MAT-269: Sesión 11
Test de hipótesis II

Felipe Osorio

fosorios.mat.utfsm.cl

Departamento de Matemática, UTFSM

Considere la partición:

$$\boldsymbol{x}_i = \begin{pmatrix} \boldsymbol{x}_{i1} \\ \boldsymbol{x}_{i2} \end{pmatrix}, \qquad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}.$$

Nuestro interés es probar la hipótesis $H_0 : \boldsymbol{\Sigma}_{12} = \boldsymbol{0}$.

Considere $\widehat{\boldsymbol{\mu}} = \overline{\boldsymbol{x}}$ y $\widehat{\boldsymbol{\Sigma}} = \frac{1}{n}\boldsymbol{Q}$ los MLE de $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$ particionadas como:

$$\overline{\boldsymbol{x}} = \begin{pmatrix} \overline{\boldsymbol{x}}_1 \\ \overline{\boldsymbol{x}}_2 \end{pmatrix}, \qquad \boldsymbol{Q} = \begin{pmatrix} \boldsymbol{Q}_{11} & \boldsymbol{Q}_{12} \\ \boldsymbol{Q}_{21} & \boldsymbol{Q}_{22} \end{pmatrix},$$

donde

$$\boldsymbol{Q}_{11} = \sum_{i=1}^{n} (\boldsymbol{x}_{i1} - \overline{\boldsymbol{x}}_1)(\boldsymbol{x}_{i1} - \overline{\boldsymbol{x}}_1)^\top \boldsymbol{Q}_{22} = \sum_{i=1}^{n} (\boldsymbol{x}_{i2} - \overline{\boldsymbol{x}}_2)(\boldsymbol{x}_{i2} - \overline{\boldsymbol{x}}_2)^\top$$

$$\boldsymbol{Q}_{12} = \boldsymbol{Q}_{21}^\top = \sum_{i=1}^{n} (\boldsymbol{x}_{i1} - \overline{\boldsymbol{x}}_1)(\boldsymbol{x}_{i2} - \overline{\boldsymbol{x}}_2)^\top$$

Test de independencia

Bajo H_0 se tienen dos muestras independientes $\boldsymbol{x}_{11},\dots,\boldsymbol{x}_{n1}$ desde $\mathsf{N}_{p_1}(\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_{11})$ y $\boldsymbol{x}_{12},\dots,\boldsymbol{x}_{n2}$ desde $\mathsf{N}_{p_2}(\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_{22})$. Además,

$$L(\boldsymbol{\mu},\boldsymbol{\Sigma}) = L_1(\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_{11})L_2(\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_{22}).$$

De modo que el LRT para H_0 es

$$\begin{aligned}
\Lambda &= \frac{\max_{\mu_1,\mu_2,\Sigma_{11},\Sigma_{22}} L_1(\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_{11})L_2(\boldsymbol{\mu}_2,\boldsymbol{\Sigma}_{22})}{\max_{\mu,\Sigma} L(\boldsymbol{\mu},\boldsymbol{\Sigma})} \\
&= \frac{L_1(\widehat{\boldsymbol{\mu}}_1,\widehat{\boldsymbol{\Sigma}}_{11})L_2(\widehat{\boldsymbol{\mu}}_2,\widehat{\boldsymbol{\Sigma}}_{22})}{L(\widehat{\boldsymbol{\mu}},\widehat{\boldsymbol{\Sigma}})} \\
&= \frac{(2\pi)^{-np_1/2}|\widehat{\boldsymbol{\Sigma}}_{11}|^{-n/2}e^{-np_1/2}(2\pi)^{-np_2/2}|\widehat{\boldsymbol{\Sigma}}_{22}|^{-n/2}e^{-np_2/2}}{(2\pi)^{-np/2}|\widehat{\boldsymbol{\Sigma}}|^{-n/2}e^{-np/2}} \\
&= \Big(\frac{|\widehat{\boldsymbol{\Sigma}}_{11}||\widehat{\boldsymbol{\Sigma}}_{22}|}{|\widehat{\boldsymbol{\Sigma}}|}\Big)^{-n/2} = \Big(\frac{|\boldsymbol{Q}|}{|\boldsymbol{Q}_{11}||\boldsymbol{Q}_{22}|}\Big)^{n/2},
\end{aligned}$$

donde $\widehat{\boldsymbol{\Sigma}}_{ii} = \boldsymbol{Q}_{ii}/n$, para $i = 1,2$.

