Pertemuan 1 Konsep Matriks Dalam TPG

Teras Matriks

$$tr(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}$$

Determinan Matriks

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$\det(A) = ad - bc$$

Invers Matriks

Matriks invers dari suatu matriks A adalah matriks B yang apabila dikalikan dengan matriks A memberikan satuan I.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Matriks Kovarian dan Korelasi

Misalkan matriks X memiliki matriks kovarian Σ berikut:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix}$$

Maka untuk mencari nilai ρij dalam matriks ρ sebagai berikut:

$$\rho_{ij} = \frac{cov(x_i, x_j)}{\sqrt{var(x_i)\ var(x_j)}} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}$$

$$\boldsymbol{\rho} = \begin{bmatrix} \frac{\sigma_{11}}{\sqrt{\sigma_{11}\sigma_{11}}} & \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}} & \frac{\sigma_{13}}{\sqrt{\sigma_{11}\sigma_{33}}} \\ \frac{\sigma_{21}}{\sqrt{\sigma_{22}\sigma_{11}}} & \frac{\sigma_{22}}{\sqrt{\sigma_{22}\sigma_{22}}} & \frac{\sigma_{23}}{\sqrt{\sigma_{22}\sigma_{33}}} \\ \frac{\sigma_{31}}{\sqrt{\sigma_{33}\sigma_{11}}} & \frac{\sigma_{32}}{\sqrt{\sigma_{33}\sigma_{22}}} & \frac{\sigma_{33}}{\sqrt{\sigma_{33}\sigma_{33}}} \end{bmatrix}$$

Mencari Simpangan Baku dari Matriks Kovarian

Misal X memiliki matriks kovarian yakni:

$$\boldsymbol{\Sigma} = \begin{bmatrix} 25 & -2 & 4 \\ -2 & 4 & 1 \\ 4 & 1 & 9 \end{bmatrix}$$

Untuk mencari matriks $V^{1/2}$ digunakan persamaan sebagai berikut:

$$\boldsymbol{V}^{1/2} = \begin{bmatrix} \sqrt{\sigma_{11}} & 0 & 0 \\ 0 & \sqrt{\sigma_{22}} & 0 \\ 0 & 0 & \sqrt{\sigma_{33}} \end{bmatrix} = \begin{bmatrix} \sqrt{25} & 0 & 0 \\ 0 & \sqrt{4} & 0 \\ 0 & 0 & \sqrt{9} \end{bmatrix} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Akar Ciri & Vektor Ciri

Sebuah matriks A dikatakan positif definit jika $x'Ax > 0$ untuk sembarang vektor $x \neq 0$.

Karena vektor x tidak diketahui, maka untuk menunjukkan A merupakan definit positif maka dicari vektor cirinya terlebih dahulu menggunakan persamaan berikut: $|A - \lambda I| = 0$.

Latihan Soal 1

a, Tentukan akar ciri dan vektor ciri

b. Apakah matriks tsb definit posistif?

$$\mathbf{A} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Jawab:

a. $|A - \lambda I| = 0$

$$\left|\begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix} - \lambda\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right| = 0$$

$$\left|\begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}\right| = 0$$

$$\left|\begin{pmatrix} 2-\lambda & 0 & 0 \\ 0 & 4-\lambda & 0 \\ 0 & 0 & 3-\lambda \end{pmatrix}\right| = 0$$

$$(2-\lambda)(4-\lambda)(3-\lambda) = 0$$

Vektor ciri akan diperoleh dengan menyelesaikan persamaan $\boldsymbol{Ax} = \lambda\boldsymbol{x}$

- $\lambda_1 = 2$

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 2\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$2x_1 = 2x_1$

$4x_2 = 2x_2$ → diperoleh $x_1 = t, t \in R$, $x_2 = 0$ dan $x_3 = 0$ sehingga

$3x_3 = 2x_3$

$$\mathbf{x} = \begin{pmatrix} t \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} t, t \in R$$

vektor ciri untuk $\lambda_1 = 2$ adalah :

$$\boldsymbol{e} = \frac{\boldsymbol{x}}{\sqrt{\boldsymbol{x}'\boldsymbol{x}}} \qquad e = \frac{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}{\sqrt{(1\ \ 0\ \ 0)\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

- $\lambda_2 = 4$

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 4\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$2x_1 = 4x_1$

$4x_2 = 4x_2$ → diperoleh $x_1 = 0$, $x_2 = t$, $t \in R$ dan $x_3 = 0$ sehingga

$3x_3 = 4x_3$

$$\mathbf{x} = \begin{pmatrix} 0 \\ t \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} t, t \in R$$

vektor ciri untuk $\lambda_2 = 4$ adalah :

$$\boldsymbol{e} = \frac{\boldsymbol{x}}{\sqrt{\boldsymbol{x}'\boldsymbol{x}}}$$

$$\boldsymbol{e} = \frac{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}{\sqrt{(0\ \ 1\ \ 0)\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

- $\lambda_3 = 3$

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 3\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$2x_1 = 3x_1$

$4x_2 = 3x_2$ → diperoleh $x_1 = 0$, $x_2 = 0$ dan $x_3 = t, t \in R$ sehingga

$3x_3 = 3x_3$

$$\mathbf{x} = \begin{pmatrix} 0 \\ 0 \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} t, t \in R$$

vektor ciri untuk $\lambda_3 = 3$ adalah :

$$e = \frac{x}{\sqrt{x'x}}$$

$$e = \frac{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}{\sqrt{(0\ \ 0\ \ 1)\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

b.
- ✓ untuk matriks $\boldsymbol{A}$: $\lambda_1.\lambda_2.\lambda_3 = 2.4.3 > 0$, maka matriks **A** definit positif
- ✓ untuk matriks $\boldsymbol{B}$: $\lambda_1.\lambda_2.\lambda_3 = 2.4.0 \geq 0$, maka matriks **B** semi definit positif
- ✓ untuk matriks $\boldsymbol{C}$: $\lambda_1.\lambda_2.\lambda_3 = 1.1.1 > 0$, maka matriks **C** definit positif

Latihan Soal 2

Diketahui A adalah matriks simetris 2x2.. Tunjukkan bahwa determinan A adalah sama dengan akar ciri pertama kali akar ciri kedua.

Jawab:

Diketahui: $\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

- Mencari akar ciri dari matriks **A**:

$|\mathbf{A} - \boldsymbol{\lambda}\mathbf{I}| = 0$

$\begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = 0$

$(a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = 0$

$a_{11}a_{22} - a_{11}\lambda - a_{22}\lambda + \lambda^2 - a_{12}a_{21} = 0$

$\lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21} = 0$

Bentuk ini analog dengan rumus abc untuk persamaan kuadrat $a\lambda^2 + b\lambda + c = 0$, dimana

$\lambda_1\lambda_2 = \frac{c}{a}$

Sehingga didapat:

$\lambda_1\lambda_2 = \frac{a_{11}a_{22} - a_{12}a_{21}}{1} = a_{11}a_{22} - a_{12}a_{21}$

- Mencari determinan dari matriks **A**:

$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$

∴ terbukti bahwa $|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \lambda_1\lambda_2$.

Latihan Soal 3

Suatu matriks ragam peragam $\sum = \begin{pmatrix} 25 & -2 & 4 \\ -2 & 4 & 1 \\ 4 & 1 & 9 \end{pmatrix}$

a. Tentukan ρ_{13}

b. Tentukan korelasi X1 dan ½ X2+ ½ X3

Jawab:

a.

