



# ANALISIS PEUBAH GANDA (MULTIVARIATE ANALYSIS)

---

## Pengantar Analisis Peubah Ganda

Dr.Ir. I Made Sumertajaya, MSi  
Departemen Statistika-FMIPA IPB  
Email : kuliah\_apg@yahoo.com  
Password: pakmade

# Materi APG

No	Materi
I	Pendahuluan (Introduction)
II	Sebaran Normal Ganda (The Multivariate Normal Distribution)
III	Inferensia Vektor Rataan (Inferences about a Mean Vector)
IV	Perbandingan Dua Vektor Rataan (Comparison of two mean vectors)
V	MANOVA
VI	Analisis Profil (Profile Analysis)
VII	UTS
VIII	Analisis Komponen Utama (Principal Component Analysis)
IX	Analisis Faktor (Factor Analysis)
X	Analisis Korelasi Kanonik (Canonical Correlation Analysis)
XI	Analisis Gerombol (Cluster Analysis)
XII	Analisis Diskriminan (Discriminant Analysis)
XIII	Analisis Biplot (Biplot Analysis)
XIV	Analisis Korespondensi (Correspondency Analysis)
XV	Analisis Regresi Multivariate (Path analysis dan SEM)
XVI	UAS

# Pengamatan Peubah Ganda

---

- Menggambarkan suatu objek tidak cukup menggunakan satu peubah saja
- Kasus pengamatan peubah ganda dijumpai di seluruh bidang terapan
- Perlu analisis lebih canggih, jika antar peubah tidak saling bebas

# Beberapa Notasi

- Misal  $p$  peubah diamati dari  $n$  objek
- Skala pengukuran peubah bisa nominal, ordinal, interval, atau rasio
- Dapat berupa peubah bebas maupun peubah tak bebas

Obs	X1	X2	...	X <sub>p</sub>
1	x <sub>11</sub>	x <sub>21</sub>	...	x <sub>p1</sub>
2	x <sub>12</sub>	x <sub>22</sub>	...	x <sub>p2</sub>
3	x <sub>13</sub>	x <sub>23</sub>	...	x <sub>p3</sub>
4	x <sub>14</sub>	x <sub>24</sub>	...	x <sub>p4</sub>
...	...	...	...	...
n	x <sub>1n</sub>	x <sub>2n</sub>	...	x <sub>pn</sub>

# Notasi berikut khusus untuk peubah-peubah berskala interval atau rasio

---

- Vektor peubah acak:

$${}_p\underline{X}_1 = \begin{bmatrix} X_1 \\ X_2 \\ \dots \\ X_p \end{bmatrix}$$

- Nilai harapan vektor peubah acak

$$E({}_p\underline{X}_1) = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \dots \\ E(X_p) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \dots \\ \mu_p \end{bmatrix} = \underline{\mu}$$

## Matriks Ragam-peragam (Variance Covariance Matrix)

$$Cov(\underline{X}) = \sum_p = \begin{bmatrix} \text{var}(x_1) & \text{cov}(x_1, x_2) & \cdots & \text{cov}(x_1, x_p) \\ \text{cov}(x_2, x_1) & \text{var}(x_2) & \cdots & \text{cov}(x_2, x_p) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(x_p, x_1) & \text{cov}(x_p, x_2) & \cdots & \text{var}(x_p) \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}$$

dimana:

$$Cov(X_i, X_i) = Var(X_i) = \sigma_{ii} = \sigma_i^2 = E(X_i - E(X_i))^2$$

$$Cov(X_i, X_j) = \sigma_{ij} = E(X_i - E(X_i))(X_j - E(X_j))$$

- Matriks korelasi berukuran  $p \times p$

$${}_p R_p = \begin{bmatrix} 1 & \rho_{12} & \dots & \rho_{1p} \\ \rho_{21} & 1 & & \rho_{2p} \\ & & \dots & \\ \rho_{p1} & \rho_{p2} & & 1 \end{bmatrix}$$

- Hubungan matriks ragam peragam dengan matriks korelasi

$${}_p R_p = D \left( \frac{1}{\sigma_i} \right) \Sigma D \left( \frac{1}{\sigma_i} \right)$$

