



# Kajian tentang Jumlah Kuadrat dan Hasil Kali dalam Analisis Varians Multivariat

Marzuki<sup>1</sup>, Hafnani<sup>1</sup>

<sup>1</sup>Jurusan Matematika FMIPA, Universitas Syiah Kuala, Indonesia

Diterima 24 Mei 2014; Disetujui 12 Juni 2014

**Abstract** – Multivariate analysis of variance is the generalization analysis of variance. Two-way multivariate analysis of variance is to know two-way effects of  $p$  criteria variables with  $p \geq 2$ . The matrix of sum of squares and cross-product or matrix of JKHK total is rather same with sum of squares total at univariate situation divided four JKHK matrices that has mean, each JKHK factor A, factor B, interaction between factor A and factor B, and JKHK error. With pass comparative determinant of JKHK error matrix to determinant of sum of matrices JKHK factor that we want to know the effects and JKHK error, it can be found a statistic  $A^*$  that is know as Wilks' Lambda.

**Keywords** : multivariate analysis, sums of squares and cross products, two factors, variance

## 1. Pendahuluan

Data yang diperoleh dari suatu individu anggota sampel dalam suatu penelitian kadang-kadang terdiri dari beberapa peubah yang diukur. Jika korelasi antarpeubah itu kemungkinan ada, maka kelompok peubah itu diperlakukan dalam suatu sistem karena korelasi itu sangat penting untuk diperhatikan. Kumpulan data yang lebih dari satu ini biasa dikenal dengan data multivariat [1].

Korelasi diperlukan untuk mengukur eratnya hubungan linier antara dua peubah [2]. Peubah respons yang memiliki beberapa komponen memerlukan keberadaan model regresi pengaruh campuran untuk mendapatkan statistik uji yang bagus [3]. Jika pengukuran beberapa kriteria untuk melihat pengaruh faktor-faktor terhadapnya maka selanjutnya akan dilihat pengaruh interaksi antarfaktor dan pengaruh utama. Analisis varians multivariat digunakan pada kasus seperti di atas untuk data berskala minimal interval dan peubah faktor-faktornya berskala nominal atau ordinal. Tujuan analisis ini adalah menganalisis data tentang perbedaan pengaruh beberapa faktor terhadap sekelompok peubah kriteria yang saling berkorelasi sebagai satu sistem dengan memperhitungkan korelasi antarpeubah

tersebut. Peubah kriteria yang dikaji dalam tulisan ini dibatasi untuk dua peubah saja. Tujuan dari kajian ini adalah memperkenalkan suatu model tetap pengaruh dua faktor dalam analisis varians multivariat, khususnya untuk rancangan acak lengkap sehingga dapat dipakai untuk mengetahui ada tidaknya pengaruh interaksi faktor atau pengaruh faktor utama dari suatu masalah.

## 2. Model Rancangan Acak Lengkap

Pengamatan untuk Rancangan Acak Lengkap dengan dua faktor  $x_{ijk}$  dapat diurai menjadi jumlah dari dua komponen, yaitu komponen rata-rata  $\mu_{ij}$  dan komponen acak  $\varepsilon_{ijk}$  [4]. Model linier Rancangan Acak Lengkap untuk dua faktor dapat ditulis

$$x_{ijk} = \mu_{ij} + \varepsilon_{ijk} \quad (1)$$

dengan  $i, j$ , dan  $k$  adalah bilangan bulat,  $1 \leq i \leq a$ ,  $1 \leq j \leq b$ , dan  $1 \leq k \leq n$ . Komponen acak  $\varepsilon_{ijk}$  adalah penyimpangan nilai pengamatan  $x_{ijk}$  pada sel ke- $ij$  dari rata-rata populasi  $\mu_{ij}$ .

Rata-rata populasi dapat ditulis sebagai

$$\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} \quad (2)$$

dengan  $\mu$  merupakan rata-rata keseluruhan,  $\alpha_i, \beta_j$ , dan  $\gamma_{ij}$  berturut-turut merupakan pengaruh faktor A, faktor B, dan faktor interaksi [5]. Sehingga model (1) dapat ditulis

$$x_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}. \quad (3)$$

Pengaruh  $\alpha_i, \beta_j$ , dan  $\gamma_{ij}$  dalam model ini didefinisikan sebagai deviasi dari rata-rata [6]. Dengan demikian, batasannya diperoleh sebagai  $\sum_{i=1}^a \alpha_i = \sum_{j=1}^b \beta_j = \sum_{i=1}^a \sum_{j=1}^b \gamma_{ij} = \sum_{j=1}^b \sum_{i=1}^a \gamma_{ij} = \mathbf{0}$  dan komponen acak bebas satu sama lain serta berdistribusi normal dengan rata-rata nol dan varians konstan.

