13_trend_model_estimating_and_residuals.R

felixreichel

2021-11-05

```
# Course: Time series analysis
# Exercise: 13th / Fitting linear and quadratic trend models,
#           estimating and residual analysis
# Author: Felix Reichel
```

```
require(astsa)
```

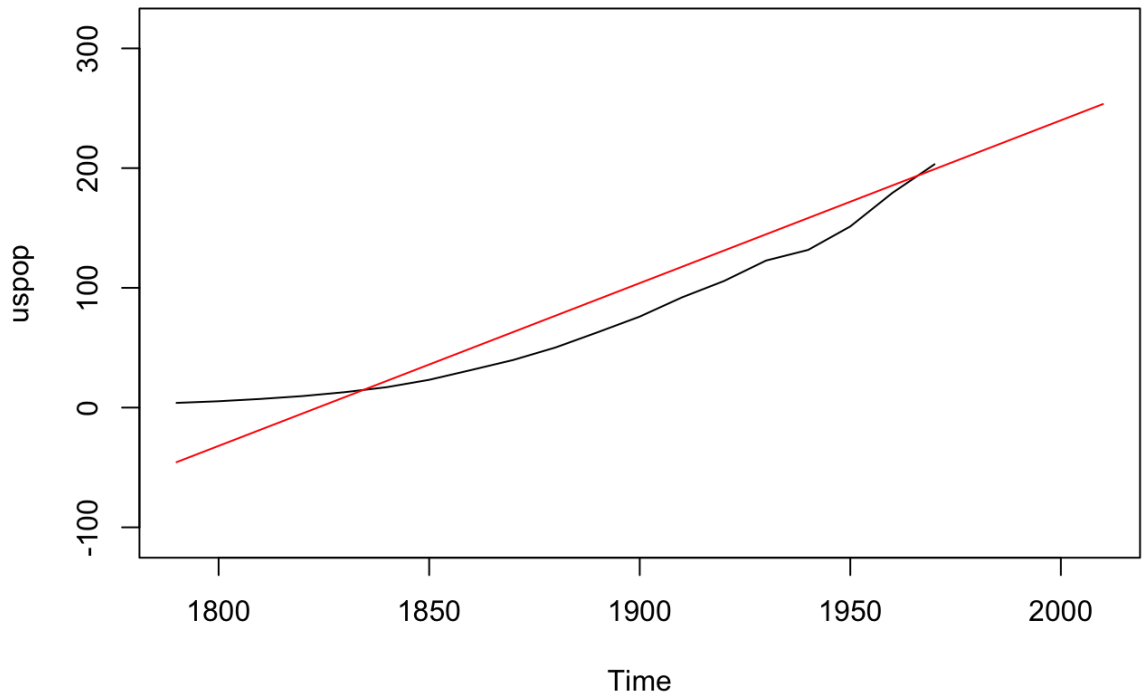
```
## Loading required package: astsa
```

```
require(tseries)
```

