO INSTITUCIONAL IMPLEMENTADO

En este capítulo se pretende lograr nuestro segundo objetivo, que es observar y analizar el proceso seguido en la enseñanza de las medidas de tendencia central en el curso matemáticas en EEGGLL. Para ello, por medio de las observaciones de las sesiones de clase, se describe la enseñanza tal como es llevado a cabo. Para tal fin se usó la técnica de la "observación participante" y toda la información se recogió en un diario de observación diseñado para ese propósito (anexo3)

En la sección 6.1, se presenta la información recolectada en cinco aulas (cada una con un profesor distinto y con el apoyo de dos asistentes de docencia), para luego realizar un análisis y presentar en 6.2 los elementos de significado puestos en juego al estudiar el tema.

El tema de medidas de tendencia central fue abordado en el orden que se presenta en el libro de texto "Matemáticas para no matemáticos" (moda, mediana y media), dentro de una sesión de clase y en un tiempo aproximado de una hora.

6.1 Descripción de la secuencia de clase relativo al tema de medidas de tendencia central.

TABLA 18. Descripción la apertura de la clase relativa a las medidas de tendencia central.

APERTURA			
ASPECTO A OBSERVAR	**SI**	**NO**	**DESCRIPCIÓN**
a. Explica la importancia del tema, contextualizándolo en situaciones próximas al estudiante.	1		Se preguntó cuál es la *moda* de las marcas de las prendas de vestir que usan los alumnos.
		4	
b. Recoge saberes previos de los alumnos	4		Se plantearon preguntas como: ¿Qué es la moda? ¿Cuál es la definición de mediana? ¿Conocen las M.T.C?
		1	
c. Presenta una situación	2		Presentaron las siguientes características:

problema relacionado al concepto matemático "Moda" para que los alumnos intenten resolver			• Ejercicios sencillos de cálculo directo • Situaciones planteadas en el texto *matemáticas para no matemáticos* (situación 14 *p.121*)
		3	Dos profesores definieron la *moda* de forma directa y desarrollaron ejemplos prácticos sin brindar un tiempo para que los alumnos intenten resolverlos. El tercer profesor hizo lo mismo, pero desarrolló la situación 14 *p.121*.
d. Presenta una situación problema relacionado al concepto matemático "Mediana" para que los alumnos intenten resolver	2		Presentaron las siguientes características: • Ejercicios sencillos de cálculo directo • Situaciones planteadas en el texto *matemáticas para no matemáticos* (situación 15 *p.122*)
		3	Dos profesores definieron la *mediana* de forma directa y desarrollaron ejemplos prácticos y la situación 16 *p.125* sin brindar un tiempo para que los alumnos intenten resolverlos. El tercer profesor resolvió un ejercicio sencillo y luego pasó a definir el concepto de mediana.
e. Presenta una situación problema relacionado al concepto matemático "Media" para que los alumnos intenten resolver	2		Presentaron las siguientes características: • Ejercicios sencillos de cálculo directo • Situaciones planteadas en el texto *matemáticas para no matemáticos* (situación 18 *p.129* y problema 1 *p.133*)
		3	Dos profesores definieron la *media* de forma directa y desarrollaron ejemplos prácticos sin brindar un tiempo para que los alumnos intenten resolverlos. El tercer profesor presentó la situación 18 *p.129,* pero la resolvió inmediatamente.

Observaciones:

- Solo un profesor contextualiza las M.T.C mediante una pregunta relativa a la moda. Sin embargo, tampoco explica la importancia, en general, del conocimiento y uso adecuado de estos conceptos.
- En general, los alumnos saben cuáles son las M.T.C. Los conocimientos sobre la media es a nivel de algoritmo, mientras que para la mediana es más intuitivo pues lo conocen como *"el dato que está en medio"* o *"el que divide al grupo en*

dos partes iguales”. Respecto a la moda la mayoría recuerda bien, al definirlo como *“el dato que más se repite”.*

- La mayoría de los profesores usan las situaciones problemas presentadas en el texto “Matemáticas para no matemáticos” en el desarrollo de su clase. Las situaciones escogidas aparecen como introductorios al tema de estudio o dentro de los problemas propuestos.
- Los ejercicios sencillos son usados principalmente para visualizar el algoritmo de cálculo más que para dar una interpretación del mismo.

TABLA 19. Descripción del desarrollo de la clase relativa a las medidas de tendencia central.

DESARROLLO DE LA CLASE			
ASPECTO A OBSERVAR	**SI**	**NO**	**DESCRIPCIÓN**
f. Brinda pautas (ideas) a los alumnos en la búsqueda de solución del problema planteado	2		• Piden a los alumnos que identifiquen el tipo de variable • Aclaran que para calcular la mediana los datos deben estar ordenados • Verbalmente se aclara la definición para la moda (en general, recuerdan bien). Así también, se les dice cómo se calcula la media y mediana. Por ejemplo: “*para la media, sumar todos los datos y dividir entre el número de datos”* • Los asistentes de docencia absuelven algunas dudas respecto al cálculo de estas medidas.
		3	
g. Da un tiempo necesario para que los alumnos traten de resolver el problema.	2		Para la moda, la situación se resuelve bastante rápido. Algo similar ocurre con la media. No ocurre lo mismo con la mediana, donde confunden la posición de la mediana con su valor.
		3	
h. Recoge las ideas de los alumnos y a partir de ellas arma la explicación.	2		Los alumnos recuerdan bien la definición de moda. Para la media y la mediana el profesor procede a explicar de forma detenida.

		3	Se dan por lo descrito en los puntos *c, d y e.*
i. Da una explicación general, de la solución del problema, identificando claramente los conceptos relacionados(moda, mediana y media)	5	0	Ya sea que el problema se planteó para que los alumnos intenten resolver o simplemente para iniciar inmediatamente una explicación, en ambos casos, se expone su solución para todos.
j. Se explica los conceptos de moda, mediana y media, relacionados en la solución del problema.	5	0	• Todas las definiciones que se dan sobre la media, la mediana y la moda están basadas en el texto (la presentan en el ecram). Todos aclaran el criterio que se establece respecto a la existencia de la moda. • Se explican los cálculos de la media, mediana y moda a partir de tablas con datos sin agrupar, así como de la media para datos agrupados.
k. Se considera los aspectos más importantes de la media, mediana y moda	2		Respecto a sus propiedades, dos profesores mencionan que la mediana puede o no coincidir con uno de los datos de la distribución de frecuencia.
		3	
l. Explican en qué casos es pertinente usar cada una de las medidas de posición central	2		• Respecto a la moda, explican que se usa tanto para variables cualitativas como cuantitativas. • Respecto a la mediana, explican que tiene sentido solo en variables cuantitativas como cualitativas ordinales. • Respecto a la media, solo uno aclara que éste tiene sentido para variables cuantitativas.
		3	
m. Brinda una espacio para las dudas que puedan surgir de parte de los alumnos, así como también para abordar ideas equivocadas (concepciones erróneas) o cuestionables.	1		Un alumno formuló las siguientes preguntas: *¿si el número de datos es par, la mediana tiene dos valores? ¿si es impar, es una sola?* (este, además, confundía el valor de la mediana con su posición). El profesor aclara que, en una

			distribución, el valor de la mediana siempre es único y le explica una vez más el procedimiento, haciendo notar que el valor de mediana "no es" la posición de la misma.
		4	

Observaciones:

- Cuando se brinda un espacio a los alumnos para que se enfrenten a una situación, éstos intercambian ideas, se ven más motivados a indagar, recurriendo al profesor y/o a los asistentes docencia.
- Respecto a los aspectos más importante de las M.T.C sólo dos profesores mencionan una característica de la mediana. No se menciona ni se hace una reflexión respecto a las demás características y propiedades de estas medidas, necesarias para su interpretación y uso posterior.
- El uso pertinente de las MTC no es abordado con la importancia necesaria que éste merece. Solo se menciona los tipos de variables en las cuales adquieren sentido, pero no se distingue, la conveniencia, entre el uso de uno y otro.
- No se dio uso de calculadoras ni de ningún otro software estadístico en el estudio de este tema.
- No se resuelven problemas de cálculo a partir de gráficos.
- El tiempo que se usó para el desarrollo del tema fue la principal razón por la cual no hubo un espacio para absolver las dudas y otras cuestiones.

TABLA 20. Descripción del cierre de la clase relativa a las medidas de tendencia central.

CIERRE			
ASPECTO A OBSERVAR	**SI**	**NO**	**DESCRIPCIÓN**
n. Se procede a evaluar	5	0	Se procede con una evaluación continua, que aborda aspectos básicos sobre los conceptos de la media, mediana y moda.

6.2 Elementos del significado implementado.

A continuación, presentamos los elementos puestos en juego en el desarrollo del tema de medidas de tendencia central.

TABLA 21. Resumen de los elementos de significado implementado.

ELEMENTOS DEL SIGNIFICADO IMPLEMENTADO							
	Situaciones/problemas	**14**	**15**	**16**	**18**	**1**	**IPA[4]**
LENGUAJE	Términos y expresiones verbales	x	x	x	X	x	x
	Notaciones y símbolos		x	x	X		x
	Tablas estadísticas y gráficos	x	x			x	x
SITUACIONES PROBLEMA	PM3 Media en distribuciones simétricas				X	x	
	PMO1 Moda como valor más frecuente	x					x
	PMO2 Moda para datos cualitativos nominales	x					x
CONCEPTOS	DM2 Media como promedio				X	x	x
	DME2 Mediana como valor que divide al conjunto en dos partes iguales		x				x
	DMO1 Moda, valor más frecuente	x					x
PROPIEDADES	**NUMÉRICAS**						
	N1. Pertenecen al rango de la variable		x	x	X	x	
	N2. La moda coincide con un valor de los datos, la media y mediana no siempre.		x	x	X		x
	N3. La media toma en cuenta todos los valores del conjunto, no así la mediana ni la moda						
	N4. El valor de la media cambia cuando se modifica cualquier dato del conjunto					x	
	ALGEBRAICAS						
	A4. Son operaciones conmutativas, excepto la mediana.				X	x	

[4] Información y práctica adicional desarrollada por el docente.

	A5. Conservan cambios de origen y escala						x
	A7. La moda puede no existir o, si existe, no ser única (a lo más, dos modas)						x
	ESTADISTICAS						
	E1. Representan a un colectivo				x	x	
	E3. Las tres medidas coinciden en distribuciones simétricas						
	E4. La media es bastante sensible a la variación de los datos del conjunto					x	
	E8. Existe la moda para variables cualitativas y cuantitativas.	x					x

ELEMENTOS DEL SIGNIFICADO IMPLEMENTADO							
	Situaciones/problemas	**14**	**15**	**16**	**18**	**1**	**IPA**
ALGORITMOS Y PROCEDIMIENTOS	AM1. Media para datos sin agrupar				x	x	
	AM2. Media para datos presentados en tablas de frecuencias						x
	AM3. Media para datos agrupados						x
	AM6. Uso del algoritmo en sentido inverso				x		
	AM7. Dada la media, construir una distribución						
	AME1. Mediana para datos aislados (n impar)		x	x			
	AME2. Mediana para datos aislados (n par)		x	x			
	AME3. Mediana para datos presentados en tabla de frecuencias (n impar)						X
	AME4. Mediana para datos presentados		x				X

	en tabla de frecuencias (n par)						
	AMO1. Moda para datos sin agrupar	x					
	AMO2. Moda para datos en tabla de frecuencias	x					
	AMO4. Cálculo de la moda a partir de un gráfico.						
ARGUMENTOS	ARG1. Justificación con ejemplos	x	x	x	x		
	ARG3. Razonamiento algebraico deductivo				x		
	ARG4. Razonamiento verbal deductivo	x	x	x			

CAPITULO VII

SIGNIFICADO PERSONAL

Nuestro tercer objetivo es describir el significado personal que los estudiantes asignan a las medidas de tendencia central luego de finalizado el proceso de estudio. Esto resulta relevante pues, según el EOS, para analizar la comprensión de los estudiantes, se requiere observar las prácticas personales debido a que la comprensión no se puede observar de manera directa. Luego, para completar el estudio de la comprensión, se estudiará la correspondencia entre los significados personales logrados y los significados institucionales implementados.

Para conocer los elementos del significado personal, usamos un cuestionario de evaluación (parte del cuestionario desarrollado por Cobo, 2003) adaptado a nuestro contexto y con ítems modificados en algunos casos.

7.1. DESCRIPCIÓN Y ANÁLISIS DE CONTENIDO DEL CUESTIONARIO

El cuestionario que se usó para evaluar a los estudiantes estuvo compuesto de siete ítems. Varias de las preguntas fueron adaptadas no solo en el contexto, sino también en la presentación, para lograr una lectura más fluida.

A continuación presentamos una propuesta de solución a las preguntas del cuestionario y mostramos, en diversos cuadros, a los elementos que emergen en la búsqueda de solución a las mismas. El rigor de estas soluciones corresponden a lo que se espera deberían responder alumnos que no siguen la especialidad de estadística, pero respetando los conceptos teóricos involucrados.

7.1.1. Soluciones expertas a los ítems del cuestionario y elementos de significado intervinientes.

Ítem 1.

Diez estudiantes han pesado un objeto en la clase de ciencias, usando el mismo instrumento. Los pesos registrados por cada estudiante (en gramos) son los siguientes:

6,1	6,3	6,0	6,1	6,2	6,5	6,1	6,2	6,1	6,0

Determine un número que represente mejor el peso real del objeto medido. Explique su respuesta.

Comentario

Una manera de encontrar una buena estimación del peso real sería obteniendo la media aritmética de los pesos registrados. Observando que todos los datos son próximos entre sí, también resulta pertinente calcular la moda y la mediana y tomarlas como representantes del conjunto de datos registrados.

Solución:

Solución1:

Calculamos la media:

$$\bar{x} = \frac{6{,}1 + 6{,}3 + 6{,}0 + 6{,}1 + 6{,}2 + 6{,}5 + 6{,}1 + 6{,}2 + 6{,}1 + 6{,}0}{10} = \frac{61{,}6}{10} = 6{,}16$$

Dado que los datos no están muy dispersos, la media representa bien el peso real del objeto.

Solución2:

Calculamos la mediana:

Procedemos a ordenar los datos:

6,0	6,0	6,1	6,1	6,1	6,1	6,2	6,2	6,3	6,5

Como el número de datos "n" es par $(n = 10)$, la mediana M_e será la media aritmética de los valores que ocupan las posiciones $\frac{n}{2}(posición\ 5)\ y\ \frac{n}{2} + 1(posición\ 6)$. Así, $M_e = \frac{6{,}1+6{,}1}{2} = 6{,}1$.

Este valor que se ubica en el centro de la distribución, también puede representar al peso real del objeto.

Solución3:

Calculamos la moda:

Observamos que 6,1 es el único dato que se repite un mayor número de veces (cuatro veces). Entonces, $M_o = 6{,}1$.

Este valor, que es el “peso” que más se repite, también puede representar al peso real del objeto.

Históricamente, en este tipo de problemas, con datos no muy dispersos (obtenidos al medir un objeto con un mismo instrumento), el representante que se adopta es el valor de la media; sin embargo, como hemos visto, los valores de la mediana y la moda también pueden representarlos adecuadamente y sus valores son todos próximos entre sí.

Esta situación involucra los siguientes elementos de significado:

Elementos de significado	***Especificaciones***
LENGUAJE	❖ *Términos y expresiones: verbal y numérica.* ❖ *Representaciones: notaciones de la media, mediana y moda* $(\bar{x}\,, M_e\ y\ \ M_o)$.
SITUACIÓN	❖ *Cálculo de un representante de un conjunto de datos no agrupados con variable cuantitativa.*
DEFINICIONES	❖ *Media.* ❖ *Mediana.* ❖ *Moda.*
PROPIEDADES	❖ *La media, mediana y moda son representantes de un conjunto de datos.* ❖ *Un representante de un conjunto de datos, puede ser uno de ellos (caso de la moda o de la mediana si hubiera un número impar de datos) o algún valor, que no necesariamente coincide con uno de los datos, obtenido por un promedio adecuado (la media aritmética o la mediana, cuando se tiene un número par de datos).* ❖ *Son operaciones conmutativas, excepto la mediana.*
ALGORITMOS Y PROCEDIMIENTOS	**Para el cálculo de la media:** ❖ *Sumar los datos.* ❖ *Dividir la suma de datos entre el número de datos.* **Para el cálculo de la mediana:** – *Ordenar los datos.* – *Observar si el número de datos es par o impar.* – *Obtener la mediana al calcular la semisuma de los datos que ocupan las posiciones 5 y 6.* **Para el cálculo de la moda:** ❖ *Observar e identificar el dato que se repite el mayor número de veces.*
ARGUMENTO	❖ *Como todos los datos son cercanos entre sí, tanto la media como la moda y la mediana son buenos representantes del conjunto de datos. En este caso particular los tres prácticamente coinciden.*

TABLA 22. Configuración de la solución experta al ítem1.

Ítem 2.

En cierta provincia de un país, que tiene cuatro distritos, se han hecho estudios sobre el número promedio de hijos por familia.

