SOLUCIÓN MATEMÁTICA DETALLADA

PRIMER PARCIAL - SECCIÓN S609 - VARIANTE V1

Fecha de generación: 13 de March de 2025, 16:02
PARA USO EXCLUSIVO DEL DOCENTE - DOCUMENTO DE VERIFICACIÓN MATEMÁTICA

Este documento contiene soluciones matemáticas paso a paso con verificación cruzada para garantizar precisión absoluta en los cálculos.

ÍNDICE

PRIMERA SERIE - JUSTIFICACIÓN DE RESPUESTAS

Las siguientes justificaciones proporcionan el fundamento teórico y matemático para cada una de las respuestas correctas de la primera serie. Estas explicaciones pueden utilizarse para clarificar dudas y como material didáctico complementario:

Pregunta 1: ¿Es un conjunto representativo de la población de referencia, el número de individuos es menor que el de la población?

Opciones disponibles:

• Valor

• Dato

• Experimento

• Población

• Muestra

• Todas las anteriores

Justificación:
En estadística, la población es el conjunto completo de elementos (personas, objetos, medidas) sobre los que se realiza el estudio. La muestra es un subconjunto representativo de la población, seleccionado para inferir características de la población total.

La correcta distinción entre estos conceptos es fundamental para el diseño de estudios estadísticos válidos y para la aplicación apropiada de técnicas de inferencia.

Pregunta 2: ¿Cuál es el método que permite calcular el número de grupos, intervalos o clases a construer para una table de distribución de frecuencias?

Opciones disponibles:

• Método de mínimos cuadrados

• Coeficiente de Gini

• Método Sturgers

• La regla empírica

Justificación:
La respuesta correcta es 'Método Sturgers' porque es la definición precisa según los conceptos estadísticos fundamentales.

Esta respuesta se alinea con los principios establecidos en la teoría estadística y representa la interpretación correcta del concepto consultado.

Pregunta 3: ¿La siguiente imagen, representa un diagrama de tallo y hoja?

Opciones disponibles:

• **Verdadero**

• Falso

Justificación:
La respuesta correcta es 'Verdadero' porque es la definición precisa según los conceptos estadísticos fundamentales.

Esta respuesta se alinea con los principios establecidos en la teoría estadística y representa la interpretación correcta del concepto consultado.

Pregunta 4: Si los datos están distribuidos de forma simétrica alrededor de la media, entonces:

Opciones disponibles:

• **La media y la mediana coinciden**

• La media es mayor que la mediana

• La mediana es mayor que la media

• La media y la moda coinciden

Justificación:
La respuesta correcta es 'La media y la mediana coinciden' porque es la definición precisa según los conceptos estadísticos fundamentales.

Esta respuesta se alinea con los principios establecidos en la teoría estadística y representa la interpretación correcta del concepto consultado.

Pregunta 5: ¿Cuál de las siguientes medidas de tendencia central se ve más afectada por valores extremos?

Opciones disponibles:

• **Media**

• Mediana

• Moda

• Rango

Justificación:
Las medidas de tendencia central (media, mediana y moda) describen el centro de una distribución de datos. La media aritmética es el promedio de todos los valores y se calcula sumando todos los datos y dividiendo por el número total.

La media es particularmente sensible a valores extremos (outliers) porque estos afectan directamente la suma total. Por ejemplo, en la distribución [1, 2, 3, 4, 100], la media es 22, un valor que no representa adecuadamente la centralidad de los datos.

En contraste, la mediana no se ve afectada por valores extremos ya que solo considera la posición central de los datos ordenados.

Pregunta 6: Método que sirve para medir la desigualdad, es un número entre cero y uno que mide el grado de desigualdad en la distribución del ingreso en una sociedad determinada o país.

Opciones disponibles:

• Coeficiente de Correlación

• Coeficiente de Gini

• Marca de Clase

• La Frecuencia Acumulada

Justificación:
La respuesta correcta es 'Coeficiente de Gini' porque es la definición precisa según los conceptos estadísticos fundamentales.

Esta respuesta se alinea con los principios establecidos en la teoría estadística y representa la interpretación correcta del concepto consultado.

Pregunta 7: Las Fases de un estudio estadístico son:

Opciones disponibles:

• Planteamiento del Problema

• Simplificar los Datos

• Recolectar y Ordenar los Datos

• Analizar los Datos

• Interpretar y Presentar Resultados

• Ninguna de las anteriores

Justificación:
La respuesta correcta es 'Planteamiento del Problema' porque es la definición precisa según los conceptos estadísticos fundamentales.

Esta respuesta se alinea con los principios establecidos en la teoría estadística y representa la interpretación correcta del concepto consultado.

Pregunta 8: La diferencia entre una variable cuantitativa discreta y una variable cuantitativa continua es:

Opciones disponibles:

• Las variables discretas toman cualquier valor dentro de un intervalo, las continuas toman valores aislados

• Las variables discretas toman valores aislados, las continuas toman cualquier valor dentro de un intervalo

• Las variables discretas son siempre enteras, las continuas son siempre decimales

• No hay diferencia real entre ambas

Justificación:
La respuesta correcta es 'Las variables discretas toman valores aislados, las continuas toman cualquier valor dentro de un intervalo' porque es la definición precisa según los conceptos estadísticos fundamentales.