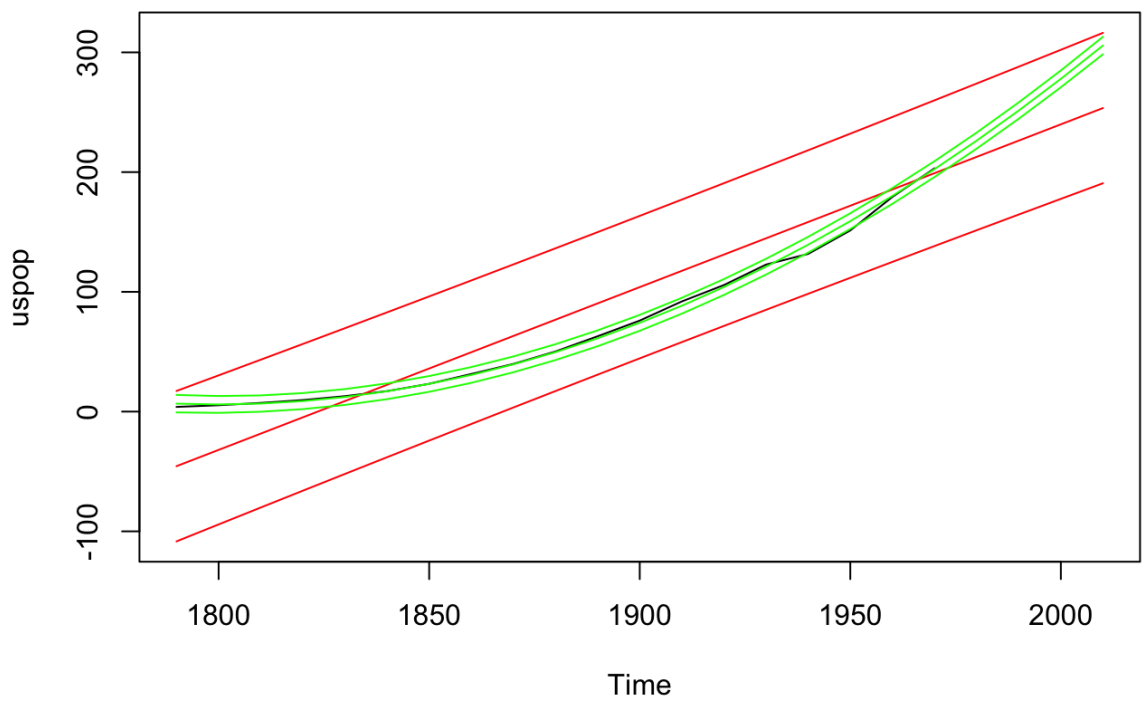
```
## Loading required package: tseries
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```
# set path
# current_path = rstudioapi::getActiveDocumentContext()$path
# setwd(dirname(current_path))

# uspop_1971_2020 <- c(data =
#  read.csv(file = "13_data/uspop_worldbank_1971_2020_Data.csv")[1,])[5:54]

# uspop2_world_bank_data <- c(
#  round(as.numeric(uspop_1971_2020[10])*10^-6, 1),
#  round(as.numeric(uspop_1971_2020[20])*10^-6, 1),
#  round(as.numeric(uspop_1971_2020[30])*10^-6, 1),
#  round(as.numeric(uspop_1971_2020[40])*10^-6, 1)
# )
uspop2 <- c(226.5, 248.7, 281.4, 308.7)

# add observations to existing time series uspop
uspop_mod <- ts(data = c(uspop, uspop2), start = start(uspop), frequency = frequency(uspop))

# linear trend model
fc_steps <- 2
Tf <- length(uspop_mod) + fc_steps
t <- (time(uspop_mod) - start(uspop_mod)) * frequency(uspop_mod) + 1

linear_model <- lm(uspop_mod ~ t)
newdata_lm = data.frame(t=1:Tf)

lm_fc = predict(linear_model, newdata = newdata_lm, interval="predict")
lm_fc.ts = ts(lm_fc, start = start(uspop_mod), end = end(uspop_mod), frequency = frequency(uspo
p_mod))

# plot uspop and linear trend model forecast
plot(uspop, xlim=c(start(uspop),end(uspop_mod)),ylim=c(min(lm_fc.ts),max(lm_fc.ts)))
lines(lm_fc.ts[,1],col="red")
```



