DEMOSTRACIONES

$$\bar{x} = \sum_{j=1}^{p} f_{\cdot j}\, \bar{x}_j$$

$$\bar{x} = \sum_{i=1}^{k} f_{i\cdot}\, x_i = \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}\, x_i = \sum_{i=1}^{k}\sum_{j=1}^{p} f_{\cdot j}\, f_i^j\, x_i = \sum_{j=1}^{p} f_{\cdot j} \sum_{i=1}^{k} f_i^j\, x_i = \sum_{j=1}^{p} f_{\cdot j}\, \bar{x}_j$$

$$\bar{y} = \sum_{i=1}^{k} f_{i\cdot}\, \bar{y}_i$$

$$\bar{y} = \sum_{j=1}^{p} f_{\cdot j}\, y_j = \sum_{j=1}^{p}\sum_{i=1}^{k} f_{ij} \cdot y_j = \sum_{j=1}^{p}\sum_{i=1}^{k} f_{i\cdot} f_j^i \cdot y_j = \sum_{i=1}^{k} f_{i\cdot} \sum_{j=1}^{p} f_j^i\, y_j = \sum_{i=1}^{k} f_{i\cdot}\, \bar{y}_i$$

$$\sigma_x^2 = \sum_{j=1}^{p} f_{\cdot j}\, \sigma_{x,j}^2 + \sum_{j=1}^{p} f_{\cdot j}\, (\bar{x}_j - \bar{x})^2$$

$$\sigma_x^2 = \sum_{j=1}^{p} f_{i\cdot}\, (x_i - \bar{x})^2 = \sum_{i=1}^{k}\sum_{j=1}^{p} f_i^j\, f_{\cdot j}\, (x_i - \bar{x})^2 = \sum_{j=1}^{p}\sum_{i=1}^{k} f_{\cdot j}\, f_i^j\, (x_i - \bar{x})^2 =$$

$$= \sum_{j=1}^{p} f_{\cdot j} \left[\sum_{i=1}^{k} f_i^j\, (x_i - \bar{x})^2\right] = \sum_{j=1}^{p} f_{\cdot j} \left[\sum_{i=1}^{k} f_i^j\, (x_i - \bar{x}_j + \bar{x}_j - \bar{x})^2\right] =$$

$$= \sum_{j=1}^{p} f_{\cdot j} \left[\sum_{i=1}^{k} f_i^j\, (x_i - \bar{x}_j)^2 + \sum_{i=1}^{k} f_i^j\, (\bar{x}_j - \bar{x})^2 + \cancelto{0}{2 \sum_{i=1}^{k} f_i^j\, (x_i - \bar{x}_j)(\bar{x}_j - \bar{x})}\right] =$$

$$= \sum_{j=1}^{p} f_{\cdot j} \left[\sigma_{x,j}^2 + (\bar{x}_j - \bar{x})^2\right]$$

■ X indep. Y $\Longleftrightarrow$ Y indep X

$\Rightarrow$

¿Si $f_i^j = f_{i\cdot} \Rightarrow f_j^i = f_{\cdot j}$?

$$f_{ij} = f_{i\cdot}\, f_j^i = f_{\cdot j}\, f_i^j$$

$$\cancel{f_{i\cdot}}\, f_j^i = f_{\cdot j}\, \cancel{f_{i\cdot}}$$

$$f_j^i = f_{\cdot j} \Longrightarrow \text{Y indep X}$$

$\Leftarrow$

¿Si $f_j^i = f_{\cdot j} \Rightarrow f_i^j = f_{i\cdot}$?

$$f_{ij} = f_{\cdot j}\, f_i^j = f_{i\cdot}\, f_j^i$$

$$\cancel{f_{\cdot j}}\, f_i^j = f_{i\cdot}\, \cancel{f_{\cdot j}}$$

$$f_i^j = f_{i\cdot} \Longrightarrow \text{X indep Y}$$

■ $f_{i/j} = f_{i\cdot}$

$$f_{i\cdot} = \sum_{j=1}^{p} f_{ij} = \sum_{j=1}^{p} f_{i/j}\, f_{\cdot j} = f_{i/j} \sum_{j=1}^{p} f_{\cdot j} = f_{i/j}$$

