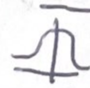
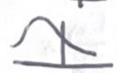
TEMA 1.-

- Medidas posición:
  - Tendencia central { Media aritmética, Mediana, Moda
  - Tendencia no central { Cuartiles, Cuantiles
- Medidas de dispersión:
  - Absoluta { Varianza, Cuasivarianza
  - Relativa { Coeficiente de variación. $\frac{S}{\bar{x}}$ no depende de las unidades.
- Medidas de forma:
  - Coeficiente de asimetría. $\gamma F = \frac{1}{S^3} \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^3$

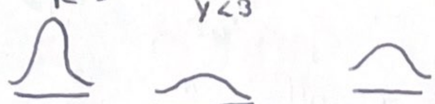
$\gamma F < 0$  más valores por ~~encima~~ debajo de la media.

$\gamma F = 0$  simétrica.

$\gamma F > 0$  mas valores por encima de la media.

  - Coeficiente de curtosis. $\gamma C = \frac{1}{S^4} \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^4$

$\gamma > 3$ $\gamma < 3$



• Diagrama de caja

$$LI = \max\{\min\{x\}, Q_1 - 1'5(Q_3 - Q_1)\}$$
$$LS = \min\{\max, Q_3 + 1'5(Q_3 - Q_1)\}$$

• Regresión:

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

$S_{xy} > 0$ rel. directa

$S_{xy} < 0$ rel. inversa

$S_{xy} = 0$ incorrelados

- Coef correlación.

$r_{xy} = \frac{S_{xy}}{S_x S_y} \in [-1,1]$ no depende de las unidades.

- Recta:

$y = a + bx$ $b = \frac{S_{xy}}{S_x^2}$ $a = \bar{y} - b\bar{x}$

$y = \bar{y} - b\bar{x} + \frac{S_{xy}}{S_x^2}x$ $y = \bar{y} - \frac{S_{xy}}{S_x^2}\bar{x} + \frac{S_{xy}}{S_x^2}x$

$$y = \bar{y} - \frac{S_{xy}}{S_x^2}(\bar{x} - x)$$

- Coef determinación

$R^2 = r_{xy}^2 = \frac{S_{xy}^2}{S_x^2 S_y^2}$ $[0,1]$

Si se acerca a 1 es un buen ajuste.

TEMA 2:

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

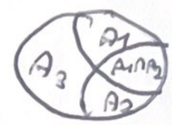
$$P(\overline{A \cup B}) = P(\bar{A} \cap \bar{B})$$

$$P(\overline{A \cap B}) = P(\bar{A} \cup \bar{B})$$

$$A \setminus B = A \cap \bar{B}$$

Sucesos incompatibles $A \cap B = \emptyset$

Sistema exhaustivo $\bigcup_{i=1}^{n} A_i = \Omega$



Sistema completo $A_i \cap A_j = \emptyset$

y

$\bigcup_{i=1}^{n} A_i = \Omega$

$P(\emptyset) = 0$

$P(\Omega) = 1$

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

Prob. condicionada

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

Sucesos independientes

$$P(A \cap B) = P(A) \cdot P(B)$$

$P(A \cap B \cap C)$

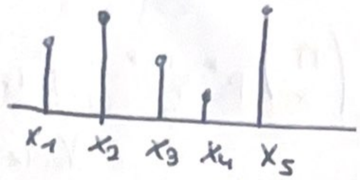
$= P(A) \cdot P(B|A) \cdot P(C|A \cap B)$

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

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

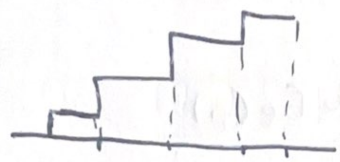
• Tema 3.

Masa probabilidad $P(X=x_i)$



$P(X=x) \begin{cases} x_1 = 0'2 \\ x_2 = 0'4 \\ x_3 = 0'4 \end{cases}$

Función distribución $P(X \leq x)$



$F(X\leq x) \begin{cases} 0 & \text{si } x<0 \\ 0'25 & \text{si } x\in[0,1) \\ 0'5 & \text{si } x\in[1,2) \\ 1 & \text{si } x\geq 2 \end{cases}$

$E(x) = \mu = \sum_{i=1}^{k} x_i p_i \begin{cases} E(aX+b) = aE(x)+b \\ E(X+Y) = E(x)+E(y) \end{cases}$

$Var(x) = \sigma^2 = E[(x-E(x))^2]$

$Var(aX+b) = a^2 Var(x)$

$\boxed{Var(x) = E(x^2) - E^2(x)}$

Bernoulli -> 2 resultados (éxito fracaso)

$X \sim Ber(p)$ X=" prob de exito"

