Ricardo Michel MALLQUI BAÑOS

Elementos de la estadística

estadística descriptiva y probabilidades


PRIMVN VIVERE DE INDE PHILOSOPHARI
universitas + guamangensis + sancti + christophori
1677

Índice general

Índice de tablas ix

Índice de figuras xi

Resumen xiii

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
2.2. 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 16

4. Gráficos estadísticos 17
4.1. Histograma de frecuencias 17
4.2. Circulares 17
4.3. Histograma de frecuencias 20

5. Medidas de tendencia central 21
5.1. La media 21
5.1.1. Media de datos no agrupados 21
5.1.2. Media de datos agrupados 21
5.1.3. Ejemplo 22

5.2. La moda (Mo) 23
5.2.1. Moda de datos no tabulados 23
5.2.2. Moda de datos tabulados 23
5.2.3. Ejemplo 24
5.3. La mediana (Me) 25
5.3.1. Mediana de datos no tabulados 25
5.3.2. Mediana de datos tabulados 25
5.3.3. Ejemplo 26
5.4. Asignación 27

6. Medidas de dispersión 29
6.1. Rango 29
6.2. Varianza 29
6.2.1. Datos no tabulados 29
6.2.2. Datos tabulados 30
6.3. Desviación típica 31
6.4. Desviación media absoluta 32
6.4.1. Datos no tabulados 32
6.4.2. Datos tabulados 32
6.5. Desviación mediana absoluta 33
6.5.1. Datos no tabulados 33
6.5.2. Datos tabulados 33
6.6. Coeficiente de variacion 34
6.7. Asignación 35

7. Medidas de posicion (cuantiles) 37
7.1. Cuartiles 37
7.1.1. Datos no agrupados 37
7.1.2. Datos agrupados 38
7.2. Quintiles 39
7.3. Deciles 39
7.3.1. Datos no agrupados 39
7.3.2. Datos agrupados 40
7.4. Percentiles 40
7.4.1. Datos no agrupados 41
7.4.2. Datos agrupados 41

8. Medidas de asimetría 43
8.1. Normal: Simétrica 44
8.2. Normal: Simetría positiva 45

8.3. Normal: Simetría negativa 46

9. Medidas de curtosis o apuntamiento **49**

II Probabilidades **51**

1. Experimento aleatorio **53**

2. Álgebra de eventos **55**

3. Técnicas de conteo **57**

4. Definición de probabilidad **59**

5. Probabilidad condicional **61**

6. Teorema de Bayes **63**

7. Eventos independientes y secuencias de experimentos **65**

8. Probabilidad en espacio **67**

III Inferencia estadística **69**

1. Variables aleatorias **71**

1.1. Clases de variables aleatorias 73

1.1.1. Variable aleatoria discreta 73

1.1.2. Variable aleatoria continua 73

1.1.3. Variable aleatoria mixta 73

1.2. Función de probabilidad de una variable aleatoria 73

1.2.1. Función de probabilidad de una variable aleatorias discreta 73

1.2.2. Función de probabilidad de una variable aleatoria continua 74

1.3. Función de distribución de una variable aleatoria 75

1.3.1. Función de distribución de una variable aleatoria discreta 75

1.3.2. Función de distribución de una variable aleatoria continua 76

2. Parámetros de una variable aleatoria **77**

2.1. Esperanza matemática 77

2.2. Medidas de variación 77

3. Variables aleatorias bidimensionales **79**
3.1. Distribución bidimensional discreta 79
3.1.1. Distribuciones marginales 80
3.1.2. Variables aleatorias independientes 80
3.1.3. Distribuciones de probabilidad condicional 80
3.2. Distribución bidimensional continua 80

4. Distribuciones discreta importantes **81**
4.1. Variable aleatoria discreta binomial 81
4.2. Variable aleatoria discreta Poisson 81

5. Distribuciones continuas importantes **83**
5.1. Variable aleatoria continua normal 83
5.2. Variable aleatoria continua gamma 83

6. Distribuciones muestrales **85**

7. Estimación **87**

8. Prueba de hipótesis **89**

Apéndice **89**

A. Sumatorias **91**
A.1. ee 91
A.1.1. eeeee 91

B. Matrices **93**
B.1. Algebra de matrices 93
B.2. Matrices particulares 95
B.2.1. Matriz triangular 95
B.2.2. Matriz particular de una matriz cuadrada 96
B.2.3. Matriz transpuesta 96
B.2.4. Matriz simetrica 96
B.2.5. Matriz conjugada 96
B.2.6. Matriz hermitica 96
B.2.7. Matriz escalonada 96

Bibliografía **99**

Índice alfabético **101**

Índice de tablas

3.1. Caption 9
3.2. Datos cuantitativos (intervalos) 10
3.4. Datos cualitativos 15
3.6. Datos cuantitativos (intervalos) 16

5.1. Captionww 22
5.3. wwwww 24
5.5. Mediana 26

B.1. Caption 97

Índice de figuras

8.1. Medidas de asimetría 43

9.1. wwwwwwwwwwwwwwwww 49

B.1. Regresión lineal 98

Resumen

La estadística es la ciencia que manipula datos las analiza e interpreta para poder sacar concluciones 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

Matemáticas básicas

2

Variables

Definición 2.1 (Variable)**.** Es una **característica** de personas cosas u objetos que son propensos a ser medidas o cualificadas

Definición 2.2. 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

Definición 2.3. 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).

2.0.2. Ordinales

Definición 2.4. Son características que que si están propensos a ser jerarquizados tales como: Nivel de instrucción (inicial, primaria, secundaria, superior).

2.1. Variables cuantitativas

Definición 2.5. Son aquellas variables que están propensos a ser medidas mediante números ya sean números enteros o reales.

2.1.1. Discretas

Definición 2.6. Aquellas que solo son medidos mediante números enteros por ejemplo: Número de hijos, número de habitaciones.

2.1.2. Continuas

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

2.2. Asignación

1. Reconozca **5** variables **cualitativas** de una persona, admosfera, una pintura

    ▪ Persona

        •Color (moreno, blanco, trigueo)
        •Religión (catolico, evangelico, pentecostal, etc)
        •wwwwwww
        •wwww
        •wwwwww

    ▪ Admosfera

        •www
        •wwwww
        •wwwwwww
        •wwww
        •wwwwww

    ▪ Pintura

        •www
        •wwwww
        •wwwwwww
        •wwww
        •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)
        •wwwwwww
        •wwww

- wwwwww

- Tela
  - www
  - wwwww
  - wwwwwww
  - wwww
  - wwwwww

- Celular
  - www
  - wwwww
  - wwwwwww
  - wwww
  - wwwwww

3

Organización de datos en tablas de frecuencias

3.1. Distribución de frecuencias

Definición 3.1. 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_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 ademas 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 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 interválica 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_i	F_i	F_i^*	$\frac{f_1}{n}$	$\frac{f_1}{n}$	$\frac{F_1^*}{n}$	$h_1\,\%$	$H_1\,\%$	$H_1^*\,\%$
$[y_2 - y_3)$	Y_2	f_i	F_i	F_i^*	$\frac{f_2}{n}$	$\frac{f_1}{n}$	$\frac{F_1^*}{n}$	$h_2\,\%$	$H_2\,\%$	$H_2^*\,\%$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
$[y_{r-1} - y_r)$	Y_r	f_r	F_r	F_r^*	$\frac{f_r}{n}$	$\frac{f_1}{n}$	$\frac{F_r^*}{n}$	$h_r\,\%$	$H_r\,\%$	$H_r^*\,\%$

Tenga en cuenta que n es el número de datos, es decir $n = f_1 + f_2 + \ldots + f_r = \sum_{i=1}^{r}$ 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
$$F_m = f_1 + f_2 + \ldots + f_m = \sum_{i=1}^{m} f_i$$
3. Las **frecuencias absolutas acumuladas mayor que** F_i^* obedecen a la fórmula
$$\begin{aligned} F_m^* &= f_m + f_{m+1} + \ldots + f_r \\ &= \sum_{i=m}^{r} f_i \\ &= n - \sum_{i=1}^{m-1} f_i \\ &= n - (f_1 + f_2 + \ldots + f_{m-1}) \end{aligned}$$
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 $H_i^*\,\% = 100 \cdot H_i^*$

10. Y_i marca de clase o punto medio de la clase 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 necesidad 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 entreteniendo 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 comunica-

tivas, 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.

