

UAS statistika

Bagas Arya Amirul Jawad

2023-07-24



**UAS STATISTIKA**

Oleh

**NAMA: Bagas Arya Amirul Jawad**

**NIM: 21102198**

**PROGRAM STUDI S1 INFORMATIKA  
FAKULTAS INFORMATIKA  
INSTITUT TEKNOLOGI TELKOM PURWOKERTO  
2023**

# UAS Statistika Dan Probabilitas

## Soal Nomor 1

Jelaskan apa yang dimaksud dengan variabel random

Jawaban: Variabel random adalah variabel yang nilainya tidak tetap dan dapat berubah secara acak dalam suatu eksperimen atau kejadian.

## soal nomor 2

a. buat dataset

1. ID number

```
n <- 10
ID_number <- seq(1:n)
ID_number
```

```
## [1] 1 2 3 4 5 6 7 8 9 10
```

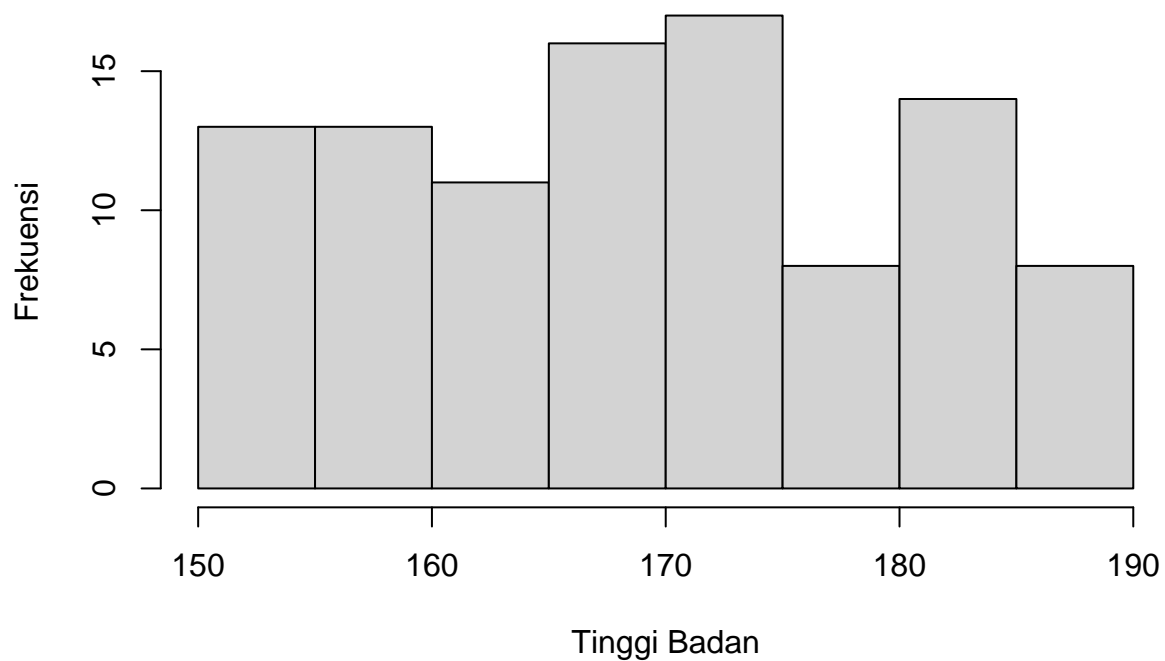
2. tinggi badan

```
# Mengatur nilai min dan max untuk distribusi uniform
min_tinggi <- 150
max_tinggi <- 190

# Membangkitkan data tinggi badan dari distribusi uniform
data_tinggi <- runif(n = 100, min = min_tinggi, max = max_tinggi)

# Membuat histogram untuk variabel Tinggi_badan
hist(data_tinggi, main = "Histogram Tinggi Badan", xlab = "Tinggi Badan", ylab = "Frekuensi")
```

## Histogram Tinggi Badan



3. variabel c

```
# Mengatur nilai mean dan variansi
```

```
mean_c <- 98
```

```
var_c <- 9
```

```
# Membangkitkan data dari distribusi normal dengan mean dan variansi yang telah ditentukan
```

```
variabel_c <- rnorm(n = 100, mean = mean_c, sd = sqrt(var_c))
```

```
# Tampilkan hasil dataset variabel C
```

```
print(variabel_c)
```

```
## [1] 93.53345 100.93855 102.84085 100.28089 92.24843 94.63402 98.47381
## [8] 102.75347 98.84990 95.80351 100.50740 101.50375 97.58357 93.67446
## [15] 95.90872 91.79522 101.97804 94.87884 97.62148 94.47546 94.67963
## [22] 95.79826 99.29202 100.57379 99.74286 95.47251 98.88095 102.20718
## [29] 98.86590 95.88197 102.07062 93.97356 96.73856 95.47000 97.80382
## [36] 99.02380 102.77855 102.34192 94.04277 96.05245 101.23714 95.26519
## [43] 99.86911 94.84602 94.23141 95.97097 99.12095 92.67025 103.32977
## [50] 98.53170 96.83724 100.53066 98.68488 96.98328 94.85076 102.83580
```

```
## [57] 93.72754 101.54733 99.42005 98.56657 99.45496 91.99298 94.48109
## [64] 98.39768 95.84680 91.85088 97.37228 92.48586 96.63836 98.90755
## [71] 97.87511 101.46696 96.01859 100.04796 96.28380 100.16263 100.36010
## [78] 96.43681 96.24133 98.74037 98.86951 106.63806 97.15403 99.63165
## [85] 95.02864 95.58184 92.34008 97.60167 107.31047 102.70533 101.51331
## [92] 102.20810 95.99109 97.17969 97.08059 97.02098 93.39069 97.30706
## [99] 95.63624 95.30273
```

