Resumen

1

Prueba de Box: m poblaciones

$$H_0: \Sigma_1 = \cdots = \Sigma_m = \Sigma$$

H_A: alguna es diferente

El paquete utiliza una aproximación más inexacta

Estimador combinado

$$S_{comb} = \frac{1}{n-m}\left[(n_1-1)S_1 + \ldots + (n_m-1)S_m\right]$$

Si H_0 es cierta el estimador combinado es un estimador insesgado de Σ

Estadístico de la prueba

$$M = -2\ln(\Lambda) \underset{H_0}{\sim} F(\nu_1, \nu_2)$$

$$\Lambda = \prod_{i=1}^{m}\left(\frac{\det(S_i)}{\det(S_{comb})}\right)^{(n_i-1)/2}$$

2

Caso 1 población multivariada. Resumen

$$X \sim n_p(\underset{\sim}{\mu}, \Sigma)$$

$$(X_i)_{i=\overline{1,n}}$$

$H_0: \underline{\mu} = \underline{\mu_0}$ $H_A: \underline{\mu} \neq \underline{\mu_0}$

Estadístico T^2 (de Hotelling)

$$T^2 = n(\bar{X} - \mu_0)^t \bar{S}^{-1} (\bar{X} - \mu_0)$$

$$T^2 \underset{H_0}{\sim} \frac{(n-1)p}{n-p} F(p, n-p)$$

Región de confianza

$$RC(X_1, \ldots, X_n) =$$

$$\left\{ n(\bar{X} - \mu)^t \bar{S}^{-1} (\bar{X} - \mu) \leq \frac{(n-1)p}{n-p} F_{1-\alpha}(p, n-p) \right\}$$

Intervalos de confianza simultáneos

$$\bar{X}_i \pm \sqrt{\frac{p_\alpha}{n} S_{ii}}$$

$$p_\alpha = \frac{(n-1)p}{n-p} F_{1-\alpha}(p, n-p)$$

3

Caso 2 poblaciones multivariadas. Varianzas y cov iguales. Resumen

$H_0: \underline{\mu_1} = \underline{\mu_2}$

$H_A: \underline{\mu_1} \neq \underline{\mu_2}$

$$T^2 = (\bar{X}_1 - \bar{X}_2)^t \tilde{S}^{-1} (\bar{X}_1 - \bar{X}_2)$$

$$\tilde{S} = \left(\frac{1}{n_1} + \frac{1}{n_2}\right) S_{comb}$$

$$T^2 \underset{H_0}{\sim} \frac{(n_1+n_2-2)p}{(n_1+n_2-p-1)} F_{p, n_1+n_2-p-1}$$

Valor-p ; $P_{H_0}(\{T^2 > T^2_{obs}\})$

Región de confianza

$$(\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2))^t \tilde{S}^{-1} (\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)) \leq C$$

$$\tilde{S} = \left(\frac{1}{n_1} + \frac{1}{n_2}\right) S_{comb}$$

$$S_{comb} = \frac{1}{n_1+n_2-2} \left[(n_1-1) S_1 + (n_2-1) S_2\right]$$

$$C = \frac{(n_1+n_2-2)p}{(n_1+n_2-p-1)} \cdot F_{1-\alpha}(p, n_1+n_2-p-1)$$

Intervalo de confianza simultáneo

$$\mu_{1i} - \mu_{2i} \qquad i = \overline{1,p}$$

$$\bar{X}_{1i} - \bar{X}_{2i} \pm \sqrt{C\, \tilde{S}[i,i]}$$

4

Caso 2 poblaciones multivariadas.
Varianzas y cov diferentes.
Resumen

$$H_0: \underline{\mu}_1 = \underline{\mu}_2$$
$$H_A: \underline{\mu}_1 \neq \underline{\mu}_2$$

James (1954) proposed a test for linear hypotheses of the population means when the variances (or the covariance matrices) are not known. Its form for two p-dimensional samples is:

$$T_u^2 = (\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_1)^T \tilde{\mathbf{S}}^{-1} (\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_1), \text{ with } \tilde{\mathbf{S}} = \tilde{\mathbf{S}_1} + \tilde{\mathbf{S}_2} = \frac{\mathbf{S}_1}{n_1} + \frac{\mathbf{S}_2}{n_2}.$$

5