

ANALISIS MULTIVARIAT

Pertemuan 3-4

Wahyu Sri Utami, S.Si.,M.Sc.

Program Studi Sains Data
Universitas Teknologi Yogyakarta

Multivariate Analysis of Varians (MANOVA)

- MANOVA adalah generalisasi dari ANOVA untuk situasi yang memiliki beberapa variabel terikat
- MANOVA memiliki kemiripan asumsi dengan ANOVA tetapi diperluas untuk kasus multivariat.
- Perbedaan antara ANOVA dan MANOVA terletak pada jumlah variabel dependennya.

Multivariate Analysis of Varians (MANOVA)

- ANOVA digunakan untuk mengetahui apakah terdapat perbedaan pengaruh perlakuan terhadap satu variabel dependen, sedangkan MANOVA digunakan untuk mengetahui apakah terdapat perbedaan pengaruh terhadap lebih dari satu variabel dependen

Perbandingan an Anova dan Manova

ANOVA <i>(Univariate Analysis of Variance)</i>	MANOVA <i>(Multivariate Analysis of Variance)</i>
Hanya mengkaji berbagai pengaruh percobaan yang dilakukan terhadap respons tunggal (satu unit variabel respons).	Mengkaji pengaruh dari berbagai perlakuan yang dicobakan terhadap respons ganda (lebih dari satu respons).
Ketergantungan di antara variabel respons tidak menjadi perhatian utama karena pada dasarnya terdapat anggapan bahwa variabel-variabel respons saling bebas satu sama lain sehingga pengkajian struktur keragaman hanya dilakukan terhadap setiap variabel respons secara terpisah.	Mempertimbangkan adanya ketergantungan antar variabel respons sehingga cocok digunakan untuk pengkajian pengaruh dari berbagai perlakuan terhadap lebih dari satu respons.

Tujuan Manova

- Tujuan dari MANOVA adalah untuk menguji apakah vektor rerata 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 dependennya

- Penggunaan MANOVA memiliki keunggulan, yaitu mampu menganalisis semua variabel terikat secara simultan, sehingga dapat memperkecil kesalahan tipe I (α) dalam pengambilan keputusan uji statistik
- MANOVA mampu mendeteksi dan mengungkapkan perbedaan yang tidak ditampilkan ANOVA pada masing-masing variabel terikat secara terpisah.

- MANOVA menguji ada tidaknya perbedaan rata-rata dari dua atau lebih variabel tak bebas secara simultan (simultaneously) berdasarkan kelompok-kelompok pada variabel bebas.
- Perlu diperhatikan bahwa pada MANOVA, variabel bebas (independent variable) bersifat non-metrik (terdiri dari beberapa kelompok/kategori), sedangkan variabel bebas bersifat metrik (interval atau rasio)

Multivariate Analysis of Variance

$$Y_1 + Y_2 + Y_3 + \cdots + Y_6 = X_1 + X_2 + X_3 + \cdots + X_n$$

Manova

Kadang pada lebih dari dua populasi diambil sampel acak, berupa kumpulan dari tiap g populasi, sebagai berikut

Populasi 1: $X_{11}, X_{12}, \dots, X_{1n_1}$

Populasi 2: $X_{21}, X_{22}, \dots, X_{2n_2}$

\vdots \vdots

Populasi g : $X_{g1}, X_{g2}, \dots, X_{gn_g}$

- MANOVA digunakan untuk meneliti apakah vektor rata-rata populasi itu sama atau tidak,
- jika tidak komponen rata-rata yang mana yang berbeda secara signifikan

Manova One Way

- Untuk membandingkan vektor rata-rata populasi g berdasarkan bentuk model Anova One Way sbb:

$$Y_{ij} = \mu + \tau_i + e_{ij}$$

Dengan $i = 1, 2, \dots, g$ dan $j = 1, 2, \dots, n$

Y_{ij} : nilai pengamatan (respon Tunggal) dari ulangan ke- j terhadap perlakuan ke- i

