

Analisis Variansi Multivariat

Muhammad Ridhwan Ramdhani - 18209019
Program Studi Sistem dan Teknologi Informasi
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
m.ridhwan.ramdhani@gmail.com

Abstrak—Makalah ini menguraikan tentang analisis multivariat yang biasa digunakan untuk menganalisis kondisi saat terdapat lebih dari satu variabel untuk dianalisis. Terdapat empat jenis tes statistik yang diuraikan dan contoh keluaran keempat tes statistik tersebut dengan menggunakan software statistik SPSS (Statistical Package for Social Sciences).

Kata kunci—Multivariat, SPSS, Tes Statistik.

I. PENDAHULUAN

Tujuan dari analisis variansi (ANOVA) adalah untuk menguji apakah rataan dua atau lebih grup sampel diambil dari sampel distribusi yang sama. MANOVA adalah singkatan dari Analisis Variansi Multivariat yang merupakan pengembangan dari ANOVA. Tujuan dari MANOVA adalah untuk menguji apakah vektor rataan dua atau lebih grup sampel diambil dari sampel distribusi yang sama.

MANOVA biasa digunakan dalam dua kondisi utama. Kondisi pertama adalah saat terdapat beberapa variabel dependen yang berkorelasi, sementara peneliti hanya menginginkan satu kali tes keseluruhan pada kumpulan variabel ini dibandingkan dengan beberapa kali tes individual. Kondisi kedua adalah saat peneliti ingin mengetahui bagaimana variabel independen mempengaruhi pola variabel dependen.

II. ANALISIS MULTIVARIAT

MANOVA baik digunakan jika pada percobaan terdapat beberapa variabel independen yang nilainya akan diubah. MANOVA memiliki beberapa kelebihan dibandingkan dengan ANOVA, salah satunya adalah mengurangi kemungkinan terjadinya galat Tipe I yang mungkin terjadi saat dilakukan perhitungan beberapa ANOVA secara independen.

Asumsi yang berlaku pada MANOVA antara lain:

1. Variabel dependen terdistribusi normal.
2. Setiap pasang variabel dependen, kovariat, dan semua pasangan variabel dependen-kovariat adalah linear.
3. Variabel dependen memiliki tingkat variansi yang sama sepanjang daerah variabel prediktor.
4. Variabel dependen maupun independen dapat

dikorelasikan satu sama lain.

Batasan-batasan penggunaan MANOVA antara lain:

1. Outliers – Outliers dapat menghasilkan galat Tipe I atau Tipe II dan tidak ada tanda yang menandakan galat tipe mana yang terjadi dalam analisis.
2. Multikolinearitas dan Singularitas – Saat terdapat korelasi yang tinggi antara variabel dependen, salah satu variabel dependen akan menjadi kombinasi linear dari variabel dependen lainnya.

Semua tes MANOVA dijabarkan sebagai $A = E^{-1}H$. Terdapat empat jenis tes statistik multivariat yang dijabarkan dalam $E^{-1}H$. Statistik yang pertama adalah *Bartlett-Pillai's Criterion* dengan rumus sebagai berikut:

$$\text{trace}[H(H + E)^{-1}] = \sum_{i=1}^q \frac{\lambda_i}{1 + \lambda_i}$$

Criterion ini dapat diubah menjadi statistika distribusi F dengan derajat kebebasan $v_1 = br$ dan

$$v_2 = r(df_e - p + r) \text{ dan persamaan berikut:}$$

$$F = \frac{U}{r - U} \frac{df_e - p + r}{b}$$

dengan $b = \max(p, df_h)$ dan $r = \min(p, df_h)$.

Statistik yang kedua adalah *Hotelling-Lawley Criterion* dengan rumus sebagai berikut:

$$\text{trace}(A) = \text{trace}(HE^{-1}) = \sum_{i=1}^q \lambda_i$$

Criterion ini dapat diubah menjadi statistika distribusi F dengan derajat kebebasan $v_1 = br$ dan

$$v_2 = r(df_e - p - 1) + 2 \text{ dimana } b = \max(p, df_h)$$

dan $r = \min(p, df_h)$ dan persamaan berikut:

$$F = V \left(\frac{r(df_e - p - 1) + 2}{r^2 b} \right).$$

Statistik yang ketiga adalah *Wilk's Lambda Criterion* dengan rumus sebagai berikut:

$$\Lambda = \frac{|E|}{|H + E|} = \prod_{i=1}^q \frac{1}{1 + \lambda_i}$$

Criterion ini dapat diubah menjadi statistika distribusi F dengan derajat kebebasan $v_1 = 2p$ dan $v_2 = 2(df_e - p + 1)$ dan persamaan berikut:

$$F = \frac{1 - \Lambda^{1/2}}{\Lambda^{1/2}} \frac{df_e - p + 1}{p}$$

Statistik yang keempat adalah *Roy Criterion* yang merupakan batas atas statistik F.

