S1 Sains Data
FMIPA UNESA

Analisis Multivariat – Materi 06

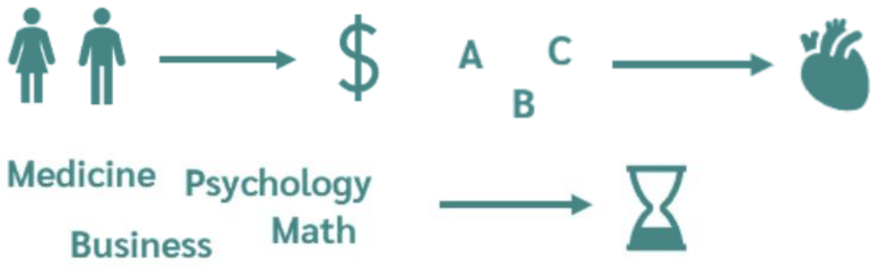
MANOVA

Prodi S1 Sains Data
Universitas Negeri Surabaya
2025

Sebelum ke MANOVA, kita mengingat kembali~

One-Way ANOVA

One-Way ANOVA is used to determine whether there are statistically significant differences between the means of **three or more independent groups** based on **one independent variable (factor).**



Assumptions

- The samples are **independent**.
- The dependent variable is **normally distributed** within each group.
- The groups have **equal variances (homoscedasticity)**.

Hypotheses

H_0: $\mu_1 = \mu_2 = \cdots = \mu_k$,

H_1: At least two of the means are not equal.

Data Structure

Treatment:	1	2	$\cdots$	i	$\cdots$	k	
	y_{11}	y_{21}	$\cdots$	y_{i1}	$\cdots$	y_{k1}	
	y_{12}	y_{22}	$\cdots$	y_{i2}	$\cdots$	y_{k2}	
	$\vdots$	$\vdots$		$\vdots$		$\vdots$	
	y_{1n}	y_{2n}	$\cdots$	y_{in}	$\cdots$	y_{kn}	
Total	$Y_{1.}$	$Y_{2.}$	$\cdots$	$Y_{i.}$	$\cdots$	$Y_{k.}$	$Y_{..}$
Mean	$\bar{y}_{1.}$	$\bar{y}_{2.}$	$\cdots$	$\bar{y}_{i.}$	$\cdots$	$\bar{y}_{k.}$	$\bar{y}_{..}$

Test Statistic *F*

Tabel ANOVA

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Treatments	SSA	$k-1$	$s_1^2 = \frac{SSA}{k-1}$	$\boxed{f = \frac{s_1^2}{s^2}}$
Error	SSE	$k(n-1)$	$s^2 = \frac{SSE}{k(n-1)}$	
Total	SST	$kn-1$		

$$\sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{..})^2 = n\sum_{i=1}^{k}(\bar{y}_{i.}-\bar{y}_{..})^2 + \sum_{i=1}^{k}\sum_{j=1}^{n}(y_{ij}-\bar{y}_{i.})^2$$

$$SST = SSA + SSE.$$

Between-Group Sum of Squares (SSA)
Within-Group Sum of Squares (SSE)

Reject *H₀* if, $f > f_\alpha[k-1, k(n-1)]$.

S1 Sains Data
FMIPA UNESA

One-Way ANOVA
(Unequal Sample Sizes)

Hypotheses

H_0: $\mu_1 = \mu_2 = \cdots = \mu_k$,

H_1: At least two of the means are not equal.

Example

Table 13.4: Serum Alkaline Phosphatase Activity Level

G-1		G-2	G-3	G-4
49.20	97.50	97.07	62.10	110.60
44.54	105.00	73.40	94.95	57.10
45.80	58.05	68.50	142.50	117.60
95.84	86.60	91.85	53.00	77.71
30.10	58.35	106.60	175.00	150.00
36.50	72.80	0.57	79.50	82.90
82.30	116.70	0.79	29.50	111.50
87.85	45.15	0.77	78.40	
105.00	70.35	0.81	127.50	
95.22	77.40			

G-1 has 20 samples
G-2 has 9 samples
G-3 has 9 samples
G-4 has 7 samples

Test Statistic *F*

Tabel ANOVA

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Treatments	SSA	$k-1$	$s_1^2 = \frac{SSA}{k-1}$	$f = \frac{s_1^2}{s^2}$
Error	SSE	$\boxed{N-k}$	$s^2 = \frac{SSE}{N-k}$	
Total	SST	$\boxed{N-1}$		

where, $N = \sum_{i=1}^{k} n_i$

The difference between equal and unequal sample size

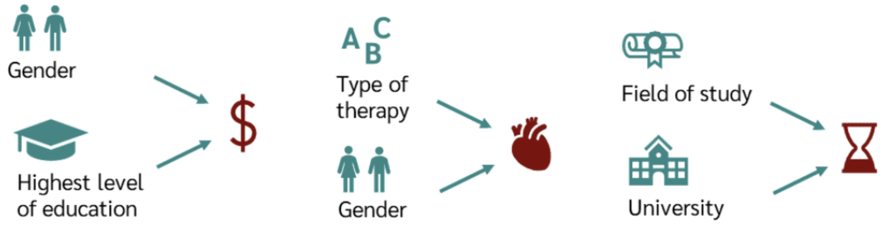
$$SST = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{..})^2, \quad \boxed{SSA = \sum_{i=1}^{k} n_i(\bar{y}_{i.} - \bar{y}_{..})^2}, \quad SSE = SST - SSA$$

Reject *H₀* if,

$f > f_\alpha[k-1,\ N-k\ ]$.

Two-Way ANOVA

Two-Way ANOVA is used to examine the effects of **two independent variables (factors)** on a dependent variable.



Data Structure (with n observations in each group)

A	B: 1	2	$\cdots$	b	Total	Mean
1	y_{111}	y_{121}	$\cdots$	y_{1b1}	$Y_{1..}$	$\bar{y}_{1..}$
	y_{112}	y_{122}	$\cdots$	y_{1b2}		
	$\vdots$	$\vdots$		$\vdots$		
	y_{11n}	y_{12n}	$\cdots$	y_{1bn}		
2	y_{211}	y_{221}	$\cdots$	y_{2b1}	$Y_{2..}$	$\bar{y}_{2..}$
	y_{212}	y_{222}	$\cdots$	y_{2b2}		
	$\vdots$	$\vdots$		$\vdots$		
	y_{21n}	y_{22n}	$\cdots$	y_{2bn}		
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
a	y_{a11}	y_{a21}	$\cdots$	y_{ab1}	$Y_{a..}$	$\bar{y}_{a..}$
	y_{a12}	y_{a22}	$\cdots$	y_{ab2}		
	$\vdots$	$\vdots$		$\vdots$		
	y_{a1n}	y_{a2n}	$\cdots$	y_{abn}		
Total	$Y_{.1.}$	$Y_{.2.}$	$\cdots$	$Y_{.b.}$	$Y_{...}$	
Mean	$\bar{y}_{.1.}$	$\bar{y}_{.2.}$	$\cdots$	$\bar{y}_{.b.}$		$\bar{y}_{...}$

Hypotheses

1. H_0': $\alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$,
   H_1': At least one of the α_i is not equal to zero. *Factor A*
2. H_0'': $\beta_1 = \beta_2 = \cdots = \beta_b = 0$,
   H_1'': At least one of the β_j is not equal to zero. *Factor B*
3. H_0''': $(\alpha\beta)_{11} = (\alpha\beta)_{12} = \cdots = (\alpha\beta)_{ab} = 0$,
   H_1''': At least one of the $(\alpha\beta)_{ij}$ is not equal to zero. *Interaction*

Jika tidak ada interaksi, tidak perlu dilakukan pengujian

Assumptions

- The **samples are independent**.
- The dependent variable is **normally distributed** within each group.
- The groups have **equal variances (homoscedasticity)**.
- The **two factors are independent.**

Two-Way ANOVA

Statistical Testing

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	Computed f
Main effect:				
A	SSA	$a-1$	$s_1^2 = \frac{SSA}{a-1}$	$f_1 = \frac{s_1^2}{s^2}$
B	SSB	$b-1$	$s_2^2 = \frac{SSB}{b-1}$	$f_2 = \frac{s_2^2}{s^2}$
Two-factor interactions:				
AB	$SS(AB)$	$(a-1)(b-1)$	$s_3^2 = \frac{SS(AB)}{(a-1)(b-1)}$	$f_3 = \frac{s_3^2}{s^2}$
Error	SSE	$ab(n-1)$	$s^2 = \frac{SSE}{ab(n-1)}$	
Total	SST	$abn-1$		

Jika tidak ada interaksi, maka ***source of variation AB tidak perlu ada.***

Jika n = 1, *maka*
df Error = (a-1)(b-1) sehingga
Mean Square Error = SSE / (a-1)(b-1)

Critical region

Reject *H0* if

$$f_1 > f_\alpha[a-1, ab(n-1)].$$
$$f_2 > f_\alpha[b-1, ab(n-1)].$$
$$f_3 > f_\alpha[(a-1)(b-1), ab(n-1)].$$

