PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

ESCUELA DE POSGRADO

PROPUESTA DIDÁCTICA A PRIORI BASADA EN CRITERIOS DE IDONEIDAD PARA LA ENSEÑANZA DEL USO DE LA MEDIA ARITMÉTICA Y LA MEDIANA EN ESTUDIANTES DE EDUCACIÓN SECUNDARIA

Tesis para optar el grado de Magíster en Enseñanza de las Matemáticas que presenta

WILMER JESÚS OYOLA VILELA

Dirigido por

AUGUSTA ROSA OSORIO GONZALES

San Miguel, 2015

A mis padres Fermín y Rosa, que me dieron la vida.

A mi esposa Flor María que la comparte conmigo.

A mis hijos Luigi, Ariana y Fabiana que son la continuación de ella.

AGRADECIMIENTOS

Al Ministerio de Educación del Perú, quien por medio del Programa Nacional de Becas y Crédito Educativo-PRONABEC, nos permitió acceder a la Beca Presidente de la República denominada "Beca Docente de Posgrado para estudios de Maestría en Ciencias de la Educación en el Perú 2014".

A la escuela de Posgrado en la mención "Maestría en enseñanza de las matemáticas" de la Pontificia Universidad Católica del Perú.

A la Mg. Augusta Rosa Osorio Gonzales, por su acertada asesoría, demostrando profesionalismo en el campo de la estadística y su preocupación constante por mejorarla, motivando a contribuir con investigaciones en este importante campo. Además por sus cualidades humanas que la hacen una excelente persona. Su comprensión y conocimientos puestos en práctica durante la asesoría; sus correcciones para avanzar en el trabajo, fueron pilares para culminar mi investigación.

A los miembros del jurado Dra. Cileda de Queiroz e Silva Coutinho y Dr. Francisco Ugarte Guerra, por el trabajo de leer y hacer las sugerencias a mi tesis, a pesar de la carga académica que demanda su trabajo y por el tiempo que le quitaron a su familia.

A los profesores becarios de la maestría en "Enseñanza de las matemáticas" de la PUCP, provenientes de diferentes regiones del país, que participaron de la muestra piloto a un cuestionario, del mismo modo a los docentes de Tumbes y Arequipa que fueron generosos en participar

Al Mg. Jaime Bravo Febres, por sus sugerencias y apoyo académico puesto a disposición para avanzar en mi tesis, demostrando su compromiso con la educación matemática y estadística en bien de los estudiantes.

Al Lic. Manuel Coveñas Naquiche, por su apoyo con material académico que permitió analizar desde diferentes aspectos la situación de la enseñanza de la estadística.

A los profesores y profesoras de la maestría por sus enseñanzas impregnadas de buen nivel de exigencia académica y por su compromiso con la didáctica de la enseñanza de la matemática.

A mis queridos padres y hermanos, por haberme apoyado incondicionalmente en los momentos que más lo necesité. A ti madre. Gracias.

A mi esposa e hijos, por su comprensión y por motivarme a seguir adelante como profesional. Sin su apoyo no hubiera sido posible lograr mis metas.

A Dios, nuestro creador, que día a día nos da la oportunidad para realizarnos y lograr las metas trazadas.

RESUMEN

En la presente investigación se realiza una propuesta didáctica a priori para la enseñanza del uso de la media aritmética y la mediana en estudiantes (14 - 16 años) de educación secundaria, aplicando algunos aspectos de los criterios de idoneidad didáctica tanto epistémica como cognitiva del Enfoque Ontosemiótico de la Cognición e Instrucción Matemática (EOS), para la elaboración de las configuraciones epistémicas de referencia y de las actividades.

La metodología de la investigación fue de tipo cualitativa y descriptiva, complementándose con información del tipo cuantitativa, al momento de presentar los resúmenes de las respuestas de los docentes a un cuestionario aplicado con el fin de reconocer su dominio acerca del uso de la media aritmética y la mediana en un conjunto de datos, así como del concepto de sesgo que manejan.

Los objetivos específicos fueron recopilar los errores que se cometen respecto al uso de la media aritmética y la mediana, declarados por otros investigadores; identificar preguntas estadísticas comunes referentes al uso y elaborar la secuencia de actividades en base a problemas contextualizados y la aplicación de propiedades estadísticas de la media aritmética y la mediana, que permita a los estudiantes ser capaces de tomar la decisión de elegir el mejor representante del conjunto de datos.

Palabras clave: Análisis exploratorio, Datos atípicos, sesgo, Toma de decisiones.

ABSTRACT

In the present investigation a didactic offer is realized a priori for the education of the use of the arithmetic mean and the median in students (14 - 16 years) of secondary education, applying some aspects of the criteria of didactic suitability so much epistemic as cognitive of the Approach Ontosemiótico of the Cognition and Mathematical Instruction (EOS), for the production of the configurations epistémicas of reference and of the activities.

The methodology of the investigation was of type qualitative and descriptive, complementing itself with quantitative information of the type, to the moment to present the summaries of the answers of the teachers to a questionnaire applied in order to recognize his domain it brings over of the use of the arithmetic mean and the median in a set of information, as well as of the concept of bias that they handle.

The specific aims were to compile the mistakes that are committed with regard to the use of the arithmetic mean and the median, declared by other investigators; to identify statistical common questions relating to the use and to elaborate the sequence of activities on the basis of problems contextualizados and the application of statistical properties of the arithmetic mean and the median, which allows to the students to be capable of taking the decision to choose the best representative of the set of information.

Keywords: Exploratory analysis, atypical Information, bias, Capture of decisions.

Lista de figuras

Lista de tablas

ÍNDICE

INTRODUCCIÓN

La estadística, según Batanero (2001) como "ciencia de los datos", va adquiriendo cada vez más una singular importancia en los currículos del mundo, esto ha motivado que muchos investigadores se preocupen por dar solución a los errores y dificultades que se presentan en su enseñanza. Por ejemplo citamos a profesionales de la Universidad de Granada en España, Universidad de Haifa en Israel y el grupo de la Universidad de Auckland de Nueva Zelanda, quienes han desarrollado reconocidos aportes a la didáctica de la estadística.

El uso de la media aritmética y la mediana en un conjunto de datos, es uno de los contenidos estadísticos incluido recientemente en el currículo del nivel secundario del Perú. La enseñanza de este contenido nos ha motivado a contribuir con una propuesta didáctica que le permita al docente desarrollar sesiones de aprendizaje, movilizando un conjunto de conceptos, procedimientos, propiedades y argumentos, que favorezcan el aprendizaje en nuestros estudiantes. En los Mapas de Progreso de Estadística y Probabilidad, se declara que:

> Los aprendizajes que logran los estudiantes a partir de la estadística, deben adquirir hoy mayor importancia de la que tenían en el pasado, pues se han constituido en herramientas que ayudan al estudiante a organizar y profundizar su conocimiento sobre la realidad que lo circunda; contribuyendo a la toma de decisiones en escenarios de cambio y abundante información. (Perú, 2013, p. 8)

Los resultados de un cuestionario que aplicamos a una muestra de 44 docentes a nivel nacional, evidencian que no manejan estrategias para desarrollar el pensamiento y razonamiento estadístico en los estudiantes, centrándose únicamente en la aplicación de algoritmos de cálculo, influenciados por los libros de texto. Los docentes imparten contenidos estadísticos en los que no todos han tenido o recibido una formación didáctica específica, ni desde su formación ni en el ejercicio de su carrera docente.

Las investigaciones sobre las medidas de tendencia central son pocas, de las cuales hemos considerado algunas que indirectamente se refieren al uso de las medidas de tendencia central, por mencionar la investigación de Mayén (2009).

La propuesta didáctica a presentar, proporciona un conjunto de actividades para ser desarrolladas por el docente con el objetivo que nuestros estudiantes estén en la capacidad de analizar un conjunto de datos y para reconocer la existencia o no de datos u observaciones atípicas que generan sesgo, y, a partir de este análisis, tomar la decisión de elegir entre la media aritmética y la mediana al mejor representante del conjunto de datos. En nuestro trabajo también se presentan las herramientas necesarias para el análisis de la futura implementación de la propuesta didáctica, como son la configuración epistémica de referencia del uso de la media aritmética y la mediana, así como de la determinación de datos u observaciones atípicas que generan sesgo y los criterios de idoneidad didáctica tanto epistémica como cognitiva, tomadas del Enfoque Ontosemiótico de la Cognición e Instrucción Matemática (EOS).

Hemos estructurado el trabajo de investigación de la siguiente manera. En el capítulo uno se expone la problemática, los antecedentes, la justificación, la pregunta de investigación y los objetivos; en el capítulo dos se explican detalladamente los marcos de trabajo: epistémico, didáctico y teórico, así como la metodología; en el capítulo tres se diseña la propuesta didáctica, atendiendo a sus características, objetivos y descripción de las actividades asociadas a sus configuraciones epistémicas; y finalmente en el capítulo cuatro se describen las conclusiones y recomendaciones.

Esta propuesta didáctica a priori, servirá de referencia para que futuros investigadores se interesen en aplicarla.

CAPÍTULO I: EL PROBLEMA

Los medios de comunicación social muestran información estadística como una forma de representar las diferentes actividades que realiza el hombre en los distintos ámbitos en que se desenvuelve, en consecuencia, es necesario que los ciudadanos estemos preparados para comprender, analizar e interpretar los datos de esta información estadística. Un aspecto importante relacionado a la información estadística recibida por estos medios de comunicación social, son los promedios. Nuestra investigación desea proponer una secuencia de actividades que permita a los docentes de matemática enseñar a sus estudiantes que en algunas distribuciones, hay una medida estadística más representativa del conjunto de datos (la media aritmética o la mediana). Estas medidas se complementan entren sí y sirven para realizar el análisis exploratorio de datos y el desarrollo del pensamiento estadístico.

Hemos encontrado pocas investigaciones relacionadas a la enseñanza del uso de las medidas de tendencia central, entre las encontradas tenemos por ejemplo las realizadas por Mayén (2009), Cobo y Batanero (2000), Garrett y García (2005) y Sayritupac (2013). En estos trabajos los investigadores destacan qué medida de tendencia central usar en una situación o problema determinado, así como el análisis del concepto, su comprensión, y las dificultades y errores que presentan los estudiantes al momento de elegir un representante para un conjunto de datos, aplicando propiedades estadísticas y justificando su decisión.

El presente capítulo recoge los aspectos más importantes de estas investigaciones, destacando la importancia del tema tanto en el área de la gestión de datos así como en situaciones de la vida diaria.

También se analizan los conocimientos de los docentes de matemática en ejercicio respecto al uso de las medidas de tendencia central y sesgo estadístico. Se observa que en Perú (2010), el Diseño Curricular Básico Nacional de formación docente (DCBN), no incluye la enseñanza del uso de las medidas de tendencia central. Se revisa el tratamiento que realizan los libros de texto respecto a las medidas de tendencia central y como focalizan la aplicación de algoritmos de cálculo. Finalmente se plantean los objetivos de investigación.

1.1. ANTECEDENTES

Actualmente la educación estadística forma parte de los sistemas educativos a nivel mundial, la inclusión de la enseñanza del uso de las medidas de tendencia central, en las instituciones educativas del nivel secundaria, ha motivado que los investigadores muestren interés por resolver la problemática que tienen los estudiantes y docentes respecto a la asimilación, comprensión y uso de estas medidas. En primer lugar es importante considerar la comprensión de los conceptos, luego el reconocimiento y aplicación de propiedades estadísticas de cada medida, y finalmente evaluar la representatividad de la media aritmética y la mediana teniendo en cuenta ciertos criterios de análisis en el conjunto de datos.

En su tesis doctoral Mayén (2009), con 518 estudiantes mexicanos de educación secundaria y bachillerato, señala que uno de los conflictos semióticos relacionado a los conceptos de la media aritmética y la mediana, es no saber determinar con precisión qué medida usar cuando el conjunto de datos presenta sesgo, considerando lo pretendido en el currículo y los textos que utilizan los estudiantes y profesores. Este trabajo contribuyó a validar el cuestionario aplicado por Cobo (2003) a estudiantes españoles, verificando errores similares y el registro de errores nuevos. La investigadora al referirse a la elección de la medida de tendencia central más representativa de un conjunto de datos, considera que se pretende evaluar si los estudiantes son capaces de elegir un buen representante estadístico para una distribución en la que aparece un valor atípico, recomendando usar la mediana cuando hay sesgo en los datos.

En la construcción de nuestra propuesta, tendremos en cuenta los siguientes aspectos tomados del trabajo de Mayén: i) los errores cometidos por los estudiantes al determinar el uso de alguna de estas medidas de tendencia central, con el propósito de evitarlos, ii) la presencia de datos u observaciones atípicas que sesgan la información y iii) la aplicación de propiedades estadísticas de la mediana, que justifican su elección para considerarla como el mejor representante del conjunto de datos. Del mismo modo adaptaremos uno de los problemas que aplicó durante su investigación, relacionado con la determinación del representante de un conjunto de datos, para aplicarlo en un cuestionario a docentes en ejercicio, que permitirá realizar un análisis de los conocimientos que tienen los docentes sobre el tema en estudio.

En el artículo “La mediana en la educación secundaria obligatoria: ¿Un concepto sencillo?”, se analiza la mediana desde el punto de vista conceptual y didáctico, y se incluyen definiciones, métodos de cálculo y propiedades.

> En primer lugar los estudiantes deben comprender que, como cualquier resumen estadístico, la mediana se refiere a todo el conjunto de datos, y no a ninguno de los individuos particulares. Comprender esto requiere un cambio de perspectiva (pasar a la perspectiva estadística), que consiste en atender a las características de los agregados y no a las de los individuos (Cobo y Batanero 2000, p.2)

Creemos que es importante realizar los cálculos de la media aritmética y la mediana incluyendo a todos los datos de la muestra. De acuerdo con Sweeney (2008), el eliminar datos, creyendo que son erróneos, puede ser un error si antes no verificamos si son reales o no.

En nuestra propuesta didáctica, usaremos tanto estrategias como técnicas. La aplicación de las propiedades estadísticas, por parte de los alumnos representará el uso de estrategias, mientras que la aplicación de algoritmos de cálculo de las medidas corresponde al uso de técnicas.

Cobo y Batanero (2004), señalan las dificultades que tienen los estudiantes de Educación Secundaria con respecto a las medidas de tendencia central. Las investigadoras “encontraron errores de cálculo y aplicación incorrecta de otras propiedades. (p. 86)

Los investigadores Garrett y García (2005) aplicaron un cuestionario que adaptaron de Garfield (2003) a estudiantes españoles y angoleños sobre los promedios. En él analizan algunas estrategias presentadas por los estudiantes en la resolución de un problema.

Uno de los problemas de este cuestionario será adaptado y tomado en cuenta para aplicarlo a los docentes que forman parte de la muestra que servirá para analizar los conocimientos y dificultades de los docentes de matemática en ejercicio, relacionados al uso de la media aritmética y la mediana de un conjunto de datos.

Es importante señalar el análisis de los significados personales e institucionales de las medidas de tendencia central que se encuentran en los textos:

> Los textos restringen los significados de referencia de estas medidas al considerarlas como medidas de resumen y utilizar solo los aspectos algorítmicos y de cálculo, sin tener en cuenta su comprensión conceptual que están presentes en los significados personales e institucionales. Indica también que el estudio de estas medidas no solo implica conocer los algoritmos de cálculo, sino que aparte de comprender los conceptos y sus características, se hace también necesario determinar su uso pertinente frente a ciertas situaciones y su interpretación en contextos específicos. (Sayritupac 2013, p.II)

Por otro lado, la representatividad de la media aritmética y la mediana, según Campbell (1981), citado por Sayritupac (2013), guarda la siguiente relación: La media aritmética es el mejor representante de un conjunto de datos si la distribución es simétrica (media aritmética igual a la mediana), y en otros casos el mejor representante de los datos es la mediana, si la distribución es asimétrica (media aritmética diferente a la mediana) por la influencia de unos cuantos valores grandes o pequeños que sesgan al conjunto de datos.

El concepto de sesgo será explicado en el marco epistémico.

En nuestra investigación tendremos en cuenta estas consideraciones de la media aritmética y la mediana. Del mismo modo realizaremos el análisis de los datos, para determinar el mejor representante en situaciones problemáticas reales, de contexto cercanos al estudiante, que aparecen en los textos, revistas, periódicos, artículos y otros medios, referentes al uso de la media aritmética y la mediana.
La propuesta didáctica, es una secuencia de actividades diseñada considerando criterios de idoneidad tanto epistémicos como cognitivos, y busca que el docente de matemática en ejercicio, realice actividades de la enseñanza del uso de la media aritmética y la mediana, creadas teniendo en cuenta los errores de aprendizaje de los alumnos.

1.2. JUSTIFICACIÓN

De acuerdo con Gal (2002), citado por Mayén (2009), “la educación estadística tiene una demanda cada vez más urgente en nuestras sociedades modernas” (p.24). En nuestro país los contenidos estadísticos se han ido incrementando considerablemente en los diferentes diseños curriculares que se han aplicado hasta la actualidad. Uno de los contenidos incorporados al currículo peruano, es el uso la media aritmética y la mediana en un conjunto de datos.

Según Perú (2009), en el Diseño Curricular Nacional (DCN), la enseñanza de la estadística se programa desde el nivel primario y considera como objetivos a lograr por los estudiantes de los últimos grados de este nivel, que formulen y resuelvan problemas que requiera de las medidas de tendencia central; de la misma manera considera que los estudiantes del nivel secundario desarrollen capacidades de resolución de problemas que involucren el cálculo de la media aritmética, la mediana y la moda en datos numéricos, establecer relaciones entre ellas, interpretar la asimetría de las medidas de tendencia central y resolver problemas que involucran el cálculo de estas medidas. Las modificaciones realizadas al DCN especialmente las declaradas en Perú (2013), mapas de progreso correspondiente a estadística y probabilidad, incluida en el área de matemática del nivel secundaria, considera como uno de los estándares nacionales de aprendizaje de los estudiantes del VI ciclo (1° y 2° de secundaria de educación básica regular, cuyas edades están entre 13, 14 y 15 años), "la interpretación y uso de las medidas de tendencia central, reconociendo la medida representativa de un conjunto de datos", y en los estudiantes del VII ciclo (3°, 4° y 5° de secundaria de educación básica regular 15, 16 y 17 años), "la interpretación del sesgo en la distribución obtenida de un conjunto de datos". (p. 9)

Veamos la siguiente tabla.

Tabla 1. Descripción de los niveles del Mapa de Progreso de Estadística y Probabilidad

V CICLO (5° y 6° de Primaria) 10-11 años	Recopila datos cualitativos o cuantitativos discretos provenientes de su entorno escolar, mediante una encuesta en las que formulan preguntas y sus posibles opciones de respuestas; selecciona e interpreta datos provenientes de fuentes indirectas, los organiza en tablas y los representa mediante gráficos de barras dobles o gráficos de líneas. Interpreta información no explícita presentada en tablas, gráficos de líneas y gráficos circulares. Interpreta y determina la media aritmética de un grupo de datos. Determina y representa todos los posibles resultados de una situación aleatoria propuesta usando distintas estrategias. Interpreta la probabilidad de un evento como el cociente entre el número de casos favorables y el total de casos posibles, la representa mediante una fracción y la explica.
VI (1° y 2° de Secundaria) 12-13 años	Recopila datos cuantitativos discretos y continuos o cualitativos ordinales y nominales provenientes de su comunidad mediante encuestas, determina la población pertinente al tema de estudio. Organiza datos provenientes de variables estadísticas y los representa mediante histogramas y polígonos de frecuencia. Infiere información de diversas fuentes presentada en tablas y gráficos, la comunica utilizando un lenguaje informal. **Interpreta y usa las medidas de tendencia central reconociendo la medida representativa de un conjunto de datos**. Interpreta el rango o recorrido como una medida de dispersión. Identifica sucesos simples o compuestos relacionados a una situación aleatoria propuesta y los representa por extensión o por

	comprensión. Determina la probabilidad a partir de la frecuencia de un suceso en una situación aleatoria.
VII (3°, 4° y 5° de Secundaria) 14 a 17 años	Recopila de forma directa e indirecta datos referidos a variables cualitativas o cuantitativas involucradas en una investigación, los organiza, representa, y describe en tablas y gráficos pertinentes al tipo de variables estadísticas. Determina la muestra representativa de una población usando criterios de pertinencia y proporcionalidad. Interpreta el sesgo en la distribución obtenida de un conjunto de datos. Infiere información del análisis de tablas y gráficos, y lo argumenta. Interpreta y determina medidas de localización y desviación estándar para representar las características de un conjunto de datos. Formula una situación aleatoria considerando el contexto, las condiciones y restricciones para la determinación de su espacio muestral y de sus sucesos.

Fuente: Perú (2013, p. 9)

Un aspecto muy importante a considerar en el logro de los aprendizajes de los estudiantes y su desempeño, es el rol de los docentes en el aula. Según Behar (2001) “cada vez se aprecia más la preocupación entre los profesores por mejorar la eficacia de sus tareas, seguramente debido a la sospecha de que las formas tradicionales no están dando resultado”, y además como señala Moreno (2006) “muchas veces se ve influenciado por la muy poca, o en algunos casos ninguna capacitación en procesos de enseñanza y aprendizaje o didácticas específicas”, tal es el caso del aprendizaje de la estadística, que presenta muchas limitaciones de contenidos en el plan de estudios del Diseño Curricular Básico Nacional (DCBN) para la carrera profesional de profesor de educación secundaria en la especialidad de matemática, que no toma en cuenta algunos contenidos estadísticos muy importantes que le servirán al docente para el efectivo ejercicio de su labor en el aula. Nos damos cuenta que el plan de estudios de los futuros profesores más están orientados a capacitarlos con herramientas estadísticas que le faciliten su trabajo de investigación, como lo podemos ver en la siguiente tabla.

Tabla 2. Contenidos del curso de estadística en la formación de profesores **(VII ciclo)**

ÁREA: ESTADÍSTICA
Permite desarrollar el pensamiento lógico matemático y la formación profesional del estudiante mediante el uso adecuado de técnicas de ordenamiento, representación gráfica, sistematización, análisis e interpretación de datos estadísticos relativos a una o más variables, sobre fenómenos y situaciones sociales, educativas, comunales, etc. que permitan la toma de decisiones de manera crítica y reflexiva.
• Relación entre la distribución binomial y normal. • Esperanza. • Estimación de parámetros. Test de hipótesis. Intervalos de confianza. • Distribución normal y normal estándar. • Prueba de hipótesis y aplicaciones. • Prueba de T de students. Aplicaciones. • Prueba Chi – cuadrado. Aplicaciones. • Regresión y correlación lineal. Aplicaciones. • Distribución normal multivariante. • Error muestral. • Muestra.

Fuente: Perú (2010, p. 58)

Es importante señalar la situación de la formación estadística en los docentes:

> La situación de la formación estadística de los profesores de matemática no es, en general satisfactoria y sigue sin ser coherente con el trabajo que van a desarrollar en las aulas ni adaptada a las demandas de la sociedad actual. Estas lagunas formativas pueden provocar a la larga un bajo interés en la materia y conducir a un efecto negativo sobre el proceso de enseñanza-aprendizaje. (Estrada, 2007, p. 3)

Una estrategia para mejorar la enseñanza de la estadística es que:

> La enseñanza de la estadística debería empezar con problemas reales mediante los cuales los estudiantes puedan desarrollar sus ideas, trabajando las diferentes etapas en la resolución de un problema real (planificar la solución, recoger y analizar los datos, comprobar las hipótesis iníciales y tomar una decisión en consecuencia). (Batanero, 2001, p. 115)

Otro factor que afecta el desarrollo de la estadística es el tiempo, pues, los docentes planifican el desarrollo de los contenidos estadísticos para el final de su programación curricular, probablemente porque en el DCN y en los textos de matemática, los temas de estadística forman parte de los últimos capítulos, con el consiguiente riesgo de no llevarse a cabo dichas clases por actividades internas como (aniversario de la IE) o actividades externas como (aniversario de la provincia o distrito, feriados largos declarados por el gobierno, huelgas o paros convocados por los gremios sindicales a los que pertenecen los docentes y fenómenos naturales, entre otros). Por otro lado el tratamiento superficial de la media aritmética y la mediana, debido a que sólo se considera el algoritmo de cálculo, es otro factor que perjudica su enseñanza.

Los libros de texto, tanto los que ofrece el Ministerio de Educación, como los que ofrecen las editoriales, es otroo de los factores relacionado a esta problemática. De lo observado, estos presentan errores, pues solo se preocupan por el algoritmo de cálculo y los pocos ejemplos evidencian que no profundizan el desarrollo del tema (la media aritmética y la mediana), perjudicando el buen desarrollo de estas medidas y otros temas de estadística dentro del aula (Ver anexo). Todo ello hace, según Sayritupac (2013), que los estudiantes que terminan la educación secundaria y continúan estudios en la universidad, tengan problemas de definición, comprensión, representación y poco interés por aprender estos conceptos estadísticos.

De la información estadística que recibimos ya sea por los medios de comunicación, en los textos, artículos u otros, existen situaciones comunes relacionadas al uso de la media aritmética y la mediana. Frente a ello, los ciudadanos en la mayoría de casos, aceptamos como una regla general que el promedio de los datos está representado por la media aritmética, sin reflexionar o cuestionar los argumentos de tal elección. En nuestra propuesta didáctica identificaremos estos tipos de problemas comunes, los usaremos en la construcción de la secuencia de actividades y los resolveremos analizando la existencia de datos u observaciones atípicas y aplicando propiedades estadísticas de la media aritmética y la mediana.

Como parte del recojo de información, hemos aplicado un cuestionario individual a los docentes del área de matemática, el mismo que nos permitió verificar si en el ejercicio de su carrera docente se han capacitado y adquirido un significado personal acerca del sesgo estadístico y el uso de la media aritmética y la mediana en un conjunto de datos. El cuestionario estuvo orientado a la evaluación del significado personal que tienen los docentes del área de matemática en ejercicio, respecto a la representatividad de la media aritmética y la mediana en un conjunto de datos estadísticos. El cuestionario está compuesto de tres preguntas, la primera pregunta corresponde a una adaptación del cuestionario que aplicó Mayen (2009) con estudiantes mexicanos, relacionada a la representatividad de la media aritmética y la mediana con datos expresados en números decimales. La segunda pregunta corresponde al cuestionario adaptado que aplicaron Garrett y García (2005) con estudiantes españoles y angoleños y la tercera pregunta del cuestionario está relacionada al significado personal del concepto de sesgo estadístico en los docentes.

El cuestionario es individual, pues nos interesa conocer las concepciones que tienen los docentes de la especialidad de matemática en ejercicio sobre la determinación de datos u observaciones atípicas, propiedades estadísticas de la media aritmética y la mediana, representatividad y el sesgo estadístico de un conjunto de datos. Las preguntas son de respuesta abierta lo que permitirá obtener diferentes respuestas de los docentes a las cuestiones planteadas.

De acuerdo con Mayén (2009), "el cuestionario es un instrumento de evaluación que tiene como finalidad proporcionar información sobre los *significados personales* de un grupo". (p. 141)

Los docentes ante dos situaciones problemáticas, que presentan datos u observaciones atípicas que sesgan la información y en virtud de ello se elige a la mediana como representante de los datos por ser resistente, el 55% de los docentes muestrales considera usar la media aritmética, argumentando su decisión por las siguientes razones: i) la media aritmética representa al promedio, ii) por el algoritmo de cálculo, iii) la media es más cercana a los datos, iv) la media es exacta y rápida, v) la media no considera datos atípicos, vi) la media es representante en variables cualitativas, vii) la media considera datos atípicos. Ninguno de los argumentos señalados es válido para elegir a la media aritmética como representante de los datos.

El 41% de los docentes recomienda usar la mediana por las siguientes razones: i) por el sesgo, ii) por el dato atípico, iii) por que el valor obtenido de la mediana es un numero entero y el valor de la media aritmética es un número decimal, iv) porque es más representativa, v) porque se usa para sacar promedios, vi) porque es un valor exacto, vii) porque es más factible, viii) porque se puede demostrar, ix) porque resulta rápido ordenar los datos y x) porque hay fallas en los datos. Aunque la respuesta es correcta y los dos primeros argumentos son válidos, es preocupante que los docentes justifiquen otras razones que son inadecuadas y no corresponden a propiedades numéricas, ni estadísticas ni algebraicas de la mediana.

