

Ricardo Michel MALLQUI BAÑOS

Elementos de la estadística

estadística descriptiva y probabilidades



Índice general

Índice de tablas	vii
Índice de figuras	ix
Resumen	xi
I Estadística descriptiva	1
1. Prerrequisitos	3
2. Variables	5
2.1. Variables cualitativas	5
2.1.1. Nominales	5
2.1.2. Ordinales	5
2.2. Variables cuantitativas	5
2.2.1. Discretas	5
2.2.2. Continuas	6
2.3. Asignación	6
3. Organización de datos en tablas de frecuencias	9
3.1. Distribución de frecuencias	9
3.2. Ejemplo sin intervalos	11
3.3. Example intervalos	15
4. Gráficos estadísticos	17
4.1. Histograma de frecuencias	17
4.2. Circulares	17
4.3. Histograma de frecuencias	17
5. Medidas de tendencia central	19
5.1. La media (\bar{x})	19
5.1.1. Media de datos no agrupados	19
5.1.2. Media de datos agrupados	19
5.2. La moda (Mo)	20
5.2.1. Moda de datos no tabulados	20
5.2.2. Moda de datos tabulados	20
5.3. La mediana (Me)	21

5.3.1. Mediana de datos no tabulados	21
5.3.2. Mediana de datos tabulados	21
6. Medidas de dispersión	23
6.1. Varianza	23
7. Medidas de asimetría	25
II Probabilidades	27
1. Experimento aleatorio	29
2. Álgebra de eventos	31
3. Técnicas de conteo	33
4. Definición de probabilidad	35
5. Probabilidad condicional	37
6. Teorema de Bayes	39
7. Eventos independientes y secuencias de experimentos	41
8. Probabilidad en espacio	43
III Inferencia estadística	45
1. Variables aleatorias	47
1.0.1. Variable aleatoria continua	48
1.0.2. Variable aleatoria mixta	48
1.1. Función de probabilidad de una variable aleatoria	49
1.1.1. Función de probabilidad de una variable aleatorias discreta	49
1.1.2. Función de probabilidad de una variable aleatoria continua	50
1.2. Función de distribución de una variable aleatoria	50
1.2.1. Función de distribución de una variable aleatoria discreta	50
1.2.2. Función de distribución de una variable aleatoria continua	51
2. Parámetros de una variable aleatoria	53
2.1. Esperanza matemática	53
2.2. Medidas de variación	53
2.3. Medidas de posición	54
2.4. Medidas de curtosis	54

Contents	v
3. Variables aleatorias bidimensionales	55
3.0.1. Distribuciones marginales	56
3.0.2. Variables aleatorias independientes	56
3.0.3. Distribuciones de probabilidad condicional	56
3.1. Distribución bidimensional continua	56
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	75
Índice alfabético	77



Índice de tablas

3.1. Caption	9
3.2. Datos cuantitativos (intervalos)	10
3.4. Datos cualitativos	14
3.6. Datos cuantitativos (intervalos)	15
5.1. Captionww	19
5.2. wwwww	21
5.3. Mediana	22
B.1. Caption	71
B.2. Figures and tables with captions will be placed in ‘figure’ . .	72



Índice de figuras

B.1. Regresión lineal	72
---------------------------------	----



Resumen

La estadística es la ciencia que manipula datos los analiza e interpreta para poder sacar conclusiones razonables de ciertos fenómenos naturales. Esta ciencia puede ser dividida en dos: estadística descriptiva y estadística inferencial. En la estadística descriptiva se procesan datos de una manera teórica y utilitaria. Estos métodos consisten en la recolección, organización, resumen, descripción y presentación de la información. Si la población está disponible entonces la estadística descriptiva es suficiente para describir ciertos fenómenos. No obstante generalmente no se dispone de toda la población sino de una muestra de ella, es en este caso que se requieren usar técnicas más sofisticadas para tomar decisiones y generalizaciones acerca de la población, 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 descriptiva y la segunda estadística inferencial. Cada una de ellas divididas en capítulos.



Parte I

Estadística descriptiva



1

Prerrequisitos

www



2

Variables

Es una característica de personas cosas u objetos que son propensos a ser medidas o cualificadas

2.1. 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.

2.1.1. Nominales

Son características que simplemente nominan y están propensos a ser jerarquizados u ordenados tales como: El estado civil (soltero, casado, divorciado, viudo), Religión (católica, evangélico, judío, etc).