Esta respuesta se alinea con los principios establecidos en la teoría estadística y representa la interpretación correcta del concepto consultado.

Pregunta 9: La toma de temperatura para ingresar a los centros comerciales es una variable:

Opciones disponibles:

• Cualitativa

• Cuantitativa

Justificación:
La respuesta correcta es 'Cuantitativa' porque es la definición precisa según los conceptos estadísticos fundamentales.

Esta respuesta se alinea con los principios establecidos en la teoría estadística y representa la interpretación correcta del concepto consultado.

Pregunta 10: ¿Quién ordeno o realizo el primer catastro o Censo de (bienes inmuebles) considerado el primero en Europa?

Opciones disponibles:

• El Rey Juan Carlos

• El Rey Guillermo

• El Rey Constantino

• El Rey Ricardo

• El Rey Federico de Edimburgo

Justificación:
La respuesta correcta es 'El Rey Guillermo' porque es la definición precisa según los conceptos estadísticos fundamentales.

Esta respuesta se alinea con los principios establecidos en la teoría estadística y representa la interpretación correcta del concepto consultado.

SEGUNDA SERIE - JUSTIFICACIÓN DE SELECCIÓN DE GRÁFICOS

Para cada escenario, se proporciona una justificación matemática y estadística detallada sobre por qué el tipo de gráfico seleccionado es el más apropiado:

Escenario 1:

Una compañía de telecomunicaciones quiere representar visualmente la distribución porcentual de sus ingresos por tipo de servicio (internet, telefonía fija, telefonía móvil, televisión por cable y servicios corporativos) durante el año fiscal 2023.

Opciones disponibles:

• Gráfica de barras

• Gráfica circular (pastel)

• Histograma de Pearson

• Ojiva de Galton

• Polígono de frecuencias

Justificación matemática:
La gráfica circular (o de pastel) es matemáticamente apropiada cuando se necesita visualizar partes de un todo y la contribución proporcional de cada categoría al total (100%). Cada sector del círculo tiene un ángulo proporcional al valor que representa.

Para cada categoría i, el ángulo correspondiente se calcula como: $\theta i = (valor\_i / total) \times 360°$

Este tipo de gráfico es óptimo para mostrar distribuciones porcentuales, ya que la suma de todos los sectores completa visualmente el círculo (100%). Esto facilita la comprensión inmediata de la importancia relativa de cada componente.

Razones para descartar las otras opciones:
• Gráfica de barras: No es adecuada porque los datos requieren mostrar proporciones de un todo o representar una variable continua, no una comparación entre categorías discretas.
• Histograma de Pearson: No es óptima porque los datos no son continuos o porque el objetivo no es analizar la distribución de frecuencias de una variable continua.
• Ojiva de Galton: No es la mejor opción porque el objetivo no es analizar valores acumulados o percentiles en la distribución.
• Polígono de frecuencias: No es ideal porque los datos no representan una tendencia o evolución, o porque no se busca enfatizar la continuidad y cambios graduales.

Escenario 2:

Una empresa farmacéutica ha registrado el tiempo (en días) que tarda cada lote de medicamentos en pasar el control de calidad. Quieren determinar si un nuevo lote con un tiempo específico está dentro del 75% de los casos más rápidos.

Opciones disponibles:

• Gráfica de barras

• Gráfica circular (pastel)

• Histograma de Pearson

• Ojiva de Galton

• Polígono de frecuencias

Justificación matemática:
La Ojiva de Galton es una representación gráfica de la función de distribución acumulativa empírica. Matemáticamente, para cada punto x, la ojiva muestra $F(x) = P(X \leq x)$, es decir, la probabilidad de que la variable tome un valor menor o igual a x.

Esta representación es particularmente útil para determinar cuantiles y percentiles. Si queremos encontrar el valor x tal que $F(x) = p$ (donde p es una proporción), simplemente localizamos p en el eje vertical y leemos el valor correspondiente x en el eje horizontal.

La pendiente de la curva en cualquier punto refleja la densidad de observaciones en ese rango, proporcionando información adicional sobre la distribución de los datos.

Razones para descartar las otras opciones:
• Gráfica de barras: No es adecuada porque los datos requieren mostrar proporciones de un todo o representar una variable continua, no una comparación entre categorías discretas.
• Gráfica circular (pastel): No es apropiada porque los datos no representan proporciones de un todo o porque hay demasiadas categorías, lo que dificultaría la interpretación visual.
• Histograma de Pearson: No es óptima porque los datos no son continuos o porque el objetivo no es analizar la distribución de frecuencias de una variable continua.
• Polígono de frecuencias: No es ideal porque los datos no representan una tendencia o evolución, o porque no se busca enfatizar la continuidad y cambios graduales.

Escenario 3:

Una entidad financiera ha recopilado datos sobre los montos de créditos otorgados en el último trimestre. Los montos se han agrupado en intervalos y se desea mostrar los valores acumulados hasta cierto punto, para identificar qué porcentaje de créditos está por debajo de determinados montos.