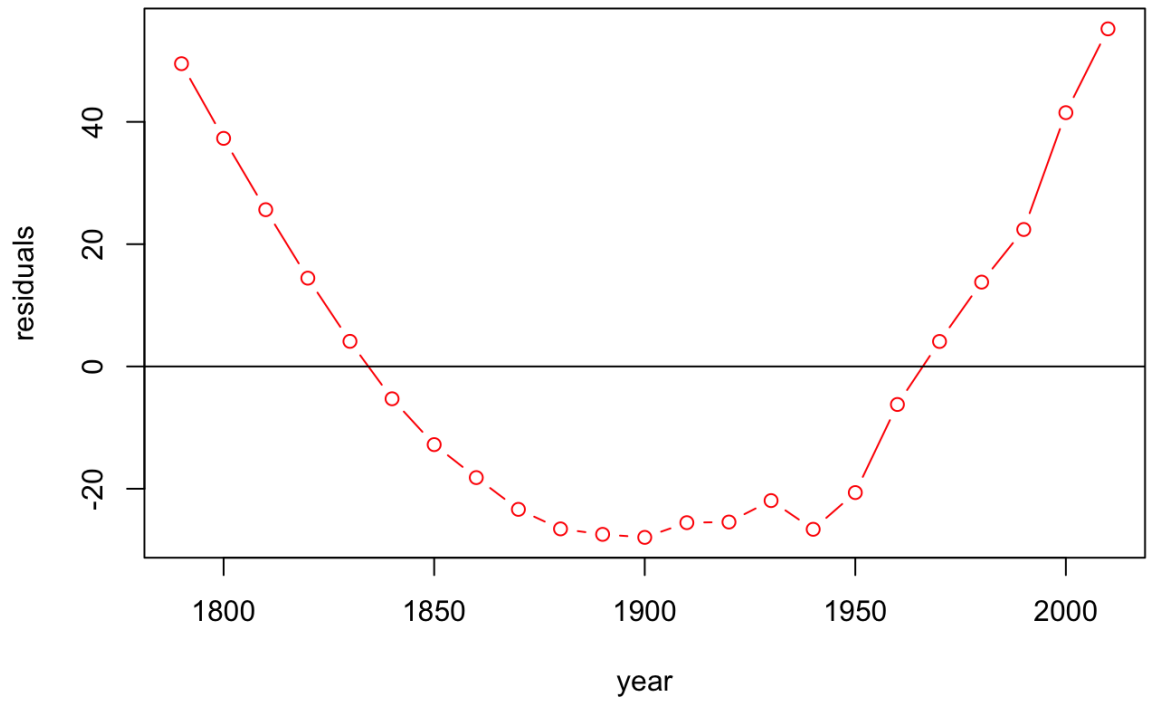

```
# quadratic trend model
tq <- t^2
quadratic_model <- lm(uspop_mod ~ t+tq)
newdata_lm_q <- data.frame(t=1:Tf, tq=(1:Tf)^2)

lm_q_fc <- predict(quadratic_model, newdata= newdata_lm_q, interval="predict")
lm_q_fc.ts <- ts(lm_q_fc, start=start(uspop), end=end(uspop_mod), frequency=frequency(uspop))

plot(uspop,
     xlim=c(start(uspop),end(uspop_mod)),
     ylim=c(min(lm_q_fc.ts,lm_fc.ts),max(lm_q_fc.ts,lm_fc.ts)))
lines(lm_fc.ts[,1],col="red") # fit
lines(lm_fc.ts[,2],col="red") # lwr
lines(lm_fc.ts[,3],col="red") # upr
lines(lm_q_fc.ts[,1],col="green") # fit
lines(lm_q_fc.ts[,2],col="green") # lwr
lines(lm_q_fc.ts[,3],col="green") # upr
```




```
# residual analysis
linear_model$residuals
```

```
##          1          2          3          4          5          6          7
##  49.514891  37.300534  25.636176  14.441818   4.107460  -5.286897 -12.781255
##          8          9         10         11         12         13         14
## -18.175613 -23.369970 -26.564328 -27.458686 -27.953043 -25.547401 -25.441759
##         15         16         17         18         19         20         21
## -21.936117 -26.630474 -20.624832  -6.219190   4.086453  13.792095  22.397737
##         22         23
##  41.503379  55.209022
```

```
plot(matrix(time(uspop_mod)),linear_model$residuals, ylab="residuals",xlab="year",
     type="b",col="red")
abline(h=mean(linear_model$residuals))
```




```
quadratic_model$residuals
```

```
##          1          2          3          4          5          6          7
## -2.6941304 -0.6696640  0.5487239  0.8810333  0.7172643  0.1374167  0.1014907
##          8          9         10         11         12         13         14
##  0.8094862  0.3614032  0.5572417  1.6970017  1.8806832  3.6082863  1.6798108
##         15         16         17         18         19         20         21
##  1.7952569 -7.6453755 -7.7420864 -0.7948758  0.6962564  0.2313100 -2.6897149
##         22         23
##  3.5331818  3.0000000
```

```
plot(matrix(time(uspop_mod)),quadratic_model$residuals, ylab="residuals",xlab="year",
     type="b",col="green")
abline(h=mean(quadratic_model$residuals))
```


```
#autocorrelations
require(lmtest)
```

```
## Loading required package: lmtest
```

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

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

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

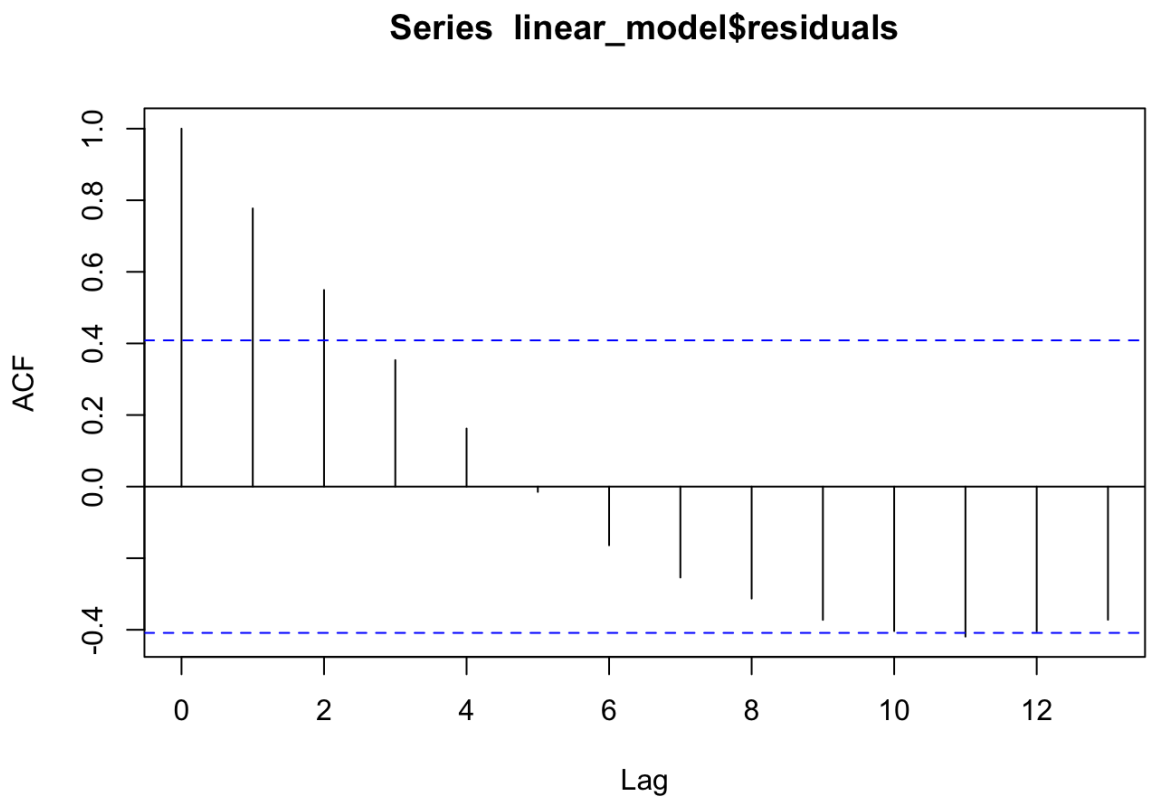
```
dwtest(linear_model) #one sided test: rho(1)>0
```

```
##
##  Durbin-Watson test
##
## data:  linear_model
## DW = 0.11174, p-value < 2.2e-16
## alternative hypothesis: true autocorrelation is greater than 0
```

```
dwtest(quadratic_model) #one sided test: rho(1)>0
```

```
##
##  Durbin-Watson test
##
## data:  quadratic_model
## DW = 1.098, p-value = 0.002006
## alternative hypothesis: true autocorrelation is greater than 0
```

