Inferencia para la normal multivariada, Parte II

Graciela Boente

Estadístico U de Rao

Sean $\mathbf{y} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ con $\boldsymbol{\Sigma} > 0$ y $\mathbf{W} \sim \mathcal{W}(\boldsymbol{\Sigma}, p, m)$ independientes entre s , $m \geq p$.

Particionemos a $\mathbf{y}$, $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ y $\mathbf{W}$ de la siguiente forma

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}^{(1)} \\ \mathbf{y}^{(2)} \end{pmatrix} \qquad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}^{(1)} \\ \boldsymbol{\mu}^{(2)} \end{pmatrix}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \qquad \mathbf{W} = \begin{pmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{pmatrix}$$

con $\mathbf{y}^{(i)}, \boldsymbol{\mu}^{(i)} \in \mathbb{R}^{p_i}$, $\boldsymbol{\Sigma}_{ii} \in \mathbb{R}^{p_i \times p_i}$, $\mathbf{W}_{ii} \in \mathbb{R}^{p_i \times p_i}$, $p_1 + p_2 = p$.

Sean

$$T^2_{p;m} = m\, \mathbf{y}^{\mathrm{T}} \mathbf{W}^{-1} \mathbf{y} \qquad T^2_{p_1;m} = m\, \mathbf{y}^{(1)^{\mathrm{T}}} \mathbf{W}^{-1}_{11} \mathbf{y}^{(1)}$$

Definamos

$$\lambda^2_p = \boldsymbol{\mu}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \qquad \lambda^2_{p_1} = \boldsymbol{\mu}^{(1)^{\mathrm{T}}} \boldsymbol{\Sigma}^{-1}_{11} \boldsymbol{\mu}^{(1)} .$$

Recordemos que

$$\frac{m-p+1}{p} \frac{T^2_{p;m}}{m} \sim \mathcal{F}_{p;m-p+1}(\lambda^2_p)$$

$$\frac{m-p_1+1}{p_1} \frac{T^2_{p_1;m}}{m} \sim \mathcal{F}_{p_1;m-p_1+1}(\lambda^2_{p_1}) .$$

Estadístico *U* de Rao

Usando que

$$\boldsymbol{\Sigma}^{-1} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{\Sigma}_{11}^{-1} + \boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{\Sigma}_{22:1}^{-1} \boldsymbol{\beta} & -\boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{\Sigma}_{22:1}^{-1} \\ -\boldsymbol{\Sigma}_{22:1}^{-1} \boldsymbol{\beta} & \boldsymbol{\Sigma}_{22:1}^{-1} \end{pmatrix}$$

con

$$\boldsymbol{\beta} = \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \in \mathbb{R}^{p_2 \times p_1}$$

es facil ver que

$$\lambda_p^2 - \lambda_{p_1}^2 = \boldsymbol{\mu}_{2:1}^{\mathrm{T}} \, \boldsymbol{\Sigma}_{22:1}^{-1} \, \boldsymbol{\mu}_{2:1}$$

con

$$\begin{aligned} \boldsymbol{\mu}_{2:1} &= \boldsymbol{\mu}^{(2)} - \boldsymbol{\beta} \boldsymbol{\mu}^{(1)} \\ \boldsymbol{\Sigma}_{22:1} &= \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12} \end{aligned}$$

Nos interesara testear $H_0 : \lambda_p^2 = \lambda_{p_1}^2$ que es equivalente a $\boldsymbol{\mu}_{2:1} = 0$.

Observemos que como

$$\lambda_p^2 - \lambda_{p_1}^2 = \boldsymbol{\mu}_{2:1}^{\mathrm{T}} \boldsymbol{\Sigma}_{22:1}^{-1} \boldsymbol{\mu}_{2:1}$$

y

$$\boldsymbol{\mu}_{2:1} = \boldsymbol{\mu}^{(2)} - \boldsymbol{\beta}\boldsymbol{\mu}^{(1)} \qquad \boldsymbol{\beta} = \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}$$

- H_0 equivale a $\boldsymbol{\mu}^{(2)} = 0$ si $\boldsymbol{\mu}^{(1)} = 0$
- H_0 equivale a $\boldsymbol{\mu}^{(2)} = 0$ si $\mathbf{y}^{(1)}$ y $\mathbf{y}^{(2)}$ son independientes, o sea, si $\boldsymbol{\Sigma}_{12} = \mathbf{0}$

Estadístico U de Rao

Para testear H_0, nos basaremos en

$$T^2_{p;m} - T^2_{p_1;m} = m\, \mathbf{y}_{2:1}^{\mathrm{T}} \mathbf{W}_{22:1}^{-1} \mathbf{y}_{2:1}$$

donde

$$\begin{aligned} \mathbf{y}_{2:1} &= \mathbf{y}^{(2)} - \mathbf{B}\mathbf{y}^{(1)} \qquad \mathbf{B} = \mathbf{W}_{21}\mathbf{W}_{11}^{-1} \in \mathbb{R}^{p_2 \times p_1} \\ \mathbf{W}_{22:1} &= \mathbf{W}_{22} - \mathbf{W}_{21}\mathbf{W}_{11}^{-1}\mathbf{W}_{12} \end{aligned}$$

El estadístico $U-$de Rao se define por

$$U = \left\{1 + \frac{T^2_{p_1;m}}{m}\right\}\left\{1 + \frac{T^2_{p;m}}{m}\right\}^{-1}$$

Como

$$\frac{T^2_{p;m} - T^2_{p_1;m}}{m + T^2_{p_1;m}} = \frac{1}{U} - 1$$

Rechazaremos H_0 si $T^2_{p;m} - T^2_{p_1;m}$ es grande o sea si U es chico.

Lema 1

Sea $\mathbf{W} \sim \mathcal{W}(\mathbf{\Sigma}, p, m)$ con $m \geq p$ y $\mathbf{\Sigma} > 0$.

a) $\mathbf{W}_{22:1} \sim \mathcal{W}(\mathbf{\Sigma}_{22:1}, p - p_1, m - p_1)$.

b) $\mathbf{W}_{22:1}$ es independiente de $(\mathbf{W}_{11}, \mathbf{W}_{12})$ y por lo tanto de $\mathbf{B}$.

c) La distribucion condicional de $\mathbf{W}_{12}$ dado $\mathbf{W}_{11} = \mathbf{w}_{11}$ es

$$\mathbf{W}_{12}\Big|_{\mathbf{W}_{11}=\mathbf{w}_{11}} \sim N(\mathbf{w}_{11}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}, \mathbf{w}_{11} \otimes \mathbf{\Sigma}_{22:1})$$

y la de $\mathbf{W}_{21}$ dado $\mathbf{W}_{11} = \mathbf{w}_{11}$ es

$$\mathbf{W}_{21}\Big|_{\mathbf{W}_{11}=\mathbf{w}_{11}} \sim N(\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{w}_{11}, \mathbf{\Sigma}_{22:1} \otimes \mathbf{w}_{11})$$

de donde la distribucion condicional de $\mathbf{B} = \mathbf{W}_{21}\mathbf{W}_{11}^{-1}$ dado $\mathbf{W}_{11} = \mathbf{w}_{11}$ es

$$\mathbf{B}\Big|_{\mathbf{W}_{11}=\mathbf{w}_{11}} \sim N(\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}, \mathbf{\Sigma}_{22:1} \otimes \mathbf{w}_{11}^{-1})$$

Corolario 1

Sea $\mathbf{W} \sim \mathcal{W}(\boldsymbol{\Sigma}, p, m)$ con $m \geq p$ y $\boldsymbol{\Sigma} > 0$ y de namos

$$\mathbf{A} = \mathbf{W}_{21}\, \mathbf{W}_{11}^{-1}\, \mathbf{W}_{12} \in \mathbb{R}^{(p-p_1)\times(p-p_1)} .$$

Entonces, si $\boldsymbol{\Sigma}_{12} = \mathbf{0}$, se cumple que

- $\mathbf{A} \sim \mathcal{W}(\boldsymbol{\Sigma}_{22}, p - p_1, p_1)$ y
- $\mathbf{A}$ es independiente de $\mathbf{W}_{22:1}$ y $\mathbf{W}_{22:1} \sim \mathcal{W}(\boldsymbol{\Sigma}_{22:1}, p - p_1, m - p_1) = \mathcal{W}(\boldsymbol{\Sigma}_{22}, p - p_1, m - p_1)$