Distrito	Número de familias	Promedio de hijos por familia
A	35	4,2
B	20	2,6
C	45	5,4
D	30	3,0

a) ¿Cuál es el promedio de hijos por familia, en toda la provincia?
b) En el distrito B hay diez niños menores de 10 años. Los pesos, en kilos, de nueve de ellos están dados en la siguiente tabla:

14	25	17	16	26	17	18	19	24

- *¿Cuál es la mediana de estos datos?*
- *¿Cuál es la mediana si incluimos el peso del décimo niño, que es 41 kg?*
  En este caso, ¿sería la media aritmética un buen representante de los 10 datos? Explique su respuesta.

Comentario

La parte a) de este ítem permite evaluar si los estudiantes advierten que, para el cálculo de la media, se requiere obtener la suma de todos los datos mediante una suma de productos y también obtener el número total de datos efectuando una suma con la información dada en la tabla. En la parte b) lo que se pide es el cálculo de la mediana, tanto para un número impar de datos como para un número par. Finalmente, se pretende saber si los estudiantes advierten el efecto que causa la presencia de valores atípicos sobre la media y la mediana.

Solución:

a) Calculamos la media:

$$\bar{x} = \frac{35(4{,}2) + 20(2{,}6) + 45(5{,}4) + 30(3{,}0)}{35 + 20 + 45 + 30} = \frac{532}{130} \approx 4{,}09$$

El número promedio de hijos por familia en toda la provincia es aproximadamente 4,09.

b)

❖ Procedemos a ordenar los datos:

14	16	17	17	18	19	24	25	26

Como el número de datos "n" es impar $(n = 9)$, la mediana M_e es el valor que ocupa la posición $\frac{n+1}{2} = \frac{9+1}{2} = 5$. Como 18 es el dato que ocupa la posición cinco, tenemos que $M_e = 18$.

- ❖ Al incluir el peso del décimo niño, los datos quedan ordenados así:

14	16	17	17	18	19	24	25	26	41

Como el número de datos "n" es par $(n = 10)$, la mediana M_e será el promedio de los valores que ocupan las posiciones $\frac{n}{2}(posición\ 5)\ y\ \frac{n}{2} + 1\ (posición\ 6)$ Así, $M_e = \frac{18+19}{2} = 18{,}5$.

La media aritmética no sería un buen representante del conjunto de datos, pues se tiene la presencia de un valor "atípico" que hace que el valor de la media se modifique significativamente.

Al calcular la media obtenemos: $\bar{x} = \frac{14+16+17(2)+18+19+24+25+26+41}{10} = \frac{217}{10} = 21{,}7$ que es un valor bastante "distante" de la mediana y también de la media considerando los nueve datos, sin el valor atípico, $(\bar{x} = 19{,}56)$.

Los elementos que intervienen en su solución son los siguientes:

Elementos de significado	***Especificaciones***
LENGUAJE	***Ítem a)*** ❖ *Términos y expresiones: media, promedio, aproximado.* ❖ *Representaciones: notación de la media* $(\bar{x})$. ***Ítem b)*** ❖ *Términos y expresiones: mediana, ordenar, par, posición, promedio, atípico, buen representante.* ❖ *Representaciones: notaciones diversas para las posiciones, notación de la mediana* (M_e).
SITUACIÓN	***Ítem a)*** ❖ *Cálculo de la media, conociendo medias de subconjuntos disjuntos de la población.* ***Ítem b)*** ❖ *Cálculo de la media y la mediana.*

DEFINICIONES	***Ítem a)*** ❖ *Media.* ***Ítem b)*** ❖ *Mediana.*
PROPIEDADES	***Ítem a)*** ❖ *La media de un conjunto de datos no es la media de las medias de cada uno de los conjuntos de una partición del conjunto.* ***Ítem b)*** ❖ *El valor de la media cambia cuando se añade un dato al conjunto y todos los otros se mantienen igual, salvo que el dato añadido sea igual al de la media anterior.* ❖ *La media es bastante sensible a la variación de cualquiera de los datos.* ❖ *La media y la mediana no siempre coinciden con el valor de uno de los datos, incluso pertenecen a un conjunto numérico distinto.* ❖ *La mediana es una medida más resistente que la media.*
ALGORITMOS Y PROCEDIMIENTOS	***Ítem a)*** ❖ *Multiplicar el número de familias por su respectivo promedio de hijos por familia.* ❖ *Sumar dichos productos.* ❖ *Dividir la suma obtenida entre el número total de familias encuestadas.* ***Ítem b)*** ❖ *Mediana del conjunto inicial:* – *Ordenar los datos.* – *Observar que el número de datos es impar (n=9).* – *La mediana es el valor que ocupa la posición 5.* ❖ *Mediana del conjunto que incluye el peso del décimo niño:* – *Ordenar los datos.* – *Observar que el número de datos es par (n=10).* – *Obtener la mediana como la semisuma de los datos que ocupan las posiciones 5 y 6.* ❖ *Media del conjunto que incluye el peso del décimo niño:* – *Sumar todos los datos.* – *Dividir la suma entre el número total de datos.*
ARGUMENTO	❖ *Cuando entre los datos de una distribución hay alguno que es atípico (bastante alejado de los otros) la media de los datos no representa bien al conjunto por verse afectado de manera significativa por el dato atípico.*

TABLA 23. Configuración de la solución experta al ítem2.

Ítem 3.

*Un profesor califica el rendimiento de sus alumnos en el curso de Educación Física del siguiente modo: **I=Insuficiente, A=Aprobado, N=Notable, S=Sobresaliente**. En la siguiente tabla tenemos las calificaciones que ha puesto a dos grupos de alumnos:*

Grupo1	I	A	A	N	N	S	S	I	I	I	A	A	A	N	S	S	I	A	A	S	S	S	S

Grupo 2	S	S	I	I	A	N	A	N	I	I	S	N	A	S	I	N	N

a) ¿Cuál es la variable observada y de qué tipo es?
b) ¿Cuál de los grupos ha obtenido mejores calificaciones? Explique su respuesta
c) Si tuviera que emplear un calificativo para el grupo, ¿Cuál sería el más representativo (y apropiado) para el grupo 1? ¿y para el grupo 2? Explique su respuesta.

Comentario:

En esta situación se pide comparar dos grupos de datos ordinales. Las medidas que se aceptan como solución son la mediana y la moda, siendo la mediana la medida más representativa pues toma en cuenta el orden de los datos y no solamente su frecuencia (como la moda). Si se cuantificaran las calificaciones podría calcularse la media; sin embargo, no resulta una medida idónea en este tipo de situaciones.

Solución:

Solución 1.

a) Variable: Calificación del rendimiento de los alumnos en el curso de Educación física.

   Tipo: Cualitativa ordinal.

b) Para comparar los grupos, procedemos a calcular el valor de la mediana en cada grupo por tratarse de una variable cualitativa ordinal con un número impar de datos. Para esto, ordenamos los datos:

   Grupo 1: I I I I I A A A A A A ***A*** *N N N S S S S S S S S* $(n = 23)$

   Grupo 2: I I I I I A A A ***N*** *N N N N S S S S* $(n = 17)$

   Como en ambos grupos el número de datos "n" es impar, las medianas M_e serán:

Grupo 1: la calificación que ocupa la posición $\frac{n+1}{2} = \frac{23+1}{2} = 12$. Así, $\boldsymbol{M_{e_1} = A}$.

Grupo 2: la calificación que ocupa la posición $\frac{n+1}{2} = \frac{17+1}{2} = 9$. Así, $\boldsymbol{M_{e_2} = N}$.

Al comparar las medianas en ambos grupos, concluimos que el grupo 2 obtuvo mejores calificaciones (notable).

c) Por tratarse de una variable cualitativa ordinal, la mediana sería el mejor representante en ambos grupos. Así, un calificativo para el grupo 1 sería $\boldsymbol{M_e = Aprobado}$ y para el grupo 2 $\boldsymbol{M_e = Notable}$.

Solución 2.

a) (se repite lo hecho en la *solución 1*)

b) (Considerando la moda). Para comparar los grupos necesitamos observar cual es el calificativo que más se repite en cada grupo (moda). En efecto,

En el *Grupo 1: I I I I I A A A A A A A N N N S S S S S S S S,* tenemos:

I se repite 5 veces

A se repite 7 veces

N se repite 3 veces

S se repite 8 veces

Así, la moda es $\boldsymbol{M_{o_1} = S}$ $(\boldsymbol{se\ repite\ más\ veces})$.

En el Grupo 2: I I I I I A A A N N N N N S S S S, tenemos*:*

I se repite 5 veces

A se repite 3 veces

N se repite 5 veces

S se repite 4 veces

Por lo tanto, tenemos las siguientes dos modas:

$\boldsymbol{M_{o_2'} = I\ y\ M_{o_2''} = N}$ $(\boldsymbol{ambos\ se\ repiten\ más\ veces})$.

Al comparar las modas en ambos grupos, concluimos que el grupo 1 obtuvo mejores calificaciones.

c) Si se adopta la moda como representante en ambos grupos, un calificativo que representa al grupo1 sería $\boldsymbol{M_{o_1} = Sobresaliente}$ y para el grupo 2 $\boldsymbol{M_{o_2'} = Insuficiente\ o\ M_{o_2''} = Notable}$.

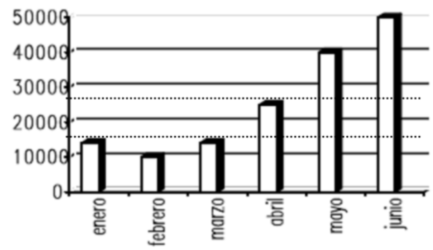
Los elementos que intervienen en su solución son los siguientes:

Elementos de significado	***Especificaciones***
LENGUAJE	❖ *Términos y expresiones: cualitativa, ordinal, comparar, ordenar, impar, mediana.* ❖ *Representaciones: notación de la mediana* (M_e), *etiquetas para los calificativos.*
SITUACIÓN	❖ *Moda y mediana en conjuntos de datos cualitativos ordinales.*
DEFINICIÓN	❖ *Variable cualitativa ordinal.* ❖ *Mediana.* ❖ *Moda.*
PROPIEDADES	❖ *Los promedios son representantes de todo el conjunto de datos.* ❖ *Existe la mediana y moda para variables cualitativas. En el caso de la mediana solo cuando la variable es ordinal y el número de datos es impar o cuando los datos centrales son iguales.* ❖ *La moda puede no ser única (a lo más dos)*
ALGORITMOS Y PROCEDIMIENTOS	***Ítem a)*** – *Reconocer la variable e identificar el tipo.* ***Ítem b)*** ❖ **Mediana:** – *Ordenar los datos en cada grupo.* – *Observar que, en ambos grupos, el número de datos es impar (n=23 y n=17).* – *La mediana para el grupo 1 es el calificativo que ocupa la posición 12* $(M_e = A)$ *y para el grupo 2 es el calificativo que ocupa la posición 9* $(M_e = N)$. ❖ **Moda:** – *En cada grupo, contar el número de veces que se repite cada uno de los calificativos.* – *En cada grupo, reconocer el calificativo que se repite más veces.* – *La moda para el grupo 1 es* $\boldsymbol{M_{o_1} = Sobresaliente}$ *y para el grupo 2* $\boldsymbol{M_{o'_2} = Insuficiente\ o\ M_{o''_2} = Notable}$. ❖ Comparar las medidas obtenidas y responder.
ARGUMENTO	❖ *(Si optó por la mediana) Por tratarse de una variable cualitativa ordinal, la mediana es un buen representante en ambos grupos. Así, un calificativo que representa al grupo 1 es A=Aprobado y para el grupo 2, N=Notable.* ❖ *(Si optó por la moda) Si se considera el calificativo que se repite más veces en cada grupo, el calificativo que representa al grupo 1 sería A=Aprobado y para el grupo 2, I=Insuficiente o N=Notable).*

TABLA 24. Configuración de la solución experta al ítem3.

Ítem4.

Observe el siguiente diagrama de barras que muestra el número de sándwiches que vendió la empresa Ambrossi durante el primer semestre del presente año:



a) *¿Cuál es la variable observada y de qué tipo es?*

b) *Dé un valor aproximado del promedio de sándwiches que la empresa vendió al mes.*

c) *Dé un valor aproximado de la moda del número de sándwiches que la empresa vendió al mes.*

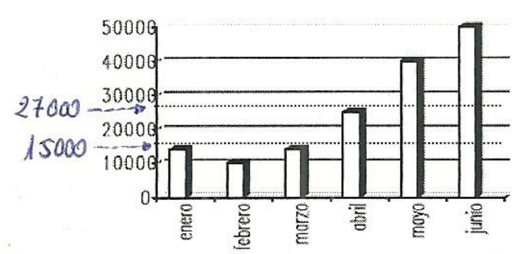
d) Dé un valor aproximado de la *mediana* del número de sándwiches que la empresa vendió al mes.

Comentario:

El formato de esta situación fue tomado de Cobo 2003. Llevada a otro contexto y agregando los ítems a) y d) pretendemos conocer si los estudiantes logran calcular, de manera correcta, los valores de la media, mediana y moda, a partir de un gráfico estadístico. Los resultados pueden diferir según el valor que se asignen a las barras (número de sándwiches vendidos por mes).

Solución:

a) Variable: Número de sándwiches
   Tipo: Cuantitativa discreta
b) Observamos que no queda clara la cantidad de sándwiches vendidos en los meses de enero, marzo y abril. Asumimos que la cantidad de enero y marzo es la misma (15000) y asignamos 27 000 para abril.



Entonces, $\bar{x} = \frac{15000+10000+15000+27000+40000+50000}{6} = \frac{157000}{6} \approx 26166.66$ (promedio de sándwiches que la empresa vendió al mes).

c) Observamos que, sólo en los meses de Enero y Marzo se vendieron la misma cantidad de sándwiches (15000). Así, la moda es $M_o = 15000$ sándwiches.

d) Procedemos a ordenar los datos

10000, 15000, 15000, 27000, 40000, 50000

Como el número de datos "n" es par ($n = 6$), la mediana M_e será el promedio de los valores que ocupan la posiciones $\frac{n}{2}(posición\ 3)\ y\ \frac{n}{2}+1(posición\ 4)$. Así, $M_e = \frac{15000+27000}{2} = 21000$.

TABLA 25. Configuración de la solución experta al ítem 4.

Elementos de significado	***Especificaciones***
LENGUAJE	❖ *Términos y expresiones: cuantitativa, discreta, comparar, par, posición, media, mediana y moda.* ❖ *Representaciones: notaciones de la media, mediana y moda* $(\bar{x}, M_e\ y\ M_o)$. ❖ *Gráficas***:** *diagrama de barras.*
SITUACIÓN	❖ *Cálculo de las medidas de tendencia central a partir de un gráfico estadístico.*
DEFINICIONES	❖ *Media.* ❖ *Mediana.* ❖ *Moda.*
PROPIEDADES	❖ *Un representante de un conjunto de datos, puede ser uno de ellos (caso de la moda o de la mediana si hubiera un número impar de datos) o algún valor, que no necesariamente coincide con uno de los datos, obtenido por un promedio adecuado (la media aritmética o la mediana, cuando se tiene un número par de datos.)*
ALGORITMOS Y PROCEDIMIENTOS	❖ *Reconocer la variable en estudio.* ❖ *Asignar valores al número de sándwiches que se vendió en los meses de enero, marzo y abril e identificar los valores para los demás meses.* ❖ *Cálculo de la media a partir de un gráfico:* – *Sumar los seis datos (número de sándwiches vendidos en cada mes).* – *Dividir la suma entre el número de datos (seis).* ❖ *Cálculo de la moda a partir de un gráfico: observar cual es el número de sándwiches vendidos al mes que más se repite en el gráfico.* ❖ *Cálculo de la mediana a partir de un grafico:* – *Reconocer los datos.*

	– *Ordenar los datos* – *Observar que el número de datos es par (n=6).* – *Obtener la mediana al calcular la semisuma de los datos que ocupan las posiciones 3 y 4.*
ARGUMENTO	• *De tipo verbal.*

Ítem5.

a) *Un informe del INEI en el 2007, dice que "en la provincia de Lima, el número medio (promedio) de hijos por familia es 1,3". Explique qué significa para usted esta expresión.*

b) *Carlos tomó información sobre el número de hijos en 10 familias de San Borja. Registró la información, pero por un error de digitación se le borraron algunos datos.*

Familia A	3
Familia B	2
Familia C	
Familia D	
Familia E	
Familia F	
Familia G	
Familia H	
Familia I	
Familia J	

- *¿Qué datos podrían haber estado en su cuadro, si se sabe que el promedio de hijos por familia es 1,2?*
- *¿Existe una única posibilidad de completar el cuadro? Justifique su respuesta.*

Comentario:

Este ítem ha sido tomado de Mayén, S. (2009), adaptado a nuestro contexto y con una presentación más legible.

En la parte a) de este ítem se espera que el alumno logre dar una interpretación adecuada, en sus propias palabras, sobre la media. En la parte b) el alumno deberá construir una distribución a partir de la media conocida. Finalmente, tendrá que responder, justificando, que existe más de una posibilidad de completar el cuadro.