Ahora, como $\boldsymbol{Q} > \mathbf{0}$ con prob. 1 y $\boldsymbol{Q}_{11}$ es no singular, sigue que

$$|\boldsymbol{Q}| = |\boldsymbol{Q}_{11}||\boldsymbol{Q}_{22} - \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}|,$$

haciendo $\boldsymbol{E} = \boldsymbol{Q}_{22} - \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}$ $(= \boldsymbol{Q}_{22\cdot 1})$ y $\boldsymbol{H} = \boldsymbol{Q}_{21}\boldsymbol{Q}_{11}^{-1}\boldsymbol{Q}_{12}$. Tenemos

$$LR = \Lambda^{2/n} = \frac{|\boldsymbol{Q}_{22\cdot 1}|}{|\boldsymbol{Q}_{22}|} = \frac{|\boldsymbol{E}|}{|\boldsymbol{E} + \boldsymbol{H}|}.$$

Dado que $\boldsymbol{Q} = (n-1)\boldsymbol{S} \sim \mathsf{W}_p(n-1, \boldsymbol{\Sigma})$. Luego, cuando $H_0 : \boldsymbol{\Sigma}_{12} = \mathbf{0}$ es verdadero $\boldsymbol{\Sigma}_{22\cdot 1} = \boldsymbol{\Sigma}_{22}$ $(= \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12})$ y las matrices $\boldsymbol{H}$ y $\boldsymbol{E}$ son independientes

$$\boldsymbol{H} \sim \mathsf{W}_{p_2}(p_1, \boldsymbol{\Sigma}_{22}), \qquad \boldsymbol{E} \sim \mathsf{W}_{p_2}(n - 1 - p_1, \boldsymbol{\Sigma}_{22}),$$

de ahí que

$$LR \sim \Lambda_{p_2}(p_1, n - p_1 - 1),$$

y H_0 es rechazado si LR es muy pequeño.

Matriz de dispersión diagonal

Un caso especial de la hipótesis anterior es $H_0 : \boldsymbol{\Sigma} = \operatorname{diag}(\sigma_{11}, \dots, \sigma_{pp})$. De este modo, tenemos:

$$LR = \frac{|\widehat{\boldsymbol{\Sigma}}|}{\widehat{\sigma}_{11} \cdots \widehat{\sigma}_{pp}} = \frac{|\boldsymbol{Q}|}{q_{11} \cdots q_{pp}},$$

donde $q_{rr} = \sum_{i=1}^{n}(x_{ir} - \overline{x}_r)^2 = n\widehat{\sigma}_{rr}$, para $r = 1, \dots, p$. De este modo, $LR \xrightarrow{\mathsf{D}} \chi^2(p(p-1)/2)$.

Note que

$$r_{jk} = \frac{q_{jk}}{\sqrt{q_{jj}q_{kk}}} = \frac{\widehat{\sigma}_{jk}}{\sqrt{\widehat{\sigma}_{jj}\widehat{\sigma}_{kk}}},$$

es la correlación muestral entre las variables j y k. Además, la matriz de correlación $\boldsymbol{R} = (r_{jk})$ está dada por

$$\boldsymbol{R} = \operatorname{diag}(q_{11}^{-1/2}, \dots, q_{pp}^{-1/2})\boldsymbol{Q}\operatorname{diag}(q_{11}^{-1/2}, \dots, q_{pp}^{-1/2}).$$

De ahí que

$$LR = \frac{|\boldsymbol{Q}|}{q_{11} \cdots q_{pp}} = |\boldsymbol{R}|,$$

es decir, el estadístico está basado en el determinante de la matriz de correlación.

