

Ricardo Michel MALLQUI BAÑOS

Elementos de la estadística



Índice general

Índice de tablas	vii
Índice de figuras	ix
Resumen	xi
I Estadística descriptiva	1
1. Prerrequisitos	3
2. Variables	5
2.0.1. Nominales	5
2.0.2. Ordinales	5
2.1. Variables cuantitativas	5
2.1.1. Discretas	5
2.1.2. Continuas	6
3. Característica de los datos	7
4. Población, muestra y estadística	9
5. Organización de datos en tablas de frecuencias	11
6. Matriz y rango	15
7. Distribución de frecuencias	17
8. Gráficos estadísticos	19
9. Medidas de tendencia central	21
9.1. La media (\bar{x})	21
9.1.1. Media de datos no agrupados	21
9.1.2. Media de datos agrupados	21
9.2. La moda (Mo)	22
9.2.1. Moda de datos no tabulados	22
9.2.2. Moda de datos tabulados	22
9.3. la mediana (Me)	23
9.3.1. Mediana de datos no tabulados	23

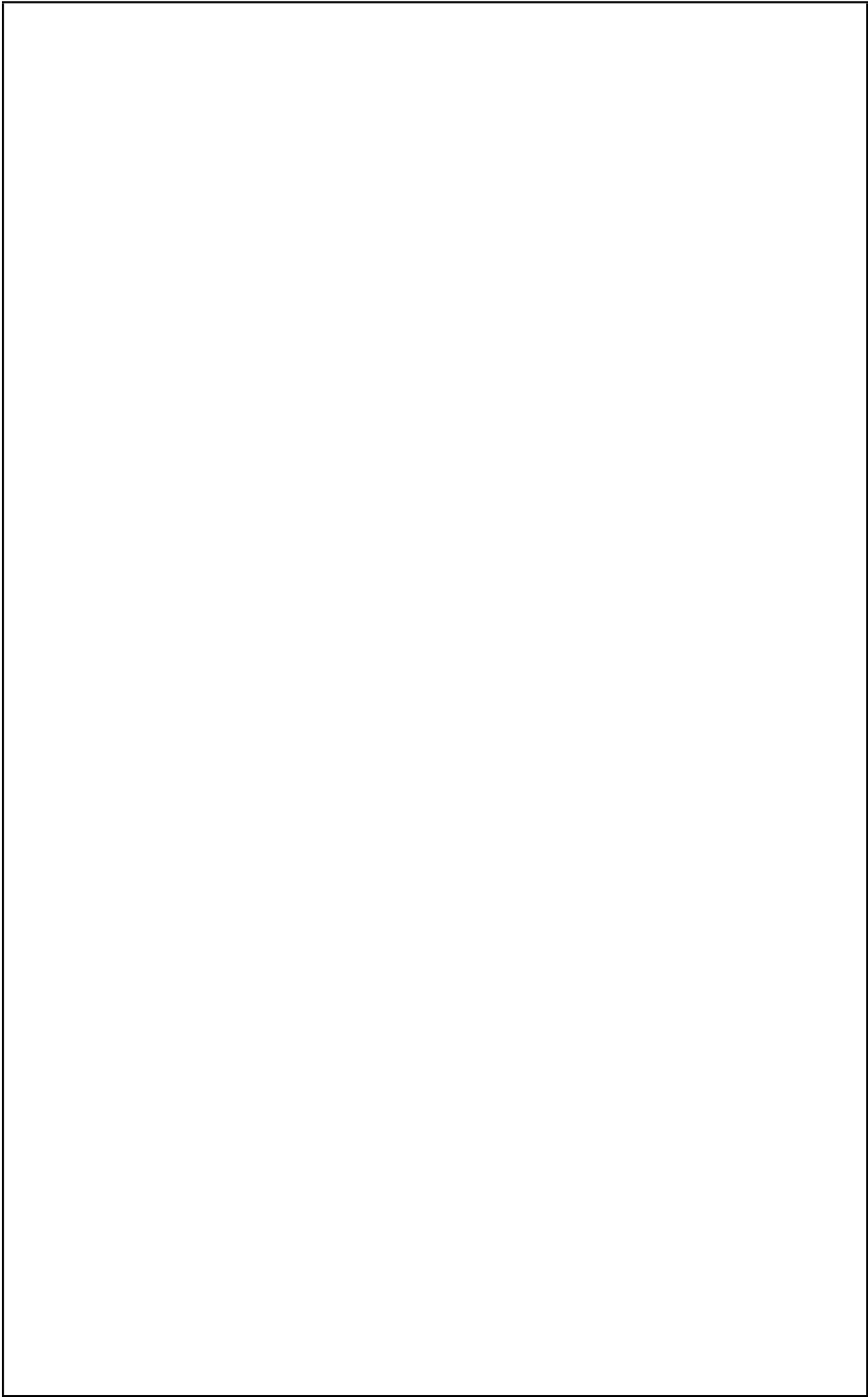
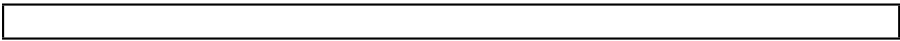
9.3.2. Mediana de datos tabulados	23
10. Medidas de dispersión	25
11. Medidas de asimetría	27
II Probabilidades	29
1. Experimento aleatorio	31
2. Álgebra de eventos	33
3. Técnicas de conteo	35
4. Definición de probabilidad	37
5. Probabilidad condicional	39
6. Teorema de Bayes	41
7. Eventos independientes y secuencias de experimentos	43
8. Probabilidad en espacio	45
III Inferencia estadística	47
1. Variables aleatorias	49
1.1. Clases de variables aleatorias	50
1.1.1. Variable aleatoria discreta	50
1.1.2. Variable aleatoria continua	51
1.1.3. Variable aleatoria mixta	51
1.2. Función de probabilidad de una variable aleatoria	51
1.2.1. Función de probabilidad de una variable aleatorias discreta .	51
1.2.2. Función de probabilidad de una variable aleatoria continua .	51
1.3. Función de distribución de una variable aleatoria	52
1.3.1. Función de distribución de una variable aleatoria discreta . .	52
1.3.2. Función de distribución de una variable aleatoria continua .	52
2. Estadígrafos de la variables aleatorias	53
2.1. esperanza matematica	53
2.2. Medidas de tendencia central	53
2.3. Medidas de variación	53
2.4. Cuantiles	53
3. Variables aleatorias bidimensionales	55
3.1. Distribuciones marginales	55
3.2. Distribuciones	55