$$\rho_{13} = \frac{\sigma_{13}}{\sigma_1\sigma_3} = \frac{4}{\sqrt{25}\sqrt{9}} = \frac{4}{(5 * 3)} = \frac{4}{15} = 0.27$$

b. Misal,

$$Y = X_1$$

$$Z = \frac{1}{2}X_2 + \frac{1}{2}X_3$$

Maka untuk mencari simpangan baku dari Z,

$$\sigma_{zz} = Var\left(\frac{1}{2}X2 + \frac{1}{2}X3\right)$$

$$= \frac{1}{4}Var(X2) + \frac{1}{4}Var(X3) + 2\left(\frac{1}{2}\right)Cov(X2, X3)\left(\frac{1}{2}\right)$$

$$= \frac{1}{4}(4) + \frac{1}{4}(9) + \frac{1}{2}(1) = \frac{15}{4} = 3.75$$

$$\sigma_z = \sqrt{\sigma_{zz}} = 1.94$$

Dan kovarian dari Y dan Z diperoleh dengan

$$\sigma_{yz} = Var\ (Y, Z) = COV\left(X1, \frac{1}{2}X2 + \frac{1}{2}X3\right)$$

$$= Cov\left(X1, \frac{1}{2}X2\right) + Cov\left(X1, \frac{1}{2}X3\right)$$

$$= \frac{1}{2}Cov(X1, X2) + \frac{1}{2}Cov(X1, X3)$$

$$= \frac{1}{2}(-2) + \frac{1}{2}(4)$$

$$= -1 + 2$$

$$= 1$$

Sehingga didapatkan

$$\rho_{yz} = \frac{\sigma_{yz}}{\sqrt{\sigma_{yy}\sigma_{zz}}} = \frac{\sigma_{yz}}{\sigma_y\sigma_z}$$

$$\rho_{yz} = \frac{1}{5(1.94)}$$

Pertemuan 2 Sebaran Normal Ganda

Sebaran Normal Univariat

Fungsi kepekatan peluang dari sebaran Normal univariat (p=1) adalah sebagai berikut:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} \qquad \text{for } -\infty < x < \infty$$

Dengan parameter:

$\mu = E(X) = \text{mean}$
$\sigma^2 = \text{var}(X) = \text{variance}$

dalam bentuk lain dapat dinyatakan:

$$\left(\frac{x-\mu}{\sigma}\right)^2 = (x-\mu)(\sigma^2)^{-1}(x-\mu)$$

Sifat sebaran normal univariat:

Plot dari fungsi sebelumnya akan menghasilkan kurva berbentuk lonceng dengan ciri sebagai berikut:

1. Simetrik terhadap nilai tengah (μ)
2. Nilai tengah, median, dan modus berada pada titik yang sama
3. Peluang amatan berada antara μ±σ adalah 68%
4. Peluang amatan berada antara μ±1.96σ adalah 95%.

Sebaran Normal univariat beserta parameternya dinotasikan dengan N(μ, σ^2).

Sebaran Normal Ganda

Fungsi kepekatan sebaran Normal ganda (multivariate normal) merupakan generalisasi dari fungsi kepekatan Normal univariat dengan $p \geq 2$ dimensi.

$$\left(\frac{x-\mu}{\sigma}\right)^2 = (x-\mu)(\sigma^2)^{-1}(x-\mu) \xrightarrow[\text{Bentuk eksponen multivariat}]{p \geq 2} (\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) = \Delta^2$$

Jarak Mahalanobis

di mana:

- x' = [x1, ..., xp] yang merupakan vektor peubah
- μ' = [μ1, ..., μp] yang merupakan nilai tengah dari vektor acak x
- Σ merupakan matriks kovarian berukuran pxp

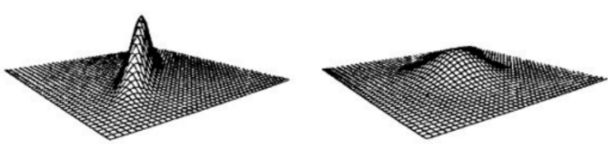
Sehingga diperoleh FKP sebaran Normal ganda sebagai berikut:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$$

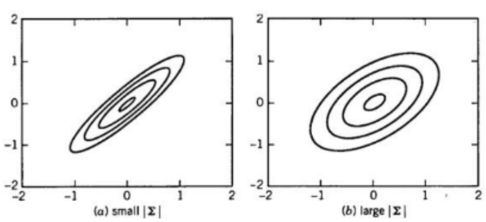
where $-\infty < x_i < \infty$ for $i = 1, \ldots, p$.

Sebaran Normal multivariat beserta parameternya dinotasikan dengan $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

Sebaran Normal Ganda – Bivariat



(a) small |Σ| (b) large |Σ|



Sifat-sifat Peubah Ganda Normal:

Sebaran Kombinasi Linier dari Peubah Ganda Normal

Kombinasi linier dari semua komponen peubah X juga menyebar Normal.

- Jika a adalah vektor konstanta, maka fungsi linier a'x = a1x1+ a2x2 + ... + apxp adalah univariat normal
  Sehingga, ketika x ~ Np(μ, Σ), maka a'x ~ N(a'μ, a'Σa)
- Jika A adalah konstanta q x p matriks pangkat q, di mana q ≤ p, q kombinasi linier di Ax memiliki sebaran multivariat normal.
  Sehingga, ketika x ~ Np(μ, Σ), maka Ax ~ Nq(Aμ, AΣA')

Kenormalan Baku

Vektor z diperoleh dengan: z = $(T)^{-1}$ (x - μ)

di mana Σ = T'T dan T diperoleh menggunakan pemfaktoran dengan prosedur Cholesky, dengan $z = (\Sigma^{1/2})^{-1}(x - \mu)$ dan $\Sigma^{1/2}$adalah matriks akar kuadrat simetris yang didefinisikan dari $\Sigma = \Sigma^{1/2}\Sigma^{1/2}$

Jika x ~ Np(μ, Σ), maka z ~ Np(0, I)

Distribusi Chi Kuadrat

Jika z merupakan vektor standar yang telah didefinisikan sebelumnya, maka z'z ~ $\chi^2_{(p)}$.

z'z juga dapat diperoleh dari: z'z = (x - μ)'Σ^{-1} (x - μ).

Maka dari itu, jika x ~ Np(μ, Σ) maka (x - μ)'Σ^{-1} (x - μ) ~ $\chi^2_{(p)}$

Normalitas Distribusi Marginal

Jika **x** ~ $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, maka setiap partisi dari **x** mengikuti sebaran normal multivariat, dengan ketentuan:

$$X = \begin{bmatrix} X_{(1)} \\ X_{(2)} \end{bmatrix}_{p\,X\,1}, \quad \mu = \begin{bmatrix} \mu_{(1)} \\ \mu_{(2)} \end{bmatrix} \ and \ \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

$$\text{where } X_{(1)} = \begin{bmatrix} X_1 \\ \vdots \\ X_q \end{bmatrix}, \quad X_{(2)} = \begin{bmatrix} X_{(q+1)} \\ \vdots \\ X_p \end{bmatrix}, \quad \mu_{(1)} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_q \end{bmatrix}, \quad \mu_{(2)} = \begin{bmatrix} \mu_{(q+1)} \\ \vdots \\ \mu_p \end{bmatrix}$$

$$\Sigma_{ij} = E\left[(X_{(i)} - \mu_{(i)})(X_{(j)} - \mu_{(j)})'\right], i,j = 1,2$$

Jika $\mathbf{X}_{(1)}$ dan $\mathbf{X}_{(2)}$ saling bebas: $\Sigma = \begin{bmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{bmatrix}$

Jika **x** ~ $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, maka $\mathbf{x}_1 \sim N_q(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$ **(tidak berlaku sebaliknya)**

Sebaran Bersyarat

Jika **x** dan **y** saling bebas, maka kovarian **x** dan **y** adalah 0.

Jika **x** dan **y** saling bebas, maka $\boldsymbol{\Sigma}_{xy} = \boldsymbol{\Sigma}_{yx} = \mathbf{0}$, dan sebaran bersyarat dari **y** jika diberikan **x**, di mana **x** ~ $N_p(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_{xx})$ dan **y** ~ $N_q(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_{yy})$, maka sebaran bersyarat (**y**|**x**) adalah

$$(\boldsymbol{y}|\boldsymbol{x}) \sim N_{p+q}\left(\begin{bmatrix} \boldsymbol{\mu}_x \\ - \\ \boldsymbol{\mu}_y \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{xx} & \mathbf{0} \\ - & - \\ \mathbf{0} & \boldsymbol{\Sigma}_{yy} \end{bmatrix}\right)$$

Jika **tidak** saling bebas, maka $\boldsymbol{\Sigma}_{xy} = \boldsymbol{\Sigma}_{yx} \neq \mathbf{0}$, nilai tengah dan ragam sebaran bersyarat diperoleh dengan:

$$E(\boldsymbol{y}|\boldsymbol{x}) = \boldsymbol{\mu}_y + \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}_x)$$

$$cov(\boldsymbol{y}|\boldsymbol{x}) = \boldsymbol{\Sigma}_{yy} - \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\Sigma}_{xy}$$