Solución:

a) Podrían ser algunas respuestas:

- *"Al sumar el número de hijos en cada familia y dividir esta suma entre el total de familias, resulta 1,2".*
- *"Las familias tienen alrededor de 1 hijo, pero algunas familias tienen poco más y otras ninguno".*

b) Dado que conocemos el valor de la media, planteamos la siguiente ecuación:

$$\frac{\left(\sum_{i=1}^{8} x_i\right) + (3+2)}{10} = 1{,}2$$

$$\sum_{i=1}^{8} x_i + 5 = 12$$

$$\sum_{i=1}^{8} x_i = 12 - 5$$

$$\sum_{i=1}^{8} x_i = 7$$

Luego, una forma de completar el cuadro sería:

Familia A	3
Familia B	2
Familia C	**1**
Familia D	**1**
Familia E	**1**
Familia F	**1**
Familia G	**1**
Familia H	**1**
Familia I	**1**
Familia J	**0**

El cuadro se puede completar de diversas formas, lo que importa es que la suma de los ocho datos que se borraron resulte 7, para que así la media sea 1,2.

Elementos de significado	Especificaciones
LENGUAJE	❖ *Términos y expresiones: media o promedio, alrededor, igualdad.* ❖ *Representaciones: notación para la suma* $\sum_{i=1}^{8} x_i$
SITUACIÓN	❖ *Conocida la media, dar una interpretación en un contexto determinado.* ❖ *Conocida la media de un cierto conjunto de datos numéricos, brindar los posibles valores de los elementos del conjunto.*
DEFINICIÓN	❖ *Media.*
PROPIEDADES	❖ *La media es un valor que pertenece al rango de la variable.* ❖ *La media no tiene por qué ser uno de los valores de los datos.* ❖ *La media es un representante del conjunto de datos.* ❖ *La media toma en cuenta todos los valores del conjunto, incluso el cero.* ❖ *Son operaciones conmutativas (los datos se pueden conmutar)*
ALGORITMOS Y PROCEDIMIENTOS	*tem a)* ❖ *Interpretar de manera correcta el concepto de media en el contexto del problema.* *Ítem b)* ❖ *Para una media dada, construir una distribución. Para esto será necesario conocer el algoritmo de cálculo para la media y aplicarlo en sentido inverso.* ❖ *Es posible encontrar los datos por ensayo y error.*
ARGUMENTO	❖ *Es de tipo verbal y algebraico, tal como se presenta en la solución.*

TABLA 26. Configuración de la solución experta al ítem 5.

Ítem6.

Cierto grupo de niños, el 20 de julio de 2009, tenía como edad promedio 13,8 años. ¿Es verdad que la edad promedio del mismo grupo de niños, el 20 de julio de 2013 será 18,2 años? Justifique su respuesta.

Comentario:

En esta situación se pretende conocer si el alumno recuerda y usa de manera adecuada la siguiente propiedad: "si se añade un valor a cada uno de los datos de la distribución, la media aumenta también en dicho valor".

Solución*:*

Es falsa. El 20 de julio del 2013, cada niño tendrá 4 años más. Al sumar todos los datos aumentados (en 4 años cada uno) y dividirlo entre el número de niños (que es el mismo) el cociente será 4 unidades mayor que el cociente correspondiente al promedio calculado para el 20 de julio del 2009.

Formalmente:

Sea "n" el número de niños cuya edad promedio es $13{,}8$ años.

$$\bar{x}_{2009} = \frac{x_1 + x_2 + \cdots + x_n}{n} = 13{,}8$$

Dentro de 4 años la edad promedio será:

$$\bar{x}_{2013} = \frac{(x_1 + 4) + (x_2 + 4) + \cdots (x_n + 4)}{n}$$

Al asociar,

$$\bar{x}_{2013} = \frac{(x_1 + x_2 + \cdots + x_n) + 4(n)}{n}$$

$$\bar{x}_{2013} = \frac{x_1 + x_2 + \cdots + x_n}{n} + \frac{4n}{n}$$

$$\bar{x}_{2013} = 13{,}8 + 4 = 17{,}8\ años.$$

Elementos de significado	*Especificaciones*
LENGUAJE	❖ *Términos y expresiones: promedio, asociar, expresión algebraica* $\frac{x_1+x_2+\cdots+x_n}{n}$. ❖ *Representaciones: notación para las dos medias* $(\bar{x}_{2009}\ y\ \bar{x}_{2013})$, *y para el número de niños (n).*
SITUACIÓN	❖ *Dada la media de cierto conjunto de datos, hallar la nueva media del conjunto cuyos elementos corresponden a los del conjunto anterior pero aumentados, cada uno, en una misma cantidad.*
DEFINICIÓN	❖ *Media.*
PROPIEDADES	❖ *La media toma en cuenta a todos los valores del conjunto de datos.* ❖ *Si sumamos una misma cantidad a cada elemento de un conjunto de datos, la media también aumentará en esa misma cantidad.* ❖ *Propiedades de la adición y multiplicación de números reales.* $\forall\, a,b,n\ \in IR, n\neq 0: \frac{a+b}{n}=\frac{a}{n}+\frac{b}{n}$ $\forall\, \alpha,n\ \in IR, n\neq 0: \frac{\alpha n}{n}=\alpha$
ALGORITMOS Y PROCEDIMIENTOS	❖ *Observar que el tiempo que transcurrirá es cuatro años.* ❖ *Se suma cuatro al promedio inicial, de donde resulta el nuevo promedio* $\bar{x}_{2013}=17{,}8\ \text{años}$ *diferente al que se da en el problema.* ❖ *Se concluye que la afirmación es falsa.*
ARGUMENTO	❖ *La afirmación es falsa, pues al calcular el nuevo promedio de edades resulta un valor distinto a 18,2.*

TABLA 27. Configuración de la solución experta al ítem6.

Ítem7.

Se realizó una encuesta a los asistentes al Coloquio de Ciencias Humanas para conocer cuál fue el primer medio por el cual se informaron de esta actividad. Los resultados fueron los siguientes:

Medio de información	Número de personas
Periódico	33
E-mail	55
Página Web	30
TV	32
Radio	40

a) ¿Cuál es la variable observada y de qué tipo es?

b) ¿Cuál de las tres medidas de posición central (media, mediana y moda) tiene sentido hallar en este contexto? ¿Por qué?

Comentario:

Mediante una tabla de frecuencia absoluta, se presenta un conjunto de datos cualitativos nominales. Debido al tipo de variable en estudio, la única medida que es posible calcular es la moda.

Solución:

a) Variable: medio de información

   Tipo: cualitativa nominal

b) La única medida que adquiere sentido en este contexto es la moda, pues no podríamos calcular la media ni la mediana de los datos en cuestión (la variable es cualitativa nominal).

   La moda es el dato que más veces se repite, en nuestro caso $M_o = e\_mail$ (se repite 55 veces).

Elementos de significado	*Especificaciones*
LENGUAJE	❖ *Términos y expresiones: cualitativa, nominal, moda.* ❖ *Representaciones: notación para la moda* (M_o). ❖ *Gráfica: Tabla con frecuencia absoluta.*
SITUACIÓN	❖ *Moda para datos cualitativos nominales.*
DEFINICIÓN	❖ *Moda.*
PROPIEDADES	❖ *La moda pertenece coincide con uno de los datos del conjunto.* ❖ *Existe la moda para variables cualitativas y cuantitativas.*
ALGORITMOS Y PROCEDIMIENTOS	❖ *Identificar el tipo de variable en estudio.* ❖ *Reconocer que la única medida que tiene sentido hallar es la moda.* ❖ *Observar el número de veces que se repiten cada dato (frecuencia)* ❖ *La moda resulta* $M_o = e\_mail$ *(el único que se repite mayor número de veces).*
ARGUMENTO	❖ *La única medida que tiene sentido hallar es la moda pues se trata de una variable cualitativa nominal.*

TABLA 28. Configuración de la solución experta al ítem7.

7.1.2 Resumen de los elementos de significado que intervienen en la solución de los ítems del cuestionario.

El siguiente cuadro resume los diversos campos de problemas, lenguaje, conceptos, propiedades, algoritmos y argumentos que intervienen en la solución de los problemas planteados. Notamos las soluciones a los 7 ítems propuestos involucran casi todos los elementos del significado pretendido. Las limitaciones del tiempo para la evaluación hicieron que no se puedan considerar preguntas que involucren *AM3, AME3 y AME4.*

TABLA 29. Resumen de los elementos de significado presentes en el cuestionario.

.	**ÍTEMS**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
LENGUAJE	Verbal	x	x	x	x	x	x	X
	Numérico	x	x		x	x	x	
	Gráfico				x			
	Simbólico	x	x	x			x	X
SITUACIÓN PROBLEMA	PM1 Media como mejor estimación	x						
	PM3 Media en distribuciones simétricas	x				x		
	PME1 Mediana, si la media no es representativa		x	x				
	PME2 Representante en datos cualitativos ordinales			x				
	PMO1 Moda como valor más frecuente	x		x	x			X
	PMO2 Moda para datos nominales							X
CONCEPTOS	DM2 Media como promedio	x	x		x	x	x	
	DME1 Mediana como valor central	x	x	x		x		
	DME2 Mediana como valor que divide al conjunto en dos partes iguales		x	x	x			
	DMO1 Moda, valor más frecuente	x		x	x			X
PROPIEDADES	NUMÉRICAS							
	N1. Pertenecen al rango de la variable	x	x	x	x	x		X
	N2. La moda coincide con un valor de los datos, la media y mediana no siempre.	x	x	x	x	x		X
	N3. La media toma en cuenta todos los valores del conjunto, no así la mediana ni la moda.		x			x	x	
	N4. El valor de la media cambia cuando se modifica cualquier dato del conjunto.		x					
	ALGEBRAICA							
	A4. Son operaciones conmutativas, excepto la mediana.	x			x	x		X

	A5. Conservan cambios de origen y escala						x	
	A6. La media de la suma de dos o más variables es la suma de las medias de dichas variables							
	A7. La moda puede no existir o, si existe, no ser única			x				
	ESTADÍSTICA							
	E1. Representan a un conjunto e datos	x	x	x	x	x	x	x
	E3. Las tres medidas coinciden en distribuciones simétricas	x						
	E4. La media es bastante sensible a la variación de los datos del conjunto		x					
	E5. La suma de las desviaciones de los datos respecto a la media es cero							
	E8. Existe la moda para variables cualitativas y cuantitativas.			x				x

	ÍTEMS	**1**	**2**	**3**	**4**	**5**	**6**	**7**
ALGORITMOS	AM1. Media para datos sin agrupar	x	x			x	x	
	AM2. Media para datos presentados en tabla de frecuencias		x					
	AM3. Media para datos agrupados							
	AM4. Cálculo de la media a partir de gráficos				x			
	AM6. Uso del algoritmo en sentido inverso					x		
	AM7. Dada la media, construir una distribución					x		
	AME1. Mediana para datos aislados (n impar)		x	x				
	AME2. Mediana para datos aislados (n par)	x	x		x			
	AME3. Mediana para datos presentados en tabla de frecuencias (n impar)							
	AME4. Mediana para datos presentados en tabla de frecuencias (n par)							
	AME6. Cálculo de la mediana a partir de un gráfico				x			
	AMO1. Moda para datos sin agrupar	x		x	x			x
	AMO2. Moda para datos presentados en tabla de							x

	frecuencias							
	AMO4. Cálculo de la moda a partir de un gráfico.				x			X
ARGUMENTO	ARG1. Justificación de ejemplos							
	ARG3. Razonamiento algebraico deductivo				x	X	x	
	ARG4. Razonamiento verbal deductivo	x	x			X	x	X

7.2. ANÁLISIS DE LAS RESPUESTAS A LOS ÍTEMS PROPUESTOS.

En esta sección, presentamos primero los resultados cuantitativos respecto a la evaluación realizada a los estudiantes sobre las medidas de tendencia central. Luego, complementando el estudio cuantitativo, realizamos un análisis cualitativo a través de las configuraciones cognitivas a las respuestas de los estudiantes a algunos ítems seleccionados por presentar un mayor porcentaje de respuestas incorrectas.

7.2.1 ESTUDIO CUANTITATIVO DE LAS RESPUESTAS. Presentamos un resumen cuantitativo de las respuestas de los alumnos a los ítems planteados en la evaluación sobre las medidas de tendencia central. Clasificamos las respuestas como correctas e incorrectas, considerando las respuestas en blanco como incorrectas. Finalmente, resumimos esta información en tablas de frecuencia de tipologías.

Ítem 1. (PGTA 2 EL EXAMEN)

Diez estudiantes han pesado un objeto en la clase de ciencias, usando el mismo instrumento. Los pesos registrados por cada estudiante (en gramos) son los siguientes:

6,1	6,3	6,0	6,1	6,2	6,5	6,1	6,2	6,1	6,0

Determine un número que represente mejor el peso real del objeto medido. Explique su respuesta.

TABLA 30. Frecuencia de tipologías de significados personales declarados en el ítem 1.

n=49						
Tipologías	**Correcta (81,6%)**		**Incorrecta (18,4%)**			**Total**
	Argumenta	**No argumenta**	**Argumenta**	**No argumenta**	**No contesta**	
Media	29	0	6	0	0	35
Moda	10	0	0	0	0	10
Mediana	0	0	0	0	0	0
Media y Mo	1	0	2	0	0	3
No responde	0	0	0	0	1	1
Total	40	0	8	0	1	**49**

Los resultados obtenidos no difieren mucho de los obtenidos por Mayén, quién en su muestra de alumnos de bachillerato (que son alumnos con edades similares a la de nuestra muestra) obtuvo 85% de respuestas correctas frente a 15% de incorrectas.

Ítem 2. (PGTA 3 DEL EXAMEN)

En cierta provincia de un país, que tiene cuatro distritos, se han hecho estudios sobre el número promedio de hijos por familia.

Distrito	Número de familias	Promedio de hijos por familia
A	35	4,2
B	20	2,6
C	45	5,4
D	30	3,0

a) ¿Cuál es el promedio de hijos por familia, en toda la provincia?

b) En el distrito B hay diez niños menores de 10 años. Los pesos, en kilos, de nueve de ellos están dados en la siguiente tabla:

14	*25*	*17*	*16*	*26*	*17*	*18*	*19*	*24*

- *¿Cuál es la mediana de estos datos?*
- *¿Cuál es la mediana si incluimos el peso del décimo niño, que es 41 kg?*
  En este caso, ¿sería la media aritmética un buen representante de los 10 datos? Explique su respuesta.

Observación: Para el análisis, hemos dividido el ítem b) en tres partes: la primera viñeta como 2b.1) y la segunda como 2b.2) y 2b.3), esto, para un mejor orden.

TABLA 31. Frecuencia de tipologías de significados personales declarados en el ítem 2a.

n=49						
Tipologías	**Correcta (59,2%)**		**Incorrecta (40,8%)**			**Total**
	Argumenta	**No argumenta**	**Argumenta**	**No argumenta**	**No contesta**	
Media	29	0	18	0	0	47
Otras respuestas	0	0	0	2	0	2
Total	29	0	18	2	0	**49**

TABLA 32. Frecuencia de tipologías de significados personales declarados en el ítem 2b.1.

n=49				
Tipologías	**Correcta (67,3%)**	**Incorrecta (32,7%)**	**No contesta**	**Total**
Mediana	33	11	0	45
Media	0	1	0	1
Otros cálculos	0	2	2	4
Total	33	14	2	**49**

TABLA 33. Frecuencia de tipologías de significados personales declarados en el ítem 2b.2.

n=49				
Tipologías	**Correcta (59,2%)**	**Incorrecta (40,8%)**	**No contesta**	**Total**
Mediana	29	15	0	44
Media	0	1	0	1
Otras justificaciones	0	2	2	4
Total	29	18	2	**49**

TABLA 34. Frecuencia de tipologías de significados personales declarados en el ítem 2b.3.

n=49						
Tipologías	**Correcta (26,5%)**		**Incorrecta (73,5%)**			**Total**
	Argumenta	**No argumenta**	**Argumenta**	**No argumenta**	**No contesta**	
Media	13	0	6	0	0	19
	0	0	19	0	11	30
Total	13	0	25	0	11	**49**

En los ítems 2b.1, 2b.2 los resultados son similares a los obtenidos por Mayén respecto a su muestra con alumnos de bachillerato. No ocurre lo mismo respecto al ítem 2b.3, donde el porcentaje de respuestas correctas de nuestra muestra está por debajo del de Mayén (36%).

Ítem 3. (PGTA 4 DEL EXAMEN)

*Un profesor califica el rendimiento de sus alumnos en el curso de Educación Física del siguiente modo: **I=Insuficiente, A=Aprobado, N=Notable, S=Sobresaliente**. En la siguiente tabla tenemos las calificaciones que ha puesto a dos grupos de alumnos:*

Grupo1	I	A	A	N	N	S	S	I	I	I	A	A	A	N	S	S	I	A	A	S	S	S	S

Grupo 2	S	S	I	I	A	N	A	N	I	I	S	N	A	S	I	N	N

a) ¿Cuál es la variable observada y de qué tipo es?

b) ¿Cuál de los grupos ha obtenido mejores calificaciones? Explique su respuesta

c) Si tuviera que emplear un calificativo para el grupo, ¿Cuál sería el más representativo (y apropiado) para el grupo 1? ¿y para el grupo 2? Explique su respuesta.

