Godinez Bravo Diego

Estadística Multivariada

Tarea 2

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

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

a) Verifica que T^2 permanece sin cambios si cada observación $x_j, j = 1, 2, 3$ es reemplazada por Cx_j, donde

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

Note que las observaciones

$$Cx_j = \begin{pmatrix} x_{j1} - x_{j2} \\ x_{j1} + x_{j2} \end{pmatrix}$$

Producen la matriz

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

Solución

Si $x \sim N_p(\mu, \Sigma)$, la variable $(x-\mu)'\Sigma^{-1}(x-\mu)$ sigue una distribución χ^2_p. Si sustituimos Σ por su estimación $\hat{\Sigma} = S$, y sea $\bar{x}$ la media muestral donde $\bar{x} \sim N_p(\mu, \frac{\Sigma}{n})$, la distribución que se obtiene se

$$n(\bar{x}-\mu)'S^{-1}(\bar{x}-\mu)$$

se denomina T^2 de Hotelling.

Haciendo uso del lenguaje de programación R se obtuvieron las estimaciones de $\bar{x}$, S.
Para el vector de medias se obtuvo el vector de medias $\bar{x} = (8, 6)$, mientras que para la matriz de covarianza muestral se obtuvo $S = \begin{pmatrix} 4 & -3 \\ -3 & 9 \end{pmatrix}$; $S^{-1} = \frac{1}{27}\begin{pmatrix} 9 & 3 \\ 3 & 4 \end{pmatrix}$. Por lo tanto

$$T^2 = 3(8-\mu_1 \;\; 6-\mu_2)\begin{pmatrix} \frac{1}{3} & \frac{1}{9} \\ \frac{1}{9} & \frac{4}{27} \end{pmatrix}\begin{pmatrix} 8-\mu_1 \\ 6-\mu_2 \end{pmatrix}$$

$$= (8-\mu_1 \;\; 6-\mu_2)\begin{pmatrix} 8-\mu_1 + \frac{6-\mu_2}{3} \\ \frac{8-\mu_1}{3} + \frac{4(6-\mu_2)}{9} \end{pmatrix}$$

$$= (8-\mu_1 \;\; 6-\mu_2)\begin{pmatrix} 8-\mu_1+2-\frac{\mu_2}{3} \\ -\frac{1}{3}\mu_1 - \frac{4}{9}\mu_2 + \frac{16}{3} \end{pmatrix}$$

$$= (8-\mu_1)(8-\mu_1+2-\frac{\mu_2}{3}) + (6-\mu_2)(-\frac{1}{3}\mu_1 - \frac{4}{9}\mu_2 + \frac{16}{3})$$

$$= 64 - 8\mu_1 + 16 - \frac{8}{3}\mu_2 - 8\mu_1 + \mu_1^2 - 2\mu_1 + \frac{1}{3}\mu_1\mu_2 - 2\mu_1 + 32 - \frac{8}{3}\mu_2 + \frac{1}{3}\mu_1\mu_2 - \frac{16}{3}\mu_2 + \frac{4}{9}\mu_2^2$$

$$= 112 - 20\mu_1 - \frac{32}{3}\mu_2 + \frac{2}{3}\mu_1\mu_2 + \mu_1^2 + \frac{4}{9}\mu_2^2$$

De esta manera

$$T^2 = 112 - 20\mu_1 - \frac{32}{3}\mu_2 + \frac{2}{3}\mu_1\mu_2 + \mu_1^2 + \frac{4}{9}\mu_2^2$$