Opciones disponibles:

• Gráfica de barras

• Gráfica circular (pastel)

• Histograma de Pearson

• Ojiva de Galton

• Polígono de frecuencias

Justificación matemática:
La Ojiva de Galton es una representación gráfica de la función de distribución acumulativa empírica. Matemáticamente, para cada punto x, la ojiva muestra $F(x) = P(X \leq x)$, es decir, la probabilidad de que la variable tome un valor menor o igual a x.

Esta representación es particularmente útil para determinar cuantiles y percentiles. Si queremos encontrar el valor x tal que $F(x) = p$ (donde p es una proporción), simplemente localizamos p en el eje vertical y leemos el valor correspondiente x en el eje horizontal.

La pendiente de la curva en cualquier punto refleja la densidad de observaciones en ese rango, proporcionando información adicional sobre la distribución de los datos.

Razones para descartar las otras opciones:
• Gráfica de barras: No es adecuada porque los datos requieren mostrar proporciones de un todo o representar una variable continua, no una comparación entre categorías discretas.
• Gráfica circular (pastel): No es apropiada porque los datos no representan proporciones de un todo o porque hay demasiadas categorías, lo que dificultaría la interpretación visual.
• Histograma de Pearson: No es óptima porque los datos no son continuos o porque el objetivo no es analizar la distribución de frecuencias de una variable continua.
• Polígono de frecuencias: No es ideal porque los datos no representan una tendencia o evolución, o porque no se busca enfatizar la continuidad y cambios graduales.

Escenario 4:

Un departamento de recursos humanos ha realizado una encuesta sobre los tiempos de transporte (en minutos) que los empleados tardan en llegar a la oficina. Los datos obtenidos son continuos y quieren mostrar cómo se distribuyen estos tiempos, identificando claramente dónde se concentra la mayoría de los casos.

Opciones disponibles:

• Gráfica de barras

• Gráfica circular (pastel)

• Histograma de Pearson

• Ojiva de Galton

• Polígono de frecuencias

Justificación matemática:
El histograma es la representación matemática idónea para variables continuas agrupadas en intervalos o clases. A diferencia del gráfico de barras, en el histograma los rectángulos son contiguos, reflejando la

continuidad de la variable subyacente.

La construcción matemática implica dividir el rango de datos [min, max] en k intervalos de clase, generalmente usando la fórmula de Sturges: $k = 1 + 3.322 \log_{10}(n)$, donde n es el número de observaciones.

La altura de cada rectángulo representa la frecuencia o densidad de observaciones en ese intervalo. El área total del histograma es proporcional al número total de datos, lo que permite visualizar la forma de la distribución y identificar características como la normalidad, asimetría o multimodalidad.

Razones para descartar las otras opciones:
• Gráfica de barras: No es adecuada porque los datos requieren mostrar proporciones de un todo o representar una variable continua, no una comparación entre categorías discretas.
• Gráfica circular (pastel): No es apropiada porque los datos no representan proporciones de un todo o porque hay demasiadas categorías, lo que dificultaría la interpretación visual.
• Ojiva de Galton: No es la mejor opción porque el objetivo no es analizar valores acumulados o percentiles en la distribución.
• Polígono de frecuencias: No es ideal porque los datos no representan una tendencia o evolución, o porque no se busca enfatizar la continuidad y cambios graduales.

Escenario 5:

Un estudio sobre calificaciones finales en un curso de estadística muestra datos que podrían seguir una distribución normal. Los investigadores quieren representar las frecuencias de cada intervalo de calificación y, al mismo tiempo, identificar visualmente si la distribución se aproxima a una curva normal.

Opciones disponibles:

• Gráfica de barras

• Gráfica circular (pastel)

• Histograma de Pearson

• Ojiva de Galton

• Polígono de frecuencias

Justificación matemática:
El histograma es la representación matemática idónea para variables continuas agrupadas en intervalos o clases. A diferencia del gráfico de barras, en el histograma los rectángulos son contiguos, reflejando la continuidad de la variable subyacente.

La construcción matemática implica dividir el rango de datos [min, max] en k intervalos de clase, generalmente usando la fórmula de Sturges: $k = 1 + 3.322 \log_{10}(n)$, donde n es el número de observaciones.

La altura de cada rectángulo representa la frecuencia o densidad de observaciones en ese intervalo. El área total del histograma es proporcional al número total de datos, lo que permite visualizar la forma de la distribución y identificar características como la normalidad, asimetría o multimodalidad.

Razones para descartar las otras opciones:
• Gráfica de barras: No es adecuada porque los datos requieren mostrar proporciones de un todo o representar una variable continua, no una comparación entre categorías discretas.
• Gráfica circular (pastel): No es apropiada porque los datos no representan proporciones de un todo o porque hay demasiadas categorías, lo que dificultaría la interpretación visual.
• Ojiva de Galton: No es la mejor opción porque el objetivo no es analizar valores acumulados o percentiles en la distribución.
• Polígono de frecuencias: No es ideal porque los datos no representan una tendencia o evolución, o porque no se busca enfatizar la continuidad y cambios graduales.

