

# ★ ANALISIS KORELASI KANONIK



NOVI HIDAYAT PUSPONEGORO

# Pendahuluan (1)

Analisis Korelasi Kanonik, bertujuan untuk

- Mengetahui dan mengukur hubungan antara suatu set variabel bebas dan satu set variabel tak bebas
- Menyederhanakan hubungan antara dua set variabel menjadi pasangan-pasangan variabel kanonik.

Korelasi Kanonik fokus pada korelasi antara kombinasi linier dalam satu set variabel dengan kombinasi linear variabel set berikutnya, sehingga korelasi antara dua set variabel tersebut maksimum.

Hal ini bertujuan untuk menentukan pasangan kombinasi linier dengan korelasi terbesar. Selanjutnya, kita menentukan pasangan yang memiliki kombinasi linier korelasi terbesar di antara semua pasangan yang pada awalnya tidak berkorelasi, dan seterusnya.

Pasangan kombinasi linier variabel disebut variabel kanonik, dan korelasinya disebut korelasi kanonik.

# Pendahuluan (2)

## *Contoh Penggunaan Analisis Korelasi Kanonik:*

Misal terdapat satu set variabel bebas, yang terdiri dari variabel besar promosi produk, harga produk, jumlah outlet yang menjual produk.

Dan terdapat satu set variabel tak bebas, yang terdiri dari variabel penjualan produk dan biaya produksi.

Maka analisis korelasi kanonik dapat digunakan untuk meneliti hubungan penjualan dan biaya produksi produk dengan promosi, harga dan jumlah outlet

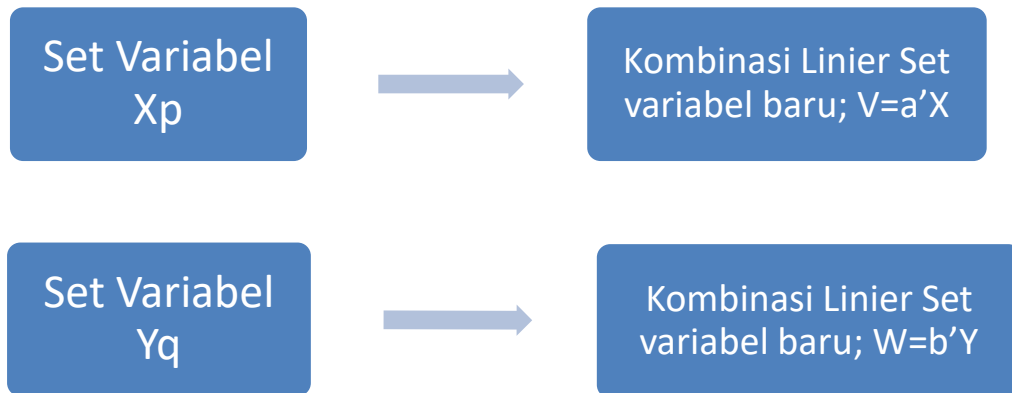
# Pendahuluan (3)

Data yang digunakan adalah jenis data metrik

Asumsi:

- Variabel dalam set variabel bebas dan set variabel tak bebas, saling berhubungan linier
- Tidak ada hubungan antar variabel dalam set yang sama (tidak terjadi multikolinieritas)
- Jika diperlukan, sifat selanjutnya adalah memenuhi sebaran multivariat normal

# Ilustrasi



Dimana:

$X$  merupakan vektor set variabel dengan ukuran  $p \times 1$

$Y$  merupakan vektor set variabel dengan ukuran  $q \times 1$ , dengan asumsi  $p \leq q$

$V$  dan  $W$  disebut dengan Variabel kanonik

$a$  dan  $b$ , merupakan vektor koefisien kanonik

Misalkan untuk kedua set variabel X dan Y, diketahui:

- $E(X)=\mu_X$                        $E(Y)=\mu_Y$
- $\text{Cov}(X) = \Sigma_{XX}$      $\text{Cov}(Y) = \Sigma_{YY}$
- $\text{Cov}(X, Y) = \Sigma_{XY} = \Sigma_{YX}$

Dengan susunan matrik kovariansinya adalah sbb:

$$\Sigma_{(p+q) \times (p+q)} = \begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix}$$

### Contoh (2)

Sedangkan:  $Var(V) = a' \Sigma_{YY} a$

$$Var(W) = b' \Sigma_{YY} b$$

$$Cov(V, W) = a' \Sigma_{XY} b$$

$$Corr(V, W) = \frac{Cov(V, W)}{\sqrt{Var(V)} \sqrt{Var(W)}}$$



Sedangkan untuk data sampel:

$$Corr(V, W) = \frac{Cov(V, W)}{\sqrt{Var(V)} \sqrt{Var(W)}} = \frac{a' S_{XY} b}{\sqrt{a' S_{XX} a} \sqrt{b' S_{YY} b}}$$

