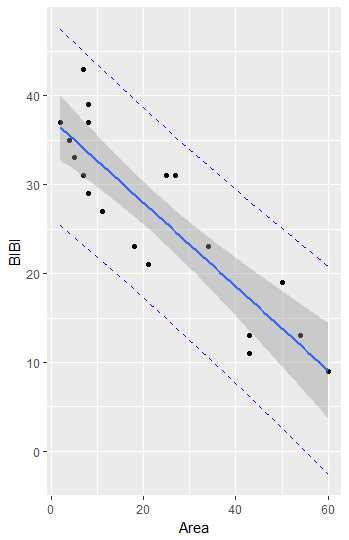
Anthony Le

ST 511 (Tu 16:00-16:50)

22 November 2016

Homework 8

Question 1:

>Streams <- read.csv("Streams.csv")

>summary(Streams)

>head(Streams)

>Streams.lm <- lm(BIBI~Area,data=Streams)

>summary(Streams.lm)

>Streams.lm$coefficients

>qplot(Area,BIBI,data=Streams)+geom\_abline(slope=-0.4710054,intercept=37.3624929)

>ggplot(Streams, aes(x=Area, y=BIBI))+

+ geom\_point()+

+ geom\_smooth(method=lm,se=FALSE)

>ggplot(Streams, aes(x=Area, y=BIBI))+

+ geom\_point()+

+ geom\_smooth(method=lm)

>pred.intervals <- data.frame(predict(Streams.lm,interval="prediction"))

>head(pred.intervals)

>Streams.df2<-cbind(Streams,pred.intervals)

>head(Streams.df2)

>ggplot(Streams.df2, aes(x=Area, y=BIBI))+

+ geom\_point()+

+ geom\_line(aes(y=lwr), color="blue", linetype="dashed")+

+ geom\_line(aes(y=upr), color="blue",linetype="dashed")+

+ geom\_smooth(method=lm)

Question 2:

> Streams.lm <- lm(BIBI~Area,data=Streams)

> summary(Streams.lm)

Question 3:

> Streams.lm <- lm(BIBI~Area,data=Streams)

> summary(Streams.lm)

> qt(0.975,17)

> 37.36249+qt(0.975,17)\*1.80228

> 37.36249-qt(0.975,17)\*1.80228

The relationship between mean measured area and BIBI approximates a straight line (p-value = 1.66e-13, from two-sided test that the intercept is zero). It is estimated that the mean BIBI of the streams is 37.36 (95% CI 33.56 to 41.16).

Question 4:

These data are not a random sample. The streams were selected to examine the effects of urbanization on the biotic integrity of the Puget Sound Lowlands and specifically represent the range of land use conditions in King, Kitsap, Snohomish, and Thurston counties so they do not necessarily represent that would result from taking measurements from other streams in other counties. This analysis assumes that there is an exact linear relationship between area and BIBI but that the measured BIBIs do not all fall exactly on the straight line because of measurement errors in BIBI measurements.