μ : nilai rata-rata

τ_i : pengaruh dari perlakuan ke- i terhadap respon

e_{ij} : pengaruh galat yang timbul pada perulangan ke- j dan perlakuan ke- i yang diasumsikan independent dan berdistribusi $N_p(0, \Sigma)$

Asumsi Dalam Manova One Way

- $X_{i1}, X_{i2}, \dots, X_{in}$ adalah sampel acak berukuran n_i dari sebuah populasi dengan rata-rata $\mu_l, l=1, 2, \dots, g$
- Sampel acak dari populasi yang beda adalah independen
- Semua populasi memiliki matrik kovarian bersama Σ
- Tiap populasi adalah normal multivariat

Hipotesis Pada Manova One Way

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_g$$

$$H_1 : \exists i \neq j, \text{ sehingga } \mu_i \neq \mu_j$$

- Terdapat beberapa statistic yang digunakan dalam Manova yaitu:
 - a. Uji Wilk's Lambda
 - b. Uji Pillai Trace
 - c. Uji Lawley-Hotelling (T^2 -Hotelling)
 - d. Uji Roy's Largest Root

Prosedur dalam Analisis Manova One Way Dengan Uji Wilk's Lambda

- Menyusun hipotesis H_0 dan H_1
- Menentukan Faktor Koreksi (FK) untuk respon Y_1 dan Y_2

➤ FK untuk Y_1 :

$$FK_{11} = \frac{\left(\sum_{l=1}^g Y_{1lj}\right)^2}{n}$$

➤ FK untuk Y_2 :

$$FK_{22} = \frac{\left(\sum_{l=1}^g Y_{2lj}\right)^2}{n}$$

➤ FK untuk Y_1 dan Y_2 :

$$FK_{12} = \frac{\left(\sum_{l=1}^g Y_{1lj}\right) \left(\sum_{l=1}^g Y_{2lj}\right)}{n}$$

Dengan
 $n = \text{banyak data total}$
 $g = \text{banyak kelompok}$

Prosedur dalam Analisis Manova One Way Dengan Uji Wilk's Lambda

- Menghitung Jumlah Kuadrat Total Terkoreksi (JKT) dan Jumlah Hasil Kali Total Terkoreksi (JHKT) untuk respon Y_1 dan Y_2

➤ JKT untuk respon Y_1 : $JKT_{11} = \sum_{l=1}^g \sum_{j=1}^{n_l} Y_{1lj}^2 - FK_{11}$

➤ JKT untuk respon Y_2 : $JKT_{22} = \sum_{l=1}^g \sum_{j=1}^{n_l} Y_{2lj}^2 - FK_{22}$

➤ JHKT untuk Y_1 dan Y_2 :

$$JHKT_{12} = \left(\sum_{l=1}^g \sum_{j=1}^{n_l} (Y_{1lj} \cdot Y_{2lj}) \right) - FK_{12}$$

Prosedur dalam Analisis Manova One Way Dengan Uji Wilk's Lambda

- Menghitung Jumlah Kuadrat Perlakuan Terkoreksi (JKP) dan Jumlah Hasil Kali Perlakuan Terkoreksi (JHKP) untuk Y_1 dan Y_2

➤ JKP untuk Y_1 :
$$JKP_{11} = \sum_{l=1}^g \left(\frac{(\sum_{j=1}^{n_l} Y_{1lj})^2}{n_l} \right) - FK_{11}$$

➤ JKP untuk Y_2 :
$$JKP_{22} = \sum_{l=1}^g \left(\frac{(\sum_{j=1}^{n_l} Y_{2lj})^2}{n_l} \right) - FK_{22}$$

➤ JHKP untuk Y_1 dan Y_2 :
$$JHKP_{12} = \sum_{l=1}^g \frac{(\sum_{j=1}^{n_l} Y_{1lj})(\sum_{j=1}^{n_l} Y_{2lj})}{n_l} - FK_{12}$$

Prosedur dalam Analisis Manova One Way Dengan Uji Wilk's Lambda

- Menghitung Jumlah Kuadrat Galat(JKG) dan Jumlah Hasil Kali Galat (JHKG) untuk Y_1 dan Y_2

