Inferencia Estadística para la Media (μ): Caso Multivariado

Pruebas de Hipótesis para $\underline{\mu}$ cuando $\boldsymbol{\Sigma}$ es Desconocida (Normal)

Ahora se considerará el problema de determinar si un vector dado $\underline{\mu}_0$ es un vector-viable para el vector de medias $\underline{\mu}$ de una población Normal-Multivariada (con $\boldsymbol{\Sigma}$-desconocida).

Se procederá de manera análoga al caso univariado.

Ahora se tendrá la distancia al cuadrado dada por:

$$T^2 = \left(\overline{\underline{\boldsymbol{x}}} - \underline{\mu}_0\right)^t \left(\frac{1}{n}\boldsymbol{S}\right)^{-1} \left(\overline{\underline{\boldsymbol{x}}} - \underline{\mu}_0\right) = n(\overline{\underline{\boldsymbol{x}}} - \underline{\mu}_0)^t \boldsymbol{S}^{-1}(\overline{\underline{\boldsymbol{x}}} - \underline{\mu}_0)$$

$$\overline{\underline{\boldsymbol{x}}} = \frac{1}{n}\sum_{i=1}^{n}\underline{\boldsymbol{x}}_i \quad , \quad \boldsymbol{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\underline{\boldsymbol{x}}_i - \overline{\underline{\boldsymbol{x}}})(\underline{\boldsymbol{x}}_i - \overline{\underline{\boldsymbol{x}}})^t \quad \text{y} \quad \underline{\mu}_0 = \begin{bmatrix} \mu_{10} \\ \mu_{20} \\ \vdots \\ \mu_{p0} \end{bmatrix}$$

La estadística T^2-se llama Estadística T^2-de Hotelling, en honor a Harold Hotelling, un pionero en Análisis Multivariado, quien fue el primero en obtener su distribución muestral.

Aquí, $\frac{\boldsymbol{S}}{n}$-es la matriz de varianzas-covarianzas estimada de $\underline{\overline{\boldsymbol{x}}}$, pues

$$Var\left[\underline{\overline{\boldsymbol{x}}}\right] = \frac{\boldsymbol{\Sigma}}{n} \quad ie. \quad \widehat{Var}\left[\underline{\overline{\boldsymbol{x}}}\right] = \frac{\hat{\boldsymbol{\Sigma}}}{n} = \frac{\boldsymbol{S}}{n}$$

Si la distancia estadística observada T^2 es mu grande, ie. $\underline{\overline{\boldsymbol{x}}}$-está demasiado lejos de $\underline{\mu}_0$, y la hipótesis $H_0 \ : \ \underline{\mu} = \underline{\mu}_0$ es rechazada.

Pero en este caso las tablas de puntos porcentuales de la distribución de la T^2-de Hotelling, no son necesarias para las pruebas formales de hipótesis.

Se utilizará el siguiente resultado:

$$T^2 \;\sim\; \frac{(n-1)p}{(n-p)} F_{p,n-p} = kF, \quad \text{con:} \quad k = \frac{(n-1)p}{(n-p)}$$

o equivalentemente: $\frac{(n-p)}{(n-1)p} T^2 = \frac{1}{k} T^2 \;\sim\; F_{p,n-p}$, donde, $F_p, n-p$-denota una v.a con distribución F con p y $n-p$ grados de libertad respectivamente.

En Resumen: Sea $\underline{\boldsymbol{x}}_1, \underline{\boldsymbol{x}}_2, \ldots, \underline{\boldsymbol{x}}_n$ una m.a de una distribución $N_p(\underline{\mu}\ ,\ \boldsymbol{\Sigma})$, ($\boldsymbol{\Sigma}$-desconocido), entonces:

$$P\left[T^2 > \frac{(n-1)p}{(n-p)} F_{\alpha;p,n-p}\right] = P\big[T^2 > kF\big]$$

$$= P\Big[n(\overline{\underline{\boldsymbol{x}}} - \underline{\mu})^t \boldsymbol{S}^{-1}(\overline{\underline{\boldsymbol{x}}} - \underline{\mu}) > kF\Big] = \alpha$$

cualesquiera sean los valores de $\underline{\mu}$ y $\boldsymbol{\Sigma}$.

