Inferencia para la normal multivariada

Graciela Boente

Propiedad

Sean $\mathbf{x}_1, \ldots, \mathbf{x}_n$ i.i.d., $\mathbf{x}_i \in \mathbb{R}^p$, $\mathbf{x}_i \sim N_p(\boldsymbol{\mu}, \mathbf{\Sigma})$, $\mathbf{\Sigma} > 0$

- La familia $N_p(\boldsymbol{\mu}, \mathbf{\Sigma})$ es una familia exponencial.
- $\mathbf{Q} = \sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^{\mathrm{T}}$ y $\bar{\mathbf{x}}$ son estadísticos suficientes y completos.
- Por lo tanto, cualquier estimador insesgado basado en $\mathbf{Q}$ y $\bar{\mathbf{x}}$ resulta IMVU.

Propiedad

- Los estimadores de máxima verosimilitud de $\boldsymbol{\mu}$ y $\mathbf{\Sigma}$ son $\widehat{\boldsymbol{\mu}} = \overline{\mathbf{x}}$ y $\widehat{\mathbf{\Sigma}} = \mathbf{Q}/n$. Además si

$$L(\boldsymbol{\mu}, \mathbf{\Sigma}) = \prod_{j=1}^{n} f_{\mathbf{x}}(\mathbf{x}_i, \boldsymbol{\mu}, \mathbf{\Sigma})$$

tenemos que

$$L(\widehat{\boldsymbol{\mu}}, \widehat{\mathbf{\Sigma}}) = (2\pi)^{-\frac{np}{2}} \left(\det(\widehat{\mathbf{\Sigma}})\right)^{-\frac{n}{2}} e^{-\frac{np}{2}}$$

- $\mathbb{E}(\overline{\mathbf{x}}) = \boldsymbol{\mu}$

$$\mathbb{E}(\widehat{\mathbf{\Sigma}}) = \frac{n-1}{n}\mathbf{\Sigma}$$

luego el estimador insesgado de $\mathbf{\Sigma}$ es

$$\mathbf{S} = \frac{\mathbf{Q}}{n-1}$$

Distribución de los estimadores

Sean $\mathbf{x}_1, \ldots, \mathbf{x}_n$ i.i.d., $\mathbf{x}_i \in \mathbb{R}^p$, $\mathbf{x}_i \sim N_p(\boldsymbol{\mu}, \mathbf{\Sigma})$, $\mathbf{\Sigma} > 0$ y $n \geq p + 1$

a) $\bar{\mathbf{x}} \sim N_p(\boldsymbol{\mu}, (1/n)\mathbf{\Sigma})$,
$\mathbf{Q} = \sum_{i=1}^{n} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^{\mathrm{T}} \sim \mathcal{W}(\mathbf{\Sigma}, p, n-1)$.

b) $\bar{\mathbf{x}}$ y $\mathbf{Q}$ son independientes.

c)

$$T^2 = n\,(n-1)\,(\bar{\mathbf{x}} - \boldsymbol{\mu})^{\mathrm{T}} \mathbf{Q}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}) = n\,(\bar{\mathbf{x}} - \boldsymbol{\mu})^{\mathrm{T}} \mathbf{S}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}) \sim T^2_{p,n-1}$$

o sea,

$$\frac{n-p}{p} \frac{T^2}{n-1} \sim \mathcal{F}_{p,n-p}$$

Una muestra

Sean $\mathbf{x}_1, \ldots, \mathbf{x}_n$ i.i.d., $\mathbf{x}_i \in \mathbb{R}^p$, $\mathbf{x}_i \sim N_p(\boldsymbol{\mu}, \mathbf{\Sigma})$, $\mathbf{\Sigma} > 0$

Un test de nivel α para

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

está dado por

$$\Phi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad n\,(\overline{\mathbf{x}} - \boldsymbol{\mu}_0)^{\mathrm{T}} \mathbf{\Sigma}^{-1} (\overline{\mathbf{x}} - \boldsymbol{\mu}_0) > \chi^2_{p,\alpha} \\ 0 & \text{si} \quad n\,(\overline{\mathbf{x}} - \boldsymbol{\mu}_0)^{\mathrm{T}} \mathbf{\Sigma}^{-1} (\overline{\mathbf{x}} - \boldsymbol{\mu}_0) \leq \chi^2_{p,\alpha} \end{cases}$$

Dos muestras

Sean $\boldsymbol{\Sigma} > 0$

- $\mathbf{x}_{1,1}, \ldots, \mathbf{x}_{1,n_1}$ i.i.d., $\mathbf{x}_{1,j} \in \mathbb{R}^p$, $\mathbf{x}_{1,j} \sim N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$
- $\mathbf{x}_{2,1}, \ldots, \mathbf{x}_{2,n_2}$ i.i.d., $\mathbf{x}_{2,j} \in \mathbb{R}^p$, $\mathbf{x}_{2,j} \sim N_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$
- Supongamos que ambas muestras son independientes entre sí

Definamos

$$\overline{\mathbf{x}}_1 = \frac{1}{n_1} \sum_{j=1}^{n_1} \mathbf{x}_{1,j} \qquad \overline{\mathbf{x}}_2 = \frac{1}{n_2} \sum_{j=1}^{n_2} \mathbf{x}_{2,j}$$

Un test de nivel α para

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \qquad H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$$

está dado por

$$\Phi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad \frac{n_1\, n_2}{n_1+n_2} (\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2)^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2) > \chi^2_{p,\alpha} \\ 0 & \text{si} \quad \frac{n_1\, n_2}{n_1+n_2} (\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2)^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2) \leq \chi^2_{p,\alpha} \end{cases}$$

Teorema

Sean $\mathbf{x}_1, \ldots, \mathbf{x}_n$ i.i.d., $\mathbf{x}_i \in \mathbb{R}^p$, $\mathbf{x}_i \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > 0$

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

a) Sea $\lambda^2 = n(\boldsymbol{\mu} - \boldsymbol{\mu}_0)^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \boldsymbol{\mu}_0)$ luego

$$\frac{n-p}{p} \frac{T_0^2}{n-1} \sim \mathcal{F}_{p,n-p}(\lambda^2)$$

b) Una región de confianza para $\boldsymbol{\mu}$ de nivel $1 - \alpha$ está dada por

$$\left\{ \boldsymbol{\mu} : n \frac{n-p}{p(n-1)} (\bar{\mathbf{x}} - \boldsymbol{\mu})^{\mathrm{T}} \mathbf{S}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}) \leq f_{p,n-p}(\alpha) \right\}$$

donde $f_{p,m}(\alpha)$ es tal que $\mathbb{P}(\mathcal{F}_{p,m} > f_{p,m}(\alpha)) = \alpha$.

c) El test de máxima verosimilitud para

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

está dado por

$$\Phi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad T_0^2 > t^2_{p,n-1}(\alpha) \\ 0 & \text{si} \quad T_0^2 \leq t^2_{p,n-1}(\alpha) \end{cases}$$

con

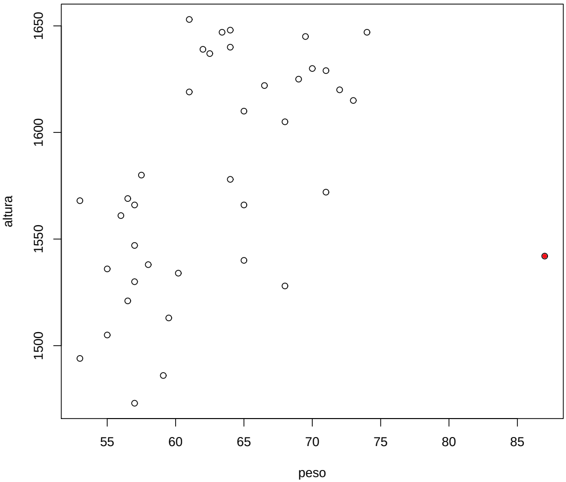
$$t^2_{p,n-1}(\alpha) = \frac{p(n-1)}{n-p}\, f_{p,n-p}(\alpha)$$