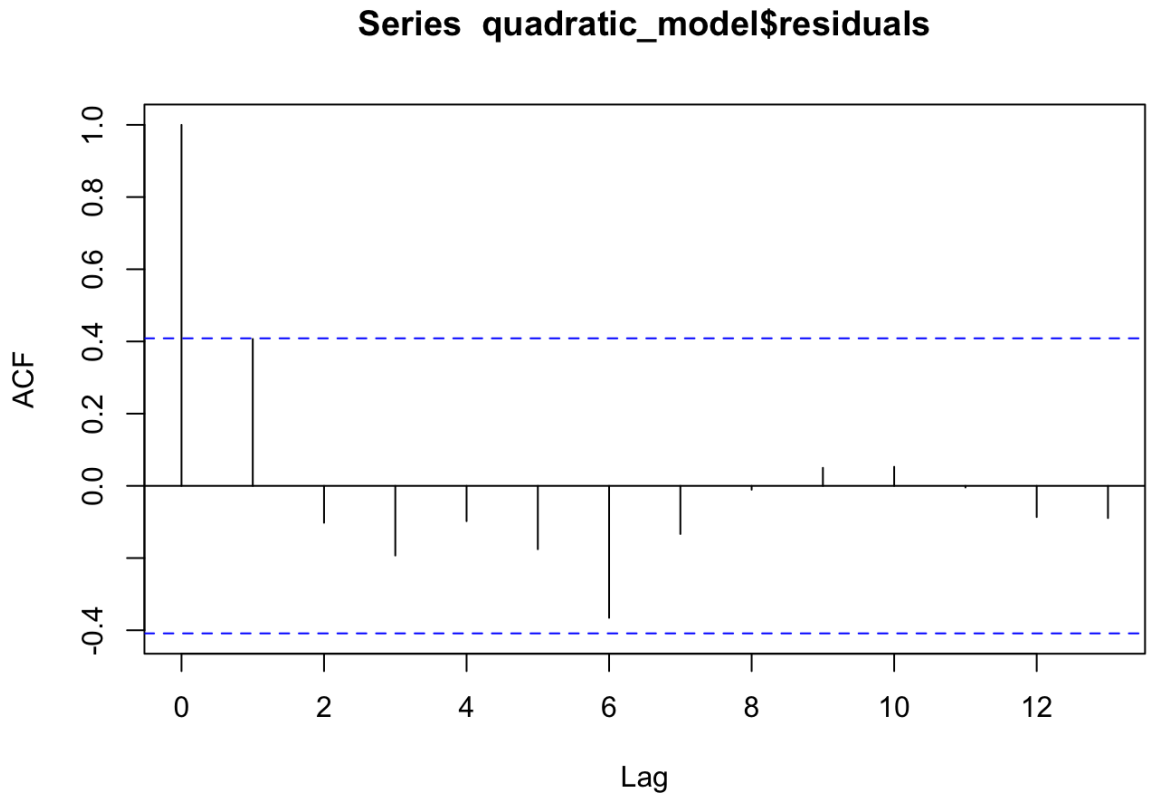
```
acf(linear_model$residuals)
```




```
Box.test(linear_model$residuals, lag = 3, type = "Ljung")
```

```
##
##  Box-Ljung test
##
## data:  linear_model$residuals
## X-squared = 27.636, df = 3, p-value = 4.33e-06
```