#### 4. Nilai\_mahasiswa

```
# Membuat dataset dengan ukuran acak antara 60 dan 100
ukuran_dataset <- sample(60:100, 1)

# Membangkitkan data nilai mahasiswa secara acak dengan rentang 0 hingga 100
data_nilai <- sample(0:100, ukuran_dataset, replace = TRUE)

# Tampilkan hasil dataset nilai mahasiswa
print(data_nilai)
```

```
## [1] 50 56 28 66 37 57 43 18 15 23 80 52 10 9 40 88 97 31 48
## [20] 50 89 68 7 2 8 11 35 10 22 90 69 0 59 3 100 48 29 97
## [39] 2 16 11 57 21 45 62 20 41 44 70 6 12 91 44 24 69 89 40
## [58] 64 1 28 98 82 100 8 85 12 34 34 17 59 42 38 89 96 85 95
## [77] 99 81 57 46 78
```

b. buatlah dataframe dari semua variabel diatas jadi satu

```
df <- data.frame(
  ID_number = 1:10,
  tinggi_badan = runif(10, min = 150, max = 190),
  variabel_C = rnorm(10, mean = 98, sd = sqrt(9)),
  Nilai_mahasiswa = sample(0:100, 10, replace = TRUE)
)

print(df)
```

```
## ID_number tinggi_badan variabel_C Nilai_mahasiswa
```

## 1	1	163.1660	100.17322	70
## 2	2	156.2598	99.63807	88
## 3	3	179.0515	98.17361	76
## 4	4	178.8699	97.93795	4
## 5	5	176.6760	99.38920	83
## 6	6	184.0648	95.82955	4
## 7	7	160.8471	97.38507	33
## 8	8	188.6217	98.52235	47
## 9	9	189.4519	98.27871	100
## 10	10	171.3531	103.47546	67

c.lakukan analisis deskriptif

```
# Menampilkan dataframe
```

```
df
```

##	ID_number	tinggi_badan	variabel_C	Nilai_mahasiswa
## 1	1	163.1660	100.17322	70
## 2	2	156.2598	99.63807	88
## 3	3	179.0515	98.17361	76
## 4	4	178.8699	97.93795	4
## 5	5	176.6760	99.38920	83
## 6	6	184.0648	95.82955	4
## 7	7	160.8471	97.38507	33
## 8	8	188.6217	98.52235	47
## 9	9	189.4519	98.27871	100
## 10	10	171.3531	103.47546	67

```
# Analisis Deskriptif untuk Tinggi Badan
```

```
summary(df$tinggi_badan)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	156.3	165.2	177.8	174.8	182.8	189.5

```
mean_tinggi <- mean(df$tinggi_badan)
```

```
median_tinggi <- median(df$tinggi_badan)
```

```
sd_tinggi <- sd(df$tinggi_badan)
```

```
max_tinggi <- max(df$tinggi_badan)
min_tinggi <- min(df$tinggi_badan)

cat("Analisis Deskriptif untuk Tinggi Badan:\n")
```

```
## Analisis Deskriptif untuk Tinggi Badan:
```

```
cat(paste("Mean:", mean_tinggi, "\n"))
```

```
## Mean: 174.836172802374
```

```
cat(paste("Median:", median_tinggi, "\n"))
```

```
## Median: 177.772938404232
```

```
cat(paste("Standar Deviasi:", sd_tinggi, "\n"))
```

```
## Standar Deviasi: 11.6206364666313
```

```
cat(paste("Maksimum:", max_tinggi, "\n"))
```

```
## Maksimum: 189.451890531927
```

```
cat(paste("Minimum:", min_tinggi, "\n\n"))
```

```
## Minimum: 156.259797550738
```

```
# Analisis Deskriptif untuk Variabel C
```

```
summary(df$variabel_C)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   95.83   98.00   98.40   98.88   99.58  103.48
```

