ISBN:1858-1382

J M S K
Jurnal Matematika, Statistika, & Komputasi
http://journal.unhas.ac.id/index.php/jmsk
Vol. 13, No. 1, 71-77, Juli 2016

Penaksiran Parameter dari Variansi Vektor pada Pengujian Hipotesis Kesamaan Matriks Kovariansi

Nasrah Sirajang*

Abstrak

Variansi vektor merupakan salah satu ukuran dispersi data, yang didefinisikan sebagai jumlah dari semua elemen diagonal matriks kovariansi. Paper ini membahas tentang penaksiran parameter dari variansi vektor yang digunakan pada pengujian hipotesis kesamaan matriks kovariansi. Yang dilakukan dalam menguji kesamaan matriks kovariansi adalah menentukan taksiran tak bias dari rataan dan variansi, dengan memamfaatkan metode momen.

Kata Kunci: Metode Momen, matriks kovariansi, variansi vector, vec operator.

1. Pendahuluan

Salah satu statistik inferensial yang digunakan dalam menarik kesimpulan mengenai suatu populasi dengan memamfaatkan sampel-sampel yang diambil dari populasi tersebut yaitu dengan menggunakan estimasi (penaksiran). Misalkan $X = \begin{pmatrix} X_{(1)} \\ X_{(2)} \end{pmatrix}$ suatu vektor acak yang dipartisi menjadi $X_{(1)}$ dan $X_{(2)}$ masing-masing berdimensi q dan p-q. Vektor mean dan matriks kovariansi masing-masing $\mu_{(i)} = E\left(X_{(i)}\right)$, dengan $i = 1, 2$, dan $\sum_{ij} = E\left(\left(X_{(i)} - \mu_{(i)}\right)\left(X_{(j)} - \mu_{(j)}\right)^t\right)$, dengan $i, j = 1, 2$. Untuk mengukur hubungan linier antara dua vektor acak $X_{(1)}$ dan $X_{(2)}$ digunakan $Tr\left(\sum_{12}\sum_{21}\right) = \left\|vec\left(\sum_{12}\right)\right\|^2$. Parameter ini disebut kovariansi vektor. Bentuk $Tr\left(\sum_{12}\sum_{21}\right)$ adalah jumlah semua elemen diagonal dari $\left(\sum_{12}\sum_{21}\right)$. Dari pengertian tersebut, dapat didefinisikan Variansi Vektor dari $X_{(1)}$ dan $X_{(2)}$ adalah $VV = \left(X_{(1)}\right) = Tr\left(\sum_{11}^2\right) = \left\|vec\left(\sum_{11}\right)\right\|^2$ dan $VV = \left(X_{(2)}\right) = Tr\left(\sum_{22}^2\right) = \left\|vec\left(\sum_{22}\right)\right\|^2$.

2. Penaksiran Sampel

Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak yang saling bebas dari distribusi $N_p\left(\mu, \Sigma\right)$. Vektor rata-rata sampel dan matriks kovariansi sampel adalah

* *Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Hasanuddin, Jl. Perintis Kemerdekaan Km.10, Tamalanrea, Makassar.*

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{dan} \quad S = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(X_i - \overline{X}\right)^t . \tag{1}$$

Jika S berukuran *m x m* maka *vec(S)* adalah representasi dari S dalam bentuk vektor berukuran *m x 1*, dengan

$$E\left\{\left\|vec\left(S\right)\right\|^2\right\} = \left\|vec\left(\Sigma\right)\right\|^2 \text{ dan } \mathrm{var}\left\{\left\|vec\left(S\right)\right\|^2\right\} = \frac{8}{n-1}\left\|vec\left(\Sigma^2\right)\right\|^2 .$$

Pada aplikasi akan diperlukan nilai taksiran tak bias dari $\left\|vec\left(\Sigma\right)\right\|^2$ dan $\left\|vec\left(\Sigma^2\right)\right\|^2$, dimana nilai tersebut berasal dari nilai sampel. Oleh karena itu, akan dibahas bagaimana menentukan taksiran tak bias dari $\left\|vec\left(\Sigma\right)\right\|^2$ dan $\left\|vec\left(\Sigma^2\right)\right\|^2$, dengan memanfaatkan sifat-sifat dari trace yang berhubungan dengan nilai eigen dari matriks kovariansi. Misalkan λ_k dan λ_k adalah nilai eigen dari matriks kovariansi sampel (S) dan matriks kovariansi populasi (∑).

