BAB III

DISTRIBUSI WISHART DAN HOTELLING'S T^2

3.1. Distribusi Wishart

Definisi 33

Jika Z_j berdistribusi $N_p(\mathbf{0}, \boldsymbol{\Sigma})$, dengan j= 1,2, . . . , n maka $\sum_{i=1}^{n} \mathbf{Z}_j \mathbf{Z}_j^T$

berdistribusi Wishart dengan derajat bebas n, dan ditunjukkan dengan :
$W_n(. \mid \boldsymbol{\Sigma})$.

Teorema 11

Jika $\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_n$ independen masing-masing berdistribusi $N_p(\mathbf{0}, \boldsymbol{\Sigma})$ dan

$A = \sum_{j=1}^{n} Z_j Z_j^T$ maka fungsi karakteristik A adalah :

$$\phi(\mathrm{A}) = |\, \mathbf{I} - 2i\theta\, \Sigma \,|^{-n/2}$$

Bukti :

Diberikan matriks simetri $\theta_{p \times p}$, maka

$$\phi(\mathrm{A}) = E[\exp(i \operatorname{tr}(A\theta))]$$

$$= E\left[\exp\left(i \operatorname{tr}\left(\sum_{j=1}^{n} \mathbf{Z}_j \mathbf{Z}_j^T \theta\right)\right)\right]$$

$$= E\left[\exp\left(i \operatorname{tr}\left(\sum_{j=1}^{n} \mathbf{Z}_j^T \theta \mathbf{Z}_j\right)\right)\right]$$

$$= E\left[\exp\left(i \sum_{j=1}^{n} \mathbf{Z}_j^T \theta \mathbf{Z}_j\right)\right]$$

$$= \prod_{j=1}^{n} E\left[\exp\left(i\ Z_j^T\,\theta\ Z_j\right)\right]$$

$$= \left[E\left[\exp\left(i\ Z_j^T\,\theta\ Z_j\right)\right]\right]^n$$

Selanjutnya jika **B** adalah matriks nonsingular sedemikian hingga $\mathbf{B}^T \boldsymbol{\Sigma}^{-1} \mathbf{B} = \mathbf{I}$ dan $\mathbf{B}^T \boldsymbol{\theta} \mathbf{B} = \mathbf{D}$ dimana **D** adalah matriks diagonal dengan elemen - elemen pada diagonalnya adalah d_{jj} dan $\mathbf{Z} = \mathbf{B}\,\mathbf{Y}$, maka :

$$\begin{aligned}
E\left[\exp\left(i\,\mathbf{Z}^T \boldsymbol{\theta}\,\mathbf{Z}\right)\right] &= E\left[\exp\left(i\,(\mathbf{B}\,\mathbf{Y})^T \boldsymbol{\theta}\,\mathbf{B}\,\mathbf{Y}\right)\right] \\
&= E\left[\exp\left(i\,\mathbf{Y}^T \mathbf{B}^T \boldsymbol{\theta}\,\mathbf{B}\,\mathbf{Y}\right)\right] \\
&= E\left[\exp\left(i\,\mathbf{Y}^T \mathbf{D}\,\mathbf{Y}\right)\right] \\
&= E\left[\exp\left(i \sum_{j=1}^{p} d_{jj}\ y_j^2\right)\right] \\
&= E\left[\prod_{j=1}^{p} \exp\left(i\ d_{jj}\ y_j^2\right)\right] \\
&= \prod_{j=1}^{p} E\left[\exp\left(i\ d_{jj}\ y_j^2\right)\right]
\end{aligned}$$

dimana y_j independen dengan distribusi N (0,1) dan y_j^2 berdistribusi χ_1^2 maka :

$$\prod_{j=1}^{p} E\left[\exp\left(i\ d_{jj}\ y_j^2\right)\right] = \prod_{j=1}^{p} \frac{1}{\sqrt{1-2id_{jj}}}$$

$$\begin{aligned}
\prod_{j=1}^{p} E\left[\exp\left(i\ d_{jj}\ y_j^2\right)\right] &= \prod_{j=1}^{p} \left(1-2id_{jj}\right)^{-1/2} \\
&= \left|\mathbf{I} - 2\,i\,\mathbf{D}\right|^{-\frac{1}{2}}
\end{aligned}$$

