



## DESCRIPTIVA

UNA VARIABLE CUALITATIVA  
UNA VARIABLE CUANTITATIVA  
DOS VARIABLES CUALITATIVAS  
UNA V. CUALI. Y UNA V. CUANTI.  
DOS VARIABLES CUANTITATIVAS

*Emilio Letón*  
*Dpto. Estadística, UC3M*

## ¿Dónde estamos?

DESCRIPTIVA

CÁLCULO  
DE  
PROBAB.

INFERENCIA

1981

## YT: SBTC

Standing in the dark  
Oh I was waiting for man to come  
I am beautiful and clean

To cut a long story short

## YT: SBTC

To cut a long story short  
I lost my mind

Questions, questions  
Give me no answers

## Frentes abiertos

Empezar

Llegar a las poblaciones

Estadística: E. Letón

## UNA VARIABLE CUALITATIVA

Frecuencias, moda  
Gráficos de barras, de tartas

UNA VARIABLE CUANTITATIVA

DOS VARIABLES CUALITATIVAS

UNA V. CUANTI. Y UNA V. CUALI.

DOS VARIABLES CUANTITATIVAS

Estadística: E. Letón

## UNA VARIABLE CUALITATIVA



Florence  
Nightingale  
(1820-1910)

Estadísticos  
vs. Parámetros

Gráficos vs.  
Gráficos poblacionales

PORTRAITS HISTORY STATISTICS

Estadística: E. Letón

## Frecuencias

1 (a)  
2 (b)  
1 (a)  
1 (a)  
3 (c)  
3 (c)  
1 (a)  
1 (a)  
1 (a)

	fa	fr	faa	fra
1	7	0,7	7	0,7
2	1	0,1	8	0,8
3	2	0,2	10	1,0

Estadística: E. Letón

## Resumen: frecuencias

---

## Moda

---

1 (a)
2 (b)
1 (a)
1 (a)
1 (a)
3 (c)
3 (c)
1 (a)
1 (a)
1 (a)

	fa	fr	faa	fra
1	7	0,7	7	0,7
2	1	0,1	8	0,8
3	2	0,2	10	1,0

## Resumen: moda

---

## Gráfico de barras

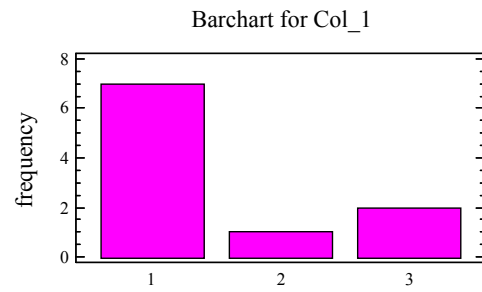
---

1 (a)
2 (b)
1 (a)
1 (a)
1 (a)
3 (c)
3 (c)
1 (a)
1 (a)
1 (a)

	fa	fr	faa	fra
1	7	0,7	7	0,7
2	1	0,1	8	0,8
3	2	0,2	10	1,0

## Sin acumular

---



---

Estadística: E. Letón

## Acumulado

---

---

Estadística: E. Letón

## Simetría

---

---

Estadística: E. Letón

## Resumen: gráfico de barras

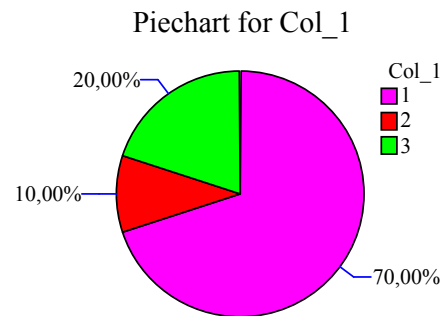
---

---

Estadística: E. Letón

## Gráfico de tartas

---



Estadística: E. Letón

## Resumen: gráfico de tartas

---

Estadística: E. Letón

**UNA VARIABLE CUALITATIVA**

---

**UNA VARIABLE CUANTITATIVA**

---

Media, mín, máx, dt, cuartiles, CAS  
Histograma, box-plot

**DOS VARIABLES CUALITATIVAS**

---

**UNA V. CUANTI. Y UNA V. CUALI.**

---

**DOS VARIABLES CUANTITATIVAS**

---

Estadística: E. Letón

**UNA VARIABLE CUANTITATIVA**

---



John Tukey  
(1915-2000)

Estadísticos  
vs. Parámetros

Gráficos vs.  
Gráficos poblacionales

Estadística: E. Letón

## Media

1
4
5
5
5
6
6
6
7
10

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \sum_{j=1}^k x_j fr(x_j)$$

Estadística: E. Letón

## Propiedades (1/3)

1
4
5
5
5
6
6
6
7
10

Medida de centralización

Estadística: E. Letón

## Propiedades (2/3)

1
4
5
5
5
6
6
6
7
10

¿Unidades?

