

TUGAS ANREG

Tubagus Achmad Aditya (G1401221006)

2024-03-05

Data

```
dataadit <- read.csv("C:\\Users\\TUBAGUS\\Downloads\\KULIAH  
ADIT\\anreg\\Anreg Individu.csv", sep = ";")  
head(dataadit)
```

```
##      X  Y  
## 1   2 54  
## 2   5 50  
## 3   7 45  
## 4  10 37  
## 5  14 35  
## 6  19 25  
## 7  26 20  
## 8  31 16  
## 9  34 18  
##10 38 13  
##11 45  8  
##12 52 11  
##13 53  8  
##14 60  4  
##15 65  6
```

Model

```
model.reg= lm(formula = Y ~ ., data = dataadit)  
model.reg
```

```
##  
## Call:  
## lm(formula = Y ~ ., data = dataadit)  
##  
## Coefficients:  
## (Intercept)          X  
##    46.4604      -0.7525
```

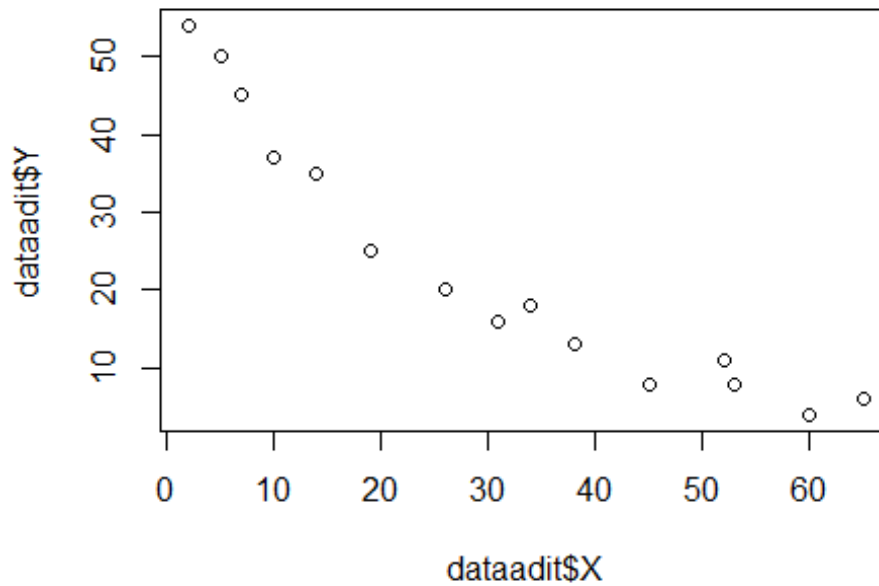
Model: $\hat{Y} = 46.4604 - 0.7525X + e$

Hasil diatas belum dipastikan menjadi model terbaik maka diperlukannya eksplorasi Gauss-Markov

Eksplorasi Kondisi Gauss-Markov

Plot X dan Y

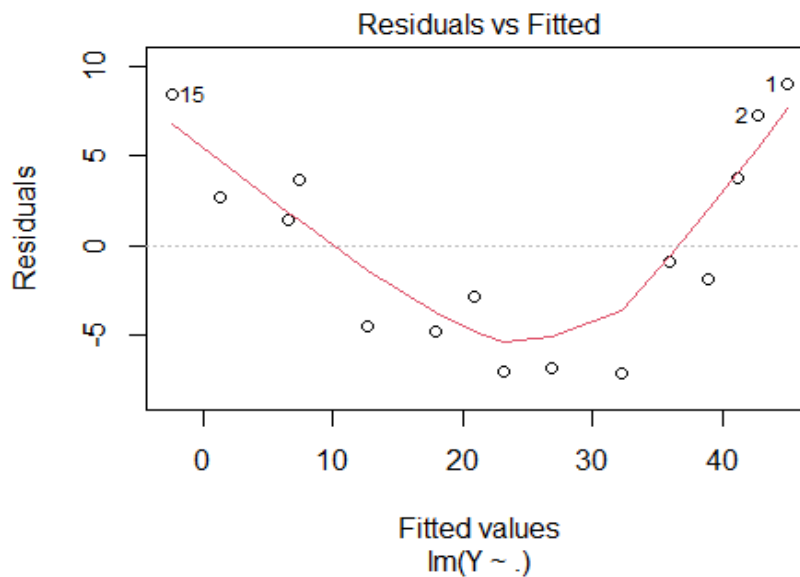
```
plot(x = dataadit$X, y = dataadit$Y)
```



Hasil plot variable X dan Y menggambarkan hubungan yang tidak linier dan cenderung membentuk parabola.