Karena **I** - 2 i **D** adalah matriks diagonal dan $B^T \Sigma^{-1} B = I$ dan $B^T \theta B = D$ maka :

$$
\begin{aligned}
| I - 2iD | &= | B^T \Sigma^{-1} B - 2i B^T \theta B | \\
&= | B^T (\Sigma^{-1} - 2i\theta) B | \\
&= | B^T | \, | (\Sigma^{-1} - 2i\theta) | \, | B | \\
&= | B |^2 \, | (\Sigma^{-1} - 2i\theta) |
\end{aligned}
$$

Karena $| B^T | \, | \Sigma^{-1} | \, | B | = | I | = 1$ maka $| B |^2 = \dfrac{1}{| \Sigma^{-1} |}$ sehingga

$$
\begin{aligned}
E[\exp[i \operatorname{tr}(A\theta)]] &= \left[\frac{1}{|\Sigma^{-1}|} |\Sigma^{-1} - 2i\theta| \right]^{-n/2} \\
&= \frac{|\Sigma^{-1}|^{n/2}}{|\Sigma^{-1} - 2i\theta|^{n/2}} = \left| \frac{\Sigma^{-1}}{\Sigma^{-1} - 2i\theta} \right|^{n/2} \\
&= \left| \frac{I}{I - 2i\theta\Sigma} \right|^{n/2} = \frac{| I |^{n/2}}{| I - 2i\theta\Sigma |^{n/2}} \\
&= \frac{1}{| I - 2i\theta\Sigma |^{n/2}} = | I - 2i\theta\Sigma |^{-n/2}
\end{aligned}
$$

Terbukti bahwa fungsi karakteristik dari A adalah $| I - 2i\theta\Sigma |^{-n/2}$. ■

Proposisi 8

Jika A_1 berdistribusi $W_{m1}(A_1 | \Sigma)$ dan A_2 berdistribusi $W_{m2}(A_2 | \Sigma)$ maka $A_1 + A_2$ berdistribusi $W_{m1+m2}(A_1 + A_2 | \Sigma)$ dengan derajat bebas $m_1 + m_2$.

Bukti :

Dibuktikan dengan menggunakan fungsi karakteristik pada teorema 11 , diperoleh:

$$\phi (A_1 + A_2) = E [\exp [i \operatorname{tr} (\theta (A_1 + A_2))]]$$

$$= E\left[\exp\left(i \operatorname{tr}\left(\theta \sum_{i=1}^{2} A_i\right)\right)\right]$$

$$= E\left[\exp\left(i \sum_{i=1}^{2} \operatorname{tr}\left(\theta A_i\right)\right)\right]$$

$$= \prod_{i=1}^{2} E\left[\exp\left(i \operatorname{tr}\left(\theta A_i\right)\right)\right]$$

$$= \prod_{i=1}^{2} | I - 2i\,\theta\Sigma |^{-m_i/2}$$

$$= | I - 2i\,\theta\Sigma |^{-m_1/2} \; | I - 2i\,\theta\Sigma |^{-m_2/2}$$

$$= | I - 2i\,\theta\Sigma |^{-\frac{1}{2}(m_1 + m_2)}$$

Bentuk khusus dari fungsi karakteristik diatas menunjukkan bahwa $A_1 + A_2$ berdistribusi Wishart dengan derajat bebas $m_1 + m_2$ dan ditulis dengan $W_{m_1+m_2}(A_1 + A_2 \mid \Sigma)$. ■

Untuk bentuk $(n-1) S = \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})^T$ dimana x_i dan $\bar{x}$ adalah vektor random dengan dimensi p x 1 , maka $(n - 1) S$ berdistribusi identik dengan

$I (Z_1 Z_1^T + Z_2 Z_2^T + \ldots + Z_{n-1} Z_{n-1}^T)$ dimana :

$Z_i = (x_i - \mu) / \Sigma$ berdistribusi $N_p(0, \Sigma)$, sehingga :

$$(n-1)\ S = I \sum_{i=1}^{n-1} Z_i\ Z_i^T \qquad (15)$$

dimana I adalah matriks identitas dengan dimensi p x p dan sesuai dengan definisi 33, maka $(n-1)\ S$ berdistribusi Wishart dengan derajat bebas n - 1.