Ejemplo

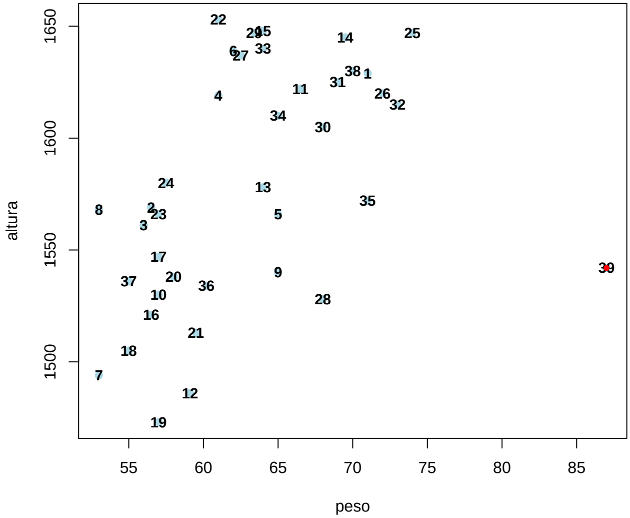
Peso (kg) y Altura (mm) de 39 Indios Peruanos

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

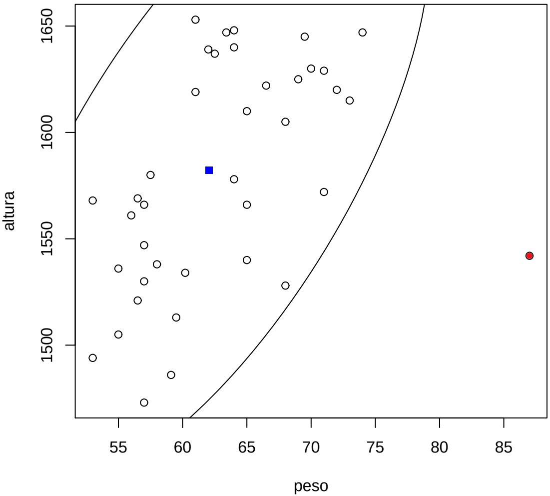


Ejemplo


altura
peso
1500
1550
1600
1650
55
60
65
70
75
80
85
22
29
15
25
14
6
27
33
38
1
31
11
26
32
4
34
30
24
13
35
8
2
23
5
3
17
39
9
37
20
36
10
28
16
21
18
7
12
19

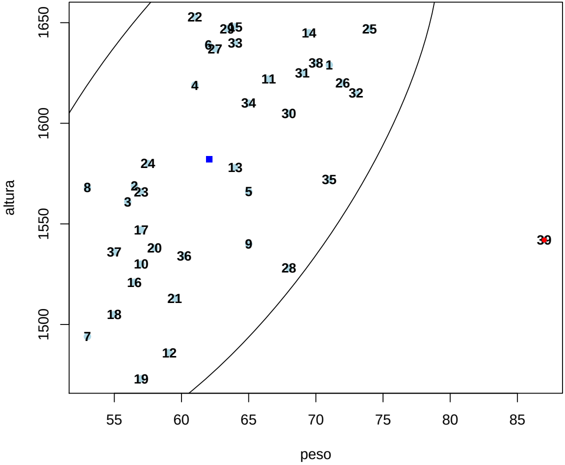
Elipse obtenida mediante un $S-$estimador $(\mathbf{m}, \mathbf{V})$

Para detectar datos atípicos construímos la elipse
$\{\mathbf{u} \in \mathbb{R}^p : (\mathbf{u} - \mathbf{m})^{\mathrm{T}} \mathbf{V}^{-1} (\mathbf{u} - \mathbf{m}) \leq \chi^2_{p,0.05}\}$



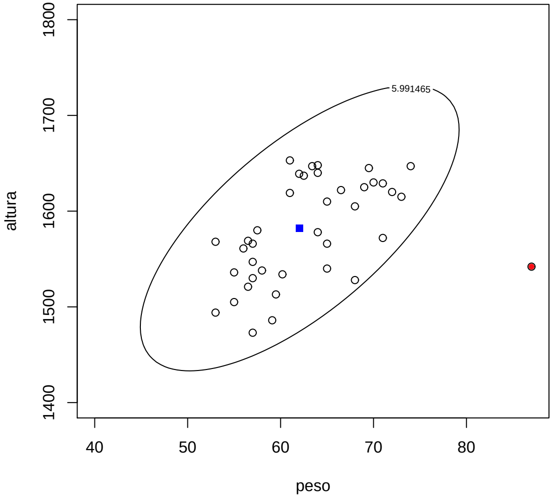
Elipse obtenida mediante un $S-$estimador $(\mathbf{m}, \mathbf{V})$

Para detectar datos atípicos construímos la elipse
$\{\mathbf{u} \in \mathbb{R}^p : (\mathbf{u}-\mathbf{m})^{\mathrm{T}}\mathbf{V}^{-1}(\mathbf{u}-\mathbf{m}) \leq \chi^2_{p,0.05}\}$



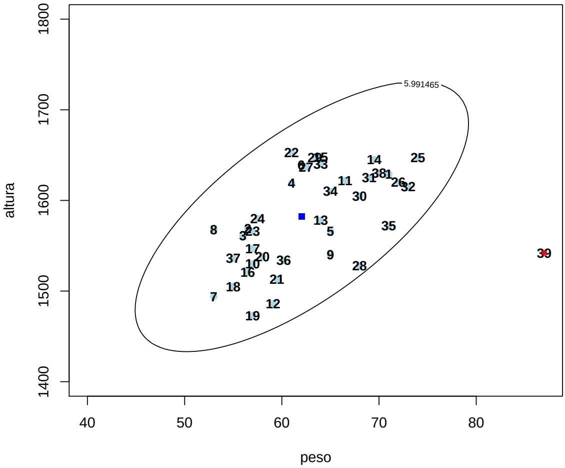
Elipse obtenida mediante un $S-$estimador $(\mathbf{m}, \mathbf{V})$

Para detectar datos atípicos construímos la elipse
$\{\mathbf{u} \in \mathbb{R}^p : (\mathbf{u}-\mathbf{m})^{\mathrm{T}}\mathbf{V}^{-1}(\mathbf{u}-\mathbf{m}) \leq \chi^2_{p,0.05}\}$

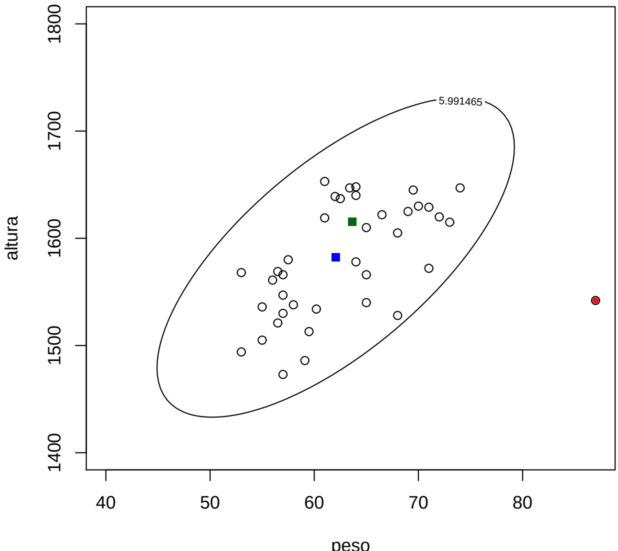


Elipse obtenida mediante un $S-$estimador $(\mathbf{m}, \mathbf{V})$

Para detectar datos atípicos construímos la elipse
$\{\mathbf{u} \in \mathbb{R}^p : (\mathbf{u} - \mathbf{m})^{\mathrm{T}} \mathbf{V}^{-1} (\mathbf{u} - \mathbf{m}) \leq \chi^2_{p,0.05}\}$

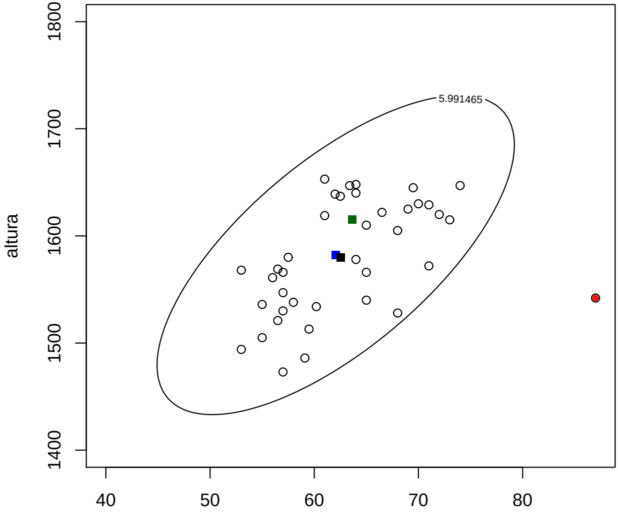


En verde $\mu_0 = (63.64, 1615.38)^{\mathrm{T}}$, En rojo el dato atípico, En azul el centro robusto. Se grafica la elipse de detección $\{\mathbf{u} \in \mathbb{R}^p : (\mathbf{u} - \mathbf{m})^{\mathrm{T}} \mathbf{V}^{-1} (\mathbf{u} - \mathbf{m}) \leq \chi^2_{p,0.05}\}$



En verde $\mu_0 = (63.64, 1615.38)^{\mathrm{T}}$, En rojo el dato atípico, En azul el centro robusto

En negro el promedio de todos los datos menos la observación $(87, 1542)^{\mathrm{T}}$. Se grafica la elipse de detección $\{\mathbf{u} \in \mathbb{R}^p : (\mathbf{u} - \mathbf{m})^{\mathrm{T}} \mathbf{V}^{-1} (\mathbf{u} - \mathbf{m}) \leq \chi^2_{p,0.05}\}$



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 observación $(87, 1542)^{\mathrm{T}}$ que es un dato atípico.