1. Plot Sisaan dan \hat{Y}

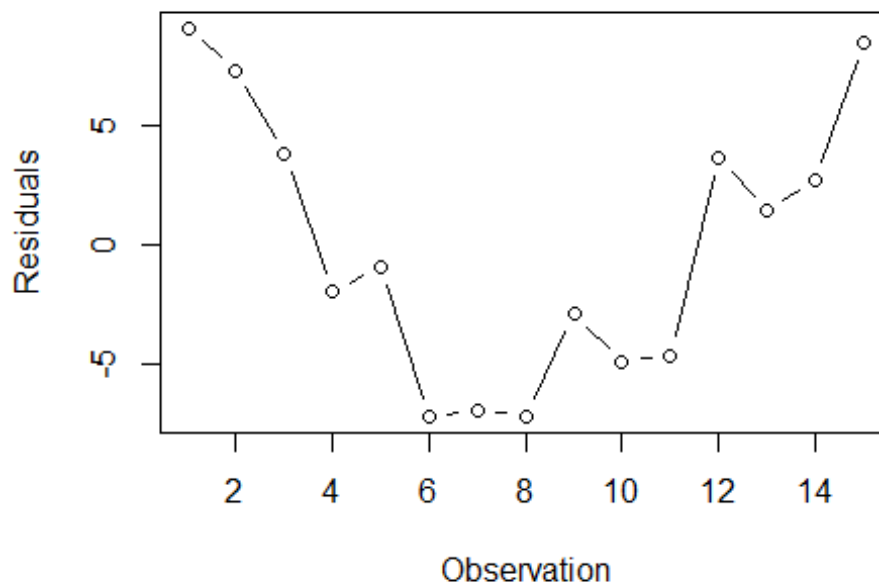
```
plot(model.reg, 1)
```



1. Sisaan menyebar di sekitar 0, sehingga nilai harapan galat sama dengan 0
2. Lebar pita sama untuk setiap nilai dugaan, sehingga ragam homogen
3. Plot sisaan vs y duga membentuk pula kurva, sehingga model tidak pas dan perlu transformasi terhadap variabel

2. Plot Sisaan vs Urutan

```
plot(x = 1:dim(dataadit)[1],
     y = model.reg$residuals,
     type = 'b',
     ylab = "Residuals",
     xlab = "Observation")
```



Tebaran berpola maka sisaan saling tidak bebas, model tidak pas.

Uji Formal Asumsi

p-value < 0.05 tolak H)

H1: Nilai harapan sisaan tidak sama dengan nol

1. Nilai harapan sisaan sama dengan nol

```
t.test(model.reg$residuals,mu = 0,conf.level = 0.95)
```

```
##
## One Sample t-test
##
## data: model.reg$residuals
## t = -4.9493e-16, df = 14, p-value = 1
```

```
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## -3.143811 3.143811
## sample estimates:
## mean of x
## -7.254614e-16
```

p-value > 0, maka tidak tolak H_0 , nilai harapan sisaan sama dengan 0 pada taraf nyata 5%. Asumsi terpenuhi.

2. Ragam sisaan homogen

$H_0 : \text{var}[e] = \sigma^2$ (ragam sisan homogen)

$H_1 : \text{var}[e] \neq \sigma^2$ (ragam sisan tidak homogen)

```
cek.homogen = lm(formula = abs(model.reg$residuals) ~ X, # y: abs residual
  data = dataadit)
summary(cek.homogen)

##
## Call:
## lm(formula = abs(model.reg$residuals) ~ X, data = dataadit)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4.2525 -1.7525  0.0235  2.0168  4.2681
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.45041     1.27241   4.284  0.00089 ***
## X           -0.01948     0.03456  -0.564  0.58266
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.714 on 13 degrees of freedom
## Multiple R-squared:  0.02385,    Adjusted R-squared:  -0.05124
## F-statistic: 0.3176 on 1 and 13 DF,  p-value: 0.5827

library(lmtest)

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##      as.Date, as.Date.numeric

bptest(model.reg)
```

```
##
## studentized Breusch-Pagan test
##
## data: model.reg
## BP = 0.52819, df = 1, p-value = 0.4674
```

```
library(car)
```

```
## Loading required package: carData
```

```
ncvTest(model.reg)
```

```
## Non-constant Variance Score Test
## Variance formula: ~ fitted.values
## Chisquare = 0.1962841, Df = 1, p = 0.65774
```

p-value > alpha, maka tak tolak H_0 , ragam sisaan homogen pada taraf nyata 5%. Asumsi terpenuhi.

