PRÁCTICA 3

(1) Considera el comportamiento de la mortalidad infantil (MI) en relación con el PIB per cápita (PIBpc).

MI = mortalidad infantil, el número de defunciones de niños menores de 5 años en un año por cada 1 000 nacidos vivos.

PIBPC = PIB per cápita en 1980 (USD)

(a) Pide a R el resumen de los estimadores del modelo de regresión lineal. Escribe los estimadores $\hat{\beta}_1$ y $\hat{\beta}_2$ para el modelo

$$MI_i = \beta_1 + \beta_2 \ PIBpc_i + u_i$$

(b) Interpreta los estimadores.

(c) Calcula los intervalos de confianza para los estimadores, con una confianza del 90%, usando las siguientes fórmulas,

$$\hat{\beta}_1 - t_{\alpha/2} e.e.(\hat{\beta}_1) \leq \beta_1 \leq \hat{\beta}_1 + t_{\alpha/2} e.e.(\hat{\beta}_1)$$

$$\hat{\beta}_2 - t_{\alpha/2} e.e.(\hat{\beta}_2) \leq \beta_2 \leq \hat{\beta}_2 + t_{\alpha/2} e.e.(\hat{\beta}_2)$$

donde t es la distribución t con $n-2$ grados de libertad.
Compara tus resultados con los intervalos de confianza que calcula R, usando la función $confint()$.

(d) Realiza el diagrama de dispersión, con la línea de regresión.

(e) Determina si el PIBpc ejerce un impacto negativo o positivo sobre la MI. Da una explicación de por qué era de esperarse este resultado.

(f) Escribe el coeficiente de determinación R^2 del modelo.

(2) Considera ahora también como variable independiente, el alfabetismo femenino, medido por la tasa de alfabetización de las mujeres (TAM). A priori, se espera que la TAM también ejerza un impacto negativo en la MI. Cuando se introducen ambas variables en el modelo, el modelo es:

$$MI_i = \beta_1 + \beta_2 \ PIBpc_i + \beta_3 TAM_i + u_i$$

Ten en cuenta que la MI es el número de muertes de niños menores de 5 años por cada 1 000 nacidos vivos, el PIBPC es el PIB per cápita en 1980 (USD) y la TAM se mide en porcentaje. La muestra se realizó en 64 países.

(a) Pide a R el resumen de los estimadores del modelo de regresión lineal. Comenta si los signos de los estimadores de los coeficientes son los esperados.

(b) Escribe los estimadores numéricos de $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$ para el modelo.

(c) Interpreta los valores numéricos de los estimadores de acuerdo a este modelo.

(d) Manteniendo constante $PIBpc$, ¿cuánto tendría que aumentar TAM para tener una reducción de una muerte de un niño menor de 5 años por cada 1 000 nacidos vivos predichos?

(e) Para un país que tiene un PIB per cápita de 680 USD, y una tasa de alfabetización de mujeres del 54% ¿Cuánto predice el modelo que, en promedio, sería el número de muertes de niños menores de 5 años por cada 1 000 nacidos vivos?

(f) Suponga que un país A tiene un PIB per cápita de 3870 USD, y una tasa de alfabetización de mujeres del 84%. Por otro lado, un país B tiene un PIB per cápita de 1100 USD, y una tasa de alfabetización de mujeres del 65% ¿Cuál es la diferencia entre B y A en muertes de niños menores de 5 años por cada 1 000 nacidos vivos?

(g) Escribe el error estándar para cada estimador.

(h) Escribe el coeficiente de determinación R^2 y el coeficiente ajustado R^2_{adj}. Calcula R^2_{adj} a partir del valor de R^2, con la fórmula vista en clase.

(i) Calcula los intervalos de confianza para los estimadores con una confianza del 95%, usando las siguientes fórmulas,

$$\hat{\beta}_1 - t_{\alpha/2} e.e.(\hat{\beta}_1) \leq \beta_1 \leq \hat{\beta}_1 + t_{\alpha/2} e.e.(\hat{\beta}_1)$$

$$\hat{\beta}_2 - t_{\alpha/2} e.e.(\hat{\beta}_2) \leq \beta_2 \leq \hat{\beta}_2 + t_{\alpha/2} e.e.(\hat{\beta}_2)$$

$$\hat{\beta}_3 - t_{\alpha/2} e.e.(\hat{\beta}_3) \leq \beta_3 \leq \hat{\beta}_3 + t_{\alpha/2} e.e.(\hat{\beta}_3)$$

donde t es la distribución t con $n-3$ grados de libertad.
Compara tus resultados con los intervalos de confianza que calcula R, usando la función $confint()$.

(j) Calcula los intervalos de confianza para los estimadores con una confianza del 90%, usando las siguientes fórmulas,

$$\hat{\beta}_1 - t_{\alpha/2} e.e.(\hat{\beta}_1) \leq \beta_1 \leq \hat{\beta}_1 + t_{\alpha/2} e.e.(\hat{\beta}_1)$$

$$\hat{\beta}_2 - t_{\alpha/2} e.e.(\hat{\beta}_2) \leq \beta_2 \leq \hat{\beta}_2 + t_{\alpha/2} e.e.(\hat{\beta}_2)$$

$$\hat{\beta}_3 - t_{\alpha/2} e.e.(\hat{\beta}_3) \leq \beta_3 \leq \hat{\beta}_3 + t_{\alpha/2} e.e.(\hat{\beta}_3)$$

donde t es la distribución t con $n-3$ grados de libertad.
Compara tus resultados con los intervalos de confianza que calcula R.