Sifat-sifat Variabel Kanonik:

$$Var(V_i) = Var(W_i) = 1$$

$$Cov(V_l, V_k) = Cor(V_l, V_k) = 0$$

$$Cov(W_l, W_k) = Cor(W_l, W_k) = 0$$

$$Cov(V_l, W_k) = Cor(V_l, W_k) = 0$$

$$Corr(V_1, W_1) > Corr(V_2, W_2) > Corr(V_3, W_3) > \dots > Corr(V_r, W_r)$$





## PROSES ANALISIS KORELASI KANONIK:

1. Menentukan variabel-variabel yang akan masuk dalam set variabel bebas dan tak bebas
2. Membentuk beberapa Fungsi Kanonik, dan memilih fungsi yang akan digunakan dengan Uji signifikan, Canonical relationship, serta Redudancy Index
3. Dari fungsi kanonik terpilih, dilakukan interpretasi hasil dengan menggunakan beberapa metode, yaitu:

Canonical Weight

Canonical Loadings atau Cross Canonical Loadings





## Persamaan ciri yang dibuat untuk mendapatkan akar ciri (2)

Sehingga:

$$V = a'X = e' \sum_{XX}^{-1/2} X$$

$$W = b'Y = f' \sum_{YY}^{-1/2} Y$$

Dengan banyaknya pasangan kanonik yang terbentuk adalah:

- Jika banyak variabel dalam set variabel bebas (p) < banyak variabel dalam set variabel tak bebas (q), maka banyaknya pasangan kanonik adalah p
- Jika banyak variabel dalam set variabel bebas (p) > banyak variabel dalam set variabel tak bebas (q), maka banyaknya pasangan kanonik adalah q

## Persamaan ciri yang dibuat untuk mendapatkan akar ciri

Dimana akar ciri  $\lambda^2$  yang diperoleh merupakan koefisien korelasi kanonik yang dikuadratkan.

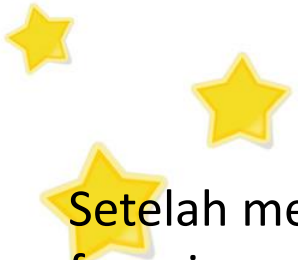
Untuk set Variabel X, persamaan cirinya adalah:

$$\left| \sum_{XX}^{-1/2} \sum_{XY} \sum_{YY}^{-1} \sum_{YX} \sum_{XX}^{-1/2} - \lambda^2 I \right| e = 0$$

Untuk set Variabel Y, persamaan cirinya adalah:

$$\left| \sum_{YY}^{-1/2} \sum_{YX} \sum_{XX}^{-1} \sum_{XY} \sum_{YY}^{-1/2} - \lambda^2 I \right| f = 0$$

[Contoh \(3\)](#)



Setelah mendapat beberapa Fungsi Kanonik, selanjutnya untuk memilih fungsi yang akan digunakan dengan Uji signifikan, Canonical relationship, serta Redudancy Index

Uji Hipotesis:  $H_0 = \sum_{yx} = 0$   
 $H_1 = \sum_{yx} \neq 0$

Dengan statistik uji:

$$\chi^2 = - \left[ (n-1) - \frac{1}{2}(p+q+1) \right] \ln \Lambda$$
$$\Lambda = \prod_{j=1}^{\min(p,q)} (1 - \lambda_j^2)$$

Tolak Hipotesis nol, jika  $\chi^2$  hitung lebih besar daripada  $\chi^2$  tabel dengan derajat bebas pq pada tingkat signifikansi tertentu.





## Canonical relationship

Untuk mengetahui apakah fungsi yang dibentuk sudah baik atau belum bisa dilihat dari nilai Canonical relationship .

Apabila Canonical relationship mendekati 1, berarti fungsi kanonik yang terbentuk sudah baik


## Redudancy Index

Merupakan ukuran (dalam persentase) untuk melihat keragaman yang dapat dijelaskan oleh variabel kanonik Y ke-j pada set variabel X

$$R^2_{(j)Y|X} = \lambda_j^2 \sum \frac{r_{Y^*Y_i}^j}{q} = \lambda_j^2 R^2_{(j)Y}$$

Apabila persentase keragaman yang dapat dijelaskan mendekati 100%, maka fungsi kanonik yang terpilih sudah baik (analog seperti Koefisien determinasi pada Analisis Regresi)





Dari fungsi kanonik terpilih, dilakukan interpretasi hasil dengan menggunakan beberapa metode, yaitu:

### Canonical Weight

Merupakan ukuran untuk mengetahui variabel dari set X dan set Y yang mendominasi variabel kanonik ke-j

### Canonical Loadings

Untuk mengukur korelasi (sederhana) antara variabel asal dengan variabel kanonik, untuk masing-masing set:

Set X: 
$$r^{(j)}_{X^*X} = \sum_{XX} a^{(j)}$$

Set Y: 
$$r^{(j)}_{Y^*Y} = \sum_{YY} b^{(j)}$$



# Contoh:

Data dikemukakan oleh Green dan Tull ( 1978. p.411) tentang efektivitas iklan untuk ban radial.