<i>Contents</i>	v
4. Distribuciones discreta importantes	57
4.1. Variable aleatoria discreta binomial	57
4.2. Variable aleatoria discreta Poisson	57
5. Distribuciones continuas importantes	59
5.1. Variable aleatoria continua normal	59
5.2. Variable aleatoria continua gamma	59
6. Distribuciones muestrales	61
7. Estimación	63
8. Prueba de hipótesis	65
Apéndice	65
A. Sumatorias	67
A.1. eeeee	67
B. Matrices	69
B.1. Algebra de matrices	69
Bibliografía	71
Índice alfabético	73



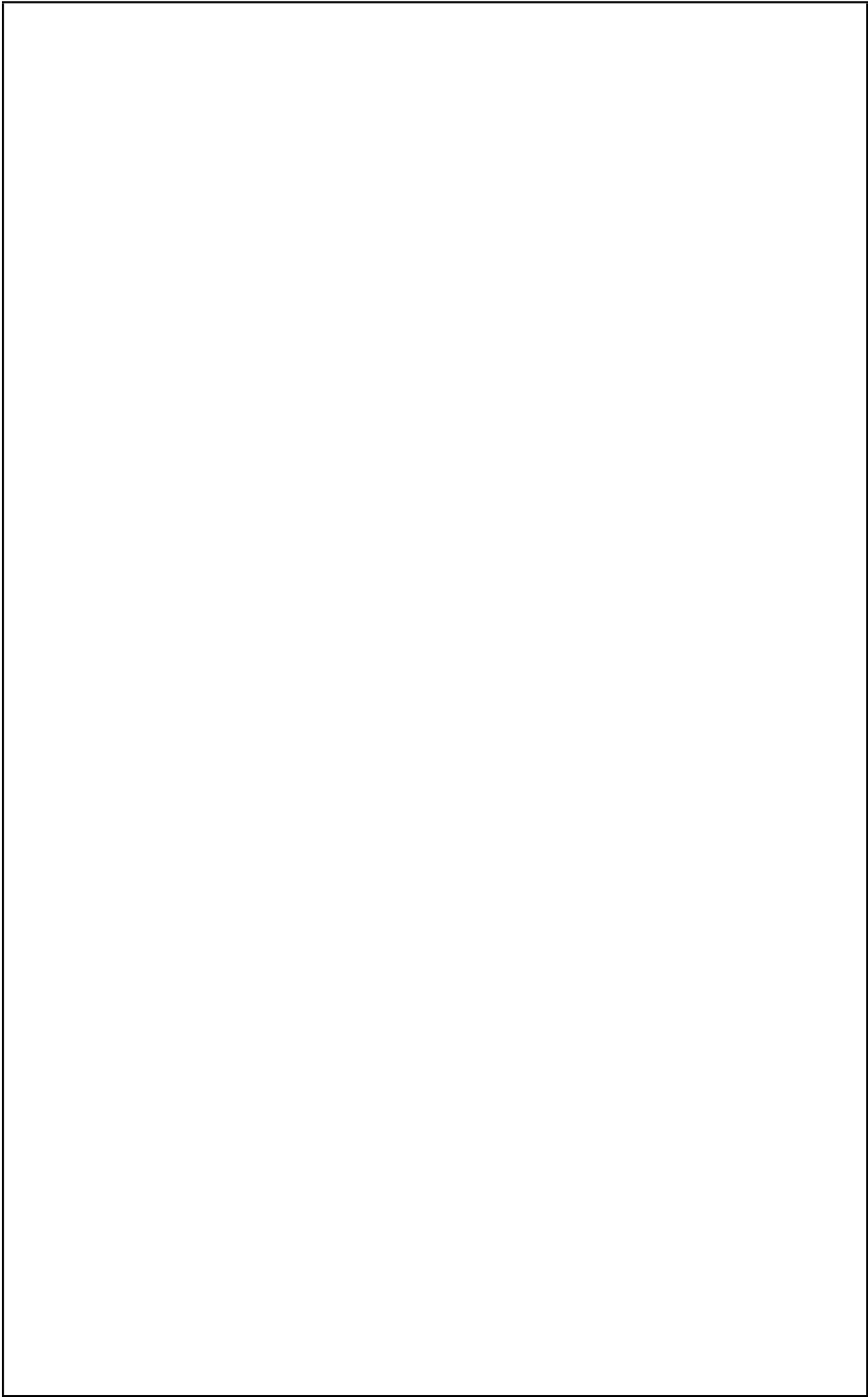
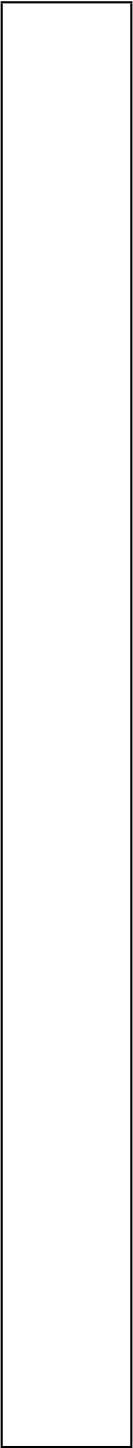
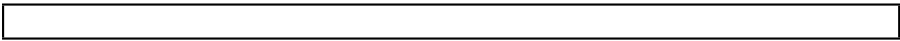
--

--



Índice de figuras

5.1. Regresión lineal	12
---------------------------------	----



Resumen

La estadística es la ciencia que manipula datos los analiza e interpreta para poder sacar conclusiones razonables de ciertos fenomenos naturales. Esta ciencia puede ser dividido en dos: **estadística descirptiva** y **estadística inferencial**. En la estadística descriptiva se procesan datos de una manera teórica y utilitaria. Estos métodos consisten en la recolección, organización, resumen, descripcion y presenatacion de la información. Si la poblacion está disponible entonces la estadística descriptiva es suficiente para describir ciertos fenomenos. No obstante generalmente no se dispone de toda la población si no de una muestra de ella, es en este caso que se requieren usar técnicas más sofisticadas para tomar decisiones y generalizaciones acerca de la poblacion, desde una pequeña muestra de información. Es cuando entra en el juego la estadística inferencial.

La base teórica de la estadística son las matemáticas

Este libro se compone de dos partes, la primera parte trata sobre la **estadística descirptiva** y la segunda **estadística inferencial**. Cada una de ellas divididas en capítulos.



Parte I

Estadística descriptiva



1

Prerrequisitos



2

Variables

Es una característica de personas cosas u objetos que son propensos a ser medidas ##
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.

2.0.1. Nominales

Son características que simplemente nominan y están propensos a ser jerarquizados u ordenados tales como: El estado civil (soltero, casado, divorciado, viudo), Religion (católico, evangelico, judio, etc).#

2.0.2. Ordinales

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

2.1. Variables cuantitativas

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

2.1.1. Discretas

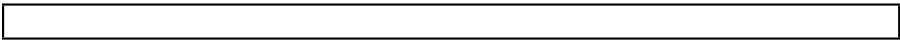
Aquellas que solo son medidos mediante numeros enteros por ejemplo: Número de hijos, número de habitaciones.