TABLA 35. Frecuencia de tipologías de significados personales declarados en el ítem 3a.

n=49				
	Correcta (77,5%)	**Incorrecta (22,5%)**	**No contesta**	**Total**
3.a	38	8	3	**49**

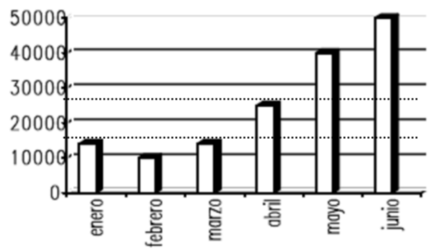
TABLA 36. Frecuencia de tipologías de significados personales declarados en el ítem 3b.

n=49						
Tipologías	**Correcta (10,2%)**		**Incorrecta (89,8%)**			**Total**
	Argumenta	**No argumenta**	**Argumenta**	**No argumenta**	**No contesta**	
Media	3	0	0	0	0	3
Moda	2	0	4	0	0	6
Mediana	0	0	0	0	0	0
Otras justificaciones	0	0	36	0	0	36
No responde	0	0	0	0	4	4
total	5	0	40	0	4	49

El porcentaje de respuestas incorrectas es mucho mayor a los obtenidos en el estudio de Mayén con alumnos de bachillerato (64%).

Ítem 4 (PGTA 5 DEL EXAMEN)

Observe el siguiente diagrama de barras que muestra el número de sándwiches que vendió la empresa Ambrossi durante el primer semestre del presente año:



e) ¿Cuál es la variable observada y de qué tipo es?
f) Dé un valor aproximado del promedio de sándwiches que la empresa vendió al mes.
g) Dé un valor aproximado de la moda del número de sándwiches que la empresa vendió al mes.
h) Dé un valor aproximado de la mediana del número de sándwiches que la empresa vendió al mes.

TABLA 37. Frecuencia de tipologías de significados personales declarados en el ítem 4.

	Correcta	Incorrecta	Total
Ítem 4.a	39 (79,6%)	10 (20,4%)	49
Ítem 4.b	43 (87,8%)	6 (12,2%)	49
Ítem 4.c	42 (85,7%)	7 (14,3%)	49
Ítem 4.d	33 (67,3%)	16 (32,7%)	49

En general, respecto a este ítem, los resultados obtenidos en nuestro estudio son positivos y, los porcentajes de respuestas correctas superan ampliamente a los de Mayén. Sin embargo, resulta importante señalar que - respecto a los ítems 4a, 4b y 4c - se obtuvo un menor porcentaje de respuestas correctas en el ítem 4d, donde se pide calcular la mediana. Tal como ocurrió en el estudio de Mayén.

Ítem5. (PGTA 6 DEL EXAMEN)

a) Un informe del INEI en el 2007, dice que "en la provincia de Lima, el número medio (promedio) de hijos por familia es 1,3". Explique qué significa para usted esta expresión.

b) Carlos tomó información sobre el número de hijos en 10 familias de San Borja. Registró la información, pero por un error de digitación se le borraron algunos datos.

Familia A	3
Familia B	2
Familia C	
Familia D	
Familia E	
Familia F	
Familia G	
Familia H	
Familia I	
Familia J	

- *¿Qué datos podrían haber estado en su cuadro, si se sabe que el promedio de hijos por familia es 1,2?*
- *¿Existe una única posibilidad de completar el cuadro? Justifique su respuesta.*

TABLA 38. Frecuencia de tipologías de significados personales declarados en el ítem 5a.

n=49				
Ítem	**Correcta (36,7%)**	**Incorrecta (63,3%)**		**Total**
			No contesta	
5a	18	26	5	**49**

El porcentaje de respuestas correctas está por debajo al obtenido por Mayen (80%)

TABLA 39. Frecuencia de tipologías de significados personales declarados en el ítem 5b.1.

n=49				
Ítem	**Correcta (63,3%)**	**Incorrecta (36,7%)**		**Total**
			No contesta	
5b.1	31	12	6	**49**

Los resultados obtenidos, respecto al ítem *5b.1,* son muy próximos a los obtenidos por Mayén (69% de correctas)

TABLA 40. Frecuencia de tipologías de significados personales declarados en el ítem 5b.2.

n=49						
Ítem	**Correcta (57,1%)**		**Incorrecta (42,9%)**			**Total**
	Argumenta	**No argumenta**	**Argumenta**	**No argumenta**	**No contesta**	
5b.2	27	1	9	2	10	**49**

Ítem6.

Cierto grupo de niños, el 20 de julio de 2009, tenía como edad promedio 13,8 años. ¿Es verdad que la edad promedio del mismo grupo de niños, el 20 de julio de 2013 será 18,2 años? Justifique su respuesta.

TABLA 41. Frecuencia de tipologías de significados personales declarados en el ítem 6.

n=49						
Ítem	**Correcta (51%)**		**Incorrecta (49%)**			**Total**
	Argumenta	**No argumenta**	**Argumenta**	**No argumenta**	**No contesta**	
6	25	0	17	0	7	**49**

Ítem7.

Se realizó una encuesta a los asistentes al Coloquio de Ciencias Humanas para conocer cuál fue el primer medio por el cual se informaron de esta actividad. Los resultados fueron los siguientes:

Medio de información	Número de personas
Periódico	33
E-mail	55
Página Web	30
TV	32
Radio	40

a) ¿Cuál es la variable observada y de qué tipo es?

b) ¿Cuál de las tres medidas de posición central (media, mediana y moda) tiene sentido hallar en este contexto? ¿Por qué?

TABLA 42. Frecuencia de tipologías de significados personales declarados en el ítem 7a.

n=49			
Ítem	**Correcta (53,1%)**	**Incorrecta (46,9%)**	**Total**
7a	26	23	**49**

TABLA 43. Frecuencia de tipologías de significados personales declarados en el ítem 7b.

n=49						
Ítem	**Correcta (59,2%)**		**Incorrecta (40,8%)**			**Total**
	Argumenta	**No argumenta**	**Argumenta**	**No argumenta**	**No contesta**	
7b	28	1	19	0	1	**49**

7.2.1.1 RESUMEN Y CONCLUSIONES DEL ESTUDIO CUANTITATIVO. Se presenta a continuación el resumen cuantitativo de las respuestas de los estudiantes a las preguntas propuestas para la evaluación.

TABLA 44. Resumen cuantitativo de las respuestas a los ítems propuestos.

Ítem	Correcta	Incorrecta
1	81,6%	18,4%
2a	59,2%	40,8%
2b.1	67,3%	32,7%
2b.2	59,2%	40,8%
2b.3	26,5%	73,5%
3a	77,5%	22,5%
3b	10,2%	89,8%
4a	79,6%	20,4%
4b	87,8%	12,2%
4c	85,7%	14,3%
4d	67,3%	32,7%
5a	36,7%	63,3%
5b.1	63,3%	36,7%
5b.2	57,1%	42,9%
6	51%	49%
7a	53,1%	46,9%
7b	59,2%	40,8%

Dentro de los resultados más relevantes, ya sea por presentar un mayor o elevado porcentaje de respuestas incorrectas, tenemos a los ítems 2a, 2b.2, 2b.3, 3b, 5a, 5b.2, 6, 7a y7b. Para el análisis cualitativo, consideraremos los ítems 2b.3, 3b y 5a ya que son los problemas donde se han presentado los porcentajes más elevados de respuestas incorrectas y, suponemos, ha habido mayor dificultad. Estos problemas están relacionados, principalmente con reconocer si la media es o no un buen representante de un conjunto de datos en presencia de valores atípicos (2b.3), comparación de dos grupos de datos ordinales (3b) y la interpretación adecuada de la media es un contexto específico (5a).

7.2.2 ESTUDIO CUALITATIVO DE LAS RESPUESTAS (configuraciones cognitivas).

En esta sección realizamos un análisis cualitativo de los significados personales declarados de los alumnos. Para ello, primero clasificamos los diferentes tipos de respuestas, para luego realizar una configuración cognitiva a algunas respuestas de los estudiantes respecto a los ítems 2b.3, 3b y 5a, y examinarlas teniendo en cuenta las configuraciones realizadas al elaborar el cuestionario. Para esto nos valemos de la metodología seguida por Ortiz y Font (2011) en su estudio sobre los significados personales de los futuros profesores de educación primaria.

Ítem2b.3 (PGTA 3 DEL EXAMEN)

En cierta provincia de un país, que tiene cuatro distritos, se han hecho estudios sobre el número promedio de hijos por familia.

Distrito	Número de familias	Promedio de hijos por familia
A	35	4,2
B	20	2,6
C	45	5,4
D	30	3,0

a) *¿Cuál es el promedio de hijos por familia, en toda la provincia?*
b) *En el distrito B hay diez niños menores de 10 años. Los pesos, en kilos, de nueve de ellos están dados en la siguiente tabla:*

14	25	17	16	26	17	18	19	24

- *¿Cuál es la mediana de estos datos?*
- *¿Cuál es la mediana si incluimos el peso del décimo niño, que es 41 kg?*
  En este caso, ¿sería la media aritmética un buen representante de los 10 datos? Explique su respuesta.

CATEGORÍA 1. AFIRMAN QUE LA MEDIA NO ES UN BUEN REPRESENTANTE DEL CONJUNTO DE DATOS. Calculan la mediana, para un número de datos pares, de forma correcta.

1.1 Justifican correctamente (A1, A2, A12, A13, A17, A21, A24, A35, A36, A40, A45, A46, A48). Trece alumnos reconocen la presencia de un peso que está bastante alejado de los demás (dato atípico) que se hace que la media no sea representativa y justifican correctamente. Se muestra una de las respuestas de los alumnos y su respectiva configuración cognitiva.

FIGURA 5. Respuesta del alumno A13 al ítem 2b.3.

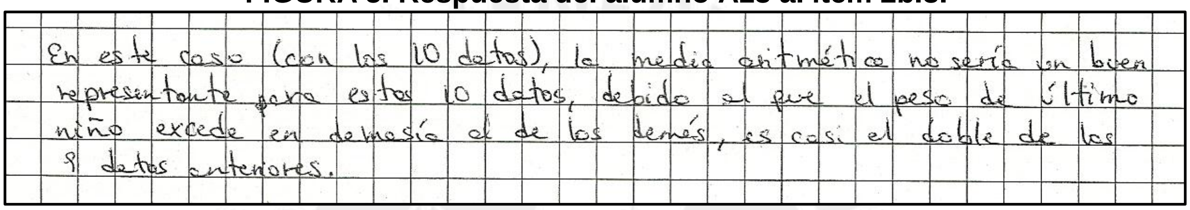

En este caso (con los 10 datos), la media aritmética no sería un buen representante para estos 10 datos, debido a que el peso del último niño excede en demasía al de los demás, es casi el doble de los 9 datos anteriores.

TABLA 45. Configuración cognitiva del alumno A13, respecto al ítem 2.b.3.

LENGUAJE " En este caso (con los 10 datos), la media aritmética no sería un buen representante para estos 10 datos, debido a que el peso del último niño excede en demasía al de los demás, es casi el doble de los anteriores"	Verbal relacionado con el contexto: peso, representante, excede, demasía, doble.
SITUACIÓN PROBLEMA	Obtención del mejor representante de un conjunto de datos numéricos, con la presencia de un valor atípico.
CONCEPTOS	❖ Media ❖ Valor atípico
PROPIEDADES/PROPOSICIONES	❖ El promedio se ve afectado (y no es buen representante) cuando, en el conjunto de datos, hay presencia de valores atípicos.

ALGORITMOS Y PROCEDIMIENTOS	❖ Observar la presencia de un valor atípico y notar que eso modifica significativamente el valor de la media.
ARGUMENTO	❖ Tesis: la media aritmética no sería un buen representante. ❖ Argumento: "debido a que el peso del último niño excede en demasía al de los demás…"

1.2 No justifican correctamente o es ambigua (A7, A18, A25, A27, A33, A49). Seis alumnos brindan diversas respuestas que afirman que la media no es un buen representante; sin embargo, no logran justificar de forma correcta. El alumno A25 responde: *"En este caso la media aritmética sería 17,6 kg, y no sería un buen representante del conjunto de los 10 datos pues se aleja del dato 41kg."* A continuación mostramos su configuración cognitiva.

TABLA 46. Configuración cognitiva del alumno A25, respecto al ítem 2.b.3.

LENGUAJE "En este caso la media aritmética sería 17,6 kg. y no sería un buen representante del conjunto de los 10 datos pues se aleja del dato 41kg."	Simbólico: kg. Verbal relacionado con el contexto: peso, representante, aleja. Numéricos: números decimales.
SITUACIÓN PROBLEMA	Obtención del mejor representante de un conjunto de datos numéricos, con la presencia de un valor atípico.
CONCEPTOS	❖ Media
PROPIEDADES/PROPOSICIONES	❖ El promedio se ve afectado (y no es buen representante) cuando, en el conjunto de datos, hay presencia de valores atípicos.
ALGORITMOS Y PROCEDIMIENTOS	❖ Calcular la media y observar que el

	resultado está bastante alejado del dato 41kg.
ARGUMENTO	❖ Tesis: la media aritmética no sería un buen representante. ❖ Argumento: "…La media aritmética sería 17,6 kg. […] se aleja del dato 41kg."

Observación: El alumno, presenta primero un error de cálculo, pues la media (incluido el peso de 41kg) resulta 21,7 kg. En seguida, observa que la media que obtuvo está alejado del valor atípico 41kg. y considera que por esa razón la media no es representativa.

CATEGORÍA 2. CONSIDERAN QUE LA MEDIA SI ES UN BUEN REPRESENTANTE DEL CONJUNTO DE DATOS.

2.1 Calculan correctamente la media pero no reflexionan sobre la presencia del valor atípico, que hace que la media no sea un buen representante (A3, A5, A6, A8, A14, A19, A32, A37, A39, A41, A43, A47). Doce alumnos, a pesar que calculan bien la media, responden de manera incorrecta a la pregunta, al considerar que la media si es un buen representante del conjunto de datos. El alumno A14 responde: *"La media aritmética o promedio representaría mejor los datos pues incluye a todos en conjunto".* A continuación mostramos su configuración cognitiva.

TABLA 47. Configuración cognitiva del alumno A14, respecto al ítem 2.b.3.

LENGUAJE “La media aritmética o promedio representaría mejor los datos pues incluye a todos en conjunto”	Verbal relacionado con el contexto: media aritmética, promedio, representante, mejor, incluye, todos, conjunto.
SITUACIÓN PROBLEMA	Obtención del mejor representante de un conjunto de datos numéricos, con la presencia de un valor atípico.
CONCEPTOS	❖ Media
PROPIEDADES/PROPOSICIONES	❖ La media aritmética sería un buen representante.
ALGORITMOS Y PROCEDIMIENTOS	❖ Cálculo del nuevo valor de la media (con los 10 datos).
ARGUMENTO	❖ Tesis: la media aritmética o promedio representaría mejor a los datos. ❖ Argumento: “ pues incluye a todos en conjunto”

Observación: El alumno reconoce que la media aritmética considera todos los valores del conjunto, sin embargo, no reflexiona en el hecho que, por esa misma razón, se ve afectada por el dato atípico, dejando así de ser representativa del conjunto de datos. Este caso no lleva a pensar en un problema relevante: varios alumnos han mecanizado el algoritmo de cálculo de la media, pero no han interiorizado el significado del mismo y sus propiedades más básicas.

2.2 Calculan mal la media y, además, no reflexionan sobre la presencia del valor atípico, que hace que la media no sea un buen representante (A6, A22). Mostramos la siguiente respuesta del alumno A6.

FIGURA 6. Respuesta del alumno A6, al ítem 2.b.3.

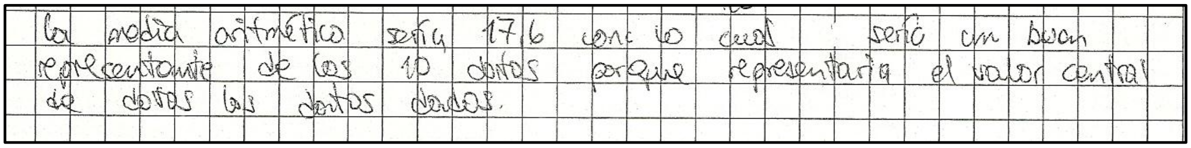

la media aritmético sería 17,6 con lo cual sería un buen representante de los 10 datos porque representaría el valor central de datos los datos dados.

TABLA 48. Configuración cognitiva del alumno A6, respecto al ítem 2.b.3.

LENGUAJE " La media aritmética sería 17,6, con lo cual sería un buen representante de los 10 datos, porque representaría el valor central de los datos dados"	Verbal relacionado con el contexto: media aritmética, representante, valor central. Numérico: números decimales.
SITUACIÓN PROBLEMA	Obtención del mejor representante de un conjunto de datos numéricos, con la presencia de un valor atípico.
CONCEPTOS	❖ Media
PROPIEDADES/PROPOSICIONES	❖ La media aritmética sería un buen representante.
ALGORITMOS Y PROCEDIMIENTOS	❖ Cálculo del nuevo valor de la media (con los 10 datos).
ARGUMENTO	❖ Tesis: la media aritmética sería un buen representante. ❖ Argumento: "porque representaría el valor central de los datos"

Observación: El alumno calcula mal el nuevo valor de la media (no muestra su procedimiento). Además no reflexiona sobre la presencia del valor atípico. También presenta conflictos semióticos al asociar el término "valor central" a la media.