Reemplazando las observaciones por Cx_j se obtuvo el vector de medias $\bar{x}_{transformed} = (2, 14)$, y la matriz de covarianza $S_{transformed} = \begin{pmatrix} 19 & -5 \\ -5 & 7 \end{pmatrix}$; $S^{-1} = \frac{1}{108}\begin{pmatrix} 7 & 5 \\ 5 & 19 \end{pmatrix}$. De manera que

$$T^2 = \frac{3}{108}(2 - (\mu_1 - \mu_2)\ 14 - (\mu_1 + \mu_2)) \begin{pmatrix} 7 & 5 \\ 5 & 19 \end{pmatrix} \begin{pmatrix} 2 - (\mu_1 - \mu_2) \\ 14 - (\mu_1 + \mu_2) \end{pmatrix}$$

$$= \frac{1}{36}(2 - (\mu_1 - \mu_2)\ 14 - (\mu_1 + \mu_2)) \begin{pmatrix} 14 - 7\mu_1 + 7\mu_2 + 70 - 5\mu_1 - 5\mu_2 \\ 10 - 5\mu_1 + 5\mu_2 + 266 - 19\mu_1 - 19\mu_2 \end{pmatrix}$$

$$= \frac{1}{36}(2 - (\mu_1 - \mu_2)\ 14 - (\mu_1 + \mu_2)) \begin{pmatrix} 84 - 12\mu_1 + 2\mu_2 \\ 276 - 24\mu_1 - 14\mu_2 \end{pmatrix}$$

$$= \frac{1}{36}(2 - (\mu_1 - \mu_2))(84 - 12\mu_1 + 2\mu_2) + (14 - (\mu_1 + \mu_2))(276 - 24\mu_1 - 14\mu_2)$$

$$= \frac{1}{36}(168 - 24\mu_1 + 4\mu_2 - 84\mu_1 12\mu_1^2 - 2\mu_1\mu_2 + 84\mu_2 - 12\mu_1\mu_2 + 2\mu_2^2)$$
$$+(3864 - 336\mu_1 - 196\mu_2 - 276\mu_1 + 24\mu_1^2 + 14\mu_1\mu_2 - 276\mu_2 + 24\mu_1\mu_2 + 14\mu_2^2)$$

$$= \frac{1}{36}(168 - 108\mu_1 + 88\mu_2 - 14\mu_1\mu_2 + 12\mu_1^2 + 2\mu_2^2 + 3864 - 612\mu_1 - 472\mu_2 + 38\mu_1\mu_2 + 24\mu_1^2 + 14\mu_2^2)$$

$$= \frac{1}{36}(4032 - 720\mu_1 - 384\mu_2 + 24\mu_1\mu_2 + 36\mu_1^2 + 16\mu_2^2)$$

$$= 112 - 20\mu_1 - \frac{32}{3}\mu_2 + \frac{2}{3}\mu_1\mu_2 + \mu_1^2 + \frac{4}{9}\mu_2^2$$

Por lo tanto, verificamos que T^2 permanece sin cambios al aplicar la transformación Cx_j.

$$T^2 = 112 - 20\mu_1 - \frac{32}{3}\mu_2 + \frac{2}{3}\mu_1\mu_2 + \mu_1^2 + \frac{4}{9}\mu_2^2$$

Ejercicio 2. Dadas la siguiente muestra de observaciones bivariadas:

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

a) Evalua T^2, para probar $H_0 : \mu' = [7, 11]$, 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 $\alpha = 0{,}05$. ¿Que conclusión se tiene?

d) Evalue T^2 utilizando la relación que tiene con la λ de Wilks.

e) Evalua Λ y la λ de Wilks.

Solución

a) Definimos las hipótesis nula y alternativa como

$$H_0 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 11 \end{bmatrix}; \qquad H_1 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \neq \begin{bmatrix} 7 \\ 11 \end{bmatrix}$$

Basándonos en el estadístico T^2, la regla de decisión se declara de la siguiente manera:

No rechazar H_0 si $T^2 \leq \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)$; de otra manera se rechaza H_0.

Una generalización de la distancia cuadrada univariada t es su análogo multivariado T^2 de Hotelling

$$T^2 = n(\bar{x} - \mu_0)' S^{-1} (\bar{x} - \mu_0)$$

donde

$$\bar{x} = \frac{1}{n} \sum_{j=1}^{n} x_j = \frac{1}{n} X' 1 = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{bmatrix}; \quad \mu_0 = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$$

con $n^{-1}S$ como la matriz de covarianza estimada de $\bar{x}$.