2.1.2. Continuas

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

3

Característica de los datos



4

Población, muestra y estadística



5

Organización de datos en tablas de frecuencias

- Es decir los elementos son demagogos y déspotas
- Es decir los elementos son demagogos y déspotas Tabla 5.1

Tabla 5.1: Caption

Option	N	w	Observation	Description
Es decir los elementos son demagogos y déspotas Es decir los elementos son demagogos y déspotas	1	w	Es decir los elementos son demagogos y déspotas	Es decir los elementos son demagogos y déspotas Es decir los elementos son demagogos y déspotas
Engine	2	w	Es decir los elementos son demagogos y déspotas $\sum_{i=1}^n f_i$	Engine to be used for processing templates. Handlebars is the default.
Es decir los elementos son demagogos y déspotas	3	w	$\sum_{i=1}^n f_i$	extension to be used for dest files.

The value of x in the Python session is Es decir los elementos son demagogos y déspotas. It is not the same x as the one in R.

$$f(x) = -34,093 + 13,987x + -1,167x^2 + 0,089x^3$$

$$f(29) = 1571$$

$$f(30) = 1750$$

$$f(39) = 4041$$

$$f(100) = 7,9107 \times 10^4$$

$$f(360) = 4,0252117 \times 10^6$$

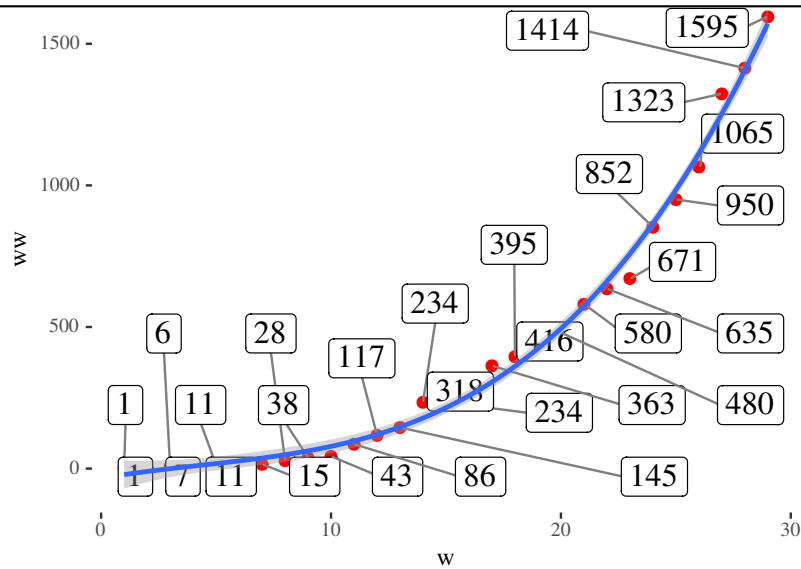


Figura 5.1 Regresión lineal

<https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6>

Sea la Tabla 5.2 Figures and tables with captions will be placed in figure and table environments, respectively. Figures and tables with captions will be placed in figure and table environments, respectively. Figures and tables with captions will be placed in figure and table environments, respectively.

Tabla 5.2: Figures and tables with captions will be placed in ‘figure’

Y_i	f_i	F_i	F_i^*	h_i	H_i	H_i^*	$h_i \%$	$H_i \%$	$H_i^* \%$
a	1	1	73	0.01	0.01	1.00	1.37	1.37	100.00
b	2	3	72	0.03	0.04	0.99	2.74	4.11	98.63
c	3	6	70	0.04	0.08	0.96	4.11	8.22	95.89
d	4	10	67	0.05	0.14	0.92	5.48	13.70	91.78
e	5	15	63	0.07	0.21	0.86	6.85	20.55	86.30
f	6	21	58	0.08	0.29	0.79	8.22	28.77	79.45
g	7	28	52	0.10	0.38	0.71	9.59	38.36	71.23
h	8	36	45	0.11	0.49	0.62	10.96	49.32	61.64
i	9	45	37	0.12	0.62	0.51	12.33	61.64	50.68
j	10	55	28	0.14	0.75	0.38	13.70	75.34	38.36
k	6	61	18	0.08	0.84	0.25	8.22	83.56	24.66
l	5	66	12	0.07	0.90	0.16	6.85	90.41	16.44

m	3	69	7	0.04	0.95	0.10	4.11	94.52	9.59
n	2	71	4	0.03	0.97	0.05	2.74	97.26	5.48
ñ	1	72	2	0.01	0.99	0.03	1.37	98.63	2.74
o	1	73	1	0.01	1.00	0.01	1.37	100.00	1.37
$\sum_{i=1}^6$	73			1.00					

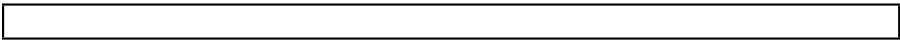
Tabla 5.3: Figures and tables with captions will be placed in ‘figure’

	\bar{x}	α	$\sum_{i=1}^n x_i$	\bar{x}	\bar{x}	\bar{x}	X7
1		IT1	IT2	O1	IT3	IT4	IT5
2	1	2	2	2	2	1	1
3	2	3	2	3	3	3	3
4	3	3	3	3	3	2	2
5	4	3	3	3	2	2	2
6	5	3	2	3	3	3	3
β_0	6	1	1	1	1	2	2
β_1	7	2	3	3	3	3	2
β_3	8	2	2	2	1	1	1
10	9	1	2	2	1	1	1
11	10	1	2	2	2	1	1
12	11	2	2	2	2	2	2
13	12	2	3	3	3	2	2
14	13	3	2	3	3	2	2
15	14	2	3	3	2	2	3
16	15	2	2	2	2	2	1
17	16	2	2	2	3	2	3
18	17	2	2	2	2	2	2
19	18	1	2	2	1	1	2
20	19	3	2	3	3	3	3
21	20	3	3	3	3	2	2
22	21	1	1	1	1	2	2
23	22	3	3	3	3	3	2
24	23	3	2	3	3	3	3
25	24	3	2	3	3	3	3
26	25	2	3	3	3	3	2
27	26	2	2	2	1	1	1
28	27	1	2	2	1	2	2
29	28	3	2	3	3	2	2
30	29	3	2	3	3	3	3
31	30	1	2	2	2	1	1

32	31	3	3	3	3	3	2
33	32	3	3	3	2	3	3
34	33	1	1	1	1	1	1
$\sum_{i=1}^n x_i$	34	1	1	1	1	1	1
36	35	3	2	3	2	2	2
37	$\sum_{i=1}^n x_i$						

6

Matriz y rango



7

Distribución de frecuencias

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 Y_i , se considera la siguiente Tabla 7.1

Tabla 7.1: Caption

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 ser 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 Sturges 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^*	$\frac{f_2}{n}$	$\frac{F_2}{n}$	$\frac{F_2^*}{n}$...