3.2 Hotelling's T^2

3.2.1. Sifat Optimal Invarian Pada Uji T^2

Pada analisis *univariate,* uji t terhadap $\mu = 0$ adalah *invarian* berkaitan dengan transformasi skala. Jika skalar variabel random X berdistribusi menurut $N(\mu, \sigma^2)$ maka:

$X^* = cX$ berdistribusi menurut $N(c\mu, c^2\sigma^2)$ sehingga hipotesa $\mathcal{E}\ X = 0 \Leftrightarrow \mathcal{E}\ X^* = \mathcal{E}\ cX = 0$, observasi x_α ditransformasi secara sama, $x_\alpha^* = c\,x_\alpha$ dengan $c > 0$, t^* dihitung dari x_α^* , hal ini sama dengan nilai t yang dihitung dari x_α . Jadi apapun unit pengukurannya hasil statistiknya adalah sama.

Pada analisis *multivariate*, uji T^2 secara umum mempunyai sifat sama, yaitu jika vektor random variabel **X** berdistribusi $N(\mu, \Sigma)$, kemudian $\mathbf{X}^* = \mathbf{C}\ \mathbf{X}$ (untuk $|\mathbf{C}| \neq 0$), maka variabel vektor random tersebut berdistribusi menurut $N(\mathbf{C}\mu, \mathbf{C}\Sigma\mathbf{C}^T)$. Hipotesa $\mathcal{E}\ \mathbf{X} = \mathbf{0} \Leftrightarrow \mathcal{E}\ \mathbf{X}^* = \mathcal{E}\ \mathbf{CX} = \mathbf{0}$. Jika x_α ditransformasi dengan cara yang sama yaitu :

$x_\alpha^* = \mathbf{C}x_\alpha$ maka T^{*2} dihitung berdasarkan x_α^* dan hasilnya adalah sama dengan T^2 yang dihitung berdasarkan pada x_α , hal ini mengikuti dari $\bar{x}^* = \mathbf{C}\bar{x}$ dan $\mathbf{A}^* = \mathbf{C}\ \mathbf{A}\ \mathbf{C}^T$

Contoh :

Diberikan bentuk :

$$\mathbf{Y}_{pxl} = \mathbf{C}_{pxp}\ \mathbf{x}_{pxl} + \mathbf{d}_{pxl}$$ C nonsingular

Maka sesuai dengan sifat ekspektasi pada definisi 21

$\bar{\mathbf{y}} = \mathbf{C}\bar{\mathbf{x}} + \mathbf{d}$ dan berdasarkan definisi 25 serta teorema 5 maka :

$$\mathbf{S}_y = \frac{1}{n-1}\sum(\mathbf{y}_j-\bar{\mathbf{y}})(\mathbf{y}_j-\bar{\mathbf{y}})^T \qquad = \mathbf{CSC}^T$$

Sehingga diperoleh :

$$\mu_Y = E(\ \mathbf{Y})$$
$$= E(\ \mathbf{Cx} + \mathbf{d}\ )$$
$$= E(\ \mathbf{Cx}\ ) + E(\ \mathbf{d}\ )$$
$$= \mathbf{C}\mu + \mathbf{d}$$
$$\mu_{Y0} = \mathbf{C}\ \mu_0 + \mathbf{d}$$

