Introducción al Análisis Multivariado

Cap-3. Inferencia Estadística
(SEMANA-8)

Raúl Alberto Pérez

Universidad Nacional de Colombia

Escuela de Estadística

Semestre 2021-I

Oficna 43-216A

Correo: raperez1@unal.edu.co

Introducción

En este tema realizaremos tareas de inferencia sobre el vector de medias ($\underline{\mu}$) y la matriz de varianzas covarianzas (Σ) de una población normal multivariada, en base a una muestra aleatoria simple extraída de ella.

También se tratarán problemas que involucren a varias poblaciones.

Muchos procedimientos resultarán ser extensiones naturales de los métodos ya conocidos para poblaciones normales univariadas, mientras que en algún caso surgirán problemas nuevos, por ejemplo, comparación entre componentes del vector de medias o cuestiones de inferencia simultánea.

Inferencia Estadística para la Media (μ): Caso Univariado

Desde el punto de vista de pruebas de hipótesis (PH), el problema se puede formular como una prueba de hipótesis que compiten:

$$\begin{cases} H_0 \ : \ \mu = \mu_0 \\ \\ H_a \ : \ \mu \neq \mu_0 \end{cases}$$

Si $X_1, X_2, \ldots, X_n$ es una m.a de una población normal, el estadístico de prueba es:

$$t = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1} \ , \ \text{con}$$

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{y} \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

Rechazamos H_0, ie. rechazamos que μ_0 es un valor viable de μ, si $|t|$-excede un punto porcentual especificado de la distribución t con $(n-1)$ grados de libertad.

Ahora, Rechazar H_0 cuando $|t|$-es grande, es equivalente a rechazar H_0 si su cuadrado también es grande, ie: rechazar H_0 si:

$$t^2 = \frac{(\overline{X} - \mu_0)^2}{S^2/n} = n(\overline{X} - \mu_0)(S^2)^{-1}(\overline{X} - \mu_0), \quad \text{es grande.}$$

Lo anterior es el cuadrado de la Distancia entre la media muestral $\overline{X}$ y el valor propuesto μ_0.

Las unidades de esta distancia son expresadas en términos de desviaciones estándar estimadas de $\overline{X}$, ie. en términos de $S/\sqrt{n}$.

Una vez que $\overline{X}$ y S^2 son observados, la prueba se convierte en:

Rechazar H_0, en favor de H_1, a un nivel de significancia del α%, si

$$t_0^2 = n(\overline{x} - \mu_0)(s^2)^{-1}(\overline{x} - \mu_0) > t^2_{\alpha/2;n-1}$$

donde, $t_{\alpha/2;n-1}$-es el percentil $(1 - \alpha/2)100\%$ de la distribución t-student con $(n-1)$-grados de libertad.

Si H_0-no es rechazada, concluimos que μ_0 es un valor viable para la media poblacional normal (μ) de donde se extrajo la muestra.

Pregunta ¿Existen otros valores de μ que sean también consistentes con los datos observados? Si.

Siempre existe un conjunto de valores viables para una media poblacional normal (μ).

A partir de la correspondencia bien-conocida entre la Región de Aceptación para la P.H:

$$\begin{cases} H_0 \ : \ \mu = \mu_0 \\ \\ H_a \ : \ \mu \neq \mu_0 \end{cases}$$

y los intervalos de confianza para μ, se tiene que:

$$\{\text{No-Rechazamos } H_0 \ : \ \mu = \mu_0 \ \text{ al nivel } \ \alpha\} \Longleftrightarrow \ |\frac{\overline{x} - \mu_0}{s/\sqrt{n}}| \leq t_{\alpha/2;n-1}$$

lo cual equivale a que:

$$\{\mu_0 \ : \ \text{caé en el I.C del } \ (1-\alpha)100\% \ : \ \overline{x} \pm t_{\alpha/2;n-1}\frac{s}{\sqrt{n}}\}$$

$$\overline{x} - t_{\alpha/2;n-1}\frac{s}{\sqrt{n}} \leq \mu_0 \leq \overline{x} + t_{\alpha/2;n-1}\frac{s}{\sqrt{n}}.$$

Este intervalo consiste de todos aquellos valores μ_0, que no serían rechazados al nivel α-para probar: $H_0 \ : \ \mu = \mu_0$.

Antes de que la muestra aleatoria sea seleccionada, el I.C anterior es un Intervalo-Aleatorio, debido a que los extremos finales dependen de las vs.as $\overline{X}$ y S, ie.

$$\overline{X} - t_{\alpha/2;n-1}\frac{S}{\sqrt{n}} \leq \mu_0 \leq \overline{X} + t_{\alpha/2;n-1}\frac{S}{\sqrt{n}}.$$

La probabilidad de que este Intervalo-Aleatorio contenga a μ es $(1-\alpha)$.

Entre un número grande de tales intervalos-aleatorios independientes, aproximadamente el $(1-\alpha)100\%$ de ellos contendrán a μ.

Inferencia Estadística para la Media (μ): Caso Multivariado

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

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

Se procederá de manera análoga al caso univariado.

Ahora se tendrá la distancia al cuadrado dada por:

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

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

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

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

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

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

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

Se utilizará el siguiente resultado:

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

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

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

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

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

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

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

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

o equivalentemente, rechazamos H_0 si:

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

en caso contrario no-rechazamos H_0.

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

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

Realizar la siguiente prueba de hipótesis:

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

De la matriz de datos se encuentra que:

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

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

Similarmente se tiene que:

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

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

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

es decir que:

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

Ahora la estadística de prueba calculada es:

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

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

Antes de que la muestra es seleccionada, la estadística T^2 tiene la siguiente distribución:

$$T^2 \sim \frac{(n-1)p}{(n-p)}F_{p,n-p} = kF = \frac{(3-1)2}{(3-2)}F_{2\ ,\ 3-1} = 4F_{2\ ,\ 1}$$

Con un nivel de significancia de $\alpha = 0.05$ se tiene que:

$F_{\alpha\ ;\ 2,1} = F_{0.05\ ;\ 2,1} = 200$, *de donde*

$kF = 4F_{\alpha\ ;\ 2,1} = 4F_{0.05\ ;\ 2,1} = 4 \times 200 = 800$,

y por lo tanto como: $T_0^2 = \frac{7}{9} < 800 = kF_{cal}$, *luego no se rechaza* H_0. *equivalentemente:*

$F_0 = \frac{1}{k}T_0^2 = \frac{1}{4}(7/9) = 3.111 < 200 = F_{tabla}$, *luego: No Rechazamos* H_0

Pruebas de Hipótesis para $\underline{\mu}$ cuando Σ es Conocida (Población Normal)

Sea $\underline{\boldsymbol{x}_1}, \underline{\boldsymbol{x}_2}, \cdots, \underline{\boldsymbol{x}_n}$ una m.a de una distribución normal p-variada con $\underline{\mu}$-desconocida y $\boldsymbol{\Sigma}$-conocida, ie. $\underline{\boldsymbol{x}_i} \sim N_p(\underline{\mu}\ ,\ \boldsymbol{\Sigma})$.

Se desea contrastar las hipótesis: $\begin{cases} H_0 \ : \ \underline{\mu} = \underline{\mu}_0 \\ \\ H_a \ : \ \underline{\mu} \neq \underline{\mu}_0 \end{cases}$

El estadístico de prueba utilizado en esta situación es:

$$\chi_0^2 = n(\underline{\overline{\boldsymbol{x}}} - \underline{\mu}_0)^t \boldsymbol{\Sigma}^{-1} (\underline{\overline{\boldsymbol{x}}} - \underline{\mu}_0) \underset{\text{Bajo } H_0}{\sim} \chi^2_{(p)}$$

La regla de decisión es: Rechazar H_0 si $\chi_0^2 > \chi^2_{\alpha;p}$, con $\chi^2_{\alpha;p}$-cuantil α-superior de la distribución chi-cuadrado con p-grados de libertad.

Pruebas de Hipótesis para $\underline{\mu}$ en Muestra Grande

Sea $\underline{\boldsymbol{x}_1}, \underline{\boldsymbol{x}_2}, \cdots, \underline{\boldsymbol{x}_n}$ una m.a de una población p-variada con vector de medias $\underline{\mu}$-desconocida y matriz de varianzas covarianzas $\boldsymbol{\Sigma}$-desconocida.

