

# Formulas de Estadística

## Estadística Descriptiva

### Frecuencias

**Tamaño muestral**  $n$  número de individuos en la muestra.

**Frecuencia Absoluta**  $n_i$  (nº de  $x_i$  en la muestra)

**Frecuencia Relativa**  $f_i = n_i/n$

**Frec. Absoluta Acumulada**  $N_i = \sum_{k=0}^i n_k$

**Frec. Relativa Acumulada**  $F_i = N_i/n$

### Estadísticos de tendencia central

**Media**  $\bar{x} = \frac{\sum x_i n_i}{n}$

**Mediana**  $me$  El valor con frec. rel. acumulada  $F_{me} = 0.5$ .

**Moda**  $mo$  El valor más frecuente.

### Estadísticos de posición

**Cuartiles**  $Q_1, Q_2, Q_3$  dividen la distribución en 4 partes iguales. Sus frec. rel. acumuladas son  $F_{Q_1} = 0.25$ ,  $F_{Q_2} = 0.5$  and  $F_{Q_3} = 0.75$ .

**Percentiles**  $P_1, P_2, \dots, P_{99}$  dividen la distribución en 100 partes iguales. Su frec. rel. acumulada es  $F_{P_i} = i/100$ .

### Estadísticos de dispersión

**Rango intercuartílico**  $IQR = Q_3 - Q_1$

**Varianza**  $s^2 = \frac{\sum (x_i - \bar{x})^2 n_i}{n} = \frac{\sum x_i^2 n_i}{n} - \bar{x}^2$

**Desviación típica**  $s = +\sqrt{s^2}$

**Coefficiente de variación**  $cv = \frac{s}{|\bar{x}|}$

### Estadísticos de forma

**Coefficiente de asimetría**  $g_1 = \frac{\sum (x_i - \bar{x})^3 f_i}{s^3}$

**Coefficiente de apuntamiento**  $g_2 = \frac{\sum (x_i - \bar{x})^4 f_i}{s^4} - 3$

### Transformaciones lineales

**Transformación lineal**  $y = a + bx$

$$\bar{y} = a + b\bar{x}$$

$$s_y = bs_x$$

**Tipificación**  $z = \frac{x - \bar{x}}{s_x}$

## Regresión y correlación

### Regresión lineal

**Covarianza**  $s_{xy} = \frac{\sum x_i y_i n_{ij}}{n} - \bar{x}\bar{y}$

**Rectas de regresión :**

$$y \text{ on } x : y = \bar{y} + \frac{s_{xy}}{s_x^2}(x - \bar{x})$$

$$x \text{ on } y : x = \bar{x} + \frac{s_{xy}}{s_y^2}(y - \bar{y})$$

**Coefficientes de regresión**

$$(y \text{ on } x) b_{yx} = \frac{s_{xy}}{s_x^2} \quad (x \text{ on } y) b_{xy} = \frac{s_{xy}}{s_y^2}$$

**Coefficiente de determinación**

$$r^2 = \frac{s_{xy}^2}{s_x^2 s_y^2} \quad 0 \leq r^2 \leq 1$$

**Coefficiente de correlación**

$$r = \frac{s_{xy}}{s_x s_y} \quad -1 \leq r \leq 1$$

### Regresión no lineal

**Modelo exponencial**  $y = e^{a+bx}$

Aplicar el logaritmo a la variable dependiente y calcular la recta de regresión  $\log y = a + bx$ .

**Modelo logarítmico**  $y = a + b \log x$

Aplicar el logaritmo a la variable independiente y calcular la recta de regresión  $y = a + b \log x$ .

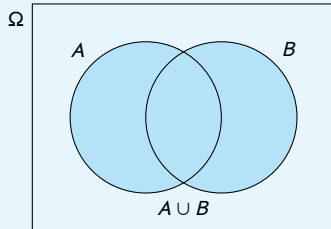
**Modelo potencial**  $y = ax^b$

Aplicar el logaritmo a ambas variables y calcular la recta de regresión  $\log y = a + b \log x$ .