Haciendo uso del lenguaje de programación R se obtuvo el cálculo del estadístico T^2. En este caso se obtuvo

$$T^2 = 13{,}6363$$

Para decidir si se rechaza la hipótesis nula H_0, comparamos el valor calculado de T^2 con el valor crítico de la distribución F. Asumiendo un nivel de significancia $\alpha = 0{,}05$, $p = 2$, $n - p = 2$ son los grados de libertad de F, entonces $F_{2,2}(0{,}05) = 19{,}00$.

De manera que

$$T^2 \leq \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)$$

$$13{,}63 \leq \frac{(4-1)2}{4-2} 19{,}00$$

$$13{,}63 \leq 57{,}00$$

Por lo tanto, no rechazamos H_0 y concluimos que no hay suficiente evidencia para afirmar que el vector de medias difiere de $[7, 11]$ basándonos en el valor obtenido de T^2.

b) Con el propósito de verificar la normalidad de los datos, se generó un gráfico de contornos de confianza con niveles de confianza del 50 y 95 %. Según la interpretación del gráfico, podemos concluir que los datos exhiben una distribución normal bivariada (**Figura 1**).

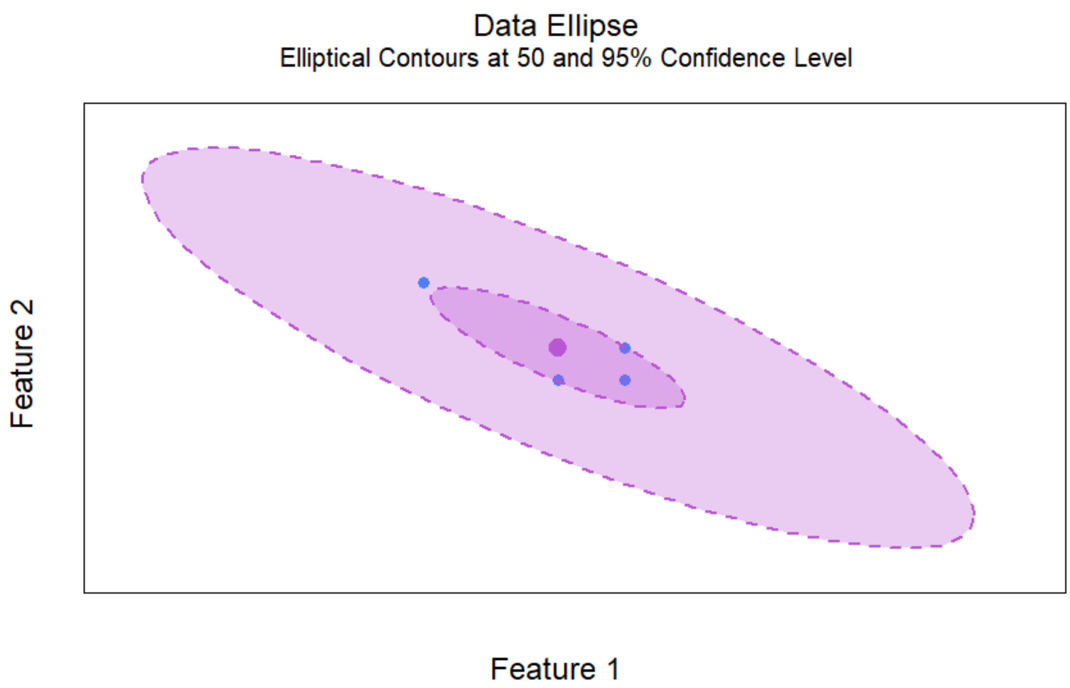

Figura 1. Contornos de confianza con un nivel de confianza del 50 y 95 %.