3. Sisaan saling bebas

$H_0 : E[e_i, e_j] = 0$ (sisaan saling bebas/tidak ada autokorelasi)

$H_1 : E[e_i, e_j] \neq 0$ (sisaan tidak saling bebas/ada autokorelasi)

```
library(randtests)
```

```
runs.test(model.reg$residuals)
```

```
##
## Runs Test
##
## data: model.reg$residuals
## statistic = -2.7817, runs = 3, n1 = 7, n2 = 7, n = 14, p-value =
## 0.005407
## alternative hypothesis: nonrandomness
```

```
library(lmtest)
```

```
dwtest(model.reg)
```

```
##
## Durbin-Watson test
##
## data: model.reg
## DW = 0.48462, p-value = 1.333e-05
## alternative hypothesis: true autocorrelation is greater than 0
```

Uji ini sering disebut dengan uji autokorelasi yang dilakukan dengan Durbin_watson. Karena p-value = 1.333e-05 (pada DW test) < alpha = 0.05, maka tolak H0, sisaan tidak saling bebas pada taraf nyata 5%, sehingga asumsi tidak terpenuhi. Dibuktikan pula pada eksplorasi sisaan bahwa nilai autokorelasi pada lag 1 bernilai 0.5 dan 0.4 pada lag 2 yang berada di luar batas kepercayaan 95%, autokorelasi signifikan.

Uji Formal Normalitas Sisaan

H0: Sisaan menyebar normal

H1: Sisaan tidak menyebar normal

```
ks.test(model.reg$residuals, "pnorm", mean=mean(model.reg$residuals),
sd=sd(model.reg$residuals))

##
## Exact one-sample Kolmogorov-Smirnov test
##
## data: model.reg$residuals
## D = 0.12432, p-value = 0.9521
## alternative hypothesis: two-sided

library(car)
shapiro.test(model.reg$residuals)

##
## Shapiro-Wilk normality test
##
## data: model.reg$residuals
## W = 0.92457, p-value = 0.226
```

