

Tugas Penanganan Kondisi Tak Standar dan Transformasi

Sintong M.N Purba - G1401221077

2024-03-05

Pembentukan Model Awal

Pembacaan Data

Data yang diberikan berisi 2 peubah X dan Y yang terdiri atas 15 amatan. Data terlebih dahulu dimasukkan kedalam excel kemudian dibaca menggunakan R.

```
library(readxl)
data_model <- read_excel("C:/Users/Poerba/Documents/Data_Anreg.xlsx")
X=data_model$X
Y=data_model$Y
data_model= data.frame(X,Y)
data_model
```

```
##      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 Awal

Dibentuk model menggunakan fungsi lm sehingga didapatkan parameter $\hat{\beta}_0$ atau intersep dan $\hat{\beta}_1$ atau slope dari garis regresi

```
model=lm(formula = Y~X, data=data_model)
summary(model)
```

```
##
## Call:
## lm(formula = Y ~ X, data = data_model)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.1628 -4.7313 -0.9253  3.7386  9.0446
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 46.46041    2.76218   16.82 3.33e-10 ***
## X           -0.75251    0.07502  -10.03 1.74e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.891 on 13 degrees of freedom
## Multiple R-squared:  0.8856, Adjusted R-squared:  0.8768
## F-statistic: 100.6 on 1 and 13 DF,  p-value: 1.736e-07
```

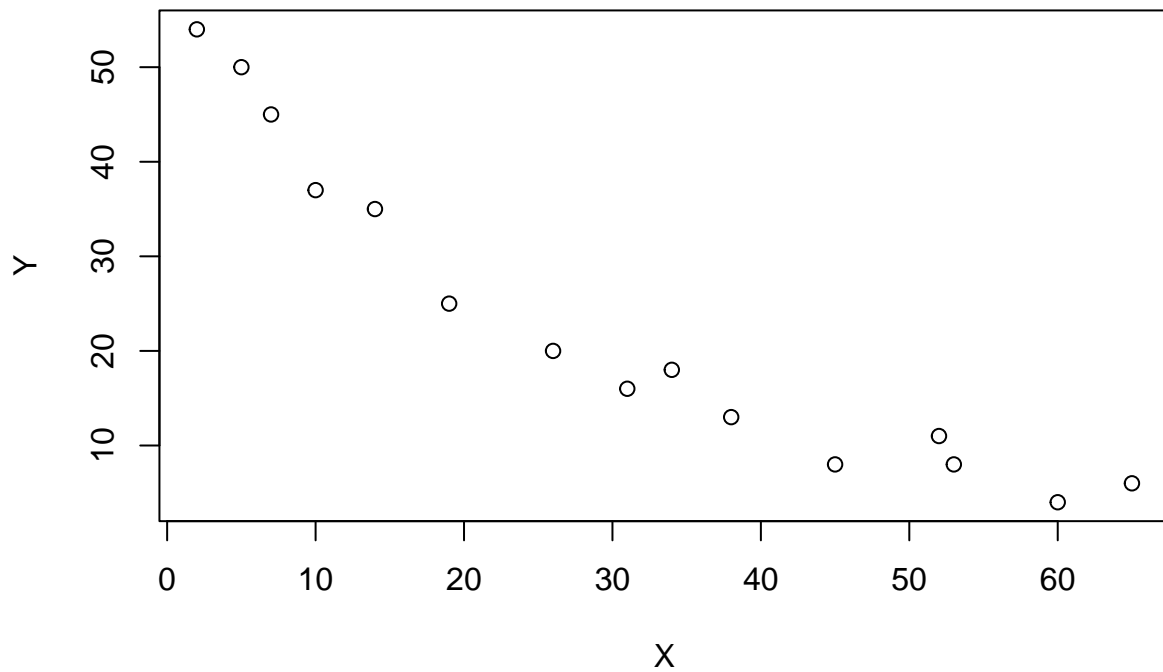
Model yang didapatkan adalah

$$\hat{Y} = 46.46041 - 0.75251X$$

Didapatkan juga nilai Adjusted R-squared: 0.8768

Scatterplot

