Inferencia Estadística Multivariada

Raúl Alberto Pérez
raperez1@unal.edu.co
Profesor Asociado - Escuela de Estadística
Universidad Nacional de Colombia, Sede Medellín

Semestre 2024-01

Inferencia Estadística para la Media (μ): Caso Multivariado

Pruebas de Hipótesis para $\underline{\mu}$ cuando $\boldsymbol{\Sigma}$ 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 $\boldsymbol{\Sigma}$-desconocida).

Se procederá de manera análoga al caso univariado.

Ahora se tendrá la distancia al cuadrado dada por:

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

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

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 $\overline{\underline{\boldsymbol{x}}}$, pues

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

Si la distancia estadística observada T^2 es mu grande, ie. $\overline{\underline{\boldsymbol{x}}}$-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:

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

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:

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

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

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 $\boldsymbol{\Sigma}$-desconocida, ie. $\underline{\boldsymbol{x}_i} \sim N_2(\underline{\mu}, \boldsymbol{\Sigma})$, para $i = 1, 2, 3$:

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

Realizar la siguiente prueba de hipótesis:

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

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:

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

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:

$$F_0 = \frac{1}{k} T_0^2 = \frac{1}{4}(7/9) = 0.19444 < 200 = F_{tabla}, \text{ luego: No Rechazamos } H$$

El Ejemplo Anterior en R, mediante una función de Usuario

```
datos<-data.frame(x1=c(6,10,8),x2=c(9,6,3) )
datos
```

```
##   x1 x2
## 1  6  9
## 2 10  6
## 3  8  3
```

```
mu_0<-c(9,5)
res_mu0<-HT2_mu0(datos,mu_0,0.05)
kable(res_mu0)
```

T2	K	F0	df1	df2	F_Tabla	Valor_p
0.777778	4	0.194444	2	1	199.5	0.848528

Resultados de esta PH utilizando la función `HottellingsT2` del paquete `ICSNP` del R.

En R existen varias funciones en distintos paquetes o librerias, las cuales se utilizan para realizar este tipo de pruebas de hipóteis. Se recomienda leer muy bien las ayudas que existen sobre estas funciones para utilizarlas de manera adecuada y definir de forma apropiada sus respectivos argumentos.

```r
mu_0<-c(9,5)
HotellingsT2(datos, mu = mu_0)
```

```
##
##  Hotelling's one sample T2-test
##
## data:  datos
## T.2 = 0.19444, df1 = 2, df2 = 1, p-value = 0.8485
## alternative hypothesis: true location is not equal to c(9,5)
```

Resultados de esta PH Utilizando la función `T2.test` del paquete `rrcov` del R.

```r
mu_0<-c(9,5)
T2.test(datos, mu = mu_0)
```

```
##
##  One-sample Hotelling test
##
## data:  datos
## T2 = 0.77778, F = 0.19444, df1 = 2, df2 = 1, p-value = 0.8485
## alternative hypothesis: true mean vector is not equal to (9, 5)'
##
## sample estimates:
##               x1 x2
## mean x-vector  8  6
```

Pruebas de Hipótesis para $\underline{\mu}$ cuando $\mathbf{\Sigma}$ 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 $\mathbf{\Sigma}$-conocida, ie. $\underline{\boldsymbol{x}_i} \sim N_p(\underline{\mu}\ ,\ \mathbf{\Sigma})$.

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:

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

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 $\chi_0^2 > \chi^2_{\alpha;p}$.

Ejemplo en R. Pruebas de Hipótesis para $\underline{\mu}$ n-Grande

```
##            X1     X2     X3 Grupos
## 1   9.6838 8.7045 4.0509      1
## 2   5.9703 4.5724 8.3056      1
## 3   1.5094 6.3666 7.5676      1
## 4   1.5706 3.5433 8.7762      1
## 5   3.7626 2.6972 3.7973      1
## 6        .      .      .      .
## 7        .      .      .      .
## 8        .      .      .      .
## 96  5.4857 6.2888 3.2312      3
## 97  4.9054 4.0601 1.4129      3
## 98  7.8864 6.0692  6.445      3
## 99  4.1332 4.5396 6.7288      3
## 100 4.4509 7.6449 8.2376      3
```

Estructura de los datos:

```
## 'data.frame':    100 obs. of  4 variables:
##  $ X1     : num  9.68 5.97 1.51 1.57 3.76 ...
##  $ X2     : num  8.7 4.57 6.37 3.54 2.7 ...
##  $ X3     : num  4.05 8.31 7.57 8.78 3.8 ...
##  $ Grupos: Factor w/ 3 levels "1","2","3": 1 1 1 1 1 1 1
```

```
## [1] 1 2 3
## Levels: 1 2 3
```

```
## .
##  1  2  3
## 27 28 45
```

Resultados de la PH usando la función de usuario `HT2_mu0_ngrande()`

Para este se creo una función de usuario llamada: `HT2_mu0_ngrande` la cual se utiliza a continuación.

```
mu_0<-c(0,0,0)
res_mu0ngrande<-HT2_mu0_ngrande(grupo1[,1:3], mu_0, 0.05)
kable(res_mu0ngrande)
```

Chi_2	df	Chi_Tabla	Valor_p
307.064	3	7.81473	0

Resultados utilizando la Función `HotellingsT2` del R

Igualmente, también se puede suar la función `HotellingsT2` del R de la siguiente forma.

```
mu_0<-c(0,0,0)
HotellingsT2(grupo1[,1:3],mu=mu_0,test="chi")
```

```
##
##  Hotelling's one sample T2-test
##
## data:  grupo1[, 1:3]
## T.2 = 307.06, df = 3, p-value < 2.2e-16
## alternative hypothesis: true location is not equal to c(0,0,0)
```

