

Praktikum Analisis Data Statistika

Analisis Ragam (ANOVA)

Tujuan Praktikum

1. Menentukan model linier sesuai rancangan percobaan yaitu Rancangan Acak Lengkap , Rancangan Acak Kelompok , atau Rancangan Bujur Sangkar Latin.
2. Menghitung komponen-komponen ANOVA menggunakan perangkat lunak R Studio
3. Menyusun tabel ANOVA dan menarik kesimpulan berdasarkan nilai- F dan nilai- p .
4. Menginterpretasikan hasil uji ANOVA dalam konteks penelitian atau eksperimen nyata.

Pendahuluan

Dalam penelitian eksperimental, rancangan percobaan (design of experiment) digunakan untuk mengatur dan mengendalikan variasi agar pengaruh perlakuan dapat diuji secara obyektif.

Tiga rancangan dasar yang umum digunakan adalah:

1. RAL (Rancangan Acak Lengkap)
2. RAK (Rancangan Acak Kelompok)
3. RBSL (Rancangan Bujur Sangkar Latin)

Masing-masing rancangan memiliki karakteristik berbeda dalam mengatasi sumber ketidakhomogenan pada unit percobaan.

Rancangan Acak Lengkap (RAL)

Rancangan Acak Lengkap digunakan ketika satuan percobaan dianggap homogen, artinya seluruh unit memiliki kondisi lingkungan yang relatif sama.

Ciri-ciri Rancangan Acak Lengkap:

- Hanya melibatkan satu faktor perlakuan.
- Semua perlakuan ditempatkan secara acak penuh pada unit percobaan.
- Cocok digunakan pada percobaan laboratorium atau kondisi lingkungan terkendali.

Model Linier Rancangan Acak Lengkap :

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

Keterangan:

- Y_{ij} : respon pada perlakuan ke- i , ulangan ke- j
- μ : rata-rata umum
- τ_i : pengaruh perlakuan ke- i
- ε_{ij} : galat acak

Ilustrasi Rancangan Acak Lengkap

Sebuah perusahaan tekstil sedang melakukan penelitian untuk mengetahui bagaimana variasi kadar serat katun dalam campuran bahan memengaruhi kekuatan kain yang dihasilkan. Dalam proses produksinya, serat katun dicampurkan dengan serat sintetis dalam proporsi tertentu. Perusahaan menduga bahwa semakin tinggi kandungan serat katun, maka kekuatan kain yang dihasilkan dapat berubah — namun belum diketahui apakah perubahan tersebut signifikan secara statistik. Percobaan dilakukan di laboratorium dengan rancangan Rancangan Acak Lengkap (RAL), di mana setiap unit percobaan dianggap homogen seperti pada Tabel 1.

- Faktor yang diuji: Persentase berat katun dalam campuran bahan.
- Taraf (level) perlakuan: 15%, 20%, 25%, 30%, dan 35%.
- Jumlah ulangan: 5 kali pengulangan untuk setiap taraf.
- Variabel respon: Kekuatan bahan tekstil (dalam lb/in²).
- Kondisi percobaan: Seluruh pengujian dilakukan di laboratorium dengan suhu dan kelembapan terkontrol.

Table 1: Kekuatan bahan (lb/in²) pada berbagai kadar katun

Persentase_Katun	U1	U2	U3	U4	U5
15	7	7	15	11	9
20	12	17	12	18	18
25	14	18	18	19	19
30	19	25	22	19	23
35	7	10	11	15	11

Penyelesaian:

Model Linier Rancangan Acak Lengkap yang digunakan:

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

Keterangan :

- Y_{ij} : nilai kekuatan kain pada perlakuan ke- i dan ulangan ke- j
- μ : rata-rata umum
- τ_i : pengaruh perlakuan ke- i
- ε_{ij} : galat acak

Hipotesis :

$$H_0 : \mu_{15} = \mu_{20} = \mu_{25} = \mu_{30} = \mu_{35}$$

(Rata-rata kekuatan kain pada kadar 15%, 20%, 25%, 30%, dan 35% serat katun adalah sama)

$$H_1 : \text{paling sedikit ada satu } \mu_i \neq \mu_j$$

(Ada perbedaan rata-rata kekuatan kain pada salah satu kadar serat katun yang diuji.)

