# Probabilidad y Estadística

FCEN – UNCuyo

Raúl MARINO
Florencia CODINA
Augusto NORTE
Marcelo E. ALBERTO

Bibliografía:

**Probabilidad Estadística para Ingeniería y Ciencias**

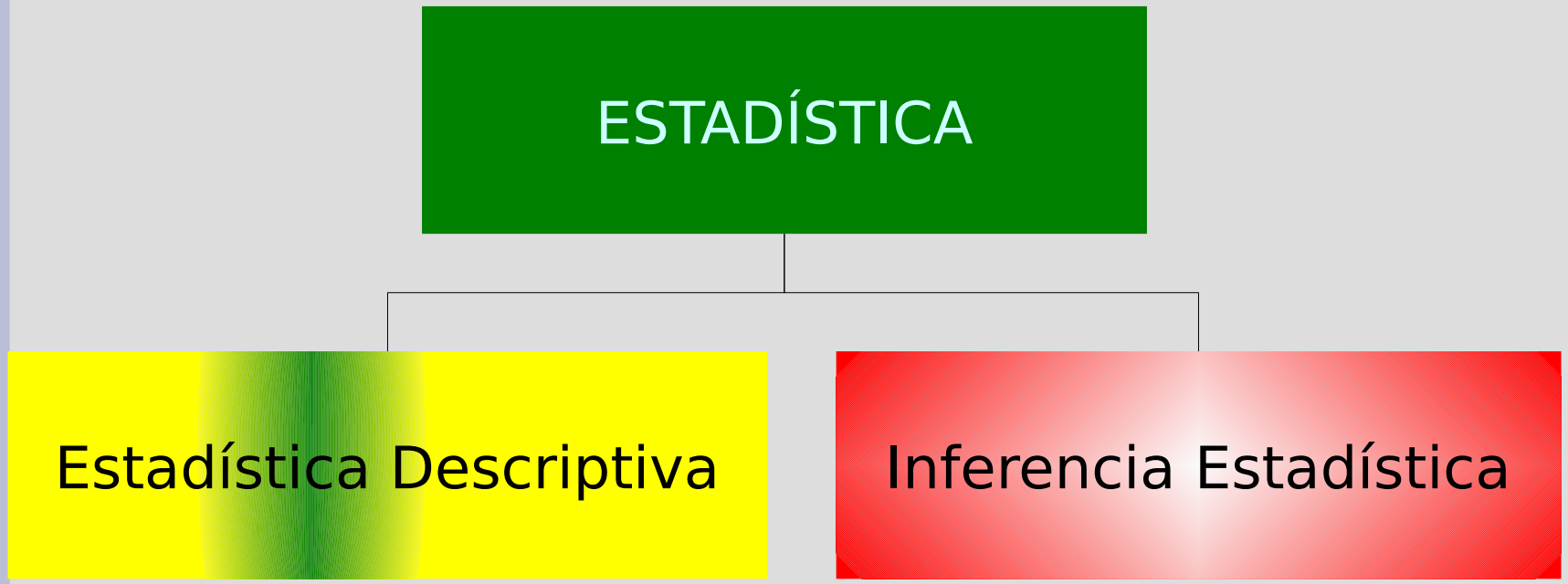
**Walpole – Myers – Myers**

**Mc Graw - Hill**

# VARIABLES ESTADÍSTICAS

# ¿Qué es la Estadística?

La Estadística se ocupa de los métodos y procedimientos para recoger, clasificar, resumir, hallar regularidades y analizar los **datos**, siempre y cuando la variabilidad e incertidumbre sean una causa intrínseca de los mismos; así como de realizar **inferencias** a partir de ellos, con la finalidad de ayudar a la toma de decisiones y en su caso formular **predicciones**.


ESTADÍSTICA
Estadística Descriptiva
Inferencia Estadística

## Estadística descriptiva

Describe, analiza y representa un grupo de datos utilizando métodos numéricos y gráficos que resumen y presentan la información contenida en ellos.

**Ejemplo**: en un estudio sobre *embarazo y diabetes* se midió el peso en kg de 30 mujeres mendocinas, con 35 semanas de gestación, de clase media, de 18 a 30 años de edad:

{75, 52, 92, 71, 51, 84, 75, 65, 81, 98, 60, 106, 113, 77, 91, 75, 65, 98, 73, 91, 80, 67, 61, 68, 85, 67, 67, 102, 73, 103}

**Peso promedio = 78,87 kg**

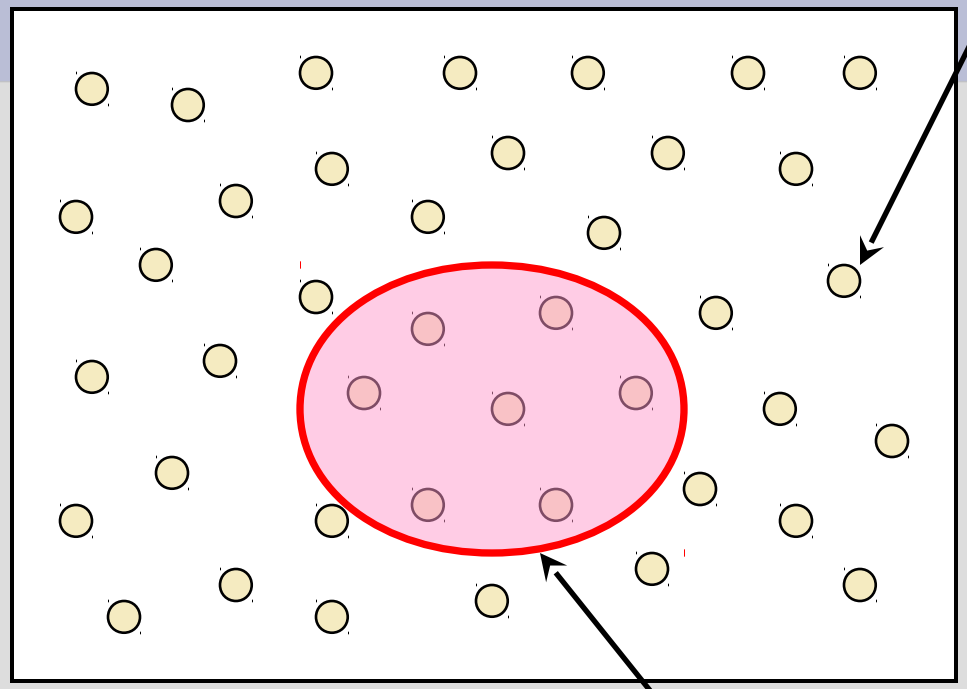
## Estadística inferencial

Apoyándose en el cálculo de probabilidades y a partir de datos muestrales, efectúa estimaciones, toma decisiones, realiza predicciones u otras generalizaciones sobre un conjunto mayor de datos.

**Ejemplo**: con base en una muestra de 30 mujeres se estableció que el peso promedio en kg de las mujeres mendocinas, embarazadas, de clase media, de 18 a 25 años de edad está en el intervalo **[72,8 ; 84,9]**, con una probabilidad de error del 5%.

Unidad de Análisis =
Elemento = Individuo

Población



Muestra

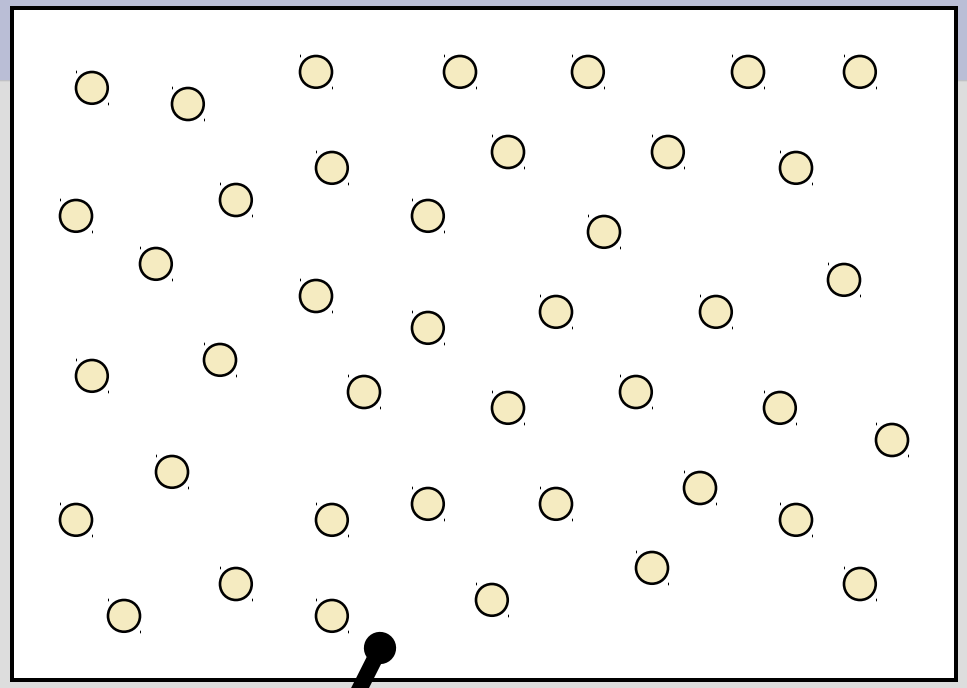
# Definiciones

- Individuos: personas u objetos que contienen la información que se desea estudiar.
- Población: conjunto de individuos que poseen ciertas propiedades comunes.
- Muestra: subconjunto representativo de la población.

Promedio Polacional:

Población

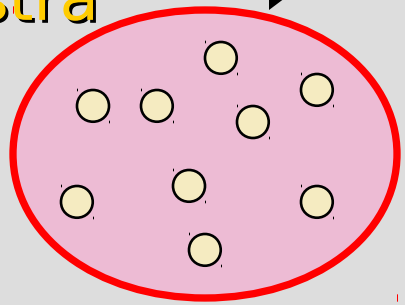
$\mu$



**Parámetro**

**Estadística**

Muestra



Promedio Muestral:

$\bar{x}$

# Definiciones

- Parámetro: función definida sobre los valores numéricos de características medibles de una población.

- Estadística (=estadígrafo): función definida sobre los valores observados de una muestra.