Model (3) dapat dianalogikan bahwa masing-masing observasi ditetapkan untuk sampel sebagai

$$\begin{aligned} x_{ijk} = & \bar{x}_{...} + (\bar{x}_{i..} - \bar{x}_{...}) + (\bar{x}_{.j.} - \bar{x}_{...}) \\ & + (\bar{x}_{ij.} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{x}_{...}) \\ & + (x_{ijk} - \bar{x}_{ij.}). \end{aligned} \quad (4)$$

Atau dalam penulisan yang sederhana dapat ditulis sebagai

$$\begin{aligned} x_{ijk} = & \bar{x} + (\bar{x}_i - \bar{x}) + (\bar{x}_j - \bar{x}) \\ & + (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x}) \\ & + (x_{ijk} - \bar{x}_{ij}) \end{aligned} \quad (5)$$

Apabila selisih dari masing-masing observasi dengan rata-rata keseluruhan dikuadratkan dan dijumlahkan maka diperoleh Teorema Kesamaan Jumlah Kuadrat [7].

**Teorema Kesamaan Jumlah Kuadrat**

$$\begin{aligned} & \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \bar{x})^2 \\ & = bn \sum_{i=1}^a (\bar{x}_i - \bar{x})^2 + an \sum_{j=1}^b (\bar{x}_j - \bar{x})^2 + \\ & n \sum_{i=1}^a \sum_{j=1}^b (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})^2 + \\ & \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \bar{x}_{ij})^2. \end{aligned}$$

Teorema ini dengan mudah dapat dibuktikan karena enam suku perkalian silang yang akan terjadi, bernilai nol [8].

**3. Analisis Varians Multivariat**

**Model tetap pengaruh dua faktor tetap**

Faktor tetap adalah faktor yang level-levelnya bukan merupakan sampel acak dari sebuah populasi dan model tetap adalah model yang melibatkan faktor tetap [9]. Model tetap untuk vektor respons yang mengandung  $p$

komponen yaitu peubah kriteria yang dipilih, dapat ditulis

$$x_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk} \quad (6)$$

dengan  $i, j$ , dan  $k$  adalah bilangan bulat,  $1 \leq i \leq a$ ,  $1 \leq j \leq b$ , dan  $1 \leq k \leq n$ , serta batasan  $\sum_{i=1}^a \alpha_i = \sum_{j=1}^b \beta_j = \sum_{i=1}^a \sum_{j=1}^b \gamma_{ij} = \sum_{j=1}^b \sum_{i=1}^a \gamma_{ij} = \mathbf{0}$ .

Model ini mempunyai asumsi bahwa vektor acak  $\epsilon_{ijk}$  berdistribusi normal  $p$ -variat dengan vektor rata-rata nol dan matriks dispersi (variens-kovarians)  $\Sigma$  konstan atau dinotasikan sebagai  $\epsilon_{ijk} \sim N_p(\mathbf{0}, \Sigma)$ . Matriks dispersi ini adalah matriks simetris yang unsurnya merupakan varians dari masing-masing peubah dan kovarians antarpeubah.

Semua vektor di persamaan (1) berorder  $px1$ . Demikian respons-respons yang mengandung  $p$  ukuran diulang sebanyak  $n$  kali pada setiap kombinasi yang mungkin dari level-level faktor A dan faktor B.

Vektor-vektor observasi  $x_{ijk}$  dapat ditetapkan sebagai berikut

$$\begin{aligned} x_{ijk} = & \bar{x}_{...} + (\bar{x}_{i..} - \bar{x}_{...}) + (\bar{x}_{.j.} - \bar{x}_{...}) \\ & + (\bar{x}_{ij.} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{x}_{...}) \\ & + (x_{ijk} - \bar{x}_{ij.}). \end{aligned} \quad (7)$$

Atau secara sederhana dapat ditulis sebagai

$$\begin{aligned} x_{ijk} = & \bar{x} + (\bar{x}_i - \bar{x}) + (\bar{x}_j - \bar{x}) \\ & + (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x}) \\ & + (x_{ijk} - \bar{x}_{ij}) \end{aligned} \quad (8)$$

dengan

$\bar{x}_{...} = \bar{x}$  adalah rata-rata keseluruhan dari vektor-vektor observasi;

$\bar{x}_{i..} = \bar{x}_i$  adalah rata-rata vektor observasi ke-  $i$  faktor A;

$\bar{x}_{.j.} = \bar{x}_j$  adalah rata-rata vektor observasi ke-  $j$  faktor B;

$\bar{x}_{ij.} = \bar{x}_{ij}$  adalah rata-rata vektor observasi ke-  $i$  faktor A dan ke-  $j$  faktor B.

