Ozone_Assignment.R

SAMSON

2023-02-14

```
#### Assignment
getwd()
```

```
## [1] "C:/Users/SAMSON"
```

```
setwd("C:/Users/SAMSON")
ozone <- read.table("ozone (1).txt", header=TR
UE)
head(ozone)
```

```
##   rad temp wind ozone
## 1 190   67  7.4    41
## 2 118   72  8.0    36
## 3 149   74 12.6    12
## 4 313   62 11.5    18
## 5 299   65  8.6    23
## 6  99   59 13.8    19
```

```
Y = ozone$ozone
X1 = ozone$rad
X2 = ozone$temp
X3 = ozone$wind
model.identity = glm(Y ~ X1 + X2 + X3, family
= gaussian(link = "identity"))
model.inverse = glm(Y ~ X1 + X2 + X3, family =
gaussian(link = "inverse"))
model.log =  glm(Y ~ X1 + X2 + X3, family = ga
ussian(link = "log"))
model.exponential = glm(Y ~ log(X1) + log(X2)
+ log(X3), family = gaussian(link = "log"))
#1a
summary(model.identity)$aic
```

```
## [1] 998.6276
```

```
summary(model.inverse)$aic
```

```
## [1] 999.0104
```

```
summary(model.log)$aic # iii Best Model AIC = 972.5169
```

```
## [1] 972.5169
```

```
summary(model.exponential)$aic
```

```
## [1] 974.9654
```

```
# 1b
var.yi = summary(model.log)$dispersion
mu = predict(model.log,type="response")
pearson = (Y - mu)/sqrt(var.yi)
model.pearson.normal = glm(pearson^2 ~ mu,family = gaussian(link ="identity"))
model.pearson.gamma = glm(pearson^2 ~ mu,family = Gamma(link ="identity"))
model.pearson.inv = glm(pearson^2 ~ mu,family = inverse.gaussian(link ="identity"))
summary(model.pearson.normal)#i insignificant at 1%
```

```
##
## Call:
## glm(formula = pearson^2 ~ mu, family = gaus
sian(link = "identity"))
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.4306  -0.7060  -0.2830   0.0626  19.4220
##
## Coefficients:
##             Estimate Std. Error t value Pr
(>|t|)
## (Intercept) -0.147945   0.405849  -0.365
0.71617
## mu           0.026197   0.008072   3.246
0.00156 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*'
0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family t
aken to be 5.255567)
##
##     Null deviance: 628.22  on 110  degrees
of freedom
## Residual deviance: 572.86  on 109  degrees
```

```
of freedom
## AIC: 503.17
##
## Number of Fisher Scoring iterations: 2
```

```
summary(model.pearson.gamma)
```

```
##
## Call:
## glm(formula = pearson^2 ~ mu, family = Gamm
a(link = "identity"))
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.8385  -1.5794  -0.6042   0.1484   4.0541
##
## Coefficients:
##             Estimate Std. Error t value Pr
(>|t|)
## (Intercept) -0.092842   0.046616  -1.992
0.0489 *
## mu           0.023788   0.004176   5.697 1.
05e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*'
0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family take
n to be 2.63651)
##
##     Null deviance: 328.61  on 110  degrees
of freedom
## Residual deviance: 256.46  on 109  degrees
```

```
of freedom
## AIC: 115.83
##
## Number of Fisher Scoring iterations: 7
```

```
summary(model.pearson.inv)
```

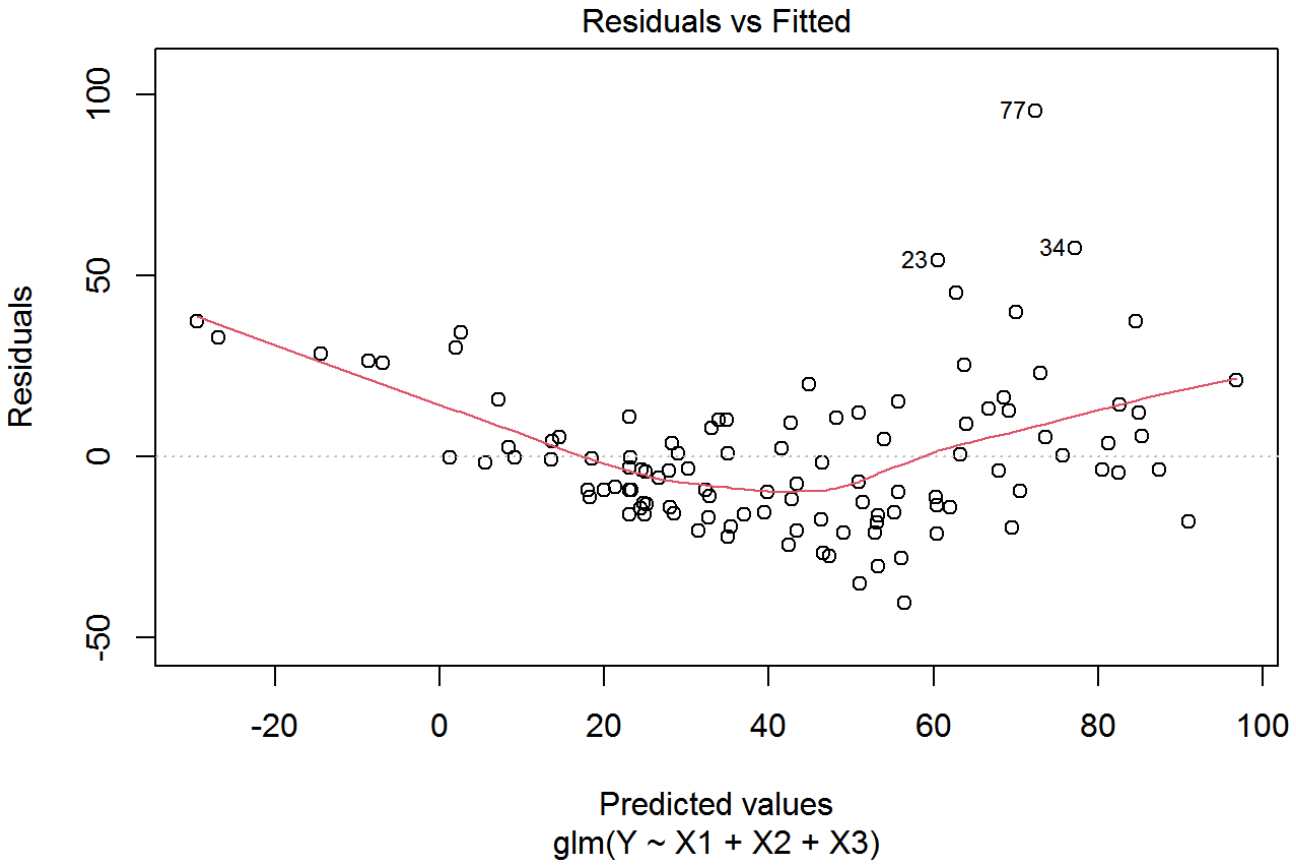
```
##
## Call:
## glm(formula = pearson^2 ~ mu, family = inve
rse.gaussian(link = "identity"))
##
## Deviance Residuals:
##      Min        1Q    Median        3Q
Max
## -136.619    -2.754    -0.782     0.177
4.780
##
## Coefficients:
##             Estimate Std. Error t value Pr
(>|t|)
## (Intercept) -0.103365   0.021468  -4.815 4.
78e-06 ***
## mu           0.025080   0.003892   6.445 3.
24e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*'
0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for inverse.gaussian
family taken to be 3.935746)
##
##     Null deviance: 36009  on 110  degrees o
```