Se tiene 16 características con los siguientes datos

Personalidad	Cantidad
ESTJ	1
ESTJ	2
ESTP	3
ESFJ	4
ESFP	5

Personalidad	Cantidad
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

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

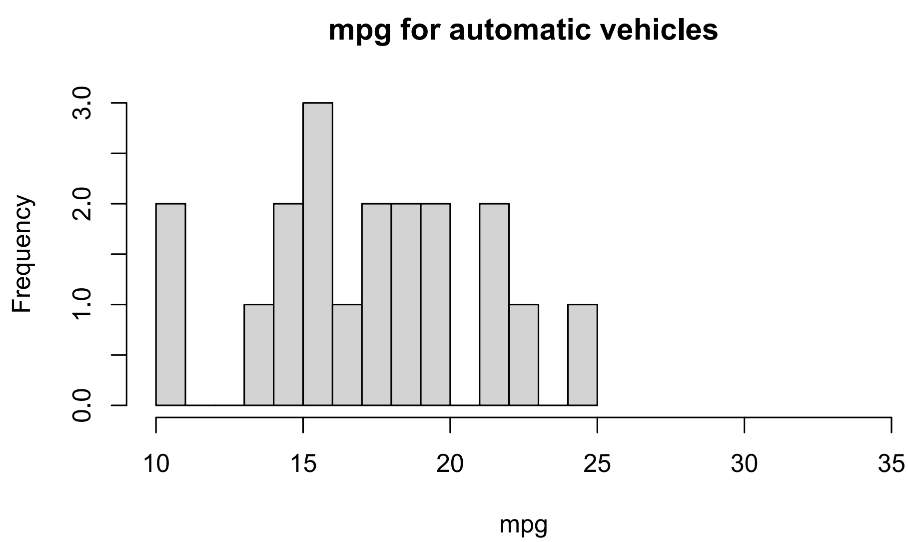
Gráficos estadísticos

4.1. Histograma de frecuencias

```
mtcars$am <- as.factor(mtcars$am)
levels(mtcars$am) <-c("Automatic", "Manual")
```

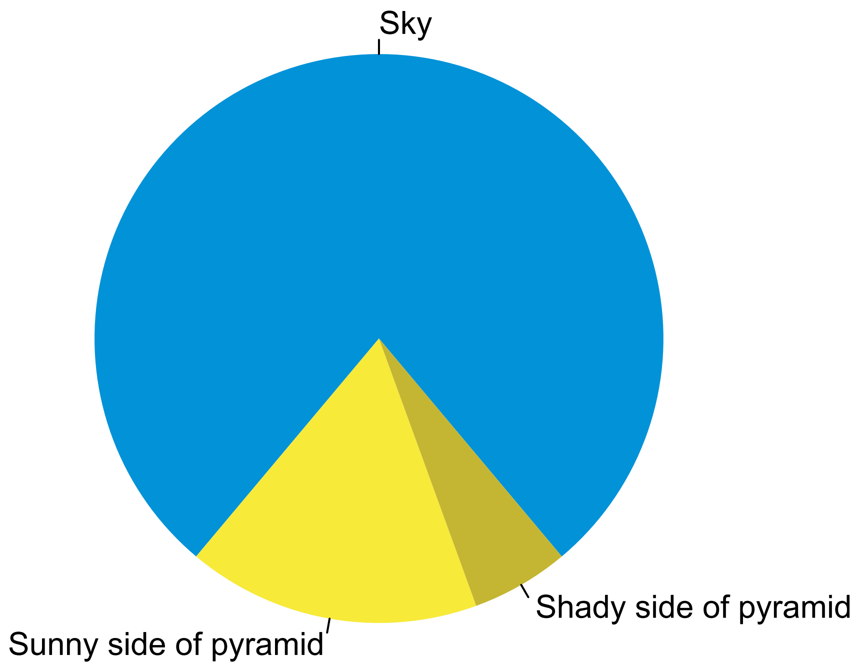
Some text

```
hist(mtcars$mpg[mtcars$am=="Automatic"], breaks=12, main="mpg for automa
```

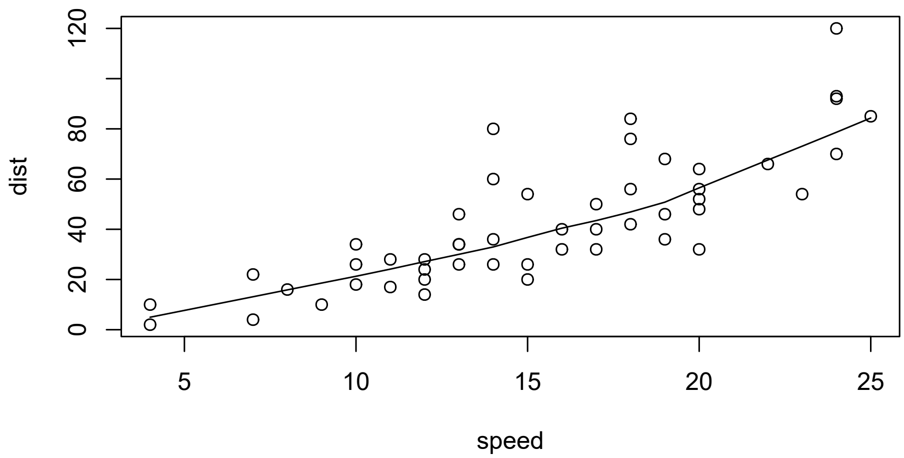


4.2. Circulares

```
par(mar = c(0, 1, 0, 1))
pie(
  c(280, 60, 20),
  c('Sky', 'Sunny side of pyramid', 'Shady side of pyramid'),
  col = c('#0292D8', '#F7EA39', '#C4B632'),
  init.angle = -50, border = NA
)
```




```
plot(cars)
lines(lowess(cars))
```

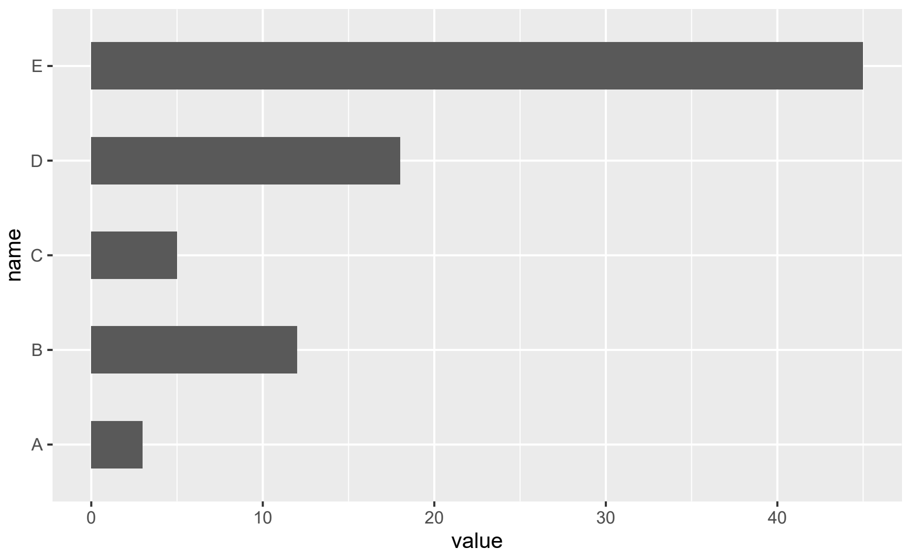



```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.1.3
```