Two-Way ANOVA

Sum-of-Squares

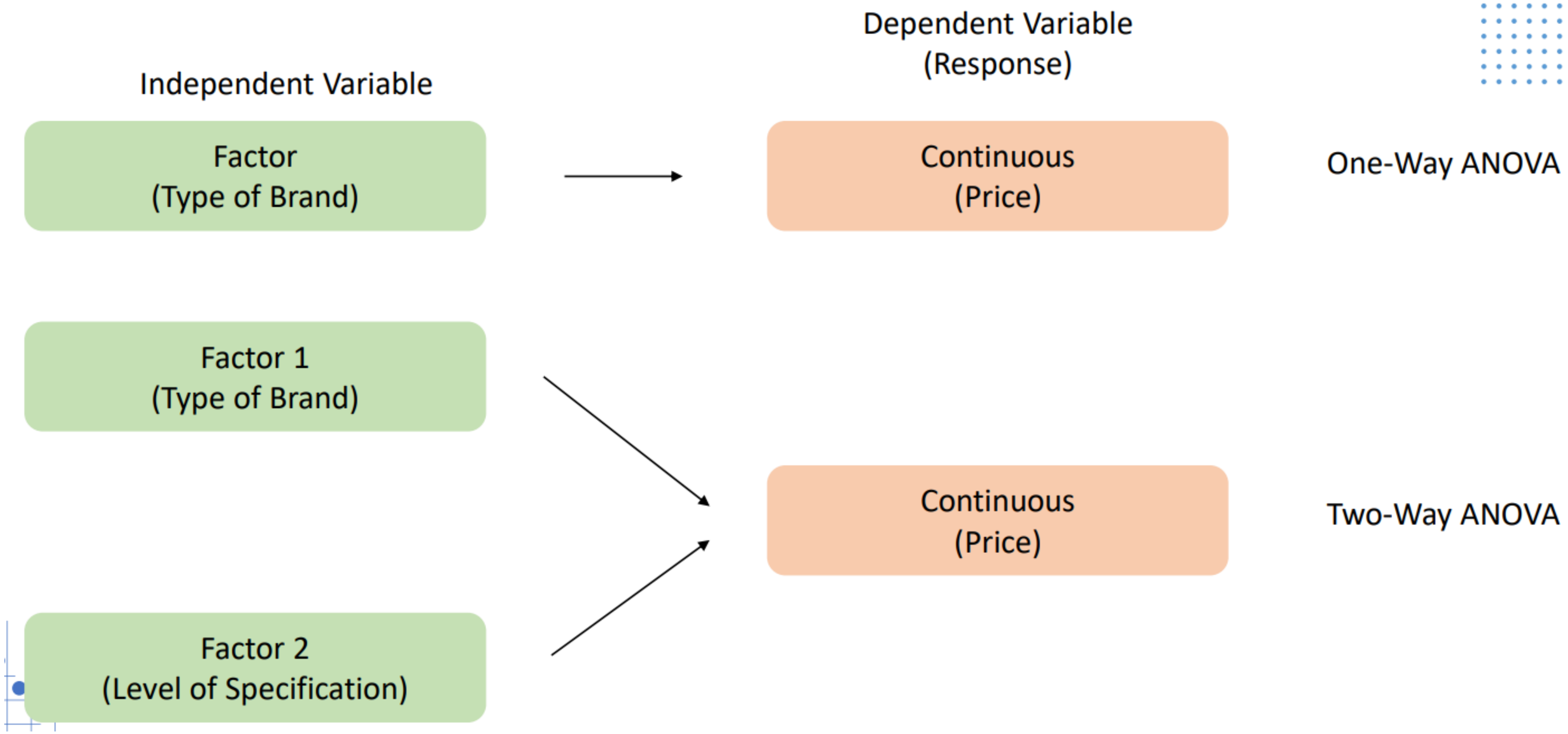
$$\underbrace{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{...})^2}_{\text{SST}} = \underbrace{bn\sum_{i=1}^{a}(\bar{y}_{i..}-\bar{y}_{...})^2}_{\text{SSA}} + \underbrace{an\sum_{j=1}^{b}(\bar{y}_{.j.}-\bar{y}_{...})^2}_{\text{SSB}} + \underbrace{n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij.}-\bar{y}_{i..}-\bar{y}_{.j.}+\bar{y}_{...})^2}_{\text{SS(AB)}} + \underbrace{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk}-\bar{y}_{ij.})^2}_{\text{SSE}}$$

Degrees of Freedom

$$abn-1=(a-1)+(b-1)+(a-1)(b-1)+ab(n-1).$$

Table of Contents

ANOVA


Independent Variable
Dependent Variable
(Response)
Factor
(Type of Brand)
Continuous
(Price)
One-Way ANOVA
Factor 1
(Type of Brand)
Factor 2
(Level of Specification)
Continuous
(Price)
Two-Way ANOVA

MANOVA

Multivariate

Independent Variable

Dependent Variable

Categorical (Factor)

Factor
(Type of Brand)

Continuous
(Price)

Continuous
(Sales)

Multivariate Analysis of Variance (MANOVA)

- The Multivariate Analysis of Variance (MANOVA) is an **extension of the ANOVA**

- While we only deal with **ONE DV in ANOVA**, **MANOVA accounts for multiple DVs at once**

- It wants to know if there are mean differences across groups on multiple DVs; it is suitable to test related DVs – e.g., **testing depression, anxiety, and stress across groups at one go**

What is MANOVA?

- A statistical method for testing whether the vector of means (variate) across groups on multiple variables are equal (i.e., the probability that any differences in the variate means across several groups are due solely to sampling error).
- Variables in ANOVA (Analysis of Variance):
  - **Dependent variables are metric**.
  - **Independent variable(s) is nominal** with two or more levels – also called treatment, manipulation, or factor.
- **One-way MANOVA**: only one independent variable with two or more levels.
- **Two-way MANOVA**: two independent variables each with two or more levels (factorial design). One or more could be control variables.

Example:
One-Way MANOVA

I am interested in ***finding out if coffee consumption affects anxiety and fatigue levels.***

To test this, I shall recruit 100 participants and randomly assign them into 2 groups: an experimental group who will drink a cup of coffee, and a control group who will drink a cup of water.

I will then ask each participant to rate their level of anxiety and fatigue.

Example:
One-Way MANOVA

In this example, we have **1 IV with 2 levels**: Coffee vs. Water

We have **2 DVs: Anxiety, Fatigue**

Thus, it is appropriate to conduct a one-way between subjects MANOVA

MANOVA ASSUMPTIONS

Assumptions Testing

1. Observations must be **independent**.
2. Variance–covariance matrices must be **equal** (or comparable) for all treatment groups.
3. The dependent variables must have a **multivariate normal** distribution.

1. Uji Dependensi

Hipotesis pengujian *Bartlett:*

$H_0 : \rho = \mathbf{I}$ (matriks korelasi adalah matriks identitas)
$H_1 : \rho \neq \mathbf{I}$ (matriks korelasi bukan matriks identitas)

Statistik ujinya :

$$T = \frac{(n-1)}{(1-\bar{r})^2}\left[\sum_{i<k}\sum (r_{ik}-\bar{r})^2 - \hat{\gamma}\sum_{k=1}^{p}(\bar{r}_k-\bar{r})^2\right]$$

dimana

$$\bar{r}_k = \frac{1}{p-1}\sum_{i=1}^{p} r_{ik} \qquad \bar{r} = \frac{2}{p(p-1)}\sum_{i<k}\sum r_{ik} \qquad \hat{\gamma} = \frac{(p-1)^2\left[1-(1-\bar{r})^2\right]}{p-(p-2)(1-\bar{r})^2}$$

$\bar{r}_k$ = rata-rata elemen diagonal pada kolom atau baris ke k dari matrik R (matriks korelasi).

$\bar{r}$ = rata-rata keseluruhan dari elemen diagonal.

Syarat penolakan yaitu tolak H_0 jika

$$T > \chi^2_{(p+1)(p-2)/2;\alpha}$$

Menguji apakah terdapat hubungan antara semua variabel dalam kasus multivariat

2. Uji Homogenitas

Hipotesis

$H_0 : \Sigma_1 = \Sigma_2 = \cdots = \Sigma_g$

H_1 : minimal terdapat satu $\Sigma_l \neq \Sigma_g \; untuk \;\; l \neq g$

Uji homogenitas

Untuk mengetahui beberapa kelompok data sampel berasal dari populasi dengan varians yang sama atau tidak.

Statistik Uji – Box's M

$$C = (1 - u)M = (1 - u)\left\{\left[\sum_\ell (n_\ell - 1)\right]\ln|\mathbf{S}_{\text{pooled}}| - \sum_\ell [(n_\ell - 1)\ln|\mathbf{S}_\ell|]\right\}$$

has an approximate χ^2 distribution with

$$\nu = g\frac{1}{2}p(p+1) - \frac{1}{2}p(p+1) = \frac{1}{2}p(p+1)(g-1)$$

where

$$u = \left[\sum_\ell \frac{1}{(n_\ell - 1)} - \frac{1}{\sum_\ell (n_\ell - 1)}\right]\left[\frac{2p^2 + 3p - 1}{6(p+1)(g-1)}\right]$$

$$M = \left[\sum_\ell (n_\ell - 1)\right]\ln|\mathbf{S}_{\text{pooled}}| - \sum_\ell [(n_\ell - 1)\ln|\mathbf{S}_\ell|]$$

$$\mathbf{S}_{\text{pooled}} = \frac{1}{\sum_\ell (n_\ell - 1)}\{(n_1 - 1)\mathbf{S}_1 + (n_2 - 1)\mathbf{S}_2 + \cdots + (n_g - 1)\mathbf{S}_g\}$$

Daerah Penolakan

reject H_0 if $C > \chi^2_{p(p+1)(g-1)/2}$

$$\Lambda = \prod_\ell \left(\frac{|\mathbf{S}_\ell|}{|\mathbf{S}_{\text{pooled}}|}\right)^{(n_\ell - 1)/2}$$

Example

The Wisconsin Department of Health and Social Services reimburses nursing homes in the state for the services provided. The department develops a set of formulas for rates for each facility, based on factors such as level of care, mean wage rate, and average wage rate in the state.

Nursing homes can be classified on the basis of ownership (**private party, nonprofit organization, and government**) and certification (**skilled nursing facility, intermediate care facility, or a combination of the two**).