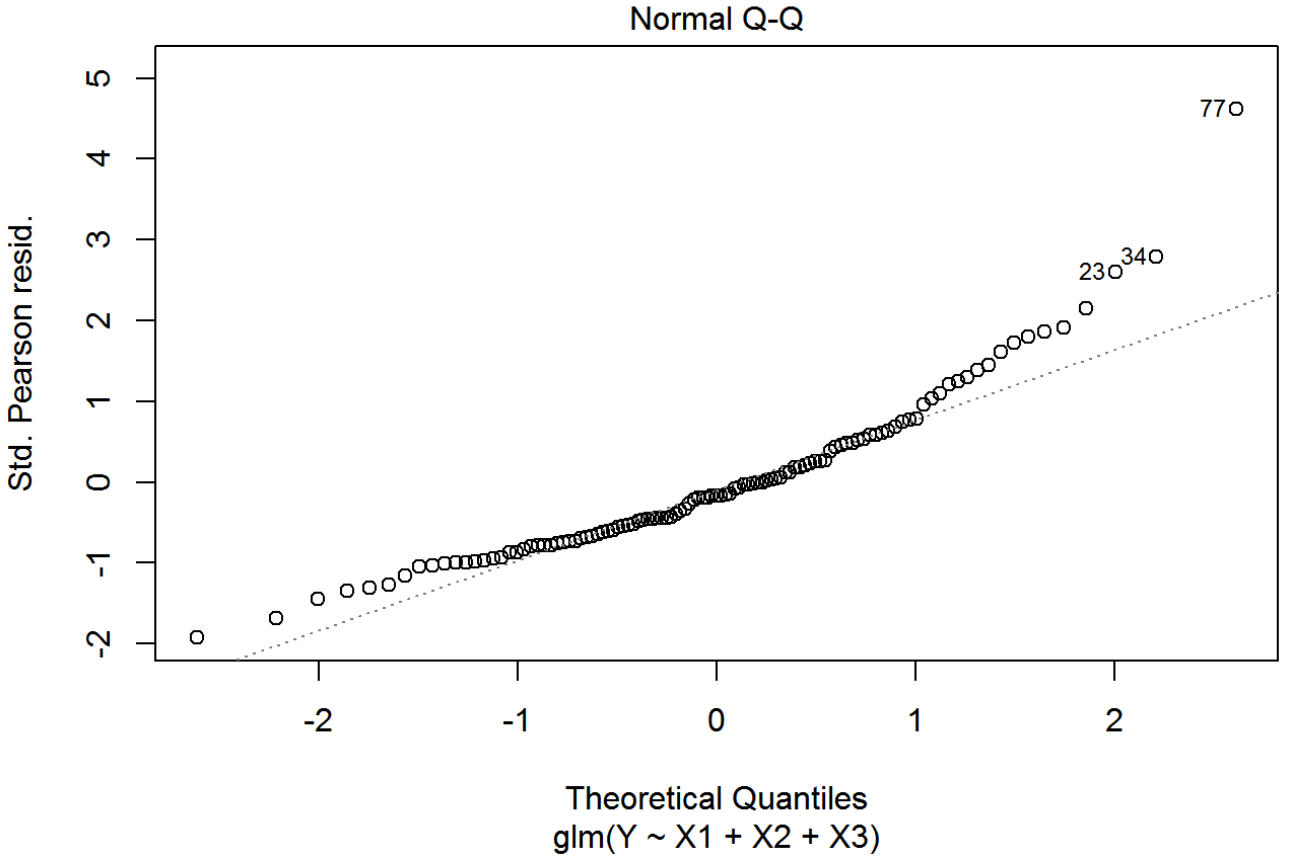
```
f freedom
## Residual deviance: 35867  on 109  degrees o
f freedom
## AIC: 457.23
##
## Number of Fisher Scoring iterations: 6
```

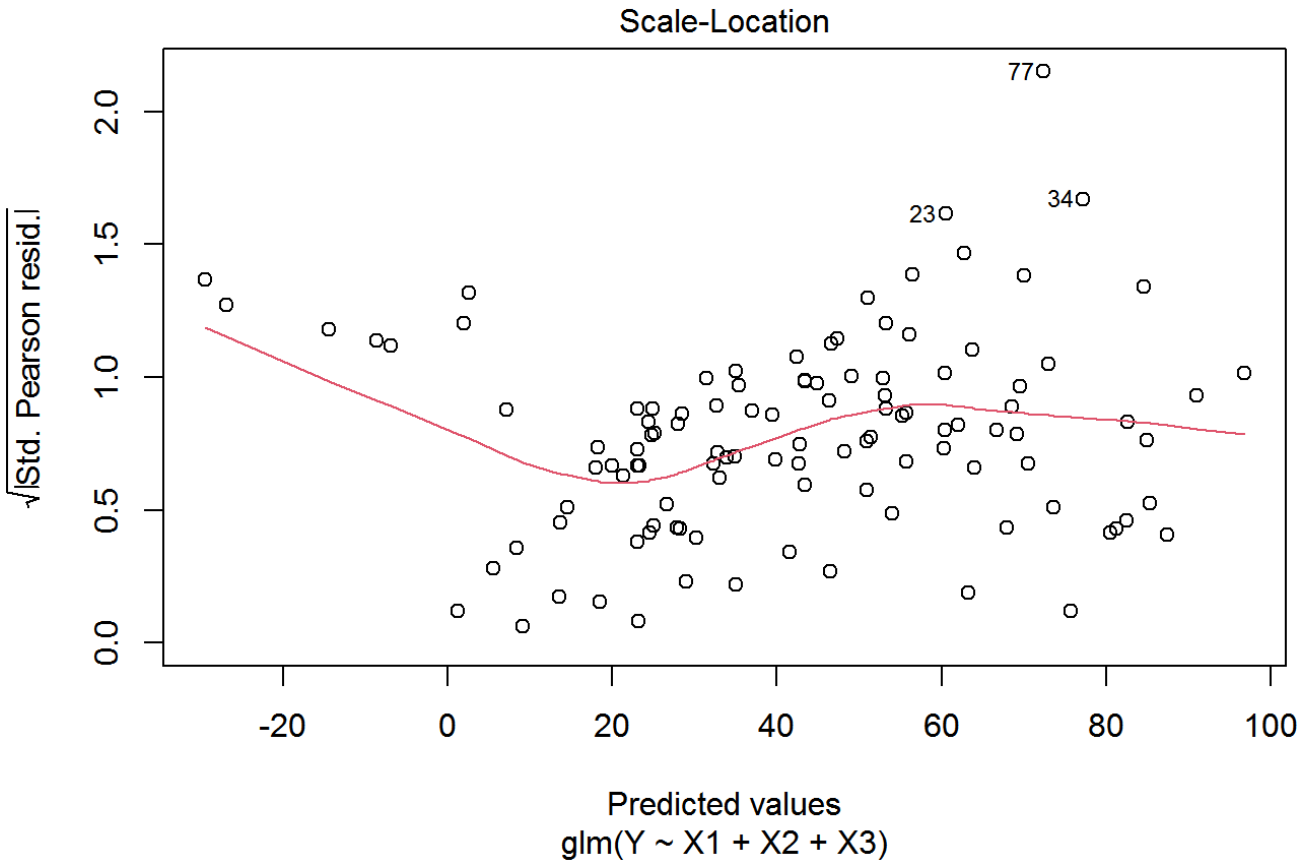
```
# 1c
betahat = summary(model.log)$coef[1:4]
betahat = matrix(betahat)
K = matrix(c(0,0,1,1),nrow = 4,ncol = 1)
W = (t(t(K)%*%betahat))*solve(t(K)%*%vcov(mode
l.log)%*%K)*(t(K)%*%betahat)
W # 10.66784
```

