MODELOS LINEALES
Tercera Relación de Problemas

1. Consideremos tres variables cualesquiera X_1, X_2 y X_3. A partir de la definición de coeficiente de correlación parcial, calcular $r_{12|3}$.

2. Consideremos el modelo de regresión lineal múltiple

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i \quad ; \quad i = 1, \ldots, N$$

en el cual se verifica:

$$\begin{cases} \mathrm{E}\left[\epsilon_i\right] = 0 \;\; ; \;\; i = 1, \ldots, N \\ \\ \mathrm{Cov}\left[\epsilon_i, \epsilon_j\right] = 0 \;\; ; i, j = 1, \ldots, N \;\; i \neq j \\ \\ \mathrm{Var}\left[\epsilon_i\right] = \sigma^2 \;\; ; \;\; i = 1, \ldots, N \end{cases}$$

y supongamos que se verifica la relación funcional

$$x_{i1} = \lambda x_{i2} \quad i = 1, \ldots, N \quad ; \quad \lambda \neq 0$$

a) Construir la matriz de diseño, X. ¿Cuál es su rango?

b) ¿Existe en este caso solución única para el sistema de ecuaciones normales? ¿Por qué?

3. Consideremos el modelo de regresión lineal múltiple univariante

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + \epsilon_i \;\; ; \;\; i = 1, \ldots, N$$

en el que se suponen las hipótesis usuales.

Se pretende estudiar el comportamiento de dicho modelo ante cambios de escala en las variables que intervienen en él. Para ello se multiplica cada variable explicativa por escalares $\lambda_1, \ldots, \lambda_k$, $(\lambda_i \neq 0)$ así como la variable dependiente por λ, $(\lambda \neq 0)$, obteniéndose así un nuevo modelo

$$y_i^* = \beta_0^* + \beta_1^* x_{i1}^* + \cdots + \beta_k^* x_{ik}^* + \epsilon_i^* \;\; ; \;\; i = 1, \ldots, N$$

donde $y_i^* = \lambda y_i$ y $x_{ij}^* = \lambda_j x_{ij}$, $i = 1, \ldots, N$; $j = 1, \ldots, k$.

a) Calcula el estimador del vector de parámetros $\beta^* = (\beta_0^*, \ldots, \beta_k^*)'$ del modelo transformado en función del estimador del vector $\beta = (\beta_0, \ldots, \beta_k)'$ del modelo original.

b) Calcula el coeficiente de determinación del modelo transformado y compáralo con el del modelo original.

4. Consideremos el modelo de regresión lineal múltiple sin término constante $\mathbf{y} = \mathbf{X}_1 \beta_* + \varepsilon$ (donde hemos supuesto las hipótesis usuales).

- ¿Es ahora cierto que $(\mathbf{y} - \mathbf{1}\overline{y})'(\mathbf{y} - \mathbf{1}\overline{y}) = (\widehat{\mathbf{y}} - \mathbf{1}\overline{y})'(\widehat{\mathbf{y}} - \mathbf{1}\overline{y}) + \mathbf{e}'\mathbf{e}$? ¿Por qué?
- ¿Qué expresiones adoptan en este caso las distintas variabilidades así como el coeficiente de determinación?

5. Consideremos el modelo de regresión lineal múltiple $\mathbf{y} = \mathbf{X}\beta + \varepsilon$ en el que se suponen las hipótesis usuales. No obstante, muchos programas de ordenador no permiten estimar el modelo sin término constante $\mathbf{y} = \mathbf{X}_1\beta_* + \varepsilon$ (donde hemos considerado $X = (\mathbf{1}|\mathbf{X}_1)$ y del vector $\beta = (\beta_0|\beta_*)')$. Para conseguir tal estimación se sugiere trabajar con el modelo con término constante $\mathbf{y} = \mathbf{X}\beta + \varepsilon$, pero con el doble de observaciones, introduciendo los valores de cada variable y sus simétricos.

   a) Empleando esta estrategia, estima el vector de parámetros del modelo $\mathbf{y} = \mathbf{X}\beta + \varepsilon$. Compara dicha estimación con la que se obtendría en el modelo sin término constante $\mathbf{y} = \mathbf{X}_1\beta_* + \varepsilon$.

   b) Obtén la expresión de la variabilidad no explicada y compárala con la del modelo sin término constante.