0R Code for Chapter 5

profsalary <- read.delim("C:/Teaching@cofc/Math 550/Chapter 5/profsalary.txt")

attach(profsalary)

plot(Salary~Experience)

sqe<-Experience^2

poly<-lm(Salary~Experience+sqe, data=profsalary)

xvalue<- seq(0, 36, 1)

predictcurve <- predict(poly,list(Experience=xvalue, sqe =xvalue^2)) #remodel

lines(xvalue, predictcurve, col = "darkgreen", lwd = 3)

# scatterplot seems good fit and now move on to diagnostic plots

plot(rstandard(poly)~Experience)

plot(sqrt(abs(rstandard(poly)))~ Experience)

plot(hatvalues(poly)~ Experience)

abline(h=0.042)

plot(poly)

summary(poly)

nexp<-10

nsqe<-100

predictpo <- predict(poly,list(Experience=nexp, sqe =nsqe),interval = "prediction", level = 0.95)

predictpo

library(readxl)

dataformath <- read\_excel("C:/Teaching@cofc/Math 550/Chapter 5/dataformath.xlsx")

View(dataformath)

lmm<-lm(y~x1+x2,data=dataformath)

|  |  |  |  |
| --- | --- | --- | --- |
| crti<-qt(0.975,2)  out<-summary(lmm)  lb<-out$coefficients[2,1]-crti\*out$coefficients[2,2]  ub<-out$coefficients[2,1]+crti\*out$coefficients[2,2]   |  | | --- | | confint(lmm,"x1",level=0.95)  # Hypothesis test could be done in the form of the confidence interval.  tn<-(out$coefficients[2,1]-1)/out$coefficients[2,2]  vcov(lmm) | |  | |  | |
|  |
| |  | | --- | |  | |

Forx<-c(1,-2, 4 ,1,-1 ,1, 1,0, 0, 1,1, 1, 1,2, 4)

Fory<-c(0,0,1,1,3)

Matrix\_y<-matrix(Fory,nrow=5,ncol=1,byrow=TRUE)

Matrix\_x<-matrix(Forx,nrow=5,ncol=3,byrow=TRUE)

Prod\_tx\_x<-t(Matrix\_x) %\*% Matrix\_x

Inv\_tx\_x=Inverse(Prod\_tx\_x)

Prod\_tx\_y= t(Matrix\_x) %\*% Matrix\_y

betahat<-Inv\_tx\_x %\*% Prod\_tx\_y

fitted\_y<- Matrix\_x %\*% betahat

Matrix\_fitted<-matrix(fitted\_y,nrow=5,ncol=1,byrow=TRUE)

residual<- Matrix\_y- Matrix\_fitted

RSS<- t(residual) %\*% residual

S\_square<- RSS/(nrow(dataformath)-(2+1))

Var\_beta\_hat<- as.vector(S\_square)\*Inv\_tx\_x

library(expm)

s\_beta\_hat<-sqrtm(diag(Var\_beta\_hat))

library(matlib)

Invmatrix<-Inverse(Matrix)

#transpose a matrix

Estimate<- Invmatrix

#assume model is valid for now and check the model assumption in next chapter.

nyc <- read.csv("C:/Teaching@cofc/Math 550/Chapter 5/nyc.csv")

lmf<-lm(Price~Food+Decor+Service+East, data=nyc)

summary(lmf)

# Food = customer rating of the food (out of 30)

*#* Décor = customer rating of the decor (out of 30)

*#* Service = customer rating of the service (out of 30)

#drop service because it is not useful with other predictors in the mdoel

# t test and F partial test

lmr<-lm(Price~Food+Decor+East, data=nyc)

summary(lmr)

sf<-summary(lmf)

sr<-summary(lmr)

rssf<-sum(sf$residuals^2)

rssr<-sum(sr$residuals^2)

F<-(rssr-rssf)/(rssf/sf$df[2])

1-pf(F, sr$df[2]- sf$df[2], sf$df[2])

anova(lmr,lmf)

travel <- read.delim("C:/Teaching@cofc/Math 550/Chapter 5/travel.txt")

attach(travel)

NC<-ifelse(Segment=="A", 1, 0)

AgeC<-Age\*C

LmA<-lm(Amount~Age, data=travel)

LmANCO<-lm(Amount~Age+C+AgeC, data=travel)

Outr<-summary(LmA)

Outf<-summary(LmANCO)

NumF<-(sum(Outr$residual^2)-sum(Outf$residual^2))/(Outr$df[2]-Outf$df[2])

denF<- sum(Outf$residual^2)/Outf$df[2]

Fcalc<-NumF/denF

Pvalc<-1-pf(Fcalc, Outr$df[2]-Outf$df[2], Outf$df[2])

anova(LmA, LmANCO)

nyc <- read.csv("C:/Teaching@cofc/Math 550/Chapter 5/nyc.csv")

attach(nyc)

Foodi<-Food\*East

Decori<-Decor\*East

Servicei<-Service\*East

Lmi<-lm(Price~Food+Decor+Service+East+Foodi+Decori+Servicei, data=nyc)

Lmwi<-lm(Price~Food+Decor+Service+East, data=nyc)

Lm<-lm(Price~Food+Decor+East, data=nyc)

Outr<-summary(Lm)

Outf<-summary(Lmi)

NumF<-(sum(Outr$residual^2)-sum(Outf$residual^2))/(Outr$df[2]-Outf$df[2])

denF<- sum(Outf$residual^2)/Outf$df[2]

Fcalc<-NumF/denF

Pvalc<-1-pf(Fcalc, Outr$df[2]-Outf$df[2], Outf$df[2])

anova(Lm,Lmi)

HoustonChronicle <- read.csv("C:/Teaching@cofc/Math 550/Chapter 5/HoustonChronicle.csv", header=TRUE)

attach(HoustonChronicle)

t<-ifelse(Year=="2004",1,0)

int<-X.Low.income.students\*t

lmfh<-lm(**X.Repeating.1st.Grade~X.Low.income.students+t+int, data=**HoustonChronicle)

lmrh<-lm(**X.Repeating.1st.Grade~X.Low.income.students, data=**HoustonChronicle)

anova(lmrh,lmfh)