■ X e Y indep. $\Leftrightarrow f_{ij} = f_{i\cdot}\, f_{\cdot j} \Leftrightarrow n_{ij} = \dfrac{n_{i\cdot}\, n_{\cdot j}}{n}$

$$f_{i/j} = f_{i\cdot} \Leftrightarrow \underbrace{f_{i/j}\, f_{\cdot j}}_{f_{ij}} = f_{i\cdot}\, f_{\cdot j} \Leftrightarrow n f_{ij} = n f_{i\cdot}\, f_{\cdot j} \Leftrightarrow n_{ij} = n_{i\cdot} \cdot \frac{n_{\cdot j}}{n}$$

■ $m_{r0} = \sum_{i=1}^{k} f_{i\cdot}\, x_i^2$

$$m_{r0} = \sum_{i=1}^{k} \left(\sum_{j=1}^{p} f_{ij}\, x_i^r \right) = \sum_{i=1}^{k} x_i^2 \left(\sum_{j=1}^{p} f_{ij} \right) = \sum_{i=1}^{k} f_{i\cdot}\, x_i^r$$

■ $m_{rs} = m_{r0}\, m_{0s}$ (X e Y indep)

$$m_{rs} = \sum_{i=1}^{k} \sum_{j=1}^{p} f_{ij}\, x_i^r\, y_j^s = \sum_{i=1}^{k} \sum_{j=1}^{p} f_{i\cdot}\, f_{\cdot j}\, x_i^r\, y_j^s = \sum_{i=1}^{k} f_{i\cdot}\, x_i^r \left(\sum_{j=1}^{p} f_{\cdot j}\, y_j^s \right) = m_{r0}\, m_{0s}$$

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

$$\sigma_{xy} = \sum_{i=1}^{k}\sum_{j=1}^{p} f_{ij}(x_i - \bar{x})(y_j - \bar{y}) = \sum_{i}^{k}\sum_{j}^{p} f_{ij}(x_i y_j - \bar{x} y_j - \bar{y} x_i + \bar{x}\bar{y}) =$$

$$= \sum_{i}^{k}\sum_{j}^{p} f_{ij} x_i y_j - \sum_{i}^{k}\sum_{j}^{p} f_{ij} \bar{x} y_j - \sum_{i}^{k}\sum_{j}^{p} f_{ij} \bar{y} x_i + \sum_{i}^{k}\sum_{j}^{p} f_{ij} \bar{x}\bar{y} =$$

$$= m_{11} - \bar{x}\bar{y} - \bar{x}\bar{y} + \bar{x}\bar{y} = m_{11} - \bar{x}\bar{y} = m_{11} - m_{10}m_{01}$$

$$\left.\begin{array}{l} \dfrac{\partial \varphi(a,b)}{\partial a} = 0 \\ \dfrac{\partial \varphi(a,b)}{\partial b} = 0 \end{array}\right\} \Rightarrow \begin{array}{l} m_{11} = a m_{20} + b m_{10} \\ m_{01} = a m_{10} + b \end{array} \Rightarrow a = \frac{\sigma_{xy}}{\sigma_x^2}, \quad b = \bar{y} - \frac{\sigma_{xy}}{\sigma_x^2} \cdot \bar{x}$$

$$\left.\begin{array}{l} \dfrac{\partial \varphi(a,b)}{\partial a} = 0 \\ \dfrac{\partial \varphi(a,b)}{\partial b} = 0 \end{array}\right\} \begin{array}{l} \sum_i\sum_j f_{ij}(y_j - a x_i - b)x_i = 0 = \sum_i\sum_j f_{ij} x_i y_j - a \sum_i\sum_j f_{ij} x_i^2 - b \sum_i\sum_j f_{ij} x_i \\ \sum_i\sum_j f_{ij}(y_j - a x_i - b) = 0 = \sum\sum f_{ij} y_j - a \sum\sum f_{ij} x_i - b \sum\sum f_{ij} \end{array} \Rightarrow$$

$$\Rightarrow \left\{\begin{array}{l} \sum\sum f_{ij} x_i y_j = a \sum\sum f_{ij} x_i^2 + b \sum\sum f_{ij} x_i \\ \sum\sum f_{ij} y_j = a \sum\sum f_{ij} x_i + b \sum\sum f_{ij} \end{array}\right\} \Rightarrow \left\{\begin{array}{l} m_{11} = a m_{20} + b m_{10} \\ m_{01} = a m_{10} + b \end{array}\right\}$$