X1 = cost of nursing labor
X2 = cost of dietary labor
X3 = cost of plant operation and maintenance labor
X4 = cost of housekeeping and laundry labor

n = 516 observations

Group	Number of observations
$\ell = 1$ (private)	$n_1 = 271$
$\ell = 2$ (nonprofit)	$n_2 = 138$
$\ell = 3$ (government)	$n_3 = 107$
	$\sum_{\ell=1}^{3} n_\ell = 516$

Example

S1 Sains Data FMIPA UNESA

$n_1 = 271$

$$\mathbf{S}_1 = \begin{bmatrix} .291 & & & \\ -.001 & .011 & & \\ .002 & .000 & .001 & \\ .010 & .003 & .000 & .010 \end{bmatrix} \quad \mathbf{S}_2 = \begin{bmatrix} .561 & & & \\ .011 & .025 & & \\ .001 & .004 & .005 & \\ .037 & .007 & .002 & .019 \end{bmatrix} \quad \mathbf{S}_3 = \begin{bmatrix} .261 & & & \\ .030 & .017 & & \\ .003 & -.000 & .004 & \\ .018 & .006 & .001 & .013 \end{bmatrix}$$

$n_2 = 138$

$|\mathbf{S}_1| = 2.783 \times 10^{-8}$, $\ln|\mathbf{S}_1| = -17.397$

$|\mathbf{S}_2| = 89.539 \times 10^{-8}$, $\ln|\mathbf{S}_2| = -13.926$

$|\mathbf{S}_3| = 14.579 \times 10^{-8}$, $\ln|\mathbf{S}_3| = -15.741$

$n_3 = 107$

$$\mathbf{S}_{\text{pooled}} = \frac{1}{\sum_\ell (n_\ell - 1)}\{(n_1 - 1)\mathbf{S}_1 + (n_2 - 1)\mathbf{S}_2 + \cdots + (n_g - 1)\mathbf{S}_g\}$$

$|\mathbf{S}_{\text{pooled}}| = 17.398 \times 10^{-8}$. $\ln|\mathbf{S}_{\text{pooled}}| = -15.564$.

$$u = \left[\sum_\ell \frac{1}{(n_\ell - 1)} - \frac{1}{\sum_\ell (n_\ell - 1)}\right]\left[\frac{2p^2 + 3p - 1}{6(p+1)(g-1)}\right]$$

$$= \left[\frac{1}{270} + \frac{1}{137} + \frac{1}{106} - \frac{1}{270 + 137 + 106}\right]\left[\frac{2(4^2) + 3(4) - 1}{6(4+1)(3-1)}\right] = .0133$$

$$M = \left[\sum_\ell (n_\ell - 1)\right]\ln|\mathbf{S}_{\text{pooled}}| - \sum_\ell [(n_\ell - 1)\ln|\mathbf{S}_\ell|]$$

$$= [270 + 137 + 106](-15.564) - [270(-17.397) + 137(-13.926) + 106(-15.741)]$$

$$= 289.3$$

$$C = (1 - u)M = (1 - .0133)289.3 = 285.5$$

$$\nu = \frac{1}{2}p(p+1)(g-1) = 4(4+1)(3-1)/2 = 20$$

$C > \chi^2_{20,0.05} = 31.41$

so *we reject H0*

We conclude that the covariance matrices of the cost variables associated with the three populations of nursing homes are not the same.

3. Uji Multivariat Normal

Hipotesis:
H_0 : data berdistribusi normal multivariat
H_1 : data tidak berdistribusi normal multivariat

Statistik:

$$r_q = \frac{\sum_{j=1}^{n}(d_{(j)}^2 - \bar{d}_{(j)}^2)(q_{(j)} - \bar{q})}{\sqrt{\sum_{j=1}^{n}(d_{(j)}^2 - \bar{d}_{(j)}^2)}\sqrt{\sum_{j=1}^{n}(q_{(j)} - \bar{q})^2}}$$

Tolak H_0 jika

$$r_q > r_{(n,\alpha)}$$

Finally… MANOVA

Multivariate Tests

Effect		Value	F
Intercept	Pillai's Trace	.983	1101.405[b]
	Wilks' Lambda	.017	1101.405[b]
	Hotelling's Trace	59.535	1101.405[b]
	Roy's Largest Root	59.535	1101.40
Condition	Pillai's Trace	.593	26.961[b]
	Wilks' Lambda	.407	26.961[b]
	Hotelling's Trace	1.457	26.961[b]
	Roy's Largest Root	1.457	26.961[b]

a. Design: Intercept + Condition

b. Exact statistic

How to choose the multivariate test?

Multivariate Test	Robustness					
	Sample Size	Levels of IVs	Uneven Cell Sizes	Unequal variance	Non-normal Data	Collinearity
Pillai's Trace	Small	> 2	Y	Y	Y	Low to medium
Wilk's Lambda	Medium to large	> 2	N	N	N	Low to medium
Hotelling's Trace	Medium to large	= 2	N	N	N	Low to medium
Roy's Largest Root	Medium to large	> 2	N	N	N	Medium to high

S1 Sains Data
FMIPA UNESA

ONE-WAY MANOVA

One-Way MANOVA

$$\text{Population 1}: \quad \mathbf{X}_{11}, \mathbf{X}_{12}, \ldots, \mathbf{X}_{1n_1}$$

$$\text{Population 2}: \quad \mathbf{X}_{21}, \mathbf{X}_{22}, \ldots, \mathbf{X}_{2n_2}$$

$$\vdots$$

$$\text{Population g}: \quad \mathbf{X}_{g1}, \mathbf{X}_{g2}, \ldots, \mathbf{X}_{gn_g}$$

Ubah X → Y

Model

$$\mathbf{X}_{\ell j} = \boldsymbol{\mu} + \boldsymbol{\tau}_\ell + \mathbf{e}_{\ell j}, \quad j = 1, 2, \ldots, n_\ell \quad \text{and} \quad \ell = 1, 2, \ldots, g$$

$$\mathbf{x}_{\ell j} = \bar{\mathbf{x}} + (\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}}) + (\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)$$

(Observation) Vektor pengamatan untuk grup ke-ℓ dan sampel ke-j.

Vektor mean keseluruhan

Efek perlakuan dari grup ke-ℓ dibandingkan dengan mean keseluruhan.

Vektor residual, yaitu deviasi setiap pengamatan dari rata-rata grupnya sendiri.

Hypotheses

$$H_0: \tau_1 = \tau_2 = \cdots = \tau_g = 0$$

$$H_1: \text{at least one } \tau_\ell \neq 0$$

Contoh

Ingin diuji, Apakah terdapat perbedaan secara signifikan, pada kualitas hasil pembelajaran (vektor mean nilai teori & praktik) diantara ketiga kelas suatu mata kuliah.

grup ke-ℓ	var dependen	sampel ke-j		
		1	2	3
kelas A (n1 = 3)	nilai teori	9	6	9
	nilai praktik	3	2	7
Kelas B (n2 = 2)	nilai teori	0	2	
	nilai praktik	4	0	
Kelas C (n3 = 3)	nilai teori	3	1	2
	nilai praktik	8	9	7

Rata-rata setiap grup **xℓ**

$$\bar{\mathbf{x}}_1 = \begin{bmatrix} 8 \\ 4 \end{bmatrix}, \quad \bar{\mathbf{x}}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \bar{\mathbf{x}}_3 = \begin{bmatrix} 2 \\ 8 \end{bmatrix}$$

Rata-rata keseluruhan $\bar{\mathbf{x}}$

$$\bar{\mathbf{x}} = \frac{n_1\bar{\mathbf{x}}_1 + n_2\bar{\mathbf{x}}_2 + n_3\bar{\mathbf{x}}_3}{n_1 + n_2 + n_3} \qquad \bar{\mathbf{x}} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

One-Way MANOVA

Ubah X → Y

S1 Sains Data
FMIPA UNESA

MANOVA Table

Source of variation	Matrix of sum of squares and cross products (SSP)	Degrees of freedom (d.f.)
Treatment	B = $\sum_{\ell=1}^{g} n_\ell(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})'$	$g - 1$
Residual (Error)	W = $\sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell}(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)'$	$\sum_{\ell=1}^{g} n_\ell - g$
Total (corrected for the mean)	B + W = $\sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell}(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})'$	$\sum_{\ell=1}^{g} n_\ell - 1$

The within sum of squares and cross products matrix can be expressed as:

$$\mathbf{W} = \sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell}(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)'$$
$$= (n_1 - 1)\mathbf{S}_1 + (n_2 - 1)\mathbf{S}_2 + \cdots + (n_g - 1)\mathbf{S}_g$$

Test Statistic *Wilks' lambda*

$$\Lambda^* = \frac{|\mathbf{W}|}{|\mathbf{B} + \mathbf{W}|} = \frac{\left|\sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell}(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)'\right|}{\left|\sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell}(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})(\mathbf{x}_{\ell j} - \bar{\mathbf{x}})'\right|}$$

One-Way MANOVA – Statistik Uji

Distribution of Wilks' Lambda

Critical region, Reject *H0* if

No. of variables	No. of groups	Sampling distribution for multivariate normal data
$p = 1$	$g \geq 2$	$\left(\frac{\sum n_\ell - g}{g-1}\right)\left(\frac{1-\Lambda^*}{\Lambda^*}\right) \sim F_{g-1, \sum n_\ell - g}$
$p = 2$	$g \geq 2$	$\left(\frac{\sum n_\ell - g - 1}{g-1}\right)\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \sim F_{2(g-1), 2(\sum n_\ell - g - 1)}$
$p \geq 1$	$g = 2$	$\left(\frac{\sum n_\ell - p - 1}{p}\right)\left(\frac{1-\Lambda^*}{\Lambda^*}\right) \sim F_{p, \sum n_\ell - p - 1}$
$p \geq 1$	$g = 3$	$\left(\frac{\sum n_\ell - p - 2}{p}\right)\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \sim F_{2p, 2(\sum n_\ell - p - 2)}$

for large n

if H_0 is true and $\Sigma n_\ell = n$ is large,

we reject H_0 at significance level α if

$$-\left(n - 1 - \frac{(p+g)}{2}\right)\ln\left(\frac{|\mathbf{W}|}{|\mathbf{B}+\mathbf{W}|}\right) > \chi^2_{p(g-1)}(\alpha)$$

Kesimpulan

- Tolak *H0* artinya tidak ada efek perlakuan, atau semua kelompok memiliki rata-rata yang sama.
- Gagal tolak *H0* artinya, setidaknya ada satu kelompok yang memiliki efek perlakuan berbeda secara signifikan dari yang lain.