Medias para cada Grupo

```r
descriptivos<-resumen_xgrupos(datos[1:3],datos[4])
descriptivos[1]
```

```
## $Medias
##   Grupos       X1       X2       X3
## 1      1 4.504767 4.987719 5.178689
## 2      2 4.928179 5.102939 4.947521
## 3      3 5.690567 5.277362 4.760589
```

```r
mu_0<-c(4.5,5,5)
HotellingsT2(grupo1[,1:3],mu=mu_0,test="chi")
```

```
##
##  Hotelling's one sample T2-test
##
## data:  grupo1[, 1:3]
## T.2 = 0.22328, df = 3, p-value = 0.9737
## alternative hypothesis: true location is not equal to c(4.5,5,5)
```

Con la función de Usuario `HT2_mu0_ngrande`

```r
mu_0<-c(4.5,5,5)
res_mu0ngrande<-HT2_mu0_ngrande(grupo1[,1:3], mu_0, 0.05)
kable(res_mu0ngrande)
```

Chi_2	df	Chi_Tabla	Valor_p
0.223285	3	7.81473	0.973746

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.

1. $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$-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:

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

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

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 $\boldsymbol{\Sigma}$-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:

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

Otro Ejemplo Usando R: $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$-Desconocida, Pob. Normal

Resultados usando la función de usuario de Usuario `HT2_sigmas_iguales()`

```
x<-grupo1[,1:3]
y<-grupo2[,1:3]
mu_0<-c(0,0,0)
res2p<-HT2_sigmas_iguales(x,y,mu_0,0.05)
kable(res2p)
```

T2	k	F0	df1	df2	F_Tabla	Valor_p
0.387667	3.11765	0.124346	3	51	2.78623	0.945298

Resultados Utilizando la función `HotellingsT2` del paquete `ICSNP` del R.

Existen varias funciones en distintos paquetes del R que se utilizan para esta prueba de hipótesis cuando $\Sigma_1 = \Sigma_2 = \Sigma$-Desconocida, Pob. Normal.

```
res_1<-HotellingsT2(x,y)
res_1
```

```
##
##  Hotelling's two sample T2-test
##
## data:  x and y
## T.2 = 0.12435, df1 = 3, df2 = 51, p-value = 0.9453
## alternative hypothesis: true location difference is not equal to c(0,0,0)
```

Estadístico$= F_0 \sim F$	gl_{num}	gl_{den}	p-Valor
0.12435	3	51	0.9453

Resultados Utilizando la función `T2.test` del paquete `rrcov` del R.

```
resT2<-T2.test(x,y)
resT2
```

```
##
##  Two-sample Hotelling test
##
## data:  x and y
## T2 = 0.38767, F = 0.12435, df1 = 3, df2 = 51, p-value = 0.9453
## alternative hypothesis: true difference in mean vectors is not equal to (0,0
## sample estimates:
##                      X1       X2       X3
## mean x-vector 4.504767 4.987719 5.178689
## mean y-vector 4.928179 5.102939 4.947521
```

$T2$	F	df_1	df_2	$Valor - p$
0.38767	0.12435	3	51	0.9453

Resultados Utilizando la función `hotelling.test` del paquete `Hotelling` del R.

```
resh<-hotelling.test(x, y)
resh
```

```
## Test stat:  0.38767
## Numerator df:  3
## Denominator df:  51
## P-value:  0.9453
```

$T2$	$1/K$	df_1	df_2	n_x	n_y	p
0.387666735927368	0.320754716981132	3	51	27	28	3

2. $\boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2$-Desconocida (**Aproximación de: Nel and Van Der Merwe-1986**)

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:

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

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

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

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:

$$T^2 = (\underline{\bar{\mathbf{x}}} - \underline{\bar{\mathbf{y}}} - \underline{\boldsymbol{\delta}}_0)^t \left[\frac{S_1}{n} + \frac{S_2}{m}\right]^{-1} (\underline{\bar{\mathbf{x}}} - \underline{\bar{\mathbf{y}}} - \underline{\boldsymbol{\delta}}_0) = 96.8178$$

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

$$kF = \frac{vp}{v-p+1} F_{\alpha\ :\ p\ ,\ v-p+1} = \frac{13(3)}{11} \times 3.59 = 3.54 \times 3.59 = 12.7083,$$

como $T_0^2 = 96.8178 > 12.73 = kF_{tabla}$,

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:

$$F_0 = \frac{1}{k} T_0^2 = \frac{11}{13(3)}(96.18) = 27.1276 > 3.59 = F_{tabla}$$

Usando R. $\mathbf{\Sigma}_1 \neq \mathbf{\Sigma}_2$-Desconocidas, Pob. Normal

```
##    x1  x2 x3
## 1   8  60 58
## 2   8 156 68
## 3   8  90 37
```

```
##     x1 x2 x3
## 1 6.2 49 30
## 2 5.6 31 23
## 3 5.8 42 22
```

Resultados usando la función de usuario `HT2_sigmas_diferentes()` (**Aproximación de: Nel and Van Der Merwe-1986**).

