**Lab 6: Linear Regression Validation**

**Problem statement:**

Calculate following parameter of Simple Liner Regression (drug2.csv) by implementing the formula from scratch.

* R2
* Adjusted- R2
* Residual standard error
* S-value (for slope only)
* P-value (for slope only)

**Source Code:**

#Author: Ashish Upadhyay

#Branch: Computer Science and Engineering

#Semester: 6th

#Dr. SP Mukherjee International Institute of Information Technology, Naya Raipur

#Subject: Machine Learning Lab 6

#Task: Linear Regression Validation

setwd("C:/Users/Ashish Upadhyay/Documents/Semester6/MachineLearning/Lab Programs")

getwd()

drug = read.csv("drug2.csv")

attach(drug)

head(drug)

nrow(drug)

model1 = lm(response~dose)

summary(model1)

err1 = residuals(model1)

plot(model1$fitted.values, err1)

hist(err1)

plot(model1$fitted.values, drug$response)

mean <- mean(response)

sst <- sum((response-mean)\*\*2)

sse <- sum((response-model1$fitted.values)\*\*2)

rsq <- 1 - (sse/sst)

fuv <- sse/sst

fuv

rsq

a <- (nrow(drug)-1)\*(1-rsq)

b <- nrow(drug)-2

rad <- 1-(a/b)

rad

**Output:**

> #Author: Ashish Upadhyay

> #Branch: Computer Science and Engineering

> #Semester: 6th

> #Dr. SP Mukherjee International Institute of Information Technology, Naya Raipur

> #Subject: Machine Learning Lab 6

> #Task: Linear Regression Validation

>

> setwd("C:/Users/Ashish Upadhyay/Documents/Semester6/MachineLearning/Lab Programs")

> getwd()

[1] "C:/Users/Ashish Upadhyay/Documents/Semester6/MachineLearning/Lab Programs"

>

> drug = read.csv("drug2.csv")

> attach(drug)

> head(drug)

sex dose response

1 1 0.1 13.75

2 1 0.2 12.90

3 1 0.3 19.26

4 1 0.4 20.34

5 1 0.5 19.97

6 1 0.6 26.80

> nrow(drug)

[1] 3200

>

> model1 = lm(response~dose)

> summary(model1)

Call:

lm(formula = response ~ dose)

Residuals:

Min 1Q Median 3Q Max

-123.514 -62.764 0.401 63.669 124.707

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 7.2534 2.5778 2.814 0.00493 \*\*

dose 15.0020 0.4432 33.852 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

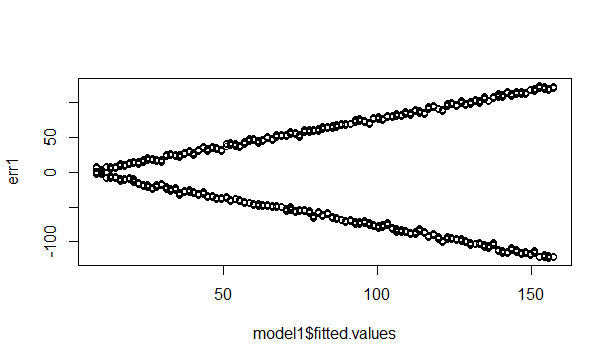
Residual standard error: 72.36 on 3198 degrees of freedom

Multiple R-squared: 0.2638, Adjusted R-squared: 0.2636

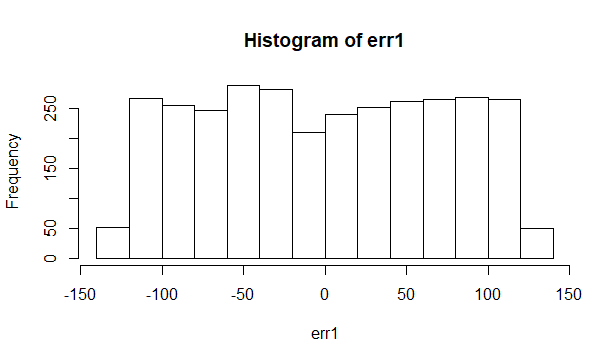
F-statistic: 1146 on 1 and 3198 DF, p-value: < 2.2e-16

> err1 = residuals(model1)

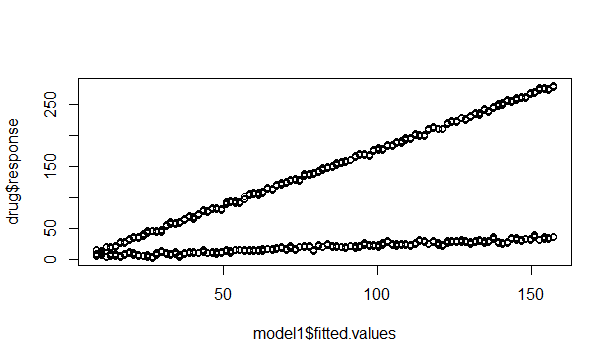
> plot(model1$fitted.values, err1)



> hist(err1)



> plot(model1$fitted.values, drug$response)



> mean <- mean(response)

> sst <- sum((response-mean)\*\*2)

> sse <- sum((response-model1$fitted.values)\*\*2)

> rsq <- 1 - (sse/sst)

> fuv <- sse/sst

> fuv

[1] 0.7361931

> rsq

[1] 0.2638069

> a <- (nrow(drug)-1)\*(1-rsq)

> b <- nrow(drug)-2

> rad <- 1-(a/b)

> rad

[1] 0.2635767