$P(X=x) = p^x (1-p)^{1-x}$

$E(x) = p$ $Var(x) = p \cdot q$

• Binomial → $X \sim Bi(n,p)$

$X =$ " ~~Prob de obtener~~ X éxitos en n pruebas de Bernoulli "

$$P(X=x) = \binom{n}{x} p^x q^{n-x}$$

$E(x) = n \cdot p$

$Var(x) = npq$

$$\left.\begin{array}{l} X \in Bi(n_1,p) \\ Y \in Bi(n_2,p) \end{array}\right\} X+Y \sim Bi(n_1+n_2,p)$$

• Geométrica → $X \sim Ge(p)$

$X =$ " nº fracasos hasta el primer éxito. "

$$P(X=x) = q^x p$$

$E(x) = \frac{q}{p}$

$Var(x) = \frac{q}{p^2}$

• Binomial Negativa → $X \sim BN(n,p)$

$X =$ " nº fracasos hasta el n-ésimo éxito".

$$P(X=x) = \binom{n+x-1}{x} q^x p^n$$

$E(x) = \frac{nq}{p}$

$Var(x) = \frac{nq}{p^2}$

• Poisson → $X \sim Pois(\lambda)$

$X =$ " nº sucesos en un intervalo ".

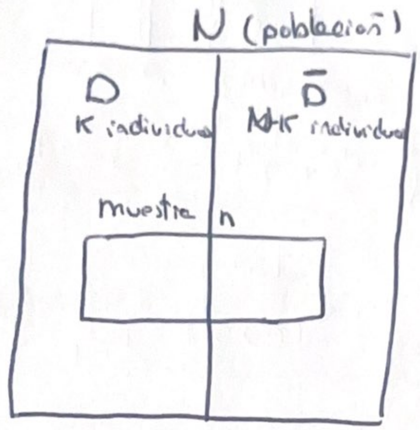
$\lambda =$ nº medio de sucesos en el intervalo.

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

$E(x) = Var(x) = \lambda$

- Hipergeométrica $X \sim H(N, n, k)$

X = nº de elementos de la clase D en la muestra de tamaño n



$$P(X=x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

~~P(X=x) = ((K x)(N-~~

$$E(X) = \frac{nk}{N}$$

$$Var(X) = npq\,\frac{N-n}{N-1}$$

Hipergeométrico no tiene reemplazamiento.
Binomial sí que tiene reemplazo.

- Uniforme discreto $X \sim U(x_1, \dots x_k)$

$$P(X=x) = \frac{1}{k}$$

$$E(X) = \frac{1}{k}\sum_{i=1}^{k} x_i$$

$$Var(X) = \frac{1}{k}\sum_{i=1}^{k} (x_i-\mu)^2$$

Tema 4-

- Funcion densidad $F(x) = \int_{-\infty}^{x} F(t)\,dt$

$P(X=x)$

- Función distribución $F(x) = P(X \leq x)$

$$\int_{-\infty}^{\infty} f(t)\,dt = 1 \qquad E(x) = \int_{-\infty}^{\infty} x f(x)\,dx$$

$$Var(x) = \int_{-\infty}^{\infty} x^2 f(x)\,dx - E^2(x)$$

- Uniforme continua $X \sim U(a,b)$

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & x \geq b \end{cases}$$

$$E(x) = \mu = \frac{a+b}{2} \qquad Var(x) = \sigma^2 = \frac{(b-a)^2}{12}$$

- Normal $X \sim N(\mu, \sigma^2)$

$$Z = \frac{X-\mu}{\sigma} \in N(0,1) \qquad (X+Y) \in N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

$$q(p_0) = x_0 \qquad (X-Y) \in N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$$

· Exponencial $\quad X \sim Exp(\lambda)$

$X =$ " tiempo entre 2 sucesos consecutivos "

$F(x) = 1 - e^{-\lambda x}$ $\qquad \lambda = \dfrac{\text{sucesos}}{\text{unidad de tiempo}}$

$E(x) = \dfrac{1}{\lambda} \qquad Var(x) = \dfrac{1}{\lambda^2}$

· Th central del límite.

