

# Analisis Variansi Dua Arah pada Dataset Diving

Sahda Huwaidah Estiningtyas

2024-04-21

The Genesis Diving Institute of Florida menyatakan penyelam scuba menggunakan berbagai sistem pendidikan yang berbeda. Organisasi ini telah mempelajari waktu yang dihabiskan di bawah air oleh penyelam scuba dalam menjelajahi gua dan perairan terbuka. Setiap penyelam juga diklasifikasikan berdasarkan kelompok umur. Penyelaman mandiri dipilih secara acak untuk setiap kombinasi kelompok umur dan tipe penyelaman. Data (dalam menit) diberikan dalam sebuah dataset.

Lakukan analisis variansi dua arah. Interpretasikan hasilnya. Gunakan  $\alpha = 0.001$  untuk setiap pengujian hipotesis.

## Menyiapkan Dataset

```
# Membaca data
diving <- read.csv("C:/Users/LENOVO/OneDrive/Documents/UNY/Semester 2/diving.csv", head = TRUE)
head(diving)
```

```
##   time dive.type age.group
## 1   39      cave    20-30
## 2   38      cave    20-30
## 3   41      cave    20-30
## 4   39      cave    20-30
## 5   41      cave    20-30
## 6   42      cave    20-30
```

```
str(diving)
```

```
## 'data.frame':   64 obs. of  3 variables:
## $ time      : int  39 38 41 39 41 42 40 40 42 40 ...
## $ dive.type: chr   "cave" "cave" "cave" "cave" ...
## $ age.group: chr   "20-30" "20-30" "20-30" "20-30" ...
```

```
diving$int <- with(diving, interaction(age.group, dive.type), drop=TRUE)
head(diving$int)
```

```
## [1] 20-30.cave 20-30.cave 20-30.cave 20-30.cave 20-30.cave 20-30.cave
## 8 Levels: ≥ 50.cave 20-30.cave 30-40.cave 40-50.cave ... 40-50.open.water
```

```
attach(diving)
```

## Model Linear

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

Dengan:

- $i = 1, 2, 3, 4$  mewakili kelompok umur (1: 20-30, 2: 30-40, 3: 40-50, 4:  $\geq 50$ )
- $j = 1, 2$  mewakili tipe penyelaman (1: cave, 2: open water)
- $k = 1, 2, 3, 4, 5, 6, 7, 8$  mewakili penyelam  $k$

Parameter model:

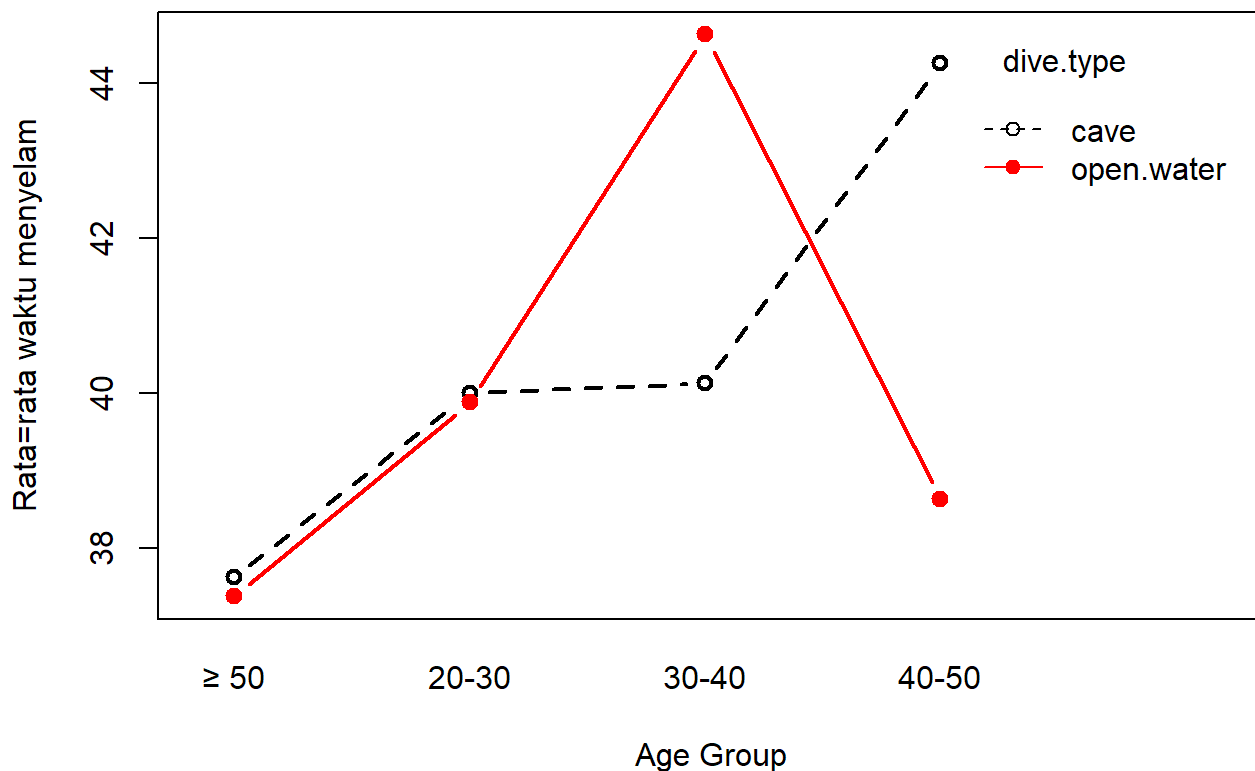
- $\mu$  adalah rata-rata waktu yang dihabiskan oleh penyelam secara keseluruhan.
- $\alpha_i$  adalah pengaruh kelompok umur ke- $i$ .
- $\beta_j$  adalah pengaruh tipe penyelaman ke- $j$ .
- $(\alpha\beta)_{ij}$  adalah pengaruh interaksi antara kelompok umur ke- $i$  dan tipe penyelaman ke- $j$ .
- $\epsilon_{ijk}$  adalah galat dari interaksi kelompok umur dan tipe penyelam pada penyelam  $k$ .
- Asumsi:  $\sum_{i=1}^4 \alpha_i = 0$ ,  $\sum_{j=1}^2 \beta_j = 0$ ,  $\sum_{i=1}^4 \sum_{j=1}^2 (\alpha\beta)_{ij} = 0$ ,  $\sum_{j=1}^2 (\alpha\beta)_{ij} = 0$
- Galat percobaan diasumsikan independen serta  $\epsilon_{ijk} \sim N(0, \sigma^2)$

Interpretasi hasil:

- Koefisien  $\alpha_i$  menunjukkan perbedaan waktu yang dihabiskan di bawah air antara kelompok umur.
- Koefisien  $\beta_j$  menunjukkan perbedaan waktu yang dihabiskan di bawah air antara tipe penyelaman.
- Interaksi  $(\alpha\beta)_{ij}$  menunjukkan apakah efek kelompok umur berbeda antara tipe penyelaman.

## Plot Interaksi

```
interaction.plot(x.factor = age.group, trace.factor = dive.type, response = time, fun = mean,
               type = "b", legend=TRUE, xlab="Age Group", ylab = "Rata-rata waktu menyelam",
               pch = c(1,19), col = c("black", "red", "blue"), lwd=2)
```



- Berdasarkan plot di atas, terlihat bahwa terdapat pengaruh interaksi antara umur dan tipe menyelam.
- Terdapat perbedaan yang signifikan dalam rata-rata waktu bagi semua kelompok umur antara tipe menyelam cave dan tipe menyelam open water, kecuali pada kelompok umur 30-40. Pada kelompok umur tersebut, rata-rata waktu yang lebih tinggi terlihat pada tipe menyelam open water dibandingkan dengan tipe menyelam cave.
- Perbandingan antara kelompok umur 20-30 dan ≥ 50 menunjukkan perbedaan rata-rata waktu yang kecil dan hampir sama. Namun, perbedaan yang cukup signifikan terlihat antara kelompok umur 30-40 dan 40-50. Dari hasil ini, dapat disimpulkan bahwa tidak ada pengaruh tipe menyelam terhadap waktu.
- Ketika membandingkan tipe menyelam cave dan open water, terlihat perbedaan yang cukup besar dalam rata-rata waktu di antara keempat kelompok umur. Hal ini menunjukkan bahwa terdapat pengaruh kelompok umur terhadap waktu.