Maka :

$$T^2 = n(\ \bar{\mathbf{y}} - \mu_{Y0}\ )^T\mathbf{S}_y^{-1}(\ \bar{\mathbf{y}} - \mu_{Y0}\ )$$

$$= n\ (\ \mathbf{C}\bar{\mathbf{x}} + \mathbf{d} - (\ \mathbf{C}\mu_0 + \mathbf{d}\ ))^T(\ \mathbf{CSC}^T)^{-1}(\ \mathbf{C}\bar{\mathbf{x}} + \mathbf{d} - (\ \mathbf{C}\mu_0 + \mathbf{d}\ )$$

$$= n(\bar{\mathbf{x}} - \mu_0)^T\ \mathbf{C}^T(\ \mathbf{C}^T)^{-1}\ \mathbf{S}^{-1}\ \mathbf{C}^{-1}\mathbf{C}\ (\bar{\mathbf{x}} - \mu_0)$$

$$= n\ (\bar{\mathbf{x}} - \mu_0)^T\ \mathbf{S}^{-1}\ (\bar{\mathbf{x}} - \mu_0)$$

3.2.2. Hotelling's T^2 untuk Uji Rata - Rata Satu Populasi Normal

Pada teori *univariate*, untuk memutuskan apakah sebuah harga tertentu μ_0 dapat diterima sebagai harga rata - rata populasi μ , dalam hal ini tinjauan uji hipotesa dapat dirumuskan sebagai : $H_0 : \mu = \mu_0$ melawan $H_1 : \mu \neq \mu_0$, statistik yang biasa digunakan adalah perbedaan antara rata - rata sampel $\bar{x}$ dan hipotesa rata - rata populasi μ_0 dibagi dengan standar deviasi sampel s. Jika x_1 , x_2 , ...x_n adalah sampel random dari populasi berdistribusi normal $N(\mu, \Sigma^2)$ maka digunakan :

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

sesuai dengan proposisi 6 statistik ini berdistribusi t dengan derajat bebas n-1

$$t^2 = \frac{(\bar{x} - \mu_0)^2}{s^2 / n}$$

$$t^2 = n(\bar{x} - \mu_0)(s^2)^{-1}(\bar{x} - \mu_0)$$

dimana $\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$ dan $s^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_j - \bar{x})^2$

Selanjutnya dalam masalah *Multivariate* jika px1 vektor μ_0 adalah harga yang mungkin untuk rata - rata populasi berdistribusi normal multivariate, analog dengan teori *univariate* diatas maka:

$$T^2 = n(\mathbf{X} - \mu_0)^T (\mathbf{S}^{-1})(\mathbf{X} - \mu_0) \qquad (16)$$

T^2 berdistribusi $\frac{(n-1)p}{(n-p)} F_{p,\,n-p}$ (17)

adalah variabel random dengan distribusi F dan derajat kebebasan p dan n-p.

Dimana :

$$\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j \qquad S_{(p\times p)} = \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \bar{X})(X_j - \bar{X})^T \qquad \mu_0 = \begin{bmatrix} \mu_{10} \\ \mu_{20} \\ \cdot \\ \cdot \\ \cdot \\ \mu_{p0} \end{bmatrix}$$

Statistik T^2 ini disebut Hotellling's T^2 .

Selanjutnya jika diberikan $X_1, X_2, \ldots, X_n$ sampel random dari populasi $N_p(\mu, \Sigma)$ dengan :

$$\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$$

$$S = \frac{1}{(n-1)}\sum_{j=1}^{n}(X_j - \bar{X})(X_j - \bar{X})^T$$

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

$$= P\left[n(\bar{X}-\mu)^T S^{-1}(\bar{X}-\mu) > \frac{(n-1)p}{(n-p)} F_{p,\,n-p}(\alpha)\right] \qquad (18)$$

maka uji hipotesa $H_0 : \mu = \mu_0$ melawan $H_1 : \mu \neq \mu_0$, tolak H_0 pada taraf signifikansi α Jika :

$$T^2 = n\,(\bar{X}-\mu_0)^T S^{-1} (\bar{X}-\mu_0) \;>\; \frac{(n-1)p}{(n-p)} F_{p,\,n-p}(\alpha) \qquad (19)$$

$$T^2 = \sqrt{n}\,(\bar{X}-\mu_0)^T \left(\frac{\sum_{j=1}^{n} (X_j-\bar{X})(X_j-\bar{X})^T}{n-1} \right) \sqrt{n}(\bar{X}-\mu_0) \qquad (20)$$