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 $\alpha = 0.001$.

Más aún, el $p-$valor es 0.00038.

Programa

```
x = cbind(peso,altura)
rownames(x)=1:length(peso)
plot(peso,altura, col="lightblue", pch=19)
text(peso, altura, labels=rownames(x), cex=0.9, font=2)
points(peso[39],altura[39],col="red",pch=20)
y = x[-39,]
oy = apply(y,2,mean)
wsigma = cov(y)
mu0 = cbind(63.64,1615.38)
n = dim(y)[1]
p = dim(y)[2]
T0 = n*(oy-mu0)%*% solve(wsigma)%*%t(oy-mu0)
F0 = T0*(n-p)/(p*(n-1))
falpha = qf(0.001, p, n-p, ncp=0, lower.tail = FALSE, log.p = FALSE)
pvalor = pf(F0, p, n-p, ncp=0, lower.tail = FALSE, log.p = FALSE)
```

Potencia del test

Hemos visto que

$$\frac{n-p}{p}\frac{T_0^2}{n-1} \sim \mathcal{F}_{p,n-p}(\lambda^2)$$

con $\lambda^2 = n(\boldsymbol{\mu}-\boldsymbol{\mu}_0)^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\boldsymbol{\mu}-\boldsymbol{\mu}_0)$.

Luego, si $\boldsymbol{\mu} \neq \mu_0$ es una alternativa fija, la potencia del test Φ e está dada por

$$\mathbb{P}\left(\frac{n-p}{p}\frac{T_0^2}{n-1} \geq f_{p,n-p}(\alpha)\right) = \mathbb{P}\left(\mathcal{F}_{p,n-p}(\lambda^2) \geq f_{p,n-p}(\alpha)\right)$$

que es una función de p, $n-p$ y de $\delta = \left(\lambda^2/(p+1)\right)^{1/2}$.

Propiedades del test

El test de Hotelling es

- invariante por transformaciones lineales no singulares
- UMP de nivel α entre los invariantes por transformaciones lineales.
- UMP de nivel α entre los tests cuya potencia depende de $(\boldsymbol{\mu} - \boldsymbol{\mu}_0)^{\mathrm{T}} \mathbf{\Sigma}^{-1} (\boldsymbol{\mu} - \boldsymbol{\mu}_0)$
- admisible

Propiedades del test

- El test de Hotelling es el test derivado del principio de unión intersección

$$T_0^2 = \sup_{\mathbf{a}\neq 0} t_0^2(\mathbf{a}) \quad \text{con} \quad t_0(\mathbf{a}) = \sqrt{n}\frac{\mathbf{a}^{\mathrm{T}}(\overline{\mathbf{x}} - \boldsymbol{\mu}_0)}{(\mathbf{a}^{\mathrm{T}}\mathbf{S}\mathbf{a})^{\frac{1}{2}}}$$

- Esto permite obtener intervalos simultáneos. Si llamamos

$$t_{p,N}^2(\alpha) = \frac{p\,N}{N-p+1}\,f_{p,N-p+1}(\alpha)$$

tenemos que $\forall\ (\boldsymbol{\mu}, \boldsymbol{\Sigma})$

$$\mathbb{P}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}\left(\mathbf{a}^{\mathrm{T}}\overline{\mathbf{x}} - \frac{(\mathbf{a}^{\mathrm{T}}\mathbf{S}\mathbf{a})^{\frac{1}{2}}}{\sqrt{n}}\,t_{p,n-1}(\alpha) \leq \mathbf{a}^{\mathrm{T}}\boldsymbol{\mu} \leq \mathbf{a}^{\mathrm{T}}\overline{\mathbf{x}} + \frac{(\mathbf{a}^{\mathrm{T}}\mathbf{S}\mathbf{a})^{\frac{1}{2}}}{\sqrt{n}}\,t_{p,n-1}(\alpha)\right) = 1-\alpha$$

lo que provee un método análogo al de Scheffe para contrastes.

Dos muestras

- $\mathbf{x}_{1,1}, \ldots, \mathbf{x}_{1,n_1}$ i.i.d., $\mathbf{x}_{1,j} \in \mathbb{R}^p$, $\mathbf{x}_{1,j} \sim N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$
- $\mathbf{x}_{2,1}, \ldots, \mathbf{x}_{2,n_2}$ i.i.d., $\mathbf{x}_{2,j} \in \mathbb{R}^p$, $\mathbf{x}_{2,j} \sim N_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$
- Supongamos que ambas muestras son independientes entre sí

Definamos

$$\overline{\mathbf{x}}_1 = \frac{1}{n_1}\sum_{j=1}^{n_1} \mathbf{x}_{1,j} \qquad \overline{\mathbf{x}}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2} \mathbf{x}_{2,j}$$

Queremos testear

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \qquad H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$$

cuando $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma} > 0$ desconocida.

Dos muestras

- $\mathbf{x}_{1,1}, \ldots, \mathbf{x}_{1,n_1}$ i.i.d., $\mathbf{x}_{1,j} \in \mathbb{R}^p$, $\mathbf{x}_{1,j} \sim N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma})$
- $\mathbf{x}_{2,1}, \ldots, \mathbf{x}_{2,n_2}$ i.i.d., $\mathbf{x}_{2,j} \in \mathbb{R}^p$, $\mathbf{x}_{2,j} \sim N_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma})$
- Supongamos que ambas muestras son independientes entre sí

Definamos

$$\mathbf{S}_i = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (\mathbf{x}_{i,j} - \overline{\mathbf{x}}_i)(\mathbf{x}_{i,j} - \overline{\mathbf{x}}_i)^{\mathrm{T}} \qquad i = 1, 2$$

y sea

$$\mathbf{S} = \frac{(n_1 - 1)\mathbf{S}_1 + (n_2 - 1)\mathbf{S}_2}{n_1 + n_2 - 2}$$

Dos muestras: Teorema

a) Sea

$$T_0^2 = \frac{n_1\, n_2}{n_1 + n_2} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^{\mathrm{T}} \mathbf{S}^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)$$

entonces

$$\frac{n_1 + n_2 - p - 1}{(n_1 + n_2 - 2)p}\, T_0^2 \sim \mathcal{F}_{p, n_1+n_2-p-1}(\lambda^2)$$

con $\lambda^2 = \frac{n_1\, n_2}{n_1+n_2} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$.

b) Un test para $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \qquad H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$ está dado por

$$\Phi(\mathbf{X}) = \begin{cases} 1 & \text{si} \quad \frac{n_1\, n_2}{n_1+n_2} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^{\mathrm{T}} \mathbf{S}^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) > t^2_{p, n_1+n_2-2}(\alpha) \\ 0 & \text{si} \quad \frac{n_1\, n_2}{n_1+n_2} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^{\mathrm{T}} \mathbf{S}^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \leq t^2_{p, n_1+n_2-2}(\alpha) \end{cases}$$

con $t^2_{p,N}(\alpha) = \frac{p\, N}{N-p+1}\, f_{p, N-p+1}(\alpha)$

Dos muestras: Propiedades

El test de Hotelling es

- el test de cociente de verosimilitud
- el test derivado del principio de unión intersección

$$T_0^2 = \sup_{\mathbf{a} \neq 0} t_0^2(\mathbf{a}) \quad \text{con} \quad t_0(\mathbf{a}) = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \, \frac{\mathbf{a}^{\mathrm{T}}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2)}{(\mathbf{a}^{\mathrm{T}} \mathbf{S} \mathbf{a})^{\frac{1}{2}}}$$

El supremo se alcanza en

$$\widehat{\mathbf{a}} = \mathbf{S}^{-1}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2)$$

A la función $H(\mathbf{u}) = \widehat{\mathbf{a}}^{\mathrm{T}} \mathbf{u}$ se la llama función discriminante lineal.

