#A Chernobyl Accident

Ho: B1 = 0, there is no relationship between distance and radiation.

Ha: B1 ≠ 0, there is a relationship between distance and radiation.

D<-c(10,20,50,100,200,500,1000)  
Rad<-c(155,121,110,84,45,10,4)  
A<-data.frame(D,Rad)

#1  
AModel<- lm(Rad~D,data = A)  
summary(AModel)

##   
## Call:  
## lm(formula = Rad ~ D, data = A)  
##   
## Residuals:  
## 1 2 3 4 5 6 7   
## 45.019 12.350 5.342 -14.004 -39.696 -34.774 25.763   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 111.3112 16.5231 6.737 0.00109 \*\*  
## D -0.1331 0.0383 -3.475 0.01776 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 34.21 on 5 degrees of freedom  
## Multiple R-squared: 0.7072, Adjusted R-squared: 0.6486   
## F-statistic: 12.07 on 1 and 5 DF, p-value: 0.01776

#2  
qf(.95,1,5)

## [1] 6.607891

pf(12.07,1,5,lower.tail = F)

## [1] 0.0177692

#Reject null hypothesis due to P value being smaller than the alpha.

#3  
#70.72% of the variation in the response can be explained by the regeresion model.

#4  
#34.21 is the residual standard error and it is small. The smaller it is the more predictive the model is.

#5  
radhat<-111.3112+-0.1331\*D  
radhat

## [1] 109.9802 108.6492 104.6562 98.0012 84.6912 44.7612 -21.7888

radhat<-AModel$coefficients[1]+AModel$coefficients[2]\*D  
radhat

## [1] 109.98051 108.64977 104.65756 98.00387 84.69649 44.77435 -21.76254

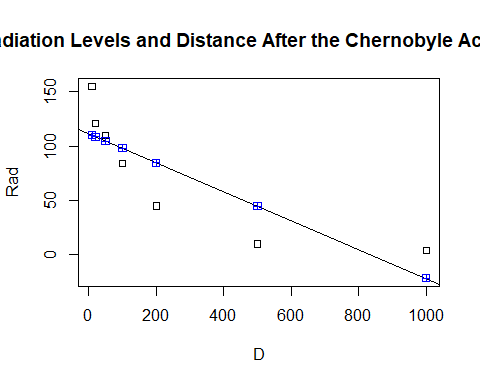
ymin<-min(c(Rad,radhat))  
ymax<-max(c(Rad,radhat))  
ymin

## [1] -21.76254

ymax

## [1] 155

plot(A,main = "Radiation Levels and Distance After the Chernobyle Accident",pch=22,ylim=(c(ymin,ymax)))  
abline(coef = coef(AModel))  
points(D, radhat,pch=12,col="blue")



#6  
RadLog = log(Rad)  
DLog = log(D)  
RDLog = data.frame(DLog,RadLog)  
summary(RDLog)

## DLog RadLog   
## Min. :2.303 Min. :1.386   
## 1st Qu.:3.454 1st Qu.:3.055   
## Median :4.605 Median :4.431   
## Mean :4.605 Mean :3.781   
## 3rd Qu.:5.756 3rd Qu.:4.748   
## Max. :6.908 Max. :5.043

RDLin<- lm(RadLog~DLog,data = RDLog)  
summary(RDLin)

##   
## Call:  
## lm(formula = RadLog ~ DLog, data = RDLog)  
##   
## Residuals:  
## 1 2 3 4 5 6 7   
## -0.5323 -0.2396 0.3793 0.6500 0.5661 -0.2237 -0.5997   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 7.3706 0.6793 10.850 0.000115 \*\*\*  
## DLog -0.7795 0.1398 -5.574 0.002559 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5722 on 5 degrees of freedom  
## Multiple R-squared: 0.8614, Adjusted R-squared: 0.8337   
## F-statistic: 31.07 on 1 and 5 DF, p-value: 0.002559

# .5722 is the residual standard error and it is very small. The smaller it is the more predictive the model is.

#7  
Rhatlog<-7.3706+-0.7795\*DLog  
Rhatlog

## [1] 5.575735 5.035427 4.321178 3.780870 3.240562 2.526313 1.986005

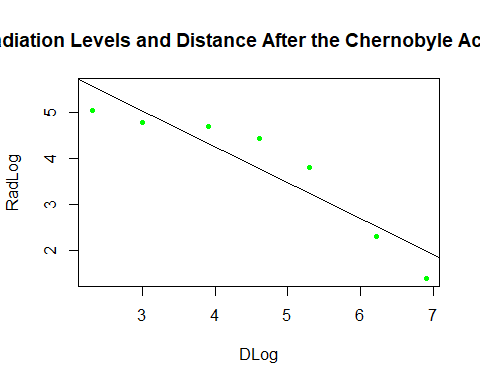
ymin<-min(c(RadLog,Rhatlog))  
ymax<-max(c(RadLog,Rhatlog))  
ymin