Suatu skalar  $(x_{ijk} - \bar{x})^2$  dari kasus data univariat digeneralisasikan ke kasus multivariat menjadi  $(x_{ijk} - \bar{x})(x_{ijk} - \bar{x})'$  yaitu matriks yang berkorespondensi dan menghasilkan Teorema Kesamaan Jumlah Kuadrat dan Hasil Kali (JKHK) [7] seperti berikut ini.

**Teorema Kesamaan JKHK**

$$\begin{aligned} & \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \bar{x})(x_{ijk} - \bar{x})' \\ &= bn \sum_{i=1}^a (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})' + an \sum_{j=1}^b (\bar{x}_j - \bar{x})(\bar{x}_j - \bar{x})' + n \sum_{i=1}^a \sum_{j=1}^b (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})(\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})' + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \bar{x}_{ij})(x_{ijk} - \bar{x}_{ij})'. \end{aligned}$$

Pembuktian teorema kesamaan JKHK dapat dilakukan sebagai berikut.

$$\begin{aligned} & \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \bar{x})(x_{ijk} - \bar{x})' \\ &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n [(\bar{x}_i - \bar{x}) + (\bar{x}_j - \bar{x}) + (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x}) + (x_{ijk} - \bar{x}_{ij})][(\bar{x}_i - \bar{x}) + (\bar{x}_j - \bar{x}) + (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x}) + (x_{ijk} - \bar{x}_{ij})]' \\ &= \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n [(\bar{x}_i - \bar{x}) + (\bar{x}_j - \bar{x}) + (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x}) + (x_{ijk} - \bar{x}_{ij})][(\bar{x}_i - \bar{x})' + (\bar{x}_j - \bar{x})' + (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})' + (x_{ijk} - \bar{x}_{ij})]' \\ &= bn \sum_{i=1}^a (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})' + an \sum_{j=1}^b (\bar{x}_j - \bar{x})(\bar{x}_j - \bar{x})' + n \sum_{i=1}^a \sum_{j=1}^b (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})(\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})' + \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \bar{x}_{ij})(x_{ijk} - \bar{x}_{ij})' \end{aligned}$$

Hasil akhir ini diperoleh karena dua belas suku perkalian silangnya sama dengan nol [8].

Korespondensi derajat bebas untuk masing-masing matriks pada pemisahan JKHK dapat diurai sebagai berikut

$$abn - 1 = (a - 1) + (b - 1) + (a - 1)(b - 1) + ab(n - 1) \tag{9}$$

Suku kesamaan JKHK tersebut dapat ditulis sebagai berikut

$$JKHK_{total} = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \bar{x})(x_{ijk} - \bar{x})'$$

$$JKHK_{faktor A} = bn \sum_{i=1}^a (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})'$$

$$JKHK_{faktor B} = an \sum_{j=1}^b (\bar{x}_j - \bar{x})(\bar{x}_j - \bar{x})'$$

$$JKHK_{interaksi} = n \sum_{i=1}^a \sum_{j=1}^b (\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})(\bar{x}_{ij} - \bar{x}_i - \bar{x}_j + \bar{x})'$$

$$JKHK_{error} = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \bar{x}_{ij})(x_{ijk} - \bar{x}_{ij})'$$

sehingga diperoleh

$$\begin{aligned} JKHK_{total} &= JKHK_{faktor A} + JKHK_{faktor B} \\ &+ JKHK_{interaksi} \\ &+ JKHK_{error} \end{aligned} \tag{10}$$

Perhitungan masalah Analisis Varians Multivariat biasanya diringkas dalam bentuk tabel berikut ini

Tabel 1. Analisis varians multivariat untuk dua faktor

Sumber Variasi	Matriks JKHK	Derajat Bebas
Faktor A	$JKHK_{faktor A}$	$(a - 1)$
Faktor B	$JKHK_{faktor B}$	$(b - 1)$
Interaksi	$JKHK_{interaksi}$	$(a - 1)(b - 1)$
Error	$JKHK_{error}$	$ab(n - 1)$
Total	$JKHK_{total}$	$abn - 1$

