TRANSFORMACIÓN AFÍN

$Y = aX + b$

$$\begin{cases} \bar{y} = a\bar{x} + b \\ \sigma_y^2 = a^2 \sigma_x^2 \end{cases} \Rightarrow \sigma_y = |a|\sigma_x$$

Desv. típica
No afectada por cambios de origen.
→ Si escala.

MOMENTOS NO CENTRALES.

$$m_r = \frac{1}{n} \sum_{i=1}^{k} n_i x_i^r$$

MOMENTOS CENTRALES

$$\mu_r = \frac{1}{n} \sum_{i=1}^{k} n_i (x_i - \bar{x})^r$$

~~MOMENTOS~~

TEMA 2

TEMA 2:

D. CONDICIONADAS

Distribución del carácter X condicionado a la modalidad y_j de Y:

$$\boxed{X/Y = y_j}$$

Notación

Distribución del carácter Y condicionado a la modalidad x_i de X:

$$\boxed{Y/X = x_i}$$

Notación

Con ceros "0" de frecuencia NO HAY INDEPENDENCIA

INDEPENDENCIA ESTADÍSTICA

Distribuciones $X/Y = y_j \quad \forall j = 1, 2, \dots$ identicas (proporcionales)
↳ identicas a distrib. marg. X.

$$\boxed{f_i^j = f_{i/j} = f_{i\cdot} \quad \forall i = 1, 2, \dots, k \qquad X \text{ indep. de } Y}$$

Idem. Y indep. X. ; 1 Indep ⇒ 2 indep.

DEPENDENCIA FUNCIONAL

X dep. funcionalmente de Y cuando a cada modalidad y_j d Y le corresp. una única modalidad de X.

↳ Frec. absoluta n_{ij} nula excepto 1 valor $i = \varphi(j)$

" Y (1dem).

MOMENTOS Bidimensionales

no centrales → $m_{rs} = \sum_{i=1}^{k} \sum_{j=1}^{p} f_{ij} x_i^r \cdot y_j^s$

centrales → $\mu_{rs} = \sum_{i=1}^{k} \sum_{j=1}^{p} f_{ij} (x_i - \bar{x})^r (y_j - \bar{y})^s$

⊛ $\boxed{\mu_2 = m_2 - m_1^2}$

$\mu_{20} = \sigma_x^2$

$\mu_{02} = \sigma_y^2$

μ_{11} = COVARIANZA = $\sigma_{xy} = \mu_{11} = m_{11} - m_{10} m_{01}$

Cov. = σ_{xy}

→ $m_{11} - m_{10} \cdot m_{01} = \frac{1}{n} \left[\sum_{i=1}^{k} \sum_{j=1}^{p} n_{ij} \cdot x_i \cdot y_j - \sum_{i=1}^{k} x_i n_i \cdot \sum_{j=1}^{p} y_j \cdot n_j \right]$

→ $\boxed{\frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{p} x_i \cdot y_j \, n_{ij} - \bar{x} \cdot \bar{y}}$

Varianza σ_x^2

→ $\mu_2 = m_2 - m_1^2 = \frac{1}{n} \sum_{k=1}^{k} n_i x_i^2 - \left(\frac{1}{n} \sum_{i=1}^{k} n_i x_i \right)^2$

→ $\frac{1}{n} \sum_{i=1}^{k} n_i \cdot (x_i - \bar{x})^2$

Desv. típ. $+\sqrt{\sigma^2}$

Prop

→ X, Y independientes est.

$\begin{cases} m_{rs} = m_{r0} m_{0s} \\ \mu_{rs} = \mu_{r0} \mu_{0s} \end{cases}$

⇆ σ_{xy} covarianza $= 0$

$\not\Leftarrow$

REGRESIÓN

- **Regresión**

Determinar estructura de dependencia que mejor expresa la relación de y con el resto.

- **CORRELACIÓN**

Estudio de grado de dependencia entre variables.

⇒ CURVA REGRESIÓN tipo I: (mayor η^2)

Y/X → curva pasa por $(x_i, \bar{y}_i)$ $i = 1, 2, \ldots, k$

X/Y → curva pasa por $(\bar{x}_j, y_j)$ $j = 1, 2, \ldots, p$.

→ curva que mejor se ajusta a datos (pero no predicción)

⇒ RECTA REGRESIÓN:

$\boxed{\text{Curvas Regresión Tipo II}}$

Y/X (y sobre X)

$$y = ax + b \Rightarrow y - \bar{y} = \frac{\sigma_{xy}}{\sigma^2_x}(x - \bar{x})$$

$$\Downarrow$$

$$\boxed{y = \underbrace{\frac{\sigma_{xy}}{\sigma^2_x}}_{a} x + \underbrace{\bar{y} - \frac{\sigma_{xy}}{\sigma^2_x}\bar{x}}_{b}}$$

X/Y Idem $x - \bar{x} = \frac{\sigma_{xy}}{\sigma^2_y}(y - \bar{y}) \ldots$

- coeficiente de regresión lineal: (Y sobre X = $\frac{\sigma_{xy}}{\sigma_x^2}$)

→ Y/X $\frac{\sigma_{xy}}{\sigma^2_x}$

- mismo signo ambas rectas ⇒ pendiente ↗ ↘
- mismo signo covarianza.

