Tema: Distribución normal multivariante

1. Encuentre las estimaciones de máxima verosimilitud del vector media 2×1 μ y la covarianza 2×2 matriz Σ basada en la muestra aleatoria

$$\mathbf{X} = \begin{bmatrix} 3 & 6 \\ 4 & 4 \\ 5 & 7 \\ 4 & 7 \end{bmatrix}$$

de una población normal bivariada.

2. Sea $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{20}$ una muestra aleatoria de tamaño $n = 20$ de una población $N_6(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Especifique completamente cada uno de los siguientes.

   (a) La distribución de $(\mathbf{X}_1 - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X}_1 - \boldsymbol{\mu})$
   (b) Las distribuciones de $\overline{\mathbf{X}}$ y $\sqrt{n}(\overline{\mathbf{X}} - \boldsymbol{\mu})$
   (c) La distribución de $(n-1)\mathbf{S}$
   (d) Especifique la distribución de $\mathbf{B}(19\mathbf{S})\mathbf{B}'$ en cada caso.

      i. $\mathbf{B} = \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix}$

      ii. $\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$

3. Sea $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{75}$ una muestra aleatoria de una distribución de población con media $\boldsymbol{\mu}$ y matriz de covarianza $\boldsymbol{\Sigma}$. ¿Cuál es la distribución aproximada de cada uno de los siguientes?

   (a) $\overline{\mathbf{X}}$
   (b) $n(\overline{\mathbf{X}} - \boldsymbol{\mu})'\mathbf{S}^{-1}(\overline{\mathbf{X}} - \boldsymbol{\mu})$

Tema: Inferencias acerca del vector de medias

4. Utilizando los siguientes datos

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

   (a) Evaluar T^2, para la prueba $H_0 : \boldsymbol{\mu}' = \begin{bmatrix} 7, & 11 \end{bmatrix}$
   (b) Especifique la distribución de T^2 para la situación en (a).
   (c) Utilizando (a) y (b), pruebe H_0 al nivel $\alpha = .05$. ¿A qué conclusión se llega?

(d) USe

$$T^2 = \frac{(n-1)\left|\hat{\mathbf{\Sigma}}_0\right|}{|\hat{\mathbf{\Sigma}}|} - (n-1) = \frac{(n-1)\left|\sum_{j=1}^n (\mathbf{x}_j - \boldsymbol{\mu}_0)(\mathbf{x}_j - \boldsymbol{\mu}_0)'\right|}{\left|\sum_{j=1}^n (\mathbf{x}_j - \overline{\mathbf{x}})(\mathbf{x}_j - \overline{\mathbf{x}})'\right|} - (n-1)$$

para evaluar T^2.

(e) Evalúe Λ y el lambda de Wilk usando la expresión

$$\Lambda = \left(\frac{|\hat{\mathbf{\Sigma}}|}{\left|\hat{\mathbf{\Sigma}}_0\right|}\right)^{n/2} = \left(\frac{\left|\sum_{j=1}^n (\mathbf{x}_j - \overline{\mathbf{x}})(\mathbf{x}_j - \overline{\mathbf{x}})'\right|}{\left|\sum_{j=1}^n (\mathbf{x}_j - \boldsymbol{\mu}_0)(\mathbf{x}_j - \boldsymbol{\mu}_0)'\right|}\right)^{n/2}$$

5. Para los datos de *sweat* (visto en clase):

(a) Determine los ejes del elipsoide de confianza 90% para $\boldsymbol{\mu}$. Determine las longitudes de estos ejes.

(b) Encuentre intervalos de confianza simultáneos 95% T^2 para μ_1, μ_2, y μ_3. Construya los intervalos de Bonferroni 95%. Compare los dos conjuntos de intervalos.

(c) Determine la distribución aproximada de $-n \ln\left(|\hat{\mathbf{\Sigma}}| / \left|\hat{\mathbf{\Sigma}}_0\right|\right)$.