$\sigma_x^2 = m_{20} - m_{10}^2 \neq 0 \Rightarrow$ Tiene sol y es única

$$a = \frac{m_{11} - m_{10}m_{01}}{m_{20} - m_{10}^2} = \frac{\sigma_{xy}}{\sigma_x^2}$$

$$b = m_{01} - \frac{\sigma_{xy}}{\sigma_x^2} m_{10} = \bar{y} - \frac{\sigma_{xy}}{\sigma_x^2}\bar{x}$$

Buscar $f(x_1), \ldots, f(x_k)$: $\sum_i^k \sum_j^p f_{ij} (y_j - f(x_i))^2$ sea mínima

$$\sum_i^k \sum_j^p f_{ij} (y_j - f(x_i))^2 = \sum \sum f_{j/i} f_{i\cdot} (y_j - f(x_i))^2 = \sum_{i=1}^k f_{i\cdot} \left[\sum_{j=1}^p f_{j/i} (y_j - f(x_i))^2 \right]$$

Minimizar $\sum_j^p f_{j/i} (y_j - f(x_i))^2$, $\forall i = 1, \ldots, k$

↳ Media de desviaciones cuadr. de Y respecto de $f(x_i)$ en las desv. cuadr. $Y/X = x_i$

↓

$$f(x_i) = \bar{y}_i = \sum_{j=1}^p f_{j/i} y_j$$

$$\sum_{i=1}^k \bar{y}_i f_{i\cdot} = \bar{y}$$

$$0 = \sum_i^k \sum_j^p f_{ij} (y_j - \bar{y}_i) = \sum_i^k \sum_j^p f_{ij} y_j - \sum_i^k \sum_j^p f_{ij} \bar{y}_i = m_{01} - \sum_i^k \bar{y}_i \left(\sum_j^p f_{ij} \right) = \bar{y} - \sum_{i=1}^k \bar{y}_i f_{i\cdot}$$

$$\sigma_{r_y}^2 = \sum_i^k \sum_j^p f_{ij} (r_{ij}^{Ly})^2 \quad \Big| \quad \sum^k f_{i\cdot} \left(\bar{y} + \frac{\sigma_{xy}}{\sigma_x^2}(x_i - \bar{x})\right) = \sum^k f_{i\cdot} \bar{y} + \sum^k f_{i\cdot} \frac{\sigma_{xy}}{\sigma_x^2} (x_i - \bar{x}) = \bar{y}$$

$$(r_{ij}^y = y_i - \bar{y}_i \ , \quad \forall i = 1, \ldots, k \ , \quad \forall j = 1, \ldots, p)$$

$$\sigma_{r_y}^2 = \sum_i^k \sum_j^p f_{ij} (r_{ij}^y)^2 \overset{\text{error cuadr. medio}}{=} \sum_i^k \sum_j^p f_{ij} (y_j - \bar{y}_i)^2 = \sum_i^k \sum_j^p f_{j/i} f_{i\cdot} (y_j - \bar{y}_i)^2 =$$

$$= \sum_i^k f_{i\cdot} \underbrace{\sum_j^p f_{j/i} (y_j - \bar{y}_i)^2}_{\text{Varianza condic}} = \underbrace{\sum_{i=1}^k f_{i\cdot} \sigma_{y,i}^2}_{\text{Media de las varianzas condic}}$$

Los resid. lin. mín. cuad. tienen media 0 y su varianza es el error cuad. medio

$$\sum_i^k \sum_j^p f_{ij} r_{ij}^{Ly} = \sum_i^k \sum_j^p f_{ij} \left(y_j - \bar{y} + \frac{\sigma_{xy}}{\sigma_x^2}(x_i - \bar{x})\right) = \sum^k \sum^p f_{ij} y_j - \sum^k \sum^p f_{ij} \frac{\sigma_{xy}}{\sigma_x^2} (x_i - \bar{x}) - \bar{y} \sum^k \sum^p f_{ij} =$$

$$= \bar{y} - \bar{y} - \frac{\sigma_{xy}}{\sigma_x^2} \sum^k \sum^p f_{ij} \overbrace{(x_i - \bar{x})}^{0 \text{ (media)}} = 0$$

$$\Rightarrow \underset{\uparrow}{\sigma_{r_y^L}^2} = \sum_i^k \sum_j^p f_{ij} (r_{ij}^{Ly} \underset{\uparrow}{- 0})^2 = \sum^k \sum^p f_{ij} (r_{ij}^{Ly})^2 = ECM$$

varianza de resid.