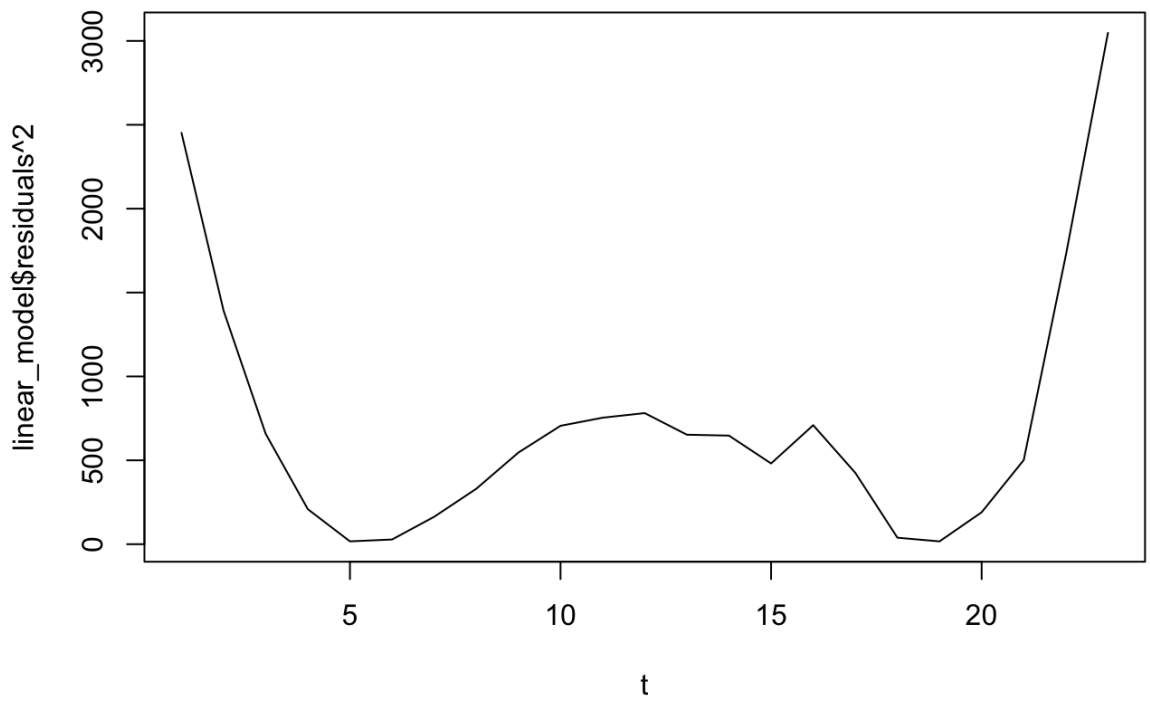
```
acf(quadratic_model$residuals)
```




```
Box.test(quadratic_model$residuals, lag = 3, type = "Ljung")
```

```
##
##  Box-Ljung test
##
## data:  quadratic_model$residuals
## X-squared = 5.6842, df = 3, p-value = 0.128
```