Keterangan :

$\sqrt{n}\,(\bar{X}-\mu_0)^T =$ Vektor random normal *multivariate*

$\left(\frac{\sum_{j=1}^{n} (X_j-\bar{X})(X_j-\bar{X})^T}{n-1} \right) =$ Matriks random wishart / derajat bebas

$\sqrt{n}\,(\bar{X}-\mu_0) =$ vektor random normal *multivariate*

Untuk melihat permasalahan ini maka diberikan contoh data matriks untuk sampel random ukuran n=3 dari populasi normal *bivariate.*

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

Bagaimana distribusi sampling T^2 ? Jika diketahui $\mu_0{}^T = [9\ 5]$
jawab : berdasarkan (18) :

$$\bar{X} = \begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \end{bmatrix} = \begin{bmatrix} (6+10+8)/3 \\ (9+6+3)/3 \end{bmatrix} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}$$

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

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

$$S_{12} = S_{21}$$

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

$$\mathbf{S} = \begin{bmatrix} 4 & -3 \\ -3 & 9 \end{bmatrix}$$

$$\mathbf{S}^{-1} = \frac{1}{4(9)-(-3)(-3)} \begin{bmatrix} 9 & 3 \\ 3 & 4 \end{bmatrix} = \frac{1}{27} \begin{bmatrix} 9 & 3 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{9} \\ \frac{1}{9} & \frac{4}{27} \end{bmatrix}$$

$$T^2 = 3 \begin{bmatrix} 8-9 & 6-5 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{1}{9} \\ \frac{1}{9} & \frac{4}{27} \end{bmatrix} \begin{bmatrix} 8-9 \\ 6-5 \end{bmatrix} = 7/9$$

T^2 berdistribusi $\frac{(3-1)2}{(3-2)} F_{2,3-2} = 4F_{2,1} = 198$

3.2.3. Hotelling's T^2 Untuk Membandingkan Vektor Rata - Rata Dari 2(dua) Populasi Normal

Statistik T^2 untuk uji persamaan vektor rata - ratadari 2 (dua) populasi multivariate, dapat dibangun dengan prosedur univariate. Misalnya sampel random ukuran n_1 dari populasi 1 dan sampel random n_2 dari populasi 2, observasi p variabel dapat dituliś sebagai berikut :

Populasi 1 :

x_{11} , x_{12} , ... , x_{1n_1}

$$\bar{x}_1 = \frac{1}{n_1}\sum_{j=1}^{n_1} x_{1j}$$

$$S_1 = \frac{1}{n_1-1}\sum_{j=1}^{n_1} (x_{1j} - \bar{x}_1)(x_{1j} - \bar{x}_1)^T$$

Populasi 2 :

x_{21} , x_{22} , ... , x_{2n_2}

$$\bar{x}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2} x_{2j}$$

$$S_2 = \frac{1}{(n_2-1)}\sum_{j=1}^{n_2} (x_{2j} - \bar{x}_2)(x_{2j} - \bar{x}_2)^T$$

Jika ingin dilakukan inferensi tentang (vektor rata - rata populasi 1) - (vektor rata - rata populasi 2) $\mu_1 - \mu_2$ . .Dalam hal ini, misalnya ingin diketahui apakah $\mu_1 = \mu_2$ atau $(\mu_1 - \mu_2) = \mathbf{0}$.