xq la media de los resid. es 0

Varianza valores observ. = varianza val. ajustados + varianza resid.

$$\sigma_y^2 = \sigma_{ey^*}^2 + \sigma_{ry^*}^2$$

- $\sigma^2_{ey^*} = \sum_{i=1}^{k} f_{i\cdot}\left(\bar y + \frac{\sigma_{xy}}{\sigma_x^2}(x_i-\bar x) - \bar y\right)^2 = \sum_i f_{i\cdot}\frac{\sigma_{xy}^2}{(\sigma_x^2)^2}(x_i-\bar x)^2 =$

$$= \frac{\sigma_{xy}^2}{(\sigma_x^2)^2}\sum_i^k f_{i\cdot}(x_i-\bar x)^2 = \frac{\sigma_{xy}^2}{(\sigma_x^2)^{\cancel{2}}}\cancel{\sigma_x^2} = \frac{\sigma_{xy}^2}{\sigma_x^2} \Rightarrow \sigma^2_{ey^*} = \frac{\sigma^2_{xy}}{\sigma^2_x}$$

- $\sigma^2_{ry^*} = \sum_i^k\sum_j^r f_{ij}\left(y_j - \bar y - \frac{\sigma_{xy}}{\sigma_x^2}(x_i-\bar x)\right)^2 =$

$$= \sum^k\sum^r f_{ij}\left[(y_j-\bar y)^2 + \left(\frac{\sigma_{xy}}{\sigma_x^2}(x_i-\bar x)\right)^2 - 2(y_j-\bar y)\left(\frac{\sigma_{xy}}{\sigma_x^2}(x_i-\bar x)\right)\right] =$$

$$= \sum^k\sum^r f_{ij}(y_j-\bar y)^2 + \sum_i^k\sum_j^r f_{ij}\left(\frac{\sigma_{xy}}{\sigma_x^2}\right)^2(x_i-\bar x)^2 - 2\sum^k\sum^r f_{ij}(y_j-\bar y)\left(\frac{\sigma_{xy}}{\sigma_x^2}(x_i-\bar x)\right) =$$

$$= \sigma_y^2 + \frac{\sigma_{xy}^2}{(\sigma_x^2)^2}\sigma_x^2 - 2\frac{\sigma_{xy}}{\sigma_x^2}\sigma_{xy} = \sigma_y^2 + \frac{\sigma_{xy}^2}{\sigma_x^2} - 2\frac{\sigma_{xy}^2}{\sigma_x^2} = \sigma_y^2 - \frac{\sigma_{xy}^2}{\sigma_x^2} \Rightarrow$$

$$\Rightarrow \sigma^2_{xy^*} = \sigma_y^2 - \frac{\sigma^2_{xy}}{\sigma^2_x}$$

X, Y indep $\Rightarrow \sigma_{xy} = 0$

$$\sigma_{xy} = \sum\sum f_{i\cdot}f_{\cdot j}(x_i-\bar x)(y_j-\bar y) = \sum f_{i\cdot}\cancelto{0}{(x_i-\bar x)}\left[\sum f_{\cdot j}\cancelto{0}{(y_j-\bar y)}\right] = 0$$

$\sigma_y^2 = \sigma_{res}^2 + \sigma_{expl}^2$ (Caso lineal)

$$\sigma_y^2 = \sum_j f_{\cdot j}(y_j-\bar y)^2 = \sum_j f_{\cdot j}(\underbrace{y_j - ax_i - b}_{r_{ij}} + \underbrace{ax_i + b - \bar y}_{\hat y_j - \bar y})^2 =$$

$$= \sum f_{\cdot j}(y_j-\hat y_j)^2 + \sum f_{\cdot j}(\hat y_j-\bar y)^2 + \cancelto{0}{2\sum f_{\cdot j}(y_j-\hat y_j)(\hat y_j-\bar y)} =$$

$$= \sum f_{\cdot j}(r_{ij}-0)^2 + \sigma^2_{expl} = \sigma^2_{res} + \sigma^2_{expl}$$