Dado que $x \sim N_p(\mu, \Sigma)$, la variable $(x-\mu)'\Sigma^{-1}(x-\mu)$ sigue una distribución χ^2_p. Sustituyendo Σ por su estimación $\hat{\Sigma} = S$ entonces $(n-1)S \sim W_p(S|\Sigma)$, por lo que la distribución de la variable escalar

$$T^2 = (x-\mu)'S^{-1}(x-\mu)$$

se denomina distribución T^2 de Hotelling con p y $n-1$ grados de libertad, es decir $T^2 \sim T^2(p, n-1)$.

c) Considerando un nivel de significancia $\alpha = 0{,}05$, $F_{2,2}(0{,}05) = 19{,}00$, entonces

$$\underbrace{13{,}63}_{T^2} \leq \underbrace{57{,}00}_{\frac{(n-1)p}{n-p}F_{p,n-p}(\alpha)}$$

De manera que no se rechaza la hipótesis nula H_0 y concluimos que no hay suficiente evidencia para afirmar que el vector de medias difiere de $[7, 11]$ basándonos en el valor obtenido del estadístico T^2.

d) Suponga que $x_1, ..., x_n$ es una muestra aleatoria de una población $N_p(\mu, \Sigma)$. Entonces la prueba basada en T^2 es equivalente a la prueba de la razón de verosimilitud para probar que $H_0 : \mu = \mu_0 vs H_1 : \mu \neq \mu_0$ debido a que

$$\Lambda^{\frac{2}{n}} = \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} = \left(1 + \frac{T^2}{n-1}\right)^{-1}$$

Al sustituir el valor crítico apropiado de T^2 en la ecuación anterior podemos encontrar el valor crítico de la razón de verosimilitud.

La ecuación anterior tiene la ventaja de demostrar que T^2 se puede calcular como el cociente de dos determinantes. Al resolver para T^2 se obtiene

$$T^2 = \frac{(n-1)|\hat{\Sigma}_0|}{|\hat{\Sigma}|} - (n-1) = \frac{(n-1)|\sum_{j=1}^{n}(x_j-\mu_0)(x_j-\mu_0)'|}{|\sum_{j=1}^{n}(x_j-\bar{x})(x_j-\bar{x})'|} - (n-1)$$

Haciendo uso de la expresión anterior y con ayude del lenguaje de programación R se obtuvo el calculo de T^2. Al igual que el caso anterior el valor que se obtuvo fue de

$$T^2 = 13{,}6363$$

De esta manera se comprueba que el valor de T^2 se puede calcular como el cociente de dos determinantes, evitando así el calculo de S^{-1}.

e) Para determinar si μ_0 es un valor plausible de μ comparamos el máximo de $L(\mu_0, \Sigma)$ con el máximo de $L(\mu, \Sigma)$ cuando no se imponen restricciones sobre μ y Σ.

La relación resultante se denomina Estadístico de la Razón de Verosimilitud

$$\Lambda = \frac{maxL(\mu_0, \Sigma)}{maxL(\mu, \Sigma)} = \frac{\frac{1}{(2\pi)^{np/2}|\hat{\Sigma}_0|^{n/2}}e^{-\frac{np}{2}}}{\frac{1}{(2\pi)^{np/2}|\hat{\Sigma}|^{n/2}}e^{-\frac{np}{2}}} = \left(\frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|}\right)^{\frac{n}{2}}$$

De manera equivalente

$$\Lambda^{\frac{2}{n}} = \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} = \frac{|\sum_{j=1}^{n}(x_j-\bar{x})(x_j-\bar{x})'|}{|\sum_{j=1}^{n}(x_j-\mu_0)(x_j-\mu_0)'|}$$

Lambda de Wilks

Utilizando los calculos obtenidos de $\hat{\Sigma}$, $\hat{\Sigma}_0$ en el inciso d) y las ecuaciones anteriores, realizamos el cálculo de ambos estadísticos.

$$\Lambda = \left(\frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|}\right)^{\frac{n}{2}} = \left(\frac{44}{244}\right)^{\frac{4}{2}} = 0{,}0325$$

Razón de verosimilitud

$$\Lambda = 0{,}0325$$

Por otro lado

$$\Lambda^{\frac{2}{n}} = \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} = \frac{44}{244} = 0{,}1803$$

