

1. Datos básicos

- **Frecuencias acumuladas:** la suma de todas las frecuencias hasta un valor dado. **No puede ser mayor que 1.**
- **Moda (Mo):** valor más repetido, solo aplicable a distribuciones discretas.
- **Cuantil:** división proporcional de todos los datos. Si son pares, se tendrá que hacer la media entre ambos y usar ese valor como separación. Ejemplo: la mediana es el cuartil 2 o el decil 5.
- **Coefficiente de curtosis:** coeficiente que indica cómo de puntiaguda es una gráfica.

2. Medidas a distribuciones

2.1. Unidimensionales

- **Media:** $\bar{x} = \frac{\sum_{i=1}^N [x_i]}{N}$
- **Momentos de orden r respecto del origen:** $a_r = \frac{\sum_{i=1}^N [x_i^r]}{N} \rightarrow \bar{x} = a_1$
- **Mediana (Me):** valor de la muestra que deja un 50% de datos por debajo del mismo. Si son pares se debe hacer la media de los dos valores colindantes y tomar ese valor como mediana, aunque sea relativamente ficticio.
- **Varianza:** $s^2 = \frac{\sum_{i=1}^N [(x_i - \bar{x})^2]}{N} = a_2 - a_1^2$
- **Momentos desde la media de grado r :** $m_r = \frac{\sum_{i=1}^N [(x_i - \bar{x})^r]}{N} \rightarrow s^2 = m_2$
- **Coefficiente de variación de Pearson:** $g = \frac{s}{|\bar{X}|}$
- **Desigualdad de Chebyshev:**
 $k > 1, [\bar{x} - k \cdot s, \bar{x} + k \cdot s]$ recoge al menos $100 \cdot \left(1 - \frac{1}{k^2}\right)$ datos.
- **Coefficiente de asimetría de Fisher:** $g_1 = \frac{m_3}{s^3}$.
 $g_1 > 0$ asimetría a la derecha. $g_1 = 0$ simétrica. $g_1 < 0$ asimetría a la izquierda.
- **Cambio de variable:** $Y = a \cdot X + b$
 1. $\bar{y} = a\bar{x} + b$
 2. $s_Y^2 = a^2 \cdot s_X^2$
 3. $s_Y = |a| \cdot s_X$
 4. $Mo_Y = a \cdot Mo_X + b$
 5. $|g_{1Y}| = |g_{1X}|$
- **Tipificación:** muy importante para poder aplicar el teorema central del límite. $Y = \frac{X - \bar{x}}{s_X}$. Esta transformación centra la media en cero y hace la desviación típica 1.

2.2. Bidimensionales

- **Frecuencias absolutas de la clase z_i :**
 $n_{i.} = \sum_{j=1}^p [n_{ij}] \quad n_{.j} = \sum_{i=1}^p [n_{ij}]$
- **Momentos respecto del origen de orden (r, s) :**

$$a_{r,s} = \frac{\sum_{i=1}^N [x_i^r \cdot y_i^s]}{N}$$

- **Momentos respecto de la media de orden (r, s) :**
 $m_{r,s} = \frac{\sum_{i=1}^N [(x_i - \bar{x})^r \cdot (y_i - \bar{y})^s]}{N}$. **Covarianza:** $m_{1,1} = a_{1,1} - a_{1,0} \cdot a_{0,1}$
- **Coefficiente de correlación de Pearson:** $\rho = r = \frac{m_{1,1}}{s_x \cdot s_y}$. Se usa para independizar de la escala de medida. Indica si hay relación lineal entre las variables, dada si $|\rho| > 0.5$. El signo indica si es directo o inverso.
- **Dependencia lineal:** se ha de cumplir y es suficiente para que dos variables sean dependientes linealmente $f_{ij} = f_{i.} \cdot f_{.j}$