Corolario 2

Sea $\mathbf{W} \sim \mathcal{W}(\boldsymbol{\Sigma}, p, m)$ con $m \geq p$ y $\boldsymbol{\Sigma} > 0$ y de namos

$$\mathbf{A} = \mathbf{W}_{12} \mathbf{W}_{22}^{-1} \mathbf{W}_{21} \in \mathbb{R}^{(p-p_2)\times(p-p_2)} .$$

Entonces, si $\boldsymbol{\Sigma}_{21} = \mathbf{0}$, se cumple que

- $\mathbf{A} \sim \mathcal{W}(\boldsymbol{\Sigma}_{11}, p - p_2, p_2)$ y
- $\mathbf{A}$ es independiente de $\mathbf{W}_{11:2}$ y $\mathbf{W}_{11:2} \sim \mathcal{W}(\boldsymbol{\Sigma}_{11:2}, p - p_2, m - p_2) = \mathcal{W}(\boldsymbol{\Sigma}_{11}, p - p_2, m - p_2)$

Teorema 1

Sean $\mathbf{y} \sim N(\boldsymbol{\mu}, \mathbf{\Sigma})$ con $\mathbf{\Sigma} > 0$ y $\mathbf{W} \sim \mathcal{W}(\mathbf{\Sigma}, p, m)$ independientes entre s , $m \geq p$.

$$T^2_{2:1} = (m - p_1)\,\frac{T^2_{p;m} - T^2_{p_1;m}}{m + T^2_{p_1;m}}$$

a) Bajo $H_0 : \lambda^2_p = \lambda^2_{p_1}$ que es equivalente a $H_0 : \boldsymbol{\mu}_{2:1} = 0$ se tiene

$$\frac{m-p+1}{p-p_1}\,\frac{T^2_{2:1}}{m-p_1} = \frac{m-p+1}{p-p_1}\,\frac{T^2_{p;m} - T^2_{p_1;m}}{m + T^2_{p_1;m}} \sim \mathcal{F}_{p-p_1;m-p+1}$$

y $T^2_{2:1}$ es independiente de $T^2_{p_1}$

El factor $m + T^2_{p_1;m}$ aparece pues la $\mathrm{VAR}(\mathbf{y}_{2:1}\big|_{\mathbf{y}^{(1)};\mathbf{W}_{11}}) = \mathbf{\Sigma}_{22:1}\,(1 + T^2_{p_1;m}/m)$

Teorema 1

b) Si $\lambda_p^2 \neq \lambda_{p_1}^2$, la distribucion de $T_{2:1}^2$ condicional a $T_{p_1}^2$ es un Hotelling no central, o sea,

$$\frac{m-p+1}{p-p_1} \frac{T_{2:1}^2}{m-p_1}\Big|_{T_{p_1}^2} \sim \mathcal{F}_{p-p_1;m-p+1}(\nu)$$

con

$$\nu = \frac{\lambda_p^2 - \lambda_{p_1}^2}{1 + \frac{T_{p_1}^2}{m}}$$

Observación

El estad stico anterior se aplicara cuando $\mathbf{y} = \sqrt{n}\overline{\mathbf{x}} \sim N(\sqrt{n}\,\mu, \mathbf{\Sigma})$ y $\mathbf{W} = \mathbf{Q} \sim \mathcal{W}(\mathbf{\Sigma}, p, m)$, $m = n - 1$.

Supongamos que $p_1 = p - 1$ luego $p_2 = 1$.

Por lo tanto

- $\mathbf{y}^{(2)} = \sqrt{n}\,\overline{x}_p$,
- $\mathbf{y}^{(1)} = \sqrt{n}\,\overline{\mathbf{x}}_{(p-1)} = \sqrt{n}\,(\overline{x}_1, \ldots, \overline{x}_{p-1})^{\mathrm{T}}$.
- $\mathbf{W} = \mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{(p-1;p-1)} & \mathbf{q}_{(p-1;p)} \\ \mathbf{q}^{\mathrm{T}}_{(p-1;p)} & q_{pp} \end{pmatrix}$ con $\mathbf{q}_{(p-1;p)} \in \mathbb{R}^{p-1}$
- $\mathbf{B} = \mathbf{q}^{\mathrm{T}}_{(p-1;p)}\mathbf{Q}^{-1}_{(p-1;p-1)}$, $\mathbf{B}^{\mathrm{T}} \in \mathbb{R}^{p-1}$

Luego

$$T^2_{p;n-1} - T^2_{p-1;n-1} = (n-1)\, \mathbf{y}_{2:1}^{\mathrm{T}} \mathbf{W}_{22:1}^{-1} \mathbf{y}_{2:1}$$

donde ahora

$$\begin{aligned} \mathbf{y}_{2:1} &= \sqrt{n}\left(\overline{x}_p - \mathbf{B}\overline{\mathbf{x}}_{(p-1)}\right) \\ \mathbf{W}_{22:1} &= q_{pp} - \mathbf{q}_{(p-1;p)}^{\mathrm{T}} \mathbf{Q}_{(p-1;p-1)}^{-1} \mathbf{q}_{(p-1;p)} = (n-1) s_{p:(p-1)} \end{aligned}$$

Es decir,

$$T^2_{p;n-1} - T^2_{p-1;n-1} = n\, \frac{\left(\overline{x}_p - \mathbf{B}\overline{\mathbf{x}}_{(p-1)}\right)^2}{s_{p:(p-1)}} = T^2_{p|1;\ldots;p-1}$$

Como

$$T^2_{p;n-1} - T^2_{p-1;n-1} = n\,\frac{\left(\overline{x}_p - \mathbf{B}\overline{\mathbf{x}}_{(p-1)}\right)^2}{s_{p:(p-1)}} = T^2_{p|1;\ldots;p-1}$$

obtenemos

$$T^2_{p;n-1} = T^2_{p-1;n-1} + n\,\frac{\left(\overline{x}_p - \mathbf{B}\overline{\mathbf{x}}_{(p-1)}\right)^2}{s_{p:(p-1)}} = T^2_{p-1;n-1} + T^2_{p|1;\ldots;p-1}$$

que si se aplica sucesivamente sobre todas las variables da la descomposicion de Mason, Young and Tracy (MYT) (1995, 1997, 1999) y permite entender el procedimiento Step down que daremos.

$$T^2_{p;n-1} = T^2_{1;n-1} + T^2_{2|1} + T^2_{3|1;2} + \cdots + T^2_{p-1|1;\ldots;p-2} + T^2_{p|1;\ldots;p-1}$$

Aplicación del test de Rao

Una aplicacion del test de Rao es el metodo *step-down* para testear $H_0 : \boldsymbol{\mu} = \mathbf{0}$ basandonos en una muestra $\mathbf{x}_1, \ldots, \mathbf{x}_n$ i.i.d., $\mathbf{x}_i \in \mathbb{R}^p$, $\mathbf{x}_i \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > 0$, $n > p$.