Se desea contrastar las hipótesis: $\begin{cases} H_0 & : \ \underline{\mu} = \underline{\mu}_0 \\ & \\ H_a & : \ \underline{\mu} \neq \underline{\mu}_0 \end{cases}$

El estadístico de prueba utilizado en esta situación es:

$$\chi_0^2 = n(\underline{\overline{\boldsymbol{x}}} - \underline{\mu}_0)^t \boldsymbol{S}^{-1} (\underline{\overline{\boldsymbol{x}}} - \underline{\mu}_0) \underset{\text{Bajo } H_0}{\sim} \chi^2_{(p)}$$

La regla de decisión es: Rechazar H_0 si $\chi_0^2 > \chi^2_{\alpha;p}$.

PH para igualdad de vectores de medias de dos Poblaciones $\underline{\mu}_1$, $\underline{\mu}_2$

Sea $\underline{\boldsymbol{x}_1}, \underline{\boldsymbol{x}_2}, \cdots, \underline{\boldsymbol{x}_n}$ una m.a de una población normal p-variada con vector de medias $\underline{\mu}_1$-desconocida y matriz de varianzas-covarianzas $\boldsymbol{\Sigma}_1$, ie. $\underline{\boldsymbol{x}_i} \sim N_p(\underline{\mu}_1\ ,\ \boldsymbol{\Sigma}_1)$ y sean $\underline{\boldsymbol{y}_1}, \underline{\boldsymbol{y}_2}, \cdots, \underline{\boldsymbol{y}_m}$ una m.a de una población normal p-variada con vector de medias $\underline{\mu}_2$-desconocida y matriz de varianzaz covarianzas $\boldsymbol{\Sigma}_2$, ie. $\underline{\boldsymbol{y}_i} \sim N_p(\underline{\mu}_2\ ,\ \boldsymbol{\Sigma}_2)$. Ambas m.a son independientes entre si.

- $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$-Desconocida

  Se desea contrastar las hipótesis: $\begin{cases} H_0 \ :\ \underline{\mu}_1 - \underline{\mu}_2 = \underline{\boldsymbol{\delta}}_0 \\ \\ H_a \ :\ \underline{\mu}_1 - \underline{\mu}_2 \neq \underline{\boldsymbol{\delta}}_0 \end{cases}$

  En este caso, se usa como estimador de $\boldsymbol{\Sigma}$ a la varianza ponderada dada por:

$$S_p = \hat{\boldsymbol{\Sigma}} = \frac{(n-1)S_1 + (m-1)S_2}{n+m-2},$$

Bajo H_0-cierto, el estadístico de prueba es:

$$T^2 = \frac{nm}{n+m}\left(\underline{\overline{\boldsymbol{x}}} - \underline{\overline{\boldsymbol{y}}} - \underline{\boldsymbol{\delta}}_0\right)^t \boldsymbol{S}_p^{-1}\left(\underline{\overline{\boldsymbol{x}}} - \underline{\overline{\boldsymbol{y}}} - \underline{\boldsymbol{\delta}}_0\right) \sim \frac{(n+m-2)p}{n+m-p-1} F_{p;\ n+m-p-1}$$

$$= kF$$

con:

$$k = \frac{(n+m-2)p}{n+m-p-1}$$

Rechazamos H_0 si:

$$T_0^2 > kF = \frac{(n+m-2)p}{n+m-p-1} F_{\alpha\ :\ p\ ,\ n+m-p-1}$$

O equivalentemente, rechazamos H_0 si:

$$F_0 = \frac{n+m-p-1}{(n+m-2)p} T_0^2 = \frac{1}{k} T_0^2 > F_{tabla} = F_{\alpha\ :\ p\ ,\ n+m-p-1}$$

Ejemplo 2 .

Cuatro pruebas psicológicas fueron aplicadas a 32 hombres y 32 mujeres. Las variables consideradas en dicha prueba fueron: X_1*: Inconsistencias pictóricas,* X_2*: Reconocimiento de herramientas,* X_3*: Forma de emplear el papel y* X_4*: Vocabulario. Se pretende establecer si la respuesta media para las cuatro variables es similar en hombres y mujeres. El experimento fue llevado a cabo de tal manera que las observaciones para ambos grupos fueron independientes. Se asume además que cada grupo de personas es una m.a de una distribución normal 4-variada, con vectores de media* $\underline{\mu}_H$ *y* $\underline{\mu}_M$ *respectivamente, y matriz de varianzas-covarianzas* $\mathbf{\Sigma}$*-común pero desconocida.*

Se desea probar las hipótesis:

$$\begin{cases} H_0 \ : \ \underline{\mu}_H = \underline{\mu}_M \\ \\ H_a \ : \ \underline{\mu}_H \neq \underline{\mu}_M \end{cases} \iff \begin{cases} H_0 \ : \ \underline{\mu}_H - \underline{\mu}_M = \underline{\mathbf{0}} \\ \\ H_a \ : \ \underline{\mu}_H - \underline{\mu}_M \neq \underline{\mathbf{0}} \end{cases} , \quad \underline{\boldsymbol{\delta}}_0 = \underline{\mathbf{0}}$$

La información recopilada permite obtener los siguientes cálculos resúmenes:

$$n = 32 \; , \quad \underline{\overline{\boldsymbol{x}}}_H = \begin{bmatrix} 15.97 \\ 15.91 \\ 27.19 \\ 22.75 \end{bmatrix} \; , \quad \boldsymbol{S}_H = \begin{bmatrix} 5.192 & 4.545 & 6.522 & 5.250 \\ & 13.18 & 6.760 & 6.266 \\ & & 28.67 & 14.47 \\ & & & 16.65 \end{bmatrix}$$

$$m = 32 \; , \quad \underline{\overline{\boldsymbol{x}}}_M = \begin{bmatrix} 12.34 \\ 13.91 \\ 16.59 \\ 21.94 \end{bmatrix} \; , \quad \boldsymbol{S}_M = \begin{bmatrix} 9.136 & 7.549 & 5.531 & 4.151 \\ & 18.60 & 5.446 & 5.466 \\ & & 13.55 & 13.55 \\ & & & 28.00 \end{bmatrix}$$

Usando estos resultados se tiene que la matriz de var-cov ponderada es:

$$S_p = \hat{\boldsymbol{\Sigma}} = \frac{(n-1)S_1 + (m-1)S_2}{n+m-2} = \begin{bmatrix} 7.164 & 6.047 & 6.027 & 4.701 \\ & 15.89 & 6.103 & 5.866 \\ & & 21.11 & 14.01 \\ & & & 22.325 \end{bmatrix}$$

Bajo H_0-cierto, el estadístico de prueba es:

$$T_0^2 = \left[\begin{pmatrix}15.97\\15.91\\27.19\\22.75\end{pmatrix} - \begin{pmatrix}12.34\\13.91\\16.59\\21.94\end{pmatrix}\right]^t$$

$$\left[\left(\frac{1}{32}+\frac{1}{32}\right)\begin{pmatrix}7.164 & 6.047 & 6.027 & 4.701\\ & 15.89 & 6.103 & 5.866\\ & & 21.11 & 14.01\\ & & & 22.325\end{pmatrix}\right]^{-1}\left[\begin{pmatrix}15.97\\15.91\\27.19\\22.75\end{pmatrix} - \begin{pmatrix}12.34\\13.91\\16.59\\21.94\end{pmatrix}\right]$$

$$= 136.67$$

Usando $\alpha = 0.05$, se tiene que:

$$F_{tabla} = F_{\alpha\ ;\ p,n+m-p-1} = F_{0.05\ ;\ 4,32+32-4-1} = F_{0.05\ ;\ 4,59} = 2.53,$$

luego:

$$kF = \frac{(n+m-2)p}{n+m-p-1} F_{\alpha\ ;\ p,n+m-p-1} = \frac{(62)4}{59} F_{0.05\ ;\ 4,59} = 10.635,$$

de donde, como:

$$T_0^2 = 136.67 > 10.635 = kF = \frac{(n+m-2)p}{n+m-p-1} F_{\alpha\ ;\ p,n+m-p-1},$$

entonces se rechaza H_0 y se concluye que los resultados medios son diferentes para hombres y mujeres a un nivel de significancia del 5%.

equivalentemente:

$$F_0 = \frac{1}{k} T_0^2 = \frac{59}{(62)4}(136.67) > 2.53 = F_{tabla}$$

- $\Sigma_1 \neq \Sigma_2$-Desconocida

  Se desea contrastar las hipótesis:

$$\begin{cases} H_0 & : \ \underline{\mu}_1 - \underline{\mu}_2 = \underline{\boldsymbol{\delta}}_0 \\ H_a & : \ \underline{\mu}_1 - \underline{\mu}_2 \neq \underline{\boldsymbol{\delta}}_0 \end{cases}$$

