

Estadística 1

Jorge Miguel Alvarado Reyes

16 Agosto 2023

Índice

1. 16 de agosto 2023	3
1.1. Medidas de Tendencia Central	3
1.1.1. Media	3
1.1.2. Mediana	3
1.1.3. Moda	4
1.2. Media para una serie de frecuencias	4
1.3. Media para datos agrupados	4
2. 18 Agosto 2023	6
2.1. Breve introduccion a latex	6
2.1.1. Principales clases de documentos	6
2.1.2. Paquetes	6
2.1.3. Estructura de un documento	6
2.1.4. LaTeX en linea	7
2.1.5. Partes de un documento	7
2.1.6. Tamaños de fuente	7
2.1.7. Listas numeradas y viñetas	7
2.1.8. Alineacion de texto	7
2.1.9. Composicion de ecuaciones	7
2.1.10. Alinear expresion con algun elemento	8
2.1.11. Tablas	8
2.1.12. Como insertar una imagen	8
2.2. Clase	9
3. 21 Agosto 2023	9
3.1. Medidas de posición (1.6.4)	9
3.1.1. Cuantiles	9
3.1.2. Desviacion estandar	12
3.1.3. Desviacion media	12
3.1.4. Rango	12
3.2. Medidas de forma (1.6.5)	13
3.2.1. Asimetria	13

1. 16 de agosto 2023

Las medidas de tendencia central son valores de un conjunto de datos que se encuentran en el centro de los datos ordenados.

1.1. Medidas de Tendencia Central

1.1.1. Media

Existen dos tipos de media: la aritmética y la ponderada.

La **media aritmética** se calcula sumando todos los valores y dividiendo por la cantidad de valores:

$$\text{Media}(x) = \frac{\sum_{i=1}^n x_i}{n}$$

Propiedades:

1. $\text{Media}(cx) = c \cdot \text{Media}(x)$
2. $\text{Media}(x + c) = \text{Media}(x) + c$

Ejemplo 4:

Mostrar que $\text{Media}(x + c) = \text{Media}(x) + c$.

Demostración:

$$\begin{aligned}\text{Media}(x + c) &= \frac{x_1 + c + \dots + x_n + c}{n} \\ &= \frac{x_1 + \dots + x_n + n \cdot c}{n} \\ &= \frac{x_1 + \dots + x_n}{n} + c \\ &= \text{Media}(x) + c\end{aligned}$$

Ejemplo 5:

Mostrar que $\text{Media}(cx) = c \cdot \text{Media}(x)$.

La **media ponderada** se define como:

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

1.1.2. Mediana

La mediana es el valor central que divide a un conjunto de datos ordenados en dos partes iguales. Si n es par, se calcula como:

$$\text{Mediana}(x) = \frac{x(\frac{n}{2}) + x(\frac{n}{2} + 1)}{2}$$

1.1.3. Moda

Es el valor que mas se repite en un conjunto de observaciones.

Ejemplo 6:

1. $[1, 2, 3, 4, 5]$ Aqui no existe moda
2. $[3, 4, 4, 5, 5, 6]$ Moda = 4.5
3. $[3, 3, 4, 5, 6, 6]$ Moda = 3 y 6
4. $[2, 7, 7, 7, 9]$ Moda = 7

1.2. Media para una serie de frecuencias

Si f_1, \dots, f_n son las frecuencias de la variable x . Entonces.

$$Mediana(x) = \frac{\sum_{i=1}^n f_i \cdot x_i}{\sum_{i=1}^n f_i}$$

Ejemplo 7:

Calcula la media para los siguientes valores

x	f_i
2	4
5	1
6	3
8	4

$$Mediana(x) = \frac{(2 \cdot 4) + (5 \cdot 1) + (6 \cdot 3) + (8 \cdot 4)}{4 + 1 + 3 + 4} = \frac{8 + 5 + 18 + 32}{12} = \frac{63}{12}$$

1.3. Media para datos agrupados

Sean f_1, \dots, f_n las frecuencias de la variable x y c_1, \dots, c_n las marcas de clase, entonces: (Marca de clase es un representante)

$$Media(x) = \frac{\sum_{i=1}^n f_i \cdot c_i}{\sum_{i=1}^n f_i}$$

Ejemplo 8:

Calcula la edad promedio para el siguiente conjunto de datos

Adulto	25
Adulto de la tercera edad	10

Adulto, edad = $[20, 65]$, $c_1 = 43$
Adulto tercera edad, edad = $[65, 100]$, $c_1 = 83$

25 veces 43 y 10 veces 83

$$Mediana(x) = L_i t \left(\frac{\frac{n}{2} - \sum f_i}{f_{mediana}} \right) \cdot c$$

L_i = limite inferior de la clase que contiene la mediana

n = frecuencia total

$\sum f_i$ = suma de las frecuencias menores a la mediana

$f_{mediana}$ = Frecuencia de la clase que contiene la mediana

c = longitud del intervalo que contiene la mediana

2. 18 Agosto 2023

Sitio del curso: <https://plaza.com/unam.mx/other/ei20241>
codigo de acceso: 150621

2.1. Breve introduccion a latex

LaTeX es una herramienta para crear documentos de una gran calidad tipográfica, en donde los usuarios se ocupan en mayor medida del contenido del texto en lugar del formato.