3. Probabilidad

- **Operaciones:** con A y B
 1. Unión (ó): $A \cup B$.
 2. Intersección (y): $A \cap B$. Si $A \cap B = \emptyset$ entonces son sucesos incompatibles.
 3. Diferencia: $A - B$.
 4. Complementario (o negado): $A \rightarrow \bar{A}$.
 5. Contenido: A contenido en $B \Rightarrow A \subset B$.
- **Propiedades:** $P[\]$ denota probabilidad
 1. $A \cup B = B \cup A$ y $A \cap B = B \cap A$
 2. $\overline{(A \cup B)} = \bar{A} \cap \bar{B}$ y $\overline{(A \cap B)} = \bar{A} \cup \bar{B}$
 3. Si $A \cap B = \emptyset \rightarrow P[A \cup B] = P[A] + P[B]$
 4. $P[A] = 1 - P[\bar{A}]$
 5. $P[A \cup B] = P[A] + P[B] - P[A \cap B]$
- **Probabilidad:** $P[A] = \frac{\text{Resultados de } A}{\text{Resultados posibles}}$
- **Combinaciones sin repetición:** $\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$
- **Probabilidad condicionada, A cuando se cumple B :** $P\left[\frac{A}{B}\right] = \frac{P[A \cap B]}{P[B]}$. Haciendo un poco de desarrollo: $P[A \cap B] = P[B] \cdot P\left[\frac{A}{B}\right] = P[A] \cdot P\left[\frac{B}{A}\right]$. También se cumple: $P\left[\frac{A}{B}\right] = 1 - P\left[\frac{\bar{A}}{B}\right]$. **Para sucesos independientes:** $P\left[\frac{A}{B}\right] = P[A]$
- **Teorema de la probabilidad total con condicionadas:** $P[B] = \sum_{i=1}^n \left(P[A_i] \cdot P\left[\frac{B}{A_i}\right]\right)$
- **Teorema de Bayes:** $P\left[\frac{A_k}{B}\right] = \frac{P[A_k \cap B]}{P[B]} = \frac{P[A_k] \cdot P\left[\frac{B}{A_k}\right]}{\sum_{i=1}^n \left(P[A_i] \cdot P\left[\frac{B}{A_i}\right]\right)}$

4. Modelos para sistemas discretos

- **Función de cuantía o probabilidad:** $P_X(x_i) = P[X = x_i]$
- **Función de distribución:** $F(x) = P[X \leq x] = \sum_{x_i \leq x} P_X(x_i)$
- **Esperanza (media):** $E(X) = \sum_{i=1}^n [x_i \cdot P_X(x_i)] = \mu_x$.
Para funciones: $E(g(X)) = \sum_{i=1}^n [g(x_i) \cdot P_X(x_i)]$

- **Varianza:** $Var(X) = \sum_{i=1}^n [(x_i - E(X))^2 \cdot P_X(x_i)] = E(X^2) - (E(X))^2$
- **Distribuciones importante:**
 1. *Bernoulli*: dos posibilidades, éxito o fracaso. A la probabilidad del éxito se la nombrará p .
 - a) Función de cuantía: $P_X(0) = P[X = 0] = 1 - p$
 $P_X(1) = P[X = 1] = p$
 - b) Esperanza: $E(X) = \sum_{i=0}^n [i \cdot P_X(i)] = p$
 - c) Varianza: $Var(X) = p(1 - p)$
 2. *Binomial*: experiencia de *Bernoulli* repetida n veces.
 - a) Función de cuantía: $P_X(i) = P[X = i] = \binom{n}{i} \cdot p^i \cdot (1 - p)^{n-i}$
 - b) Esperanza: $E(X) = n \cdot p$
 - c) Varianza: $Var(X) = n \cdot p \cdot (1 - p)$
 3. *Poisson*: se modelan con un parámetro λ que es la media en una constante de tiempo. Usada en el caso de la binomial cuando $n > 30$ y $p < 0.1$ siendo $\lambda = n \cdot p$.
 - a) Función de cuantía: $P_X(i) = P[X = i] = \frac{\lambda^i \cdot e^{-\lambda}}{i!}$
 - b) Esperanza: $E(X) = \lambda$
 - c) Varianza: $Var(X) = \lambda$

5. Modelos para sistemas continuos

- **Función de densidad:** $P[a \leq X \leq b] = \int_a^b f_X(x) \cdot dx \quad \forall a \leq b \in \mathbb{R}$. Propiedad, al ser continua se cumple: $P[a \leq X \leq b] = P[a \leq X < b] = P[a < X \leq b] = P[a < X < b]$.
- **Función de distribución:** $F_X(x) = P[X \leq x] = \int_{-\infty}^x f_X(x) \cdot dx$. Para un intervalo dado $P[a \leq X \leq b] = F_X(b) - F_X(a)$
- **Esperanza:** $E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x) \cdot dx$. Generalizando:

$$E(g(X)) = \int_{-\infty}^{\infty} g(x) \cdot f_X(x) \cdot dx$$
- **Varianza:** $Var(X) = \int_{-\infty}^{\infty} (x - E(X))^2 \cdot f_X(x) \cdot dx = E(X^2) - (E(X))^2$
- **Distribuciones:** notación: $X \sim Y$ significa que X es sigue una distribución Y
 1. *Uniforme* $X \sim U(a, b)$:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{si } x \notin [a, b] \end{cases}$$

$$F_X(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } x > b \end{cases}$$

$$E(X) = \frac{a+b}{2}$$

$$Var(X) = \frac{(b-a)^2}{12}$$
 2. *Exponencial* $T \sim E(\lambda)$: suele usarse para el fallo de componentes electrónicos, usa el parámetro de escala λ . Solo vale para valores positivos, por

ejemplo tiempo. Posee ausencia de memoria.

$$f_T(t) = \begin{cases} 0 & \text{si } t < 0 \\ \lambda \cdot e^{-\lambda \cdot t} & \text{si } t \geq 0 \end{cases}$$

$$F_T(t) = \begin{cases} 0 & \text{si } t < 0 \\ 1 - \lambda \cdot e^{-\lambda \cdot t} & \text{si } t \geq 0 \end{cases}$$

$$E(T) = \frac{1}{\lambda}$$

$$Var(T) = \frac{1}{\lambda^2}$$

- a) Usando la función de distribución se cumple $P[T > t] = e^{-\lambda \cdot t}$
 - b) Por la ausencia de memoria se da $P[T > t + s | T > s] = P[T > t] = e^{-\lambda \cdot t}$
 - c) Localización en parámetro a :

$$f_T(t) = \begin{cases} 0 & \text{si } t < 0 \\ \lambda \cdot e^{-\lambda \cdot (t-a)} & \text{si } t \geq 0 \end{cases}$$
3. *Normal* $X \sim N(\mu, \sigma^2)$:

$$f_X(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2}$$

$$F_X(x) = \int_{-\infty}^x \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2} \left(\frac{s-\mu}{\sigma} \right)^2} \cdot ds$$

$$E(X) = \mu$$

$$Var(X) = \sigma^2$$

$$X \sim N(\mu, \sigma^2) \rightarrow Y = a \cdot X + b \Rightarrow$$

$$Y \sim N(a \cdot \mu + b, a^2 \sigma^2)$$

$$X_1 \sim N(\mu_1, \sigma_1^2) \quad X_2 \sim N(\mu_2, \sigma_2^2) \rightarrow$$

$$Y = X_1 + X_2 \Rightarrow Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

- a) *Distribución normal estándar* $Z \sim N(0, 1)$: es con la que se trabaja normalmente y es la que sale en las tablas de probabilidades. Para tener una normal estándar de una distribución cualquiera se ha de tipificar. Necesaria para el teorema central del límite.
 - b) *Cálculo de probabilidades:* $\Phi(z) = P[Z \leq z]$, esta definición acumulativa. Para hallar la probabilidad de las colas se haría $\Phi(-z) = 1 - \Phi(z)$. $P[a \leq Z \leq b] = \Phi(b) - \Phi(a)$ si $a, b \geq 0$; $\Phi(-a) - \Phi(-b)$ si $a, b < 0$
 - c) *Aproximación normal de la distribución binomial:* para poder hacer esta aproximación se han de cumplir las siguientes condiciones: $X \sim B(n, p)$ & $n \geq 30$ & $n \cdot p \geq 5$ & $n \cdot (1 - p) \geq 5$. Entonces X se puede aproximar por $W \sim N(\mu, \sigma^2)$ $\mu = n \cdot p$ & $\sigma^2 = n \cdot p \cdot (1 - p)$
4. *Log-normal* $X \sim LN(\mu, \sigma^2)$: son datos que al ser transformados por un neperiano se transforma en una distribución normal pura.

$$E(X) = e^{\mu + \frac{1}{2}\sigma^2}$$

$$Var(X) = e^{2\mu + 2\sigma^2} \cdot (e^{\sigma^2} - 1)$$

$$Me(X) = e^{\mu} \quad Mo(X) = e^{\mu - \sigma^2}$$

$$P[a \leq X \leq b] = \Phi\left(\frac{\ln(b) - \mu}{\sigma}\right) - \Phi\left(\frac{\ln(a) - \mu}{\sigma}\right)$$

6. Fiabilidad

T indicará el tiempo hasta el fallo de un objeto.

