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 valores y usar el resultado como separación. Ejemplo: la mediana es el cuartil 2 o el decil 5.

2. Medidas a distribuciones

2.1. Unidimensionales

- **Media:** $\bar{x} = \frac{\sum_{i=1}^{N}[x_1]}{N}$
- **Momentos de orden r respecto del origen:** $a_r = \frac{\sum_{i=1}^{N}[x_i^r]}{N} \to \bar{x} = a_1$
- **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 tomas ese valor como mediana, aunque sea relativamente ficticio.
- **Varianza:** $s^2 = \frac{\sum_{i=1}^{N}\left[(x_i - \bar{x})^2\right]}{N} = a_2 - a_1^2$
- **Momentos desde la media de grado r:** $m_r = \frac{\sum_{i=1}^{N}[(x_i - \bar{x})^r]}{N} \to s^2 = m_2$
- **Coeficiente de variación de Pearson:** $g = \frac{s}{|\bar{x}|}$. Indica cómo de relevante es la media, si $g \leq 1$ la media es representativa.
- **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.
- **Coeficiente de asimetría de Fisher:** $g_1 = \frac{m_3}{s^3}$. $g_1 > 0$ asimétría a la derecha. $g_1 = 0$ simétrica. $g_1 < 0$ asimetría a la izquierda.
- **Coeficiente de curtosis:** coeficiente que indica cómo de puntiaguda es una gráfica $g_2 = \frac{m_4}{s^4} - 3$. $g_2 > 0$ leptocúrtica, $g_2 = 0$ mesocúrtica, $g_2 < 0$ platicúrtica.
- **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. $|g_{1_Y}| = |g_{1_X}|$
- **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

- **Número de valores:** $\sum_{i=1}^{k}\sum_{j=1}^{p}[n_{ij}] = N = \sum_{i=1}^{k}[n_{i\cdot}] = \sum_{j=1}^{p}[n_{\cdot j}]$
- **Frecuencias absolutas de la clase z_i:** $n_{i\cdot} = \sum_{j=1}^{p}[n_{ij}] \qquad n_{\cdot j} = \sum_{i=1}^{p}[n_{ij}]$
- **Frecuencias marginales:** $f_{i\cdot} = \frac{n_{i\cdot}}{N} \leq 1$. Se puede entender como la frecuencia de una característica de valores respecto del total.
- **Momentos respecto del origen de orden (r, s):** $a_{r,s} = \frac{\sum_{i=1}^{N}[x_i^r \cdot y_i^s]}{N}$. El más importante es el $m_{1,1}$.
- **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:** $m_{1,1} = a_{1,1} - a_{1,0} \cdot a_{0,1}$
- **Coeficiente 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. El signo indica si es directa o inversa. Se entiende que hay relación lineal notable si $|\rho| \geq 0{,}8$.
- **Dependencia:** se ha de cumplir y es suficiente para que dos variables sean independientes $f_{ij} = f_{i\cdot} \cdot f_{\cdot j}$

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): $A \to \bar{A}$.
  5. Contenido: A contenido en $B \Rightarrow A \subset B$.
- **Propiedades:** $P[\quad]$ 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 $A \cap B = \emptyset \to P[A \cup B] = P[A] + P[B]$
  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:** $\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$. Seleccionar k elementos de un total de n sin importar el orden.
- **Probabilidad condicionada, A cuando se cumple B:** $P\left[{}^{A}\!/_{B}\right] = \frac{P[A \cap B]}{P[B]}$. Haciendo un poco de desarrollo: $P[A \cap B] = P[B] \cdot P\left[{}^{A}\!/_{B}\right] = P[A] \cdot P\left[{}^{B}\!/_{A}\right]$. También se cumple: $P\left[{}^{A}\!/_{B}\right] = 1 - P\left[{}^{\bar{A}}\!/_{B}\right]$.

  Para sucesos independientes: $P\left[{}^{A}\!/_{B}\right] = P[A]$
- **Teorema de la probabilidad total con condicionadas:** $P[B] = \sum_{i=1}^{n}\left(P[A_i] \cdot P\left[{}^{B}\!/_{A_i}\right]\right)$. Se tiene que cumplir que los A_i compongan todo el espacio y su intersección sea nula.