Eksplorasi Sebaran Normal Ganda

1. Untuk mengevaluasi apakah data yang dimiliki menyebar normal ganda dapat ditelusuri secara eksplorasi
2. Seperti halnya untuk kasus univariate penelusuran sebaran normal gandadapat juga memanfaatkan plot quantil-quantil -> quantil khi-kuadrat

Pertemuan 3 Uji Vektor Nilai Tengah 1 Populasi

Uji Nilai Tengah

- Jika ingin mengambil keputusan valid mengenai rataan dari suatu populasi berdasarkan contoh yang diperoleh maka dilakukan uji nilai tengah.
- Dari satu populasi diambil sejumlah contoh, di mana contoh tersebut terdiri dari beberapa peubah yang saling berkorelasi sebanyak p → maka harus dilakukan analisis secara bersama-sama.
- Analisis beberapa peubah yang dilakukan secara bersama-sama akan Memberikan hasil uji yang lebih valid.
- Uji ini didasarkan pada hipotesis berikut: $H_0: \boldsymbol{\mu} = \boldsymbol{\mu_0}$

  $H_1: \boldsymbol{\mu} \neq \boldsymbol{\mu_0}$
- Di mana μ_{px1} merupakan vektor nilai tengah populasi dan $\mu_{0,px1}$ merupakan beberapa nilai tertentu di bawah hipotesis nol.
- **p** merupakan banyak peubah yang diuji nilai tengahnya. Jika **p = 1**, maka pengujian dilakukan dengan statistik t.
- Jika **p > 1**, maka pengujian dilakukan dengan statistik T^2 Hotelling.

Kasus Univariat

- Jika kita punya contoh acak dari *n* amatan dari suatu populasi, di mana:
  - Amatan saling bebas
  - Amatan berasal dari populasi yang sama, $E(X_i) = \mu$ untuk semua *i*
  - Jika ukuran contoh kecil, diasumsikan $X_i \sim N(\mu, \sigma^2)$

  Maka :

  $$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$
- Jika ingin dilakukan pengujian terhadap nilai tengah, digunakan statistik uji berikut:

  $$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} \sim t_{n-1}$$

Pada Kasus p Dimensi, p > 1

- Jika dicari kuadrat dari statistik uji *t*:

  $$t^2 = \frac{(\bar{x} - \mu_0)^2}{\frac{s^2}{n}} = n(\bar{x} - \mu_0)(s^2)^{-1}(\bar{x} - \mu_0)$$
- Diperoleh bahwa *t2* adalah jarak kuadrat statistik antara rataan contoh dan nilai hipotesis $\boldsymbol{\mu}_0$
- Ingat bahwa $t^2_{df} = F_{1,df}$
- Maka,

  $$t^2 = n(\bar{x} - \mu_0)(s^2)^{-1}(\bar{x} - \mu_0) \sim F_{1,df}$$
- Di mana ***df = n - p***

Kasus Multivariat

$$H_0: \boldsymbol{\mu} = \boldsymbol{\mu_0}$$

$$H_1: \boldsymbol{\mu} \neq \boldsymbol{\mu_0}$$

- Untuk kasus multivariat, ganti skalar dengan vektor dan matriks sebagai berikut:

  $$T^2 = n(\bar{\boldsymbol{x}} - \boldsymbol{\mu}_0)'\boldsymbol{S}^{-1}(\bar{\boldsymbol{x}} - \boldsymbol{\mu}_0) \sim \frac{(n-1)p}{n-p}F_{p,(n-p)} = c^2$$
- Di mana:

  $$\bar{\boldsymbol{x}}_{p\times1} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{x}_i \qquad \boldsymbol{\mu_0}_{p\times1} = \begin{pmatrix} \mu_{1_0} \\ \mu_{2_0} \\ \vdots \\ \mu_{p_0} \end{pmatrix}$$

  $$\boldsymbol{S}_{p\times p} = \frac{1}{n-1}\sum_{i=1}^{n}(\boldsymbol{x}_i - \bar{\boldsymbol{x}})(\boldsymbol{x}_i - \boldsymbol{x})'$$
- Tolak H_0 jika:

  $$T^2 > c^2 = \frac{(n-1)p}{(n-p)}F_{(p,n-p)}(\alpha)$$

Selang Kepercayaan Multivariat

- Untuk kasus univariat, selang kepercayaan untuk parameter θ adalah daerah yang memuat nilai sebenarnya dari parameter tersebut dengan peluang 1 - α.

- Untuk kasus multivariat berdimensi p, daerah kepercayaan untuk parameter θ adalah daerah yang memuat nilai sebenarnya dari parameter tersebut dengan peluang sebesar 1 - α.
- Beberapa selang kepercayaan yang dapat digunakan pada kasus multivariat antara lain selang kepercayaan simultan, Bonferroni, dan ellips.

Selang Kepercayaan Simultan

- Digunakan untuk mencari selang kepercayaan pada setiap parameter peubah ke-*i*.
- Batas-batas selang kepercayaan **(1 - α)100% bagi μ** diperoleh dengan rumus:

$$\bar{x}_i \pm c\sqrt{\frac{Sii}{n}} \quad \text{atau} \quad \boldsymbol{a'\bar{x}} \pm c\sqrt{\frac{\boldsymbol{a'Sa}}{n}} \quad \text{dengan} \quad c^2 = \frac{(n-1)p}{n-p}F_{(p,n-p)}(\alpha)$$

- Peluang nilai parameter akan berada di dalam selang adalah:

$$P\left(\bar{x}_i - \sqrt{\frac{(n-1)p}{n-p}F_{(p,n-p)}(\alpha)}\sqrt{\frac{Sii}{n}} \le \mu_i \le \bar{x}_i + \sqrt{\frac{(n-1)p}{n-p}F_{(p,n-p)}(\alpha)}\sqrt{\frac{Sii}{n}}\right) = 1-\alpha$$

- Di mana *i* = 1, ..., *p*

Selang Kepercayaan Ellips

- Selang ini memuat nilai $\boldsymbol{\mu}_0$ yang tidak akan ditolak oleh T^2 Hotelling pada taraf nyata α. Selang ini dinyatakan dengan:

$$n(\bar{\boldsymbol{x}} - \boldsymbol{\mu})'\boldsymbol{S}^{-1}(\bar{\boldsymbol{x}} - \boldsymbol{\mu}) \le c^2 = \frac{(n-1)p}{n-p}F_{p,(n-p)}(\alpha)$$

$$P\left(n(\bar{\boldsymbol{x}} - \boldsymbol{\mu})'\boldsymbol{S}^{-1}(\bar{\boldsymbol{x}} - \boldsymbol{\mu}) \le \frac{(n-1)p}{n-p}F_{p,(n-p)}(\alpha)\right) = 1-\alpha$$

- Selang ini dibentuk dengan:
  - Panjang ½ sumbu mayor = $\frac{\sqrt{\lambda_1}}{\sqrt{n}}c$
  - Panjang ½ sumbu minor = $\frac{\sqrt{\lambda_2}}{\sqrt{n}}c$
  - Daerah kepercayaan dinyatakan dengan

$$\bar{\mathbf{x}} \pm \frac{\sqrt{\lambda_i}}{\sqrt{n}}ce_i$$

  Di mana e_i adalah vektor eigen dari eigen matriks S

  Dan nilai c diperoleh dari $c^2 = \frac{(n-1)p}{n-p}F_{p,(n-p)}(\alpha)$

Selang Kepercayaan Bonferroni

- Jika ingin membentuk sebanyak m selang kepercayaan dengan masing-masing selang memuat (1 - α)100% amatan, maka peluang dari amatan berada pada selang secara simultan atau keseluruhan akan berkurang menjadi sebesar (1 - mα)100%.
- Maka sebaliknya, jika ingin memperoleh selang simultan dengan peluang (1 -α)100%, maka pada setiap selang individu dipilih taraf sebesar α/m dan menghasilkan peluang masing-masing selang sebesar (1 - α/m)100%.
- Metode dalam membentuk selang kepercayaan simultan ini disebut Metode Bonferroni, di mana akan menghasilkan selang yang lebih lebar dibandingkan selang individualnya.
- Selang ini lemah ketika nilai m nya besar, karena dianggap terlalu konservatif.
- Selang ini dinyatakan dengan:

$$P(\textit{Nilai parameter berada pada selang } C_i) = 1 - \alpha_i$$

$$P(\textit{Semua nilai parameter berada pada selang } C) = 1 - (\alpha_1 + \cdots + \alpha_p)$$

