Ejercicios

Lino AA Notarantonio

8/6/2019

**Ejercicio 1** Usa la función *which()* y determina la entrada de la observación con el valor máximo (resp., mínimo) de la variable en el dataset *cleve.csv*. Después, haz lo mismo para la variable .

**Ejercicio 2** Determina las observaciones de la variable en el dataset *cleve.csv* que sean mayores a

1. su media
2. su mediana

¿Cuántas observaciones hay en cada inciso?

**Ejercicio 3** Usa la función *apply()* para calcular la desviación estándar del dataframe formado por las variables , , , . Después, calcula la matriz de varianza-covarianza de estas variables, usando la función *var()*

**Ejercicio 4** Calcula una tabla de contingencia con los valores observados de las variables categóricas , . Usen una de las funciones que vimos en la sesión del 5 de agosto. Después, usen otra de las funciones para calcular una tabla de contingencia con las frecuencias relativas de estas variables.

**Ejercicio 5** Grafiquen las tablas de contingencias en el inciso anterior usando *barplot*, con la opción de gráficas agrupadas (*grouped barplots*). Después, grafiquen también un *mosaicplot* de estas variables.

**Ejercicio 6** Determina con un boxplot con *notch* si se puede detectar una diferencia entre, (edad de las mujeres con diagnóstico positivo) y (edad de los varones con diagnóstico positivo).

**Ejercicio 7** Prueba la hipótesis

(nivel de significancia: .)

Después, realiza una prueba de potencia con . Finalmente, calcula el tamaño muestral necesario con estos datos para obtener una potencia de .90.

**Ejercicio 8** Considera el modelo de regresión lineal (con datos *cleve.csv*)

Estima el modelo y después valídalo eligiendo de manera aleatoria el 80% de los datos para la estimación y el 20% de los datos para la validación *out of sample*.

**Ejercicio 9** Estima el modelo logit

Después, determina la bondad de la estimación comparando los valores de probabilidad observados con los valores predichos.

**Ejercicio 10** Los datos *AirPassengers* están cargados por defecto en R. Replica el análisis con el método de Holt-Winters para obtener una estimación de la serie de tiempo. Comparen los valores observados con los valores estimados usando los métodos RMSE, MAE, MAPE.

**Tip** para la comparación, tomen en cuenta que en la estimación de la serie de tiempo se pierden las primeras 12 observaciones de *AirPassengers*.