  Bajo H_0-cierto, el estadístico de prueba es:

$$T^2 = \left(\underline{\overline{\boldsymbol{x}}} - \underline{\overline{\boldsymbol{y}}} - \underline{\boldsymbol{\delta}}_0\right)^t \left[\frac{S_1}{n} + \frac{S_2}{m}\right]^{-1} \left(\underline{\overline{\boldsymbol{x}}} - \underline{\overline{\boldsymbol{y}}} - \underline{\boldsymbol{\delta}}_0\right) \sim \frac{vp}{v-p+1} F_{p\ ;\ v-p+1} = kF$$

$$\text{con:} \quad k = \frac{vp}{v-p+1} \quad \text{y} \quad v = \frac{tr(S_e) + \left[tr(S_e)\right]^2}{\sum_{i=1}^{2} \frac{1}{n_i - 1}\left\{tr(V_i) + [tr(V_i)]^2\right\}}$$

$$V_i = \frac{S_i}{n_i} \quad \text{y} \quad S_e = V_1 + V_2 = \frac{S_1}{n} + \frac{S_2}{m}.$$

Rechazamos H_0 si: $T_0^2 > kF$ ó $F_0 = \frac{1}{k} T_0^2 > F_{tabla}$

Ejemplo 3 *Se compararon dos tipos de suelos, uno de los cuales tiene un tipo de bacteria y el otro no. Las variables aleatorias de interés que fueron medidas son: X_1: PH del suleo, X_2: Cantidad de fosfato y X_3: Contenido de nitrógeno.*

Los resultados de las mediciones se muestran en la siguiente tabla. Se desea verificar si hay similitud en ambos suelos, en relación con los vectores de medias asociados a estas tres variables medidas. ¿Cuál es la conclusión usando un $\alpha = 0.05$? $n = 13$ y $m = 10$.

X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3
8.0	60	58	8.3	57	60	5.7	42	14
8.0	156	68	7.0	94	43	6.2	40	23
8.0	90	37	8.5	86	40	6.4	49	18
6.1	44	27	8.4	52	48	5.8	31	17
7.4	207	31	7.9	146	52	6.4	31	19
7.4	120	32	6.2	49	30	5.4	62	26
8.4	65	43	5.6	31	23	5.4	42	16
8.1	237	45	5.8	42	22			

Se desea probar las hipótesis:

$$\begin{cases} H_0 \ : \ \underline{\mu}_1 = \underline{\mu}_2 \\ \\ H_a \ : \ \underline{\mu}_1 \neq \underline{\mu}_2 \end{cases} \iff \begin{cases} H_0 \ : \ \underline{\mu}_1 - \underline{\mu}_2 = \underline{\mathbf{0}} \\ \\ H_a \ : \ \underline{\mu}_1 - \underline{\mu}_2 \neq \underline{\mathbf{0}} \end{cases}$$

es decir que: $\underline{\boldsymbol{\delta}}_0 = \underline{\mathbf{0}}$.

Realizando los cálculos respectivos, se obtiene que bajo H_0-cierto, el estadístico de prueba es:

$$T^2 = \left(\underline{\overline{\boldsymbol{x}}} - \underline{\overline{\boldsymbol{y}}} - \underline{\boldsymbol{\delta}}_0\right)^t \left[\frac{S_1}{n} + \frac{S_2}{m}\right]^{-1} \left(\underline{\overline{\boldsymbol{x}}} - \underline{\overline{\boldsymbol{y}}} - \underline{\boldsymbol{\delta}}_0\right) = 96.18$$

Después de ciertos cálculos, se obtiene que el valor de v es:

$$v = \frac{tr(S_e) + \left[tr(S_e)\right]^2}{\sum_{i=1}^{2} \frac{1}{n_i - 1}\left\{tr(V_i) + [tr(V_i)]^2\right\}} = 12.874 \approx 13$$

Para $\alpha = 0.05$ se tiene que:

$$F_{tabla} = F_{\alpha\,:\,p\,,\,v-p+1} = F_{0.05;3,13-3+1} = F_{0.05;3,11} = 3.59$$

y

$$kF = \frac{vp}{v-p+1} F_{\alpha\,:\,p\,,\,v-p+1} = \frac{13(3)}{11} \times 3.59 = 12.73,$$

como $T_0^2 = 96.18 > 12.73 = kF_{tabla}$,

luego rechazamos H_0 y se cuncluye que las características medias son diferentes para ambos tipos de suelo, aun nivel de significancia del 5%.

equivalentemente:

$$F_0 = \frac{1}{k} T_0^2 = \frac{11}{13(3)}(96.18) > 3.59 = F_{tabla}$$

- $\Sigma_1 = \Sigma_2 = \Sigma$-Conocida

Se desea contrastar las hipótesis: $\begin{cases} H_0 \ : \ \underline{\mu}_1 - \underline{\mu}_2 = \underline{\boldsymbol{\delta}}_0 \\ \\ H_a \ : \ \underline{\mu}_1 - \underline{\mu}_2 \neq \underline{\boldsymbol{\delta}}_0 \end{cases}$

Como ambas m.a provienen de poblaciones normales p-variadas, entonces:

$$\underline{\overline{\boldsymbol{x}}} \sim N_p\left(\underline{\mu}_1 \ , \ \frac{\boldsymbol{\Sigma}}{n}\right) \quad \text{y} \quad \underline{\overline{\boldsymbol{y}}} \sim N_p\left(\underline{\mu}_2 \ , \ \frac{\boldsymbol{\Sigma}}{m}\right)$$

y como además, ambas m.a provienen de poblaciones independientes, luego:

$$\left(\underline{\overline{\boldsymbol{x}}} - \underline{\overline{\boldsymbol{y}}}\right) \sim N_p\left(\underline{\mu}_1 - \underline{\mu}_2 \ , \ \frac{\boldsymbol{\Sigma}}{n} + \frac{\boldsymbol{\Sigma}}{m}\right) \quad \text{es decir,}$$

$$Var\left[\underline{\overline{\boldsymbol{x}}} - \underline{\overline{\boldsymbol{y}}}\right] = \left(\frac{1}{n} + \frac{1}{m}\right)\boldsymbol{\Sigma} = \frac{n+m}{nm}\boldsymbol{\Sigma}.$$

Luego, bajo H_0 : $\underline{\mu}_1 - \underline{\mu}_2 = \underline{\boldsymbol{\delta}}_0$-cierto, es decir: $\underline{\mu}_1 - \underline{\mu}_2 = \underline{\boldsymbol{\delta}}_0$, se tiene que el estadístico de prueba:

$$\chi_0^2 = \left(\frac{nm}{n+m}\right)\left(\underline{\overline{\boldsymbol{x}}} - \underline{\overline{\boldsymbol{y}}} - \underline{\boldsymbol{\delta}}_0\right)^t \boldsymbol{\Sigma}^{-1}\left(\underline{\overline{\boldsymbol{x}}} - \underline{\overline{\boldsymbol{y}}} - \underline{\boldsymbol{\delta}}_0\right) \sim \chi^2_{(p)}.$$

La regla de decisión es: Rechazamos H_0 si $\chi_0^2 > \chi^2_{\alpha;p}$.

Ejemplo 4 .

*Los Biólogos **Grojan** y **Wirth** (1981) descubrieron dos nuevas especies de insectos **Ameroheleafasciata** (AF) y **Apseudofasciata** (APF). Puesto que las especies son similares en apariencia, resulta útil para el biólogo estar en capacidad de clasificar un especimen como AF o APF basado en características externas que son fáciles de medir. Entre alguna de las características que distinguen los AF de los APF, los biólogos reportan medidas de la longitud de las antenas (X) y la longitud de las alas (Y), ambas medidas en milímetros, de nueve insectos AF y de seis insectos APF.*

Suponga que las observaciones bivariadas para ambas poblaciones de insectos, provienen de distribuciones normales bivariadas, con igual matriz de covarianzas.

Sea $\underline{\boldsymbol{x}_1}, \underline{\boldsymbol{x}_2}, \cdots, \underline{\boldsymbol{x}_9}$ *una m.a que representa las observaciones bivariadas para la población de insectos* AF*, ie.* $\underline{\boldsymbol{x}_i} \sim N_p(\underline{\mu}_{AF}\ ,\ \Sigma)$ *para* $i = 1, 2, \cdots, 9$*. Análogamente, Sea* $\underline{\boldsymbol{y}_1}, \underline{\boldsymbol{y}_2}, \cdots, \underline{\boldsymbol{x}_6}$ *una m.a que representa las observaciones bivariadas para la población de insectos* APF*, ie.* $\underline{\boldsymbol{y}_i} \sim N_p(\underline{\mu}_{APF}\ ,\ \Sigma)$ *para* $i = 1, 2, \cdots, 6$*.*

Se asume que ambas m.a son independientes entre si.