2.3 Identifican la presencia de un valor atípico, pero escoge la media como representante (A9). El alumnos responde: *"Una media aritmética sería mejor* *representante porque el último peso posee una gran diferencia con los pesos anteriores, así que necesitamos equilibrar los pesos antes que solo ubicar el punto medio de los valores".* Si bien notamos que el alumno reconoce la presencia de un peso que se encuentra alejado de los demás, insiste en que la media es el mejor representante, y considera a la media como el punto de "equilibrio", que si bien es correcto, no determina que la media sea buen representante. El alumno intenta comparar el valor de la nueva media con el valor de la mediana, sin notar que, precisamente en este caso, la mediana es mejor representante que la media.

2.4 Compara el valor de la nueva media de los pesos, con la media o con la mediana hallada en los ítems anteriores (A11, A16) El alumno A16 responde: *"En este caso la media aritmética sería un buen representante de los 10 niños, ya que se encuentra muy cerca del valor de la mediana"*

2.5 Justificación carece de sentido (A28, A29). Responden de manera incorrecta y, además, brindan justificaciones que carecen de sentido.

CATEGORÍA 3. NO RESPONDE O SOLO CALCULA EL VALOR DE LA NUEVA MEDIA (A4, A10, A15, A20, A23, A26, A31, A34, A38, A42, A44). Fueron 11 alumnos los que no respondieron a este ítem.

Ítem 3b. (PGTA 4 DEL EXAMEN)

*Un profesor califica el rendimiento de sus alumnos en el curso de Educación Física del siguiente modo: **I=Insuficiente, A=Aprobado, N=Notable, S=Sobresaliente**. En la siguiente tabla tenemos las calificaciones que ha puesto a dos grupos de alumnos:*

Grupo1	I	A	A	N	N	S	S	I	I	I	A	A	A	N	S	S	I	A	A	S	S	S	S

Grupo 2	S	S	I	I	A	N	A	N	I	I	S	N	A	S	I	N	N

a) ¿Cuál es la variable observada y de qué tipo es?

b) ¿Cuál de los grupos ha obtenido mejores calificaciones? Explique su respuesta

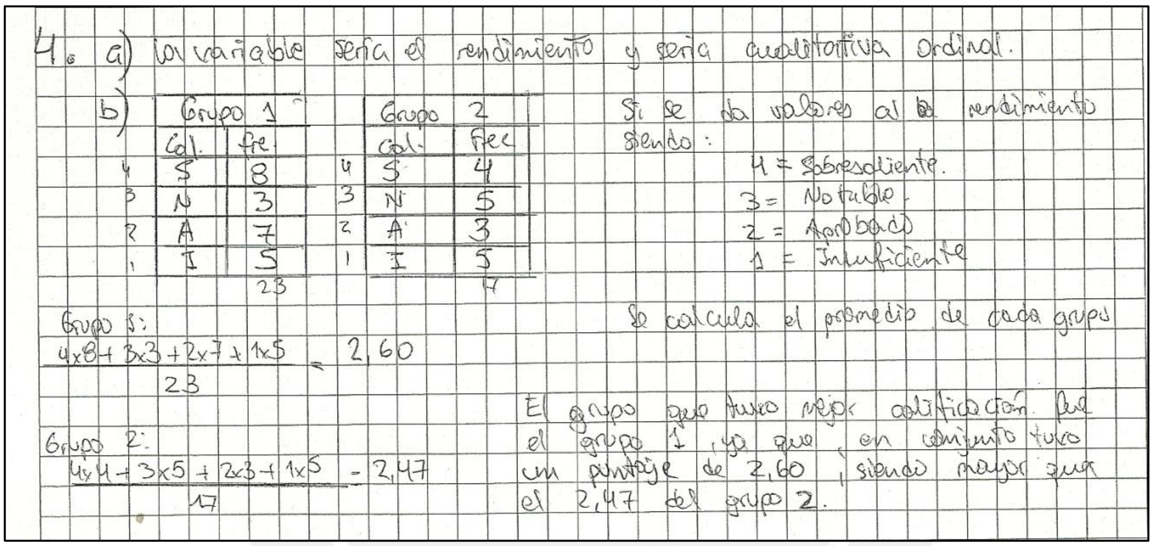
c) Si tuviera que emplear un calificativo para el grupo, ¿Cuál sería el más representativo (y apropiado) para el grupo 1? ¿y para el grupo 2? Explique su respuesta.

Item 3a. Un 77,5% lo responde correctamente.

Item 3b

CATEGORIA 1. RESPUESTAS BASADAS EN LA MEDIA ARITMÉTICA (A6, A13, A23). Convierten la variable cualitativa ordinal en una escala numérica y calculan la correspondiente media aritmética. Tres alumnos mostraron este proceso. El alumno A6 presentó la siguiente solución.

FIGURA 7. Respuesta del alumno A6, al ítem 3b.


4. a) La variable sería el rendimiento y sería cualitativa ordinal.

b)

	Grupo 1	
	Cal.	fre.
4	S	8
3	N	3
2	A	7
1	I	5
		23

	Grupo 2	
	Cal.	Frec
4	S	4
3	N	5
2	A	3
1	I	5
		17

Si se da valores al rendimiento siendo:
4 = Sobresaliente.
3 = Notable.
2 = Aprobado
1 = Insuficiente

Se calcula el promedio de cada grupo

Grupo 1:
$$\frac{4\times8+3\times3+2\times7+1\times5}{23}=2,60$$

Grupo 2:
$$\frac{4\times4+3\times5+2\times3+1\times5}{17}=2,47$$

El grupo que tuvo mejor calificación fue el grupo 1, ya que en conjunto tuvo un puntaje de 2,60, siendo mayor que el 2,47 del grupo 2.

TABLA 49. Configuración cognitiva del alumno A6, respecto al ítem 3b.

LENGUAJE Si se da valores al rendimiento, siendo: 4=Sobresaliente 3=Notable 2=Aprobado 1= Insuficiente	Verbal relacionado con el contexto: valores, rendimiento, promedio, mejor, calificación. Numérico: números decimales.
SITUACIÓN PROBLEMA	Comparación de dos grupos cuyos datos son variables cualitativas ordinales.
CONCEPTOS	❖ Media ❖ Mediana ❖ Moda
PROPIEDADES/PROPOSICIONES	❖ El grupo que presenta mayor

	promedio, obtuvo mayor calificación.
ALGORITMOS Y PROCEDIMIENTOS	❖ Cuantifica los datos de cada grupo. ❖ Calcula la media para cada grupo. ❖ Compara los promedios obtenidos en cada grupo. ❖ Concluye luego de comparar los promedios obtenidos en cada grupo.
ARGUMENTO	❖ Tesis: El grupo que obtuvo mejor calificación fue el grupo 1. ❖ Argumento: “ya que, en conjunto, obtuvo de 2,6, siendo mayor que 2,47 del grupo 2 ”

Observación: Esta solución no es la ideal, pues, aunque reconoce que se trata de una variable cualitativa ordinal, no le da el tratamiento correcto luego de cuantificar los valores de los conjuntos. Hubiera sido correcto si, luego de cuantificar los valores, hubiera calculado la mediana o la moda. Sin embargo, consideramos esta respuesta como parcialmente correcta, dado que los alumnos no tuvieron mucha práctica con este tipo de problemas y variables.

CATEGORÍA 2. RESPUESTAS BASADAS EN LA MODA. En esta categoría, hubo respuestas correctas e incorrectas:

3.1. CORRECTAS: (A4, A44). Dos alumnos usaron la moda para comparar las calificaciones en ambos grupos, aunque no completaron la respuesta. La siguiente solución corresponde al alumno A4.

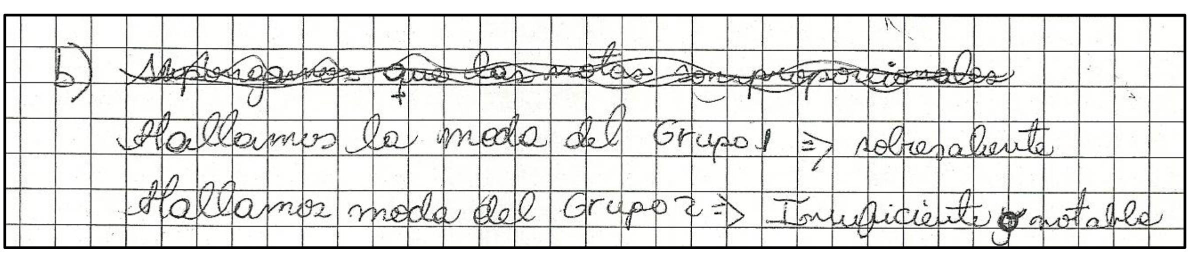


FIGURA 8. Respuesta del alumno A4, al ítem 3b.

3.2. CON ERROR:(A22, A42, A47, A49). Cuatro alumnos usaron el término “mayor cantidad”, “más alumnos” para hacer referencia a la moda; sin embargo, no tienen clara la definición, pues toman como referencia a varias de las calificaciones, incluso, aquella que tiene menor frecuencia (Notable). El alumno A47 presentó la siguiente respuesta.

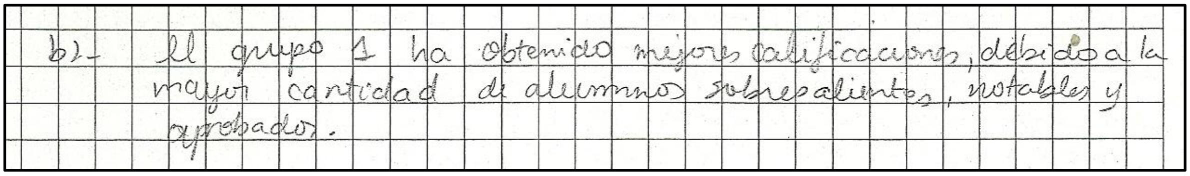

b)- El grupo 1 ha obtenido mejores calificaciones, debido a la mayor cantidad de alumnos sobresalientes, notables y aprobados.

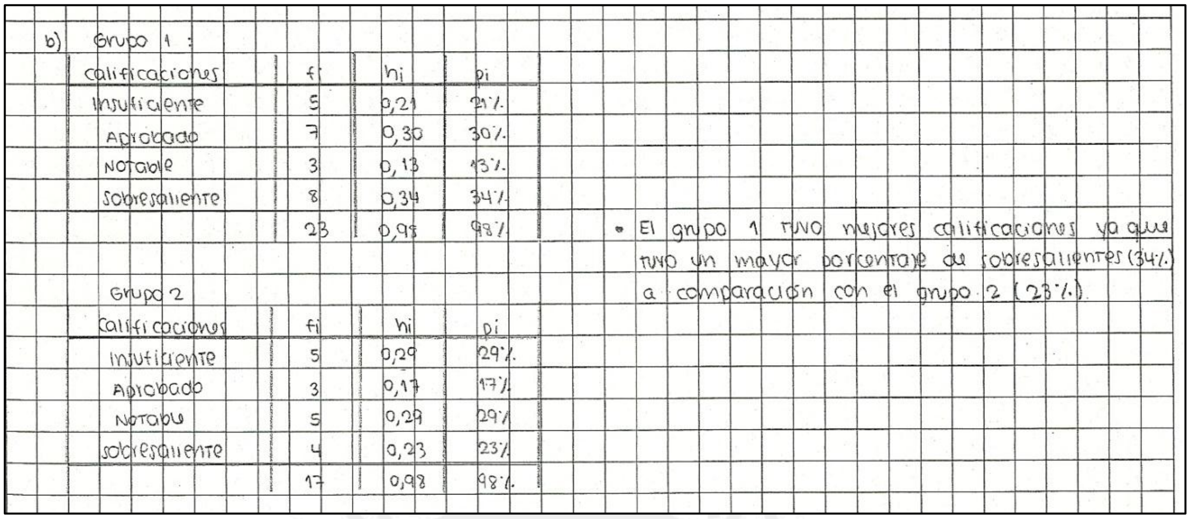
FIGURA 19. Respuesta del alumno A47, al ítem 3b.

Observaciones: Es importante señalar que la moda solo mide la frecuencia con que se repiten los datos, más no su orden, como lo hace la mediana (razón por la cual no resulta tan representativo para poder luego hacer las comparaciones entre ambos grupos); sin embargo, también la consideramos como una respuesta válida pues la moda es aplicable en variables cualitativas ordinales.

CATEGORÍA 3. COMPARAN USANDO PORCENTAJES, FRECUENCIA RELATIVA, ANALIZA DE MANERA INDEPENDIENTE O REALIZA OTRAS OPERACIONES. (A1, A2, A3, A5, A7, A8, A9, A11, A12, A14, A16, A17, A19, A20, A21, A22, A24, A25, A26, A28, A29, A30, A31, A32, A33, A34, A35, A36, A37, A38, A39, A40, A41, A43, A45, A46, A48)

La gran mayoría (37 alumnos) optaron por usar los porcentajes o las frecuencias relativas para comparar las calificaciones en ambos grupos. Consideramos estas respuestas como incorrectas ya que analizan los datos de manera aislada. El alumno A36 mostró la siguiente respuesta:

FIGURA 20. Respuesta del alumno A36, al ítem 3b.



b) Grupo 1 :

calificaciones	fi	hi	pi
Insuficiente	5	0,21	21%
Aprobado	7	0,30	30%
Notable	3	0,13	13%
Sobresaliente	8	0,34	34%
	23	0,98	98%

• El grupo 1 tuvo mejores calificaciones ya que tuvo un mayor porcentaje de sobresalientes (34%) a comparación con el grupo 2 (23%).

Grupo 2

Calificaciones	fi	hi	pi
Insuficiente	5	0,29	29%
Aprobado	3	0,17	17%
Notable	5	0,29	29%
Sobresaliente	4	0,23	23%
	17	0,98	98%

TABLA 50. Configuración cognitiva del alumno A36, respecto al ítem 3b.

LENGUAJE	Verbal relacionado con el contexto: Mejor, calificación, porcentaje. Simbólico: f_i, h_i, $p_{i,}$ %, tablas. Numérico: números decimales y porcentajes.
SITUACIÓN PROBLEMA	Comparación de dos grupos cuyos datos son variables cualitativas ordinales.
CONCEPTOS	❖ Tabla de frecuencias ❖ Frecuencia absoluta ❖ Frecuencia relativa ❖ Frecuencia relativa en %
PROPIEDADES/PROPOSICIONES	❖ El dato que tenga mayor % representa a cada grupo.
ALGORITMOS Y PROCEDIMIENTOS	❖ Diseña y completa una tabla de frecuencias para cada grupo. ❖ Identifica el dato que tienen mayor porcentaje en el grupo 1. ❖ Observa que porcentaje tiene en el grupo 2, el dato obtenido en el paso

	anterior. ❖ Responde luego de comparar los porcentajes de ese dato en ambos grupos.
ARGUMENTO	❖ Tesis: El grupo 1 tuvo mejores calificaciones. ❖ Argumento: “ya que tuvo un mayor porcentaje de sobresalientes (34%) a comparación con el grupo 2 (23%).

Observación: Esta respuesta sólo compara datos aislados en ambos grupos, es decir, no considera los demás datos que también están en juego. En consecuencia, el dato obtenido no es representativo de los conjuntos. El alumno no identifica el problema como una situación donde se hace necesario el uso de las medidas de tendencia central.

CATEGORÍA 5. NO RESPONDEN (A10, A15, A18, A27). Cuatro alumnos no respondieron a este ítem.

Observación: A pesar de que la mediana se presenta como la medida más idónea para este tipo de problemas, por tratarse de una variable cualitativa ordinal, sorprende que ningún alumno lo haya usado para comparar ambos grupos.

Ítem5a. (PGTA 6 DEL EXAMEN)

a) Un informe del INEI en el 2007, dice que "en la provincia de Lima, el número medio (promedio) de hijos por familia es 1,3". Explique qué significa para usted esta expresión.

CATEGORIA 1. ARGUMENTAN ADECUADAMENTE.

1.1 Recurren a la definición (A2, A4, A7, A8, A9, A14, A16, A17, A23, A26, A28, A38, A41, A44). Catorce alumnos responden usando la definición de media. Una respuesta es el que presentó el alumno A4: *"La suma total del número de hijos de cada familia sobre el número total de familias es 1,3".* El alumno A30 responde con la operación inversa: *"1,3 es el número de hijos promedio en la provincia de Lima, es decir, ese promedio multiplicado por el número de familias dará como resultado el total de hijos en Lima".*

1.2 El número de hijos por familia están alrededor de la media (A24, A39, A46). Tres alumnos usan esta explicación. El alumno A24 respondió: *"Que oscila entre 1 y 2, en promedio, pero este número es propenso a acercarse más a 1 que a 2. Entiendo que hay más personas con 1 que con dos hijos"*

CATEGORIA2. ARGUMENTAN ERRÓNEAMENTE O REPITEN EL ENUNCIADO DEL PROBLEMA.

