Tarea 3

Econometría I (ECO3404)
Primer Semestre 2025
Universidad Anáhuac

Analiza cómo varía el peso de los niños (variable dependiente) con su altura (variable independiente) utilizando una regresión lineal simple y la base de datos del libro *Introducción a la Ecología Experimental* de T. Lewis y L.R. Taylor (1967). En la base, el peso está en libras (1 libra = 0.45 kg), la altura en pulgadas (1 pulgada = 2.54 cm) y la edad en meses.

I. En una hoja de Excel realiza lo siguiente:[2]

1. En celdas. Calcula el promedio de las variables dependiente e independiente: $\bar{y}$, $\bar{x}$.
2. En columna. Calcula la desviación de cada observación de la variable dependiente con respecto a su media: $y_i - \bar{y}$.
3. En columna. Calcula la desviación de cada observación de la variable independiente con respecto a su media: $x_i - \bar{x}$.
4. En columna. Calcula el producto de las desviaciones de las variables dependiente e independiente con respecto a sus medias: $(y_i - \bar{y})(x_i - \bar{x})$.
5. En columna. Calcula el cuadrado de la desviación de cada observación de la variable independiente con respecto a su media: $(x_i - \bar{x})^2$.
6. En celda. Calcula el estimado para β_1 dividiendo la suma de la columna del paso 4 entre la suma de la columna del paso 5: $\hat{\beta}_1 = \sum (y_i - \bar{y})(x_i - \bar{x}) / \sum (x_i - \bar{x})^2$.
7. En celda. Calcula el estimado para β_0 utilizando los valores obtenidos en los pasos 1 y 6: $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$.
8. En columna. Obtén la predicción para la variable dependiente $\hat{y}$ utilizando los estimados $\hat{\beta}_0$ y $\hat{\beta}_1$ obtenidos en los pasos 6 y 7: $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$.
9. En columna. Obtén los residuales $\hat{u}$ como la resta de la variable dependiente observada y la ajustada obtenida en el paso 8: $\hat{u}_i = y_i - \hat{y}_i$.
10. En columna. Obtén el cuadrado de la desviación de la variable dependiente ajustada (paso 8) con respecto a la media de la variable dependiente (paso 1): $(\hat{y}_i - \bar{y})^2$.
11. En columna. Obtén el cuadrado de cada uno de los residuales: $\hat{u}_i^2$.
12. En columna. Obtén el cuadrado de la desviación de la variable dependiente observada con respecto a su media (paso 1): $(y_i - \bar{y})^2$.

[2] Revisa las fórmulas vistas en clase para confirmar cómo se obtienen los diferentes valores.

13. En columna. Obtén el cuadrado de cada observación de la variable independiente: x_i^2.
14. En celda. Obtén la suma de cuadrados explicada (SCE) sumando la columna del paso 10.
15. En celda. Obtén la suma de cuadrados de los residuales (SCR) sumando la columna del paso 11.
16. En celda. Obtén la suma de cuadrados total (SCT) sumando la columna del paso 12. En otra celda, verifica que se obtiene lo mismo al sumar las celdas de los pasos 14 y 15.
17. En celda. Calcula el coeficiente de determinación R^2 dividiendo los valores de los pasos 14 y 16: $R^2 = SCE/SCT$. En otra celda, verifica que se obtiene el mismo valor si se calcula utilizando los valores de los pasos 15 y 16: $R^2 = 1 - SCR/SCT$.
18. En celda. Calcula el número de observaciones: n.
19. En celda. Calcula los grados de libertad: $n - 2$.
20. En celda. Calcula la R^2 ajustada: $R_{aj}^2 = 1 - (1 - R^2) * (n - 1)/(n - 2)$.
21. En celda. Estima la varianza del error σ^2 dividiendo los valores en los pasos 15 y 19: $\hat{\sigma}^2 = SCR/(n - 2)$.
22. En celda. Calcula el error estándar de la regresión (también conocido como la raíz del error cuadrático medio) como la raíz del valor obtenido en el paso 21: $\hat{\sigma} = \sqrt{\hat{\sigma}^2}$.
23. En celda. Estima la varianza de $\hat{\beta}_0$ dividiendo dos productos. El numerador es el producto de $\hat{\sigma}^2$ (paso 21) y la suma de la columna del paso 13. El denominador es el producto del número de observaciones (paso 18) y la suma de la columna del paso 5.
24. En celda. Estima la varianza de $\hat{\beta}_1$ dividiendo $\hat{\sigma}^2$ (paso 21) entre la suma de la columna del paso 5.
25. En celdas. Obtén los errores estándar de $\hat{\beta}_0$ y $\hat{\beta}_1$ como la raíz cuadrada de los valores en los pasos 23 y 24.

II. Una vez que obtengas los valores anteriores, verifica que tus resultados:

- Son iguales a lo que genera el comando LINEST o ESTIMACION.LINEAL de Excel.
- Satisfacen las siguientes propiedades: $\bar{y} = \bar{\hat{y}}$, $\sum \hat{u}_i = 0$, $corr(x, \hat{u}) = 0$, $\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}$.

III. Gráfica $\hat{u}$ del paso 9 (eje y) contra los valores ajustados $\hat{y}$ del paso 8 (eje x) usando un diagrama de dispersión, ¿la varianza de los residuales parece constante?

IV. Sin repetir los pasos anteriores, ¿cómo cambiarían $\hat{\beta}_0$, $\hat{\beta}_1$ y sus errores estándar si quisiéramos expresar los resultados en kilogramos y centímetros? ¿Cómo cambia el peso en kg con 1 cm adicional?