## [1] 1.386294

ymax

## [1] 5.575735

plot(RDLog,main = "Radiation Levels and Distance After the Chernobyle Accident",pch=20,col="green",ylim=(c(ymin,ymax)))  
abline(coef = coef(RDLin))  
points(D, Rhatlog,pch=10,col="green")



#B IQ and College Students

Ho: B1=B2=…=Bp=0

Ha: At least one B is not zero

getwd()

## [1] "C:/Users/erika/Desktop/Fall 2018/Stats Lab/Labs"

IQdata<- read.table("C:/Users/erika/Desktop/Fall 2018/Stats Lab/Labs/IQ\_scores.txt",header = T)

IQ35 <- head(IQdata, n=35)  
IQ35

## IQ Brain Height Weight  
## 1 147 85.65 70.5 155  
## 2 90 87.89 66.0 146  
## 3 96 86.54 68.0 135  
## 4 120 85.22 68.5 127  
## 5 102 94.51 73.5 178  
## 6 84 80.80 66.3 136  
## 7 86 88.91 70.0 180  
## 8 84 90.59 76.5 186  
## 9 134 79.06 62.0 122  
## 10 128 95.50 68.0 132  
## 11 102 83.18 63.0 114  
## 12 131 93.55 72.0 171  
## 13 88 92.59 74.5 178  
## 14 123 80.69 63.5 117  
## 15 149 102.84 72.3 142  
## 16 127 95.54 67.8 171  
## 17 133 94.15 64.0 146  
## 18 109 91.88 68.0 145  
## 19 130 98.13 63.5 137  
## 20 97 84.43 65.0 174  
## 21 89 93.59 75.5 179  
## 22 124 81.69 64.5 118  
## 23 150 103.84 73.3 143  
## 24 128 96.54 68.8 172  
## 25 134 95.15 65.0 147  
## 26 110 92.88 69.0 146  
## 27 131 99.13 64.5 138  
## 28 98 85.43 66.0 175  
## 29 84 90.49 66.3 134  
## 30 147 95.55 68.8 172  
## 31 124 83.39 64.5 118  
## 32 128 107.95 70.0 151  
## 33 124 92.41 69.0 155  
## 34 147 85.65 70.5 155  
## 35 90 87.89 66.0 146

IQ<-c(147,90,96,120,102,84,86,84,134,128,102,131,88,123,149,127,133,109,130,97,89,124,150,128,134,110,131,98,84,147,124,128,124,147,90)  
Brain<-c(85.65,87.89,86.54,85.22,94.51,80.80,88.91,90.59,79.06,95.50,83.18,93.55,92.59,80.69,102.84,95.54,94.15,91.88,98.13,84.43,93.59,81.69,103.84,96.54,95.15,92.88,99.13,85.43,90.49,95.55,83.39,107.95,92.41,85.65,87.89)  
Height<-c(70.5,66.0,68.0,68.5,73.5,66.3,70.0,76.5,62.0,68.0,63.0,72.0,74.5,63.5,72.3,67.8,64.0,68.0,63.5,65.0,75.5,64.5,73.3,68.8,65.0,69.0,64.5,66.0,66.3,68.8,64.5,70.0,69.0,70.5,66.0)  
Weight<-c(155,146,135,127,178,136,180,186,122,132,114,171,178,117,142,171,146,145,137,174,179,118,143,172,147,146,138,175,134,172,118,151,155,155,146)

IQDF<-data.frame(IQ,Brain,Height,Weight)  
IQDF

## IQ Brain Height Weight  
## 1 147 85.65 70.5 155  
## 2 90 87.89 66.0 146  
## 3 96 86.54 68.0 135  
## 4 120 85.22 68.5 127  
## 5 102 94.51 73.5 178  
## 6 84 80.80 66.3 136  
## 7 86 88.91 70.0 180  
## 8 84 90.59 76.5 186  
## 9 134 79.06 62.0 122  
## 10 128 95.50 68.0 132  
## 11 102 83.18 63.0 114  
## 12 131 93.55 72.0 171  
## 13 88 92.59 74.5 178  
## 14 123 80.69 63.5 117  
## 15 149 102.84 72.3 142  
## 16 127 95.54 67.8 171  
## 17 133 94.15 64.0 146  
## 18 109 91.88 68.0 145  
## 19 130 98.13 63.5 137  
## 20 97 84.43 65.0 174  
## 21 89 93.59 75.5 179  
## 22 124 81.69 64.5 118  
## 23 150 103.84 73.3 143  
## 24 128 96.54 68.8 172  
## 25 134 95.15 65.0 147  
## 26 110 92.88 69.0 146  
## 27 131 99.13 64.5 138  
## 28 98 85.43 66.0 175  
## 29 84 90.49 66.3 134  
## 30 147 95.55 68.8 172  
## 31 124 83.39 64.5 118  
## 32 128 107.95 70.0 151  
## 33 124 92.41 69.0 155  
## 34 147 85.65 70.5 155  
## 35 90 87.89 66.0 146

#1&2  
all<-lm(IQDF)  
summary(all)

##   
## Call:  
## lm(formula = IQDF)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -37.779 -15.568 -2.157 10.112 41.846   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 60.4928 68.1490 0.888 0.38156   
## Brain 1.5357 0.5360 2.865 0.00742 \*\*  
## Height -0.4818 1.2546 -0.384 0.70360   
## Weight -0.3414 0.2131 -1.602 0.11939   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 19.54 on 31 degrees of freedom  
## Multiple R-squared: 0.2568, Adjusted R-squared: 0.1848   
## F-statistic: 3.57 on 3 and 31 DF, p-value: 0.02508

qf(.95,3,31)

## [1] 2.911334

pf(3.57,3,31,lower.tail = F)

## [1] 0.02507167

#due to the p value being smaller than the alpha we will reject the null hypothesis.