2.1.2. Ordinales

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

2.2. Variables cuantitativas

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

2.2.1. Discretas

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

2.2.2. Continuas

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

2.3. Asignación

1. Reconosca 5 variables cualitativas de una persona, admosfera, una pintura

- Persona

- Color (moreno, blanco, triguelo)
- Religión (catolico, evangelico, pentecostal, etc)
- wwwwww
- www
- wwwwww

- Admosfera

- www
- wwwwww
- wwwwww
- www
- wwwwww

- Pintura

- www
- wwwwww
- wwwwww
- www
- wwwwww

2. Reconosca 5 variables cuantitativas de una video, tela, un celular.

- Video

- Duracion (x segundos)
- Numero video en youtube a la semana (n cantidades)
- wwwwww
- www
- wwwwww

- Tela

- www
- wwwwww
- wwwwww
- www
- wwwwww

- Celular

2.3 Asignación

7

- www
- wwwwww
- wwwwwwwww
- wwwwww
- wwwwwwwww



3

Organización de datos en tablas de frecuencias

3.1. 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 3.1

Tabla 3.1: 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_2^*
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_3^*
\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_r^*

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

1. Criterio del investigador r no puede ser más de 20 ni menos de 5
2. $r = \sqrt{n}$ donde n es el número de datos
3. La regla de Stargess 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 $l = \frac{L}{r}$

Tabla 3.2: Datos cuantitativos (intervalos)

Clase	Y_i	f_i	F_i	F_i^*	h_i	H_i	H_i^*	$h_i\%$	$H_i\%$	$H_i^*\%$
$[y_1 - y_2)$	Y_1	f_1	F_1	F_1^*	$\frac{f_1}{n}$	$\frac{f_1}{n}$	$\frac{F_1^*}{n}$	$h_1\%$	$H_1\%$	$H_1^*\%$
$[y_2 - y_3)$	Y_2	f_2	F_2	F_2^*	$\frac{f_2}{n}$	$\frac{f_2}{n}$	$\frac{F_2^*}{n}$	$h_2\%$	$H_2\%$	$H_2^*\%$
...
$[y_{r-1} - y_r)$	Y_r	f_r	F_r	F_r^*	$\frac{f_r}{n}$	$\frac{f_r}{n}$	$\frac{F_r^*}{n}$	$h_r\%$	$H_r\%$	$H_r^*\%$

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.

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

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

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

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

4. Las frecuencias absolutas relativas obedecen a la fórmula

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

5. Las frecuencias absolutas relativas menor que obedecen a la fórmula

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

6. Las frecuencias absolutas relativas mayor que obedecen a la fórmula

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

7. Las frecuencias absolutas relativas porcentuales obedecen a la fórmula $h_i \% = 100 \cdot h_i$
 8. Las frecuencias absolutas relativas menor que porcentuales obedecen a la fórmula $H_i \% = 100 \cdot H_i$
 9. Las frecuencias absolutas relativas mayor que porcentuales obedecen a la fórmula $H_i^* \% = 100 \cdot H_i^*$
-

3.2. Ejemplo sin intervalos

Ejercicio 3.1. Sean Los 16 tipos de personalidad en un grupo social encuestado.

1. ESTJ (Extraverted Sensing Thinking Judging) Personas a las que les gusta tener el control sobre lo que ocurre a su alrededor, siempre buscan la manera de que todo funcione como debe y, si es necesario, la implementan ellos mismos.

2. ESTP ((Extraverted Sensing Thinking Perceiving) Las personas que pertenecen a esta categoría son espontáneas, alegres y activas, pero al igual que lo que ocurre con los ESTJ, tienden a ejercer dominio sobre los demás, en este caso a través de su capacidad de observación y su carisma.

3. ESFJ (Extraverted Sensing Feeling Judging) Se trata de personas muy volcadas en la atención de las necesidades de los demás, especialmente si forman parte de su círculo cercano: familia y amistades. Por eso siempre que pueden prestan su ayuda y procuran que sus círculos sociales cercanos permanezcan siempre estables y con buena salud. Por eso tienden a evitar que aparezcan conflictos fuertes y se muestran diplomáticas cuando hay choques de intereses.

4. ESFP (Extraverted Sensing Feeling Perceiving) Se trata de personas alegres y espontáneas que disfrutan entreteniéndose y entreteniéndolo a los demás. La diversión es uno de los pilares más importantes de sus vidas, y son de trato cercano y temperamento cálido. Aman la novedad y hablar acerca de experiencias personales.