Sean

- $\boldsymbol{\mu}_{(k)} = (\mu_1, \ldots, \mu_k)^{\mathrm{T}}$, $k = 1, \ldots, p$ y $\boldsymbol{\mu}_{(0)} = 0$.
- $\mathbf{x}_{i;(k)} = (x_{i;1}, \ldots, x_{i;k})^{\mathrm{T}}$ o sea, consideramos las primeras k componentes de $\mathbf{x}_i$.
- $T^2_{0;(k)} = n\,\overline{\mathbf{x}}^{\mathrm{T}}_{(k)}\mathbf{S}^{-1}_{(k)}\overline{\mathbf{x}}_{(k)}$ el estad stico de Hotelling basado en $\mathbf{x}_{1;(k)}, \ldots, \mathbf{x}_{n;(k)}$ para testear $\boldsymbol{\mu}_{(k)} = \mathbf{0}$ donde

$$\overline{\mathbf{x}}_{(k)} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_{i;(k)} \quad \mathbf{S}_{(k)} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_{i;(k)} - \overline{\mathbf{x}}_{(k)})((\mathbf{x}_{i;(k)} - \overline{\mathbf{x}}_{(k)})^{\mathrm{T}}$$

- Sea $H_{0;k} : \mu_k - \boldsymbol{\beta}_{(k-1)}\boldsymbol{\mu}_{(k-1)} = 0$
- Para $k = 1, \ldots p$ de namos $T_{0;(0)} = 0$ y

$$F_k = \frac{T^2_{0;(k)} - T^2_{0;(k-1)}}{(n-1) + T^2_{0;(k-1)}}(n-k)$$

- $\boldsymbol{\beta}_{(k-1)} = \sigma^{\mathrm{T}}_{(k;k-1)}\boldsymbol{\Sigma}^{-1}_{(k-1)}$ con
  - ⋆ $\boldsymbol{\Sigma}_{(k)} = \mathrm{Var}(\mathbf{x}_{(k)})$
  - ⋆ $\sigma_{(k,k-1)} = \mathrm{Cov}(x_k, \mathbf{x}_{(k-1)})$

Por el Teorema 1, si $H_{0;k}$ es cierta,

- $F_k \sim \mathcal{F}_{1;n-k}$ y
- F_k es independiente de $T^2_{0;(k-1)}$, mas aun
- F_k es independiente de $\{F_j : j < k\}$.

Aplicación del test de Rao

$H_{0;k} : \mu_k - \boldsymbol{\beta}_{(k-1)}\boldsymbol{\mu}_{(k-1)} = 0$

Si $\cap_{k=1}^{r-1} H_{0;k}$ es cierta, entonces

a) $\boldsymbol{\mu}_{(r-1)} = \mathbf{0}$

b) ademas, $H_{0;r}$ sera cierta si y solo si $\mu_r = 0$.

Con lo cual $H_0 : \boldsymbol{\mu} = \mathbf{0}$ puede escribirse como

$$H_0 = \bigcap_{k=1}^{p} H_{0;k}$$

Es decir, H_0 puede testearse sucesivamente con los estad sticos $F_1, \ldots, F_k$ en este orden.

H_0 se acepta si $F_j < f_{1;n-j}(\alpha_j)$ para todo $j = 1, \ldots, p$.

Luego el test tendra nivel α si

$$1 - \alpha = \prod_{k=1}^{p}(1 - \alpha_k)$$

Este test es un procedimiento alternativo al test de Hotelling que dimos antes y se usa si hay un orden a priori entre las medias μ_k.

Aplicación del test de Rao: Observaciones

Para que el test tenga nivel α necesitamos que

$$1-\alpha = \prod_{k=1}^{p}(1-\alpha_k)$$

Supongamos que $\alpha_1 = \cdots = \alpha_k = \alpha_0$ entonces la condicion es

$$1-\alpha = (1-\alpha_0)^p$$

de donde

$$\alpha_0 = 1-(1-\alpha)^{\frac{1}{p}}$$

p	α		
	0.05	0.01	0.001
2	0.0253	0.0050	0.00050
4	0.0127	0.0025	0.00025
5	0.0102	0.0020	0.00020
10	0.0051	0.0010	0.00010

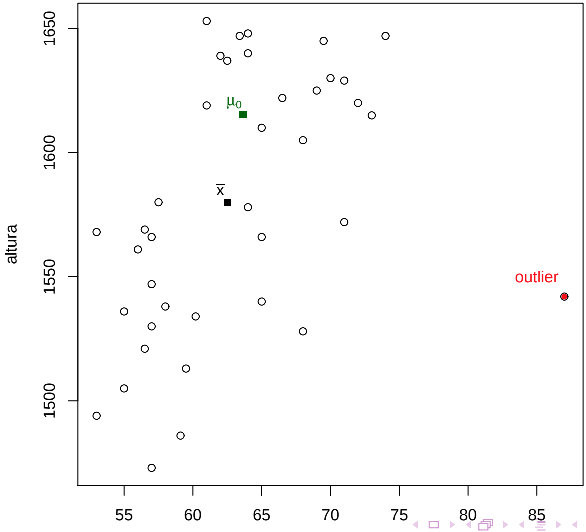
Valores de α_0

Ejemplo

Peso (kg) y Altura (mm) de 39 Indios Peruanos

peso	71	56.5	56	61	65	62	53	53
altura	1629	1569.0	1561	1619	1566	1639	1494	1568
peso	65	57	66.5	59.1	64	69.5	64	56.5
altura	1540	1530	1622.0	1486.0	1578	1645.0	1648	1521.0
peso	57	55	57	58	59.5	61	57	57.5
altura	1547	1505	1473	1538	1513.0	1653	1566	1580.0
peso	74	72	62.5	68	63.4	68	69	73
altura	1647	1620	1637.0	1528	1647.0	1605	1625	1615
peso	64	65	71	60.2	55	70	87	
altura	1640	1610	1572	1534.0	1536	1630	1542	

Ejemplo



altura
1500
1550
1600
1650
55
60
65
70
75
80
85
μ_0
$\bar{x}$
outlier

Ejemplo

Queremos testear

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}_0 \qquad H_1 : \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$$

con $\boldsymbol{\mu}_0 = (63.64, 1615.38)^{\mathrm{T}}$.

Haremos el test con todos los datos menos la observacion $(87, 1542)^{\mathrm{T}}$ que es un dato at pico.

Tenemos que

$$T_0^2 = n(\overline{\mathbf{x}} - \boldsymbol{\mu}_0)^{\mathrm{T}}\mathbf{S}^{-1}(\overline{\mathbf{x}} - \boldsymbol{\mu}_0) = 20.27881$$

y

$$F_0 = \frac{n-p}{p}\frac{T_0^2}{n-1} = 9.865$$

Como $f_{2;36}(0.001) = 8.420$ rechazo con nivel $\alpha = 0.001$.

Mas aun, el $p-$valor es 0.00038.

Método Step Down

Tenemos que

$$T^2_{(1)} = n\,\frac{(\bar{\mathbf{x}}_1 - \mu_{0;1})^2}{s_{11}} \quad \text{y} \quad T^2_{(2)} = n\,(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)^{\mathrm{T}}\mathbf{S}^{-1}(\bar{\mathbf{x}} - \boldsymbol{\mu}_0)$$

y

$$F_1 = T^2_{(1)} \sim \mathcal{F}_{1;n-1} \quad \text{y} \quad F_2 = (n-2)\frac{T^2_{(2)} - T^2_{(1)}}{n - 1 + T^2_{(1)}} \sim \mathcal{F}_{1;n-2}$$

Luego, si $\alpha = 0.001$ obtenemos $\alpha_0 = 0.00050$ y

	F_j	$f_{1;n-j}(\alpha_0)$	Decision
$j = 1$	1.297221	14.55683	No rechazo
$j = 2$	17.84299	14.63394	Rechazo

Valores de F_j y $f_{1;n-j}(\alpha_0)$, $j = 1, 2$

$$R_{1;2} = \max_{\mathbf{a}\neq 0} \frac{\mathrm{Cov}(x_1, \mathbf{a}^{\mathrm{T}}\mathbf{x}^{(2)})}{\left(\mathrm{Var}(x_1)\mathrm{Var}(\mathbf{a}^{\mathrm{T}}\mathbf{x}^{(2)})\right)^{\frac{1}{2}}} = \max_{\mathbf{a}\neq 0} \frac{\mathbf{a}^{\mathrm{T}}\boldsymbol{\sigma}_{21}}{(\sigma_{11}\mathbf{a}^{\mathrm{T}}\boldsymbol{\Sigma}_{22}\mathbf{a})^{\frac{1}{2}}}$$

Usando Cauchy-Schwartz se obtiene que

$$R_{1;2}^2 = \frac{\boldsymbol{\sigma}_{21}^{\mathrm{T}}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\sigma}_{21}}{\sigma_{11}}$$

y se alcanza en

$$\mathbf{a}^{\mathrm{T}} = \boldsymbol{\sigma}_{21}^{\mathrm{T}}\boldsymbol{\Sigma}_{22}^{-1} = \boldsymbol{\sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}$$

que era el coe ciente de la regresion de x_1 en $\mathbf{x}_2$ ya que

$$x_1|\mathbf{x}^{(2)} = \mathbf{x}_0 \sim N\left(\mu_1 + \boldsymbol{\sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_0 - \boldsymbol{\mu}^{(2)}), \boldsymbol{\Sigma}_{11:2}\right)$$

con $\boldsymbol{\Sigma}_{11:2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} = 1/\sigma^{11}$.