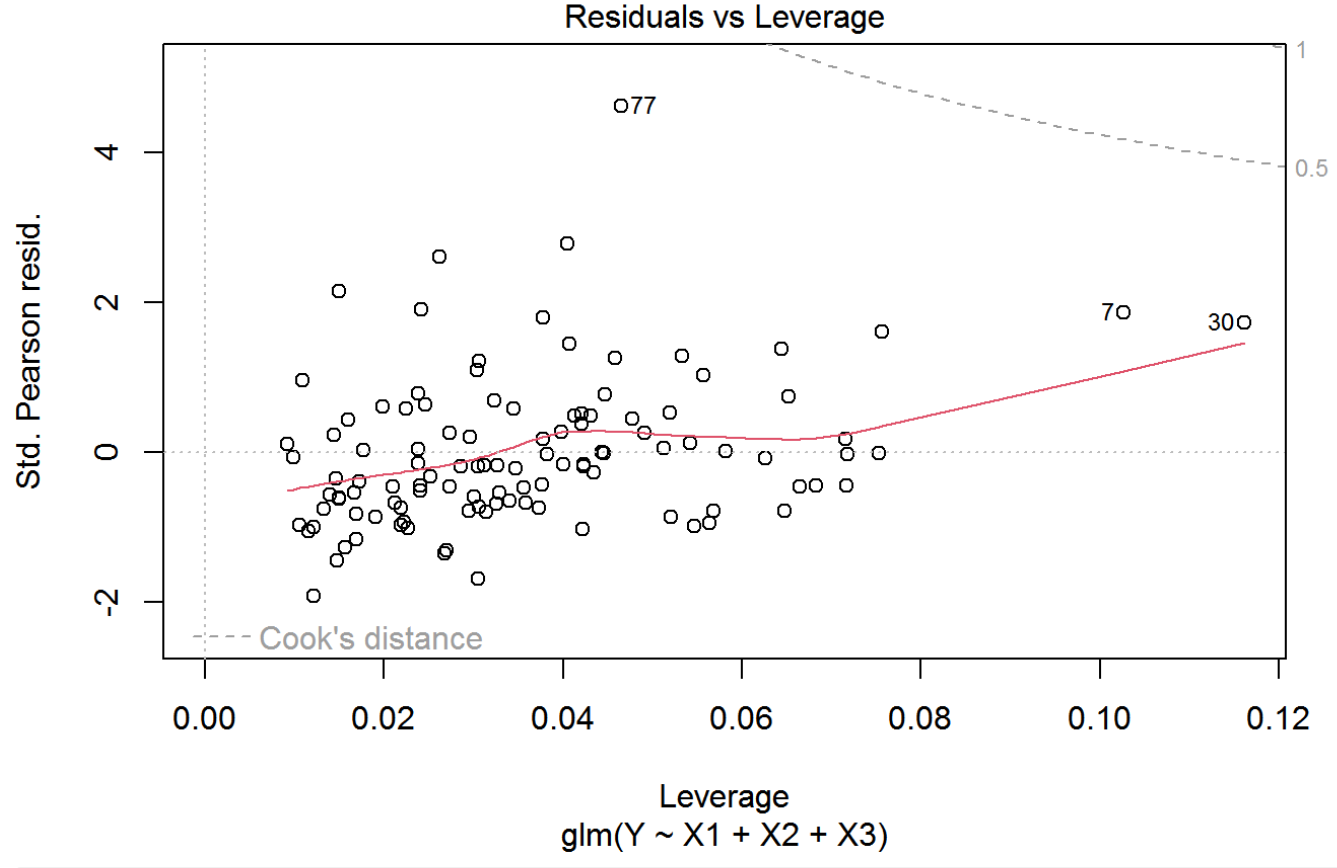
```
##          [,1]
## [1,] 10.66784
```

```
plot(model.identity)
```

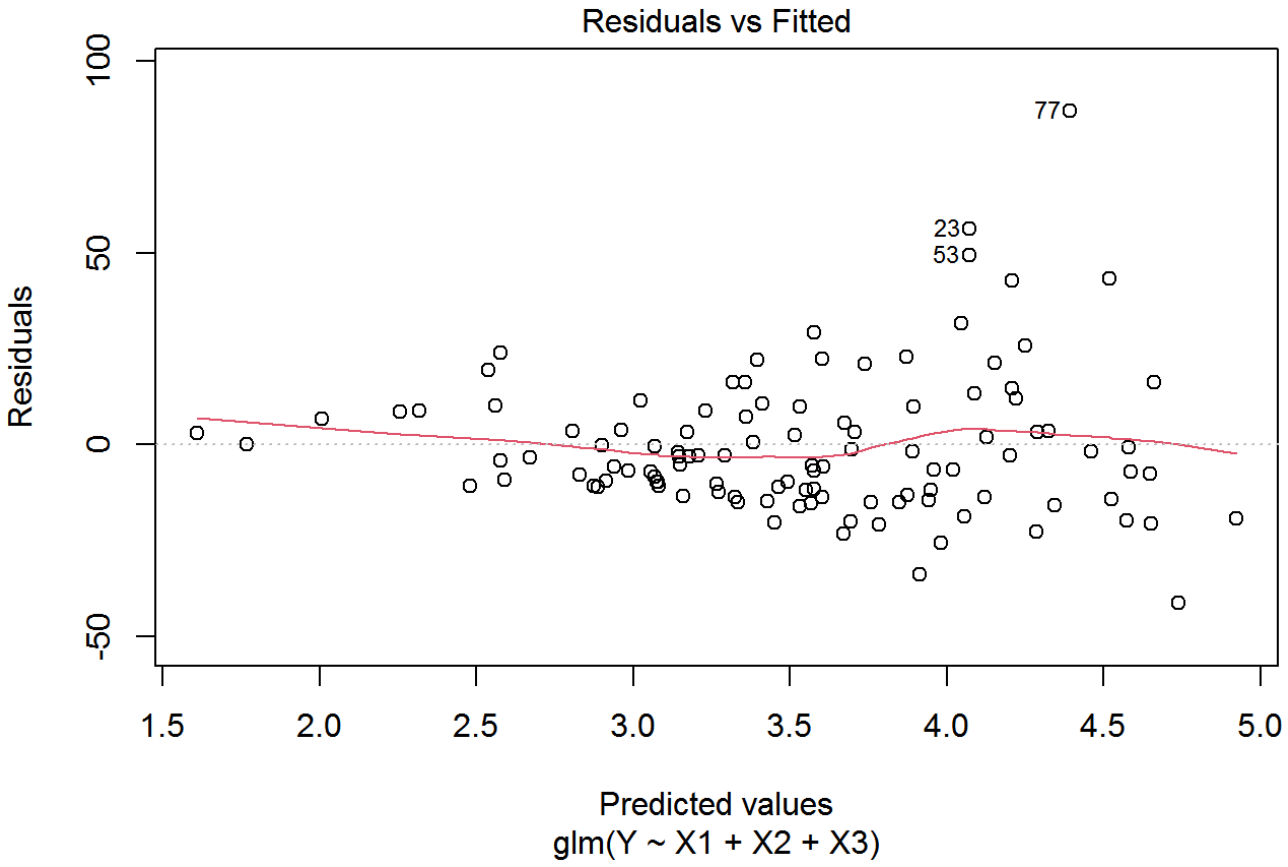

Residuals vs Fitted
Residuals
Predicted values
glm(Y ~ X1 + X2 + X3)
77
23
34
100
50
0
-50
-20
0
20
40
60
80
100