La tercera pregunta del cuestionario estuvo orientada al significado personal del concepto de sesgo estadístico que manejan los docentes, a partir de una situación que muestra datos que están “alejados” del resto. Las respuestas de los docentes es como sigue: i) el 27% considera el sesgo como datos alejados o sesgados, ii) el 2% considera que sesgo es la presencia de datos atípicos, iii) el 4% son datos que están más cerca, iv) el 16% cree que se debe a datos que no cumplen con el parámetro normal, v) el 7% considera que es una diferencia entre la esperanza matemática y los parámetros del desarrollo, vi) para el 23% de los docentes , el sesgo son datos con error, vii) para el 5% es una variación de las frecuencias, viii) otro 5% piensa que es la asimetría de las medidas de tendencia central y ix) el 11% de docentes no sabe o no responde. El análisis de los resultados de las respuestas de los 44 docentes que conforman la muestra se encuentra en el numeral 4 del anexo 8.

Podemos concluir con esta información que la mayoría de los docentes en ejercicio no reconocen la necesidad del uso de la mediana, ni cuándo se debe tomar esta medida en cuenta como representante de un conjunto de datos, y este resultado es importante porque indica que nuestra propuesta didáctica servirá como una herramienta de apoyo para los docentes en la enseñanza de este tema.

La secuencia de actividades de nuestra propuesta, tendrá como objetivo facilitar a los docentes de matemática en ejercicio un conjunto de estrategias metodológicas que le permitan desarrollar sesiones de aprendizaje orientadas a elegir a la media aritmética o la mediana como el mejor representante de un conjunto de datos, para ello utiliza ciertos criterios de análisis para determinar datos u observaciones atípicas y aplica propiedades estadísticas, que le permitan a los estudiantes argumentar la decisión tomada con fundamentos válidos y reconociendo que no siempre la media aritmética es el representante de los datos.

Esta propuesta didáctica es a priori porque está basada en consideraciones previas, que tiene en cuenta todos los aspectos de una sesión de aprendizaje ideal antes de su realización.

En conclusión , las demandas actuales de la educación estadística, la inclusión en el currículo la enseñanza del uso de la media aritmética y la mediana, la preparación limitada de los docentes en temas de estadística, influenciada por los planes de estudio de los centros de formación docente, la programación de los contenidos estadísticos al final del año escolar, los libros de texto inadecuados y los resultados obtenidos sobre sesgo que manejan los docentes en ejercicio, son consideraciones importantes para preocuparnos en construir una propuesta didáctica que permita mejorar la enseñanza y aprendizaje del uso de la media aritmética y la mediana como el mejor representante de un conjunto de datos.

1.3. PREGUNTA DE INVESTIGACIÓN

¿Qué características debe tener una propuesta didáctica a priori para la enseñanza del uso de la media aritmética y la mediana, considerando la idoneidad didáctica del Enfoque Ontosemiótico de la cognición e instrucción matemática (EOS)?.

1.4. OBJETIVOS

1.4.1. OBJETIVO GENERAL

Caracterizar una propuesta didáctica a priori para la enseñanza del uso de la media aritmética y la mediana, teniendo en cuenta los aspectos de idoneidad del Enfoque Ontosemiótico de la Cognición e Instrucción Matemática (EOS).

1.4.2. OBJETIVOS ESPECÍFICOS

1. Recopilar los errores, declarados por otros investigadores, que cometen los estudiantes, al usar la media aritmética y la mediana, para considerarlos durante la elaboración de la propuesta.
2. Identificar las preguntas estadísticas comunes que aparecen en los textos, artículos y otros medios referentes al uso de la media aritmética y la mediana.
3. Elaborar la secuencia de actividades.

CAPÍTULO II: MARCOS DE TRABAJO

En este capítulo nos referiremos a los tres marcos de trabajo que guían nuestra investigación, el ***marco epistémico,*** que describe los conceptos, algoritmos de cálculo y propiedades más importantes de la media aritmética y la mediana, el significado de sesgo estadístico en un conjunto de datos, el algoritmo de cálculo de los cuartiles y su uso en la construcción del diagrama de caja o bigotes, también lo llamaremos intervalo para determinar datos u observaciones atípicas. El ***marco didáctico***, relacionado con la situación actual de la enseñanza de la estadística especialmente sobre el uso de la media aritmética y la mediana por parte de los docentes en el aula, las reglas, las consideraciones a tener en cuenta, las ventajas y desventajas de cada medida estadística. Y por último el ***marco teórico***, que describe la teoría que respalda nuestro trabajo, en la línea de investigación de la didáctica para la enseñanza de las matemáticas, consideramos algunos aspectos del Enfoque Ontosemiótico de la Cognición e Instrucción Matemática (EOS), porque es el más pertinente a nuestra propuesta al proporcionarnos dos herramientas teóricas fundamentales como son las configuraciones epistémicas y los criterios de idoneidad de las prácticas matemáticas.

2.1. MARCO EPISTÉMICO

En esta sección nos referiremos a la media aritmética y la mediana en su aspecto conceptual, procedimental, propiedades estadísticas más importantes, así como las ventajas y desventajas de cada una de ellas. Este análisis permitirá comprenderlas tanto cuando se presentan en su forma simple y en datos agrupados. También nos referiremos al sesgo, los cuartiles y el procedimiento para construir el diagrama de caja y bigotes o un intervalo que permite reconocer datos u observaciones atípicas.

2.1.1. La media aritmética.

La definición según Sweeney (2008):

> Es la medida de localización más importante, llamada también media o promedio de una variable, que representa al conjunto de datos de una distribución. Se calcula sumando todos los valores observados de la variable, dividido por el número de observaciones. Es una medida de representación central que cumple tres requisitos fundamentales:
>
> - Para su obtención se utilizan todas las observaciones.
> - Es un valor comprendido entre el menor y el mayor de los valores de la distribución.
> - Viene expresada en la misma unidad que los datos. (p. 83)

La definición según Novaes (2011):

> Es un valor alrededor del cual los demás valores se distribuyen (o concentran), o sea, es un valor de referencia para el conjunto analizado. Una de las posibles interpretaciones para el valor medio de un conjunto de datos sería como punto de equilibrio para los valores de la distribución. (p. 90)

Se denota por $(\bar{x})$.

Algunas características para el análisis de la idoneidad del uso, ventajas y desventajas de la media aritmética.

Tabla 3. Características para el análisis de la idoneidad del uso de la media aritmética.

USAR MEDIA ARITMÉTICA	En datos numéricos y distribuciones simétricas, es decir sin ningún tipo de sesgo, en este caso la media es igual que la mediana o muy cercanos.

Fuente: Pinzón (2012, p.p. 19-20)

Tabla 4. Ventajas y desventajas de la media aritmética.

	VENTAJAS	DESVENTAJAS
LA MEDIA ARITMÉTICA	Su concepto es familiar a la mayoría de la gente e intuitivamente claro. Cada conjunto de datos tiene una y solo una media. Cada observación en el conjunto de datos es tomada en cuenta cuando se calcula la media. Es útil para realizar procedimientos estadísticos como comparar las medias de varios conjuntos de datos.	Su sensibilidad es un problema, ya que puede ser fácilmente distorsionada por uno o algunos valores aislados y no representativos del conjunto de datos. Los valores aislados o atípicos, sesgan la información y por tanto distorsionan la media aritmética.

Fuente: Pinzón (2012, p.25)

2.1.2. Propiedades más importantes de la media aritmética

1) Si un conjunto de datos está formado por la repetición de un mismo dato, la media aritmética es ese dato constante.
2) La media aritmética está comprendida entre el valor máximo y el valor mínimo del conjunto de datos.

3) La media aritmética no es el mejor representante cuando hay datos extremos que toman valores muy altos o muy bajos.
4) Si en un conjunto de datos, se cambia uno o más datos, el valor de la media aritmética también cambia, es decir la media aritmética es muy sensible a la variación de los datos.

Por ejemplo, la media aritmética de los grupos:

a) 55, 56, 57, 58, 59, 60 es igual a $\bar{x}_1$ = 345/6 = 57,5

b) 55, 56, 57, 58, 59, 100 es igual a $\bar{x}_2$ = 385/6 = 64,2

c) 55, 56, 57, 58, 59, 0 es igual a $\bar{x}_3$ = 285/6 = 47,5

Como se puede observar la media aritmética cambió, al cambiar el dato 60 por el valor: 100 en el grupo b) y 0 en el grupo c).

2.1.3. Cálculo de la media aritmética de datos no agrupados o no tabulados

Según Córdova (2003):

> La media de n valores x_1, x_2,..., x_n, de la variable cuantitativa x, observados en una muestra es el número:

$$\boxed{\bar{x} = \frac{\sum x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}}$$

Figura 1. Cálculo de la media aritmética en datos no agrupados.

Fuente: Córdova (2003, p.44)

Donde:

$\bar{x}$, es la representación simbólica de media aritmética o promedio aritmético.

Σ, es la letra griega sigma que se lee suma.

x_i, representa a todos los datos.

n, es el total de datos.

$x_1 + x_2 + \cdots + x_n$, es la suma de todos los datos.

Ejemplo:

Determinar el promedio de las calificaciones de un estudiante en el área de matemáticas:

14, 18, 10, 16, 08, 20, 12, 05.

Solución:

$$\bar{x} = \frac{14 + 18 + 10 + 16 + 8 + 20 + 12 + 5}{8} = \frac{103}{8} = 12{,}875$$

El promedio de notas redondeado al entero es 13.

2.1.4. Media aritmética para datos agrupados o tabulados

De acuerdo con Córdova (2003) se presentan dos casos:

a) Media para datos agrupados de variable discreta.

Si *n* valores de una variable estadística discreta *x* se clasifican en *k* valores distintos x_1, x_2,..., x_k con frecuencias absolutas respectivas f_1,f_2,...,f_k, entonces su media aritmética es el número:

$$\boxed{\bar{x} = \frac{\sum_{i=1}^{k} f_i . x_{i.}}{n}}$$

Figura 2. Cálculo de la media aritmética en variables discretas.

Fuente: Córdova (2003, p.44)

Ejemplo: Determinar la media en la siguiente tabla.

Tabla 5. Cálculo de la media del número de hijos por familia.

Valores de x	frecuencias
x_i	f_i
0	1
1	4
2	7
3	6
4	2
Total	20

Fuente: Córdova (2003, p.45)

Solución:

Agregamos una columna de productos $x_i.f_i$

Tabla 6. Cálculo de la media del número de hijos por familia.

Valores de x	frecuencias	Productos
x_i	f_i	$x_i.\ f_i$
0	1	0
1	4	4
2	7	14
3	6	18
4	2	8
Total	20	44

Fuente: Córdova (2003, p.45)

Se obtiene: $\bar{x} = \frac{\sum_{i=1}^{5} x_{i.} f_{i.}}{n} = \frac{44}{20} = 2{,}2$

De acuerdo a los datos de la tabla de frecuencias absolutas y redondeando el resultado obtenido, el número promedio de hijos por familia es 2.

b) Media para datos agrupados o tabulados por intervalos

Si *n* valores de alguna variable *x* están tabulados en una distribución de frecuencias de *k* intervalos, donde**:**

$m_1, m_2, \ldots, m_k$ son las marcas de clase, y

$f_1, f_2, \ldots, f_k$ son las frecuencias absolutas respectivas, entonces, su media aritmética es el número:

$$\boxed{\bar{x} = \frac{\sum_{i=1}^{k} f_{i.} m_{i.}}{n}}$$

Figura 3. Cálculo de la media aritmética en datos agrupados.

Fuente: Córdova (2003, p.45)

Ejemplo:

Tabla 7. Cálculo de la media aritmética de los ingresos quincenales de una muestra de 45 personas, en dólares.

Ingresos I_i	N° de personas f_i
[260, 340[	1
[340, 420[	2
[420, 500[	4
[500, 580[	10
[580, 660[	16
[660, 740[	8
[740, 820[	3
[820, 900]	1
Total	45

Fuente: Córdova (2003, p.46)

Solución:

Determinamos las marcas de clase (m_i) y agregamos dos columnas en la tabla, una para las marcas de clase y otra para los productos de la marca de clase por la frecuencia ($f_i.m_i$).

Tabla 8. Cálculo de la media aritmética de los ingresos quincenales de una muestra de 45 personas, en dólares.

Ingresos I_i	Marcas m_i	N° de personas f_i	Productos $f_i.m_i$
[260, 340[	300	1	300
[340, 420[	380	2	760
[420, 500[	460	4	1840
[500, 580[	540	10	5400
[580, 660[	620	16	9920
[660, 740[	700	8	5600
[740, 820[	780	3	2340
[820, 900]	860	1	860
Total		45	27020

Fuente: Córdova (2003, p.46)

Resultando, $\bar{x} = \frac{\sum_{i=1}^{8} f_i.m_i}{n} = \frac{27020}{45} = \$600{,}44$

La media aritmética puede ser calculada también en distribución de frecuencias por intervalos de amplitud diferentes, siempre que puedan determinarse los puntos medios (marcas de clase) de los intervalos, sin embargo no puede ser calculada a partir de una distribución de frecuencias con intervalos abiertos, como se muestra en la siguiente tabla.

Tabla 9. Frecuencias con intervalos de diferente amplitud.

Intervalos	Frecuencias
Menor o igual que 20	50
[20 , 25[	20
25 o más	10

Fuente: Córdova (2003, p.51)

2.1.5. La Mediana

La definición según Novaes (2011), "Es el valor que divide el conjunto ordenado de valores observados en otros dos conjuntos con el mismo número de elementos". (p. 91)

La definición según Córdova (2003), "Es la medida promedio que depende del número de datos ordenados y no de los valores de estos datos". (p. 35)

Una tercera definición menciona que:

> La mediana es el número que deja al menos la mitad de los datos a ese número por debajo, y la otra mitad de los datos a ese número, por encima; si existe más de uno de esos números, la mediana es el punto medio de este intervalo.(Langford, 2006, p. 5)

Un ejemplo se muestra en el cálculo de la mediana en datos no agrupados. (p. 33)

Se denota por *Me*.

Algunas Características para el análisis de la idoneidad del uso y ventajas de la mediana.

Tabla 10. Características para el análisis de la idoneidad del uso de la mediana.

USAR MEDIANA	En datos numéricos con distribución sesgada a la derecha (la media es mayor que la mediana), y con distribución sesgada a la izquierda (la media es menor que la mediana).

Fuente: Pinzón (2012, p.p. 19-20)

Tabla 11. Ventajas de la mediana.

	VENTAJAS
LA MEDIANA	No es afectada por las observaciones extremas, ya que no depende de los valores que toma la variable, sino del orden de las mismas.

Fuente: Pinzón (2012, p.25)

2.1.6. Propiedades más importantes de la mediana.

1) La mediana, solo depende del número de datos ordenados y no del valor de los datos. Por lo tanto no es sesgada por algún valor grande o pequeño.
2) La mediana puede ser calculada para variables con valores en escala ordinal.
3) La mediana es el mejor representante cuando hay datos extremos que toman valores muy altos o muy bajos.

2.1.7. Cálculo de la mediana de datos no agrupados o no tabulados

Según Córdova (2003):

> Para calcular la mediana de *n* valores no tabulados de alguna variable cuantitativa *x*, se sigue el siguiente proceso.
>
> 1) Se ordenan los datos en forma creciente
> 2) Luego, se ubica el valor central *Me*. Si *n* es impar la mediana es un dato observado. Si *n* es par la mediana es la semisuma de los dos valores centrales.

Ejemplo:

Calcular la mediana para las siguientes series de datos.

a) 120, 3, 14, 1, 99, 7, 30, 2 000, 16

b) 30, 77, 3, 300, 36, 11, 10 000, 29

Solución:

a) Como el número de datos es impar, la mediana de la serie ordenada de los 9 datos corresponde al quinto dato (x_5):

1, 3, 7, 14, **16**, 30, 99, 120, 2 000.

Me = 16.

La mediana divide a la serie en 2 grupos de igual cantidad de datos.

b) Como el número de datos es par, la mediana de la serie ordenada de los 8 datos corresponde a la semisuma de los datos centrales (x_4 y x_5):

3, 11, 29, **30, 36**, 77, 300, 10 000.

Esto es *Me* = (30 + 36) / 2 = 33.

La mediana en este caso, puede ser cualquier número situado entre 30 y 36, ya que dividirá a los datos en dos grupos de 4 datos cada uno. Se elige el valor 33 que corresponde a la semisuma, para evitar la infinidad de valores.

Observar pues, que la mediana no depende de la magnitud de los datos, depende solo del número de ellos.

c) Las edades de 25 personas son:

63	89	36	49	56	64	59	35	78	43
53	70	57	62	43	68	62	26	64	72
52	51	62	61	71					

Solución:

Ordenando los 25 datos se verifica que la mediana corresponde al dato decimotercero (x_{13}).

26	35	36	43	43	49	51	52	53	56
57	59	**61**	62	62	62	63	64	64	68
70	71	72	78	89					

$$Me = 61$$

A continuación se muestra el procedimiento para el caso donde el valor de la mediana se repite:

> Del conjunto de datos (1, 2, 2, 3, 4). Nadie niega que la mediana es 2. Pero es el segundo "2" del todo y no es el primer "2", que tiene la propiedad: Si se repiten los valores de los datos, tenemos que reemplazar "mayor que" para "la derecha de" y, del mismo modo, "menor que", "mayor que o igual a" y "menor o igual a". (Los dos son iguales, pero algunos son más iguales que otros). En este caso 2, depende sólo de su valor numérico y no la ocurrencia particular de ese valor. (Langford, 2006, p. 4)

2.1.8. Mediana de datos agrupados o tabulados

a) Mediana para datos agrupados de variable discreta.

Si los valores de una variable discreta se tabulan en una distribución de frecuencias de la forma “dato ---> frecuencia”, el cálculo de la mediana se hace siguiendo el procedimiento de ordenar el conjunto de datos.

Ejemplo d.

La mediana del número de hijos por familia (variable x) de una muestra de 20 hogares, marcó las siguientes respuestas:

2, 1, 2, 4, 1, 3, 2, 3, 2, 0, 3, 2, 1, 3, 2, 3, 3, 1, 2, 4

Construir la distribución de frecuencias de la variable x y graficar.

Solución:

Al ordenar estos datos en forma ascendente, se obtienen cinco valores distintos 0, 1, 2, 3, 4 que se repiten respectivamente 1, 4, 7, 6, 2 veces. La distribución de frecuencias de *x* se da en el siguiente cuadro.

Tabla 12. Distribución de frecuencias del número de hijos por familia

Número de hijos x_i	Frecuencias absolutas f_i	Frecuencias absolutas acumuladas F_i	Frecuencias relativas h_i	Frecuencias porcentajes p_i
0	1	1	0.05	5
1	4	5	0.20	20
2	7	12	0.35	35
3	6	18	0.30	30
4	2	20	0.10	10
Total	n=20		1.00	100

Fuente: Córdova (2003, p. 13)

Si n = 20, entonces $n/2$ = 10 (ubicamos la frecuencia acumulada que contiene al valor 10).

Por lo tanto la mediana corresponde a 2 hijos por familia.

b) Mediana para datos agrupados o tabulados por intervalos.

Hemos considerado los procedimientos de Suarez (2012):

i) Por interpolación:

Ejemplo: Calcular la mediana de las 25 edades tabuladas del ejemplo c).

Tabla 13. Cálculo de la mediana

Edades I_i	Número de personas f_i
[26, 39[	3
[39, 52[	4
[52, 65[	12
[65, 78[	5
[78, 91]	1
Total	25

Solución:

Primero se calcula n/2 y después se averigua el intervalo en el que está la mediana, este intervalo recibe el nombre de intervalo o clase de la mediana. Para averiguar el intervalo en el que está la mediana se aconseja calcular la frecuencia acumulada.

$$\frac{n}{2} = \frac{25}{2} = 12{,}5$$

Tabla 14. Cálculo de la mediana

I_i	f_i	F_i
[26, 39[	3	3
[39, 52[	4	7
[52, 65[	12	19
[65, 78[	5	24
[78, 91]	1	25
Total	25	

En este ejemplo el intervalo de la mediana es [52, 65[. Se observa que 7 valores están por debajo del valor 52. Los 5,5 que faltan para llegar a 13 se interpolan en el ancho del intervalo de la mediana que en este ejemplo es 13.

12 corresponde a 13

1 corresponde a 13/12

5,5 corresponde a $5{,}5\left(\frac{13}{12}\right) = 5{,}96$

Por lo tanto la Mediana es igual a 52 + 5,96 = 57,96

ii) Empleando la ecuación:

$$\boxed{M_e = L_i + \frac{\frac{n}{2} - F_{i-1}}{f_i} . A}$$

Figura 4. Cálculo de la mediana en datos agrupados por intervalos.

Fuente: Córdova (2003, p.53)

En donde:

L_i es el límite inferior del intervalo de la mediana.

n es el número de datos observados.

F_{i-1} es la frecuencia acumulada absoluta del intervalo inmediatamente anterior al intervalo de la mediana.

f_i es la frecuencia absoluta del intervalo de la mediana.

A es la amplitud del intervalo de la mediana.

Ejemplo: Calcular la mediana del ejemplo anterior usando la fórmula.

Solución:

Se calcula la frecuencia acumulada como se muestra en la siguiente tabla.

Tabla 15. Cálculo de la mediana

I_i	f_i	F_i
[26, 39[	3	3
[39, 52[	4	7
[52, 65[	12	19
[65, 78[	5	24
[78, 91]	1	25
Total	25	

Se calcula la posición de la mediana de la siguiente manera:

$$\frac{n}{2} = \frac{25}{2} = 12{,}5$$

por lo tanto el intervalo o clase de la mediana es [52, 65[.

Al aplicar la ecuación respectiva se obtiene:

$$\boxed{M_e = L_i + \frac{\frac{n}{2} - F_{i-1}}{f_i}.A = 52 + \left(\frac{12.5 - 7}{12}\right)x13 = 52 + 5.96 = 57.96}$$

iii) De manera gráfica usando proporcionalidad

Se construye un histograma de las frecuencias acumuladas.

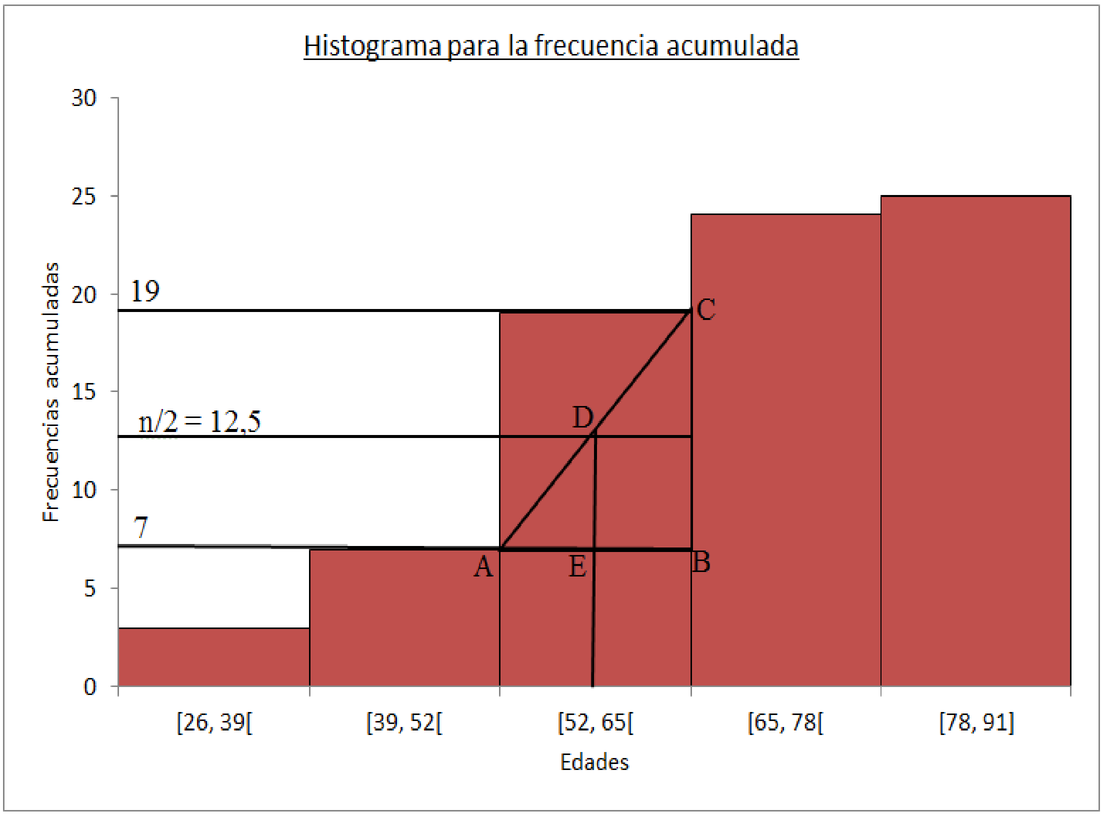


Figura 5. Cálculo de la mediana en datos agrupados por intervalos.

Fuente: Suarez (2012, p.15)

Observando el gráfico se determina que *Me* = 52 + AE

$$\frac{AB}{CB} = \frac{AE}{DE}$$

$$\frac{65-52}{19-7} = \frac{AE}{12{,}5-7} \rightarrow \frac{13}{12} = \frac{AE}{5{,}5}$$

Despejando AE se obtiene:

$$AE = \frac{13\ x\ 5{,}5}{12} = 5{,}96$$

Entonces *Me* = 52 + AE = 52 + 5,96 = 57,96

Para determinar la medida más representativa del conjunto de datos, no basta con saber el algoritmo de cálculo de la media aritmética y la mediana, sino que se debe determinar la presencia o no de datos atípicos que sesgan la información, para ello necesitamos construir un intervalo de referencia usando el cuartil 1 (Q_1) y el cuartil 3 (Q_3). A continuación describimos los cuartiles.

2.1.9. Cuartiles

Según la definición de Sweeney (2008), son medidas de posición que permiten determinar las características de un conjunto de datos, dividiendo los datos en cuatro partes iguales a través de tres cortes y cada parte contiene el 25% de las observaciones.

Primer cuartil: Denotado por Q_1, es el menor valor que es mayor que una cuarta parte de los datos; es decir, aquel valor de la variable que supera el 25% y es superado por el 75% de las observaciones.

Tercer cuartil: Denotado por Q_3, es el menor valor que es mayor que tres cuartas partes de los datos, es decir aquel valor de la variable que supera el 75% y es superado por el 25% de las observaciones se interpreta como que el 75% de los datos es menor que el Q_3 obtenido.

La siguiente figura es un ejemplo de la localización de los cuartiles en una distribución simétrica con intervalos de igual magnitud.

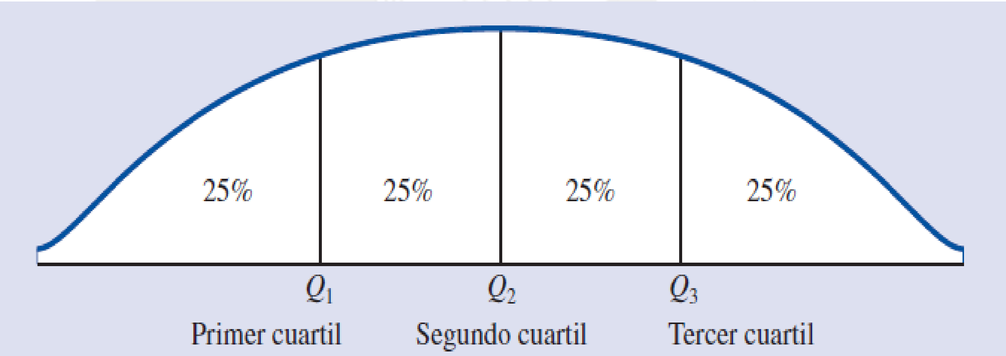


Figura 6. Localización de los cuartiles en distribuciones simétricas.

Fuente: Sweeney (2008, p. 87)

NOTA: El cuartil 2 equivale al valor de la mediana ($Q_2 = Me$)

Sin embargo aclaramos que existen otros tipos de distribuciones.

2.1.10. Cálculo de los cuartiles

Tendremos en cuenta el siguiente procedimiento de cálculo:

> Los cuartiles han sido definidos como el percentil 25, el percentil 50 y el percentil 75. Por lo que los cuartiles se calculan de la misma manera que los percentiles. Sin embargo, algunas veces se siguen otras convenciones para calcular los cuartiles, por ello los valores que se dan para los cuartiles varían ligeramente, dependiendo de la convención que se siga. De cualquier manera, el objetivo de calcular los cuartiles siempre es dividir los datos en cuatro partes iguales. (Sweeney, 2008, p. 88)

- **Cálculo del primer cuartil (Q_1):**

Primero se ordenan los datos.