Estadística: E. Letón

## Propiedades (3/3)

1	1
4	4
5	5
5	5
5	5
6	6
6	6
6	6
7	7
10	100

¿Sensible a atípicos?

Estadística: E. Letón

## Resumen: media

---

## Mediana

---

1
4
5
5
5
6
6
6
7

1
4
5
5
5
6
6
6
7
10

---

Estadística: E. Letón

---

Estadística: E. Letón

## Propiedades (1/3)

---

1
4
5
5
5
6
6
6
7
10

Medida de centralización

## Propiedades (2/3)

---

1
4
5
5
5
6
6
6
7
10

¿Unidades?

---

Estadística: E. Letón

---

Estadística: E. Letón

## Propiedades (3/3)

1	1
4	4
5	5
5	5
5	5
6	6
6	6
6	6
7	7
10	100

¿Sensible a atípicos?

Estadística: E. Letón

## Resumen: mediana

Estadística: E. Letón

## Mínimo, máximo y amplitud

1
4
5
5
5
6
6
6
7
10

Estadística: E. Letón

## Propiedades

1
4
5
5
5
6
6
6
7
10

Medida de dispersión

¿Unidades?

¿Sensible a atípicos?

Estadística: E. Letón



## Resumen: min, máx, amplitud

---

## Varianza

---

1	1	1
4	2	1
5	3	1
5	4	1
5	5	1
6	6	10
6	7	10
6	8	10
7	9	10
10	10	10

---

Estadística: E. Letón

---

Estadística: E. Letón

## Origen (1/2)

---

## Origen (2/2)

---

	D1	D2	D3
$s^2$	4,65	8,25	20,25
$s^2_{n-1}$	5,17	9,17	22,50

---

Estadística: E. Letón

---

Estadística: E. Letón

## Propiedades (1/3)

1
4
5
5
5
6
6
6
7
10

Medida de dispersión

Estadística: E. Letón

## Propiedades (2/3)

1
4
5
5
5
6
6
6
7
10

¿Unidades?

Estadística: E. Letón

## Propiedades (3/3)

1
4
5
5
5
6
6
6
7
10

1
4
5
5
5
6
6
6
7
100

¿Sensible a atípicos?

Estadística: E. Letón

## Resumen: varianza

Estadística: E. Letón

## Desviación típica

1	1	1
4	2	1
5	3	1
5	4	1
5	5	1
6	6	10
6	7	10
6	8	10
7	9	10
10	10	10

Estadística: E. Letón

## Origen

	D1	D2	D3
s	2,16	2,87	4,50
s <sub>n-1</sub>	2,27	3,03	4,74

Estadística: E. Letón

## Propiedades

1
4
5
5
5
6
6
6
7
10

Medida de dispersión

¿Unidades?

¿Sensible a atípicos?

Estadística: E. Letón

## Resumen: desviación típica

Estadística: E. Letón

## Desigualdad de Chebyshev (1/2)



Pafnuty  
Chebyshev  
(1821-1894)

Alrededor de la media  
muestral +/- k veces la dt  
muestral hay al menos ...

$$\left(1 - \frac{1}{k^2}\right)\%$$

## Desigualdad de Chebyshev (2/2)

k	%min.obs.
1	0
1,5	55,6%
2	75,0%
3	88,9%
4	93,8%
10	96,0%

## Resumen: des. de Chebyshev

## Cuartiles

1
4
5
5
5
6
6
6
7
10

## Propiedades

1
4
5
5
5
6
6
6
7
10

Medida de posición

¿Unidades?

¿Sensible a atípicos?

Estadística: E. Letón

## Resumen: cuartiles

Estadística: E. Letón

## Rango intercuartílico

1	1	1
4	2	1
5	3	1
5	4	1
5	5	1
6	6	10
6	7	10
6	8	10
7	9	10
10	10	10

Estadística: E. Letón

## Propiedades

1
4
5
5
5
6
6
6
7
10

Medida de dispersión

¿Unidades?