Queremos testear que x_1 es independiente de $(x_2, \ldots, x_p)$ basados en una muestra $\mathbf{x}_1, \ldots, \mathbf{x}_n$ i.i.d., $\mathbf{x}_i \in \mathbb{R}^p$, $\mathbf{x}_i \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > 0$, $n > p$.

$$H_0 : R_{1;2} = 0 \quad \text{o equivalentemente} \quad H_0 : \boldsymbol{\sigma}_{21} = \mathbf{0}$$

El test de cociente de maxima verosimilitud rechaza si $(1 - \widehat{R}^2_{1:2})^{n=2}$ es chico, donde

$$\widehat{R}^2_{1;2} = \frac{\mathbf{s}_{21}^{\mathrm{T}} \mathbf{S}_{22}^{-1} \mathbf{s}_{21}}{s_{11}}$$

y

$$\mathbf{S} = \frac{1}{n-1}\mathbf{Q} = \begin{pmatrix} s_{11} & \mathbf{s}_{12} \\ \mathbf{s}_{21} & \mathbf{S}_{22} \end{pmatrix} = \begin{pmatrix} s_{11} & \mathbf{s}_{21}^{\mathrm{T}} \\ \mathbf{s}_{21} & \mathbf{S}_{22} \end{pmatrix}$$

$$\widehat{R}^2_{1;2} = \frac{\mathbf{s}_{21}^{\mathrm{T}} \mathbf{S}_{22}^{-1} \mathbf{s}_{21}}{s_{11}}$$

Ahora bien

$$\frac{\widehat{R}^2_{1;2}}{1 - \widehat{R}^2_{1;2}} = \frac{\mathbf{s}_{21}^{\mathrm{T}} \mathbf{S}_{22}^{-1} \mathbf{s}_{21}}{s_{11:2}} = \frac{\mathbf{q}_{21}^{\mathrm{T}} \mathbf{Q}_{22}^{-1} \mathbf{q}_{21}}{q_{11:2}}$$

con $s_{11:2} = s_{11} - \mathbf{s}_{21}^{\mathrm{T}} \mathbf{S}_{22}^{-1} \mathbf{s}_{21} = q_{11:2}/(n-1)$.

Recordemos el Corolario 2 que enunciamos

Corolario 2

Sea $\mathbf{W} \sim \mathcal{W}(\mathbf{\Sigma}, p, m)$ con $m \geq p$ y $\mathbf{\Sigma} > 0$ y de namos $\mathbf{A} = \mathbf{W}_{12}\mathbf{W}_{22}^{-1}\mathbf{W}_{21} \in \mathbb{R}^{(p-p_2)\times(p-p_2)}$. Entonces, si $\mathbf{\Sigma}_{21} = \mathbf{0}$, se cumple que

- $\mathbf{A} \sim \mathcal{W}(\mathbf{\Sigma}_{11}, p-p_2, p_2)$ y
- $\mathbf{A}$ es independiente de $\mathbf{W}_{11.2} \sim \mathcal{W}(\mathbf{\Sigma}_{11}, p-p_2, m-p_2)$

Por lo tanto, como $\mathbf{Q} \sim \mathcal{W}(\mathbf{\Sigma}, p, n-1)$, $p_1 = 1$, $p_2 = p-1$ obtenemos que, bajo $H_0 : \boldsymbol{\sigma}_{21} = \mathbf{0}$

- $\mathbf{q}_{21}^{\mathrm{T}}\mathbf{Q}_{22}^{-1}\mathbf{q}_{21} \sim \mathcal{W}(\sigma_{11}, 1, p-1) = \sigma_{11}\,\chi^2_{p-1}$
- $\mathbf{q}_{21}^{\mathrm{T}}\mathbf{Q}_{22}^{-1}\mathbf{q}_{21}$ es independiente de $q_{11:2}$
- $q_{11:2} \sim \mathcal{W}(\mathbf{\Sigma}_{11}, 1, (n-1)-(p-1)) = \sigma_{11}\,\chi^2_{n-p}$

$$H_0 : R_{1,2} = 0 \quad \text{o equivalentemente} \quad H_0 : \boldsymbol{\sigma}_{21} = \mathbf{0}$$

Luego, como

$$\frac{\widehat{R}^2_{1;2}}{1-\widehat{R}^2_{1;2}} = \frac{\mathbf{q}_{21}^{\mathrm{T}}\mathbf{Q}_{22}^{-1}\mathbf{q}_{21}}{q_{11:2}} \text{ tenemos que } \frac{\widehat{R}^2_{1;2}}{1-\widehat{R}^2_{1;2}}\,\frac{n-p}{p-1} \sim \mathcal{F}_{p-1;n-p}$$

El test de cociente de maxima verosimilitud para testear H_0(o sea, x_1 es independiente de $(x_2, \ldots, x_p)$) resulta entonces

$$\phi(\mathbf{X}) = \begin{cases} 1 \quad \text{si} \quad \dfrac{\widehat{R}^2_{1;2}}{1-\widehat{R}^2_{1;2}}\,\dfrac{n-p}{p-1} > f_{p-1;n-p}(\alpha) \\ \\ 0 \quad \text{si} \quad \dfrac{\widehat{R}^2_{1;2}}{1-\widehat{R}^2_{1;2}}\,\dfrac{n-p}{p-1} \leq f_{p-1;n-p}(\alpha) \end{cases}$$

Test para el vector de coeficientes de regresión

De igual forma usando el Lema 1 podemos obtener un test para

$$H_{0;r} : \boldsymbol{\beta}_r = \boldsymbol{\beta}_{0;r}$$

donde $\boldsymbol{\beta}_{0;r} \in \mathbb{R}^{p_2}$ es un vector jo y $\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ con

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}^{(1)} \\ \mathbf{x}^{(2)} \end{pmatrix} \qquad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}^{(1)} \\ \boldsymbol{\mu}^{(2)} \end{pmatrix}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix} \qquad \boldsymbol{\beta} = \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_{p_1}) \in \mathbb{R}^{p_2 \times p_1}$$

basandonos en la $r-$esima columna de la matriz

$$\mathbf{B} = \mathbf{S}_{21}\mathbf{S}_{11}^{-1} = (\mathbf{b}_1, \ldots, \mathbf{b}_{p_1})$$

Test para el vector de coeficientes de regresión

Un test para H_0 esta dado por

$$\phi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad \dfrac{T^2}{n-1-p_1}\dfrac{n-p}{p-p_1} > f_{p-p_1;n-p}(\alpha) \\[2ex] 0 & \text{si} \quad \dfrac{T^2}{n-1-p_1}\dfrac{n-p}{p-p_1} \leq f_{p-p_1;n-p}(\alpha) \end{cases}$$

donde $\mathbf{S}^{-1} = (s^{ij})$ y

$$T^2 = (n-1-p_1)\frac{1}{s^{rr}}(\mathbf{b}_r - \boldsymbol{\beta}_{0;r})^{\mathrm{T}}\mathbf{S}_{22:1}^{-1}(\mathbf{b}_r - \boldsymbol{\beta}_{0;r})$$

Test sobre la matriz $\boldsymbol{\Sigma}$

Consideraremos los siguientes test sobre la matriz $\boldsymbol{\Sigma}$

- *Independencia de a bloques*, o sea

  $H_{01}: \boldsymbol{\Sigma}_{ij} = \mathbf{0}\ 1 \leq i < j \leq k$ o sea $H_0 : \boldsymbol{\Sigma} = \operatorname{diag}(\boldsymbol{\Sigma}_{11}, \ldots, \boldsymbol{\Sigma}_{kk})$

  Es decir, $x^{(1)}, \ldots, x^{(k)}$ son independientes.
  En particular,

  - ⋆ si $k = 2$, $H_0 : \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$, o sea $x^{(1)}$ y $x^{(2)}$ son independientes.
  - ⋆ si $k = 2$ y $p_1 = 1$, es el caso que consideramos anteriormente, tomando $x^{(1)} = x_1$ y $x^{(2)} = (x_2, \ldots, x_p)^{\mathrm{T}}$.
  - ⋆ si $p_r = 1$ para todo $1 \leq r \leq k$, tenemos que H_0 es equivalente a $H_{01} : \boldsymbol{\Sigma} = \operatorname{diag}(\sigma_{11}, \ldots, \sigma_{pp})$.