Lambda de Wilks

$$\Lambda^{\frac{2}{n}} = 0{,}1803$$

Ejercicio 3. El departamento de control de calidad de un fabricante de hornos de microondas es requerido por el gobierno federal para monitorear la cantidad de radiación emitida por los hornos que fabrican. Se realizaron mediciones de la radiación emitida por 42 hornos seleccionados al azar con las puertas cerradas y abiertas. Los datos están en el archivo **datosradiacion**.

a) Construye un elipse de confianza del 95 % para μ, considerando la transformación de las variables:

$$x_1 = \sqrt[4]{mediciones\ de\ la\ radiacion\ con\ puerta\ cerrada}$$

$$x_2 = \sqrt[4]{mediciones\ de\ la\ radiacion\ con\ puerta\ abierta}$$

b) Prueba si $\mu' = (0{,}562, 0{,}589)$ está en la región de confianza

c) Calcula los valores y vectores propios de S y obten la gráfica del elipsoide de confianza

d) Realiza una prueba para la hipótesis $H_0 : \mu' = (0{,}55, 0{,}60)$ en un nivel de confianza de $\alpha = 0{,}05$. Es consistente el resultado con la gráfica de la elipse de confianza del 95 % para μ obtenida en el inciso anterior.? Explica

Solución

a) El estadístico para probar $H_0 : \mu = \mu_0$ se encuentra dado por

$$T^2 = n(\bar{x} - \mu_0)' S^{-1}(\bar{x} - \mu_0)$$

Recordando que no se rechaza H_0 si $T^2 \leq \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)$, es decir

$$n(\bar{x} - \mu_0)' S^{-1}(\bar{x} - \mu_0) \leq \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)$$

Donde la región de confianza para μ de una población normal p-variada está dado por

$$P\left(n(\bar{x} - \mu_0)' S^{-1}(\bar{x} - \mu_0) \leq \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)\right) = 1 - \alpha$$

De manera formal, una región de confianza $R(x)$ del $100(1-\alpha)$ % para el vector de medias μ de una distribución normal p-dimensional es el elipsoide determinado por todos los puntos posibles μ que satisfacen

$$n(\bar{x} - \mu_0)' S^{-1}(\bar{x} - \mu_0) \leq \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)$$

con

$$\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j \qquad S = \frac{1}{n-1}\sum_{j=1}^{n}(x_j - \bar{x})(x_j - \bar{x})'$$

y $x_1, ..., x_n$ como las muestras de observaciones.

Haciendo uso del lenguaje de programación R se obtuvo el cálculo del vector de medias $\bar{x}$ y la matriz de coviarianzas S. Donde

$$\bar{x} = (0{,}5642\ 0{,}6029) \qquad S^{-1} = \begin{pmatrix} 203{,}49 & -163{,}90 \\ -163{,}90 & 200{,}76 \end{pmatrix}$$

Considerando un nivel de significancia $\alpha = 0{,}05$, $p = 2$, $n - p = 40$ son los grados de libertad de F, entonces $F_{2,40}(0{,}05) = 3{,}23$.

Por lo tanto, nuestra región de confianza $R(X)$ son todos los puntos (μ_1, μ_2) que satisfacen

$$42\begin{pmatrix}0{,}56-\mu_1 & 0{,}60-\mu_2\end{pmatrix}\begin{pmatrix}203{,}49 & -163{,}90\\ -163{,}90 & 200{,}76\end{pmatrix}\begin{pmatrix}0{,}56-\mu_1\\ 0{,}60-\mu_2\end{pmatrix}\leq 6{,}63$$

$$\begin{pmatrix}0{,}56-\mu_1 & 0{,}60-\mu_2\end{pmatrix}\begin{pmatrix}114{,}82-203{,}49\mu_1-98{,}83+163{,}90\mu_2\\ -92{,}48+163{,}90\mu_1+121{,}06-200{,}76\mu_2\end{pmatrix}\leq 0{,}15$$