Definisi 1.
Misalkan A adalah matriks berukuran m x m dengan nilai eigen $\lambda_1, \lambda_2, ..., \lambda_m$, *maka* $Tr\left(A\right) = \sum_{k=1}^{m} \lambda_k$.

Hubungan dari Variansi Vektor dengan nilai eigen dari matriks kovariansi sampel dan matriks kovariansi populasi, sebagai berikut:

$$\left\|vec\left(S\right)\right\|^2 = Tr\left(S^2\right) = \sum_{k=1}^{m} \lambda_k^2 \text{ dan } \left\|vec\left(\Sigma\right)\right\|^2 = Tr\left(\Sigma^2\right) = \sum_{k=1}^{m} \lambda_k^2 \tag{2}$$

2.1 Satu Sampel

Misalkan $X_1, X_2, ..., X_n$ sampel acak dari distribusi $N_p\left(\mu, \Sigma\right)$ dan misalkan λ_k dan λ_k nilai-nilai eigen dari S dan ∑, *(n-1)S* berdistribusi Wishart yang dinotasikan $\left(n-1\right)S \sim W_p\left(\Sigma, n-1\right)$. Misalkan λ_k dan λ_k nilai-nilai eigen dari S dan ∑. Untuk n yang berukuran besar maka distribusi asimptotik dari λ_k adalah normal dengan rataan λ_k dan variansi $\frac{2\lambda_k^2}{n-1}$ ditulis $\lambda_k \sim N\left(\lambda_k, \frac{2\lambda_k^2}{n-1}\right)$. Untuk memperoleh taksiran dari λ_k^2 dan λ_k^4 digunakan metode momen, dimana momen kedua dan keempat dari λ_k yang digunakan untuk mendapatkan nilai taksiran rataan dan variansi dari variansi vektor.

Definisi 2.
Jika X adalah vektor acak multinormal dengan rataan μ dan matriks kovariansi ∑, maka fungsi pembangkit momen dari X adalah

$$M_x\left(t\right) = \exp\left(t^{'}\mu + \frac{t^{'}\Sigma t}{2}\right), \tag{3}$$

$$E\left(X^r\right)=\frac{d^r M_x(t)}{dt^r}, \tag{4}$$

Dengan mengambil t = 0, maka $E\left(X^r\right)$ menyatakan momen ke–r dari vektor acak X . Dari persamaan (3) , untuk rataan λ_k dan variansi $\frac{2\lambda_k^2}{n-1}$, maka

$$M_x(t)=\exp\left(\lambda_k t+\frac{t^2\lambda_k^2}{n-1}\right) \tag{5}$$

Untuk menemukan momen kedua dan keempat dari λ_k, maka persamaan (4) diterapkan pada persamaan (5), sehingga diperoleh

a. $E\left(\lambda_k\right)=\frac{dM_{\lambda_k}(t)}{dt}$; $\frac{dM_{\lambda_k}(t)}{dt}\Big|_{t=0}=\lambda_k$

b. $E\left(\lambda_k^2\right)=\frac{d^2M_{\lambda_k}(t)}{dt^2}$; $\frac{d^2M_{\lambda_k}(t)}{dt^2}\Big|_{t=0}=\lambda_k^2+\left(\frac{2\lambda_k^2}{n-1}\right)$

c. $E\left(\lambda_k^3\right)=\frac{d^3M_{\lambda_k}(t)}{dt^3}$; $\frac{d^3M_{\lambda_k}(t)}{dt^3}\Big|_{t=0}=\left(1+\frac{6}{n-1}\right)\lambda_k^3$

d. $E\left(\lambda_k^4\right)=\frac{d^4M_{\lambda_k}(t)}{dt^4}$; $\frac{d^4M_{\lambda_k}(t)}{dt^4}\Big|_{t=0}=\left(1+\frac{12}{n-1}+\frac{12}{(n-1)^2}\right)\lambda_k^4$