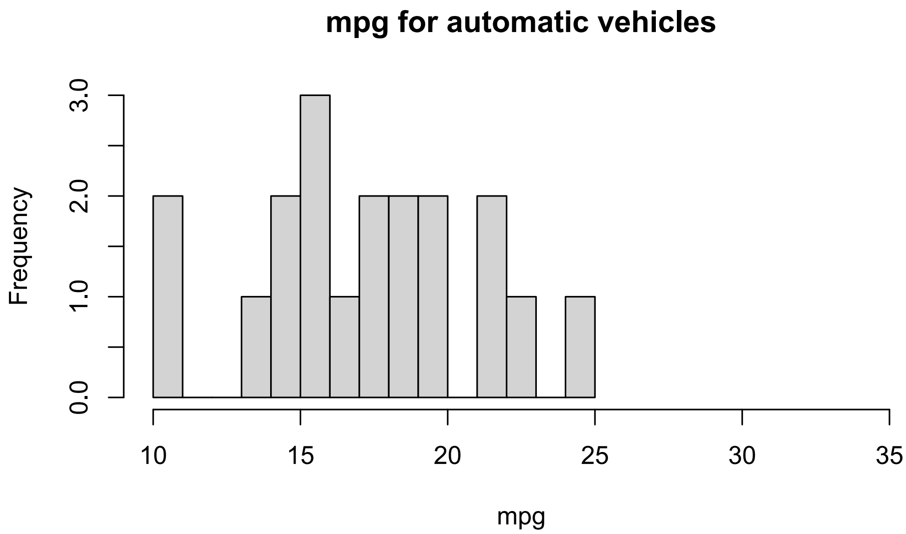
```
# Create data
data <- data.frame(
  name=c("A","B","C","D","E") ,
  value=c(3,12,5,18,45)
  )

# Barplot
ggplot(data, aes(x=name, y=value)) +
  geom_bar(stat = "identity", width=0.5) +
  coord_flip()
```



4.3. Histograma de frecuencias

```
hist(mtcars$mpg[mtcars$am=="Automatic"], breaks=12, main="mpg for automa
```



5

Medidas de tendencia central

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

5.1. La media

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

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

5.1.2. Media de datos agrupados

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

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

Tabla 5.1: Captionww

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

5.1.3. Ejemplo

Clase	Y_i	f_i	$Y_i * f_i$
$[10, 15)$	12.5	1	12.5
$[15, 20)$	17.5	2	35
$[20, 25)$	22.5	5	112.5
$[25, 30)$	27.5	3	82.5
$[30, 35]$	32.5	2	65
$\sum$		13	307.5

$$\overline{x} = \frac{12{,}5 + 35 + +112{,}5 + 82{,}5 + 65}{13} = \frac{307{,}5}{13} = 23{,}65$$

Ejercicio 5.1. Si el promedio de n datos es $\overline{x}$ entonces el promedio del conjunto inicial más un dato adicional x_{n+1} es

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

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

$$\overline{x}' = \frac{n\overline{x} + y_1 + y_2 + \ldots + y_r}{n+r}$$

Solución. En efecto sea el promedio

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

5.2. La moda (Mo)

- La moda es el valor que tiene mayor frecuencia absoluta.
- Se representa por Mo
- Si en un grupo hay dos o varias puntuaciones con la misma frecuencia y esa frecuencia es la máxima, entonces la distribución es bimodal es decir, tiene varias modas.
- Cuando todas las puntuaciones de un grupo tienen la misma frecuencia, no hay moda.
- Se puede hallar la moda para variables cualitativas y cuantitativas.
- Cuando todas las puntuaciones de un grupo tienen la misma frecuencia, no hay moda.
- Si dos puntuaciones adyacentes tienen la frecuencia máxima, la moda es el promedio de las dos puntuaciones adyacentes.
- Si dos puntuaciones adyacentes tienen la frecuencia máxima, la moda es el promedio de las dos puntuaciones adyacentes.Ejemplos de ejercicios de moda

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 $= x_2$

Halle la moda de los siguinetes datos 3, 5,3,6,7,3,4,5,5 ya que hay hay precencia de dotas que se repiten dos veces en tonces este conjunto de datos recibe el nombre de datos bimodal Mo=3 y Mo=5

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

$$M_o = L_i + \frac{f_i - f_{i-1}}{(f_i - f_{i-1}) + (f_i - f_{i+1})} \cdot a_i$$

- L_i es el límite inferior de la clase modal
- f_i es la frecuencia absoluta de la clase modal
- f_{i-1} es la frecuencia absoluta inmediatamente inferior a la clase modal
- f_{i+1} es la frecuencia absoluta inmediatamente posterior a la clase modal
- a_i es la amplitud de la clase

Tabla 5.3: wwwww

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

5.2.3. Ejemplo

Clase	f_i
$[10, 15)$	2
$[15, 20)$	5
$[20, 25)$	10
$[25, 30)$	3
$[30, 35]$	1

Primeramente la mayor frecuencia absoluta es 10 y corresponde $f_3 = 10$ por tanto $i = 3$. $L_i = 20$

$$M_o = L_i + \frac{f_i - f_{i-1}}{(f_i - f_{i-1}) + (f_i - f_{i+1})} \cdot a_i =$$

$$= 20 + \frac{10 - 5}{(10 - 5) + (10 - 3)} \cdot 5 = 20 + \frac{5}{12} * 5 = 22{,}08$$

Más información

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

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

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

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

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 $x_1 = 0$, $x_2 = 1$, …, $x_{11} = 14$ en este caso el número de datos es impar entonces el dato $x_{\frac{11+1}{2}} = x_6 = 6$ el la mediana. De otro lado el segundo conjunto de datos al ser ordenados 5, 5, 6, 19, 20, 25 ademas considerando que $x_1 = 5$, $x_2 = 5$, …, $x_6 = 25$ conducen a obtener la mediana $\mathrm{Me} = \frac{x_{\frac{6}{2}}+x_{\frac{6}{2}+1}}{2} = \frac{6+19}{2} = 12{,}5$.

5.3.2. Mediana de datos tabulados

La mediana se encuentra en el intervalo donde la frecuencia acumulada llega hasta la mitad de la suma de las frecuencias absolutas.

Es decir tenemos que buscar el intervalo en el que se encuentre.

$$M_e = L_i + \frac{\frac{N}{2} - F_{i-1}}{f_i} \cdot a_i$$

L_i es el límite inferior de la clase donde se encuentra la mediana

$\frac{N}{2}$ es la semisuma de las frecuencias absolutas

f_i es la frecuencia absoluta de la clase mediana

F_{i-1} es la frecuencia acumulada anterior a la clase mediana

a_i es la amplitud de la clase

La mediana es independiente de las amplitudes de los intervalos

Tabla 5.5: Mediana

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

Más información

5.3.3. Ejemplo

Clase	f_i	F_i
$[10, 15)$	1	1
$[15, 20)$	2	3
$[20, 25)$	5	8
$[25, 30)$	3	11
$[30, 35]$	1	12
$\sum$	12	

$$\frac{N}{2} = 12/2 = 6$$

ubicando en las frecuencias absolutas acumuladas que corresponde al intervalo $[20, 25)$

$$\begin{aligned} M_e &= L_i + \frac{\frac{N}{2} - F_{i-1}}{f_i} \cdot a_i \\ &= 20 + \frac{\frac{12}{2} - 3}{5} \cdot 5 = 23 \end{aligned}$$

por lo tanto la mediana de este conjunto de datos tabulados (agrupados) es $Me = 23$