➤ **JKG untuk Y_1** : $JKG_{11} = JKT_{11} - JKP_{11}$

➤ **JKG untuk Y_2** : $JKP_{22} = JKT_{22} - JKP_{22}$

➤ **JHKG untuk Y_1 dan Y_2** : $JHKG_{12} = JHKT_{12} - JHKG_{12}$

Prosedur dalam Analisis Manova One Way Dengan Uji Wilk's Lambda

- Tabel Manova

Sumber Keragaman	Derajat Kebebasan	Matriks Jumlah Kuadrat Hasil Kali
Perlakuan (Treatment)	$g - 1$	$P = \sum_{i=1}^g n_l (y_i - \bar{y})(y_i - \bar{y})^2$
Galat (Error)	$\sum_{i=1}^g n_l - g$	$G = \sum_{i=1}^g \sum_{j=1}^{n_l} (y_{ij} - \bar{y})(y_{ij} - \bar{y})^2$
Total (Teroreksi dengan nilai Tengah)	$\sum_{i=1}^g n_l - 1$	$P + G = \sum_{i=1}^g \sum_{j=1}^{n_l} (y_{ij} - \bar{y})(y_{ij} - \bar{y})$

Prosedur dalam Analisis Manova One Way Dengan Uji Wilk's Lambda

- Statistik Uji Wilk's Lambda ($\Lambda - Wilk$)

$$\Lambda = \frac{|G|}{|G + P|} = \frac{|G|}{|T|}$$

Dengan $|G|$ menyatakan determinan matriks G

$|T|$ menyatakan determinan matriks R .

Prosedur dalam Analisis Manova One Way Dengan Uji Wilk's Lambda

- Transformasi dari Besaran Wilk's Lambda ke nilai distribusi F

Banyak Variabel	Banyak grup	Tranformasi F	Derajat Bebas
$p = 1$	$g \geq 2$	$\left(\frac{1 - \Lambda}{\Lambda}\right) \left(\frac{\sum n_l - g}{g - 1}\right)$	$g - 1; \sum n_l - g$
$p = 2$	$g \geq 2$	$\left(\frac{1 - \sqrt{\Lambda}}{\sqrt{\Lambda}}\right) \left(\frac{\sum n_l - g - 1}{g - 1}\right)$	$2(g - 1); 2(\sum n_l - g - 1)$
$p \geq 1$	$g = 2$	$\left(\frac{1 - \Lambda}{\Lambda}\right) \left(\frac{\sum n_l - p - 1}{p}\right)$	$p; \sum n_l - p - 1$
$p \geq 1$	$g = 3$	$\left(\frac{1 - \sqrt{\Lambda}}{\sqrt{\Lambda}}\right) \left(\frac{\sum n_l - p - 2}{p}\right)$	$2p ; 2(\sum n_l - p - 2)$

Prosedur dalam Analisis Manova One Way Dengan Uji Wilk's Lambda

- Kriteria Pengujian: H_0 ditolak jika $F_{hitung} > F_{tabel}$
- Kesimpulan

Contoh

Seorang peneliti bidang kesehatan melakukan percobaan untuk melihat hubungan antara aktifitas metabolis di antara kelinci-kelinci percobaan dan daya tahan terhadap kuman TBC. Peneliti menetapkan 4 perlakuan sbb:

P_1 : Kontrol (tidak divaksin)

P_2 : Diinfeksi kuman TBC selama aktifitas metabolic rendah

P_3 : Diinfeksi kuman TBC selama aktifitas metabolic tinggi

P_4 : Diinfeksi kuman TBC selama aktifitas metabolic normal, tetapi diirradiasi dengan 400 rontgens terlebih dahulu

Perlakuan P_1 dan P_2 diulang sampai 7 kali ($n_1 = n_2 = 7$), perlakuan P_3 diulang sampai 5 kali ($n_3 = 5$) dan P_4 diulang sebanyak 2 kali ($n_4 = 2$).

Misalkan Y_1 adalah banyaknya basil yang hidup per tubercle formed (mm);
 Y_2 adalah banyaknya basil yang hidup per tubercle size (mm).

