```
################## Problem 4 ################################



# Author: Md Salman Rahman
# Course: MATH 6364 Statistical Methods
# Course Instructor: Dr. George Yanev


library("readxl")
library(ggplot2)
squid_data<- read_excel("C:/Users/User/OneDrive - The University of Texas-Rio
Grande Valley/Course_video/Statistical Methods/HW_and R/Midterm Exam/
Squid_data.xlsx")


s_data_frame<- as.data.frame(squid_data)

attach(s_data_frame)




########### (a) generating the residual for the multiple regression model
###########

regg_1<-lm(y~ x1+x2+x3+x4+x5, data=s_data_frame)

summary(regg_1)




# residuals plot

residuals(regg_1)
plot(residuals(regg_1),xlab=" Observation",ylab="residual",main="Residual vs
Obs No.")






########### (b) computing the 95% confidence interval on the mean response
###########

# confidence interval
CI <- predict(regg_1, newdata = s_data_frame, interval = 'confidence')
```

```
CI_2 <- confint(regg_1, data=s_data_frame, interval ="confidence", level=0.95)

# prediction interval
PI <- predict(regg_1, newdata = s_data_frame, interval = 'prediction')




########### (c) new multiple regression using regressor x2,x4, and x5
###########

regg_2<-lm(y~ x2+x4+x5, data=s_data_frame)

summary(regg_2)


anova(regg_1,regg_2)


# confidence interval

CI_22 <- predict(regg_2, newdata = s_data_frame, interval = 'confidence')


# hypothesis testing


##########

model_restricted <- lm(y~x4+x5, data=s_data_frame)
summary(model_restricted)


anova(regg_1,model_restricted)


# hypothesis testing 2


n= 22

beta_hat_1 = 1.9994

beta_hat_2 = -3.6751

beta_hat_3= 2.5245

se_beta_hat_1 = 2.5733

se_beta_hat_2 = 2.7737

se_beta_hat_3 = 6.3475
```

```
t_statistics = qt(1-0.05/2, df=n-3)  # 95% CI

c(beta_hat_1-t_statistics*se_beta_hat_1,
beta_hat_1+t_statistics*se_beta_hat_1)

2 * pt(2.093024, df = n-3)
```