

Ricardo Michel MALLQUI BAÑOS

Estadística



Índice general

| | |
|--|------------|
| Índice de cuadros | v |
| Índice de figuras | vii |
| Resumen | ix |
| Introducción | xi |
| 1. Variables | 1 |
| 1.1. Variables cualitativas | 1 |
| 1.1.1. Nominale | 1 |
| 1.1.2. Ordinale | 1 |
| 1.2. Variables cuantitativas | 1 |
| 1.2.1. Discretas | 1 |
| 1.2.2. Continuas | 2 |
| 2. Tabulación | 3 |
| 3. Medidas de tendencia central | 5 |
| 3.1. La media (\bar{x}) | 5 |
| 3.1.1. Media de datos no agrupados | 5 |
| 3.1.2. Media de datos agrupados | 5 |
| 3.2. La moda (Mo) | 6 |
| 3.2.1. Moda de datos no tabulados | 6 |
| 3.2.2. Moda de datos tabulados | 6 |
| 3.3. la mediana (Me) | 7 |
| 3.3.1. Mediana de datos no tabulados | 7 |
| 3.3.2. Mediana de datos tabulados | 7 |
| 4. traductor | 9 |
| 5. traducto2 | 11 |
| 6. estadística | 13 |
| 7. eee | 15 |
| Apéndice | 15 |

| | |
|---|-----------|
| A. Sumatorias | 17 |
| A.1. ee | 17 |
| A.2. eeeee | 17 |
| B. eeeeeeee | 19 |
| B.1. ww | 19 |
| B.2. www | 19 |
| B.3. Forma matricial de una ecuación lineal | 19 |
| Bibliografía | 21 |
| Índice alfabético | 23 |

Índice de cuadros



Índice de figuras



Resumen

Este libro sobre la estadística descriptiva, cuyo objetivo es demostrar resultados básicos muy útiles en el desarrollo de investigaciones.

$$\sum_1^2$$



Introducción

$$\sum_1^2$$

$$\vec{u} = (1, 1) - \rho \int_2^3$$

Debido a la poca información estructurada de estadística descriptiva se propone escribir este libro con un enfoque demostrativo.

$$x^2 + y^2$$



1

Variables

1.1. Variables cualitativas

Denotan cualidades de objetos personas o animales tales como características inherentes que no son medibles por números, tenemos dos casos de esta variable.

1.1.1. Nominales

Son características que simplemente nominan y están propensas a ser jerarquizadas u ordenadas tales como: El estado civil (soltero, casado, divorciado, viudo), Religión (católico, evangélico, judío, etc).

1.1.2. Ordinales

Son características que si están propensas a ser jerarquizadas tales como: Nivel de instrucción (primaria, secundaria, superior).

1.2. Variables cuantitativas

Son aquellas variables que están propensas a ser medidas mediante números ya sean números enteros o reales.

1.2.1. Discretas

Aquellas que solo son medidas mediante números enteros por ejemplo: Número de hijos, número de habitaciones.

1.2.2. Continuas

Aquellas que solo son medidos mediante numeros reales es decir este incluye a los numeros racionales e irracionales. Estatura, volumen, peso.

2

Tabulación

La tabulación es un proceso en el cual los datos son ordenados en grupos llamados *clases* para un análisis más eficaz de estos, los datos podrían estar clasificados mediante una variable cualitativa o cuantitativa en el caso de las variables cualitativas se considera la siguiente tabla

| Y_i | f_i | F_i | F_i^* | h_i | H_i | H_i^* | $h_i \%$ | $H_i \%$ | $H_i^* \%$ |
|----------|----------|----------|----------|-----------------|-----------------|-------------------|----------|----------|------------|
| Y_1 | f_1 | F_1 | F_1^* | $\frac{f_1}{n}$ | $\frac{F_1}{n}$ | $\frac{F_1^*}{n}$ | h_1 | H_1 | H_1^* |
| Y_2 | f_2 | F_2 | F_2^* | $\frac{f_2}{n}$ | $\frac{F_2}{n}$ | $\frac{F_2^*}{n}$ | h_2 | H_2 | H_1^* |
| Y_3 | f_3 | F_3 | F_3^* | $\frac{f_3}{n}$ | $\frac{F_3}{n}$ | $\frac{F_3^*}{n}$ | h_3 | H_3 | H_1^* |
| \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots |
| Y_r | f_r | F_r | F_r^* | $\frac{f_r}{n}$ | $\frac{F_r}{n}$ | $\frac{F_r^*}{n}$ | h_r | H_r | H_1^* |