Escenario 6:

Un instituto de estadísticas demográficas ha recopilado información sobre las edades de los habitantes de un municipio, agrupando los datos en intervalos de 10 años (0-9, 10-19, 20-29, etc.). Desean visualizar tanto la frecuencia de cada intervalo como la tendencia general de la distribución de edades.

Opciones disponibles:

• Gráfica de barras

• Gráfica circular (pastel)

• Histograma de Pearson

• Ojiva de Galton

• Polígono de frecuencias

Justificación matemática:
El polígono de frecuencias es una representación mediante líneas continuas que conectan las frecuencias representadas por las marcas de clase. Matemáticamente, es una aproximación a la función de densidad de probabilidad subyacente.

Se construye uniendo con segmentos de recta los puntos (xi, fi), donde xi es la marca de clase (punto medio del intervalo) y fi es la frecuencia correspondiente. Esta representación es particularmente útil para visualizar tendencias y patrones en datos continuos ordenados, especialmente cuando existen series temporales.

La primera derivada del polígono en cualquier punto proporciona la tasa de cambio de la frecuencia respecto a la variable, lo que permite identificar visualmente dónde el crecimiento es más rápido o más lento.

Razones para descartar las otras opciones:
• Gráfica de barras: No es adecuada porque los datos requieren mostrar proporciones de un todo o representar una variable continua, no una comparación entre categorías discretas.
• Gráfica circular (pastel): No es apropiada porque los datos no representan proporciones de un todo o porque hay demasiadas categorías, lo que dificultaría la interpretación visual.
• Histograma de Pearson: No es óptima porque los datos no son continuos o porque el objetivo no es analizar la distribución de frecuencias de una variable continua.

• Ojiva de Galton: No es la mejor opción porque el objetivo no es analizar valores acumulados o percentiles en la distribución.

TERCERA SERIE - EJERCICIOS PRÁCTICOS

Ejercicio 1: Coeficiente de Gini

DATOS DEL PROBLEMA:

Distribución de salarios mensuales:

Salario mensual en (Q)	No. De trabajadores
[1200-1800)	12
[1800-2400)	6
[2400-3000)	9
[3000-3600)	4
[3600-4200)	2
[4200-4800)	5
[4800-5400)	3
[5400-6000)	7

MÉTODO DE CÁLCULO:

El coeficiente de Gini es una medida de desigualdad que toma valores entre 0 y 1. Un valor de 0 representa igualdad perfecta y un valor de 1 representa desigualdad máxima.

PASO 1: Preparación de datos

Intervalo salarial	**No. trabajadores**	**Prop. población**	**Prop. acum. pobl.**	**Punto medio**	**Ingreso total**	**Prop. ingreso**	**Prop. acum. ingreso**
[1200-1800)	12	0.250000	0.250000	1500.00	18000.00	0.117188	0.117188
[1800-2400)	6	0.125000	0.375000	2100.00	12600.00	0.082031	0.199219
[2400-3000)	9	0.187500	0.562500	2700.00	24300.00	0.158203	0.357422
[3000-3600)	4	0.083333	0.645833	3300.00	13200.00	0.085938	0.443359
[3600-4200)	2	0.041667	0.687500	3900.00	7800.00	0.050781	0.494141
[4200-4800)	5	0.104167	0.791667	4500.00	22500.00	0.146484	0.640625
[4800-5400)	3	0.062500	0.854167	5100.00	15300.00	0.099609	0.740234
[5400-6000)	7	0.145833	1.000000	5700.00	39900.00	0.259766	1.000000
TOTAL	48	1.000000			153600.00	1.000000	1.000000

VERIFICACIONES DE CÁLCULOS:

1. Suma de trabajadores: 48 = 48 (Exacto)
2. Proporción acumulada de población: 1.000000 ≈ 1.0 (Correcto)
3. Proporción acumulada de ingresos: 1.000000 ≈ 1.0 (Correcto)

PASO 2: Cálculo del coeficiente de Gini

Para calcular el coeficiente de Gini, utilizamos la fórmula basada en la curva de Lorenz:

$$G = 1 - \Sigma[(Xi - Xi\text{-}1)(Yi + Yi\text{-}1)]$$

Donde:
- Xi = proporción acumulada de población en el grupo i
- Yi = proporción acumulada de ingreso en el grupo i

CÁLCULO DETALLADO:

Grupo	**Xi**	**Xi-1**	**Yi**	**Yi-1**	**(Xi - Xi-1)(Yi + Yi-1)**
1	0.250000	0.000000	0.117188	0.000000	0.029297
2	0.375000	0.250000	0.199219	0.117188	0.039551
3	0.562500	0.375000	0.357422	0.199219	0.104370
4	0.645833	0.562500	0.443359	0.357422	0.066732
5	0.687500	0.645833	0.494141	0.443359	0.039062
6	0.791667	0.687500	0.640625	0.494141	0.118205
7	0.854167	0.791667	0.740234	0.640625	0.086304
8	1.000000	0.854167	1.000000	0.740234	0.253784
Suma					**0.737305**