Segundo se divide el total de datos entre 4 y se multiplica por 1. Si el resultado es un número entero, se le deberá sumar 0,5. En cambio si el resultado no es un número entero, este se deberá tomar como el siguiente entero más grande. El valor obtenido corresponde a la posición del cuartil.

- **Cálculo del tercer cuartil (Q_3):**

Primero se ordenan los datos.

Segundo se divide el total de datos entre 4 y se multiplica por 3. Si el resultado es un número entero, se le deberá sumar 0,5. En cambio si el resultado no es un número entero, este se deberá tomar como el siguiente entero más grande. El valor obtenido corresponde a la posición del cuartil.

El segundo cuartil (Q_2), se calcula de la misma forma que la mediana, teniendo en cuenta que se presentan dos casos: cuando la cantidad de datos es un número impar y cuando la cantidad de datos es un número par. Esto ya ha sido señalado anteriormente (ver pág. 31).

Otro procedimiento:

> Dividir el conjunto de los datos en dos mitades, una superior y otra inferior. Si el número de datos de cada conjunto es impar, añadir o eliminar la mediana en ambas mitades, por lo que cada mitad tendrá un número par de elementos. Los cuartiles superior e inferior serán entonces la mediana de las mitades superior e inferior, respectivamente. (Langford, 2006, p. 23)

Como estrategia didáctica, se recomienda ubicar primero el segundo cuartil, luego el cuartil 1 (Q_1) y cuartil 3 (Q_3), teniendo en cuenta que entre cada una de estas medidas deben estar ubicados la misma cantidad de datos.

Ejemplo:

Determinar los cuartiles correspondientes a las notas obtenidas por 10 estudiantes en el área de matemática: 05, 08, 09, 12, 14, 15, 18, 19, 20, 20.

Solución:

- Primero ordenamos los datos.

  05, 08, 09, 12, 14, 15, 18, 19, 20, 20

- Luego calculamos el Q_2 o mediana de los datos, teniendo en cuenta que cuando la cantidad de datos es par, la mediana es igual a la semisuma de los valores centrales, así:

$$Q_2 = M_e = \frac{x_5 + x_6}{2} = \frac{14 + 15}{2} = 14{,}5$$

05, 08, 09, 12, 14, | 15, 18, 19, 20, 20

$Q_2 = Me = 14{,}5$

- Ahora localizamos los cuartiles 1 (Q_1) y cuartil 3 (Q_3).
- $Q_1 = \frac{n}{4}.1 = \frac{10}{4}.1 = 2{,}5$

Como el resultado no es un número entero, tomamos el dato que ocupa el lugar 3, como posición del primer cuartil. En este caso Q_1 = 09.

- $Q_3 = \frac{n}{4}.3 = \frac{10}{4}.3 = 7{,}5$

Como el resultado no es un número entero, tomamos el dato que ocupa el lugar 8, como posición del tercer cuartil. En este caso Q_3 = 19

$$\underbrace{05,\ 08}, \ \mathbf{09},\ \underbrace{12,\ 14} \qquad \underbrace{15,\ 18},\ \mathbf{19},\ \underbrace{20,\ 20}$$

Q_1= 09 Me=14,5 Q_3=19

Observamos que antes y después de cada cuartil hay dos valores.

Los cuartiles nos van a ser útiles para determinar datos u observaciones atípicas, pero antes de ello expliquemos lo relacionado al sesgo en la información.

2.1.11. El sesgo en la distribución de los datos

Según Sweeney (2008), entendemos por sesgo a una medida de la forma de distribución de los datos. El sesgo es una consecuencia por la presencia de datos u observaciones atípicas, que son valores muy grandes o muy pequeños que hacen cambiar de posición a la media aritmética en relación a la mediana.

Cuando el conjunto de datos no presenta sesgo, se dice que la distribución es ***simétrica***; en este caso, la media aritmética coincide con la mediana o ambos valores son muy cercanos.

Sin embargo, cuando el conjunto de datos presenta sesgo, se dice que la distribución es ***asimétrica***; en este caso, la media aritmética es diferente a la mediana.

Se presentan dos casos de asimetría:

i) Si la media aritmética es menor que la mediana, se dice que el sesgo es negativo, o sea la gráfica que es una curva presenta una cola hacia la izquierda (asimétrica hacia la izquierda).

ii) Si la media aritmética es mayor que la mediana, se dice que el sesgo es positivo, o sea la curva presenta una cola hacia la derecha (asimétrica hacia la derecha).

La siguiente figura muestra las diferentes formas de la distribución.

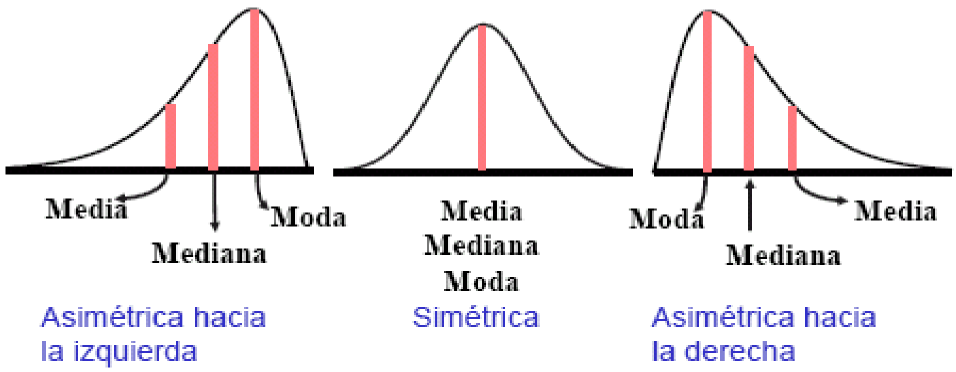
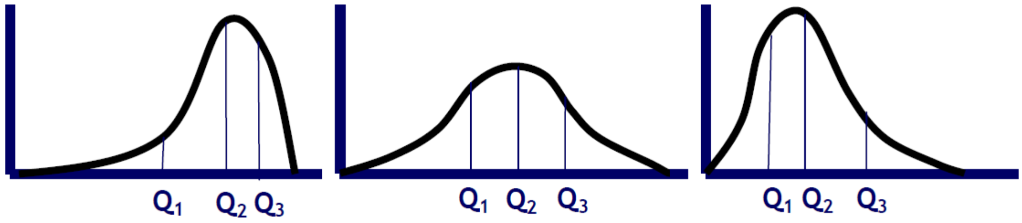


Figura 7. Distribución simétrica y asimétrica.

Fuente: Perú (2006, p.57)

También se puede mostrar la relación entre los cuartiles respecto a la simetría o asimetría del conjunto de datos.



$Q_3 - Q_2 < Q_2 - Q_1$ Asimetría negativa

$Q_3 - Q_2 = Q_2 - Q_1$ Simetría=0

$Q_3 - Q_2 > Q_2 - Q_1$ Asimetría positiva

Figura 8. Distribuciones simétricas y asimétricas en relación a los cuartiles.

Fuente: Estadística descriptiva en Psicología (p. 4).

2.1.12. Datos u observaciones atípicas.

Conocidos también como datos muy grandes o muy pequeños que están alejados del resto de datos. Generan sesgo, hacia la derecha o hacia la izquierda, estos valores influyen considerablemente en el cálculo de la media aritmética y la hacen poco representativa.

> Es una observación que es inusual en relación con el resto de datos; en otras palabras, una observación atípica no sigue el patrón del resto de los datos. Una observación extraña quizá sea el valor de un dato que se anotó de modo incorrecto. Si es así puede corregirse antes de continuar con el análisis. Una observación atípica tal vez provenga, también, de una observación que se incluyó indebidamente en el conjunto de datos; si es así se puede eliminar. Por último, una observación atípica quizá es un dato con un valor inusual, anotado correctamente y que sí pertenece al conjunto de datos. En tal caso debe conservarse. (Sweeney, 2008, p.p. 102, 659)

2.1.13. ¿Cómo determinar datos u observaciones atípicas?

Según Sweeney (2008), “se realiza el análisis exploratorio de datos que permite usar operaciones aritméticas sencillas, y representaciones gráficas fáciles de dibujar para resumir datos, considerando los resúmenes de cinco números y los diagramas de caja”. (p. 105)

> En el **resumen de cinco números** se usan los cinco números siguientes para resumir los datos.
> **1.** El valor menor.

2. El primer cuartil (Q_1).

3. La mediana (Q_2).

4. El tercer cuartil (Q_3).

5. El valor mayor.

La manera más fácil de elaborar un resumen de cinco números es, primero, colocar los datos en orden ascendente. Hecho esto, es fácil identificar el valor menor, los tres cuartiles y el valor mayor. Por ejemplo los salarios de 12 administradores (en dólares):

$$\underbrace{3310,\ 3355,\ 3450}, \quad \underbrace{3480,\ 3480,\ 3490}, \quad \underbrace{3520,\ 3540,\ 3550} \ \underbrace{3650,\ 3730,\ 3925}$$

Q_1=3465 Q_2=Me=3505 Q_3=3600

Así, el resumen de cinco números correspondiente a los datos de los salarios iniciales es 3310, 3465, 3505, 3600 y 3925. Entre cada dos números adyacentes del resumen de los cinco números se encuentran aproximadamente el 25% de los datos.

Diagrama de caja y bigotes

Es un resumen gráfico de los datos con base en el resumen de cinco números.

La clave para la elaboración de un diagrama de caja es el cálculo de la mediana y de los cuartiles Q_1 y Q_3 También se necesita el rango intercuartílico, RIC=Q_3 – Q_1. En la figura 9 se presenta el diagrama de caja de los datos de los salarios mensuales iniciales. Los pasos para elaborar un diagrama de caja son los siguientes.

1 Se dibuja una caja cuyos extremos se localicen en el primer y tercer cuartil. En los datos de los salarios iniciales Q_1=3465 y Q_3=3600. Esta caja contiene 50% de los datos centrales.

2. En el punto donde se localiza la mediana (3505 en los datos de los salarios) se traza una línea vertical.

3. Usando el rango intercuartílico, RIC = Q_3 – Q_1, se localizan los *límites*. En un diagrama de caja los límites se encuentran 1,5 (RIC) abajo del Q_1 y 1,5 (RIC) arriba del Q_3.

En el caso de los salarios, RIC = Q_3 – Q_1 = 3600 – 3465 =135

Por tanto, los límites del intervalo son:

Límite inferior = 3465 – 1,5(135) = 3262,5

Límite superior = 3600 + 1,5(135) = 3802,5

El intervalo para determinar datos atípicos es [3262,5; 3802,5].

Los datos que quedan fuera de estos límites se consideran ***observaciones atípicas***.

4. A las líneas punteadas que se observan en la figura 9 se les llama *bigotes*. Los bigotes van desde los extremos de la caja hasta los valores menor y mayor *de los límites* calculados en el paso 3. Por tanto, los bigotes terminan en los salarios cuyos valores son 3310 y 3730.

5. Por último mediante un asterisco se indica la localización de las observaciones atípicas. (Sweeney, 2008, p.p. 105-106)

En la figura 8 se observa que 3925 es una observación atípica.

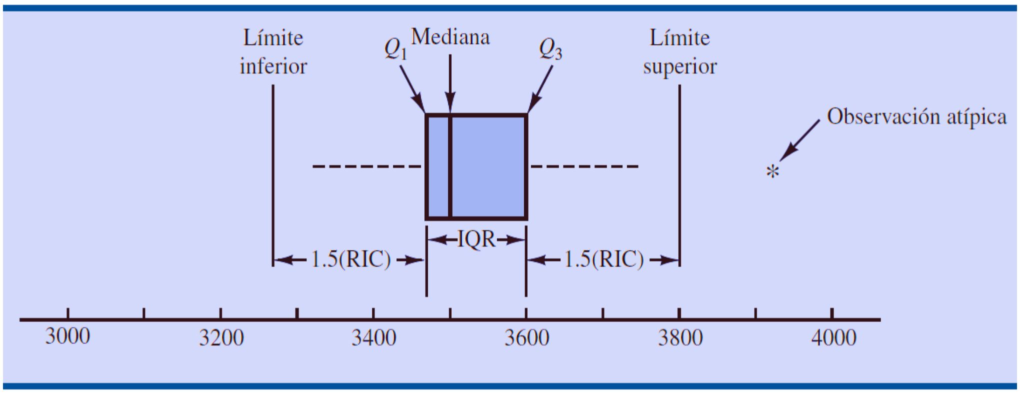


Figura 9. Diagrama de caja de los salarios iniciales en dólares

Fuente: Sweeney (2008, p.106)

Una forma sencilla que resume el diagrama de caja y bigotes para determinar datos u observaciones atípicas, consiste en construir el intervalo de la forma $[Q_1 - D, Q_3 + D]$.

Donde:

Q_1 y Q_3 son los cuartiles primero y tercero respectivamente.

D, es el producto del rango intercuartílico (RIC=Q_3–Q_1) multiplicado por el factor 1,5.

$$D = 1{,}5(Q_3 - Q_1)$$

Utilizando esta estrategia no es necesario dibujar el diagrama de caja y bigotes, pues, solo basta construir el intervalo usando los cuartiles y analizar los datos de la muestra, teniendo en cuenta que:

- Si todos los datos pertenecen al intervalo, no existen datos u observaciones atípicas.
- Si algún dato no pertenece al intervalo, diremos que es dato u observación atípica y por lo tanto existe sesgo en el conjunto de datos.

Veamos el siguiente ejemplo:

Determinar si existen datos atípicos en las evaluaciones obtenidas por 9 estudiantes en el área de matemática.

5, 6, 7, 9, 9, 10, 12, 14, 20.

Solución:

Como los datos ya están ordenados, ubicamos primero la mediana (Me = Q_2) y luego los cuartiles Q_1 y Q_3.

$$\underbrace{5,\quad 6,\quad 8},\quad \underbrace{9,\quad 9,\quad 10},\quad \underbrace{10,\quad 11,\quad 11},\quad \underbrace{12,\quad 14,\quad 20}$$

Q_1=8,5 $\quad$ Q_2=Me=10 $\quad$ Q_3=11,5

$$D = 1{,}5(Q_3 - Q_1)$$

$$D = 1{,}5(11{,}5 - 8{,}5) = 1{,}5(3) = 4{,}5$$

El intervalo de determinación de datos u observaciones atípicas será de la forma:

$$[Q_1 - D, Q_3 + D]$$

Reemplazando valores, tenemos:

$$[8{,}5 - 4{,}5;\ 11{,}5 + 4{,}5]$$

Finalmente se obtiene el intervalo $[4;\ 16]$

Procedemos a analizar la pertenencia o no de los datos del conjunto muestral, respecto al intervalo construido $[4;\ 16]$ y nos damos cuenta que el dato 20 no pertenece al intervalo [4; 6], por lo que concluimos que es un dato u observación atípica.

En conclusión:

> Una ventaja de los procedimientos del análisis exploratorio de datos es que son fáciles de usar; son necesarios pocos cálculos. Simplemente se ordenan los datos de menor a mayor y se identifican los cinco números del resumen de cinco números. Después se construye el diagrama de caja. No es necesario calcular la media ni la desviación estándar de los datos. (Sweeney, 2008, p. 107)

Es importante resaltar que este procedimiento nos ayuda a justificar la decisión de elegir a la media aritmética o la mediana como el mejor representante del conjunto de datos. La sencillez de su construcción facilita el análisis.

2.2. MARCO DIDÁCTICO

2.2.1. Situación actual de la enseñanza de las medidas de tendencia central

Es conveniente indicar las opiniones respecto a la enseñanza de la estadística ante el creciente interés por incluir la estadística dentro de la educación matemática en las instituciones educativas del nivel secundario, la enseñanza de la misma se ve afectada por diferentes razones. Según Batanero (2001), “Una primera dificultad proviene de los cambios progresivos que la estadística está experimentando en nuestros días, tanto desde el punto de vista de su contenido, como del punto de vista de las demandas de formación. (p. 6)

Esto se refiere a la inserción, en el currículo, de contenidos estadísticos, como por ejemplo, el interés por incluir el uso de las medidas de tendencia central, la media aritmética y la mediana. Para su enseñanza, los docentes, deben desarrollar estos contenidos estadísticos usando estrategias y propiedades estadísticas que ayuden a la comprensión e interpretación de la información.

Por otro lado destacamos que:

> La realidad docente indica que estos contenidos no se enseñan con la profundidad que merecen. En el mejor de los casos, la enseñanza de la estadística es un pretexto para aplicar otros temas matemáticos y ejercitar la capacidad de cálculo o representación gráfica, olvidando el trabajo con datos reales y los aspectos de razonamiento estadístico. (Mayen, 2009, p.11)

Según Batanero (2001), “El aprendizaje conceptual (definiciones, propiedades y sus significados) permiten que el docente desarrolle las habilidades necesarias de los estudiantes”, para el análisis de la media aritmética y la mediana como representante de un conjunto de datos.

La estadística dentro del currículo cada vez adquiere un protagonismo más importante, lo que implica contar con docentes bien preparados en esta materia:

> La estadística como ciencia, atraviesa un periodo de notable expansión, siendo cada vez más numerosos los procedimientos disponibles, alejándose cada vez más de la matemática pura y convirtiéndose en una "ciencia de los datos", lo que implica la dificultad de enseñar un tema en continuo cambio y crecimiento. (Batanero, 2001, p. 6)

De los resultados obtenidos en el cuestionario aplicado a los docentes de matemática (ver anexo 8), podemos decir que la mayoría de la muestra piensa que la media aritmética es la medida más representativa de un conjunto de datos, por considerarla más exacta que la mediana. Al observar la presencia de datos que no son similares al resto, creen que se debe a errores de medición o información inexacta, procediendo en algunos casos a eliminarlos sin antes analizarlos. En este trabajo creemos que los datos recogidos de un estudio deben ser considerados en su totalidad, dejar de lado uno o más datos por considerarlos erróneos sin antes analizarlos; hace que la media aritmética no sea realmente la que representa al conjunto de datos. Son los datos u observaciones atípicas los que sesgan la información y los que permiten decidir finalmente, qué medida usar como promedio.

En nuestra propuesta didáctica, tendremos en cuenta a todos los datos del conjunto. Ninguna actividad muestra datos erróneos, más bien realizaremos el análisis de los datos, usando estrategias para determinar datos u observaciones atípicas (ver pág. 27). Después de ello haciendo uso de las propiedades estadísticas de la media aritmética y la mediana (ver pág. 14 y 19), tendremos argumentos para justificar la elección de la media aritmética o la mediana como medida representativa. Los casos de los salarios de los trabajadores de una empresa, son algunos de los ejemplos a considerar en nuestra propuesta didáctica, en la que, la elección de la mediana resulta ser la medida estadística más representativa.

Actualmente la elección del representante del conjunto de datos se constituye como uno de los logros de aprendizaje en la competencia relacionada a gestión de datos, declarados en los mapas de progreso, "L***a interpretación y uso de las medidas de tendencia central reconociendo la medida representativa y la interpretación del sesgo en la distribución obtenida de un conjunto de datos"***. (Perú, 2013, p.10)

Considerando que es necesario que los docentes de matemática en ejercicio tengan acceso a información relacionada con la estadística, estamos seguros que una secuencia de actividades, para la enseñanza del uso de la media aritmética y la mediana en un conjunto de datos, deberían y podrían ser propuestos para acercarse a otros conceptos estadísticos que se enseñan en la escuela, e incluso proponer secuencias para el estudio de la variabilidad en un conjunto de datos (el eje central del pensamiento estadístico a desarrollar por los estudiantes).

.

2.3. MARCO TEÓRICO

Se ha considerado como marco teórico el Enfoque Ontosemiótico de la Cognición e Instrucción Matemática (EOS), porque “proporciona un conjunto de nociones teóricas que hacen posible analizar con detalle los procesos de enseñanza y aprendizaje de temas específicos relacionados a los sistemas de prácticas matemáticas”. Explicamos los niveles y aspectos o facetas que utiliza esta teoría en el tratamiento de objetos matemáticos, en nuestro caso la enseñanza del uso de la media aritmética y la mediana. Del mismo modo, los diversos conflictos semióticos, llamado así por tratarse de los problemas que manifiesta un sujeto respecto a los signos en el campo del conocimiento de este concepto matemático, que sin lugar a dudas suceden influenciados por diversos factores en el intento de aprender.

2.3.1. Enfoque Ontosemiótico de la Cognición e Instrucción Matemática (EOS)

El EOS como teoría para la investigación en educación matemática, propone a la comunidad de investigadores un conjunto de nociones teóricas, basadas en las teorías didácticas para la enseñanza de las ciencias de Peirce, Wittgenstein, Vygotsky, Habermas y Morín. Según Godino, Batanero y Font. (2009), el EOS se compone de cinco facetas o etapas, cada una de las cuales permite un nivel de análisis de los procesos de enseñanza y aprendizaje de temas específicos de matemáticas. Según Godino et al., este marco teórico integrativo ha surgido en el seno de la didáctica de las matemáticas, con el propósito de articular diferentes puntos de vista y nociones teóricas sobre el conocimiento matemático, su enseñanza y aprendizaje. El EOS, adopta una perspectiva global, teniendo en cuenta las diversas dimensiones implicadas y las interacciones entre las mismas. Propone un modelo epistemológico sobre las matemáticas basadas en presupuestos antropológicos/socioculturales, un modelo de cognición matemática sobre bases semióticas, un modelo instruccional sobre bases socio-constructivistas, un modelo sistémico – ecológico que relaciona las anteriores dimensiones entre sí y con el trasfondo biológico, material y sociocultural en que tiene lugar la actividad de estudio y comunicación matemática.

El EOS se apoya y nutre de las aportaciones de las diversas disciplinas y tecnologías interesadas en la cognición humana: Epistemología, Sociología, Semiótica, Ciencias de la educación, entre otras. También permite definir una “agenda de investigación” para abordar de manera sistemática cuestiones profesionales básicas que respondan ¿Qué matemáticas enseñar?, y ¿Cómo enseñar esas matemáticas de modo que se aprendan de la mejor manera posible?

Los cinco niveles o facetas propuestos por el EOS, para la investigación de la enseñanza de las matemáticas son: Sistema de prácticas, Configuración de objetos y procesos, Trayectorias didácticas, Dimensión normativa e Idoneidad didáctica. A continuación explicamos las facetas de las cuales usaremos en nuestra propuesta, las configuraciones de objetos y procesos y las idoneidades didácticas.

2.3.2. Sistemas de prácticas.

Según Godino (2003), el EOS la define como acciones que realiza el sujeto resolutor frente a problemas matemáticos, utiliza para ello la Socioepistemología de (Cantoral/Farfán), la Teoría Antropológica de (Chevalard), la Etnomatemática de (D'Ambrosio) y la Fenomenología de (Freudenthal) entre otras. Del mismo modo, este sistema de prácticas analiza los significados personales tipificados como (global, declarado y logrado) e institucionales (referencial, pretendido, implementado y evaluado). Este es un sistema de prácticas operativas y discursivas que tiene como trasfondo ecológico de las prácticas (material, biológico y social).

2.3.3. Configuración de objetos y procesos.

El EOS la define como la adopción de una noción interaccionista de objeto y pragmatista del significado (contenido de funciones semióticas). Utiliza para ello la Teoría APOS de (Dubinsky), la Teoría de Campos Conceptuales de (Vergnaud), la Teoría de Registros de Representaciones Semióticas de (Duval) entre otras. Esta configuración de objetos y procesos distingue según el EOS, seis tipos primarios de entidades intervinientes:

Situaciones, acciones, lenguaje, conceptos, propiedades y argumentos, en una segunda fase, estos objetos se relacionan entre sí formando "configuraciones", definidas como las redes de objetos intervinientes y emergentes de los sistemas de prácticas. Estas configuraciones pueden ser epistémicas (redes de objetos institucionales) o cognitivas (redes de objetos personales), definidas como las redes de objetos intervinientes y emergentes de los sistemas de prácticas.

Las prácticas matemáticas según el EOS, son actuaciones o expresiones realizadas por personas relacionadas con el aprendizaje, dichas acciones tienen propósitos muy específicos, son intencionales y las realiza el docente para lograr los aprendizajes esperados en sus estudiantes. Estas prácticas según el EOS se pueden clasificar en tres tipos: i) *Las practicas operativas o actuativas,* que tienen que ver con el desempeño docente en el aula, cuyo fin es la resolución de situaciones problemáticas que se plantean al inicio de un tema específico, y para ello demanda realizar acciones y argumentaciones; ii) *Las practicas discursivas o* *comunicativas*, encargadas de validar las acciones, guarda relación con el uso del lenguaje al momento de realizar las argumentaciones; y iii) *Las practicas regulativas o normativas* encargadas de establecer las propiedades y definiciones de objetos matemáticos.

En la propuesta didáctica presentamos una ***configuración epistémica*** de referencia y otras dos configuraciones epistémicas de las actividades, las mismas que guardan relación y han sido elaboradas con la intención de proporcionarle al docente herramientas que contribuyan al logro de los aprendizajes esperados con relación al uso de la media aritmética y la mediana, como representante de un conjunto de datos.

2.3.4. Trayectorias Didácticas.

En este nivel se considera que un proceso de estudio está conformado por seis dimensiones que se entrelazan: Epistémica (de los significados institucionales), docente (de las funciones del profesor), discente (de las funciones de los estudiantes), mediacional (de los medios o recursos materiales), cognitiva (de los significados personales de los estudiantes) y emotiva (de los estados emocionales de los estudiantes).

2.3.5. Dimensión normativa.

Basada en las normas sociomatemáticas de Cobb y Volgt, según el EOS, se refiere al conjunto de normas, hábitos y convenciones generalmente implícitas que regulan el funcionamiento de la clase de matemáticas. Llamadas también contrato didáctico entre los estudiantes y el profesor, que son las reglas de juego que regulan el proceso enseñanza y aprendizaje dentro y fuera del aula.

Estas normas son de cuatro tipos: según el origen (Administración, sociedad, escuela, aula y disciplina), según la faceta (epistémicas, cognitivas, afectivas, interacciónales, mediacionales y ecológicas), según el momento (curricular, planificación, implementación y evaluación) y según el tipo y grado de coerción (en el aspecto social son: leyes, decretos, ordenes, resoluciones, hábitos y costumbres y en el aspecto disciplinar son: teoremas, definiciones y convenciones).

Según Godino, Font y Wilhelmi (2008), en el aula hay normas de convivencia, dentro de ellas se desprenden las normas sociomatemáticas. Las normas sociomatemáticas son diferentes de las normas sociales generales que rigen el comportamiento en las aulas, en el sentido de que son específicas de los aspectos matemáticos de la actividad de los estudiantes. En este contexto, se da mucha importancia a las normas sociomatemáticas debido a que el desarrollo del razonamiento y los procesos de dotar de sentido a los estudiantes no puede ser separado de su participación en la constitución interactiva del significado matemático.

2.3.6. Idoneidad Didáctica.

El término **Idoneidad,** según el diccionario de la real academia española y otros, lo refieren como un sustantivo femenino que deriva del término en latín, ***idoneitate***, que expresa la calidad de lo **idóneo**, lo adecuado, y también significa capacidad, aptitud, calificación, habilidad y competencia.

La característica de alguien que es idóneo, revela a alguien conveniente, apto, capaz, útil, apropiado y adecuado, que tiene ciertas condiciones para desempeñar determinados cargos o funciones o realizar determinadas obras.

La idoneidad a la que nos referimos en este trabajo está relacionada con las características más significativas de nuestra propuesta en el desarrollo del uso de la media aritmética y la mediana.

Basada en la Educación Matemática crítica de Skovsmose (1999), tiene como objetivo mejorar los procesos de enseñanza y aprendizaje de las matemáticas.

> La noción de idoneidad didáctica, sus dimensiones, criterios, y un desglose operativo de dicha noción, ha sido introducida en ella, como herramienta que permite el paso de una didáctica descriptiva – explicativa a una didáctica normativa, esto es, una didáctica que se orienta hacia la intervención efectiva en el aula. Consideramos que esta noción puede servir de punto de partida para una teoría de diseño instruccional (Teoría de la Idoneidad Didáctica) que tenga en cuenta, de manera sistémica, las dimensiones epistémica – ecológica, cognitiva – afectiva e interaccional – mediacional, implicadas en los procesos de estudio de las áreas curriculares específicas. La idoneidad didáctica de un proceso de instrucción se define como la articulación coherente y sistémica de las seis componentes siguientes:
>
> - *Idoneidad epistémica*, se refiere al grado de representatividad de los significados institucionales implementados (o pretendidos), respecto de un significado de referencia.
> - *Idoneidad cognitiva*, expresa el grado en que los significados pretendidos/implementados estén en la zona de desarrollo potencial de los estudiantes, así como la proximidad de los significados personales logrados a los significados pretendidos/implementados.
> - *Idoneidad interaccional*. Un proceso de enseñanza-aprendizaje tendrá mayor idoneidad desde el punto de vista interaccional si las configuraciones y trayectorias didácticas permiten, por una parte, identificar conflictos semióticos potenciales (que se puedan detectar a priori), y por otra parte permitan resolver los conflictos que se producen durante el proceso de instrucción.