One-Way MANOVA – SSCP (B)

The within sum of squares and cross products matrix ***(B)***

$$\mathbf{B} = \sum_{\ell=1}^{g} n_\ell (\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})'$$

$$\bar{\mathbf{x}}_1 = \begin{bmatrix} 8 \\ 4 \end{bmatrix} \qquad \bar{\mathbf{x}}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \bar{\mathbf{x}}_3 = \begin{bmatrix} 2 \\ 8 \end{bmatrix} \qquad \bar{\mathbf{x}} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

$$\bar{\mathbf{x}}_1 - \bar{\mathbf{x}} = \begin{bmatrix} 8 \\ 4 \end{bmatrix} - \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

$$\bar{\mathbf{x}}_2 - \bar{\mathbf{x}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} -3 \\ -3 \end{bmatrix}$$

$$\bar{\mathbf{x}}_3 - \bar{\mathbf{x}} = \begin{bmatrix} 2 \\ 8 \end{bmatrix} - \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

$$\mathbf{B} = n_1(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}})(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}})^T + n_2(\bar{\mathbf{x}}_2 - \bar{\mathbf{x}})(\bar{\mathbf{x}}_2 - \bar{\mathbf{x}})^T + n_3(\bar{\mathbf{x}}_3 - \bar{\mathbf{x}})(\bar{\mathbf{x}}_3 - \bar{\mathbf{x}})^T$$

$$= 3 \times \begin{bmatrix} 4 \\ -1 \end{bmatrix} \begin{bmatrix} 4 & -1 \end{bmatrix} + 2 \times \begin{bmatrix} -3 \\ -3 \end{bmatrix} \begin{bmatrix} -3 & -3 \end{bmatrix} + 3 \times \begin{bmatrix} -2 \\ 3 \end{bmatrix} \begin{bmatrix} -2 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 78 & -12 \\ -12 & 48 \end{bmatrix}$$

One-Way MANOVA – SSCP (W)

*The within sum of squares and cross products matrix **(W)***

W = $\sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell} (\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)'$

$$\bar{\mathbf{x}}_1 = \begin{bmatrix} 8 \\ 4 \end{bmatrix} \quad \bar{\mathbf{x}}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \bar{\mathbf{x}}_3 = \begin{bmatrix} 2 \\ 8 \end{bmatrix} \quad \bar{\mathbf{x}} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

grup ke-ℓ	var dependen	sampel ke-j 1	2	3
kelas A (n1 = 3)	nilai teori	9	6	9
	nilai praktik	3	2	7
Kelas B (n2 = 2)	nilai teori	0	2	
	nilai praktik	4	0	
Kelas C (n3 = 3)	nilai teori	3	1	2
	nilai praktik	8	9	7

Kelas A

$$\mathbf{x}_{11} - \bar{\mathbf{x}}_1 = \begin{bmatrix} 9 \\ 3 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mathbf{x}_{12} - \bar{\mathbf{x}}_1 = \begin{bmatrix} 6 \\ 2 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$$

$$\mathbf{x}_{13} - \bar{\mathbf{x}}_1 = \begin{bmatrix} 9 \\ 7 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

$$\sum (\mathbf{x}_{1j} - \bar{\mathbf{x}}_1)(\mathbf{x}_{1j} - \bar{\mathbf{x}}_1)^T = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 3 \\ 3 & 9 \end{bmatrix}$$

$$= \begin{bmatrix} 6 & 6 \\ 6 & 14 \end{bmatrix}$$

One-Way MANOVA – SSCP (W)

*The within sum of squares and cross products matrix **(W)***

$$W = \sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell} (\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)'$$

$$\bar{\mathbf{x}}_1 = \begin{bmatrix} 8 \\ 4 \end{bmatrix} \quad \bar{\mathbf{x}}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \bar{\mathbf{x}}_3 = \begin{bmatrix} 2 \\ 8 \end{bmatrix} \quad \bar{\mathbf{x}} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

grup ke-ℓ	var dependen	sampel ke-j		
		1	2	3
kelas A (n1 = 3)	nilai teori	9	6	9
	nilai praktik	3	2	7
Kelas B (n2 = 2)	nilai teori	0	2	
	nilai praktik	4	0	
Kelas C (n3 = 3)	nilai teori	3	1	2
	nilai praktik	8	9	7

Kelas B

$$(x_{21} - \bar{x}_2) = \begin{bmatrix} 0 \\ 4 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$(x_{22} - \bar{x}_2) = \begin{bmatrix} 2 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$\sum (\mathbf{x}_{2j} - \bar{\mathbf{x}}_2)(\mathbf{x}_{2j} - \bar{\mathbf{x}}_2)^T = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} + \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & -4 \\ -4 & 8 \end{bmatrix}$$

One-Way MANOVA – SSCP (W)

*The within sum of squares and cross products matrix **(W)***

W = $\sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell} (\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j} - \bar{\mathbf{x}}_\ell)'$

$$\bar{\mathbf{x}}_1 = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

$$\bar{\mathbf{x}}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \bar{\mathbf{x}} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

$$\bar{\mathbf{x}}_3 = \begin{bmatrix} 2 \\ 8 \end{bmatrix}$$

grup ke-ℓ	var dependen	sampel ke-j		
		1	**2**	**3**
kelas A (n1 = 3)	nilai teori	9	6	9
	nilai praktik	3	2	7
Kelas B (n2 = 2)	nilai teori	0	2	
	nilai praktik	4	0	
Kelas C (n3 = 3)	nilai teori	3	1	2
	nilai praktik	8	9	7

Kelas C

$$\mathbf{x}_{31} - \bar{\mathbf{x}}_3 = \begin{bmatrix} 3 \\ 8 \end{bmatrix} - \begin{bmatrix} 2 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\mathbf{x}_{32} - \bar{\mathbf{x}}_3 = \begin{bmatrix} 1 \\ 9 \end{bmatrix} - \begin{bmatrix} 2 \\ 8 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$\mathbf{x}_{33} - \bar{\mathbf{x}}_3 = \begin{bmatrix} 2 \\ 7 \end{bmatrix} - \begin{bmatrix} 2 \\ 8 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

$$\sum (\mathbf{x}_{3j} - \bar{\mathbf{x}}_3)(\mathbf{x}_{3j} - \bar{\mathbf{x}}_3)^T = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

One-Way MANOVA – SSCP (W)

The within sum of squares and cross products matrix ***(W)***

W = $\sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\mathbf{x}_{\ell j}-\bar{\mathbf{x}}_\ell)(\mathbf{x}_{\ell j}-\bar{\mathbf{x}}_\ell)'$

$$\bar{\mathbf{x}}_1 = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

$$\bar{\mathbf{x}}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \bar{\mathbf{x}} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

$$\bar{\mathbf{x}}_3 = \begin{bmatrix} 2 \\ 8 \end{bmatrix}$$

grup ke-ℓ	var dependen	sampel ke-j		
		1	**2**	**3**
kelas A (n1 = 3)	nilai teori	9	6	9
	nilai praktik	3	2	7
Kelas B (n2 = 2)	nilai teori	0	2	
	nilai praktik	4	0	
Kelas C (n3 = 3)	nilai teori	3	1	2
	nilai praktik	8	9	7

$$\mathbf{W} = \begin{bmatrix} 6 & 6 \\ 6 & 14 \end{bmatrix} + \begin{bmatrix} 2 & -4 \\ -4 & 8 \end{bmatrix} + \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 10 & 1 \\ 1 & 24 \end{bmatrix}$$

One-Way MANOVA – Statistik Uji

Calculate Statistics Test ***Wilks' lambda***

$$\Lambda^* = \frac{|\mathbf{W}|}{|\mathbf{B}+\mathbf{W}|} \qquad \mathbf{B} = \begin{bmatrix} 78 & -12 \\ -12 & 48 \end{bmatrix} \qquad W = \begin{bmatrix} 10 & 1 \\ 1 & 24 \end{bmatrix}$$

$$\Lambda^* = \frac{|\mathbf{W}|}{|\mathbf{B}+\mathbf{W}|} = \frac{\begin{vmatrix} 10 & 1 \\ 1 & 24 \end{vmatrix}}{\begin{vmatrix} 88 & -11 \\ -11 & 72 \end{vmatrix}} = \frac{239}{6215} = \boxed{0.0385}$$

```
> summary(a,test="W")
          Df    Wilks approx F num Df den Df   Pr(>F)
group      2 0.038455   8.1989      4      8 0.006234 **
Residuals  5
```

Approximate F Stat

$$\textbf{\textit{F}} = \left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \times \left(\frac{\sum n_\ell - g - 1}{g-1}\right) = \left(\frac{1-\sqrt{0.0385}}{\sqrt{0.0385}}\right) \times \left(\frac{8-3-1}{3-1}\right) = \boxed{8.19}$$

Compare 8.19 with F-distribution having
v1 = 2(g-1)=4 and v2 = $2(\sum n\ell - g - 1) = 8$

Conclusion

Since 8.19 > $F_{4,8}(0.01) = 7.01$, we reject *H₀* at the α = 0.01 and conclude that **treatment differences exist.**

One-Way MANOVA – Partial Eta-Squared

ukuran efek dalam analisis varians (MANOVA) yang menunjukkan proporsi variabilitas dependen yang dijelaskan oleh variabel independen, setelah mengontrol variabilitas error.

$$\eta^2 = \frac{\text{between sum of squares}}{\text{total sum of squares}}.$$

Example

The Wisconsin Department of Health and Social Services reimburses nursing homes in the state for the services provided. The department develops a set of formulas for rates for each facility, based on factors such as level of care, mean wage rate, and average wage rate in the state.

Nursing homes can be classified on the basis of ownership (**private party, nonprofit organization, and government**) and certification (**skilled nursing facility, intermediate care facility, or a combination of the two**).