Hal ini menunjukkan bahwa dari momen kedua dan momen keempat diperoleh:

$$E\left(\lambda_k^2\right)=\left(1+\frac{2}{n-1}\right)\lambda_k^2 \Leftrightarrow E\left[\left(1+\frac{2}{n-1}\right)^{-1}\left\|vec(S)\right\|^2\right]=\left\|vec(\Sigma)\right\|^2$$

$$\text{dan } E\left(\lambda_k^4\right)=\left(1+\frac{12}{n-1}+\frac{12}{(n-1)^2}\right)\lambda_k^4$$

$$\Leftrightarrow E\left[\left(1+\frac{12}{n-1}+\frac{12}{(n-1)^2}\right)^{-1}\left\|vec\left(S^2\right)\right\|^2\right]=\left\|vec\left(\Sigma^2\right)\right\|^2$$

Maka statistik $\left(1+\frac{2}{n-1}\right)^{-1}\left\|vec(S)\right\|^2$, penaksir tak bias dari $\left\|vec(\Sigma)\right\|^2$ dan $\left\{1+\frac{12}{n-1}+\frac{12}{(n-1)^2}\right\}^{-1}\left\|vec\left(S^2\right)\right\|^2$, penaksir tak bias dari $\left\|vec\left(\Sigma^2\right)\right\|^2$. Karena $\left\|vec(S)\right\|^2=Tr\left(S^2\right)$, maka $E\left(Tr\left(S^2\right)\right)=\left(1+\frac{2}{n-1}\right)^{-1}Tr\left(S^2\right)$ dan

$$\operatorname{var}\left(Tr\left(S^2\right)\right)=\frac{8}{n-1}\left\{1+\frac{12}{n-1}+\frac{12}{(n-1)^2}\right\}^{-1}Tr\left(S^4\right).$$

2.2 Beberapa Sampel

Misalkan *m* sampel acak yang saling bebas dari suatu distribusi $N_p(\mu,\Sigma)$, masing-masing sampel berukuran:

$$\begin{array}{ll} \text{Sampel-1} : & X_1, X_2, \ldots, X_{n_1} \\ \text{Sampel-2} : & Y_1, Y_2, \ldots, Y_{n_2} \\ \vdots & \vdots \\ \text{Sampel-}m : & Z_1, Z_2, \ldots, Z_{n_m} \end{array}$$

Kemudian S_i matriks kovariansi sampel ke-*i*, $i = 1, 2, \ldots, m$ dan $(n_i - 1)S_i \sim W_p\left(\sum, (n_i - 1)\right)$. Ini berarti $(n_1 - 1)S_1 \sim W_p(\sum, n_1 - 1)$, $(n_2 -)S_2 \sim W_p(\sum, n_2 - 1)$,...,$(n_m - 1)S_m \sim W_p(\sum, n_m - 1)$. Misalkan $A_1 = (n_1 - 1)S_1$, $A_2 = (n_2 - 1)S_2$, ... , $A_m = (n_m - 1)S_m$ maka $A = \sum_{i=1}^{m} A_i$ berdistribusi Wishart dengan derajat bebas *n* – *m*, dengan kata lain distribusi dari S_{gab} adalah $(n - m)S_{gab} \sim W_p(\sum, (n - m))$. Rata-rata matriks kovariansi gabungan adalah $\overline{S}_{gab} = \frac{\sum_{i=1}^{m}(n_i - 1)S_i}{\sum_{i=1}^{m}(n_i - 1)} = \frac{1}{n - m}\sum_{i=1}^{m}(n_i - 1)S_i$, dimana $n = \sum_{i=1}^{m} n_i$. Misalkan $\overline{\lambda}_k$ rataan nilai eigen dari S_{gab}. Untuk *n* yang berukuran besar distribusi asimptotik dari λ_k adalah normal dengan rataan λ_k dan variansi $\frac{2\lambda_k^2}{n - m}$, ditulis

$$\lambda_k \sim N\left(\lambda_k, \frac{2\lambda_k^2}{n - m}\right).$$

Seperti halnya pada satu sampel. Untuk memperoleh taksiran λ_k^2 dan λ_k^4 digunakan metode momen, sehingga diperoleh

$$E\left(\lambda_k^2\right) = \frac{d^2 M_{\lambda_k}(t)}{dt^2} \; ; \quad \frac{d^2 M_{\lambda_k}(t)}{dt^2}\Big|_{t=0} = \lambda_k^2 + \left(\frac{2\lambda_k^2}{n-m}\right)$$

$$E\left(\lambda_k^4\right) = \frac{d^4 M_{\lambda_k}(t)}{dt^4} \; ; \quad \frac{d^4 M_{\lambda_k}(t)}{dt^4}\Big|_{t=0} = \left(1 + \frac{12}{n-m} + \frac{12}{(n-m)^2}\right)\lambda_k^4$$

