

Práctica 6: Influencia y covariables categóricas

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

Miguel A. Martinez-Beneito

Tareas

1. Para el modelo que ajustaste en la Tarea 1 de la Práctica 4 (modelo de regresión lineal múltiple para `mpg` en función de todas las variables cuantitativas del banco de datos)
 - ¿Observas algún valor particularmente influyente para el modelo?
 - ¿Los valores más influyentes que observas tienen un `mpg` inferior o superior al que se esperaría de acuerdo con el modelo?
 - ¿Observas la existencia de algún vehículo potencialmente influyente en el modelo de regresión ajustado?
 - El modelo estudiado tiene un buen número de covariable no significativas ¿Crees que éstas pueden alterar de alguna manera el análisis que acabamos de hacer? ¿De qué forma?
2. Para el banco de datos `Auto`, considera el modelo que estimaras más adecuado de los que ajustaras en la práctica 4.
 - Incluye ahora en el modelo de regresión `origin` como variable categórica.
 - ¿Consideras que hay diferencias significativas en cuanto a `mpg` entre los coches europeos y japoneses?
 - Valora la presencia de interacción entre el origen de los vehículos y las variables que estuvieran ya anteriormente en el modelo. Interpreta los resultados obtenidos: ¿Qué región ha tenido una mejor evolución temporal en cuanto al consumo de sus coches? ¿En que región el peso de los coches tiene un efecto más importante en el consumo?
 - Sobre este último modelo que has ajustado ¿Qué vehículos consideras que han podido tener más influencia sobre el ajuste de la recta de regresión?
 - Por último, sobre el último modelo que has ajustado, incluye como variable explicativa el número de cilindros de cada coche junto a su interacción con la región de procedencia del vehículo ¿Consideras que la interacción con la región de fabricación tienen un efecto significativo en cuanto a la explicación de `mpg`?