2.1 Repiten el mismo enunciado del problema (A11, A20, A29, A32). Una respuesta que reproducimos es la del alumno A29: *"Esta expresión significa que en toda la provincia de Lima se estima que las familias que viven ahí tienen 1,3 hijos por familia"*

2.2 La media debe ser uno de los datos (A13, A22, A25, A31, A35, A45) El alumno A13 respondió lo siguiente: *"Se refiere que al sumar el número de hijos y dividirlo por el número de familias dio 1,3. Pero no se puede tener 1,3 sino sería un hijo (que sería el valor más cercano). Así que el promedio aproximadamente de hijos por familia sería 1".*

TABLA 51.Configuración cognitiva del alumno A13, respecto al ítem 5a.

LENGUAJE “Se refiere que al sumar el número de hijos y dividirlo por el número de familias dio 1,3. Pero no se puede tener 1,3, sino sería un hijo (que sería el valor más cercano). Así que el promedio aproximadamente de hijos por familia sería 1”.	Verbal, relacionado con el contexto: Sumar, dividir, cercano, promedio de hijos por familia. Numérico: números decimales.
SITUACIÓN PROBLEMA	Explicación intuitiva del significado de la media de valores enteros, cuando ésta no es entera.
CONCEPTOS	❖ MEDIA
PROPIEDADES/PROPOSICIONES	❖ El número de hijos por familia, en cada caso, es un número entero.
ALGORITMOS Y PROCEDIMIENTOS	❖ Explicación verbal.
ARGUMENTO	❖ Tesis: El promedio aproximado de hijos por familia es 1. ❖ Argumento: Si los datos son números enteros, la media debe ser un número entero, haciendo el redondeo si fuera necesario.

Observación: El alumno piensa que, al estar trabajando con valores enteros (número de hijos por familia), la media debería de ser también un valor entero. Esto no es cierto, pues la media no necesariamente es uno de los valores de la distribución. El redondeo puede distorsionar la información.

2.3 ***Respuesta que carece de sentido para la variable usada*** (A5, A18, A19, A34, A36, A37, A40, A49). Ocho alumnos no contextualizan de forma adecuada el problema. Mostramos una de las respuestas, alumno A18: *“El número que representa la cantidad de hijos por familia en total”*

2.4 Confunde el concepto de media con moda (A1, A21, A33). El alumno A1 responde: *" Significa que una familia en el 2007 en su mayoría tiene 1 o 2 hijos en la provincia de Lima"*

2.5 Confunde el concepto de media con mediana (A6, A42) El alumno A6 responde: *"Quiere decir que 1, 3 hijos por familia sería el valor que represente el número central de la cantidad de hijos por familia…"*

Observación: En estos dos últimos casos, se presentan confusiones de la media con la mediana y la moda. Usan las expresiones "mayoría" y "número central" para intentar explicar la media del número de hijos en la provincia de Lima. No se están identificando bien los términos relacionados con cada uno de los conceptos.

2.6 Brinda una respuesta que no es acorde con lo que se pide *(A2, A3, A48)* Tres alumnos no dieron una respuesta en el contexto del problema ni con lo que pedía.

2.7 No responde *(A10, A15, A27, A43, A47).* Cinco alumnos no respondieron a este ítem.

7.2.2.3 CONCLUSIONES DEL ESTUDIO CUALITATIVO DE LAS RESPUESTAS DE LOS ALUMNOS.

Se analizaron los ítems 2b.3, 3b y 5a, que presentaron bajos porcentajes de respuestas correctas. Clasificamos, en cada uno de ellos, los diversos tipos de respuestas que presentaron los alumnos. Así también, se mostraron para cada clasificación uno o dos ejemplos tipos. A varias de las respuestas se le realizó una configuración cognitiva, producto de las cuales, tenemos las siguientes conclusiones:

- Respecto al ítem 2.b.3, sólo el 26,5% (13 alumnos) respondió de manera correcta, identificando que la presencia de un valor atípico hace que la media ya no sea representativa. Otros doce alumnos calculan bien la media, pero no reflexionan sobre el valor atípico y responden erróneamente. Esto nos lleva a pensar que la parte algorítmica, por lo general, los alumnos la han aprendido bien; no así, la parte conceptual y sus propiedades más básicas. Por ello, será importante enfatizar en la comprensión de estos conceptos, así como en el comportamiento de las mismas.

- Respecto al ítem 3b, donde sólo el 10,2% (5 alumnos) responden de manera correcta, llama la atención que ningún alumno haya usado la mediana como medida representativa para comparar dos conjuntos con datos cualitativos ordinales, siendo ésta la más idónea. Por el contrario la mayoría (37 alumnos) acudió al cálculo de frecuencias relativas para comparar los conjuntos, sin darse cuenta que esto solo compara datos de manera aislada. Ninguno de estos 37 alumnos recurrió a las medidas de posición central para responder a esta pregunta. Consideramos que será necesario revisar estos tipos de problemas con mayor detenimiento.
- En lo que respecta al ítem 5a, un 34,7% responde de forma correcta, interpretando el algoritmo de la media. Dentro de los errores más frecuentes se encuentra al pensar que la media debe ser necesariamente uno de los datos del conjunto. Otro conflicto es la verbalización del concepto de media y su confusión con la mediana y la moda, presentándose así un conflicto de tipo semiótico.

CAPITULO VIII

CONCLUSIONES Y RECOMENDACIONES

8.1 CONCLUSIONES

1. En los libros de texto y en las investigaciones sobre didáctica de las medidas de tendencia central, predominan la perspectiva de estos conceptos como medidas de resumen y no como estimadores de parámetros de una población. Consideramos importante que esta perspectiva esté presente en los cursos de estadística básica y se dé una visión de su uso en el marco de la estadística inferencial.
2. En el capítulo 4 se analizó y se describió el significado institucional de referencia de las medidas de tendencia central, cumpliendo así nuestro primer objetivo. Las conclusiones se presentaron en la p.58 (Sección 4.1.2). En general, se puede concluir que los textos analizados no son del todo adecuados, a pesar de tener una variedad de problemas contextualizados, pues enfatizan en el algoritmo de cálculo y no en el concepto y sus propiedades. Esto se hace evidente por las siguientes observaciones: Uno de los textos, casi en su totalidad, presenta problemas resueltos y ello no contribuye a estimular el razonamiento y la creatividad del alumno en la búsqueda de sus propias soluciones. No se hace mención a los tipos de variables y tampoco se presentan problemas que involucren a las variables cualitativas, dejando de lado un aspecto importante en cuanto al uso adecuado de los promedios según el tipo de variable en estudio.
3. En el capítulo 5 se analizó y describió el significado institucional pretendido de las medidas de tendencia central, cumpliendo así nuestro segundo objetivo. Se obtuvieron resultados del análisis al texto guía y de la entrevista a tres profesores del curso que se presentan en la sección 5.1. Respecto al texto guía, podemos mencionar que, en general, está diseñado en un contexto cercano al estudiante, lo que hace que su lectura y estudio sea más ágil y motivador. Sin embargo,

como se ha hecho notar, existen algunos vacíos en la introducción del concepto de media, el estudio de sus propiedades, el planteamiento de algunas situaciones problemáticas, así como en el uso de argumentos.

Respecto a los profesores entrevistados, todos hacen mención a la importancia de las medidas de tendencia central ya sea en la vida cotidiana, profesional o académica. Esperan que sus alumnos aprendan a resolver problemas contextualizados, saber cuándo usar qué medida de tendencia central, interpretarlas correctamente, usar la calculadora, entre otros logros. Todos coinciden en que el libro es idóneo para desarrollar el curso y brindan para ello diversas razones. Es interesante hacer notar que varios de los logros que desean alcanzar con sus estudiantes no se encuentran en el texto.

4. En el capítulo 6, se describió el significado institucional implementado, logrando cumplir con ello nuestro tercer objetivo. En general se puede resumir lo siguiente: en la apertura de la clase solo un profesor contextualiza las medidas de tendencia central mediante una pregunta relativa a la moda, sin embargo, en ninguna de las clases observadas se hizo mención a la importancia de estos conceptos. Esto se contrapone, en cierta medida, a los logros que esperaban alcanzar con sus alumnos (significados pretendidos), y que se mencionaron en la conclusión anterior. En el desarrollo de la clase, lo profesores usan las situaciones presentadas en el texto guía y los ejercicios sencillos son usados, principalmente, para visualizar los algoritmos de cálculo más que para su interpretación del mismo. Se cierra la clase con una pequeña prueba, como parte de la evaluación continua en el curso.

   El tiempo que se usó para estudiar estos conceptos (una sesión de 110 minutos) no fue suficiente para abordar con mayor profundidad los conceptos de media, mediana y moda.

5. En el capítulo 7, se analizó y describió el significado personal logrado, de las medidas de tendencia central de los alumnos de las carreras de humanidades de EEGGLL. Con esto se cumplió nuestro cuarto objetivo. Las mayores dificultades se presentaron, principalmente, con problemas que están relacionados con saber reconocer si la media es o no un buen representante de un conjunto de

datos en presencia de valores atípicos (2b.3); obtener conclusiones a partir de dos grupos de datos cualitativos ordinales (3b); y la interpretación adecuada de la media en un contexto específico (5a).

Las respuestas al ítem 2b.3 evidenciaron que los alumnos no identifican la presencia de un valor atípico en una distribución, o si lo hacen, no reflexionan sobre el efecto que causa la presencia del mismo sobre la media. Los alumnos aprendieron el algoritmo de cálculo pero no analizan su resultado en el contexto en el que fue obtenido.

Para el ítem 3b ningún alumno usó la mediana para comparar dos grupos de datos ordinales, siendo ésta la medida más idónea. Un 10,2% optó por la moda o la media (cuantificando los datos). El 90,8% de los alumnos usaron frecuencias relativas para comparar ambos grupos, esto es incorrecto, pues solo compara datos de manera aislada.

Para el ítem 5a un error frecuente de parte de los alumnos consiste en pensar que la media debe ser necesariamente uno de los datos de la distribución. Se muestra también algunos conflictos semióticos, básicamente cuando intentan interpretar la media en un contexto determinado y terminan confundiendo con la mediana o la moda, pues usan términos como “”número central” o “mayoría”.

8.2 RESUMEN DE LAS RELACIONES ENTRE LOS DIVERSOS SIGNIFICADOS DE LAS MEDIDAS DE TENDENCIA CENTRAL.

Nuestro objetivo general fue analizar los significados personales e institucionales de las medidas de tendencia central. Hemos logrado describir los diversos significados institucionales y personales. Queda entonces presentar un breve resumen acerca de cómo se encuentran relacionados estos significados, implicaciones entre ellas, así como el acoplamiento (en términos del EOS) entre el significado implementado y el logrado.

8.2.1 Fundamentos teóricos y los significados institucionales de referencia.

En comparación con los fundamentos teóricos presentados en nuestra investigación, se mostró que las diversas investigaciones sobre didáctica de la estadística, así como los libros de texto analizados abordan el estudio de la media, mediana y moda como una medida de resumen de un conjunto de datos y no como un estimador de parámetros de una población bajo el concepto de variable aleatoria.

8.2.2 Significado institucional de referencia y el pretendido. En los libros de texto analizados para el significado de referencia, así como para el pretendido, encontramos que, para el estudio de las media, mediana y moda, se enfatiza en el algoritmo de cálculo y no en el concepto y sus propiedades. Un aspecto importante que no se toma en cuenta es el uso pertinente de los promedios según el tipo de variable, así como la poca presencia de problemas que involucren a las variables cualitativas. El lenguaje usado es amplio, lo que demuestra que, si bien las medidas de tendencia central suele considerarse como un concepto sencillo, tienen gran riqueza en lenguaje tanto verbal como simbólico, lo cual hace que muchos alumnos confundan los términos y representaciones, presentándose así conflictos de tipo semiótico.

Cabe resaltar, que los profesores que impartieron este curso (matemáticas), consideraron necesario, entre otros logros, que los alumnos, más allá de aprender los algoritmos de cálculo, comprendieran los conceptos y sepan interpretarlos.

8.2.3 Significado institucional pretendido y el implementado. De lo anterior, se hubiera esperado que, los profesores que dictaron el curso de matemáticas, hubieran cubierto esos vacíos que dejaron los libros de texto. Sin embargo, en las clases observadas, no se enfatizó en la comprensión de estos conceptos y su interpretación, así como el uso pertinente de la media, mediana y moda no se abordó con la importancia necesaria que éstos merecen.

Consideramos que uno de los factores se debió al tiempo usado en la sesión dedicada a este tema, el cual fue bastante limitado.

8.2.4 Significado institucional implementado y el significado personal logrado. Se hizo evidente los aspectos no considerados en los diversos significados. Hubo porcentajes bastante bajos de respuestas correctas. Podemos mencionar, entre otros, problemas relacionados con saber reconocer si la media es o no un buen representante de un conjunto de datos en presencia de valores atípicos, esto debido a la poca información que se brindó sobre las características de las medidas de tendencia central. Así también, los alumnos no lograron interpretar la media en un contexto específico, siendo todos los datos números enteros. Se advirtió bastante confusión en el uso de términos. Finalmente, no supieron establecer conclusiones a partir de dos grupos de datos cualitativos ordinales.

8.2.5 UN BREVE COMENTARIO QUE RELACIONA LOS DIVERSOS SIGNIFICADOS. A manera de resumen, podemos manifestar que los significados de referencia reflejados en los textos analizados, por una parte son restringidos a considerar las medidas de tendencia central como medidas de resumen, sin dar una perspectiva de la media como un estimador del parámetro μ de la población; y por otra parte, enfatizan los aspectos algorítmicos y de cálculo, y no la comprensión conceptual de estas medidas; sin embargo, esto está presente en los significados pretendidos, como se refleja en las entrevistas realizadas a los docentes. A pesar de ello, no se encuentra entre los significados institucionales implementados y como consecuencia, tampoco en los significados personales logrados de los estudiantes.

8.3 RECOMENDACIONES

1. Es importante que los profesores que imparten estos temas, asuman una visión más amplia sobre los conceptos de medidas de tendencia central, considerando situaciones que relacionen estos conceptos con la inferencia estadística, al menos, en su aspecto más básico.
2. Es necesario profundizar en los conceptos de las medidas de tendencia central, sus propiedades y la discriminación en el uso de las media, mediana y moda (reconocer cuándo es más significativo o pertinente usar cierta medida tendencia central). El uso correcto de los algoritmos de cálculo por sí solo no basta, y se

hace infructuoso, pues se hace necesario saber interpretar y tener, al menos, el conocimiento básico de su uso en la estadística inferencial.

3. Se debe incluir o potenciar, en el texto guía, problemas que logren suplir los vacíos que han evidenciado los alumnos en nuestro análisis.

4. Teniendo en cuenta los resultados y conclusiones de este trabajo, diseñar una propuesta para la inclusión de las MTC en el marco de los cursos de formación y capacitación docente de educación básica.

5. Desarrollar una investigación similar con estudiantes de ciencias e ingeniería

REFERENCIAS

Alvarado, H. (2007). *Significado institucional y personal del teorema central del límite en la enseñanza de estadística en ingeniería*. Tesis doctoral. Universidad de Granada.

Batanero, C., Godino, J. D. Green, D. R., Holmes, P. y Vallecillos, A. (1994)*. Errores y dificultades en la comprensión de los conceptos estadísticos elementales. International Journal of Mathematics Education in Science and Technology,* 25(4), 527-547. Disponible en: http://www.ugr.es/~batanero/publicaciones%20index.htm

Batanero, C., Godino, J. D. y Navas, F. (1997). *Concepciones de maestros de primaria en formación sobre los promedios.* En H. Salmerón (Ed.), *VII Jornadas LOGSE: Evaluación Educativa* (pp. 310-304). Universidad de Granada.

Batanero, C. (2000). *Significado y comprensión de las medidas de posición central*. *UNO, 2000, 25, 41-58. Disponible en:*

http://www.ugr.es/~batanero/publicaciones%20index.htm

Batanero, C. (2001). *Didáctica de la Estadística.* Granada: Grupo de Investigación en Educación Estadística. Universidad de Granada.

Batanero, C. Godino, J. D. (2001). *Análisis de datos y su didáctica.* Materiales para la asignatura. Departamento de Didáctica de la Matemática. Universidad de Granada. ISBN 84-699-4296-6.

Bazán, J. (2006). *La estadística llega a la escuela en el Perú*. En Gonzales, M., Bazán, J. L., Sánchez, R. (eds). Coloquios sobre Matemática Educativa 2005, parte 2., 87-109. Reporte de Investigación 19. Serie C. Sección Matemática. Pontificia Universidad Católica del Perú.

Campbell, S.K. (1981). *Equívocos y falacias en la interpretación de estadísticas*. Mexico: Limusa.

Cobo, B. (2003). *Significado de las medidas de posición central para estudiantes de secundaria.* Tesis doctoral. Universidad de Granada.

Chan, C. (2009) *Una propuesta didáctica sobre la media aritmética, la mediana y su representatividad*. Tesis de licenciatura en enseñanza de la matemática. Universidad Autónoma de Yucatán, México.
http://www.uady.mx/~matemati/dme/docs/tesis/Tesis_CarlosChan.pdf

Godino, J.D. y Batanero, C. (1994) Significado *institucional y personal de los objetos matemáticos*. *Recherches en Didactique des Mathématiques, 14 (3): 325-355.*Disponible en: http://www.ugr.es/local/jgodino/indice_eos.htm

Godino, J.D. y Batanero, C. y Font, V. (2008). *Un enfoque Ontosemiótico del conocimiento y la instrucción matemática.* Departamento de la Didáctica de la Matemática. Universidad de Granada. Disponible en:

http://www.ugr.es/local/jgodino/indice_eos.htm

Johnson, R. y Kuby, P. (2004). *Estadística Elemental, lo esencial,* 3ra edición. México: Internacional Thomson Editores.