X1 = cost of nursing labor
X2 = cost of dietary labor
X3 = cost of plant operation and maintenance labor
X4 = cost of housekeeping and laundry labor

n = 516 observations

Group	Number of observations
$\ell = 1$ (private)	$n_1 = 271$
$\ell = 2$ (nonprofit)	$n_2 = 138$
$\ell = 3$ (government)	$n_3 = 107$
	$\sum_{\ell=1}^{3} n_\ell = 516$

Mean for each group

$$\bar{\mathbf{x}}_1 = \begin{bmatrix} 2.066 \\ .480 \\ .082 \\ .360 \end{bmatrix}; \quad \bar{\mathbf{x}}_2 = \begin{bmatrix} 2.167 \\ .596 \\ .124 \\ .418 \end{bmatrix}; \quad \bar{\mathbf{x}}_3 = \begin{bmatrix} 2.273 \\ .521 \\ .125 \\ .383 \end{bmatrix}$$

Grand Mean

$$\bar{\mathbf{x}} = \frac{n_1\bar{\mathbf{x}}_1 + n_2\bar{\mathbf{x}}_2 + n_3\bar{\mathbf{x}}_3}{n_1 + n_2 + n_3} = \begin{bmatrix} 2.136 \\ .519 \\ .102 \\ .380 \end{bmatrix}$$

Covariance for each group

$$\mathbf{S}_1 = \begin{bmatrix} .291 & & & \\ -.001 & .011 & & \\ .002 & .000 & .001 & \\ .010 & .003 & .000 & .010 \end{bmatrix}$$

$$\mathbf{S}_2 = \begin{bmatrix} .561 & & & \\ .011 & .025 & & \\ .001 & .004 & .005 & \\ .037 & .007 & .002 & .019 \end{bmatrix}$$

$$\mathbf{S}_3 = \begin{bmatrix} .261 & & & \\ .030 & .017 & & \\ .003 & -.000 & .004 & \\ .018 & .006 & .001 & .013 \end{bmatrix}$$

SSCP (W)

$$\mathbf{W} = (n_1 - 1)\mathbf{S}_1 + (n_2 - 1)\mathbf{S}_2 + (n_3 - 1)\mathbf{S}_3$$

$$= \begin{bmatrix} 182.962 & & & \\ 4.408 & 8.200 & & \\ 1.695 & .633 & 1.484 & \\ 9.581 & 2.428 & .394 & 6.538 \end{bmatrix}$$

SSCP (B)

$$\mathbf{B} = \sum_{\ell=1}^{3} n_\ell(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})(\bar{\mathbf{x}}_\ell - \bar{\mathbf{x}})'$$

$$= \begin{bmatrix} 3.475 & & & \\ 1.111 & 1.225 & & \\ .821 & .453 & .235 & \\ .584 & .610 & .230 & .304 \end{bmatrix}$$

Uji Hipotesis

$$H_0 : \tau_1 = \tau_2 = \cdots = \tau_g = 0$$

$$H_1 : \text{at least one } \tau_\ell \neq 0$$

Statistik Uji

$$\Lambda^* = \frac{|\mathbf{W}|}{|\mathbf{B} + \mathbf{W}|} = .7714$$

$$\left(\frac{\sum n_\ell - p - 2}{p}\right)\left(\frac{1 - \sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) = \left(\frac{516 - 4 - 2}{4}\right)\left(\frac{1 - \sqrt{.7714}}{\sqrt{.7714}}\right) = 17.67$$

$$17.67 > F_{8,1020}(.01) \doteq 2.51$$

$$-(n - 1 - (p + g)/2) \ln\left(\frac{|\mathbf{W}|}{|\mathbf{B} + \mathbf{W}|}\right) = -511.5 \ln(.7714) = 132.76$$

$$\chi^2_{p(g-1)}(.01) = \chi^2_8(.01) = 20.09. \quad 132.76 > \chi^2_8(.01) = 20.09,$$

Tolak H0

TWO-WAY MANOVA

Two-Way MANOVA

Model

$$X_{\ell kr} = \mu + \tau_\ell + \beta_k + \gamma_{\ell k} + e_{\ell kr}$$

$$\ell = 1, 2, \ldots, g$$
$$k = 1, 2, \ldots, b$$
$$r = 1, 2, \ldots, n$$

$$\begin{array}{ccccccccc} E(X_{\ell kr}) & = & \mu & + & \tau_\ell & + & \beta_k & + & \gamma_{\ell k} \\ \begin{pmatrix} \text{mean} \\ \text{response} \end{pmatrix} & = & \begin{pmatrix} \text{overall} \\ \text{level} \end{pmatrix} & + & \begin{pmatrix} \text{effect of} \\ \text{factor 1} \end{pmatrix} & + & \begin{pmatrix} \text{effect of} \\ \text{factor 2} \end{pmatrix} & + & \begin{pmatrix} \text{factor 1-factor 2} \\ \text{interaction} \end{pmatrix} \end{array}$$

$$x_{\ell kr} = \bar{x} + (\bar{x}_{\ell\cdot} - \bar{x}) + (\bar{x}_{\cdot k} - \bar{x}) + (\bar{x}_{\ell k} - \bar{x}_{\ell\cdot} - \bar{x}_{\cdot k} + \bar{x}) + (x_{\ell kr} - \bar{x}_{\ell k})$$

Two-Way MANOVA

Sum-of-Square

$$\sum_{\ell=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}(\mathbf{x}_{\ell kr}-\bar{\mathbf{x}})(\mathbf{x}_{\ell kr}-\bar{\mathbf{x}})' = \sum_{\ell=1}^{g} bn(\bar{\mathbf{x}}_{\ell\cdot}-\bar{\mathbf{x}})(\bar{\mathbf{x}}_{\ell\cdot}-\bar{\mathbf{x}})' \quad \rightarrow \text{SSP Factor 1}$$

$$+\sum_{k=1}^{b} gn(\bar{\mathbf{x}}_{\cdot k}-\bar{\mathbf{x}})(\bar{\mathbf{x}}_{\cdot k}-\bar{\mathbf{x}})' \quad \rightarrow \text{SSP Factor 2}$$

$$+\sum_{\ell=1}^{g}\sum_{k=1}^{b} n(\bar{\mathbf{x}}_{\ell k}-\bar{\mathbf{x}}_{\ell\cdot}-\bar{\mathbf{x}}_{\cdot k}+\bar{\mathbf{x}})(\bar{\mathbf{x}}_{\ell k}-\bar{\mathbf{x}}_{\ell\cdot}-\bar{\mathbf{x}}_{\cdot k}+\bar{\mathbf{x}})' \quad \rightarrow \text{SSP Interaksi}$$

$$+\sum_{\ell=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}(\mathbf{x}_{\ell kr}-\bar{\mathbf{x}}_{\ell k})(\mathbf{x}_{\ell kr}-\bar{\mathbf{x}}_{\ell k})' \quad \rightarrow \text{SSP Res / error}$$

Degree of freedom (df)

$$gbn - 1 = (g-1) + (b-1) + (g-1)(b-1) + gb(n-1)$$

Two-Way MANOVA

MANOVA Table

Source of Variation	*Matrix of Sum of Squares and Cross Product* (SSCP)	*Degrees of freedom* (df)
Faktor 1	$\mathrm{SSP}_{\mathrm{fak1}} = \sum_{l=1}^{g} bn(\bar{x}_{l.} - \bar{x})(\bar{x}_{l.} - \bar{x})^T$	$g - 1$
Faktor 2	$\mathrm{SSP}_{\mathrm{fak2}} = \sum_{k=1}^{b} gn(\bar{x}_{.k} - \bar{x})(\bar{x}_{.k} - \bar{x})^T$	$b - 1$
Interaksi	$\mathrm{SSP}_{\mathrm{int}} = \sum_{l=1}^{g}\sum_{k=1}^{b} n(\bar{x}_{lk} - \bar{x}_{l.} - \bar{x}_{.k} + \bar{x})(\bar{x}_{lk} - \bar{x}_{l.} - \bar{x}_{.k} + \bar{x})^T$	$(g-1)(b-1)$
Residual	$\mathrm{SSP}_{\mathrm{res}} = \sum_{l=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n} (x_{lkr} - \bar{x}_{lk})(x_{lkr} - \bar{x}_{lk})^T$	$gb(n-1)$
Total (Terkoreksi)	$\mathrm{SSP}_{\mathrm{cor}} = \sum_{l=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n} (x_{lkr} - \bar{x})(x_{lkr} - \bar{x})^T$	$gbn - 1$

Two-Way MANOVA – Statistik Uji

Factor 1

Hipotesis

$$H_0: \boldsymbol{\tau}_1 = \boldsymbol{\tau}_2 = \cdots = \boldsymbol{\tau}_g = \mathbf{0}$$
$$H_1: \text{at least one } \boldsymbol{\tau}_\ell \neq \mathbf{0}.$$

Statistik Uji

$$\Lambda^* = \frac{|\mathrm{SSP}_{\mathrm{res}}|}{|\mathrm{SSP}_{\mathrm{fac1}} + \mathrm{SSP}_{\mathrm{res}}|}$$

$$F_1 = \left(\frac{1 - \Lambda_1^*}{\Lambda_1^*}\right)\frac{(gb(n-1) - p + 1)/2}{(|(g-1) - p| + 1)/2}$$

Daerah Penolakan

Tolak H0, jika F1 > Ftabel

$$\nu_1 = |(g-1) - p| + 1, \quad \nu_2 = gb(n-1) - p + 1$$

For large n

$$-\left[gb(n-1) - \frac{p + 1 - (g-1)}{2}\right]\ln \Lambda^* > \chi^2_{(g-1)p}(\alpha)$$

Factor 2

Hipotesis

$$H_0: \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \cdots = \boldsymbol{\beta}_b = \mathbf{0}$$
$$H_1: \text{at least one } \boldsymbol{\beta}_k \neq \mathbf{0}.$$

Statistik Uji

$$\Lambda^* = \frac{|\mathrm{SSP}_{\mathrm{res}}|}{|\mathrm{SSP}_{\mathrm{fac2}} + \mathrm{SSP}_{\mathrm{res}}|}$$

$$F_2 = \left(\frac{1 - \Lambda_2^*}{\Lambda_2^*}\right)\frac{(gb(n-1) - p + 1)/2}{(|(b-1) - p| + 1)/2}$$