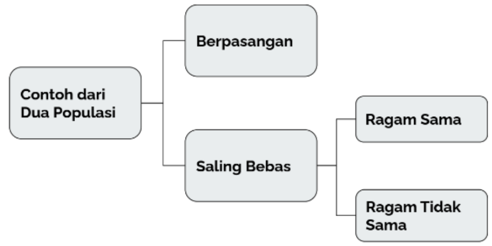
- Adapun batas wilayah dan peluang selang Bonferroni pada setiap selang dinyatakan dengan:

$$\bar{x}_i \pm t_{(n-1)}\left(\frac{\alpha}{2p}\right)\sqrt{\frac{S_{ii}}{n}}$$

$$P\left(\bar{x}_i - t_{(n-1)}\left(\frac{\alpha}{2p}\right)\sqrt{\frac{S_{ii}}{n}} \le \mu_i \le \bar{x}_i + t_{(n-1)}\left(\frac{\alpha}{2p}\right)\sqrt{\frac{S_{ii}}{n}}\right) = 1-\alpha$$

Pertemuan 4 Uji Vektor Nilai Tengah 2 Populasi

Uji Hipotesis terhadap Nilai Tengah Dua Populasi



- Uji ini didasarkan pada:

$$\boxed{\begin{aligned} H_0&: \boldsymbol{\mu_1} = \boldsymbol{\mu_2} \\ H_1&: \boldsymbol{\mu_1} \ne \boldsymbol{\mu_2}\end{aligned}}$$

- Di mana:

$$\boldsymbol{\mu_1} = \begin{bmatrix}\mu_{11}\\ \mu_{12}\\ \mu_{13}\\ \vdots\\ \mu_{1p}\end{bmatrix} \text{ dan } \boldsymbol{\mu_2} = \begin{bmatrix}\mu_{21}\\ \mu_{22}\\ \mu_{23}\\ \vdots\\ \mu_{2p}\end{bmatrix}$$

Biasanya uji berpasangan dilakukan ketika populasi 1 merupakan populasi yg muncul sebelum populasi 2.

Misal: nilai 5 mata kuliah yang diambil mahasiswa sebelum dan setelah pelatihan kompetensi R.

- Uji ini didasarkan pada:

$$H_0: \boldsymbol{\delta} = \boldsymbol{\mu_1} - \boldsymbol{\mu_2} = \boldsymbol{0}$$
$$H_1: \boldsymbol{\delta} = \boldsymbol{\mu_1} - \boldsymbol{\mu_2} \ne \boldsymbol{0}$$

- Statistik Uji:

$$\boxed{T^2 = n(\bar{\boldsymbol{d}} - \boldsymbol{\delta})'\boldsymbol{S}_d^{-1}(\bar{\boldsymbol{d}} - \boldsymbol{\delta}) \sim \frac{(n-1)p}{n-p}F_{(p,n-p)}}$$

$$\bar{\boldsymbol{d}} = \frac{1}{n}\sum_{j=1}^{n} d_j$$

$$\boldsymbol{S}_d = \frac{1}{n-1}\sum_{j=1}^{n}(\boldsymbol{d}_j - \bar{\boldsymbol{d}})(\boldsymbol{d}_j - \bar{\boldsymbol{d}})'$$

- Tolak H0 ketika:

$$T^2 > c^2 = \frac{(n-1)p}{n-p}F_{(p,n-p)(\alpha)}$$

Dengan **n** adalah ukuran contoh, dan **p** adalah banyak peubah

Selang Kepercayaan Dua Populasi Saling Berpasangan

- **Simultan**

$$\bar{d}_i - \sqrt{\frac{(n-1)p}{n-p}F_{(p,n-p)(\alpha)}}\sqrt{\frac{S_{d_i}^2}{n_i}} < \delta_i < \bar{d}_i + \sqrt{\frac{(n-1)p}{n-p}F_{(p,n-p)(\alpha)}}\sqrt{\frac{S_{d_i}^2}{n_i}}$$

- **Bonferroni**

$$\bar{d}_i - t_{n-1\left(\frac{\alpha}{2p}\right)}\sqrt{\frac{S_{d_i}^2}{n_i}} < \delta_i < \bar{d}_i + t_{n-1\left(\frac{\alpha}{2p}\right)}\sqrt{\frac{S_{d_i}^2}{n_i}}$$

Uji Hipotesis terhadap Nilai Tengah Dua Populasi Saling Bebas

- Setiap populasi diambil contoh acak berukuran tertentu (bisa sama, bisa beda. Anggaplah ukurannya n1 dan n2)
- Pengambilan contoh dari masing-masing populasi saling bebas.
- Dilakukan untuk menguji apakah parameter nilai tengah pada kedua populasi sama.

Uji Hipotesis Dua Populasi Saling Bebas (Σ1 = Σ2)

- Uji ini didasarkan pada:

$$H_0: \boldsymbol{\mu_1} = \boldsymbol{\mu_2}$$
$$H_1: \boldsymbol{\mu_1} \ne \boldsymbol{\mu_2}$$

- Statistik Uji:

$$\boldsymbol{S}_{gab} = \frac{(n_1-1)\boldsymbol{S}_1 + (n_2-1)\boldsymbol{S}_2}{n_1+n_2-2}$$

$$T^2 = (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2)'\left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\boldsymbol{S}_{gab}\right]^{-1}(\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2) \sim \frac{(n_1+n_2-2)p}{n_1+n_2-p-1}F_{(p,n_1+n_2-p-1)}$$

- Tolak H0 ketika:

$$T^2 > c^2 = \frac{(n_1+n_2-2)p}{n_1+n_2-p-1}F_{(p,n_1+n_2-p-1)(\alpha)}$$

Dengan **n** adalah ukuran contoh, dan **p** adalah banyak peubah

Selang Kepercayaan Dua Populasi Saling Bebas (Σ1= Σ2)

- Simultan

$$\boldsymbol{a}'(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) \le \boldsymbol{a}'(\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2) \pm \sqrt{\frac{(n_1+n_2-2)p}{n_1+n_2-p-1}F_{(p,n_1+n_2-p-1)(\alpha)}}\sqrt{\boldsymbol{a}'\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\boldsymbol{S}_{gab}\boldsymbol{a}}$$

- Bonferroni

$$(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) \le (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2) \pm t_{(n_1+n_2-2),\left(\frac{\alpha}{2p}\right)}\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)S_{ii}}$$

Uji Hipotesis Dua Populasi Saling Bebas (Σ1 ≠ Σ2)

- Uji ini didasarkan pada:

$$H_0: \boldsymbol{\mu_1} = \boldsymbol{\mu_2}$$
$$H_1: \boldsymbol{\mu_1} \ne \boldsymbol{\mu_2}$$

- Statistik Uji:

$$T^2 = (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2)'\left(\frac{\boldsymbol{S}_1}{n_1}+\frac{\boldsymbol{S}_2}{n_2}\right)^{-1}(\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2) \sim \chi^2_p$$

- Tolak H0 ketika:

$$T^2 > \chi^2_{\alpha,p}$$

Dengan **p** merupakan banyak peubah

Selang Kepercayaan Dua Populasi Saling Bebas (Σ1 ≠ Σ2)

- **Simultan**

$$\boldsymbol{a}'(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) \le \boldsymbol{a}'(\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2) \pm \sqrt{\chi^2_{\alpha,p}\boldsymbol{a}'\left(\frac{\boldsymbol{S}_1}{n_1}+\frac{\boldsymbol{S}_2}{n_2}\right)\boldsymbol{a}}$$

Pertemuan 5 Analisis Ragam Peubah Ganda (MANOVA)

Pengantar Multivariate ANOVA (MANOVA)

- Merupakan generalisasi dari ANOVA pada respon multivariat (p > 1)
- Pada kasus multivariat, misal terdapat sekumpulan contoh acak yang diambil dari *g* populasi:

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

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

  ...