Clase	Clase	f_i	F_i	F_i^*	h_i	H_i	H_i^*	...
$< y_r - y_r >$	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 $h_i \% = 100 \cdot h_i$

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

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

Ejercicio 7.1. Sean los datos

Solución. Entonces

8

Gráficos estadísticos

The value of x in the Python session is Es decir los elementos son demagogos y déspotas. It is not the same x as the one in R.



9

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.

9.1. La media (\bar{x})

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

9.1.1. 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

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i \quad (9.1)$$

$$\frac{d[P; F_1]}{d[P; \mathcal{L}_1]} = e = \frac{d[P; F_2]}{d[P; \mathcal{L}_2]} \quad (9.2)$$

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

9.1.2. 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 9.1. 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

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

9.2. La moda (Mo)

9.2.1. 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 $Mo = x_2$

9.2.2. 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
-

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

can be found on the Pandoc website <http://pandoc.org>.

$$\sum$$

> 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 $B_{r \times s}$

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

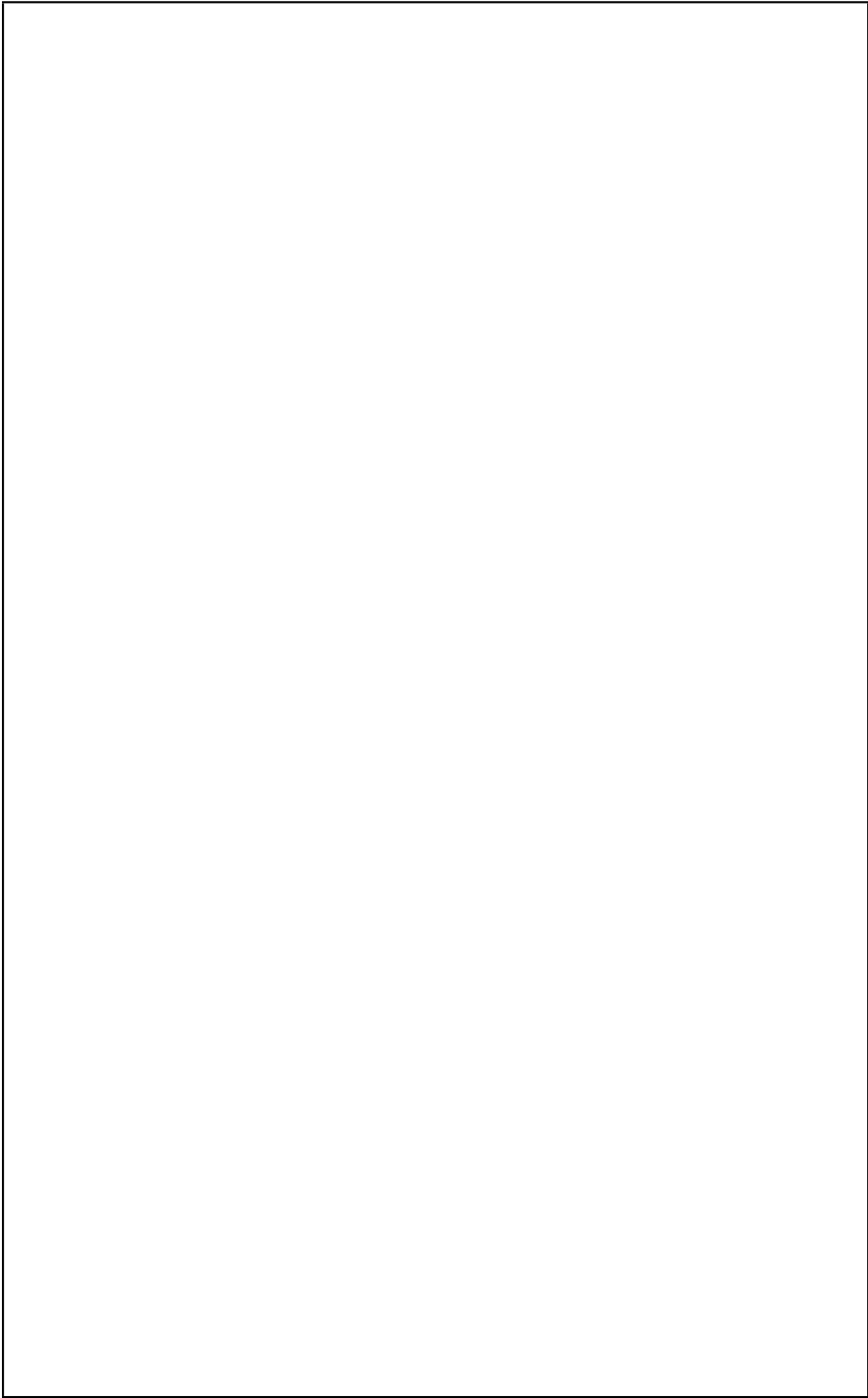
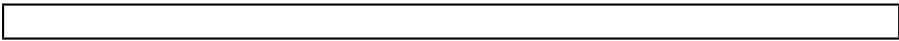
$$X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{bmatrix}$$