#2  
IQBrain<-(lm(IQ~Brain,IQDF))  
IQBrain

##   
## Call:  
## lm(formula = IQ ~ Brain, data = IQDF)  
##   
## Coefficients:  
## (Intercept) Brain   
## 16.501 1.097

summary(IQBrain)

##   
## Call:  
## lm(formula = IQ ~ Brain, data = IQDF)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -31.834 -16.763 5.738 14.661 36.582   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 16.5012 46.4531 0.355 0.7247   
## Brain 1.0965 0.5093 2.153 0.0387 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 20.57 on 33 degrees of freedom  
## Multiple R-squared: 0.1232, Adjusted R-squared: 0.09658   
## F-statistic: 4.635 on 1 and 33 DF, p-value: 0.03873

#2  
IQHeight<-(lm(IQ~Height,IQDF))  
IQHeight

##   
## Call:  
## lm(formula = IQ ~ Height, data = IQDF)  
##   
## Coefficients:  
## (Intercept) Height   
## 152.5669 -0.5334

summary(IQHeight)

##   
## Call:  
## lm(formula = IQ ~ Height, data = IQDF)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -33.205 -20.599 5.835 13.668 36.528   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 152.5669 68.3158 2.233 0.0324 \*  
## Height -0.5334 1.0012 -0.533 0.5978   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 21.87 on 33 degrees of freedom  
## Multiple R-squared: 0.008526, Adjusted R-squared: -0.02152   
## F-statistic: 0.2838 on 1 and 33 DF, p-value: 0.5978

#2  
IQWeight<-(lm(IQ~Weight,IQDF))  
IQWeight

##   
## Call:  
## lm(formula = IQ ~ Weight, data = IQDF)  
##   
## Coefficients:  
## (Intercept) Weight   
## 151.8752 -0.2381

summary(IQWeight)

##   
## Call:  
## lm(formula = IQ ~ Weight, data = IQDF)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -35.976 -20.883 0.215 15.856 36.070   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 151.8752 26.3739 5.759 1.96e-06 \*\*\*  
## Weight -0.2381 0.1745 -1.364 0.182   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 21.37 on 33 degrees of freedom  
## Multiple R-squared: 0.0534, Adjusted R-squared: 0.02472   
## F-statistic: 1.862 on 1 and 33 DF, p-value: 0.1817

#3 19.54 is the residual standard error and it is small. The smaller it is the more predictive the model is. Due to this we reject the null hypothesis.

#4  
IQCombo<-(step(lm(IQ~Brain+Height+Weight,IQDF),direction = "both",test="F"))

## Start: AIC=211.81  
## IQ ~ Brain + Height + Weight  
##   
## Df Sum of Sq RSS AIC F value Pr(>F)   
## - Height 1 56.28 11887 209.98 0.1475 0.703599   
## <none> 11831 211.81   
## - Weight 1 978.97 12810 212.59 2.5651 0.119386   
## - Brain 1 3132.62 14964 218.03 8.2082 0.007423 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Step: AIC=209.98  
## IQ ~ Brain + Weight  
##   
## Df Sum of Sq RSS AIC F value Pr(>F)   
## <none> 11887 209.98   
## + Height 1 56.3 11831 211.81 0.1475 0.703599   
## - Weight 1 2070.5 13958 213.60 5.5738 0.024488 \*   
## - Brain 1 3180.9 15068 216.28 8.5627 0.006265 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#5  
IQBW<-lm(IQ~Brain+Weight,IQDF)  
summary(IQBW)

##   
## Call:  
## lm(formula = IQ ~ Brain + Weight, data = IQDF)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -37.72 -16.08 -0.21 11.37 40.63   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 40.9619 44.7499 0.915 0.36685   
## Brain 1.4724 0.5032 2.926 0.00627 \*\*  
## Weight -0.3916 0.1659 -2.361 0.02449 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 19.27 on 32 degrees of freedom  
## Multiple R-squared: 0.2532, Adjusted R-squared: 0.2066   
## F-statistic: 5.426 on 2 and 32 DF, p-value: 0.009354

qf(.95,2,32)

## [1] 3.294537

pf(5.426,2,32,lower.tail = F)

## [1] 0.009351087

#The p value is smaller than the alpha so we reject the null hypothesis. There are large improvements with the p-value which is more accurate and the critical value which is larger. This showed us which one was more significant, which was brain and weight. By only doing the ones that are more significant we were able to get rid of noise that was making the data less acurate according to the linear regression.