Teorema

- Una región de confianza de nivel $1-\alpha$ para $\boldsymbol{\theta} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$ está dada por

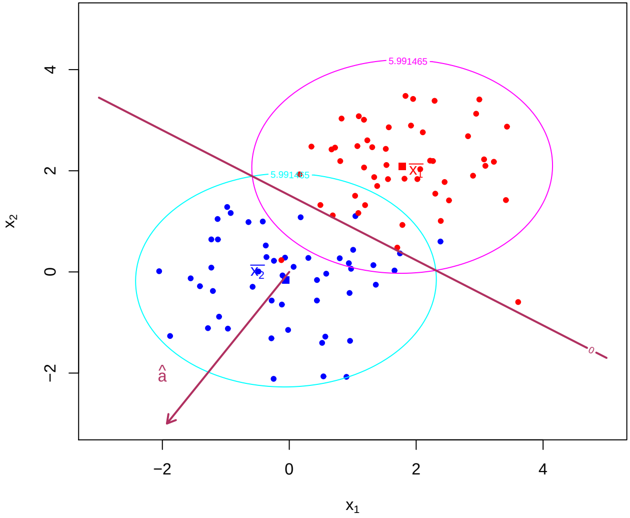
$$\left\{ \boldsymbol{\theta} : \quad \frac{n_1\, n_2}{n_1 + n_2} (\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2 - \boldsymbol{\theta})^{\mathrm{T}} \mathbf{S}^{-1} (\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2 - \boldsymbol{\theta}) \leq t^2_{p, n_1+n_2-2}(\alpha) \right\}$$

donde

$$t^2_{p,N}(\alpha) = \frac{p\,N}{N-p+1}\, f_{p,N-p+1}(\alpha)$$

- Un conjunto de intervalos simultáneos es

$$\mathbf{a}^{\mathrm{T}}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2) \pm \left(\left(\frac{1}{n_1} + \frac{1}{n_2} \right) \mathbf{a}^{\mathrm{T}} \mathbf{S} \mathbf{a} \right)^{\frac{1}{2}} \; t_{p, n_1+n_2-2}(\alpha)$$

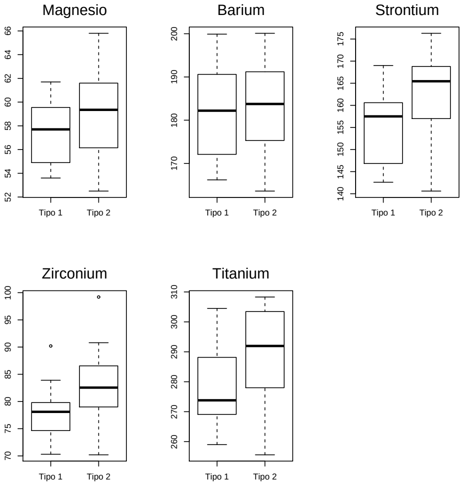

5.991465
5.991465
$\overline{x_1}$
$\overline{x_2}$
$\hat{a}$
0
−2
0
2
4
x_1
x_2

Ejemplo

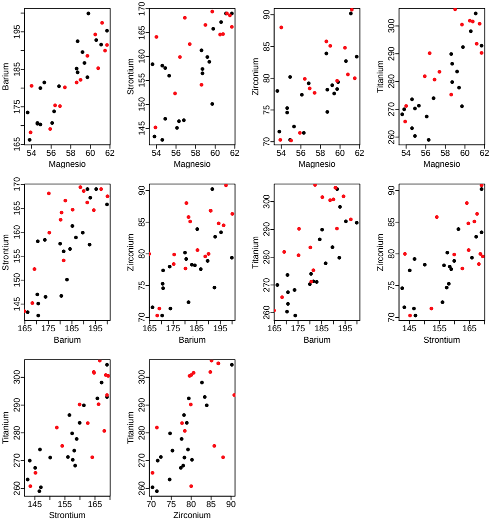
Concentración de cinco elementos (Manganese, Barium, Strontium, Zirconium, and Titanium) en muestras de vidrio tomadas de botellas de seis tipos de cerveza Heineken.

		Tipo 1					**Tipo 2**		
Mg	Ba	Sr	Zr	Ti	Mg	Ba	Sr	Zr	Ti
56.1	170.7	145.1	77.4	267.4	58.6	181.5	154.1	85.8	275.3
53.8	166.2	143.3	71.6	270.0	52.5	164.8	143.5	80.0	260.8
58.7	184.2	156.5	78.2	286.4	61.2	197.4	169.0	90.8	293.6
54.6	170.5	158.1	75.3	273.6	54.0	180.6	164.1	88.0	271.2
58.6	185.2	161.3	83.9	289.9	61.7	191.5	166.2	99.2	290.3
56.8	180.5	146.7	79.2	274.0	59.0	182.2	166.6	85.1	306.0
54.6	170.7	142.6	74.6	263.2	53.9	168.2	145.2	70.3	265.6
59.2	189.6	159.9	78.9	283.6	61.5	190.0	168.6	80.0	300.8
54.9	180.0	157.6	80.2	270.3	57.4	180.2	162.6	77.7	283.5
61.1	191.6	169.0	90.2	304.5	60.5	194.3	164.6	84.8	301.9
59.4	186.7	158.9	77.6	277.8	65.8	200.1	167.5	86.3	307.0
54.9	170.3	147.0	70.3	260.4	55.0	163.6	140.6	70.2	255.6
58.7	192.5	157.4	74.7	279.8	60.8	185.3	164.7	80.6	301.6
55.3	181.5	156.0	72.4	271.3	56.4	175.4	159.9	79.9	290.2
61.7	195.3	169.0	83.4	292.9	63.4	190.9	176.3	86.8	305.0
59.7	182.9	150.1	78.3	271.1	65.4	196.4	174.0	84.5	308.3
56.3	173.8	146.5	71.4	259.0	55.9	169.1	152.3	71.4	281.9
59.8	199.9	165.8	79.4	292.4	59.7	188.6	169.4	79.6	300.5
53.6	173.5	158.4	78.0	268.2	56.9	175.2	168.1	78.4	280.7
60.6	192.8	167.2	82.7	298.1	63.4	190.9	176.3	86.8	305.0

Ejemplo: Boxplots

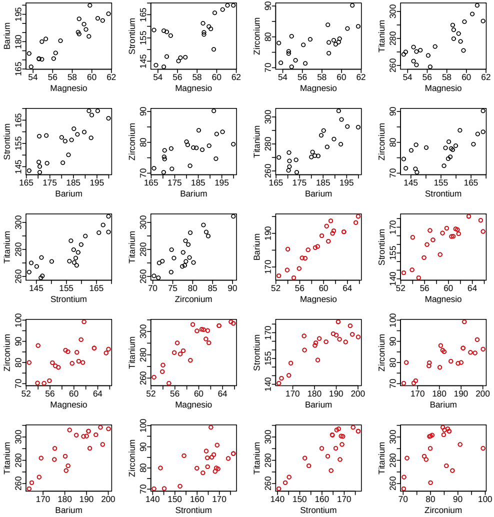


Ejemplo



Barium
Strontium
Zirconium
Titanium
Magnesio
Barium
Strontium
Zirconium

Ejemplo


Barium
Strontium
Zirconium
Titanium
Magnesio
Barium
Strontium
Zirconium

Ejemplo

Queremos testear

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \qquad\qquad H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$$

$$\overline{\mathbf{x}}_1 = \begin{pmatrix} 57.860 \\ 182.892 \\ 155.467 \\ 77.845 \\ 277.396 \end{pmatrix} \qquad \overline{\mathbf{x}}_2 = \begin{pmatrix} 59.554 \\ 183.832 \\ 163.299 \\ 81.4887 \\ 291.625 \end{pmatrix} \qquad \overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2 = \begin{pmatrix} -1.730 \\ -1.390 \\ -6.860 \\ -4.425 \\ -11.545 \end{pmatrix}$$

$$\mathbf{S} = \begin{pmatrix} 10.825 & 30.636 & 23.410 & 11.915 & 41.361 \\ & 109.865 & 82.385 & 43.337 & 129.349 \\ & & 90.452 & 37.707 & 121.731 \\ & & & 36.826 & 53.564 \\ & & & & 219.991 \end{pmatrix}$$

Ejemplo

Queremos testear

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \qquad\qquad H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$$

Obtenemos

$$T_0^2 = \frac{n_1\, n_2}{n_1 + n_2}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2)^{\mathrm{T}}\mathbf{S}^{-1}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2) = 25.1697$$

y

$$F_0 = \frac{n_1 + n_2 - p - 1}{(n_1 + n_2 - 2)p}\, T_0^2 = 4.504052$$

Como $f_{p,n_1+n_2-p-1}(0.01) = f_{5,34}(0.01) = 3.610562$ rechazo con nivel $\alpha = 0.01$.