¿Sensible a atípicos?

Estadística: E. Letón

## Resumen: rango intercuartílico

---

## Deciles y percentiles

---

$$P_p = \begin{cases} \frac{1}{2}(x_{(i)} + x_{(i+1)}) & , f = 0 \\ x_{(i+1)} & , f > 0 \end{cases}$$

$$i = [np] \quad f = np - [np]$$

## Propiedades

---

1
4
5
5
5
6
6
6
7
10

Medida de posición

¿Unidades?

¿Sensible a atípicos?

## Resumen: deciles y percentiles

---

## Coeficiente de asimetría

4
5
9

¿Puede haber dos conjuntos de datos, de igual tamaño, con todos los valores distintos entre sí, con igual media y con igual dt?

3
7

## Origen (1/3)

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

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

## Origen (2/3)

$$\frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{s^3} = CAS$$

## Origen (3/3)

$$CAS = \frac{n}{(n-1)(n-2)} \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{\hat{s}^3}$$

## Propiedades (1/3)

1
4
5
5
5
6
6
6
7
10

Medida de simetría

¿Unidades?

¿Sensible a atípicos?

## Propiedades (2/3)

Simetría implica  
CAS=0

CAS=0 no implica  
simetría

Simetría implica que media=mediana

media=mediana no implica simetría

## Propiedades (3/3)

CAS>0

CAS<0

## Resumen: coeficiente de asimetría



## Des. Cheb. (sim. y unimodal) (1/2)

Alrededor de la media muestral +/- k veces la dt muestral hay al menos ...

$$\left(1 - \frac{4/9}{k^2}\right)\%$$

## Des. Cheb. (sim. y unimodal) (2/2)

k	%min.obs.	%min.obs. (sim+unim)
1	0	55,6%
1,5	55,6%	80,2%
2	75,0%	88,9%
3	88,9%	95,1%
4	93,8%	97,2%
10	96,0%	99,6%

## Resumen: des. Che. (sim. y unim.)

## Coeficiente de apuntamiento

¿Puede haber dos conjuntos de datos, de igual tamaño, con todos los valores distintos entre sí, con igual media, con igual dt y con igual CAS?

## Origen

---

$$\frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{s^4} - 3 = CAP$$

## Propiedades (1/2)

---

1
4
5
5
5
6
6
6
7
10

Medida de apuntamiento

¿Unidades?

¿Sensible a atípicos?

