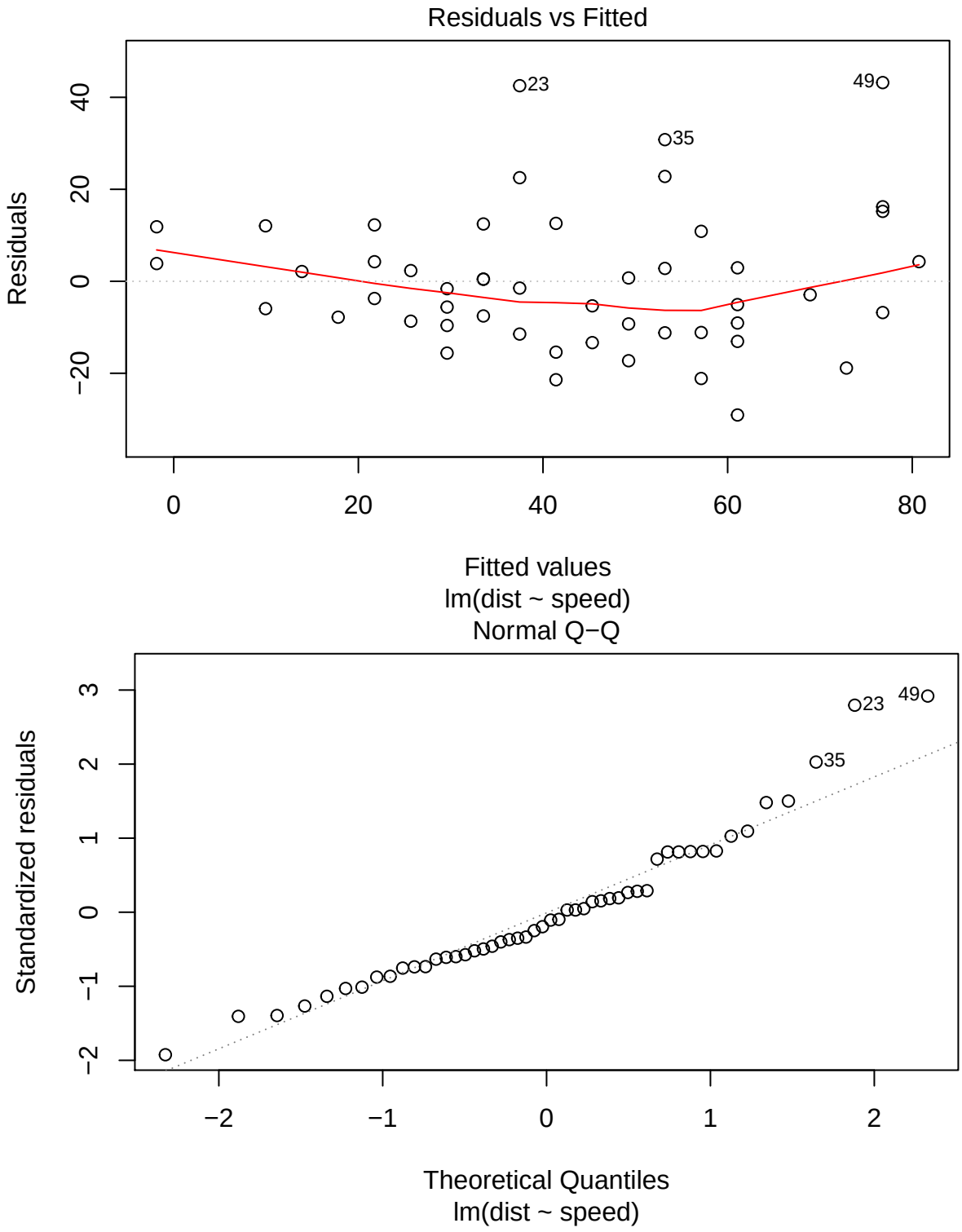
Stat 230: Linear Models
Homework 4
Professor Ding
Lev Golod