6. Harry Roberts, naturalista del Departamento de Pesca y Caza de Alaska, estudia a los osos pardos con el objetivo de mantener una población sana. Las mediciones en $n = 61$ osos proporcionaron las siguientes estadísticas resumidas:

Variable	Peso (kg)	Longitud del cuerpo (cm)	Cuello (cm)	Circunferencia (cm)	Longitud de la cabeza (cm)	Ancho de la cabeza (cm)
Media muestral $\bar{x}$	95.52	164.38	55.69	93.39	17.98	31.13

Matriz de covarianza:

$$\mathbf{S} = \begin{bmatrix} 3266.46 & 1343.97 & 731.54 & 1175.50 & 162.68 & 238.37 \\ 1343.97 & 721.91 & 324.25 & 537.35 & 80.17 & 117.73 \\ 731.54 & 324.25 & 179.28 & 281.17 & 39.15 & 56.80 \\ 1175.50 & 537.35 & 281.17 & 474.98 & 63.73 & 94.85 \\ 162.68 & 80.17 & 39.15 & 63.73 & 9.95 & 13.88 \\ 238.37 & 117.73 & 56.80 & 94.85 & 13.88 & 21.26 \end{bmatrix}$$

(a) Obtenga los intervalos de confianza de 95% simultáneos para las seis mediciones corporales medias de la población, considerando que es una muestra grande.

(b) Obtenga la region simultánea elíptica de confianza del 95% para el peso medio y la circunferencia media.

(c) Obtenga los 95% intervalos de confianza de Bonferroni para las seis medias de la Parte a.

(d) Consulte la Parte b. Construya el rectángulo de confianza de Bonferroni 95% para el peso medio y la circunferencia media usando $m = 6$. Compare este rectángulo con la elipse de confianza de la Parte b.

(e) Obtenga el 95% intervalo de confianza de Bonferroni para

$$\text{ancho medio de la cabeza} - \text{longitud media de la cabeza}$$

usando $m = 6 + 1 = 7$.

Tema: Comparaciones emparejadas y diseño de medidas repetidas

7. Construya y dibuje una región de confianza conjunta del 95% para el vector de diferencia de medias $\boldsymbol{\delta}$ utilizando los datos de aguas (visto en clase). Tenga en cuenta que el punto $\boldsymbol{\delta} = \mathbf{0}$ queda fuera del contorno 95%. ¿Es este resultado coherente con la prueba de $H_0 : \boldsymbol{\delta} = \mathbf{0}$? Explique.

8. Utilizando los datos de aguas, construya los intervalos simultáneos de Bonferroni 95% para los componentes del vector de diferencia de medias $\boldsymbol{\delta}$. Compare las longitudes de estos intervalos con las de los intervalos simultáneos construidos en el ejemplo.

9. Los datos correspondientes a la muestra 8 en los datos aguas parecen inusualmente grandes. Retire la muestra 8. Construya una región de confianza conjunta 95% para el vector de diferencia de medias $\boldsymbol{\delta}$ y los intervalos simultáneos de Bonferroni de 95% para las componentes del vector de diferencia de medias ¿Los resultados son consistentes con una prueba de $H_0 : \boldsymbol{\delta} = \mathbf{0}$? Discuta. ¿El "valor atípico" hace una diferencia en el análisis de estos datos?

10. Un investigador consideró tres índices que miden la gravedad de los ataques cardíacos. Los valores de estos índices para $n = 40$ pacientes con ataque cardíaco que llegaron a la sala de emergencias de un hospital produjeron las estadísticas resumidas

$$\overline{\mathbf{x}} = \begin{bmatrix} 46.1 \\ 57.3 \\ 50.4 \end{bmatrix} \text{ y } \mathbf{S} = \begin{bmatrix} 101.3 & 63.0 & 71.0 \\ 63.0 & 80.2 & 55.6 \\ 71.0 & 55.6 & 97.4 \end{bmatrix}$$

(a) Los tres índices se evalúan para cada paciente. Pruebe la igualdad de los índices medios utilizando la siguiente expresión con $\alpha = .05$.
H_0 se rechaza si: $T^2 = n(\mathbf{C}\overline{\mathbf{x}})' \left(\mathbf{CSC}'\right)^{-1} \mathbf{C}\overline{\mathbf{x}} > \frac{(n-1)(q-1)}{(n-q+1)} F_{q-1,n-q+1}(\alpha)$

(b) Juzgue las diferencias en pares de índices medios utilizando 95% intervalos de confianza simultáneos. Utilice $\mathbf{c}'\boldsymbol{\mu}$: $\quad \mathbf{c}'\overline{\mathbf{x}} \pm \sqrt{\frac{(n-1)(q-1)}{(n-q+1)} F_{q-1,n-q+1}(\alpha)} \sqrt{\frac{\mathbf{c}'\mathbf{S}\mathbf{c}}{n}}$.