Untuk uji hipotesa ini diperlukan asumsi sebagai berikut :

Asumsi mengenai struktur data :

1. Sampel x_{11} , x_{12} , ... , x_{1n1} , adalah sampel random ukuran n_1 dari populasi p *variate* dengan vektor mean μ_1 dan matriks kovariansi Σ_1 .
2. Sampel x_{21} , x_{22} , ... , x_{2n2} , adalah sampel random ukuran n_2 dari populasi p *variate* dengan vektor mean μ_2 dan matriks kovariansi Σ_2 .
3. x_{11} , x_{12} , ... , x_{1n1} independen dengan x_{21} , x_{22} , ... , x_{2n2} .

Asumsi tersebut diperlukan untuk inferensi vektor $\mu_1 - \mu_2$ dalam hal sampel besar.

Untuk sampel kecil diperlukan asumsi tambahan sebagai berikut :

1. Kedua populasi normal multivariate.
2. $\Sigma_1 = \Sigma_2 = \Sigma$

Jika $\Sigma_1 = \Sigma_2 = \Sigma$ maka :

$\sum_{j=1}^{n_1} (\mathbf{x}_{1j} - \bar{\mathbf{x}}_1)(\mathbf{x}_{1j} - \bar{\mathbf{x}}_1)^T$ adalah penduga untuk $(n_1 - 1)\Sigma$

$\sum_{j=1}^{n_2} (\mathbf{x}_{2j} - \bar{\mathbf{x}}_2)(\mathbf{x}_{2j} - \bar{\mathbf{x}}_2)^T$ adalah penduga untuk $(n_2 - 1)\Sigma$

sehingga keduanya dapat digabung menjadi :

$$\mathbf{S}_{pooled} = \frac{\sum_{j=1}^{n_1} (\mathbf{x}_{1j} - \bar{\mathbf{x}}_1)(\mathbf{x}_{1j} - \bar{\mathbf{x}}_1)^T + \sum_{j=1}^{n_2} (\mathbf{x}_{2j} - \bar{\mathbf{x}}_2)(\mathbf{x}_{2j} - \bar{\mathbf{x}}_2)^T}{n_1 + n_2 - 2} \qquad (21)$$

Keterangan :

$\sum_{j=1}^{n_1} (\mathbf{x}_{1j} - \bar{\mathbf{x}}_1)(\mathbf{x}_{1j} - \bar{\mathbf{x}}_1)^T$ mempunyai n_1-1 derajat bebas dan

$\sum_{j=1}^{n_2} (\mathbf{x}_{2j} - \bar{\mathbf{x}}_2)(\mathbf{x}_{2j} - \bar{\mathbf{x}}_2)^T$ mempunyai $n_2 - 1$ derajat bebas

Analog dengan persamaan (14) maka persamaan (21) dapat pula ditulis dalam bentuk :

$$\mathbf{S}_{pooled} = \frac{(n_1-1)\mathbf{S}_1 + (n_2-1)\mathbf{S}_2}{n_1+n_2-2} \tag{22}$$

Teorema 12

Jika $\mathbf{X}_{11}, \mathbf{X}_{12}, \ldots, \mathbf{X}_{1n_1}$ adalah sampel random ukuran n_1 dari $N_p(\mu_1, \Sigma)$ dan $\mathbf{X}_{21}, \mathbf{X}_{22}, \ldots, \mathbf{X}_{2n_2}$ adalah sampel random independen ukuran n_2 dari $N_p(\mu_2, \Sigma)$ maka:

$$T^2 = [\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2 - (\mu_1 - \mu_2)]^T \left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\mathbf{S}_{pooled}\right]^{-1} [\bar{\mathbf{X}}_1 - \bar{\mathbf{X}}_2 - (\mu_1 - \mu_2)] \tag{23}$$

berdistribusi : $\frac{(n_1+n_2-2)\,p}{(n_1+n_2-p-1)} F_{p,\, n_1+n_2-p-1}(\alpha)$

akibatnya :

$$P[\, T^2 \le c^2 \,] = 1-\alpha \tag{24}$$

dimana :

$$c^2 = \frac{(n_1+n_2-2)\,p}{(n_1+n_2-p-1)} F_{p,\, n_1+n_2-p-1}(\alpha) \tag{25}$$

Bukti :