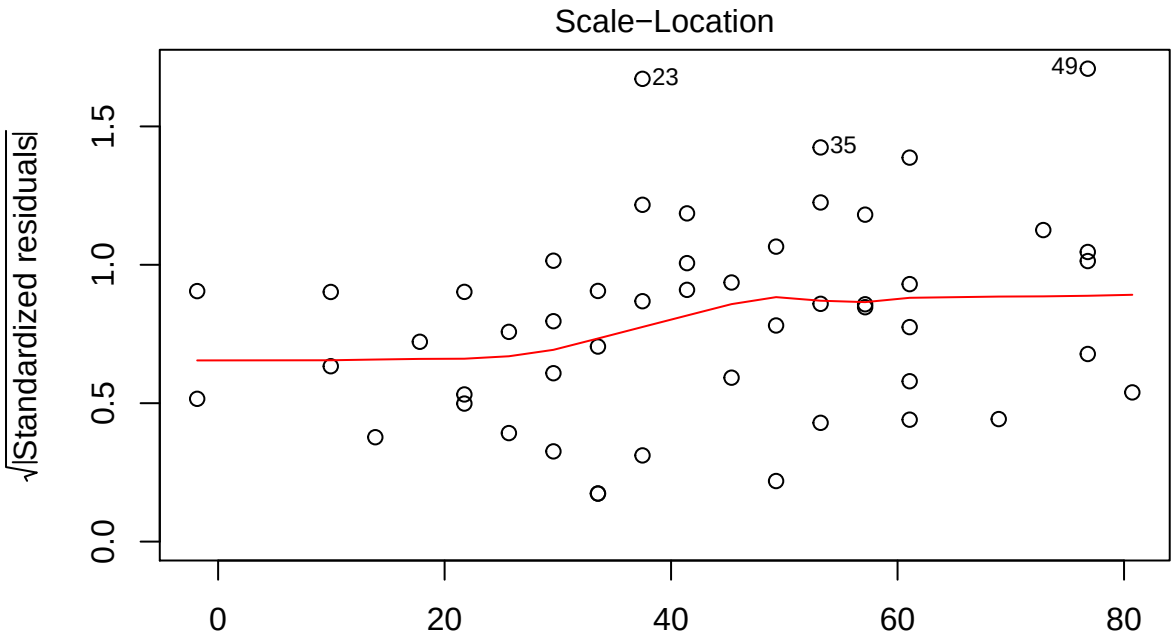
QUESTION 3, PART (A)

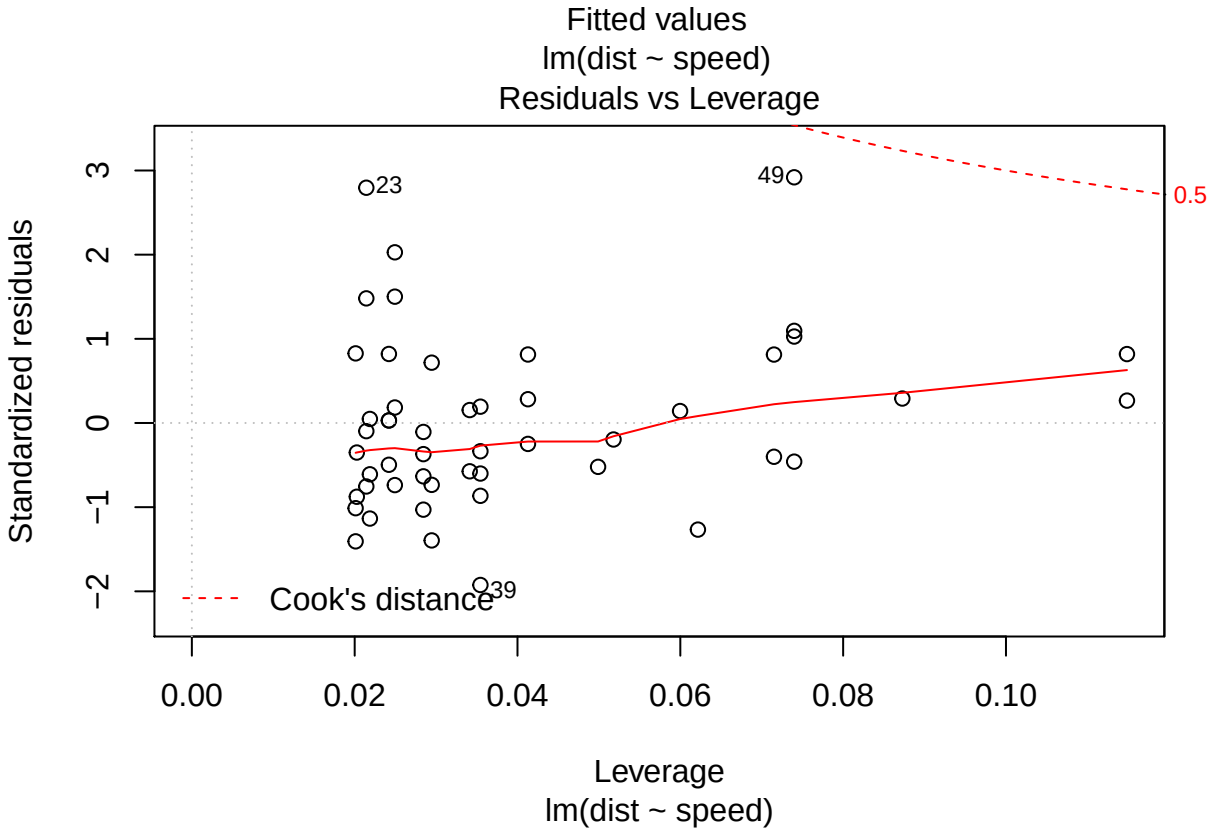
We see a strong cone shape in the residual plot, and a weaker cone in the plot ofsquare roots of studentized residuals. This is evidence of heteroskedasticity.

```
lm_model <- lm(dist ~ speed, data = cars)

# ## Residuals vs Fitted Values
# plt1a_resid <- ggplot(data.frame(x=lm_model$fitted.values, y=lm_model$residuals),
#                       aes(x,y)) +
#   geom_point() +
#   xlab('Fitted Values') +
#   ylab('Residuals') +
#   ggtitle('Question 3 (A) : Residual vs. Fitted')
# plt1a_resid
#
# ## Square Root of Studentized Residuals vs Fitted Values
# plt1a_stdresid <- ggplot(data.frame(x=lm_model$fitted.values,
#                                     y=abs(rstandard(lm_model))^0.5),
#                       aes(x,y)) +
#   geom_point() +
#   xlab('Fitted Values') +
#   ylab('Square Root of Studentized Residuals') +
#   ggtitle('Question 3 (A) : Square Root of Studentized Residuals vs. Fitted')
# plt1a_stdresid
plot(lm_model)
```


