

Lista de ejercicios 2 - Análisis de regresión

Ciclo nivelación 2025-2

Mg. Sc. J. Eduardo Gamboa U.

En un curso universitario de Estadística General, se busca analizar qué factores influyen en el tiempo que un estudiante tarda en resolver un examen.

Antes del examen, los estudiantes rindieron una prueba diagnóstica para medir conocimientos previos. Además, se registraron las horas de estudio realizadas la semana anterior al examen, y se aplicó un breve cuestionario de ansiedad ante evaluaciones.

Durante el examen final se registró el tiempo de resolución en minutos. Los datos se encuentran en el archivo **Lista2_datos.csv**

Las variables son:

- `time_minutes`: tiempo de resolución del examen final (minutos)
- `diagnostic_score`: puntaje en la prueba diagnóstica
- `prior_study_hours`: tiempo de estudio (horas)
- `anxiety_level`: nivel de ansiedad (escala ordinal de 1 a 10)

1. Presentar y comentar un gráfico que permite ver las asociaciones entre las variables en estudio.
2. Formular el modelo de regresión lineal.
3. Presentar el modelo de regresión lineal estimado.
4. Analizar el cumplimiento del supuesto de normalidad de errores.
5. Analizar el cumplimiento del supuesto de homocedasticidad de errores.
6. Analizar el cumplimiento del supuesto de independencia de errores.
7. En caso de incumplimiento de modelo, aplicar las siguientes transformaciones y comparar sus efectos: Raíz cuadrada, logaritmo, Box Cox, Mínimos cuadrados ponderados.
8. Presentar un cuadro comparativo y elegir el mejor modelo. Resolver las preguntas que siguen utilizando dicho mejor modelo

9. Escribir la ecuación estimada del mejor modelo e interpretar sus coeficientes.
10. Ejecutar la prueba de hipótesis global del modelo, presentando el cuadro ANVA.
11. Ejecutar las pruebas de hipótesis individuales
12. Interpretar el coeficiente de determinación ajustado
13. Estimar el tiempo mediano del examen final si el puntaje diagnóstico fue de 65 puntos, estudió 10 horas y el nivel de ansiedad es 5.5, e indicar si coincide con el tiempo medio. Además, reportar el intervalo de confianza.
14. Predecir el tiempo que tardará en resolver el examen final un estudiante con puntaje diagnóstico de 50 puntos, que estudió 4.5 horas y con nivel de ansiedad 8. Compararlo con uno que tuvo 85 puntos diagnósticos, 12 horas de estudio y nivel de ansiedad 2.3