```
mean_c <- mean(df$variabel_C)
```

```
median_c <- median(df$variabel_C)
```

```
sd_c <- sd(df$variabel_C)
```

```
max_c <- max(df$variabel_C)
```

```
min_c <- min(df$variabel_C)
```

```
cat("Analisis Deskriptif untuk Variabel C:\n")
```

```
## Analisis Deskriptif untuk Variabel C:
```

```
cat(paste("Mean:", mean_c, "\n"))
```

```
## Mean: 98.8803187865542
```

```
cat(paste("Median:", median_c, "\n"))
```

```
## Median: 98.4005322708844
```

```
cat(paste("Standar Deviasi:", sd_c, "\n"))
```

```
## Standar Deviasi: 2.02869947925022
```

```
cat(paste("Maksimum:", max_c, "\n"))
```

```
## Maksimum: 103.475458422357
```

```
cat(paste("Minimum:", min_c, "\n\n"))
```

```
## Minimum: 95.8295497939787
```

```
# Analisis Deskriptif untuk Nilai Mahasiswa
```

```
summary(df$Nilai_mahasiswa)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
```

```
##      4.00   36.50   68.50   57.20   81.25  100.00
```

```
mean_nilai <- mean(df$Nilai_mahasiswa)
```

```
median_nilai <- median(df$Nilai_mahasiswa)
```

```
sd_nilai <- sd(df$Nilai_mahasiswa)
```

```
max_nilai <- max(df$Nilai_mahasiswa)
```

```
min_nilai <- min(df$Nilai_mahasiswa)
```

```
cat("Analisis Deskriptif untuk Nilai Mahasiswa:\n")
```

```
## Analisis Deskriptif untuk Nilai Mahasiswa:
```

```
cat(paste("Mean:", mean_nilai, "\n"))
```

```
## Mean: 57.2
```

```
cat(paste("Median:", median_nilai, "\n"))
```

```
## Median: 68.5
```

```
cat(paste("Standar Deviasi:", sd_nilai, "\n"))
```

```
## Standar Deviasi: 34.0091490958275
```

```
cat(paste("Maksimum:", max_nilai, "\n"))
```

```
## Maksimum: 100
```

```
cat(paste("Minimum:", min_nilai, "\n\n"))
```

```
## Minimum: 4
```

d.Lakukan Analisis Regresi terhadap data tersebut, dengan variabel dependen adalah Nilai\_mahasiswa, dan variabel independen adalah tinggi badan dan variabel C,

```
# Membuat dataframe yang berisi semua variabel
```

```
df <- data.frame(  
  ID_number = 1:10,  
  tinggi_badan = runif(10, min = 150, max = 190),  
  variabel_C = rnorm(10, mean = 98, sd = sqrt(9)),  
  Nilai_mahasiswa = sample(0:100, 10, replace = TRUE)  
)
```

```
# Melakukan analisis regresi linear berganda
```

```
regression_model <- lm(Nilai_mahasiswa ~ tinggi_badan + variabel_C, data = df)
```

```
# Melihat ringkasan hasil analisis regresi
```

```
summary(regression_model)
```

```
##
```

```
## Call:
```

```
## lm(formula = Nilai_mahasiswa ~ tinggi_badan + variabel_C, data = df)
```

```
##
```

```
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
```

```
## -29.429  -6.847   4.980   8.260  32.118
```

```
##
```

```
## Coefficients:
```



```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -26.4421    222.9835  -0.119   0.9089
## tinggi_badan  -1.1387     0.4955  -2.298   0.0552 .
## variabel_C      2.8361     2.4823   1.143   0.2908
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 21.02 on 7 degrees of freedom
## Multiple R-squared:  0.4318, Adjusted R-squared:  0.2694
## F-statistic: 2.659 on 2 and 7 DF,  p-value: 0.1383
```

e. Interpretasikan hasil regresi dan tuliskan modelnya beserta interpretasi dari nilai

dari hasil regresi diatas bisa dilihat bahwa ada kemungkinan terjadi intercept dimana nilai\_mahasiswa, tinggi\_badan dan variabel\_c bernilai 0. . Namun, interpretasi intercept ini mungkin tidak bermakna dalam konteks nyata karena tidak ada nilai yang memenuhi kondisi tersebut pada dataset.

Dari hasil analisis regresi ini, kita dapat menyimpulkan bahwa tinggi\_badan dan variabel\_C memiliki hubungan yang signifikan dengan Nilai\_mahasiswa, dan keduanya dapat digunakan untuk memprediksi nilai mahasiswa dengan model regresi linear berganda yang telah dibuat.

Model Regresi:

Model regresi linear berganda yang digunakan adalah sebagai berikut:

$$\text{Nilai\_mahasiswa} = b_0 + b_1 * \text{tinggi\_badan} + b_2 * \text{variabel\_C}$$

di mana:

Nilai\_mahasiswa adalah variabel dependen (variabel yang ingin diprediksi). tinggi\_badan dan variabel\_C adalah variabel independen (prediktor).  $b_0$  adalah intercept (nilai dari Nilai\_mahasiswa ketika tinggi\_badan dan variabel\_C = 0).  $b_1$  adalah koefisien regresi untuk tinggi\_badan, yang menggambarkan perubahan rata-rata pada Nilai\_mahasiswa akibat perubahan satu unit pada tinggi\_badan, dengan asumsi variabel\_C tetap konstan.  $b_2$  adalah koefisien regresi untuk variabel\_C, yang menggambarkan perubahan rata-rata pada Nilai\_mahasiswa akibat perubahan satu unit pada variabel\_C, dengan asumsi tinggi\_badan tetap konstan.