Test sobre la matriz $\mathbf{\Sigma}$

- *Test de esfericidad* Nos interesa testear
  - $H_{02} : \mathbf{\Sigma} = \sigma^2 \mathbf{I}_p$ con σ^2 desconocido.
  - $H_{03} : \mathbf{\Sigma} = \mathbf{I}_p$
- *Igualdad de bloques diagonales* $H_{04} : \mathbf{\Sigma}_{11} = \mathbf{\Sigma}_{22} = \cdots = \mathbf{\Sigma}_{kk}$ cuando $p_1 = p_2 = \cdots = p_k = p_0$
- *Iguales correlaciones e iguales varianzas*

$$H_{05} : \mathbf{\Sigma} = \sigma^2 \begin{pmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \vdots & \vdots & \vdots & \vdots \\ \rho & \rho & \dots & 1 \end{pmatrix}$$

que es un supuesto que aparece en modelos mixtos o con mediciones repetidas.

Criterio de Wilks

Tanto en el caso del test de independencia como en el modelo lineal multivariado es posible hallar dos matrices

- $\mathbf{U} \sim \mathcal{W}(\boldsymbol{\Sigma}, p, N-r)$ y
- $\mathbf{H} \sim \mathcal{W}(\boldsymbol{\Sigma}, p, r)$

independientes bajo la hipotesis nula de interes.

Mas generalmente, tendremos

- $\mathbf{U} \sim \mathcal{W}(\boldsymbol{\Sigma}, p, N-r)$ y
- $\mathbf{z}_j \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $1 \leq j \leq r$, $\mathbf{H} = \sum_{j=1}^{r} \mathbf{z}_j \mathbf{z}_j^{\mathrm{T}}$

donde muchas veces $r < p$ pero $N - r > p$.

El estad stico de Wilks se utiliza para testear cualquier hipotesis equivalente a $\boldsymbol{\mu} = \mathbf{0}$. Rechazaremos si el Wilks es pequeno.

Criterio de Wilks

En analog a con el caso univariado Wilks (1932) de nio el estad stico de Wilks.

Definición Sean $\mathbf{U} \sim \mathcal{W}(\boldsymbol{\Sigma}, p, N-r)$ y $\mathbf{z}_j \sim N(\mathbf{0}, \boldsymbol{\Sigma})$, $1 \leq j \leq r$ independientes entre s , el criterio de Wilks se de ne como

$$(N, p, r) = \frac{|\mathbf{U}|}{|\mathbf{U}+\mathbf{H}|}$$

donde $|\mathbf{A}|$ indica el determinante de la matriz $\mathbf{A} \in \mathbb{R}^{p \times p}$ y $\mathbf{H} = \sum_{j=1}^{r} \mathbf{z}_j \mathbf{z}_j^{\mathrm{T}} \sim \mathcal{W}(\boldsymbol{\Sigma}, p, r)$

N, p y r son los parametros del Wilks y corresponden respectivamente a los grados de libertad de $\mathbf{U}+\mathbf{H}$, la dimension de las matrices y los grados de libertad de $\mathbf{H}$.

Por otra parte, si $\mathbf{U}$ es inversible (N, p, r) depende solo de los autovalores de $\mathbf{H}\mathbf{U}^{-1}$.

Distribución del criterio de Wilks

a) Si $r \geq p$, $\mathbf{H} \sim \mathcal{W}(\mathbf{\Sigma}, p, r)$ tiene densidad y

$$(N, p, r) \sim \prod_{j=1}^{p} b_{ii}^2$$

con $b_{11}^2, \ldots, b_{pp}^2$ son independientes $b_{ii}^2 \sim \mathcal{B}\left(\frac{N-r+1-i}{2}, \frac{r}{2}\right)$

b) Si $r < p$,

$$(N, p, r) \sim \prod_{j=1}^{r} b_{ii}^2$$

con $b_{11}^2, \ldots, b_{rr}^2$ son independientes $b_{ii}^2 \sim \mathcal{B}\left(\frac{N-p+1-i}{2}, \frac{p}{2}\right)$

Es decir, $(N, p, r) \sim \quad (N, r, p)$

Corolario

a) Si $p = 1$,

$$\frac{1 - \quad (N,1,r)}{(N,1,r)} \, \frac{N-r}{r} \sim \mathcal{F}_{r;N-r}$$

b) Si $r = 1$,

$$\frac{1 - \quad (N,p,1)}{(N,p,1)} \, \frac{N-p}{p} \sim \mathcal{F}_{p;N-p}$$

c) Si $p = 2$,

$$\frac{1 - \quad (N,2,r)^{\frac{1}{2}}}{(N,2,r)^{\frac{1}{2}}} \, \frac{N-r-1}{r} \sim \mathcal{F}_{2r;2(N-r-1)}$$

d) Si $r = 2$,

$$\frac{1 - \quad (N,p,2)^{\frac{1}{2}}}{(N,p,2)^{\frac{1}{2}}} \, \frac{N-p-1}{p} \sim \mathcal{F}_{2p;2(N-p-1)}$$

Aproximaciones a la Distribución del criterio de Wilks

a) Barlett (1938) mostro que

$$\mathbb{P}\left(-f \log(\ (N,p,r)) > C_\alpha \chi^2_{pr,\alpha}\right) \approx \alpha$$

donde

$$f = N - r - \frac{1}{2}(p - r + 1) = N - \frac{1}{2}(p + r + 1)$$

Los valores C_α para esta aproximac on estan en el Apendice D13 de Seber (1984).

b) Rao (1951) mostro que

$$\frac{(fs + 2\lambda)}{2m} \frac{(1 - \ (N,p,r)^{\frac{1}{s}})}{(N,p,r)^{\frac{1}{s}}} \approx \mathcal{F}_{2m, fs+2\lambda}$$

donde $f = N - (p + r + 1)/2$

$$m = \frac{pr}{2} \quad \lambda = -\frac{(pr - 2)}{4} \quad s = \frac{(p^2 r^2 - 4)^{\frac{1}{2}}}{(p^2 + r^2 - 5)^{\frac{1}{2}}}$$

Otros criterios

Supongamos que $\mathbf{U}$ es inversible y sean $\lambda_1 \geq \cdots \geq \lambda_p$ los autovalores de $\mathbf{HU}^{-1}$.

$$(N, p, r) = \frac{1}{\prod_{j=1}^{p}(1 + \lambda_j)}$$

Hay otros criterios que se utilizan

a) **Criterio de Lawley Hotelling**

$$\mathrm{tr}(\mathbf{HU}^{-1}) = \sum_{j=1}^{p} \lambda_j$$

Los percentiles de la distribucion de este estad stico estan dados en el Apendice D15 de Seber (1984).

Otros criterios

b) **Criterio de Pillai**

$$\mathrm{tr}\left(\mathbf{H}(\mathbf{U}+\mathbf{H})^{-1}\right) = \sum_{j=1}^{p} \frac{\lambda_j}{1+\lambda_j}$$

Los percentiles de la distribucion de este estad stico estan dados en el Apendice D16 de Seber (1984).

c) **Criterio de Roy** o de la maxima ra z. Considera la maxima ra z $\theta_{\max}$ de $|\mathbf{H} - \theta(\mathbf{U}+\mathbf{H})| = 0$. Luego

$$\theta_{\max} = \frac{\lambda_1}{1+\lambda_1}$$

y rechazo si $\theta_{\max}$ es grande. Los percentiles de la distribucion de $\theta_{\max}$ estan dados en el Apendice D14 de Seber (1984).

Si $r = 1$, todos los criterios son equivalentes, en particular, tenemos que $\Lambda(N, p, r) = 1 - \theta_{\max}$.

