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 $\mathbf{\Sigma}$.

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:

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

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

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:

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

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%.