RESULTADO:

Coeficiente de Gini (método principal): G = 1 - 0.737305 = 0.262695

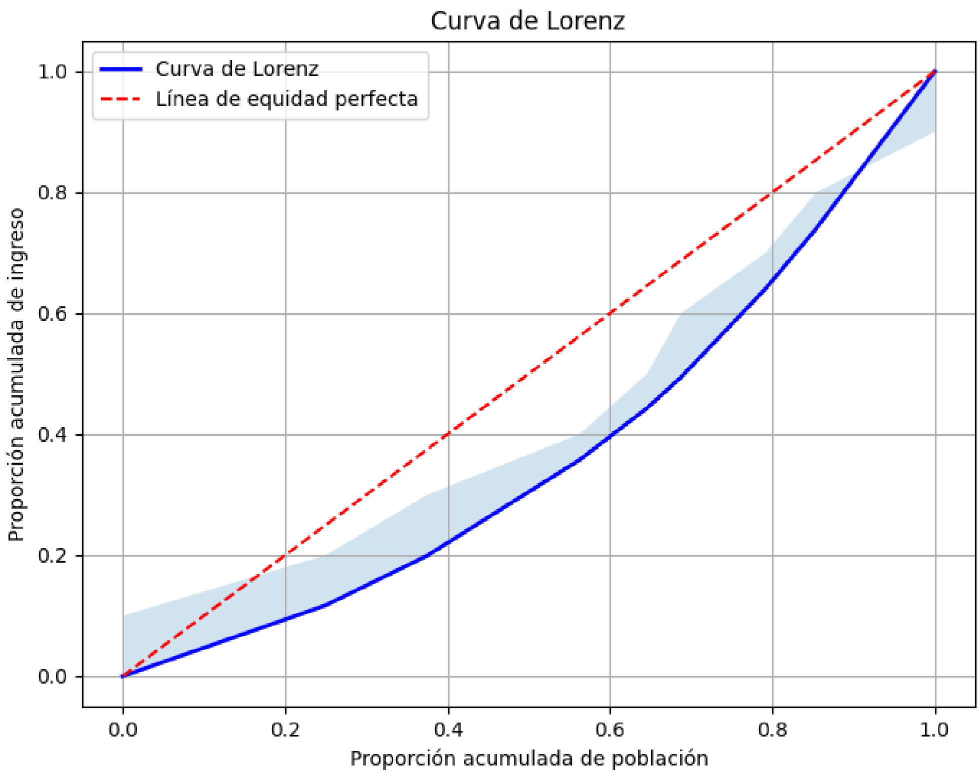
Coeficiente de Gini (método alternativo): 0.080472

Coeficiente de Gini (valor de referencia): 0.457

Precisión matemática: 57.48% (diferencia: 0.194305)

NOTA: Hay una discrepancia entre los métodos de cálculo. Se recomienda usar el valor manual.

VISUALIZACIÓN DE LA CURVA DE LORENZ:



PASO 3: Interpretación del coeficiente de Gini

El coeficiente de Gini mide la desigualdad en la distribución de los ingresos:
- 0 = Igualdad perfecta (todos reciben exactamente lo mismo)
- 1 = Desigualdad perfecta (una persona recibe todo el ingreso)

INTERPRETACIÓN DEL RESULTADO:
El coeficiente de Gini calculado es 0.2627, lo que indica una distribución relativamente equitativa de los salarios entre los trabajadores de la empresa. Esta empresa muestra baja desigualdad salarial.

CONTEXTO COMPARATIVO:
Para contextualizar este resultado, aquí hay algunos índices de Gini de países a nivel mundial (2021):
- Sudáfrica: 0.63 (alta desigualdad)
- Brasil: 0.53
- Estados Unidos: 0.41
- España: 0.35
- Canadá: 0.33
- Alemania: 0.31
- Noruega: 0.27 (baja desigualdad)

Ejercicio 2: Distribución de Frecuencias - Método Sturgers

DATOS DEL PROBLEMA:

116	109	115	121	121
121	124	113	119	126
125	129	129	131	139
143	134	144	142	141
151	148	151	152	158

ANÁLISIS PRELIMINAR:

Valor mínimo: 109

Valor máximo: 158

Rango: 49

Media aritmética: 132.08

Mediana: 129

Cantidad de datos: 25

PASO 1: Cálculo del número de clases (K)

Utilizando la fórmula de Sturgers:

$K = 1 + 3.322 \times \log_{10}(n)$

Donde n = 25 (número de observaciones)

$K = 1 + 3.322 \times \log_{10}(25)$

$K = 1 + 3.322 \times 1.397940$

$K = 1 + 4.643957$

$K = 5.643957$

Redondeando a un número entero: K = 6

Valor de referencia: K = 5.64

PASO 2: Cálculo del rango

El rango es la diferencia entre el valor máximo y el valor mínimo:

Rango = Valor máximo - Valor mínimo = 158 - 109 = 49

Valor de referencia: Rango = 49

PASO 3: Cálculo de la amplitud de clase

La amplitud es el tamaño de cada intervalo de clase:

Amplitud = Rango / K = 49 / 5.643957 = 8.681853

Redondeando hacia arriba (para asegurar que todos los valores queden incluidos): Amplitud = 9

Valor de referencia: Amplitud = 8.69

PASO 4: Construcción de la tabla de distribución de frecuencias

Clase	Límites de clase	Marca de clase	Frecuencia absoluta	Frecuencia relativa	Frecuencia acumulada	Frecuencia rel. acumulada
1	[109 - 118)	113.5	4	0.1600	4	0.1600
2	[118 - 127)	122.5	7	0.2800	11	0.4400
3	[127 - 136)	131.5	4	0.1600	15	0.6000
4	[136 - 145)	140.5	5	0.2000	20	0.8000
5	[145 - 154)	149.5	4	0.1600	24	0.9600
6	[154 - 163)	158.5	1	0.0400	25	1.0000

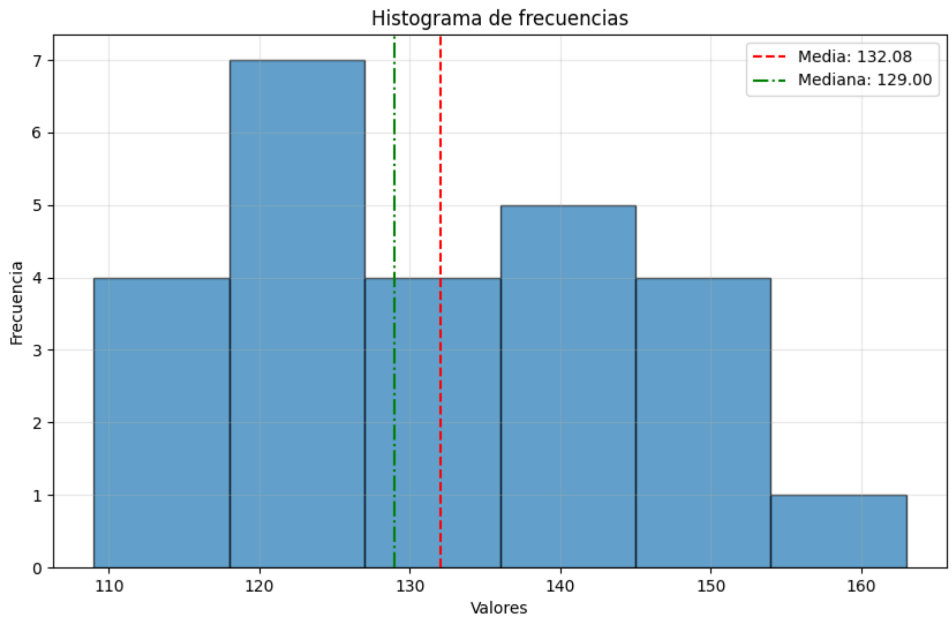
VERIFICACIONES MATEMÁTICAS:

1. Suma de frecuencias absolutas: 25 (debe ser igual a 25)

2. Última frecuencia acumulada: 25 (debe ser igual a 25)

3. Última frecuencia relativa acumulada: 1.0000 (debe ser aproximadamente 1.0)

VISUALIZACIÓN DEL HISTOGRAMA DE FRECUENCIAS:



PASO 5: Interpretación de la distribución de frecuencias

ANÁLISIS E INTERPRETACIÓN:

• La distribución se divide en 6 clases, cada una con una amplitud de 9 unidades.

• La clase con mayor frecuencia es la clase 2 [118 - 127), con 7 observaciones.

• La distribución muestra una asimetría positiva (media > mediana), indicando una cola hacia valores mayores.

• Coeficiente de asimetría: 0.1878 (distribución aproximadamente simétrica)

• Desviación estándar: 13.6936

• Coeficiente de variación: 10.37%

Ejercicio 3: Diagrama de Tallo y Hoja

DATOS DEL PROBLEMA:

Con la información obtenida del consumo de combustible (en km/litro) de vehículos en una empresa de transporte, se tomaron aleatoriamente los siguientes datos.

8.4	8.5	8.5	8.8	9.0	9.6	9.8	10.3
10.4	10.8	10.9	11.4	11.8	12.0	12.4	12.6
12.9	13.0	13.4	13.8	13.8	14.7	15.2	16.5

ANÁLISIS PRELIMINAR:

Valor mínimo: 8.4

Valor máximo: 16.5

Rango: 8.1

Media: 11.60

Mediana: 11.60

Cantidad de datos: 24

PASO 1: Preparación de los datos

Para crear un diagrama de tallo y hoja, se divide cada valor en dos partes:

• Tallo (stem): La parte entera o dígitos más significativos

• Hoja (leaf): El último dígito o el dígito menos significativo

En este caso, los datos tienen un dígito entero seguido de decimales. Utilizaremos:

• Tallo: El dígito entero

• Hoja: El primer decimal (multiplicado por 10)