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

- Sekumpulan acak tersebut memerlukan asumsi dasar berikut:

$$\boldsymbol{X}_l(l = 1,2,\ldots,g) \sim N_p(\boldsymbol{\mu}_l, \boldsymbol{\Sigma})$$

Di mana:

- $\mathbf{X}_l$ merupakan sampel acak berukuran n_l dari suatu populasi l dengan rata-rata $\boldsymbol{\mu}_l$
- Matriks kovarians antara *g* populasi sama
- Setiap populasi menyebar normal multivariat

Uji Asumsi Homogenitas Matriks

$$H_0: \boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \cdots = \boldsymbol{\Sigma}_g = \boldsymbol{\Sigma}_0$$

H_1: Setidaknya satu pasang $\boldsymbol{\Sigma}_i$ yang berbeda

$$\boldsymbol{S} = \frac{1}{N}\sum_{l=1}^{g}(n_l - 1)\boldsymbol{S}_l \quad \text{dengan} \quad N = \sum_{l=1}^{g} n_l - g$$

Statistik uji:

H_0 ditolak ketika $MC^{-1} > \chi^2_{\left(\frac{1}{2}(g-1)p(p+1)\right)}(\alpha)$

$$M = \sum_{l=1}^{g}(n_l-1)\ln|\boldsymbol{S}| - \sum_{l=1}^{g}(n_l-1)\ln|\boldsymbol{S}_l|$$

$$C^{-1} = 1 - \frac{2p^2+3p-1}{g(p+1)(g-1)}\left(\sum_{l=1}^{g}\frac{1}{n_l-1} - \frac{1}{\sum_{l=1}^{g}(n_l-1)}\right)$$

Uji Asumsi Normalitas Multivariat

$\mathrm{H_0}$: Data berdistribusi normal multivariat

$\mathrm{H_1}$: Data tidak berdistribusi normal multivariat.

Jika $\boldsymbol{X}_1, \boldsymbol{X}_2, \dots, \boldsymbol{X}_g$ berdistribusi normal multivariat maka:

$(\boldsymbol{x}-\boldsymbol{\mu})^t \boldsymbol{S}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) \sim \chi^2_p$

Berdasarkan sifat ini, pemeriksaan distribusi normal multivariat dapat dengan membuat qq plot.

Statistik Uji Keputusan dalam MANOVA

Pillai's Trace

Statistik uji ini paling cocok digunakan jika asumsi homogenitas matriks varians-kovarians tidak dipenuhi, ukuran-ukuran sampel kecil, dan jika hasil-hasil dari pengujian bertentangan satu sama lain yaitu jika ada beberapa vektor rata-rata yang berbeda sedang yang lain tidak. Semakin tinggi nilai statistik Pillai's Trace, pengaruh terhadap model semakin besar.

$$P = \sum_{i=1}^{p}\left(\frac{\lambda_i}{1+\lambda_i}\right) = tr\ \lambda_i(1+\lambda_i)^{-1} = tr\ \frac{|B|}{|B+W|}$$

dimana $\lambda_1, \lambda_2, \dots, \lambda_p$ adalah akar-akar karakteristik dari $(W)^{-1}(B)$.

(W) = matriks varians-kovarians galat pada MANOVA

(B) = matriks varians-kovarians perlakuan pada MANOVA

Wilks' Lambda

Statistik uji digunakan jika terdapat lebih dari dua kelompok variabel independen dan asumsi homogenitas matriks varians-kovarians dipenuhi. Semakin rendah nilai statistik Wilk's Lambda, pengaruh terhadap model semakin besar. Nilai Wilk's Lambda berkisar antara 0-1.

$$U = \prod_{i=1}^{p}(1+\lambda_i)^{-1} = \frac{|W|}{|B+W|}$$

Hotelling's Trace

Statistik uji ini cocok digunakan jika hanya terdapat dua kelompok variabel independen. Semakin tinggi nilai statistik Hotelling's Trace, pengaruh terhadap model semakin besar.

$$T = \sum_{i=1}^{p}\lambda_i = tr\ \lambda_i = tr\ (W)^{-1}(B)$$

Roy's Largest Root.

Statistik uji ini hanya digunakan jika asumsi homogenitas varians-kovarians dipenuhi. Semakin tinggi nilai statistik Roy's Largest Root, pengaruh terhadap model semakin besar. Nilai Roy's Largest Root > Hotelling's Trace > Pillai's Trace. Dalam hal pelanggaran asumsi normalitas multivariat, statistik ini kurang robust (kekar) dibandingkan dengan statistik uji yang lainnya.

$$R = \lambda_{maks} = maks\left(\lambda_1, \lambda_2, \dots, \lambda_p\right)$$

$= $ akar karakteristik maksimum dari $(W)^{-1}(B)$

One Way MANOVA - RAL

- Ciri-ciri: Keragaman berasal dari satu arah (perlakuan) dari data yang diamati.
- Mode linier:

$$\boldsymbol{X}_{lj} = \boldsymbol{\mu} + \boldsymbol{\tau}_l + \boldsymbol{\varepsilon}_{lj},\quad l = 1, 2, \dots, g \quad dan \quad j = 1, 2, \dots, n_l$$

$\boldsymbol{X}_{lj}$: Pengamatan pada perlakuan ke-*l*, ulangan ke-*j*

$\boldsymbol{\mu}$: Vektor nilai tengah umum

$\boldsymbol{\tau}_l$: Pengaruh perlakuan ke-*l*

$\boldsymbol{\varepsilon}_{lj}$: Peubah acak $N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$

- Hipotesis: $H_0: \boldsymbol{\tau}_1 = \boldsymbol{\tau}_2 = \cdots = \boldsymbol{\tau}_g = \boldsymbol{\tau}_0 = \boldsymbol{0}$ $\quad \boldsymbol{\tau}_l = \begin{bmatrix}\boldsymbol{\mu}_{l1}\\ \boldsymbol{\mu}_{l2}\\ \dots\\ \boldsymbol{\mu}_{lp}\end{bmatrix}, l = 1, 2, \dots, g$

  H_1: Minimal ada satu $\boldsymbol{\tau}_l \neq \boldsymbol{0}$

Sumber Variansi	Matriks jumlah dari kuadrat dan hasil kali	Derajat bebas
Perlakuan	$B = \sum_{l=1}^{g} n_l(\bar{x}_{l.} - \bar{x})(\bar{x}_{l.} - \bar{x})^t$	$g-1$
Galat (sisa)	$W = \sum_{l=1}^{g}\sum_{j=1}^{n_l}(x_{lj} - \bar{x}_{l.})(x_{lj} - \bar{x}_{l.})^t$	$\sum_{l=1}^{g} n_l - g$
total	$\sum_{l=1}^{g}\sum_{j=1}^{n_l}(x_{lj} - \bar{x})(x_{lj} - \bar{x})^t$	$\sum_{l=1}^{g} n_l - 1$

Statistik Uji diperoleh dengan statistik Wilks' Lambda: $\Lambda^* = \frac{|W|}{|W+B|}$

Tolak H0 jika $\Lambda^* < \Lambda(p, df1, df2)(\alpha)$

Di mana df1 = df hipotesis, df2 = df galat

Statistik Uji sebelumnya akan dibandingkan dengan titik kritis berikut:

Variabel	Grup	Distribusi sampling untuk data normal multivariat
$p = 1$	$g \geq 2$	$\left(\frac{\sum_{l=1}^{g} n_l - g}{g-1}\right)\left(\frac{1-\Lambda^*}{\Lambda^*}\right) \sim F_{g-1, n_l - g}$
$p = 2$	$g \geq 2$	$\left(\frac{\sum_{l=1}^{g} n_l - g - 1}{g-1}\right)\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \sim F_{2(g-1), 2(\sum n_l - g - 1)}$
$p \geq 1$	$g = 2$	$\left(\frac{\sum_{l=1}^{g} n_l - p - 1}{p}\right)\left(\frac{1-\Lambda^*}{\Lambda^*}\right) \sim F_{p, \sum n_l - p - 1}$
$p \geq 1$	$g = 3$	$\left(\frac{\sum_{l=1}^{g} n_l - p - 2}{p}\right)\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \sim F_{2p, 2(\sum n_l - p - 2)}$

One Way MANOVA - RAK

- Model linier:

$$\boldsymbol{X}_{lj} = \boldsymbol{\mu} + \boldsymbol{\tau}_l + \boldsymbol{\beta}_j + \boldsymbol{\varepsilon}_{lj},\quad l = 1, 2, \dots, g \quad dan \quad j = 1, 2, \dots, n_l$$

$\boldsymbol{X}_{lj}$: Pengamatan pada perlakuan ke-*l*, ulangan ke-*j*

$\boldsymbol{\mu}$: Vektor nilai tengah umum

$\boldsymbol{\tau}_l$: Pengaruh perlakuan ke-*l*

$\boldsymbol{\beta}_j$: Pengaruh kelompok ke-*j*

$\boldsymbol{\varepsilon}_{lj}$: Peubah acak $N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$