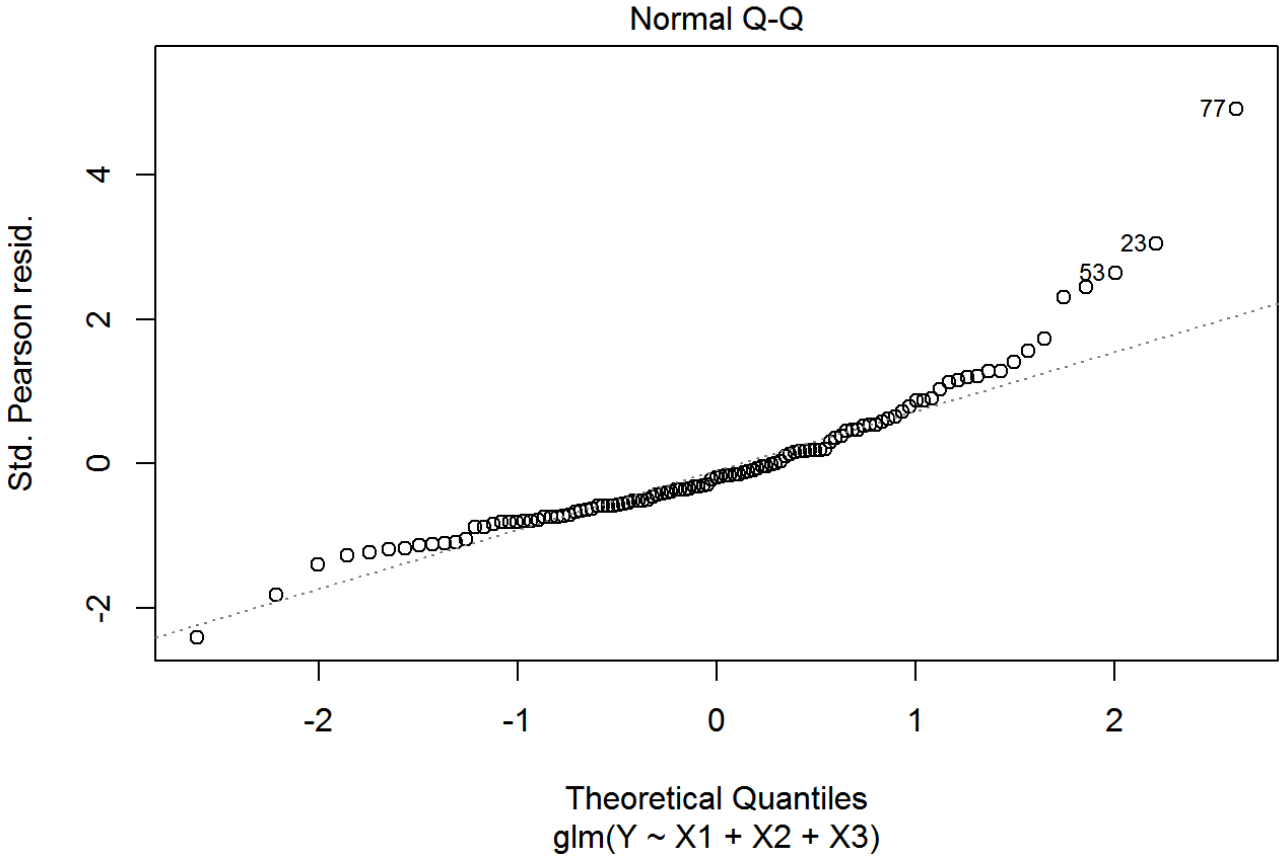

Normal Q-Q
Std. Pearson resid.
Theoretical Quantiles
glm(Y ~ X1 + X2 + X3)
77
23
34
5
4
3
2
1
0
-1
-2
-2
-1
0
1
2

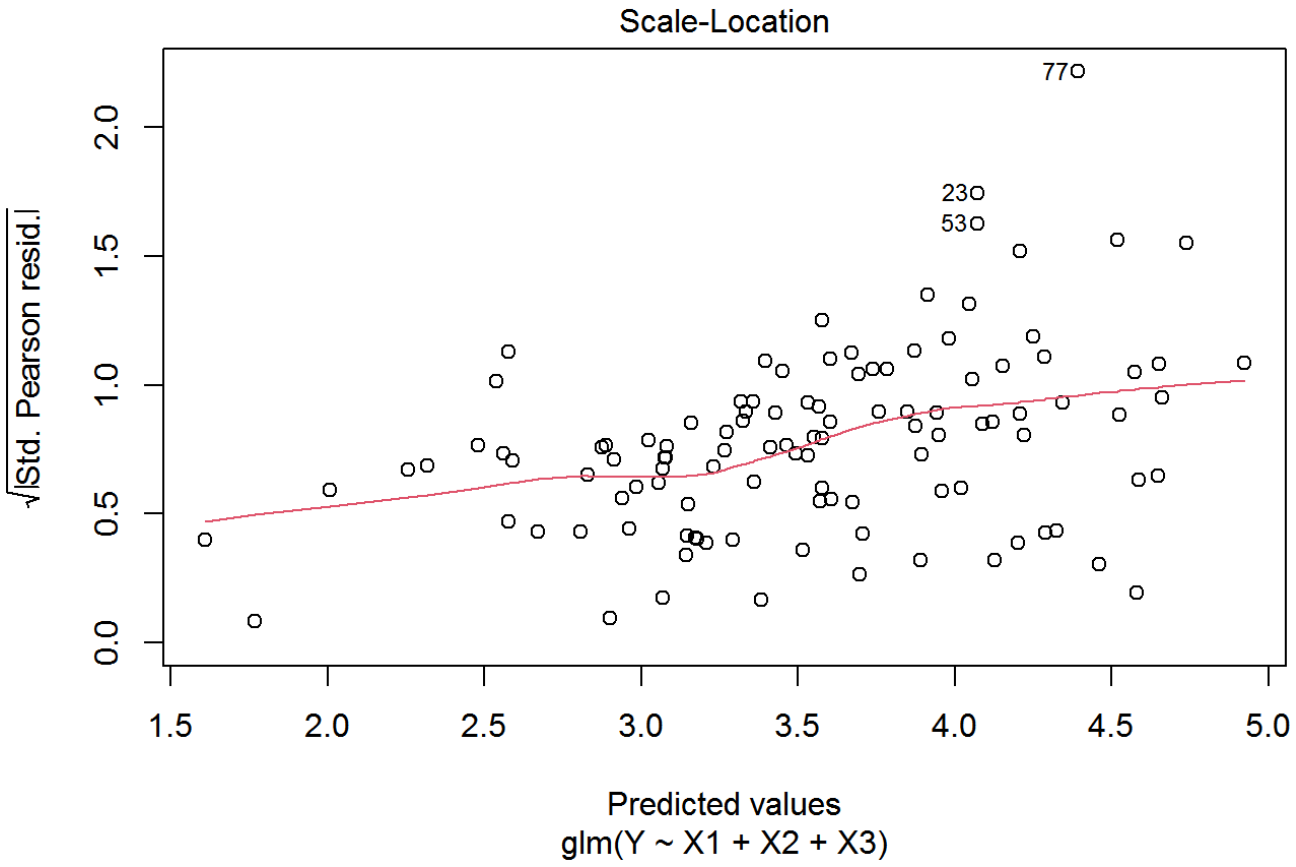

Scale-Location
77
34
23
√|Std. Pearson resid.|
0.0
0.5
1.0
1.5
2.0
-20
0
20
40
60
80
100
Predicted values
glm(Y ~ X1 + X2 + X3)