$$\begin{pmatrix}0{,}56-\mu_1 & 0{,}60-\mu_2\end{pmatrix}\begin{pmatrix}15{,}99-203{,}49\mu_1+163{,}90\mu_2\\ 28{,}57+163{,}90\mu_1-200{,}76\mu_2\end{pmatrix}\leq 0{,}15$$

$$\begin{gathered}9{,}02-114{,}82\mu_1+92{,}48\mu_2-15{,}99\mu_1+203{,}49\mu_1^2+163{,}90\mu_1\mu_2+17{,}22+98{,}83\mu_1\\ -121{,}06\mu_2-28{,}57\mu_2-163{,}90\mu_2\mu_1+200{,}76\mu_2^2\leq 0{,}15\end{gathered}$$

Por lo tanto el elipsoide de confianza del 95 % de confianza está dado de la siguiente manera

$$203{,}49\mu_1^2+200{,}76\mu_2^2-327{,}81\mu_1\mu_2-31{,}98\mu_1-57{,}14\mu_2+26{,}25\leq 0{,}15$$

b) Para corroborar si $\mu=(0{,}562,0{,}589)$ se encuentra en el intervalo de confianza evaluamos la desigualdad anterior

$$203{,}49(0{,}562)^2+200{,}76(0{,}589)^2-327{,}81(0{,}562)(0{,}589)-31{,}98(0{,}562)-57{,}14(0{,}589)+26{,}24\leq 0{,}15$$

$$64{,}27+69{,}64-108{,}51-17{,}97-33{,}65+26{,}25\leq 0{,}15$$

$$0{,}03\leq 0{,}15$$

Dado que la desigualdad se cumple podemos afirmar que el vector de medias $\mu'=(0{,}562,0{,}589)$ se encuentra en el intervalo de confianza.

c) Haciendo uso del lenguaje de programación R se obtuvo el cálculo de los valores y vectores propios de la matrix S.

Valores propios obtenidos

$$(\lambda_1,\lambda_2)=(0{,}0261,0{,}0027)$$

Vectores propios obtenidos

$$(e_1,e_2)=\begin{pmatrix}0{,}7041 & -0{,}7100\\ 0{,}7100 & 0{,}7041\end{pmatrix}$$

Observamos que el vector de medias $\mu'=(0{,}562,0{,}589)$ se encuentra dentro de la región de confianza, lo que sugiere que los datos están bien ajustados a la distribución establecida por las elipses generadas (**Figura 2**).

