

Clase 1: Datos descriptivos

Denise Laroze

24 de agosto de 2018

CESS - Universidad de Santiago
denise.laroze@usach.cl

¿Para qué sirven los métodos cuantitativos?

Repaso

Estadísticas descriptivas

**¿Para qué sirven los métodos
cuantitativos?**

Uso de métodos cuantitativos

Entre otras cosas, sirve para:

- Resumir información de números “grandes” de personas.
- Identificar patrones/tendencias en la/s poblaciones de interés
- Establecer similitudes y diferencias entre distintos grupos
- Identificar cuando dos variables co-varían
- (Sólo en casos muy específicos) identificar causas de el/los fenómenos de interés
- Permite responder preguntas del tipo: ¿Cuándo ocurre Y? ¿Hay una asociación entre X e Y? ¿Si pasa X, que se espera que pase en Y?

En cambio, es muy difícil

Entender el por qué de las cosas.

- El por qué está dado por la **teoría**
- Los datos nos dan evidencia a favor o en contra de las asociaciones que nuestra teoría nos dice que probemos, pero son sólo tan buenas como nuestra teoría.
- Los resultados estadísticos son sobre el **promedio**, siempre el promedio
- Los datos **NO** nos dicen qué ocurrió en cada caso específico.

Repaso

Nomenclatura básica:

X = Variable independiente

Y = Variable dependiente

$X \rightarrow Y$ Una hipótesis

$Y = \alpha + \beta_1 X + \epsilon$ Una regresión lineal

Tipos de Variables:

¿Cuáles son?

Tipos de Variables:

¿Cuáles son?

1. Continuas (ej. edad, tiempo, calor, PIB)

¿Cuáles son?

1. Continuas (ej. edad, tiempo, calor, PIB)
2. Dicotómicas (ej. sí, no)

¿Cuáles son?

1. Continuas (ej. edad, tiempo, calor, PIB)
2. Dicotómicas (ej. sí, no)
3. Categóricas (ej. género, partidos políticos, compañías. drogas, tipos de estafas)

Tipos de Variables:

¿Qué tipo de variables son estas?

Tipos de Variables:

¿Qué tipo de variables son estas?

1. $v1 = (1; 0; 1; 0; 0; 0; 0; 0; 1; 1; 1; 0)$

¿Qué tipo de variables son estas?

1. $v1 = (1; 0; 1; 0; 0; 0; 0; 1; 1; 1; 0)$

2. $v2 = (1; 2; 3; 4; 3; 1; 2; 3; 4; 1; 2; 4; 3; 2; 4)$

¿Qué tipo de variables son estas?

1. $v1 = (1; 0; 1; 0; 0; 0; 0; 1; 1; 1; 0)$
2. $v2 = (1; 2; 3; 4; 3; 1; 2; 3; 4; 1; 2; 4; 3; 2; 4)$
3. $v3 = (1; 2; 3; 4; 99; 1; 2; 3; 4; 1; 999; 4; 3; 2; 4)$

¿Qué tipo de variables son estas?

1. $v1 = (1; 0; 1; 0; 0; 0; 0; 1; 1; 1; 0)$
2. $v2 = (1; 2; 3; 4; 3; 1; 2; 3; 4; 1; 2; 4; 3; 2; 4)$
3. $v3 = (1; 2; 3; 4; 99; 1; 2; 3; 4; 1; 999; 4; 3; 2; 4)$
4. $v4 = (1,23; 1.63; 10,95; 5,89; \dots ; n)$

¿Por qué importa?

- Por que la selección de tipos de estadísticas que debemos utilizar depende directamente del tipo de variable que tenemos
 - Como variables dependientes: el tipo de variable define el tipo de método de estimación debemos usar (OLS, Logit, Poisson, mlogit, ologit, etc)
 - Como variables independientes: el tipo de variable define cómo debemos interpretar los coeficientes

¿Por qué importa?

- Por que la selección de tipos de estadísticas que debemos utilizar depende directamente del tipo de variable que tenemos
 - Como variables dependientes: el tipo de variable define el tipo de método de estimación debemos usar (OLS, Logit, Poisson, mlogit, ologit, etc)
 - Como variables independientes: el tipo de variable define cómo debemos interpretar los coeficientes
- ¿Se puede sacar un promedio de variables categóricas?