- **Teorema de Bayes:** $P\left[{}^{A_k}\!/_{B}\right] = \frac{P[A_k \cap B]}{P[B]} = \frac{P[A_k] \cdot P\left[{}^{B}\!/_{A_k}\right]}{P[B]} = \frac{P[A_k] \cdot P\left[{}^{B}\!/_{A_k}\right]}{\sum_{i=1}^{n}\left(P[A_i] \cdot P\left[{}^{B}\!/_{A_i}\right]\right)}$

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:** $Var(X) = \sum_{i=1}^{n} \left[(x_i - E(X))^2 \cdot P_X(x_i)\right] = E(X^2) - (E(X))^2 = \sum_{i=1}^{n} \left[x_i^2 \cdot P_X(x_i)\right] - (E(X))^2$

Distribuciones importante:

1. *Bernoulli:* dos posibilidades, éxito o fracaso. A la probabilidad del éxito se la nombrará p.
   a) Función de cuantía: $P_X(1) = P[X = 1] = p$
   b) Esperanza: $E(X) = \sum_{i=0}^{1} [i \cdot P_X(i)] = p$
   c) Varianza: $Var(X) = p(1-p)$
2. *Binomial:* experiencia de *Bernoulli* repetida n veces. Se denota por $\sim B(n, p)$, siendo n el número de repeticiones y p la probabilidad de éxito.
   a) Función de cuantía: $P_X(i) = P[X = i] = \binom{n}{i} \cdot p^i \cdot (1-p)^{n-1}$. i denota la cantidad de éxitos.
   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: $Var(X) = \lambda$

Prestar especial atención a los límites de probabilidades indicado y a sus cambios (complementarios, siguiente, etc.) ya que los $\leq, \geq$ hay que respetarlos.

5. Modelos para sistemas continuos

- **Función de densidad:** $f_X(x)$ es la función que define la probabilidad.
- **Función de densidad 2:** $P[a \leq X \leq b] = \int_a^b f_X(x) \cdot dx \quad \forall a \leq b \in \mathbb{R}$. 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:** $F_X(x) = P[X \leq x] = \int_{-\infty}^{x} f_X(x) \cdot dx$. Define la probabilidad acumulada hasta el valor dado.
- **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 $X \sim U(a, b)$:*
  $$f_X(x) = \begin{cases} \frac{1}{b-a} & si \quad a \leq x \leq b \\ 0 & si \quad x \notin [a, b] \end{cases}$$
  $$F_X(x) = \begin{cases} 0 & si \quad x < a \\ \frac{x-a}{b-a} & si \quad a \leq x \leq b \\ 0 & si \quad x > b \end{cases}$$
  $$E(X) = \frac{a+b}{2}$$
  $$Var(X) = \frac{(b-a)^2}{12}$$
  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.
  $$f_T(t) = \begin{cases} 0 & si \quad t < 0 \\ \lambda \cdot e^{-\lambda \cdot t} & si \quad t \geq 0 \end{cases}$$
  $$F_T(t) = \begin{cases} 0 & si \quad t < 0 \\ 1 - \lambda \cdot e^{-\lambda \cdot t} & si \quad t \geq 0 \end{cases}$$
  $$E(T) = \frac{1}{\lambda}$$
  $$Var(T) = \frac{1}{\lambda^2}$$
     a) Usando la función de distribución se cumple $P[T > t] = e^{-\lambda \cdot t}$
     b) Por la ausencia de memoria se da $P\left[{}^{T > t+s}\!/_{T > s}\right] = P[T > t] = e^{-\lambda \cdot t}$
     c) Localización en parámetro a:
  $$f_T(t) = \begin{cases} 0 & si \quad t < 0 \\ \lambda \cdot e^{-\lambda \cdot (t-a)} & si \quad t \geq 0 \end{cases}$$
  3. *Normal $X \sim N(\mu, \sigma^2)$:*
  $$f_X(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$
  $$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{s-\mu}{\sigma}\right)^2} \cdot ds$$
  $$E(X) = \mu$$
  $$Var(X) = \sigma^2$$
  $$X \sim N(\mu, \sigma^2) \rightarrow Y = a \cdot X + b \Rightarrow Y \sim N(a \cdot \mu + b, a^2\sigma^2)$$
  $$X_1 \sim N(\mu_1, \sigma_1^2) \quad X_2 \sim N(\mu_2, \sigma_2^2) \rightarrow Y = X_1 \pm X_2 \Rightarrow Y \sim N(\mu_1 \pm \mu_2, \sigma_1^2 + \sigma_2^2)$$