Daerah Penolakan

Tolak H0, jika F2 > Ftabel

$$\nu_1 = |(b-1) - p| + 1, \nu_2 = gb(n-1) - p + 1$$

For large n

$$-\left[gb(n-1) - \frac{p + 1 - (b-1)}{2}\right]\ln \Lambda^* > \chi^2_{(b-1)p}(\alpha)$$

Two-Way MANOVA – Statistik Uji

Interaksi

Hipotesis

$$H_0: \gamma_{11} = \gamma_{12} = \cdots = \gamma_{gb} = \mathbf{0}$$

$$H_1: \text{At least one } \gamma_{\ell k} \neq \mathbf{0}$$

Statistik Uji

$$\Lambda^* = \frac{|SSP_{res}|}{|SSP_{int} + SSP_{res}|}$$

$$F = \left(\frac{1-\Lambda^*}{\Lambda^*}\right)\frac{(gb(n-1)-p+1)/2}{(|(g-1)(b-1)-p|+1)/2}$$

For large n

$$-\left[gb(n-1) - \frac{p+1-(g-1)(b-1)}{2}\right]\ln \Lambda^* > \chi^2_{(g-1)(b-1)p}(\alpha)$$

Daerah Penolakan

Tolak H0, jika F > Ftabel

$$\nu_1 = |(g-1)(b-1)-p|+1$$

$$\nu_2 = gb(n-1)-p+1$$

Two-Way MANOVA – Example

S1 Sains Data FMIPA UNESA

- The optimum conditions for extruding plastic film have been examined using a technique called Evolutionary Operation.
- In the course of the study that was done, three responses X1 = tear resistance, X2 = gloss, and *X3* = opacity-were measured at two levels of the factors, *rate of extrusion* and *amount of an additive.*
- The measurements were repeated *n* = 5 times at each combination of the factor levels.

Ubah X → Y

Table 6.4 Plastic Film Data

x_1 = tear resistance, x_2 = gloss, and x_3 = opacity

		Factor 2: Amount of additive					
		Low (1.0%)			High (1.5%)		
		x_1	x_2	x_3	x_1	x_2	x_3
Factor 1: Change in rate of extrusion	Low (−10)%	[6.5	9.5	4.4]	[6.9	9.1	5.7]
		[6.2	9.9	6.4]	[7.2	10.0	2.0]
		[5.8	9.6	3.0]	[6.9	9.9	3.9]
		[6.5	9.6	4.1]	[6.1	9.5	1.9]
		[6.5	9.2	0.8]	[6.3	9.4	5.7]
	High (10%)	[6.7	9.1	2.8]	[7.1	9.2	8.4]
		[6.6	9.3	4.1]	[7.0	8.8	5.2]
		[7.2	8.3	3.8]	[7.2	9.7	6.9]
		[7.1	8.4	1.6]	[7.5	10.1	2.7]
		[6.8	8.5	3.4]	[7.6	9.2	1.9]

f1

$$\bar{x}_{1.} = \begin{bmatrix} 6,49 \\ 9,57 \\ 3,79 \end{bmatrix} \qquad \bar{x}_{2.} = \begin{bmatrix} 7,08 \\ 9,06 \\ 4,08 \end{bmatrix} \qquad \bar{x} = \begin{bmatrix} 6,785 \\ 9,315 \\ 3,935 \end{bmatrix}$$

$$SSP_{fac1} = \sum_{l=1}^{g} bn \, (\bar{x}_{l.} - \bar{x})(\bar{x}_{l.} - \bar{x})'$$

$$= 2.5 \left[\left(\begin{bmatrix} 6,49 \\ 9,57 \\ 3,79 \end{bmatrix} - \begin{bmatrix} 6,785 \\ 9,315 \\ 3,935 \end{bmatrix} \right) \left(\begin{bmatrix} 6,49 \\ 9,57 \\ 3,79 \end{bmatrix} - \begin{bmatrix} 6,785 \\ 9,315 \\ 3,935 \end{bmatrix} \right)' + \left(\begin{bmatrix} 7,08 \\ 9,06 \\ 4,08 \end{bmatrix} - \begin{bmatrix} 6,785 \\ 9,315 \\ 3,935 \end{bmatrix} \right) \left(\begin{bmatrix} 7,08 \\ 9,06 \\ 4,08 \end{bmatrix} - \begin{bmatrix} 6,785 \\ 9,315 \\ 3,935 \end{bmatrix} \right)' \right]$$

$$= \begin{bmatrix} 1,7405 & -1,5405 & 0,855 \\ -1,5405 & 1,3005 & -0,7395 \\ 0,8555 & -0,7395 & 0,4205 \end{bmatrix}$$

f2

$$\bar{x}_{\cdot 1} = \begin{bmatrix} 6,59 \\ 9,14 \\ 3,44 \end{bmatrix} \qquad \bar{x}_{\cdot 2} = \begin{bmatrix} 6,98 \\ 9,49 \\ 4,43 \end{bmatrix} \qquad \bar{x} = \begin{bmatrix} 6,785 \\ 9,315 \\ 3,935 \end{bmatrix}$$

$$SSP_{fac2} = \sum_{k=1}^{b} gn \, (\bar{x}_{\cdot k} - \bar{x})(\bar{x}_{\cdot k} - \bar{x})'$$

$$= 2 \cdot 5 \left[\left(\begin{bmatrix} 6,59 \\ 9,14 \\ 3,44 \end{bmatrix} - \begin{bmatrix} 6,785 \\ 9,315 \\ 3,935 \end{bmatrix} \right) \left(\begin{bmatrix} 6,59 \\ 9,14 \\ 3,44 \end{bmatrix} - \begin{bmatrix} 6,785 \\ 9,315 \\ 3,935 \end{bmatrix} \right)' + \left(\begin{bmatrix} 6,98 \\ 9,49 \\ 4,43 \end{bmatrix} - \begin{bmatrix} 6,785 \\ 9,315 \\ 3,935 \end{bmatrix} \right) \left(\begin{bmatrix} 6,98 \\ 9,49 \\ 4,43 \end{bmatrix} - \begin{bmatrix} 6,785 \\ 9,315 \\ 3,935 \end{bmatrix} \right)' \right]$$

$$= \begin{bmatrix} 0,7605 & 0,6825 & 1,9305 \\ 0,6825 & 0,6125 & 1,7325 \\ 1,9305 & 1,7325 & 4,9005 \end{bmatrix} //$$

$$SSP\ Total = \sum_{l=1}^{g} \sum_{k=1}^{b} \sum_{r=1}^{n} (x_{lkr} - \bar{x})(x_{lkr} - \bar{x})'$$

$$\bar{x} = \begin{bmatrix} 6,785 \\ 9,315 \\ 3,935 \end{bmatrix}$$

$$SSP_{tot} = \left(\begin{bmatrix} 6,5 & 6,2 & \cdots & 7,6 \\ 9,5 & 9,9 & \cdots & 9,2 \\ 4,4 & 6,4 & \cdots & 1,9 \end{bmatrix} - \begin{bmatrix} 6,785 \\ 9,315 \\ 3,935 \end{bmatrix} \right) \left(\begin{bmatrix} 6,5 & 6,2 & \cdots & 7,6 \\ 9,5 & 9,9 & \cdots & 9,2 \\ 4,4 & 6,4 & \cdots & 1,9 \end{bmatrix} - \begin{bmatrix} 6,785 \\ 9,315 \\ 3,935 \end{bmatrix} \right)'$$

$$= \begin{bmatrix} 4,2655 & -0,7855 & -0,2395 \\ -0,7855 & 5,0855 & 1,9095 \\ -0,2395 & 1,9895 & 74,2055 \end{bmatrix} //$$

Error : $(X_{11r} - \bar{X}_{11})(X_{11r} - \bar{X}_{11})'$

$$\left(\begin{bmatrix} 6,5 & 6,2 & 5,8 & 6,5 & 6,5 \\ 9,5 & 9,9 & 9,6 & 9,6 & 9,2 \\ 4,4 & 6,4 & 3 & 4,1 & 0,8 \end{bmatrix} - \begin{bmatrix} 6,3 \\ 9,56 \\ 3,74 \end{bmatrix}\right)\left(\begin{bmatrix} 6,5 & 6,2 & 5,8 & 6,5 & 6,5 \\ 9,5 & 9,9 & 9,6 & 9,6 & 9,2 \\ 4,4 & 6,4 & 3 & 4,1 & 0,8 \end{bmatrix} - \begin{bmatrix} 6,3 \\ 9,56 \\ 3,74 \end{bmatrix}\right)'$$

$$(X_{11r} - \bar{X}_{11})(X_{11r} - \bar{X}_{11})' = \begin{bmatrix} 0,38 & 0,13 & -0,28 \\ -0,13 & 0,252 & 1,908 \\ -0,28 & 1,908 & 16,832 \end{bmatrix}$$

factor 1 low, factor 2 low

$$(X_{12r} - \bar{X}_{12})(X_{12r} - \bar{X}_{12})' = \begin{bmatrix} 0,848 & 0,298 & -0,116 \\ 0,298 & 0,548 & -1,826 \\ -0,116 & -1,826 & 14,072 \end{bmatrix}$$

factor 1 low, factor 2 high

S1 Sains Data FMIPA UNESA

$$(X_{21r} - \bar{X}_{21})(X_{21r} - \bar{X}_{21})' = \begin{bmatrix} 0{,}268 & -0{,}418 & -0{,}356 \\ -0{,}418 & 0{,}808 & 0{,}586 \\ -0{,}356 & 0{,}586 & 3{,}912 \end{bmatrix}$$

factor 1 high, factor 2 low

$$(X_{22r} - \bar{X}_{22})(X_{22r} - \bar{X}_{22})' = \begin{bmatrix} 0{,}268 & 0{,}27 & -2{,}318 \\ 0{,}27 & 1{,}02 & -1{,}22 \\ -2{,}318 & -1{,}22 & 30{,}108 \end{bmatrix}$$

factor 1 high, factor 2 high

$$SSP_{error} = \sum_{l=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n} (X_{lkr} - \bar{X}_{lk})(X_{lkr} - \bar{X}_{lk})'$$

$$= \begin{bmatrix} 0{,}38 & 0{,}13 & -0{,}28 \\ -0{,}13 & 0{,}252 & 1{,}908 \\ -0{,}28 & 1{,}908 & 16{,}832 \end{bmatrix} + \begin{bmatrix} 0{,}848 & 0{,}298 & -0{,}116 \\ 0{,}298 & 0{,}548 & -1{,}826 \\ -0{,}116 & -1{,}826 & 14{,}072 \end{bmatrix} + \begin{bmatrix} 0{,}268 & -0{,}418 & -0{,}356 \\ -0{,}418 & 0{,}808 & 0{,}586 \\ -0{,}356 & 0{,}586 & 3{,}912 \end{bmatrix} + \begin{bmatrix} 0{,}268 & 0{,}27 & -2{,}318 \\ 0{,}27 & 1{,}02 & -1{,}22 \\ -2{,}318 & -1{,}22 & 30{,}108 \end{bmatrix}$$

$$= \begin{bmatrix} 1{,}764 & 0{,}02 & -3{,}07 \\ 0{,}02 & 2{,}628 & -0{,}552 \\ -3{,}07 & -0{,}552 & 64{,}924 \end{bmatrix}$$