Analisis ANOVA

```
library(tidyverse)
library(dplyr)
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##   filter, lag
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
# Data
data <- data.frame(
  Persentase = c(15, 20, 25, 30, 35),
  U1 = c(7, 12, 14, 19, 7),
  U2 = c(7, 17, 18, 25, 10),
  U3 = c(15, 12, 18, 22, 11),
  U4 = c(11, 18, 19, 19, 15),
  U5 = c(9, 18, 19, 23, 11)
)
head(data)
```

```
##   Persentase U1 U2 U3 U4 U5
## 1         15  7  7 15 11  9
## 2         20 12 17 12 18 18
## 3         25 14 18 18 19 19
## 4         30 19 25 22 19 23
## 5         35  7 10 11 15 11
```

Kode diatas digunakan untuk membentuk data awal dalam format tabel lebar (wide format).

```
# Ubah ke format long
data_long <- pivot_longer(
  data,
  cols = starts_with("U"),
  names_to = "Ulangan",
  values_to = "Kekuatan"
)

# Ubah ke faktor
data_long$Persentase <- factor(data_long$Persentase)
head(data_long)
```

```
## # A tibble: 6 x 3
##   Persentase Ulangan Kekuatan
##   <fct>      <chr>      <dbl>
## 1 15        U1          7
## 2 15        U2          7
## 3 15        U3         15
## 4 15        U4         11
## 5 15        U5          9
## 6 20        U1         12
```

Penjelasan:

1. `pivot_longer()` mengubah data dari format lebar ke format panjang.
2. `cols = starts_with("U")` berarti semua kolom yang dimulai dengan huruf `U` ($U1 \sim U5$) dijadikan satu kolom baru bernama Kekuatan.
3. Kolom Persentase diubah menjadi faktor karena merupakan variabel perlakuan kategorikal.