- *Idoneidad mediacional*, grado de disponibilidad y adecuación de los recursos materiales y temporales necesarios para el desarrollo del proceso de enseñanza-aprendizaje.
- *Idoneidad afectiva,* grado de implicación (interés, motivación,…) del alumnado en el proceso de estudio. La idoneidad afectiva está relacionada tanto con factores que dependen de la institución como con factores que dependen básicamente del alumno y de su historia escolar previa.
- *Idoneidad ecológica*, grado en que el proceso de estudio se ajusta al proyecto educativo del centro, la escuela y la sociedad y a los condicionamientos del entorno en que se desarrolla. (Godino, Batanero y Font (2007, p. 5-6)

Según el EOS, estos seis aspectos de la idoneidad se complementan y hacen posible la valoración de una secuencia didáctica con indicadores claros, que pretenda no solo dar al docente una herramienta valiosa para validar una sesión de aprendizaje, sino que se constituye como el patrón encargado de su fiabilidad.

Es importante señalar que:

> La teoría de la idoneidad didáctica trata de interrelacionar las distintas facetas que intervienen en el diseño, implementación y evaluación de procesos de enseñanza – aprendizaje de las matemáticas. Las nociones de idoneidad epistémica y ecológica y el sistema de indicadores asociados constituyen el germen de una teoría curricular, mientras que los correspondientes a las facetas cognitiva – afectiva lo constituye para una teoría del aprendizaje. Las facetas interaccional y mediacional contienen, a su vez, el germen de una teoría de la enseñanza. (Godino, 2011, p.18)

Tanto la idoneidad cognitiva como la epistémica, implican un acoplamiento, participación y apropiación de los significados personales de los estudiantes. Del mismo modo la idoneidad emocional, implica las actitudes, los afectos y las motivaciones de los estudiantes frente al proceso de estudio.

Según el EOS, la *idoneidad epistémica*, es la adaptación entre los significados institucionales *implementados* y de *referencia*, que en particular, supondría la elaboración de una transposición didáctica viable (capaz de adaptar el significado *implementado* al *pretendido*) y pertinente (capaz de adaptar el significado *pretendido* al de *referencia*).
En nuestro trabajo diseñaremos las configuraciones epistémicas de referencia de la enseñanza del uso de la media aritmética y la mediana (ver tabla 16), así como en cada una de las actividades que forman parte de la secuencia (ver tablas 22 y 23). También aplicaremos algunos criterios de idoneidad epistémica y cognitiva (ver tablas 17 y 18), que fueron

seleccionados porque permiten realizar, la planificación adecuada por parte del docente de la enseñanza del uso de la media aritmética y la mediana y por parte del estudiante, la manifestación de los logros de aprendizaje que se obtienen en la realización de las actividades.

Estos niveles o fases del EOS, se pueden observar en la siguiente figura presentada por Godino, Batanero y Font (2008). En ella se aprecia que las fases están organizadas como elementos que tienen la misma jerarquía, asociándole a cada una las teorías que se sustentan, así como el procedimiento que deben seguir los investigadores para su desarrollo.

Con el propósito de contribuir con los docentes del área de matemática y lograr el objetivo general de caracterizar una propuesta didáctica a priori para la enseñanza del uso de la media aritmética y la mediana, que permita en los estudiantes desarrollar la capacidad de comprensión, análisis y elección con éxito, elaboramos la configuración epistémica de la enseñanza del uso de estas medidas, que corresponde al segundo nivel propuesto en el EOS y la aplicación de algunos aspectos de la idoneidad didáctica que corresponde al quinto nivel.

El EOS: Un Marco Teórico Integrativo para la Educación Matemática

Juan D. Godino, Carmen Batanero y Vicenc Font

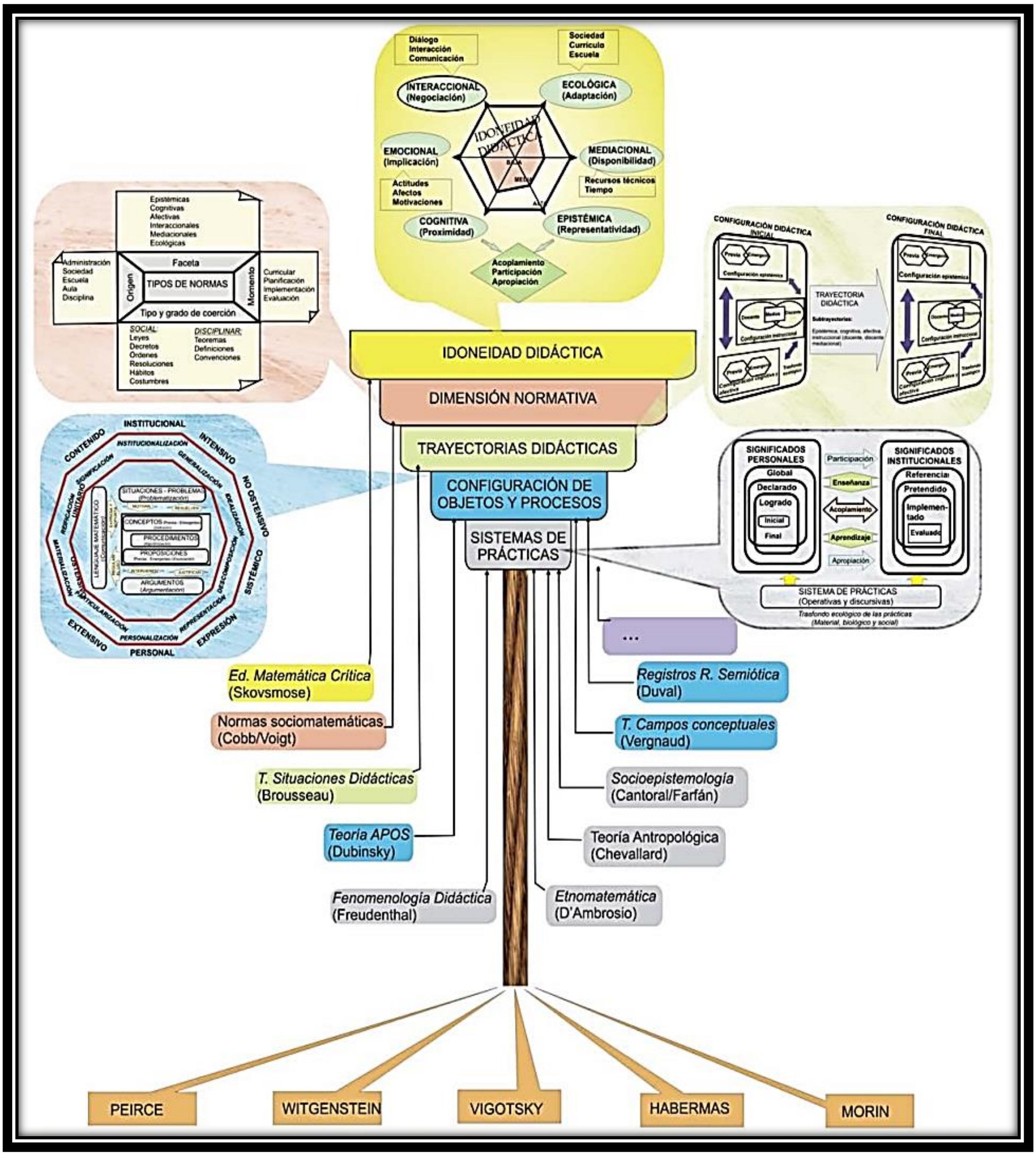


Figura 10. Síntesis del EOS

Fuente: Godino (2008, p. 2)

2.3.7. Configuración epistémica del uso de la media aritmética y la mediana.

El interés por la enseñanza del uso de la media aritmética y la mediana con estudiantes del VI ciclo de educación secundaria, hace necesario plantear la configuración epistémica de referencia.

El aprendizaje de los conceptos matemáticos depende de muchos factores como: buen docente, considerar al alumno como eje central del sistema educativo, reconocer sus talentos, adaptarse a los cambios y avance de la tecnología, despertar la curiosidad, aprender equivocándose, comunicación adecuada y oportuna, el docente como motivador de sus estudiantes y apasionado con su profesión entre otros.

Tendremos en cuenta para nuestra configuración, lo que el EOS define como una ontología formada por los siguientes elementos:

> **Lenguaje matemático**, es un conjunto de expresiones que usaremos en la configuración que tiene relación con el objeto de estudio, expresa y soporta reglas dentro de la configuración y su uso depende de las definiciones, procedimientos y proposiciones.
>
> **Situaciones – Problemas,** son actividades matemáticas motivadoras extraídas del contexto en que se desenvuelve el alumno y que tiene estrecha relación con el objeto de estudio. Se denomina también situación problemática que puede presentarse mediante texto o láminas, que involucra interrogantes, lo que llamamos conflicto cognitivo. Como implica resolver un problema, al final de la configuración se resuelve la situación planteada inicialmente.
>
> **Conceptos,** son las definiciones o significados de cada uno de los términos en la configuración, deben ser claros y sencillos.
>
> **Procedimientos, técnicas, algoritmos, operaciones,** son acciones o pasos que permiten avanzar en la solución de un problema. Se pueden obtener valores a partir de la utilización de una fórmula o la utilización de una simbología.
>
> **Proposiciones, propiedades, teoremas, etc.,** conjunto de enunciados que tienen la cualidad de ser verdaderos y que al ser utilizados, sirven para resolver una situación problemática.
>
> **Argumentaciones,** son las razones que justifican las tesis, inherentes a los significados trabajados como verdaderas. (Godino, Contreras y Font, 2006, p. 69)

Tabla 16. Configuración epistémica de la enseñanza del uso de la media aritmética y la mediana.

Situaciones – Problema	Las situaciones problemáticas deben ser de contexto real, trascendentes en el diario vivir de los estudiantes, las calificaciones, las propinas que reciben, su edad, así como situaciones que vinculen a la matemática y la estadística con la economía, la educación, los negocios, los problemas sociales, la política, entre otros. Estas situaciones deben ser oportunas y que estimulen el interés por determinar y elegir el mejor representante de un conjunto de datos cuantitativo.
Lenguaje	Dato, conjunto de datos, adición, sustracción, multiplicación, división, media aritmética, mediana, primer cuartil, tercer cuartil, dato u observación atípica, el sesgo de una distribución de datos, el rango intercuartílico, el intervalo para determinar datos u observaciones atípicas, elección del representante del conjunto de datos.
Conceptos previos	Dato: Son cifras que se recogen, analizan y resumen de una variable cuantitativa, para su presentación e interpretación. Conjunto de datos: Se les llama así a todos los datos reunidos para un determinado estudio. Adición: Operación matemática que resulta al reunir dos o más cantidades en una sola varias cantidades. Se denota con el signo +. Sustracción: Operación matemática que consiste en restar una cantidad (el sustraendo) de otra (el minuendo) para averiguar la diferencia entre las dos. Se denota con el signo –. Multiplicación: Operación binaria que consiste en hacer corresponder a dos números reales llamados factores, un tercer número real llamado producto. Se denota con el símbolo *x*. División: Operación matemática que consiste en repartir una cantidad fija en otra dada. La división se denota con el símbolo ÷ o con /. Media aritmética: Es un valor alrededor del cual los demás valores se distribuyen (o concentran), o sea, es un valor de referencia para el conjunto analizado. Una de las posibles interpretaciones para el valor medio de un conjunto de datos sería como punto de equilibrio para los

	valores de la distribución. Se denota por $(\bar{x})$. Mediana.- Es el valor que divide el conjunto ordenado de valores observados en otros dos conjuntos con el mismo número de elementos. Se denota por (Me). Primer cuartil: Denotado por Q_1, es el menor valor que es mayor que una cuarta parte de los datos; es decir, aquel valor de la variable que supera el 25% y es superado por el 75% de las observaciones. Tercer cuartil: Denotado por Q_3, es el menor valor que es mayor que tres cuartas partes de los datos, es decir aquel valor de la variable que supera el 75% y es superado por el 25% de las observaciones. Dato u observación atípica: Es una observación que es inusual en relación con el resto de datos; en otras palabras, una observación atípica no sigue el patrón del resto de los datos. Son observaciones que son sospechosas y que requieren un análisis cuidadoso. Rango intercuartílico: Es una medida de dispersión. Su valor se obtiene como la diferencia del tercer cuartil (Q_3) menos el primer cuartil (Q_1), definido por la expresión RIC = $Q_3 - Q_1$. (Perú, 2006, 54) Sesgo: **S**e denomina así a la asimetría que presenta una distribución de frecuencias. Puede ser sesgo negativo o a la izquierda y sesgo positivo o a la derecha (Perú, 2006, 57)
Procedimientos, técnicas,…	I. Para determinar la media aritmética, se suman todos los datos y se divide entre el total de datos. II. Para determinar la mediana, se ordenan los datos, pueden darse dos casos: • Si el número de datos es impar, la mediana está representada por el valor central. • Si el número de datos es par, la mediana está representada por la media aritmética de los dos valores centrales. *III.* El valor del primer cuartil se obtiene calculando la mediana del 50% de los datos inferiores a la mediana o Q_2. IV. El valor del tercer cuartil se obtiene calculando la mediana del

	50% de los datos superiores a la mediana o Q_2. V. La determinación de datos u observaciones atípicas es el proceso mediante el cual usando los cuartiles primero y tercero, se determina un intervalo que servirá para reconocer si alguno de los datos forman o no parte del mismo. VI. El intervalo para determinar datos u observaciones atípicas es un intervalo cuyos límites inferior y superior están formados por las expresiones ($Q_1 – D$) y ($Q_3 + D$) respectivamente, donde: Q_1 es el primer cuartil. Q_3 es el tercer cuartil. $D = 1{,}5(Q_3 - Q_1)$
Proposiciones, propiedades, teoremas, etc.	• Para calcular la media aritmética, intervienen todos los valores de los datos. • La media aritmética se ve afectada por valores muy grandes o muy pequeños, mientras que la mediana no. • Si algún dato no pertenece al intervalo para determinar datos u observaciones atípicas [$Q_1 – D; Q_3 + D$], entonces es atípico, que genera sesgo en la distribución de los datos, haciéndola asimétrica. • En distribuciones asimétricas, el mejor representante de los datos es la mediana. • Si en el conjunto de datos no se observa la presencia de datos u observaciones atípicas, la distribución es simétrica y el mejor representante de los datos es la media aritmética.
Argumentos	Tesis: Todos los datos del conjunto pertenecen al intervalo [$Q_1 – D; Q_3 + D$], por tanto no existen datos u observaciones atípicas. Justificación: Es verdad porque un valor para que sea considerado dato u observación atípica, no debe pertenecer al intervalo [$Q_1 – D; Q_3 + D$]. Tesis: La situación presenta información de un conjunto de datos

	numéricos y no existe sesgo, por tanto la elección del mejor representante recae en la media aritmética. Justificación: En distribuciones simétricas, la media aritmética es la medida más representativa, porque es un valor central igual o cercano a la mediana. Tesis: La situación presenta información de un conjunto de datos numéricos y presenta datos u observaciones atípicas que sesgan la información, por lo tanto la elección del mejor representante recae en la mediana. Justificación: La distribución de datos es asimétrica y la mediana es la medida más representativa, porque considera igual cantidad de elementos hacia arriba y hacia abajo de ella.

2.3.8. Idoneidad de procesos de enseñanza aprendizaje de la matemática.

Aquí señalamos las pautas de análisis y valoración de la idoneidad de procesos de enseñanza y aprendizaje de la matemática, que según Godino (2011) va de la enseñanza. (p.1) pueden ser el punto de partida de una teoría de la instrucción matemática orientada hacia la mejora progresiva.

Al referirse a los indicadores de idoneidad señala que:

> El EOS proporciona herramientas para hacer operativa la noción de idoneidad de las configuraciones y trayectorias didácticas en que se puede descomponer un proceso de estudio matemático. […] presentamos algunos indicadores de las distintas idoneidades parciales y de las interacciones entre las mismas, los cuales pueden servir de pauta o guía para el diseño y valoración de acciones formativas planificadas o efectivamente implementadas. (Godino et al. 2006, p.8)

La valoración de la idoneidad epistémica la haremos teniendo en cuenta, según Wilhelmi, Bencomo y Godino (2003), "la representatividad de los significados implementados" para la enseñanza del uso de la media aritmética y la mediana como mejor representante del conjunto de datos, en correspondencia al significado de referencia. Otro aspecto importante que demanda la propuesta didáctica a priori, es establecer las condiciones que permitan superar los obstáculos cognitivos o epistemológicos de los estudiantes, que según Bachelar (1976), son cinco: i) la experiencia básica o conocimientos previos, ii) el obstáculo verbal, iii) el peligro de la explicación por la utilidad, iv) el conocimiento general y v) el obstáculo animista.

De acuerdo con Wilhelmi, Bencomo y Godino (2003), "Las nociones, proposiciones, lenguaje, argumentos, acciones y problemas, consideradas como entidades constituyentes de los significados institucionales y personales, son los observables que permiten hacer operativos los criterios de idoneidad y, por lo tanto, valorar un proceso instruccional".

Nuestra propuesta didáctica, está basada en la aplicación de algunos criterios de idoneidad, y para ello hemos considerado tomar en cuenta las idoneidades epistémicas y cognitivas del EOS, que representan por un lado la labor del docente y por otro lado la labor del estudiante en el proceso de enseñanza aprendizaje.

Según Godino (2011), "Un programa formativo, o un proceso de estudio matemático, tiene mayor idoneidad epistémica en la medida en que los contenidos implementados (o pretendidos) representan bien a los contenidos de referencia".

La presente tabla, muestra los componentes y los indicadores más pertinentes a nuestra propuesta que han sido tomados del EOS, seguidamente describimos cada una de las acciones para su realización.

Tabla 17. Componentes e indicadores de idoneidad epistémica (matemática)

COMPONENTE	INDICADOR	DESCRIPTOR
Situaciones-Problemas	• Se presentan situaciones de contextualización, y aplicación.	• La secuencia de actividades (ver anexos 4, 5, 6 y 7), inician con situaciones problemáticas contextualizadas cercanas al estudiante, con el propósito de promover el aprendizaje y razonamiento teniendo en cuenta que el contexto y las circunstancias sociales interactúan con las características individuales. • Las situaciones servirán para aplicar estrategias al determinar datos u observaciones atípicas.

		• Cada situación implica analizar los datos para tomar la decisión de elegir la medida estadística más representativa del conjunto.
Lenguajes	• Uso de diferentes modos de expresión matemática verbal y gráfica. • Nivel del lenguaje adecuado a los estudiantes a que se dirige. • Se proponen situaciones de elección de una medida estadística.	• Debe existir información de las situaciones problema presentadas como enunciados verbales y en tablas. • Se emplea un lenguaje acorde a la edad de los estudiantes, teniendo en cuenta no usar expresiones que puedan confundirlos, ni términos que no comprendan. • La elección del mejor representante de un conjunto de datos, es una decisión que realiza el estudiante después de haber analizado los datos y aplicado propiedades estadísticas para argumentar con fundamentos válidos la medida elegida. Previamente deben hacer cálculos matemáticos. • El uso del software Geogebra, para representar el conjunto de datos, la caja y la media aritmética también son suficientes para decidir la idoneidad de cualquiera de las medidas, sin necesidad de cálculos.
Reglas (Definiciones, proposiciones, procedimientos)	• Las definiciones y procedimientos son claros y correctos, y están adaptados al nivel educativo al que se dirigen.	• Los procedimientos de cálculo de la media aritmética, la mediana, los cuartiles y el intervalo para determinar datos u observaciones atípicas, son sencillos y hacen uso de operaciones aritméticas básicas. Las actividades consideran un conjunto de datos

		simples y además la enseñanza del uso es una actividad que se realiza después que los estudiantes ya estudiaron los algoritmos de cálculo de estas medidas, sin embargo, no estaría demás hacer una revisión ligera.
	• Se presentan los enunciados y procedimientos fundamentales del tema para el nivel educativo dado.	• Los enunciados de los problemas de cada una de las actividades, se presentan mediante situaciones contextualizadas que contienen información cuantitativa. Cada una de las actividades tiene una intencionalidad pedagógica sobre el uso de la media aritmética y la mediana como representante de un conjunto de datos. Las actividades inician con la construcción del intervalo para reconocer datos u observaciones atípicas y las otras presentan situaciones para la toma de decisiones al elegir el representante del conjunto, previo análisis de los datos.
Argumentos	• Las explicaciones, son adecuadas al nivel educativo a que se dirigen. • Se promueven situaciones donde el alumno tenga que argumentar.	• Las explicaciones de cada situación se realiza teniendo en cuenta el lenguaje y nivel de los estudiantes. • Los estudiantes deben justificar la decisión que tomaron al momento de elegir el mejor representante del conjunto de datos.

Relaciones	• Los objetos matemáticos (problemas, definiciones, proposiciones, etc.) se relacionan y conectan entre sí. • Se identifican y articulan los diversos significados de los objetos que intervienen en las prácticas matemáticas.	• La mediana, los cuartiles y el intervalo para determinar datos u observaciones atípicas están relacionados entre sí, para determinar datos u observaciones atípicas, mientras que la aplicación de propiedades estadísticas de la media aritmética y la mediana, se relacionan para tomar la decisión de elegir el mejor representante del conjunto de datos. • Para elegir el representante del conjunto de datos, previamente se construye el intervalo para reconocer datos u observaciones atípicas. Aquí se articulan los significados de los objetos mediana, cuartiles, dato atípico, rango intercuartílico, sesgo, intervalo, límites y operaciones aritméticas básicas. El siguiente procedimiento para elegir el representante del conjunto de datos, articula los algoritmos de cálculo de la media aritmética y la mediana, las propiedades estadísticas y la simetría o asimetría de la distribución.

Tabla 18. Componentes e indicadores de idoneidad cognitiva

COMPONENTE	INDICADOR	DESCRIPTOR
Conocimientos previos (se tienen en cuenta los mismos elementos que para la	• Los estudiantes tienen los conocimientos previos necesarios para el estudio del tema	• La media aritmética y la mediana son contenidos que los estudiantes ya han aprendido. (Ver mapa de progreso de Estadística y

idoneidad epistémica)	(bien se han estudiado anteriormente o el profesor planifica su estudio). • Los contenidos pretendidos se pueden alcanzar (tienen una dificultad manejable) en sus diversas componentes.	Probabilidad) • Los cuartiles y el sesgo serán planificados por el docente a modo un repaso breve en la pizarra sobre los algoritmos de cálculo. • La determinación de datos u observaciones atípicas por medio de un intervalo tiene una dificultad manejable, pues aparte de saber el algoritmo de cálculo de los cuartiles, implica también el uso de la adición, sustracción y multiplicación de números racionales.
Adaptaciones curriculares a las diferencias individuales	• Se incluyen actividades de ampliación y de refuerzo. • Se promueve el acceso y el logro de todos los estudiantes.	• La tercera actividad de la secuencia didáctica servirá para reforzar en los estudiantes el uso de la media aritmética y la mediana en un conjunto de datos, desarrollando en dúos una práctica calificada, antes de una evaluación. • La secuencia de actividades da los medios pero no garantiza que todos los estudiantes logren los aprendizajes esperados. Esto depende de otros factores que influyen en el rendimiento académico de los estudiantes. De la misma manera se debe comprender que la elección del representante del conjunto de

		datos no es una tarea simple.
Aprendizaje: Se tienen en cuenta los mismos elementos que para la idoneidad epistémica)	• Los diversos modos de evaluación indican que los estudiantes logran la apropiación de los conocimientos, comprensiones y competencias pretendidas: -Comprensión conceptual y proposicional. -Competencia comunicativa y argumentativa. -Fluencia procedimental. -Comprensión situacional. -Competencia metacognitiva. • La evaluación tiene en cuenta distintos niveles de comprensión y competencia. • Los resultados de las evaluaciones se difunden y usan para tomar decisiones.	• Se evalúa mediante un problema con similares características de lo trabajado en clase y en la práctica en dúos. • Se tienen en cuenta los ritmos y estilos de aprendizaje. La evaluación es para todo el grupo de la clase. • Los algoritmos de cálculo, la determinación de datos u observaciones atípicas y la elección del mejor representante del conjunto de datos, son indicadores que evidencian el logro del objetivo de esta actividad. • Los resultados se utilizan como motivación para seguir aprendiendo. Cada estudiante evalúa su desempeño (Autoevaluación).

2.4. METODOLOGÍA DE LA INVESTIGACIÓN

Para efectos de realizar el estudio, hemos planificado organizar el trabajo, teniendo en cuenta a Latorre y Cols (1996), quienes proponen seis fases a seguir en el proceso de la investigación cualitativa:

Tabla 19. Fases de la investigación cualitativa

FASES	ACCIONES
Fase exploratoria/de reflexión	Identifico el problema.
	Planteo el problema de investigación.
	Planteo las cuestiones de investigación.
	Realizo la revisión de antecedentes
	Formulo hipótesis de investigación.
Fase de planificación	Determino la Institución Educativa.
	Delimito y ajusto el problema y pregunta de investigación.
	Reformulo y ajusto los objetivos de investigación.
	Selecciono textos de material a analizar.
	Elijo el marco teórico para la investigación.
	Selecciono la estrategia a seguir.
Fase de entrada en el escenario	Selecciono el público objetivo.
	Defino el papel del investigador.
Fase de recogida y análisis de la información	Recolecto datos.
	Describo y analizo la información.
Fase de retirada del escenario	Finalizo la recogida de la información de los libros de texto.
	Realizo un análisis comparativo y reflexivo de la información y se dieron los resultados.
Fase de elaboración del informe	Redacto las conclusiones de la investigación.
	Elaboro la redacción y retiramiento de dicho informe.

Teniendo en cuenta estas fases, describimos los aspectos que se relacionan con nuestro trabajo de investigación.

Tabla 20. Acciones realizadas

FASE	ACCIONES
EXPLORATORIA	Se determinó el problema de nuestra investigación y se estableció la pregunta de investigación. Para ello fue necesario declarar la importancia de la enseñanza del uso de la media aritmética y la mediana, presentar los trabajos realizados por otros investigadores que tienen relación con el nuestro y considerar que ésta investigación permitirá contribuir a una práctica educativa que mejore los procesos de enseñanza en el aula y aporte estrategias a los docentes en el desarrollo del contenido.
DE PLANIFICACIÓN	Para el **primer objetivo específico**, identificar los errores que cometen los estudiantes al determinar y usar la media aritmética y la mediana, para evitarlos, realizamos los siguientes pasos: Paso 1: Analizar investigaciones y artículos relacionados con las medidas de tendencia central. Paso 2: Recopilar los errores, ya declarados en investigaciones y artículos, que cometen los estudiantes respecto al uso de las medidas de tendencia central en trabajos que ya han sido realizados por Mayén (2009) con estudiantes mexicanos, a quienes se les aplicó unos cuestionarios y de cuyas respuestas un buen porcentaje de errores son coincidentes y la aparición de otros errores considerados como nuevos. (ver pág. 68)
	Para el **segundo objetivo específico**, analizar las preguntas comunes que aparecen en los textos, artículos y otros medios referentes a la media aritmética y la mediana, para determinar los significados o usos, se realizaron los siguientes pasos: Paso 1: Recopilar información donde aparecen los tipos de problemas pertinentes al uso de las medidas de tendencia central, en textos, trabajos de investigación estadística, artículos y otros medios. Paso 2: Se revisaron cuestionarios aplicados a estudiantes españoles, angoleños y mexicanos, del mismo modo se observó un video en la web, donde se analiza y

	discute el representante de los sueldos de los trabajadores de una empresa. Esto sirvió como modelo para la creación de situaciones problemáticas en nuestra propuesta. Paso 3: Establecer las situaciones de contexto real que servirán de apoyo para proponer casos que motiven la discusión y análisis de la determinación del uso de la media aritmética y la mediana en un conjunto de datos.
	Para describir el objeto matemático de nuestra investigación "El uso de las medidas de tendencia central la media aritmética o la mediana", los conceptos, procedimientos y propiedades han sido extraídos del libro de Córdova (2003), Sweeney (2008) y Novaes (2011). Asimismo las reglas de uso, ventajas y desventajas de estas medidas fueron tomadas de la investigación de Pinzón (2012). De estos trabajos, hemos considerado los aspectos más importantes de estas medidas que permitan determinar la medida más representativa de un conjunto de datos; del mismo modo hemos elaborado la configuración epistémica de referencia de la media aritmética y la mediana, como una herramienta fundamental de nuestra propuesta, y para finalizar esta etapa, la metodología que usaremos es de tipo cualitativa descriptiva con aspectos interpretativos.
	Se ha creído conveniente utilizar algunos aspectos del Enfoque Ontosemiótico de la Cognición e Instrucción Matemática como la teoría que guiará nuestra investigación (configuraciones epistémicas y aplicación de algunos criterios de idoneidad epistémica y cognitiva), por considerarla la más pertinente y que nos servirá de soporte.
RECOGIDA DE LA INFORMACIÓN	Inicialmente se aplicó un cuestionario (ver anexo 1, 2 y 3) a una muestra piloto de 10 docentes de la especialidad de matemática del nivel secundaria pertenecientes a diferentes regiones del país, que realizan estudios de maestría con mención en enseñanza de las matemáticas en la Pontificia Universidad Católica del Perú (PUCP). En este cuestionario se busca recoger información acerca de los conocimientos de los docentes sobre el uso de la media aritmética y la mediana en un conjunto de datos. Luego de algunos reajustes, este cuestionario fue aplicado a docentes de la especialidad de matemática en ejercicio que laboran en instituciones educativas públicas y privadas de las regiones de Tumbes (28) y Arequipa (6).