SSP Interaksi

$$SSP_{tot} = SSP_{fac1} + SSP_{fac2} + SSP_{int} + SSP_{error}$$

$$SSP_{int} = \begin{bmatrix} 0,0005 & 0,0165 & 0,0445 \\ 0,0165 & -0,5445 & 1,4685 \\ 0,0445 & 1,4685 & 3,9605 \end{bmatrix} //$$

S1 Sains Data
FMIPA UNESA

Faktor 1

$H_0 : \tau_1 = \tau_2 = 0$

H_1 : at least one $\tau_l \neq 0$ l= 1, 2

Stat uji

$$\Lambda_1^* = \frac{|SSP_{error}|}{|SSP_{fac1} + SSP_{error}|}$$

$$= \frac{275{,}71}{722{,}02}$$

$$= 0{,}3818$$

$$F_1 = \left(\frac{1-\Lambda_1^*}{\Lambda_1^*}\right)\frac{(gb(n-1)-p+1)/2}{(|(g-1)-p|+1)/2}$$

$$= \frac{1-0{,}3818}{0{,}3818}\left(\frac{(2.2(5-1)-3+1)/2}{(|(2-1)-3|+1)/2}\right)$$

$$= 7{,}55 > F_{3,14(0{,}05)} \rightarrow \text{tolak } H_0$$

$$F_{3,14(0.05)} = 3{,}34$$

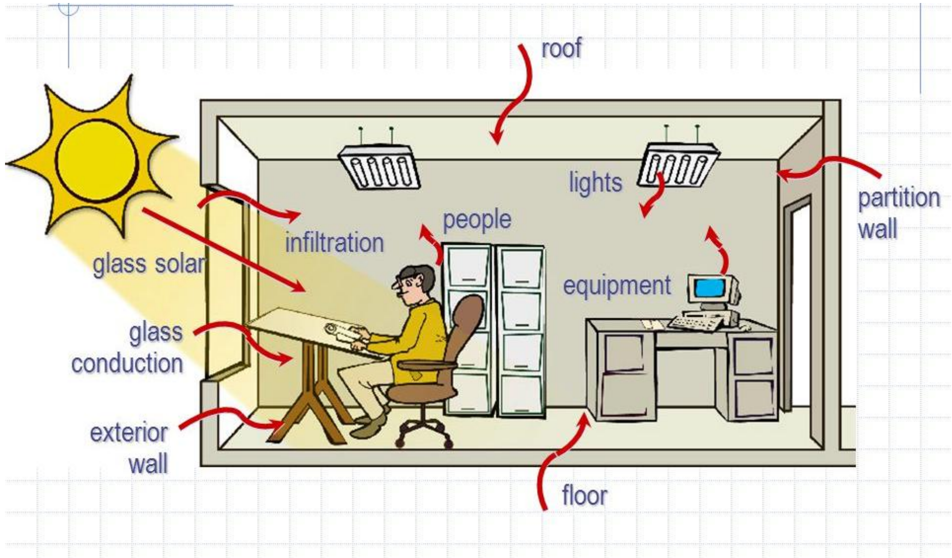
$$v_1 = |(g-1)-p|+1 = |(2-1)-3|+1 = 3$$

$$v_2 = gb(n-1)-p+1 = 2.2(5-1)-3+1 = 14$$

CONTOH KASUS

Contoh

Simbol	Variabel	
X_1	*Glazing Area*	0,00 0,10 0,25 0,40
X_2	*Orientation*	2 3 4 5
Y_1	*Heating load*	
Y_2	*Cooling load*	



Dilakukan analisis pada faktor-faktor yang berpengaruh terhadap efisiensi energi

Uji distribusi normal multivariat

normal univariat

```
  Shapiro-Wilk Normality Test (alpha = 0.05)
--------------------------------------------------
  data : Y1 + Y2 and X1

  Level Statistic      p.value   Normality
1     0 0.8828207 1.834680e-04      Reject
2   0.1 0.8653184 1.093988e-13      Reject
3  0.25 0.8547543 2.908999e-14      Reject
4   0.4 0.8495001 1.543426e-14      Reject
--------------------------------------------------
```

```
  Shapiro-Wilk Normality Test (alpha = 0.05)
--------------------------------------------------
  data : Y1 + Y2 and X2

  Level Statistic      p.value   Normality
1     2 0.9072441 1.330382e-09      Reject
2     3 0.9089738 1.730971e-09      Reject
3     4 0.9094403 1.859334e-09      Reject
4     5 0.9082059 1.539486e-09      Reject
--------------------------------------------------
```

normal multivariat

```
> mvnormtest::mshapiro.test(t(data[, c("Y1", "Y2")]))

        Shapiro-Wilk normality test

data:  Z
W = 0.91501, p-value < 2.2e-16
```

diasumsikan berdistribusi normal multivariat

Uji dependensi

Uji Bartlet

```
$chisq
[1] 2329.382

$p.value
[1] 0

$df
[1] 1
```

Nilai tersebut kurang dari α (0,05) yang berarti tolak H0 sehingga antar variabel heating load dan cooling load terjadi dependensi dan dapat dilanjutkan pengujian asumsi berikutnya

Uji Homogenitas

Uji Box's M

```
> boxM(cbind(Y1, Y2) ~ X1, data = data)

	Box's M-test for Homogeneity of Covariance Matrices

data:  Y
Chi-Sq (approx.) = 19.424, df = 9, p-value = 0.02182

> boxM(cbind(Y1, Y2) ~ X2, data = data)

	Box's M-test for Homogeneity of Covariance Matrices

data:  Y
Chi-Sq (approx.) = 17.494, df = 9, p-value = 0.04152

> boxM(cbind(Y1, Y2) ~ X1_X2, data = data)

	Box's M-test for Homogeneity of Covariance Matrices

data:  Y
Chi-Sq (approx.) = 60.931, df = 45, p-value = 0.05676
```

- P-value untuk data *one-way* MANOVA kurang dari α (0,05). Hal ini menunjukkan bahwa data *heating load* dan *cooling load* yang dipengaruhi oleh setiap level pada variabel *glazing area* dan *orientation* bersifat tidak homogen. *Namun*, data tersebut diasumsikan homogen agar dapat dianalisis lebih lanjut dengan menggunakan *one-way* MANOVA.

- Data *two-way* MANOVA, diperoleh nilai *p-value* sebesar 0,057. Hal ini menunjukkan bahwa data *heating load* dan *cooling load* bersifat homogen. Dengan demikian, analisis dapat dilanjutkan pada *two-way* MANOVA.

- Nilai Wilks Lambda mendekati 0 (**pada value Wilks**)
    - semakin mendekati 0 semakin berpengaruh (>0,5 mendekati 1 dan <0,5 mendekati 0)
- Partial Eta Mendekati 1
    - semakin mendekati 1 semakin berpengaruh
- Signifikansi kurang dari 0,05
    - untuk melihat pengaruh saja
- Maka ada pengaruh treatment **terhadap variable dependen.**

One-Way MANOVA

```
> summary(one_way_manova_x1,test="W")
           Df   Wilks approx F num Df den Df    Pr(>F)
X1          3 0.80641   28.887      6   1526 < 2.2e-16 ***
Residuals 764
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
> etasq(one_way_manova_x1, test="Wilks")
       eta^2
X1 0.1019964
```

Wilks' Lambda -> mendekati 1
Partial Eta Squared -> mendekati 0

Bertolak belakang dengan
Sig. < alpha (0,05)

Diindikasikan karena adanya asumsi MANOVA yang tidak terpenuhi

Perlakuan *glazing area* dengan 4 level yang berbeda memiliki pengaruh signifikan terhadap *heating load* dan *cooling load.*

One-Way MANOVA


```
> summary(one_way_manova_x2,test="W")
             Df   Wilks approx F num Df den Df   Pr(>F)  
X2            3 0.98386   2.0774      6   1526 0.05304 .
Residuals   764                                         
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Wilks' Lambda -> mendekati 1
Partial Eta Squared -> mendekati 0

Sesuai dengan
Sig. > alpha (0,05)

```
> etasq(one_way_manova_x2, test="Wilks")
         eta^2
X2 0.008101778
```

Perlakuan *orientation* dengan 4 level yang berbeda tidak memiliki pengaruh signifikan terhadap *heating load* dan *cooling load*

Two-Way MANOVA

```
> summary(two_way_manova,test="W")
           Df   Wilks approx F num Df den Df   Pr(>F)    
X1          3 0.80341  28.9533      6   1502 < 2e-16 ***
X2          3 0.98134   2.3688      6   1502 0.02785 *  
X1:X2       9 0.98975   0.4310     18   1502 0.98173    
Residuals 752                                           
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