5. ISTJ (Introverted Sensing Thinking Perceiving) Un tipo de personalidad definido por su fuerte sentido de la moralidad y del deber. Les gusta planear e implementar sistemas de reglas que permitan que equipos y organizaciones funcionen con una clara lógica y orden. Dan un gran valor a las normas y a la necesidad de que la realidad se corresponda con cómo deberían ser las cosas. Aunque son personas introvertidas, no rehuyen la interacción con los demás.

6. ISTP (Introverted Sensing Thinking Perceiving) Se trata de personas reservadas, orientadas a la acción y a las soluciones prácticas ante problemas del día a día. También son definidas por su tendencia hacia el pensamiento lógico y su espontaneidad y autonomía. Les gusta explorar entornos y descubrir modos en los que se puede interactuar con ellos.

7. ISFJ (Introverted Sensing Feeling Judging) Son personas definidas principalmente por sus ganas de proteger y ayudar a los demás y, en definitiva, de resultar confiables para los otros. Se esfuerzan por hacer todo lo que se espera de ellas, pero no tienen grandes aspiraciones ni se muestran muy ambiciosas. Tienden a pensar que es malo pedir compensaciones o aumentos a cambio de los sacrificios que realizan a la hora de trabajar, ya que este debería ser una meta en sí.

8. ISFP (Introverted Sensing Feeling Perceiving) Personas que viven totalmente en el aquí y el ahora, en constante búsqueda de la novedad y de las situaciones sensorialmente estimulantes. Son reservadas, pero también alegres, espontáneas y cálidas con sus amistades. Tienen un especial talento en el mundo de las artes.

9. ENTJ (Extraverted Intuitive Thinking Judging) Este es uno de los 16 tipos de personalidad más relacionados con el liderazgo y la asertividad. Las personas descritas por esta categoría son comunicativas, de pensamiento ágil y analítico y predispuestas a encabezar equipos y organizaciones. Se adaptan bien al cambio y hacen que sus estrategias también se amolden cada vez que el entorno varía. Además, casi siempre saben cómo explicar sus proyectos o historias de manera que resulten de interés para el resto, lo cual los convierte en comerciales muy aptos.

10. ENTP (Extraverted Intuitive Thinking Perceiving) Personas especialmente movidas por la curiosidad y por los retos que para ser resueltos requieren afrontar preguntas intelectualmente estimulantes. Su agilidad mental y su facilidad para detectar inconsistencias lógicas hace de ellas personas predispuestas a interesarse por la ciencia o la filosofía. Además, su tendencia a mostrarse competitivas las vuelve personas muy activas durante el día, siempre intentando llegar a soluciones innovadoras a problemas complejos.

11. ENFJ (Extraverted Intuitive Feeling Judging) Personas que aprenden constantemente acerca de todos los ámbitos del conocimiento (o una buena parte de ellas) y ayudan a aprender a las demás, guiándolas en su propia evolución. Les gusta ofrecer tutela y consejo, y son muy buenas influyendo en la conducta de los demás. Se centran en sus valores e ideales y hacen lo posible por mejorar

el bienestar del mayor número de personas a través de sus ideas y sus acciones.

12. ENFP (Extraverted Intuitive Feeling Perceiving) Uno de los 16 tipos de personalidad con mayor propensión al pensamiento creativo, las artes y a la sociabilidad. Son alegres, disfrutan de la interacción con otras personas, y actúan teniendo en mente su posición como parte de un “todo” formado por la humanidad, y no se muestran individualistas. De hecho, suelen involucrarse en tareas colectivas para ayudar a los demás, pensando en el impacto social de sus acciones. Sin embargo, también se distraen fácilmente y es frecuente que posterguen tareas que consideran aburridas o demasiado simples y rutinarias.

13. INTJ (Introverted Intuitive Thinking Judging) Un tipo de personalidad orientado hacia la resolución de problemas específicos a partir del razonamiento analítico. Las descritas por esta categoría son personas muy centradas en sus propias ideas y teorías acerca del funcionamiento del mundo, lo cual significa que analizan su entorno centrándose en sus ideas sobre cómo opera este. Son conocedoras de sus propias capacidades y confían en su propio criterio, aunque este vaya en contra de algunos superiores.

Es muy frecuente que lleguen a ser expertas en un ámbito de conocimiento muy específico, ya que les gusta tener el suficiente conocimiento sobre algo como para poder tener en cuenta todos los factores que entran en juego en su funcionamiento y, a partir de ahí, saber lo que se puede hacer o lo que pasará en el futuro.