T_2	v	k	F0	df1	df2	F_Tabla	Valor_p
96.8178	13	3.54545	27.3076	3	11	3.58743	2.14103e-05

Resultados usando la función de usuario `HT2_sigmas_diferentes_texto_guia()` (**Aproximación de: Krishnamoorty and Yu-2004**)

T_2	v	k	F0	df1	df2	F_Tabla	Valor_p
96.8178	70	3.08824	31.3505	3	68	2.7395	7.69607e-13

Otro Ejemplo Usando R. $\mathbf{\Sigma}_1 \neq \mathbf{\Sigma}_2$-Desconocidas, Pob. Normal

Resultados usando la función de usuario `HT2_sigmas_diferentes()` (**Aproximación de: Nel and Van Der Merwe-1986**).

```
mu_0<-c(0,0,0)
res_sigmasdif<-HT2_sigmas_diferentes(x,y,mu_0,0.05)
kable(res_sigmasdif)
```

T_2	v	k	F0	df1	df2	F_Tabla	Valor_p
0.386428	53	3.11765	0.123949	3	51	2.78623	0.94554

Otro Ejemplo Usando R. $\boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2$-Desconocidas, Pob. Normal

Resultados usando la función de usuario `HT2_sigmas_diferentes_texto_guia()` (**Aproximación de: Krishnamoorty and Yu-2004**)

Para este caso el v se calcula como sigue:

$$v = \frac{p+p^2}{\frac{1}{n_1}\left\{tr\left[(V_1 S_e^{-1})^2\right] + [tr(V_1 S_e^{-1})]^2\right\} + \frac{1}{n_2}\left\{tr\left[(V_2 S_e^{-1})^2\right] + [tr(V_2 S_e^{-1})]^2\right\}}$$

```
mu_0<-c(0,0,0)
res_sigmasdif1<-HT2_sigmas_diferentes_texto_guia(x,y,mu_0,0.05)
kable(res_sigmasdif1)
```

T_2	v	k	F0	df1	df2	F_Tabla	Valor_p
0.386428	211	3.02871	0.127589	3	209	2.6478	0.943665

3. $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2 = \mathbf{\Sigma}$-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:

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

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

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

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

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. $\underline{\boldsymbol{x}_i} \sim N_p(\underline{\mu}_{AF}\ ,\ \boldsymbol{\Sigma})$ 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. $\underline{\boldsymbol{y}_i} \sim N_p(\underline{\mu}_{APF}\ ,\ \boldsymbol{\Sigma})$ 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

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:

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

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.

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

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

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

y el estadístico de prueba es:

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

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

2. $\boldsymbol{\Sigma}_1 \neq \boldsymbol{\Sigma}_2$-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:

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

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

3. $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$-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 $\chi_0^2 > \chi_{\alpha,p}$.

Ejemplos Usando R. $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \Sigma$-Desconocida para n-grande

Resultados usando la función de usuario
`HT2_sigmas_iguales_ngrande()`

```
mu_0<-c(0,0,0)
res2p_ngrande<-HT2_sigmas_iguales_ngrande(x,y,mu_0,0.05)
kable(res2p_ngrande)
```

Chi2	df	Chi_Tabla	Valor_p
0.387667	3	7.81473	0.942778

Resultados Utilizando la función `HotellingsT2` del paquete `ICSNP` del R:

```
res_1A<-HotellingsT2(x,y,test="chi")
res_1A
```

```
##
##  Hotelling's two sample T2-test
##
## data:  x and y
## T.2 = 0.38767, df = 3, p-value = 0.9428
## alternative hypothesis: true location difference is not equal to c(0,0,0)
```

Estadístico= $\chi_0 \sim \chi^2$	gl	p-Valor
0.38767	3	0.94278

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:

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 $\boldsymbol{\Sigma}$.

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:

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

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

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.

	Esmalte-1		Esmalte-2		Diferencia	
Localidad	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 $N_p(\underline{\mu}\ ,\ \mathbf{\Sigma})\}$

a) $\mathbf{\Sigma}$-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 $\mathbf{\Sigma}$- desconocida, ie. $\underline{\boldsymbol{x}_i} \sim N_p(\underline{\mu}\ ,\ \mathbf{\Sigma})$.

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:

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

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

b) $\boldsymbol{\Sigma}$-es conocida

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

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 $\boldsymbol{\Sigma}$, 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:

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

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

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

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

De igual forma se tiene que:

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

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.

Un ejemplo en R. Usando Función de Usuario.

Se desea contrastar las hipótesis:

$$\begin{cases} H_0 \;:\; C\underline{\mu} = \underline{\delta}_0 \\ \\ H_0 \;:\; C\underline{\mu} \neq \underline{\delta}_0 \end{cases}$$

$$C\underline{\mu} = \underline{\delta}_0 \Longleftrightarrow C\underline{\mu} = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Longleftrightarrow \begin{bmatrix} \mu_1 - 2\mu_3 \\ \mu_1 + \mu_2 - \mu_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Es decir, se desea probar:

$\mu_1 - 2\mu_3 = 0, \;\; \mu_1 + \mu_2 - \mu_3 = 0 \;\Longleftrightarrow\; \mu_1 = 2\mu_3, \;\; \mu_1 + \mu_2 = \mu_3$

Resultados usando la función de usuario: `HT2_CU`

```r
C<-matrix(c(1,0,2,1,1,-1),byrow=T,ncol=3)
delta_0<-c(0,0)
res_CU<-HT2_CU(grupo1[,1:3],C,delta_0,0.05)
kable(res_CU)
```

T2	c	F0	df1	df2	F_Tabla	Valor_p
290.25	2.08	139.543	2	25	3.38519	2.73115e-14

Un ejemplo en R. Usando Función de Usuario para n-grande

Resultados usando la función de usuario: `HT2_CU_ngrande`

```
C<-matrix(c(1,0,2,1,1,-1),byrow=T,ncol=3)
delta_0<-c(0,0)
res_CUng<-HT2_CU_ngrande(grupo1[,1:3],C,delta_0,0.05)
kable(res_CUng)
```

Chi2	df1	Chi_Tabla	Valor_p
290.25	2	5.99146	0

Prueba de Razón de Verosimilitud ($\boldsymbol{\Sigma}$-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:

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

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:

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

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

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

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

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

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

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

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

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

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

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

En Resumen:

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

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

$$\lambda^{2/n} = \frac{|\hat{\mathbf{\Sigma}}|}{|\hat{\mathbf{\Sigma}}_0|}$$

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:

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

y

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

se tiene que:

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

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

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

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

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

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

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

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

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:

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

$$\text{y por tanto:} \quad T_0^2 = (n-1)\left[\frac{1-\lambda_0^{2/n}}{\lambda_0^{2/n}}\right]$$

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

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

donde C_α- es el percentil α de la distribución de λ, llamada Distribución 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:

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

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

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

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 Razón de 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: Cuando n-es grande, bajo H_0-cierto, se cumple que:

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

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. $N_p(\underline{\mu}\ ,\ \mathbf{\Sigma})$, 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:

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

Cuando $H_0\ :\ \underline{\mu} = \underline{\mu}_0$ es cierto y $\mathbf{\Sigma}$ es no especificada, ie. cuando $\underline{\mu}$ es conocida por el valor de $\underline{\mu}_0$ y $\mathbf{\Sigma}$ 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:

$$v_0 = \dim(\underline{\Theta}_0) = 0 + \frac{p(p+1)}{2} = \frac{p(p+1)}{2} \quad \text{y} \quad v^\star = v - v_0 = p$$

Región de Confianza

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

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

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 $\boldsymbol{\Sigma}$.

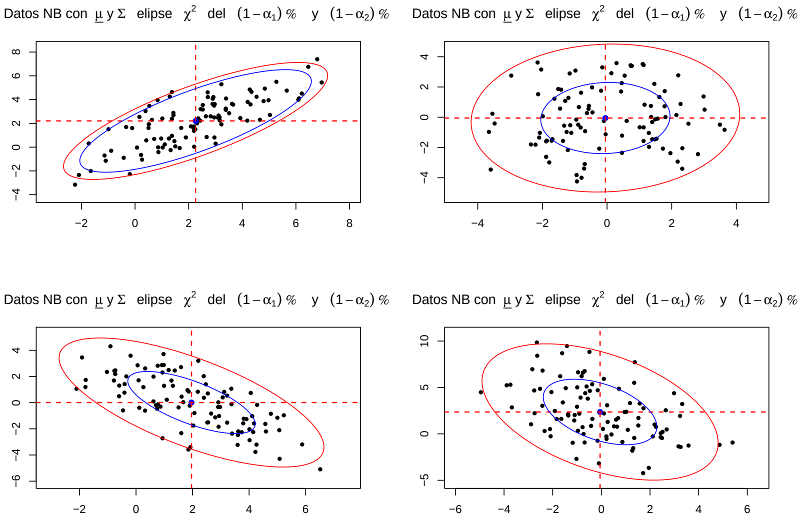
Por ejemplo, si $\boldsymbol{\Sigma} = \boldsymbol{I_p}$ entonces la región de confianza es una esfera.

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

Función de Usuario en R que representa datos bivariados, junto con elipses (regiones) de confianza

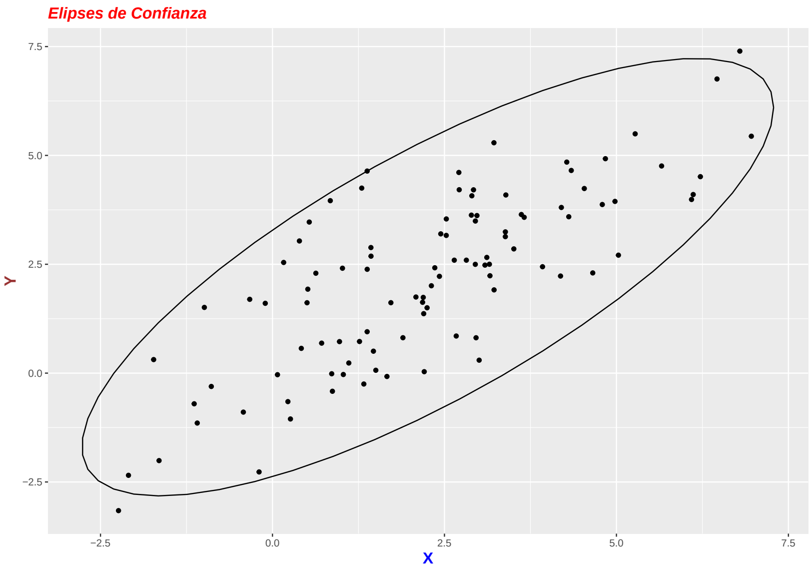
Con esta función se realzia un gráfico de dispersión de dos variables junto con dos elipses de confianza del $(1-\alpha_1)100\%$ y $(1-\alpha_2)100\%$. El centro de la elipse está representado por un punto de color azul.

```
par(mfrow=c(2,2))
mu=c(2,2)
sigma<-matrix(c(5,4,4,5),ncol=2)
datos2<-round(mvrnorm(n=100, mu,sigma),3)
representa2c_ng(datos2,0.9,0.95)
mu=c(0,0)
sigma<-matrix(c(3,0,0,3),ncol=2)
datos2<-round(mvrnorm(n=100, mu,sigma),3)
representa2c_ng(datos2,0.5,0.95)
mu=c(2,0)
sigma<-matrix(c(5,-4,-4,5),ncol=2)
datos2<-round(mvrnorm(n=100, mu,sigma),3)
representa2c_ng(datos2,0.5,0.95)
mu=c(0,2)
sigma<-matrix(c(5,-4,-4,10),ncol=2)
datos2<-round(mvrnorm(n=100, mu,sigma),3)
representa2c_ng(datos2,0.5,0.95)
```