Test de Independencia H_{01}

Dada muestra $\mathbf{x}_1, \ldots, \mathbf{x}_n$ i.i.d., $\mathbf{x}_i \in \mathbb{R}^p$, $\mathbf{x}_i \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > 0$, $n > p$. Particionemos a $\mathbf{x}_i$, $\boldsymbol{\mu}$, $\bar{\mathbf{x}}$, $\boldsymbol{\Sigma}$ y a $\mathbf{Q} = (n-1)\mathbf{S}$ como

$$\mathbf{x}_i = \begin{pmatrix} \mathbf{x}_{i1} \\ \vdots \\ \mathbf{x}_{ik} \end{pmatrix} \qquad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \vdots \\ \boldsymbol{\mu}_k \end{pmatrix} \quad \bar{\mathbf{x}} = \begin{pmatrix} \bar{\mathbf{x}}_1 \\ \vdots \\ \bar{\mathbf{x}}_k \end{pmatrix}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} & \ldots & \boldsymbol{\Sigma}_{1k} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} & \ldots & \boldsymbol{\Sigma}_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{\Sigma}_{k1} & \boldsymbol{\Sigma}_{k2} & \ldots & \boldsymbol{\Sigma}_{kk} \end{pmatrix} \qquad \mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} & \ldots & \mathbf{Q}_{1k} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} & \ldots & \mathbf{Q}_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{Q}_{k1} & \mathbf{Q}_{k2} & \ldots & \mathbf{Q}_{kk} \end{pmatrix}$$

donde $\boldsymbol{\mu}_r \in \mathbb{R}^{p_r}$ y $\boldsymbol{\Sigma}_{rr} \in \mathbb{R}^{p_r \times p_r}$.

Observemos que si $n > p$, $\mathbb{P}(\mathbf{Q} > 0) = 1$ luego $\mathbb{P}(\mathbf{Q}_{jj} > 0) = 1$, $j = 1, \ldots k$.

Test de Independencia $H_{01} : \boldsymbol{\Sigma} = \mathrm{diag}(\boldsymbol{\Sigma}_{11}, \ldots, \boldsymbol{\Sigma}_{kk})$

El test de cociente de maxima verosimilitud para H_0 esta basado en

$$\gamma^{?} = \left(\frac{|\mathbf{Q}|}{\prod_{j=1}^{k} |\mathbf{Q}_{jj}|} \right)^{\frac{n}{2}}$$

donde $-2\log(\gamma^{?}) \xrightarrow{D} \chi^2$ con $\nu = \frac{1}{2}\left(p^2 - \sum_{j=1}^{k} p_j^2\right)$.

Por lo tanto, el test sera

$$\phi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad \frac{|\mathbf{Q}|}{\prod_{j=1}^{k} |\mathbf{Q}_{jj}|} \leq k \\ 0 & \text{si} \quad \frac{|\mathbf{Q}|}{\prod_{j=1}^{k} |\mathbf{Q}_{jj}|} > k \end{cases}$$

donde para tener un test de nivel asintotico α,

$$k = \exp\left\{ -\frac{1}{n}\chi^2_{;} \right\}$$

Test de Independencia $H_{01} : \boldsymbol{\Sigma} = \mathrm{diag}(\boldsymbol{\Sigma}_{11}, \ldots, \boldsymbol{\Sigma}_{kk})$

En particular, si $k = p$, $H_{01} : \boldsymbol{\Sigma} = \mathrm{diag}(\sigma_{11}, \ldots, \sigma_{pp})$, estamos testeando la independencia de las p variables.

El test de cociente de maxima verosimilitud para H_{01} esta basado en

$$\gamma = \frac{|\mathbf{Q}|}{\prod_{j=1}^{p} |q_{jj}|} = |\mathbf{R}|$$

donde

$$\mathbf{R} = \mathrm{diag}\left(s_{11}^{-\frac{1}{2}}, \ldots, s_{pp}^{-\frac{1}{2}}\right) \mathbf{S}\, \mathrm{diag}\left(s_{11}^{-\frac{1}{2}}, \ldots, s_{pp}^{-\frac{1}{2}}\right)$$

es la matriz de correlacion muestral.

En este caso,

$$-n \log(|\mathbf{R}|) \xrightarrow{D} \chi^2 \quad \text{con} \quad \nu = \frac{p\,(p-1)}{2}$$

Test de Independencia $H_{01} : \mathbf{\Sigma} = \mathrm{diag}(\mathbf{\Sigma}_{11}, \mathbf{\Sigma}_{22})$

Si $k = 2$, estamos testeando $\mathbf{\Sigma}_{12} = 0$.

El test de cociente de maxima verosimilitud para H_0 esta basado en

$$\gamma = \frac{|\mathbf{Q}|}{|\mathbf{Q}_{11}|\,|\mathbf{Q}_{22}|}\,.$$

Pero

$$\begin{aligned} |\mathbf{Q}| &= |\mathbf{Q}_{11}|\,|\mathbf{Q}_{22} - \mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}\mathbf{Q}_{12}| = |\mathbf{Q}_{11}|\,|\mathbf{Q}_{22.1}| \\ &= |\mathbf{Q}_{11}|\,|\mathbf{Q}_{22}|\,|\mathbf{I}_{p_2} - \mathbf{Q}_{22}^{-1}\mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}\mathbf{Q}_{12}| \end{aligned}$$

con lo cual

$$\gamma = |\mathbf{I}_{p_2} - \mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}\mathbf{Q}_{12}\mathbf{Q}_{22}^{-1}| = \prod_{j=1}^{s}(1 - r_i^2)$$

donde $s = \min(p_1, p_2)$ y r_j^2 son los autovalores de $\mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}\mathbf{Q}_{12}\mathbf{Q}_{22}^{-1}$.

Test de Independencia $H_{01} : \mathbf{\Sigma} = \mathrm{diag}(\mathbf{\Sigma}_{11}, \mathbf{\Sigma}_{22})$

Por otra parte,

$$\gamma = \frac{|\mathbf{Q}_{22:1}|}{|\mathbf{Q}_{22}|} = \frac{|\mathbf{U}|}{|\mathbf{U} + \mathbf{H}|}$$

Ahora bien,

- $\mathbf{U} = \mathbf{Q}_{22:1} \sim \mathcal{W}(\mathbf{\Sigma}_{22:1}, p_2, n-1-p_1) = \mathcal{W}(\mathbf{\Sigma}_{22}, p_2, n-1-p_1)$ bajo H_{01}
- $\mathbf{H} = \mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}\mathbf{Q}_{12} \sim \mathcal{W}(\mathbf{\Sigma}_{22:1}, p_2, p_1) = \mathcal{W}(\mathbf{\Sigma}_{22}, p_2, p_1)$ bajo H_{01}
- $\mathbf{H}$ y $\mathbf{U}$ son independientes bajo H_{01}.

Luego,

$$\gamma = \quad (n-1, p_1, p_2)$$

Test de Independencia $H_{01} : \boldsymbol{\Sigma} = \text{diag}(\boldsymbol{\Sigma}_{11}, \boldsymbol{\Sigma}_{22})$

Queremos aplicar el principio de union intersecion de Roy para testear $H_{01} : \boldsymbol{\Sigma}_{12} = 0$.

Recordemos que

$$\mathbf{x}_i = \begin{pmatrix} \mathbf{x}_{i1} \\ \mathbf{x}_{i2} \end{pmatrix}$$

entonces $\mathrm{Cov}(\mathbf{a}^{\mathrm{T}}\mathbf{x}_{11}, \mathbf{b}^{\mathrm{T}}\mathbf{x}_{12}) = \mathbf{a}^{\mathrm{T}}\boldsymbol{\Sigma}_{12}\mathbf{b}$. De namos

$$H_{0;\mathbf{ab}} : \mathbf{a}^{\mathrm{T}}\boldsymbol{\Sigma}_{12}\mathbf{b} = 0$$

Luego

$$H_{01} = \cap_{\mathbf{a}\neq\mathbf{0}} \cap_{\mathbf{b}\neq\mathbf{0}} H_{0;\mathbf{ab}}$$

Test de Independencia $H_{01} : \mathbf{\Sigma} = \text{diag}(\mathbf{\Sigma}_{11}, \mathbf{\Sigma}_{22})$

Aplicando el principio de union interseccion, se obtiene el criterio de Roy.