→ X/Y $\frac{\sigma_{xy}}{\sigma^2_y}$

⊛ Covarianza nula ⇒ rectas regresión paralelas ejes

- Propiedades rectas regresión:
  - → Ambas (X/y, Y/x) pasan por $(\bar{x}, \bar{y})$
  - → Media de los valores ajustados = Media valores observados
  - → Media de residuos = 0
  - → Media de los productos de residuos por valores var. explicativa. = 0.
  - → Media de los productos de residuos por valores ajustados = 0.

⇒ OTRO TIPO DE FUNCIÓN

⇒ TRANSFORMACIÓN A AJUSTE LINEAL → recta

1. Pasar a "recta"
2. Deshacer cambio → Sacar $\eta^2_{y/x}$ de la recta (r^2)

Cambio a recta → expresión otras variables (ó Z, W, ...) $\begin{cases} X' = f(x) \\ Y' = g(y) \end{cases}$

Hipérbola equilátera:

→ $y = a + b\frac{1}{x}$ ⇒ $\boxed{x' = \frac{1}{x}}$ → $\boxed{y = a + bx'}$

Curva potencial:

→ $y = ax^b$ ⇒ $\boxed{\underbrace{\log y}_{W} = b \underbrace{\log x}_{Z} + \underbrace{\log a}_{a'}}$ → $\boxed{W = bZ + a'}$

Curva exponencial:

→ $y = ab^x$ ⇒ $\boxed{\log y = x \log b + \log a}$ → $\boxed{y' = b'x + a'}$

(Cálculos y Fórmulas)

$\boxed{\text{MOMENTOS}}$ (unidimensionales)

$$m_r = \frac{1}{n}\sum_{i=1}^{k} n_i x_i^r \qquad \mu_r = \frac{1}{n}\sum_{i=1}^{k} n_i (x_i - \bar{x})$$

NO CENTRALES — CENTRALES

$\boxed{\text{MEDIA}}$ ARITMÉTICA $\bar{x}$

momento 1 m_1

$$\frac{1}{n}\sum_{i=1}^{k} n_i x_i = m_1$$

$\boxed{\text{VARIANZA}}$ σ_x^2

$$\boxed{\mu_2 = m_2 - m_1^2} \qquad \boxed{\sigma_x^2 = \frac{1}{n}\sum_{i=1}^{k} n_i x_i^2 - (\bar{x})^2}$$

$$\hookrightarrow \frac{1}{n}\sum_{i=1}^{k} n_i (x-\bar{x})^2$$

$\boxed{\text{DESV. TÍPICA}}$ σ_x

$$+\sqrt{\sigma_x^2}$$

$\boxed{\text{COVARIANZA}}$ $x, y \to \sigma_{xy}$

$$\boxed{\sigma_{xy} = m_{11} - m_{10}\cdot m_{01}} = \boxed{\frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{p} n_{ij} x_i y_j - \bar{x}\bar{y} = \sigma_{xy}}$$

Momentos bidim

$$m_{rs} = \frac{1}{n}\sum^{k}\sum^{p} n_{ij}\, x_i^r y_j^s$$

m_{11}

Escaneado con CamScanner

¿Cómo de bien expresa g la rela?

→ Min cuad → error cometido

① Coeficiente de determinación $\boxed{\eta^2}$ caso lineal $\approx r^2$

" razón de correlación"

② Coeficiente de correlación lineal caso lineal r

CORRELACIÓN

VARIANZA RESIDUAL $\sigma^2_{res\,y}$

$$\frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{p} n_{ij}(y_j-\hat{y})^2 - \text{media resid.}^2 = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{p} n_{ij}[y_j-g(x_i)]^2 - \text{med res}^2$$

- En funciones lineales parámetros → media residuos $=0$

$$\Rightarrow \sigma^2_{res} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{p} n_{ij}(y_j-g(x_i))^2$$

Varianza residual → Varianza de variable no explicada por $g()$. (regresión)

VARIANZA EXPLICADA σ^2_{ey}

$$\sigma^2_{ey} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{p} n_{ij}(\hat{y}-\bar{y})^2$$

→ Varianza explicada por la regresión

$$\boxed{\sigma^2_y = \sigma^2_{ey} + \sigma^2_{res.y}}$$

COEFICIENTE DETERMINACIÓN (¿razón de correlación?)

- $$\eta^2_{y/x} = \frac{\sigma^2_{ey}}{\sigma^2_y} = 1 - \frac{\sigma^2_{res.y}}{\sigma^2_y}$$

- $0 \leq \eta^2_{y/x} \leq 1$

⊛ CASO LINEAL

En caso lineal $y = ax + b$

$$\eta^2_{Y/X} = \eta^2_{X/Y} = \frac{\sigma^2_{xy}}{\sigma^2_x \cdot \sigma^2_y} = r^2$$ razón de correlación lineal

→ r^2 = producto pendientes rectas $\left\{ \begin{matrix} Y/X \\ X/Y \end{matrix} \right.$

$\boxed{\text{COEFICIENTE CORRELACIÓN LINEAL}}$ "r"

$$r = \pm\sqrt{r^2} = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y}$$ (signo de covarianza σ_{xy})

→ Determina grado de dependencia lineal de la variable dependiente ante la independiente

→ Si $\sigma_{xy} = 0 \Rightarrow r = 0 \Rightarrow$ no dependencia entre variables tipo lineal.

→ $-1 \leq r \leq 1$

→ $\left\{ \begin{array}{l} \ominus \rightarrow \text{relación inversa (pendientes 2 negativas)} \\ \oplus \rightarrow \text{relación directa (pendientes 2 positivas).} \end{array} \right.$