# Problema Estadístico

- Una población de interés y un procedimiento de muestreo
- Los datos y su análisis matemático
- Las inferencias estadísticas que resulten del análisis
- La probabilidad de que esas inferencias sean erróneas

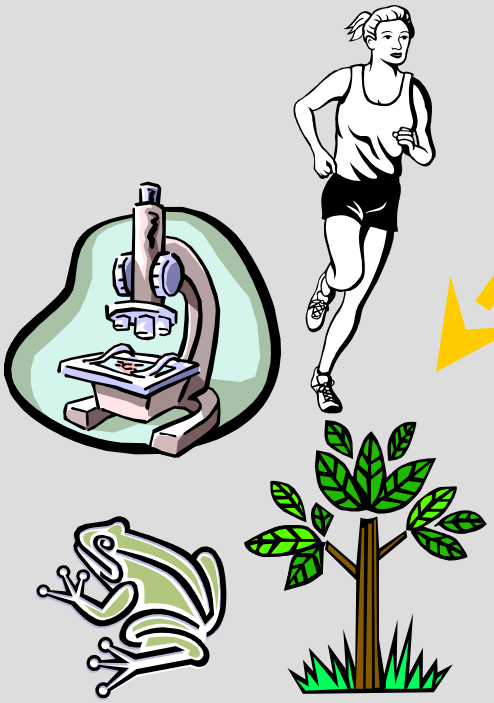
George Canavos

# Organización de Datos

Esta faena es la ciencia: como se ve, consiste en dos operaciones distintas. Una puramente imaginativa, que el hombre pone de su propia y libérrima substancia; otra confrontadora con lo que no es el hombre, con lo que lo rodea, con los hechos, con los datos.

José Ortega y Gasset

# Observación = Medición



Peso Corporal

Porcentaje de hojas atacadas

Número de leucocitos por $mm^3$

Velocidad de crecimiento

Volumen de Expiración Forzada

**VARIABLES ESTADÍSTICAS**

↓

**DATOS**

**Unidad de Análisis**

**[Unidad Experimental]**

# Definiciones

- Variable Estadística (carácter): propiedad, rasgo o cualidad de los elementos de la población.

- Modalidad: uno de los posibles valores que puede tomar una variable estadística.

# Definiciones

- Soporte = Recorrido = Rango: conjunto de todas las posibles modalidades o valores.

- Clase: conjunto de modalidades en la que cada modalidad pertenece a una y sólo una de las clases.

# Definiciones

Ejemplo: en el ejemplo de las embarazadas tenemos:

Variable: peso

Modalidades: 75, 52, 92, 71

Recorrido: $[0,\infty)$

Clase: [51 , 67]

# Tipos de Variables
## (o tipos de datos)

- **Cualitativas**: modalidades de tipo nominal (etiquetas).
  - dicotómicos = binarios
- **Ordinales**: modalidades de tipo nominal naturalmente ordenadas.

- **Numéricas**: modalidades numéricas que admiten operaciones aritméticas.
  i. **Discretas**: número enteros
  ii. **Continuas**: números reales

# Tabla Estadística

Sea un grupo de *n* individuos en el que se observa una variable *X* con modalidades agrupadas en *k* clases: $c_1$, $c_2$,…,$c_k$. Para cada clase se definen:

Frecuencia Absoluta: el número $n_i$ de observaciones con modalidades pertenecientes a la clase $c_i$.

Frecuencia Relativa: el número $f_i$ definido como

$$f_i = n_i / n$$

# Tabla Estadística

Frecuencia Absoluta Acumulada: el número $N_i$ de observaciones con modalidades pertenecientes a cualquiera de las clases $c_j$ con $j \leq i$

$$N_i = \sum_{j=1}^{i} n_j$$

Frecuencia Relativa Acumulada: el número $F_i$ definido como

$$F_i = \sum_{j=1}^{i} \frac{n_j}{n}$$

# Tabla Estadística

- Distribución de Frecuencias: función que asigna a cada clase una y sólo una frecuencia.

- Tabla estadística: una forma de representación ordenada de la distribución de frecuencias

# Tabla Estadística

| Modali. | Frec. Abs. | Frec. Rel. | Frec. Abs. Acumu. | Frec. Rel. Acumu. |
|---|---|---|---|---|
| $C$ | $n_i$ | $f_i$ | $N_i$ | $F_i$ |
| $c_1$ | $n_1$ | $f_1 = \frac{n_1}{n}$ | $N_1 = n_1$ | $F_1 = \frac{N_1}{n} = f_1$ |
| $\dots$ | $\dots$ | $\dots$ | $\dots$ | $\dots$ |
| $c_j$ | $n_j$ | $f_j = \frac{n_j}{n}$ | $N_j = n_1 + \dots + n_j$ | $F_j = \frac{N_j}{n} = f_1 + \dots + f_j$ |
| $\dots$ | $\dots$ | $\dots$ | $\dots$ | $\dots$ |
| $c_k$ | $n_k$ | $f_k = \frac{n_k}{n}$ | $N_k = n$ | $F_k = 1$ |
| | $n$ | $1$ | | |

Ejemplo 1: en una muestra de 400 insectos se observo el sexo de cada uno y se contaron 275 hembras y 125 machos.

La variable observada es:

X:=sexo del insecto

Es de tipo “cualitativa” (binaria o dicotómica)

Las clases son “hembra” y “macho”

Las frecuencias absolutas son 275 y 125 respectivamente

Las frecuencias relativas son 0,6875 y 0,3125 respectivamente

# Tabla Estadística

X: “Sexo de los insectos”

| Clase | Frecuencia Absoluta | Frecuencia Relativa |
|---|---|---|
| Hembra | 275 | 0,6875 |
| Macho | 125 | 0,3125 |
| Total | 400 | 1,0000 |

Ejemplo 2: en una muestra de 80 flores se observo el color de la corola de cada una y se contaron 29 rojas 37 blancas y 14 rosadas.

La variable observada es:

X:=color de la corola

Es de tipo “cualitativa” (politómica)

Las clases son “roja”, “rosada” y “blanca”

Las frecuencias absolutas son 29, 37 y 14 respectivamente

Las frecuencias relativas son 0,3625; 0,4625 y 0,1750 respectivamente

# Tabla Estadística

X: “color de la corola”

| Clase | F. A. | F. R. |
|---|---|---|
| Roja | 29 | 0,3625 |
| Blanca | 37 | 0,4625 |
| Rosada | 14 | 0,1750 |
| Total | 80 | 1,0000 |

Ejemplo 3: en 40 sitios se observó la presencia de cierta especie de gramínea. Según la cantidad de plantas se anotó como “predominante”, “abundante”, “moderada”, “escasa” y “ausente”. Se contaron 9, 21, 6, 3 y 1 sitios respectivamente.

X:=presencia de la especie de gramínea

Es de tipo “ordinal” (cuasicuantitativa)

Las clases son “predominante”, “abundante”, “moderada”, “escasa” y “ausente”

Las frecuencias absolutas son 9, 21, 6, 3 y 1 respectivamente

Las frecuencias relativas son 0,225; 0,525; 0,150; 0,075 y 0,025 respectivamente

# Tabla Estadística

X: “Presencia de una especie de gramínea en distintos sitios”

| Clase | F. A. | F.A. Acum. | F. R. | F. R.Acum. |
|---|---|---|---|---|
| Predominante | 9 | 9 | 0,225 | 0,225 |
| Abundante | 21 | 30 | 0,525 | 0,750 |
| Moderada | 6 | 36 | 0,150 | 0,900 |
| Escasa | 3 | 39 | 0,075 | 0,975 |
| Ausente | 1 | 40 | 0,025 | 1,000 |
| Total | 40 | – | 1,000 | – |

Ejemplo 4: en 30 preparados de cierto tejido se contaron las células de tipo HL.

X:=número de células HL

Es de tipo “cuantitativa discreta”

Las clases son .....

Las frecuencias absolutas son.....

Las frecuencias relativas son.....

# Tabla Estadística

X: “número de células HL”

| Clase | F. A. | F.A. Acum. | F. R. | F.R. Acum. |
|---|---|---|---|---|
| 2 | 12 | 12 | 0,400 | 0,400 |
| 3 | 9 | 21 | 0,300 | 0,700 |
| 4 | 5 | 26 | 0,167 | 0,867 |
| 5 | 3 | 29 | 0,100 | 0,967 |
| 6 | 1 | 30 | 0,033 | 1,000 |
| Total | 30 | – | 1,000 | – |

Ejemplo 5: a través de una cliserie se tomaron muestras de suelo en 40 sitios y se midió el contenido de nitrógeno del suelo expresado en gramos de $N_2O_5$ cada 100g de suelo.

X:= contenido de nitrógeno del suelo expresado en gramos de $N_2O_5$ cada 100g de suelo

Es de tipo “cuantitativa continua”

Las clases son intervalos de números reales no negativos……..

Las frecuencias absolutas son……..

Las frecuencias relativas son……..

# Tabla Estadística

X: “contenido de Nitrógeno en el suelo expresado como g de $N_2O_5$ cada 100g de suelo”

| Clase | Punto Medio | F. A. | F.A. Acumulada | F. R. | F.R. Acumulada |
|---|---|---|---|---|---|
| [0.0;0.3] | 0,15 | 31 | 31 | 0,775 | 0,775 |
| (0.3;0.6] | 0,45 | 7 | 38 | 0,175 | 0,950 |
| (0.6;0.9] | 0,75 | 1 | 39 | 0,025 | 0,975 |
| (0.9;1.2] | 1,05 | 0 | 39 | 0,000 | 0,975 |
| (1.2;1.5] | 1,35 | 1 | 40 | 0,025 | 1,000 |
| Total | - | 40 | – | 1,000 | - |

# Representación Gráfica

- El tipo de gráfico utilizado debe ajustarse al tipo de variable estadística en estudio.