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.
$$E(X) = e^{\mu + \frac{1}{2}\sigma^2}$$
$$Var(X) = e^{2\cdot\mu+\sigma^2} \cdot (e^{\sigma^2} - 1)$$
$$Me(X) = e^{\mu} \qquad Mo(X) = e^{\mu - \sigma^2}$$
$$P\left[a \leq X \leq b\right] = \Phi\left(\frac{ln(b) - \mu}{\sigma}\right) - \Phi\left(\frac{ln(a) - \mu}{\sigma}\right)$$
5. *Weibull* $X \sim W(\beta, \eta, \gamma)$*:* flexibiliza la distribución exponencial. Da probabilidad para valores superiores al parámetro γ; $\beta > 0, \eta > 0$ son parámetros de forma y escala respectivamente.
$$f_X(x) = \frac{\beta}{\eta^\beta}(x-\gamma)^{\beta - 1} \cdot e^{-\left(\frac{x-\gamma}{\eta}\right)^\beta}$$
$$F_X(x) = 1 - e^{-\left(\frac{x-\gamma}{\eta}\right)^\beta}$$
$$E(X) = \gamma + \eta \cdot \Gamma\left(1 + \frac{1}{\beta}\right)$$
$$Var(X) = \eta^2 \cdot \Gamma\left(1 + \frac{2}{\beta}\right) - \left(\eta \cdot \Gamma\left(1 + \frac{1}{\beta}\right)\right)^2$$

Teorema central del límite: este teorema nos indica que una suma numerosa (>30) de datos, sus mendias, se distribuyen como una normal $\sim N\left(\frac{\sum \mu}{n}, \frac{\sum \sigma^2}{n^2}\right)$. Si todos los datos provienen de una misma distribución entonces el TCL se simplifica a $\sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

Funciones de distribución: las más importantes son la Chi-cuadrado χ^2_n de n grados de libertad, la T de Student t_n de n grados de libertad; y la F de Snedecor $F_{n,m}$ de n, m grados de libertad $\left(F_a(m,n) = \frac{1}{F_{1-\alpha}(n,m)}\right)$. Si la T de Student ($t_n$) es de una n muy grande, se la puede aproximar por una normal normalizada.

6. Fiabilidad

T indicará el tiempo hasta el fallo de un objeto.

- **Función de infiabilidad (función de distribución):** da la probabilidad de que el sistema haya fallado en el tiempo t. $F_T(t) = P\left[T \leq t\right] \quad \forall t \geq 0$
- **Función de supervivencia:** es la función complementaria a la de infiabilidad. Suele indicarse como $S_T(t) = R_T(t) = P\left[T > t\right]$
- **Función de densidad:** se la suele denominar *densidad de fallos* y a su complementaria (esperanza) *duración media.*
- **Función de riesgo o tasa de fallos:** $h(t) = \lambda(t) = \lim_{\Delta \to 0} \frac{P\left[t \leq T \leq \Delta t / T > t\right]}{\Delta t} \quad \forall t \geq 0$
- **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. $h_T(t) = -\frac{R'(t)}{R(t)} = -\frac{d}{dt}\left[ln(R(t))\right]$
  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. *Distibución log-normal:*
  $$T \sim LN(\mu, \sigma^2) \Leftrightarrow$$
  $$h_T(t) = \frac{e^{-\frac{1}{2}\left(\frac{ln(t)-\mu}{\sigma}\right)^2}}{t \cdot \int_t^\infty \frac{1}{s} \cdot e^{-\frac{1}{2}\left(\frac{ln(s)-\mu}{\sigma}\right)^2} \cdot ds}$$
  $$t > 0$$

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:** $x_1, x_2, \ldots, x_n$ son los distintos valores observados de una variable X.
- **Muestra aleatoria simple (m.a.s):** conjunto de n variables aleatorias: $X_1, X_2, \ldots, X_n$.
- **Distribución de probabilidad de una m.a.s:** $F(x_1, x_2, \ldots, x_n) = P\left[X_1 \leq x_1 \cap X_2 \leq x_2 \cap \ldots \cap X_n \leq x_n\right] = F_X(x_1) \cdot F_X(x_2) \cdot \ldots F_X(x_n) = \prod_{i=1}^n F_X(x_i)$. Siendo F_X la función de distribución de nuestra población X. $f(x_1, x_2, \ldots, x_n) = P\left[X_1 = x_1 \cap X_2 = x_2 \cap \ldots \cap X_n = x_n\right] = \prod_{i=1}^n f_X(x_i, \theta)$. 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:** $S_X^2 = \frac{\sum_{i=1}^n \left(X_i - \bar{X}\right)^2}{n-1}$.
- **Estimadores**
  1. **Criterio de analogía:** se tomas 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. $P(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} P_X(x_i, \theta)$. Dada la función de verosimilitud $L(X;\theta)$ siendo esta la función de probabilidad se extrae su neperiano $l(X;\theta) = ln\left[L(X;\theta)\right]$ 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:* $n \cdot p$.
   d) *Distribución de Poisson:* $E(X) = \hat{\lambda} = \bar{x}$, $V(\hat{\lambda}) = \frac{\lambda}{n} \to V(X) = \lambda \cdot n$.
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 muestral 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 Cuadrático Medio (ECM):** $(E(\theta^*) - \theta)^2 + V(\theta^*)$. Mejor ECM, mejor estimador (si no es muy sesgado).