Más aún, el $p-$valor es 0.002949.

Ejemplo

Un conjunto de intervalos simultáneos de nivel $1-\alpha$ con $\alpha = 0.01$

$$\begin{aligned} & \mathbf{a}^{\mathrm{T}}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2) \pm \left(\left(\frac{1}{20}+\frac{1}{20}\right)\mathbf{a}^{\mathrm{T}}\mathbf{S}\mathbf{a}\right)^{\frac{1}{2}} t_{5,38}(\alpha) \\ = \quad & \mathbf{a}^{\mathrm{T}}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2) \pm \left(\mathbf{a}^{\mathrm{T}}\mathbf{S}\mathbf{a}\right)^{\frac{1}{2}} \sqrt{2.017667} = \mathbf{a}^{\mathrm{T}}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2) \pm 1.4204 \left(\mathbf{a}^{\mathrm{T}}\mathbf{S}\mathbf{a}\right)^{\frac{1}{2}} \end{aligned}$$

$\mathbf{a}$	$\mathbf{a}^{\mathrm{T}}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2)$	$\mathbf{a}^{\mathrm{T}}\mathbf{S}\mathbf{a}$	$(\mathbf{a}^{\mathrm{T}}\mathbf{S}\mathbf{a})^{1/2}$	Intervalo
$\mathbf{e}_1 = (1,0,0,0,0)$	-1.730	10.825	3.290	(-6.404, 2.944)
$\mathbf{e}_2 = (0,1,0,0,0)$	-1.390	109.865	10.482	(-16.279, 13.499)
$\mathbf{e}_3 = (0,0,1,0,0)$	-6.860	90.452	9.511	(-20.369, 6.649)
$\mathbf{e}_4 = (0,0,0,1,0)$	-4.425	36.826	6.068	(-13.045, 4.195)
$\mathbf{e}_5 = (0,0,0,0,1)$	-11.545	219.991	14.832	(-32.613, 9.523)
$(1,-1,0,0,0)/\sqrt{2}$	-0.240	29.709	5.451	(-7.983, 7.502)
$\widehat{\mathbf{a}}$	2.517	2.517	1.586	(0.263, 4.771)

Si usamos Bonferroni los intervalos de nivel por lo menos $1-\alpha$ son $\mathbf{a}^{\mathrm{T}}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2) \pm \left(\mathbf{a}^{\mathrm{T}}\mathbf{S}\mathbf{a}\right)^{\frac{1}{2}} \xi$ donde $\xi = 1.0496$ si consideramos solamente los vectores $\mathbf{a} = \mathbf{e}_1, \ldots, \mathbf{e}_5$ y $\xi = 1.0877$ para los 7 vectores dados.

Efecto de no cumplimiento de la hipótesis $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$

Supongamos que n_1 y n_2 son suficientemente grandes como para que $\mathbf{S}_j \simeq \mathbf{\Sigma}_j$, entonces

$$
\begin{aligned}
T_0^2 &= \frac{n_1\, n_2}{n_1+n_2}(\overline{\mathbf{x}}_1-\overline{\mathbf{x}}_2)^{\mathrm{T}}\left(\frac{(n_1-1)\mathbf{S}_1+(n_2-1)\mathbf{S}_2}{n_1+n_2-2}\right)^{-1}(\overline{\mathbf{x}}_1-\overline{\mathbf{x}}_2) \\
&\simeq (\overline{\mathbf{x}}_1-\overline{\mathbf{x}}_2)^{\mathrm{T}}\left(\frac{1}{n_2}\mathbf{\Sigma}_1+\frac{1}{n_1}\mathbf{\Sigma}_2\right)^{-1}(\overline{\mathbf{x}}_1-\overline{\mathbf{x}}_2) = D_0^2
\end{aligned}
$$

Pero $(\overline{\mathbf{x}}_1-\overline{\mathbf{x}}_2) \sim N(\boldsymbol{\theta}, (1/n_1)\mathbf{\Sigma}_1 + (1/n_2)\mathbf{\Sigma}_2)$, luego

$$
D_1^2 = (\overline{\mathbf{x}}_1-\overline{\mathbf{x}}_2)^{\mathrm{T}}\left(\frac{1}{n_1}\mathbf{\Sigma}_1+\frac{1}{n_2}\mathbf{\Sigma}_2\right)^{-1}(\overline{\mathbf{x}}_1-\overline{\mathbf{x}}_2) \sim \chi_p^2(\lambda^2)
$$

con

$$
\lambda^2 = \boldsymbol{\theta}^{\mathrm{T}}\left(\frac{1}{n_1}\mathbf{\Sigma}_1+\frac{1}{n_2}\mathbf{\Sigma}_2\right)^{-1}\boldsymbol{\theta} \quad \text{con} \quad \boldsymbol{\theta} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2\,.
$$

Luego, bajo H_0, $\boldsymbol{\theta} = 0$, o sea, $D_1^2 \sim \chi_p^2$.

Efecto de no cumplimiento de la hipótesis $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$

- Si $n_1 = n_2$, $D_0^2 = D_1^2$, o sea, en este caso la distribución de T_0^2 será poco sensible al alejamiento del supuesto $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$
- Si $n_1 \simeq n_2$, n_1 y n_2 son grandes, hay poco efecto ver Ito and Schull (1964, Biometrika)
- Si n_1 y n_2 son muy diferentes, la diferencia entre $\boldsymbol{\Sigma}_1$ y $\boldsymbol{\Sigma}_2$ tiene un efecto importante en nivel y potencia.

Problema de Fisher–Behrens: Caso $n_1 = n_2 = n$

Definamos $\mathbf{u}_i = \mathbf{x}_{1,i} - \mathbf{x}_{2,i}$, $1 \leq i \leq n$,

$$\mathbf{u}_i \sim N(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\,.$$

Para testear $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 \qquad H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$, definamos

$$T^2_{0,\mathbf{u}} = n\,\overline{\mathbf{u}}^{\mathrm{T}} \mathbf{S}_{\mathbf{u}}^{-1} \overline{\mathbf{u}}$$

Luego, bajo H_0, $T^2_{0,\mathbf{u}} \sim T^2_{p,n-1}$, por lo tanto,

$$\frac{(n-p)\,T^2_{0,\mathbf{u}}}{p\,(n-1)} \sim \mathcal{F}_{p,n-p}$$

por lo tanto, rechazo H_0 si

$$T^2_{0,\mathbf{u}} > \frac{(n-1)\,p}{n-p}\, f_{p,n-p}(\alpha)\,.$$

Problema de Fisher–Behrens: Caso $n_1 = n_2 = n$

Problema: Este test tiene pérdida de potencia si $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$ pues está basado en n observaciones, obteniendo un Hotelling con $n-1$ grados de libertad,

$$T^2_{0,\mathbf{u}} \sim T^2_{p,n-1}$$

mientras que

$$T_0^2 = \frac{n_1\, n_2}{n_1 + n_2}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2)^{\mathrm{T}}\mathbf{S}^{-1}(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2) \sim T^2_{p,2(n-1)}$$

tendrá $2(n-1)$ grados de libertad.