2.1.1. Principales clases de documentos

Clase	Proposito
article	Articulos de revista
report	Textos largos como tesis o reportes
book	Libros o documentos con una estructura similar
lette	cartas

2.1.2. Paquetes

Nombre	Funcion
amsmath, amssymb, amsfont	Permiten el uso de símbolos matemáticos.
babel	Escribir en diversos idiomas.
inputec	Codificacion de entradas.

2.1.3. Estructura de un documento

```
\documentclass[11pt, a4paper]{report}
\usepackage[utf8]{inputec}
\usepackage[spanish]{babel}
\usepackage{amsmath}
\usepackage{amssymb}
\usepackage{amsfont}

\title{Titulo}
\author{Nombre}
\date{\today}
\begin{document}
\maketitle
...
\end{document}
```

2.1.4. LaTeX en línea

Crear cuenta en <https://es.overleaf.com>

New project → Blank project → Escribir nombre del documento → Create
Menu → spell check spanish

2.1.5. Partes de un documento

```
\section*{title}
\subsection*{title}
\subsubsection*{title}

\part*{title}
\chapter*{title}
```

2.1.6. Tamaños de fuente

```
\huge
\Huge
\LARGE
\Large
\large
\normalsize
\small
\tiny
```

2.1.7. Listas numeradas y viñetas

```
\begin{itemize}[a]
  \item
  \item
\end{itemize}

\begin{enumerate}
  \item
  \item
\end{enumerate}
```

2.1.8. Alineacion de texto

```
\begin{center}
...
\end{center}
```

2.1.9. Composicion de ecuaciones

```

$$x^2+2x+3=0$$

```

2.1.10. Alinear expresion con algun elemento

```
\begin{align*}
c^2 &= a^2 + b^2 \\
&= 2^2 + 3^2 \\
&= 13
\end{align*}
```

2.1.11. Tablas

```
\begin{table}[h]
\centering
\begin{tabular}{c | c c}
a & b & c \\
a & b & c \\
a & b & c
\end{tabular}
\end{table}
```

```
\begin{table}[h]
\centering
\begin{tabular}{| c c c |}
\hline
a & b & c \\
\hline
a & b & c \\
\hline
a & b & c \\
\hline
\end{tabular}
\end{table}
```

2.1.12. Como insertar una imagen

```
\usepackage{graphicx}
\includegraphics[width = , height = ]{archivo.jpg,png,etc.}
```


2.2. Clase

Ejemplo 9:

Encuentran la mediana para las siguientes observaciones

Intervalo	Frecuencia	Frecuencia acumulada
(118.5,126.5]	3	3
(126.5,135.5]	5	8
(135.5,144.5]	9	17
(144.5,153.5]	12	29
(153.5,162.5]	5	34
(162.5,171.5]	4	38
(171.5, 180.5]	2	40

L_i = limite inferior de la clase que contiene la mediana

n = frecuencia total

$\sum f_i$ = suma de las frecuencias menores a la mediana

$f_{mediana}$ = Frecuencia de la clase que contiene la mediana

c = longitud del intervalo que contiene la mediana

$$L_i = 144.5$$

$$n = 40$$

$$\sum f_i = 17$$

$$f_{mediana} = 12$$

$$c = 153.5 - 144.5 = 9$$

$$\begin{aligned} Mediana(x) &= \frac{x_{20} + x_{21}}{2} = 146,75 \\ &= L_i + \left(\frac{\frac{n}{2} - \sum f_i}{f_{mediana}} \right) \cdot c \end{aligned}$$

3. 21 Agosto 2023

3.1. Medidas de posición (1.6.4)

3.1.1. Cuantiles

Sean x_1, \dots, x_n observaciones de una variable aleatoria x y $p \in (0, 1)$. Un cuantil al 100 p % es el numero c que cumple con las siguientes condiciones.

$$\begin{aligned} \blacksquare \quad \frac{\#\{x_i | x_i \leq c\}}{n} &\geq p \\ \blacksquare \quad \frac{\#\{x_i | x_i \leq c\}}{n} &\geq 1 - p \end{aligned}$$

Casos particulares

- Deciles: si $p = \{0,1, \dots, 0,9\}$
- Cuartiles: si $p = \{0,25, 0,50, 0,75\}$
- Percentiles: si $p = \{0,01, 0,02, \dots, 0,99\}$