Jika $\bar{X}_1$ adalah rata-rata dari sampel random ukuran n_1 dan $\bar{X}_2$ adalah rata-rata dari sampel random ukuran n_2 maka :

$$\bar{X}_1 - \bar{X}_2 = \frac{1}{n_1} X_{11} + \ldots + \frac{1}{n_1} X_{1n_1} - \frac{1}{n_2} X_{21} - \ldots - \frac{1}{n_2} X_{2n_2}$$

dimana E ($\bar{X}_1 - \bar{X}_2$) = E ($\bar{X}_1$) - E ($\bar{X}_2$) = $\mu_1 - \mu_1$ dan

$$\begin{aligned} \text{var}(\bar{X}_1 - \bar{X}_2) &= E[(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)]^2 \\ &= E[(\bar{X}_1 - \mu_1') - (\bar{X}_2 - \mu_2)]^2 \\ &= E[(\bar{X}_1 - \mu_1)^2 + (\bar{X}_2 - \mu_2)^2 - 2(\bar{X}_1 - \mu_1)(\bar{X}_2 - \mu_2)] \\ &= \text{var}(\bar{X}_1) + \text{var}(\bar{X}_1) - 2\,\text{cov}(\bar{X}_1, \bar{X}_2). \end{aligned}$$

Karena $\bar{X}_1$ dan $\bar{X}_2$ independen maka:

$$\begin{aligned} \text{cov}(\bar{X}_1, \bar{X}_2) &= E[(\bar{X}_1 - \mu_1)(\bar{X}_2 - \mu_2)] \\ &= E[\bar{X}_1 \bar{X}_2] - \mu_1 \mu_1 \\ &= 0 \text{ sehingga :} \end{aligned}$$

var ($\bar{X}_1 - \bar{X}_2$) = var ($\bar{X}_1$) + var ($\bar{X}_1$) dan berdasarkan proposisi 4 maka :

$\bar{X}_1 - \bar{X}_2$ berdistribusi :

$$N_p\left(\mu_1 - \mu_2, \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\Sigma\right)$$

Selanjutnya diambil :

$$C_1 = C_2 = \ldots = C_{n_1} = \frac{1}{n_1} \quad \text{dan} \quad C_{n_{1+1}} = C_{n_{1+2}} = \ldots = C_{n_1+n_2} = \frac{1}{n_2}$$

Menurut (15) ($n_1 - 1$) S_1 berdistribusi Wn_1-1 (- | Σ) dan ($n_2 - 1$) S_2 berdistribusi Wn_2-1 (-| Σ) dengan asumsi X_{1j} dan X_{2j} independent maka (n_1 - 1)S_1 dan (n_2 - 1) S_2 juga independent . Berdasarkan proposisi 8 ($n_1 - 1$) S_1 + ($n_2 - 1$) S_2 berdistribusi Wn_1+n_2-2 (-| Σ). Oleh karena itu :

$$T^2 = [\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)]^T \left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S_{pooled}\right]^{-1} [\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)]$$

$$= \left(\frac{1}{n_1} + \frac{1}{n_2}\right)^{-1/2} (\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2))^T S_{pooled}{}^{-1} \left(\frac{1}{n_1} + \frac{1}{n_2}\right)^{-1/2} (\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2))$$

$$= N_p(\mathbf{0}, \boldsymbol{\Sigma}) \left[\frac{W_{n_1 + n_2 - 2}(\cdot \mid \boldsymbol{\Sigma})}{n_1 + n_2 - 2}\right]^{-1} N_p(\mathbf{0}, \boldsymbol{\Sigma})$$

dimana :

$\left(\frac{1}{n_1} + \frac{1}{n_2}\right)^{-1/2} (\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2))^T$ = (Vektor random normal *multivariate* $)^T$

$S_{pooled}{}^{-1}$ = (Matriks Random wishart / derajat bebas $)^{-1}$

$\left(\frac{1}{n_1} + \frac{1}{n_2}\right)^{-1/2} (\bar{X}_1 - \bar{X}_2 - \mu_1 - \mu_2))$ = (Vektor random normal *multivariate*)

T^2 ini adalah distribusi khusus dari (16) dengan mengganti n dengan $n_1 + n_2 - 1$.