En el caso de variables cuantitativas además si los datos son muy variados, que para se clasificados adecuadamente, necesitan generarse particiones de longitudes semejantes entonces se utiliza el siguiente proceso; el **número de las particiones** r se consideran de acuerdo a **tres criterios**

- Criterio del investigador r no puede ser más de 20 ni menos de 5
- $r = \sqrt{n}$ donde n es el número de datos
- La regla de Starges que consiste en considerar la fórmula $r = 3,322 \cdot \log_{10} n$ Una vez establecido el número de particiones se procede a generar los límites laterales de cada una de las particiones, sea L la longitud de todo el conjunto es decir $L = x_{\max} - x_{\min}$ entonces la longitud de las particiones o amplitud intervállica se obtiene con $l = \frac{L}{r}$

| Clase | Clase | f_i | F_i | F_i^* | h_i | H_i | H_i^* | ... |
|--------------------|----------|----------|----------|----------|-----------------|-----------------|-------------------|----------|
| $[y_1 - y_2 >$ | y_1 | f_1 | F_1 | F_1^* | $\frac{f_1}{n}$ | $\frac{F_1}{n}$ | $\frac{F_1^*}{n}$ | ... |
| $< y_1 - y_2 >$ | y_2 | f_2 | F_2 | F_2^* | $\frac{f_2}{n}$ | $\frac{F_2}{n}$ | $\frac{F_2^*}{n}$ | ... |
| $< y_r - y_r >$ | y_3 | f_3 | F_3 | F_3^* | $\frac{f_3}{n}$ | $\frac{F_3}{n}$ | $\frac{F_3^*}{n}$ | ... |
| \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots |
| $< y_{r-1} - y_r]$ | y_r | f_r | F_r | F_r^* | $\frac{f_r}{n}$ | $\frac{F_r}{n}$ | $\frac{F_r^*}{n}$ | ... |

Tenga en cuenta que n es el número de datos, es decir $n = f_1 + f_2 + \dots + f_r = \sum_{i=1}^r f_i$ donde f_i es número de datos en la partición X_i , una de las r particiones del conjunto total de datos.

- Las **frecuencias absolutas** f_i indican el número de datos con la característica X_i .
- Las **frecuencias absolutas acumuladas menor que** F_i obedecen a la fórmula

$$F_m = f_1 + f_2 + \dots + f_m = \sum_{i=1}^m f_i$$

- Las **frecuencias absolutas acumuladas mayor que** F_i^* obedecen a la fórmula

$$F_m^* = f_m + f_{m+1} + \dots + f_r = \sum_{i=m}^r f_i = n - \sum_{i=1}^{m-1} f_i = n - (f_1 + f_2 + \dots + f_{m-1})$$

- Las **frecuencias absolutas relativas** obedecen a la fórmula

$$h_m = \frac{f_m}{n}$$

- Las **frecuencias absolutas relativas menor que** obedecen a la fórmula

$$H_m = \frac{f_m}{n}$$

- Las **frecuencias absolutas relativas mayor que** obedecen a la fórmula

$$H_m^* = \frac{F_m}{n}$$

- Las **frecuencias absolutas relativas porcentuales** obedecen a la fórmula $h_m \% = 100 \cdot h_i$
- Las **frecuencias absolutas relativas menor que porcentuales** obedecen a la fórmula $H_m \% = 100 \cdot H_i$
- Las **frecuencias absolutas relativas mayor que porcentuales** obedecen a la fórmula $H_m^* \% = 100 \cdot H_i^*$

Ejercicio 2.1. Sean los datos

Solución. Entonces

3

Medidas de tendencia central

Son aquellas medidas que buscan un dato representativo central de un conjunto de datos tales como la media, la moda y la mediana.

3.1. La media (\bar{x})

A veces llamada *promedio aritmético*, es la medida de tendencia central que pondera los datos.