- Matriks D, matriks diagonal berukuran  $p \times p$

$$D \left( \frac{1}{\sigma_i} \right) = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & \dots & 0 \\ 0 & \frac{1}{\sigma_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{1}{\sigma_p} \end{bmatrix}$$

# Beberapa notasi untuk data sampel

- Vektor rata-rata berukuran  $p \times 1$ , merupakan penduga bagi vektor  $\underline{\mu}$

$${}_p \bar{x}_1 = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \dots \\ \bar{x}_p \end{bmatrix} = \begin{bmatrix} \frac{\sum_{i=1}^n x_{1i}}{n} \\ \frac{\sum_{i=1}^n x_{2i}}{n} \\ \dots \\ \frac{\sum_{i=1}^n x_{pi}}{n} \end{bmatrix} = \underline{\hat{\mu}}$$

- Matriks ragam peragam berukuran  $p \times p$ , merupakan penduga bagi matriks  $\Sigma$

$${}_p S_p = \begin{bmatrix} s_{11} & s_{12} & \dots & s_{1p} \\ s_{21} & s_{22} & \dots & s_{2p} \\ \dots & \dots & \dots & \dots \\ s_{p1} & s_{p2} & \dots & s_{pp} \end{bmatrix} = \hat{\Sigma}$$

$$s_{ii} = s_i^2 = \frac{\sum_{j=1}^n (x_{ij} - \bar{x}_i)^2}{n-1}$$

$$s_{ij} = \frac{\sum_{k=1}^n (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)}{n-1}$$



- Matriks korelasi berukuran  $p \times p$ , merupakan penduga bagi matriks  $\rho$

$${}_p R_p = \begin{bmatrix} 1 & r_{12} & \dots & r_{1p} \\ r_{21} & 1 & \dots & r_{2p} \\ \dots & \dots & \dots & \dots \\ r_{p1} & r_{p2} & \dots & 1 \end{bmatrix} = \hat{\rho}$$

$$r_{ii} = 1$$

$$r_{ij} = \frac{\sum_{k=1}^n (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ik} - \bar{x}_i)^2 \sum_{k=1}^n (x_{jk} - \bar{x}_j)^2}}$$

# Konsep Jarak

---

- Jarak pengamatan ke titik pusat
- Jarak antar pengamatan (i,j)
  - Jarak Euclidean
  - Jarak Mahalanobis
  - Jarak Minkowski (City Block)

$$d_i = \sqrt{(\underline{x}_i - \bar{x})'(\underline{x}_i - \bar{x})}$$

$$d_{i,j} = \sqrt{(\underline{x}_i - \underline{x}_j)'(\underline{x}_i - \underline{x}_j)}$$

$$d_{i,j} = \sqrt{(\underline{x}_i - \underline{x}_j)'S^{-1}(\underline{x}_i - \underline{x}_j)}$$

$$d_{i,j} = \sqrt[k]{(\underline{x}_i - \underline{x}_j)'(\underline{x}_i - \underline{x}_j)}, k = 2,3,4,\dots$$



# Ringkasan Teknik Analisis Peubah Ganda

---

# MANOVA

---

- MANOVA = Multivariate Analysis of Variance
- Analog dengan ANOVA (analysis of variance) pada analisis peubah tunggal
- Digunakan untuk membandingkan rata-rata populasi (misal membandingkan rata-rata perlakuan dalam percobaan)

# MANOVA

---

- Mengasumsikan data berasal dari populasi multi-normal
- Model yang ada bervariasi dari Satu Arah (one-way) hingga model Banyak Arah (multi-way) termasuk adanya interaksi

# MANOVA

---

## Ilustrasi

Sebuah studi ingin membandingkan apakah ada perbedaan hasil panen tiga buah varietas padi. Respon yang diamati adalah tinggi tanaman ketika panen, bobot 100 butir padi, umur panen, panjang galur.