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

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

10

Medidas de dispersión



11

Medidas de asimetría



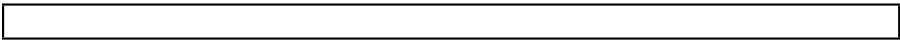
Parte II

Probabilidades



1

Experimento aleatorio



2

Álgebra de eventos



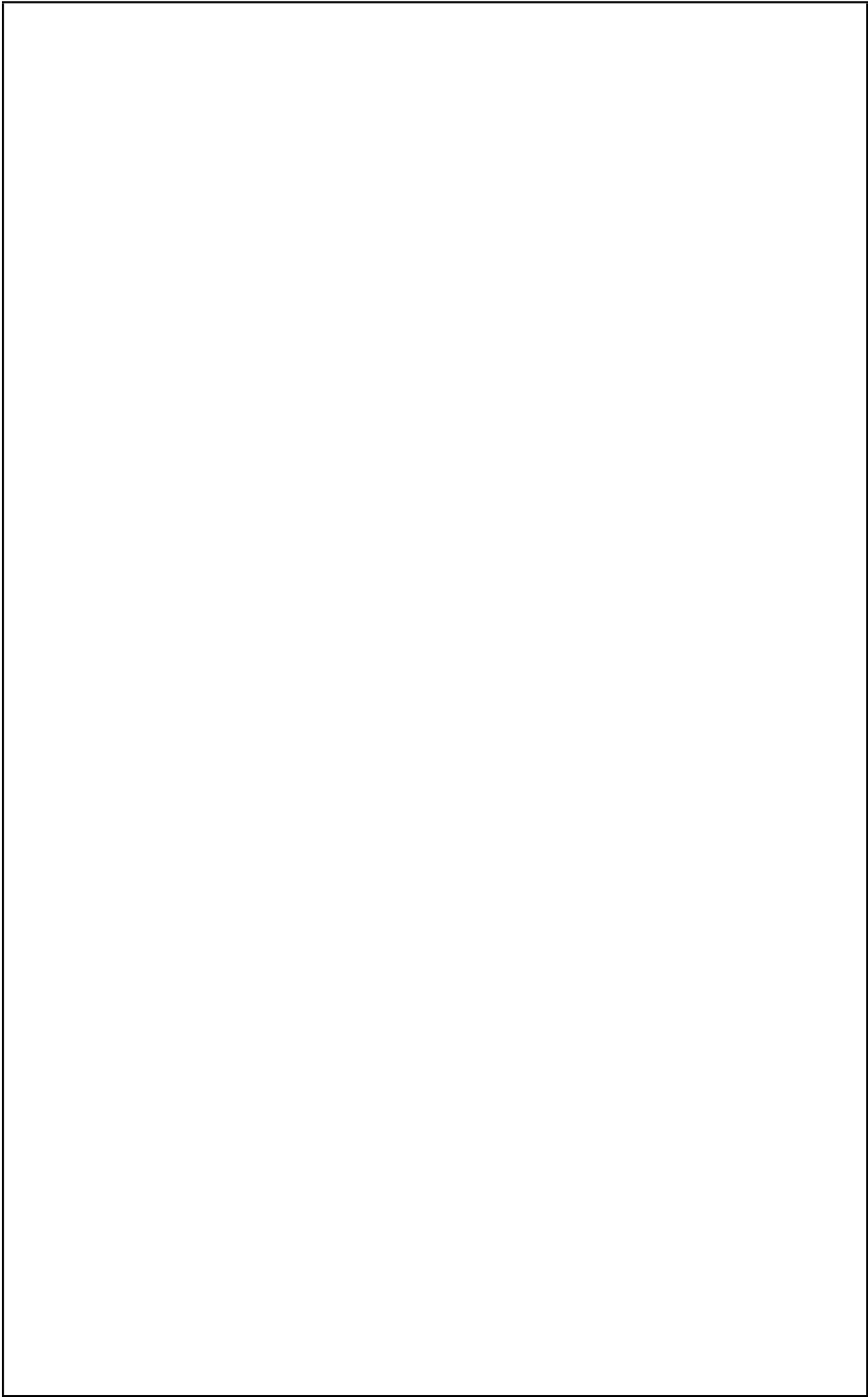
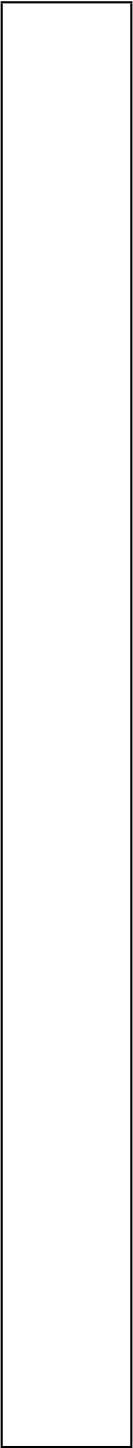
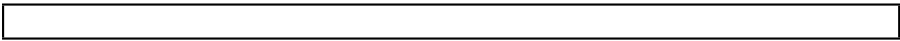
3

Técnicas de conteo



4

Definición de probabilidad



5

Probabilidad condicional



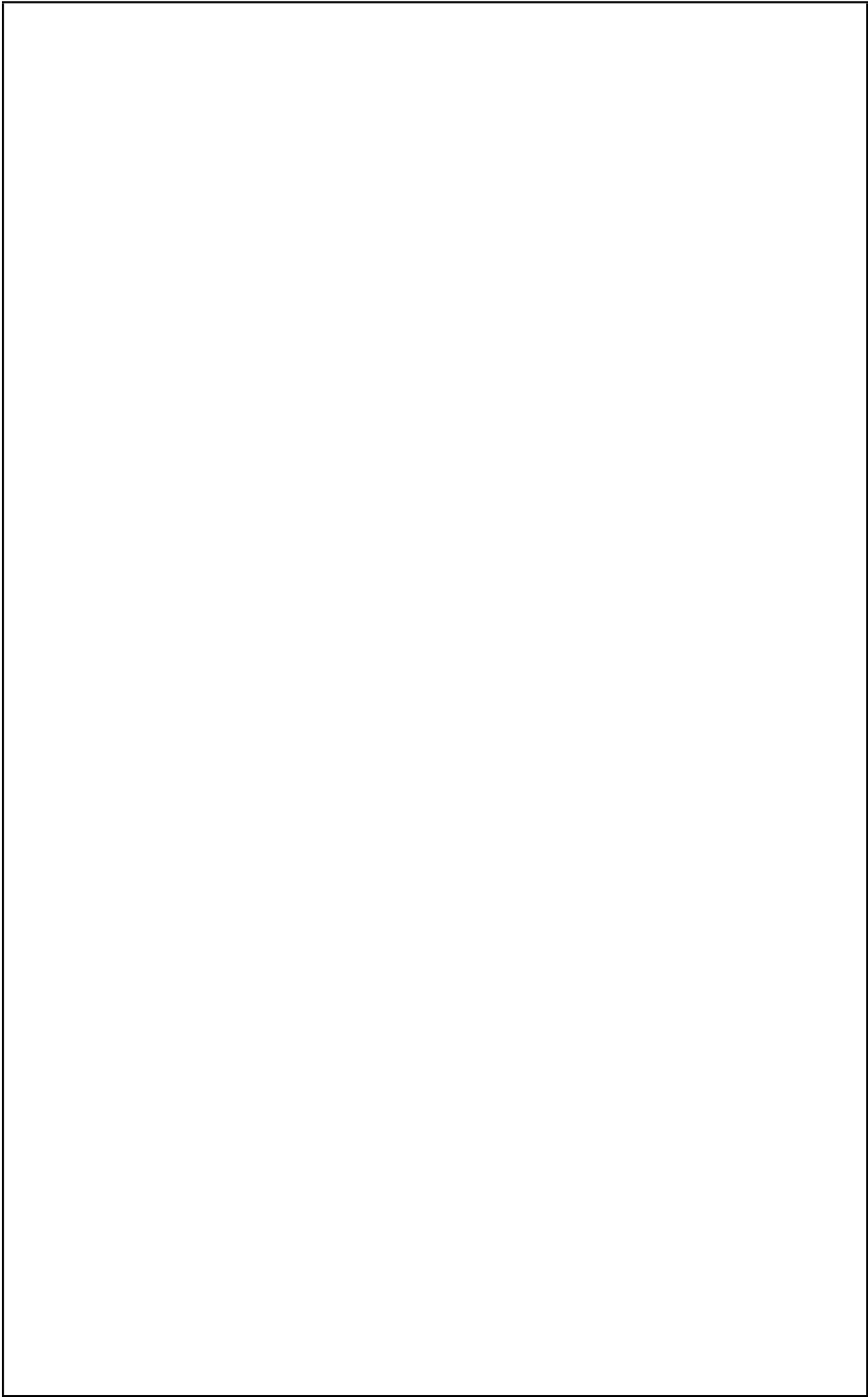
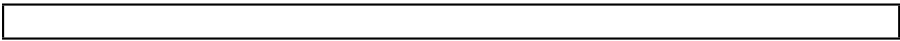
6