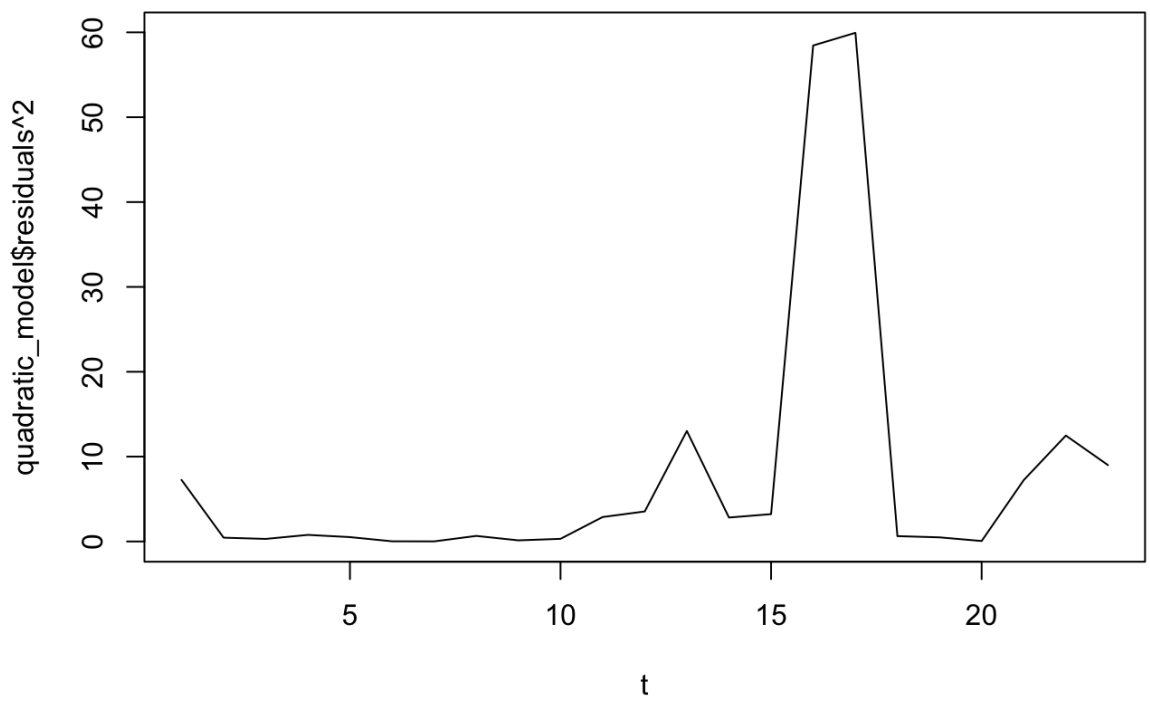
```
# heteroscedasticity
plot(linear_model$residuals^2,type="l",xlab="t")
```




```
bptest(linear_model)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  linear_model
## BP = 0.20313, df = 1, p-value = 0.6522
```

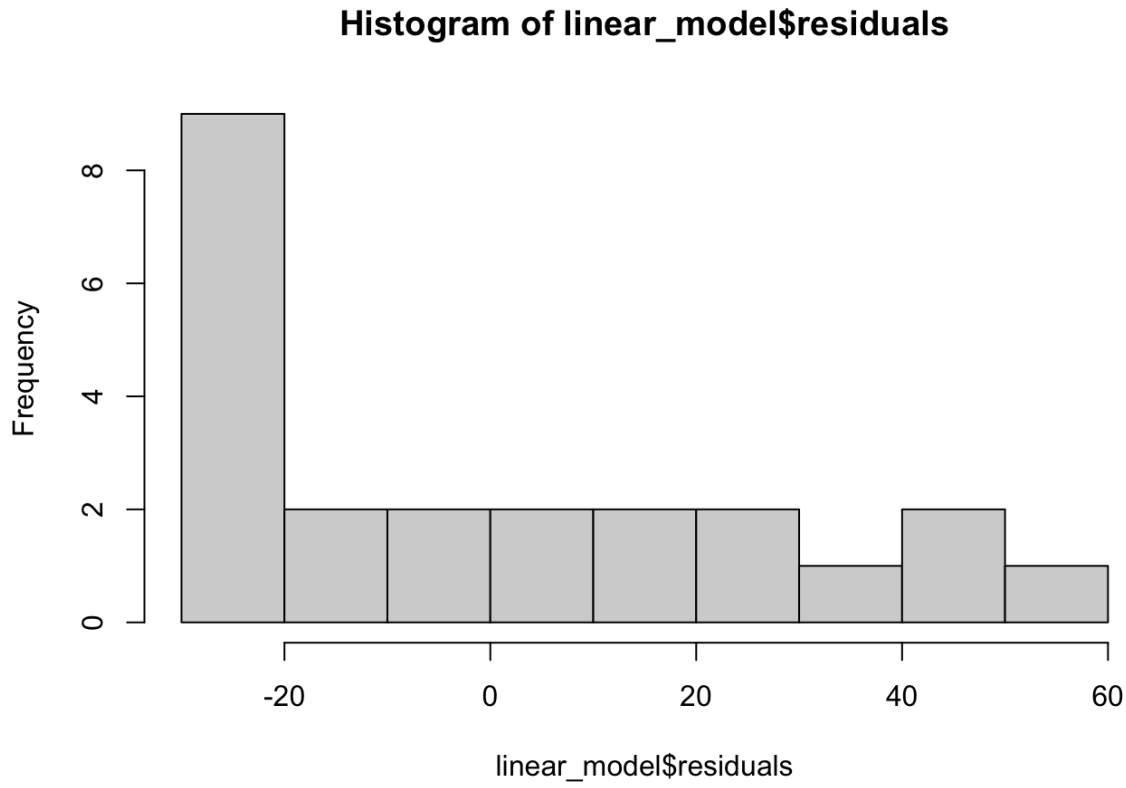
```
plot(quadratic_model$residuals^2,type="l",xlab="t")
```