p-value > alpha, maka tak tolak H0, sisaan menyebar normal pada taraf nyata 5%

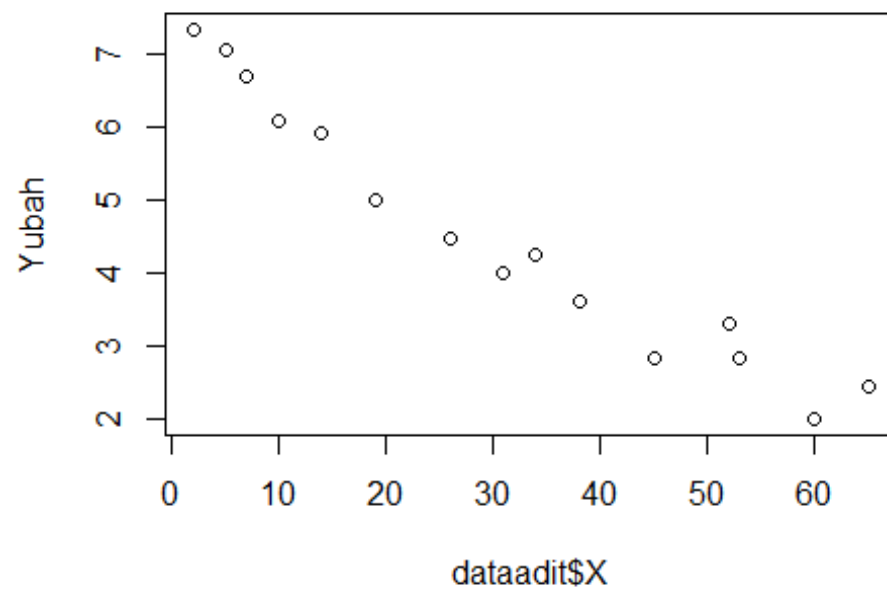
Penyesuaian Data

Transformasi Data

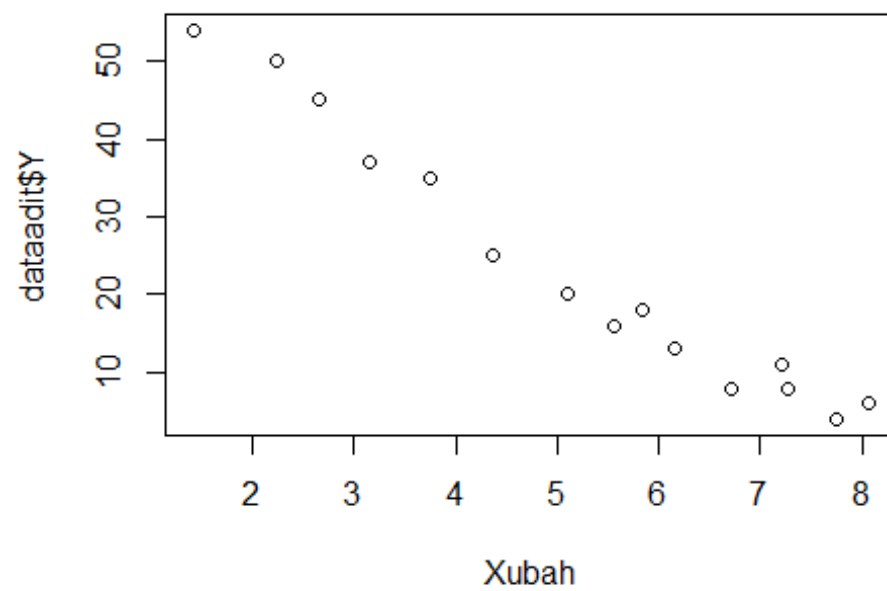
```
Yubah = sqrt(dataadit$Y)
```

```
Xubah = sqrt(dataadit$X)
```

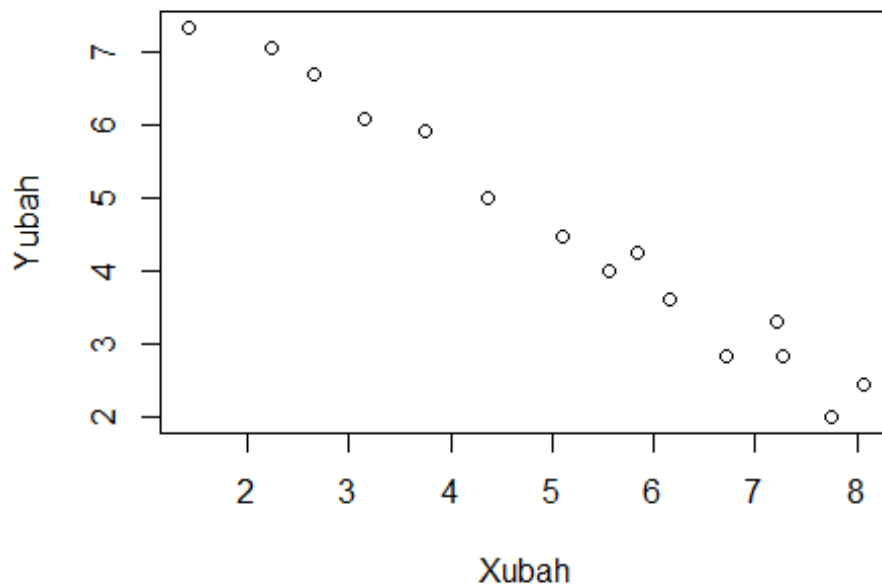
```
plot(x = dataadit$X, y = Yubah)
```



```
plot(x = Xubah, y = dataadit$Y)
```



```
plot(x = Xubah, y = Yubah)
```



Karena hubungan X dan Y cenderung membentuk sebuah parabola dan nilai $B1 < 0$, maka data dapat ditransformasi dengan mengecilkan nilai X dan/atau Y dengan membentuknya menjadi pangkat setengah atau akar dari data asli. Terdapat perbedaan antara hasil plot hubungan Xubah dengan Y, X dengan Yubah, dan Xubah dengan Yubah sehingga perlu ditelusuri lebih lanjut untuk memperoleh model terbaik melalui pemeriksaan asumsi pada data dengan sisaan paling bebas.