Es decir, el test rechaza si

$$\theta_{\max} > k$$

donde $\theta_{\max}$ es la maxima ra z de $|\mathbf{H} - \theta(\mathbf{U} + \mathbf{H})| = 0$, o sea, $\theta_{\max}$ es el maximo autovalor de $\mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}\mathbf{Q}_{12}\mathbf{Q}_{22}^{-1}$.

Test de esfericidad $H_{02} : \mathbf{\Sigma} = \sigma^2 \mathbf{I}_p,\ \sigma^2 > 0$

El test de cociente de maxima verosimilitud para H_{02} esta basado en

$$\gamma^{?} = \left(\frac{|\mathbf{Q}|^{\frac{1}{p}}}{\frac{1}{p}\mathrm{tr}(\mathbf{Q})} \right)^{\frac{np}{2}}$$

donde $-2\log(\gamma^{?}) \xrightarrow{D} \chi^2$ con $\nu = \frac{p(p+1)}{2} - 1$. Por lo tanto, el test sera

$$\phi(\mathbf{X}) = \begin{cases} 1 \quad \text{si} \quad \frac{|\mathbf{Q}|^{\frac{1}{p}}}{\frac{1}{p}\mathrm{tr}(\mathbf{Q})} \leq k \\ \\ 0 \quad \text{si} \quad \frac{|\mathbf{Q}|^{\frac{1}{p}}}{\frac{1}{p}\mathrm{tr}(\mathbf{Q})} > k \end{cases}$$

donde para tener un test de nivel asintotico α,

$$k = \exp\left\{ -\frac{1}{np}\chi^2_{;} \right\}$$

Test para varias muestras

Tenemos ahora k poblaciones normales independientes.
Supongamos que la $i-$esima poblacion es $N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$.

Nos interesara testear

- $H_1 : \boldsymbol{\Sigma}_1 = \cdots = \boldsymbol{\Sigma}_k$
- $H_2 : \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_k \,, \boldsymbol{\Sigma}_1 = \cdots = \boldsymbol{\Sigma}_k$
- $H_3 : \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_k$ cuando sabemos que $\boldsymbol{\Sigma}_1 = \cdots = \boldsymbol{\Sigma}_k$, o sea, un analisis de la varianza de un factor multivariado.

Supongamos tener k muestras $\mathbf{x}_{i;1}, \ldots, \mathbf{x}_{i;n_i}$, $1 \leq i \leq k$ independientes. $\mathbf{x}_{i;j} \sim N(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$

Test para varias muestras

$$\mathbf{Q}_i = \sum_{j=1}^{n_i} (\mathbf{x}_{i;j} - \bar{\mathbf{x}}_i)(\mathbf{x}_{i;j} - \bar{\mathbf{x}}_i)^{\mathrm{T}} \qquad \widehat{\mathbf{\Sigma}}_i = \frac{\mathbf{Q}_i}{n_i} \qquad \bar{\mathbf{x}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{x}_{i;j}$$

La suma de cuadrados dentro de grupos es

$$\mathbf{U} = \mathbf{Q}_1 + \cdots + \mathbf{Q}_k = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (\mathbf{x}_{i;j} - \bar{\mathbf{x}}_i)(\mathbf{x}_{i;j} - \bar{\mathbf{x}}_i)^{\mathrm{T}}$$

Tenemos que

- $\mathbf{Q}_i \sim \mathcal{W}(\mathbf{\Sigma}_i, p, n_i - 1)$
- $\mathbf{Q}_1, \ldots, \mathbf{Q}_k$ son independientes.

Luego, bajo $H_1 : \mathbf{\Sigma}_1 = \cdots = \mathbf{\Sigma}_k$, $\mathbf{U} \sim \mathcal{W}(\mathbf{\Sigma}, p, n - k)$ con $n = \sum_{i=1}^{k} n_i$.

Test para varias muestras

Por otra parte, sea **H** la suma de cuadrados entre poblaciones.

$$\mathbf{H} = \sum_{i=1}^{k} n_i (\overline{\mathbf{x}}_i - \overline{\mathbf{x}})(\overline{\mathbf{x}}_i - \overline{\mathbf{x}})^{\mathrm{T}}$$

donde

$$\overline{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^{k} n_i \overline{\mathbf{x}}_i$$

Veremos que bajo H_3

- $\mathbf{H} \sim \mathcal{W}(\mathbf{\Sigma}, p, k-1)$
- **H** es independiente de **U**.

Teorema

Sean $\nu = pk + \frac{p(p+1)}{2}k$, $\nu_1 = pk + \frac{p(p+1)}{2}$ y $\nu_2 = p + \frac{p(p+1)}{2}$.

a) El criterio de cociente de verosimilitud para $H_1 : \boldsymbol{\Sigma}_1 = \cdots = \boldsymbol{\Sigma}_k$, se basa en

$$\gamma_1^? = \frac{\prod_{j=1}^{k} \left| \frac{\mathbf{Q}_i}{n_i} \right|^{\frac{n_i}{2}}}{\left| \frac{\mathbf{U}}{n} \right|^{\frac{n}{2}}}$$

Rechazo si $\gamma_1^? < k_{1;}$.

Ademas, bajo H_1, se tiene que $-2\log(\gamma_1^?) \xrightarrow{D} \chi^2_{-\ 1}$

Teorema

b) El criterio de cociente de verosimilitud para
$H_2 : \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_k \,, \boldsymbol{\Sigma}_1 = \cdots = \boldsymbol{\Sigma}_k$

$$\gamma_2^{?} = \frac{\prod_{j=1}^{k} \left| \frac{\mathbf{Q}_i}{n_i} \right|^{\frac{n_i}{2}}}{\left| \frac{\mathbf{U} + \mathbf{H}}{n} \right|^{\frac{n}{2}}}$$

Rechazo si $\gamma_2^{?} < k_{2;}$.

Ademas, bajo H_2, se tiene que $-2\log(\gamma_2^{?}) \xrightarrow{D} \chi^2_{-\ 2}$

Teorema

c) El criterio de cociente de verosimilitud para $H_3 : \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_k$ cuando $\boldsymbol{\Sigma}_1 = \cdots = \boldsymbol{\Sigma}_k$

$$\gamma_3^{?} = \left(\frac{|\mathbf{U}|}{|\mathbf{U} + \mathbf{H}|} \right)^{\frac{n}{2}}$$

Rechazo si $\gamma_3^{?} < k_{3;}$.

Ademas, bajo H_3, se tiene que $-2\log(\gamma_3^{?}) \xrightarrow{D} \chi^2{}_{1-\ \ 2}$. Mas aun, bajo H_3, $\mathbf{U} \sim \mathcal{W}(\boldsymbol{\Sigma}, p, n-k)$, $\mathbf{H} \sim \mathcal{W}(\boldsymbol{\Sigma}, p, k-1)$ independientes de donde

$$\gamma_3^{?\,\frac{2}{n}} = \frac{|\mathbf{U}|}{|\mathbf{U} + \mathbf{H}|} = \quad (n-1, p, k-1)$$

Observemos que rango($\mathbf{U}$) = min($p, n-k$) y rango($\mathbf{H}$) = min($p, k-1$).
Si $k = 2$, rango($\mathbf{H}$) = 1 y obtenemos el test de Hotelling para dos muestras.

Propiedad

Sea $\mathbf{x} \in \mathbb{R}^p$ un vector aleatorio y G una variable aleatoria que indica la pertencia al grupo, tales que para $1 \leq j \leq k$

$$\mathbb{P}(G = j) = \pi_j \qquad \mathbb{E}(\mathbf{x}|G = j) = \boldsymbol{\mu}_j \qquad \mathrm{VAR}(\mathbf{x}|G = j) = \boldsymbol{\Sigma}_j$$

entonces si $\overline{\boldsymbol{\mu}} = \mathbb{E}(\mathbf{x})$ y $\overline{\boldsymbol{\Sigma}} = \mathrm{VAR}(\mathbf{x})$ se cumple que

$$\overline{\boldsymbol{\mu}} = \sum_{j=1}^{k} \pi_j \boldsymbol{\mu}_j \qquad \overline{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}_{\mathrm{W}} + \boldsymbol{\Sigma}_{\mathrm{B}}$$

donde

$$\boldsymbol{\Sigma}_{\mathrm{W}} = \sum_{j=1}^{k} \pi_j \boldsymbol{\Sigma}_j \qquad \boldsymbol{\Sigma}_{\mathrm{B}} = \sum_{j=1}^{k} \pi_j (\boldsymbol{\mu}_j - \overline{\boldsymbol{\mu}})(\boldsymbol{\mu}_j - \overline{\boldsymbol{\mu}})^{\mathrm{T}}$$

Por lo tanto, si $\boldsymbol{\Sigma}_j = \boldsymbol{\Sigma}$ para $1 \leq j \leq k$ tenemos que

- $\boldsymbol{\Sigma}_{\mathrm{W}} = \sum_{j=1}^{k} \pi_j \boldsymbol{\Sigma}_j = \boldsymbol{\Sigma}$ mide la variabilidad dentro de grupos

- $\boldsymbol{\Sigma}_{\mathrm{B}}$ mide la variabilidad entre grupos.