4. Struktur ini memudahkan fungsi `aov()` dalam membaca model statistik.

```
# Analisis ANOVA
anova_model <- aov(Kekuatan ~ Persentase, data = data_long)
summary(anova_model)

##              Df Sum Sq Mean Sq F value    Pr(>F)
## Persentase    4  475.8   118.94    14.76 9.13e-06 ***
## Residuals    20   161.2     8.06
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Penjelasan:

1. Fungsi `aov()` digunakan untuk menjalankan uji ANOVA satu arah.
2. Sintaks `Kekuatan ~ Persentase` menyatakan bahwa Kekuatan bahan adalah variabel respon, sedangkan Persentase adalah faktor perlakuan.
3. Fungsi `summary()` menampilkan tabel ANOVA yang berisi:
Df: derajat bebas (degree of freedom) Sum Sq: jumlah kuadrat (Sum of Squares) Mean Sq: kuadrat tengah (Mean Square) F value: nilai statistik uji F Pr(>F): nilai-p untuk pengujian hipotesis
4. Hasil `summary(anova_model)` menunjukkan apakah ada perbedaan rata-rata kekuatan bahan yang signifikan antar taraf persentase katun.

Interpretasi Hasil :

Nilai F hitung = 14.76

Nilai p ($\Pr(>F)$) = 9.13×10^{-6}

Taraf signifikansi (α) = 0.05

Karena nilai-p jauh lebih kecil dari 0.05, maka: $p < 0.05 \Rightarrow$ Tolak H_0 . Artinya,

- Terdapat perbedaan yang signifikan secara statistik antara rata-rata kekuatan bahan kain pada berbagai kadar serat katun.
Dengan kata lain, persentase serat katun berpengaruh signifikan terhadap kekuatan bahan tekstil.
- Perbedaan kekuatan kain antara kadar katun 15%, 20%, 25%, 30%, dan 35% tidak hanya disebabkan oleh variasi acak, tetapi benar-benar karena pengaruh kadar serat katun yang berbeda. Hal ini menunjukkan bahwa komposisi serat katun dalam bahan merupakan faktor penting yang menentukan tingkat kekuatan kain. Semakin tepat kadar serat katun yang digunakan, semakin optimal pula kualitas kekuatan kain yang dihasilkan oleh perusahaan.

Rancangan Acak Kelompok (RAK)

Rancangan Acak Kelompok digunakan ketika terdapat satu sumber ketidakhomogenan utama yang dapat dikendalikan dengan pengelompokan.

Ciri-ciri Rancangan Acak Kelompok:

- Kondisi lingkungan tidak homogen satu arah (misal: perbedaan kesuburan tanah antar baris).
- Unit percobaan dikelompokkan menjadi blok/kelompok yang relatif homogen.
- Setiap kelompok memuat seluruh perlakuan secara acak.

Model Linier Rancangan Acak Kelompok :

$$Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$$

Simbol	Keterangan
Y_{ij}	Nilai pengamatan atau respon pada perlakuan ke- i dan kelompok ke- j .
μ	Rataan umum dari seluruh pengamatan.
τ_i	Pengaruh dari perlakuan ke- i , yaitu perbedaan rata-rata akibat faktor perlakuan yang diuji.
β_j	Pengaruh dari kelompok (blok) ke- j , yang mencerminkan variasi kondisi lingkungan antar kelompok.
ε_{ij}	Galat acak pada perlakuan ke- i dan kelompok ke- j , dengan asumsi: $\varepsilon_{ij} \sim N(0, \sigma^2)$.

Ilustrasi Rancangan Acak Kelompok:

Seorang peneliti ingin membandingkan pengaruh tiga jenis insektisida (A, B, dan C) terhadap pertumbuhan tanaman kacang tanah. Lahan percobaan diketahui tidak homogen satu arah — misalnya antar lajur memiliki tingkat kesuburan yang berbeda. Oleh karena itu, dilakukan pengelompokan berdasarkan lajur untuk mengendalikan variasi lingkungan. Rancangan yang digunakan adalah Rancangan Acak Kelompok Lengkap (RAKL) seperti pada Tabel 3. Desain percobaan sebagai berikut

- Perlakuan: 3 jenis insektisida (A, B, C)
- Kelompok/blok: 4 lajur (Kel1, Kel2, Kel3, Kel4)
- Total unit percobaan: $3 \times 4 = 12$
- Respon: Bobot kering kacang tanah (gram)
- Pengacakan: Pada setiap kelompok, ketiga perlakuan diacak dan masing-masing muncul satu kali.

Table 3: Bobot kering kacang tanah (gram) pada berbagai jenis insektisida dan kelompok lajur.

Insektisida	Kel1	Kel2	Kel3	Kel4	Total
A	56	48	66	62	232
B	83	78	94	93	348
C	80	72	83	85	320
Total	219	198	243	240	900

Penyelesaian:

Model linier Rancangan Acak Kelompok yang digunakan:

$$Y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$$

Simbol	Keterangan
Y_{ij}	Respon atau nilai pengamatan pada perlakuan ke- i dan kelompok ke- j .
μ	Rataan umum dari seluruh pengamatan.
τ_i	Pengaruh perlakuan ke- i (misalnya jenis insektisida ke- i).
β_j	Pengaruh kelompok (blok) ke- j (misalnya perbedaan antar lajur).

Simbol	Keterangan
ε_{ij}	Galat acak (error) pada perlakuan ke- i dan kelompok ke- j , dengan asumsi $\varepsilon_{ij} \sim N(0, \sigma^2)$.

Hipotesis :

1. Hipotesis untuk pengaruh jenis insektisida

$$H_0 : \tau_1 = \tau_2 = \tau_3 = 0 \quad (\text{tidak ada pengaruh jenis insektisida})$$

$$H_1 : \text{paling sedikit ada satu } \tau_i \neq 0$$

2. Hipotesis untuk Kelompok

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0 \quad (\text{kelompok tidak berpengaruh signifikan})$$

$$H_1 : \text{paling sedikit ada satu } \beta_j \neq 0$$

Analisis ANOVA