3.1.1. Media de datos no agrupados

Los datos no están agrupados cuando no están ordenados sobre una tabla de distribución de frecuencias. Sean los n datos x_1, x_2, \dots, x_n entonces la media o promedio aritmético se define como

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i \quad (3.1)$$

$$\frac{d[P; F_1]}{d[P; \mathcal{L}_1]} = e = \frac{d[P; F_2]}{d[P; \mathcal{L}_2]} \quad (3.2)$$

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

3.1.2. Media de datos agrupados

Considérese la siguiente tabla de distribución de frecuencias entonces el promedio es

$$\bar{x} = \frac{y_1 f_1 + y_2 f_2 + \dots + y_n f_n}{n} = \frac{1}{n} \sum_{i=1}^n y_i f_i$$

| Clase | Clase | f_i | F_i | F_i^* | h_i | H_i | H_i^* | ... |
|--------------------|----------|----------|----------|----------|-----------------|-----------------|-------------------|----------|
| $< y_1 - y_2 >$ | y_1 | f_1 | F_1 | F_1^* | $\frac{f_1}{n}$ | $\frac{F_1}{n}$ | $\frac{F_1^*}{n}$ | ... |
| $< y_2 - y_3 >$ | y_2 | f_2 | F_2 | F_2^* | $\frac{f_2}{n}$ | $\frac{F_2}{n}$ | $\frac{F_2^*}{n}$ | ... |
| $< y_3 - y_4 >$ | y_3 | f_3 | F_3 | F_3^* | $\frac{f_3}{n}$ | $\frac{F_3}{n}$ | $\frac{F_3^*}{n}$ | ... |
| \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots |
| $< y_{r-1} - y_r]$ | y_r | f_r | F_r | F_r^* | $\frac{f_r}{n}$ | $\frac{F_r}{n}$ | $\frac{F_r^*}{n}$ | ... |

Ejercicio 3.1. Si el promedio de n datos es \bar{x} entonces el promedio del conjunto inicial más un dato adicional x_{n+1} es

$$\bar{x}' = \frac{n\bar{x} + x_{n+1}}{n + 1}$$

en general si se adicionan r datos y_1, y_2, \dots, y_r entonces el nuevo promedio será

$$\bar{x}' = \frac{n\bar{x} + y_1 + y_2 + \dots + y_r}{n + r}$$

Solución. En efecto sea el promedio

$$\begin{aligned} \bar{x}' &= \frac{x_1 + x_2 + \dots + x_{n+1}}{n + 1} \\ &= \frac{n \frac{x_1 + x_2 + \dots + x_n}{n} + x_{n+1}}{n + 1} \\ &= \frac{n\bar{x} + x_{n+1}}{n + 1} \end{aligned}$$

3.2. La moda (Mo)

3.2.1. Moda de datos no tabulados

En este caso es dato que más repite en un conjunto de datos dados.

La moda es el dato que más se repite por ejemplo sea el conjunto de datos x_1, x_2, x_2, x_2, x_3 entonces la moda $Mo = x_2$

3.2.2. Moda de datos tabulados

La moda es el dato que más se repite por ejemplo sea el conjunto de datos x_1, x_2, x_2, x_2, x_3 entonces la moda $Mo = Li + \frac{Li - Ls}{Li + Ls}r$

| Clase | Clase | f_i | F_i | F_i^* | h_i | H_i | H_i^* | ... |
|--------------------|-------|-------|-------|---------|-----------------|-----------------|-------------------|-----|
| $< y_{r-1} - y_r]$ | y_r | f_r | F_r | F_r^* | $\frac{f_r}{n}$ | $\frac{F_r}{n}$ | $\frac{F_r^*}{n}$ | ... |

Los pasos son:

- Se halla $\frac{n}{2}$ luego
- x_n
-

4

traductor



5

traducto2



6

estadística

Teorema 6.1. *En la elipse se verifican las siguientes igualdades*

1. $d [B_1; F_i] = d [B_2; F_i] = a$
2. $d [V_1; C] = d [V_2; C] = a$
3. $d [C; \mathcal{L}_1] = d [C; \mathcal{L}_2] = \frac{c}{e}$
4. $c = d [P; F_1] = d [P; F_2]$ entonces $c = ae$

Demostración. 1. Ya que $d [B_1; F_1] + d [B_1; F_2] = 2a = d [B_2; F_1] + d [B_2; F_2]$ es decir $2d [B_1; F_i] = 2a = 2d [B_2; F_i]$ entonces $d [B_1; F_i] = a = d [B_2; F_i]$ $i = 1, 2$.