# Datos Binarios

X: “Sexo de los insectos”

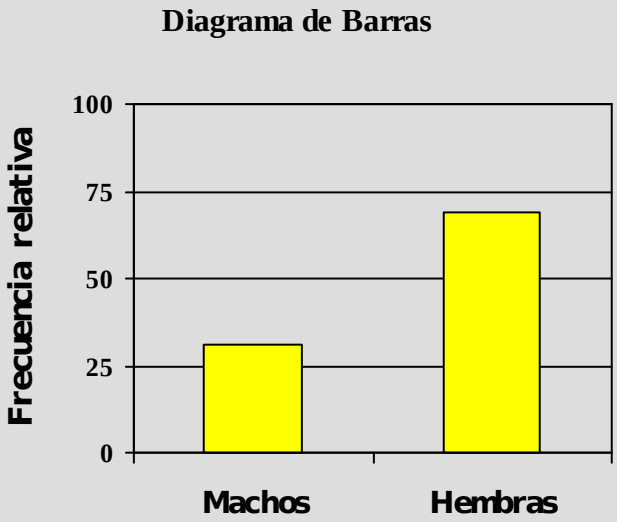



Diagrama de Barras

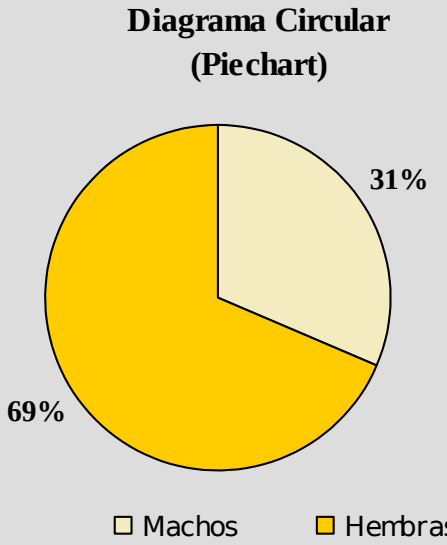



Diagrama de Sectores

# Datos Nominales

X: “color de la corola”

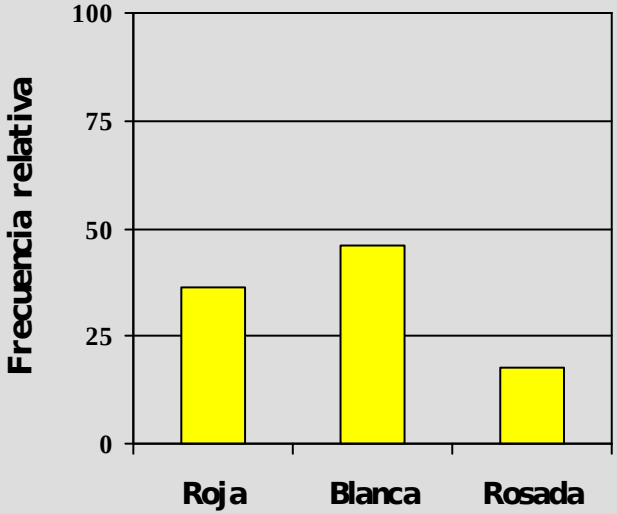



Diagrama de Barras

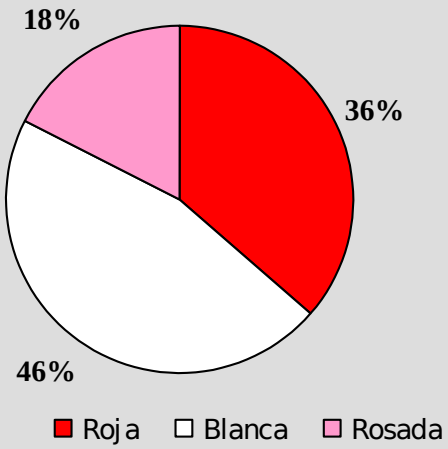
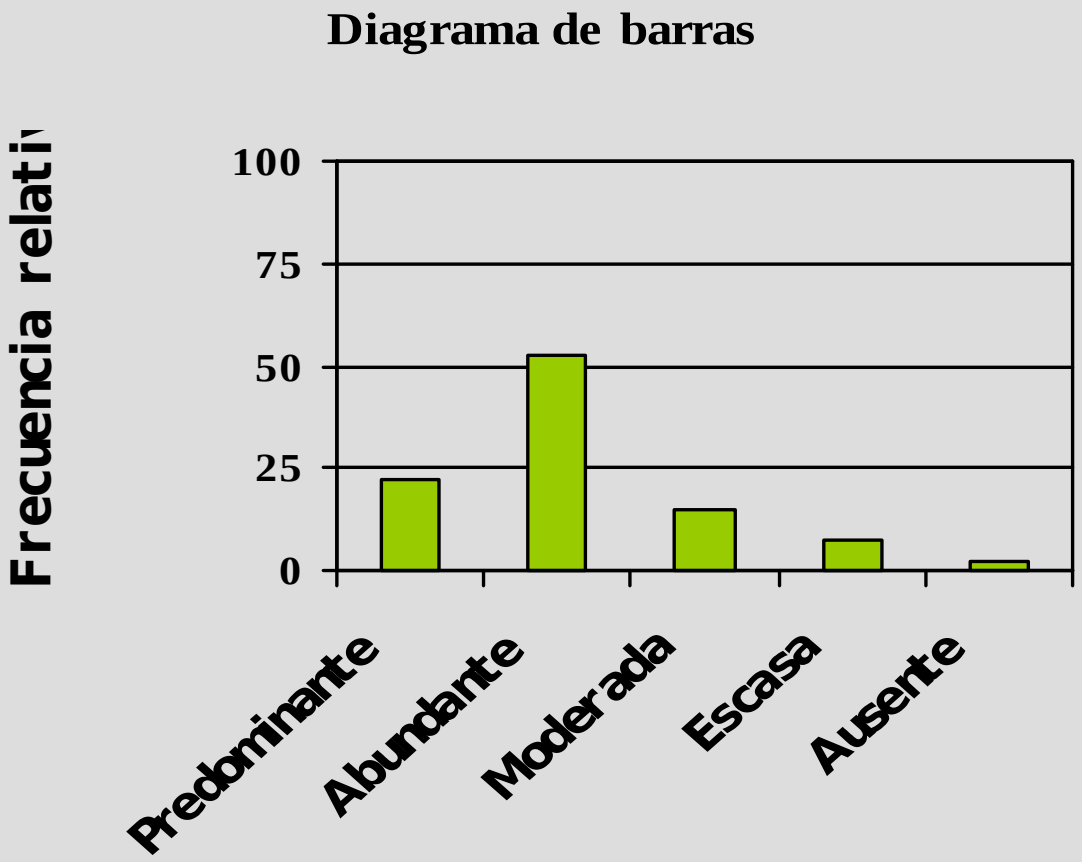



Diagrama de Sectores

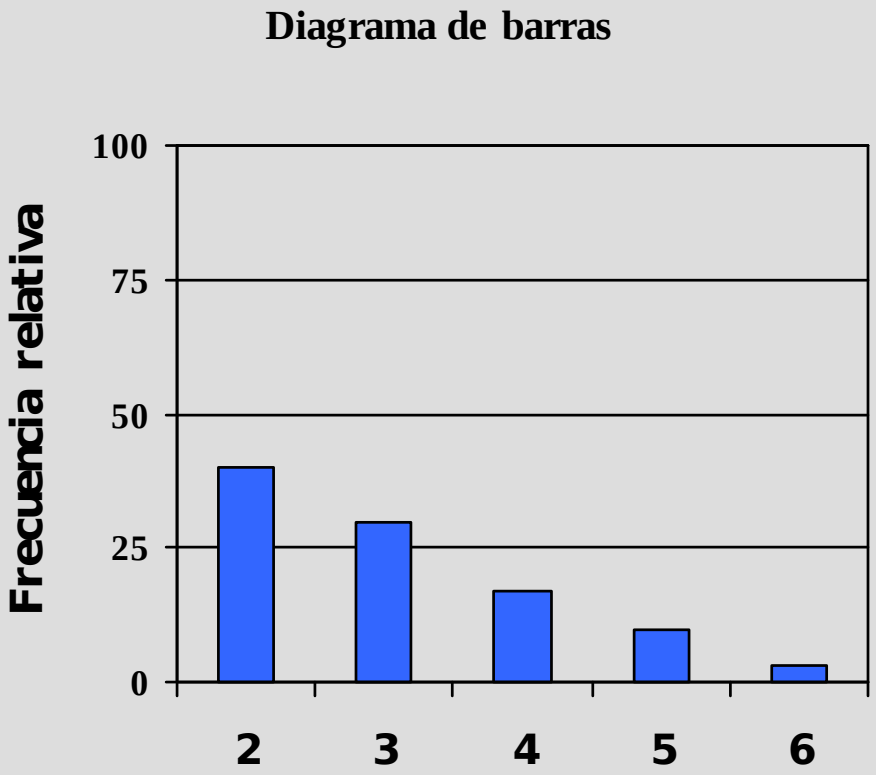
# Datos Ordinales

X: “Presencia de una especie de gramínea en distintos sitios”