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 $\xi \in \left[t_I(X^0); t_S(X^0)\right]_\gamma$. Esto es la probabilidad talque $P\left[t_I(X) \leq \xi \leq t_S(X)\right] = \gamma$. 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 $P\left[k_1 \leq t(X;\theta) \leq k_2\right] = \gamma$, 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. Si H_0 resulta ser muy poco probable se pasará a la hipótesis alternativa H_1. La H_0 ha de ser una hipótesis de igualdad, H_1 no tiene porqué. 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, $\alpha = 1 - \gamma$ (a α se le denomina significación de contraste); *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 y haciendo el contraste con la distribución en cuestión (sacada de las hojas). **Cuidado:** el p-valor es el α que tendríamos con la "medición" realizada, hay que mirar cómo es el α para cada caso. **RC:** región de rechazo.
- **Contrastes no paramétricos:** nos permiten comprobar si una muestra sigue un modelo conocido mediante hipótesis. **IMPORTANTÍSIMO:** las hipótesis de contraste se presuponen completamente dadas (se dan todos los parámetros), si esto no fuera así, se tendrían que "inferir"mediante método de los momentos o similar.
  1. *Chi-cuadrado:* discrimina en un número finito r de clases cuantitativas, categóricas o intervalos. **Es un modelo discreto.** Se realizará mediante una tabla.
     a) Primera columna: clases o intervalos en orden creciente.
     b) Segunda: frecuencia absoluta para cada clase o intervalo.
     c) Tercera: probabilidades de tales clases o intervalos bajo el contraste H_0, la hipótesis. Para intervalos, la probabilidad es $P\left[min < X < max\right]$.
     d) Cuarta: frecuencias teóricas para cada intervalo o clase según H_0. Esto es, la probabilidad de la columna anterior c) por la frecuencia absoluta de la columna b). **IMPORTANTE:** si la frecuencia teórica de esta columna es menor que 5, se tienen que reagrupar los intervalos o las clases de tal manera que todas las frecuencias sean mayores que 5.

     ---

     e) Id_0, que será nuestro valor de contraste. $Id_0 = \sum_{i=0}^{k} \frac{(n_i - E_i)^2}{E_i}$. E_i es la frecuencia teórica (última columna) de cada valor; n_i es la frecuencia absoluta, columna b).
     f) Id_0 se distribuye como una χ^2_{k-r-1}. Siendo k el número de intervalos o clases, r el número de parámetros estimados en la hipótesis nula.
     g) El p-valor de nuestra hipótesis es por lo tanto: $P[\chi^2 \geq \chi^2_{k-r-1}]$
  2. *Kolmogorov-Smirnov:* **no puede aplicarse a variables discretas.** A continuación se hace una tabla con las siguientes columnas:
     a) Se ordenan los datos de forma ascendente.
     b) Se crea la **función de distribución** empírica $F_n(x)$, que es la frecuencia acumulada entre el número n de datos (es la escalonada).
     c) Función de distribución para cada valor usando la función del contraste.
     d) Resta de las dos columnas anteriores en valor absoluto.
     e) La función de distribución de la segunda columna de un valor anterior (empieza con 0 y asciende igual que la segunda columna).
     f) Resta de la columna anterior con la tercera columna (función de distribución del contraste).

     ---

g) Se halla el valor máximo de las columnas de valores absolutos, d) y f). Este valor será el valor de Id.

h) Buscamos en las tablas para una α dada y el número n de muestras recoginas.

i) Si Id es mayor que el valor de la tabla para el valor de significación dado (α), la hipótesis nula ha de rechazarse.