

ANALYSIS OF VARIANCE (ANOVA)

Prof. Dr. Dyah Erny Herwindiati, S.Si, M.Si.

Lely Hiryanto, S.T., M.Sc., Ph.D.

Anova

- Anova adalah satu metode analisis statistika yang digolongkan ke dalam kelompok statistik inferensial.
 - Anova Satu Arah: satu independent variable
 - Anova Dua Arah: dua independent variables
- **Anova Satu Arah (One Way Anova):** digunakan untuk melakukan pengujian perbandingan rata-rata beberapa kelompok, biasanya terdiri dari lebih dari dua kelompok.
- Penggunaan Anova kelompok yang berasal dari sampel yang berbeda antar kelompok.

Hipotesis Anova Satu Arah

- Hipotesis yang digunakan dalam Anova satu arah adalah sebagai berikut:
 - Hipotesis nol (H_0) dalam uji ANOVA adalah bahwa semua (minimal 3) populasi yang sedang dikaji memiliki rata-rata hitung (*mean*) sama
 - $H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_n$, Tidak terdapat perbedaan signifikan antara rata-rata hitung dari n kelompok.
 - $H_1: \mu_1 \neq \mu_2 \neq \mu_3 \neq \dots \neq \mu_n$, Ada perbedaan yang signifikan antara rata-rata hitung dari n kelompok
- Hipotesis yang digunakan untuk Anova satu arah dan Anova dua arah adalah sama.
- Analisis ragam Anova tidak dapat menentukan mana kelompok yang benar-benar berbeda.
- Kemampuan analisis ragam Anova hanya mampu mendeteksi apakah ada perbedaan rata-rata dari beberapa kelompok tersebut.

Anova Satu Arah

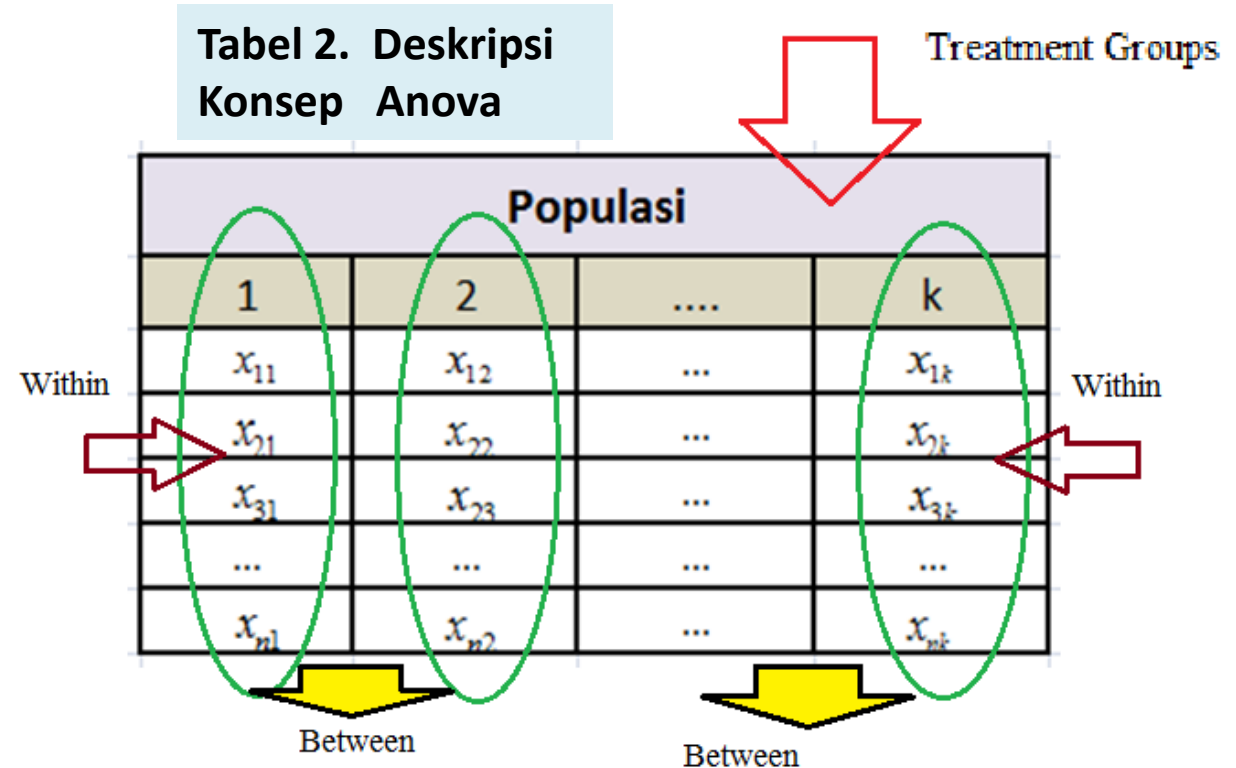
- **Tujuan:** untuk membandingkan nilai mean (rata-rata) atau varians dari beberapa populasi (yang biasanya lebih dari 2 populasi).
- **Contoh:** jika terdapat k populasi , $k = 1, 2, \dots, k$ dan di setiap populasi terdapat sejumlah n sampel seperti pada tabel 1 berikut. Kita ingin membandingkan apakah sekurang-kurangnya ada satu perbedaan rata-rata atau varians dari k populasi tersebut.