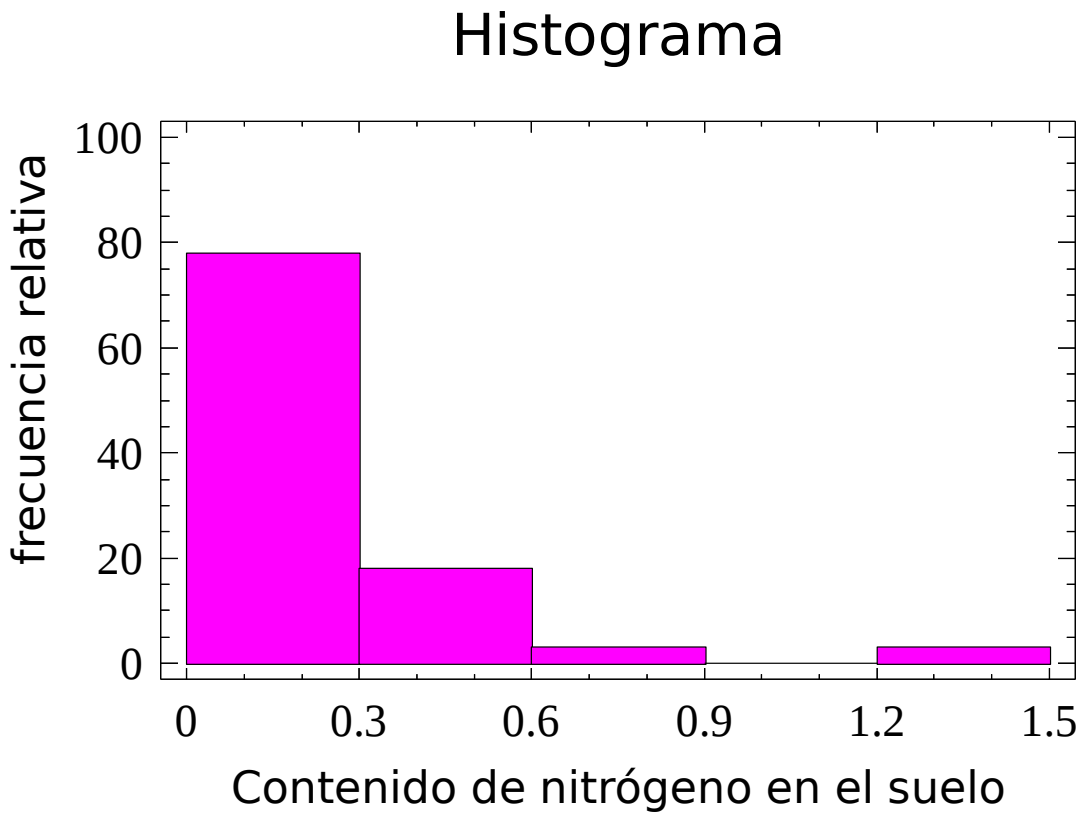


# Datos Numéricos Discretos

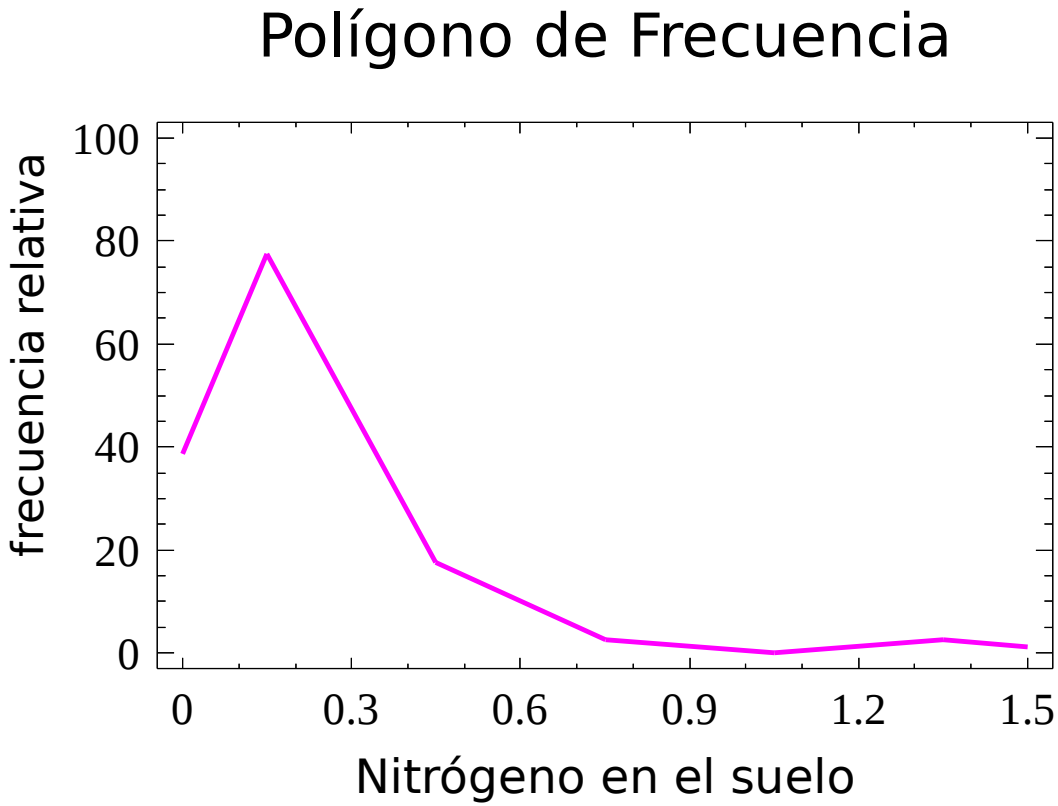
X: “Recuento de células en microscopio”



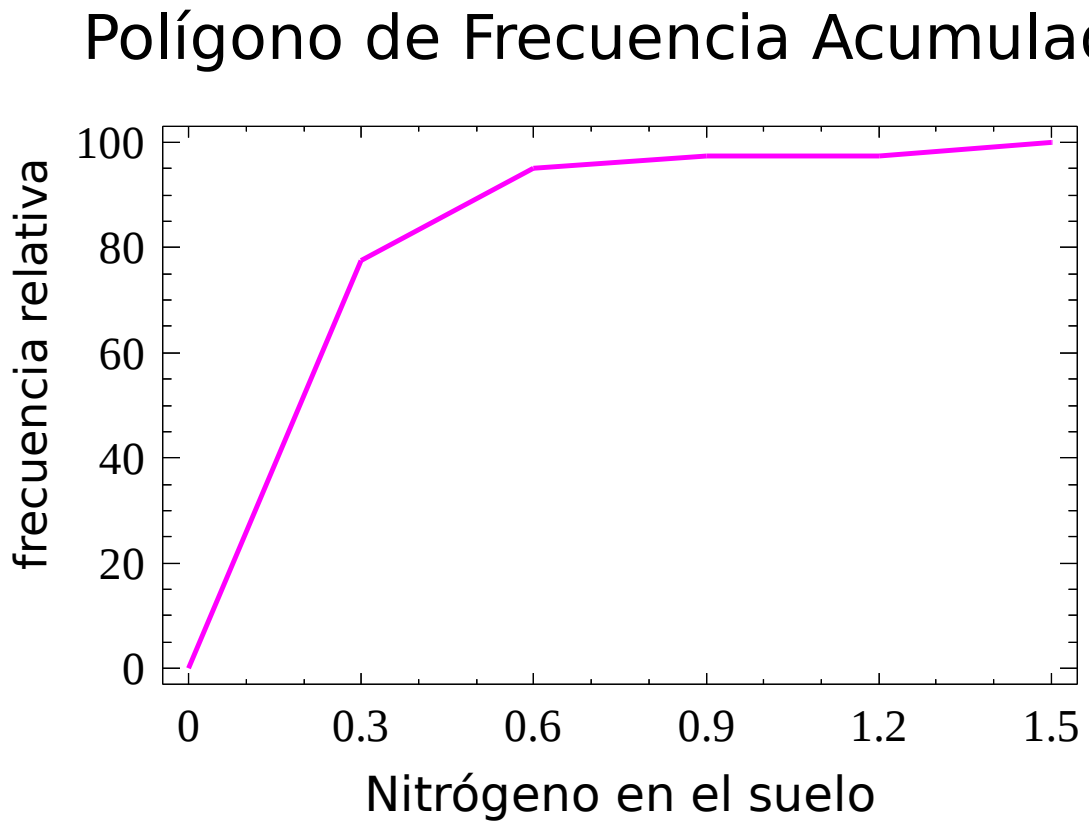
# Datos Numéricos Continuos



# Datos Numéricos Continuos



# Datos Numéricos Continuos

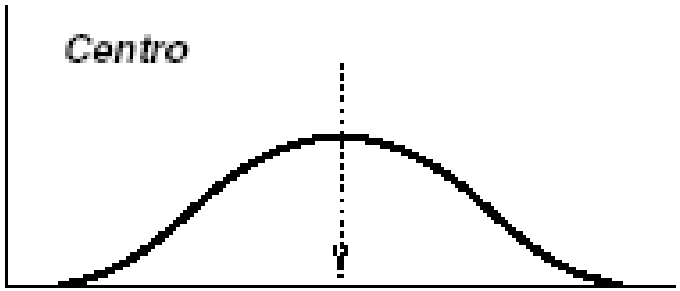


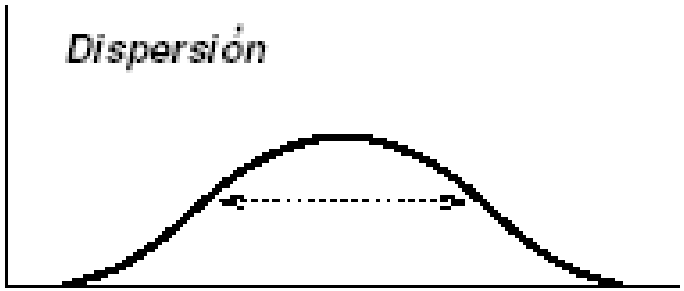
# En síntesis:

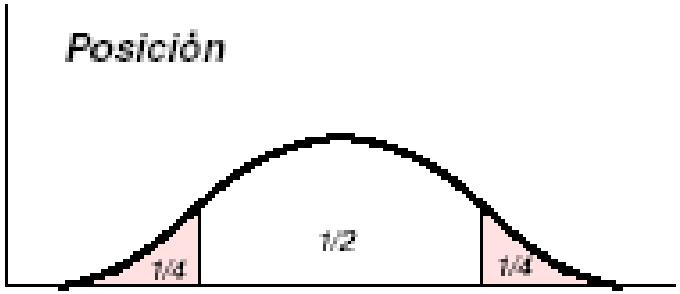
| Tipo de variable | Diagrama |
|---|---|
| V. Cualitativa | Barras, sectores, pictogramas |
| V. Discreta | Diferencial (barras)<br>Integral (en escalera) |
| V. Continua | Diferencial (histograma, polígono de frecuencias)<br>Integral (diagramas acumulados) |

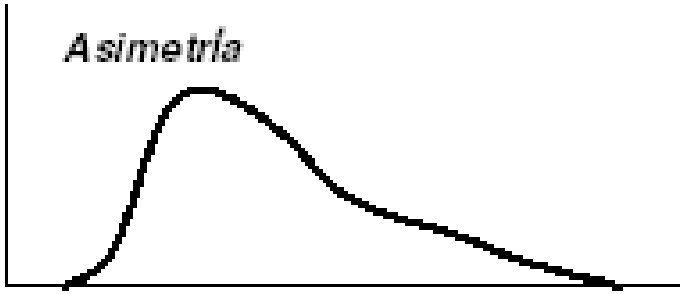
# Medidas de Resumen

- Tendencia Central
- Dispersión = Variación
- Posición
- Asimetría = Sesgo
- Curtosis = Apuntamiento


Centro


Dispersión


Posición
1/4
1/2
1/4


Asimetría

# Tendencia Central

Para un conjunto de ***n*** datos organizados en ***k*** clases se definen:

- **Media aritmética**: $\bar{x} = \sum_{i=1}^{k} x_i f_i = \frac{1}{n} \sum_{i=1}^{k} x_i n_i$

- **Mediana**: dato que ocupa la posición central en el conjunto ordenado de los datos.

- **Moda**: el valor con la mayor frecuencia.

# Dispersión

- **Rango**

$$X_{max} - X_{min}$$

- **Varianza**:

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

- **Desviación Estándar** = **Desviación Típica**:

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

# Dispersión

- **Error estándar**

$$ee = \frac{s}{\sqrt{n}}$$

- **Coeficiente de Variación**

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

# Medidas de Posición (Estadísticas de Orden)

- **Percentilo**: el valor que deja por debajo de sí el k% de los datos.