14. INTP (Introverted Intuitive Thinking Perceiving) Uno de los 16 tipos de personalidad más definido por la propensión a la reflexión. A estas personas les gustan las teorías con capacidad para explicar todo lo que puede ocurrir en un sistema, y su tendencia hacia el perfeccionismo hace que corrijan a los demás en múltiples ocasiones. Valoran más la exactitud en términos teóricos que el pragmatismo y la resolución de problemas concretos.

15. INFJ (Introverted Intuitive Feeling Judging) Personas muy sensibles, reservadas y movidas por unos ideales muy definidos y que, además, sienten la necesidad de hacer que los demás también se beneficien de estos ideales. Esto hace que sean propensas tanto a la reflexión como a la acción, lo cual puede llegar a suponer tanto trabajo que se sobrecargan por tener demasiadas responsabilidades. Muestran una gran capacidad para interpretar exitosamente los estados mentales de los demás y tratan de utilizar esta información para ayudarlas antes de que la otra persona se lo pida.

16. INFP (Introverted Intuitive Feeling Perceiving) Menos moralistas que los INFJ, los INFP también se preocupan mucho por ayudar a los demás desde su posición de personas reservadas. Muestran una sensibilidad estética y artística que las vuelve creativas.

Solución. Se tiene 16 características con los siguientes datos

Personalidad	Cantidad
ESTJ	1
ESTJ	2
ESTP	3
ESFJ	4
ESFP	5
ISTJ	6
ISTP	7
ISFJ	8
ISFP	9
ENTJ	10
ENTP	6
ENFJ	5
ENFP	3
INTJ	2
INTP	1
INFJ	1
INFP	1

Tabla 3.4: Datos cualitativos

Y_i	f_i	F_i	F_i^*	h_i	H_i	H_i^*	$h_i \%$	$H_i \%$	$H_i^* \%$
ESTJ	1	1	75	0.01	0.01	1.00	1.33	1.33	100.00
ESTJ	2	3	150	0.03	0.04	2.00	2.67	4.00	200.00
ESTP	3	6	150	0.04	0.08	2.00	4.00	8.00	200.00
ESFJ	4	10	150	0.05	0.13	2.00	5.33	13.33	200.00
ESFP	6	16	150	0.08	0.21	2.00	8.00	21.33	200.00
ISTJ	6	22	155	0.08	0.29	2.07	8.00	29.33	206.67
ISTP	7	29	157	0.09	0.39	2.09	9.33	38.67	209.33
ISFJ	8	37	158	0.11	0.49	2.11	10.67	49.33	210.67
ISFP	9	46	166	0.12	0.61	2.21	12.00	61.33	221.33
ENTJ	10	56	166	0.13	0.75	2.21	13.33	74.67	221.33
ENTP	6	62	166	0.08	0.83	2.21	8.00	82.67	221.33
ENFJ	5	67	166	0.07	0.89	2.21	6.67	89.33	221.33
ENFP	3	70	166	0.04	0.93	2.21	4.00	93.33	221.33
INTJ	2	72	167	0.03	0.96	2.23	2.67	96.00	222.67

INTP	1	73	169	0.01	0.97	2.25	1.33	97.33	225.33
INFJ	1	74	172	0.01	0.99	2.29	1.33	98.67	229.33
INFP	1	75	176	0.01	1.00	2.35	1.33	100.00	234.67
TOTAL	75						100.00		

3.3. Example intervalos

Solución. Edades de cierta comunidad

25 35 38 45 47 48 51 52 53 55 60 62 63 66 67 70 71 72 75 77 78 81 88 89 90 99

Clase	f_i
[20 – 30 >	1
[30 – 40 >	2
[40 – 50 >	3
[50 – 60 >	4
[60 – 70 >	5
[70 – 80 >	6
[80 – 90 >	3
[90 – 100]	2

Tabulando

Tabla 3.6: Datos cuantitativos (intervalos)

Clase	f_i	F_i	F_i^*	h_i	H_i	H_i^*	$h_i\%$	$H_i\%$	$H_i^*\%$
[20 – 30)	1	1	26	0.02	0.01	0.35	2.38	1.33	34.67
[30 – 40)	2	3	68	0.05	0.04	0.91	4.76	4.00	90.67
[40 – 50)	3	6	68	0.07	0.08	0.91	7.14	8.00	90.67
[50 – 60)	4	10	68	0.10	0.13	0.91	9.52	13.33	90.67
[60 – 70)	5	15	68	0.12	0.20	0.91	11.90	20.00	90.67
[70 – 80)	6	21	84	0.14	0.28	1.12	14.29	28.00	112.00
[80 – 90)	3	24	84	0.07	0.32	1.12	7.14	32.00	112.00
[90 – 100]	2	26	84	0.05	0.35	1.12	4.76	34.67	112.00
TOTAL	42						61.90		



4

Gráficos estadísticos

4.1. Histograma de frecuencias

4.2. Circulares

4.3. Histograma de frecuencias



5

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.