$$r^2 = \frac{\sigma_{xy}^2}{\sigma_x^2 \sigma_y^2}$$

$$\sigma_{res}^2 = \frac{1}{n} \sum\sum n_{ij} (y_j - ax_i - b)^2 = \sum\sum f_{ij} \left(y_j - \frac{\sigma_{xy}}{\sigma_x^2} x_i - \bar{y} + \frac{\sigma_{xy}}{\sigma_x^2} \bar{x}\right)^2 =$$

$$= \sum\sum f_{ij} \left[(y_j - \bar{y}) - \frac{\sigma_{xy}}{\sigma_x^2}(x_i - \bar{x})\right]^2 =$$

$$= \sum\sum f_{ij} (y_j - \bar{y})^2 + \sum\sum f_{ij} \frac{\sigma_{xy}^2}{(\sigma_x^2)^2} (x_i - \bar{x})^2 - 2 \sum\sum f_{ij} \frac{\sigma_{xy}}{\sigma_x^2} (x_i - \bar{x})(y_j - \bar{y}) =$$

$$= \sigma_y^2 + \frac{\sigma_{xy}^2}{(\sigma_x^2)^2} \sigma_x^2 - 2 \frac{\sigma_{xy}^2}{\sigma_x^2} = \sigma_y^2 - \frac{\sigma_{xy}^2}{\sigma_x^2}$$

$$\eta^2_{Y/X} = 1 - \frac{\sigma_{res}^2}{\sigma_y^2} = 1 + \frac{\sigma_y^2 + \frac{\sigma_{xy}^2}{\sigma_x^2}}{\sigma_y^2} = \frac{\sigma_{xy}^2}{\sigma_x^2 \sigma_y^2} = r^2$$

$$\shortparallel$$

$$\eta^2_{X/Y}$$

¿μ_{rs} invariante en cambios de escala y origen?

Como $\mu_{rs} = \sum\sum f_{ij} (x_i - \bar{x})^r (y_j - \bar{y})^s$

$$x' = \frac{x - x_0}{a}$$

$$y' = \frac{y - y_0}{b}$$

Tenemos que:

$$\boxed{\mu_{r's'}} = \sum\sum f_{ij} (x_i' - \bar{x}')^r (y_j' - \bar{y}')^s = \sum\sum f_{ij} \left(\frac{x_i - x_0}{a} - \frac{\bar{x} - x_0}{a}\right)^r \left(\frac{y_j - y_0}{b} - \frac{\bar{y} - y_0}{b}\right)^s$$

$$= \sum\sum f_{ij} \left(\frac{x_i - \bar{x}}{a}\right)^r \left(\frac{y_j - \bar{y}}{b}\right)^s = \frac{1}{a^r b^s} \sum\sum f_{ij} (x_i - \bar{x})^r (y_j - \bar{y})^s =$$

$$= \boxed{\frac{1}{a^r b^s} \mu_{rs}}$$

¿r invariante ante cambios de escala y origen?

Como $r_{xy} = \dfrac{\sigma_{xy}}{\sigma_x \sigma_y} = \dfrac{\mu_{11}}{\sqrt{\mu_{20}\mu_{02}}}$

$$x' = \frac{x - x_0}{a}$$

$$y' = \frac{y - y_0}{b}$$

Tenemos que:

$$\mu_{r's'} = \sum \sum f_{ij} (x_i' - \bar{x}')^r (y_j - \bar{y}')^s = \frac{1}{a^r b^s} \mu_{rs}$$

$$\sigma_{x'y'} = \mu_{11} = \frac{1}{ab} \sigma_{xy}$$

Por lo que:

$$r_{x'y'} = \frac{\sigma_{x'y'}}{\sigma_{x'} \sigma_{y'}} = \frac{\frac{1}{ab} \sigma_{xy}}{+\sqrt{\frac{1}{a^2} \sigma_x^2} \cdot +\sqrt{\frac{1}{b^2} \sigma_y^2}} = \frac{\frac{1}{ab} \sigma_{xy}}{\frac{1}{|a|} \sigma_x \cdot \frac{1}{|b|} \sigma_y}$$

Si $a, b > 0 \Rightarrow$ r no varía

σ_{xy} sí varía

Si a ó $b < 0 \Rightarrow r$ varía (solo el signo)