Problema de Fisher–Behrens: Solución Bennet (1951)

Supongamos $n_1 \leq n_2$. Definamos

$$\mathbf{u}_i = \mathbf{x}_{1,i} - \left(\frac{n_1}{n_2}\right)^{1/2} \mathbf{x}_{2,i}\,, \quad 1 \leq i \leq n_1$$

Sean

$$\mathbf{S}_{\mathbf{u}} = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (\mathbf{u}_i - \overline{\mathbf{u}})(\mathbf{u}_i - \overline{\mathbf{u}})^{\mathrm{T}}$$

$$T_0^2 = n_1 \left(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2\right)^{\mathrm{T}} \mathbf{S}_{\mathbf{u}}^{-1} \left(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2\right)$$

Entonces, bajo $H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$,

$$T_0^2 \sim T^2_{p, n_1 - 1}\,.$$

Luego, el test rechaza H_0 si

$$T_0^2 > \frac{(n_1 - 1)\, p}{n_1 - p}\, f_{p, n_1 - p}(\alpha)$$

Problema de Fisher–Behrens: Solución James (1954)

Sabemos que $(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \sim N(\boldsymbol{\theta}, (1/n_1)\boldsymbol{\Sigma}_1 + (1/n_2)\boldsymbol{\Sigma}_2)$ luego

$$D_1^2 = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^{\mathrm{T}} \left(\frac{1}{n_1}\boldsymbol{\Sigma}_1 + \frac{1}{n_2}\boldsymbol{\Sigma}_2 \right)^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \sim \chi_p^2(\lambda^2)$$

con

$$\lambda^2 = \boldsymbol{\theta}^{\mathrm{T}} \left(\frac{1}{n_1}\boldsymbol{\Sigma}_1 + \frac{1}{n_2}\boldsymbol{\Sigma}_2 \right)^{-1} \boldsymbol{\theta} \quad \text{con} \quad \boldsymbol{\theta} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 \,.$$

Un estadístico para testear $H_0 : \boldsymbol{\theta} = \mathbf{0}$ es

$$V = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^{\mathrm{T}} \mathbf{S}_{\mathrm{W}}^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \quad \text{con} \quad \mathbf{S}_{\mathrm{W}} = \frac{1}{n_1}\mathbf{S}_1 + \frac{1}{n_2}\mathbf{S}_2$$

Problema de Fisher–Behrens: Solución James (1954)

$$V = (\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2)^{\mathrm{T}} \mathbf{S}_{\mathrm{W}}^{-1} (\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2) \quad \text{con} \quad \mathbf{S}_{\mathrm{W}} = \frac{1}{n_1}\mathbf{S}_1 + \frac{1}{n_2}\mathbf{S}_2$$

Rechazo H_0 si

$$\mathbf{V} > \mathbf{k}_\alpha$$

donde

$$\begin{aligned}
k_\alpha &= \chi^2_{p,\alpha}(A + B\,\chi^2_{p,\alpha}) \\
A &= 1 + \frac{1}{2p}\sum_{j=1}^{2}\frac{1}{n_j - 1}\left(\mathrm{tr}\left(\frac{\mathbf{S}_{\mathrm{W}}^{-1}\mathbf{S}_j}{n_j}\right)\right)^2 \\
B &= \frac{1}{p(p+2)}\sum_{j=1}^{2}\frac{1}{n_j - 1}\left\{\frac{1}{2}\left(\mathrm{tr}\left(\frac{\mathbf{S}_{\mathrm{W}}^{-1}\mathbf{S}_j}{n_j}\right)\right)^2 + \mathrm{tr}\left[\left(\frac{\mathbf{S}_{\mathrm{W}}^{-1}\mathbf{S}_j}{n_j}\right)^2\right]\right\}
\end{aligned}$$

Problema de Fisher–Behrens: Aproximación Yao (1965)

$$V = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^{\mathrm{T}} \mathbf{S}_{\mathrm{W}}^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \quad \text{con} \quad \mathbf{S}_{\mathrm{W}} = \frac{1}{n_1}\mathbf{S}_1 + \frac{1}{n_2}\mathbf{S}_2$$

Yao aproximó la distribución de V bajo H_0 por un Hotelling

$$V \sim T^2_{p,m}$$

donde

$$\frac{1}{m} = \sum_{j=1}^{2} \frac{1}{n_j - 1} \left\{ \frac{(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^{\mathrm{T}} \mathbf{S}_{\mathrm{W}}^{-1} \frac{\mathbf{S}_j}{n_j} \mathbf{S}_{\mathrm{W}}^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)}{(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^{\mathrm{T}} \mathbf{S}_{\mathrm{W}}^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)} \right\}^2$$

y $\min(n_1 - 1, n_2 - 1) \leq m \leq n_1 + n_2 - 2$. Luego, rechazo H_0 si

$$\mathbf{V} > \frac{\mathbf{pm}}{\mathbf{m - p + 1}} \mathbf{f_{p,m-p+1}}(\boldsymbol{\alpha})$$

pero esta solución no es mejor que la de Bennet ya que da niveles empíricos más altos sobre todo cuando crece p.

Otra opción: Bootstrap Paramétrico

$$V = (\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2)^{\mathrm{T}} \mathbf{S}_{\mathrm{W}}^{-1} (\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2) \quad \text{con} \quad \mathbf{S}_{\mathrm{W}} = \frac{1}{n_1}\mathbf{S}_1 + \frac{1}{n_2}\mathbf{S}_2$$

a) Genero
   - $\mathbf{x}^\star_{1,1}, \ldots, \mathbf{x}^\star_{1,n_1}$ i.i.d., $\mathbf{x}^\star_{1,j} \in \mathbb{R}^p$, $\mathbf{x}^\star_{1,j} \sim N_p(\overline{\mathbf{x}}_1, \mathbf{S}_1)$
   - $\mathbf{x}^\star_{2,1}, \ldots, \mathbf{x}^\star_{2,n_2}$ i.i.d., $\mathbf{x}^\star_{2,j} \in \mathbb{R}^p$, $\mathbf{x}^\star_{2,j} \sim N_p(\overline{\mathbf{x}}_1, \mathbf{S}_2)$
   - $\mathbf{x}^\star_{1,1}, \ldots, \mathbf{x}^\star_{1,n_1}$ independientes de $\mathbf{x}^\star_{2,1}, \ldots, \mathbf{x}^\star_{2,n_2}$

b) Calculo

$$\overline{\mathbf{x}}^\star_i = \frac{1}{n_i}\sum_{j=1}^{n_i} \mathbf{x}^\star_{i,j} \qquad \mathbf{S}^\star_i = \frac{1}{n_i - 1}\sum_{j=1}^{n_i} (\mathbf{x}^\star_{i,j} - \overline{\mathbf{x}}^\star_i)(\mathbf{x}^\star_{i,j} - \overline{\mathbf{x}}^\star_i)^{\mathrm{T}} \qquad i = 1, 2$$

y definamos

$$V^\star = (\overline{\mathbf{x}}^\star_1 - \overline{\mathbf{x}}^\star_2)^{\mathrm{T}} (\mathbf{S}^\star_{\mathrm{W}})^{-1} (\overline{\mathbf{x}}^\star_1 - \overline{\mathbf{x}}^\star_2) \quad \text{con} \quad \mathbf{S}^\star_{\mathrm{W}} = \frac{1}{n_1}\mathbf{S}^\star_1 + \frac{1}{n_2}\mathbf{S}^\star_2$$

Otra opción: Bootstrap Paramétrico

c) Repitamos a) y b) *Nboot* veces obteniendo $V_1^\star, \ldots, V_{Nboot}^\star$.

d) El $p-$valor se estima como

$$p = \frac{k}{Nboot + 1}$$

donde k es la cantidad de $V_j^\star$ que son mayores o iguales que V.

$$V = (\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2)^{\mathrm{T}} \mathbf{S}_{\mathrm{W}}^{-1} (\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_2) \quad \text{con} \quad \mathbf{S}_{\mathrm{W}} = \frac{1}{n_1}\mathbf{S}_1 + \frac{1}{n_2}\mathbf{S}_2$$

Supongamos tener una batería de p tests psicológicos que se aplican a n_1 individuos de una cierta población. $\mathbf{x}_1, \ldots, \mathbf{x}_{n_1}$ son los resultados obtenidos.

$$\mathbf{x}_i \sim N(\boldsymbol{\eta}, \mathbf{\Sigma})$$

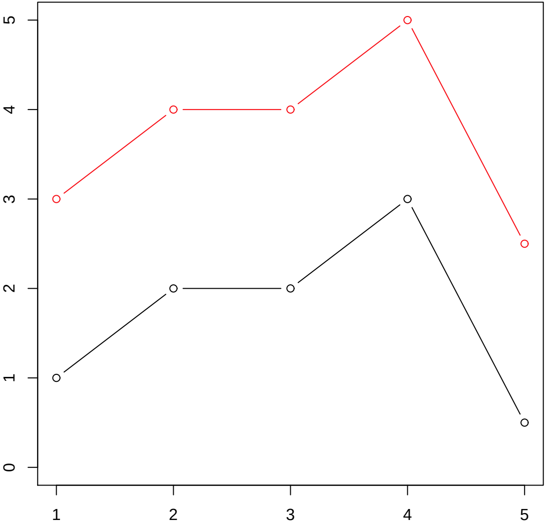
El gráfico obtenido uniéndo los puntos $(1, \eta_1), \ldots, (p, \eta_p)$ se llama el perfil de la población.