ANÁLISIS DE LA INFORMACIÓN	Una vez aplicado el cuestionario, se procedió a la descripción y análisis cualitativo de la información. Se revisó cada una de las respuestas, se identificaron respuestas similares a un mismo ítem y se analizaron las respuestas de 6 docentes. Se esperaba que los docentes recomienden usar la mediana por la existencia de datos u observaciones atípicas que generan sesgo en la información en las preguntas 1 y 2 del cuestionario, sin embargo 2 docentes de este grupo recomiendan usar la media aritmética con argumentos que no son adecuados; otros 2 recomiendan usar la mediana sin argumentar la representatividad por la presencia de datos atípicos y finalmente los otros 2 docentes, recomiendan usar la media aritmética en el problema 1 y la mediana en el problema 2 con argumentos inapropiados frente a situaciones de respuesta semejante.
ELABORACIÓN DE LA SECUENCIA DE ACTIVIDADES E INFORME	Para el **tercer objetivo específico**, elaborar la secuencia de actividades, se realizaron los siguientes pasos: Paso 1: Determinar el total de actividades y el tiempo para cada una de ellas. Paso 2: Plantear los objetivos de cada actividad y los errores que levanta. Paso 3: En la primera actividad se construye el intervalo para reconocer datos u observaciones atípicas. Paso 4: En la segunda actividad se analizan los datos y se aplican propiedades estadísticas para elegir el mejor representante del conjunto de datos. Paso 5: La tercera actividad es una evaluación en dúos que recoge lo que el estudiante conoce acerca de la elección del representante y permite realizar los ajustes necesarios. Paso 6: la cuarta y última actividad se evalúa la capacidad de los estudiantes para tomar decisiones y argumentar la elección de la media aritmética o la mediana como representante del conjunto de datos. Para presentar los resultados de la investigación cualitativa, situaremos al lector en el escenario de distintos problemas contextualizados, donde se aprecie la necesidad de elegir al mejor representante del conjunto de datos, especialmente en casos donde el conjunto de datos tiene inmerso datos u observaciones atípicas (datos muy grandes o muy pequeños) que sesgan la información.

CAPÍTULO III: DISEÑO DE LA SECUENCIA DIDÁCTICA

En este capítulo se presentan las características de la propuesta didáctica, señalando el número de actividades y la duración de cada una de ellas. Del mismo modo se establecen los requisitos mínimos de los estudiantes para desarrollar la enseñanza del uso de la media aritmética y la mediana en un conjunto de datos, los objetivos y desempeños de los estudiantes. Adicionalmente muestra los errores señalados por Mayén (2009) y una descripción de cada una de las actividades.

Destacamos la importancia de estas actividades por diferentes razones:
Primero porque las actividades relacionadas a las medidas de tendencia central son trabajadas desde el punto de vista estadístico y no matemático, dando más interés a la aplicación de propiedades estadísticas de la media aritmética y la mediana, teniendo como apoyo los algoritmos de cálculo de estas medidas, así como de los cuartiles. Con estas actividades creemos que el docente tendrá herramientas que le permitan enseñar el uso de la media aritmética y la mediana en un conjunto de datos.
Segundo porque los docentes en su gran mayoría desconocen procedimientos para determinar de modo seguro la presencia de datos u observaciones atípicas que sesgan la información. En muchas ocasiones hay duda en reconocer si un dato es atípico o no. Esto sucede porque en la formación del docente, no incluye el aprendizaje del uso, y además porque los currículos anteriores tampoco tenían en cuenta como contenido la enseñanza del uso de las medidas de tendencia central en estadística. Sin embargo esto no debe ser una justificación para que el docente no se haya preocupado por conocerlo.
Tercero porque actualmente el Ministerio de Educación con el propósito de realizar mejoras en el currículo, incluye en el DCN la enseñanza del uso del mejor representante de un conjunto de datos.

Consideramos que para determinar el mejor representante de los datos, es insuficiente comparar la media aritmética con la mediana para tomar una decisión, sin embargo es un primer paso. El siguiente procedimiento consiste en hallar un método o estrategia que permita determinar con seguridad la presencia de datos u observaciones atípicas que sesgan la información, debido a la presencia de algunos datos que no aseguran a simple vista que sean observaciones atípicas, y como investigadores preocupados por resolver esta situación, encontramos que según Sweeney (2008) existe un método llamado “Diagrama de caja y bigotes" para determinar datos u observaciones atípicas (ver pág. 40). Hemos simplificado este método construyendo el intervalo $[Q_1 – D; Q_3 + D]$ (ver pág. 41).

La propuesta didáctica de la enseñanza del uso de la media aritmética y la mediana a trabajar en el aula basada en situaciones problemáticas contextualizadas, tiene la intencionalidad de brindarle a los docentes del área de matemática una secuencia de actividades que les permitan:

i) Conocer los errores y dificultades que ponen de manifiesto los estudiantes en el aprendizaje de las medidas de tendencia central para evitarlos.
ii) Determinar datos u observaciones atípicas que sesgan la información, aplicando estrategias.
iii) Considerar todos los datos del conjunto sin ninguna exclusión.
iv) Aplicar las propiedades estadísticas de la media aritmética y la mediana.
v) Elegir la medida más adecuada como el mejor representante de los datos y justificar su elección.

3.1. Características de la propuesta

La presente propuesta didáctica está dirigida a los docentes del área de matemática para ser aplicada con estudiantes del nivel secundario. Tiene como propósito contribuir al aprendizaje del uso de la media aritmética y la mediana en cuatro actividades que hacen un total de 4 horas. Ver tabla:

Tabla 21. Distribución de las actividades

Construcción del intervalo para reconocer datos u observaciones atípicas	1 h
Elección del representante del conjunto de datos	1 h
Práctica calificada: Elegimos analizando los datos	1h
Evaluación: Evaluando la elección del mejor representante	1 h

Esta propuesta didáctica tiene como prerrequisito los siguientes conocimientos de los estudiantes antes de su realización:

- Determinación de la media aritmética.
- Cálculo de la mediana.
- Determinación de cuartiles.

A continuación describimos las actividades.

La propuesta didáctica plantea situaciones problemáticas contextualizadas, que promuevan en los estudiantes la necesidad de determinar la presencia de datos u observaciones atípicas que generan sesgo, aplicar propiedades estadísticas, tomar decisiones y justificar su decisión con argumentos válidos.

En ella se trabajará el procedimiento para determinar datos atípicos en un conjunto de datos, y aplicar propiedades estadísticas de la media aritmética y la mediana que permitan elegir a una de ellas como el mejor representante del conjunto de datos.

Esta propuesta busca superar los errores identificados por otros investigadores y que cometen los estudiantes respecto a la comprensión y aplicación de propiedades estadísticas.

Los ***errores*** a considerar tomados de Mayen (2009) son:

a) El alumno considera a la media aritmética o la mediana como el mejor representante del conjunto de datos, sin apoyar su conclusión con argumentos o razones.

b) Aplicación indebida y confusión de propiedades estadísticas, haciendo generalizaciones abusivas de ellas, por ejemplo, la media es menos resistente que la mediana a la presencia de datos u observaciones atípicas.

c) El suponer que cualquier promedio es representativo.

d) Aplicación incorrecta en la propiedad de los promedios de coincidir en distribuciones simétricas.

e) Escasez de estrategias para identificar valores atípicos.

Estos errores suceden por desconocimiento del tema, lo que trae como consecuencia la utilización de argumentos inadecuados.

De acuerdo con Chan (2009):

> "Los errores y dificultades que se han reportado con anterioridad muestran que no se está llevando a cabo una culturización estadística hacia la sociedad, porque, uno de los objetivos que persigue la educación estadística, es la capacidad de argumentación por parte de los

> estudiantes. […] La causa de esto podría ser el miedo, la inseguridad, la falta de conocimiento y preparación, los cuales se deben de combatir en el aula de clases mediante la participación activa del estudiante y su exposición con situaciones reales el cual le lleve a tener que recolectar, analizar, interpretar, dar a conocer y validar el trabajo que se encuentra realizando”. (p. 60)

En nuestra propuesta se considera que la participación de los estudiantes debe ser activa en la solución de las situaciones planteadas, su rol protagónico en la construcción de su aprendizaje, guiado por el docente, mediante preguntas relacionadas a la elección del representante del conjunto de datos, aplicando propiedades estadísticas de la media aritmética y la mediana, teniendo en cuenta la presencia o no de datos u observaciones atípicas en muestras de variables cuantitativas y justificando las razones que condujeron a tomar una decisión.

3.2. Objetivos de la propuesta

- Facilitar al docente del área de matemática un conjunto de estrategias metodológicas y procedimentales que le permitan desarrollar sesiones de aprendizaje orientadas a elegir entre el uso de la media aritmética o la mediana como el mejor representante de un conjunto de datos.
- Desarrollar sesiones de aprendizaje para que los estudiantes aprendan a elegir el representante de un conjunto de datos justificando su decisión con argumentos válidos.
- Que el alumno realice adecuadamente el cálculo de la media aritmética, la mediana y los cuartiles en datos simples, de ser necesario se puede usar calculadoras o software.
- Que el alumno reconozca que no siempre la media aritmética es el representante de un conjunto de datos, sino que aplicando estrategias se debe evaluar la existencia de datos u observaciones atípicas que sesgan la información.
- Que el alumno pueda construir el intervalo que permite reconocer la existencia de datos atípicos.
- Que el alumno tome la decisión de elegir al mejor representante del conjunto de datos, aplicando las propiedades estadísticas de la media aritmética y la mediana.

3.3. Descripción de las actividades

La secuencia de actividades comprende 4 sesiones (2 clases guiadas por el docente, 1 práctica calificada y 1 evaluación), la primera actividad se elabora para que los estudiantes aprendan a reconocer datos u observaciones atípicas y las otras tres tienen como objetivo lograr que los estudiantes estén en la capacidad de elegir al mejor representante de un conjunto de datos, con la variante que una se desarrolla guiados por el docente, otra en dúos y la ultima de manera individual.

ACTIVIDAD 01: SESIÓN DE APRENDIZAJE

- Nombre de la actividad: "Construcción del intervalo para reconocer datos u observaciones atípicas"
- Lugar: El salón de clases.
- Duración: Una hora pedagógica (45 min).

Esta actividad está diseñada con el propósito de construir el intervalo para reconocer datos u observaciones atípicas. Se presenta una situación problemática que se analiza en un primer momento y luego se analiza por segunda vez al cambiar uno de los datos.

La intencionalidad pedagógica de esta primera actividad es que el estudiante reconozca la construcción de un intervalo como estrategia para determinar datos u observaciones atípicas que generan sesgo en la distribución.

Del mismo modo esta actividad evidencia que en el primer caso no existen datos u observaciones atípicas, mientras que en el segundo caso sí.

Esta actividad está diseñada para evitar que los estudiantes cometan errores relacionados a:

- Suponer que cualquier promedio es representativo (error c).
- La escasez de estrategias para identificar valores atípicos. (Error e), que fueron señalados anteriormente por otros investigadores.

Los detalles del desarrollo de esta actividad se muestran en el anexo N° 4.

El docente realiza las siguientes acciones:

- ➢ Recojo de los saberes previos de los estudiantes sobre los algoritmos de cálculo de la mediana, los cuartiles y la relación que existe entre la mediana y el segundo cuartil.
- ➢ Como la situación problemática propuesta solicita que el estudiante determine la existencia o no de datos u observaciones atípicas, el docente puede preguntar ¿qué

entienden por dato atípico?, como una manera de ir acercando al estudiante a comprender que se trata de datos que no son comunes a los demás.

- También puede preguntar a los estudiantes: ¿conocen alguna estrategia para determinar datos u observaciones atípicas?, ¿qué es un intervalo?, ¿qué es el rango intercuartílico?
- Se presentan los datos de la situación problemática y se pide a los estudiantes que observen los datos detenidamente, tratando de reconocer datos que están muy distantes del resto, de ser así, por intuición diremos que son datos u observaciones atípicas.
- El docente con la participación de los estudiantes, define la mediana, los cuartiles 1 y 3, dato u observación atípica, el rango intercuartilico y el sesgo. (ver tabla 20, pág. 75)
- Para ello los estudiantes, guiados por el docente, ordenan los datos de menor a mayor y localizan la mediana aplicando su algoritmo de cálculo.
- El docente indica que la mediana equivale al segundo cuartil $Me=Q_2$.
- Seguidamente se localizan los cuartiles 1 y 3, aplicando los algoritmos de cálculo.
- A partir de estas medidas ya localizadas en el conjunto de datos, el docente muestra el intervalo que servirá para determinar datos u observaciones atípicas, el mismo que tiene la siguiente forma: $[Q_1 - 1{,}5\ RIC;\ Q_3 + 1{,}5\ RIC]$.
- El docente procede a construir el intervalo que llamamos "Intervalo para determinar datos u observaciones atípicas", que representa al "diagrama de caja y bigotes", utilizando la mediana, los cuartiles 1 y 3 y el rango intercuartilico.
- Llamemos D a la expresión 1,5 RIC, que representa 1,5 veces el rango intercuartílico, $RIC = (Q_3 - Q_1)$, que finalmente será $D = 1{,}5\ (Q_3 - Q_1)$.
- Una vez construido el intervalo se analizan los datos, y al observar que todos pertenecen al intervalo, se concluye que no existen datos u observaciones atípicas (ver pág. 77).
- La segunda parte de esta actividad consiste en cambiar uno de los datos.
- El docente conjuntamente con los estudiantes construye el "Intervalo para determinar datos u observaciones atípicas".
- Luego procede al análisis de los datos, y al observar que uno de los datos no pertenece al intervalo, se concluye que existen datos u observaciones atípicas (ver pág. 77).
- Las afirmaciones anteriormente expuestas se justifican porque según Sweeney (2008), el intervalo para determinar datos u observaciones atípicas es una manera de realizar el análisis exploratorio de datos.

ACTIVIDAD 02: SESIÓN DE APRENDIZAJE

- Nombre de la actividad: "Eligiendo el mejor representante"
- Lugar: El salón de clases.
- Duración: Una hora pedagógica (45 min).

Esta actividad está diseñada para elegir al mejor representante del conjunto de datos (la media aritmética o la mediana), se plantea una situación problemática de contexto real, relacionada con los sueldos de los trabajadores de una institución educativa, se realiza el análisis exploratorio de los datos para reconocer datos u observaciones atípicas.

La intencionalidad pedagógica de esta segunda actividad está orientada a que el estudiante elija a la media aritmética o la mediana como representante del conjunto de datos, analizándolos para determinar la existencia de datos u observaciones atípicas que generan sesgo. También los estudiantes estarán en la capacidad de argumentar las razones de su elección, sobre la base de la aplicación de propiedades estadísticas y reglas de uso de la media aritmética y la mediana. (Ver anexo N° 5).

Esta actividad es diseñada para evitar que los estudiantes cometan los errores correspondientes a:

- Considerar a la media aritmética o la mediana como el mejor representante del conjunto de datos, sin apoyar su conclusión con argumentos o razones. (error a)
- Aplicación indebida y confusión de propiedades estadísticas, haciendo generalizaciones abusivas de ellas, la media es menos resistente que la mediana a la presencia de valores atípicos. (error b).
- Aplicación incorrecta en la propiedad de los promedios de coincidir en distribuciones simétricas. (error d).

Los detalles del desarrollo de esta actividad se muestran en el anexo N° 5.

El docente realiza las siguientes acciones:

- Recojo de los saberes previos de media aritmética y el procedimiento para construir el intervalo de determinación de datos u observaciones atípicas.
- El docente con la participación de los estudiantes, define la media aritmética y el sesgo. (ver tabla 21, pág. 78)

- Se calcula la media aritmética y la mediana de los datos, teniendo en cuenta a todos los valores de los datos y reconociendo que la media aritmética y la mediana, pueden ser representantes de un conjunto de datos.
- Se cuestiona ¿qué sucede si la media aritmética no es igual que la mediana? ¿esto se debe a que hay datos u observaciones atípicas que hace que sean diferentes?
- El docente con la participación de los estudiantes, construye el intervalo para determinar datos u observaciones atípicas y procede a realizar el análisis exploratorio de los datos, teniendo en cuenta las siguientes consideraciones:
    - Si todos los datos pertenecen al intervalo, entonces no hay datos u observaciones atípicas, y el mejor representante de los datos es la media aritmética. Esto se justifica por la regla de uso de la media aritmética que es representativa de los datos en distribuciones simétricas.
    - Si algún dato no pertenece al intervalo, entonces existen datos u observaciones atípicas que sesgan la información, en este caso, el mejor representante del conjunto de datos es la mediana. Esto se justifica por la regla de uso de la mediana que es representativa en distribuciones sesgadas y por la propiedad que no se ve afectada por algún valor grande o pequeño, mientras que la media aritmética sí.

ACTIVIDAD 03: PRÁCTICA CALIFICADA

- Nombre de la actividad : “Elegimos analizando los datos”
- Lugar: El salón de clases.
- Duración: Una hora pedagógica (45 min).

Esta actividad se resuelve en dúos y plantea dos situaciones problemáticas contextualizadas, la primera situación relacionada con las propinas que reciben a diario los estudiantes y la segunda situación relacionada con las edades del grupo de estudiantes. Se sugiere que los dúos estén formados por estudiantes de nivel alto o medio con estudiantes de nivel bajo, para que la efectividad del aprendizaje sea mejor.

La intencionalidad pedagógica de esta tercera actividad está orientada a que el estudiante refuerce los conocimientos para elegir el mejor representante del conjunto de datos, realice el análisis de los datos, aplique propiedades estadísticas y reglas de uso de la media aritmética y la mediana, y argumente las razones de su elección. El estudiante elegirá a la mediana como el mejor representante del conjunto de datos, en la primera situación problemática y a la media aritmética en la segunda situación.

Esta actividad es diseñada para evitar que los estudiantes cometan los 5 errores señalados por Mayén (2009). (Ver pág. 68 y anexo N° 6).

Los estudiantes deben realizar las siguientes acciones:

- Aplicar el algoritmo de cálculo de la media aritmética.
- Aplicar el algoritmo de cálculo de la mediana.
- Comparar la media aritmética y la mediana. Si son iguales la media aritmética es el representante de los datos, pero si son diferentes, se debe realizar el análisis exploratorio de los datos.
- Calcular los cuartiles 1 y 3.
- Construir el intervalo $[Q_1 - D; Q_3 + D]$, basado en el primer cuartil, tercer cuartil y la expresión $D = 1{,}5\,[Q_3 - Q_1]$.
- Una vez construido el intervalo, los estudiantes proceden a realizar el análisis exploratorio de los datos que consiste en determinar la existencia o no de datos u observaciones atípicas que sesgan la información.

- ➢ Finalmente, los estudiantes aplican las reglas de uso y propiedades estadísticas de la media aritmética y la mediana, para tomar la decisión de elegir el representante del conjunto de datos y justificar su elección.

ACTIVIDAD 04: EVALUACIÓN ESCRITA

- Nombre : “Evaluando la elección del mejor representante”
- Lugar: El salón de clase.
- Duración: Una hora pedagógica (45 min).

Esta actividad es individual y plantea una situación problemática contextualizada relacionada con las calificaciones obtenidas por los estudiantes en el área de matemática. Los estudiantes deben responder que medida estadística (la media aritmética o la mediana) es el mejor representante del conjunto de datos.

Esta actividad tiene la intención pedagógica de verificar el aprendizaje de los estudiantes al elegir a la mediana como el mejor representante del conjunto de datos, usando para ello estrategias para determinar datos u observaciones atípicas y la aplicación de propiedades estadísticas y reglas de uso de la media aritmética y la mediana, como sustento para justificar su decisión. (Ver anexo N° 7)

Para ello deben realizar las siguientes acciones:

- ➢ Aplicar el algoritmo de cálculo de la media aritmética.
- ➢ Aplicar el algoritmo de cálculo de la mediana.
- ➢ Comparar la media aritmética y la mediana. Si son iguales la media aritmética es el representante de los datos, pero si son diferentes, se debe realizar el análisis exploratorio de los datos.
- ➢ Calcular los cuartiles 1 y 3, aplicando su algoritmo de cálculo.
- ➢ Construir el intervalo para determinar la existencia o no de datos u observaciones atípicas, basado en el primer cuartil, tercer cuartil y 1,5 veces el rango intercuartílico.
- ➢ Analizar si en el conjunto de datos hay observaciones que pertenecen o no a al intervalo.
- ➢ Tomar la decisión de elegir el representante más idóneo del conjunto de datos, aplicando las propiedades estadísticas y las reglas de uso de la media aritmética y la mediana, teniendo en cuenta la existencia o no de datos u observaciones atípicas que generan sesgo en los datos.

- Justifican su decisión en base a la aplicación de propiedades y reglas de uso de la media aritmética y la mediana.

A continuación presentamos las configuraciones epistémicas de las actividades, que según el EOS, son redes de objetos institucionales, que se proponen como herramientas para describir los conocimientos matemáticos y sirven para orientar la mejora de un tema específico, en nuestro caso de la enseñanza del uso de la media aritmética y la mediana.

Las configuraciones epistémicas de las actividades 1 y 2 que mostramos, toman en cuenta los aspectos de la configuración epistémica de referencia. La diferencia entre ellas radica en que pretenden lograr objetivos distintos. No consideramos necesario elaborar las configuraciones epistémicas de las actividades 3 y 4 por la similitud con la actividad 2.

3.4. Configuraciones epistémicas de las actividades

Tabla 22. Configuración epistémica de la actividad 1.

Situaciones – Problema	El problema ha sido extraído de la realidad de los estudiantes teniendo en cuenta el contexto real, muestra información acerca de las calificaciones obtenidas por los estudiantes de segundo año de secundaria, en el área de matemática. Esta actividad tiene dos momentos para el análisis de los datos, uno con los datos normales y otro al cambiar un dato ante el reclamo de un estudiante. Se pide determinar la existencia de datos u observaciones atípicas.
Lenguaje	Dato, conjunto de datos, adición, sustracción, multiplicación, división, mediana, primer cuartil, tercer cuartil, dato u observación atípica, el sesgo de una distribución de datos, el rango intercuartílico, el intervalo para determinar datos u observaciones atípicas.
Conceptos o definiciones	<u>Dato:</u> Son cifras que se recogen, analizan y resumen de una variable cuantitativa, para su presentación e interpretación. <u>Conjunto de datos</u>: Se les llama así a todos los datos reunidos para un determinado estudio. <u>Adición:</u> Operación matemática que resulta al reunir dos o más

	cantidades en una sola varias cantidades. Se denota con el signo +. Sustracción: Operación matemática que consiste en restar una cantidad (el sustraendo) de otra (el minuendo) para averiguar la diferencia entre las dos. Se denota con el signo –. Multiplicación: Operación binaria que consiste en una abreviación de la suma repetida de un mismo número varias veces. Se denota con el símbolo x. División: Operación matemática que consiste en repartir una cantidad fija en otra dada. La división se denota con el símbolo ÷ o con /. Mediana.- Es el valor que divide el conjunto ordenado de valores observados en otros dos conjuntos con el mismo número de elementos. Se denota por (Me). Primer cuartil: Denotado por Q_1, es el menor valor que es mayor que una cuarta parte de los datos; es decir, aquel valor de la variable que supera el 25% y es superado por el 75% de las observaciones. Tercer cuartil: Denotado por Q_3, es el menor valor que es mayor que tres cuartas partes de los datos, es decir aquel valor de la variable que supera el 75% y es superado por el 25% de las observaciones. Dato u observación atípica: Es una observación que es inusual en relación con el resto de datos; en otras palabras, una observación atípica no sigue el patrón del resto de los datos. Son observaciones que son sospechosas y que requieren un análisis cuidadoso. Rango intercuartílico: Es una medida de dispersión. Su valor se obtiene como la diferencia del tercer cuartil (Q_3) menos el primer cuartil (Q_1), definido por la expresión RIC = Q_3 - Q_1. (Perú, 2006, 54) Sesgo: **S**e denomina así a la asimetría que presenta una distribución de frecuencias. Puede ser sesgo negativo o a la

	izquierda y sesgo positivo o a la derecha (Perú, 2006, 57)
Procedimientos, técnicas,…	I. Para determinar la mediana, se ordenan los datos, pueden darse dos casos: o Si el número de datos es impar, la mediana está representada por el valor central. o Si el número de datos es par, la mediana está representada por la media aritmética de los dos valores centrales. II. El valor del primer cuartil se obtiene calculando la mediana del 50% de los datos inferiores a la mediana o Q_2. III. El valor del tercer cuartil se obtiene calculando la mediana del 50% de los datos superiores a la mediana o Q_2. IV. La determinación de datos u observaciones atípicas es el proceso mediante el cual usando los cuartiles primero y tercero, se determina un intervalo que servirá para reconocer si alguno de los datos forman o no parte del mismo. V. El intervalo para determinar datos u observaciones atípicas es un intervalo cuyos límites inferior y superior están formados por las expresiones (Q_1 – D) y (Q_3 + D) respectivamente, donde: Q_1 es el primer cuartil. Q_3 es el tercer cuartil. $D = 1{,}5(Q_3 - Q_1)$.
Proposiciones, propiedades, teoremas, etc.	Una vez construido el intervalo para determinar datos u observaciones atípicas se procede de la siguiente manera: • Si algún dato no pertenece al intervalo [Q_1 – D; Q_3 + D] para determinar datos u observaciones atípicas, entonces es dato u observación atípica, que genera sesgo en la distribución de los datos, haciéndola asimétrica. • Si en el conjunto de datos no se observa la presencia de datos u observaciones atípicas, la distribución es simétrica

Argumentos	Tesis: Todos los datos del conjunto pertenecen al intervalo [Q_1 – D; Q_3 + D], por tanto no existen datos u observaciones atípicas. Justificación: Es verdad porque los datos del conjunto deben pertenecer al intervalo, en caso no pertenezcan o estén fuera del intervalo, serán datos u observaciones atípicas que sesgan la distribución de datos.

Tabla 23. Configuración epistémica de la actividad 2.