```
plot(X,Y)
```

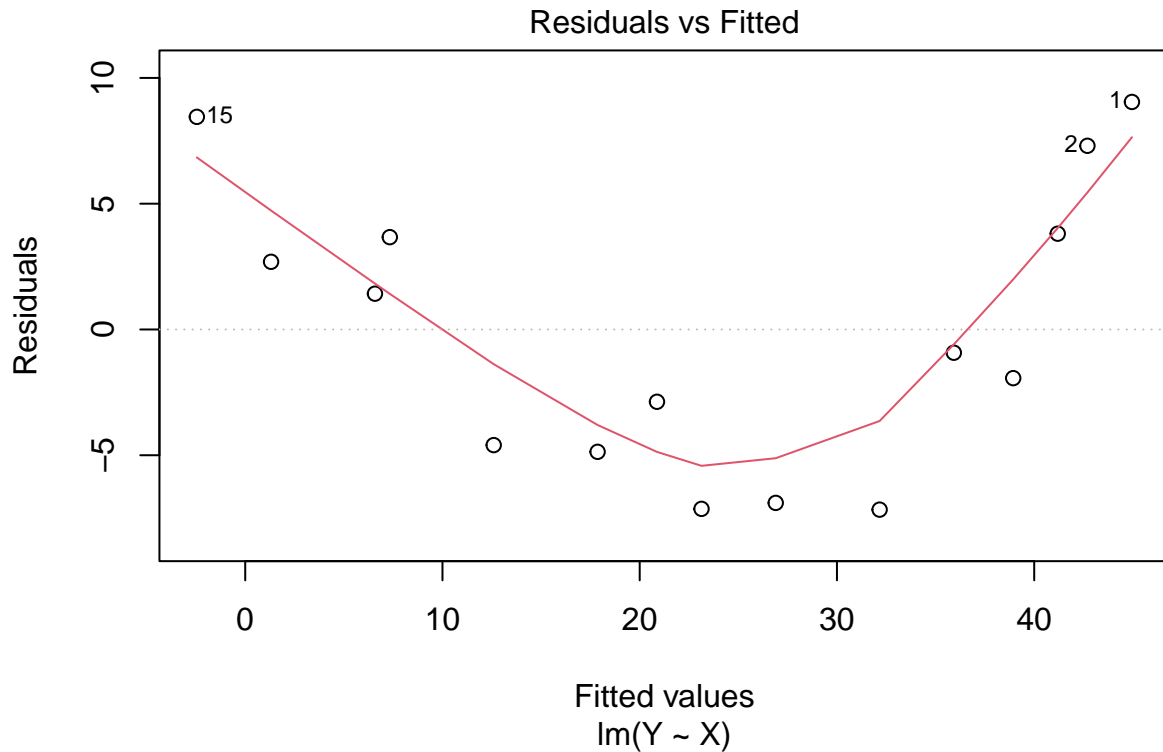


Pada scatter plot antara X dan Y dapat dilihat bahwa data tidak mengikuti garis linear, melainkan sedikit melengkung dimana pada model didapatkan bahwa $\hat{\beta}_1 < 0$, sehingga perlu transformasi dengan memperkecil variabel Y maupun variabel X.

Pengujian Asumsi

Pengujian nilai harapan $\text{erorr} = 0$

```
plot(model,1)
```



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

```
##
## One Sample t-test
##
## data: model$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
```

Karena nilai $p\text{-value} > 0.05$ maka asumsi $E[e] = 0$ terpenuhi

Pengujian Homoscedasticity

Uji ini memiliki hipotesis sebagai berikut,

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

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

```
library(lmtest)
```

```
## Warning: package 'lmtest' was built under R version 4.3.2
```

```
## Loading required package: zoo
```

```
## Warning: package 'zoo' was built under R version 4.3.2
```

```
##
```

```
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      as.Date, as.Date.numeric
```

```
bptest(model)
```

```
##
```

```
## studentized Breusch-Pagan test
```

```
##
```

```
## data: model
```

```
## BP = 0.52819, df = 1, p-value = 0.4674
```

Karena nilai p-value > 0.05 maka tak tolak H_0 , artinya tidak cukup bukti untuk menyatakan bahwa ragam sisaan tidak homogen atau asumsi Homoscedasticity terpenuhi

Pengujian Autokorelasi Sisaan

Autokorelasi sisaan dideteksi menggunakan uji Durbin-Watson dengan hipotesis :

$H_0 : E[e_i, e_j] = 0$ (Tidak ada Autokorelasi)

$H_1 : E[e_i, e_j] \neq 0$ (Ada Autokorelasi)

```
(uji_autokol <- dwtest(model,
                        alternative="two.sided"))
```

```
##
```

```
## Durbin-Watson test
```

```
##
```

```
## data: model
```

```
## DW = 0.48462, p-value = 2.665e-05
```

```
## alternative hypothesis: true autocorrelation is not 0
```

```
ifelse(uji_autokol$p < 0.05, "Ada Autokorelasi", "Tidak Ada Autokorelasi")
```

