

Curso: Estadística Multivariada

Tarea 3

Fecha de entrega: 2 de marzo de 2018

Instrucciones

1. Los primeros 2 ejercicios se entregan en **pdf**
2. Los ejercicios restantes se entregarán con el script de **R** que generen y con un reporte en pdf que incluya los resultados obtenidos y las respuestas solicitadas.
3. El script debe correrse de tal manera que los datos (si ocupa) estén situados en la carpeta donde se encuentra el script
4. Los ejercicios de **R** se revisarán en base a los resultados provenientes de correr el script, por lo que verifiquen y marquen la parte donde se encuentra el resultado.
5. Los 2 primeros ejercicios, el script y el reporte se subirán a la plataforma

1. La matriz de datos para una muestra aleatoria de tamaño $n = 3$ de una población normal bivariada está dada por:

$$\mathbf{X} = \begin{pmatrix} 6 & 9 \\ 10 & 6 \\ 8 & 3 \end{pmatrix}$$

- (a) Evalua T^2
- (b) Verifica que T^2 permanece sin cambios si cada observación \mathbf{x}_j , $j = 1, 2, 3$ es reemplazada por $\mathbf{C}\mathbf{x}_j$, donde

$$\mathbf{C} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

Note que las observaciones

$$\mathbf{C}\mathbf{x}_j = \begin{pmatrix} x_{j1} - x_{j2} \\ x_{j1} + x_{j2} \end{pmatrix}$$

Producen la matriz

$$\begin{pmatrix} (6-9) & (10-6) & (8-3) \\ (6+9) & (10+6) & (8+3) \end{pmatrix}'$$

2. Dadas la siguiente muestra de observaciones bivariadas:

$$\mathbf{X} = \begin{pmatrix} 2 & 12 \\ 8 & 9 \\ 6 & 9 \\ 8 & 10 \end{pmatrix}$$

- (a) Evalúa T^2 , para probar $H_0 : \boldsymbol{\mu}' = [7, 11]$, usando los datos.
 - (b) Especifica la distribución de T^2 (verificando la normalidad de los datos)
 - (c) Usando (a) y (b) prueba H_0 en $\alpha = .05$ ¿Qué conclusión se tiene?
 - (d) Evalúa T^2 utilizando la relación que tiene con la lambda de Wilks
 - (e) Evalúa Λ y la lambda de Wilks.
3. Considera los datos de radiación con la puerta cerrada (archivo: *raddata*). Construye un Q-Q plot para los logaritmos naturales de estos datos. [Notar que la transformación logaritmo natural corresponde al valor $\lambda = 0$ en la transformación de Box-Cox]. Los datos transformados parecen ser normalmente distribuidos? Compara los resultados con los obtenidos si transformamos los datos con $\lambda = 1/4$. Existe mucha diferencia al usar $\lambda = 0$ ó $\lambda = 1/4$ en este caso?.
4. Considera los datos sobre contaminación del aire (archivo: *datoscontaminacion*).
- (a) Construye un Q-Q plot para las mediciones de radiación solar y realiza una prueba de normalidad basada en el coeficiente de correlación r_Q , para $\alpha = .05$.
 - (b) Examina el par $x_3 = CO$ y $x_4 = NO$ como una distribución bivariada. Construye un diagrama Chi-cuadrado y aplica la prueba de normalidad basada en r_Q . Se puede concluir normalidad bivariada en este conjunto de datos?
5. Examina la normalidad marginal y bivariada de las observaciones en las variables x_1, x_2, x_3 y x_4 para los datos del archivo *datosrigidez*.
6. Considera los datos sobre parques nacionales (*datosparquesnac*)
- (a) Comenta sobre cualquier posible outlier construyendo scatter plot para las variables originales
 - (b) Determina la transformación de potencia $\hat{\lambda}_1$, que hace a los valores de x_1 aproximadamente normales. Construye un Q-Q plot sobre los datos transformados
 - (c) Determina la transformación de potencia $\hat{\lambda}_2$, que hace a los valores de x_2 aproximadamente normales. Construye un Q-Q plot sobre los datos transformados
 - (d) Determina la transformación de potencia $\boldsymbol{\lambda} = [\hat{\lambda}_1, \hat{\lambda}_2]$ que hace a los valores (x_1, x_2) conjuntamente normales. Compara los resultados con los obtenidos en b) y c).