Situaciones – Problema	El problema es propio del entorno de los estudiantes relacionado con los sueldos de los trabajadores de una institución educativa privada. Se pide determinar el mejor representante del conjunto de datos.
Lenguaje	Dato, conjunto de datos, adición, sustracción, multiplicación, división, media aritmética, mediana, primer cuartil, tercer cuartil, dato u observación atípica, el sesgo de una distribución de datos, el rango intercuartílico, el intervalo para determinar datos u observaciones atípicas, elección del representante del conjunto de datos.
Conceptos o definiciones	Dato: Son cifras que se recogen, analizan y resumen de una variable cuantitativa, para su presentación e interpretación. Conjunto de datos: Se les llama así a todos los datos reunidos para un determinado estudio. Adición: Operación matemática que resulta al reunir dos o más cantidades en una sola varias cantidades. Se denota con el signo +. Sustracción: Operación matemática que consiste en restar una cantidad (el sustraendo) de otra (el minuendo) para averiguar la diferencia entre las dos. Se denota con el signo –. Multiplicación: Operación binaria que consiste en hacer corresponder a dos números reales llamados factores, un tercer

	número real llamado producto. Se denota con el símbolo *x*. División: Operación matemática que consiste en repartir una cantidad fija en otra dada. La división se denota con el símbolo ÷ o con /. Media aritmética: Es un valor alrededor del cual los demás valores se distribuyen (o concentran), o sea, es un valor de referencia para el conjunto analizado. Una de las posibles interpretaciones para el valor medio de un conjunto de datos sería como punto de equilibrio para los valores de la distribución. Se denota por $(\bar{x})$. Mediana**.-** Es el valor que divide el conjunto ordenado de valores observados en otros dos conjuntos con el mismo número de elementos. Se denota por (Me). Primer cuartil: Denotado por Q_1, es el menor valor que es mayor que una cuarta parte de los datos; es decir, aquel valor de la variable que supera el 25% y es superado por el 75% de las observaciones. Tercer cuartil: Denotado por Q_3, es el menor valor que es mayor que tres cuartas partes de los datos, es decir aquel valor de la variable que supera el 75% y es superado por el 25% de las observaciones. Dato u observación atípica: Es una observación que es inusual en relación con el resto de datos; en otras palabras, una observación atípica no sigue el patrón del resto de los datos. Son observaciones que son sospechosas y que requieren un análisis cuidadoso. Rango intercuartílico: Es una medida de dispersión. Su valor se obtiene como la diferencia del tercer cuartil (Q_3) menos el primer cuartil (Q_1), definido por la expresión RIC = Q_3 - Q_1. (Perú, 2006, 54) Sesgo: **S**e denomina así a la asimetría que presenta una distribución de frecuencias. Puede ser sesgo negativo o a la izquierda y sesgo positivo o a la derecha (Perú, 2006, 57)

Procedimientos, técnicas,…	I. Para determinar la media aritmética, se suman todos los datos y se divide entre el total de datos. II. Para determinar la mediana, se ordenan los datos, pueden darse dos casos: Si el número de datos es impar, la mediana está representada por el valor central. Si el número de datos es par, la mediana está representada por la media aritmética de los dos valores centrales. III. El valor del primer cuartil se obtiene calculando la mediana del 50% de los datos inferiores a la mediana o Q_2. IV. El valor del tercer cuartil se obtiene calculando la mediana del 50% de los datos superiores a la mediana o Q_2. V. La determinación de datos u observaciones atípicas es el proceso mediante el cual usando los cuartiles primero y tercero, se determina un intervalo que servirá para reconocer si alguno de los datos forman o no parte del mismo. VI. El intervalo para determinar datos u observaciones atípicas es un intervalo cuyos límites inferior y superior están formados por las expresiones (Q_1 – D) y (Q_3 + D) respectivamente, donde: Q_1 es el primer cuartil. Q_3 es el tercer cuartil. $D = 1{,}5(Q_3 - Q_1)$. $D = 1{,}5(Q_3 - Q_1)$.
Proposiciones, propiedades, teoremas, etc.	• Para calcular la media aritmética, intervienen todos los valores de los datos. • La media y la mediana pueden ser representante de un conjunto de datos. • La media aritmética se ve afectada por valores muy grandes o muy pequeños, mientras que la mediana no. • En presencia de datos u observaciones atípicas que generan sesgo en la distribución de los datos, el mejor representante de

	los datos es la mediana. • Si no se observa la presencia de datos u observaciones atípicas, la distribución es simétrica y el mejor representante de los datos es la media aritmética.
Argumentos	Tesis: La situación presenta información de un conjunto de datos numéricos y no existe sesgo, por tanto la elección del mejor representante recae en la media aritmética. Justificación: La distribución de datos es simétrica y la media aritmética es la medida más representativa, porque es un valor central igual o cercano a la mediana. Tesis: La situación presenta información de un conjunto de datos numéricos y presenta datos u observaciones atípicas que sesgan la información, por lo tanto la elección del mejor representante recae en la mediana. Justificación: Si el conjunto de datos presenta sesgo por la existencia de datos atípicos, la distribución es asimétrica y la mediana es la medida más representativa del conjunto de datos.

CAPÍTULO IV: CONCLUSIONES Y RECOMENDACIONES

4.1. CONCLUSIONES

El diseño de la propuesta didáctica a priori para la enseñanza del uso de la media aritmética y la mediana en un conjunto de datos, exige realizar el análisis exploratorio de los datos para determinar datos u observaciones atípicas que generan sesgo. Los estudiantes deben trabajar los contenidos de mediana, cuartiles y el rango intercuartilico (definiciones y estrategias de cálculo), para construir un intervalo que permita reconocer datos atípicos. A partir de este análisis, la utilización de las propiedades estadísticas y teniendo en cuenta las características para el análisis de la idoneidad del uso de la media aritmética y la mediana, el docente tiene herramientas para guiar a sus alumnos a identificar y elegir la medida más representativa de un conjunto de datos.

La propuesta didáctica contiene una secuencia de actividades elaboradas por configuraciones epistémicas y la aplicación de algunos aspectos de los criterios de idoneidad epistémica y cognitiva del EOS. Esta propuesta didáctica está dirigida a los docentes, al conocimiento, uso y manejo apropiado de la media aritmética y la mediana en un conjunto de datos para que ellos a su vez dirijan adecuadamente situaciones de aprendizaje en donde utilicen en forma coherente estos conceptos.

Los errores han sido recopilados de trabajos de investigación, relacionados a las medidas de tendencia central, que aplicaron cuestionarios a estudiantes españoles, mexicanos y angoleños. El conocimiento de estos errores a permitido desarrollar la secuencia de actividades de la propuesta didáctica con el propósito de evitarlos.

El tipo de preguntas que se identificaron han sido aquellas que muestran datos que no son comunes al resto. Tienen la intención de crear la necesidad de realizar el análisis exploratorio de los datos para averiguar la existencia de datos u observaciones atípicas que generan sesgo. Estas preguntas identificadas están incluidas en la secuencia de actividades, así como en el cuestionario que se aplicó a los docentes.

4.2. RECOMENDACIONES

Invitar a las instituciones de formación inicial docente a que incrementen los conocimientos teóricos adecuados al uso de la media aritmética y la mediana en un conjunto de datos. De esta manera los docentes estarán en la capacidad de realizar el análisis exploratorio, para determinar datos u observaciones atípicas que generan sesgo en el conjunto de datos, con el propósito de mejorar la deficiencia de elegir arbitrariamente la media aritmética o la mediana como medida representativa.

Si se quiere tener un estudio real es importante no quitar los datos extremos de un conjunto de datos, sin antes analizarlos, porque se corre el riesgo de no trabajar con la verdadera media aritmética y mediana.

La propuesta didáctica presenta las herramientas necesarias para que futuros investigadores la implementen y realicen el análisis de los resultados de la misma, como son la configuración epistémica de referencia del uso de la media aritmética y la mediana, así como la determinación de datos u observaciones atípicas que generan sesgo y los criterios de idoneidad didáctica tanto epistémica como cognitiva, tomadas del Enfoque Ontosemiótico de la Cognición e Instrucción Matemática (EOS).

REFERENCIAS

Bachelard, G. (1976). *La formación del espíritu científico (*5ta ed.). Editorial Siglo Veintiuno S.A. México

Batanero, C. (2001). *Didáctica de la estadística*. Departamento de Didáctica de la Matemática. Universidad de Granada.

Behar R. y Grima P. (2001). "Mil y una dimensiones del aprendizaje de la estadística". *Revista española de estadística.*

Chan, C. (2009). *Una propuesta didáctica sobre la media aritmética, la mediana y su representatividad.* (Tesis de licenciatura en enseñanza de las matemáticas. Universidad Autónoma de Yucatán-México).

Cobo, B. (2003). *Significado de las medidas de posición central para estudiantes de secundaria. (*Tesis doctoral. Universidad de Granada-España).

Cobo, B. y Batanero, C. (2000). "La mediana en la educación secundaria ¿Un concepto sencillo?", *Uno*, 23, 85-94.

Cobo, B. y Batanero, C. (2004). "Razonamiento numérico en problemas de promedios". *Suma*. 45, pp. 79-86.

Córdova, M. (2003). *Estadística descriptiva e inferencial* (5ta ed.). Editorial Moshera. Lima-Perú.

Estrada, A. (2007). "Evaluación del conocimiento estadístico en la formación inicial del profesorado". *UNO*, 45, 80-98.

García, M. (1992). *Métodos y Técnicas de investigación*. Alianza Universidad. Madrid.

García I. y García J. (2003). "Algunos resultados sobre la actuación de los estudiantes en las cuestiones de estadística en la PAU", *Actas de las XI Jornadas sobre el Aprendizaje y la Enseñanza de las Matemáticas (JAEM)*, pp. 733-738.

Garrett, A. y García, J. (2005). "Caracterización de la comprensión de algunos aspectos de la media aritmética: Un estudio con estudiantes de secundaria y universitarios". *Investigación*, 1(17), pp. 31-57.

Godino, J. (2003). *Teoría de las funciones semióticas. Un enfoque ontológico-semiótico de la cognición e instrucción matemática.* Departamento de Didáctica de la Matemática. Universidad de Granada-España.

Godino, J. (2011). Indicadores de la idoneidad didáctica de procesos de enseñanza y aprendizaje de las matemáticas. *Conferencia presentada en la XIII CIAEM-IACME*, Recife-Brasil.

Godino, J., Batanero, C. y Font, V. (2007). "El enfoque ontosemiótico a la investigación en educación". *Revista Internacional de Educación Matemática*. 39 (1-2), pp.127-135.

Godino, J., Batanero, C. y Font, V. (2008). *El Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemática (EOS)*. Departamento de Didáctica de la Matemática. Universidad de Granada-España.

Godino, J., Batanero, C. y Font, V. (2009). *Un enfoque ontosemiótico del conocimiento y la instrucción matemática*. Departamento de Didáctica de la Matemática. Universidad de Granada-España.

Godino, J., Contreras, A. y Font, V. (2006). "Análisis de procesos de instrucción basado en el enfoque ontológico-semiótico de la cognición matemática". *Recherches en Didactiques des Mathématiques*, 26 (1), 39-88.

Godino, J., Font, V. y Wilhelmi, M. (2008). "Análisis didáctico de procesos de estudio matemático basado en el enfoque ontosemiótico". *Publicaciones*. 38, 25-49.

Langford, E. (2006). "Quartiles in elementary statistics". *Journal of Statistics Education*. 14 (3), 1-27.

Latorre, A. y Cols. (1996). *Bases metodológicas de la investigación educativa* (1era ed.). GR92. Monterrey-México.

Mayén, S. (2009). *Comprensión de las medidas de tendencia central en estudiantes mexicanos de educación secundaria y bachillerato.* (Tesis Doctoral. Universidad de Granada-España).

Molina, G. y Rodrigo, M. (2009). Estadística descriptiva en psicología Estadísticos de forma de la distribución. *Documentos de curso*, 5, p.1-9. Universidad de Valencia-España.

Moreno, L. (2006). Profesorado de secundaria y calidad de la educación: Un marco de opciones políticas para la formación y desarrollo profesional docente. *Revista de curriculum y formación del profesorado*, 10, p.1-21.

Novaes, D. (2011). *Conceptos de professores de Educacao Básica sobre variabilidad de estadística.* (Tesis Doctoral. Universidad Católica de Sao Paulo-Brasil).

Perú (2006). *Glosario básico de términos estadísticos.*(INEI), Lima-Perú.

Perú (2009). *Diseño Curricular Nacional*. Ministerio de Educación del Perú. Lima.

Perú (2010). *Diseño Curricular Básico Nacional para la carrera profesional de profesor de educación secundaria en la especialidad de matemática.* Ministerio de Educación del Perú. Lima.

Perú (2013). *Mapas de progreso*. Ministerio de Educación del Perú. Lima.

Pinzón, L. (2012). *Propuesta didáctica para el aprendizaje de la media aritmética, la mediana y la moda, para estudiantes del programa de psicología*. (Tesis de maestría en Enseñanza de las Ciencias Exactas y Naturales. Universidad Nacional de Colombia.

Sayritupac, J. (2013). *Significados de las medidas de tendencia central. Un estudio con estudiantes universitarios de carreras de humanidades.* Tesis de maestría en enseñanza de las matemáticas. Pontificia Universidad Católica del Perú. Lima.

Skovsmose, O. (1999). *Hacia una filosofía de la educación matemática critica*. Universidad de Los Andes, Bogotá-Colombia.

Suárez, M. (2012). *Interaprendizaje de Estadística Básica.* Universidad Técnica del Norte. Ibarra-Ecuador.

Sweeney, D. (2008). *Estadística para administración y economía* (10ma. ed.). Editorial Thomson. México.

Wilhelmi, M., Bencomo, D. y Godino, J. (2003). C*riterios de idoneidad de un proceso de instrucción matemática.* Universidad de Granada-España.

ANEXOS

ANEXO 1:

Este cuestionario es adaptado del cuestionario aplicado por García y García (2003).

Estimado colega:

Con el propósito de mejorar algunos aspectos importantes de la estadística, nos dirigimos a Ud. para que conteste las preguntas planteadas en el presente cuestionario, esto nos permitirá hacer un análisis y proponer mejoras que despierten por un lado, el interés por aprender y por otro caracterizar la enseñanza del uso de la media aritmética y la mediana, que se trabaja dentro del aula para desarrollarla exitosamente, así que te pedimos responder con sinceridad las siguientes preguntas.

DATOS PERSONALES

TIPO DE IE DONDE TRABAJA : ESTATAL: ◯ PRIVADO: ◯

TIEMPO DE SERVICIOS : ☐ AÑOS

GRADO O AÑO QUE ENSEÑA : ______________________________

INSTITUCIÓN SUPERIOR DONDE SE FORMÓ: ______________________________

1. Nueve estudiantes en una sesión de aprendizaje de CTA, pesaron un objeto obteniendo los resultados siguientes: 6.2; 6.0; 6.0; 9.3; 6.1; 6.3; 6.2; 6.15; 6.2. Los estudiantes esperan determinar si la media aritmética o la mediana es la más representativa del conjunto de datos. ¿Cuál le recomendaría usar?

   a) Usar la media aritmética

   b) Usar la mediana

   Justifique su respuesta.

ANEXO 2:

2. La profesora María, registra la participación de los estudiantes quienes formulan preguntas referentes al tema tratado. Un registro del número de preguntas hechas por sus 8 estudiantes durante una clase se muestra en la siguiente tabla:

	INICIALES DEL ALUMNO							
Estudiantes	AA	RF	AG	JG	CK	NK	JL	AW
Número de preguntas	0	5	2	10	3	2	1	2

La profesora María, quiere determinar si la media aritmética o la mediana es la más representativa de los datos. ¿Cuál le recomendaría usar?

a) Usar la media aritmética

b) Usar la mediana

Justifique su respuesta.

ANEXO 3:

3. Se han obtenido un conjunto de datos de una muestra estadística, y se observa que algunos datos están “alejados” del resto.

Según lo anterior, ¿Qué entiende Ud. por sesgo estadístico?

GRACIAS POR TU PARTICIPACIÓN…

ANEXO 4: ACTIVIDAD 1

NOMBRE: CONSTRUCCION DEL INTERVALO PARA RECONOCER DATOS U OBSERVACIONES ATIPICAS

Datos informativos

1.1 Institución Educativa	:	______________________________
1.2 Director (a)	:	______________________________
1.3 Grado y sección	:	2° ____
1.4 Duración	:	1 hora (45 min)
1.5 Nivel	:	Secundaria
1.6 Ciclo	:	VI
1.7 Docente	:	_____________________________
1.8 Área	:	Matemática
1.9 Fecha	:	________

Aprendizajes esperados

Construir el intervalo para reconocer datos u observaciones atípicas.

Área	Competencia	Capacidad	Indicadores
Matemática	ESTADISTICA Y PROBABILIDAD Selecciona la medida de tendencia central apropiada para representar un conjunto de datos al resolver problemas. Argumenta procedimientos para hallar la media, mediana y moda de datos agrupados y no agrupados; determina la medida más representativa de un conjunto de datos y su importancia en la toma de decisiones.	• Elabora y usa estrategias • Razona y argumenta usando ideas matemáticas	Usa estrategias para construir el intervalo para determinar datos u observaciones atípicas

INICIO (10 min)	**SITUACION DE CONTEXTO**: Las calificaciones obtenidas por los estudiantes de segundo año de secundaria en el área de matemática. Se les invita a los estudiantes a observar y leer el siguiente problema: **MATEMATIZA** El siguiente cuadro muestra los resultados que obtuvieron los estudiantes del segundo año de educación secundaria en el área de matemática correspondiente al tercer bimestre del año escolar 2014 en la I.E.P "Carmelitas" de Tumbes: 15, 17, 12, 13, 10, 17, 18, 11, 10, 09, 14, 16, 14, 12, 18, 15, 12, 14, 15, 13, 11, 10, 17, 14, 12. A partir de la información determine si existen datos u observaciones atípicas. **COMUNICA** Recogemos los saberes previos: • ¿Qué es la mediana y como se calcula? • ¿Qué es un cuartil y como se calcula? • ¿Qué relación tiene la mediana con el segundo cuartil? Conflicto cognitivo. • ¿Qué entienden por dato atípico? • ¿Conocen alguna estrategia para determinar datos u observaciones atípicas? • ¿Qué es un intervalo? • ¿Qué son límites de un intervalo? • ¿Qué es el rango intercuartílico? **ELABORA LA ESTRATEGIA** El docente presenta a los estudiantes la estrategia para responder la pregunta del problema.

<table>
<tr><td>DESARROLLO (25 min)</td><td>

REPRESENTA

El docente muestra como estrategia “El diagrama de caja y bigotes” o “Intervalo para determinar datos u observaciones atípicas”.
El docente con la participación de los estudiantes ordena los datos y calcula la mediana, indicando que la mediana es equivalente al segundo cuartil (Me=Q_2).
Luego localiza el primer y tercer cuartil.

- El docente escribe la forma del intervalo a construir: $[Q_1 - 1{,}5, RIC, Q_3 + 1{,}5\, RIC]$, el mismo que se puede escribir también como $[Q_1 - D, Q_3 + D]$, haciendo D = 1,5 (Q_3 – Q_1), que representa el rango intercuartílico (RIC = Q_3 – Q_1) multiplicado por el factor 1,5.
- Después de efectuar las operaciones aritméticas entre los cuartiles y la expresión D, se ha obtenido el intervalo que nos permitirá realizar el análisis de los datos para determinar la existencia de datos u observaciones atípicas.
- Se presentan dos casos:
  - Si los datos pertenecen al intervalo: $[Q_1 - D, Q_3 + D]$, entonces no existen datos u observaciones atípicas, ni sesgo en los datos.
  - Si algún dato no pertenece al intervalo, entonces es dato u observación s atípica que sesga al conjunto de datos.

ARGUMENTA

Justifican el “Diagrama de caja y bigotes” o el “Intervalo para determinar datos u observaciones atípicas” como estrategia que permite determinar con seguridad datos u observaciones atípicas.

</td></tr>
<tr><td>CERRE (10 min)</td><td>

Luego preguntamos:

¿Cómo construyeron el intervalo para determinar datos u observaciones atípicas?

¿Qué conceptos se tuvieron en cuenta para construir el intervalo?

¿Por qué es necesario determinar datos u observaciones atípicas?

El docente plantea que ante el reclamo de un estudiante que había obtenido 17 de nota, se realiza la rectificación y obtiene ahora de nota 20. Con este cambio de dato analizar si existen datos u observaciones atípicas.

Resuelven una ficha de autoevaluación y metacognición.

</td></tr>
</table>

EVALUACIÓN

TÉCNICA	INSTRUMENTO
Cuestionario	Ficha de aplicación

Fecha.

D I R E C T O R (A) D O C E N T E

FICHA TÉCNICA DE LA ACTIVIDAD 01

SITUACIÓN PROBLEMÁTICA

El siguiente cuadro muestra los resultados que obtuvieron 25 estudiantes del segundo año de educación secundaria, en el área de matemática correspondiente al tercer bimestre del año escolar 2014 en la I.E.P "Carmelitas" de Tumbes:

14,	17,	12,	13,	10,	17,	18,	11,	10,	09,
14,	15,	14,	12,	17,	14,	12,	14,	14,	13,
11,	10,	16,	14,	12.					

A partir de la información determine si existen datos u observaciones atípicas.

SOLUCIÓN:

a) Ordenamos los datos de menor a mayor.

09, 10,10, 10, 11, 11, 12, 12, 12, 12, 13, 13, 14, 14, 14, 14, 14, 14, 14, 15, 16, 17, 17, 17, 18.

b) Calculamos la mediana de los datos ordenados, teniendo en cuenta que la cantidad de datos es un número impar.

09, 10,10, 10, 11, 11, 12, 12, 12, 12, 13, 13, 14, 14, 14, 14, 14, 14, 14, 15, 16, 17, 17, 17, 18.

Me= (n+1)/2 = (25+1)/2 = 13 (posición de la mediana)

Luego el dato de la posición número 13 lo ocupa el valor 14.

Me=14

c) Como la media aritmética es diferente a la mediana, averiguamos si existen datos u observaciones atípicas en el conjunto de datos.

Para ello construimos el intervalo para determinar datos u observaciones atípicas en el conjunto de datos

Localizamos los cuartiles Q_1 y Q_3 respectivamente.

$$\underbrace{09, 10,10, 10, 11, 11}\ \Big|\ \underbrace{12, 12, 12, 12, 13, 13}\ \mathbf{14}\ \underbrace{14, 14, 14, 14, 14, 14}\ \Big|\ \underbrace{15, 16, 17, 17, 17, 18}$$

Q_1 =11,5 Me=Q_2= 14 Q_3=14,5

D, es la diferencia del rango intercuartílico (RIC = $Q_3 - Q_1$) multiplicado por el factor 1,5.

$D = 1{,}5(Q_3 - Q_1) \Longrightarrow$ D = 1,5 (14,5 – 11,5) = 1,5 (3) = 4,5

Luego los datos deben pertenecer al intervalo:

$[Q_1 - D, Q_3 + D] \quad = \quad [1\,1{,}5 - 4{,}5;\ 14{,}5 + 4{,}5] = \ [7;\ 19]$

Como vemos todos los datos pertenecen al intervalo $[7;\ 19]$.

CONCLUSIÓN:

Deducimos que no existen datos u observaciones atípicas.

Un estudiante inconforme con su nota 17, presenta un reclamo al docente, quién después de corregir cambia la nota del estudiante a 20.

A partir de la información determine si existen datos u observaciones atípicas.

14, 17, 12, 13, 10, 17, 18, 11, 10, 09,

14, 15, 14, 12, 17, 14, 12, 14, 14, 13,

11, 10, 16, 14, 12.

SOLUCIÓN:

a) Ordenamos los datos de menor a mayor.

09, 10,10, 10, 11, 11, 12, 12, 12, 12, 13, 13, 14, 14, 14, 14, 14, 14, 14, 15, 16, 17, 17, 18, 20.

b) Calculamos la media aritmética.

$$\bar{x} = \frac{9 + 3(10) + 2(11) + 4(12) + 2(13) + 7(14) + 15 + 16 + 2(17) + 18 + 20}{25}$$

$$\bar{x} = \frac{9 + 30 + 22 + 48 + 26 + 98 + 15 + 16 + 34 + 18 + 20}{25}$$

$$\bar{x} = \frac{336}{25} = 13{,}44$$

c) Calculamos la mediana de los datos ordenados.

09, 10, 10, 10, 11, 11, 12, 12, 12, 12, 13, 13,

14, 14, 14, 14, 15, 15, 15, 16, 17, 17, 18, 18,

20.

Me= 14

d) Como la media aritmética es diferente a la mediana, averiguamos si existen datos u observaciones atípicas en el conjunto de datos.

Para ello construimos el intervalo para determinar datos u observaciones atípicas en el conjunto de datos

Localizamos los cuartiles Q_1 y Q_3 respectivamente.

$$\underbrace{09, 10, 10, 10, 11, 11}, \underbrace{12, 12, 12, 12, 13, 13}, \mathbf{14}, \underbrace{14, 14, 14, 14, 14, 14}, \underbrace{15, 16, 17, 17, 18, 20}$$

Q_1 =11,5 Me=Q_2= 14 Q_3=14,5

D, es la diferencia del rango intercuartílico (RIC = $Q_3 - Q_1$) multiplicado por el factor 1,5.

$$D = 1{,}5(Q_3 - Q_1) \Longrightarrow D = 1{,}5\ (14{,}5 - 11{,}5) = 1{,}5\ (3) = 4{,}5$$

Luego los datos deben pertenecer al intervalo:

$[Q_1 - D, Q_3 + D] = [1\ 1{,}5 - 4{,}5;\ 14{,}5 + 4{,}5] = [7;\ 19]$

Como vemos todos el dato 20 no pertenece al intervalo $[7;\ 19]$.

CONCLUSIÓN:

Deducimos que existen datos u observaciones atípicas.

ANEXO 5: ACTIVIDAD 2

NOMBRE: ELIGIENDO EL MEJOR REPRESENTANTE

Datos informativos

1.1 Institución Educativa : ________________________________

1.2 Director (a) : ________________________________

1.3 Grado y sección : 2° ____

1.4 Duración : 2 horas (90 min)

1.5 Nivel : Secundaria

1.6 Ciclo : VI

1.7 Docente : _______________________________

1.8 Área : Matemática

1.9 Fecha : ________

Aprendizajes esperados

Determinar datos u observaciones atípicas y elegir el mejor representante del conjunto de datos, aplicando propiedades y criterios de la media aritmética y la mediana.

Área	Competencia	Capacidad	Indicadores
Matemática	ESTADISTICA Y PROBABILIDAD Argumenta procedimientos para determinar datos u observaciones atípicas que generan sesgo, para determina la medida más representativa de un conjunto de datos y su importancia en la toma de decisiones.	• Elabora y usa estrategias • Razona y argumenta usando ideas matemáticas	Usa estrategias para determinar datos u observaciones atípicas que sesgan la información, así como propiedades y criterios para tomar decisiones y elegir a la media aritmética o la mediana como el mejor representante del conjunto de datos.

INICIO (10 min)

.

SITUACION DE CONTEXTO: Los sueldos de los trabajadores de la I.E.P. "Carmelitas" de Tumbes.

Se les invita a los estudiantes a observar y leer el siguiente problema:

MATEMATIZA

El siguiente cuadro muestra el número de trabajadores y el sueldo que ganan en la Institución Educativa privada "Carmelitas" de la región Tumbes:

1	Promotor	3 000
1	Director	2 500
17	Docentes de aula	1 500
2	Secretarias	1 200
2	Trabajadores de servicio	1 000
4	Auxiliares de educación	1 000
1	Guardián	1 000

¿Cuál sería la medida estadística más representativa del conjunto de datos?

COMUNICA

A partir de la situación planteada, recogemos los saberes previos con preguntas como:
¿Qué es la media aritmética y como se calcula?
¿Si la media aritmética no es igual a la mediana, esto se debe a que hay datos que son atípicos que hace que sean diferentes?
¿Qué es una distribución simétrica?
¿Qué es una distribución asimétrica?

ELABORA LA ESTRATEGIA

El docente presenta a los estudiantes las propiedades estadísticas para responder la pregunta del problema.
Conflicto cognitivo.
¿Cuándo usar media aritmética o mediana, como representante de un conjunto de datos?

<table>
<tr><td>DESARROLLO (25 min)</td><td>

REPRESENTA

- El docente pide a los estudiantes que calculen la media aritmética y la mediana.

Luego les pregunta:

- ¿Qué diferencia hay entre la media aritmética y la mediana?
- Si existe diferencia entre la media aritmética y la mediana, entonces averigüemos la existencia datos u observaciones atípicas.
- El docente pide a los estudiantes que construyan el intervalo para determinar la existencia de datos u observaciones atípicas.
- Luego de haber construido con los estudiantes, el intervalo para determinar datos u observaciones atípicas y analiza los datos aplicando la siguiente propiedad:
    - Si los datos pertenecen al intervalo: $[Q_1 - D, Q_3 + D]$, entonces no existen datos u observaciones atípicas ni sesgo en el conjunto de datos.
    - Si algún dato no pertenece al intervalo, entonces existen datos u observaciones atípicas y por tanto existe sesgo en el conjunto de datos.
- Después de averiguar la existencia o no de datos u observaciones atípicas, los estudiantes están en la capacidad de determinar la medida estadística que mejor representa al conjunto de datos, aplicando la siguiente propiedad:
  "En presencia de datos u observaciones atípicas, la mediana es el mejor representante de los datos, en caso contrario si no se verifica la existencia de datos u observaciones atípicas, la media aritmética es el mejor representante del conjunto de datos.