Residuals vs Fitted
Residuals
Fitted values
lm(dist ~ speed)
23
49
35
Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(dist ~ speed)





QUESTION 3, PART (B)

The p-value > 0.05; we fail to reject the null hypothesis of homoskedasticity at the level of alpha $= 0.05$. However, based on the graph in Part (A) I think that there is probably some heteroskedasticity after all.

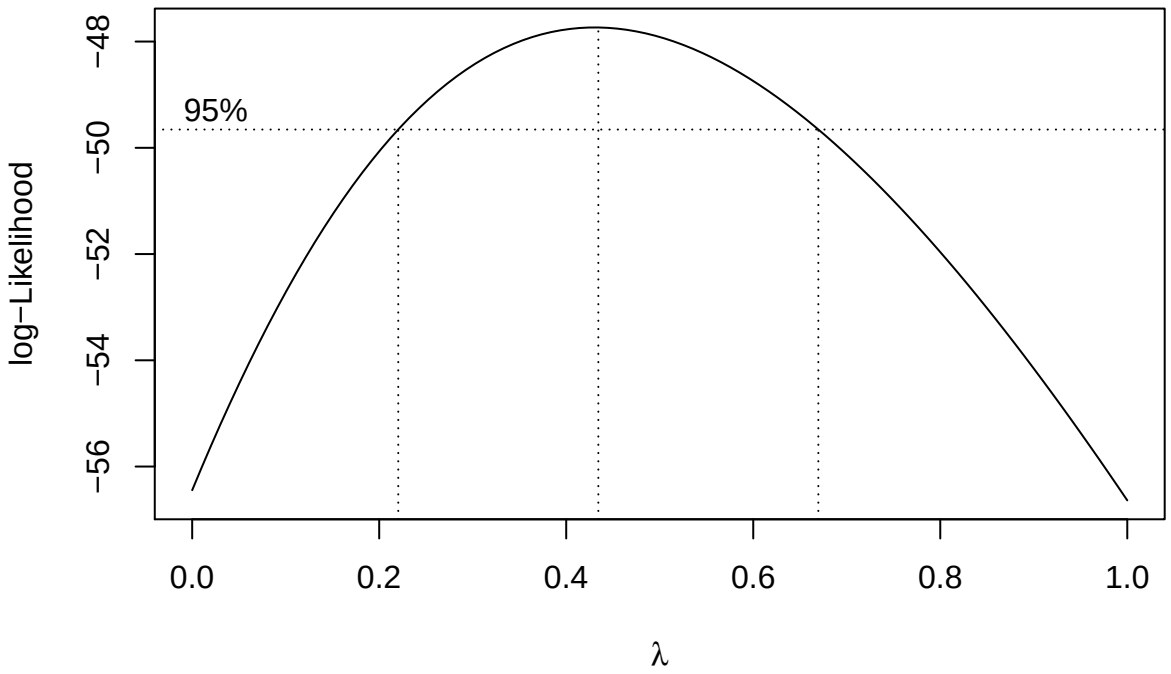
```
bptest(lm_model)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  lm_model
## BP = 3.21, df = 1, p-value = 0.073
```