ESPECIE	X	Y
AF	1.38	1.64
AF	1.40	1.20
AF	1.24	1.72
AF	1.36	1.74
AF	1.38	1.82
AF	1.48	1.82
AF	1.54	1.82
AF	1.38	1.90
AF	1.56	2.08

ESPECIE	X	Y
APF	1.14	1.78
APF	1.20	1.86
APF	1.18	1.96
APF	1.30	1.96
APF	1.26	2.00
APF	1.28	2.00

Además se asume que:

$$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma} = \begin{bmatrix} 0.022 & 0.006 \\ 0.006 & 0.043 \end{bmatrix}$$

Se desea contrastar las hipótesis:

$$\begin{cases} H_0 : \underline{\mu}_{AF} = \underline{\mu}_{APF} \\ \\ H_a : \underline{\mu}_{AF} \neq \underline{\mu}_{APF} \end{cases} \iff \begin{cases} H_0 : \underline{\mu}_{AF} - \underline{\mu}_{APF} = \underline{\mathbf{0}} \\ \\ H_a : \underline{\mu}_{AF} - \underline{\mu}_{APF} \neq \underline{\mathbf{0}} \end{cases} .$$

De los datos muestrales se tiene que:

$$\overline{\underline{\boldsymbol{x}}} = \begin{bmatrix} 1.41 \\ 1.80 \end{bmatrix} \quad \text{y} \quad \overline{\underline{\boldsymbol{y}}} = \begin{bmatrix} 1.23 \\ 1.93 \end{bmatrix},$$

luego bajo H_0-cierto, el estadístico de prueba es:

$$\chi_0^2 = \left(\frac{nm}{n+m}\right)(\underline{\overline{\boldsymbol{x}}} - \underline{\overline{\boldsymbol{y}}})^t \boldsymbol{\Sigma}^{-1}(\underline{\overline{\boldsymbol{x}}} - \underline{\overline{\boldsymbol{y}}})$$

$$= \left(\frac{9 \times 6}{9+6}\right)[1.41 - 1.23 \quad 1.80 - 1.93] \begin{bmatrix} 0.022 & 0.006 \\ 0.006 & 0.043 \end{bmatrix}^{-1} \begin{bmatrix} 1.41 - 1.23 \\ 1.80 - 1.93 \end{bmatrix}$$

$$= \frac{54}{15}[0.18 \quad -0.13] \begin{bmatrix} 48.62 & -7.29 \\ -7.29 & 24.31 \end{bmatrix} \begin{bmatrix} 0.18 \\ -0.13 \end{bmatrix} = \frac{54}{15} \times 2.33 = 8.39$$

Para $\alpha = 0.05$, se tiene que: $\chi^2_{\alpha;p} = \chi^2_{0.05;2} = 5.99$, de donde, como $\chi_0^2 > \chi^2_{\alpha;p}$, luego se rechaza H_0 y se concluye que la información registrada indica que los vectores de medias para la longitud de alas y antenas de las dos especies de insectos son diferentes a un nivel de significancia del 5%.

Pruebas de Hipótesis Acerca de dos vectores de Medias Poblacionales $\underline{\mu}_1$ y $\underline{\mu}_2$, para muestras grandes

En este caso se desconoce la distribución de la cual provienen ambas muestras. Se supone que los tamaños de ambas m.a son grandes, $n-p$ y $m-p$ grandes.

Sea $\underline{\boldsymbol{x}_1}, \underline{\boldsymbol{x}_2}, \cdots, \underline{\boldsymbol{x}_n}$ una m.a de una población p-variada con vector de medias $\underline{\mu}_1$-desconocida y matriz de varianzas-covarianzas $\boldsymbol{\Sigma}_1$ y sean $\underline{\boldsymbol{y}_1}, \underline{\boldsymbol{y}_2}, \cdots, \underline{\boldsymbol{y}_m}$ una m.a de una población p-variada con vector de medias $\underline{\mu}_2$-desconocida y matriz de varianzas-covarianzas $\boldsymbol{\Sigma}_2$. Ambas m.a son independientes entre si.

$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$-Desconocida

Se usa como estimador de $\boldsymbol{\Sigma}$ a S_P, ie.

$$\widehat{\boldsymbol{\Sigma}} = S_P = \frac{(n-1)S_1 + (m-1)S_2}{n-m-2}$$

y el estadístico de prueba es:

$$\chi_0^2 = \frac{nm}{n+m}\left(\overline{\boldsymbol{x}} - \overline{\boldsymbol{y}} - \underline{\gamma}_0\right)^t \boldsymbol{S}_p^{-1}\left(\overline{\boldsymbol{x}} - \overline{\boldsymbol{y}} - \underline{\gamma}_0\right) \sim \chi_p^2.$$

Rechazamos H_0 si $\chi_0^2 > \chi_{\alpha,p}$.

$\Sigma_1 \neq \Sigma_2$-Desconocidas

Se usa como estimador de $\boldsymbol{\Sigma}$ a S_e, ie.

$$\hat{\boldsymbol{\Sigma}} = S_e = \frac{S_1}{n} + \frac{S_2}{m}$$

y el estadístico de prueba es:

$$\chi_0^2 = \left(\overline{\underline{\boldsymbol{x}}} - \overline{\underline{\boldsymbol{y}}} - \underline{\gamma}_0\right)^t \boldsymbol{S}_e^{-1} \left(\overline{\underline{\boldsymbol{x}}} - \overline{\underline{\boldsymbol{y}}} - \underline{\gamma}_0\right) \sim \chi_p^2.$$

Rechazamos H_0 si $T_0^2 > \chi_{\alpha,p}$.

$\Sigma_1 = \Sigma_2 = \Sigma$-Conocida

Se desea contrastar las hipótesis:

$$\begin{cases} H_0 \ : \ \underline{\mu}_1 - \underline{\mu}_2 = \underline{\gamma}_0 \\ \\ H_a \ : \ \underline{\mu}_1 - \underline{\mu}_2 \neq \underline{\gamma}_0 \end{cases}$$

En este caso se utiliza como estadístico de prueba a:

$$\chi_0^2 = \frac{nm}{n+m}\left(\underline{\overline{\boldsymbol{x}}} - \underline{\overline{\boldsymbol{y}}} - \underline{\gamma}_0\right)^t \boldsymbol{\Sigma}^{-1}\left(\underline{\overline{\boldsymbol{x}}} - \underline{\overline{\boldsymbol{y}}} - \underline{\gamma}_0\right) \sim \chi_p^2$$

Rechazamos H_0 si $\chi_0^2 > \chi_{\alpha,p}$.

Pruebas de Hipótesis Acerca de dos vectores de Medias Poblacionales $\underline{\mu}_1$ y $\underline{\mu}_2$, Observaciones Pareadas

Sea $\underline{\boldsymbol{x}}_1, \underline{\boldsymbol{x}}_2, \cdots, \underline{\boldsymbol{x}}_n$ y $\underline{\boldsymbol{y}}_1, \underline{\boldsymbol{y}}_2, \cdots, \underline{\boldsymbol{y}}_n$ dos m.a de una población p-variada con vectores de medias $\underline{\mu}_{\underline{\boldsymbol{x}}}$, $\underline{\mu}_{\underline{\boldsymbol{y}}}$ desconocidos y matriz de varianzas-covarianzas $\boldsymbol{\Sigma}$- Desconocida.

Suponga además que $Cov[\underline{\boldsymbol{x}}_i, \underline{\boldsymbol{y}}_i] = \boldsymbol{\Sigma} \neq \boldsymbol{0}$, $i = 1, 2, \ldots, n$, ie. las dos m.a son correladas (Muestras no independientes o dependientes).

Para contrastar las hipótesis:: $\begin{cases} H_0 : \underline{\mu}_{\underline{\boldsymbol{x}}} - \underline{\mu}_{\underline{\boldsymbol{y}}} = \underline{0} \\ \\ H_a : \underline{\mu}_{\underline{\boldsymbol{x}}} - \underline{\mu}_{\underline{\boldsymbol{y}}} \neq \underline{0} \end{cases}$

Se trabaja con las diferencias de cada par de observaciones multivariadas, definidas como:

$$\underline{\boldsymbol{D}}_i = \underline{\boldsymbol{x}}_i - \underline{\boldsymbol{y}}_i.$$

Se asume que estas n-diferencias tienen una distribución normal-multivariada con vector de medias $\underline{\mathbf{0}}$ y matriz de Var-Cov dada por $\mathbf{\Sigma}$.