ARGUMENTA

Justifican su decisión basándose en propiedades estadísticas de la media aritmética y la mediana.
El docente verifica las estrategias y resultados.
Se les pide que comparen sus respuestas, que las analicen y que expliquen porque tomaron esa decisión.
</td></tr>
<tr><td>CERRE (10 min)</td><td>

Luego preguntamos:
¿Cómo decidieron la medida más representativa?
¿Qué propiedades estadísticas se tuvieron en cuenta para tomar la decisión?
¿Por qué la determinación de datos u observaciones atípicas les llevó a tomar esa decisión?
¿Cómo se sintieron al resolver la situación planteada?
¿Fue fácil o difícil tomar la decisión? ¿Por qué?
¿Qué otros problemas parecidos podemos crear?
Resuelven una ficha de autoevaluación y metacognición.
Analizan una situación parecida usando el texto del MED.
</td></tr>
</table>

EVALUACIÓN

TÉCNICA	INSTRUMENTO
Cuestionario	Ficha de aplicación

Fecha.

________________________ ________________________

D I R E C T O R (A) D O C E N T E

FICHA TÉCNICA DE LA ACTIVIDAD 02

SITUACIÓN PROBLEMÁTICA

El siguiente cuadro muestra el número de trabajadores y el sueldo que ganan en la Institución Educativa privada "Carmelitas" de la región Tumbes:

1	Promotor	3 000
1	Director	2 500
17	Docentes de aula	1 500
2	Secretarias	1 200
2	Trabajadores de servicio	1 000
4	Auxiliares de educación	1 000
1	Guardián	1 000

¿Cuál sería el mejor representante del conjunto de datos?

SOLUCIÓN:

a) Calculamos la media aritmética.

$$\bar{x} = \frac{1000 + 4(1000) + 2(1000) + 2(1200) + 17(1500) + 2500 + 3000}{28}$$

$$\bar{x} = \frac{1000 + 4000 + 2000 + 2400 + 25500 + 2500 + 3000}{28}$$

$$\bar{x} = \frac{40400}{28} = 1442{,}86$$

b) Calculamos la mediana, ordenando los datos.

1 000, 1 000, 1 000, 1 000, 1 000, 1 000, 1 000, 1 200, 1 200, 1 500, 1 500, 1 500, 1 500,

1 500, 1 500, 1 500, 1 500, 1 500, 1 500, 1 500, 1 500, 1 500, 1 500, 1 500, 1 500, 1 500,

2 500, 3 500.

$$Me = \frac{1500 + 1500}{2} = \frac{3000}{2} = 1\,500$$

c) Como la media aritmética es diferente a la mediana, averiguamos si existen datos atípicos en el conjunto de datos.

Para ello aplicamos el siguiente procedimiento:

- Si los datos pertenecen al intervalo $[Q_1 - D, Q_3 + D]$, significa que no existen datos atípicos, y por tanto la media aritmética es el mejor representante del conjunto de datos.
- Si algún dato no pertenece al intervalo $[Q_1 - D, Q_3 + D]$, entonces se dice que hay datos atípicos que sesgan la información y por lo tanto la mediana es el mejor representante del conjunto de datos.

Q_1 y Q_3 son los cuartiles primero y tercero respectivamente.

1 000, 1 000, 1 000, 1 000, 1 000, 1 000, 1 000, | 1 200, 1 200, 1 500, 1 500, 1 500, 1 500,

Q_1=1 100

1 500, | 1 500, 1 500, 1 500, 1 500, 1 500, 1 500, 1 500, | 1 500, 1 500, 1 500, 1 500, 1 500,

Q_2= 1500 Q_3= 1500

2 500, 3 000.

$n = 28 \qquad \frac{n}{2} = \frac{28}{2} = 14 \quad \Longrightarrow \quad i_{14} = 1500, \quad i_{15} = 1500$

$$Me = Q_2 = \frac{1500 + 1500}{2} = 1500$$

D, es la diferencia del $Q_3 - Q_1$ multiplicado por el factor 1,5.

$D = 1,5(Q_3 - Q_1) \quad \Longrightarrow$ D = 1,5 (1 500 – 1100) = 1,5 (400) = 600

Luego los datos deben pertenecer al intervalo:

$[Q_1 - D, Q_3 + D] \quad = \quad [1\,100 - 600;\ 1\,500 + 600] = [500;\ 2\,100]$

Como vemos los datos 2 500 y 3 000 no pertenecen al intervalo $[500;\ 2\,100]$, lo que significa que son atípicos y por tanto existe sesgo en los datos.

CONCLUSIÓN:

Deducimos que el mejor representante del conjunto de datos es la mediana, Me = 1 500.

ANEXO 6: ACTIVIDAD 3

NOMBRE: ELIGIENDO EL MEJOR REPRESENTANTE

Datos informativos

1.1 Institución Educativa	:	________________________________
1.2 Director (a)	:	
1.3 Grado y sección	:	2° ____
1.4 Duración	:	1 hora (45 min)
1.5 Nivel	:	Secundaria
1.6 Docente	:	_______________________________
1.7 Área	:	Matemática
1.8 Fecha	:	________

Aprendizajes esperados

Que los estudiantes reconozcan y elijan el mejor representante de un conjunto de datos, aplicando criterios.

Área	Competencia	Capacidad	Indicadores
Matemática	ESTADISTICA Y PROBABILIDAD Selecciona la medida de tendencia central apropiada para representar un conjunto de datos al resolver problemas. Argumenta procedimientos para hallar la media y la mediana de datos agrupados y no agrupados; determina la medida más representativa de un conjunto de datos y su importancia en la toma de decisiones.	• Elabora y usa estrategias • Razona y argumenta usando ideas matemáticas	Usa criterios para determinar datos atípicos que sesgan la información, para tomar decisiones y elegir el mejor representante del conjunto de datos.

SITUACION DE CONTEXTO 1: Las propinas que recibieron los alumnos de 2do de secundaria durante una semana.
SITUACION DE CONTEXTO 2: Las edades de los alumnos de 2do de secundaria.

Se les entrega a los una hoja fotocopiada con 2 situaciones a resolver.

El siguiente cuadro muestra las propinas que recibieron los 30 alumnos de 2do A, durante una semana.

S/. 1.00	12
S/. 1.50	6
S/. 2.00	10
S/. 2.50	1
S/. 5.00	1

¿Cuál sería el mejor representante del conjunto de datos?
El siguiente cuadro muestra las edades de los 30 alumnos de 2do A, durante una semana.

12 años	10
13 años	15
14 años	4
15 años	1

¿Cuál sería el mejor representante del conjunto de datos?

A partir de las situaciones planteadas, y trabajando en dúos los estudiantes elaboran estrategias de solución del problema.
Representan matemáticamente cada situación.
Comunica y representa ideas matemáticas.
Buscan y elaboran una estrategia de solución.
Justifica y argumenta cada uno de los procedimientos de cálculo realizados y propiedades aplicadas.
Se apoya en el libro del MED y el ejemplo resuelto en clase.

Fecha.

DIRECTOR (A)   DOCENTE

FICHA TÉCNICA DE LA ACTIVIDAD 03

SOLUCIÓN DE LA SITUACIÓN PROBLEMÁTICA 1:

1) El siguiente cuadro muestra las propinas que recibieron los 30 estudiantes de 2do A, durante una semana.

S/. 1.00	12
S/. 1.50	6
S/. 2.00	10
S/. 2.50	1
S/. 5.00	1

¿Cuál es el mejor representante del conjunto de datos?

SOLUCIÓN:

a) Calculamos la media aritmética de los datos.

$$\bar{x} = \frac{12(1) + 6(1{,}50) + 10(2) + 2{,}50 + 5}{30}$$

$$\bar{x} = \frac{12 + 9 + 20 + 2{,}50 + 5}{30}$$

$$\bar{x} = \frac{48{,}50}{30} = 1{,}61666.. \approx 1{,}60$$

b) Calculamos la mediana, ordenando los datos.

1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1,50; 1,50; **1,50;**

1,50; 1,50; 1,50; 2; 2 ; 2; 2; 2; 2; 2; 2; 2; 2; 2,50; 5

$$Me = \frac{1{,}50 + 1{,}50}{2} = \frac{3}{2} = 1{,}50$$

c) Como la media aritmética es diferente a la mediana, averiguamos si existen datos atípicos en el conjunto de datos.

Vale indicar que no se ha establecido que tan diferentes deben ser estas medidas para proceder a determinar datos atípicos.

Aplicamos el procedimiento para determinar datos u observaciones atípicas que sesgan la información.

Localizamos los cuartiles Q_1 y Q_3 respectivamente en el conjunto de datos.

$$\underbrace{1;\ 1;\ 1;\ 1;\ 1;\ 1;\ 1}_{},\quad 1\quad \underbrace{1;\ 1;\ 1;\ 1;\ 1{,}50;\ 1{,}50;\ 1{,}50}_{},\ \downarrow\quad \underbrace{1{,}50;\ 1{,}50;\ 1{,}50;\ 2;\ 2;\ 2;\ 2}_{}$$

$Q_1 = 1$ Me = 1,50

$$2;\quad \underbrace{2\quad 2\quad 2\quad 2\quad 2\quad 2{,}50\quad 5}_{}$$

$Q_3 = 2$

Determinamos el valor de la expresión que representa 1,5 RIC.

$D = 1{,}5(Q_3 - Q_1)$ $\Longrightarrow$ D = 1,5 (2 – 1) = 1,5 (1) = 1,50

Reemplazamos valores en el intervalo $[Q_1 - D, Q_3 + D]$, para determinar datos u observaciones atípicas, quedando como sigue:

$$[Q_1 - D, Q_3 + D] \quad = \quad [1{,}00 - 1{,}50;\ 2{,}00 + 1{,}50] = [-0{,}50;\ 3{,}50]$$

Luego analizamos los datos y como vemos el dato 5,00 no pertenecen al intervalo $[-0{,}50;\ 3{,}50]$, lo que significa que es atípico y por tanto existe sesgo en los datos.

Por propiedad, "Si alguno de los datos no pertenece al intervalo $[Q_1 - D, Q_3 + D]$, significa que es atípico, y por tanto la mediana es el mejor representante del conjunto de datos.

CONCLUSIÓN:

Deducimos que el mejor representante del conjunto de datos es la mediana, Me= 1,50.

SOLUCIÓN DE LA SITUACIÓN PROBLEMÁTICA 2:

2) El siguiente cuadro muestra las edades de los 30 estudiantes de 2do A, durante una semana.

12 años	10
13 años	15
14 años	4
15 años	1

¿Cuál es el mejor representante del conjunto de datos?

SOLUCIÓN:

i) Calculamos la media aritmética de los datos.

$$\bar{x} = \frac{10(12) + 15(13) + 4(14) + 15}{30}$$

$$\bar{x} = \frac{120 + 195 + 56 + 15}{30}$$

$$\bar{x} = \frac{386}{30} = 12{,}8666.. \approx 13$$

ii) Calculamos la mediana, ordenando los datos.

12; 12; 12; 12; 12; 12; 12; 12; 12; 12; 13; 13; 13; 13; **13;** **13;** 13; 13; 13; 13; 13; 13; 13; 13; 13; 14; 14; 14; 14; 15.

$$Me = \frac{13+13}{2} = \frac{26}{2} = 13$$

iii) Como la media aritmética es igual a la mediana, la distribución es simétrica y no es necesario averiguar si existen datos u observaciones atípicas en el conjunto de datos.

CONCLUSIÓN:

Por la propiedad “en distribuciones simétricas, el mejor representante del conjunto de datos es la media aritmética” $\bar{x} = 13$.

ANEXO 7: ACTIVIDAD 4

NOMBRE: EVALUANDO LA ELECCIÓN DEL MEJOR REPRESENTANTE

Datos informativos

1.1 Institución Educativa : ______________________________

1.2 Director (a) :

1.3 Grado y sección : 2° ____

1.4 Duración : 1 hora

1.5 Nivel : Secundaria

1.6 Docente : ______________________________

1.7 Área : Matemática

1.8 Fecha : ________

Objetivo

Que los estudiantes demuestren capacidad de elegir el mejor representante de un conjunto de datos, teniendo en cuenta datos atípicos y aplicando propiedades

Estadísticas.

Área	Competencia	Capacidad	Indicadores
Matemática	ESTADISTICA Y PROBABILIDAD Selecciona la medida de tendencia central apropiada para representar un conjunto de datos al resolver problemas. Argumenta procedimientos para hallar la media y la mediana de datos agrupados y no agrupados; determina la medida más representativa de un conjunto de datos y su importancia en la toma de decisiones.	• Elabora y usa estrategias • Razona y argumenta usando ideas matemáticas	Usa criterios para determinar datos atípicos que sesgan la información, para tomar decisiones y elegir el mejor representante del conjunto de datos.

INICIO (5 min)	El docente distribuye dentro del aula a los estudiantes por orden alfabético, da las instrucciones del examen y entrega una hoja A4 fotocopiada que contiene una pregunta que ha sido tomada de una situación problemática contextualizada, relacionada con datos estadísticos numéricos, en la que se pide determinar el mejor representante del conjunto de datos. *Las siguientes notas son los resultados de la evaluación aplicada a 30 estudiantes del 2do A de nuestra Institución Educativa, correspondiente al I bimestre en el presente año:* 05 08 08 09 09 10 10 11 11 12 12 12 12 13 13 13 13 14 14 14 14 14 14 14 15 15 15 15 15 20 ¿Cuál es el mejor representante del conjunto de datos?
DESARROLLO (35 min)	Los estudiantes inician el desarrollo del problema y responden la pregunta planteada, teniendo en cuenta que la decisión a tomar está relacionada con el análisis de la presencia o no de datos atípicos, aplican para ello las estrategias aprendidas en la sesión de clase anterior. Justifican los razonamientos aplicados. El docente atiende dudas de los estudiantes evitando dar ayudas que influyen directamente con la respuesta al problema planteado.
CERRE (5 min)	Luego preguntamos: ¿Cómo se sintieron al resolver el problema? ¿Te fue fácil o difícil tomar la decisión? ¿Por qué?

Fecha.

D I R E C T O R (A)   D O C E N T E

FICHA TÉCNICA DE LA ACTIVIDAD 04

SITUACIÓN PROBLEMÁTICA

Las siguientes notas son los resultados de la evaluación aplicada a 30 estudiantes del 2do A de nuestra Institución Educativa, correspondiente al I bimestre en el presente año:

05 08 08 09 09 10 10 11 11 12

12 12 12 13 13 13 13 14 14 14

14 14 14 14 15 15 15 15 15 20

¿Cuál sería el mejor representante del conjunto de datos?

SOLUCIÓN:

i) Calculamos la media aritmética de los datos.

$$\bar{x} = \frac{5 + 2(8) + 2(9) + 2(10) + 2(11) + 4(12) + 4(13) + 7(14) + 5(15) + 20}{30}$$

$$\bar{x} = \frac{5 + 16 + 18 + 20 + 22 + 48 + 52 + 98 + 75 + 20}{30}$$

$$\bar{x} = \frac{374}{30} = 12{,}4666 \ldots \approx 12$$

ii) Calculamos la mediana, aprovechando que ya están ordenados los datos.

05 08 08 09 09 10 10 11 11 12 12 12 12 13 **13**

13 13 14 14 14 14 14 14 14 15 15 15 15 15 20

$$Me = \frac{13 + 13}{2} = \frac{26}{2} = 13$$

iii) Como la media aritmética es diferente a la mediana, averiguamos si existen datos atípicos en el conjunto de datos.

iv) Construimos el intervalo para determinar datos u observaciones atípicas, localizando la mediana que equivale al cuartil 2 los cuartiles 1 y 3.

05	08	08	09	09	10	10	11	11	12

Q_1=11

12	12	12	13	13	13	13	14	14	14

Q_2= 13

14	14	14	14	15	15	15	15	15	20

Q_3=14

Una vez localizados los cuartiles, determinamos la expresión D que es 1,5 RIC.

Reemplazando valores en $D = 1{,}5(Q_3 - Q_1)$ ⟹ D = 1,5 (14 – 11) = 1,5 (3) = 4,50

y luego en el intervalo $[Q_1 - D, Q_3 + D]$ = $[11 - 4{,}50;\ 14 + 4{,}50]$ = $[6{,}50;\ 18{,}50]$

Al analizar los datos, nos damos cuenta que los datos 05 y 20 no pertenecen al intervalo $[6{,}50;\ 18{,}50]$, lo que significa que son datos u observaciones atípicas y por lo tanto existe sesgo en los datos.

Luego de este análisis de datos, aplicamos la siguiente propiedad estadística: “En presencia de datos u observaciones atípicas que sesgan la información, la mediana es el mejor representante del conjunto de datos”.

CONCLUSIÓN:

Tomamos la decisión de elegir como representante del conjunto de datos es la mediana.

ANEXO 8:

ESTUDIO CUALITATIVO DEL CUESTIONARIO

1. Introducción

En este capítulo presentamos los resultados obtenidos de un cuestionario sobre el uso de la media aritmética y la mediana en un conjunto de datos estadísticos y el concepto de sesgo, que se realizó con una muestra de 44 docentes en ejercicio del área de matemática que trabajan en el sector estatal y/o privado, con un rango de tiempo de servicios diferente, así como los diferentes grados o años que enseña y la institución superior que lo formó como docente (Instituto Superior Pedagógico o Universidad).

Empezamos el capítulo con el detalle de las características de los docentes que forman parte de la muestra, se describe el cuestionario aplicado, que es una adaptación de algunas preguntas que aplicó Mayén (2009) y Garrett y García (2005) y que utilizaron con estudiantes. Estos cuestionarios fueron validados por docentes expertos en matemática y estadística. Nuestro aporte consiste en el análisis de las respuestas como un diagnóstico de los conocimientos de los docentes de matemática en ejercicio.

2. Descripción de la muestra

La muestra está compuesta por 44 docentes del área de matemática en ejercicio que pertenecen a la zona norte, centro y sur del Perú, correspondiendo el 64% a la región Tumbes, tal como mostramos en la siguiente tabla:

Tabla 24. Distribución de la muestra por procedencia

ZONA	REGIÓN	TOTAL
NORTE	TUMBES	28
	LA LIBERTAD	1
	CAJAMARCA	1
CENTRO	JUNIN	1
	LIMA	4
SUR	AREQUIPA	6
	PUNO	1
	TACNA	1
	AYACUCHO	1
	TOTAL	44

Esta muestra representa un conjunto de docentes que laboran en distintas regiones del país y en distintos tipos de instituciones educativas. El 80% trabajan en la educación pública, mientras que el 30% trabaja para el sector privado y un 10% del total trabaja en ambos sectores.

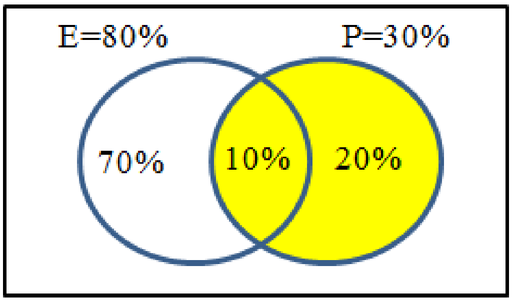


Figura 11. Distribución de los docentes por sectores

El tiempo de servicios de los docentes que conforman la muestra es muy similar, agrupados en rangos de 10 años, observamos que, aunque la diferencia no es muy significativa es un buen referente para afirmar que el magisterio actual encargado de la enseñanza de la matemática es joven 36,4%, comparado con los docentes de más experiencia 34%, y el 30% representa un significativo grupo de docentes con experiencia de 11 a 20 años de servicios.

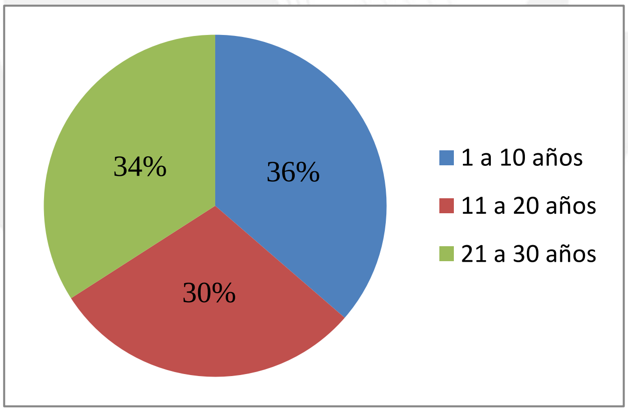


Figura 12. Distribución de los docentes por tiempo de servicios

Otro aspecto que describimos de la muestra está referido al grado que enseñan, observamos que la mayoría distribuyen su jornada de trabajo en varios grados como se muestra en la figura.

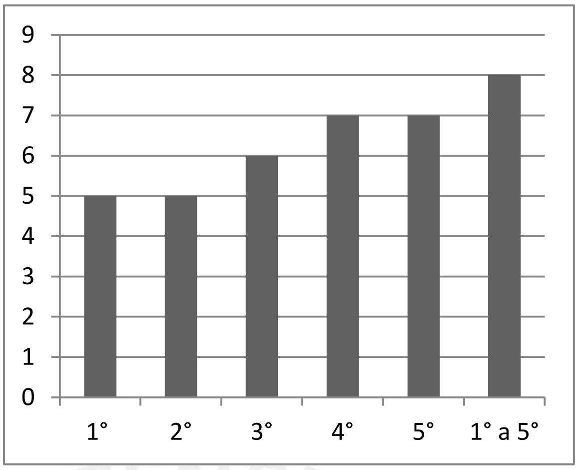


Figura 13. Distribución por grados que enseña

La mayoría de docentes recibieron formación profesional en Institutos Superiores Pedagógicos (ISP) 59%, mientras que el 41% lo hizo en una universidad.

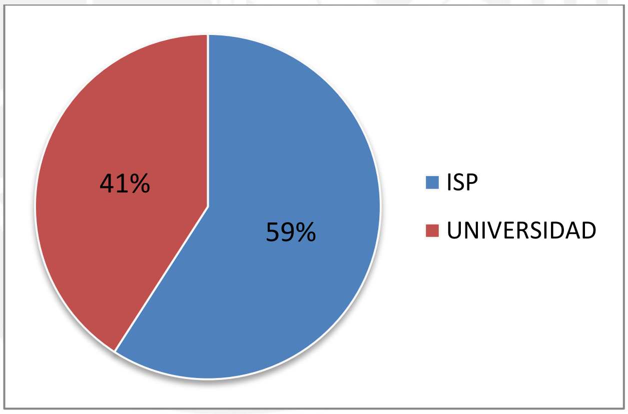


Figura 14. Institución superior de formación docente

3. Descripción del instrumento

El instrumento que se aplicó a los docentes es un cuestionario que según García (1992) lo define como:

> Un conjunto de preguntas sobre los hechos o aspectos que interesan en una investigación y son contestados por los encuestados. Se trata de un instrumento fundamental para la obtención de datos. [...] las preguntas que se hagan deben responder a la información que se desea obtener. No debe precipitarse el investigador en la confección del cuestionario porque es la pieza esencial en la obtención de los

> fines propuestos. [..] El cuestionario hace que todos los encuestados se encuentren en la misma situación psicológica, y además, que sus respuestas pueden ser comparadas. Para hacer un buen cuestionario la experiencia juega un gran papel ya que se ha considerado como un "arte" la confección de un cuestionario.
>
> El cuestionario se puede presentar bajo dos esquemas:
>
> - Cuestionario individual. Es el que el encuestado contesta de forma individual por escrito y sin que intervenga para nada el encuestador. Se suele enviar por correo y se presenta en forma de boletín o cuadernillo en donde se enumeran las preguntas dejando espacio para cada contestación.
> - Cuestionario-lista. Este cuestionario es preguntado por el encuestador en una entrevista por uno de los especialistas de la investigación. En una entrevista el encuestador va preguntando al encuestado, anotando las respuestas en unas hojas que contienen una especie de cuadriculas, reservando una columna cada pregunta y una fila a cada uno de los encuestados. (p.2)

El cuestionario está orientado a la evaluación del significado personal que tienen los docentes del área de matemática en ejercicio, respecto a la representatividad de la media aritmética y la mediana en un conjunto de datos estadísticos.

El cuestionario está compuesto de tres preguntas, la primera pregunta corresponde al cuestionario que aplicó Mayen (2009) con estudiantes mexicanos, relacionada a la representatividad de la media aritmética y la mediana con datos expresados en números decimales. El contexto del problema original se refiere a una clase de ciencias, nosotros lo hemos adaptado en una clase de ciencia tecnología y ambiente (CTA), asimismo el problema original incluye un valor muy grande respecto a los demás datos, lo que hace que a simple vista se reconozca como dato u observación atípica, en nuestro cuestionario este valor ha sido modificado por otro de menos medida con el propósito de crear una duda en el docente respecto a si es atípico o no y conlleve al docente a realizar el análisis de los datos que le permita determinar si el dato u observación es atípica o no. Otro cambio realizado al problema original ha sido reemplazar las alternativas a, b, c y d que inducen a determinar un método para determinar con mayor precisión el peso posible del objeto (moda o media aritmética), observamos que en la alternativa d, se sugiere desechar el valor muy grande y determinar la media de los datos que quedan. En nuestro cuestionario se plantean dos alternativas: a) recomendar usar la media aritmética y b) recomendar usar la mediana como medida representativa del conjunto de datos.

La segunda pregunta corresponde al cuestionario que aplicaron Garrett y García (2005) con estudiantes españoles y angoleños, la modificación que realizamos ha sido cambiar el dato 22 por el dato 10, para que no se aprecie a simple vista que el dato u observación es atípica, y esto permita realizar el análisis de los datos, por otro lado en se realizaron modificaciones en las alternativas que al igual que el primer problema induce a determinar el método que recomendaría usar: moda en la alternativa (a), usar media tomando en cuenta todos los datos en la alternativa (b) y usar media en las alternativas (c) y (d), después de descartar el dato muy pequeño o muy grande respectivamente. Nosotros hemos adaptado a dos alternativas que plantean qué medida estadística usar como mejor representante de los datos.

La tercera pregunta del cuestionario está relacionada al significado personal del concepto de sesgo estadístico en los docentes.

El cuestionario es individual, pues nos interesa conocer las concepciones que tienen los docentes de la especialidad de matemática en ejercicio sobre la determinación de datos u observaciones atípicos, propiedades estadísticas de la media aritmética y la mediana, representatividad y el sesgo estadístico de un conjunto de datos. Las preguntas son de respuesta abierta lo que permitirá obtener diferentes respuestas de los docentes a las cuestiones planteadas.

De acuerdo con Mayén (2009), el cuestionario es un instrumento de evaluación que tiene como finalidad proporcionar información sobre los significados personales de un grupo.

A continuación presentamos los problemas:

Problema 1:

Nueve estudiantes en una sesión de aprendizaje de CTA, pesaron un objeto (en gramos), obteniendo los resultados siguientes: 6.2; 6.0; 6.0; 9.3; 6.1; 6.3; 6.2; 6.15; 6.2. Los estudiantes esperan determinar si la media aritmética o la mediana es la más representativa del conjunto de datos. ¿Cuál le recomendaría usar?

a) Usar la media aritmética

b) Usar la mediana

Justifique su respuesta.

Es un problema extraído del trabajo de Mayén (2009), al que realizamos algunas adaptaciones. El problema inicial se refiere a una clase de ciencias, nosotros lo hemos adaptado a una sesión de aprendizaje de CTA, se había considerado como dato muy grande el número decimal 15,2. Nosotros lo hemos cambiado por el número decimal 9,3 con el propósito de generar la reflexión del docente y determine si el dato es atípico.

Solución del problema N° 1: Nueve estudiantes en una sesión de aprendizaje de CTA, pesaron un objeto (en gramos), obteniendo los resultados siguientes: 6.2; 6.0; 6.0; 9.3; 6.1; 6.3; 6.2; 6.15; 6.2. Los estudiantes esperan determinar si la media aritmética o la mediana es la más representativa del conjunto de datos. ¿Cuál le recomendaría usar?

i) Calculamos la media aritmética.

$$\bar{x} = \frac{2(6,0) + 6,1 + 6,15 + 3(6,2) + 6,3 + 9,3}{9}$$

$$\bar{x} = \frac{12 + 6,1 + 6,15 + 18,6 + 6,3 + 9,3}{9}$$

$$\bar{x} = \frac{58,45}{9} = 6,49444 \ldots g$$

ii) Calculamos la mediana, ordenando los datos.