$X_1, \dots X_i$

$$\sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2) \iff \sum_{i=1}^{n} \frac{X_i - n\mu}{\sigma^2\sqrt{n}} \sim N(0,1)$$

$Bi(n,p) \qquad n \geq 30 \qquad 0'1 < p < 0'9 \qquad N(np, npq)$

$Pois(\lambda) \qquad \lambda \geq 10 \qquad N(\lambda, \lambda)$

- Tema 5 -

Proporción
$$\begin{cases} \hat{p} \sim N\left(p, \frac{pq}{n}\right) \\ \hat{p}_x - \hat{p}_y \sim N\left(p_x - p_y, \frac{p_x q_x}{n} + \frac{p_y q_y}{m}\right) \end{cases}$$

Media
$$\begin{cases} \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \\ \bar{X} - \bar{Y} \sim N\left(\mu_x - \mu_y, \frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{m}\right) \\ \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1) \end{cases}$$

Varianza
$$\begin{cases} \frac{ns^2}{\sigma^2} \sim \chi^2_{n-1} \quad E(x) = n \quad Var(x) = 2n \\ \frac{(n-1)S_c^2}{\sigma^2} \sim \chi^2_{n-1} \end{cases}$$

Media varianza desconocida
$$\left\{ \frac{\bar{X} - \mu}{\frac{S_c}{\sqrt{n-1}}} \sim t_{n-1} \quad o \quad \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1} \right.$$

Intervalo de confianza → intervalo que nos indica que en un cierto % nuestros valores se encontrarán a él. Si lo tenemos al 99% de confianza sobre la proporción, el 99% de las ocasiones la proporción estará en el intervalo.

$$IC = \left(\hat{p} \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{pq}{n}} \right)$$ Proporcion

$$IC = \left(\bar{x} \pm z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right)$$ Media

$$I_C = \left(\bar{x} \pm t_{n-1, 1-\frac{\alpha}{2}} \frac{S_c}{\sqrt{n-1}} \right)$$ Media varianza desconocida.

$$I_C = \left(\frac{ns^2}{\chi^2_{n-1, 1-\frac{\alpha}{2}}}, \frac{ns^2}{\chi^2_{n-1, \frac{\alpha}{2}}} \right)$$ Varianza.

Longitud es $2 \cdot z_{1-\frac{\alpha}{2}} \sqrt{\frac{pq}{n}}$

Tema 6:

- P-valor:

El p-valor representa la probabilidad de observar un valor o resultado tan extremo como el observado.

Al contrastar hipótesis se pueden dar cuatro situaciones:

- H_0 es verdad y la aceptamos
- H_0 es falsa y la rechazamos.
- Error tipo I: H_0 es verdad y la rechazamos
- Error tipo II: H_0 es falsa y la aceptamos.

Para el error de tipo I tenemos el nivel de significación:

$\alpha \rightarrow$ probabilidad de cometer Error tipo I.

$$\alpha = P(\text{error tipo I}) = P(\text{rechazar } H_0 / H_0 \text{ cierta})$$

Para el de tipo II tenemos la potencia de un contraste β

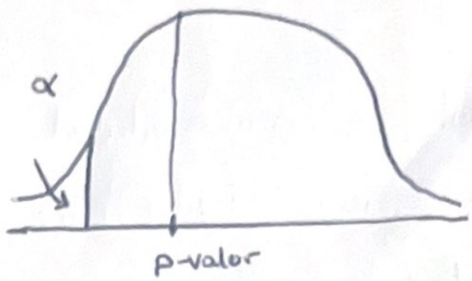
$$1-\beta = \text{potencia} = P(\text{rechazar } H_0 \mid H_0 \text{ falsa})$$

Esto ~~no~~ esta mal en los apuntes.

Resumidamente:

	Decisión: Rechazar H_0	Aceptar H_0
H_0 verdad	Error tipo (I) $P(\text{err tipo I}) = \alpha$ nivel de significación.	$P(\text{aceptar } H_0 / H_0 \text{ cierta}) = 1-\alpha$
H_0 Falsa	$P(\text{Rechazar } H_0, H_0 \text{ Falsa}) = 1-\beta$ Potencia de contraste	Error tipo (II) $P(\text{err tipo II}) = \beta$

$H_0: \mu \geq 0$
$H_1: \mu < 0$



α -> prob de rechazar siendo verdad

p-valor -> prob ver un resultado tan extremo como el observado

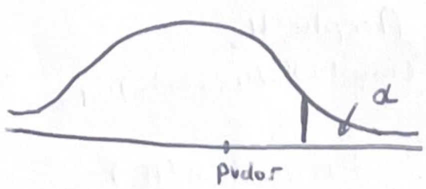
Si $\alpha < P(\text{p-valor})$ -> La probabilidad de nuestro suceso observado es mayor que la probabilidad de error. Por lo que no hay suficiente evidencia como para rechazar Ho.

Si $\alpha > P(\text{p-valor})$ -> La probabilidad de ver un resultado tan extremo es menor que la probabilidad de error. Por lo que el dato observado es improbable que sea producto del ~~fa~~ azar. Debemos rechazar Ho.

$P(z < \text{p-valor})$

$H_0: \mu \leq 0$
$H_1: \mu > 0$

$P(z > \text{pvalor})$



$H_0: \mu = 0$
$H_1: \mu \neq 0$

$P(z < \text{pvalor}) + P(z > \text{p-valor})$