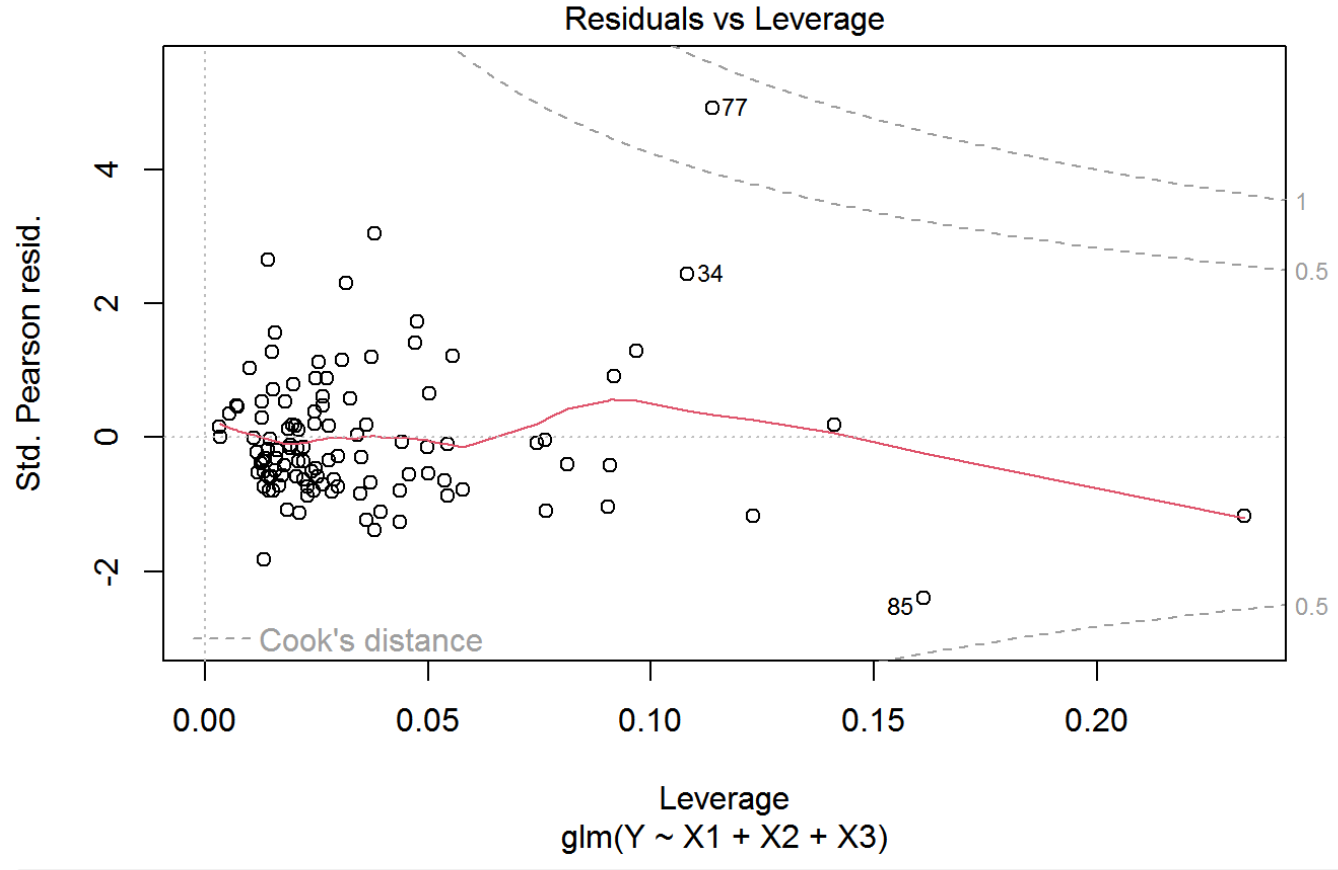

Residuals vs Leverage
77
70
30
1
0.5
Cook's distance
Std. Pearson resid.
-2
0
2
4
0.00
0.02
0.04
0.06
0.08
0.10
0.12
Leverage
glm(Y ~ X1 + X2 + X3)

```
plot(model.log)
```

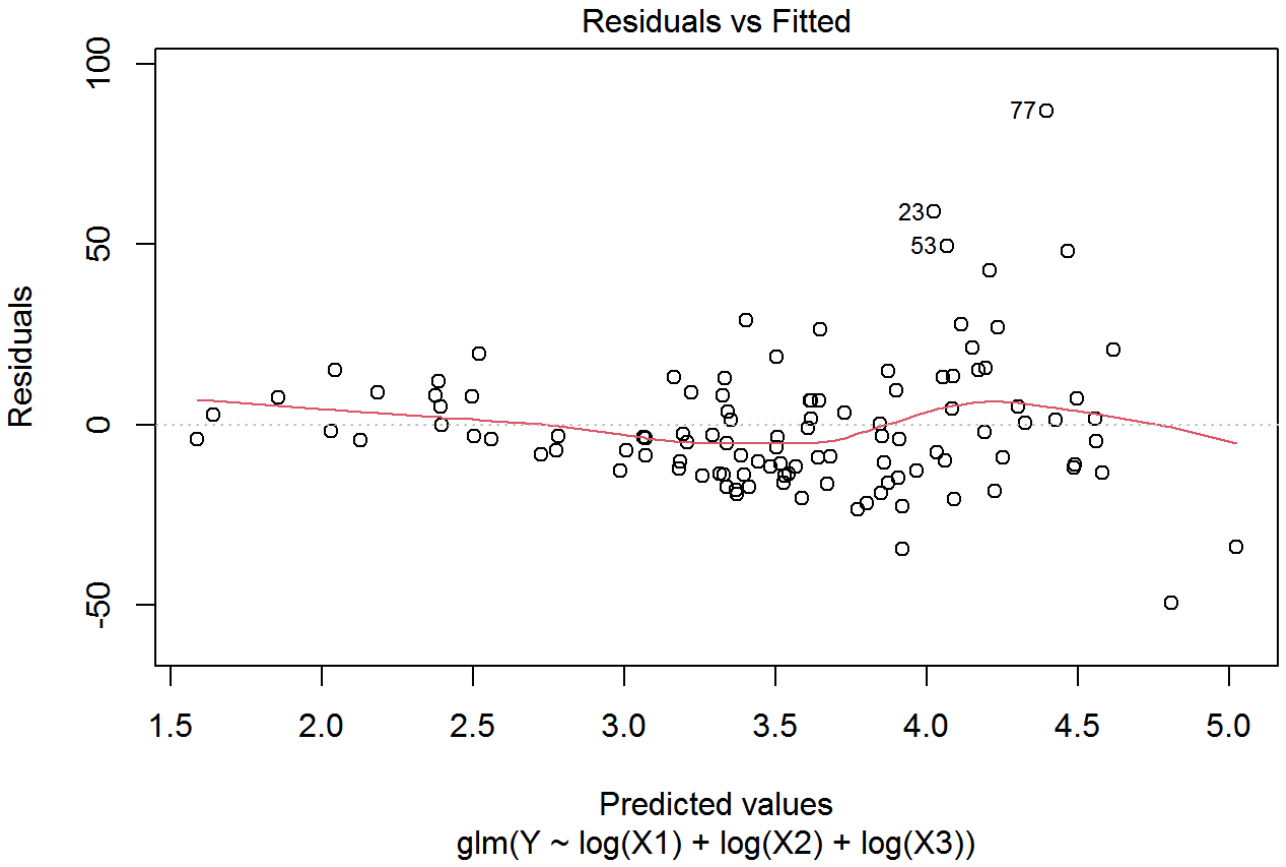

Residuals vs Fitted
Residuals
Predicted values
glm(Y ~ X1 + X2 + X3)
77
23
53
100
50
0
-50
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0

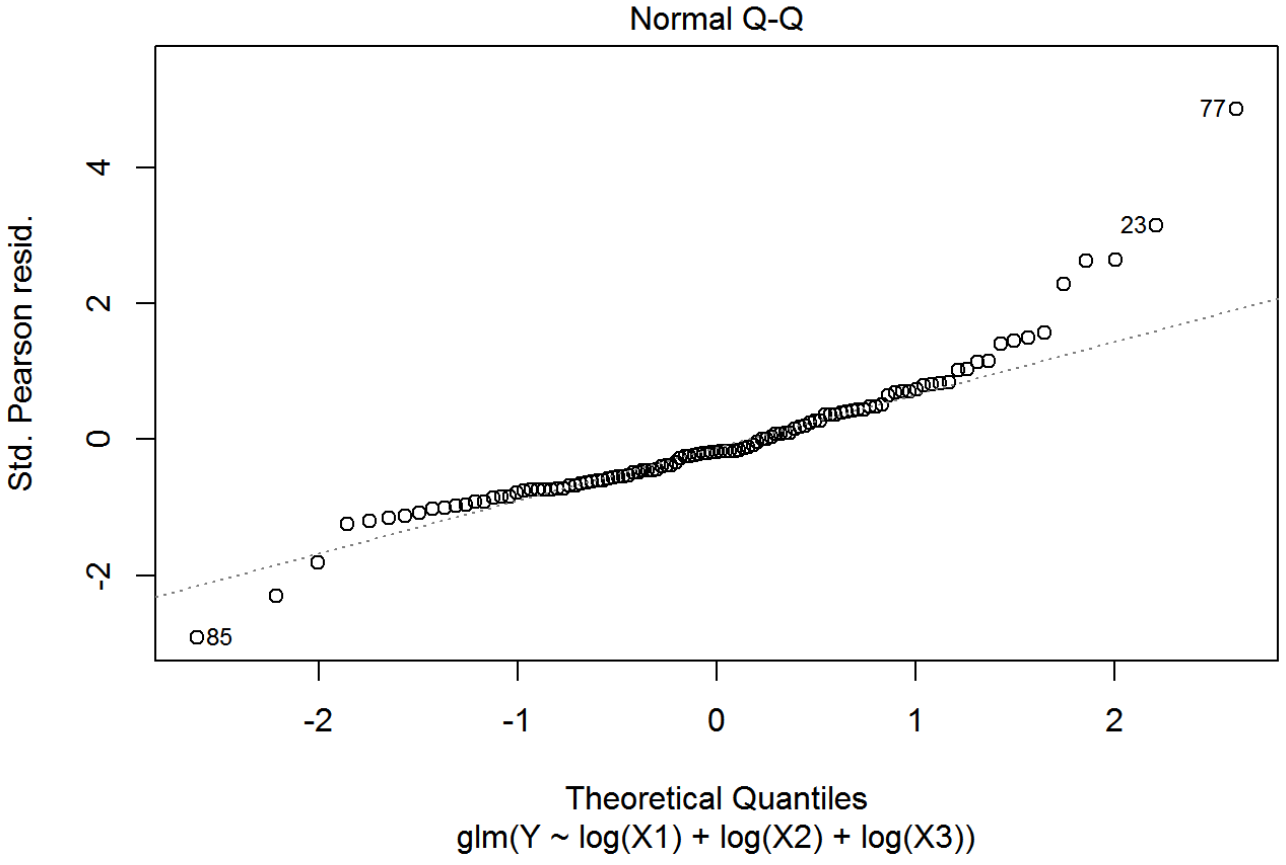

Normal Q-Q
Std. Pearson resid.
Theoretical Quantiles
glm(Y ~ X1 + X2 + X3)
77
23
53
4
2
0
-2
-2
-1
0
1
2