$$\Theta = \frac{\lambda_i}{1 + \lambda_i}$$

Roy Criterion dalam statistika distribusi F adalah

$$F = \frac{(N - b - 1)\lambda_1}{b}$$

III. PENERAPAN PADA KOMPUTER

Berikut adalah contoh analisis statistika univariat dan multivariat pada komputer dengan menggunakan SPSS (*Statistical Package for the Social Sciences*):

Syntax:

```
MANOVA Y1 Y2
/PRINT=CELLINFO(MEANS SSCP COV)
HOMOGENEITY(BOX) ERROR(SSCP COR)
SIGNIV(MULTIV).
```

Nilai Y1 dan Y2 sesuai pada tabel berikut: (semua data dalam tabel diambil dari buku *Applied MANOVA and Discriminant Analysis* karangan Carl J. Huberty, 2006)

TABEL 3.1 Nilai Y₁ (*Error Detection Task*) dan Y₂ (*Degrees of Reading Power*) untuk kelompok TA (*Think Aloud*)

Y ₁	Y ₂
4	43
4	34
4	45
3	39
8	40
1	27
7	46
7	39
9	31
6	39
4	40
12	52
14	53

12	53
7	41
5	41
9	46
13	52
11	55
5	36
11	50
15	54

Output:

Cell Means and Standard Deviations
Variable . . Y1

	Mean	Std.Dev.	N	95 percent	Conv. Interval
For entire sample	7.773	3.927	22	6.032	9.514

Variable . . Y2

	Mean	Std.Dev.	N	95 percent	Conv. Interval
For entire sample	43.455	7.860	22	39.970	46.940

WITHIN CELLS Correlations with Std. Devs. on Diagonal

	Y1	Y2
Y1	3.92710	
Y2	.79952	7.86025

Statistics for WITHIN CELLS correlations

Log(Determinant) = -1.01954
Bartlett test of sphericity = 19.88095 with 1 D. F.
Significance = .000

F(max) criterion = 4.00618 with (2,21) D. F.

WITHIN CELLS Sum-of-Squares and Cross-Products

	Y1	Y2
Y1	323.86364	
Y2	518.27273	1297.45455

***** Analysis of Variance *****
 EFFECT .. CONSTANT
 Multivariate Test of Significance (S = 1, M = 0, N = 9)

Test Name	Value	Exact F	Hypoth. DF
Pillais	.98013	493.18834	2.00
Hottelings	49.31883	493.18834	2.00
Wilks	.01987	493.18834	2.00
Roys	.98013		

Note.. F statistics are exact.

 EFFECT .. CONSTANT (Cont.)
 Univariate F-test with (1,21) D. F.

Variable	Hypoth. SS	Error SS	Hypoth. MS
Y1	1329.13636	323.86364	1329.13636
Y2	41542.54545	1297.45455	41542.54545

Variable	Error MS	F	Sig. of F
Y1	15.42208	86.18400	.000
Y2	61.78355	672.38845	.000

TABEL 3.2 Nilai Y_1 (Error Detection Task) dan Y_2 (Degrees of Reading Power) untuk kelompok DRA (Directed Reading Activity)

Y_1	Y_2
5	34
9	36
5	42
7	37
4	44
9	49
3	38
4	38
2	38
5	50
7	31
8	49
10	54
9	52
12	50
5	35
8	36
12	46

4	42
8	47
6	39
5	38

Output:

Cell Means and Standard Deviations
 Variable . . Y1

	Mean	Std.Dev.	N	95 percent	Conv. Interval
For entire sample	6.682	2.767	22	5.455	7.909

Variable . . Y2

	Mean	Std.Dev.	N	95 percent	Conv. Interval
For entire sample	42.045	6.615	22	39.112	44.978

 WITHIN CELLS Correlations with Std. Devs. on Diagonal

	Y1	Y2
Y1	2.76692	
Y2	.49774	6.61511

 Statistics for WITHIN CELLS correlations

Log(Determinant) = -.28468
 Bartlett test of sphericity = 5.55128 with 1 D. F.
 Significance = .018
 F(max) criterion = 5.71586 with (2,21) D. F.