5.4. Asignación

Halle le media, la moda y la mediana de los siguientes datos tabulados

Clase	Y_i	f_i	F_i
$[100, 150)$		1	1
$[150, 200)$		2	3
$[200, 250)$		5	
$[250, 300)$		7	
$[300, 350]$		10	
$[350, 400]$		6	
$[400, 450]$		5	
$[450, 500]$		2	
$[500, 550]$		1	

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

$$M_o = L_i + \frac{f_i - f_{i-1}}{(f_i - f_{i-1}) + (f_i - f_{i+1})} \cdot a_i$$

$$M_e = L_i + \frac{\frac{N}{2} - F_{i-1}}{f_i} \cdot a_i$$

6

Medidas de dispersión

6.1. Rango

$R = x_{max} - x_{min}$

Sean los datos 2, 5, 6, 1, 7, 5, 8, 6

$x_{max} = 8 \; x_{min} = 1$

por lo tanto $R = x_{max} - x_{min} = 8 - 1 = 7$

6.2. Varianza

6.2.1. Datos no tabulados

$$s^2 = \frac{\sum (x_i - \overline{x})^2}{n-1}$$

Sean los datos 2, 5, 6, 1, 7, 5, 8, 6

$$\overline{x} = 40/8 = 5$$

$$
\begin{aligned}
s^2 &= \frac{\sum (x_i - \overline{x})^2}{n-1} \\
&= \frac{(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + (x_3 - \overline{x})^2 + (x_4 - \overline{x})^2 + (x_5 - \overline{x})^2 + (x_6 - \overline{x})^2 + (x_7 - \overline{x})^2 + (x_8 - \overline{x})^2}{8-1} \\
&= \frac{(2-5)^2 + (5-5)^2 + (6-5)^2 + (1-5)^2 + (7-5)^2 + (5-5)^2 + (8-5)^2 + (6-5)^2}{8-1} \\
&= \frac{9+0+1+16+4+0+9+1}{7} \\
&= \frac{40}{7} = 5{,}71
\end{aligned}
$$

6.2.2. Datos tabulados

$$s^2 = \frac{\sum f_i \left(Y_i - \overline{x}\right)^2}{n-1}$$

$\overline{x} = \frac{\sum Y_i * f_i}{n} = 747{,}5/31 = 24{,}11$

Clase	Y_i	f_i	$Y_i * f_i$
$[5, 10)$	7.5	1	7.5
$[10, 15)$	12.5	2	25
$[15, 20)$	17.5	5	87.5
$[20, 25)$	22.5	7	157.5
$[25, 30]$	27.5	10	275
$[30, 35]$	32.5	6	195
$\sum$		31	

Por lo tanto la varianza para dotos agrupados es

$$
\begin{aligned}
s^2 &= \frac{\sum f_i\,(Y_i - \overline{x})^2}{n-1} \\
&= \frac{f_1\,(Y_1 - \overline{x})^2 + f_2\,(Y_2 - \overline{x})^2 + f_3\,(Y_3 - \overline{x})^2 + f_4\,(Y_4 - \overline{x})^2 + f_5\,(Y_5 - \overline{x})^2 + f_6\,(Y_6 - \overline{x})^2}{31-1} \\
&= \frac{1\,(7{,}5 - 24{,}11)^2 + 2\,(12{,}5 - 24{,}11)^2 + 5\,(17{,}5 - 24{,}11)^2 + 7\,(22{,}5 - 24{,}11)^2 + 10\,(27{,}5 - 24{,}11)^2 +}{31-1} \\
&= \frac{1 * 275{,}89 + 2 * 134{,}79 + 5 * 43{,}69 + 7 * 2{,}59 + 10 * 11{,}49 + 6 * 7{,}39}{31-1} \\
&= \frac{275{,}89 + 269{,}58 + 218{,}45 + 18{,}13 + 114{,}9 + 44{,}34}{30} \\
&= \frac{941{,}29}{30} = 31{,}38
\end{aligned}
$$

Por lo tanto

$$s^2 = 31{,}38$$

6.3. Desviación típica

$$s = \sqrt{s^2}$$

La desviacion típica o estandar del siguiente conjunto de datos tabulados

Clase	Y_i	f_i	$Y_i * f_i$
$[5, 10)$	7.5	1	7.5
$[10, 15)$	12.5	2	25
$[15, 20)$	17.5	5	87.5
$[20, 25)$	22.5	7	157.5
$[25, 30]$	27.5	10	275
$[30, 35]$	32.5	6	195
$\sum$		31	

es

$$s = \sqrt{s^2} = \sqrt{31{,}38} = 5{,}60$$

6.4. Desviación media absoluta

6.4.1. Datos no tabulados

$$DM = \frac{1}{n}\sum |x_i - \overline{x}|$$

Sean los datos 2, 5, 6, 1, 7, 5, 8, 6

$$\overline{x} = 40/8 = 5$$

Entonces

$$\begin{aligned} DM &= \frac{\sum |x_i - \overline{x}|}{n} \\ &= ...Resolver \end{aligned}$$

6.4.2. Datos tabulados

$$DM = \frac{1}{n}\sum f_i\,|Y_i - \overline{x}|$$

y_i marca de clase o punto medio de la clase i

$\overline{x} = \frac{\sum Y_i * f_i}{n} = 747{,}5/31 = 24{,}11$

Clase	Y_i	f_i	$Y_i * f_i$
$[5, 10)$	7.5	1	7.5
$[10, 15)$	12.5	2	25
$[15, 20)$	17.5	5	87.5
$[20, 25)$	22.5	7	157.5
$[25, 30]$	27.5	10	275
$[30, 35]$	32.5	6	195
$\sum$		31	

Por lo tanto la desviación media absoluta es

$$
\begin{aligned}
DM &= \frac{\sum f_i \left|Y_i - \overline{x}\right|}{n} \\
&= \frac{f_1 \left|Y_1 - \overline{x}\right| + f_2 \left|Y_2 - \overline{x}\right| + f_3 \left|Y_3 - \overline{x}\right| + f_4 \left|Y_4 - \overline{x}\right| + f_5 \left|Y_5 - \overline{x}\right| + f_6 \left|Y_6 - \overline{x}\right|}{31} \\
&= \frac{1\left|7{,}5 - 24{,}11\right| + 2\left|12{,}5 - 24{,}11\right| + 5\left|17{,}5 - 24{,}11\right| + 7\left|22{,}5 - 24{,}11\right| + 10\left|27{,}5 - 24{,}11\right| + 6\,|3}{31} \\
&= \frac{1 * 16{,}61 + 2 * 11{,}61 + 5 * 6{,}61 + 7 * 1{,}61 + 10 * 3{,}39 + 6 * 8{,}39}{31} \\
&= \frac{277{,}33}{31} = 8{,}94
\end{aligned}
$$

por lo tanto

$$DM = 8{,}94$$

6.5. Desviación mediana absoluta

6.5.1. Datos no tabulados

$$DMe = \frac{1}{n}\sum \left|Y_i - Me\right|$$

Sean los datos 2, 5, 6, 1, 7, 5, 8, 6 (Ejercicio)

6.5.2. Datos tabulados

$$DMe = \frac{1}{n}\sum f_i \left|Y_i - Me\right|$$

$Me =$? (Ejercicio)

Clase	Y_i	f_i	$Y_i * f_i$
$[5, 10)$	7.5	2	7.5
$[10, 15)$	12.5	3	25
$[15, 20)$		4	87.5

Clase	Y_i	f_i	$Y_i * f_i$
$[20, 25)$	22.5	7	157.5
$[25, 30]$	27.5	10	275
$[30, 35]$		8	195
$\sum$			