La hipótesis a probar, es equivalente a probar: $\begin{cases} H_0 \ : \ \underline{\mu}_D = \underline{0} \\ \\ H_0 \ : \ \underline{\mu}_D \neq \underline{0} \end{cases}$

Bajo H_0-cierta el estadístico de prueba a usar es:

$$T^2 = n\overline{\underline{\boldsymbol{D}}}^t \boldsymbol{S}_D^{-1} \overline{\underline{\boldsymbol{D}}} \sim\approx \frac{(n-1)p}{n-p} F_{p,n-p}, \text{ donde,}$$

$$\overline{\underline{\boldsymbol{D}}} = \frac{1}{n}\sum_{i=1}^{n} \underline{\boldsymbol{D}}_i \quad \text{y} \quad \boldsymbol{S}_D = \frac{1}{n-1}\sum_{i=1}^{n} (\underline{\boldsymbol{D}}_i - \overline{\underline{\boldsymbol{D}}})(\underline{\boldsymbol{D}}_i - \overline{\underline{\boldsymbol{D}}})^t,$$

son el vector de medias y la matriz de Var-Cov de las diferencias muestrales.

Se rechaza H_0 si: $\quad T_0^2 > \frac{(n-1)p}{n-p} F_{\alpha;p,n-p}$.

Ejemplo 5 *Se desea comparar dos tipos de esmalte para la resistencia a la corrosión, 15 piezas de tubería fueron cubiertas con cada tipo de esmalte. Dos tuberías, cada una con un esmalte diferente, se enterraron y se dejaron durante el mismo período de tiempo en 15 lugares distintos; esto corresponde a un par de observaciones en condiciones similares, excepto por el tipo de cubrimiento. El efecto por la corrosión fue medido a través de dos variables:* X_1*-Profundidad máxima de la picadura por corrosión (milésimas de pulgadas) y* X_2*-Número de picaduras por corrosión. Los datos y las respectivas diferencias aparecen en la siguiente tabla:*

La hipótesis a probar es: $\begin{cases} H_0 \ : \ \underline{\mu}_D = \underline{0} \\ \\ H_0 \ : \ \underline{\mu}_D \neq \underline{0} \end{cases}$

donde $\underline{\mu}_D$*-representa el vector de medias para las diferencias.*

Localidad	Esmalte-1		Esmalte-2		Diferencia	
	X_1	X_2	Y_1	Y_2	D_{E1}	D_{E2}
1	73	31	51	35	22	-4
2	43	19	41	14	2	5
3	47	22	43	19	4	3
4	53	26	41	29	42	-3
5	58	36	47	34	11	2
6	47	30	32	26	15	4
7	52	29	24	19	28	10
8	38	36	43	37	-5	-1
9	61	34	53	24	8	10
10	56	33	52	27	4	6
11	56	19	57	14	-1	5
12	34	19	44	19	-10	0
13	55	26	57	30	-2	-4
14	65	15	40	7	25	8
15	75	18	68	13	7	5

Los resultados muestrales son:

$$\underline{\boldsymbol{d}} = \begin{bmatrix} 8.000 \\ 3.067 \end{bmatrix}, \quad \boldsymbol{S}_D = \begin{bmatrix} 121.571 & 17.071 \\ 17.171 & 21.781 \end{bmatrix}$$

Bajo H_0-cierta, el estadístico de prueba es:

$$T_0^2 = (15) \begin{bmatrix} 8.000 & 3.067 \end{bmatrix} \begin{bmatrix} 121.571 & 17.071 \\ 17.171 & 21.781 \end{bmatrix}^{-1} \begin{bmatrix} 8.000 \\ 3.067 \end{bmatrix} = 10.815$$

Para $\alpha = 0.05$, se tiene que $F_{\alpha;p,n-p} = F_{0.05;2,13} = 3.806$, de donde,

$$\frac{(n-1)p}{n-p} F_{\alpha;p,n-p} = \frac{14(2)}{13} \times 3.806 = 8.198,$$

luego, como $T_0^2 > 8.198$ entonces se rechaza H_0 y se concluye que la evidencia muestral no apoya la hipótesis nula, ie. que hay diferencia entre los dos tipos de esmalte, con un nivel de significancia del 5%.

PH de Contrastes para el vector de medias Poblacional $\underline{\mu}$, de una $N_p(\underline{\mu}\ ,\ \Sigma)$

- Σ-desconocida

  Sea $\underline{\boldsymbol{x}_1}, \underline{\boldsymbol{x}_2}, \cdots, \underline{\boldsymbol{x}_n}$ una m.a de una población normal p-variada con vector de medias $\underline{\mu}$-desconocido y matriz de varianzas-covarianzas $\boldsymbol{\Sigma}$- desconocida, ie. $\underline{\boldsymbol{x}_i} \sim N_p(\underline{\mu}\ ,\ \boldsymbol{\Sigma})$.

  Sea $C_{k\times p}$-una matriz de constantes. C-contiene los coeficientes para k-combinaciones lineales simultáneas de las componentes de $\underline{\mu} = (\mu_1, \mu_2, \ldots, \mu_p)^t$, es decir:

$$C\underline{\mu} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{k1} & c_{k2} & \cdots & c_{kp} \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} = \begin{bmatrix} c_{11}\mu_1 + c_{12}\mu_2 + \cdots + c_{1p}\mu_p \\ c_{21}\mu_1 + c_{22}\mu_2 + \cdots + c_{2p}\mu_p \\ \\ c_{k1}\mu_1 + c_{k2}\mu_2 + \cdots + c_{kp}\mu_p \end{bmatrix}_{k\times 1}$$

Se desea contrastar las hipótesis:

$$\begin{cases} H_0 \ : \ C\underline{\mu} = \underline{\gamma} \\ \\ H_0 \ : \ C\underline{\mu} \neq \underline{\gamma} \end{cases} , \quad \text{con} \quad \underline{\gamma} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_k \end{pmatrix} - \text{Vector de Constantes.}$$

Un estimador insesgado para $C\underline{\mu}$ es: $C\underline{\overline{\boldsymbol{x}}}$, el cual tiene la siguiente distribución:

$$C\underline{\overline{\boldsymbol{x}}} \sim N_k\Big(C\underline{\mu} \ , \ C\boldsymbol{\Sigma}_{\underline{\overline{\boldsymbol{x}}}}C^t\Big) \quad , \quad \text{es decir :}$$

$$C\underline{\overline{\boldsymbol{x}}} \sim N_k\Big(C\underline{\mu} \ , \ \frac{1}{n}C\boldsymbol{\Sigma}C^t\Big), \quad \text{pues :} \quad \boldsymbol{\Sigma}_{\underline{\overline{\boldsymbol{x}}}} = \frac{\boldsymbol{\Sigma}}{n}.$$

Como $\boldsymbol{\Sigma}$-es desconocida se usa la estadística de prueba:

$$T_0^2 = n\Big(C\underline{\overline{\boldsymbol{x}}} - \underline{\gamma}\Big)^t\Big[C\boldsymbol{S}C^t\Big]^{-1}\Big(C\underline{\overline{\boldsymbol{x}}} - \underline{\gamma}\Big) \sim \frac{(n-1)k}{n-k}F_{k,n-k} = cF$$

Se rechaza H_0 si: $T_0^2 > cF$ ó $F_0 = \frac{1}{c}T_0^2 > F_{tabla}$, $c = \frac{(n-1)k}{n-k}$

- Σ-es conocida

  En este caso se utiliza: $\chi_0^2 = n\left(C\underline{\overline{\boldsymbol{x}}} - \underline{\gamma}\right)^t \left[C\boldsymbol{\Sigma}C^t\right]^{-1}\left(C\underline{\overline{\boldsymbol{x}}} - \underline{\gamma}\right) \sim \chi_k^2$.

  Se rechaza H_0 : si $\chi_0^2 > \chi_{\alpha;k}^2$.

Ejemplo 6 .