Malaspina, U. (2009). *Intuición y rigor en la resolución de problemas de optimización. Un análisis desde el enfoque Ontosemiótico de la cognición e instrucción matemática.* Tesis doctoral. Pontificia Universidad católica del Perú.

Mayén, S., Cobo, B., Batanero, C. y Balderas, P. (2007). *Comprensión de las medidas de posición central en estudiantes mexicanos de bachillerato. UNION,* 9.

http://www.ugr.es/~batanero/publicaciones%20index.htm

Mayén, S. (2009). *Comprensión de las medidas de tendencia central por estudiantes mexicanos de Educación Secundaria y Bachillerato*. Tesis doctoral. Universidad de Granada.

Murray R. Spiegel y Larry J. Stephens (2009). *Estadística,* 4ta edición. México: Mc Graw Hill

Ortiz de Haro, J. J. (1999). *Significado de los conceptos probabilísticos en los libros de texto de Bachillerato. Tesis doctoral. Universidad de* Granada.

Ortiz, J. y Font, V (2011). *Significados personales de los futuros profesores de educación primaria sobre la media*. Educación Matemática, vol. 23, núm.2, agosto, 2011, pp. 91-109. Grupo Santillana México.
http://www.redalyc.org/articulo.oa?id=40521146005

Oviedo, S. (2013). *Significado de la asimetría estadística en los alumnos de economía de la UNAC. Tesis no publicada.* Maestría en Enseñanza de las Matemática de la PUCP.

Sarabia, J., Gómez, E. y Vásquez, F. (2007). *Estadística actuarial. Teoría y aplicaciones*. PEARSON EDUCACIÓN, S.A. Madrid.

APÉNDICE

FUNDAMENTO TEÓRICO DE LAS MEDIDAS DE TENDENCIA CENTRAL.

Definimos, de manera rigurosa, los conceptos de media, mediana y moda. Todas las definiciones presentadas fueron tomadas de Oviedo, S. (2013), a menos que se indique lo contrario y, en tal caso, corresponden a Sarabia, Gómez, y Vásquez. (2007). Presentamos primero los conceptos preliminares, mediante una lista de definiciones. Luego, pasamos a definir el concepto de esperanza matemática de una variable aleatoria, mediana y moda. Presentamos también, algunas propiedades elementales de la esperanza matemática.

CONCEPTOS PREVIOS.

Experimentos, espacio muestral y eventos.

Definición 1. Un experimento es un modo de obtener observaciones. Los experimentos se dividen en deterministas y aleatorios. Se denomina experimento aleatorio al que puede dar lugar a resultados diferentes, bajo las mismas condiciones experimentales. *Tomado de Sarabia (2007, p.11).*

Definición 2. Se denomina espacio muestral (Ω) , al conjunto de todos los posibles resultados asociados a un experimento aleatorio. Todo subconjunto $A \subset \Omega$ será llamado evento. $A = \varnothing$ es el evento imposible y $A = \Omega$ es el evento seguro. Si $\omega \in \Omega$, el evento $\{\omega\}$ es llamado simple o elemental.

Definición 3. Un evento A al cual atribuimos una probabilidad será llamado evento aleatorio.

Variables aleatorias.

Definición 4. Sea Ω un conjunto no vacío. Una clase de subconjuntos de $\mathbf{A}$ que satisface A1, A2 y A3 se llama álgebra de subconjuntos de $\mathbf{A}$.

A1. $\Omega \in \mathbf{A}$

A.2. Si $A \in \mathbf{A}$, entonces $A^c \in \mathbf{A}$

A.3. Si $A \in \mathbf{A}$ y $B \in \mathbf{A}$, entonces $A \cup B \in \mathbf{A}$ (i.e., si atribuimos una probabilidad a A y otra a B entonces atribuiremos una probabilidad a "A o B".

Sin pérdida de generalidad, se va a suponer que la clase de los eventos aleatorios también satisface:

A.3*. Si $A_n \in \mathbf{A}$ para $n = 1,2,3,\ldots$, entonces $\bigcup_{n=1}^{\infty} A_n \in \mathbf{A}$

Definición 5. Una clase $\mathbf{A}$ de subconjuntos de un conjunto no vacío Ω satisfaciendo A1, A2 y A3* se llama σ-álgebra de subconjuntos de Ω.

Definición 6. Una función P definida en un σ-álgebra $\mathbf{A}$ y que satisface los siguientes axiomas, se llama medida de probabilidad en $\mathbf{A}$ o simplemente probabilidad en $\mathbf{A}$.

Axioma1. $P(A) \geq 0$ para todo *A* de $\mathbf{A}$

Axioma2. $P(\Omega) = 1$

Axioma3. (σ-aditividad). Si $A_1, A_2, \ldots \in \mathbf{A}$ son disjuntos, entonces

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

Definición 7. Un modelo matemático para un experimento, o modelo probabilístico está constituido por:

1. Un subconjunto no vacío Ω de resultados posibles, el espacio muestral.
2. Un σ-álgebra $\mathbf{A}$ de eventos aleatorios.
3. Una probabilidad P definida en $\mathbf{A}$.

Definición 8. Un espacio de probabilidad es una terna $(\Omega, \mathbf{A}, P)$, donde

1. Ω es un conjunto no vacío.
2. $\mathbf{A}$ es un σ-álgebra de subconjuntos de Ω, y
3. P es una probabilidad en $\mathbf{A}$.

Definición 9. Una variable aleatoria X en un espacio de probabilidad $(\Omega, \mathbf{A}, P)$ es una función real definida en el conjunto Ω tal que $[X \le x]$ es un evento aleatorio para todo $x \in IR$. Es decir, $X : \Omega \to IR$ es una variable aleatoria si $[X \le x] \in \mathbf{A}, \forall x \in IR$.

Definición 10. Una variable aleatoria X se llama discreta, si su rango R_x es un conjunto discreto de números reales. Se llama continua, si su rango R_x es un intervalo o unión de intervalos sobre la línea de los reales y tiene probabilidad cero de igualar a cualquier valor aislado en R_x.

Si es X es una variable aleatoria discreta, se puede usar la medida de probabilidad definida en los subconjuntos de Ω para definir la función de probabilidad de X. Esto se puede usar para evaluar las declaraciones de probabilidad de X, lo que se define como sigue:

Definición 11. La función de probabilidad para P es una función (denotada por $p_X(x)$) de una variable real x y se define como:

$$p_X(x) = P\left(X(\omega) = x\right), \text{ para todo } x \text{ real.}$$

Si se tiene un experimento dado y un espacio muestral Ω para tal experimento y se ha definido la variable aleatoria discreta X en los elementos de Ω, entonces se puede encontrar $p_X(x)$ como sigue. Sea

$$A(x) = \{\omega : X(\omega) = x\}$$

Entonces $p_X(x) = P\left(A(x)\right)$ para toda x real. Queda claro que $A(x)$ puede ser igual a ϕ para muchas x y, por lo tanto, $p_X(x) = P\left(A(x)\right)$ es 0 para esas x.

Funciones de distribución y densidad.

Definición 12. La función de distribución para una variable aleatoria X (denotada por $F_X(t)$) es una función de una variable t tal que:

1. El dominio de la función F_X es la línea completa de los reales
2. Para toda t real $F_X(t) = P(X \leq t)$

Se puede obtener de inmediato la función de distribución de una variable aleatoria discreta X si se conoce la función de probabilidad de X. Resulta evidente que:

$$F_X(t) = \sum_{x \leq t} p_X(x)$$

En que la sumatoria es sobre todas las x en el rango de X que satisface la condición $x \leq t$.

Definición 13. La función de densidad de la probabilidad para una variable aleatoria continua Y (denotada por $f_Y(y)$) es una función de una variable real y talque:

1. El dominio de $f_Y(y)$ es la línea completa de los reales
2. Para todo número real t

$$F_Y(t) = \int_{-\infty}^{t} f_Y(y)dy$$

El rango de Y es $R_Y(y) = \left\{y : f_Y(y) \geq 0\right\}$

Cuantil de orden *p*.

La siguiente definición es tomada de Sarabia (2007, p.41)

Definición 14. El cuantil de orden *p* es aquel valor de la distribución que deja a la izquierda una probabilidad acumulada de *p.*

El *p*-cuantil x_p verifica la relación:

$$P(X \leq x_p) \leq p \leq P(X \leq x_p)$$

De este modo, si es una función de distribución continua, el *p*-cuantil x_p se obtiene fácilmente como la solución en x_p de la ecuación:

$$F(x_p) = p$$

ESPERANZA MATEMÁTICA, MEDIANA Y MODA.

Definición 14. Si X es una variable aleatoria discreta con función de probabilidad $p_X(x)$, el valor esperado de $H(X)$ o esperanza matemática (que se escribe como $E\left[H(X)\right]$) se define como:

$$E\left[H(X)\right] = \sum_{R_X} H(x) p_X(x),$$

Siempre y cuando la suma sea absolutamente convergente.

Definición 15. Si X es una variable aleatoria continua con función de densidad $f_X(x)$, el valor esperado de $H(X)$ o como:

$$E\left[H(X)\right] = \int_{-\infty}^{\infty} H(x) f_X(x) dx$$

Siempre y cuando la integral sea absolutamente convergente.

De las dos últimas definiciones, si la suma o la integral no es absolutamente convergente, se dice simplemente que no existe el valor esperado.

Definición 16. Se llama media o valor promedio de X al valor esperado de X, y se denota por μ_X; es decir $\mu_X = E(X)$.

Lo siguiente lo extraemos de Sarabia (2007, p.35).

Teorema1. La esperanza matemática cumple las siguientes propiedades:

1. La media es el centro de gravedad de la variable:

$$E(X - E(x)) = 0.$$

2. La media es un operador lineal:

$$E(a + bX) = a + bE(X),$$

$$E\left(g_1(X) + g_2(X)\right) = E\left(g_1(X)\right) + E\left(g_2(X)\right).$$

3. Si $a \le g(x) \le b$ para todo x, entonces $a \le E\left(g(X)\right) \le b.$
4. Si X es una variable aleatoria simétrica respecto de *c* (es decir su función de densidad verifica $f(x+c) = f(c-x)$, para cualquier valor de x) entonces: $E(X) = c$, siempre que exista la correspondiente serie o integral.

Definición 17. Los cuartiles con *p*-cuantiles de órdenes 0,25; 0,5 y 0,75. Al cuartil de orden 0,5 se le conoce como mediana.

Definición 18. La moda de una variable aleatoria corresponde al valor de máxima probabilidad.

La moda no tiene porqué existir y puede no ser única. En el caso continuo la moda se obtiene calculando el máximo de la función de densidad. Si X es una variable aleatoria continua con función de densidad $f_X(x)$ derivable, entonces la moda, se puede obtener como la solución de la ecuación $f_X'(x_0) = 0$, y comprobando que $f_X''(x_0) \rangle 0$.

ANEXOS

Anexo 1

CUESTIONARIO INICIAL

1. Diez estudiantes han pesado un objeto en la clase de ciencias, usando el mismo instrumento. Los pesos registrados por cada estudiante (en gramos) son los siguientes:

6,1	6,3	6,0	6,1	6,2	6,5	6,1	6,2	6,1	6,0

Determine un número que represente mejor el peso real del objeto medido. Explique su respuesta.

2. En cierta provincia de un país, que tiene cuatro distritos, se han hecho estudios sobre el número promedio de hijos por familia.

Distrito	**Número de familias**	**Promedio de hijos por familia**
A	35	4,2
B	20	2,6
C	45	5,4
D	30	3,0

¿Cuál es el número promedio de hijos por familia, en toda la provincia?

3. El peso en kilos de 9 niños se presenta en la siguiente tabla:

14	25	17	16	26	17	18	19	24

a) ¿Cuál es la mediana de los pesos?

b) ¿Cuál es la mediana si incluimos el peso de otro niño que pesa 41 kg?
En este caso, ¿sería la media aritmética un buen representante de los 10 datos? Explique su respuesta.

4. Un profesor califica el rendimiento de sus alumnos en el curso de Educación Física del siguiente modo: I=Insuficiente, A=Aprobado, N=Notable, S=Sobresaliente. En la siguiente tabla tenemos las calificaciones que ha puesto a dos grupos de alumnos:

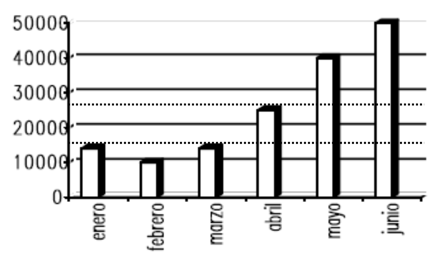
Grupo1	I	A	A	N	N	S	S	I	I	I	A	A	A	N	S	S	I	A	A	S	S	S	S

Grupo 2	S	S	I	I	A	N	A	N	I	I	S	N	A	S	I	N	N

a) ¿Cuál de los grupos ha obtenido mejores calificaciones? Explique su respuesta

b) Si tuviera que emplear un calificativo para el grupo, ¿Cuál sería el más representativo (y apropiado) para el grupo 1? ¿y para el grupo 2? Explique su respuesta.

5. Observe el siguiente diagrama de barras que muestra el número de sandwiches que vendió la empresa Ambrossi durante el primer semestre del presente año:



a) Dé un valor aproximado del *promedio* de sandwiches que la empresa vendió al mes.

b) Dé un valor aproximado de la *moda* del número de sandwiches que la empresa vendió al mes.

c) Dé un valor aproximado de la *mediana* del número de sandwiches que la empresa vendió al mes.

Marque las casillas correspondientes, según sea su caso.

	media (promedio)	**mediana**	**moda**
He estudiado anteriormente el tema, y lo recuerdo			
He estudiado anteriormente el tema, pero no lo recuerdo			
No he estudiado anteriormente el tema			

ANEXO 2.

PRUEBA DE EVALUACIÓN FINAL

1. Diez estudiantes han pesado un objeto en la clase de ciencias, usando el mismo instrumento. Los pesos registrados por cada estudiante (en gramos) son los siguientes:

6,1	6,3	6,0	6,1	6,2	6,5	6,1	6,2	6,1	6,0

Determine un número que represente mejor el peso real del objeto medido. Explique su respuesta.

2. En cierta provincia de un país, que tiene cuatro distritos, se han hecho estudios sobre el número promedio de hijos por familia.

Distrito	**Número de familias**	**Promedio de hijos por familia**
A	35	4,2
B	20	2,6
C	45	5,4
D	30	3,0

a) ¿Cuál es el promedio de hijos por familia, en toda la provincia?

b) En el distrito B hay diez niños menores de 10 años. Los pesos, en kilos, de nueve de ellos están dados en la siguiente tabla:

14	25	17	16	26	17	18	19	24

- ¿Cuál es la mediana de estos datos?
- ¿Cuál es la mediana si incluimos el peso del décimo niño, que es 41 kg?
  En este caso, ¿sería la media aritmética un buen representante de los 10 datos? Explique su respuesta.

3. Un profesor califica el rendimiento de sus alumnos en el curso de Educación Física del siguiente modo**: I=Insuficiente, A=Aprobado, N=Notable, S=Sobresaliente**. En la siguiente tabla tenemos las calificaciones que ha puesto a dos grupos de alumnos:

Grupo1	I	A	A	N	N	S	S	I	I	I	A	A	A	N	S	S	I	A	A	S	S	S	S

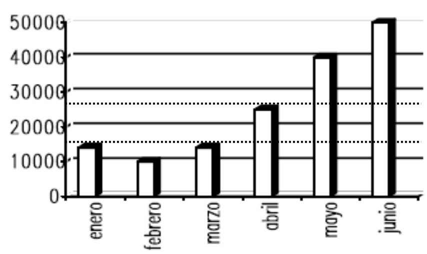
Grupo 2	S	S	I	I	A	N	A	N	I	I	S	N	A	S	I	N	N

c) ¿Cuál es la variable observada y de qué tipo es?

d) ¿Cuál de los grupos ha obtenido mejores calificaciones? Explique su respuesta

e) Si tuviera que emplear un calificativo para el grupo, ¿Cuál sería el más representativo (y apropiado) para el grupo 1? ¿y para el grupo 2? Explique su respuesta.

4. Observe el siguiente diagrama de barras que muestra el número de sandwiches que vendió la empresa Ambrossi durante el primer semestre del presente año:



a) ¿Cuál es la variable observada y de qué tipo es?

b) Dé un valor aproximado del *promedio* de sandwiches que la empresa vendió al mes.

c) Dé un valor aproximado de la *moda* del número de sandwiches que la empresa vendió al mes.
d) Dé un valor aproximado de la *mediana* del número de sandwiches que la empresa vendió al mes.

5.

a) Un informe del INEI en el 2007, dice que *"en la provincia de Lima, el número medio (promedio) de hijos por familia es 1,3"*. Explique qué significa para usted esta expresión.
b) Carlos tomó información sobre el número de hijos en 10 familias de San Borja. Registró la información, pero por un error de digitación se le borraron algunos datos.