Por lo tanto la desviacion de la mediana absoluta es

(Ejercicio)

$$\begin{aligned} s^2 &= \frac{\sum f_i \left|Y_i - Me\right|}{n} \\ &= \frac{f_1 \left(Y_1 - Me\right)^2 + f_2 \left(Y_2 - Me\right)^2 + f_3 \left(Y_3 - Me\right)^2 + f_4 \left(Y_4 - Me\right)^2 + f_5 \left(Y_5 - Me\right)^2 + f_6 \left(Y_6 - M\right.}{31} \\ &= complete \end{aligned}$$

6.6. Coeficiente de variacion

$$Cv = \frac{s}{\bar{x}} \cdot 100$$

Si $Cv > 25\,\%$ se dice que los datos estan muy dispersos Si $Cv < 25\,\%$ se dice que los datos estan muy juntos

Para el conjunto de datos

Clase	Y_i	f_i	$Y_i * f_i$
$[5, 10)$	7.5	1	7.5
$[10, 15)$	12.5	2	25
$[15, 20)$	17.5	5	87.5
$[20, 25)$	22.5	7	157.5
$[25, 30]$	27.5	10	275
$[30, 35]$	32.5	6	195
$\sum$		31	

$$Cv = \frac{5{,}60}{24{,}11} \cdot 100 = 0{,}23 \cdot 100 = 23\,\%$$

6.7. Asignación

Halle el rango, la varianza, la desviación típica, desviación media, desviación mediana absoluta y el coeficiente de variación. Grafique el hstograma y ubique estos estadigrafos

Clase	Y_i	f_i	F_i
$[50, 100)$		8	1
$[100, 150)$		20	3
$[150, 200)$		50	
$[200, 250)$		70	
$[250, 300]$		100	
$[300, 350]$		60	
$\sum$		20	

7

Medidas de posicion (cuantiles)

- Los deciles, que dividen a la distribución en diez partes;
- Los percentiles, que dividen a la distribución en cien partes.

7.1. Cuartiles

Los cuartiles, que dividen a la distribución en cuatro partes (corresponden a los cuantiles 0,25; 0,50 y 0,75); Q_1, Q_2, Q_3

7.1.1. Datos no agrupados

Sean los datos 1, 2, 5, 1 ,5, 6, 7, 8, 9, 3, 4, 5, 2, 6, 2, 5, 6, 7 Ordenar de menor a mayor (creciente)

Si

$$Q_k = \frac{k(n+1)}{4}$$

es entero entonces el cuartil es el dato de la posicion $Q_k = x_{\frac{k(n+1)}{4}}$ en caso contrario se interpola los datos extremos donde se encuentra el valor Q_k

- Ejemplo 1, 2, 5, 1 ,5, 6, 7, 8, 9, 3, 4, 5, 2, 6, 2, 5, 6, 7 ordenados de menor a mayor 1, 1, 2, 2, 2, 3, 4, 5, 5, 5, 5, 6, 6, 6, 7, 7, 8, 9

$Q_1 = \frac{1(18+1)}{4} = 4{,}75$ interpolando $Q_1 = 2 + (2-2)\cdot 0{,}75 = 2$

$Q_2 = \frac{2(18+1)}{4} = 9{,}5$ interpolando $Q_2 = 5 + (5-5)\cdot 0{,}5 = 5$

$Q_3 = \frac{3(18+1)}{4} = 14{,}25$ interpolando $Q_3 = 6 + (7-6)\cdot 0{,}25 = 6{,}25$

7.1.2. Datos agrupados

$$Q_k = L_i + A\left(\frac{\frac{kn}{4} - F_{i-1}}{F_i - F_{i-1}}\right); k = 1, 2, 3$$

- L_i limite inferior del intervalo que contiene al decil
- F_{i-1} frecuencia acumulada en la clase anterior al decil
- F_i frecuencia acumulada en la clase al decil
- A amplitud intervalica
- n numero de datos
- k indice del cuartil correspondiente

$$\begin{aligned} Q_1 &= L_i + A\left(\frac{\frac{1n}{4} - F_{i-1}}{F_i - F_{i-1}}\right) \\ &= L_i + A\left(\frac{\frac{1*39}{4} - F_{i-1}}{F_i - F_{i-1}}\right) \\ &= 20 + 5 \cdot \left(\frac{9{,}75 - 8}{15 - 8}\right) \\ &= 21{,}5 \end{aligned}$$

$$\begin{aligned} Q_2 &= L_i + A\left(\frac{\frac{2n}{4} - F_{i-1}}{F_i - F_{i-1}}\right) \\ &= L_i + A\left(\frac{\frac{2*39}{4} - F_{i-1}}{F_i - F_{i-1}}\right) \\ &= 25 + 5 \cdot \left(\frac{19{,}5 - 15}{25 - 15}\right) \\ &= 27{,}5 \end{aligned}$$

$$\begin{aligned} Q_3 &= L_i + A\left(\frac{\frac{3n}{4} - F_{i-1}}{F_i - F_{i-1}}\right) \\ &= L_i + A\left(\frac{\frac{2*39}{4} - F_{i-1}}{F_i - F_{i-1}}\right) \\ &= 30 + 5 \cdot \left(\frac{29{,}25 - 25}{31 - 25}\right) \\ &= 33{,}542 \end{aligned}$$

Clase	Y_i	f_i	F_i
$[5, 10)$	7.5	1	1
$[10, 15)$	12.5	2	3
$[15, 20)$	17.5	5	8
$[20, 25)$	22.5	7	15
$[25, 30]$	27.5	10	25
$[30, 35]$	32.5	6	31
$[35, 40]$	37.5	5	36
$[40, 45]$	42.5	3	39
$\sum$		39	

7.2. Quintiles

wwwwwwwww

7.3. Deciles

Los deciles, que dividen a la distribución en diez partes es decir $D_1, D_2, \ldots, D_9$

7.3.1. Datos no agrupados

Sean los datos 1, 2, 5, 1 ,5, 6, 7, 8, 9, 3, 4, 5, 2, 6, 2, 5, 6, 7 Ordenar de menor a mayor (creciente)

Si

$$D_k = \frac{k(n+1)}{10}$$

es entero entonces el decil es el dato de la posicion $D_k = x_{\frac{k(n+1)}{10}}; k = 1, 2, 3, \ldots, 9$ Si

$$D_k = \frac{k(n+1)}{10}$$

no es entero entonces el decil es la interpolacion lineal de de los dos valores entre las cuales se encuentra $D_k = \frac{k(n+1)}{10}$

- Ejemplo 1, 2, 5, 1 ,5, 6, 7, 8, 9, 3, 4, 5, 2, 6, 2, 5, 6, 7 ordenados de menor a mayor 1, 1, 2, 2, 2, 3, 4, 5, 5, 5, 5, 6, 6, 6, 7, 7, 8, 9 entonces

$$D_9 = \frac{9(18+1)}{10} = 17{,}1$$

interpolando el decil 9 es $D_9 = 8 + (9-8) \cdot 0{,}1 = 8{,}1$

7.3.2. Datos agrupados

$$D_k = L_i + A\left(\frac{\frac{kn}{10} - F_{i-1}}{F_i - F_{i-1}}\right)$$

Clase	Y_i	f_i	F_i
$[5,10)$	7.5	1	1
$[10,15)$	12.5	2	3
$[15,20)$	17.5	5	8
$[20,25)$	22.5	7	15
$[25,30]$	27.5	10	25
$[30,35]$	32.5	6	31
$[35,40]$	37.5	5	36
$[40,45]$	42.5	3	39
$\sum$		39	

$$D_9 = L_i + A\left(\frac{\frac{9\cdot 39}{10} - F_{i-1}}{F_i - F_{i-1}}\right)$$