Test de esfericidad

En práctica, se desea probar hipótesis del tipo $H_0 : \boldsymbol{\Sigma} = \sigma^2 \boldsymbol{I}$, donde σ^2 no es especificado. Note que bajo H_0, tenemos

$$\ell(\boldsymbol{\theta}) = -\frac{np}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \operatorname{tr}\{\boldsymbol{Q} + n(\overline{\boldsymbol{x}} - \boldsymbol{\mu})(\overline{\boldsymbol{x}} - \boldsymbol{\mu})^\top\}.$$

Además, $\widehat{\boldsymbol{\mu}} = \overline{\boldsymbol{x}}$. Por otro lado,

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \sigma^2} = -\frac{np}{2\sigma^2} + \frac{1}{2\sigma^4} \operatorname{tr}\{\boldsymbol{Q} + n(\overline{\boldsymbol{x}} - \widehat{\boldsymbol{\mu}})(\overline{\boldsymbol{x}} - \widehat{\boldsymbol{\mu}})^\top\},$$

de ahí que

$$\widehat{\sigma}^2 = \frac{1}{np} \operatorname{tr} \sum_{i=1}^{n} (\boldsymbol{x}_i - \overline{\boldsymbol{x}})(\boldsymbol{x}_i - \overline{\boldsymbol{x}})^\top = \frac{1}{p} \operatorname{tr} \widehat{\boldsymbol{\Sigma}}.$$

Por tanto,

$$LR = \Lambda^{2/n} = \frac{|\widehat{\boldsymbol{\Sigma}}|}{(\operatorname{tr} \widehat{\boldsymbol{\Sigma}}/p)^p} = \frac{|\boldsymbol{Q}|}{(\operatorname{tr} \boldsymbol{Q}/p)^p}$$

En efecto,

$$\begin{aligned}\max_{\mu,\Sigma} L(\boldsymbol{\mu},\boldsymbol{\Sigma}) &= (2\pi)^{-np/2}|\widehat{\sigma}^2\boldsymbol{I}|\exp\Big(-\frac{n}{2\widehat{\sigma}^2}\operatorname{tr}\widehat{\boldsymbol{\Sigma}}\Big)\\ &= (2\pi)^{-np/2}(\operatorname{tr}\widehat{\boldsymbol{\Sigma}}/p)^{-np/2}\exp\Big(-\frac{np}{2\operatorname{tr}\widehat{\boldsymbol{\Sigma}}}\operatorname{tr}\widehat{\boldsymbol{\Sigma}}\Big)\\ &= (2\pi)^{-np/2}(\operatorname{tr}\widehat{\boldsymbol{\Sigma}}/p)^{-np/2}e^{-np/2}.\end{aligned}$$

Note que $H_0 : \boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}$ es equivalente a probar que todos los valores propios λ_j de $\boldsymbol{\Sigma}$ son iguales, esto es

$$1 = \frac{\text{media geométrica de } \lambda_j}{\text{media aritmética de } \lambda_j} = \frac{\prod_j \lambda_j^{1/p}}{\sum_j \lambda_j/p} = \frac{|\boldsymbol{\Sigma}|^{1/p}}{\operatorname{tr}\boldsymbol{\Sigma}/p},$$

y substituímos $\boldsymbol{\Sigma}$ por su MLE $\widehat{\boldsymbol{\Sigma}}$, luego verificamos si el estadístico $LR^{1/p}$ es cercano a la unidad.

Test de equicorrelación

Una hipótesis que surge, por ejemplo, en análisis de varianza es:

$$H_0 : \mathbf{\Sigma} = \sigma^2 \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix} = \sigma^2[(1-\rho)\boldsymbol{I} + \rho \mathbf{1}\mathbf{1}^\top].$$

Si $\boldsymbol{S} = (s_{jk})$ es el estimador insesgado de $\mathbf{\Sigma}$, entonces los MLE de σ^2 y ρ son (ejercicio)

$$\widehat{\sigma}^2 = \frac{1}{p}\sum_{j=1}^{p} s_{jj}, \qquad \widehat{\sigma}^2\widehat{\rho} = \frac{1}{p(p-1)}\sum\sum_{j\neq k} s_{jk}.$$

Es estadístico de razón de verosimilitudes es

$$LR = \Lambda^{2/n} = \frac{|\boldsymbol{S}|}{\widehat{\sigma}^{2p}(1-\widehat{\rho})^{p-1}\{1+(p-1)\widehat{\rho}\}},$$

que es asintóticamente chi-cuadrado con $p(p+1)/2 - 2$ grados de libertad.

