

Práctica 3: Regresión lineal simple

Módulo de Modelos Lineales.
Máster de Bioestadística, Universitat de València.

Miguel A. Martinez-Beneito

Tareas

1. Para el banco de datos **Auto** de la librería **ISLR**, queremos estudiar el consumo de distintos vehículos, variable **mpg** (miles per galon). Queremos estudiar esta variable en función de su potencia, variable **horsepower**. Para ello llevaremos a cabo un modelo de regresión lineal simple.
 - Evalúa la existencia de relación lineal entre ambas variables ¿Encuentras evidencia de que pudiera tener sentido resumir la relación entre ambas variables de forma lineal?
 - Ajusta la recta de regresión necesaria para resumir la relación entre ambas variables.
 - Representa gráficamente la relación entre ambas variables y la recta de regresión que has ajustado.
 - ¿Es la relación entre la variable y la respuesta positiva o negativa? Interpreta dicha relación.
 - ¿Encuentras que la relación entre ambas variables es significativa(mente distinta de 0)? Halla un intervalo de confianza al 95% para el coeficiente asociado a la potencia de los vehículos.
 - ¿Qué **mpg** predecirías para un **horsepower** de 98? Halla un intervalo de confianza para $E(\text{mpg}|\text{horsepower} = 98)$ y un intervalo de predicción para el valor de **mpg** correspondiente a una potencia de 98 caballos.
 - Representa la nube de puntos junto a la recta de regresión que has ajustado, así como un intervalo de confianza para dicha recta y un intervalo de predicción para el rango de valores de la variable **horsepower**.
2. Repite la tarea anterior, pero utilizando el año de fabricación del vehículo (**year**) como variable explicativa. Valora las diferencias entre las conclusiones que extraes de ambos análisis ¿Cuál de los dos ajustes te parece más satisfactorio?
3. La relación lineal entre **mpg** y **horsepower** ajustada en la Tarea 1 no resulta del todo satisfactoria ya que la nube de puntos se arquea en sus extremos. En ese caso podría parecer más adecuada una relación lineal del tipo $\text{mpg} \sim 1/\text{horsepower}$. Crea la variable $\text{invhorsepower}=1/\text{horsepower}$ y ajusta un modelo de regresión lineal simple para **mpg** empleando esta nueva variable como covariable. Representa el ajuste obtenido y valora si la transformación que has hecho de **horsepower** mejora dicho ajuste.
4. Plantéate la veracidad o falsedad de las siguientes afirmaciones:
 - Siempre que la correlación lineal entre dos variables sea ascendente, el coeficiente de la covariable en el correspondiente modelo de regresión lineal simple será positivo.
 - Podremos reducir la amplitud del intervalo de predicción de un modelo de regresión lineal simple tanto como queramos simplemente elevando el tamaño de la muestra.
 - Para eliminar la correlación entre los coeficientes del modelo de regresión lineal simple es suficiente centrar (restar su media) la covariable del modelo.
 - La suma de los residuos del modelo de regresión valdrá necesariamente 0.