Entonces $\frac{9\cdot 39}{10} = 35{,}1$

$$D_9 = 35 + 5\left(\frac{35{,}1 - 31}{36 - 31}\right) = 39{,}1$$

7.4. Percentiles

Los percentiles, que dividen a la distribución en diez partes es decir $P_1, P_2, \ldots, P_{99}$

7.4.1. Datos no agrupados

Si

$$P_k = \frac{k(n+1)}{100}; k = 1, 2, \dots, 99$$

es entero entonces el cuartil es el dato de la posicion $P_k = x_{\frac{k(n+1)}{100}}$ Si

$$P_k = \frac{k(n+1)}{100}$$

no es entero entonces el cuartil es la interpolacion lineal de de los dos valores entre las cuales se encuentra $Q_k = \frac{k(n+1)}{100}$

- Ejemplo 1, 2, 5, 1 ,5, 6, 7, 8, 9, 3, 4, 5, 2, 6, 2, 5, 6, 7 Al ordenar de manera creciente 1, 2, 5, 1 ,5, 6, 7, 8, 9, 3, 4, 5, 2, 6, 2, 5, 6, 7 y

$$P_k = \frac{k(18+1)}{100}$$

7.4.2. Datos agrupados

$$P_k = L_i + A\left(\frac{\frac{kn}{100} - F_{i-1}}{F_i - F_{i-1}}\right) = \int_1^3 f(x)$$

Clase	Y_i	f_i	F_i
$[5, 10)$	7.5	1	1
$[10, 15)$	12.5	2	3
$[15, 20)$	17.5	5	
$[20, 25)$	22.5	7	
$[25, 30]$		10	
$[30, 35]$		6	
$[35, 40]$		5	
$[40, 45]$		3	
$\sum$		2	

8

Medidas de asimetría

Podemos decir que la asimetría indica cuánto se desvía nuestra distribución subyacente de la **distribución normal**, ya que la distribución normal tiene **asimetría 0**. Generalmente, tenemos **tres tipos de asimetría**.

1. **Desviación simétrica**: cuando la asimetría es cercana a 0 y la media es casi la misma que la mediana
2. **Desviación negativa**: cuando la cola izquierda del histograma de la distribución es más larga y la mayoría de las observaciones se concentran en la cola derecha. En este caso, también podemos utilizar el término "sesgado a la izquierda" o "cola izquierda". y la **mediana es mayor que la media**.
3. **Desviación positiva**: cuando la cola derecha del histograma de la distribución es más larga y la mayoría de las observaciones se concentran en la cola izquierda. En este caso, también podemos usar el término "sesgado a la derecha" o "cola derecha". y la **mediana es menor que la media**.

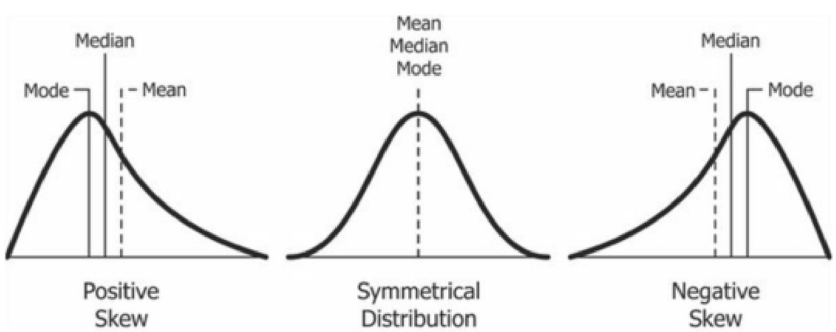


Figura 8.1 Medidas de asimetría

- Índice de simetría de **Pearson**:

$$f_1 = \frac{\overline{x} - Mo}{\sigma}$$

- Índice de simetría de **Fisher**:

$$f_2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^3}{n\sigma^3}$$

```
* Simétrico : valores entre -0,5 y 0,5
* Datos asimétricos moderados : valores entre -1 y -0,5 o entre 0,5 y 1
* Datos muy sesgados : valores menores que -1 o mayores que 1
```

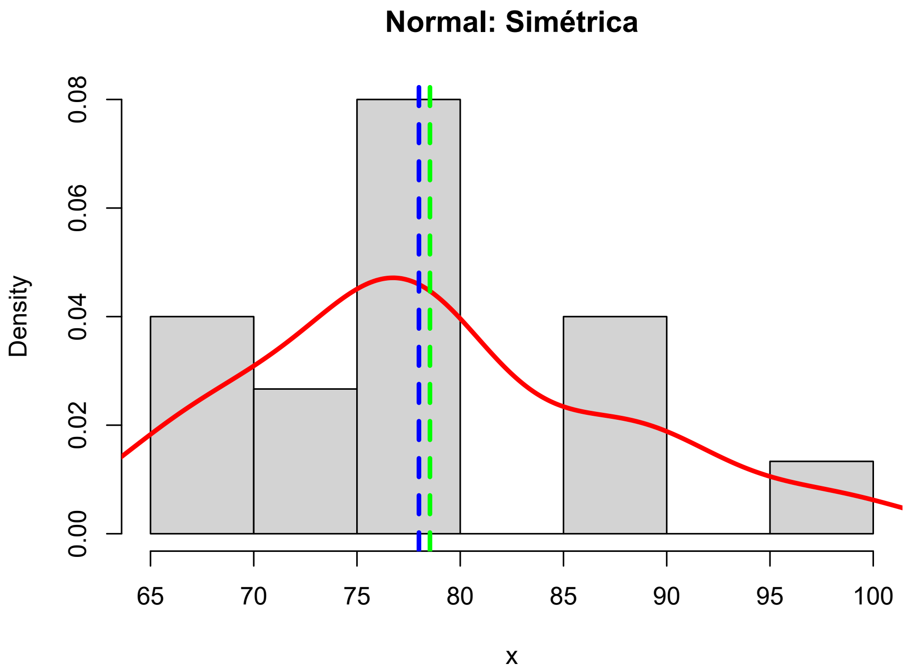
Si la distribución es simétrica, ambos índices son iguales a 0; si es asimétrica a la derecha, ambos son positivos; y si es asimétrica a la izquierda, ambos índices son negativos.

8.1. Normal: Simétrica

wwwwwwwwwwwwwwwwwwwwwwwww

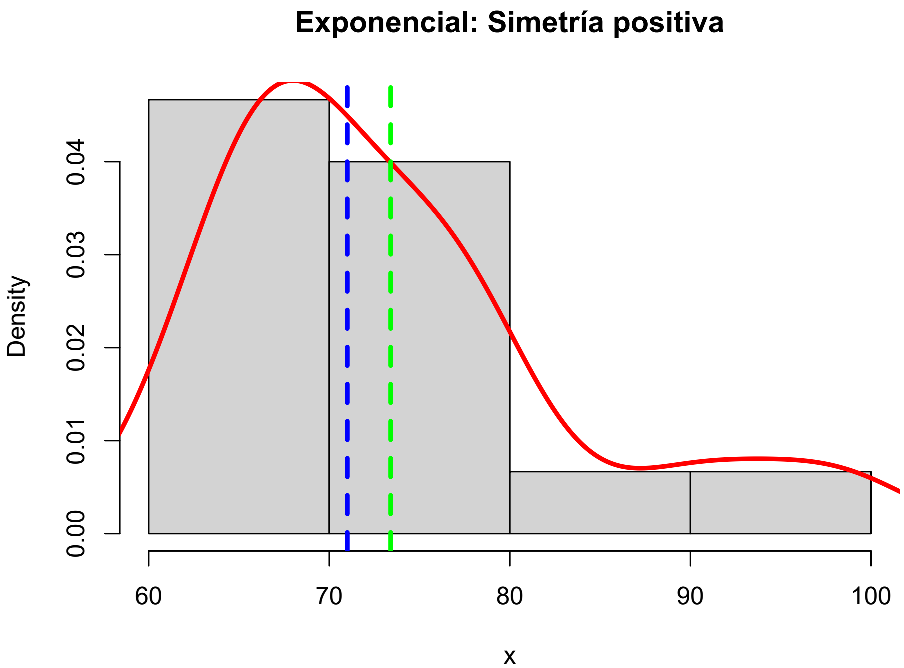
```
## [1] 0.5059805
```

```
## [1] 78.53333
```