# Analisis Komponen Utama

---

- Digunakan untuk melakukan reduksi peubah
- Menghasilkan peubah baru yang disebut KOMPONEN UTAMA dengan karakteristik : (1) informasi yang terkandung memuat hampir seluruh informasi (2) saling bebas
- Umumnya merupakan analisis antara, bukan analisis akhir

# Analisis Faktor

---

- Digunakan juga untuk mereduksi peubah
- Menghasilkan peubah baru yang disebut FAKTOR atau PEUBAH LATEN
- Banyak digunakan di bidang terapan sosial karena sulit melakukan pengukuran secara langsung terhadap peubah yang diinginkan



# Analisis Korelasi Kanonik

---

- Melakukan analisis keterkaitan antara dua gugus peubah
- Analisis korelasi tidak dilakukan antar pasangan peubah asal dari kedua gugus, namun antar peubah kanonik di kedua gugus

# Analisis Korelasi Kanonik

---

## Ilustrasi

Seorang konsultan olah gerak harus mampu melihat korelasi antara peubah anatomi tubuh manusia (lingkar lengan, berat badan, lingkar dada, panjang lengan, panjang kaki, tinggi badan, dsb) dengan berbagai aktifitas aerobik (sit-up, push-up, pull-up, scotch-jump, dsb)

# Analisis Gerombol

---

- Digunakan untuk mengelompokkan objek-objek
- Kemiripan antar objek ditentukan oleh nilai-nilai pengamatan peubah ganda
- Dikenal teknik berhirarki dan tak-berhirarki

# Analisis Gerombol

---

## Ilustrasi

Pemberian program bantuan pengembangan sekolah tidak bisa disamakan di setiap sekolah. Berdasarkan karakteristik sekolah (fasilitas, input siswa, mutu pengajar, dukungan masyarakat), DEPDIKNAS melakukan pengelompokan sekolah sehingga diperoleh 4 kelompok SMU dan 4 jenis paket program pengembangan SMU.

# Analisis Biplot

---

- Merupakan analisis eksplorasi untuk melihat (1) kedekatan antar objek (2) karakteristik atau peubah penciri setiap objek, dan (3) keterkaitan antar peubah
- Menggunakan konsep penguraian nilai singular
- Bermula di dunia pertanian, sekarang lebih banyak dipakai di riset pemasaran

# Analisis Biplot

---

## Ilustrasi

Perencanaan layanan perbankan memerlukan informasi mengenai posisi bank kita dibandingkan bank-bank pesaing. Untuk itulah dilakukan survei pasar mengenai persepsi nasabah mengenai berbagai atribut layanan perbankan, sehingga bisa kita ketahui atribut mana saja yang perlu ditingkatkan dan siapa pesaing terdekat yang harus diantisipasi.

# Analisis Korespondensi

---

- Teknik eksplorasi yang diterapkan untuk melihat asosiasi kategori peubah kategorik
- Menggunakan konsep Generalized Singular Value Decomposition
- Banyak digunakan di dunia riset pemasaran

# Analisis Korespondensi

---

## Ilustrasi

Segmentasi produk bisa dilakukan berdasarkan peubah demografi konsumen. Melalui analisis ini bisa diketahui konsumen utama produk kita dari kelompok usia berapa, jenis kelamin apa, tinggal dimana, dan sebagainya.



# Analisis Diskriminan

---

- Menghasilkan fungsi yang digunakan untuk mengklasifikasikan objek tertentu berasal dari populasi yang mana
- Linear Discriminant, Quadratic Discriminat, Canonical Discriminant, Logistic Regression, Non-Parametric Discriminant

# Analisis Diskriminan

---

## Ilustrasi

Perusahaan penyedia jasa layanan kartu kredit harus mampu membuat fungsi diskriminan yang mampu memisahkan calon pemegang kartu yang potensial melakukan transaksi dan yang tidak (idle) berdasarkan data dalam formulir aplikasi.

# Penskalaan Dimensi Ganda – Multi Dimensional Scaling

---

- Menghasilkan peta atau gambar posisi objek berdasarkan matriks jarak yang diketahui
- Peta dihasilkan pada dimensi rendah (umumnya dimensi dua) sehingga mudah menginterpretasikan kedekatan antar objek
- Metric MDS vs Non-Metric MDS, tergantung tipe peubahnya