## Propiedades (2/2)

---

CAP < 0  
platicúrtica

CAP = 0  
mesocúrtica

CAP > 0  
leptocúrtica

## Resumen: coeficiente de apuntam.

---

## Otros estadísticos

---

$$MEDA = \text{Mediana} |x_i - x_{med}|$$

$$CV = \frac{s}{|\bar{x}|}$$

$$CSR = \frac{|\bar{x}|}{s}$$

## Resumen: otros estadísticos

---

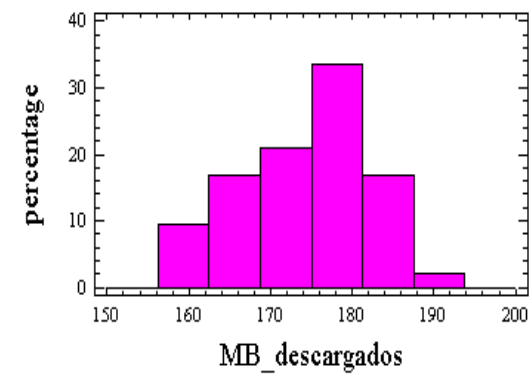
## Histograma

---

## Sin acumular

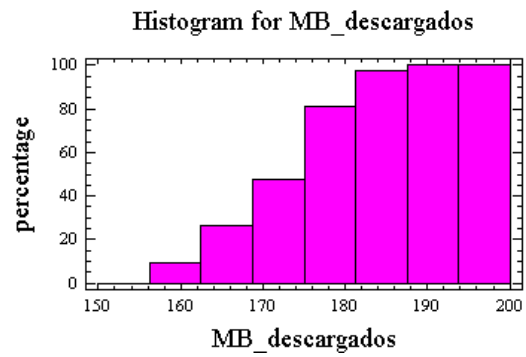
---

Histogram for MB\_descargados



## Acumulado

---



---

Estadística: E. Letón

## Simetría

---

---

Estadística: E. Letón

## Resumen: histograma

---

---

Estadística: E. Letón

## Polígono de frecuencias

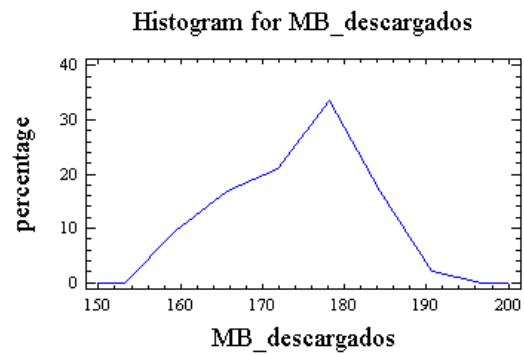
---

---

Estadística: E. Letón

## Sin acumular

---

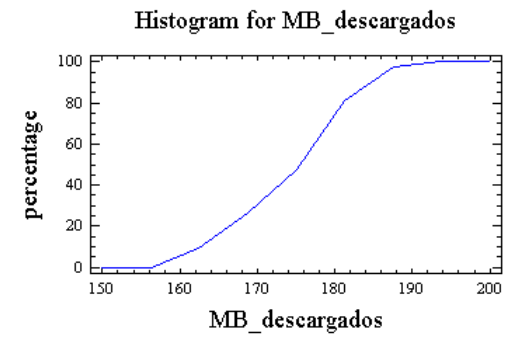


---

Estadística: E. Letón

## Acumulado

---



---

Estadística: E. Letón

## Resumen: polígono de frecuencias

---

---

Estadística: E. Letón

## Box-Plot

---

John Tukey  
(1977)

BOX-PLOT WIKIPEDIA

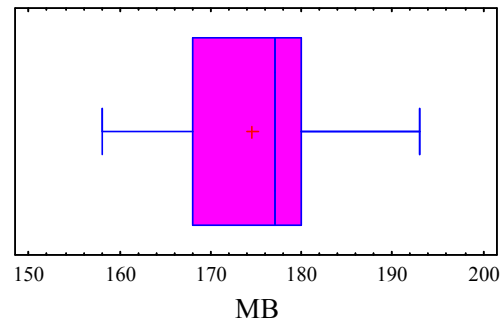
---

Estadística: E. Letón

## Sin datos atípicos

---

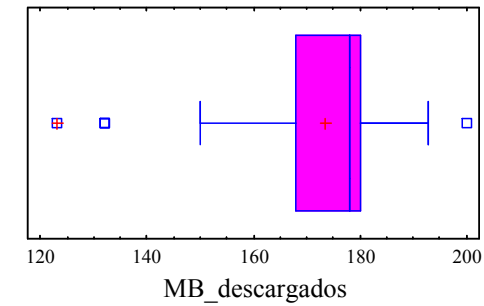
Box-and-Whisker Plot



## Con datos atípicos

---

Box-and-Whisker Plot



## Resumen: box-plot

---

UNA VARIABLE CUALITATIVA

---

UNA VARIABLE CUANTITATIVA

---

DOS VARIABLES CUALITATIVAS

---

Tabla de contingencia

Gráfico de bloques

UNA V. CUANTI. Y UNA V. CUALI.

---

DOS VARIABLES CUANTITATIVAS

---

## DOS VARIABLES CUALITATIVAS

x1	x2
1	1
1	2
1	3
1	2
2	2
2	3
2	3
2	1
2	1
2	1

Estadísticos  
vs. Parámetros

Gráficos vs.  
Gráficos poblacionales

## Tabla de contingencia

x1	x2
1	1
1	2
1	3
1	2
2	2
2	3
2	3
2	1
2	1
2	1

## Elementos

## Notación

		Var 1=x			
		Cat1	Cat 2	... Cat c	
Var 2=y	Niv1	$n_{11}$	$n_{12}$	... $n_{1c}$	$r_1$
	Niv2	$n_{21}$	$n_{22}$	... $n_{2c}$	$r_2$
	...	...	...	... ..	...
	Niv r	$n_{r1}$	$n_{r2}$	... $n_{rc}$	$r_r$
		$c_1$	$c_2$	... $c_c$	$n$