8.2. Normal: Simetría positiva

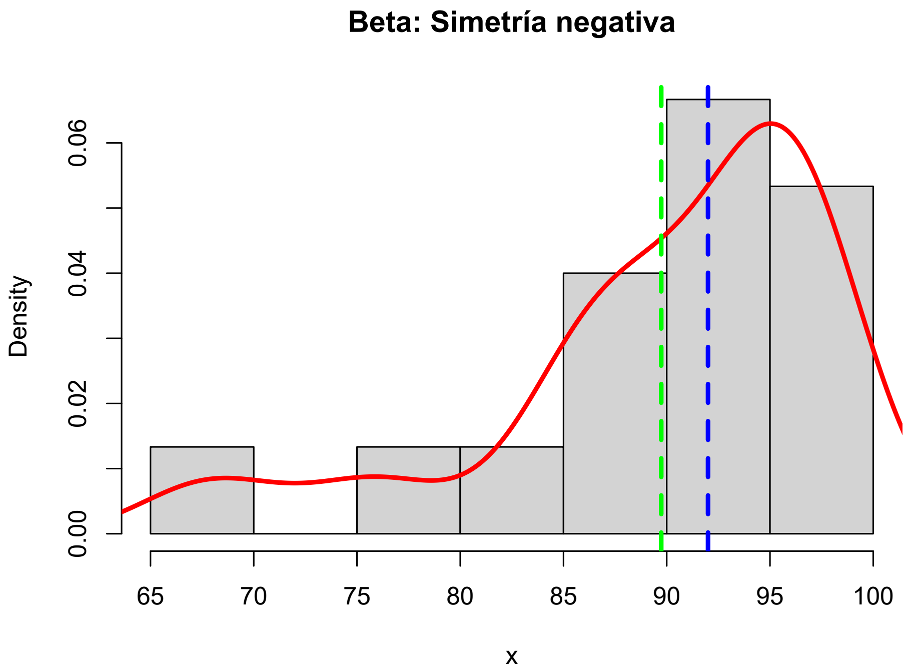
wwwwwwwwwwwwwwwwwww




```
## [1] 1.216579
```

8.3. Normal: Simetría negativa

wwwwwwwwwwwwwwwwwww




```
## [1] -1.340429
```

Clase	Y_i	f_i	$Y_i * f_i$
$[5, 10)$	7.5	2	7.5
$[10, 15)$	12.5	3	25
$[15, 20)$		4	87.5
$[20, 25)$	22.5	7	157.5
$[25, 30]$	27.5	10	275
$[30, 35]$		8	195
$\sum$			

9

Medidas de curtosis o apuntamiento

En estadística, usamos la medida de curtosis para describir la "cola" de la distribución, ya que describe la forma de la misma. También es una medida del "pico" de la distribución.

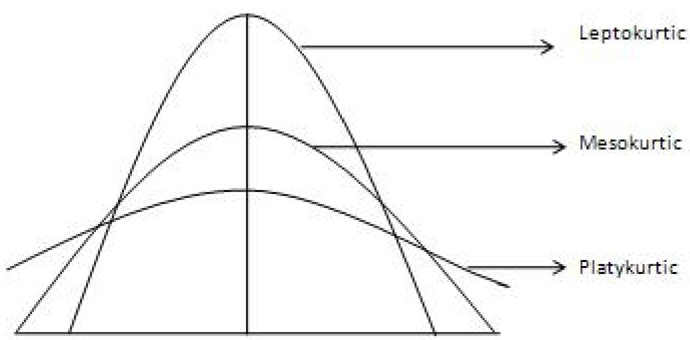


Figura 9.1 wwwwwwwwwwwwwwwww

1. **Mesocurtica** : esta es la distribución normal
2. **Leptocurtica** : esta distribución tiene colas más gruesas y un pico más afilado. La curtosis es "positiva" con un valor superior a 3
3. **Platicurtica** : La distribución tiene un pico más bajo y más ancho y colas más delgadas. La curtosis es "negativa" con un valor superior a 30.263

- En base a la media y desviación típica

$$k = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^4}{ns^4};$$

Si $k = 3$ además $k \geq 3$

- En base a percentiles
$$k = \frac{P_{75} - P_{25}}{2\,(P_{90} - P_{10})}$$
Si $k < 3$ y si $k = 3$ ademas $k \geq 0{,}263$

Si este coeficiente es nulo, la distribución se dice normal (similar a la distribución normal de Gauss) y recibe el nombre de mesocúrtica.

Si el coeficiente es positivo, la distribución se llama leptocúrtica, más puntiaguda que la anterior. Hay una mayor concentración de los datos en torno a la media.

Si el coeficiente es negativo, la distribución se llama platicúrtica y hay una menor concentración de datos en torno a la media. sería más achatada que la primera.

Parte II

Probabilidades

1

Experimento aleatorio

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

2

Álgebra de eventos

Sean A, B y C eventos entonces 1. e 2. wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww

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

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

wwwwwwwwwwwwwwwwwwww

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

6

Teorema de Bayes

Teorema 6.1 (Teorema de Bayes)**.** *Sea* $\{A_1, A_2, ..., A_i, ..., A_n\}$ *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**

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

$$f\left(k\right) = \binom{n}{k} p^{k} \left(1-p\right)^{n-k} \tag{1.1}$$

You may refer to it using (1.1)

$$\begin{aligned} 1 &= 2 \\ w &= 2 \end{aligned} \tag{1.2}$$

$$\begin{aligned} 1 &= 1 \\ &= 2 \\ &= 3 \\ &= 2 \end{aligned} \tag{1.3}$$

Definición 1.2 (wwwwwww)**.** wwwwwwwwwwwwwwww

Teorema 1.1 (wwwwwwwwwwwwwwww)**.** *wwwwwwwwwwwwwwwwwwwwwwwww*

Ejemplo 1.1 (wwwwwwww wwww)**.** wwwwwwww

Ejercicio 1.1 (wwwwwwwwwwwwwwwww). wwwwwwwwwwwwwwww

$R_X = \{x \in \mathbb{R} / \ \exists \omega \in \Omega : X(\omega) = x\}$ 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 1.2. 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

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

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

Definición 1.3 (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

Cuando el rango de la variable aleatoria X, R_X es *finito* o *infinito* contable (no necesarimente enteros) $R_X = \{x_1, x_2, \ldots, x_n \ldots\}$

1.1.2. Variable aleatoria continua

R_X abarca cualquier intervalo en la recta numerica

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

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

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

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

Ejemplo 1.3. 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

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

Ejemplo 1.4. 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

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

Ejemplo 1.5. 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

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

1.2.2. Función de probabilidad de una variable aleatoria continua

Definición 1.5 (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 1.6. For a circle with the radius `r x`, its area is `r pi * x^2`. Sea la función

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

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

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

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

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

además

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

Ejemplo 1.8. Sea $f(x) = \frac{\sigma}{\rho}$ es una funcion de densidad pues

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

además

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

1.3. Función de distribución de una variable aleatoria

1.3.1. Función de distribución de una variable aleatoria discreta

Definición 1.6 (Función de distribución)**.** Sea X una variable aleatoria con rango

$$R_X = \{x_1, x_2, \ldots x_n, \ldots\}.$$

Con función de probabilidad $p(x_i) = P[X = x_i]$, sea x cualquier número, real la función definida por