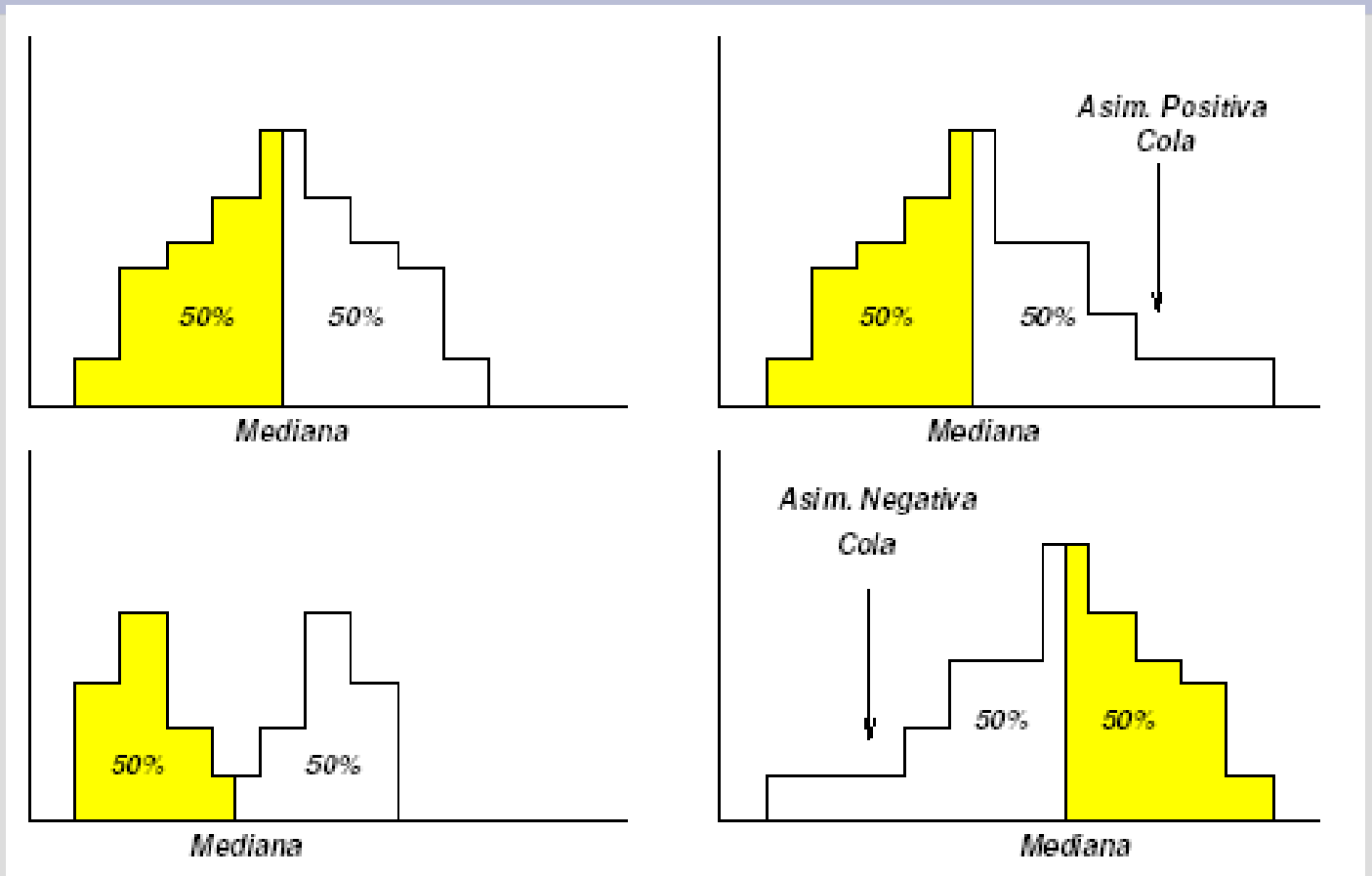
$$P_k = l_{i-1} + a_i \frac{\frac{nk}{100} - N_{i-1}}{n_i}$$

$$P_k \in \left(l_{i-1}; l_i\right]$$

- **Mediana**: el valor que deja por debajo de sí el 50% de los datos.

- **Cuartiles**: dividen a la muestra en cuatro partes con igual proporción de observaciones (25% c/u).

# Asimetría


50%
50%
Mediana
Asim. Positiva
Cola
50%
50%
Mediana
50%
50%
Mediana
Asim. Negativa
Cola
50%
50%
Mediana

# Asimetría

Momento Ordinario de orden *p*

$$\mu_p = \frac{1}{n}\sum_{i=1}^{n} x_i^p$$

Momento Centrado de orden *p*

$$m_p = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^p$$

# Asimetría

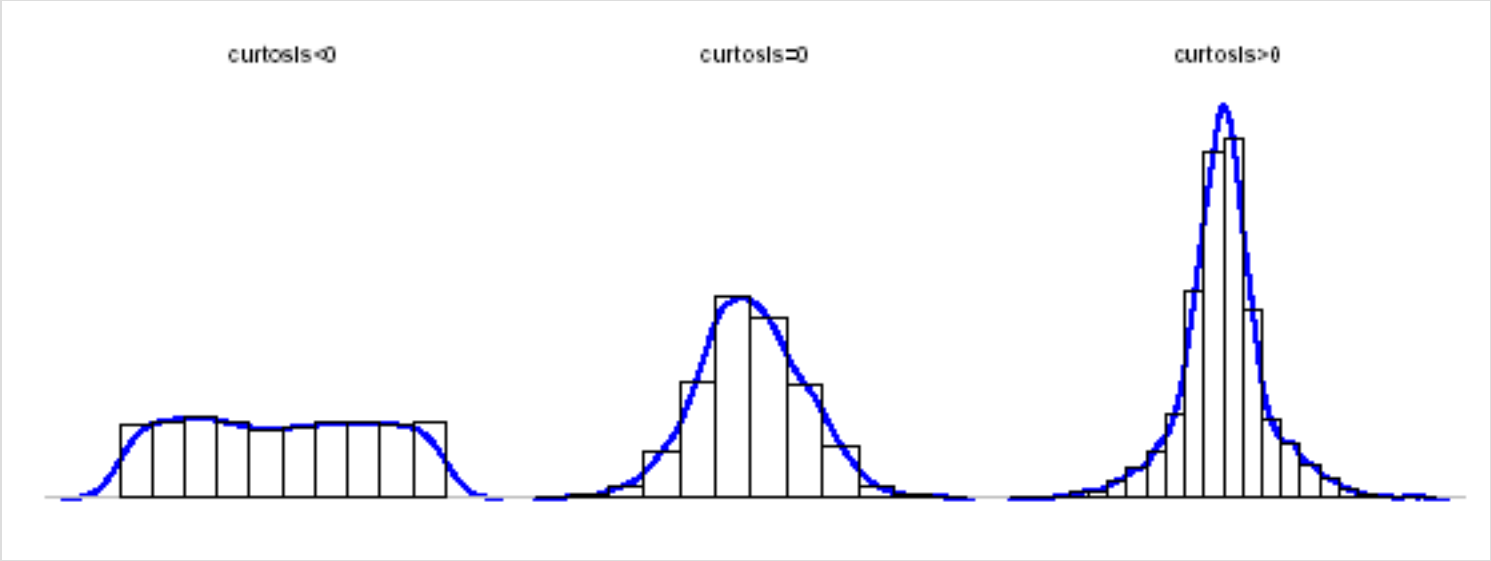
Índice de Asimetría basado en el Momento Central de tercer orden

$$m_3 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^3$$

Índice de Yule-Bowley

$$As = \frac{\left(Q_3 - Q_2\right) - \left(Q_2 - Q_1\right)}{Q_3 - Q_1}$$

# Curtosis




Platicúrtica Mesocúrtica Leptocúrtica

# Curtosis

$$\gamma_2 = \frac{m_4}{S^4} - 3$$

(tomando como patrón de referencia a la Distribución Normal o Distribución de Gauss)

**Leptocúrtica:** Cuando $\gamma_2 > 0$, o sea, si la distribución de frecuencias es más apuntada que la normal;

**Mesocúrtica:** Cuando $\gamma_2 = 0$, es decir, cuando la distribución de frecuencias es tan apuntada como la normal;

**Platicúrtica:** Cuando $\gamma_2 < 0$, o sea, si la distribución de frecuencias es menos apuntada que la normal;

# Nitrógeno en el suelo:

Count = 40

Average = 0.25175

Median = 0.175

Variance = 0.0544353

Standard deviation = 0.233314

Minimum = 0.08

Maximum = 1.43

Range = 1.35

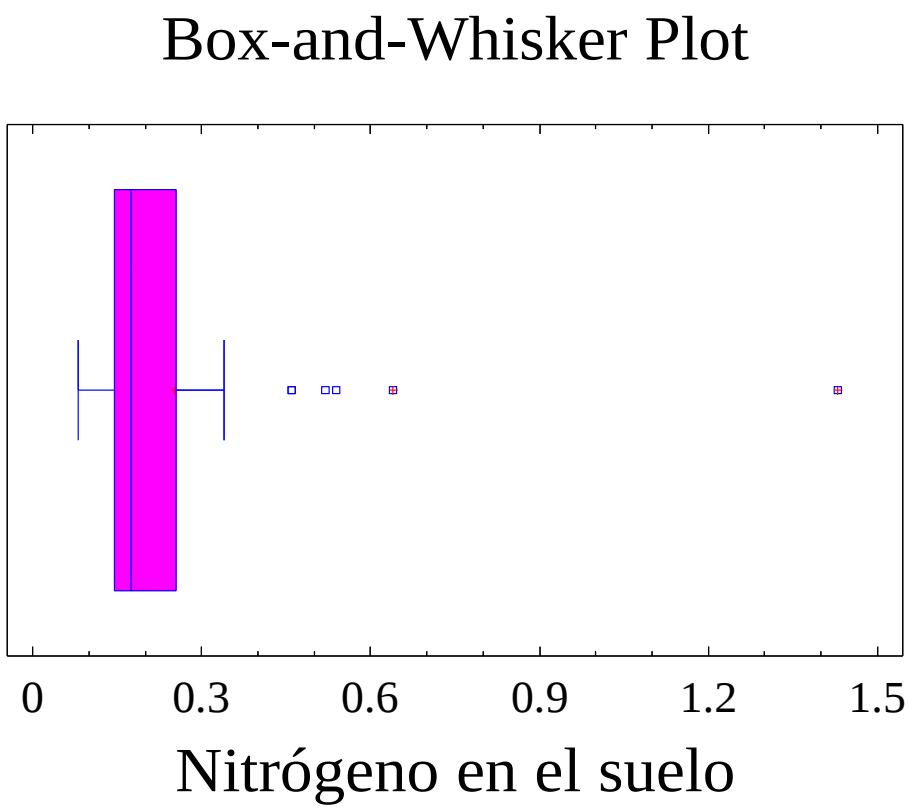
Lower quartile = 0.145

Upper quartile = 0.255

Stnd. skewness = 9.42623

Stnd. kurtosis = 21.4986

# Nitrógeno en el suelo:



# Relaciones entre la media y la desviación estándar

Distribución Normal:

| Intervalo | Porcentaje de datos incluidos el Intervalo |
|---|---|
| $\left[\bar{x}-s\,;\,\bar{x}+s\right]$ | 68% |
| $\left[\bar{x}-2s\,;\,\bar{x}+2s\right]$ | 95% |
| $\left[\bar{x}-3s\,;\,\bar{x}+3s\right]$ | 99% |