2. Por la definición (??) de la elipse se tiene

$$d [V_1; F_2] + d [V_1; F_1] = 2a \quad (6.1)$$

además la diferencia

$$d [V_1; F_2] - d [V_1; F_1] = 2c \quad (6.2)$$

restando las ecuaciones (6.1) y (6.2) se tiene

$$d [V_1; F_1] = a - c \quad (6.3)$$

entonces haciendo uso de (6.3) en $d [V_1; C] = d [V_1; F_1] + d [F_1; C] = (a - c) + c = a$; de manera similar para el vértice V_2 .

3. En efecto

$$\frac{d [B; F_i]}{d [B; \mathcal{L}_i]} = e \iff \frac{a}{d [B; \mathcal{L}_i]} = e$$

además $d [B_i; \mathcal{L}_i] = d [C; \mathcal{L}_i]$ por lo tanto $\frac{a}{d [C; \mathcal{L}_i]} = e$.

4. Pues

$$\frac{d [P; F_1]}{d [P; \mathcal{L}_1]} = e$$

implica $\frac{a-c}{\frac{a}{e}-a} = e$ es decir $c = ae$.

Por lo tanto

□



7

eee

Example text outside R code here; we know the value of pi is In this section, we give a very brief introduction to Pandoc's Markdown. Readers who are familiar with Markdown can skip this section. The comprehensive syntax of Pandoc's Markdown can be found on the Pandoc website <http://pandoc.org>.

"I thoroughly disapprove of duels. If a man should challenge me, I would take him kindly and forgivingly by the hand and lead him to a quiet place and kill him."

In this section, we give a very brief introduction to Pandoc's Markdown. Readers who are familiar with Markdown can skip this section. The comprehensive syntax of Pandoc's Markdown can be found on the Pandoc website <http://pandoc.org>. \sum_1^2

I thoroughly disapprove of duels. If a man should challenge me, I would take him kindly and forgivingly by the hand and lead him to a quiet place and kill him.

– Mark Twain

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} - \frac{2}{3} \begin{pmatrix} \alpha_1 & \beta_2 \\ \gamma & \delta \end{pmatrix}$$

* La suma de dos matrices $A_{n \times m}$ y $B_{r \times s}$

$$A_{n \times m} \pm B_{n \times m} = [a_{ij} + b_{ij}]$$

* El producto de dos matrices $A_{n \times m}$ y $B_{r \times s}$

$$A_{n \times m} \cdot B_{n \times m} = [a_{ij} + b_{ij}]$$

$$X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{bmatrix}$$

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

$$\begin{matrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{matrix}$$



A

Sumatorias

Una suma de números representados por x_1, x_2, \dots, x_n se simboliza en forma compacta mediante el simbolo \sum (sigma) es decir la suma de los números anteriores se puede escribir del siguiente modo

$$x_1 + x_2 + \dots + x_n = \sum_{i=1}^n x_i.$$

Algunas propiedades son

1. $k \sum_{i=1}^n x_i = \sum_{i=1}^n k x_i$
2. $\sum_{i=1}^n (x_i + y_i) = \sum_{i=1}^n x_i + \sum_{i=1}^n y_i$
3. $\sum_{i=1}^n x_i$

$$\int_1^3 = \lim_{n \rightarrow \infty} \sum_{i=0}^n f^i(x)$$

citado por (Xie, 2015)

A.1. ee

A.2. eeeee



B

eeeeeeee

B.1. ww

B.2. www

B.3. Forma matricial de una ecuación lineal

Sean las matrices

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{11} & a_{11} & \dots & a_{nn} \end{pmatrix}_{n \times n}$$

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{11} & a_{11} & \dots & a_{nn} \end{pmatrix}_{n \times n}$$



Bibliografía

Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.



Índice alfabético

frecuencias absolutas, 4
frecuencias absolutas acumuladas menor
que, 4
frecuencias absolutas relativas, 4
frecuencias absolutas relativas menor
que, 4