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

recie 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. $P(a \leq X \leq b) = F_X(b) - F_X(a^-)$

1.3.2. Función de distribución de una variable aleatoria continua

Definición 1.7 (Función de distribución)**.** Sea X una variable aleatoria con función de densidad $f(x)$. La función

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

Cuyas propiedades son:

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

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)**.**

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

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

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

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

o bien

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

momento de orden superior

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

3

Variables aleatorias bidimensionales

Definición 3.1 (Variable aleatoria bidmensional 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. www

Definición 3.2 (Variable aleatoria bidmensional contínua)**.**

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

Ejemplo 3.2. wwwww

3.1. Distribución bidimensional discreta

Definición 3.3 (Función de probabilidad conjunta)**.** Sea (X,Y) una variable bidimensinal 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. $\sum_{x\in R_X} \sum_{y\in R_Y} p(x,y) = 1$

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

Definición 3.4 (Función de distribución acumulada)**.**

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

3.1.1. Distribuciones marginales

3.1.2. Variables aleatorias independientes

3.1.3. Distribuciones de probabilidad condicional

3.2. 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 $x_1, x_2, \ldots, x_n$ se simboliza en forma compacta mediante el simbolo $\sum$ (sigma) es decir la suma de los números anteriores se puede escribir del siguiente modo

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

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

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

   citado por (Xie, 2015) Variable estadística variable estadística

A.1. ee

$\sum_{i=1}^{n} x_i$

$$\sum_{i=1}^{n} \frac{1}{2}$$

A.1.1. eeeee

B

Matrices

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

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

de **orden** $n \times m$ tiene **entradas** a_{ij} donde el primer subindice 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. Algebra de matrices

Sean las matrices

$$A = [a_{ij}]_{n\times m} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{11} & a_{11} & \dots & a_{nm} \end{pmatrix}_{n\times m}$$

y

$$B = [b_{ij}]_{p\times q} = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{11} & b_{11} & \dots & b_{pq} \end{pmatrix}_{p\times q}$$

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.

$$\begin{aligned} kA &= k[a_{ij}]_{n\times m} \\ &= [ka_{ij}]_{n\times m} \\ &= \begin{pmatrix} ka_{11} & ka_{12} & \dots & ka_{1n} \\ ka_{21} & ka_{22} & \dots & ka_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{11} & ka_{11} & \dots & ka_{nm} \end{pmatrix}_{n\times m} \end{aligned}$$

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

$$\begin{aligned} A \pm B &= [a_{ij} + b_{ij}]_{n\times m} \\ &= \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{11} + b_{11} & a_{11} + b_{11} & \dots & a_{nm} + b_{nm} \end{pmatrix}_{n\times m} \end{aligned}$$

donde $i = 1, 2, \dots n$ 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

$$\begin{aligned} A \cdot B &= \left[\sum_{k=1}^{p} a_{ik}b_{kj}\right]_{n\times q} \\ &= \begin{pmatrix} \sum_{k=1}^{m} a_{1k}b_{k1} & \sum_{k=1}^{m} a_{1k}b_{k2} & \dots & \sum_{k=1}^{m} a_{1k}b_{kq} \\ \sum_{k=1}^{m} a_{2k}b_{k1} & \sum_{k=1}^{m} a_{2k}b_{k2} & \dots & \sum_{k=1}^{m} a_{2k}b_{kq} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=1}^{m} a_{nk}b_{k1} & \sum_{k=1}^{m} a_{nk}b_{k2} & \dots & \sum_{k=1}^{m} a_{nk}b_{kq} \end{pmatrix}_{n\times q} \end{aligned}$$

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

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 verfican las siguientes propiedades

1. $A(B + C) = AB + AC$
2. $(A + B)C$
3. $A(BC) = (AB)C$

B.2. Matrices particulares

En esta seccion se considera las siguientes matrices: Matriz triangular, matriz particular de una matriz cuadrada, matriz transpuesta, matriz simetrica, matriz conjugada, matriz hermitica, matriz escalonada.

B.2.1. Matriz triangular

Una matriz cuadrada A cuyos elementos $a_{ij} = 0$ si $i > j$ es decir

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn} \end{pmatrix}_{n\times n}$$

se llama **matriz triangular superior**; reciprocamente si $i < j$ es decir

$$A = \begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{11} & a_{11} & \ldots & a_{nn} \end{pmatrix}_{n\times n}$$

se llama **matriz triangular inferior**. Si A es a la vez **matriz triangular superior** y **matriz triangular inferior** entonces recibe el nombre de **matriz diagonal**, representada por

$$(a_{11}, a_{22}, \ldots, a_{nn})$$

además si $a_{11} = a_{22} = \ldots = a_{nn} = k$ la matriz recibe el nombre de **matriz escalar** y si $k = 1$ la matriz recibe el nombre de **matriz unidad** representada por I_n por ejemplo

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

B.2.2. Matriz particular de una matriz cuadrada

B.2.3. Matriz transpuesta

B.2.4. Matriz simetrica

B.2.5. Matriz conjugada

B.2.6. Matriz hermitica

B.2.7. Matriz escalonada

$$\begin{pmatrix} w & warnwww & w \\ w & warnwww & w \end{pmatrix}_{4\times 3}$$

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

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

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

1. Www
$$\frac{1}{20\sqrt{2\pi}} \int_{-\infty}^{300} e^{-\frac{1}{2}\left(\frac{z-200}{20}\right)^2} dz = 0{,}9999997$$
2. 0.9500042 also Es decir los elementos son demagogos y déspotas
3. 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 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.

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 déspotas . It is not the same `x` as the one in R.

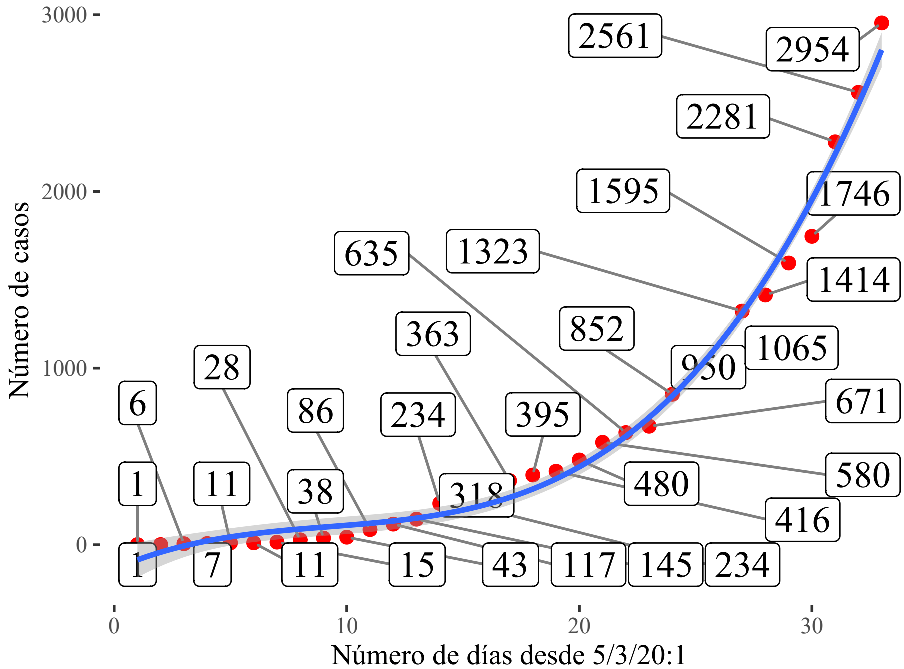


Figura B.1 Regresión lineal

```
## (Intercept)
##    12917.13
```

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, 11

traza, 93