```
## [1] "Ada Autokorelasi"
```

Dapat dilihat bahwa p-value < 0.05 sehingga tolak H_0 yang mengindikasikan terdapat autokorelasi dalam sisaan

Uji Normal

```
library(nortest)
sisaan_model <- resid(model)
(norm_model <- lillie.test(sisaan_model))

##
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  sisaan_model
## D = 0.12432, p-value = 0.7701
```

Pada uji Lilliefors (Kolmogorov-Smirnov) saat $p\text{-value} > 0.05$ maka sisaan akan menyebar normal, sehingga asumsi ini terpenuhi.

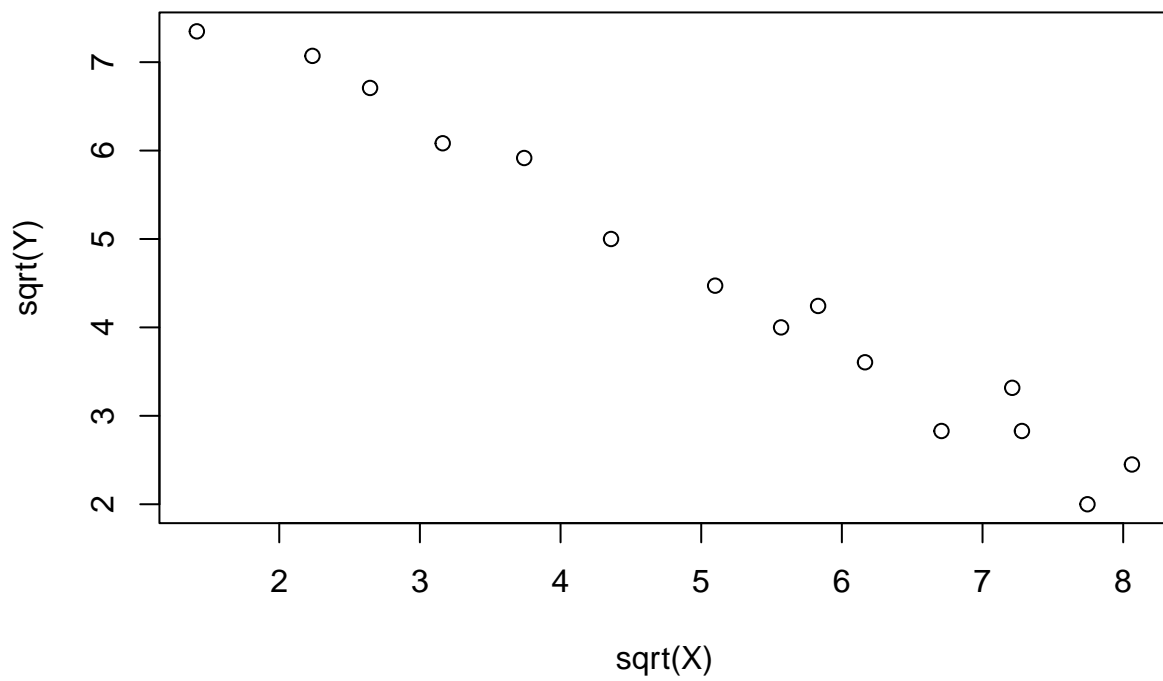
Transformasi

Transformasi yang digunakan adalah memperkecil nilai Y dengan menjadikannya \sqrt{Y} sehingga dibentuk model linearnya menggunakan fungsi lm

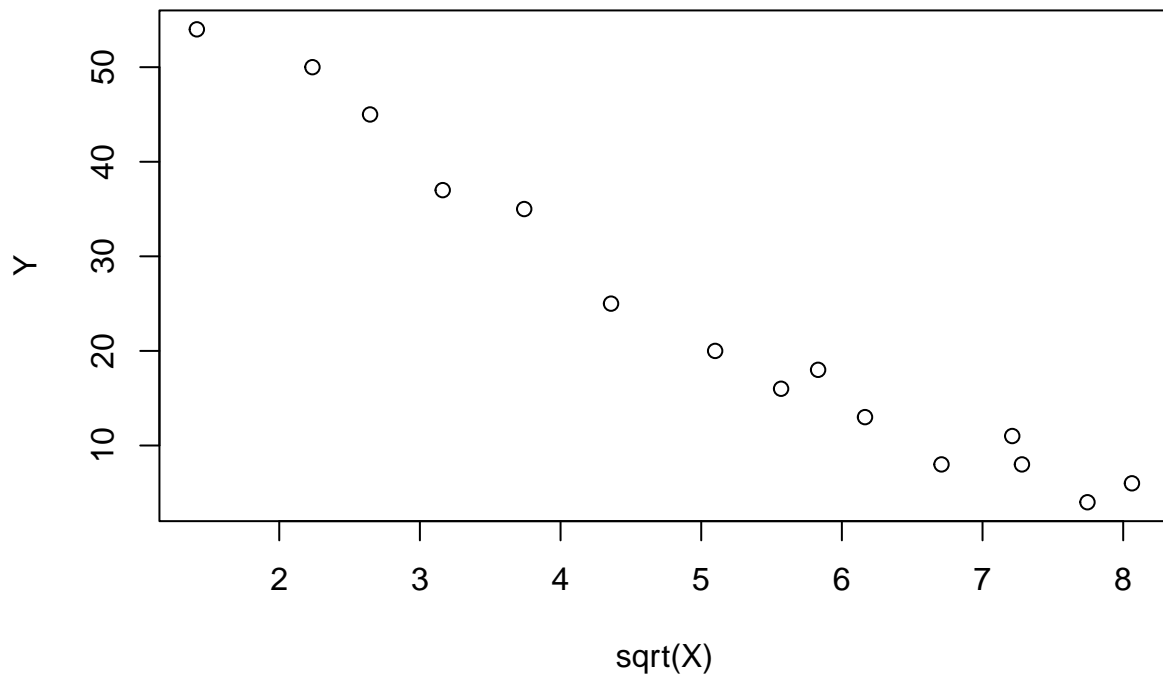
```
Model_Trans = lm(formula = Y~sqrt(X), data = data_model)
summary(Model_Trans)