5.1. La media (\bar{x})

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

5.1.1. 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

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

5.1.2. Media de datos agrupados

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

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

Tabla 5.1: Captionww

Clase	Y_i	f_i	F_i	...	H_i^* %
$[y_1, y_2)$	y_1	f_1	H_1^* %
$[y_2, y_3)$	y_2	f_2	H_1^* %
$[y_3, y_4)$	y_3	f_3	H_1^* %
\vdots	\vdots	\vdots	\vdots

Clase	Y_i	f_i	F_i	...	H_i^* %
$[y_{r-1}, y_r]$	y_r	f_r	H_1^* %

Ejercicio 5.1. 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

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

5.2. La moda (Mo)

El dato que mas se repite

5.2.1. 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 es $Mo = x_2$

5.2.2. Moda de datos tabulados

La moda es el dato que más se repite por ejemplo sea el conjunto de datos tabulados entonces la moda es

$$Mo = Li + \frac{Li - Ls}{Li + Ls}r$$

Tabla 5.2: wwwww

Clase	Y_i	f_i	F_i	...	$H_i^* \%$
$[y_1, y_2)$	y_1	f_1	$H_1^* \%$
$[y_2, y_3)$	y_2	f_2	$H_1^* \%$
$[y_3, y_4)$	y_3	f_3	$H_1^* \%$
\vdots	\vdots	\vdots	\vdots
$[y_{r-1}, y_r]$	y_r	f_r	$H_1^* \%$

5.3. La mediana (Me)

5.3.1. Mediana de datos no tabulados

Obtener la mediana consiste en ordenar los datos de menor a mayor y considerar dos casos: El primero si el número de datos es impar entonces el dato $x_{\frac{n+1}{2}}$ del conjunto ordenado será la mediana es decir

$$\text{Me} = x_{\frac{n+1}{2}}$$

de otro lado si el número de datos es par entonces la mediana es la semisuma de los dos datos intermedios es decir

$$\text{Me} = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$$

Ejercicio 5.2. Sean los conjuntos de datos 5, 6, 8, 2, 1, 5, 6, 7, 10, 0, 14 y 20, 25, 6, 5, 19, 5 obtener la mediana de estos conjuntos de datos.

Solución. Al ordenarlos se obtiene el siguiente arreglo 0, 1, 2, 5, 5, 6, 6, 7, 8, 10, 14 y considerando que $x_1 = 0, x_2 = 1, \dots, x_{11} = 14$ en este caso el número de datos es impar entonces el dato $x_{\frac{11+1}{2}} = x_6 = 6$ es la mediana. De otro lado el segundo conjunto de datos al ser ordenados 5, 5, 6, 19, 20, 25 además considerando que $x_1 = 5, x_2 = 5, \dots, x_6 = 25$ conducen a obtener la mediana $\text{Me} = \frac{x_{\frac{6}{2}} + x_{\frac{6}{2}+1}}{2} = \frac{6+19}{2} = 12,5$.

5.3.2. Mediana de datos tabulados

Tabla 5.3: Mediana

Clase	Y_i	f_i	F_i	...	H_i^* %
$[y_1, y_2)$	y_1	f_1	H_1^* %
$[y_2, y_3)$	y_2	f_2	H_1^* %
$[y_3, y_4)$	y_3	f_3	H_1^* %
\vdots	\vdots	\vdots	\vdots
$[y_{r-1}, y_r]$	y_r	f_r	H_1^* %

Los pasos son:

1. Se halla $\frac{n}{2}$ luego
2. x_n

6

Medidas de dispersión

6.1. Varianza



7

Medidas de asimetría



Parte II

Probabilidades



1

Experimento aleatorio

Definición 1.1 (Experimento aleatorio). En experimento aleatorio es un fenómeno que genera un evento