Model dan Pemeriksaan Asumsi

1. Xubah dengan Y

```
model1 = lm(formula = dataadit$Y ~ Xubah)
summary(model1)
```

```
##
## Call:
## lm(formula = dataadit$Y ~ Xubah)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-4.4518	-2.8559	0.7657	2.0035	5.2422

```
##
## Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	63.2250	2.2712	27.84	5.67e-13 ***
Xubah	-7.7481	0.4097	-18.91	7.68e-11 ***

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```



```
## Residual standard error: 3.262 on 13 degrees of freedom
## Multiple R-squared:  0.9649, Adjusted R-squared:  0.9622
## F-statistic: 357.7 on 1 and 13 DF,  p-value: 7.684e-11
```

Model: $\hat{Y} = 63.2250 - 7.7481X + e$

```
dwtest(model1)
```

```
##
## Durbin-Watson test
##
## data:  model1
## DW = 1.1236, p-value = 0.01422
## alternative hypothesis: true autocorrelation is greater than 0
```

Karena p-value < alpha, maka tolak H0, sisaan tidak saling bebas, asumsi tidak terpenuhi, bukan model terbaik

2. X dengan Yubah

```
model2 = lm(formula = Yubah ~ dataadit$X)
summary(model2)
```

```
##
## Call:
## lm(formula = Yubah ~ dataadit$X)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.53998 -0.38316 -0.01727  0.36045  0.70199
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  7.015455   0.201677   34.79 3.24e-14 ***
## dataadit$X  -0.081045   0.005477  -14.80 1.63e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4301 on 13 degrees of freedom
## Multiple R-squared:  0.9439, Adjusted R-squared:  0.9396
## F-statistic: 218.9 on 1 and 13 DF,  p-value: 1.634e-09
```

Model: $\hat{Y} = 7.015455 - 0.081045X + e$

```
dwtest(model2)
```

```
##
## Durbin-Watson test
##
## data:  model2
## DW = 1.2206, p-value = 0.02493
## alternative hypothesis: true autocorrelation is greater than 0
```

Karena p-value (pada DW test) < alpha = 0.05, maka tolak H0, sisaan tidak saling bebas, asumsi tidak terpenuhi, bukan model terbaik

3. Xubah dengan Yubah

```
model3 = lm(formula = Yubah ~ Xubah)
summary(model3)

##
## Call:
## lm(formula = Yubah ~ Xubah)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.42765 -0.17534 -0.05753  0.21223  0.46960
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  8.71245     0.19101   45.61 9.83e-16 ***
## Xubah        -0.81339     0.03445  -23.61 4.64e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2743 on 13 degrees of freedom
## Multiple R-squared:  0.9772, Adjusted R-squared:  0.9755
## F-statistic: 557.3 on 1 and 13 DF, p-value: 4.643e-12
```

Model: $\hat{Y} = 8.71245 - 0.81339X + e$

```
dwtest(model3)

##
## Durbin-Watson test
##
## data:  model3
## DW = 2.6803, p-value = 0.8629
## alternative hypothesis: true autocorrelation is greater than 0
```

Karena p-value (pada DW test) > alpha = 0.05, maka tak tolak H0, sisaan saling bebas.

Kesimpulan:

Berdasarkan hasil diatas, model terbaik didapatkan ketika X dan Y keduanya ditransformasikan ke dalam bentuk akar atau pangkat 1/2 dan memenuhi semua asumsi analisis regresi linier sederhana. Sehingga model regresi untuk kasus ini adalah

$$\hat{Y}^{1/2} = 8.71245 - 0.81339X^{1/2} + e$$

Jika ingin mengembalikan model untuk menjelaskan peubah respons sebelum ditransformasi, perlu dilakukan transformasi balik pada model yang dibentuk. Prosedur

transformasi balik dilakukan dengan proses matematika biasa dengan mengutamakan pengubahan peubah respons terlebih dahulu.

$$\hat{Y} = (8.71245 - 0.81339X^{1/2} + e)^2$$

Interpretasi terhadap model tersebut menunjukkan hubungan yang terbalik antara Y dengan $X^{1/2}$ sebagai hubungan kuadratik. Semakin besar nilai $X^{1/2}$ semakin kecil nilai dugaan rata-rata Y . Ketika $X^{1/2}$ sama dengan nol dan berada pada selang amatan, nilai dugaan rata-rata Y sebesar 8.71245 dan setiap kenaikan 1 satuannya akan menurunkan nilai Y dugaan rata-rata sebesar 0.81339.