Suponga $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n_1}$ muestra aleatoria desde $\mathsf{N}_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ y una muestra independiente $\boldsymbol{y}_1, \dots, \boldsymbol{y}_{n_2}$ desde $\mathsf{N}_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$ y considere la hipótesis $H_0 : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ $(= \boldsymbol{\Sigma})$. La función de verosimilitud puede ser escrita como:

$$L_{12}(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2) = L_1(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) L_2(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2),$$

donde

$$L_i(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) = (2\pi)^{-n_i p/2} |\boldsymbol{\Sigma}_i|^{-n_i/2} \exp\Big\{ -\frac{1}{2} \operatorname{tr} \boldsymbol{\Sigma}_i^{-1} \sum_{i=1}^{n} (\boldsymbol{x}_i - \boldsymbol{\mu}_i)(\boldsymbol{x}_i - \boldsymbol{\mu}_i)^\top \Big\},$$

maximizar L_{12} es equivalente a la maximización simultánea de cada L_i $(i = 1, 2)$, de modo que L_{12} es maximizada en

$$\widehat{\boldsymbol{\mu}}_1 = \overline{\boldsymbol{x}}, \quad \widehat{\boldsymbol{\mu}}_2 = \overline{\boldsymbol{y}}, \quad \widehat{\boldsymbol{\Sigma}}_1 = \boldsymbol{Q}_1/n_1, \quad \widehat{\boldsymbol{\Sigma}}_2 = \boldsymbol{Q}_2/n_2.$$

De este modo,

$$\begin{aligned} L_{12}(\widehat{\boldsymbol{\mu}}_1, \widehat{\boldsymbol{\mu}}_2, \widehat{\boldsymbol{\Sigma}}_1, \widehat{\boldsymbol{\Sigma}}_2) &= L_1(\widehat{\boldsymbol{\mu}}_1, \widehat{\boldsymbol{\Sigma}}_1) L_2(\widehat{\boldsymbol{\mu}}_2, \widehat{\boldsymbol{\Sigma}}_2) \\ &= (2\pi)^{-n_1 p/2} |\widehat{\boldsymbol{\Sigma}}_1|^{-n_1/2} e^{-n_1 p/2} (2\pi)^{-n_2 p/2} |\widehat{\boldsymbol{\Sigma}}_2|^{-n_2/2} e^{-n_2 p/2} \\ &= (2\pi)^{-np/2} |\widehat{\boldsymbol{\Sigma}}_1|^{-n_1/2} |\widehat{\boldsymbol{\Sigma}}_2|^{-n_2/2} e^{-np/2} \end{aligned}$$

Haciendo $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{\Sigma}$, en este caso deseamos maximizar

$$
\begin{aligned}
\log L_{12}(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \mathbf{\Sigma}) &= \log L_1(\boldsymbol{\mu}_1, \mathbf{\Sigma}) + \log L_2(\boldsymbol{\mu}_2, \mathbf{\Sigma}) \\
&= c - \frac{n}{2} \log |\mathbf{\Sigma}| - \frac{1}{2} \operatorname{tr} \mathbf{\Sigma}^{-1} \Big\{ \sum_{i=1}^{n_1} (\boldsymbol{x}_i - \boldsymbol{\mu}_1)(\boldsymbol{x}_i - \boldsymbol{\mu}_1)^\top \\
&+ \sum_{i=1}^{n_2} (\boldsymbol{y}_i - \boldsymbol{\mu}_2)(\boldsymbol{y}_i - \boldsymbol{\mu}_2)^\top \Big\} \\
&= c - \frac{n}{2} \log |\mathbf{\Sigma}| - \frac{1}{2} \operatorname{tr} \mathbf{\Sigma}^{-1} \boldsymbol{Q} + \frac{1}{2} \operatorname{tr} \mathbf{\Sigma}^{-1} \{ n_1 (\overline{\boldsymbol{x}} - \boldsymbol{\mu}_1)(\overline{\boldsymbol{x}} - \boldsymbol{\mu}_1)^\top \\
&+ n_2 (\overline{\boldsymbol{y}} - \boldsymbol{\mu}_2)(\overline{\boldsymbol{y}} - \boldsymbol{\mu}_2)^\top \},
\end{aligned}
$$

