MAT-041: Probabilidad y Estadística **Pauta de Corrección**
Certamen 1. Abril 5, 2019
Tiempo: 90 Minutos Francisca González y Felipe Osorio

---

1. Se desea calcular $r = \operatorname{cor}(\boldsymbol{x}, \boldsymbol{y})$. Notando que las observaciones centradas $u_i = x_i - \overline{x}$ y $z_i = y_i - \overline{y}$ son dadas por

$$\boldsymbol{u} = \{-5, 0, -5, 5, -5, 10, -5, 10, -5\}, \qquad \boldsymbol{z} = \{-3, -3, -3, 7, -3, 12, -3, -1, -3\},$$

luego

$$\begin{aligned}
&\sum_{i=1}^{9} (x_i - \overline{x})(y_i - \overline{y}) = \sum_{i=1}^{9} u_i z_i \\
&= (-5)(-3) + 0(-3) + (-5)(-3) + 5 \cdot 7 + (-5)(-3) + 10 \cdot 12 + (-5)(-3) + 10(-1) + (-5)(-3) \\
&= 15 + 0 + 15 + 35 + 15 + 120 + 15 - 10 + 15 = 220.
\end{aligned}$$

Sabemos que $\sum_{i=1}^{9} (x_i - \overline{x})^2 = 350$ y $\sum_{i=1}^{9} (y_i - \overline{y})^2 = 248$. De este modo,

$$\operatorname{cor}(\boldsymbol{x}, \boldsymbol{y}) = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2 \sum_{i=1}^{n} (y_i - \overline{y})^2}} = \frac{220}{\sqrt{350 \cdot 248}} = 0.7467$$

2. Sabemos que $\widehat{\alpha} = \overline{y} - \widehat{\beta}\,\overline{x}$, de este modo

$$e_i = y_i - \widehat{\alpha} - \widehat{\beta}\, x_i = y_i - \overline{y} - \widehat{\beta}(x_i - \overline{x}).$$

En efecto,

$$\sum_{i=1}^{n} e_i = \sum_{i=1}^{n} \big(y_i - \overline{y} - \widehat{\beta}(x_i - \overline{x})\big) = \sum_{i=1}^{n} (y_i - \overline{y}) - \widehat{\beta} \sum_{i=1}^{n} (x_i - \overline{x}) = 0.$$

Por otro lado,

$$\begin{aligned}
\sum_{i=1}^{n} e_i\, \widehat{y}_i &= \sum_{i=1}^{n} (y_i - \overline{y} - \widehat{\beta}(x_i - \overline{x}))(\overline{y} + \widehat{\beta}(x_i - \overline{x})) \\
&= \overline{y} \sum_{i=1}^{n} (y_i - \overline{y} - \widehat{\beta}(x_i - \overline{x})) + \widehat{\beta} \sum_{i=1}^{n} (y_i - \overline{y} - \widehat{\beta}(x_i - \overline{x}))(x_i - \overline{x}),
\end{aligned}$$

el primer término es cero pues, $\sum_{i=1}^{n} e_i = 0$. Es decir, tenemos que:

$$\sum_{i=1}^{n} e_i\, \widehat{y}_i = \widehat{\beta} \Big[\sum_{i=1}^{n} (y_i - \overline{y})(x_i - \overline{x}) - \widehat{\beta} \sum_{i=1}^{n} (x_i - \overline{x})^2 \Big].$$

Notando que $\widehat{\beta} = \operatorname{cov}(\boldsymbol{x}, \boldsymbol{y}) / s_x^2$, sigue el resultado deseado.

3. a) Sabemos

$$\mathrm{P}(A \cap B^c) = \mathrm{P}(A) - \mathrm{P}(A \cap B),$$

pero, si $B \subset A \Rightarrow A \cap B = B$, luego

$$\mathrm{P}(A \cap B^c) = \mathrm{P}(A) - \mathrm{P}(B).$$

3. b) Tenemos que

$$\frac{\mathrm{P}(A)}{\mathrm{P}(A^c)} = \frac{\mathrm{P}(A)}{1 - \mathrm{P}(A)} = \frac{a}{b},$$

es decir

$$b\,\mathrm{P}(A) = a(1 - \mathrm{P}(A)) \quad \Rightarrow \quad (a+b)\,\mathrm{P}(A) = a,$$

de donde sigue que

$$\mathrm{P}(A) = \frac{a}{a+b}.$$

4. Tenemos que

$$\mathrm{P}(C^c) = \frac{99\,994}{100\,000} = 0.99994,$$

también

$$\mathrm{P}(T|C) = 1 - \mathrm{P}(T^c|C) = 1 - 0.16 = 0.84, \qquad \mathrm{P}(T^c|C^c) = 1 - \mathrm{P}(T|C^c) = 1 - 0.10 = 0.90.$$

a) De este modo, usando el teorema de probabilidad total:

$$\begin{aligned}\mathrm{P}(T) &= \mathrm{P}(C)\,\mathrm{P}(T|C) + \mathrm{P}(C^c)\,\mathrm{P}(T|C^c) = 0.00006 \cdot 0.84 + 0.99994 \cdot 0.10 \\ &= 0.1000444\end{aligned}$$

b) Se desea

$$\begin{aligned}\mathrm{P}(C|T) &= \frac{\mathrm{P}(C \cap T)}{\mathrm{P}(T)} = \frac{\mathrm{P}(C)\,\mathrm{P}(T|C)}{\mathrm{P}(T)} \\ &= \frac{0.00006 \cdot 0.84}{0.00006 \cdot 0.84 + 0.99994 \cdot 0.10} = \frac{504}{504 + 99\,994} \\ &= 0.000504.\end{aligned}$$

Es decir, por cada 1 millón de PAPs positivos sólo 504 son casos *verdaderos* de cáncer cervicouterino.