## Penyelesaian:

### Langkah 1:

Anggap bahwa asumsi-asumsi dalam analisis variansi dua arah terpenuhi. Terdapat  $a = 4$  level faktor A (jumlah waktu),  $b = 2$  level faktor B (group age), dan  $n = 8$  pengamatan dalam masing-masing sel.

```
tapply(time, int, sum)
```

##	≥ 50.cave	20-30.cave	30-40.cave	40-50.cave
##	301	320	321	354
##	≥ 50.open.water	20-30.open.water	30-40.open.water	40-50.open.water
##	299	319	357	309

```
tapply(time, age.group, sum)
```

##	≥ 50	20-30	30-40	40-50
##	600	639	678	663

```
tapply(time, dive.type, sum)
```

##	cave	open.water
##	1296	1284

(TABEL BELUM DIBUAT)

## Langkah 2:

### Menentukan Jumlah Kuadrat:

#### Jumlah Kuadrat Total:

$$\begin{aligned}
 JKT &= \left( \sum_{i=1}^4 \sum_{j=1}^2 \sum_{k=1}^8 y_{ijk}^2 \right) - \frac{t^2}{(4)(2)(8)} \\
 &= (39^2 + 38^2 + \dots + 39^2) - \frac{2580^2}{64} \\
 &= (12812 + 12775 + 12955 + 16005 + 15772 + 12031 + 11373 + 11289) - \frac{6656400}{64} \\
 &= 105012 - 104006.25 = 1005.75
 \end{aligned}$$

#### Jumlah Kuadrat A (Faktor A):

$$\begin{aligned}
 JKA &= \frac{\sum_{i=1}^4 t_{i..}^2}{(2)(8)} - \frac{t^2}{(4)(2)(8)} \\
 &= \frac{639^2 + 678^2 + 663^2 + 600^2}{16} - \frac{2580^2}{64} \\
 &= 104223.375 - 104006.25 \\
 &= 217.125
 \end{aligned}$$

#### Jumlah Kuadrat B (Faktor B):

$$\begin{aligned}
 JKB &= \frac{\sum_{j=1}^2 t_{.j.}^2}{(4)(8)} - \frac{t^2}{(4)(2)(8)} \\
 &= \frac{1296^2 + 1284^2}{32} - \frac{2580^2}{64} \\
 &= 104008.5 - 104006.25 \\
 &= 2.25
 \end{aligned}$$

**Jumlah Kuadrat Interaksi (AB):**

$$\begin{aligned}
 JKAB &= \frac{\sum_{i=1}^4 \sum_{j=1}^2 t_{ij}^2}{(4)} - \frac{\sum_{i=1}^4 t_{i..}^2}{(2)(8)} - \frac{\sum_{j=1}^2 t_{.j.}^2}{(4)(8)} + \frac{t^2}{(4)(2)(8)} \\
 &= 104223.375 - 104008.5 + 104006.25 - 104223.375 - 104008.5 + 104006.25 \\
 &= 26055.2813 - 104221.125 \\
 &= -78165.84
 \end{aligned}$$

**Jumlah Kuadrat Galat (G):**

$$JKG = JKT - JKA - JKB - JKAB = 1005.75 - 217.125 - 2.25 - (-78165.84) = 78952.21$$

**Langkah 3:**

- **Hipotesis:**
  - $H_0 : (\alpha\beta)_{11} = (\alpha\beta)_{12} = \dots = (\alpha\beta)_{42}$  (Tidak ada pengaruh hubungan antara group age dan dive type terhadap jumlah waktu diving)
  - $H_1 : \exists(\alpha\beta)_{ij} \neq 0, i = 1, 2, 3, 4; j = 1, 2$  (Ada ada pengaruh hubungan antara group age dan dive type terhadap jumlah waktu diving)
- **Taraf signifikansi:**  $\alpha = 0.001$
- **Statistik uji:**
  - $F_{AB} = \frac{KTAB}{KTG}$
- **Kriteria keputusan:**

```

a = 4
b = 2
n = 8
qf(0.999, df1=(a-1)*(b-1), df2 = a*b*(n-1))