2

Álgebra de eventos

Sean A, B y C eventos entonces 1. e 2. [www](#)

$$\int_1^2 = \sum_2^2 x_1$$



3

Técnicas de conteo

$$P_n^m \ C_n^m$$

$$\binom{m}{n} = \frac{m}{n!(n-m)}$$



4

Definición de probabilidad



5

Probabilidad condicional

$$P(A|B) = \frac{P(B \cap A)}{P(B)}$$



6

Teorema de Bayes

Teorema 6.1 (Teorema de Bayes). Sea $\{A_1, A_2, \dots, A_i, \dots, A_n\}$ un conjunto de sucesos mutuamente excluyentes y exhaustivos, y tales que la probabilidad de cada uno de ellos es distinta de cero (0). Sea B un suceso cualquiera del que se conocen las probabilidades condicionales $P(B|A_i)$. Entonces, la probabilidad $P(A_i|B)$ viene dada por la expresión:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)}$$

donde:

1. $P(A_i)$ son las probabilidades a priori,
2. $P(B|A_i)$ es la probabilidad de B en la hipótesis A_i ,
3. $P(A_i|B)$ son las probabilidades a posteriori.



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

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

$R_X = \{x \in \mathbb{R} / \exists \omega \in \Omega : X(\omega) = x\}$ 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.

Ejemplo 1.1. 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:

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

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

2. $X(\omega) = n_c - n_s$ 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

$$\begin{aligned} P(3) &= P[ccc] = \frac{1}{8} \\ P(2) &= P[ccs] = P[csc] = P[scs] = \frac{3}{8} \\ P(1) &= P[css] = P[scs] = P[ssc] = \frac{3}{8} \\ P(0) &= P[sss] = \frac{1}{8} \end{aligned}$$

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

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

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

Clases de variables aleatorias ### Variable aleatoria discreta Cuando el rango de la variable aleatoria X , R_X es finito o infinito contable (no necesariamente enteros) $R_X = \{x_1, x_2, \dots, x_n, \dots\}$

1.0.1. Variable aleatoria continua

R_X abarca cualquier intervalo en la recta numerica

1.0.2. Variable aleatoria mixta

Discreta y continua

1.1. Función de probabilidad de una variable aleatoria

1.1.1. Función de probabilidad de una variable aleatorias discreta

Definición 1.3 (Función o ley de probabilidad). Sea X una variable aleatoria con rango R_X . Una función definida por

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

1. $p(x) > 0, x \in R_X$
2. $\sum_{x \in R_X} p(x) = P[X \in R_X] = 1$

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

Ejemplo 1.2. La variable aleatoria discreta

$$p_X(x) = p(1-p)^{x-1}, x \in \mathbb{Z}^+ \text{ y } p \in [0, 1]$$

en efecto

$$p(1-p)^{i-1} > 0, \forall x \in \mathbb{Z}^+$$

además

$$\sum_{i=1}^{\infty} p(1-p)^{i-1} = p \lim_{n \rightarrow \infty} \frac{1 - (1-p)^{n+1}}{1 - (1-p)} = 1$$

Ejemplo 1.3. La variable aleatoria discreta

$$p_X(x) = p(1-p), x \in \mathbb{Z}^+ \text{ y } p \in [0, 1]$$

en efecto

$$p(1-p)^{i-1} > 0, \forall x \in \mathbb{Z}^+$$

además

$$\sum_{i=1}^{\infty} p(1-p)^{i-1} = p \lim_{n \rightarrow \infty} \frac{1 - (1-p)^{n+1}}{1 - (1-p)} = 1$$

Ejemplo 1.4. La variable aleatoria discreta

$$p_X(x) = (1-p)^{x-1}, x \in \mathbb{Z}^+ \text{ y } p \in [0, 1]$$

en efecto

$$p(1-p)^{i-1} > 0, \forall x \in \mathbb{Z}^+$$

además

$$\sum_{i=1}^{\infty} p(1-p)^{i-1} = p \lim_{n \rightarrow \infty} \frac{1 - (1-p)^{n+1}}{1 - (1-p)} = 1$$

1.1.2. 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

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

Ejemplo 1.5. For a circle with the radius r , its area is πr^2 . Sea la función

$$\frac{1}{x^2} = 1$$

$$\frac{\sin x}{x^3} = 0,3794281$$

$$\Phi_{\mu, \sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} du$$

Ejemplo 1.6. Sea $f(x) = \frac{\alpha}{\rho}$ es una función de densidad pues

$$f(x) > 0, x \in R_X$$

además

$$\int_{R_X} f(x)dx = 1$$

Ejemplo 1.7. Sea $f(x) = \frac{\sigma}{\rho}$ es una función de densidad pues

$$f(x) > 0, x \in R_X$$

además

$$\int_{R_X} f(x)dx = 1$$

1.2. Función de distribución de una variable aleatoria

1.2.1. 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 . Cuyas propiedades son:

1. $0 \leq F_X(x) \leq 1$
2. $F_X(-\infty) = 0$
3. $F_X(\infty) = 1$
4. $P(X < x) = F_X(x^-)$
5. $P(a \leq X \leq b) = F_X(b) - F_X(a^-)$

Ejemplo 1.8. wwwwww

Ejemplo 1.9. wwwwww

Ejemplo 1.10. wwwwww

1.2.2. Función de distribución de una variable aleatoria continua

Definición 1.6 (Función de distribución). Sea X una variable aleatoria con función de densidad $f(x)$. La función

$$F_X(x) = F(x) = P[X \leq x] = \int_{-\infty}^x f(t)dt, \quad \forall x \in R_X$$

Cuyas propiedades son:

1. $0 \leq F(x) \leq 1$
2. $F(-\infty) = 0$
3. $F(\infty) = 1$

Ejemplo 1.11. wwwwww

Ejemplo 1.12. wwwwww

Ejemplo 1.13. wwwwww



2

Parámetros de una variable aleatoria

2.1. Esperanza matemática

Definición 2.1 (Esperanza matemática de una variable aleatoria discreta).

$$\mathbb{E}[X] = \sum_{i=1}^n x_i p(x_i)$$

Definición 2.2 (Esperanza matemática de una variable aleatoria continua).

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x) dx \text{ equivalentemente } \mathbb{E}[X] = \int_{\Omega} X dP$$

el valor esperado a veces se representa por $\mu = \mathbb{E}[X]$ que es el promedio o la media poblacional.

2.2. Medidas de variación

La varianza es una medida de dispersión de una variable aleatoria X respecto a su esperanza $\mathbb{E}[X]$. Se define como la esperanza de la transformación

$$\rho = \text{Var}(X) = (X - \mathbb{E}[X])^2$$

$$\sigma = \sqrt{\text{Var}(X)}$$

o bien

$$\sigma^2 = \text{Var}(X)$$

Definición 2.3 (Varianza de una variable aleatoria discreta). Sea

Definición 2.4 (Varianza matemática de una variable aleatoria continua). Sea

2.3. Medidas de posición

Definición 2.5 (Cuantiles de una variable aleatoria discreta). Sea

Definición 2.6 (Cuantiles matemática de una variable aleatoria continua). Sea

2.4. Medidas de curtosis

Definición 2.7 (Curtosis de una variable aleatoria discreta). Sea

Definición 2.8 (Curtosis de una variable aleatoria continua). Sea

momento de orden superior

$$M_X^{(n)} = \mathbb{E}[X^n] = \int_{\mathbb{R}} x^n f_X(x) \, dx$$

3

Variables aleatorias bidimensionales

Definición 3.1 (Variable aleatoria bidimensional discreta).

$$F(x, y) = P[X \leq x, Y \leq y] = \sum_{u=-\infty}^x \sum_{v=-\infty}^y p(u, v)$$

Ejemplo 3.1.

Definición 3.2 (Variable aleatoria bidimensional continua).

$$F(x, y) = P[X \leq x, Y \leq y] = \sum_{u=-\infty}^x \sum_{v=-\infty}^y p(u, v)$$

Ejemplo 3.2.

Distribución bidimensional discreta

Definición 3.3 (Función de probabilidad conjunta). Sea (X, Y) una variable bidimensional discreta con rango $R_{X \times Y}$. A cada posible resultado le asociamos un numero

$$p(x, y) = P[X = x, Y = y]$$

que cumple la siguientes condiciones

1. $1 > p(x, y) > 0, (x, y) \in R_{X \times Y} \in$
2. $\sum_{x \in R_X} \sum_{y \in R_Y} p(x, y) = 1$

A los pares ordenados $((x, y), p(x, y))$ se le llama distribución de probabilidad conjunta

Definición 3.4 (Función de distribución acumulada).

$$F(x, y) = P[X \leq x, Y \leq y] = \sum_{u=-\infty}^x \sum_{v=-\infty}^y p(u, v)$$

- 3.0.1. Distribuciones marginales
 - 3.0.2. Variables aleatorias independientes
 - 3.0.3. Distribuciones de probabilidad condicional
-

- 3.1. Distribución bidimensional continua

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. $k \sum_{i=1}^n x_i = \sum_{i=1}^n kx_i$
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