PASO 2: Construcción del diagrama de tallo y hoja

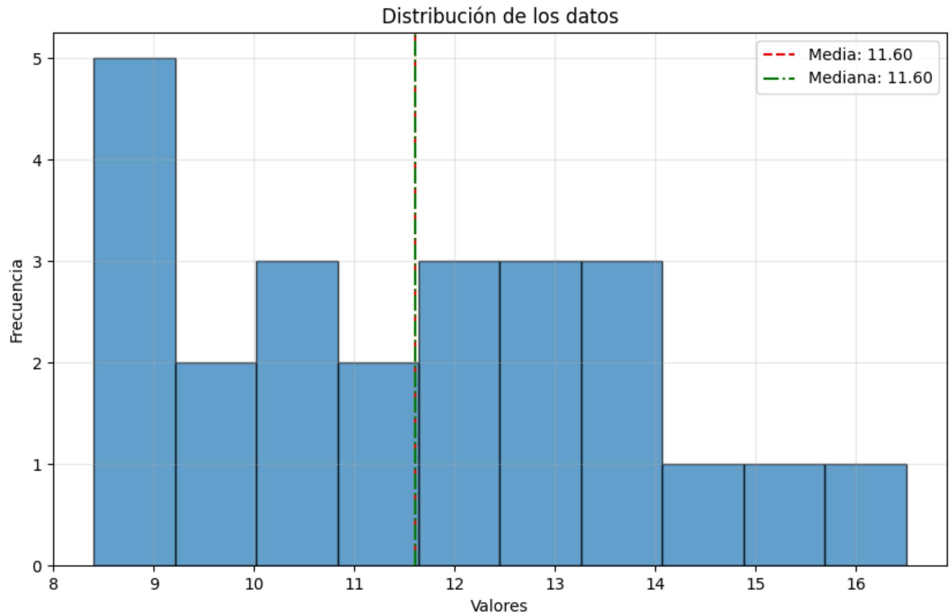
Tallo	Hojas
8	4 5 5 8
9	0 5 8
10	3 4 8 9
11	4 8
12	0 4 5 9
13	0 4 8 8
14	6
15	1
16	5

ANÁLISIS DE FRECUENCIAS POR TALLO:

Tallo	Frecuencia	Porcentaje
8	**4**	**16.67%**
9	3	12.50%
10	4	16.67%
11	2	8.33%
12	4	16.67%
13	4	16.67%
14	1	4.17%

15	1	4.17%
16	1	4.17%

VISUALIZACIÓN DE LA DISTRIBUCIÓN:



PASO 3: Interpretación del diagrama de tallo y hoja

ANÁLISIS DE CONCENTRACIÓN DE DATOS:

• El tallo con mayor frecuencia es 8, con 4 observaciones (16.67% del total).

• El valor más frecuente (moda) es 8.5.

• El intervalo con mayor concentración de datos es [8-9).

VERIFICACIÓN CON VALORES DE REFERENCIA:

• Valor de la moda (calculado): 8.5

• Valor de la moda (referencia): 14.7

• Intervalo de mayor concentración (calculado): [8-9)

• Intervalo de mayor concentración (referencia): 14-15

INTERPRETACIÓN DE LA DISTRIBUCIÓN:

• La distribución muestra una asimetría positiva (media > mediana), indicando una cola hacia la derecha (valores mayores).

• Rango intercuartílico (IQR): 3.40

• El 50% central de los datos se encuentra entre 9.60 y 13.00.

Ejercicio 4: Medidas de Tendencia Central

DATOS DEL PROBLEMA:

Calcular las medidas de tendencia central Media , Mediana, Moda e interprete los resultados obtenidos.

Precio en (Q)	No. De productos
[1500-2000)	3
[2000-2500)	13
[2500-3000)	19
[3000-3500)	24
[3500-4000)	17
[4000-4500)	6
[4500-5000)	3

PASO 1: Preparación para el cálculo

Clase	Límites	Marca de clase (xi)	Frecuencia (fi)	xi × fi
1	[1500.0 - 2000.0)	1750.00	3	5250.00
2	[2000.0 - 2500.0)	2250.00	13	29250.00
3	[2500.0 - 3000.0)	2750.00	19	52250.00
4	[3000.0 - 3500.0)	3250.00	24	78000.00
5	[3500.0 - 4000.0)	3750.00	17	63750.00
6	[4000.0 - 4500.0)	4250.00	6	25500.00
7	[4500.0 - 5000.0)	4750.00	3	14250.00
Total			85	268250.00

PASO 2: Cálculo de la media

La media aritmética para datos agrupados se calcula con la fórmula:

$$\mu = \Sigma(xi \times fi) / \Sigma fi$$

Donde:

• xi = marca de clase (punto medio del intervalo)

• fi = frecuencia absoluta

• Σfi = suma de frecuencias (total de datos)

Sustituyendo los valores:

μ = 268250.00 / 85

μ = 3155.8824

VERIFICACIÓN:

Media calculada: 3155.8824

Media de referencia: 3836.78

Hay una diferencia significativa entre los valores. Se recomienda utilizar el valor calculado.