3.2.4. Hotelling's T^2 Pada Rancangan Pengukuran Berulang Untuk Membandingkan Perlakuan

Secara umum statistik t pada pasangan *univariate* timbul dalam situasi dimana q perlakuan dibandingkan, berkaitan dengan variabel respons tunggal. Setiap subyek atau unit eksperimen menerima masing - masing satu perlakuan, sedangkan semuanya terdapat n unit eksperimen.

Diberikan observasi ke-j :

$$\mathbf{X}_j = \begin{bmatrix} X_{1j} \\ X_{2j} \\ . \\ . \\ . \\ X_{qj} \end{bmatrix} \qquad j = 1, 2, 3, \ldots, n$$

X_{ij} adalah respons perlakuan ke-i terhadap unit ke-j, disebut pengukuran berulang karena semua perlakuan dikenakan pada tiap unit. Didefinisikan :

$$\mathbf{C}_1 = \begin{bmatrix} 1 & -1 & 0 & . & . & . & 0 \\ 1 & 0 & -1 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 1 & 0 & 0 & . & . & . & -1 \end{bmatrix} \qquad \mathbf{C}_2 = \begin{bmatrix} -1 & 1 & 0 & . & . & . & 0 & 0 \\ 0 & -1 & 1 & . & . & . & 0 & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & . & -1 & 1 \end{bmatrix}$$

Maka :

$$\mathbf{C}_1\mu = \begin{bmatrix} \mu_1 - \mu_2 \\ \mu_1 - \mu_3 \\ . \\ . \\ . \\ \mu_1 - \mu_q \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & . & . & . & 0 \\ 1 & 0 & -1 & . & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 1 & 0 & 0 & . & . & . & -1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ . \\ . \\ . \\ \mu_q \end{bmatrix}$$

atau

$$\mathbf{C}_2\mu = \begin{bmatrix} \mu_2 - \mu_1 \\ \mu_3 - \mu_2 \\ . \\ . \\ . \\ \mu_q - \mu_{q-1} \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 & . & . & . & 0 & 0 \\ 0 & -1 & 1 & . & . & . & 0 & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & . & -1 & 1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ . \\ . \\ . \\ \mu_q \end{bmatrix}$$

C_1 dan C_2 disebut matriks kontras , karena q-1 barisnya independen linier dan masing masing merupakan vektor kontras.

Bila rata - rata perlakuan sama $\mathbf{C}_1\mu = \mathbf{C}_2\mu = \mathbf{0}$, secara umum untuk uji hipotesa tidak ada perbedaan rata - rata perlakuan menjadi $\mathbf{C}\,\mu = \mathbf{0}$ untuk sebarang matriks kontras. Maka uji untuk persamaan perlakuan dalam rancangan pengukuran berulang dengan memperhatikan populasi normal $N_q\,(\mu\,,\,\Sigma)$ adalah H_0 : $\mathbf{C}\,\mu = \mathbf{0}$ melawan H_1 : $\mathbf{C}\,\mu \neq \mathbf{0}$. Tolak H_0 pada taraf signifikansi α jika :

$$T^2 = n\,(\,C\bar{x}\,)^T\,(\,C\,S\,C^T\,)^{-1}\,C\bar{x} > \frac{(n-1)\,(q-1)}{n-q+1}\,F_{q-1,\,n-q+1}(\alpha) \quad \text{...... (26)}$$

dimana $F_{q-1\,,\,n-q+1}$ adalah (100 α) persentil dari distribusi F dengan derajat bebas q - 1 dan n - q + 1 . $\bar{x}$ dan S adalah rata - rata sampel dan matriks kovariansi yang didefinisikan dengan :

$$\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j \qquad S = \frac{1}{n-1}\sum_{j=1}^{n}(x_j - \bar{x})\,(x_j - \bar{x})^T$$