```

```
## [1] 6.229585
```

```

p.value <- 1-pf(-18.48, df1 = (a-1)*(b-1), df2 = a*b*(n-1))
p.value

```

```
## [1] 1
```

$$F_{0.001((a-1)(b-1), ab(n-1))} = F_{0.001(3,56)} = 6.229585$$

H0 ditolak jika  $FAB > 6.22$  atau H0 ditolak jika  $p - value < 0.001$

- **Hitungan:**
  - $F_{AB} = \frac{KTAB}{KTG} = \frac{-78165.84/3}{78952.21/56} = -18.48$
- **Kesimpulan:**
  - Oleh karena  $F_{AB} = -18.48 < 6.229585$  atau  $p\text{-value} = 1 > 0.001$ , maka  $H_0$  tidak ditolak. Jadi pada taraf signifikansi 0.001 dapat disimpulkan bahwa tidak ada hubungan antara age group dan dive

type terhadap jumlah waktu diving.

## Langkah 4:

### A. Cek untuk Pengaruh Utama Group Age (Faktor A)

- **Hipotesis:**
  - $H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$  (Tidak ada pengaruh hubungan antara group age dan jumlah waktu diving)
  - $H_1 : \exists \alpha_i \neq 0, i = 1, 2, 3, 4$  (Ada pengaruh hubungan antara group age dan jumlah waktu diving)
- **Taraf signifikansi:**  $\alpha = 0.001$
- **Statistik uji:**  $F_A = \frac{KTA}{KTG}$
- **Kriteria keputusan:**

```
a = 4
b = 2
n = 8
qf(0.999, df1=(a-1), df2 = a*b*(n-1))
```

```
## [1] 6.229585
```

```
p.value <- 1-pf(0.051, df1 = (a-1), df2 = a*b*(n-1))
p.value
```

```
## [1] 0.9846284
```

$$F_{\alpha}((a-1), a \cdot b \cdot (n-1)) = F_{0.001}(3, 56) = 6.229585$$

$H_0$  ditolak jika  $F_A > 6.22$  atau  $p\text{-value} < 0.001$

- **Hitungan:**
  - $F_A = \frac{KTA}{KTG} = \frac{217.125}{78952.21} = 0.051$
- **Kesimpulan:**
  - $F_A = 0.051$ ,  $p\text{-value} = 0.984$
  - Karena  $F_A = 0.051 < 6.229585$  atau  $p\text{-value} = 0.984 > 0.001$ , maka  $H_0$  tidak ditolak. Jadi pada taraf signifikansi 0.001 dapat disimpulkan bahwa tidak ada hubungan antara age group terhadap jumlah waktu diving.

### B. Cek untuk Pengaruh Utama Dive Type (Faktor B)

- **Hipotesis:**
  - $H_0 : \beta_1 = \beta_2 = 0$  (Tidak ada pengaruh hubungan antara dive type dan jumlah waktu diving)
  - $H_1 : \exists \beta_i \neq 0, i = 1, 2$  (Ada pengaruh hubungan antara dive type dan jumlah waktu diving)
- **Taraf signifikansi:**  $\alpha = 0.001$
- **Statistik uji:**  $F_B = \frac{KTB}{KTG}$
- **Kriteria keputusan:**

```
a = 4
b = 2
n = 8
qf(0.999, df1=(b-1), df2 = a*b*(n-1))
```

```
## [1] 12.06115
```

```
p.value <- 1-pf(0.0015, df1 = (b-1), df2 = a*b*(n-1))
p.value
```

```
## [1] 0.9692435
```

$$F_{\alpha}((b-1), a \cdot b \cdot (n-1)) = F_{0.001}(1, 56) = 12.06115$$

$H_0$  ditolak jika  $F_B > 12.06115$  atau  $p\text{-value} < 0.001$

- **Hitungan:**

- $F_B = \frac{KTB}{KTG} = \frac{2.25}{78952.21} = 0.0015$

- **Kesimpulan:**

- Karena  $F_B = 0.0015 < 12.06115$  atau  $p\text{-value} = 0.9692435 > 0.001$ , maka  $H_0$  tidak ditolak. Jadi pada taraf signifikasi 0.001 dapat disimpulkan bahwa tidak ada hubungan antara dive type terhadap jumlah waktu diving.