Tabel 1. Contoh Konsep Dasar Anova

| Populasi | | | |
|----------|----------|------|----------|
| 1 | 2 | | k |
| x_{11} | x_{12} | ... | x_{1k} |
| x_{21} | x_{22} | ... | x_{2k} |
| x_{31} | x_{32} | ... | x_{3k} |
| ... | ... | ... | ... |
| x_{n1} | x_{n2} | ... | x_{nk} |

Ide Utama Anova

- Mengukur “**within**” variabilitas dan juga “**between**” variabilitas dari “treatment groups” .
 - “**Within**” treatment adalah variabilitas karena random error setiap observasi
 - “**Between**” treatment adalah Variabilitas karena perbedaan k treatment dan karena random error setiap observasi



Tabel 3. Tabel Anova

Keterangan:

$$SS_T = \sum_{t=1}^k n_t (\bar{y}_t - \bar{y})^2$$

$$SS_R = \sum_{t=1}^k \sum_{i=1}^{n_t} (y_{ti} - \bar{y}_t)^2$$

$$SS_D = \sum_{t=1}^k \sum_{i=1}^{n_t} (y_{ti} - \bar{y})^2$$

| Sumber Variasi | Kuadrat Jumlah | Derajat Bebas | Kuadrat Mean |
|-----------------------------|----------------|-------------------|---------------------|
| Between Treatmen | SS_T | $v_T = k - 1$ | $MS_T = SS_T / v_T$ |
| Within Treatmen | SS_R | $v_R = N - k$ | $MS_R = SS_R / v_R$ |
| Total (terhadap Grand Mean) | SS_D | $N_G = v_T + v_R$ | |

k = jumlah populasi

N = total jumlah observasi untuk semua populasi

Contoh,

Seorang manajer akan membandingkan output dari 3 buah mesin.

Manajer tersebut mencatat output dari mesin ke 1 dalam durasi 1 menit (dipilih secara random), dan diulangi sebanyak 5 kali. Dengan cara yang sama, dilakukan pula pencatatan untuk mesin ke 2 sebanyak 10 kali dan 6 kali untuk mesin ke 3.

| Mesin 1 | | Mesin 2 | | Mesin 3 | |
|---------|--|---------|--|---------|--|
| Y_1 | | Y_2 | | Y_3 | |
| 10 | | 6 | | 11 | |
| 6 | | 7 | | 8 | |
| 8 | | 9 | | 13 | |
| 12 | | 4 | | 10 | |
| 6 | | 6 | | 10 | |
| | | 10 | | 12 | |
| | | 5 | | | |
| | | 6 | | | |
| | | 8 | | | |
| | | 6 | | | |
| | | | | | |
| | | | | | |

Kita hitung mean pada masing-masing mesin .