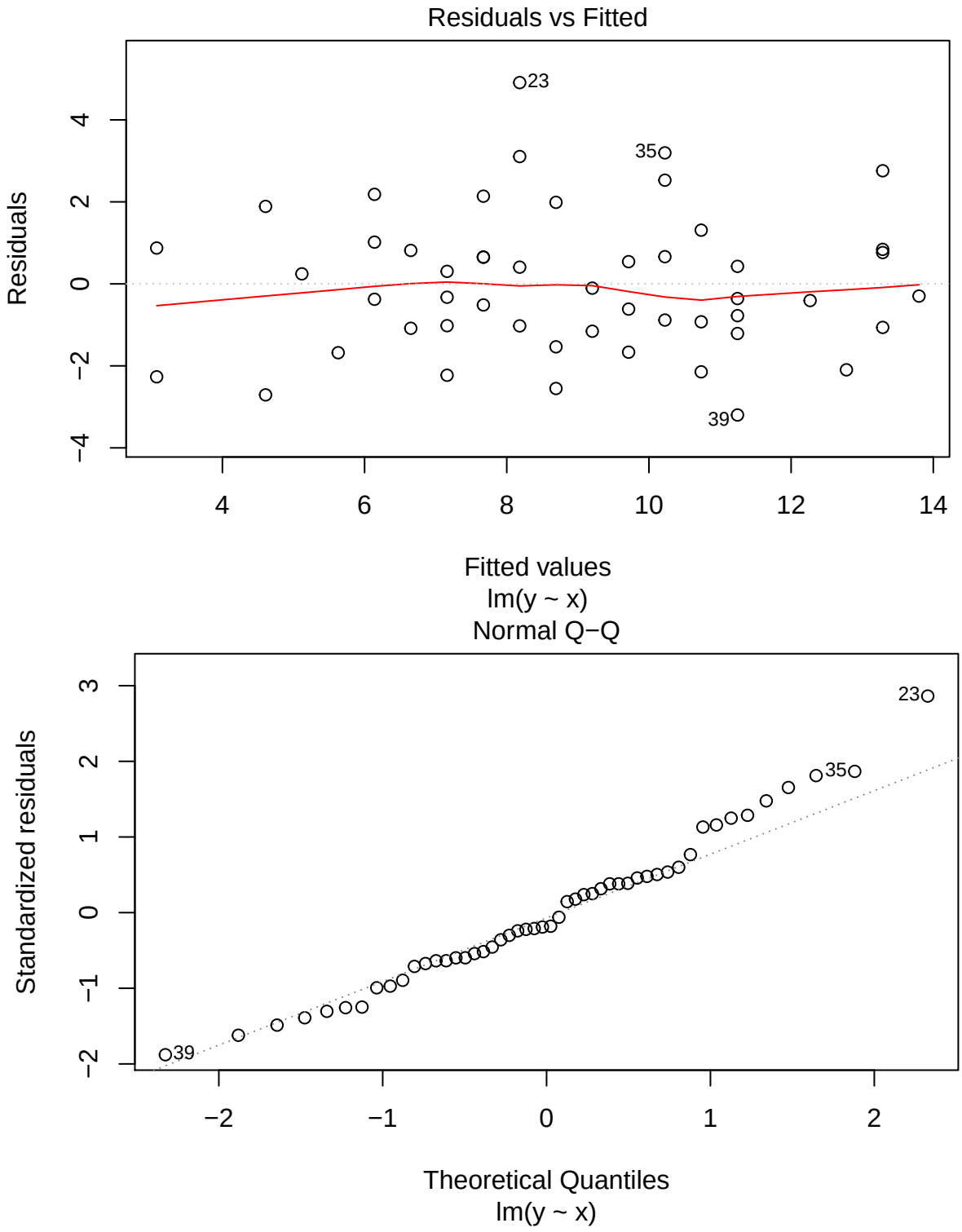
Question 3, Part (C)