Al nivel de significancia del $\alpha\%$, rechazamos $H_0 \ : \ \underline{\mu} = \underline{\mu}_0$, en favor de: $H_a \ : \ \underline{\mu} \neq \underline{\mu}_0$, si el valor de:

$$T_0^2 = n(\underline{\overline{\boldsymbol{x}}} - \underline{\mu}_0)^t \boldsymbol{S}^{-1}(\underline{\overline{\boldsymbol{x}}} - \underline{\mu}_0) > kF = \frac{(n-1)p}{(n-p)} F_{\alpha;p,n-p},$$

o equivalentemente, rechazamos H_0 si:

$$F_0 = \frac{(n-p)}{(n-1)p} T^2 = \frac{1}{k} T_0^2 > F_{\alpha;p,n-p},$$

en caso contrario no-rechazamos H_0.

Ejemplo-1 Considere la siguiente matriz de datos asociada a una ma de tamaño $n = 3$ de una población normal bi-variada con vector de medias $\underline{\mu}$ y matriz de Var-Cov $\boldsymbol{\Sigma}$-desconocida, ie. $\underline{\boldsymbol{x}_i} \sim N_2(\underline{\mu}, \boldsymbol{\Sigma})$, para $i = 1, 2, 3$:

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

Realizar la siguiente prueba de hipótesis:

$$\begin{cases} H_0 \ : \ \underline{\mu} = \underline{\mu_0} \\ \\ H_a \ : \ \underline{\mu} \neq \underline{\mu_0} \end{cases} \iff \begin{cases} H_0 \ : \ \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} \mu_{10} \\ \mu_{20} \end{bmatrix} \\ \\ H_a \ : \ \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \neq \begin{bmatrix} \mu_{10} \\ \mu_{20} \end{bmatrix} \end{cases} \iff \begin{cases} H_0 \ : \ \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 9 \\ 5 \end{bmatrix} \\ \\ H_a \ : \ \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \neq \begin{bmatrix} 9 \\ 5 \end{bmatrix} \end{cases}$$

De la matriz de datos se encuentra que:

$$\underline{\overline{\boldsymbol{x}}} = \frac{1}{n}\sum_{i=1}^{n} \underline{\boldsymbol{x}}_i = \frac{1}{3}\left[\begin{pmatrix}6\\9\end{pmatrix} + \begin{pmatrix}10\\6\end{pmatrix} + \begin{pmatrix}8\\3\end{pmatrix}\right] = \frac{1}{3}\begin{bmatrix}6+10+8\\9+6+3\end{bmatrix}$$

$$= \frac{1}{3}\begin{bmatrix}24\\18\end{bmatrix} = \begin{bmatrix}8\\6\end{bmatrix} = \begin{bmatrix}\overline{x}_1\\\overline{x}_2\end{bmatrix}$$

Similarmente se tiene que:

$$s_{11} = \frac{(6-8)^2 + (10-8)^2 + (8-8)^2}{2} = 4$$

$$s_{22} = \frac{(9-6)^2 + (6-6)^2 + (3-6)^2}{2} = 9$$

$$s_{12} = \frac{(6-8)(9-6) + (10-8)(6-6) + (8-8)(3-6)}{2} = -3$$

es decir que:

$$\boldsymbol{S} = \begin{bmatrix} 4 & -3 \\ -3 & 9 \end{bmatrix}, \quad \text{luego :} \quad \boldsymbol{S}^{-1} = \frac{1}{27}\begin{bmatrix} 9 & 3 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{9} \\ \\ \frac{1}{9} & \frac{4}{27} \end{bmatrix}$$

Ahora la estadística de prueba calculada es:

$$T_0^2 = n(\underline{\overline{\boldsymbol{x}}} - \underline{\mu}_0)^t \boldsymbol{S}^{-1}(\underline{\overline{\boldsymbol{x}}} - \underline{\mu}_0) = 3\left[\begin{pmatrix} 8 \\ 6 \end{pmatrix} - \begin{pmatrix} 9 \\ 5 \end{pmatrix}\right]^t \begin{bmatrix} \frac{1}{3} & \frac{1}{9} \\ \\ \frac{1}{9} & \frac{4}{27} \end{bmatrix} \left[\begin{pmatrix} 8 \\ 6 \end{pmatrix} - \begin{pmatrix} 9 \\ 5 \end{pmatrix}\right]$$

$$= 3\begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{1}{9} \\ \\ \frac{1}{9} & \frac{4}{27} \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \frac{7}{9}$$