$$Y1=A13/count(A3:A12)$$

Grand mean dari semua mesin adalah $(42+67+64)/21 = 8.238$.

$$\bar{Y}_1 = \frac{42}{5} = 8.4$$

$$\bar{Y}_2 = \frac{67}{10} = 6.7$$

$$\bar{Y}_3 = \frac{64}{6} = 10.67$$

| Mesin 1 | | Mesin 2 | | Mesin 3 | |
|---------|--|---------|--|---------|--|
| Y_1 | | Y_2 | | Y_3 | |
| 10 | | 6 | | 11 | |
| 6 | | 7 | | 8 | |
| 8 | | 9 | | 13 | |
| 12 | | 4 | | 10 | |
| 6 | | 6 | | 10 | |
| | | 10 | | 12 | |
| | | 5 | | | |
| | | 6 | | | |
| | | 8 | | | |
| | | 6 | | | |
| 42 | | 67 | | 64 | |
| | | | | | |

Kita hitung variabilitas tiap mesin (Within Treatment)

| Mesin1 | | Mesin 2 | | Mesin 3 | |
|--------|-----------------------|---------|-----------------------|---------|-----------------------|
| Y_1 | $(Y_1 - \bar{Y}_1)^2$ | Y_2 | $(Y_2 - \bar{Y}_2)^2$ | Y_3 | $(Y_3 - \bar{Y}_3)^2$ |
| 10 | 2.56 | 6 | 0.49 | 11 | 0.1111 |
| 6 | 5.76 | 7 | 0.09 | 8 | 7.1111 |
| 8 | 0.16 | 9 | 5.29 | 13 | 5.4444 |
| 12 | 12.96 | 4 | 7.29 | 10 | 0.4444 |
| 6 | 5.76 | 6 | 0.49 | 10 | 0.4444 |
| | | 10 | 10.89 | 12 | 1.7778 |
| | | 5 | 2.89 | | |
| | | 6 | 0.49 | | |
| | | 8 | 1.69 | | |
| | | 6 | 0.49 | | |
| 42 | 27.2 | 67 | 30.10 | 64 | 15.33 |
| | | | | | |

$$\bar{Y}_1 = \frac{42}{5} = 8.4$$

$$\bar{Y}_2 = \frac{67}{10} = 6.7$$

$$\bar{Y}_3 = \frac{64}{6} = 10.67$$

Kuadrat Jumlah variasi dari within groups SS_R adalah jumlah kuadrat dari variabilitas output setiap mesin = $27.2 + 30.10 + 15.33 = 72.63$.

$$\text{Rumus : } SS_R = \sum_{t=1}^k \sum_{i=1}^{n_t} (y_{ti} - \bar{y}_t)^2$$

$$\bar{Y}_1 = \frac{42}{5} = 8.4$$

$$\bar{Y}_2 = \frac{67}{10} = 6.7$$

$$\bar{Y}_3 = \frac{64}{6} = 10.67$$

| Machine 1 | | Machine 2 | | Machine 3 | |
|-----------|-----------------------|-----------|-----------------------|-----------|-----------------------|
| Y_1 | $(Y_1 - \bar{Y}_1)^2$ | Y_2 | $(Y_2 - \bar{Y}_2)^2$ | Y_3 | $(Y_3 - \bar{Y}_3)^2$ |
| 10 | 2.56 | 6 | 0.49 | 11 | 0.1111 |
| 6 | 5.76 | 7 | 0.09 | 8 | 7.1111 |
| 8 | 0.16 | 9 | 5.29 | 13 | 5.4444 |
| 12 | 12.96 | 4 | 7.29 | 10 | 0.4444 |
| 6 | 5.76 | 6 | 0.49 | 10 | 0.4444 |
| | | 10 | 10.89 | 12 | 1.7778 |
| | | 5 | 2.89 | | |
| | | 6 | 0.49 | | |
| | | 8 | 1.69 | | |
| | | 6 | 0.49 | | |
| 42 | 27.2 | 67 | 30.10 | 64 | 15.33 |
| | | | | | |

Jika grand mean dari semua mesin adalah $(42+67+64)/21 = 8.238$.