# Relaciones entre la media y la desviación estándar

Desigualdad de Chebyshev:

$$\left| x - \bar{x} \right| \leq k \; s$$

$$\downarrow$$

$$100\left( 1 - \frac{1}{k^2} \right)\%$$

# Variables Bidimensionales

# Ejemplo 6

Sobre 22 muestras de agua de pozo tomadas en el departamento de Lavalle se midieron las siguientes variables

X:=contenido de catión sodio

Y:=contenido de anión sulfato

Los datos son los siguientes:

| obs. | Y(SO4) | X(Na) | obs. | Y(SO4) | X(Na) |
|---|---|---|---|---|---|
| 1 | 384.00 | 103.50 | 12 | 268.80 | 161.00 |
| 2 | 3880.80 | 2530.00 | 13 | 974.40 | 632.50 |
| 3 | 2356.80 | 920.00 | 14 | 662.40 | 15.64 |
| 4 | 643.20 | 345.00 | 15 | 1488.00 | 805.00 |
| 5 | 1972.80 | 517.50 | 16 | 278.40 | 184.00 |
| 6 | 744.00 | 230.00 | 17 | 2059.20 | 230.00 |
| 7 | 998.40 | 402.50 | 18 | 984.00 | 460.00 |
| 8 | 6000.00 | 3220.00 | 19 | 1656.00 | 287.50 |
| 9 | 3840.00 | 1840.00 | 20 | 264.00 | 57.50 |
| 10 | 326.00 | 230.00 | 21 | 1593.60 | 575.00 |
| 11 | 825.60 | 402.50 | 22 | 484.80 | 276.00 |

Se tienen las siguientes medidas de resumen:

| Medida | Y(SO4) | X(Na) |
|---|---|---|
| Media | 1485.69 | 655.69 |
| Desviación Estándar | 1449.78 | 823.34 |
| Min | 264.00 | 15.64 |
| Q1 | 524.40 | 230.00 |
| Mediana | 979.20 | 373.75 |
| Q3 | 1893.60 | 618.13 |
| Max | 6000.00 | 3220.00 |
| IQR | 1369.20 | 388.13 |

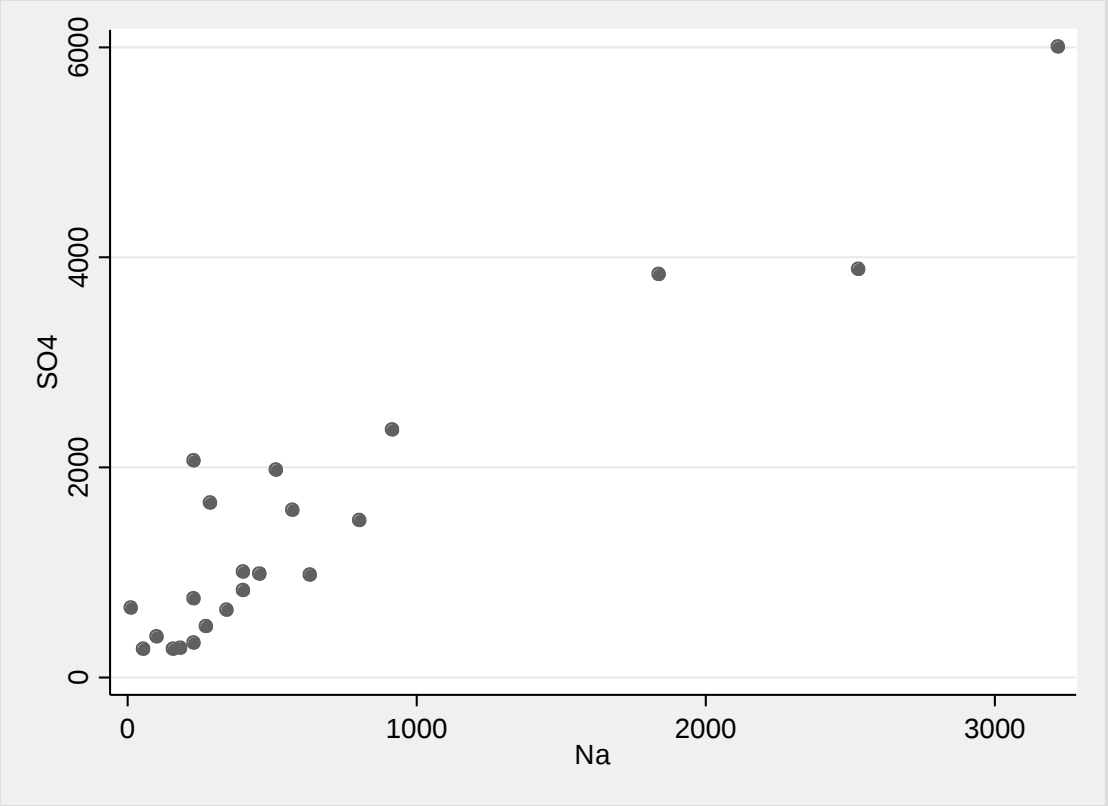
Ambas variables constituyen un “vector” o variable bidimensional:

$$(X,Y)$$

Cada individuo está caracterizado por un “par ordenado” de observaciones.

$$\text{Individuo } i \rightarrow (x_i, y_i)$$

Un gráfico adecuado es el “diagrama de dispersión” (scatterplot)



Puede medirse la variación conjunta de ambas variables mediante la "covarianza":

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

Con el inconveniente de que sus unidades son el producto de las de $X$ por las de $Y$.

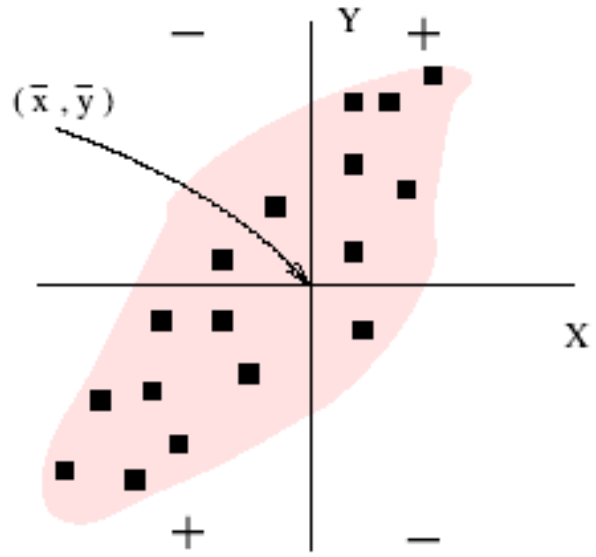
$S_{xy}$ puede tomar cualquier valor real

Si $S_{xy} = 0$ las variables X e Y son *no correlacionadas.*

La covarianza mide la *relación lineal* entre ambas variables.

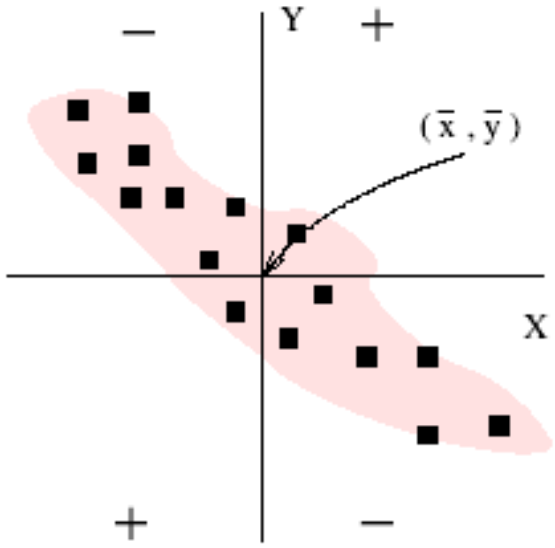
$S_{xy} > 0$ indica relación lineal directa

$S_{xy} < 0$ indica relación lineal inversa



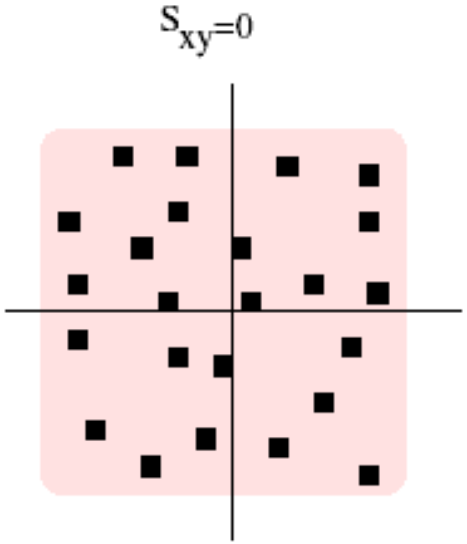

*Cuando X crece, Y crece*

*Casi todos los puntos pertenecen a los cuadrantes primero y tercero*

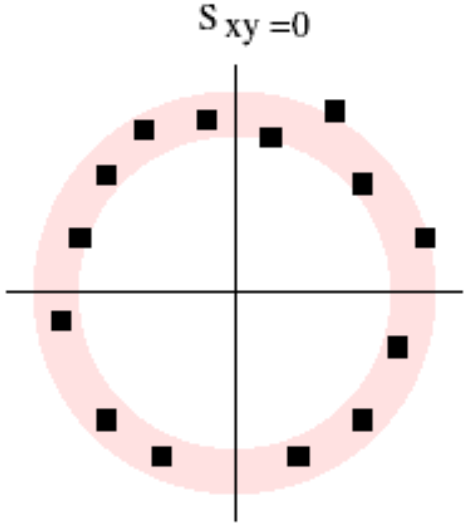



*Cuando X crece, Y decrece*

*Casi todos los puntos pertenecen a los cuadrantes segundo y cuarto*




*Las dos variables son independientes.*




*Hay dependencia entre las dos variables, aunque la covarianza sea nula.*

El coeficiente de correlación lineal de Pearson soluciona el problema de las unidades de medida mediante la estandarización de las variables