Familia A	3
Familia B	2
Familia C	
Familia D	
Familia E	
Familia F	
Familia G	
Familia H	
Familia I	
Familia J	

- ¿Qué datos podrían haber estado en su cuadro, si se sabe que el promedio de hijos por familia es 1,2?
- ¿Existe una única posibilidad de completar el cuadro? Justifique su respuesta.

6. Cierto grupo de niños, el 20 de julio de 2009, tenía como edad promedio 13, 8 años. ¿Es verdad que la edad promedio del mismo grupo de niños, el 20 de julio de 2013 será 18, 2 años? Justifique su respuesta.

7. Se realizó una encuesta a los asistentes al Coloquio de Ciencias Humanas para conocer cuál fue el primer medio por el cual se informaron de esta actividad. Los resultados fueron los siguientes:

Medio de información	Número de personas
Periódico	33
E-mail	55
Página Web	30
TV	32
Radio	40

a) ¿Cuál es la variable observada y de qué tipo es?

b) ¿Cuál de las tres medidas de posición central (media, mediana y moda) tiene sentido hallar en este contexto? ¿Por qué?

ANEXO 3

DIARIO DE OBSERVACIÓN DE CLASE

<table>
<tr><td colspan="3">Universidad: Pontificia Universidad Católica del Perú</td></tr>
<tr><td colspan="2">Unidad: Estudios Generales Letras</td><td>Semestre: 2009-2</td></tr>
<tr><td colspan="2">Curso: Matemáticas</td><td>Profesor(a) a cargo:</td></tr>
<tr><td>Horario:</td><td>N° de alumnos:</td><td>Tema de la clase: Medidas de Tendencia Central</td></tr>
<tr><td>Fecha:</td><td>Duración:</td><td>Observador:</td></tr>
</table>

ESTRUCTURA GENERAL DE LA CLASE:

<table>
<tr><th colspan="4">Apertura de la clase</th></tr>
<tr><th>Aspecto a observar</th><th>SI</th><th>NO</th><th>Observaciones</th></tr>
<tr><td>Explica la importancia del tema, contextualizándolo en situaciones próximas al estudiante.</td><td></td><td></td><td></td></tr>
<tr><td>Recoge saberes previos de los alumnos</td><td></td><td></td><td></td></tr>
</table>

Presenta una situación problema relacionado al concepto matemático "Moda" para que los alumnos intenten resolver			
Presenta una situación problema relacionado al concepto matemático "Mediana" para que los alumnos intenten resolver			
Presenta una situación problema relacionado al concepto matemático "Media" para que los alumnos intenten resolver			

<u>Otras observaciones y/o sugerencias:</u>

....

....

Desarrollo de la clase			
Aspecto a observar	**SI**	**NO**	**Observaciones**
Brinda pautas (ideas) a los alumnos en la búsqueda de solución del problema planteado.			¿Qué ideas brinda el profesor? Relacionado a la moda: Relacionado a la mediana: Relacionado a la media:
Da un tiempo necesario para que los alumnos traten de resolver el problema.			

Recoge las ideas de los alumnos y a partir de ellas arma la explicación.			¿Qué ideas surgen del alumno? Relacionado a la moda: Relacionado a la mediana: Relacionado a la media:

Otras observaciones y/o sugerencias:

....

....

....

Desarrollo de la clase (continuación)			
Aspecto a observar	**SI**	**NO**	**Observaciones**
Da una explicación general, de la solución del problema, identificando claramente los conceptos relacionados(moda , mediana y media)			Respecto a la moda: Respecto a la mediana: Respecto a la media:

Se explica los conceptos de moda, mediana y media, relacionados en la solución del problema, considerando sus aspectos más importantes.			La definición que presenta (para la moda, mediana y media), ¿está basada en el texto? y sino, ¿qué definición da? MODA: , , MEDIANA: , , MEDIA: Según la explicación, ¿en qué casos es pertinente usar cada una de las medidas de posición central? MODA: , , MEDIANA: , , MEDIA:

Desarrollo de la clase (continuación)			
Aspecto a observar	**SI**	**NO**	**Observaciones**

Brinda una espacio para las dudas que puedan surgir de parte de los alumnos, así como también para abordar ideas equivocadas (concepciones erróneas) o cuestionables.			¿Cuáles son las preguntas que formulan los estudiantes respecto a sus dudas? ¿Qué conflictos presentan los estudiantes ante estos conceptos? ¿Cuáles son las respuestas que da el profesor ante las preguntas?

<u>Otras observaciones y sugerencias:</u>

....

……

……

……

………

……

Cierre de la clase			
Aspecto a observar	**SI**	**NO**	**Observaciones**
Se procede a evaluar			

Otras observaciones y sugerencias:

……

……

……

……

………

………

………

………

………

………

Anexo 4

Entrevista semiestructurada al profesor del curso Matemáticas en EE.GG.LL.

1. ¿Por qué considera usted importante que se enseñe el tema de medidas de tendencia central?

2. ¿Qué espera usted que los alumnos comprendan como mínimo al estudiar este tema?

3. ¿Cómo cree usted que se puede conseguir ese objetivo?

4. ¿Considera usted que el libro "Matemática para no matemáticos" contribuye eficazmente a conseguir este objetivo? ¿Por qué?

5. ¿Cuáles son las dificultades más frecuentes que usted ha encontrado en sus alumnos para aprender estos conceptos?

Consideraciones para repreguntas:

a. ¿Los alumnos logran distinguir el uso de moda, mediana y media?
b. ¿Qué experiencias concretas puede narrar respecto a sus esfuerzos para lograr que los alumnos comprendan bien estos conceptos?

ANEXO 5.

5.1 Ejemplos de los elementos que se presentan de forma explícita *(propiedades y algoritmos)* en el libro de texto "Matemáticas para no matemáticos".

Dada la variedad de propiedades y algoritmos, así como los diversos ejemplos encontrados para cada una de ellas, decidimos presentar cada uno de estos ejemplos en una sección aparte. A continuación presentamos a los mismos.

PROPIEDADES

Numéricas

N2 (p.127). Para responder a la pregunta que se formula en el siguiente problema será necesario recordar el hecho que la moda siempre tiene que ser uno de los datos de la distribución. En muchos otros ejemplos se ve claramente que la media y la mediana pueden no coincidir con ninguno de los datos.

1. El ingeniero responsable de la construcción de un edificio debe planear el espacio destinado a estacionamientos para un nuevo complejo habitacional que tendrá 80 departamentos. Se pide que la propuesta se base en el dato estadístico «una medida de tendencia central del número de vehículos por departamento es 0,9». ¿Puede ser 0,9 la mediana de la variable estadística número de vehículos? ¿Puede ser la moda? ¿Por qué?

N3 y N4 (p.130). Se encontró el siguiente párrafo donde se explica claramente estas propiedades.

La media aritmética (a diferencia de la moda y la mediana) involucra a todos los datos y, por lo tanto, es más sensible a cambiar cuando se modifican los datos. Es la medida de tendencia central que posee las mejores propiedades para trabajar situaciones en la estadística inferencial.

Algebraicas

A5 (p.136). Los siguientes ítems corresponden a una situación que pide analizar la veracidad de las afirmaciones. Se espera que el alumno demuestre que son ciertas.

d) Si cada uno de los valores x_1, x_2, x_3, x_4, x_5, y x_6 de una variable estadística es disminuido en una constante c, entonces la nueva media aritmética es igual a la media aritmética de los datos originales, incrementada en la constante *c*.

e) Si a cada valor x_1, x_2, x_3, x_4, x_5, y x_6 de una variable estadística se le multiplica por una constante c, entonces la nueva media aritmética corresponderá a la media aritmética de los datos originales, multiplicada por *c*.

A7 (p.122). El convenio que se establece hace que, la moda puede no existir o, si existe, no ser única (a lo más dos).

Se adoptarán las siguientes convenciones:
- Si hay dos valores o categorías que se repiten igual número de veces y dicha cantidad de veces es la más alta, diremos que hay dos modas y que la distribución de datos es bimodal.
- Si ninguna categoría o valor se repite más veces que los otros; o si hay más de dos que se repiten más veces, admitiremos que no hay moda.

Estadísticas

E1 (p.121). Las tres medidas representan a un colectivo y no a un individuo.

Hablar de medidas de tendencia central significa, para decirlo con pocas palabras, encontrar valores que sean representativos de una muestra. Por ejemplo, cuál es en

E3 (p.136). El siguiente ítem corresponde a la misma situación presentada en *A5*. Hace notar que las tres medidas coinciden cuando los datos son iguales.

a) Si 6 alumnas tienen la misma nota en un examen, entonces la media, la mediana y la moda de estas 6 notas son iguales.

E4. Se presenta explícitamente en el párrafo usado para *N3 y N4.*

E8 (p.124). El siguiente ítem corresponde a la solución de una situación que involucra una variable cualitativa nominal. En ella se pide analizar si es posible hallar la mediana para una variable de ese tipo. Naturalmente, la única medida que tiene sentido hallar es la moda.

d) Para los datos de la situación 14 no se puede calcular el valor de la mediana, porque la variable no es cuantitativa sino cualitativa nominal.

ALGORTIMOS Y PROCEDIMIENTOS

Cálculo de la media

AM1 y AM6 (p.129). Para resolver la siguiente situación se tendrá que hacer uso de dos algoritmos. Primero, calcular la media para datos sin agrupar y luego hacer el uso del algoritmo de la media en sentido inverso.

Situación 18

Jocelyn registra las notas de sus 7 primeras actividades en el curso de Matemáticas:

14; 10; 13; 18; 12; 15; 15

a) Calcule la media aritmética de las notas de las actividades de Jocelyn.

b) Si Jocelyn desea que el promedio aritmético de sus 8 primeras notas de actividades en el curso de Matemáticas sea 14, ¿qué nota deberá obtener en la próxima actividad?

Solución propuesta

a) La media aritmética de las 7 notas está dada por la suma de las 7 notas dividida entre el número de notas.
Es decir, la media aritmética es $\frac{14 + 10 + 13 + 18 + 12 + 15 + 15}{7}$, que equivale a $\frac{97}{7}$, aproximadamente 13,86.

b) Si se llama x a la nota que desea obtener Jocelyn en la octava actividad, entonces la suma de las 8 notas dividida entre 8 debe dar 14. Es decir: $\frac{97 + x}{8} = 14$. Ello será posible si $97 + x = 112$. Luego, necesitará obtener $112 - 97 = 15$ en su siguiente actividad.

AM2 (p.130). En el siguiente cuadro se presenta la fórmula para calcular la media cuando los datos (sin agrupar) son presentados en una tabla de frecuencias.

¿Cómo se halla la media aritmética para datos sin agrupar?

Si se tienen n datos en una distribución de frecuencias simple que consta de k clases.

x_1	f_1
x_1	f_1
x_2	f_2
...	...
x_k	f_k
Total	n

Dado que cada uno de los datos se repite tantas veces como lo indica su frecuencia absoluta y la suma de dichas frecuencias absolutas coincide con el número de datos, entonces si se tienen los datos agrupados en k clases, la fórmula será:

$$\bar{x} = \frac{\sum_{i=1}^{k} (x_i f_i)}{n} = \frac{\sum_{i=1}^{k} (x_i f_i)}{\sum_{i=1}^{k} (f_i)}$$

AM3 (p.131). Se presenta el algoritmo para el cálculo de la media con datos agrupados en intervalos.

¿Cómo se halla la media aritmética para datos agrupados?

Cuando los datos con los que se cuenta han sido dados en intervalos, se requiere elegir un representante de cada intervalo para realizar el cálculo de la media aritmética. Adoptaremos como representante a la marca de clase de cada clase o intervalo.
Así, se empleará la siguiente fórmula:

$$\overline{x} = \frac{\sum_{i=1}^{k} (x_i' f_i)}{n} = \frac{\sum_{i=1}^{k} (x_i' f_i)}{\sum_{i=1}^{k} (fi)}$$

Es importante señalar que el resultado que se obtenga será un valor aproximado a la media aritmética que se obtendría si se consideraran los n datos originales, es decir, si se consideraran los datos sin agrupar.

Observación:

Para ser coherentes con la observación anterior, donde se especifica que el resultado para la media es solo una estimación o aproximación, se debe usar para ello en la primera igualdad el símbolo $\approx$.

Cálculo de la mediana.

AME1, AME2 (p.125). En la solución de la situación que se presenta (situación16), se explica los algoritmos para el cálculo de la mediana para datos sin agrupar. Primero cuando se tiene un número impar de datos y luego cuando se tiene un número par de datos.

Situación 16

a) Halle la mediana de las siguientes notas obtenidas por un estudiante a lo largo de un curso de Matemáticas: 08; 14; 10; 18; 10; 15; 16.

b) Halle la mediana de las siguientes notas: 120; 100; 200; 250; 150; 200.

Solución propuesta

a) Para hallar la mediana de las notas 08; 14; 10; 18; 10; 15; 16, primero es necesario ordenarlas.

Así, ordenándolas en forma creciente, se tiene: 08; 10; 10; 14; 15; 16; 18.

Dado que hay n = 7 datos, entonces el término central ocupará la posición:

$$\frac{7+1}{2} = 4.$$

La mediana será el dato que ocupa la cuarta posición, es decir, $M_e = 4$.

b) Para hallar la mediana de las notas: 120; 100; 200; 250; 150; 200, primero es necesario ordenarlos, obteniéndose 100; 120; 150; 200; 200; 250.

Como hay 6 datos, entonces se consideran los dos términos centrales, es decir, los que ocupan las posiciones 3.ª y 4.ª. La mediana es la semisuma del tercer y cuarto dato. Luego,

$$M_e = \frac{150 + 200}{2} = 175$$

AME3, AME4 (p.125). Se explica el algoritmo para el cálculo de la mediana cuando se tiene un número par de datos sin agrupar presentados en tablas de frecuencias.

Situación 17

A continuación, se muestra información sobre el número de hijos por familia en una muestra de 50 familias de cierto asentamiento humano:

Número de hijos	0	1	2	3	4	5
Cantidad de familias	7	8	10	11	9	5

Halle la mediana de los datos mostrados en la tabla.

Solución propuesta

Para hallar la mediana de los datos de una variable discreta cuya información se presenta agrupada en una tabla, se pueden emplear, como antes, las frecuencias acumuladas.

Clases (x_i)	Frecuencia absoluta (f_i)	Frecuencia acumulada (F_i)
0	7	7
1	8	15
2	10	25
3	11	36
4	9	45
5	5	50
Total	**50**	

Dado que hay 50 datos, la mediana será el promedio aritmético de los valores que ocupan los lugares 25.° y 26. °.
De las frecuencias acumuladas se obtiene que el dato 25.° es 2 hijos y el dato 26.° es 3 hijos. Luego, $M_e = \frac{2+3}{2} = 2,5$ hijos por familia.

Cálculo de la moda

AMO1 (p.122). Moda para datos sin agrupar.

Así, por ejemplo, si se tienen los siguientes datos acerca de la escala de pago de 20 alumnos de Estudios Generales Letras:

3 ; 1 ; 2 ; 1 ; 5 ; 5 ; 4 ; 3 ; 4 ; 5
2 ; 4 ; 3 ; 3 ; 2 ; 5 ; 3 ; 2 ; 4 ; 5

se observa que la escala 3 y la escala 5 ocurren cinco veces cada una y esa es la mayor cantidad de veces que se repite uno de los datos. Es decir, dicho conjunto de datos tiene dos modas: escala 3 y escala 5.

AMO2 (p. 134). Moda para datos sin agrupar presentados en tabla de frecuencias.

6. En la unidad de emergencias de un hospital se registró el motivo de atención de los pacientes atendidos durante una semana, y se obtuvieron los siguientes resultados:

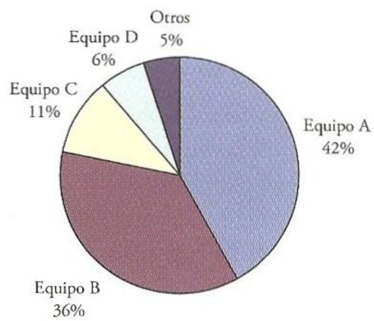
Motivo de la atención	Cantidad de pacientes
Caídas	35
Heridas provocadas por arma blanca	30
Atropellos	45
Cólicos estomacales	55
Picaduras de insectos	10
Intoxicaciones	25
Total	**200**

a) ¿Cuál es la variable observada y de qué tipo es?
b) ¿Con qué gráfico se representarían mejor estos datos? Dibuje usted dicho gráfico.
c) De las tres medidas de tendencia central media, mediana y moda, ¿cuáles tienen sentido para estos datos?

AMO4 (p.127). Cálculo de la moda a partir de un gráfico

.

2. A continuación, se muestran los resultados obtenidos en una encuesta realizada a un grupo de aficionados al fútbol con la finalidad de determinar la popularidad de los equipos que intervienen en el torneo de primera división.



Si se sabe que 200 aficionados respondieron que eran hinchas de otros equipos, responda a las siguientes preguntas:
a) ¿Cuántos aficionados fueron encuestados?
b) ¿Cuántos aficionados respondieron que eran hinchas del equipo B?
c) Construya una distribución de frecuencias para la información presentada.
d) ¿Es posible determinar la moda y la mediana de los datos? De ser posible, determine dichos valores e interprete su significado.