Perbandingan determinan dari JKHK error dengan determinan dari JKHK error ditambah JKHK pengaruh yang ingin diuji dikenal dengan statistik Wilks' Lambda  $A^*$ . Statistik ini diperkenalkan oleh Wilks'. Statistik Wilks' Lambda untuk pengaruh interaksi dirumuskan sebagai berikut:

$$A^* = \frac{|JKHK_{error}|}{|JKHK_{interaksi} + JKHK_{error}|}$$

Sedangkan statistik Wilks' Lambda untuk pengaruh faktor A dan faktor B masing-masing dirumuskan sebagai berikut:

$$A_1^* = \frac{|JKHK_{error}|}{|JKHK_{faktor A} + JKHK_{error}|} \text{ dan}$$

$$A_2^* = \frac{|JKHK_{error}|}{|JKHK_{faktor B} + JKHK_{error}|}$$

Nilai  $A^*$ ,  $A_1^*$ , dan  $A_2^*$  tersebut berturut-turut digunakan untuk menguji hipotesis  $H_0$  di bawah ini.

$$H_0 : \boldsymbol{\gamma}_{ij} = \mathbf{0}$$

$H_1$  : paling sedikit satu dari  $\gamma_{ij} \neq 0$

$H_0$  :  $\alpha_i = 0$

$H_1$  : paling sedikit satu dari  $\alpha_i \neq 0$

$H_0$  :  $\beta_j = 0$

$H_1$  : paling sedikit satu dari  $\beta_j \neq 0$

Penolakan hipotesis  $H_0$  akan terjadi pada saat nilai statistik uji itu sangat kecil.

#### 4. Kesimpulan

Kajian terhadap matriks jumlah kuadrat hasil kali dalam analisis varians multivariat ini menghasilkan beberapa kesimpulan, di antaranya:

1. Apabila dari suatu individu diperoleh  $p$  peubah kriteria dengan  $p \geq 2$  maka sebaiknya data itu dianalisis dengan analisis multivariat yang tidak mengabaikan korelasi antarpeubah kriteria sehingga kesimpulannya lebih mencerminkan ke hal yang sebenarnya.
2. Korelasi antarpeubah kriteria terdapat pada kovarians antara dua peubah kriteria atau pada semua elemen dari matriks dispersi  $\Sigma$ . Matriks ini merupakan generalisasi dari varians  $\sigma^2$  pada situasi univariat. Jika  $p = 3$  maka matriks dispersinya dapat ditulis

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ \rho_{21}\sigma_2\sigma_1 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ \rho_{31}\sigma_3\sigma_1 & \rho_{32}\sigma_3\sigma_2 & \sigma_3^2 \end{bmatrix}$$

Matriks ini merupakan matriks simetris berorde  $3 \times 3$ .

3. Jumlah kuadrat hasil kali (JKHK) dari suatu variasi pada analisis varians multivariat mirip dengan jumlah kuadrat (JK) pada situasi univariat. Perkalian vektor deviasi berorde  $p \times 1$  dengan transposnya menghasilkan matriks simetris berorde  $p \times p$ .

4. Generalisasi model tetap rancangan acak lengkap dari univariat ke multivariat berbeda pada ukuran suku-suku pada model tersebut. Model tetap multivariat, suku-sukunya berorde  $p \times 1$ . Sedangkan pada situasi univariat, suku-sukunya berukuran skalar. Penaksiran parameter pada model tersebut dilakukan secara serentak untuk semua peubah kriteria. Penaksiran parameter ini berdasarkan analogi dari situasi univariat.

#### Daftar Pustaka

- [1] Suryanto, 1988, *Metode Statistika ultivariat*, Departemen Pendidikan dan Kebudayaan RI, Jakarta
- [2] R.K. Sembiring, 1995, *Analisis Regresi*, ITB, Bandung
- [3] S. Aryala, D.K. Bhaumik, T. Mathew, dan R.D. Gibbons, 2014, An Optimal Test for Variance Components of Multivariate Mixed-effects Linear Models, *Journal of Multivariate Analysis*, Volume 124, February 2014, Pages 166–178
- [4] F. Minurita, 1997, *Analisis Rancangan Acak Lengkap Dua Faktor*, Skripsi, Universitas Syiah Kuala, Banda Aceh
- [5] R.E. Walpole dan R.H. Myers, 1986, *Ilmu Peluang dan Statistik untuk Insinyur dan Ilmuwan*, ITB, Bandung
- [6] W.W. Hines dan D.C. Montgomery, 1990, *Probabilita dan Statistik dalam Ilmu Rekayasa dan Manajemen*, Edisi Kedua, Universitas Indonesia, Jakarta
- [7] R.A. Johnson dan D.W. Wichern, 1995, *Applied Multivariate Statistical Analysis*, Edisi Ketiga, Prentice Hall, New Jersey
- [8] Marzuki, 2000, *Analisis Varians Multivariat dengan Dua Faktor Tetap*, Skripsi, Universitas Syiah Kuala, Banda Aceh
- [9] G.A. Milliken dan D.E. Johnson, 1984, *Analysis of Messy Data*, Van Nostrand Reinhold, New York