Los datos contenidos en la siguiente tabla, corresponden a los pesos (en gramos) del corcho encontrado en muestras tomadas de 28 árboles cultivados en una parcela experimental. Estas muestras fueron tomadas en 4-direcciones: Norte (N), este (E), sur (S) y oeste (O). En este caso se tienen 4-variables, que corresponden a las mediciones tomadas en cada dirección. Se quiere verificar si los pesos medios de corcho son iguales en la dirección norte-sur y en la dirección este-oeste. Se asume que estas cuatro mediciones tienen una distribución normal 4-variada, con vector de medias $\underline{\mu} = (\mu_N,\ \mu_E,\ \mu_S,\ \mu_O)$ *y matriz de varianzas-covarianzas* Σ, *ambas cantidades desconocidas. Es decir se desea probar:* $H_0\ :\ \mu_1 = \mu_3$ *y* $\mu_2 = \mu_4$.

N	E	S	O	N	E	S	O
X_1	X_2	X_3	X_4	X_1	X_2	X_3	X_4
72	66	76	77	91	79	100	75
60	53	66	63	56	68	47	50
56	57	64	58	79	65	70	61
41	29	36	38	81	80	68	58
32	32	35	36	78	55	67	60
30	35	34	26	46	38	37	38
39	39	31	27	39	35	34	37
42	43	31	25	32	30	30	32
37	40	31	25	60	50	67	54
33	29	27	36	35	37	48	39
32	30	34	28	39	36	39	31
63	45	74	63	50	34	37	40
54	46	60	52	43	37	39	50
47	51	52	43	48	54	57	43

Esta hipótesis se puede expresar como:

$$\begin{cases} H_0 \ : \ C\underline{\mu} = \underline{\gamma} \\ \\ H_0 \ : \ C\underline{\mu} \neq \underline{\gamma} \end{cases} , \quad \text{con} \quad \underline{\gamma} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} .$$

$$\text{y con :} \quad C = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} , \ \ \text{y} \quad \underline{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} .$$

Después de realizar algunos cálculos previos, se obtiene el vector de medias y la matriz de var-cov muestrales dadas por:

$$\underline{\boldsymbol{x}} = \begin{bmatrix} 50.535 \\ 46.179 \\ 49.679 \\ 45.179 \end{bmatrix} , \quad \boldsymbol{S} = \begin{bmatrix} 290.41 & 233.75 & 288.44 & 226.27 \\ & 219.93 & 229.06 & 171.37 \\ & & 350.00 & 259.54 \\ & & & 226.00 \end{bmatrix}$$

De igual forma se tiene que:

$$\boldsymbol{C}\underline{\overline{\boldsymbol{x}}} = \begin{bmatrix} 0.857 \\ 1.00 \end{bmatrix}, \qquad \boldsymbol{CSC}^t = \begin{bmatrix} 63.53 & 27.96 \\ 27.96 & 103.19 \end{bmatrix}$$

Luego, bajo H_0-cierto y con: $\underline{\gamma} = \underline{0}$, se tiene que el estadístico de prueba es:

$$T_0^2 = n\left(C\underline{\overline{\boldsymbol{x}}} - \underline{\gamma}\right)^t \left[C\boldsymbol{S}C^t\right]^{-1} \left(C\underline{\overline{\boldsymbol{x}}} - \underline{\gamma}\right) \sim \frac{(n-1)k}{n-k} F_{k,n-k}$$

$$T_0^2 = 28 \begin{bmatrix} 0.857 & 1.00 \end{bmatrix} \begin{bmatrix} 63.53 & 27.96 \\ 27.96 & 103.19 \end{bmatrix}^{-1} \begin{bmatrix} 0.857 \\ 1.00 \end{bmatrix} = 28(0.0158) = 0.443$$

Para $\alpha = 0.05$, se tiene: $F_{\alpha;\ k,n-k} = F_{0.05;\ 2,26} = 3.369$. Ahora,

$$cF_{tabla} = \frac{(n-1)k}{n-k} F_{\alpha;k,n-k} = \frac{(28-1)2}{28-2} \times 3.69 = \frac{27}{13} \times 3.369 = 6.7$$

Luego, como $T_0^2 = 0.4426 < 6.7 = cF_{tabla}$, entonces no se rechaza a H_0 y se concluye que la evidencia muestral parece ser coherente con la hipótesis planteada a un nivel de significancia del 5%, es decir, el contenido medio de corcho en la direcciones indicadas no son significativamente diferentes.

Prueba de Razón de Verosimilitud (Σ-Desconocida)

Sea $\underline{\boldsymbol{x}}_1, \underline{\boldsymbol{x}}_2, \ldots, \underline{\boldsymbol{x}}_n$ una m.a de una distribución $N_p(\underline{\mu}, \boldsymbol{\Sigma})$, donde tanto μ como $\boldsymbol{\Sigma}$ son desconocidas.

Estadística de razón de verosimilitud

Se desea contrastar las hipótesis:

$$\begin{cases} H_0 & : \ \underline{\mu} = \underline{\mu_0} \\ \\ H_a & : \ \underline{\mu} \neq \underline{\mu_0} \end{cases} \quad \Longleftrightarrow \begin{cases} H_0 \ : \ \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} = \begin{bmatrix} \mu_{10} \\ \mu_{20} \\ \vdots \\ \mu_{0p} \end{bmatrix} \quad v.s \quad H_a \ : \ \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} \neq \begin{bmatrix} \mu_{10} \\ \mu_{20} \\ \vdots \\ \mu_{0p} \end{bmatrix} \end{cases}$$

Ya se tiene una prueba para esto basada en la estadística T^2-de Hotelling, ahora se construirá otra prueba que será equivalente a la T^2-de Hotelling, pero que tiene algunas propiedades estadísticas deseables.

La estadística de Razón de Verosimilitud se define como sigue:

$$\lambda := \frac{\underset{\mathbf{\Sigma}}{\text{Máx}}\ L(\underline{\mu}_0\ ,\ \mathbf{\Sigma})}{\underset{\underline{\boldsymbol{\mu}}\ ,\ \mathbf{\Sigma}}{\text{Máx}}\ L(\underline{\mu}\ ,\ \mathbf{\Sigma})} = \frac{\text{Máximo de L-Restingida}}{\text{Máximo de L-No Restingida}}$$

$$= \frac{\text{Máximo de L-Bajo } H_0\text{-Cierta}}{\text{Máximo de L-General}}$$

$$= \frac{\text{Máximo de L-Sobre el Espacio de Parámetros Restringido}}{\text{Máximo de L-Sobre todo el espacio de parámetros}}$$

donde, L-es la función de Verosimilitud de los Datos.

La Estadística λ toma valores entre cero y uno, ie. $0 \leq \lambda \leq 1$.

Para valores grandes de λ, ie. λ-cercanos a 1, No se Rechaza H_0, mientras que para valores pequeños se Rechaza H_0.

Ahora se hallará una expresión para la Estadística de Razón de Verosimilitud y su respectiva distribución de probabilidad.

Primero recordemos que la función de verosimilitud asociada a una m.a de una distribución normal p-variada es:

$$L\Big(\underline{\mu}\,,\ \mathbf{\Sigma}\Big) = L\Big(\ \underline{\mu}\,,\ \mathbf{\Sigma}\ \Big|\ \underline{x}_1, \underline{x}_1, , \ldots, \underline{x}_n\ \Big)$$

$$= \frac{1}{(2\pi)^{np/2}}\frac{1}{|\mathbf{\Sigma}|^{n/2}}\mathrm{e}^{-\frac{1}{2}\sum_{j=1}^{n}(\underline{x}_j-\underline{\mu})'\mathbf{\Sigma}^{-1}(\underline{x}_j-\underline{\mu})}$$

y que los estimadores de Máxima-Verosimilitud de $\underline{\mu}$ y $\mathbf{\Sigma}$ son, respectivamente: $\underline{\widehat{\mu}} = \underline{\overline{\boldsymbol{x}}}$ y

$$\hat{\mathbf{\Sigma}} = \boldsymbol{S}_n = \frac{1}{n}\sum_{i=1}^{n}(\underline{\boldsymbol{x}}_i - \underline{\widehat{\mu}})(\underline{\boldsymbol{x}}_i - \underline{\widehat{\mu}})^t = \frac{1}{n}\sum_{i=1}^{n}(\underline{\boldsymbol{x}}_i - \underline{\overline{\boldsymbol{x}}})(\underline{\boldsymbol{x}}_i - \underline{\overline{\boldsymbol{x}}})^t = \frac{(n-1)}{n}\boldsymbol{S}$$