Hal ini menunjukkan bahwa dari momen kedua dan keempat diperoleh,

$$E\left(\lambda_k^2\right)=\left(1+\frac{2}{n-m}\right)\lambda_k^2 \qquad \Leftrightarrow E\left[\left(1+\frac{2}{n-m}\right)^{-1}\left\|vec\left(\overline{S}_{gab}\right)\right\|^2\right]=\left\|vec\left(\Sigma\right)\right\|^2$$

$$E\left(\lambda_k^4\right)=\left(1+\frac{12}{n-m}+\frac{12}{\left(n-m\right)^2}\right)\lambda_k^4 \Leftrightarrow E\left[\left(1+\frac{12}{n-m}+\frac{12}{\left(n-m\right)^2}\right)^{-1}\left\|vec\left(\overline{S}_{gab}{}^2\right)\right\|^2\right]=\left\|vec\left(\Sigma^2\right)\right\|^2$$

Maka statistik $\left(1+\frac{2}{n-m}\right)^{-1}\left\|vec\left(\overline{S}_{gab}\right)\right\|^2$ merupakan penaksir tak bias dari $\left\|vec\left(\Sigma\right)\right\|^2$ dan $\left(1+\frac{12}{n-m}+\frac{12}{\left(n-m\right)^2}\right)^{-1}\left\|vec\left(\overline{S}_{gab}{}^2\right)\right\|^2$ penaksir tak bias dari $\left\|vec\left(\Sigma^2\right)\right\|^2$. Sehingga diperoleh $E\left[Tr\left(\overline{S}_{gab}{}^2\right)\right]=\left(1+\frac{2}{n-m}\right)^{-1}Tr\left(\overline{S}_{gab}{}^2\right)$ dan $\text{var}\left\{Tr\left(\overline{S}_{gab}{}^2\right)\right\}=\frac{8}{n-1}\left\{1+\frac{12}{n-m}+\frac{12}{\left(n-m\right)^2}\right\}^{-1}Tr\left(\overline{S}_{gab}{}^4\right)$. Bila $n_1=n_2=\cdots=n_m=n_0$, $\overline{S}=\frac{1}{m}\sum_{i=1}^{m}S_i$, dimana S_i matriks kovariansi sampel ke-*i*; $i=1,2,...,m$ dan $(n_i-1)S_i\sim W_p\left(\sum, (n_i-1)\right)$. Ini berarti

$$(n_0-1)S_1\sim W_p\left(\sum, (n_0-1)\right)$$
$$(n_0-1)S_2\sim W_p\left(\sum, (n_0-1)\right)$$
$$\vdots$$
$$(n_0-1)S_m\sim W_p\left(\sum, (n_0-1)\right) \ .$$

Misalkan

$$A_1=\left(n_0-1\right)S_1$$
$$A_2=\left(n_0-1\right)S_2$$
$$\vdots$$
$$A_m=\left(n_0-1\right)S_m$$

maka $A=\sum_{i=1}^{m}A_i$ berdistribusi Wishart dengan derajat bebas $m\left(n_0-1\right)$ (Anderson, 1984). Dengan kata lain distribusi dari $\overline{S}$ adalah

$$[m(n_0-1)]S\sim W_p[\sum, m(n_0-1].$$

Misalkan λ_k nilai eigen dari $\overline{S}$. Untuk *n* yang berukuran besar distribusi asimptotik dari λ_k adalah normal dengan rataan λ_k dan variansi $\frac{2\lambda_k^2}{m\left(n_0-1\right)}$. Dengan cara yang sama pada *n* yang berbeda, maka untuk momen kedua dan keempat diperoleh

$$E\left(\lambda_k^2\right)=\left(1+\frac{2}{m\left(n_0-1\right)}\right)\lambda_k^2$$

$$E\left(\lambda_k^4\right)=\left(1+\frac{12}{m\left(n_0-1\right)}+\frac{12}{\left(m\left(n_0-1\right)\right)^2}\right)\lambda_k^4$$