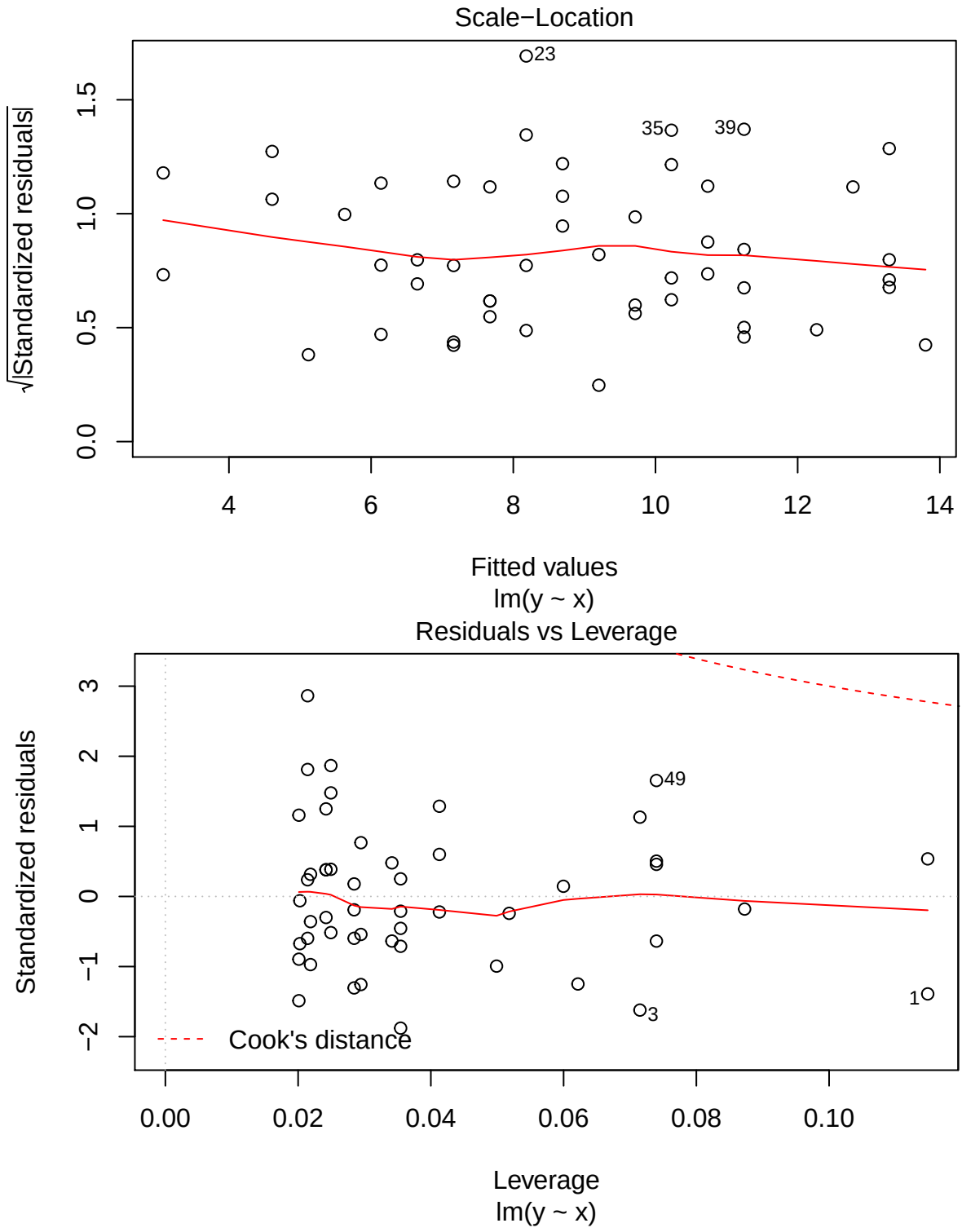
Looking at the Residual and Square-Root of Studentized Residuals plots, we see that there is now less of a fan shape. This means less heteroskedasticity.

```
boxcox(lm_model,lambda=seq(0,1,by=.05))
```




```
l <- 0.433
y_transform <- (cars$dist^l - 1)/l
newdat <- data.frame(y = y_transform, x = cars$speed)
lm2<- lm(y ~ x, data = newdat)
plot(lm2)
```


Residuals vs Fitted
Residuals
Fitted values
lm(y ~ x)
Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(y ~ x)
23
35
39


Scale−Location
√|Standardized residuals|
Fitted values
lm(y ~ x)
Residuals vs Leverage
Standardized residuals
Cook's distance
0.5
Leverage
lm(y ~ x)