Supongamos tener una muestra de otra población de n_2 individuos. $\mathbf{y}_1, \ldots, \mathbf{y}_{n_2}$ son los resultados obtenidos. Supongamos que

$$\mathbf{y}_i \sim N(\boldsymbol{\nu}, \mathbf{\Sigma})$$

Queremos comparar los perfiles
$(1, \eta_1), \ldots, (p, \eta_p)$ y $(1, \nu_1), \ldots, (p, \nu_p)$.

Hipótesis H_{01} Paralelismo



Hipótesis H_{01} Paralelismo o no interacción

$$H_{01} : \eta_k - \eta_{k-1} = \nu_k - \nu_{k-1} \quad k = 2, \ldots, p$$

Sea

$$\mathbf{C}_1 = \begin{pmatrix} 1 & -1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & -1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & 1 & -1 \end{pmatrix} \in \mathbb{R}^{(p-1)\times p}, \quad \text{rango}(\mathbf{C}_1) = p-1$$

Luego testear H_{01} es equivalente a testear

$$\mathbf{C}_1\boldsymbol{\eta} = \mathbf{C}_1\boldsymbol{\nu}$$

Hipótesis H_{01} Paralelismo o no interacción

$$H_{01} : \mathbf{C}_1\boldsymbol{\eta} = \mathbf{C}_1\boldsymbol{\nu}$$

Definamos

$$\begin{aligned} \mathbf{v}_i &= \mathbf{C}_1\mathbf{x}_i \sim N_{p-1}(\mathbf{C}_1\boldsymbol{\eta}, \mathbf{C}_1\boldsymbol{\Sigma}\mathbf{C}_1^{\mathrm{T}}) \\ \mathbf{u}_i &= \mathbf{C}_1\mathbf{y}_i \sim N_{p-1}(\mathbf{C}_1\boldsymbol{\nu}, \mathbf{C}_1\boldsymbol{\Sigma}\mathbf{C}_1^{\mathrm{T}}) \end{aligned}$$

Sea

$$\begin{aligned} \mathbf{S} &= \frac{(n_1-1)\mathbf{S}_1 + (n_2-1)\mathbf{S}_2}{n_1+n_2-2} \\ &= \frac{\sum_{j=1}^{n_1}(\mathbf{x}_j - \overline{\mathbf{x}})(\mathbf{x}_j - \overline{\mathbf{x}})^{\mathrm{T}} + \sum_{j=1}^{n_2}(\mathbf{y}_j - \overline{\mathbf{y}})(\mathbf{y}_j - \overline{\mathbf{y}})^{\mathrm{T}}}{n_1+n_2-2} \end{aligned}$$

Luego, bajo $H_{0,1}$

$$T_0^2 = \frac{n_1\, n_2}{n_1+n_2}(\overline{\mathbf{x}} - \overline{\mathbf{y}})^{\mathrm{T}}\mathbf{C}_1^{\mathrm{T}}\left(\mathbf{C}_1\mathbf{S}\mathbf{C}_1^{\mathrm{T}}\right)^{-1}\mathbf{C}_1(\overline{\mathbf{x}} - \overline{\mathbf{y}}) \sim T^2_{p-1, n_1+n_2-2}$$

Hipótesis H_{01} Paralelismo o no interacción

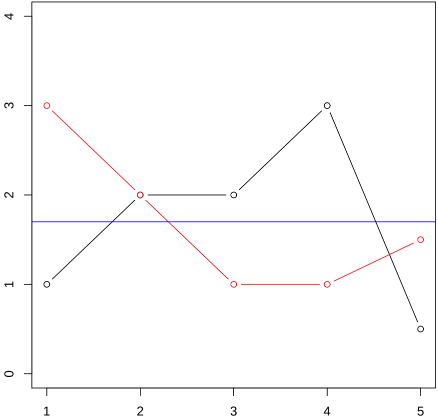
$$H_{01} : \mathbf{C}_1\boldsymbol{\eta} = \mathbf{C}_1\boldsymbol{\nu}$$

Por lo tanto rechazo $H_{0,1}$ si

$$T_0^2 > \frac{(n_1 + n_2 - 2)(p-1)}{n_1 + n_2 - p} f_{p-1, n_1+n_2-p}(\alpha)$$

Hipótesis H_{02} No hay efecto principal por población

$$H_{02} : \frac{1}{p}\sum_{j=1}^{p}\eta_j = \frac{1}{p}\sum_{j=1}^{p}\nu_j$$



Hipótesis H_{02} No hay efecto principal por población

$$H_{02} : \frac{1}{p}\sum_{j=1}^{p}\eta_j = \frac{1}{p}\sum_{j=1}^{p}\nu_j$$

Si H_{01} y H_{02} son ambas ciertas, los perfiles coinciden. Sea $\mathbf{1}_p = (1, \ldots, 1)^{\mathrm{T}} \in \mathbb{R}^p$, luego

$$H_{02} : \mathbf{1}_p^{\mathrm{T}}\boldsymbol{\eta} = \mathbf{1}_p^{\mathrm{T}}\boldsymbol{\nu}$$

Sea

$$\begin{aligned} v_i &= \mathbf{1}_p^{\mathrm{T}}\mathbf{x}_i \sim N_1(\mathbf{1}_p^{\mathrm{T}}\boldsymbol{\eta}, \mathbf{1}_p^{\mathrm{T}}\mathbf{\Sigma}\mathbf{1}) \\ u_i &= \mathbf{1}_p^{\mathrm{T}}\mathbf{y}_i \sim N_1(\mathbf{1}_p^{\mathrm{T}}\boldsymbol{\nu}, \mathbf{1}_p^{\mathrm{T}}\mathbf{\Sigma}\mathbf{1}) \end{aligned}$$

Observemos que $\mathbf{1}_p^{\mathrm{T}}\mathbf{\Sigma}\mathbf{1} = \sum_{i,j}\sigma_{ij}$.

Hipótesis H_{02}

Luego, bajo $H_{0,2}$

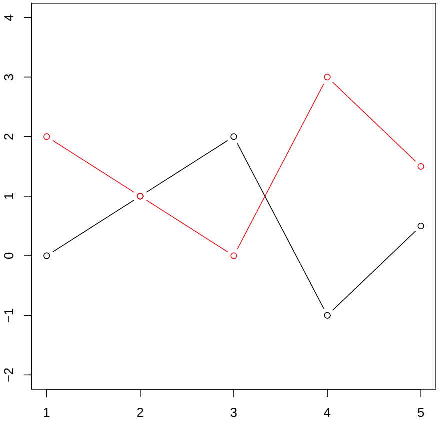
$$
\begin{aligned}
T_0^2 &= \frac{n_1\, n_2}{n_1+n_2} \mathbf{1}_p^{\mathrm{T}}(\overline{\mathbf{x}}-\overline{\mathbf{y}}) \left(\mathbf{1}_p^{\mathrm{T}} \mathbf{S} \mathbf{1}_p\right)^{-1} \mathbf{1}_p^{\mathrm{T}}(\overline{\mathbf{x}}-\overline{\mathbf{y}}) \\
&= \left\{ \frac{\mathbf{1}_p^{\mathrm{T}}(\overline{\mathbf{x}}-\overline{\mathbf{y}})}{\left[\mathbf{1}_p^{\mathrm{T}} \mathbf{S} \mathbf{1}_p \left(\frac{1}{n_1}+\frac{1}{n_2}\right)\right]^{1/2}} \right\}^2 \sim T^2_{1,n_1+n_2-2} = \mathcal{F}_{1,n_1+n_2-2}
\end{aligned}
$$

O sea, rechazo H_{02} si

$$
\frac{\left|\sum_{j=1}^{p}(\overline{x}_j-\overline{y}_j)\right|}{\left[\sum_{i,j} s_{ij}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\right]^{1/2}} > t_{n_1+n_2-2,\alpha}
$$

Hipótesis H_{03} No hay efecto principal por test

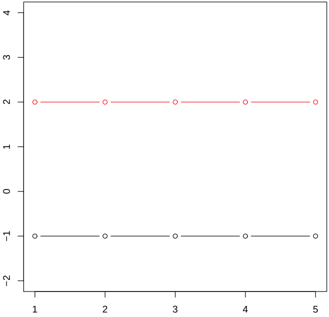
$$H_{02} : \frac{1}{2}(\eta_1 + \nu_1) = \frac{1}{2}(\eta_2 + \nu_2) = \cdots = \frac{1}{2}(\eta_p + \nu_p)$$



Hipótesis H_{03} No hay efecto principal por test

$$H_{03}: \frac{1}{2}(\eta_1 + \nu_1) = \frac{1}{2}(\eta_2 + \nu_2) = \cdots = \frac{1}{2}(\eta_p + \nu_p)$$

Si H_{01} y H_{03} son ambas ciertas, los perfiles son constantes, o sea, $\eta_1 = \eta_2 = \cdots = \eta_p$ y $\nu_1 = \nu_2 = \cdots = \nu_p$.