sehingga, statistik $\left(1+\frac{2}{m\left(n_0-1\right)}\right)^{-1}\left\|vec\left(\overline{S}\right)\right\|^2$ merupakan penaksir tak bias dari $\left\|vec\left(\Sigma\right)\right\|^2$ dan $\left(1+\frac{12}{m\left(n_0-1\right)}+\frac{12}{\left(m\left(n_0-1\right)\right)^2}\right)^{-1}\left\|vec\left(\overline{S}^2\right)\right\|^2$ merupakan penaksir tak bias dari $\left\|vec\left(\Sigma^2\right)\right\|^2$. Jadi $E\left[Tr\left(\overline{S}^2\right)\right]=\left(1+\frac{2}{m\left(n_0-1\right)}\right)^{-1}Tr\left(\overline{S}^2\right)$ penaksir tak bias dari rataan dan $\operatorname{var}\left\{Tr\left(\overline{S}^2\right)\right\}=\frac{8}{n_0-1}\left(1+\frac{12}{m\left(n_0-1\right)}+\frac{12}{\left(m\left(n_0-1\right)\right)^2}\right)^{-1}Tr\left(\overline{S}^4\right)$ penaksir tak bias dari variansi.

5. Kesimpulan

Berikut beberapa kesimpulan yang dapat diberikan dalam penelitian ini adalah sebagai berikut.

1. Momen kedua dan keempat dari λ_k yang digunakan untuk mendapatkan nilai taksiran rataan dan variansi dari variansi vektor.
2. Taksiran dari $\left\|vec\left(\Sigma\right)\right\|^2$ = $Tr\left(\Sigma^2\right)$ dan $\left\|vec\left(\Sigma^2\right)\right\|^2$ = $Tr\left(\Sigma^4\right)$ merupakan taksiran tak bias dari rataan dan variansi berdasarkan variansi vektor .
3. Rataan dan variansi berdasarkan variansi vektor adalah sebagai berikut:
   a. Untuk satu sampel,
   $$E\left(Tr\left(S^2\right)\right)=\left(1+\frac{2}{n-1}\right)^{-1}Tr\left(S^2\right),$$
   $$\operatorname{var}\left(Tr\left(S^2\right)\right)=\frac{8}{n-1}\left\{1+\frac{12}{n-1}+\frac{12}{\left(n-1\right)^2}\right\}^{-1}Tr\left(S^4\right)$$
   b. Untuk beberapa sampel dengan $n_1,n_2,...,n_m$ berbeda,
   $$E\left[Tr\left(\overline{S}_{gab}{}^2\right)\right]=\left(1+\frac{2}{n-m}\right)^{-1}Tr\left(\overline{S}_{gab}{}^2\right),$$
   $$\operatorname{var}\left\{Tr\left(\overline{S}_{gab}{}^2\right)\right\}=\frac{8}{n-1}\left\{1+\frac{12}{n-m}+\frac{12}{\left(n-m\right)^2}\right\}^{-1}Tr\left(\overline{S}_{gab}{}^4\right)$$

c. Untuk beberapa sampel dengan $n_1 = n_2 = \cdots = n_m = n_0$,

$$E\left[Tr\left(\overline{S}^2\right)\right] = \left(1 + \frac{2}{m\left(n_0 - 1\right)}\right)^{-1} Tr\left(\overline{S}^2\right) \quad ,$$

$$\text{var}\left\{Tr\left(\overline{S}^2\right)\right\} = \frac{8}{n_0 - 1}\left(1 + \frac{12}{m\left(n_0 - 1\right)} + \frac{12}{\left(m\left(n_0 - 1\right)\right)^2}\right)^{-1} Tr\left(\overline{S}^4\right)$$

Daftar Pustaka

Anderson, T.W., 1984. *An Introduction to Multivariate Statistical Analysis*. John Wiley & Sons Inc., New York .

Djauhari, M.A., 2007. A measure of data concentration. *Journal of Applied Probability & Statistics*, Vol 2 No.2, pp. 139-155.

Guritno, S., 2007. *Statistik Multivariat Terapan, Teori dan Aplikasi*. UGM, Yokyakarta.

Rencher, A.C., 2002. *Methods of Multivariate Analysis*. John Wiley & Sons Inc., New York.

Schott, J.R., 1997. *Matrix Analysis for Statistics*. John Wiley & Sons Inc., New York