Luego, bajo $H_0\ :\ \underline{\widehat{\mu}} = \underline{\mu_0}$-Cierta, se tiene que el MLE de $\mathbf{\Sigma}$ es:

$$\hat{\mathbf{\Sigma}}_0 = \frac{1}{n}\sum_{i=1}^{n}(\underline{\boldsymbol{x}}_i - \underline{\mu}_0)(\underline{\boldsymbol{x}}_i - \underline{\mu}_0)^t$$

luego, reemplazando las expresiones anteriores en la definición de λ se tiene que:

$$\lambda := \frac{\underset{\boldsymbol{\Sigma}}{\text{Máx}}\ L(\underline{\mu}_0\ ,\ \boldsymbol{\Sigma})}{\underset{\underline{\boldsymbol{\mu}}\ ,\ \boldsymbol{\Sigma}}{\text{Máx}}\ L(\underline{\mu}\ ,\ \boldsymbol{\Sigma})} = \frac{\text{Máx}\ L(\underline{\mu}_0\ ,\ \hat{\boldsymbol{\Sigma}}_0)}{\text{Máx}\ L(\hat{\underline{\mu}}\ ,\ \hat{\boldsymbol{\Sigma}})} = \frac{\text{Máx}\ L(\underline{\mu}_0\ ,\ \hat{\boldsymbol{\Sigma}}_0)}{\text{Máx}\ L(\underline{\overline{\boldsymbol{x}}}\ ,\ \hat{\boldsymbol{\Sigma}})}$$

$$= \frac{\text{Máx}\ \frac{1}{(2\pi)^{np/2}}\frac{1}{|\hat{\boldsymbol{\Sigma}}_0|^{n/2}}\mathrm{e}^{-\frac{1}{2}\sum_{j=1}^{n}(\underline{x}_j-\underline{\mu}_0)'\hat{\boldsymbol{\Sigma}}_0^{-1}(\underline{x}_j-\underline{\mu}_0)}}{\text{Máx}\ \frac{1}{(2\pi)^{np/2}}\frac{1}{|\hat{\boldsymbol{\Sigma}}|^{n/2}}\mathrm{e}^{-\frac{1}{2}\sum_{j=1}^{n}(\underline{x}_j-\underline{\overline{\boldsymbol{x}}})'\hat{\boldsymbol{\Sigma}}^{-1}(\underline{x}_j-\underline{\overline{\boldsymbol{x}}})}}$$

$$= \frac{\text{Máx}\ \frac{1}{|\hat{\boldsymbol{\Sigma}}_0|^{n/2}}\mathrm{e}^{-\frac{1}{2}\sum_{j=1}^{n}(\underline{x}_j-\underline{\mu}_0)'\hat{\boldsymbol{\Sigma}}_0^{-1}(\underline{x}_j-\underline{\mu}_0)}}{\text{Máx}\ \frac{1}{|\hat{\boldsymbol{\Sigma}}|^{n/2}}\mathrm{e}^{-\frac{1}{2}\sum_{j=1}^{n}(\underline{x}_j-\underline{\overline{\boldsymbol{x}}})'\hat{\boldsymbol{\Sigma}}^{-1}(\underline{x}_j-\underline{\overline{\boldsymbol{x}}})}}$$

$$= \frac{\text{Máx}\ \frac{1}{|\hat{\boldsymbol{\Sigma}}_0|^{n/2}}\mathrm{e}^{-\frac{1}{2}\text{tr}\left\{\hat{\boldsymbol{\Sigma}}_0^{-1}\sum_{j=1}^{n}\left[(\underline{x}_j-\underline{\mu}_0)(\underline{x}_j-\underline{\mu}_0)'\right]\right\}}}{\text{Máx}\ \frac{1}{|\hat{\boldsymbol{\Sigma}}|^{n/2}}\mathrm{e}^{-\frac{1}{2}\text{tr}\left\{\hat{\boldsymbol{\Sigma}}^{-1}\sum_{j=1}^{n}\left[(\underline{x}_j-\underline{\overline{\boldsymbol{x}}})(\underline{x}_j-\underline{\overline{\boldsymbol{x}}})'\right]\right\}}} = \frac{\frac{1}{|\hat{\boldsymbol{\Sigma}}_0|^{n/2}}\mathrm{e}^{-\frac{np}{2}}}{\frac{1}{|\hat{\boldsymbol{\Sigma}}|^{n/2}}\mathrm{e}^{-\frac{np}{2}}} = \frac{|\hat{\boldsymbol{\Sigma}}|^{n/2}}{|\hat{\boldsymbol{\Sigma}}_0|^{n/2}}$$

En Resumen:

$$\lambda = \frac{|\widehat{\boldsymbol{\Sigma}}|^{n/2}}{|\widehat{\boldsymbol{\Sigma}}_0|^{n/2}} = \frac{|\frac{1}{n}\sum_{i=1}^{n}(\underline{\boldsymbol{x}}_i - \overline{\underline{\boldsymbol{x}}})(\underline{\boldsymbol{x}}_i - \overline{\underline{\boldsymbol{x}}})^t|^{n/2}}{|\frac{1}{n}\sum_{i=1}^{n}(\underline{\boldsymbol{x}}_i - \underline{\mu}_0)(\underline{\boldsymbol{x}}_i - \underline{\mu}_0)^t|^{n/2}}$$

$$= \frac{|\sum_{i=1}^{n}(\underline{\boldsymbol{x}}_i - \overline{\underline{\boldsymbol{x}}})(\underline{\boldsymbol{x}}_i - \overline{\underline{\boldsymbol{x}}})^t|^{n/2}}{|\sum_{i=1}^{n}(\underline{\boldsymbol{x}}_i - \underline{\mu}_0)(\underline{\boldsymbol{x}}_i - \underline{\mu}_0)^t|^{n/2}}$$

$$\lambda^{2/n} = \frac{|\widehat{\boldsymbol{\Sigma}}|}{|\widehat{\boldsymbol{\Sigma}}_0|}$$

Haciendo, $A = \sum_{i=1}^{n}(\underline{\boldsymbol{x}}_i - \overline{\underline{\boldsymbol{x}}})(\underline{\boldsymbol{x}}_i - \overline{\underline{\boldsymbol{x}}})^t$, se tiene que:

$$\sum_{i=1}^{n}(\underline{\boldsymbol{x}}_i - \underline{\mu}_0)(\underline{\boldsymbol{x}}_i - \underline{\mu}_0)^t = \sum_{i=1}^{n}(\underline{\boldsymbol{x}}_i - \overline{\underline{\boldsymbol{x}}})(\underline{\boldsymbol{x}}_i - \overline{\underline{\boldsymbol{x}}})^t + n(\overline{\underline{\boldsymbol{x}}} - \underline{\mu}_0)(\overline{\underline{\boldsymbol{x}}} - \underline{\mu}_0)^t$$

$$= A + n(\overline{\underline{\boldsymbol{x}}} - \underline{\mu}_0)(\overline{\underline{\boldsymbol{x}}} - \underline{\mu}_0)^t$$

es decir, que

$$\lambda = \frac{|A|^{n/2}}{|A + n(\underline{\overline{\boldsymbol{x}}} - \underline{\mu}_0)(\underline{\overline{\boldsymbol{x}}} - \underline{\mu}_0)^t|^{n/2}}$$

de donde,

$$\lambda^{2/n} = \frac{|A|}{|A + n(\underline{\overline{\boldsymbol{x}}} - \underline{\mu}_0)(\underline{\overline{\boldsymbol{x}}} - \underline{\mu}_0)^t|}$$

$$= \frac{|A|}{|A + \underbrace{\sqrt{n}(\underline{\overline{\boldsymbol{x}}} - \underline{\mu}_0)}_{\underline{b}}\underbrace{\sqrt{n}(\underline{\overline{\boldsymbol{x}}} - \underline{\mu}_0)^t}_{\underline{b}}|}$$

$$\lambda = \frac{|A|}{|A + \underline{b}\underline{b}^t|}$$

Propiedad de Determinantes:

$$|A + \underline{b}\underline{b}^t| = |A| \left(1 + \underline{b}^t A^{-1} \underline{b}\right) = |A| + |A|\, \underline{b}^t A^{-1} \underline{b}$$