Berikut data hasil pengamatan:

Ulangan	P_1		P_2		P_3		P_4	
	Y_1	Y_2	Y_1	Y_2	Y_1	Y_2	Y_1	Y_2
1	24	3,5	7,4	3,5	16,4	3,2	25,1	2,7
2	13,3	3,5	13,2	3	24	2,5	5,9	2,3
3	12,2	4	8,5	3	53	1,5		
4	14	4	10,1	3	32,7	2,6		
5	22,2	3,6	9,3	2	42,8	2		
6	16,1	4,3	8,5	2,5				
7	27,9	5,2	4,3	1,5				

Penyelesaian

- Hipotesis

H_0 : tidak ada perbedaan antara banyaknya basil yang hidup per tubercle formed (Y_1) dan banyaknya basil yang hidup per tubercle size (Y_2) dengan respon P_1, P_2, P_3 dan P_4

H_1 : ada perbedaan antara banyaknya basil yang hidup per tubercle formed (Y_1) dan banyaknya basil yang hidup per tubercle size (Y_2) dengan respon P_1, P_2, P_3 dan P_4

H_0 : tidak ada perbedaan antara banyaknya basil yang hidup per tubercle formed (Y_1) yang diakibatkan oleh pemberian vaksin kumat TBC pada kelinci

H_1 : ada perbedaan antara banyaknya basil yang hidup per tubercle formed (Y_1) yang diakibatkan oleh pemberian vaksin kumat TBC pada kelinci

H_0 : tidak ada perbedaan antara banyaknya basil yang hidup per tubercle size (Y_2) yang diakibatkan oleh pemberian vaksin kumat TBC pada kelinci

H_1 : ada perbedaan antara banyaknya basil yang hidup per tubercle size (Y_2) yang diakibatkan oleh pemberian vaksin kumat TBC pada kelinci

- Tingkat signifikansi: $\alpha = 0,01$
- Menentukan Faktor Koreksi

Ulangan	P1					P2					P3					P4				
	Y_1	Y_1^2	Y_2	Y_2^2	Y_1Y_2	Y_1	Y_1^2	Y_2	Y_2^2	Y_1Y_2	Y_1	Y_1^2	Y_2	Y_2^2	Y_1Y_2	Y_1	Y_1^2	Y_2	Y_2^2	Y_1Y_2
1	24	576	3.5	12.25	84	7.4	54.76	3.5	12.25	25.9	16.4	268.96	3.2	10.2	52.48	25	630.01	2.7	7.29	67.77
2	13.3	176.89	3.5	12.25	46.55	13.2	174.24	3	9	39.6	24	576	2.5	6.25	60	5.9	34.81	2.3	5.29	13.57
3	12.2	148.84	4	16	48.8	8.5	72.25	3	9	25.5	53	2809	1.5	2.25	79.5					
4	14	196	4	16	56	10.1	102.01	3	9	30.3	32.7	1069.29	2.6	6.76	85.02					
5	22.2	492.84	3.6	12.96	79.92	9.3	86.49	2	4	18.6	42.8	1831.84	2	4	85.6					
6	16.1	259.21	4.3	18.49	69.23	8.5	72.25	2.5	6.25	21.25										
7	27.9	778.41	5.2	27.04	145.1	4.3	18.49	1.5	2.25	6.45										
Total	129.7	2628.19	28.1	114.99	529.6	61.3	580.49	18.5	51.75	167.6	168.9	6555.09	11.8	29.5	362.6	31	664.82	5	12.58	81.34

➤ Faktor koreksi untuk Y_1

$$FK_{11} = \frac{\sum (Y_{1ij})^2}{n} = \frac{(129,7 + 61,3 + 168,9 + 31)^2}{21} = \frac{390,9^2}{21} = 7276,3243$$

➤ Faktor koreksi untuk Y_2

$$FK_{22} = \frac{\sum (Y_{2ij})^2}{n} = \frac{(28,1 + 18,5 + 11,8 + 5)^2}{21} = \frac{63,4^2}{21} = 191,4076$$