Es decir, descompusimos
la variabilidad total como la variabilidad dentro de grupos más la variabilidad entre grupos.

Observemos que

- $\boldsymbol{\Sigma}_{\mathrm{B}}$ es de nida no{negativa de rango $s \leq \min(k-1, p)$.
- Bajo H_3, $\boldsymbol{\Sigma}_{\mathrm{B}} = \mathbf{0}$

Luego, tenemos que

$$\overline{\boldsymbol{\Sigma}} \geq \boldsymbol{\Sigma}_{\mathrm{W}} \qquad \text{y bajo } H_3 \qquad \overline{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}_{\mathrm{W}}$$

o sea, para realizar un test para H_3 basta comparar cuan distinta es $\overline{\boldsymbol{\Sigma}}$ de $\boldsymbol{\Sigma}_{\mathrm{W}}$.

Por otra parte, si las poblaciones son normales $\widehat{\boldsymbol{\Sigma}}_j = \mathbf{Q}_j / n_j$ es el EMV de $\boldsymbol{\Sigma}_j$ y $\bar{\mathbf{x}}_j$ es el de $\boldsymbol{\mu}_j$. Luego, si $\widehat{\pi}_j = n_j / n$

$$\widehat{\overline{\boldsymbol{\mu}}} = \sum_{j=1}^{k} \widehat{\pi}_j \widehat{\boldsymbol{\mu}}_j \qquad \widehat{\overline{\boldsymbol{\Sigma}}} = \overline{\overline{\boldsymbol{\Sigma}}}$$

Ejemplo

Variables medidas sobre árboles de manzana de 6 injertos. Para cada injerto hay 8 árboles. Las variables son:
x_1 =Diámetro del tronco a los 4 años en unidades de 10cm,
x_2 =Largo a los 4 años,
x_3 =Diámetro del tronco a los 15 años en unidades de 10cm,
x_4 =Peso del árbol a los 15 años, en unidades de 1000 libras.

Inj.	1	1	1	1	1	1	1	1	2	2	2	2
x_1	1.11	1.19	1.09	1.25	1.11	1.08	1.11	1.16	1.05	1.17	1.11	1.25
x_2	2.569	2.928	2.865	3.844	3.027	2.336	3.211	3.037	2.074	2.885	3.378	3.906
x_3	3.58	3.75	3.93	3.94	3.60	3.51	3.98	3.62	4.09	4.06	4.87	4.98
x_4	0.760	0.821	0.928	1.009	0.766	0.726	1.209	0.750	1.036	1.094	1.635	1.517
Inj.	2	2	2	2	3	3	3	3	3	3	3	3
x_1	1.17	1.15	1.17	1.19	1.07	0.99	1.06	1.02	1.15	1.20	1.20	1.17
x_2	2.782	3.018	3.383	3.447	2.505	2.315	2.667	2.390	3.021	3.085	3.308	3.231
x_3	4.38	4.65	4.69	4.40	3.76	4.44	4.38	4.67	4.48	4.78	4.57	4.56
x_4	1.197	1.244	1.495	1.026	0.912	1.398	1.197	1.613	1.476	1.571	1.506	1.458
Inj.	4	4	4	4	4	4	4	4	5	5	5	5
x_1	1.22	1.03	1.14	1.01	0.99	1.11	1.20	1.08	0.91	1.15	1.14	1.05
x_2	2.838	2.351	3.001	2.439	2.199	3.318	3.601	3.291	1.532	2.552	3.083	2.330
x_3	3.89	4.05	4.05	3.92	3.27	3.95	4.27	3.85	4.04	4.16	4.79	4.42
x_4	0.944	1.241	1.023	1.067	0.693	1.085	1.242	1.017	1.084	1.151	1.381	1.242
Inj.	5	5	5	5	6	6	6	6	6	6	6	6
x_1	0.99	1.22	1.05	1.13	1.11	0.75	1.05	1.02	1.05	1.07	1.13	1.11
x_2	2.079	3.366	2.416	3.100	2.813	0.840	2.199	2.132	1.949	2.251	3.064	2.469
x_3	3.47	4.41	4.64	4.57	3.76	3.14	3.75	3.99	3.34	3.21	3.63	3.95
x_4	0.673	1.137	1.455	1.325	0.800	0.606	0.790	0.853	0.610	0.562	0.707	0.952

Ejemplo

Se desea estudiar si las medias de los distintos injertos son iguales. Nosotros consideraremos solamente los Injertos 1, 2 y 3.

Primero estudiaremos si las poblaciones correspondientes a los 3 injertos tienen iguales matrices de covarianza, o sea, testearemos

$$H_1 : \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_3$$

A continuacion se dan las matrices de covarianza estimadas

$$\mathbf{S}_1 = \begin{pmatrix} 0.0034 & 0.0203 & 0.0037 & 0.0018 \\ 0.0203 & 0.2007 & 0.0580 & 0.0458 \\ 0.0037 & 0.0580 & 0.0352 & 0.0285 \\ 0.0018 & 0.0458 & 0.0285 & 0.0283 \end{pmatrix} \quad \mathbf{S}_2 = \begin{pmatrix} 0.0034 & 0.0258 & 0.0088 & 0.0032 \\ 0.0258 & 0.3048 & 0.1498 & 0.0832 \\ 0.0088 & 0.1498 & 0.1157 & 0.0711 \\ 0.0032 & 0.0832 & 0.0711 & 0.0565 \end{pmatrix}$$

$$\mathbf{S}_3 = \begin{pmatrix} 0.0068 & 0.0314 & 0.0087 & 0.0060 \\ 0.0314 & 0.1543 & 0.0480 & 0.0329 \\ 0.0087 & 0.0480 & 0.0951 & 0.0680 \\ 0.0060 & 0.0329 & 0.0680 & 0.0534 \end{pmatrix}$$

con lo cual $-2\log(\gamma_1^?) = 25.80706$ y $\chi^2_{0:99;20} = 37.56623$.

No rechazamos H_1 y el $p-$valor es 0.1723.

Ejemplo

Supongamos entonces que las poblaciones correspondientes a los injertos 1, 2 y 3 tienen la misma matriz de covarianza y estudiemos si las medias son iguales, o sea, queremos testear H_3.

El estad stico del test es

$$V = \gamma_3^{?\frac{2}{n}} = \frac{|\mathbf{U}|}{|\mathbf{U}+\mathbf{H}|} = \quad (n-1, p, k-1)$$

En nuestro caso, $p = 4$, $k = 3$ y $n = 24$. Hemos visto que

$$\frac{1 - \quad (N, p, 2)^{\frac{1}{2}}}{(N, p, 2)^{\frac{1}{2}}} \frac{N-p-1}{p} \sim \mathcal{F}_{2p;2(N-p-1)}$$

Luego, rechazaremos H_3 si

$$\frac{1 - V^{\frac{1}{2}}}{V^{\frac{1}{2}}} \frac{(n-1)-p-1}{p} > f_{2p;2(n-p-2)}(\alpha)$$

Ejemplo

En el ejemplo que nos interesa $V = 0.1447022$, $n - p - 2 = 18$ luego $f_{8;36}(0.01) = 3.051726$ y

$$\frac{1 - V^{\frac{1}{2}}}{V^{\frac{1}{2}}} = 7.329734$$

con lo que rechazamos H_3 y el $p-$valor es 10^{-5}.