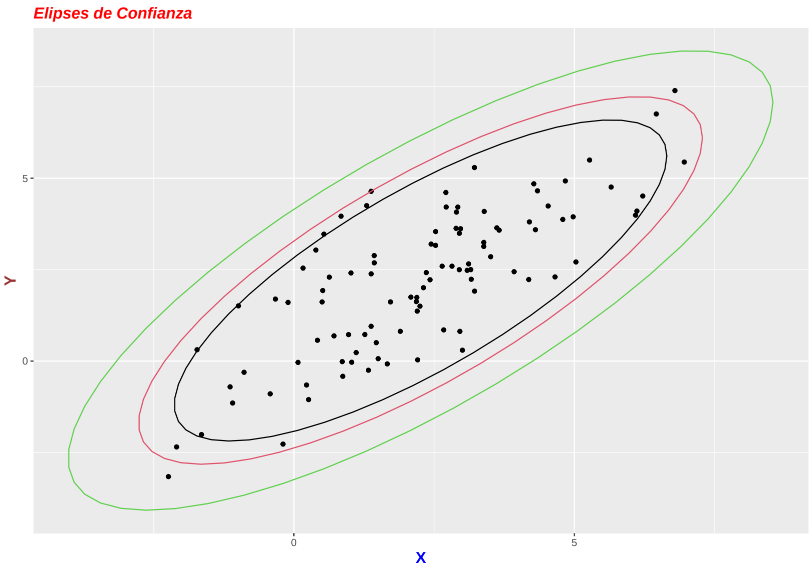

Datos NB con $\underline{\mu}$ y Σ elipse χ^2 del $(1-\alpha_1)\,\%$ y $(1-\alpha_2)\,\%$
Datos NB con $\underline{\mu}$ y Σ elipse χ^2 del $(1-\alpha_1)\,\%$ y $(1-\alpha_2)\,\%$
Datos NB con $\underline{\mu}$ y Σ elipse χ^2 del $(1-\alpha_1)\,\%$ y $(1-\alpha_2)\,\%$
Datos NB con $\underline{\mu}$ y Σ elipse χ^2 del $(1-\alpha_1)\,\%$ y $(1-\alpha_2)\,\%$

Elipses con GGplot2

```
mu=c(2,2)
sigma<-matrix(c(5,4,4,5),ncol=2)
set.seed(1)
datos2<-as.data.frame(mvrnorm(n=100, mu,sigma))
ggplot(datos2, aes(x = datos2[,1], y = datos2[,2]))  +
  geom_point() +
  stat_ellipse(type = "norm") +
  ggtitle("Elipses de Confianza") +
  xlab("X") + ylab("Y") +
  theme(plot.title = element_text(color="red", size=14, fac
```


Elipses de Confianza
Y
7.5
5.0
2.5
0.0
−2.5
−2.5
0.0
2.5
5.0
7.5
X

```
mu=c(2,2)
sigma<-matrix(c(5,4,4,5),ncol=2)
set.seed(1)
datos2<-as.data.frame(mvrnorm(n=100, mu,sigma))
ggplot(datos2, aes(x = datos2[,1], y = datos2[,2])) +
    geom_point() +
stat_ellipse(type = "norm",level = 0.9) +
  stat_ellipse(type = "norm",level = 0.95, color = 2) +
  stat_ellipse(type = "norm",level = 0.99, color = 3) +
  ggtitle("Elipses de Confianza") +
  xlab("X") + ylab("Y") +
  theme(plot.title = element_text(color="red", size=14, fac
```


Elipses de Confianza
Y
X
5
0
0
5

Inferencia para la Matriz de Covarianza

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 $\underline{\boldsymbol{\mu}}$-desconocida y matriz de varianzas-covarianzas $\boldsymbol{\Sigma}$-Desconocida, ie. $\underline{\boldsymbol{x}_i} \sim N_p(\underline{\boldsymbol{\mu}}\ ,\ \boldsymbol{\Sigma})$.

Se tiene interés en la siguiente PH:

$$\begin{cases} H_0 & : \ \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0 \\ H_a & : \ \boldsymbol{\Sigma} \neq \boldsymbol{\Sigma}_0 \end{cases}$$

donde, $\boldsymbol{\Sigma}_0$-es una matriz de valores fijos conocida para $\boldsymbol{\Sigma}$ (usualmente planteada por experiencia previa).

Dependiendo de la forma particular de $\boldsymbol{\Sigma}_0$, existen distintos nombres para la PH asociada:

1. $\boldsymbol{\Sigma}_0$-Cualquier matriz de valor fijo. Prueba general.
2. $\boldsymbol{\Sigma}_0 = \Delta$-Diagonal. Prueba de independencia de variables.
3. $\boldsymbol{\Sigma}_0 = \sigma^2 \boldsymbol{I}_p$. Prueba de homocedasticidad e independencia.
4. $\boldsymbol{\Sigma}_0 = \boldsymbol{I}_p$. Prueba de Esfericidad, ie. variables con varianzas unitarias e incorreladas.
5. $\boldsymbol{\Sigma}_0 = \boldsymbol{B}_m + \sigma^2 \boldsymbol{I}_p$, con $\boldsymbol{B}$-de rango $m < p$. Prueba de homocedasticidad e independencia parcial. Con $m = 0$-se tiene la rueba de homocedasticidad e independencia.
6. $\boldsymbol{\Sigma}_0 = \boldsymbol{B}_m + \boldsymbol{I}_p$, con $\boldsymbol{B}$-de rango $m < p$. Prueba de Esfericidad parcial. Con $m = 0$-se tiene la Prueba de Esfericidad.