➤ Faktor koreksi untuk Y_1 dan Y_2

$$FK_{12} = \frac{\sum (Y_{1ij}) \sum (Y_{2ij})}{n} = \frac{(390,9) \cdot (63,4)}{21} = 1180,146$$

Menentukan JKT dan JHKT

➤ JKT untuk respon Y_1 :

$$JKT_{11} = \sum_{l=1}^4 \sum_{j=1}^{n_l} Y_{1ij}^2 - FK_{11} = (2628,19 + 580,49 + 6555,09 + 664,82) - 7276,3243 = 3152,2657$$

➤ JKT untuk respon Y_2 :

$$JKT_{22} = \sum_{i=1}^4 \sum_{j=1}^{n_l} Y_{2ij}^2 - FK_{22} = (114,99 + 51,75 + 29,5 + 12,58) - 191,4076 = 17,4124$$

➤ JHKT untuk Y_1 dan Y_2 :

$$JHKT_{12} = \left(\sum_{i=1}^4 \sum_{j=1}^{n_l} (Y_{1ij} \cdot Y_{2ij}) \right) - FK_{12} = (529,6 + 167,6 + 362,6 + 81,34) - 1180,146 = -39,03$$

Menentukan JKP dan JHKP

Perlakuan	ΣY_1	$\frac{(\Sigma Y_1)^2}{n_i}$	ΣY_2	$\frac{(\Sigma Y_2)^2}{n_i}$	$\frac{\Sigma Y_1 \Sigma Y_2}{n_i}$
P1	129.7	2403.16	28.1	112.801	520.7
P2	61.3	536.813	18.5	48.8929	162
P3	168.9	5705.44	11.8	27.848	398.6
P4	31	480.5	5	12.5	77.5
Σ		9125.91		202.042	1159

➤ **JKP untuk Y_1** : $JKP_{11} = \sum_{l=1}^g \left(\frac{(\sum_{j=1}^{n_l} Y_{1lj})^2}{n_l} \right) - FK_{11} = 9125,91 - 7276,3243 = 1849,6$

➤ **JKP untuk Y_2** : $JKP_{22} = \sum_{l=1}^g \left(\frac{(\sum_{j=1}^{n_l} Y_{2lj})^2}{n_l} \right) - FK_{22} = 202,042 - 191,4076 = 10,635$

➤ **JHKP untuk Y_1 dan Y_2** : $JHKP_{12} = \sum_{l=1}^g \frac{(\sum_{j=1}^{n_l} Y_{1lj})(\sum_{j=1}^{n_l} Y_{2lj})}{n_l} - FK_{12} = 1159 - 1180,146 = -21,38$

Hitung JKG dan JHKG

➤ **JKG untuk Y_1** : $JKG_{11} = JKT_{11} - JKP_{11} = 3152,2657 - 1849,6 = 1302,7$

➤ **JKG untuk Y_2** : $JKG_{22} = JKT_{22} - JKP_{22} = 17,412 - 10,635 = 6,78$

➤ **JHKG untuk Y_1 dan Y_2** : $JHKG_{12} = JHKT_{12} - JHKP_{12} = -39,03 - (-21,38) = -17,64$

Tabel Manova

Sumber Keragaman	Derajat Kebebasan	Matriks Jumlah Kuadrat Hasil Kali
Perlakuan (Treatment)	$g - 1 = 4 - 1 = 3$	$P = \sum_{l=1}^g n_l (y_l - \bar{y})(y_l - \bar{y})^2 = \begin{pmatrix} JKP_{11} & JHKP_{12} \\ JHKP_{21} & JKP_{22} \end{pmatrix}$ $= \begin{pmatrix} 1849,6 & -21,38 \\ -21,38 & 10,635 \end{pmatrix}$
Galat (Error)	$\sum_{i=1}^g n_l - g = 21 - 4 = 17$	$G = \sum_{l=1}^g \sum_{j=1}^{n_l} (y_{lj} - \bar{y})(y_{lj} - \bar{y})^2 = \begin{pmatrix} JKG_{11} & JHKG_{12} \\ JHKG_{21} & JKG_{22} \end{pmatrix}$ $= \begin{pmatrix} 1302,7 & -17,64 \\ -17,64 & 6,78 \end{pmatrix}$
Total (Teroreksi dengan nilai Tengah)	$\sum_{i=1}^g n_l - 1 = 21 - 1 = 20$	$T = P + G = \sum_{l=1}^g \sum_{j=1}^{n_l} (y_{lj} - \bar{y})(y_{lj} - \bar{y}) = \begin{pmatrix} JKT_{11} & JHKT_{12} \\ JHKT_{21} & JKT_{22} \end{pmatrix}$ $= \begin{pmatrix} 3152,3 & -39,03 \\ -39,03 & 17,412 \end{pmatrix}$