 WITHIN CELLS Sum-of-Squares and Cross-Products

	Y1	Y2
Y1	160.77273	
Y2	191.31818	918.95455

***** Analysis of Variance *****

EFFECT .. CONSTANT
 Multivariate Test of Significance (S = 1, M = 0, N = 9)

Test Name	Value	Exact F	Hypoth. DF
Pillais	.97733	431.02762	2.00

Hottelings 43.10276 431.02762 2.00
 Wilks .02267 431.02762 2.00
 Roys .97733
 Note.. F statistics are exact.

Output:

Cell Means and Standard Deviations
 Variable . . Y1
 Mean Std.Dev. N 95 Conv.
 percent Interval

For entire sample 6.227 2.092 22 5.300 7.155

EFFECT .. CONSTANT (Cont.)
 Univariate F-test with (1,21) D. F.

Variable Hypoth. SS Error SS Hypoth. MS
 Y1 982.22727 160.77273 982.22727
 Y2 38892.04545 918.95455 38892.04545

Variable . . Y2

Mean Std.Dev. N 95 Conv.
 percent Interval
 For entire sample 46.636 7.644 22 43.247 50.026

Variable Error MS F Sig. of F
 Y1 7.65584 128.29771 .000
 Y2 43.75974 888.76317 .000

TABEL 3.3 Nilai Y₁ (Error Detection Task) dan Y₂ (Degrees of Reading Power) untuk kelompok DRTA (Directed Reading and Think Aloud)

Y ₁	Y ₂
6	27
6	36
5	51
5	51
0	50
6	55
6	52
11	48
6	53
8	45
8	47
3	51
7	30
7	50
6	55
9	48
7	52
6	46
7	36
6	45
6	49
6	49

WITHIN CELLS Correlations with Std. Devs. on Diagonal

Y1 Y2
 Y1 2.09152
 Y2 -.15542 7.64414

Statistics for WITHIN CELLS correlations

Log(Determinant) = -.02445
 Bartlett test of sphericity = .47682 with 1 D. F.
 Significance = .490
 F(max) criterion = 13.35774 with (2,21) D. F.

WITHIN CELLS Sum-of-Squares and Cross-Products

Y1 Y2
 Y1 91.86364
 Y2 -52.18182 1227.09091

***** Analysis of Variance *****

EFFECT .. CONSTANT
 Multivariate Test of Significance (S = 1, M = 0, N = 9)

Test Name Value Exact F Hypoth. DF
 Pillais .98231 555.37571 2.00
 Hottelings 55.53757 555.37571 2.00
 Wilks .01769 555.37571 2.00
 Roys .98231
 Note.. F statistics are exact.

EFFECT .. CONSTANT (Cont.)

Univariate F-test with (1,21) D. F.

Variable	Hypoth. SS	Error SS	Hypoth. MS
Y1	853.13636	91.86364	853.13636
Y2	47848.90909	1227.09091	47848.90909

Variable	Error MS	F	Sig. of F
Y1	4.37446	195.02672	.000
Y2	58.43290	818.86931	.000

IV. KESIMPULAN

Statistika Bartlett-Pillai, Hotelling-Lawley, Wilk, dan Roy akan memberikan hasil yang tidak jauh berbeda. Metode statistika yang paling baik adalah yang paling sesuai dengan data yang diolah. Urutan tes statistik menurut Rencher (2002, pp. 176-178) untuk konstruksi tunggal adalah Roy > Hotelling-Lawley > Wilks > Bartlett-Pillai, sementara untuk konstruksi ganda adalah Bartlett-Pillai > Wilks > Hotelling-Lawley > Roy.

REFERENSI

- [1] Carl J. Huberty, *Applied MANOVA and Discriminant Analysis*, 2nd ed. New York: A John Willey & Sons, 2006.
- [2] http://en.wikipedia.org/wiki/Multivariate_analysis_of_variance
16 Desember 2010
- [3] http://en.wikiversity.org/wiki/Advanced_ANOVA/MANOVA
16 Desember 2010
- [4] <http://www.statisticssolutions.com/methods-chapter/statistical-tests/manova/>
16 Desember 2010

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 17 Desember 2010

ttd



Muhammad Ridhwan Ramdhani (18209019)