> etasq(two_way_manova, test="Wilks")
           eta^2
X1    0.10366878
X2    0.00937393
X1:X2 0.00513914
```

- Perlakuan *glazing area* dan *orientation* dengan 4 level yang berbeda memiliki pengaruh signifikan terhadap *heating load* dan *cooling load* karena nilai *p-value* yang didapatkan kurang dari alpha (0,05) namun bertolak belakang dengan nilai *Wilks' Lambda* dan *partial eta squared* (diindikasikan karena asumsi MANOVA tidak terpenuhi).

- Pengaruh interaksi *glazing area* dan *orientation* diperoleh kesimpulan tidak berpengaruh terhadap *heating load* dan *cooling load*

Something to note…

This example only contained 1 IV with 2 levels

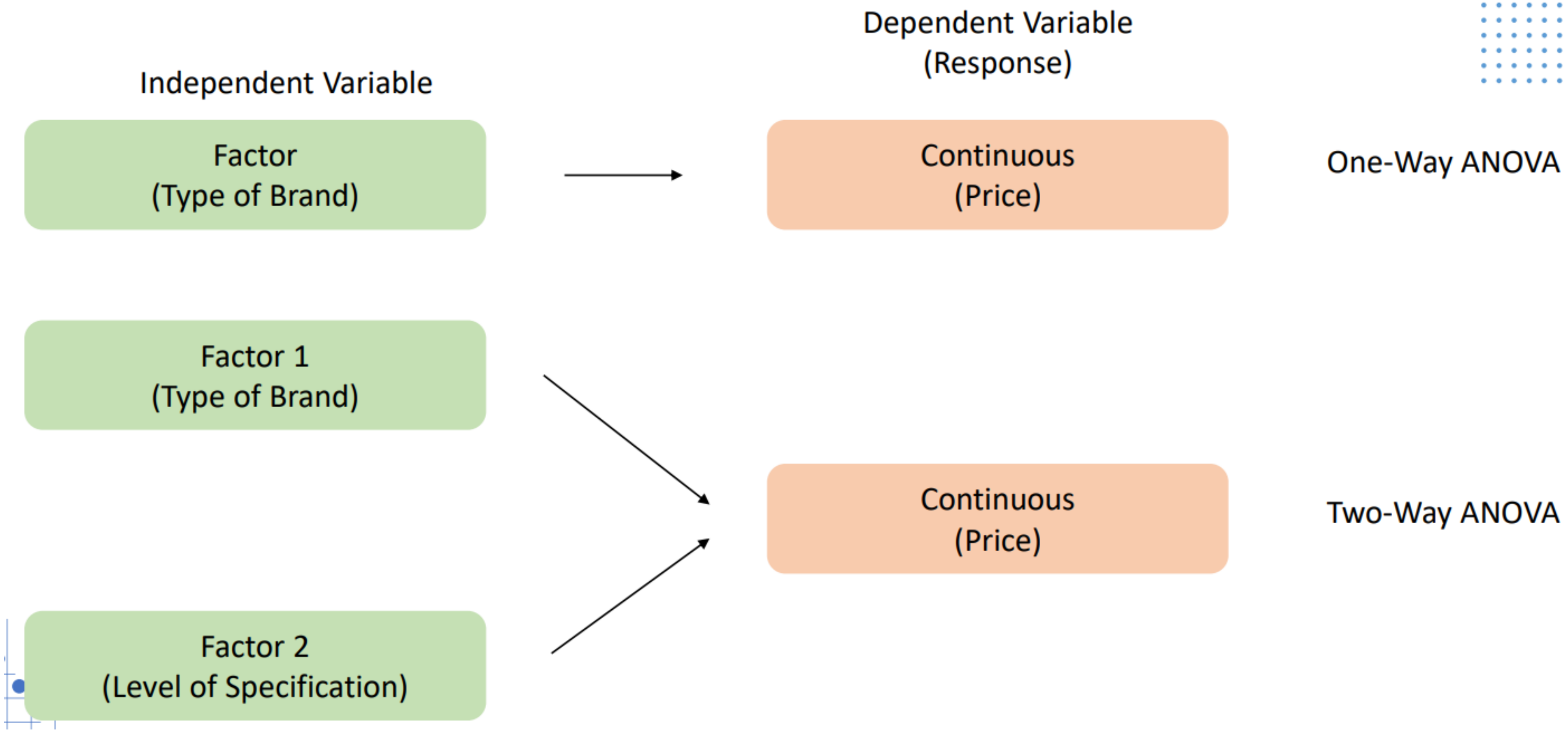
If we had 3 levels (e.g., 1 cup coffee, 3 cups coffee, 1 cup water), we would have needed to conduct a pairwise comparison test to investigate which level of the IV significantly affected the DV?

This can be done by going to

-> Analyse -> General linear model -> Multivariate -> Post-Hoc -> Moving the IV to 'Post Hoc Tests for:' -> Selecting a preferred post hoc test (common test is Tukey)

ANOVA vs MANOVA vs ANCOVA vs MANCOVA

ANOVA

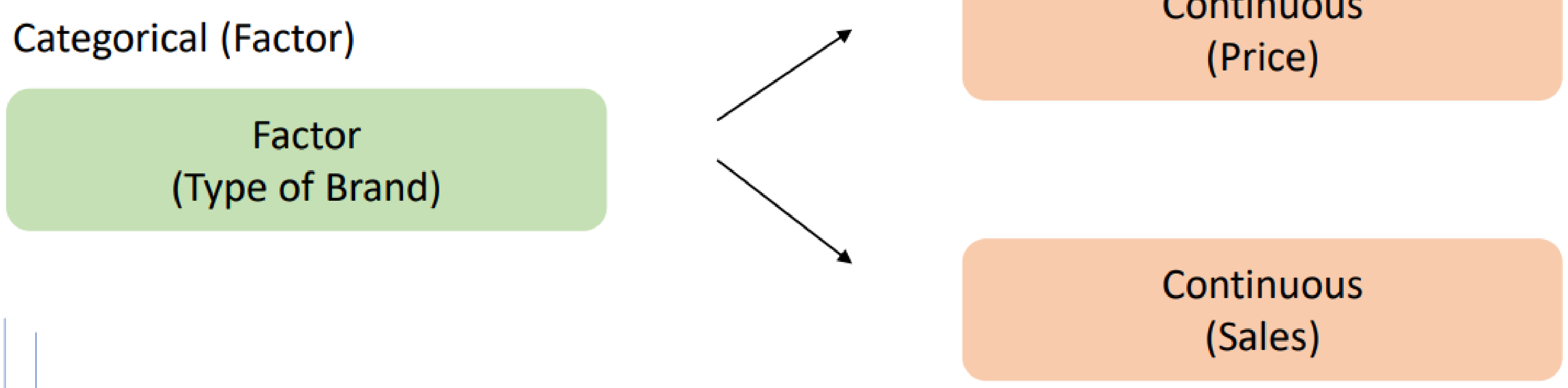

Dependent Variable
(Response)
Independent Variable
Factor
(Type of Brand)
Continuous
(Price)
One-Way ANOVA
Factor 1
(Type of Brand)
Factor 2
(Level of Specification)
Continuous
(Price)
Two-Way ANOVA

MANOVA

Multivariate

Independent Variable

Dependent Variable



ANCOVA

Covariates

Independent Variable		Dependent Variable
Categorical (Factor)		
Factor (Type of Brand)	→	Continuous (Price)
Continuous		
Covariates (Number of Spent Time)	↗	

Covariates

MANCOVA

Multivariate

Independent Variable

Dependent Variable

Categorical (Factor)

Factor
(Type of Brand)

Continuous
(Price)

Continuous

Covariates
(Number of Spent Time)

Continuous
(Sales)

ANCOVA

ANCOVA

The ANCOVA model assumes a linear relationship between the response (DV) and covariate (CV):

$$y_{ij} = \mu + \tau_i + \mathrm{B}(x_{ij} - \overline{x}) + \epsilon_{ij}.$$

In this equation, the DV, y_{ij} is the jth observation under the ith categorical group; the CV, x_{ij} is the *j*th observation of the covariate under the *i*th group. Variables in the model that are derived from the observed data are μ (the grand mean) and $\overline{x}$ (the global mean for covariate x). The variables to be fitted are τ_i (the effect of the *i*th level of the categorical IV), B (the slope of the line) and ϵ_{ij} (the associated unobserved error term for the *j*th observation in the *i*th group).

Assumptions

- Assumption 1: linearity of regression

The regression relationship between the dependent variable and concomitant variables must be linear.

- Assumption 2: homogeneity of error variances

The error is a random variable with conditional zero mean and equal variances for different treatment classes and observations.

- Assumption 3: independence of error terms

The errors are uncorrelated. That is, the error covariance matrix is diagonal.

- Assumption 4: normality of error terms

The residuals (error terms) should be normally distributed $\epsilon_{ij} \sim N(0,\sigma^2)$.

- Assumption 5: homogeneity of regression slopes

The slopes of the different regression lines should be equivalent, i.e., regression lines should be parallel among groups.

File Excel manual calculation ANCOVA terdapat di sindig

Exercise

- A study was conducted to see the impact of social-economic class (rich, middle, poor) and gender (male, female) on kindness and optimism using a sample of 24 people based on the data in Figure 1.

	A	B	C	D
3	gender	economic	kindness	optimism
4	male	wealthy	5	3
5	male	wealthy	4	6
6	male	wealthy	3	4
7	male	wealthy	2	4
8	male	middle	4	6
9	male	middle	3	6
10	male	middle	5	4
11	male	middle	5	5
12	male	poor	7	5
13	male	poor	4	3
14	male	poor	3	1
15	male	poor	7	2
16	female	wealthy	2	3
17	female	wealthy	3	5
18	female	wealthy	5	3
19	female	wealthy	4	2
20	female	middle	9	8
21	female	middle	6	5
22	female	middle	7	6
23	female	middle	8	9
24	female	poor	8	9
25	female	poor	9	8
26	female	poor	3	7
27	female	poor	5	7