Teorema de Bayes



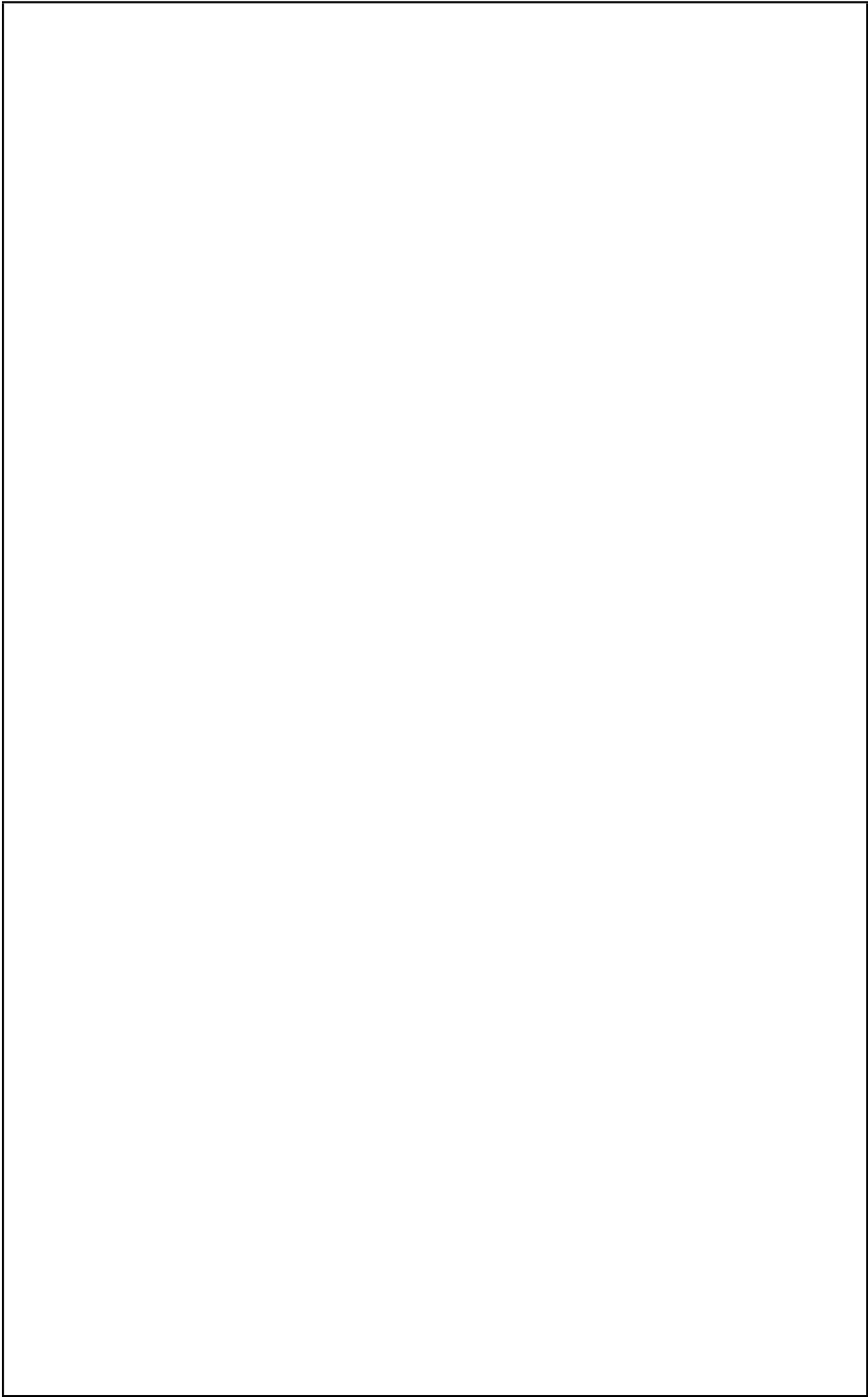
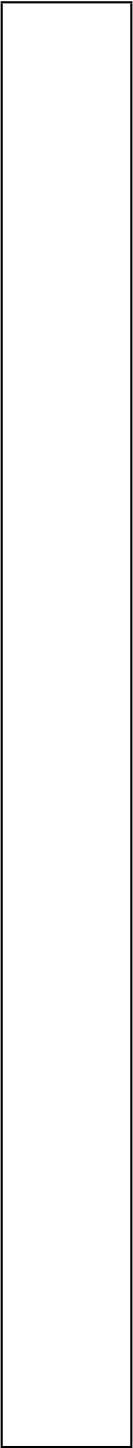
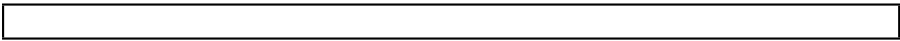
7

Eventos independientes y secuencias de experimentos



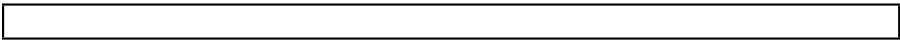
8

Probabilidad en espacio



Parte III

Inferencia estadística



1

Variables aleatorias

Definición 1.1 (Variable aleatoria). Sea Ω un espacio muestral asociado a una experimento aleatorio ϵ y $\omega \in \Omega$, entonces se genera la función **variable aleatoria**

$$\begin{aligned} X : \Omega &\longrightarrow \mathbb{R} \\ \omega &\longmapsto X(\omega) \end{aligned}$$

es decir a cada elemento de Ω se le asocia un número real \mathbb{R} , además la probabilidad de $x \in \mathbb{R}$ es $P[x] = \sum_{i=1}^n P[\omega_i]$ donde $\omega_i \in X^{-1}(x)$. La definición indica por otro lado que un espacio muestral Ω puede genera diferentes variables aleatorias. Por ejemplo el espacio muestral de lanzar una monedas tres veces es $\omega = \{ccc, ccs, csc, scc, css, scs, ssc, sss\}$ además sea n_c es número de caras y n_s el número de sellos, es posibles generar dos o mas variables aleatorias por ejemplo:

- $X(\omega) = n_c$ entonces el rango de X es $R_X \{3, 2, 1, 0\}$ pues

$$\begin{aligned} 3 &= X(ccc) \\ 2 &= X(ccs) = X(csc) = X(scc) \\ 1 &= X(css) = X(scs) = X(ssc) \\ 0 &= X(sss) \end{aligned}$$

- $X(\omega) = n_c - n_s$ entonces las imagenes de X son $R_X = \{3, 1, -1, -3\}$ en efecto

$$\begin{aligned} 3 &= X(ccc) \\ 1 &= X(ccs) = X(csc) = X(scc) \\ -1 &= X(css) = X(scs) = X(ssc) \\ -3 &= X(sss). \end{aligned}$$

Estos subconjuntos de \mathbb{R} también son espacios muestrales pues el conjunto de elementos de Ω con imagen dentro de estos valores reales x en \mathbb{R} es un elemento de 2^Ω es decir un evento por lo tanto tiene una determinada probabilidad $P[x]$, en el primer caso $X(\omega) = n_c$ tienen probabilidades

$$P(3) = P[ccc] = \frac{1}{8}$$

$$P(2) = P[ccs] = P[scs] = P[scs] = \frac{3}{8}$$

$$P(1) = P[css] = P[scs] = P[ssc] = \frac{3}{8}$$

$$P(0) = P[sss] = \frac{1}{8}$$

que es lo mismo para el segundo caso $X(\omega) = n_c - n_s$.

Definición 1.2 (Eventos equivalentes). Sea Ω un espacio muestral asociado a una experimento aleatorio ϵ y X una variable aleatoria con rango R_X definida sobre Ω . Dos eventos $W \in \Omega$ y $E_X \in R_X$ son **eventos equivalentes** si existe la relación

$$W = \{\omega \in \Omega / X(\omega) = E_X\}$$

es decir E_X consta de todos los elementos en Ω para los cuales $X(\omega) \in W$

1.1. Clases de variables aleatorias

1.1.1. Variable aleatoria discreta

R_X es *finito* o *infinito* P contable (no necesariamente enteros) $R_X = \{x_1, x_2, \dots, x_n \dots\}$

$$p_X(x) = p(1-p)^{x-1}$$

$$x \in \mathbb{Z}^+ \quad p \in [0, 1]$$

$$F_X(x) = P[X = x]$$

- $0 \leq F_X(x) \leq 1$
- $F_X(-\infty) = 0$
- $F_X(\infty) = 1$
- $P(X < x) = F_X(x^-)$
- $P(a \leq X \leq b) = F_X(b) - F_X(a^-)$

1.1.2. Variable aleatoria continua

R_X abarca cualquier intervalo en la recta numerica

Sea X una VA continua con rango R_X La funcion de densidad de probabilidad asociado a una variable aleatoria $f(x)$

- $f(x) \geq 0, x \in R_X$
- $\int_{R_X} f(x)dx = 1$ o $\int_{-\infty}^{\infty} f(x)dx = 1$

1.1.3. Variable aleatoria mixta

Discreta y continua

1.2. Función de probabilidad de una variable aleatoria**1.2.1. Función de probabilidad de una variable aleatorias discreta**

Definición 1.3 (Funcion de probabilidad). Sea X una variable aleatoria con rango R_X . Una funcion definida por

$$p(x) = P[X = x] = \sum_{\{\omega \in \Omega: X(\omega)=x\}} P[\{\omega\}]$$

- $p(x) \geq 0, x \in R_X$
- $\sum_{x \in R_X} p(x) = P[X = x] = 1$

El conjunto de pares ordenados $(x, p(x)), x \in R_X$ recibe el nombre de *distribución de probabilidad de X*

1.2.2. Función de probabilidad de una variable aleatoria continua

Definición 1.4 (Función de densidad de probabilidad). Sea X una variable aleatoria con rango R_X . La función $f(x)$ definida sobre R_X

- $f(x) \geq 0, x \in R_X$ o $f(x) \geq 0, x \in \mathbb{R}$
- $\int_{R_X} f(x)dx = 1$ o $\int_{-\infty}^{\infty} f(x)dx = 1$

1.3. Función de distribución de una variable aleatoria

1.3.1. Función de distribución de una variable aleatoria discreta

Definición 1.5 (Función de distribución). Sea X una variable aleatoria con rango $R_X = \{x_1, x_2, \dots, x_n, \dots\}$. Con función de probabilidad $p(x_i) = P[X = x_i]$, sea x cualquier número, real la función definida por

$$F(x) = P[X \leq x] = \sum_{x_i \leq x} p(x_i) = \sum_{x_i \leq x} P[X = x_i]$$

recibe el nombre de función de distribución de X .