La estadística de Razón de Verosimilitud para la PH considerada está dado por:

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

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

Luego de realizar los cálculos y simplificaciones necesarias se obtiene que:

$$\lambda = \left[\left(\frac{n-1}{n}\right)^{p} \frac{|\boldsymbol{S}|}{|\boldsymbol{\Sigma}_0|}\right]^{\frac{n}{2}} \text{Exp}\left\{-\frac{1}{2}\left[(n-1)\text{tr}(\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}) - np\right]\right\}$$

haciendo: $\frac{n-1}{n} \approx 1$, ie. $n-1 \approx n = v$, ie. $v = n-1 = n$, se tiene:

$$\lambda = \frac{|\boldsymbol{S}|^{\frac{v}{2}}}{|\boldsymbol{\Sigma}_0|^{\frac{v}{2}}} \text{Exp}\left\{-\frac{1}{2}\left[v\text{tr}(\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}) - vp\right]\right\}$$

y haciendo $\lambda^\star = -2\log\lambda$, se tiene que:

$$\lambda^\star = v\left[\text{Log}|\boldsymbol{\Sigma}_0| - \text{Log}|\boldsymbol{S}| + \text{tr}(\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}) - p\right]$$

Bajo H_0-cierta, se tiene que:

$$\lambda^\star \sim \chi^2_k \ , \quad \textit{para} \quad n-1 \ \text{ grande}$$

con $k = p + \frac{p(p+1)}{2} - p = \frac{p(p+1)}{2}$, ie.

k=Parámetros en Θ menos parámetros en Θ_0 (ie. menos p).

Rechazamos H_0 si.

$$\lambda^{\star} > \chi^2_{\alpha\ ;\ k}$$

Ahora, otra forma alterna de $\lambda^{\star}$ es como sigue:

Si $\lambda_1, \lambda_2, \ldots, \lambda_p$ son los valores propios de $\boldsymbol{S\Sigma_0^{-1}}$, entonces se sabe que:

$$\operatorname{tr}\left(\boldsymbol{S\Sigma_0^{-1}}\right) = \sum_{i=1}^{p} \lambda_i$$

y usando propiedades de determinantes se tiene que:

$$\operatorname{Log}|\boldsymbol{\Sigma}_0| - \operatorname{Log}|\boldsymbol{S}| = -\operatorname{Log}|\boldsymbol{S\Sigma_0^{-1}}| = -\operatorname{Log}\left(\prod_{i=1}^{p} \lambda_i\right),$$

luego,

$$\lambda^{\star} = v\left[\mathrm{Log}|\boldsymbol{\Sigma}_0| - \mathrm{Log}|\boldsymbol{S}| + \mathrm{tr}(\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}) - p\right]$$

$$= v\left[-\mathrm{Log}\left(\prod_{i=1}^{p}\lambda_i\right) + \sum_{i=1}^{p}\lambda_i - p\right]$$

$$\lambda^{\star} = v\left[\sum_{i=1}^{p}[\lambda_i - \mathrm{Log}\lambda_i] - p\right] \sim \chi_k^2$$

y rechazamos H_0 si $\lambda^{\star} > \chi^2_{\alpha\ ;\ k}$

con $k = \frac{p(p+1)}{2}$.

Una Modificación para $\lambda^{\star}$-fue propuesta por Bartlet, (**para el caso de muestras pequeñas**) la cual es:

$$\lambda_1^{\star} = \left\{1 - \frac{1}{6(n-1)}\left[2p + 1 - \frac{2}{p+1}\right]\right\}\lambda^{\star} \sim \chi_k^2$$

es decir,

$$\lambda_1^{\star} = c\lambda^{\star} \sim \chi_k^2$$

con

$$c = 1 - \frac{1}{6(n-1)}\left[2p + 1 - \frac{2}{p+1}\right]$$

que puede usarse para tamaños de muestras moderadamente pequeños.

EJEMPLO: Se tomaron 20 sujetos y se les midió los tiempos de reacción ante un estímulo en centésimas de segundo. A cada individuo se le midieron estos tiempos en 3 intervalos de tiempos distintos. Se asume que estas mediciones tienen una distribución $N_3(\underline{\boldsymbol{\mu}}\ ,\ \boldsymbol{\Sigma})$. Pruebe la hipótesis:

$$H_0\ :\ \boldsymbol{\Sigma} = \begin{bmatrix} 4 & 3 & 2 \\ 3 & 6 & 5 \\ 2 & 5 & 10 \end{bmatrix} \quad v.s \quad H_a\ :\ \boldsymbol{\Sigma} \neq \begin{bmatrix} 4 & 3 & 2 \\ 3 & 6 & 5 \\ 2 & 5 & 10 \end{bmatrix}$$

Solución: Como

$$\boldsymbol{\Sigma}_0 = \begin{bmatrix} 4 & 3 & 2 \\ 3 & 6 & 5 \\ 2 & 5 & 10 \end{bmatrix}\ ,\quad \text{luego} \quad \boldsymbol{\Sigma}_0^{-1} = \begin{bmatrix} 0.41 & -0.23 & 0.03 \\ -0.23 & 0.42 & -0.16 \\ 0.03 & -0.16 & 0.17 \end{bmatrix}$$

$$\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1} = \begin{bmatrix} 0.85 & -0.01 & 0.03 \\ -0.58 & 1.68 & -0.08 \\ -0.41 & 0.72 & 0.68 \end{bmatrix}.$$