Dua kelompok variabel terdiri dari :

Variabel Kriteria :

Y 1 : believability of the claims made in a TV commercial

Y2 : postexposure interest in the new brand were measured

Variabel prediktor :

X1 : general interest in the product

X2 : previous choice of tire brand

X3 : preexposure interest in the new brand



# Contoh (2)

Table 9.3-3. Pearson Product-Moment Correlation of Variables from Advertising-Effectiveness Study\*

|       | $Y_1$ | $Y_2$ |          |          | $X_1$ | $X_2$ | $X_3$ |
|-------|-------|-------|----------|----------|-------|-------|-------|
| $Y_1$ | 1.000 | .659  |          |          | .202  | .097  | .321  |
| $Y_2$ | .659  | 1.000 | $R_{yy}$ | $R_{yx}$ | .315  | .218  | .539  |
| $X_1$ | .202  | .315  | $R_{xy}$ | $R_{xx}$ | 1.000 | .086  | .226  |
| $X_2$ | .097  | .218  |          |          | .086  | 1.000 | .258  |
| $X_3$ | .321  | .539  |          |          | .226  | .258  | 1.000 |

\*Source: Green and Tull (1978, p. 411)



# Contoh (3)

Table 9.3-4. Canonical Analysis of Data in Table 9.3-3

|  | Variate<br>Number | Eigenvalue | Canonical<br>Correlation |
|--|-------------------|------------|--------------------------|
|  | 1                 | .337       | .581                     |
|  | 2                 | .002       | .049                     |

| Weights for the Y-set variables |                   |       | Loadings for the Y-set variables |                   |      |
|---------------------------------|-------------------|-------|----------------------------------|-------------------|------|
|                                 | Canonical Variate |       |                                  | Canonical Variate |      |
|                                 | 1                 | 2     |                                  | 1                 | 2    |
| $Y_1$                           | -.110             | 1.324 | $Y_1$                            | .594              | .804 |
| $Y_2$                           | 1.069             | -.789 | $Y_2$                            | .997              | .087 |
|                                 |                   |       | Percent of<br>variance           | 67.3              | 32.7 |

| Weights for the X-set variables |                   |       | Loadings for the X-set variables |                   |       |
|---------------------------------|-------------------|-------|----------------------------------|-------------------|-------|
|                                 | Canonical Variate |       |                                  | Canonical Variate |       |
|                                 | 1                 | 2     |                                  | 1                 | 2     |
| $X_1$                           | .344              | .430  | $X_1$                            | .540              | .380  |
| $X_2$                           | .142              | -.952 | $X_2$                            | .382              | -.879 |
| $X_3$                           | .815              | .141  | $X_3$                            | .929              | -.007 |
|                                 |                   |       | Percent of<br>variance           | 43.4              | 30.6  |





# Contoh (4)

Table 9.3-5. Components of Redundancy: Calculations for the Data of Example 9.3-2

|       | (1)<br>Variate<br>$i$ | (2)<br>Canonical<br>Correlation<br>$\sqrt{\lambda_{(i)}}$ | (3)<br>Squared<br>Canonical<br>Correlation<br>$\lambda_{(i)}$ | (4)                                 |                                     | (5)<br>[(4)/<br>No. of<br>Variables] | (6)<br>[(3) × (5)] |
|-------|-----------------------|---|---|-------------------------------------|-------------------------------------|--------------------------------------|--------------------|
|       |                       |   |   | $\sum_{i=1}^p (r_{y^*y_i}^{(j)})^2$ | $\sum_{i=1}^m (r_{x^*x_i}^{(j)})^2$ |                                      |                    |
| Y-set | 1                     | .581  | .337  | 1.346                               |                                     | .673                                 | .227               |
|       | 2                     | .049  | .002  | 0.654                               |                                     | .327                                 | .001               |
|       |                       |   |   |                                     |                                     |                                      | .228               |
| X-set | 1                     | .581  | .337  |                                     | 1.302                               | .434                                 | .146               |
|       | 2                     | .049  | .002  |                                     | 0.918                               | .306                                 | .001               |
|       |                       |   |   |                                     |                                     |                                      | .147               |



Referensi:

Johnson. R, Wichern. D, *Applied Multivariate Statistical Analysis*.  
Pearson Prentice Hall, 2007

Sharma. S, *Applied Multivariate Techniques*. John Willey and Sons,  
1996.