## Independencia (1/2)

---

$$fr(y_j | x_i) \stackrel{ind.}{=} fr(y_j)$$

$$fr(y_j | x_i) = \frac{n_{ji}}{c_i} = \frac{fr(y_j \cap x_i)}{fr(x_i)}$$

$$\stackrel{ind.}{=} fr(y_j)$$

## Independencia (2/2)

---

$$\Leftrightarrow fr(y_j \cap x_i) = fr(y_j)fr(x_i)$$

$$\Leftrightarrow \frac{n_{ji}}{n} = \frac{r_j}{n} \cdot \frac{c_i}{n} \Leftrightarrow n_{ji} = \frac{r_j c_i}{n}$$

## Resumen: tabla de contingencia

---

## Gráfico de bloques

---



## Resumen: gráfico de bloques

---

UNA VARIABLE CUALITATIVA

---

UNA VARIABLE CUANTITATIVA

---

DOS VARIABLES CUALITATIVAS

---

UNA V. CUANTI. Y UNA V. CUALI.

---

Estadísticos por grupo  
Gráficos por grupo

DOS VARIABLES CUANTITATIVAS

---

## UNA V. CUALI. Y UNA V. CUANTI.

---

x1	x2
1	12
1	21
1	33
1	21
2	21
2	31
2	32
2	15
2	17
2	18

Estadísticos por grupos

Gráficos por grupos

## Estadísticos por grupo

---

## Resumen: estadísticos por grupo

---

## Gráficos por grupo

---



## Resumen: gráficos por grupo

---

UNA VARIABLE CUALITATIVA

---

UNA VARIABLE CUANTITATIVA

---

DOS VARIABLES CUALITATIVAS

---

UNA V. CUANTI. Y UNA V. CUALI.

---

**DOS VARIABLES CUANTITATIVAS**

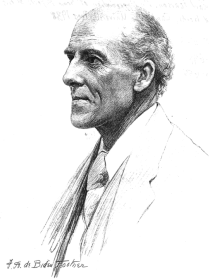
---

Covarianza, correlación

Diagrama de dispersión

## DOS VARIABLES CUANTITATIVAS

x	y
2	1
6	10
1	1
5	5
7	14
7	13
1	3
3	1
4	7
5	6



Pearson  
(1857-1936)

Estadísticos

Gráficos

Estadística: E. Letón

## Covarianza

x	y
2	1
6	10
1	1
5	5
7	14
7	13
1	3
3	1
4	7
5	6

Estadística: E. Letón

## Origen (1/2)

$$s_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

Estadística: E. Letón

## Origen (2/2)

$$\hat{s}_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

Estadística: E. Letón

## Propiedades

Medida de variabilidad conjunta

¿Unidades?

¿Sensible a atípicos?

## Resumen: covarianza

## Correlación

x	y
2	1
6	10
1	1
5	5
7	14
7	13
1	3
3	1
4	7
5	6

$$r = \frac{s_{xy}}{s_x s_y} = \frac{\hat{s}_{xy}}{\hat{s}_x \hat{s}_y} = 0,9$$

$$r^2 = 81\%$$

## Propiedades

Medida de relación lineal

¿Unidades?

¿Rango de valores?

¿=0?

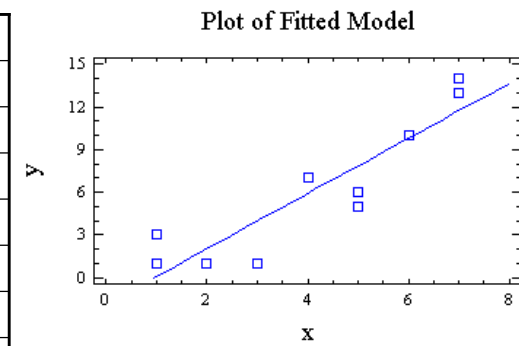
## Resumen: correlación

---

## Gráfico de dispersión

---

x	y
2	1
6	10
1	1
5	5
7	14
7	13
1	3
3	1
4	7
5	6



## Tipos (1/2)

---

## Tipos (2/2)

---

x	y
-4	16
-3	9
-2	4
-1	1
0	0
1	1
2	4
3	9
4	16

## Resumen: gráfico de dispersión

## Causalidad

x	y
2	4
6	20
8	60
7	30
0	7
1	6
2	14
2	20
3	18

$$r = 0,846$$

## Resumen: causalidad

## Webgrafía: web de la asignatura



*Software; Prácticas; ABP; Autoevaluación;  
Ejercicios; Mini-Vídeos; CPC; Tutorías; Webgrafía*