Los valores propios de $\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}$ son: $\lambda_1 = 1.61$, $\lambda_2 = 0.87$ y $\lambda_3 = 0.73$, luego: $\mathrm{tr}(\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}) = \sum \lambda_i = 3.2216$, $|\boldsymbol{\Sigma}_0| = 86$, $|\boldsymbol{S}| = 88.6355$, de donde:

$$\lambda^\star = v\left[\mathrm{Log}|\boldsymbol{\Sigma}_0| - \mathrm{Log}|\boldsymbol{S}| + \mathrm{tr}(\boldsymbol{S}\boldsymbol{\Sigma}_0^{-1}) - p\right]$$

$$= 20\,[\mathrm{Log}(86) - \mathrm{Log}(88.6355) + 3.2216 - 3]$$

$$\lambda^\star = 3.63$$

$$\lambda_1^\star = \left\{1 - \frac{1}{6(n-1)}\left[2p + 1 - \frac{2}{p+1}\right]\right\}\lambda^\star$$

$$= \left\{1 - \frac{1}{6(19)}\left[2(3) + 1 - \frac{2}{3+1}\right]\right\}3.63$$

$$\lambda_1^\star = 3.42$$

Para $\alpha = 0.05$, se tiene que: $\chi_{\alpha;v} = \chi_{0.05;6} = 12.592$.

En este caso, $v = p(p+1)/2 = 3(4)/2 = 6$, y como:

$$\lambda_1^\star = 3.42 < 12.592 = \chi^2_{\alpha;v},$$

,

entonces, no se rechaza H_0 y se concluye que la evidencia muestral apoya la hipótesis:

$$H_0 \ : \ \mathbf{\Sigma} = \begin{bmatrix} 4 & 3 & 2 \\ 3 & 6 & 5 \\ 2 & 5 & 10 \end{bmatrix},$$

a un nivel de significancia del 5%.

Ejemplo Usando R. Para n-grande

Se desea contrstar las hipótesis:

$$H_0 \ : \ \boldsymbol{\Sigma} = \begin{bmatrix} 6.5 & -0.05 & -2.5 \\ -0.05 & 5.5 & 0.5 \\ -2.5 & 0.5 & 7.5 \end{bmatrix}$$

a un nivel de significancia del 5%.

Resultados usando la función de usuario: `sigma_sigma0_ngrande`

```
Sigma_0<-matrix(c(6.5, -0.05, -2.5,-0.05, 5.5, 0.5, -2.5 ,0.5, 7.5),byrow=TRUE,
res_sigma0<-sigma_sigma0_ngrande(grupo1[,1:3],Sigma_0,0.05)
kable(res_sigma0)
```

Lamda_est	df	Chi_Tabla	Valor_P
21.7799	6	12.5916	0.00132724

Ejemplo Usando R. Para n-pequeña (Modificación de Bartlet)

Resultados usando la función de usuario: `sigma_sigma0_npqna`

```
Sigma_0<-matrix(c(6.5, -0.05, -2.5,-0.05, 5.5, 0.5, -2.5 ,0.5, 7.5),byrow=TRUE,
res_sigma0npqna<-sigma_sigma0_npqna(grupo1[,1:3],Sigma_0,0.05)
kable(res_sigma0npqna)
```

Lamda1_est	c	df	Chi_Tabla	Valor_P
20.872	0.95833	6	12.592	0.001934

Dos o más Matrices de Covarianzas

Recordar que uno de los supuestos cuando se comparan dos o más vectores de medias, es que las respectivas matrices de Var-Cov asociadas a cada población diferente, sean iguales. Un test muy común para probar la igualdad de matrices de Var-Cov es el M-Test de Box.

Suponga que se tienen g-poblaciones diferentes, con matrices de Var-Cov asociadas dadas por: $\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \ldots, \boldsymbol{\Sigma}_g$, respectivamente. Se desea contrastar las hipótesis:

$$\begin{cases} H_0 & : \ \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \cdots = \boldsymbol{\Sigma}_g = \boldsymbol{\Sigma} \\ H_a & : \ \boldsymbol{\Sigma}_i \neq \boldsymbol{\Sigma}_j \quad p.a \quad i, j \end{cases}$$

donde $\boldsymbol{\Sigma}$-es una matriz de Var-Cov común.

Si se tienen g-muestras aleatorias, una para cada población, de tamaños $n_1, n_2, \ldots, n_g$, respectivamente y además se asume que cada muestra proviene de una población con distribución normal p-variada, entonces el estadístico de prueba de razón de verosimilitud para esta PH es:

$$\lambda = \prod_{i=1}^{g} \left(\frac{|\boldsymbol{S}_i|}{|\boldsymbol{S}_p|} \right)^{\frac{n_i - 1}{2}},$$

donde, $\boldsymbol{S}_i$-es la matriz de Var-Cov muestral asociada a la m.a de la i-ésima población, $i = 1, 2, \ldots, g$ y

$$\boldsymbol{S}_p = \frac{1}{\sum_{i=1}^{g} (n_i - 1)} \left[\sum_{i=1}^{g} (n_i - 1) \boldsymbol{S}_i \right]$$

haciendo, $\{v_i = n_i - 1\}$ y $\{v = \sum_{i=1}^{g} v_i = \sum_{i=1}^{g}(n_i - 1)\}$, se tiene que:

$$\boldsymbol{S}_p = \frac{1}{v}\left[\sum_{i=1}^{g} v_i \boldsymbol{S}_i\right]$$

La **Estadística M de Box** se define como:
$M = -2\text{Log}\lambda \sim \chi^2$, (n-grande).

$$M = \left[\sum_{i=1}^{g}(n_i - 1)\right]\text{Log}|\boldsymbol{S}_p| - \sum_{i=1}^{g}(n_i - 1)\text{Log}|\boldsymbol{S}_i|$$

$$= v\text{Log}|\boldsymbol{S}_p| - \sum_{i=1}^{g} v_i \text{Log}|\boldsymbol{S}_i|$$

Bajo H_0-cierto, se espera que las matrices de Var-Cov muestrales no sean muy diferentes, en cuyo caso, el valor de λ estaría cerca a uno y por lo tanto M-sería pequeño.