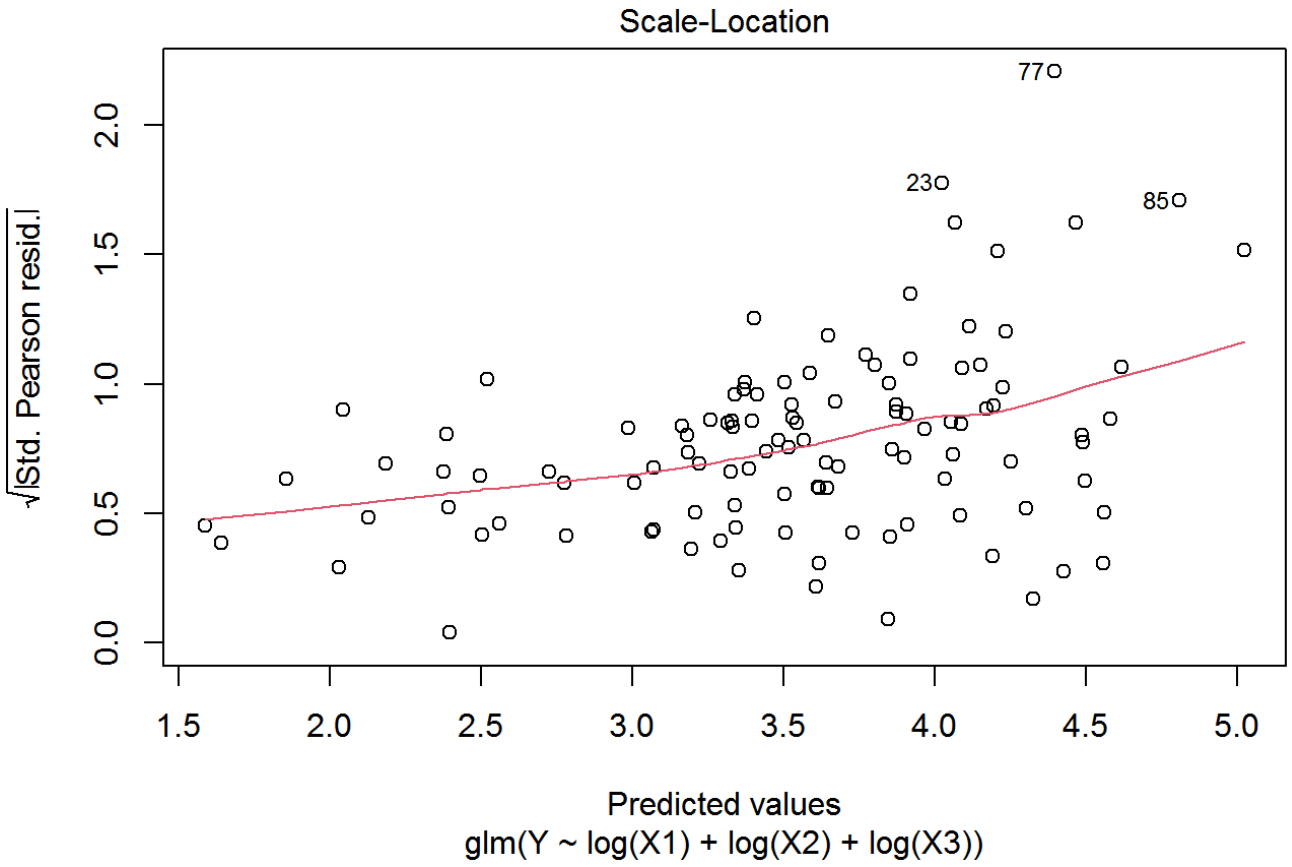

Scale-Location
77
23
53
√|Std. Pearson resid.|
0.0
0.5
1.0
1.5
2.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Predicted values
glm(Y ~ X1 + X2 + X3)

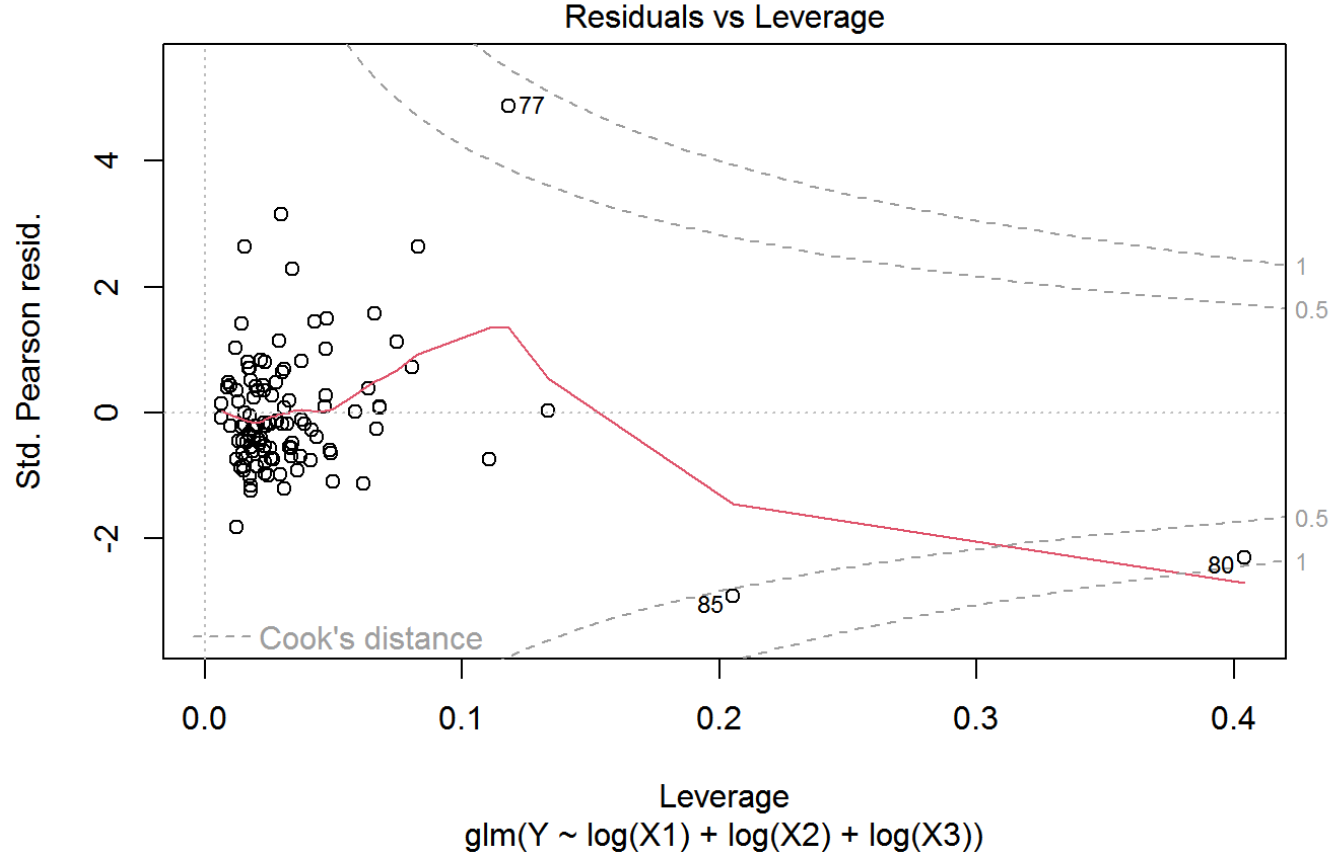

Residuals vs Leverage
77
34
85
1
0.5
0.5
Cook's distance
Std. Pearson resid.
-2
0
2
4
0.00
0.05
0.10
0.15
0.20
Leverage
glm(Y ~ X1 + X2 + X3)

```
plot(model.exponential)
```