##
## Call:
## lm(formula = Y ~ sqrt(X), data = data_model)
##
## 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 ***
## sqrt(X)       -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

plot(sqrt(X),sqrt(Y))
```



```
plot(sqrt(X),Y)
```

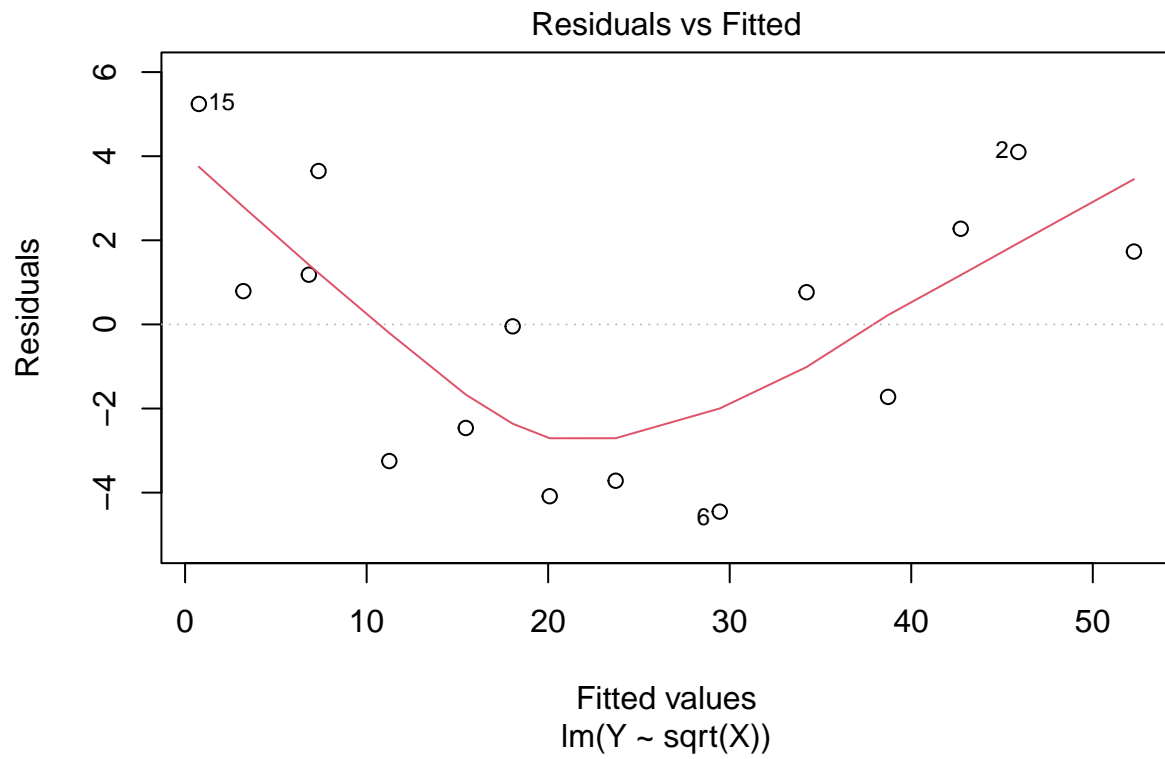


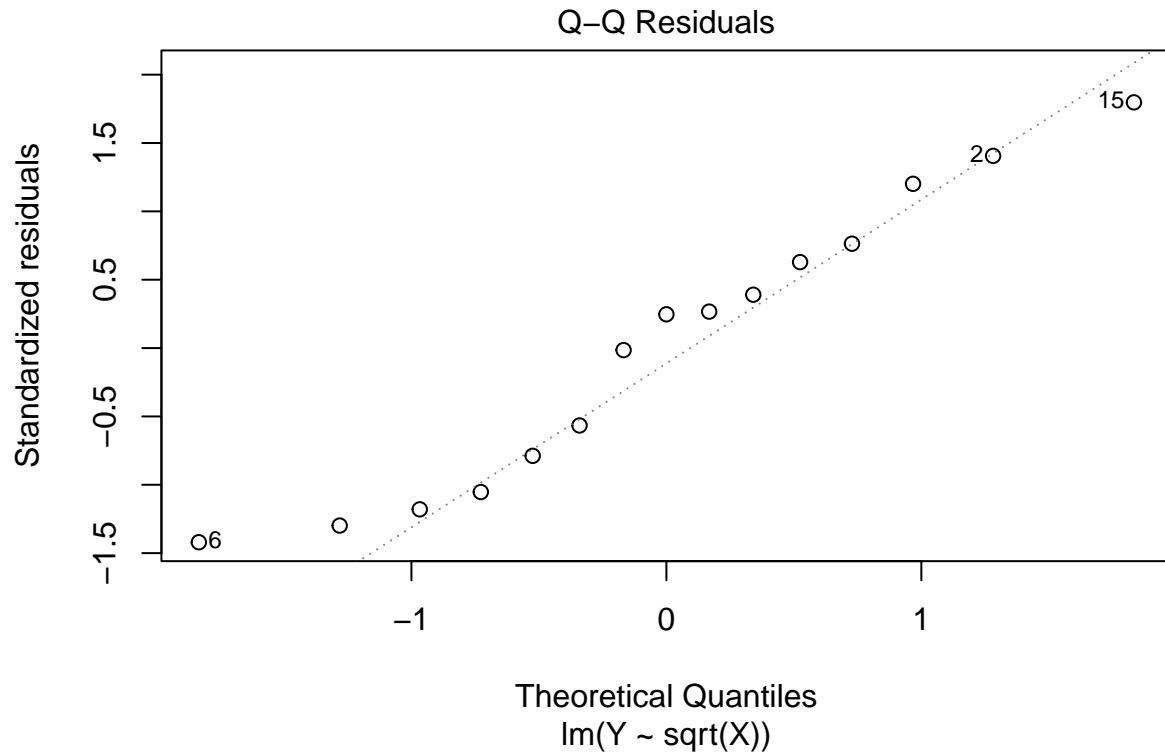
Model yang dihasilkan melalui transformasi adalah

$$\hat{Y} = 63.2250 - 7.7481\sqrt{X}$$

Dari model didapatkan nilai Adjusted R-Squared = 0.9622 yang lebih baik dibandingkan sebelum transformasi sehingga model lebih baik dalam menjelaskan keragaman dari peubah Y. dari plot antara \sqrt{X} juga terlihat data sudah linear tidak membentuk pola melengkung seperti sebelumnya.

```
plot(Model_Trans, 1:2)
```



Sisaan di sekitar 0 → Nilai harapan galat sama dengan nol

Lebar pita hampir sama untuk setiap nilai dugaan → ragam homogen

Pada QQplot juga terlihat data menyebar normal disekitar garis.

```
library(lmtest)
bptest(Model_Trans)
```

```
##
## studentized Breusch-Pagan test
##
## data: Model_Trans
## BP = 0.49234, df = 1, p-value = 0.4829
```

```
(uji_autokol <- dwtest(Model_Trans,
                        alternative="two.sided"))
```

```
##
## Durbin-Watson test
##
## data: Model_Trans
## DW = 1.1236, p-value = 0.02845
## alternative hypothesis: true autocorrelation is not 0
```

```
ifelse(uji_autokol$p < 0.05, "Ada Autokorelasi", "Tidak Ada Autokorelasi")
```

```
## [1] "Ada Autokorelasi"
```

Interpretasi :

β_0 = saat X bernilai 0 maka dugaan rata-rata dari nilai Y adalah sebesar 63.2250.

β_1 = Saat nilai akar X bertambah 1 satuan maka nilai dugaan dari rata-rata Y akan berkurang sebesar 7.7481.