Hipótesis H_{03} No hay efecto principal por test

$$H_{03}: \frac{1}{2}(\eta_1+\nu_1) = \frac{1}{2}(\eta_2+\nu_2) = \cdots = \frac{1}{2}(\eta_p+\nu_p)$$

Sea

$$\mathbf{C}_1 = \begin{pmatrix} 1 & -1 & 0 & 0 & \dots & 0 \\ 0 & 1 & -1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 & -1 \end{pmatrix} \in \mathbb{R}^{(p-1)\times p}, \quad \text{rango}(\mathbf{C}_1) = p-1$$

Luego testear H_{03} es equivalente a testear

$$\mathbf{C}_1(\boldsymbol{\eta}+\boldsymbol{\nu}) = 0 \quad \text{o sea} \quad \mathbf{C}_1\boldsymbol{\eta} = \mathbf{C}_1(-\boldsymbol{\nu})$$

Hipótesis H_{03} No hay efecto principal por test

$$H_{03}:\frac{1}{2}(\eta_1+\nu_1)=\frac{1}{2}(\eta_2+\nu_2)=\cdots=\frac{1}{2}(\eta_p+\nu_p)\ \text{ o sea }\ \mathbf{C}_1\boldsymbol{\eta}=\mathbf{C}_1(-\boldsymbol{\nu})$$

Definamos

$$\begin{aligned}\mathbf{v}_i &= \mathbf{C}_1\mathbf{x}_i\sim N_{p-1}(\mathbf{C}_1\boldsymbol{\eta},\mathbf{C}_1\boldsymbol{\Sigma}\mathbf{C}_1^{\mathrm{T}})\\ \mathbf{u}_i &= \mathbf{C}_1(-\mathbf{y}_i)\sim N_{p-1}(-\mathbf{C}_1\boldsymbol{\nu},\mathbf{C}_1\boldsymbol{\Sigma}\mathbf{C}_1^{\mathrm{T}})\end{aligned}$$

Sea

$$\begin{aligned}\mathbf{S} &= \frac{(n_1-1)\mathbf{S}_1+(n_2-1)\mathbf{S}_2}{n_1+n_2-2}\\ &= \frac{\sum_{j=1}^{n_1}(\mathbf{x}_j-\overline{\mathbf{x}})(\mathbf{x}_j-\overline{\mathbf{x}})^{\mathrm{T}}+\sum_{j=1}^{n_2}(\mathbf{y}_j-\overline{\mathbf{y}})(\mathbf{y}_j-\overline{\mathbf{y}})^{\mathrm{T}}}{n_1+n_2-2}\end{aligned}$$

Hipótesis H_{03} No hay efecto principal por test

$$H_{03} : \frac{1}{2}(\eta_1+\nu_1) = \frac{1}{2}(\eta_2+\nu_2) = \cdots = \frac{1}{2}(\eta_p+\nu_p) \ \text{ o sea } \ \mathbf{C}_1\boldsymbol{\eta} = \mathbf{C}_1(-\boldsymbol{\nu})$$

Luego, bajo H_{03}

$$T_0^2 = \frac{n_1\, n_2}{n_1+n_2}(\overline{\mathbf{x}}+\overline{\mathbf{y}})^{\mathrm{T}}\mathbf{C}_1^{\mathrm{T}}\left(\mathbf{C}_1\mathbf{S}\mathbf{C}_1^{\mathrm{T}}\right)^{-1}\mathbf{C}_1(\overline{\mathbf{x}}+\overline{\mathbf{y}}) \sim T^2_{p-1,n_1+n_2-2}$$

Por lo tanto rechazo H_{03} si

$$T_0^2 > \frac{(n_1+n_2-2)(p-1)}{n_1+n_2-p} f_{p-1,n_1+n_2-p}(\alpha)$$

Hipótesis H_{03} No hay efecto principal por test

$$H_{03} : \frac{1}{2}(\eta_1+\nu_1) = \frac{1}{2}(\eta_2+\nu_2) = \cdots = \frac{1}{2}(\eta_p+\nu_p) \text{ o sea } \mathbf{C}_1\boldsymbol{\eta} = \mathbf{C}_1(-\boldsymbol{\nu})$$

Queremos testear H_{03} dado que H_{01} es cierto, o sea, cuando sabemos que $\mathbf{C}_1\boldsymbol{\eta} = \mathbf{C}_1\boldsymbol{\nu}$. Definamos

$$\begin{aligned}
\bar{\mathbf{z}} &= \frac{n_1\bar{\mathbf{x}} + n_2\bar{\mathbf{y}}}{n_1+n_2} \sim N_p\left(\frac{n_1\boldsymbol{\eta} + n_2\boldsymbol{\nu}}{n_1+n_2}, \frac{1}{n_1+n_2}\boldsymbol{\Sigma}\right) \\
\mathbf{S} &= \frac{(n_1-1)\mathbf{S}_1 + (n_2-1)\mathbf{S}_2}{n_1+n_2-2} \\
\mathbf{Q} &= (n_1+n_2-2)\mathbf{S} \sim \mathcal{W}(\boldsymbol{\Sigma}, p, n_1+n_2-2)
\end{aligned}$$

de donde

$$\begin{aligned}
\mathbf{C}_1\bar{\mathbf{z}} &\sim N_p\left(\frac{n_1\mathbf{C}_1\boldsymbol{\eta} + n_2\mathbf{C}_1\boldsymbol{\nu}}{n_1+n_2}, \frac{1}{n_1+n_2}\mathbf{C}_1\boldsymbol{\Sigma}\mathbf{C}_1^{\mathrm{T}}\right) \\
\mathbf{C}_1\mathbf{Q}\mathbf{C}_1^{\mathrm{T}} &\sim \mathcal{W}(\mathbf{C}_1\boldsymbol{\Sigma}\mathbf{C}_1^{\mathrm{T}}, p-1, n_1+n_2-2)
\end{aligned}$$

Hipótesis H_{03} No hay efecto principal por test

$$H_{03}: \frac{1}{2}(\eta_1+\nu_1) = \frac{1}{2}(\eta_2+\nu_2) = \cdots = \frac{1}{2}(\eta_p+\nu_p) \text{ o sea } \mathbf{C}_1(\boldsymbol{\eta}+\boldsymbol{\nu}) = \mathbf{0}$$

Queremos testear H_{03} dado que H_{01} es cierto, o sea, cuando sabemos que $\mathbf{C}_1\boldsymbol{\eta} = \mathbf{C}_1\boldsymbol{\nu}$. Por lo tanto, $\mathbf{C}_1\boldsymbol{\eta} = \mathbf{C}_1\boldsymbol{\nu} = \mathbf{0}$

Bajo $H_{03} \cap H_{01}$, tenemos que

$$\mathbf{C}_1\overline{\mathbf{z}} \sim N_p(\mathbf{0}, \frac{1}{n_1+n_2}\mathbf{C}_1\mathbf{\Sigma}\mathbf{C}_1^{\mathrm{T}})$$

$$T_0^2 = (n_1+n_2)\left(\mathbf{C}_1\overline{\mathbf{z}}\right)^{\mathrm{T}}\left(\mathbf{C}_1\mathbf{S}\mathbf{C}_1^{\mathrm{T}}\right)^{-1}\mathbf{C}_1\overline{\mathbf{z}} \sim T^2_{p-1,n_1+n_2-2}$$

Por lo tanto, rechazo H_{03} cuando H_{01} es cierta, si

$$T_0^2 > \frac{(n_1+n_2-2)(p-1)}{n_1+n_2-p} f_{p-1,n_1+n_2-p}(\alpha)$$