$$R = \frac{S_{XY}}{S_X S_Y}$$

$R$ puede tomar valores en $[0;1]$

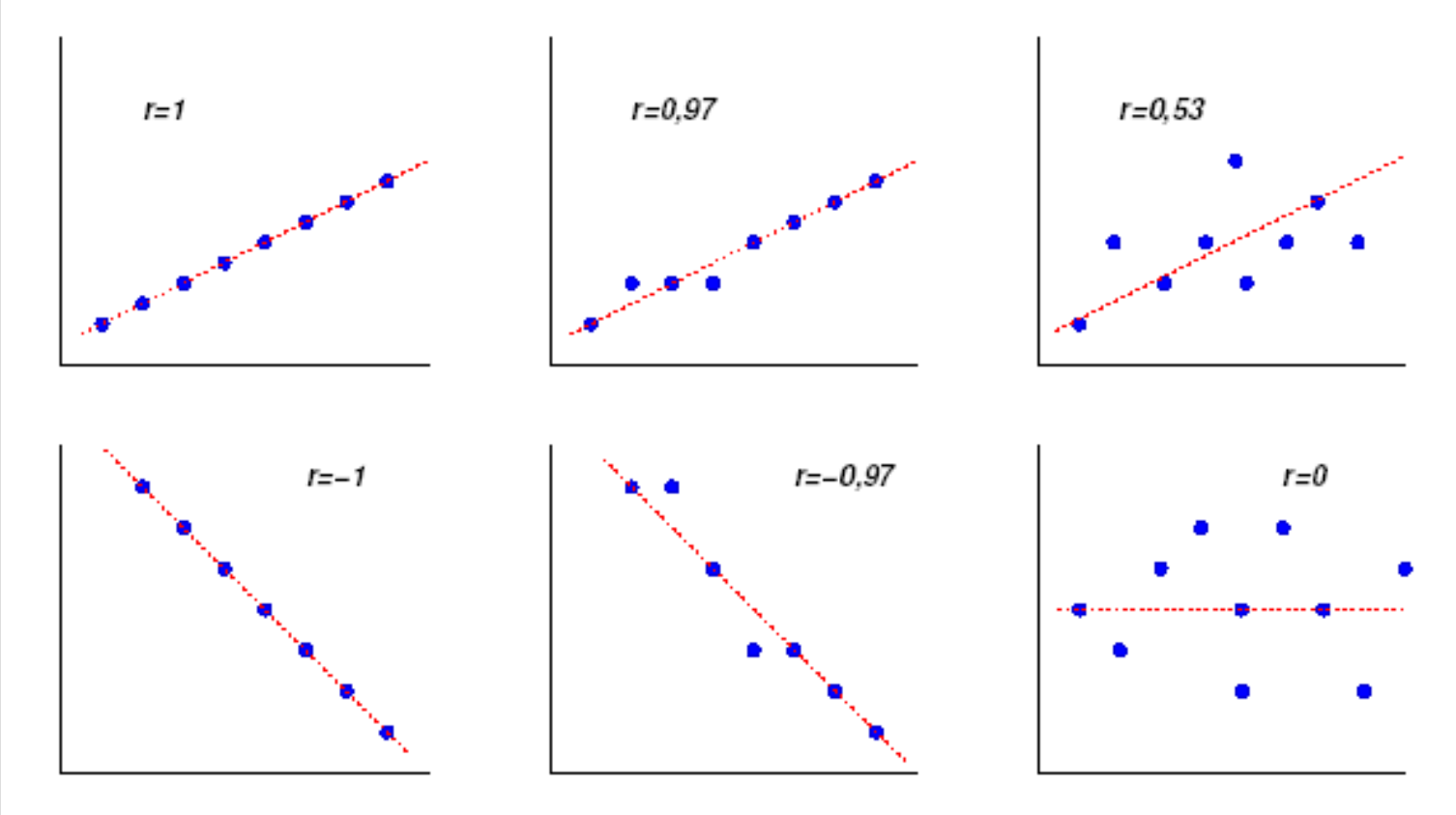
Si $R = 0$ las variables X e Y son *no correlacionadas.*

El coef. de correlación lineal de Pearson mide la *relación lineal* entre ambas variables.

$0 < R \leq 1$ indica relación lineal directa

$-1 \leq R < 0$ indica relación lineal inversa

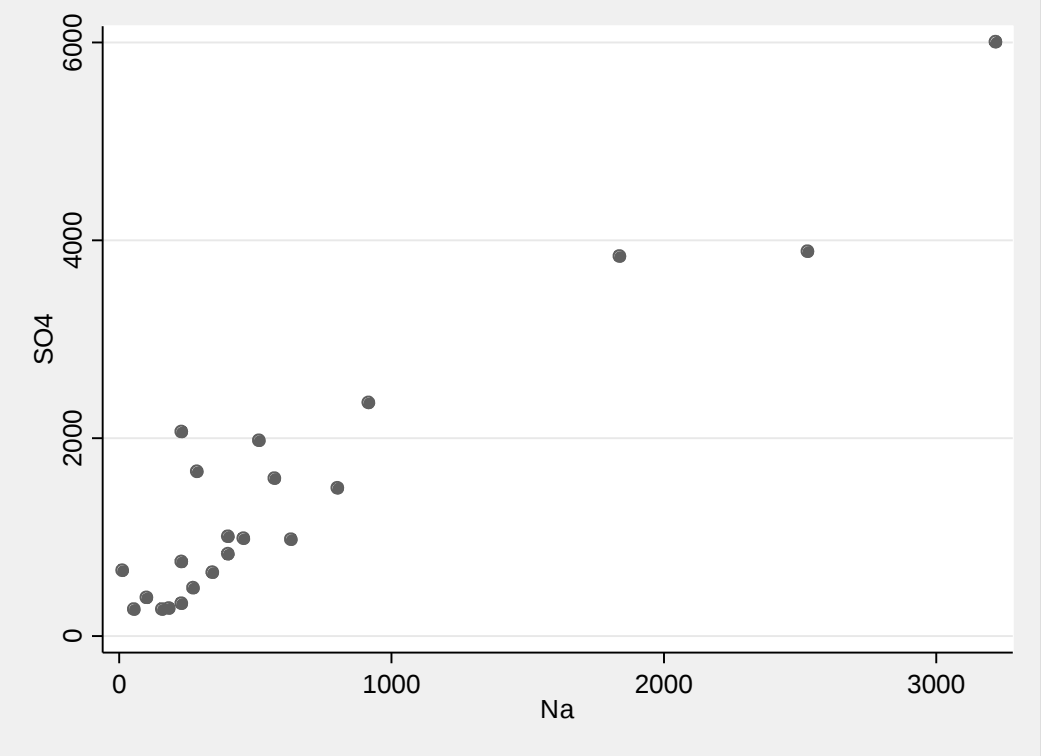
# Relación Lineal



Para el ejemplo se tiene

$$S_{XY} = 1.1 \times 10^{06}$$

$$R = 0.9411$$



# Relación Funcional

Y := variable dependiente

X := variable independiente

$$Y = f(X)$$

Consideremos el caso de *f( )* lineal

# Función Lineal

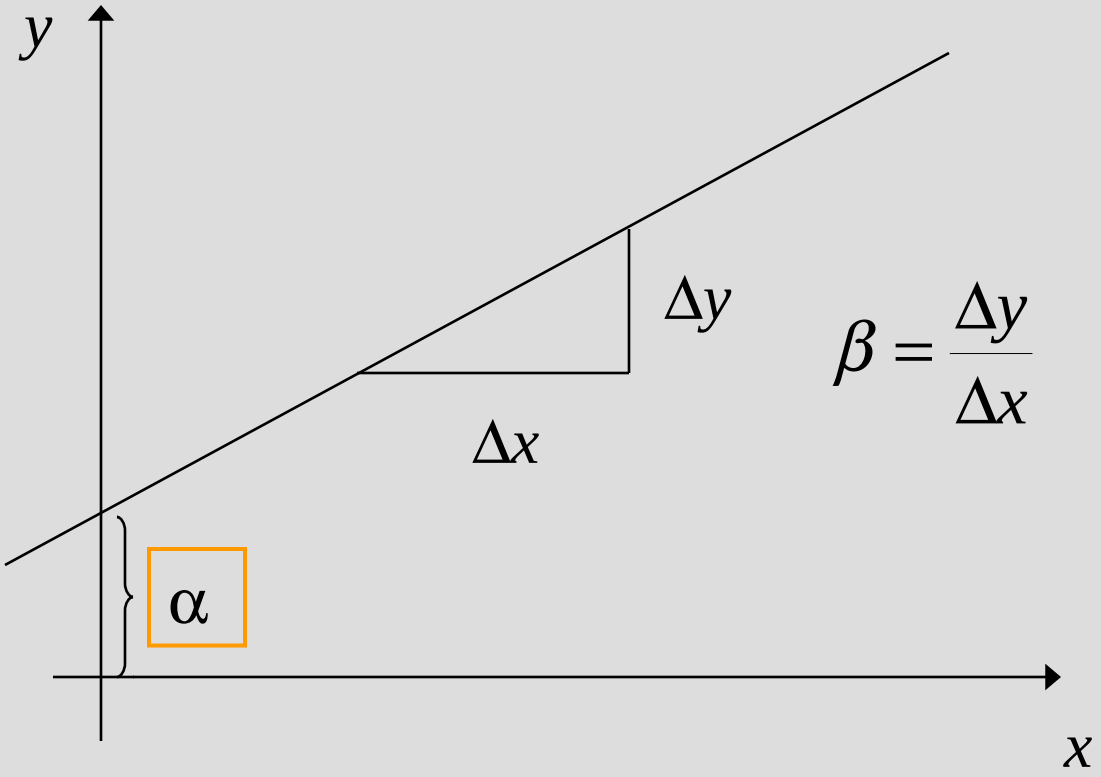
Ecuación de la línea recta

$$Y = \alpha + \beta X$$

$\alpha :=$ ordenada al origen (intercepto)