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

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

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

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

```
xw = 'Es decir los elementos son demagogos y déspotas'
x1 = 'Es decir los elementos son demagogos y déspotas'
```

$$\frac{\sin x}{x^3} = 0,3794281$$

$$\Phi_{\mu, \sigma^2}(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(u-\mu)^2}{2\sigma^2}} du$$

*

$$\frac{1}{20\sqrt{2\pi}} \int_{-\infty}^{300} e^{-\frac{1}{2} \left(\frac{z-200}{20} \right)^2} dz = 0,9999997$$

* 0.9500042 also Es decir los elementos son demagogos y déspotas * Es decir los elementos son demagogos y déspotas Tabla [B.1](#)

Tabla B.1: Caption

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

variable aleatoria Variable aleatoria entonces
2.7182818 0.9750021 0.7881446
2561
The value of x in the Python session is Es decir los elementos son demagogos y despotas . It is not the same x as the one in R.
(Intercept)
12917.13

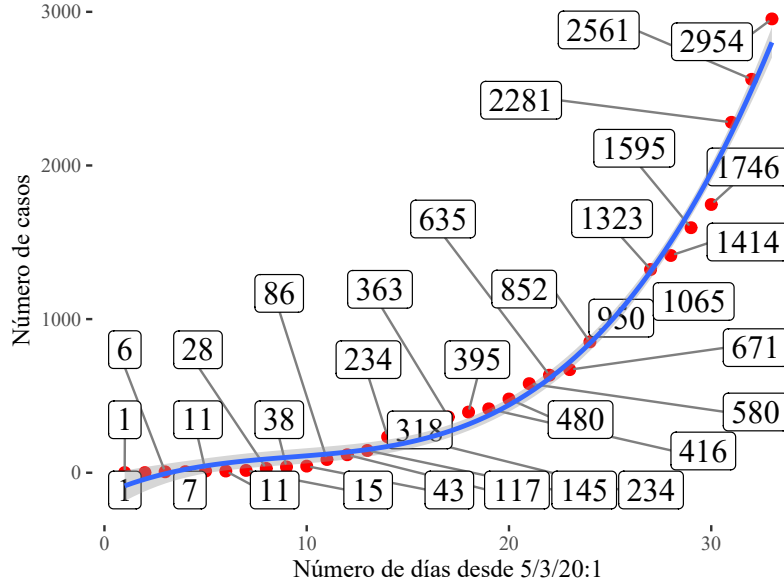


Figura B.1 Regresión lineal

Sea la Tabla B.2 Figures and tables with captions will be placed in figure and table environments, respectively.

Tabla B.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^*	$h_i\%$	$H_i\%$	$H_i^*\%$
ESTJ	1	1	75	0.01	0.01	1.00	1.33	1.33	100.00
ESTJ	2	3	150	0.03	0.04	2.00	2.67	4.00	200.00
ESTP	3	6	150	0.04	0.08	2.00	4.00	8.00	200.00
ESFJ	4	10	150	0.05	0.13	2.00	5.33	13.33	200.00
ESFP	6	16	150	0.08	0.21	2.00	8.00	21.33	200.00
ISTJ	6	22	155	0.08	0.29	2.07	8.00	29.33	206.67
ISTP	7	29	157	0.09	0.39	2.09	9.33	38.67	209.33
ISFJ	8	37	158	0.11	0.49	2.11	10.67	49.33	210.67
ISFP	9	46	166	0.12	0.61	2.21	12.00	61.33	221.33
ENTJ	10	56	166	0.13	0.75	2.21	13.33	74.67	221.33
ENTP	6	62	166	0.08	0.83	2.21	8.00	82.67	221.33
ENFJ	5	67	166	0.07	0.89	2.21	6.67	89.33	221.33
ENFP	3	70	166	0.04	0.93	2.21	4.00	93.33	221.33
INTJ	2	72	167	0.03	0.96	2.23	2.67	96.00	222.67
INTP	1	73	169	0.01	0.97	2.25	1.33	97.33	225.33
INFJ	1	74	172	0.01	0.99	2.29	1.33	98.67	229.33

INFP	1	75	176	0.01	1.00	2.35	1.33	100.00	234.67
------	---	----	-----	------	------	------	------	--------	--------



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, 10
frecuencias absolutas acumuladas
 menor que, 10
frecuencias absolutas relativas, 10
frecuencias absolutas relativas menor
 que, 10
traza, 69