Hipotesis pengaruh perlakuan:

$H_0: \boldsymbol{\tau}_1 = \boldsymbol{\tau}_2 = \cdots = \boldsymbol{\tau}_g = \boldsymbol{\tau}_0 = \boldsymbol{0}$

H_1: Minimal ada satu $\boldsymbol{\tau}_l \neq \boldsymbol{0}$

Hipotesis pengaruh kelompok:

$H_0: \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \cdots = \boldsymbol{\beta}_b = \boldsymbol{\beta}_0 = \boldsymbol{0}$

H_1: Minimal ada satu $\boldsymbol{\beta}_j \neq \boldsymbol{0}$

Sumber Variansi	Matriks jumlah dari kuadrat dan hasil kali
Perlakuan	$B = \sum_{l=1}^{g} n_l(\bar{x}_{l.} - \bar{x})(\bar{x}_{l.} - \bar{x})^t$
Kelompok	$K = \sum_{l=1}^{g} n_g(\bar{x}_{.j} - \bar{x})(\bar{x}_{.j} - \bar{x})^t$
Galat (sisa)	$W = \sum_{l=1}^{g}\sum_{j=1}^{n_l}(x_{lj} - \bar{x}_{l.})(x_{lj} - \bar{x}_{l.})^t$
total	$\sum_{l=1}^{g}\sum_{j=1}^{n_l}(x_{lj} - \bar{x})(x_{lj} - \bar{x})^t$

Statistik Uji diperoleh dengan statistik Wilks' Lambda:

$$\Lambda^*(\text{perlakuan}) = \frac{|W|}{|W+B|}$$

$$\Lambda^*(\text{kelompok}) = \frac{|W|}{|W+K|}$$

Statistik Uji sebelumnya akan dibandingkan dengan titik kritis berikut:

Variabel	Grup	Distribusi sampling untuk data normal multivariat
$p = 1$	$g \geq 2$	$\left(\frac{\sum_{l=1}^{g} n_l - g}{g-1}\right)\left(\frac{1-\Lambda^*}{\Lambda^*}\right) \sim F_{g-1, n_l - g}$
$p = 2$	$g \geq 2$	$\left(\frac{\sum_{l=1}^{g} n_l - g - 1}{g-1}\right)\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \sim F_{2(g-1), 2(\sum n_l - g - 1)}$
$p \geq 1$	$g = 2$	$\left(\frac{\sum_{l=1}^{g} n_l - p - 1}{p}\right)\left(\frac{1-\Lambda^*}{\Lambda^*}\right) \sim F_{p, \sum n_l - p - 1}$
$p \geq 1$	$g = 3$	$\left(\frac{\sum_{l=1}^{g} n_l - p - 2}{p}\right)\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \sim F_{2p, 2(\sum n_l - p - 2)}$

Two Ways MANOVA

- Ciri-ciri: Keragaman berasal dari dua arah (perlakuan) dan dapat terjadi interaksi antar faktor.
- Mode linier:

$$\boldsymbol{X}_{lj} = \boldsymbol{\mu} + \boldsymbol{\tau}_l + \boldsymbol{\beta}_j + \boldsymbol{\gamma}_{lk} + \boldsymbol{\varepsilon}_{lkj},\quad l = 1, 2, \dots, g;\quad k = 1, 2, \dots, b \quad dan \quad j = 1, 2, \dots, n_l$$

$\boldsymbol{X}_{lj}$: Pengamatan pada perlakuan ke-*l*, ulangan ke-*j*

$\boldsymbol{\mu}$: Vektor nilai tengah umum

$\boldsymbol{\tau}_l$: Pengaruh perlakuan ke-*l*

$\boldsymbol{\beta}_j$: Pengaruh kelompok ke-*j*

$\boldsymbol{\gamma}_{lk}$: Komponen interaksi

$\boldsymbol{\varepsilon}_{lkj}$: Peubah acak $N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$

- Uji Hipotesis interaksi:

$H_0: \boldsymbol{\gamma}_{11} = \boldsymbol{\gamma}_{12} = \cdots = \boldsymbol{\gamma}_{gb} = \boldsymbol{0}$

H_1: At least one $\boldsymbol{\gamma}_{\ell k} \neq \boldsymbol{0}$

Untuk ukuran contoh kecil gunakan:

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

Untuk ukuran contoh besar dengan Bartlett's Multiplier untuk meningkatkan aproksimasi:

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

- Uji Hipotesis Faktor 1:

$H_0: \boldsymbol{\tau}_1 = \boldsymbol{\tau}_2 = \cdots = \boldsymbol{\tau}_g = \boldsymbol{\tau}_0 = \boldsymbol{0}$

H_1: Minimal ada satu $\boldsymbol{\tau}_l \neq \boldsymbol{0}$

Untuk ukuran contoh kecil gunakan:

$$\Lambda^* = \frac{|\mathrm{SSP_{res}}|}{|\mathrm{SSP_{fac\,1}} + \mathrm{SSP_{res}}|}$$

Untuk ukuran contoh besar dengan Bartlett's Multiplier untuk meningkatkan aproksimasi:

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

- Uji Hipotesis Faktor 2:

$H_0: \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \cdots = \boldsymbol{\beta}_b = \boldsymbol{\beta}_0 = \mathbf{0}$

H_1: Minimal ada satu $\boldsymbol{\beta}_j \neq \mathbf{0}$

Untuk ukuran contoh kecil gunakan:

$$\Lambda^* = \frac{|\,\mathrm{SSP_{res}}\,|}{|\,\mathrm{SSP_{fac\,2}} + \mathrm{SSP_{res}}\,|}$$

Untuk ukuran contoh besar dengan Bartlett's Multiplier untuk meningkatkan aproksimasi:

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

Pertemuan 6 Analisis Profil

Analisis Profil

- Sebuah metode dalam statistika multivariat yang menganalisis rangkaian perlakuan (sebanyak p) yang dikenakan pada dua populasi/grup/kelompok atau lebih.
- Analisis profil merupakan perluasan dari pengukuran MANOVA yang berulang.
- Tujuannya untuk mengetahui pengaruh perlakuan yang satu dibandingkan dengan yang lain pada setiap populasi. Populasi yang berbeda-beda bisa berupa waktu, tempat, grup, atau lainnya.
- Beberapa asumsi yang mendasari:
  - Semua respon diukur dalam unit/satuan/skala yang sama agar dapat dijumlah/dibandingkan
  - Antar respon perlakuan pada kelompok/populasi yang berbeda saling bebas
  - Galat menyebar normal dengan rataan 0 dan simpangan baku σ.
- Beberapa contoh kasus yang dapat dilakukan analisis profil:
  - X1, X2, ..., X6 adalah hasil pengukuran detak jantung seseorang sebanyak 6 kali selama 24 jam
  - X1, X2, ..., X8 adalah total volume susu yang dihasilkan seekor sapi pada laktasi yang ke 1, 2, .., hingga 8 minggu
- Tiga hipotesis yang diuji:
  - Uji hipotesis kesejajaran (parallel test) $\rightarrow H_{01}$
  - Uji hipotesis keberhimpitan (coincident test) $\rightarrow H_{02}$
  - Uji hipotesis kesamaan (level test) $\rightarrow H_{03}$

Uji Hipotesis Analisis Profil

- Notasi analisis profil dalam persamaan matriks adalah sebagai berikut:

$$\begin{pmatrix} y_{11} \\ \cdots \\ y_{1n_1} \\ y_{21} \\ \cdots \\ y_{2n_2} \\ \cdots \\ y_{i1} \\ \cdots \\ y_{in_i} \end{pmatrix} = \begin{pmatrix} 1 & 0 & .. & 0 \\ .. & .. & .. & .. \\ 1 & 0 & .. & 0 \\ 0 & 1 & .. & 0 \\ .. & .. & .. & .. \\ 0 & 1 & .. & 0 \\ .. & .. & .. & .. \\ 0 & 0 & .. & 1 \\ .. & .. & .. & .. \\ 0 & 0 & .. & 1 \end{pmatrix} \begin{pmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1p} \\ \mu_{21} & \mu_{22} & \cdots & \mu_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{i1} & \mu_{i2} & \cdots & \mu_{ip} \end{pmatrix} + \begin{pmatrix} e_{11} \\ \cdots \\ e_{1n_1} \\ e_{21} \\ \cdots \\ e_{2n_2} \\ \cdots \\ e_{i1} \\ \cdots \\ e_{in_i} \end{pmatrix}$$

$$\boldsymbol{Y}_{N\times p} = \boldsymbol{X}_{N\times I}\boldsymbol{\beta}_{I\times p} + \boldsymbol{\varepsilon}_{N\times p}$$

X : matriks rancangan

β : matriks parameter

ε : matriks galat

p : banyak peubah tak bebas/respon

I : banyak perlakuan (populasi)

n_i : banyak pengamatan pada perlakuan ke-*i*

N : banyak total pengamatan

Menurut Morisson (1991) analisis profil merupakan suatu bagian dari pengujian hipotesis terhadap nilai tengah dari peubah ganda (multivariate) dengan menggunakan prinsip grafik.