$\beta :=$ pediente

# Función Lineal

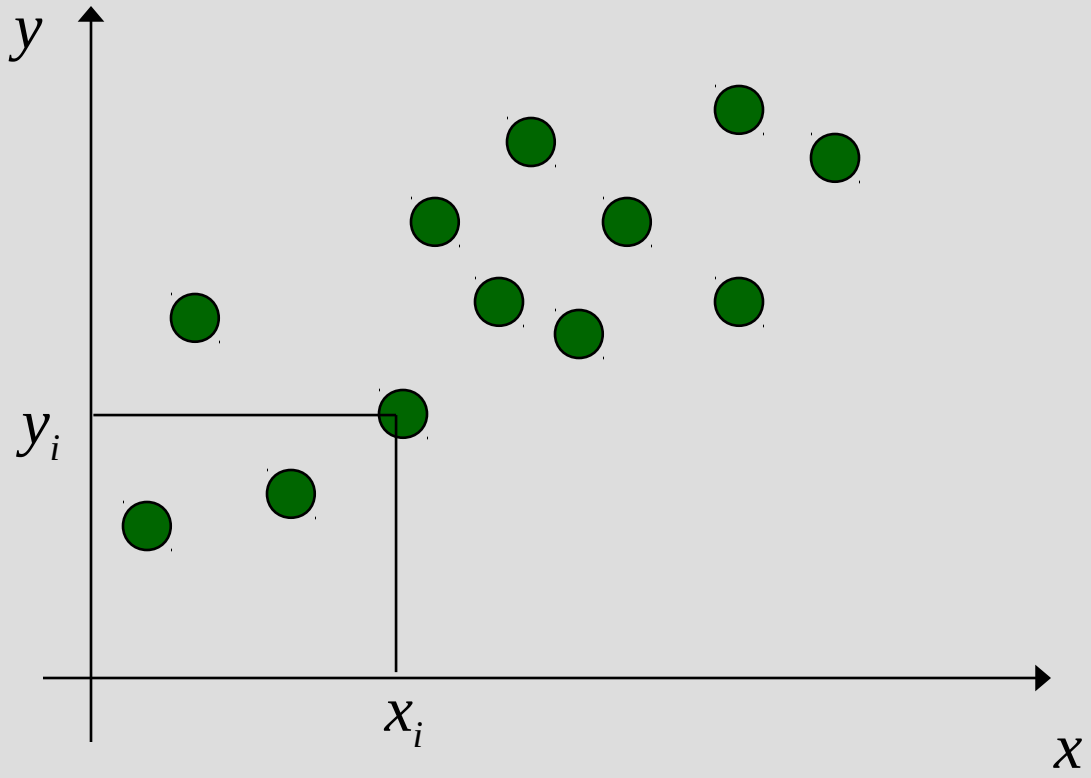



$$\beta = \frac{\Delta y}{\Delta x}$$

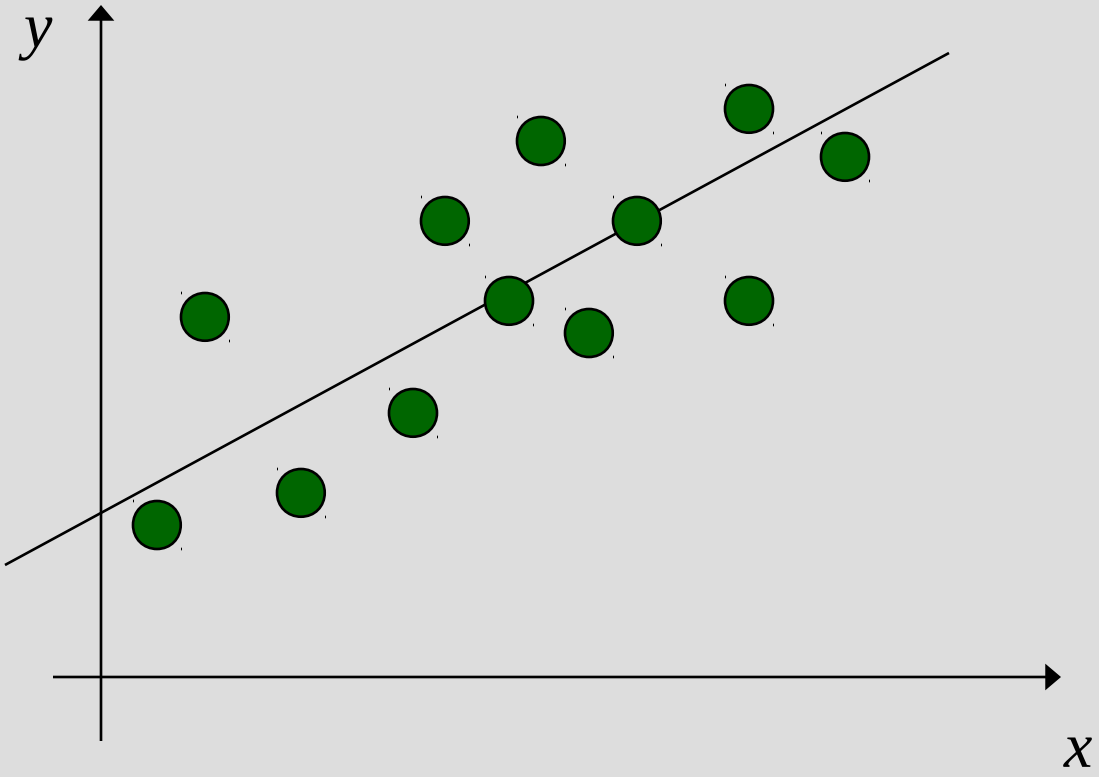
# Datos

| Observación | X | Y |
|---|---|---|
| 1 | $x_1$ | $y_1$ |
| 2 | $x_2$ | $y_2$ |
| ... | ... | ... |
| i | $x_i$ | $Y_i$ |
| ... | ... | ... |
| n | $x_n$ | $y_n$ |

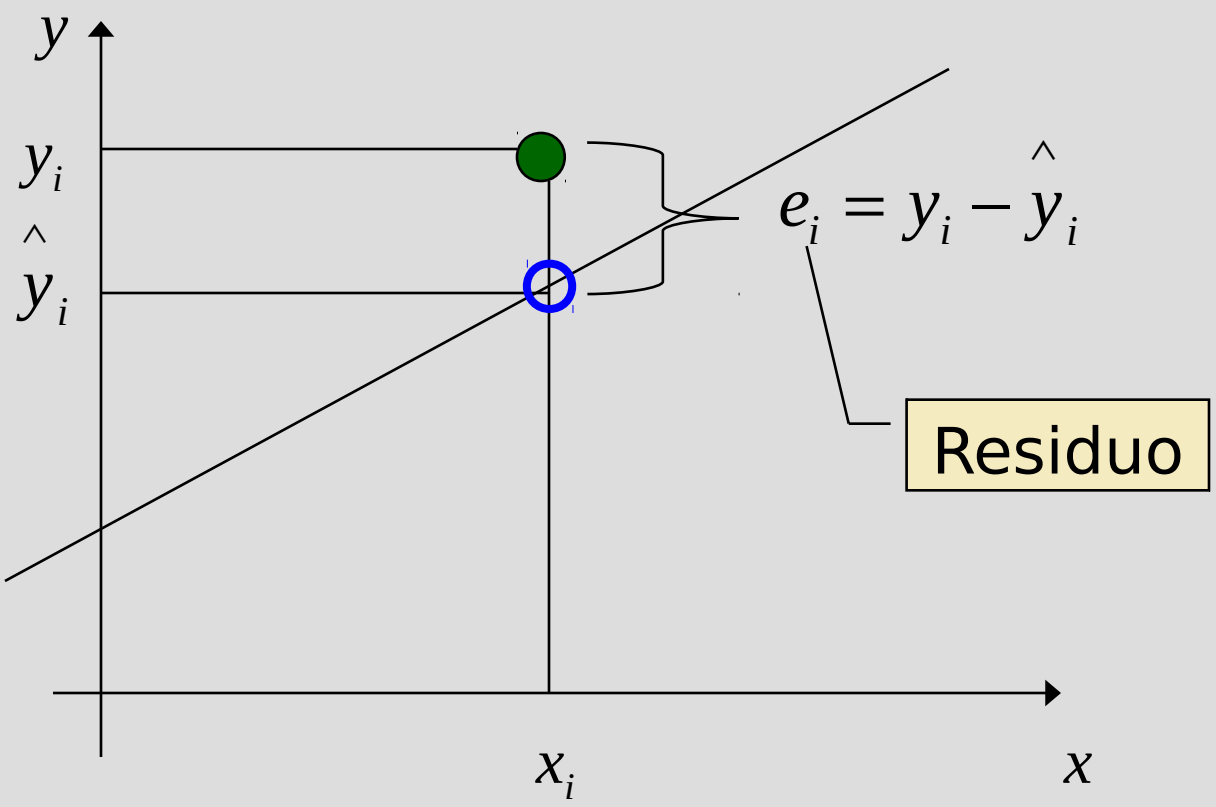
# Datos



# Recta de Ajuste



# Para un punto particular



# Residuos

Se define el "residuo" para la i-ésima observación como:

$$e_i = y_i - \hat{y}_i$$

Para puntos por encima de la recta $e_i > 0$ y por debajo $e_i < 0$.

# Criterio de Mínimos Cuadrados

Se quiere encontrar $\alpha$ y $\beta$ tales que se *minimice* la suma de cuadrados de residuos para los $n$ puntos:

$$Q = \sum_{i=1}^{n} e_i^{\,2} = \sum_{i=1}^{n} \left( y_i - \hat{y}_i \right)^2$$

# Criterio de Mínimos Cuadrados

Dado que $\hat{y}_i = \alpha + \beta x_i$ resulta que se debe **minimizar**

$$Q = \sum_{i=1}^{n} \left( y_i - \alpha - \beta x_i \right)^2$$

respecto a $\alpha$ y $\beta$.

# Criterio de Mínimos Cuadrados

Resolviendo se obtienen los **estimadores de mínimos cuadrados** cuyas formas son:

$$\hat{\beta} = \frac{\sum_{i=1}^{n} y_i \left( x_i - \bar{x} \right)}{\sum_{i=1}^{n} \left( x_i - \bar{x} \right)^2} = \frac{S_{XY}}{S_{XX}}$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\,\bar{x}$$

# Bondad de Ajuste

La bondad del ajuste se evalúa por el *coeficiente de determinación*

$$R^2 = \frac{\text{Variabilidad explicada por el modelo}}{\text{Variabilidad Total}}$$

# Bondad de Ajuste

La variabilidad total $\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2$ se descompon

en la variabilidad explicada por el modelo

$$\sum_{i=1}^{n}\left(\hat{y}_i-\bar{y}\right)^2$$

y la variabilidad residual

$$\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2$$

# Bondad de Ajuste

De modo que:

$$\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2 = \sum_{i=1}^{n}\left(\hat{y}_i-\bar{y}\right)^2 + \sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2$$

$$SC(Total) = SC(Reg) + SC(Res)$$

# Bondad de Ajuste

Así

$$R^2 = \frac{\sum_{i=1}^{n}\left(\hat{y}_i - \bar{y}\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2} = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}$$

# Bondad de Ajuste

Así

$$R^2 = \frac{SC(Regr)}{SC(Total)} = 1 - \frac{SC(Res)}{SC(Total)}$$

# Bondad de Ajuste

El coeficiente de determinación contiene la propiedad de tomar valores entre 0 y 1.

$$0 \leq R^2 \leq 1$$

$0 \rightarrow$ Ausencia de Ajuste Lineal

$1 \rightarrow$ Ajuste Lineal Perfecto

# Bondad de Ajuste

En el caso del modelo lineal el coeficiente de determinación, $R^2$, coincide con el cuadrado del coeficiente de correlación lineal de Pearson, $R$.

En el caso de ajuste no lineal esto no es así !!!