¿Por qué importa?

- Por que la selección de tipos de estadísticas que debemos utilizar depende directamente del tipo de variable que tenemos
 - Como variables dependientes: el tipo de variable define el tipo de método de estimación debemos usar (OLS, Logit, Poisson, mlogit, ologit, etc)
 - Como variables independientes: el tipo de variable define cómo debemos interpretar los coeficientes
- ¿Se puede sacar un promedio de variables categóricas?
- ¿Qué estadística sí se puede sacar de una variable categórica?

Estadísticas descriptivas

Ejemplos de datos

N	Nombre	Edad	Horas Trabajo (semanal)
1	Daniel	35	51
2	Loretto	33	44
3	José	30	44
4	Milton	30	43
5	Francisco	38	46
6	Gustavo	38	48
7	Juan	29	43
8	Felipe	31	38
9	Alina	34	48

¿Cómo se llama 1?

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{x_1 + x_2 + x_3 \dots x_n}{n}$$

¿Cómo se llama 2?

$$\left(\frac{n+1}{2}\right)^{ava} \quad \text{o} \quad \bar{x}\left[\left(\frac{n}{2}\right)^{ava}, \left(\frac{n}{2} + 1\right)^{ava}\right]$$

¿Cómo se llama 3? Valor más común en un conjunto de datos

¿Cómo se llama 4?

$$\sigma^2 = \frac{\sum (x - \bar{x})^2}{n} \quad \text{o} \quad \frac{\sum (x - \bar{x})^2}{(n - 1)}$$

¿Cómo se llama 5?

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} \quad \text{o} \quad \sqrt{\frac{\sum (x - \bar{x})^2}{(n - 1)}}$$

Ejercicio

Calcular la media, la mediana, la moda, varianza y desviación estándar de la edad en el cuadro anterior.

Ejercicio

Calcular la media, la mediana, la moda, varianza y desviación estándar de la edad en el cuadro anterior.

Respuesta:

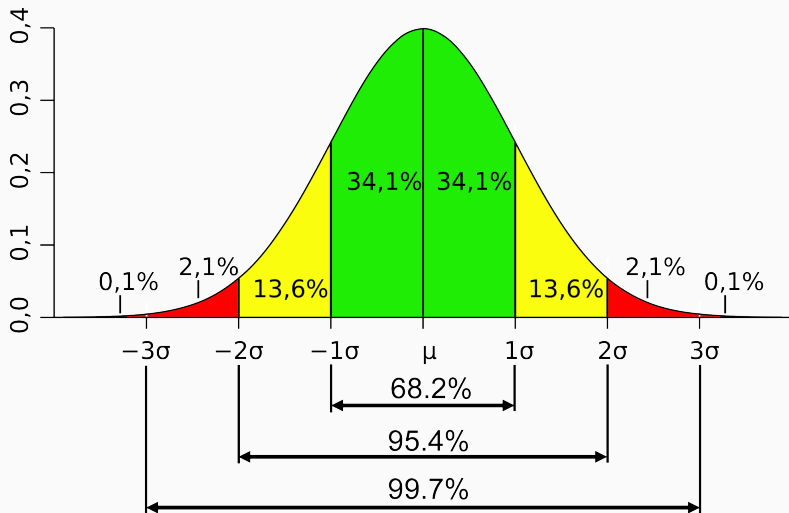
$$\text{Media} = \frac{35+33+30+30+38+38+29+31+34}{9} = 33,1111$$

$$\text{Mediana} = 29, 30, 30, 31, \textcolor{red}{33}, 34, 35, 38, 38 = 33$$

$$\text{Moda} = 30 \text{ y } 38 = \text{es bi-modal}$$

$$\begin{aligned} \text{Varianza} = \\ \frac{(35-33,1)^2 + (33-33,1)^2 + (30-33,1)^2 + \dots + (31-33,1)^2 + (34-33,1)^2}{9} = 11,61111 \end{aligned}$$

$$\text{Desviación estándar} = \sqrt{\text{var}} = 3,407508$$



Ahora pasemos a R