- **Función de infiadabilidad (función de distribución):** da la probabilidad de que el sistema haya fallado en el tiempo t . $F_T(t) = P[T \leq t] \quad \forall t \geq 0$
- **Función de supervivencia:** es la función complementaria

taria a la de infiability. Suele indicarse como $S_T(t) = R_T(t) = P[T > t]$

- **Función de densidad:** se la suele denominar *densidad de fallos* y a su complementaria (esperanza) *duración media*.

- **Función de riesgo o tasa de fallos:** $h(t) = \lambda(t) = \frac{P[t \leq T \leq \Delta t / T > t]}{\Delta t} \quad \forall t \geq 0$

$$\lim_{\Delta \rightarrow 0} \frac{\Delta t}{\Delta t} \quad \forall t \geq 0$$

- **Relaciones entre las funciones y los tipos más comunes:**

1. $R_T(t) = 1 - F_T(t)$

2. $R_t(t) = \int_t^\infty f(s) \cdot ds$

3. $f_T(t) = -R'(t) = -\frac{dR(t)}{dt}$

4. $R_T(t) = e^{-\int_0^t h(s) \cdot ds}$

5. $h_T(t) = -\frac{R'(t)}{R(t)} = -\frac{d}{dt} [\ln(R(t))]$

6. $f_T(t) = h(t) \cdot e^{-\int_0^t h(s) \cdot ds}$

7. $h_T(t) = \frac{f(t)}{\int_t^\infty f(s) \cdot ds}$

8. *Distribución exponencial:*

$$T \sim E(\lambda) \Leftrightarrow h_T(t) = \lambda \quad ; \quad t \geq 0$$

9. *Distribución log-normal:*

$$T \sim LN(\mu, \sigma^2) \Leftrightarrow$$

$$h_T(t) = \frac{e^{-\frac{1}{2} \left(\frac{\ln(t) - \mu}{\sigma} \right)^2}}{t \cdot \int_t^\infty \frac{1}{s} \cdot e^{-\frac{1}{2} \left(\frac{\ln(s) - \mu}{\sigma} \right)^2} \cdot ds} \quad t > 0$$

6.1. Sistemas

Se tratan de diagramas de cajas que indican la probabilidad de éxito de un evento. Las **cajas en serie** indican intersección, se han de cumplir todas, por lo tanto, la probabilidad será la intersección de todas ellas (multiplicación de probabilidades). Las **cajas en paralelo** indican que con un solo éxito se pasa al siguiente bloque. Por lo tanto, la probabilidad será la unión de todas ellas, pero cuidado, hay que restarles la intersección entre todas ellas.

7. Muestreo

- **Muestras observada:** x_1, x_2, \dots, x_n son los distintos valores observados de una variable X .

- **Muestra aleatoria simple (m.a.s):** conjunto de n variables aleatorias: X_1, X_2, \dots, X_n .

- **Distribución de probabilidad de una m.a.s:** $F(x_1, x_2, \dots, x_n) = P[X_1 \leq x_1 \cap X_2 \leq x_2 \cap \dots \cap X_n \leq x_n] =$

$$F_X(x_1) \cdot F_X(x_2) \cdot \dots \cdot F_X(x_n) = \prod_{i=1}^n F_X(x_i).$$

Siendo F_X la función de distribución de nuestra población X . $f(x_1, x_2, \dots, x_n) = P[X_1 = x_1 \cap X_2 = x_2 \cap \dots \cap X_n = x_n] =$

$\prod_{i=1}^n f_X(x_i, \theta)$. Siendo θ los parámetros de la distribución.

- **Función de distribución empírica:** $F_n^*(X) = \frac{\sum_{x_i \leq X} 1}{n}$. Se trata de una función escalonada con saltos de altura $\frac{1}{n}$.

- **Estadísticos:**

1. **Media muestral:** $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$.

2. **Cuasivarianza muestral:** $s_X^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$.

- **Estimadores**

1. **Criterio de analogía:** se toman las medidas muestrales como si fueran poblacionales.

2. **Método de los momentos:** toma como momentos poblacionales los momentos muestrales. De ahí se tendrá que deducir los valores que definen la población.

3. **Máxima verosimilitud:** obtiene los parámetros que hacen máxima la función que obtiene la probabilidad de la muestra obtenida.

$$P(x_1, x_2, \dots, x_n) = \prod_{i=1}^n P_X(x_i, \theta).$$

Dada la función de verosimilitud $L(X; \theta)$ siendo esta la función de probabilidad se extrae su neperiano $l(X; \theta) = \ln[L(X; \theta)]$ siendo esta la función a maximizar.