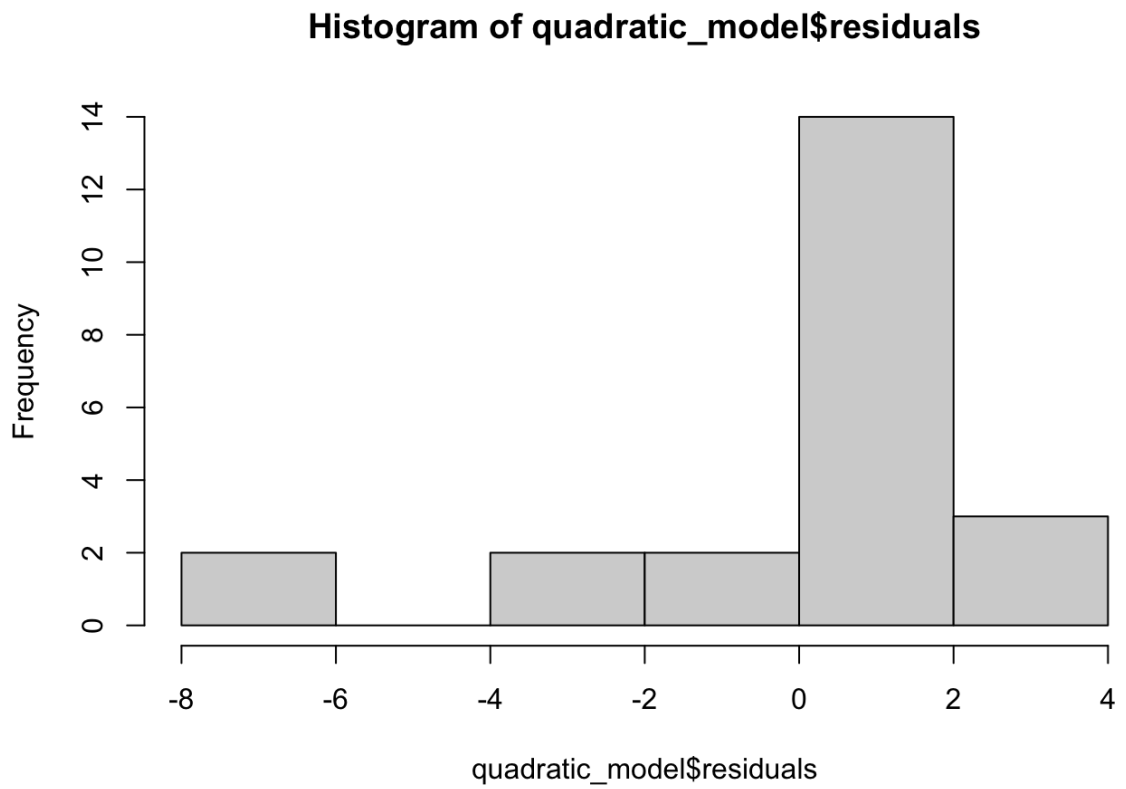

```
bptest(quadratic_model)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  quadratic_model
## BP = 2.4022, df = 2, p-value = 0.3009
```

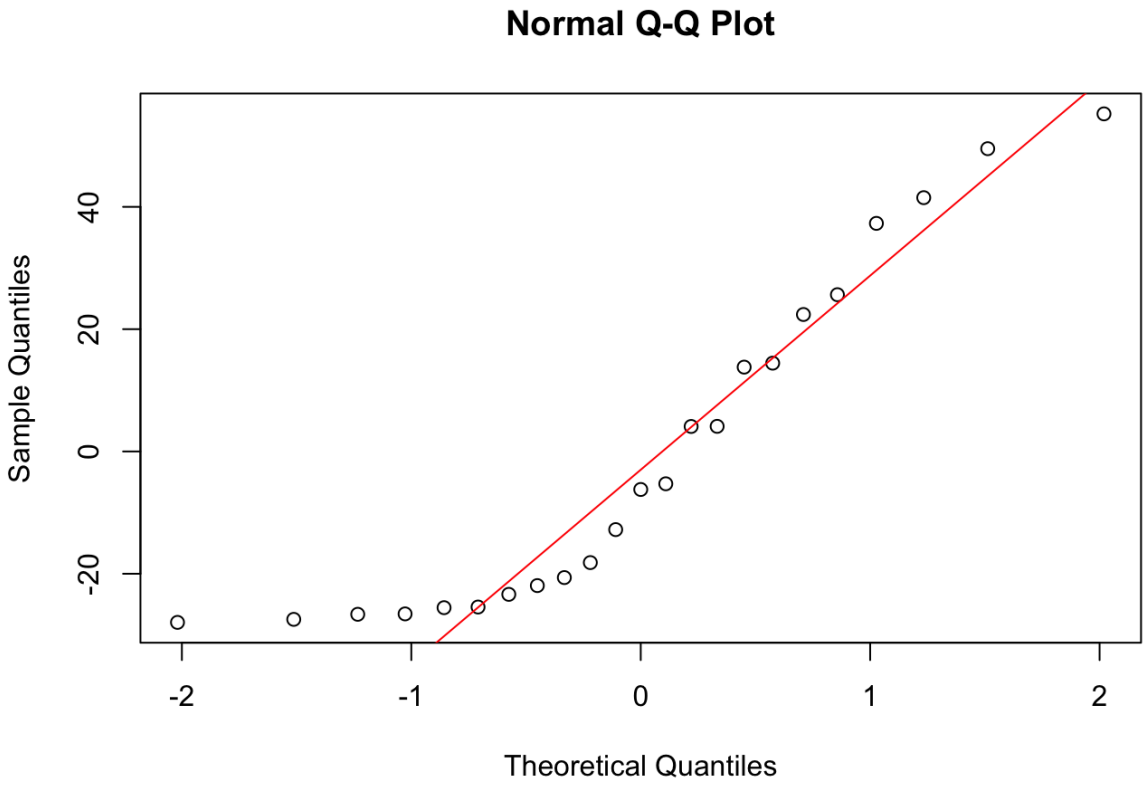
```
# normal distribution
hist(linear_model$residuals)
```