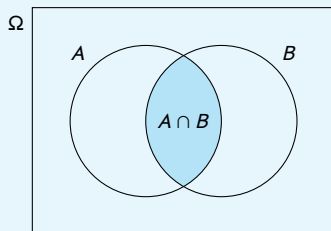
## Probabilidad

### Operaciones de sucesos

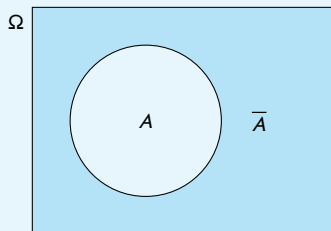
**Unión**



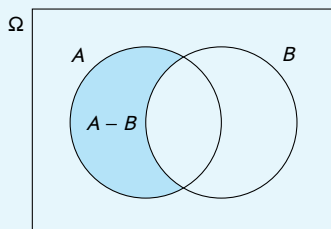
**Intersección**



**Contrario**



**Diferencia**



### Álgebra de sucesos

**Idempotencia**  $A \cup A = A$ ,  $A \cap A = A$

**Conmutativa**  $A \cup B = B \cup A$ ,  $A \cap B = B \cap A$

**Asociativa**  $(A \cup B) \cup C = A \cup (B \cup C)$ ,  $(A \cap B) \cap C = A \cap (B \cap C)$

**Distributiva**  $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$ ,  $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$

**Elemento neutro**  $A \cup \emptyset = A$ ,  $A \cap \Omega = A$

**Elemento absorbente**  $A \cup \Omega = \Omega$ ,  $A \cap \emptyset = \emptyset$

**Elemento simétrico complementario**  $A \cup \bar{A} = \Omega$ ,  $A \cap \bar{A} = \emptyset$

**Doble contrario**  $\bar{\bar{A}} = A$

**Leyes de Morgan**  $\overline{A \cup B} = \bar{A} \cap \bar{B}$ ,  $\overline{A \cap B} = \bar{A} \cup \bar{B}$

### Probabilidad básica

**Unión**  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

**Intersección**  $P(A \cap B) = P(A)P(B|A)$

**Diferencia**  $P(A - B) = P(A) - P(A \cap B)$

**Contrario**  $P(\bar{A}) = 1 - P(A)$

### Probabilidad condicionada

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

**Sucesos independientes**  $P(A|B) = P(A)$

**Teorema de la probabilidad total**

$$P(B) = \sum_{i=1}^n P(A_i)P(B|A_i)$$

**Teorema de Bayes**

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^n P(A_i)P(B|A_i)}$$

### Test diagnósticos

**Prevalencia** Proporción de individuos con la enfermedad  $P(E)$

**Sensibilidad**  $P(+|E)$

**Especificidad**  $P(-|\bar{E})$

**Valor predictivo positivo (VPP)**  $P(E|+)$

**Valor predictivo negativo (VPN)**  $P(\bar{E}|-)$

**Razón de verosimilitud positiva (RV+)**  $\frac{P(+|E)}{P(+|\bar{E})}$

**Razón de verosimilitud negativa (RV-)**  $\frac{P(-|E)}{P(-|\bar{E})}$

## Variables Aleatorias

### Discretas

**Función de probabilidad Binomial**  $B(n, p)$

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

**Función de probabilidad Poisson**  $P(\lambda)$

$$f(x) = e^{-\lambda} \frac{\lambda^x}{x!}$$

**Ley de los casos raros**  $B(n, p) \approx P(np)$  para  $n \geq 30$  y  $p \leq 0.1$ .

## Continuas

**Normal**  $N(\mu, \sigma)$

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

**Normal Estándar**  $N(0, 1)$

**Chi-cuadrado**  $\chi^2(n)$

$$X = Z_1^2 + \dots + Z_n^2,$$

donde  $Z_i \sim N(0, 1)$ .

**T de Student**  $T(n)$

$$T = \frac{Z}{\sqrt{X/n}},$$

donde  $Z \sim N(0, 1)$  y  $X \sim \chi^2(n)$ .

**F de Fisher**  $F(n, m)$

$$F = \frac{X/m}{Y/n},$$

donde  $X \sim \chi^2(m)$  y  $Y \sim \chi^2(n)$ .