Los estimadores para las distintas distribuciones son:

a) *Distribución normal:* $\mu = \bar{X}$, $\sigma^2 = s_X^2$

b) *Bernoulli:* \hat{p} , la p de la muestra.

c) *Distribución Binomial:*

d) *Distribución de Poisson:* $E(X) = \hat{\lambda} = \bar{x}$, $V(\hat{\lambda}) = \frac{\lambda}{n} \rightarrow V(X) = \lambda \cdot n$.

4. **Sesgo:** diferencia, en valor absoluto, de la media de nuestra muestra respecto a la media poblacional, viene dado por $b(\theta^*) = E(\theta^*) - \theta$, siendo θ^* la media muestra y θ la media poblacional.

5. **Eficiencia:** un estimador con poco sesgo (insesgado) que además tenga poca varianza. También se suele referir como error estándar $e = \sigma/\sqrt{n}$. $V(\bar{X}) = e^2$

6. **Error Medio Cuadrático (EMC):** $(E(\theta^*) - \theta)^2 + V(\theta^*)$.

7.1. Estimación por intervalos

- **Estimación:** se estima un parámetro ξ para una muestra X^0 con unos límites inferiores y superiores con una confianza γ . Esto se expresa como $\xi \in [t_I(X^0); t_S(X^0)]_\gamma$. Esto es la probabilidad tal que $P[t_I(X) \leq \xi \leq t_S(X)] = \gamma$. γ expresa la precisión del modelo a usar, y los t indican los valores entre los que se encuentra ξ . Se puede expresar más explícitamente como $P[k_1 \leq t(X; \theta) \leq k_2] = \gamma$, siendo los k los límites del modelo. El estimador a usar se ha de sacar de las hojas.

- **Amplitud y radio:** $Amplitud = t_S(X) - t_I(X) = 2 \cdot \varepsilon$. $Radio = \varepsilon$.

8. Inferencia estadística

- **Hipótesis:** es una suposición que haremos respecto a nuestro estudio. Suele haber dos, H_0 y H_1 , esto se debe a que si H_0 no es verdadera se pasará a la(s) siguiente(s), siendo común haber solo una alternativa. Si H_0 resulta ser muy poco probable se pasará a la siguiente hipótesis. Hay hipótesis de igualdad, comparativas, etc.
- **Errores:** en todo análisis habrá cierta precisión, por lo que es posible cometer una imprudencia. Se distinguen dos errores: *Tipo I* (α): cuando se rechaza H_0 cuando es verdadera, $\alpha = 1 - \gamma$; *Tipo II* (β): cuando se acepta H_0 cuando esta es falsa. La más grave es el *Tipo I*.
- **P-Valor:** es el valor que nos indica la probabilidad de que en una muestra nos salga cierta medición (media, varianza, probabilidad, etc), respecto cierta suposición. Su cálculo se tendrá que hacer con los modelos en las tablas. Igualando la suposición con el valor a hallar.
- **Notas importantes:** si la inferencia se realiza mediante intervalos y la hipótesis es de comparación, hay que ver si la hipótesis pudiera entrar o no en el intervalo, importante.
- **Contrastes no paramétricos:** nos permiten comprobar si una muestra sigue un modelo conocido mediante hipótesis:
 1. *Chi-cuadrado:* discrimina en un número finito r de clases cuantitativas o categóricas.
 - a) Se ordenan los datos y se agrupan en categorías.
 - b) Se plantea la hipótesis, modelo y su(s) variable(s).
 - c) Tabla que contenga por columnas las clases, los intervalos, los valores y sus frecuencias absolutas.
 - d) Se crea con el modelo teórico las frecuencias esperadas. Para ello multiplicamos la probabilidad esperada por la H_0 y la multiplicamos por n , nuestro tamaño muestral.
 - e) Se crea el estadístico de contraste con las diferencias obtenidas teóricas y empíricas.
 2. *Kolmogorov:* **no puede aplicarse a variables discretas.**
 - a) Se ordenan los datos de forma ascendente. Se crea la **función de distribución** empírica $F_n(x)$, que es la frecuencia acumulada entre el número n de datos.
 - b) Se elabora la hipótesis.
 - c) Construcción del estadístico de contraste:
$$Id = \max(|F_n(u_i) - F_0(u_i)|, |F_n(u_{i-1}) - F_0(u_i)|)$$
 - d) Crear la región crítica.