Maka Kuadrat jumlah dari between treatment/ Group :

$$SS_T = \sum_{t=1}^k n_t (\bar{y}_t - \bar{y})^2$$

$$= 5(8.4 - 8.238)^2 + 10(6.7 - 8.238)^2$$

$$+ 6(10.67 - 8.238)^2$$

$$= 5(0.026) + 10(2.365) + 6(5.898)$$

$$= 59.18$$

$$\bar{Y}_1 = \frac{42}{5} = 8.4$$

$$\bar{Y}_2 = \frac{67}{10} = 6.7$$

$$\bar{Y}_3 = \frac{64}{6} = 10.67$$

| Mesin 1 | | Mesin 2 | | Mesin 3 | |
|----------------|--|----------------|--|----------------|--|
| Y ₁ | | Y ₂ | | Y ₃ | |
| 10 | | 6 | | 11 | |
| 6 | | 7 | | 8 | |
| 8 | | 9 | | 13 | |
| 12 | | 4 | | 10 | |
| 6 | | 6 | | 10 | |
| | | 10 | | 12 | |
| | | 5 | | | |
| | | 6 | | | |
| | | 8 | | | |
| | | 6 | | | |
| 42 | | 67 | | 64 | |
| | | | | | |

Keterangan:

$$SS_T = \sum_{t=1}^k n_t (\bar{y}_t - \bar{y})^2$$

$$SS_R = \sum_{t=1}^k \sum_{i=1}^{n_t} (y_{ti} - \bar{y}_t)^2$$

$$SS_D = \sum_{t=1}^k \sum_{i=1}^{n_t} (y_{ti} - \bar{y})^2$$

Tabel 3. Tabel Anova

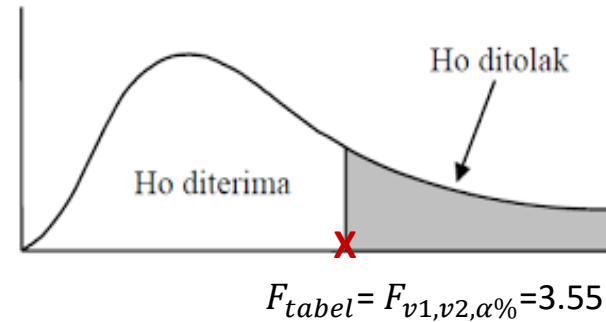
| Sumber Variasi | Kuadrat Jumlah | Derajat Bebas | Kuadrat Mean |
|-----------------------------|----------------|---------------|---------------------|
| Between Treatmen | SS_T | $v_T = k - 1$ | $MS_T = SS_T / v_T$ |
| Within Treatmen | SS_R | $v_R = N - k$ | $MS_R = SS_R / v_R$ |
| Total (terhadap Grand Mean) | SS_D | N_G | |

$$N_G = v_T + v_R$$

| Sumber Variasi | Jumlah Kuadrat | Derajat Bebas | Kudrat Mean | F-hitung |
|------------------|----------------|---------------|------------------|---|
| Between Treatmen | 59.18 | 2 | 59.18/2 = 29.588 | $\frac{MS_T}{MS_R} = \frac{29.588}{4.035}$ =7.33 |
| Within Treatmen | 72.63 | 18 | 72.63/18 = 4.035 | |
| Total | 131.81 | 20 | | |

Distribusi F dan F_{tabel}

$$F_{2,18,5\%} = 3.55$$



Jika kita bandingkan F-hitung = 7.33 dengan $F_{tabel} = 3.55$

Berarti F-hitung terletak di daerah kritis, daerah penolakan Hipotesis H_0

- $H_0: \mu_1 = \mu_2 = \mu_3$ = Tidak terdapat perbedaan signifikan antara rata-rata Mesin 1, Mesin 2 dan Mesin3.
- **$H_1: \mu_1 \neq \mu_2 \neq \mu_3$, Ada perbedaan yang signifikan antara rata-rata Mesin 1, Mesin 2 dan Mesin3**