```
# Membentuk data long
insektisida <- data.frame(
  Insektisida = c("A", "B", "C"),
  Kel1 = c(56, 83, 80),
  Kel2 = c(48, 78, 72),
  Kel3 = c(66, 94, 83),
  Kel4 = c(62, 93, 85)
)

# Tambahkan total tiap baris dan kolom
insektisida$Total <- rowSums(insektisida[, 2:5])
col_total <- colSums(insektisida[, 2:6])
insektisida <- rbind(insektisida, c("Total", as.numeric(col_total)))
library(tidyr)
data_long <- insektisida[-4, ] # hapus baris total
data_long <- pivot_longer(data_long, cols = starts_with("Kel"),
  names_to = "Kelompok", values_to = "Bobot")
```

Penjelasan :

1. Fungsi `pivot_longer()` dari paket `tidyr` digunakan untuk mengubah bentuk data dari wide (lebar) menjadi long. Dalam bentuk wide, setiap kolom mewakili satu kelompok (Kel1, Kel2, dst). Dalam bentuk long, seluruh nilai pengamatan disatukan dalam satu kolom (Bobot) dengan kolom tambahan (Kelompok) yang menunjukkan asal kelompoknya.
2. Argumen:
 - `cols = starts_with("Kel")`: memilih semua kolom yang diawali kata "Kel".
 - `names_to = "Kelompok"`: nama kolom baru yang menyimpan label kelompok.
 - `values_to = "Bobot"`: nama kolom baru yang menyimpan nilai respon (bobot kering kacang tanah).
3. Baris `insektisida[-4,]` digunakan untuk menghapus baris ke-4 karena berisi total keseluruhan (bukan data pengamatan).

```
# Konversi tipe data
data_long$Insektisida <- factor(data_long$Insektisida)
data_long$Kelompok <- factor(data_long$Kelompok)
```

Penjelasan :

1. Kolom Insektisida dan Kelompok dikonversi menjadi faktor, karena kedua variabel tersebut adalah variabel kategorik:

Insektisida: mewakili tiga jenis insektisida (A, B, C).

Kelompok: mewakili empat blok/lajur percobaan.

2. Konversi ini penting karena fungsi `aov()` mengharuskan faktor independen bertipe factor, bukan numerik.

```
# Model ANOVA RAK
anova_rakl <- aov(Bobot ~ Insektisida + Kelompok, data = data_long)
summary(anova_rakl)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## Insektisida   2   1832    916.0   211.38 2.74e-06 ***
## Kelompok      3    438    146.0    33.69 0.000377 ***
## Residuals     6     26     4.3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Penjelasan :

- Faktor Insektisida memiliki nilai $F = 211.38$ dengan p-value = 2.74×10^{-6} , yang jauh lebih kecil dari 0.05. Artinya, terdapat perbedaan yang sangat signifikan antar jenis insektisida terhadap bobot hasil. Dengan kata lain, jenis insektisida yang digunakan berpengaruh nyata terhadap bobot tanaman.
- Faktor Kelompok memiliki nilai $F = 33.69$ dengan p-value = 0.000377, juga lebih kecil dari 0.05. → Hal ini menunjukkan bahwa kelompok (misalnya lokasi, waktu, atau kondisi lingkungan) memberikan pengaruh signifikan terhadap bobot hasil.
- Residuals (galat) menunjukkan nilai Mean Sq sebesar 4.3 yang relatif kecil, menandakan bahwa variasi antar pengamatan di dalam perlakuan cukup rendah.

Kesimpulan:

Faktor Insektisida maupun Kelompok berpengaruh signifikan terhadap bobot hasil ($p < 0.001$).