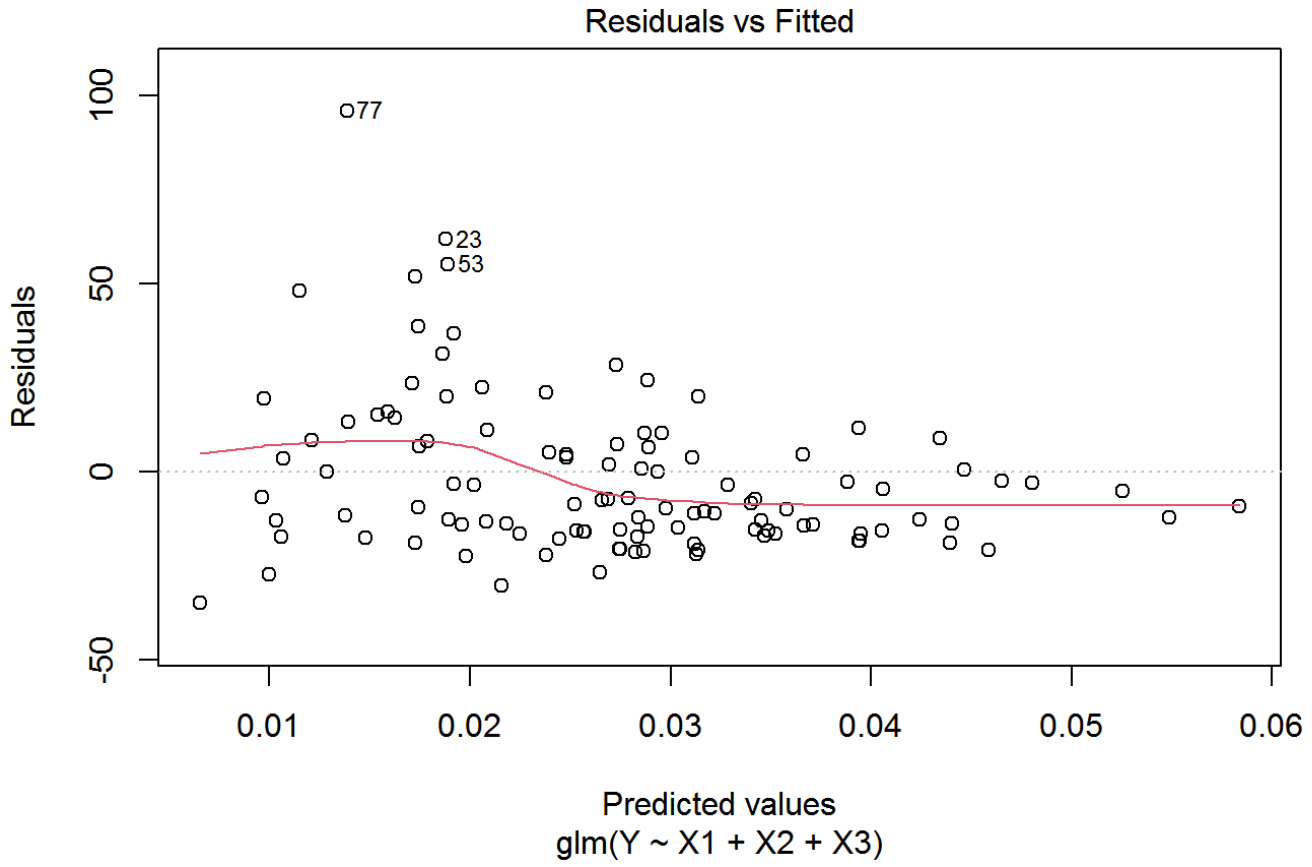

Residuals vs Fitted
Residuals
100
50
0
-50
77
23
53
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Predicted values
glm(Y ~ log(X1) + log(X2) + log(X3))

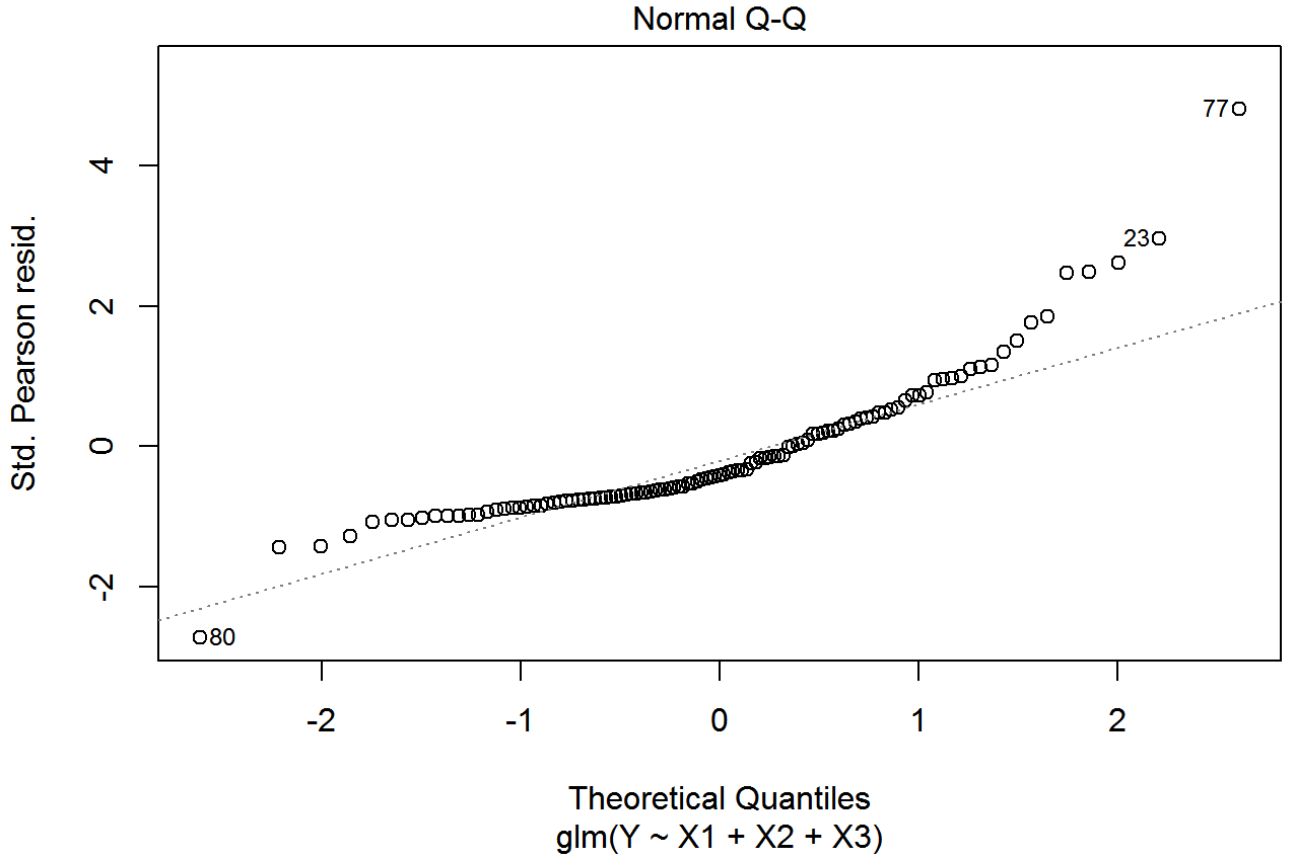

Normal Q-Q
Std. Pearson resid.
4
2
0
-2
77
23
85
-2
-1
0
1
2
Theoretical Quantiles
glm(Y ~ log(X1) + log(X2) + log(X3))