6,0; 6,0; 6,1; 6,15; **6,2**; 6,2; 6,2; 6,3; 9,3

$$Me = 6,2 \text{ g}$$

Como la media aritmética es diferente a la mediana, averiguamos si existen datos atípicos en el conjunto de datos.

Para ello aplicamos el siguiente procedimiento:

- Si los datos pertenecen al intervalo $[Q_1 - D, Q_3 + D]$, significa que no existen datos atípicos, y por tanto la media aritmética es el mejor representante del conjunto de datos.
- Si algún dato no pertenece al intervalo $[Q_1 - D, Q_3 + D]$, entonces se dice que hay datos atípicos que sesgan la información y por lo tanto la mediana es el mejor representante del conjunto de datos.

Q_1 y Q_3 son los cuartiles primero y tercero respectivamente.

6,0; 6,0; 6,1; 6,15; **6,2;** 6,2; 6,2; 6,3; 9,3

⇩ ⇩ ⇩

Q_1 Q_2 Q_3

D, es la diferencia del $Q_3 - Q_1$ multiplicado por el factor 1,5.

$D = 1{,}5(Q_3 - Q_1)$ $\Rightarrow$ D = 1,5 (6,25 – 6,05) = 1,5 (0,20) = 0,30

Luego los datos deben pertenecer al intervalo:

$[Q_1 - D, Q_3 + D] \quad = \quad [6{,}15 - 0{,}30;\ 6{,}25 + 0{,}30] = \ [5{,}85;\ 6{,}55]$

Como vemos el dato 9,3 g no pertenecen al intervalo $[5{,}85\ g;\ 6{,}55\ g]$, lo que significa que es dato u observación atípica y por tanto existe sesgo en los datos.

CONCLUSIÓN:

Deducimos que el mejor representante del conjunto de datos es la mediana, Me=6,2 g.

Problema N° 2: La profesora María, registra la participación de los estudiantes quienes formulan preguntas referentes al tema tratado. Un registro del número de preguntas hechas por sus 8 estudiantes durante una clase se muestra en la siguiente tabla:

Tabla 25. Participación de los estudiantes

	INICIALES DEL ALUMNO							
Estudiantes	AA	RF	AG	JG	CK	NK	JL	AW
Número de preguntas	0	5	2	10	3	2	1	2

La profesora María, quiere determinar si la media aritmética o la mediana es la más representativa de los datos. ¿Cuál le recomendaría usar?

Solución:

i. Calculamos la media aritmética.

$$\bar{x} = \frac{0 + 1 + 3(2) + 3 + 5 + 10}{8}$$

$$\bar{x} = \frac{25}{8}$$

$$\bar{x} = 3{,}125$$

iii) Calculamos la mediana, ordenando los datos.

0; 1; 2; **2;** **2**; 3; 5; 10

$$Me = \frac{2 + 2}{2} = 2$$

Como la media aritmética es diferente a la mediana, averiguamos si existen datos atípicos en el conjunto de datos.

Para ello aplicamos el siguiente procedimiento:

- Si los datos pertenecen al intervalo $[Q_1 - D, Q_3 + D]$, significa que no existen datos atípicos, y por tanto la media aritmética es el mejor representante del conjunto de datos.
- Si algún dato no pertenece al intervalo $[Q_1 - D, Q_3 + D]$, entonces se dice que hay datos atípicos que sesgan la información y por lo tanto la mediana es el mejor representante del conjunto de datos.

Q_1 y Q_3 son los cuartiles primero y tercero respectivamente.

0; 1; 2; **2;** **2**; 3; 5; 10

$\Downarrow$ $\Downarrow$ $\Downarrow$

Q_1 Q_2 Q_3

D, es la diferencia del $Q_3 - Q_1$ multiplicado por el factor 1,5.

$D = 1{,}5(Q_3 - Q_1) \Rightarrow$ D = 1,5 (4 – 1,5) = 1,5 (2,5) = 3,75

Luego los datos deben pertenecer al intervalo:

$[Q_1 - D, Q_3 + D] \quad = \quad [1{,}5 - 3{,}75\,; 4 + 3{,}75] = [-2{,}25\,;\ 7{,}75]$

Como vemos el dato 10 no pertenece al intervalo $[-2{,}25;\ 7{,}75]$, lo que significa que hay datos atípicos y por tanto existe sesgo en los datos.

CONCLUSIÓN:

Deducimos que el mejor representante del conjunto de datos es la mediana,

Me$=2$.

Pregunta N° 3: Se han obtenido un conjunto de datos de una muestra estadística, y se observa que algunos datos están “alejados” del resto. Según lo anterior, ¿qué entiende Ud. por sesgo estadístico?

Solución:

El sesgo estadístico es la inclinación o asimetría de un conjunto de datos por la presencia de datos atípicos.

4. Resultados

A la situación problemática N° 1:

Nueve estudiantes en una sesión de aprendizaje de CTA, pesaron un objeto obteniendo los resultados siguientes: 6.2; 6.0; 6.0; 9.3; 6.1; 6.3; 6.2; 6.15; 6.2. Los estudiantes esperan determinar si la media aritmética o la mediana es la más representativa del conjunto de datos. ¿Cuál le recomendaría usar?

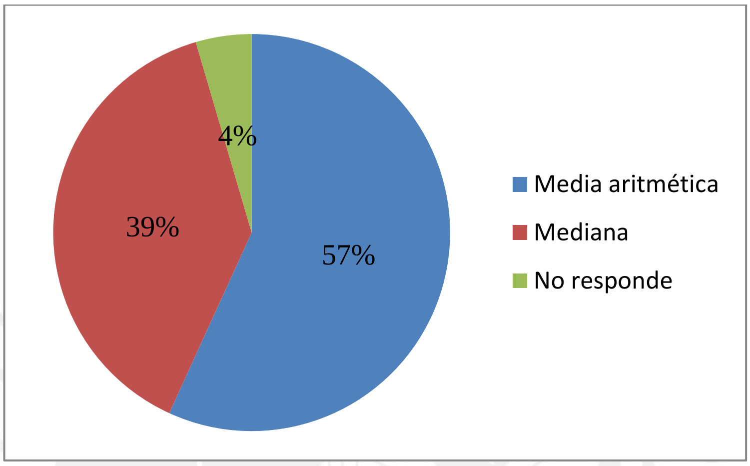


Figura 15. Respuesta de los docentes al problema 1

El 57% de los docentes de la muestra opina que la media aritmética es el mejor representante del conjunto de datos, de los cuales: (Ver figura 14)

- El 44% responde porque es un promedio más próximo a los datos, porque es lo mejor para su solución, porque pedía obtener un promedio, porque en muestras continuas la media aritmética es más representativa, porque está en el conjunto de datos o muestra, porque la media aritmética es el promedio seguro, porque es común trabajar con la media aritmética, porque los datos no son muy homogéneos, porque permite visualizar la respuesta de una manera exacta y rápida y son valores continuos
- El 4% considera que permitirá obtener un dato promedio de todos los datos al sumar y dividir la cantidad de datos.
- El 20% justifica su respuesta porque es exacta y rápida

- El 16% usa la media aritmética, porque no considera la presencia de datos atípicos, atribuyendo una propiedad que no es propia de la media aritmética.
- El 16% no justifica la elección.

Esta respuesta es incorrecta.

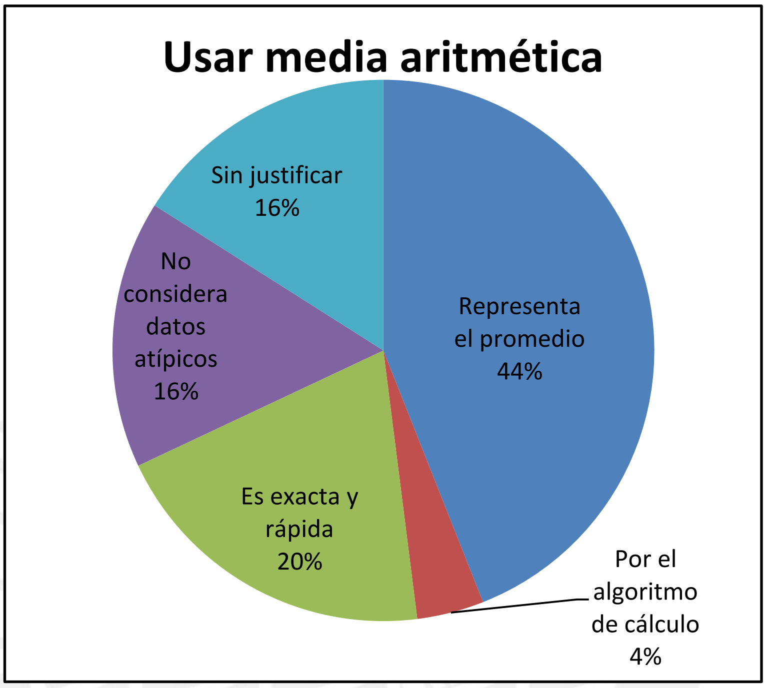


Figura 16. Justificaciones de los docentes al usar la media aritmética en el problema 1

El 39% opina que es la mediana el mejor representante, de los cuales: (Ver figura 15)

- El 37% responde que existe sesgo en los datos, sin embargo no aplica un procedimiento que lo determine con seguridad, dejándose llevar por la simple observación.
- El 13% porque la considera más factible.
- El 6% considera que se usa para sacar promedios.
- Otro 6% de los docentes la elige porque es un valor exacto.
- Un 6% más considera que resulta rápido ordenar los datos y porque es la más cercana a los datos anteriores y posteriores y además, porque es un valor que se aproxima al peso real del objeto.
- El 13% la considera por las fallas en los datos y porque la media aritmética está sesgada.
- El 13% responde que es más representativa sin decir a que se debe.

- El 6% no justifica las razones de su elección.

Aunque la respuesta es correcta, algunas de las justificaciones no son las apropiadas para afirmar tal representatividad.

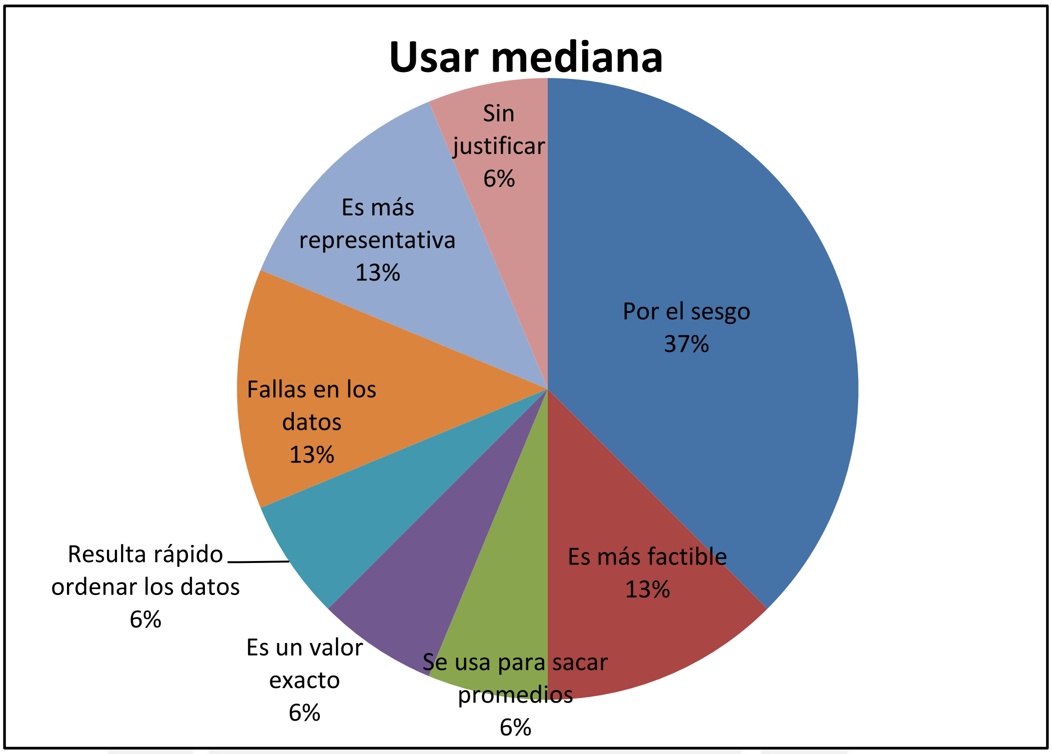


Figura 17. Justificaciones de los docentes al usar la mediana en el problema 1

Un 4% de los encuestados no respondió la pregunta.

A la pregunta N° 2:

La profesora María, registra la participación de los estudiantes quienes formulan preguntas referentes al tema tratado. Un registro del número de preguntas hechas por sus 8 estudiantes durante una clase se muestra en la siguiente tabla:

Tabla 26. Participación de los estudiantes

	INICIALES DEL ALUMNO							
Estudiantes	AA	RF	AG	JG	CK	NK	JL	AW
Número de preguntas	0	5	2	10	3	2	1	2

La profesora María, quiere determinar si la media aritmética o la mediana es la más representativa de los datos. ¿Cuál le recomendaría usar?

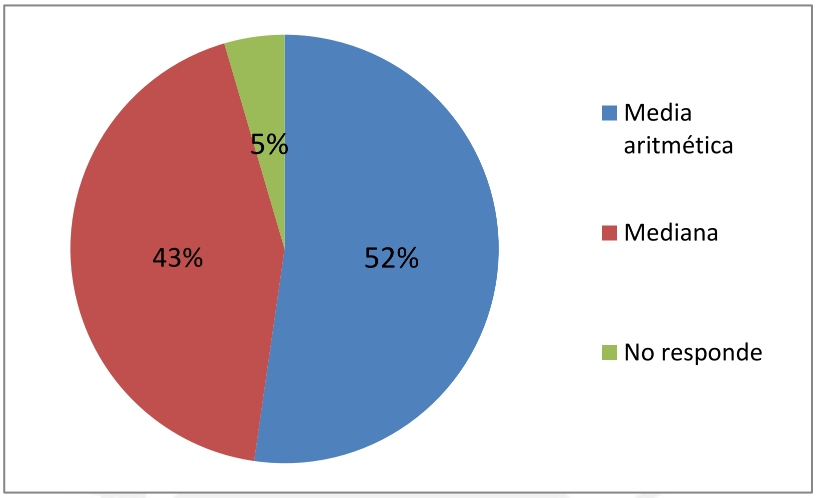


Figura 18. Respuestas de los docentes al problema 2

El 52% de los docentes de la muestra opina que la media aritmética es el mejor representante del conjunto de datos, de los cuales: (ver figura 16)

- El 27% manifiesta que siempre es la más representativa al promedio.
- 9% considera por el dato atípico.
- 14% cree es más aproximada o cercana a los datos.
- 9% porque se aproxima bastante al orden de datos y registra cantidades exactas.
- El 4% responde que es variable cualitativa.
- El 5% responde que la presencia de valores extremos o datos atípicos, no son considerados para determinar la representatividad de la media aritmética.
- El 32% no justifica la elección.

La respuesta es incorrecta y los argumentos no son los adecuados.

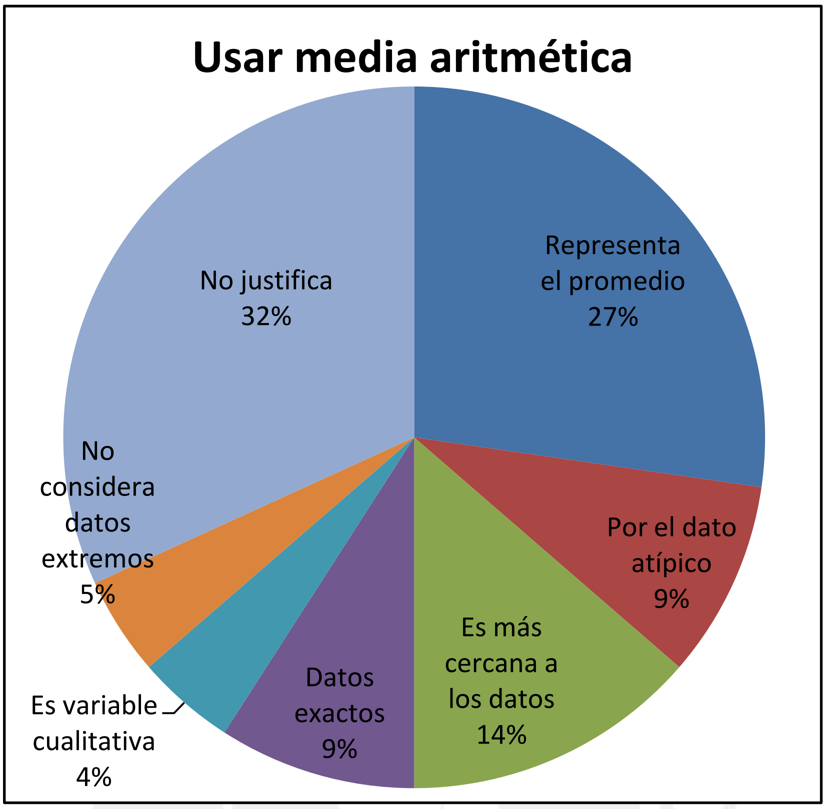


Figura 19. Justificaciones de los docentes al usar la media aritmética en el problema 2

El 43% opina que es la mediana la medida más representativa, de los cuales:

- El 29% responde que la representatividad se debe por el sesgo.
- El 6% por la presencia de datos o valores atípicos.
- El 6% por que la mediana resulta ser un valor entero y la media aritmética un valor decimal.
- El 35% considera que es representativa por los datos atípicos o sesgo, esto se da porque los datos extremos distorsionan la media aritmética, sin explicar cómo determinó esto.
- El 12% opina que la mediana es más exacta, por la heterogeneidad de los datos, por tratarse de variables cualitativas y porque el valor obtenido de la mediana está en el conjunto de datos.
- El 6% piensa que se puede demostrar.
- El 6% no contestó la pregunta.

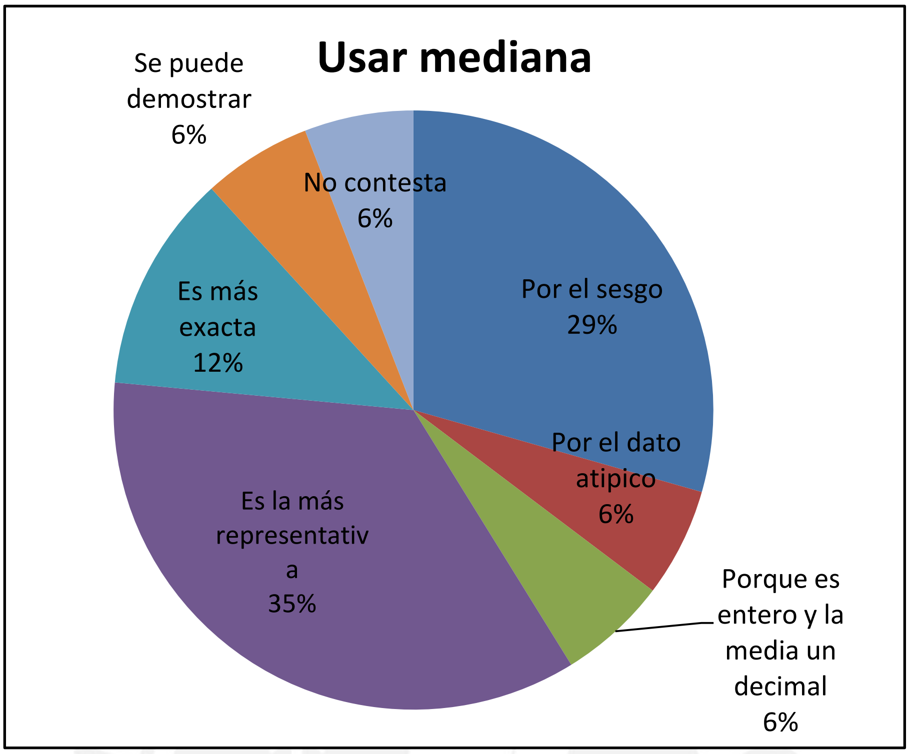


Figura 20. Justificaciones de los docentes al usar la mediana en el problema 2

El 5% de los encuestados no respondió qué medida usar, por no comprender la relación entre las condiciones del problema y los datos brindados.

En la pregunta N° 3:

Se han obtenido un conjunto de datos de una muestra estadística, y se observa que algunos datos están "alejados" del resto. Según lo anterior, ¿qué entiende Ud. por sesgo estadístico?

- El 27% c considera el sesgo como datos alejados o sesgados.
- El 2% considera que sesgo es la presencia de datos atípicos.
- El 4% son datos que están más cerca.
- El 16% cree que se debe a datos que no cumplen con el parámetro normal.
- El 7% considera que es una diferencia entre la esperanza matemática y los parámetros del desarrollo.
- Para el 23% el sesgo son datos con error.
- Para el 5% es una variación de las frecuencias.
- Otro 5% piensa que es la asimetría de las medidas de tendencia central.

- El 11% de docentes no sabe o no responde.

Uno de los docentes coincide correctamente con la idea de sesgo que representa el 2% del total de docentes encuestados.

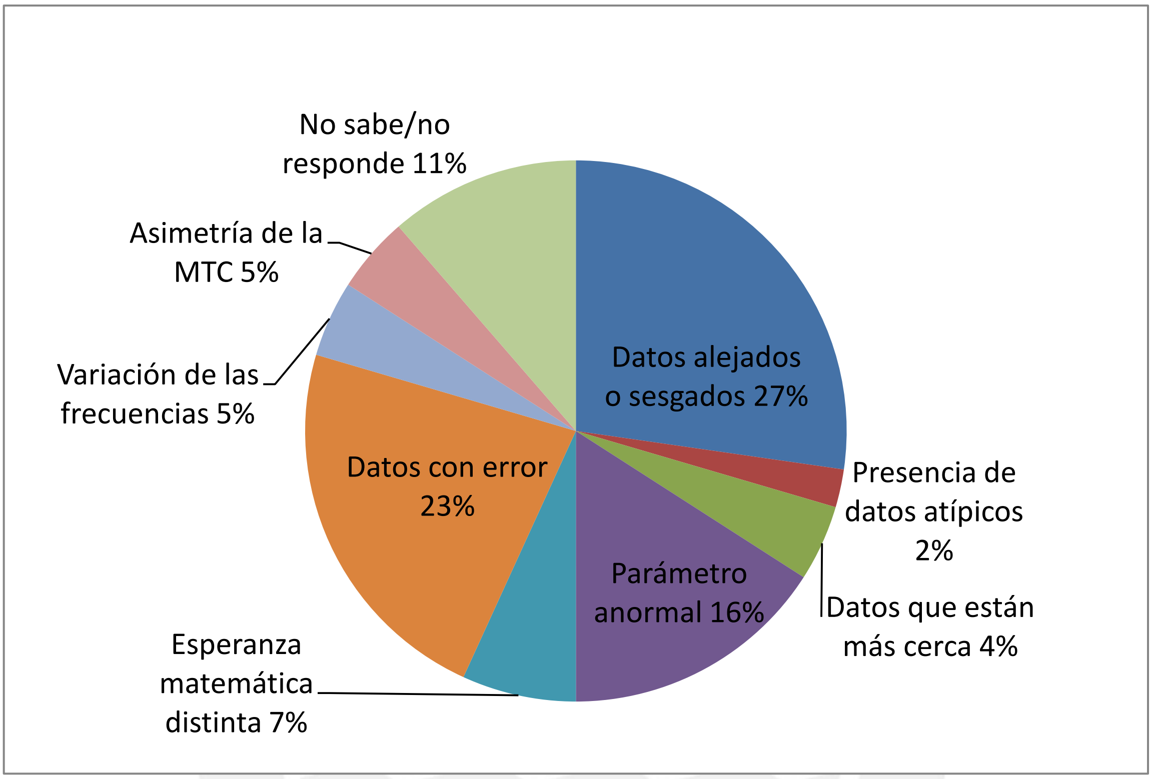


Figura 21. Respuesta de los docentes al problema 3

A continuación analizamos las respuestas de 6 docentes distribuidos de la siguiente manera:

- 2 docentes que respondieron recomendar el uso de la media aritmética como la medida más representativa del conjunto de datos.

Tabla 27. Respuestas del docente 1

Problema 1	Problema 2
Recomendaría usar la media aritmética, porque obtenemos el promedio seguro del peso del objeto, en cambio en la mediana los datos tienen una diferencia de 3,3 entre el primer y último dato por lo cual no es recomendable.	Les recomendaría usar la media aritmética, porque representa el promedio de preguntas del total de participantes. En cambio en la mediana hay una diferencia entre el primer dato y el último de 10, lo que dificulta encontrar la representación de los datos.

El docente reconoce como representante del conjunto de datos a la media aritmética, atribuyendo como propiedad de la mediana, "La diferencia de los datos extremos" no la hace representativa. Observamos que esta propiedad no corresponde a la mediana y es incorrecta. La respuesta es también incorrecta.

Tabla 28. Respuestas del docente 2

Problema 1	Problema 2
Recomendaría usar la media aritmética, ya que elle le permitirá visualizar el peso promedio del objeto y será el que más se acerque, en este caso, a lo solicitado ya que hay resultados distintos al pesar el objeto.	Les recomendaría usar la media aritmética, ya que ella representa el promedio máximo de preguntas hechas por los estudiantes. A diferencia de la mediana que representa el valor central.

El docente reconoce como representante del conjunto de datos a la media aritmética, considerándola como representante del promedio y solo reconoce a la mediana como una medida central que no es representativa de los datos.

Se cree que la media aritmética siempre representa al conjunto de datos y es una medida promedio. No reconoce que la mediana también es representante del conjunto de datos bajo ciertos criterios. La respuesta es incorrecta.

Tabla 29. Respuestas del docente 3

Problema 1	Problema 2
Recomendaría usar la mediana, por ser más representativa y más cercana a los datos anteriores y posteriores.	Les recomendaría usar la mediana por ser más representativa.

El docente reconoce que la mediana representa al conjunto de datos tomando en cuenta que antes y después de esta medida quedan igual cantidad de datos. Si bien es correcta su respuesta no argumenta la representatividad en función de la propiedad estadística de la mediana en presencia de datos atípicos.

Tabla 30. Respuestas del docente 4

Problema 1	Problema 2
Recomendaría usar la mediana, porque es la cantidad más representativa y está en el conjunto de datos o muestra.	

El docente reconoce que la mediana representa al conjunto de datos, argumentando que su valor figura como dato y no por otras razones. Esta afirmación es coherente para ambos problemas y eso es rescatable, sin embargo, si bien es correcta su respuesta los argumentos no son apropiados.

El análisis de estos resultados servirá para elaborar la secuencia didáctica.

Tabla 31. Respuestas del docente 5

Problema 1	Problema 2
Recomendaría usar la media aritmética, porque el valor obtenido representa al conjunto de datos, pues, casi todos pesaron 6 con diferencia en los décimos y centésimos.	Les recomendaría usar la mediana porque se quiere determinar el número de preguntas realizadas por los estudiantes, pudiéndose observar un promedio de 2 preguntas realizadas por cada estudiante.

El docente recomienda usar la media aritmética en el primer problema, porque observa que casi todos los datos tienen como parte entera 6 a excepción de un dato, sin embargo no se preocupó en analizar si el dato diferente es atípico. En el segundo problema justifica la elección de la mediana como representante del conjunto de datos, tomando en cuenta la moda de las preguntas hechas por los estudiantes y no por alguna propiedad estadística de la mediana. Si bien es correcta su respuesta los argumentos no son apropiados.

Tabla 32. Respuestas del docente 6

Problema 1	Problema 2
Recomendaría usar la media aritmética, porque me parece la más representativa, porque es el promedio de todos los datos, para este caso que la variable es el peso.	Les recomendaría usar la mediana, porque me parece la más representativa, ya que los datos (población) son bastante heterogéneos. Hay mínimos y máximos.

El docente recomienda usar la media aritmética en el primer problema y la mediana en el segundo problema, sin estar seguro de ello. Relaciona la media aritmética como la única medida promedio de los datos. Esta respuesta es incorrecta.

Al responder la pregunta N° 2, recomienda usar la mediana teniendo en cuenta los datos extremos que aunque no lo dice abiertamente que son atípicos los toma en cuenta para decidir su elección. Otro aspecto importante es la heterogeneidad de los datos. No expresa reconocer datos u observaciones atípicas ni tampoco propiedades estadísticas de la media aritmética y la mediana. Si bien es correcta su respuesta los argumentos carecen del uso y manejo de propiedades estadísticas.