donde $c = -\frac{np}{2} \log 2\pi$ y $\boldsymbol{Q} = \boldsymbol{Q}_1 + \boldsymbol{Q}_2$. Como

$$
\operatorname{tr} \left(n_1 \mathbf{\Sigma}^{-1} (\overline{\boldsymbol{x}} - \boldsymbol{\mu}_1)(\overline{\boldsymbol{x}} - \boldsymbol{\mu}_1)^\top \right) = n_1 (\overline{\boldsymbol{x}} - \boldsymbol{\mu}_1)^\top \mathbf{\Sigma}^{-1} (\overline{\boldsymbol{x}} - \boldsymbol{\mu}_1) \geq 0,
$$

y análogamente para $\boldsymbol{\mu}_2$. Se tiene que $\log L_{12}(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \mathbf{\Sigma})$ es maximizada para cualquier $\mathbf{\Sigma} > \mathbf{0}$ cuando $\boldsymbol{\mu}_1 = \overline{\boldsymbol{x}}$ y $\boldsymbol{\mu}_2 = \overline{\boldsymbol{y}}$.

De este modo,

$$\log L_{12}(\widehat{\boldsymbol{\mu}}_1, \widehat{\boldsymbol{\mu}}_2, \boldsymbol{\Sigma}) \geq \log L_{12}(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}),$$

y se desea maximizar

$$\log L_{12}(\widehat{\boldsymbol{\mu}}_1, \widehat{\boldsymbol{\mu}}_2, \boldsymbol{\Sigma}) = c - \frac{n}{2}\{\log|\boldsymbol{\Sigma}| + \operatorname{tr}\boldsymbol{\Sigma}^{-1}\boldsymbol{Q}/n\},$$

y es fácil notar que $\widehat{\boldsymbol{\Sigma}} = \boldsymbol{Q}/n$. En efecto,

$$\log L_{12}(\widehat{\boldsymbol{\mu}}_1, \widehat{\boldsymbol{\mu}}_2, \widehat{\boldsymbol{\Sigma}}) \geq \log L_{12}(\widehat{\boldsymbol{\mu}}_1, \widehat{\boldsymbol{\mu}}_2, \boldsymbol{\Sigma}) \geq \log L_{12}(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}),$$

de modo que $\widehat{\boldsymbol{\mu}}_1$, $\widehat{\boldsymbol{\mu}}_2$ y $\widehat{\boldsymbol{\Sigma}}$ son los MLE bajo $H_0 : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$, además

$$\log L_{12}(\widehat{\boldsymbol{\mu}}_1, \widehat{\boldsymbol{\mu}}_2, \widehat{\boldsymbol{\Sigma}}) = c - \frac{n}{2}\{\log|\widehat{\boldsymbol{\Sigma}}| + \operatorname{tr}\widehat{\boldsymbol{\Sigma}}^{-1}\widehat{\boldsymbol{\Sigma}}\},$$

o bien

$$L_{12}(\widehat{\boldsymbol{\mu}}_1, \widehat{\boldsymbol{\mu}}_2, \widehat{\boldsymbol{\Sigma}}) = (2\pi)^{-np/2}|\widehat{\boldsymbol{\Sigma}}|^{-n/2}e^{-np/2}.$$

El test de razón de verosimilitudes es

$$\begin{aligned}\Lambda &= \frac{\max_{H_0} L_{12}(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma})}{\max_{H_1} L_{12}(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)} = \frac{L_{12}(\widehat{\boldsymbol{\mu}}_1, \widehat{\boldsymbol{\mu}}_2, \widehat{\boldsymbol{\Sigma}})}{L_{12}(\widehat{\boldsymbol{\mu}}_1, \widehat{\boldsymbol{\mu}}_2, \widehat{\boldsymbol{\Sigma}}_1, \widehat{\boldsymbol{\Sigma}}_2)} \\ &= \frac{|\widehat{\boldsymbol{\Sigma}}|^{-n/2}}{|\widehat{\boldsymbol{\Sigma}}_1|^{-n_1/2}|\widehat{\boldsymbol{\Sigma}}_2|^{-n_2/2}} = c_{12}\frac{|\boldsymbol{Q}_1|^{n_1/2}|\boldsymbol{Q}_2|^{n_2/2}}{|\boldsymbol{Q}|^{n/2}},\end{aligned}$$