Rancangan Bujur Sangkar Latin (RBSL)

Rancangan Bujur Sangkar Latin diterapkan ketika unit percobaan tidak homogen dalam dua arah, misalnya baris dan kolom.

Ciri-ciri Rancangan Bujur Sangkar Latin:

- Setiap perlakuan muncul tepat satu kali pada setiap baris dan kolom.
- Mengontrol dua sumber ketidakhomogenan sekaligus.
- Digunakan untuk percobaan dengan ordo yang sama (misal: 4 perlakuan \times 4 baris \times 4 kolom).

Model Linier Rancangan Bujur Sangkar Latin

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \tau_k + \varepsilon_{ijk}$$

Keterangan:

- α_i : pengaruh baris ke- i
- β_j : pengaruh kolom ke- j

- τ_k : pengaruh perlakuan ke- k

Ilustrasi Rancangan Bujur Sangkar Latin

Seorang peneliti ingin membandingkan tingkat keausan 4 merek ban: A, B, C, dan D. Kondisi lapangan menunjukkan ada 2 sumber ketidakhomogenan: jenis kendaraan dan posisi ban (misal: depan kiri, depan kanan, belakang kiri, belakang kanan). Untuk mengendalikan kedua sumber variasi itu digunakan Rancangan Bujur Sangkar Latin (RBSL) dengan ordo 4. Dimana setiap merek ban harus muncul tepat satu kali pada setiap jenis kendaraan dan pada setiap posisi ban seperti pada Tabel 5

Table 5: Nilai Keausan Ban

Baris	X2	X4	X1	X3
3	18.52	22.46	23.66	17.68
2	17.68	24.29	18.82	22.02
4	26.04	17.49	21.06	18.91
1	22.51	19.98	17.15	22.14

Penyelesaian :

Model linier Rancangan Bujur Sangkar Latin:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

Keterangan:

- Y_{ij} : respon atau nilai keausan ban pada kombinasi perlakuan ke- i dan kelompok ke- j
- μ : rata-rata umum dari seluruh pengamatan
- α_i : pengaruh dari jenis insektisida ke- i (faktor pertama)
- β_j : pengaruh dari kelompok ke- j (faktor kedua / blok)
- ε_{ij} : galat acak (error) yang diasumsikan berdistribusi normal $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$

Hipotesis:

1. Hipotesis untuk Jenis Kendaraan

$$H_0 : \tau_1 = \tau_2 = \dots = \tau_a = 0$$

(Tidak ada perbedaan rata-rata keausan ban antar jenis kendaraan)

$$H_1 : \text{Paling sedikit satu } \tau_i \neq 0$$

(Ada paling sedikit satu jenis kendaraan yang memiliki rata-rata keausan ban berbeda.)

2. Hipotesis untuk posisi ban

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_b = 0$$

(Tidak ada perbedaan rata-rata keausan ban antar posisi ban)

$$H_1 : \text{Paling sedikit satu } \beta_j \neq 0$$

(Ada paling sedikit satu posisi ban yang memiliki rata-rata keausan ban berbeda.)

Analisis ANOVA

```
# respons dibaca baris per baris
y <- c(
  18.52, 22.46, 23.66, 17.68,
```



```

17.68, 24.29, 18.82, 22.02,
26.04, 17.49, 21.06, 18.91,
22.51, 19.98, 17.15, 22.14
)

# faktor baris (misal: jenis kendaraan)
row <- factor(rep(1:4, each = 4))

# faktor kolom (misal: posisi ban)
col <- factor(rep(1:4, times = 4))

# perlakuan sesuai pola Latin di atas
treat <- factor(c(
  "B", "C", "D", "A",
  "A", "D", "B", "C",
  "D", "A", "C", "B",
  "C", "B", "A", "D"
))

latin_dat <- data.frame(
  keausan = y,
  baris = row,
  kolom = col,
  perlakuan = treat
)

knitr::kable(latin_dat, caption = "Data Keausan Ban", align = "c")