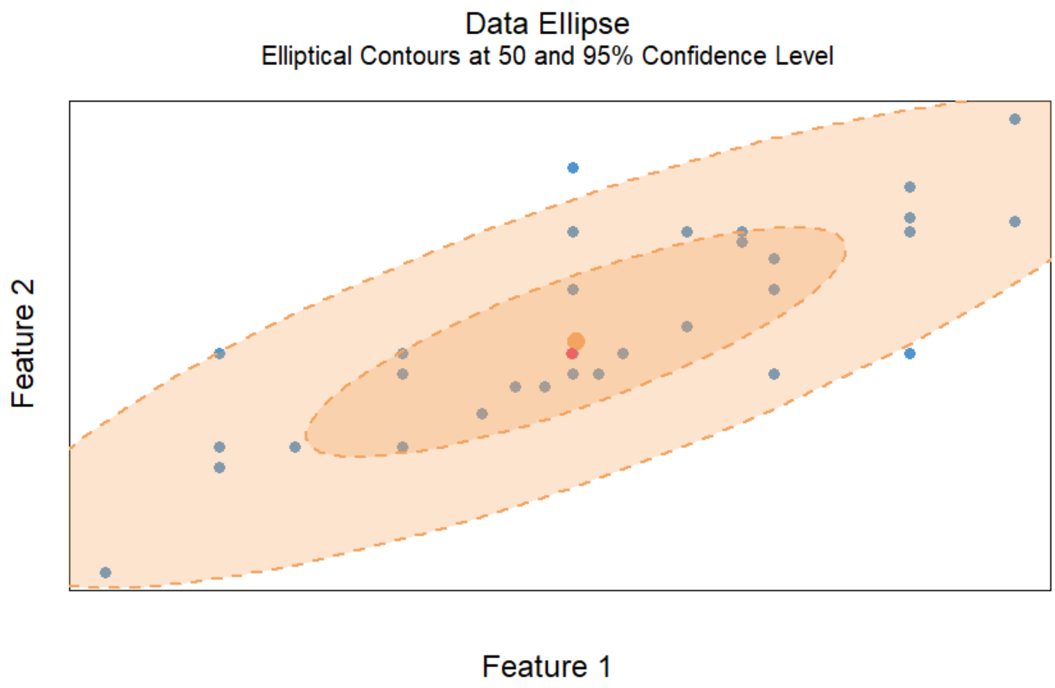

Figura 2. Contornos de confianza con un nivel de confianza del 50 y 95 %. Punto resaltado corresponde al vector de medias.

d) Utilizamos el estadístico T^2 para probar $H_0 : \mu = \mu_0$, definido de la siguiente manera

$$T^2 = n(\bar{x} - \mu_0)^{'} S^{-1}(\bar{x} - \mu_0)$$

Con la regla de desición dada por

$$T^2 \leq \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)$$

Haciendo uso del lenguaje de programación R se obtuvo el cálculo del estadístico T^2.

$$T^2 = 1{,}2271$$

Considerando un nivel de significancia $\alpha = 0{,}05$, donde $p = 2$, $n - p = 40$ son los grados de libertad de F, entonces $F_{2,40}(0{,}05) = 3{,}23$, por lo tanto

$$1{,}2271 \leq \frac{(42-1)2}{42-2} F_{2,40}(0{,}05)$$

$$1{,}2271 \leq 6{,}6215$$

Dado que $T^2 \leq \frac{(n-1)p}{n-p} F_{p,n-p}(\alpha)$ con un nivel de significancia $\alpha = 0{,}05$ no podemos rechazar la hipótesis nula H_0, no existe suficiente evidencia para afirmar que el vector de medias difiere de $H_0 : \mu' = (0{,}55, 0{,}60)$ basándonos en el valor obtenido de T^2. Este resultado coincide con lo observado en el gráfico de contornos (**Figura 2**), donde la media $\mu' = (0{,}56, 0{,}58)$ se encuentra dentro de la región de confianza, ya que en este caso se realizó la prueba para la hipótesis H_0 considerando un vector de medias μ' cercano a dichos valores.

Ejercicio 4. Sabemos que T^2 es igual al t-valor cuadrado univariado más grande construido a partir de la combinación lineal $a^{'}x_j$, con $a = S^{-1}(\bar{x} - \mu_0)$.

a) Usando los resultados del inciso anterior y la misma hipótesis nula H_0 del inciso d), evalua a para los datos transformados de radiaciones de los hornos.

b) Verifica que el valor t^2 calculando con esta a es igual a la T^2 del ejercicio anterior.

Solución

a) Haciendo uso del lenguaje de programación R se obtuvo la evaluación de a para los datos transformados.

Puerta Cerrada x_1	2.4127
Puerta Abierta x_2	-1.7383

b) Haciendo uso del lenguaje de programación R se obtuvo el cálculo de t^2 considerando la evaluación de a y la siguiente definición

$$t^2 = \frac{n(a'(\bar{x} - \mu))^2}{a'Sa}$$

Valor obtenido de t^2

$$t^2 = 1{,}2271$$

Observamos que el valor coincide con el valor del estadístico T^2 obtenido en el ejercicio anterior (i.e. $t^2 = T^2$).

Ejercicio 5. Los datos en el archivo **datososos** representan las longitudes en centímetros de siete osos hembras a los 2, 3, 4, y 5 años de edad.

a) Obtener los intervalos de confianza simultaneos T^2 del 95 % para las cuatro medias poblacionales de la longitud por año.

b) Respecto al inciso a), obtener los intervalos de confianza simultaneos T^2 del 95 % para los tres aumentos anuales sucesivos en la longitud media.