Usando esta propiedad con:

$$A = \sum_{i=1}^{n} (\underline{\boldsymbol{x}}_i - \underline{\overline{\boldsymbol{x}}})(\underline{\boldsymbol{x}}_i - \underline{\overline{\boldsymbol{x}}})^t = (n-1)\boldsymbol{S}$$

y

$$\underline{b} = \sqrt{n}(\underline{\overline{\boldsymbol{x}}} - \underline{\mu}_0)$$

se tiene que:

$$|A + \underbrace{\sqrt{n}(\overline{\boldsymbol{x}} - \underline{\mu}_0)}_{\underline{b}}\underbrace{\sqrt{n}(\overline{\boldsymbol{x}} - \underline{\mu}_0)^t}_{\underline{b}}|$$

$$= |A| + |A| \left\{ \left[\sqrt{n}(\overline{\boldsymbol{x}} - \underline{\mu}_0)\right]^t [(n-1)\boldsymbol{S}]^{-1} \left[\sqrt{n}(\overline{\boldsymbol{x}} - \underline{\mu}_0)\right] \right\}$$

$$= |A| + |A| \left(\frac{1}{n-1}\right) \underbrace{n(\overline{\boldsymbol{x}} - \underline{\mu}_0)^t \boldsymbol{S}^{-1} (\overline{\boldsymbol{x}} - \underline{\mu}_0)}$$

$$= |A| + |A| \left(\frac{1}{n-1}\right) T_0^2$$

$$
\textit{Luego}, \quad \lambda^{2/n} = \frac{|A|}{|A + n(\underline{\overline{\boldsymbol{x}}} - \underline{\mu}_0)(\underline{\overline{\boldsymbol{x}}} - \underline{\mu}_0)^t|}
$$

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

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

$$
= \frac{1}{1 + \left(\frac{1}{n-1}\right) T_0^2} = \frac{1}{1 + \frac{T_0^2}{n-1}}
$$

A la Estadística: $\lambda^{2/n} = \dfrac{1}{1 + \frac{T_0^2}{n-1}} = \dfrac{n-1}{(n-1) + T_0^2}$

se le llama **Estadística Lambda de Wilks**.

De lo anterior se tiene que:

$$\lambda = \left[\frac{n-1}{(n-1)+T_0^2}\right]^{n/2}$$

y por tanto:

$$T_0^2 = (n-1)\left[\frac{1-\lambda_0^{2/n}}{\lambda_0^{2/n}}\right]$$

De donde:

Rechazar H_0 si λ-Es pequeño, $\Longleftrightarrow$ T_0^2 $-$ Es Grande.

En la prueba de RV se rechaza H_0 si

$$\lambda = \left(\frac{|\widehat{\Sigma}|}{|\widehat{\Sigma}_0|}\right)^{n/2} < C_\alpha$$

donde C_α- es el percentil α de la distribución de λ, llamada Dsitribución Lambda de Wilks.

Pero no es necesario usar la distribución de λ, debido a la relación que existe entre el estadístico λ-de Wilks y el estadístico T^2-de Hotelling.

Una Expresión para T^2-sin necesidad de hallar $\boldsymbol{S}^{-1}$ es:

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

$$T^2 = (n-1)\left[\frac{|\widehat{\boldsymbol{\Sigma}}_0|}{|\widehat{\boldsymbol{\Sigma}}|} - 1\right]$$

$$= (n-1)\left[\frac{|\sum_{i=1}^{n}(\underline{\boldsymbol{x}}_i - \underline{\mu}_0)(\underline{\boldsymbol{x}}_i - \underline{\mu}_0)^t|}{|\sum_{i=1}^{n}(\underline{\boldsymbol{x}}_i - \underline{\overline{\boldsymbol{x}}})(\underline{\boldsymbol{x}}_i - \underline{\overline{\boldsymbol{x}}})^t|} - 1\right]$$

Los métodos de verosimilitud, producen Estadísticos de prueba que se reducen a la familia de Estadísticas F y a la familias de estadísticas t en el caso univariado.

Método de Máxima-Verosimilitud Generalizado

Sea $\underline{\theta}$ el vector de todos los parámetros poblacionales desconocidos, de una distribución multivariada. Sea $\underline{\Theta}$ el espacio del vector de parámetros $\underline{\theta}$, ie. $\underline{\theta} \in \underline{\Theta}$, con $v = \dim(\underline{\Theta})$. Sea $L(\underline{\theta})$ la función de verosimilitud obtenida al evaluar la f.d.p conjunta de $\underline{\boldsymbol{x}}_1, \underline{\boldsymbol{x}}_2, \ldots, \underline{\boldsymbol{x}}_n$ en sus valores observados $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$.

Para la prueba de hipótesis dada por: $\begin{cases} H_0 \ : \ \underline{\theta} \in \underline{\Theta}_0 \\ \\ H_a \ : \ \underline{\theta} \in \underline{\Theta}_0^c = \underline{\Theta} - \underline{\Theta}_0 \end{cases}$

La estadística de Razon de Verosimilitud es: $\lambda := \dfrac{\underset{\underline{\theta} \in \underline{\Theta}_0}{\text{Máx}} \ L(\underline{\theta})}{\underset{\underline{\theta} \in \underline{\Theta}}{\text{Máx}} \ L(\underline{\theta})}$

H_0-es rechazada cuando λ-es pequeña.

Teorema 1 *Cuando n-es grande, bajo H_0-cierto, se cumple que:*

$$-2Log\lambda = -2Log\left(\frac{\underset{\underline{\theta}\in\underline{\Theta}_0}{\text{Máx}}\ L(\underline{\theta})}{\underset{\underline{\theta}\in\underline{\Theta}}{\text{Máx}}\ L(\underline{\theta})}\right) \sim \chi^2_{v-v_0=v^\star},$$

con, $v_0 = dim(\underline{\Theta}_0)$*, ie.* $v^\star = dim(\underline{\Theta}) - dim(\underline{\Theta}_0) = v - v_0$.

v_0*-número de parámetros desconocidos bajo* H_0*-cierta.*

NOTA: *El Test de RV tiene la mayor potencia, ie.*

$$1-\beta = P\Big[\text{Rechazar}\ \ H_0\ \Big|\ H_0\ \ \text{es Falsa}\Big],$$

entre todas las pruebas con el mismo nivel de significancia α*, ie.*

$$\alpha = P\Big[\text{Rechazar}\ \ H_0\ \Big|\ H_0\ \ \text{es V}\Big]$$

Ejemplo 7 *Para el caso de una distribución Normal p-variada, ie. $N_p(\underline{\mu}\ ,\ \Sigma)$, se tiene que:*

$$\underline{\theta}^t = [\mu_1\ ,\ \cdots\ ,\ \mu_p\ ,\ \sigma_{11}\ ,\ \cdots\ ,\ \sigma_{1p}\ ,\ \sigma_{21}\ ,\ \cdots\ ,\ \sigma_{2p}\ ,\ \cdots\ \cdots\ ,\ \sigma_{p1}\ ,\ \cdots\ ,\ \sigma_{pp}]$$

$\underline{\Theta}$-consiste un un espacio p-dimensional (de p-medias), combinado con un espacio $p(p+1)/2$-dimensional (de varianzas-covarianzas, p-varianzas y $\binom{p}{2}$-covarianzas), ie. que:

$$v = dim(\underline{\Theta}) = p + \frac{p(p+1)}{2} = \frac{p(p+3)}{2}$$

Cuando $H_0\ :\ \underline{\mu} = \underline{\mu}_0$ es cierto y Σ es no especificada, ie. cuando $\underline{\mu}$ es conocida por el valor de $\underline{\mu}_0$ y Σ es desconocida, $\underline{\theta}$ caerá en un subconjunto $\underline{\Theta}_0$ de $\underline{\Theta}$, ie. Bajo H_0-cierto: $\underline{\theta} \in \underline{\Theta}_0 \subset \underline{\Theta}$.

En este caso:

$$v_0 = dim(\underline{\Theta}_0) = 0 + \frac{p(p+1)}{2} = \frac{p(p+1)}{2}$$

Región de Confianza:

En esta situación se tiene que el conjunto de todos los $\underline{\mu}$ para los cuales se cumple la desigualdad:

$$n(\underline{\overline{\boldsymbol{x}}} - \underline{\mu})^t \boldsymbol{\Sigma}^{-1} (\underline{\overline{\boldsymbol{x}}} - \underline{\mu}) \leq \chi^2_{\alpha;p},$$

se denomina una Región de Confianza del $(1-\alpha)100\%$ para el vector de medias poblacionales $\underline{\mu}$.

La expresión anterior representa el interior y la superficie de un elipsoide con centro en $\underline{\mu} = \underline{\overline{\boldsymbol{x}}}$ y cuya forma, tamaño e inclinación, dependen de $\boldsymbol{\Sigma}$ y de $\chi^2_{\alpha;p}$, particularmente de los valores y vectores propios de $\boldsymbol{\Sigma}$.

Por ejemplo, si $\boldsymbol{\Sigma} = \boldsymbol{I_p}$ entonces la región de confianza es una esfera.

Los valores de $\underline{\mu}_0$ dentro de la región de confianza apoyan la hipótesis nula $H_0 \ : \ \underline{\mu} = \underline{\mu}_0$.