con $c_{12} = n^{np/2}/(n_1^{n_1p/2} n_2^{n_2p/2})$. De este modo

$$LR = -2\log\Lambda \xrightarrow{\mathsf{D}} \chi^2(\nu),$$

con $\nu = p(p+1)/2$ (bajo H_0).

Suponga $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n_1}$ muestra aleatoria desde $\mathsf{N}_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ independiente de la muestra $\boldsymbol{y}_1, \dots, \boldsymbol{y}_{n_2}$ desde $\mathsf{N}_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$. Se desea probar $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$[1] asumiendo que $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 \ (= \boldsymbol{\Sigma})$. Sabemos que

$$\overline{\boldsymbol{x}} \sim \mathsf{N}_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}/n_1), \qquad \boldsymbol{Q}_1 = (n_1 - 1)\boldsymbol{S}_1 \sim \mathsf{W}_p(n_1 - 1, \boldsymbol{\Sigma}),$$

y $\overline{\boldsymbol{x}}$ es independiente de $\boldsymbol{Q}_1$. Análogamente $\overline{\boldsymbol{y}} \sim \mathsf{N}_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}/n_2)$ y es independiente de $\boldsymbol{Q}_2 = (n_2 - 1)\boldsymbol{S}_2 \sim \mathsf{W}_p(n_2 - 1, \boldsymbol{\Sigma})$. Dado que $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n_1}$ y $\boldsymbol{y}_1, \dots, \boldsymbol{y}_{n_2}$ son independientes, tenemos que

$$\overline{\boldsymbol{x}} - \overline{\boldsymbol{y}} \sim \mathsf{N}_p\Big(\boldsymbol{\delta}, \Big(\frac{1}{n_1} + \frac{1}{n_2}\Big)\boldsymbol{\Sigma}\Big),$$

$$\boldsymbol{Q} = \boldsymbol{Q}_1 + \boldsymbol{Q}_2 \sim \mathsf{W}_p(n_1 + n_2 - 2, \boldsymbol{\Sigma}),$$

y $\overline{\boldsymbol{x}} - \overline{\boldsymbol{y}}$ es independiente de $\boldsymbol{Q}$.

[1] o equivalentemente $H_0 : \boldsymbol{\delta} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \boldsymbol{0}$.

Defina $\boldsymbol{S}_{\mathsf{P}} = \boldsymbol{Q}/(n_1 + n_2 - 2)$, luego

$$\begin{aligned} T^2 &= \frac{n_1 n_2}{n_1 + n_2} (\overline{\boldsymbol{x}} - \overline{\boldsymbol{y}} - \boldsymbol{\delta})^\top \boldsymbol{S}_{\mathsf{P}}^{-1} (\overline{\boldsymbol{x}} - \overline{\boldsymbol{y}} - \boldsymbol{\delta}) \sim \mathsf{T}^2(p, n1 + n_2 - 2) \\ &\stackrel{\mathsf{d}}{=} \frac{p(n_1 + n_2 - 2)}{n_1 + n_2 - p - 1}\, F(p, n_1 + n_2 - p - 1). \end{aligned}$$

Para probar $H_0 : \boldsymbol{\delta} = \mathbf{0}$ usamos el estadístico

$$T_0^2 = \frac{n_1 n_2}{n_1 + n_2} (\overline{\boldsymbol{x}} - \overline{\boldsymbol{y}})^\top \boldsymbol{S}_{\mathsf{P}}^{-1} (\overline{\boldsymbol{x}} - \overline{\boldsymbol{y}}),$$

y se rechaza la hipótesis H_0 a un nivel de significancia α si

$$T_0^2 \geq \frac{p(n_1 + n_2 - 2)}{n_1 + n_2 - p - 1}\, F_{1-\alpha}(p, n_1 + n_2 - p - 1).$$