```

Table 6: Data Keausan Ban

keausan	baris	kolom	perlakuan
18.52	1	1	B
22.46	1	2	C
23.66	1	3	D
17.68	1	4	A
17.68	2	1	A
24.29	2	2	D
18.82	2	3	B
22.02	2	4	C
26.04	3	1	D
17.49	3	2	A
21.06	3	3	C
18.91	3	4	B
22.51	4	1	C
19.98	4	2	B
17.15	4	3	A
22.14	4	4	D

Penjelasan :

1. Respon (y) Vektor y berisi nilai keausan ban hasil pengamatan dari setiap kombinasi baris dan kolom dalam rancangan bujur sangkar 4×4 .
Nilai tersebut dimasukkan baris per baris, mengikuti urutan pengamatan pada tabel percobaan.

2. Faktor Baris (`row`) Variabel `row` digunakan untuk menunjukkan blok horizontal, misalnya jenis kendaraan.
Kode `rep(1:4, each = 4)` berarti setiap angka (1–4) diulang sebanyak 4 kali — sesuai jumlah kolom per baris.
3. Faktor Kolom (`col`) Variabel `col` mewakili blok vertikal, misalnya posisi ban (depan kiri, depan kanan, belakang kiri, belakang kanan).
Kode `rep(1:4, times = 4)` menghasilkan urutan 1–4 yang diulang sebanyak 4 kali — sesuai jumlah baris.
4. Perlakuan (`treat`) Variabel `treat` berisi jenis perlakuan yang diberikan, misalnya jenis ban.
Setiap huruf (A, B, C, D) muncul tepat satu kali di setiap baris dan kolom — sesuai prinsip Rancangan Bujur Sangkar Latin (Latin Square Design).
5. Membentuk Data Frame Semua komponen (`y`, `row`, `col`, `treat`) digabungkan ke dalam satu data frame bernama `latin_dat`.
Data ini berbentuk format long, yang dibutuhkan untuk analisis ANOVA di R.
6. Menampilkan Data Fungsi `knitr::kable()` digunakan untuk menampilkan tabel.

```
# Model RBSL: respons dijelaskan oleh perlakuan + 2 blok

model_rbsl <- aov(keausan ~ perlakuan + baris + kolom, data = latin_dat)
summary(model_rbsl)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## perlakuan     3 103.03   34.34   31.236 0.000466 ***
## baris         3   0.40    0.13    0.122 0.943985
## kolom         3   3.58    1.19    1.085 0.424159
## Residuals     6   6.60    1.10
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Penjelasan :

1. Keausan sebagai variabel respon (nilai yang diamati),
2. Perlakuan sebagai faktor utama yang diuji pengaruhnya,
3. Baris dan kolom sebagai dua faktor blok (pengendali keragaman).

Interpretasi Hasil ANOVA RBSL

1. Faktor Perlakuan

Nilai p-value untuk faktor perlakuan adalah 0.000466 (< 0.05), sehingga H_0 ditolak. Ini berarti terdapat perbedaan rata-rata keausan ban yang sangat signifikan antar perlakuan. Dengan kata lain, jenis perlakuan (misalnya jenis ban) memang memengaruhi tingkat keausan ban.

2. Faktor Baris

Nilai p-value untuk faktor baris adalah 0.943985 (> 0.05), sehingga H_0 tidak ditolak. Artinya, perbedaan antar baris (misalnya jenis kendaraan atau blok horizontal) tidak memberikan pengaruh yang signifikan terhadap keausan ban.

3. Faktor Kolom

Nilai p-value untuk faktor kolom adalah 0.424159 (> 0.05), sehingga H_0 tidak ditolak. Ini berarti perbedaan posisi/kolom (misalnya posisi ban) tidak berpengaruh signifikan terhadap keausan ban.

4. Kesimpulan

- Variasi utama pada data dijelaskan oleh perlakuan

- Dua faktor pengontrol (baris dan kolom) tidak signifikan, sehingga perbedaan keausan yang terjadi terutama karena perbedaan perlakuan.