Modelos no paramétricos y de regresión/Estadística II | Semestre 2019-1

Tarea 2

Fecha de entrega: 29 de agosto

1. Suponer que se ajusta un modelo RLS por MCO a partir de las observaciones (x_i, y_i), $i = 1, \ldots, n$. Verificar que se cumplen las siguientes igualdades

   a) $\sum_{i=1}^{n} e_i = 0$,

   b) $\sum_{i=1}^{n} x_i e_i = 0$,

   c) $\sum_{i=1}^{n} \hat{y}_i e_i = 0$,

   d) $\sum_{i=1}^{n} \hat{y}_i = \sum_{i=1}^{n} y_i$,

   e) $\sum_{i=1}^{n} y_i e_i = \sum_{i=1}^{n} e_i^2$,

   donde e_i es el i-ésimo residuo, es decir, $e_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$, $i = 1, \ldots, n$.

2. Considerar el modelo RLS con errores normales $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$, $i = 1, \ldots, n$, y los estimadores de MV. ¿Qué pasa con $\hat{\beta}_0$ y $\hat{\beta}_1$ cuando se aplican las siguientes transformaciones a los datos?

   a) $Y^* = Y + c$, $c \in \mathbb{R}$.

   b) $Y^* = cY$, $c \in \mathbb{R} \setminus \{0\}$.

   c) $Y^* = cY + d$, $c \in \mathbb{R} \setminus \{0\}$ y $d \in \mathbb{R}$.

   d) $X^* = X + c$, $c \in \mathbb{R}$

   e) $X^* = cX$, $c \in \mathbb{R} \setminus \{0\}$.

   f) $X^* = cX + d$, $c \in \mathbb{R} \setminus \{0\}$ y $d \in \mathbb{R}$.

3. Con los datos de desarrollo humano de las entidades del país (que pueden descargar aquí), considerar un modelo RLS para explicar la esperanza de vida con el logaritmo del ingreso:

   a) Calcular los EMV de β_0, β_1 y σ^2. Reportar los valores de las expresiones utilizadas (promedios, sumas de cuadrados o productos cruzados).

   b) Estimar las varianzas de los estimadores de β_0 y β_1 del inciso anterior.

   c) Interpretar los resultados en el contexto de los datos.

   d) Calcular los intervalos de confianza 90 % para β_0 y β_1. interpretar los intervalos calculados.

   e) Calcular un intervalo de confianza 90 % para σ^2. Reportar los cuantiles utilizados.

   - **Punto extra:** Encontrar el intervalo de confianza 90 % de menor longitud para σ^2. Describir con detalle el procedimiento a seguir para encontrar tal intervalo.

4. En el modelo RLS con errores normales mostrar que:

   a) $Cov\left(\hat{\beta}_0, \hat{\beta}_1\right) = -\sigma^2 \frac{\bar{x}_n}{S_{xx}}$.

   b) $\sum_{i=1}^{n} (y_i - \bar{y}_n)^2 = \sum_{i=1}^{n} (\hat{y}_i - \bar{y}_n)^2 + \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$.