Tetapi hanya dengan melihat grafik saja tidaklah cukup, kita juga perlu untuk mengetahui seberapa besar arti kesejajaran (kemiripan) dari populasi itu. Untuk itulah diperlukan serangkaian uji-uji yang berkaitan dengan hipotesis itu.

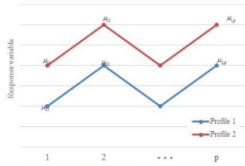


Figure 2: Illustration of a parallel profile analysis plot

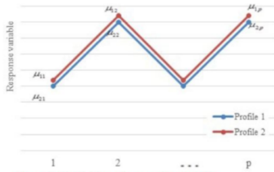


Figure 3: Illustration of a same level profile analysis plot.

Uji Hipotesis Kesejajaran (Parallel Test)

Hipotesis kesejajaran berkaitan dengan interaksi/pengaruh antar kelompok perlakuan. Jika sejajar (H0 diterima), maka interaksi/pengaruh antar perlakuan tidak ada.

$$H_{01}: \mu_{1i} - \mu_{1i-1} = \mu_{2i} - \mu_{2i-1},\ i = 2, 3, \ldots, p,$$

Atau bentuk umumnya: $\boxed{H_{01}: \mathbf{C}\boldsymbol{\mu}_1 = \mathbf{C}\boldsymbol{\mu}_2}$

Di mana **C** adalah matriks konstanta berikut:

$$\underset{((p-1)\times p)}{\mathbf{C}} = \begin{bmatrix} -1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix}$$

Uji Hipotesis Kesejajaran (Parallel Test) 2 Populasi

Jika kita mempunyai 2 populasi berukuran n_1 dan n_2, observasi pada masing-masing populasi dapat dibentuk dengan: $\mathbf{Cx}_{1j},\quad j = 1, 2, \ldots, n_1$

$\mathbf{Cx}_{2j},\quad j = 1, 2, \ldots, n_2$

Dan pengujiannya dapat dilakukan dengan:

$$T^2 = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)'\mathbf{C}'\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\mathbf{CS}_{\text{pooled}}\mathbf{C}'\right]^{-1}\mathbf{C}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) > c^2$$

S_{pooled} adalah matriks ragam-peragam gabungan dari kedua populasi

Di mana:

$$c^2 = \frac{(n_1 + n_2 - 2)(p-1)}{n_1 + n_2 - p} F_{p-1, n_1+n_2-p}(\alpha)$$

Tolak H01 ketika nilai $T^2 > c^2$

Uji Hipotesis Keberhimpitan (Coincident Test)

Jika profil dikatakan sejajar/*parallel* (H01 diterima), maka dapat diketahui bahwa $\mu_{1i} > \mu_{2i}$ untuk semua *i* atau sebaliknya. Dalam kondisi semacam ini, profil dikatakan **berhimpit** hanya jika total rataan $\mu_{11} + \mu_{12} + \cdots + \mu_{1p} = \mathbf{1}'\boldsymbol{\mu}_1$

dan

$$\mu_{21} + \mu_{22} + \cdots + \mu_{2p} = \mathbf{1}'\boldsymbol{\mu}_2$$

adalah sama. Artinya, pengaruh tiap perlakuan pada tiap kelompok sama.

Hipotesis nol pada pengujian tahap 2 ini dapat dituliskan dengan:

$$H_{02}: \mathbf{1}'\boldsymbol{\mu}_1 = \mathbf{1}'\boldsymbol{\mu}_2$$

Uji Hipotesis Keberhimpitan (Coincident Test) 2 Populasi

Jika kita mempunyai 2 populasi berukuran n_1 dan n_2, observasi pada masing-masing populasi dapat dibentuk dengan:

$$\mathbf{1}'\mathbf{x}_{1j},\ j = 1, 2, \ldots, n_1$$
$$\mathbf{1}'\mathbf{x}_{2j},\ j = 1, 2, \ldots, n_2$$

Pengujiannya dapat dilakukan dengan menghitung statistik uji berikut:

$$T^2 = \mathbf{1}'(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\mathbf{1}'\mathbf{S}_{\text{pooled}}\mathbf{1}\right]^{-1}\mathbf{1}'(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)$$

Tolak H02 jika statistik uji di atas **lebih besar** dari: $t^2_{n_1+n_2-2}\left(\frac{\alpha}{2}\right) = F_{1, n_1+n_2-2}(\alpha)$

Uji Hipotesis Kesamaan (Level Test)

Jika profil-profil dikatakan berhimpit (H02 diterima), apakah semua observasi berasal dari populasi normal yang sama? Maka dari itu ingin diuji apakah semua peubah memiliki rataan yang sama, sehingga profil dikatakan setara. Uji ini dilakukan berdasarkan hipotesis berikut:

$$H_{03}: \mathbf{C}\boldsymbol{\mu} = \mathbf{0}$$

Di mana **C** merupakan matriks konstanta sebelumnya sebagai berikut:

$$\underset{((p-1)\times p)}{\mathbf{C}} = \begin{bmatrix} -1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix}$$

Uji Hipotesis Keberhimpitan (Coincident Test) 2 Populasi

Jika kesejajaran dan keberhimpitan diterima, maka vektor rataan dari dua populasi yang berukuran n1 dan n2 dapat diduga dengan:

$$\bar{\mathbf{x}} = \frac{1}{n_1 + n_2}\left(\sum_{j=1}^{n_1}\mathbf{x}_{1j} + \sum_{j=1}^{n_2}\mathbf{x}_{2j}\right) = \frac{n_1}{(n_1 + n_2)}\bar{\mathbf{x}}_1 + \frac{n_2}{(n_1 + n_2)}\bar{\mathbf{x}}_2$$

Tolak H03 jika memenuhi ketentuan sebagai berikut:

$$(n_1 + n_2)\bar{\mathbf{x}}'\mathbf{C}'[\mathbf{CSC}']^{-1}\mathbf{C}\bar{\mathbf{x}} > c^2$$

Dimana **S** adalah matriks ragam-peragam contoh berdasarkan semua observasi (sebanyak n1+n2), dan

$$c^2 = \frac{(n_1 + n_2 - 1)(p-1)}{(n_1 + n_2 - p + 1)} F_{p-1, n_1+n_2-p+1}(\alpha)$$

Alur Pengerjaan Analisis Profil

1. Eksplorasi data menggunakan grafik
2. Uji Kesejajaran
   - Jika Uji Kesejajaran diterima, maka dilanjutkan dengan Uji Keberhimpitan
   - Jika Uji Keberhimpitan diterima, maka dilanjutkan dengan Uji Kesamaan

Pertemuan 7 Analisis Komponen Utama (AKU) / Principal Component Analysis (PCA)

Analisis Komponen Utama

- AKU/PCA adalah sebuah teknik untuk menyederhanakan suatu data yang berdimensi besar dan saling berkorelasi menjadi dimensi yang lebih kecil dan saling bebas dengan cara mentransformasi linier peubah-peubah yang diamati membentuk beberapa peubah baru yang dikenal dengan komponen utama (principal component).
- Komponen utama dibentuk berdasarkan matriks ragam-peragam atau matriks korelasi.