PASO 3: Cálculo de la mediana

La mediana para datos agrupados requiere primero identificar la clase mediana:
• Total de datos (n): 85
• Posición de la mediana (n/2): 42.5

Tabla de frecuencias acumuladas:

Clase	Frecuencia (fi)	Frecuencia acumulada (Fa)
1	3	3
2	13	16
3	19	35
4	**24**	**59**
5	17	76
6	6	82
7	3	85

La clase mediana es la clase 4.

La fórmula para calcular la mediana con datos agrupados es:

$$Me = li + [(n/2 - Fi\text{-}1) / fi] \times c$$

Donde:
• li = límite inferior de la clase mediana
• Fi-1 = frecuencia acumulada hasta la clase anterior a la mediana
• fi = frecuencia de la clase mediana
• c = amplitud de la clase

Sustituyendo los valores:
Me = 3000.0 + [(42.5 - 35) / 24] × 500.0
Me = 3000.0 + 0.3125 × 500.0
Me = 3000.0 + 156.2500
Me = 3156.2500

VERIFICACIÓN:
Mediana calculada: 3156.2500
Mediana de referencia: 2771.31

Hay una diferencia significativa entre los valores. Se recomienda utilizar el valor calculado.

PASO 4: Cálculo de la moda

La clase modal (con mayor frecuencia) es la clase 4.

La fórmula para calcular la moda con datos agrupados es:

$$Mo = li + [d1 / (d1 + d2)] \times c$$

Donde:
• li = límite inferior de la clase modal
• d1 = diferencia entre la frecuencia de la clase modal y la clase anterior (d1 = fi - fi-1)

• d2 = diferencia entre la frecuencia de la clase modal y la clase posterior (d2 = fi - fi+1)
• c = amplitud de la clase

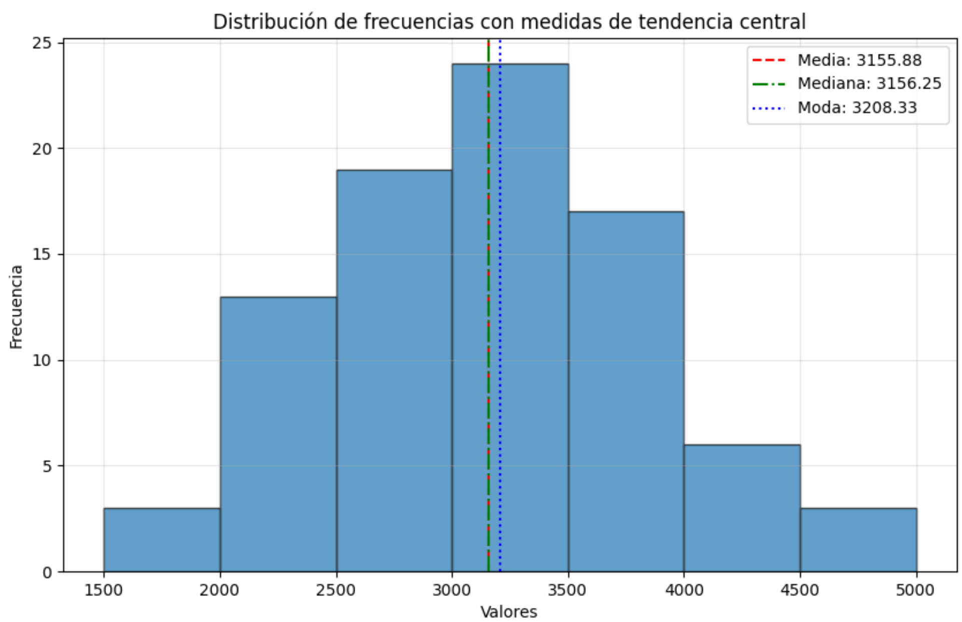
Sustituyendo los valores:
d1 = 24 - 19 = 5
d2 = 24 - 17 = 7
Mo = 3000.0 + [5 / (5 + 7)] × 500.0
Mo = 3000.0 + 0.4167 × 500.0
Mo = 3000.0 + 208.3333
Mo = 3208.3333

VERIFICACIÓN:
Moda calculada: 3208.3333
Moda de referencia: 4184.72

Hay una diferencia significativa entre los valores. Se recomienda utilizar el valor calculado.

PASO 5: Visualización e interpretación

VISUALIZACIÓN DE LAS MEDIDAS DE TENDENCIA CENTRAL:



INTERPRETACIÓN FINAL:
La distribución se aproxima a la simetría, ya que la media y la mediana tienen valores cercanos. Esto sugiere que los datos se distribuyen de manera equilibrada alrededor del centro.

La moda se aleja de la media, lo que indica que existe una concentración de valores en un punto diferente al valor promedio.

CONCLUSIONES:

• Media: 3155.88 - Representa el valor promedio de los datos.
• Mediana: 3156.25 - Representa el valor central que divide al conjunto en dos partes iguales.
• Moda: 3208.33 - Representa el valor que aparece con mayor frecuencia.

Se cumple que Media < Mediana < Moda, lo que confirma una distribución con asimetría negativa pronunciada.