1.3.2. Función de distribución de una variable aleatoria continua

2

Estadigrafos de la variables aleatorias

2.1. esperanza matematica

2.2. Medidas de tendencia central

2.3. Medidas de variación

2.4. Cuantiles



3

Variables aleatorias bidimensionales

3.1. Distribuciones marginales

3.2. Distribuciones



4

Distribuciones discreta importantes

4.1. Variable aleatoria discreta binomial

4.2. Variable aleatoria discreta Poisson

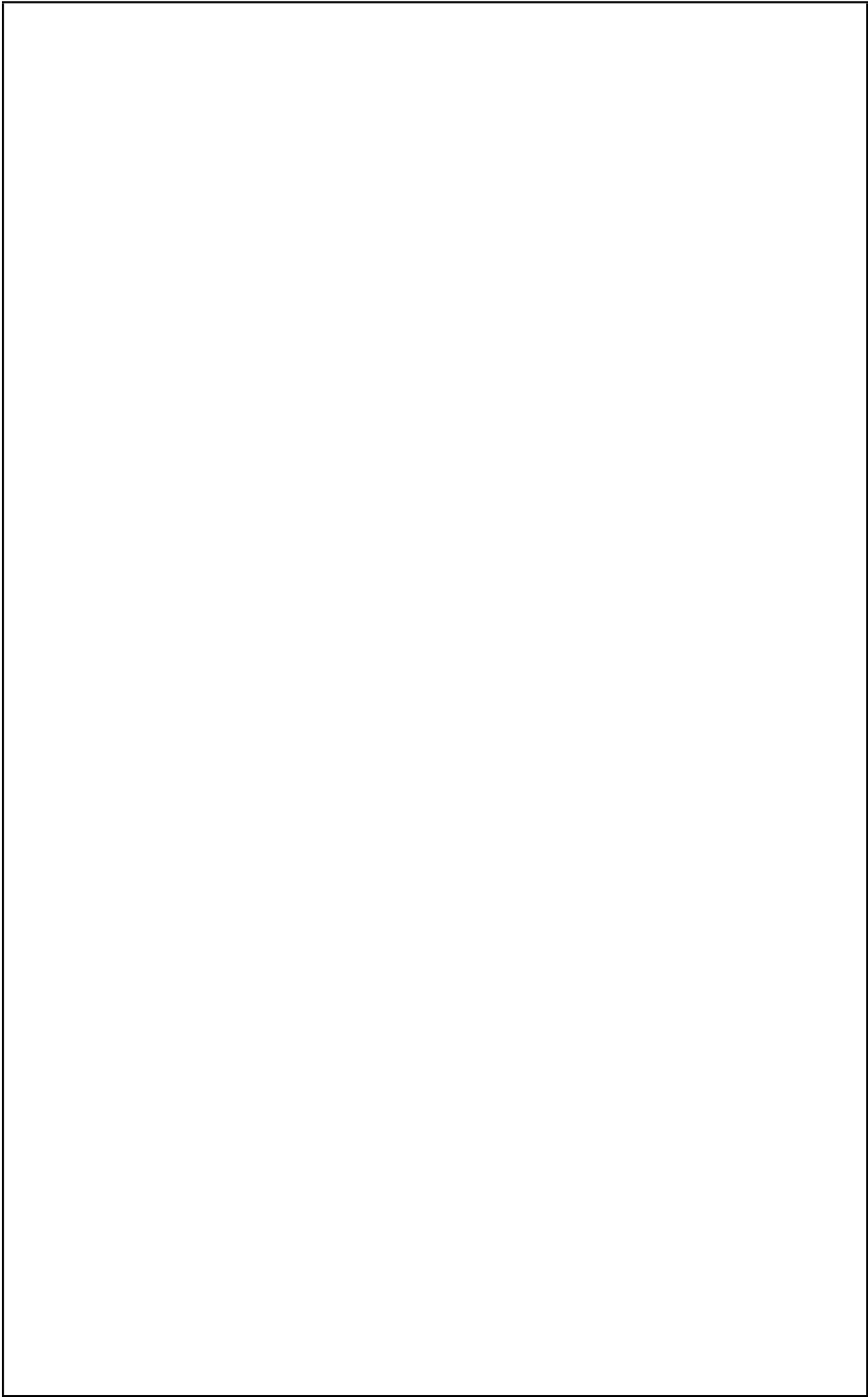
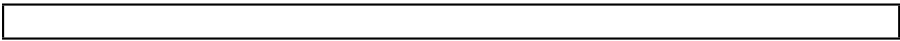


5

Distribuciones continuas importantes

5.1. Variable aleatoria continua normal

5.2. Variable aleatoria continua gamma



6

Distribuciones muestrales



7

Estimación



8

Prueba de hipótesis



A

Sumatorias

Una suma de números representados por x_1, x_2, \dots, x_n se simboliza en forma compacta mediante el símbolo \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) Variable estadística variable estadística
ee

A.1. eeeee



B

Matrices

Una matriz es un arreglo de números distribuidos en filas y columnas por ejemplo la siguiente matriz

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}_{n \times n}$$

de **orden** $n \times m$ tiene **entradas** a_{ij} donde el primer subíndice indica la fila y el segundo la columna; es usual representar por simplicidad una matriz por $A = [a_{ij}]_{n \times m}$. Si en el orden $n = m$ entonces la matriz recibe el nombre de **matriz cuadrada** la suma de los elementos de la diagonal de una matriz cuadrada $\sum_{i=1}^n a_{ii}$ se llama **traza**. Si todas las a_{ij} son cero entonces la matriz $A = 0$ recibe el nombre matriz **nula**.

Dos matrices son iguales si tienen el **mismo orden** y cada una de las entradas respectivas son iguales es decir $A = [a_{ij}]_{n \times m}$ y $B = [b_{ij}]_{n \times m}$ son iguales si $a_{ij} = b_{ij}$, $i = 1, 2, \dots, n$ y $j = 1, 2, \dots, m$

B.1. Álgebra de matrices

Sean las matrices $A = [a_{ij}]_{n \times m}$ y $B = [b_{ij}]_{p \times q}$ entonces la suma y producto de matrices se definen

1. Sea k un escalar entonces se verifica que $kA = [ka_{ij}]$, $i = 1, 2, \dots, n$ y $j = 1, 2, \dots, m$ es decir el escalar k multiplica a cada una de las entradas de la matriz.
2. La suma o diferencia es posible si $n = p$ y $m = q$ es decir los ordenes de A y B son iguales, entonces la suma o diferencia resulta $A \pm B = [a_{ij} \pm b_{ij}]_{n \times m}$, $i = 1, 2, \dots, n$ y $j = 1, 2, \dots, m$
3. El producto es posible si $m = p$ es decir el número columnas de la primera matriz es igual al número de filas de la segunda matriz, el orden de la

matriz resultante es $n \times q$ además

$$A \cdot B = \left[\sum_{k=1}^p a_{ik} b_{kj} \right]_{n \times q}$$

$$= \begin{pmatrix} \sum_{k=1}^m a_{1k} b_{k1} & \sum_{k=1}^m a_{1k} b_{k2} & \cdots & \sum_{k=1}^m a_{1k} b_{kq} \\ \sum_{k=1}^m a_{2k} b_{k1} & \sum_{k=1}^m a_{2k} b_{k2} & \cdots & \sum_{k=1}^m a_{2k} b_{kq} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^m a_{nk} b_{k1} & \sum_{k=1}^m a_{nk} b_{k2} & \cdots & \sum_{k=1}^m a_{nk} b_{kq} \end{pmatrix}_{n \times q}$$

donde $i = 1, 2, \dots, n$ y $j = 1, 2, \dots, m$

Ejemplo B.1. Sean $\begin{pmatrix} 3 & -1 & 2 \\ 2 & -1 & 2 \\ 1 & -1 & 0 \\ 5 & 0 & 0 \end{pmatrix}_{4 \times 3}$ y $\begin{pmatrix} 0 & -1 & 2 & 2 & 0 \\ 1 & -1 & -2 & 1 & 1 \\ 3 & -1 & -3 & 5 & 2 \end{pmatrix}_{3 \times 5}$ entonces $A \cdot B =$

$$\begin{pmatrix} 5 & -4 & 2 & 15 & 3 \\ 5 & -3 & 0 & 13 & 3 \\ -1 & 0 & 4 & 1 & -1 \\ 0 & -5 & 10 & 10 & 0 \end{pmatrix}_{4 \times 5}$$

En caso de ser posible la multiplicación entre A , B y C entonces se verifican las siguientes propiedades

- $A(B + C) = AB + AC$
- $(A + B)C$
- $A(BC) = (AB)C$

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, 18
frecuencias absolutas acumuladas menor
 que, 18
frecuencias absolutas relativas, 18
frecuencias absolutas relativas menor
 que, 18
traza, 69