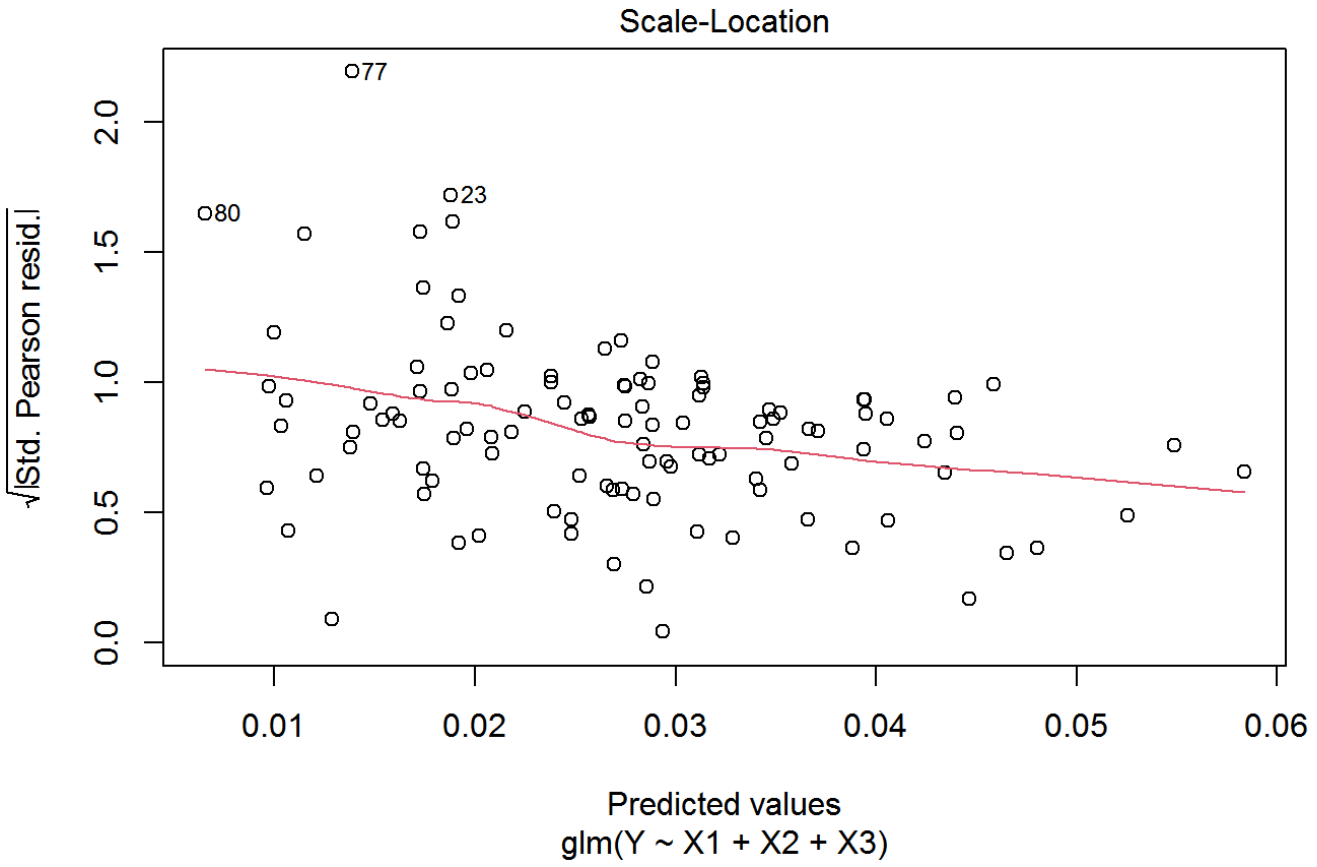

Scale-Location
77
23
85
√|Std. Pearson resid.|
0.0
0.5
1.0
1.5
2.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
Predicted values
glm(Y ~ log(X1) + log(X2) + log(X3))

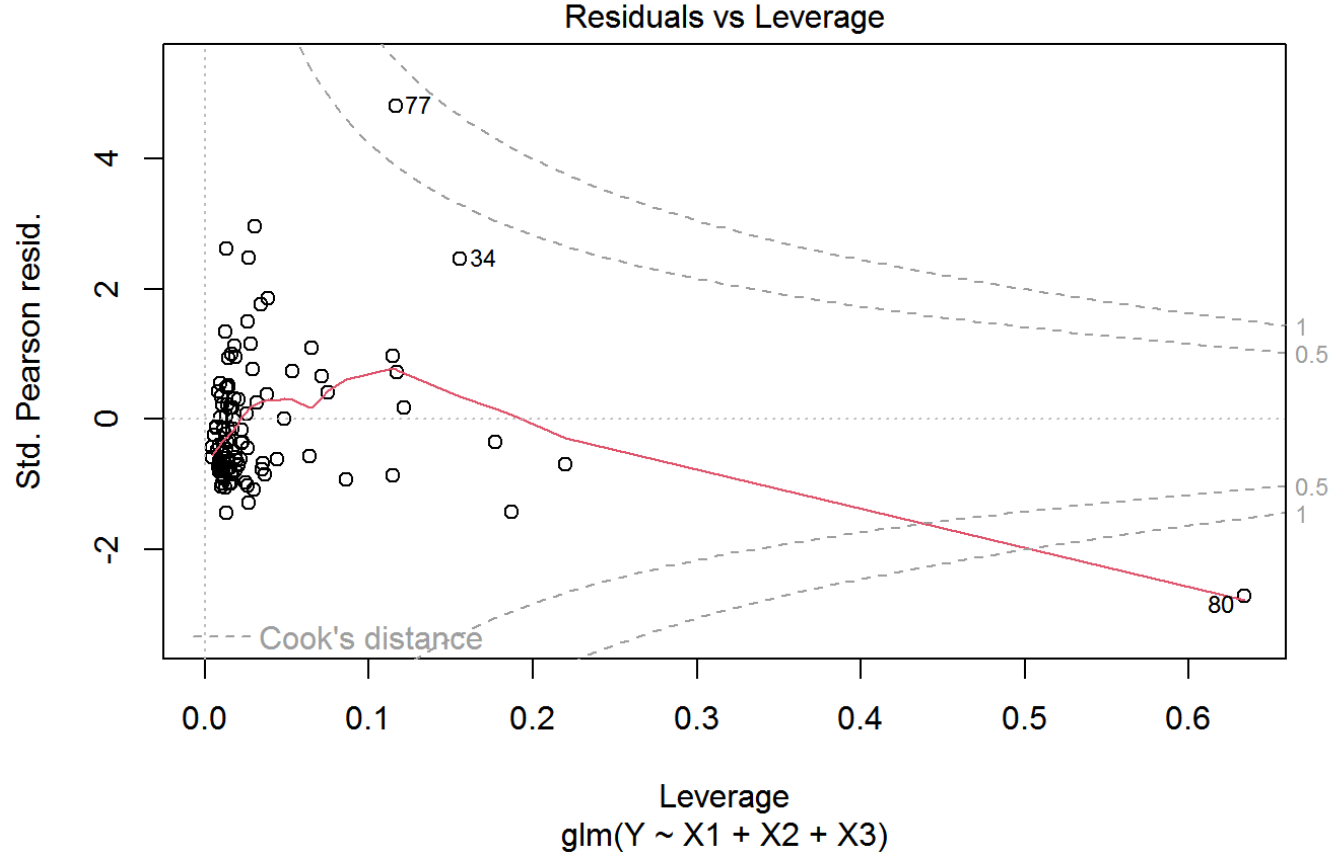

Residuals vs Leverage
77
85
80
Cook's distance
1
0.5
0.5
1
Std. Pearson resid.
-2
0
2
4
0.0
0.1
0.2
0.3
0.4
Leverage
glm(Y ~ log(X1) + log(X2) + log(X3))

```
plot(model.inverse)
```


Residuals vs Fitted
Residuals
Predicted values
glm(Y ~ X1 + X2 + X3)
77
23
53
100
50
0
-50
0.01
0.02
0.03
0.04
0.05
0.06


Normal Q-Q
Std. Pearson resid.
Theoretical Quantiles
glm(Y ~ X1 + X2 + X3)
77
23
80
4
2
0
-2
-2
-1
0
1
2


Scale-Location
77
23
80
√|Std. Pearson resid.|
0.0
0.5
1.0
1.5
2.0
0.01
0.02
0.03
0.04
0.05
0.06
Predicted values
glm(Y ~ X1 + X2 + X3)


Residuals vs Leverage
77
34
80
Std. Pearson resid.
-2
0
2
4
1
0.5
0.5
1
Cook's distance
0.0
0.1
0.2
0.3
0.4
0.5
0.6
Leverage
glm(Y ~ X1 + X2 + X3)