Si $c=0$

$$\begin{aligned} & \blacksquare \frac{\#\{x_i | x_i \leq 0\}}{n} \geq p \\ & \frac{1}{2} \geq p, p = \{0,1, 0,2, 0,3, 0,4, 0,5\} \\ & \blacksquare \frac{\#\{x_i | x_i \geq 0\}}{n} \geq 1 - p \\ & \frac{2}{2} \geq 1 - p, p = \{0,1, 0,2, 0,3, 0,4, 0,5\} \end{aligned}$$

Si $c=1$

$$\begin{aligned} & \blacksquare \frac{\#\{x_i | x_i \leq 1\}}{n} \geq p \\ & \frac{2}{2} \geq p, p = \{0,5, 0,6, 0,7, 0,8, 0,9\} \\ & \blacksquare \frac{\#\{x_i | x_i \geq 1\}}{n} \geq 1 - p \\ & \frac{1}{2} \geq 1 - p, p = \{0,5, 0,6, 0,7, 0,8, 0,9\} \end{aligned}$$

Ejemplo 10

Calcula los deciles de la variable $x = \{0,1\}$

- $q(0.1) = 0$
- $q(0.2) = 0$
- $q(0.3) = 0$
- $q(0.4) = 0$
- $q(0.5) = 0.5$
- $q(0.6) = 1$
- $q(0.7) = 1$
- $q(0.8) = 1$
- $q(0.9) = 1$

Ejemplo 11

Calcula los cuantiles cd la variable $x = \{0, 1, 2, 3\}$, $n = 4$

Si $c=0$

$$\blacksquare \frac{\#\{x_i | x_i \leq 0\}}{n} \geq p$$

$$\frac{1}{2} \geq p, p = \{0, 1, 0, 2, 0, 3, 0, 4, 0, 5\}$$

$$\blacksquare \frac{\#\{x_i | x_i \geq 0\}}{n} \geq 1 - p$$

$$\frac{2}{2} \geq 1 - p, p = \{0, 1, 0, 2, 0, 3, 0, 4, 0, 5\}$$

Si $c=1$

$$\blacksquare \frac{\#\{x_i | x_i \leq 1\}}{n} \geq p$$

$$\frac{2}{2} \geq p, p = \{0, 5, 0, 6, 0, 7, 0, 8, 0, 9\}$$

$$\blacksquare \frac{\#\{x_i | x_i \geq 1\}}{n} \geq 1 - p$$

$$\frac{1}{2} \geq 1 - p, p = \{0, 5, 0, 6, 0, 7, 0, 8, 0, 9\}$$

$$\blacksquare q(0.25) = 0.5$$

$$\blacksquare q(0.50) = 1.5$$

$$\blacksquare q(0.75) = ??? - \text{INVESTIGAR}$$

Sean x_1, \dots, x_n observaciones de la variable aleatoria x . La varianza de x es

$$var(x) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

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

Propiedades de la varianza

$$var(x + c) = var(x)$$

$$var(ax) = a^2 var(x)$$

Ejemplo 12

Muestra que $var(ax) = a^2 var(x)$

Demostracion:

$$\begin{aligned}
 var(ax) &= \frac{1}{n} \sum_{i=1}^n (ax_i - a\bar{x})^2 \\
 &= \frac{1}{n} \sum_{i=1}^n (a(x_i - \bar{x}))^2 \\
 &= a^2 \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \\
 &= a^2 var(x)
 \end{aligned}$$

3.1.2. Desviacion estandar

Sean x_1, \dots, x_n observaciones de la variable x. la desviacion estandar de x es

$$De(x) = \sqrt{var(x)}$$

Propiedades

$$De(x + c) = De(x)$$

$$De(ax) = |a|De(x)$$

3.1.3. Desviacion media

Sean x_1, \dots, x_n observaciones de la variable x. la desviacion media de x es

$$Dm(x) = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|$$

Propiedades

$$Dm(x + c) = Dm(x)$$

$$Dm(ax) = |a|Dm(x)$$

3.1.4. Rango

Sean x_1, \dots, x_n observaciones de la variable x. El rango de x es

$$R(x) = x_n - x_1$$

Propiedades

$$R(x + c) = R(x)$$

$$R(ax) = |a|R(x)$$

Ejemplo 13

Muestra que $var(x + c) = var(x)$

3.2. Medidas de forma (1.6.5)

Sean x_1, \dots, x_n observaciones de la variable x . El k -esimo momento de x es

$$m'_k(x) = \frac{1}{n} \sum_{i=1}^n x_i^k$$

Mientras que el k -esimo momento central de x es

$$m_k(x) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k$$

3.2.1. Asimetria

El coeficiente de asimetria mide la asimetria de los datos respecto a la media.

Sean x_1, \dots, x_n observaciones de la variable x . El coeficiente de asimetria de x es

$$sk(x) = \frac{1}{s^3} \left[\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3 \right]$$

Donde s es la desviacion estandar de x .