

# Lista de ejercicios 1 - Análisis de regresión

Ciclo nivelación 2025-2

Mg. Sc. J. Eduardo Gamboa U.

1. Se utilizará el dataset cars, disponible en R, que contiene observaciones de:

- speed: velocidad del automóvil (millas por hora)
- dist: distancia de frenado (pies)

Considerar  $\alpha = 0.05$  cuando sea necesario.

- a) Estimar e interpretar el coeficiente de correlación de Pearson entre la velocidad y la distancia de frenado.
- b) ¿Se puede concluir que la correlación es estadísticamente distinta de cero?
- c) Formular el modelo de regresión lineal simple que relaciona la distancia de frenado con la velocidad.
- d) Estimar de manera puntual los coeficientes del modelo e interpretarlos.
- e) Construir e interpretar un intervalo de confianza para la pendiente del modelo.
- f) ¿Existe una relación lineal significativa entre la velocidad y la distancia?
- g) ¿Puede afirmarse que por cada milla por hora adicional, al automóvil le toma más de 4 pies frenar?
- h) Construir un intervalo de confianza para la desviación estándar del error del modelo.
- i) ¿Qué porcentaje de la variabilidad de la distancia es explicada por la velocidad?
- j) Estimar la distancia media de frenado cuando la velocidad es de 18.45 millas por hora.

2. Se utilizará el dataset diamonds, disponible en el paquete ggplot2 R, que contiene observaciones de precios y otros atributos de 54 mil diamantes. Cuando sea necesario, considerar  $\alpha = 0.01$ :
- price: precio en dólares
  - carat: peso del diamante
  - cut: calidad del corte (Fair, Good, Very Good, Premium, Ideal)
  - color: color del diamante desde D (mejor) hasta J (peor)
  - clarity: medida claridad del diamante: I1 (peor), SI2, SI2, VS2, VS1, VVS2, VVS1, IF (mejor)
  - x: longitud en milímetros
  - y: ancho en milímetros
  - z: profundidad en milímetros
  - depth: porcentaje total de profundidad  $= 2 \cdot z / (x + y)$
  - table: ancho de la parte superior del diamante en relación con el punto más ancho (43–95)
- a) Estimar e interpretar el coeficiente de correlación de Pearson entre el precio del diamante y al menos dos de las variables independientes. Sugerencia: Usar una matriz y/o gráfica de correlaciones. Ver [aquí](#) algunas gráficas.
- b) ¿Se puede concluir que la correlación entre el precio y la profundidad del diamante es estadísticamente menor a cero?
- c) Formular el modelo de regresión lineal simple que relaciona el precio del diamante con su peso, longitud, ancho y profundidad.
- d) Estimar de manera puntual los coeficientes del modelo e interpretarlos.
- e) Construir e interpretar un intervalo de confianza para el coeficiente del ancho.
- f) ¿Existe una relación lineal significativa entre el precio de un diamante y su peso, longitud, ancho y profundidad?
- g) ¿Puede afirmarse que por cada milímetro adicional de longitud, el precio promedio disminuye en 800 dólares?
- h) ¿Puede afirmarse que por cada milímetro adicional de ancho, el precio promedio aumenta como máximo 200 dólares?
- i) ¿Qué porcentaje de la variabilidad del precio es explicada por el modelo?
- j) Estimar de manera intervalar el precio medio de un diamante con peso 0.25, longitud 4 mm, ancho 4.1 mm y profundidad 2.7 mm.

- k) Predecir de manera intervalar el precio de un diamante con peso 0.25, longitud 4 mm, ancho 4.1 mm y profundidad 2.7 mm.