Misalkan vektor peubah acak $\mathbf{X}' = [X_1, X_2, \ldots, X_p]$ memiliki matriks var-cov $\boldsymbol{\Sigma}$ dan akar-akar ciri $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ kemudian dilakukan kombinasi linier berikut:

$$Y_1 = \mathbf{a}_1'\mathbf{X} = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p$$
$$Y_2 = \mathbf{a}_2'\mathbf{X} = a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p$$
$$\vdots$$
$$Y_p = \mathbf{a}_p'\mathbf{X} = a_{p1}X_1 + a_{p2}X_2 + \cdots + a_{pp}X_p$$

Kemudian diperoleh:

$$\mathrm{Var}(Y_i) = \mathbf{a}_i'\boldsymbol{\Sigma}\mathbf{a}_i \qquad i = 1, 2, \ldots, p$$
$$\mathrm{Cov}(Y_i, Y_k) = \mathbf{a}_i'\boldsymbol{\Sigma}\mathbf{a}_k \qquad i, k = 1, 2, \ldots, p$$

- Komponen Utama / Principal Component (PC) dari X adalah kombinasi linier yang saling bebas dari yang memiliki ragam (dari matriks var-cov sebelumnya) yang bernilai sebesar-besarnya.

Manfaat AKU/PC

- **Eksplorasi posisi objek dan penangan masalah kolinear antar peubah**. Eksplorasi posisi objek diperlukan sebagai alat bantu dalam analisis gerombol. Eksplorasi ini dapat dilakukan dengan membuat plot skor komponen utama pertama dengan kedua.
- **Komponen utama merupakan salah satu solusi dalam mengatasi masalah kolinear**. Penerapan ini relevan dengan sifat dari komponen utama yang dibangun yaitu antar komponen utama bersifat saling orthogonal atau saling bebas.

Pembentukan PC dengan Matriks Ragam-Peragam

Misalkan $\mathbf{\Sigma}$ merupakan matriks ragam peragam dari $\mathbf{X}' = [X_1, X_2, \ldots, X_p]$ yang memiliki pasangan akar dan vektor ciri $(\lambda_1, \mathbf{e}_1), (\lambda_2, \mathbf{e}_2), \ldots, (\lambda_p, \mathbf{e}_p)$na $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_p \ge 0$ponen utama ke-*i* dibentuk dengan:

$$\boxed{Y_i = \mathbf{e}_i'\mathbf{X} = e_{i1}X_1 + e_{i2}X_2 + \cdots + e_{ip}X_p, \qquad i = 1, 2, \ldots, p}$$

$$\boxed{\begin{aligned} \mathrm{Var}(Y_i) &= \mathbf{e}_i'\mathbf{\Sigma}\mathbf{e}_i = \lambda_i \qquad & i = 1, 2, \ldots, p \\ \mathrm{Cov}(Y_i, Y_k) &= \mathbf{e}_i'\mathbf{\Sigma}\mathbf{e}_k = 0 \qquad & i \neq k \end{aligned}}$$

Dengan ketentuan bahwa pembentukkan komponen utama berdasarkan matriks ragam peragam dapat dilakukan jika **satuan pengukuran setiap peubah sama.**

Total keragaman populasi diperoleh dengan:

$$\sum_{i=1}^{p} \mathrm{Var}(X_i) = \mathrm{tr}(\mathbf{\Sigma}) = \mathrm{tr}(\mathbf{\Lambda}) = \sum_{i=1}^{p} \mathrm{Var}(Y_i)$$

$$\begin{aligned} \text{Total population variance} &= \sigma_{11} + \sigma_{22} + \cdots + \sigma_{pp} \\ &= \lambda_1 + \lambda_2 + \cdots + \lambda_p \end{aligned}$$

Sedangkan proporsi keragaman kontribusi komponen utama ke-*i* diperoleh dengan:

$$\begin{pmatrix} \text{Proportion of total} \\ \text{population variance} \\ \text{due to } k\text{th principal} \\ \text{component} \end{pmatrix} = \frac{\lambda_k}{\lambda_1 + \lambda_2 + \cdots + \lambda_p} \qquad k = 1, 2, \ldots, p$$

Pembentukan PC dengan Matriks Korelasi

Misalkan $\mathbf{\Sigma}$ merupakan matriks ragam peragam dari $\mathbf{Z}' = [Z_1, Z_2, \ldots, Z_p]$ dengan $\mathrm{Cov}(\mathbf{Z}) = \boldsymbol{\rho}$, dan memiliki pasangan akar-vektor ciri $(\lambda_1, \mathbf{e}_1), (\lambda_2, \mathbf{e}_2), \ldots, (\lambda_p, \mathbf{e}_p)$ dengan $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_p \ge 0$

maka komponen utama ke-*i* diberikan:

$$\boxed{Y_i = \mathbf{e}_i'\mathbf{Z} = \mathbf{e}_i'(\mathbf{V}^{1/2})^{-1}(\mathbf{X} - \boldsymbol{\mu}), \qquad i = 1, 2, \ldots, p}$$

Dengan ketentuan bahwa pembentukkan komponen utama berdasarkan matriks korelasi dapat dilakukan jika **satuan pengukuran setiap peubah berbeda**. Standardiasi dilakukan dengan:

$$\boxed{Z_1 = \frac{(X_1 - \mu_1)}{\sqrt{\sigma_{11}}}} \quad \boxed{Z_2 = \frac{(X_2 - \mu_2)}{\sqrt{\sigma_{22}}}} \quad \cdots \quad \boxed{Z_p = \frac{(X_p - \mu_p)}{\sqrt{\sigma_{pp}}}}$$

Total keragaman populasi diperoleh dengan:

$$\boxed{\sum_{i=1}^{p} \mathrm{Var}(Y_i) = \sum_{i=1}^{p} \mathrm{Var}(Z_i) = p}$$

Sedangkan proporsi keragaman kontribusi komponen utama ke-*i* diperoleh dengan:

$$\begin{pmatrix} \text{Proportion of (standardized)} \\ \text{population variance due} \\ \text{to } k\text{th principal component} \end{pmatrix} = \frac{\lambda_k}{p}, \qquad k = 1, 2, \ldots, p$$

Tahap Pengerjaan AKU/PC

- Sediakan set data X
- Membuat matriks varian kovarian atau matriks korelasi dari data X
- Menghitung akar ciri dan vektor ciri dari matriks data yang ditentukan poin 2
- Menentukan proporsi keragaman mengunakan akar ciri yang diperoleh pada poin 3
- Menghitung komponen utama dari vektor ciri yang diperoleh di poin 3 sekaligus menentukan banyaknya komponen yang digunakan melalui proporsi keragaman yang telah diperoleh pada poin 4
- Menghitung skor komponen utama sebanyak komponen utama yang telah ditentukan pada poin 5

Penentuan Banyak Komponen Utama

<u>Metode 1</u>

- Didasarkan pada kumulatif proporsi keragaman total yang mampu dijelaskan
- Dapat diterapkan pada matriks korelasi maupun ragam peragam
- Minimum persentase keragaman yang mampu dijelaskan ditentukan terlebih dahulu, dan selanjutnya banyaknya komponen yang paling kecil hingga batas itu terpenuhi dijadikan sebagai banyaknya komponen utama yang digunakan.
- Tidak ada patokan baku berapa batas minimum tersebut, sebagian buku menyebutkan 70%, 80%, bahkan ada yang 90%.

<u>Metode 2</u>

- Hanya dapat diterapkan pada penggunaan matriks korelasi. Saat menggunakan matriks ini, peubah asal ditransformasi menjadi peubah yang memiliki ragam sama yaitu satu.
- Didasarkan pada ragam komponen utama, yang tidak lain adalah akar ciri. Metode ini disarankan oleh Kaiser (1960) yang berargumen bahwa jika peubah asal saling bebas maka komponen utama tidak lain adalah peubah asal, dan setiap komponen utama akan memiliki ragam satu.
- Dengan cara ini, komponen yang berpadanan dengan akar ciri kurang dari satu tidak digunakan. Jollife (1972) setelah melakukan studi mengatakan bahwa cut off yang lebih baik adalah 0.7.

<u>Metode 3</u>

- Metode ini menggunakan grafik yang disebut plot scree.
- Cara ini dapat digunakan ketika titik awalnya matriks korelasi maupun ragam peragam.
- Plot scree merupakan plot antara akar ciri λk dengan k
- Dengan menggunakan metode ini, banyaknya komponen utama yang dipilih, yaitu k, adalah jika pada titik k tersebut plotnya curam ke kiri tapi tidak curam di kanan. Ide yang ada di belakang metode ini adalah bahwa banyaknya komponen utama yang dipilih sedemikian rupa sehingga selisih antara akar ciri yang berurutan sudah tidak besar lagi. Interpretasi terhadap plot ini sangat subjektif.