Ahora, sea

EJEMPLO: En el departamento de salud y servicios sociales de Wisconsin se realizó un estudio para investigar el efecto de la propiedad o la certificación (o ambas) sobre los costos. Cuatro costos fueron seleccionados para el análisis; estos fueron calculados diariamente por paciente y fueron medidos en horas por paciente diario. Las variables fueron: X_1-Costo de la enfermería, X_2-Costo de alimentación, X_4-Costo de operación y mantenimiento y X_4-Costo de administración y lavandería. Se registraron un total de $n = 516$-observaciones de cada una de las cuatro variables, separadas previamente en tres grupos de interés: Privados ($n_1 = 271$), Públicos ($n_2 = 138$) y Gubernamentales ($n_3 = 107$). Se asume que el vector $\underline{\boldsymbol{x}} = (X_1, X_2, X_3, X_4)^t$ tiene una distribución $N_4(\underline{\boldsymbol{\mu}_i}\ ,\ \boldsymbol{\Sigma}_i)$, para $i = 1, 2, 3$. Se desea probar la hipótesis:

$$\begin{cases} H_0 \ : \ \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_3 = \boldsymbol{\Sigma} \\ H_a \ : \ \boldsymbol{\Sigma}_i \neq \boldsymbol{\Sigma}_j \ \ p.a \ \ i \neq j = 1, 2, 3 \end{cases}$$

$i = 1$ Privados

$$n_1 = 271 \ , \quad \bar{\mathbf{x}}_1 = \begin{pmatrix} 2.066 \\ 0.480 \\ 0.082 \\ 0.36 \end{pmatrix} \quad , \quad S_1 = \begin{pmatrix} 0.291 & & & \\ -0.01 & 0.011 & & \\ 0.02 & 0.000 & 0.001 & \\ 0.010 & 0.003 & 0.000 & 0.10 \end{pmatrix}$$

$i = 2$ Públicos

$$n_2 = 138 \ , \quad \bar{\mathbf{x}}_1 = \begin{pmatrix} 2.167 \\ 0.596 \\ 0.124 \\ 0.418 \end{pmatrix} \quad , \quad S_2 = \begin{pmatrix} 0.561 & & & \\ 0.011 & 0.025 & & \\ 0.001 & 0.004 & 0.005 & \\ 0.037 & 0.007 & 0.002 & 0.019 \end{pmatrix}$$

$i = 3$ Gubernamentales

$$n_3 = 107 \ , \quad \bar{\mathbf{x}}_1 = \begin{pmatrix} 2.273 \\ 0.521 \\ 0.125 \\ 0.383 \end{pmatrix} \quad , \quad S_3 = \begin{pmatrix} 0.261 & & & \\ 0.030 & 0.017 & & \\ 0.003 & 0.000 & 0.004 & \\ 0.018 & 0.006 & 0.001 & 0.013 \end{pmatrix}$$

Con la información anterior se obtiene:

$$|\boldsymbol{S}_1| = 2.783 \times 10^{-8} \ , \qquad Log|\boldsymbol{S}_1| = -17.397$$

$$|\boldsymbol{S}_2| = 89.539 \times 10^{-8} \ , \qquad Log|\boldsymbol{S}_2| = -13.926$$

$$|\boldsymbol{S}_3| = 14.579 \times 10^{-8} \ , \qquad Log|\boldsymbol{S}_3| = -15.741$$

$$|\boldsymbol{S}_p| = 17.398 \times 10^{-8} \ , \qquad Log|\boldsymbol{S}_1| = -15.564$$

$$u = \left[\frac{1}{270} + \frac{1}{137} + \frac{1}{106} - \frac{1}{270+137+106}\right]\left[\frac{2(4)^2+3(4)-1}{6(4+1)(3-1)}\right] = 0.0133$$

$$= (270+137+106)\,(-15.564) - [270(-17.397) + 137(-13.926) + 106(-15.741)] = 289.3$$

$$C = (1-u)M = (1-0.0133)289.3 = 285.5$$

Ahora, para $\alpha = 0.05$, se tiene que:

$$\chi^2_{\alpha;k} = \chi^2_{0.05\ ;\ \frac{p(p+1)}{2}(g-1)} = \chi^2_{0.05\ ;\ 20} = 31.34,$$

Como $C > \chi^2_{\alpha;k}$, entonces se rechaza H_0 y se concluye que las matrices de Var-Cov asociadas a los vectores de costos para las las tres poblaciones consideradas son diferentes a un nivel de significancia del 5%.

Ejemplo Usando R. Prueba M de Box (n-grande)

Resultados de esta PH usando la función de usuario: `prueba_M_Box2()`

```
x<-grupo1[,1:3]
y<-grupo2[,1:3]
res_box<-prueba_M_Box2(x,y,0.05)
kable(res_box)
```

M	U	C	df	Chi_Tabla	Valor_p
4.16033	0.0613499	3.9051	6	12.5916	0.689518

Resultados de esta PH utilizando la función `boxM` del paquete `biotools` del R.

```r
mbox<-boxM(datos[1:55,1:3],datos$Grupos[1:55])
mbox
```

```
##
##  Box's M-test for Homogeneity of Covariance Matrices
##
## data:  datos[1:55, 1:3]
## Chi-Sq (approx.) = 3.9051, df = 6, p-value = 0.6895
```

Estadístico$= C \sim \chi^2$	gl	p-Valor
3.9051	6	0.68952