Solución

a) Sea $x_1, ..., x_n$ una muestra aleatoria obtenida de una población $N_p(\mu, \Sigma)$ positiva definida. Entonces, simultáneamente para todo a, el intervalo

$$\left(a'\bar{X} - \sqrt{\frac{(n-1)p}{(n-1)}F_{p,n-p}(\alpha)a'Sa},\ a'\bar{X} + \sqrt{\frac{(n-1)p}{(n-1)}F_{p,n-p}(\alpha)a'Sa}\right)$$

contendrá $a'\mu$ con probabilidad $1-\alpha$.

Estos intervalos simultáneos se denominan intervalos T^2, ya que la probabilidad de cobertura es determinada por la distribución T^2. Notose que las elecciones sucesivas de $a' = [1000..,0]$, $a' = [0100..,0]$, ..., $a' = [0000..,1]$ para los intervalos T^2, nos permiten obtener los intervalos de confianza para las medias de los componentes, $\mu_1, \mu_2, ..., \mu_p$, esto es

$$\bar{x_1} - \sqrt{\frac{(n-1)p}{(n-1)}F_{p,n-p}(\alpha)}\sqrt{\frac{S_{11}}{n}} \leq \mu_1 \leq \bar{x_1} + \sqrt{\frac{(n-1)p}{(n-1)}F_{p,n-p}(\alpha)}\sqrt{\frac{S_{11}}{n}}$$
$$\bar{x_2} - \sqrt{\frac{(n-1)p}{(n-1)}F_{p,n-p}(\alpha)}\sqrt{\frac{S_{22}}{n}} \leq \mu_2 \leq \bar{x_2} + \sqrt{\frac{(n-1)p}{(n-1)}F_{p,n-p}(\alpha)}\sqrt{\frac{S_{22}}{n}}$$
$$\vdots$$
$$\bar{x_p} - \sqrt{\frac{(n-1)p}{(n-1)}F_{p,n-p}(\alpha)}\sqrt{\frac{S_{pp}}{n}} \leq \mu_p \leq \bar{x_p} + \sqrt{\frac{(n-1)p}{(n-1)}F_{p,n-p}(\alpha)}\sqrt{\frac{S_{pp}}{n}}$$

Haciendo uso del lenguaje de programación R se obtuvieron los intervalos de confianza simultáneos para las medias poblacionales μ_i, para $i = (1, 2, 3, 4)$.

Considerando un nivel de significancia $\alpha = 0{,}05$, donde $p = 4$, $n - p = 3$ son los grados de libertad de F, entonces $F_{4,3}(0{,}05) = 9{,}11$.

Intervalos de confianza simultaneos T^2 obtenidos.

Limite Inferior	μ_i	Limite Superior
130.6851	$\leq$ 143.2857 $\leq$	155.8863
127.0216	$\leq$ 159.2857 $\leq$	191.5498
160.3082	$\leq$ 173.1429 $\leq$	185.9776
155.3749	$\leq$ 177.1429 $\leq$	198.9108

b) Al considerar los aumentos anuales sucesivos en la longitud de la media buscamos resaltar las diferencias entre las medias.

Haciendo uso del lenguaje de programación R se obtuvieron los intervalos de confianza simultaneos considerando $a' = (-1\ 1\ 0\ 0)$; $a' = (0\ -1\ 1\ 0)$; y $a' = (0\ 0\ -1\ 1)$.

Intervalos de confianza simultaneos T^2 obtenidos para los aumentos anuales sucesivos en la longitud media.

Limite Inferior	Aumento Anual ($\mu_{i+1} - \mu_i$)	Limite Superior
-21.2264	$\leq$ 16.0000 $\leq$	53.2264
-22.7307	$\leq$ 13.8571 $\leq$	50.4450
-20.6538	$\leq$ 4.0000 $\leq$	28.6538