Statistik Uji Wilk's

Diperhatikan bahwa

$$G = \begin{vmatrix} 1302,7 & -17,64 \\ -17,64 & 6,78 \end{vmatrix} = 8517,9$$

$$T = \begin{vmatrix} 3152,3 & -39,03 \\ -39,03 & 17,412 \end{vmatrix} = 53365$$

Dengan demikian diperoleh

$$\Lambda = \frac{|G|}{|G + P|} = \frac{|G|}{|T|} = \frac{8517,9}{53365} = 0,1596$$

Transformasi Uji F

Karena $p = 2$ dan $g = 4 \geq 2$ maka digunakan transformasi F sbb

$$\left(\frac{1 - \sqrt{\Lambda}}{\sqrt{\Lambda}} \right) \left(\frac{\Sigma n_l - g - 1}{g - 1} \right) = \left(\frac{1 - \sqrt{0,1596}}{\sqrt{0,1596}} \right) \left(\frac{21 - 4 - 1}{4 - 1} \right) = 8,0161$$

Dengan

$$df_1 = 2(g - 1) = 2(4 - 1) = 6$$

$$df_2 = 2(\Sigma n_l - g - 1) = 2(21 - 4 - 1) = 32$$

Menggunakan Tabel F diperoleh $F_{\alpha; df_1, df_2} = F_{0,01; 6; 32} = 3,43$

Kriteria Penolakan H_0 & Kesimpulan

- H_0 ditolak jika $F_{hitung} > F_{tabel}$.
- Karena $F_{hitung} = 8,0161 > F_{tabel} = 3,43$ maka H_0 ditolak.

- Kesimpulan:

Jadi, ada perbedaan antara banyaknya basil yang hidup per tubercle formed (Y_1) dan banyaknya basil yang hidup per tubercle size (Y_2) dengan respon P_1, P_2, P_3 dan P_4 . Selain itu, ada perbedaan antara banyaknya basil yang hidup per tubercle formed (Y_1) dan per tubercle size (Y_2) yang diakibatkan oleh pemberian vaksin kumat TBC pada kelinci.

Latihan

An experiment was performed to investigate four different methods for teaching school children multiplication (M) and addition (A) of two four-digit numbers. The data for four independent groups of students are summarized in Table 4.5.3. Are there any significant differences between addition and multiplication skills within the various groups?

TABLE 4.5.3. Teaching Methods

Group 1		Group 2		Group 3		Group 4	
A	M	A	M	A	M	A	M
97	66	76	29	66	34	100	79
94	61	60	22	60	32	96	64
96	52	84	18	58	27	90	80
84	55	86	32	52	33	90	90
90	50	70	33	56	34	87	82
88	43	70	32	42	28	83	72
82	46	73	17	55	32	85	67
65	41	85	29	41	28	85	77
95	58	58	21	56	32	78	68
90	56	65	25	55	29	86	70
95	55	89	20	40	33	67	67
84	40	75	16	50	30	57	57
71	46	74	21	42	29	83	79
76	32	84	30	46	33	60	50
90	44	62	32	32	34	89	77
77	39	71	23	30	31	92	81
61	37	71	19	47	27	86	86
91	50	75	18	50	28	47	45
93	64	92	23	35	28	90	85
88	68	70	27	47	27	86	65