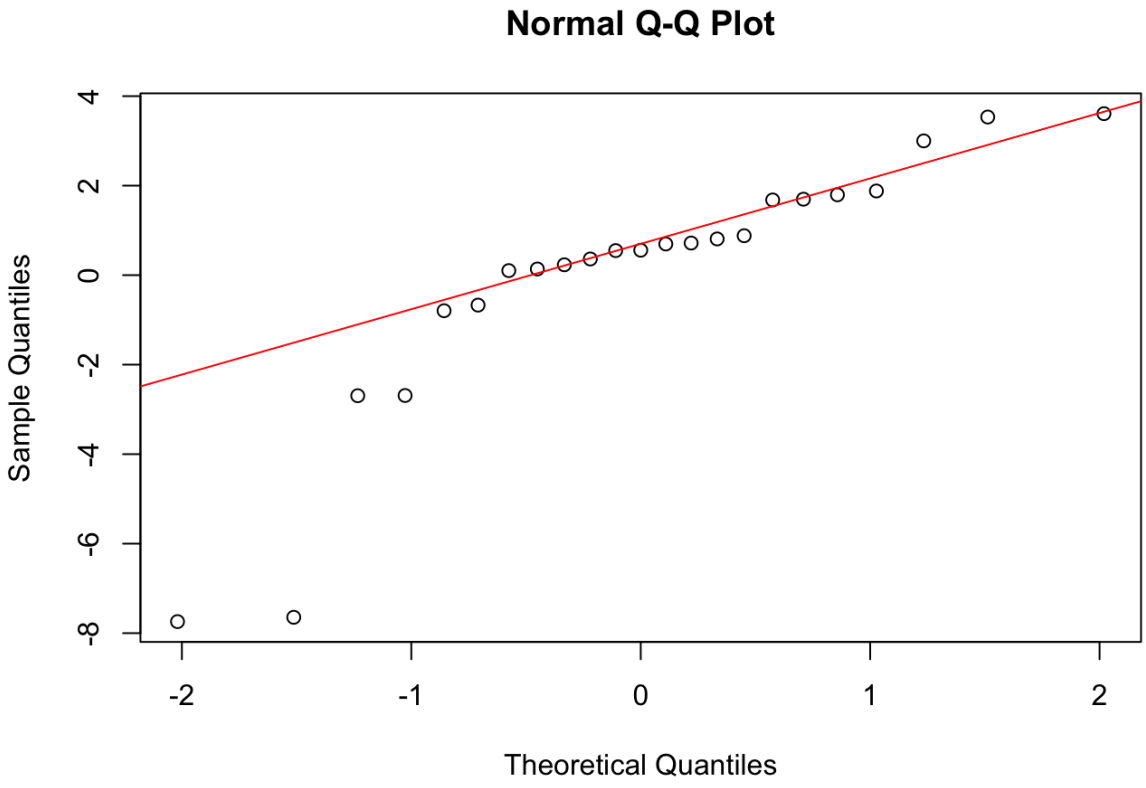

```
hist(quadratic_model$residuals)
```




```
qqnorm(linear_model$residuals)
qqline(linear_model$residuals,col="red")
```




```
qqnorm(quadratic_model$residuals)
qqline(quadratic_model$residuals,col="red")
```




```
jarque.bera.test(linear_model$residuals)
```

```
##
##  Jarque Bera Test
##
## data:  linear_model$residuals
## X-squared = 2.4596, df = 2, p-value = 0.2923
```

```
jarque.bera.test(quadratic_model$residuals)
```

```
##
##  Jarque Bera Test
##
## data:  quadratic_model$residuals
## X-squared = 13.555, df = 2, p-value = 0.001139
```