Week9

Lara Burford

Monday, July 28, 2014

##Brett's code  
library(RSQLite)

## Warning: package 'RSQLite' was built under R version 3.1.1

## Loading required package: DBI

setwd('C:/Users/Lara/Downloads')  
sqlite <- dbDriver("SQLite")  
workingdb <- dbConnect(sqlite, "cunyweek9.sqlite")  
dbListTables(workingdb)

## [1] "I" "II" "III" "IV"

I <- dbSendQuery(workingdb, "SELECT \* FROM I")  
Idf <- data.frame(fetch(I, -1))  
II <- dbSendQuery(workingdb, "SELECT \* FROM II")  
IIdf <- data.frame(fetch(II, -1))  
III <- dbSendQuery(workingdb, "SELECT \* FROM III")  
IIIdf <- data.frame(fetch(III, -1))  
IV <- dbSendQuery(workingdb, "SELECT \* FROM IV")  
IVdf <- data.frame(fetch(IV, -1))  
dbDisconnect(workingdb)

## Warning: RS-DBI driver warning: (closing pending result sets before  
## closing this connection)

## [1] TRUE

remove(sqlite)  
remove(workingdb)  
remove(I, II, III, IV)  
par(mfrow=c(2,2))  
plot(Idf, main="I")  
abline(lm(Idf$y ~ Idf$x))  
plot(IIdf, main="II")  
abline(lm(IIdf$y ~ IIdf$x))  
plot(IIIdf, main="III")  
abline(lm(IIIdf$y ~ IIIdf$x))  
plot(IVdf, main="IV")  
abline(lm(IVdf$y ~ IVdf$x))

plot of chunk - week9

##Sonja's code   
summary(Idf)

## x y   
## Min. : 4.0 Min. : 4.26   
## 1st Qu.: 6.5 1st Qu.: 6.32   
## Median : 9.0 Median : 7.58   
## Mean : 9.0 Mean : 7.50   
## 3rd Qu.:11.5 3rd Qu.: 8.57   
## Max. :14.0 Max. :10.84

summary(IIdf)

## x y   
## Min. : 4.0 Min. :3.10   
## 1st Qu.: 6.5 1st Qu.:6.70   
## Median : 9.0 Median :8.14   
## Mean : 9.0 Mean :7.50   
## 3rd Qu.:11.5 3rd Qu.:8.95   
## Max. :14.0 Max. :9.26

summary(IIIdf)

## x y   
## Min. : 4.0 Min. : 5.39   
## 1st Qu.: 6.5 1st Qu.: 6.25   
## Median : 9.0 Median : 7.11   
## Mean : 9.0 Mean : 7.50   
## 3rd Qu.:11.5 3rd Qu.: 7.98   
## Max. :14.0 Max. :12.74

summary(IVdf)

## x y   
## Min. : 8 Min. : 5.25   
## 1st Qu.: 8 1st Qu.: 6.17   
## Median : 8 Median : 7.04   
## Mean : 9 Mean : 7.50   
## 3rd Qu.: 8 3rd Qu.: 8.19   
## Max. :19 Max. :12.50

par(mfrow=c(2,2))  
Ilm <- lm(formula = y~x,data = Idf)  
plot(Ilm)

plot of chunk - week9

par(mfrow=c(2,2))  
IIlm <- lm(formula = y~x,data = IIdf)  
plot(IIlm)

plot of chunk - week9

par(mfrow=c(2,2))  
IIIlm <- lm(formula = y~x,data = IIIdf)  
plot(IIIlm)

plot of chunk - week9

par(mfrow=c(2,2))  
IVlm <- lm(formula = y~x,data = IVdf)  
plot(IVlm)

## Warning: not plotting observations with leverage one:  
## 8  
## Warning: not plotting observations with leverage one:  
## 8

plot of chunk - week9

##Tom's code  
week9df <- as.data.frame(cbind(Idf,IIdf,IIIdf,IVdf))  
str(week9df)

## 'data.frame': 11 obs. of 8 variables:  
## $ x: int 10 8 13 9 11 14 6 4 12 7 ...  
## $ y: num 8.04 6.95 7.58 8.81 8.33 ...  
## $ x: int 10 8 13 9 11 14 6 4 12 7 ...  
## $ y: num 9.14 8.14 8.74 8.77 9.26 8.1 6.13 3.1 9.13 7.26 ...  
## $ x: int 10 8 13 9 11 14 6 4 12 7 ...  
## $ y: num 7.46 6.77 12.74 7.11 7.81 ...  
## $ x: int 8 8 8 8 8 8 8 19 8 8 ...  
## $ y: num 6.58 5.76 7.71 8.84 8.47 7.04 5.25 12.5 5.56 7.91 ...

cuny9 <- week9df  
head(cuny9)

## x y x y x y x y  
## 1 10 8.04 10 9.14 10 7.46 8 6.58  
## 2 8 6.95 8 8.14 8 6.77 8 5.76  
## 3 13 7.58 13 8.74 13 12.74 8 7.71  
## 4 9 8.81 9 8.77 9 7.11 8 8.84  
## 5 11 8.33 11 9.26 11 7.81 8 8.47  
## 6 14 9.96 14 8.10 14 8.84 8 7.04

names(cuny9) <- (c("x1","y1","x2","y2","x3","y3","x4","y4"))  
pairs(cuny9)  
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.1.1

plot of chunk - week9

with(cuny9,qplot(x1,y1))

plot of chunk - week9

with(cuny9,qplot(x2,y2))

plot of chunk - week9

with(cuny9,qplot(x3,y3))

plot of chunk - week9

with(cuny9,qplot(x4,y4))

plot of chunk - week9

##Riguel's code  
influence.measures(Ilm)

## Influence measures of  
## lm(formula = y ~ x, data = Idf) :  
##   
## dfb.1\_ dfb.x dffit cov.r cook.d hat inf  
## 1 0.00033 3.15e-03 0.0104 1.406 6.14e-05 0.1000   
## 2 -0.00818 4.11e-03 -0.0136 1.406 1.04e-04 0.1000   
## 3 0.61888 -9.08e-01 -1.1578 0.698 4.89e-01 0.2364   
## 4 0.11812 1.40e-17 0.3563 1.037 6.16e-02 0.0909   
## 5 0.01197 -2.85e-02 -0.0534 1.443 1.60e-03 0.1273   
## 6 0.01618 -2.21e-02 -0.0261 1.856 3.83e-04 0.3182 \*  
## 7 0.45415 -3.51e-01 0.5104 1.145 1.27e-01 0.1727   
## 8 -0.46907 4.07e-01 -0.4813 1.646 1.23e-01 0.3182   
## 9 -0.34342 5.78e-01 0.8400 0.756 2.79e-01 0.1727   
## 10 -0.46987 3.20e-01 -0.5990 0.848 1.54e-01 0.1273   
## 11 0.08250 -6.84e-02 0.0872 1.647 4.27e-03 0.2364

influence.measures(IIlm)

## Influence measures of  
## lm(formula = y ~ x, data = IIdf) :  
##   
## dfb.1\_ dfb.x dffit cov.r cook.d hat inf  
## 1 0.0102 9.72e-02 0.3223 1.127 0.05233 0.1000   
## 2 0.1936 -9.72e-02 0.3223 1.127 0.05233 0.1000   
## 3 0.2030 -2.98e-01 -0.3798 1.480 0.07666 0.2364   
## 4 0.1139 9.91e-18 0.3436 1.057 0.05787 0.0909   
## 5 -0.0543 1.30e-01 0.2423 1.315 0.03145 0.1273   
## 6 0.9475 -1.29e+00 -1.5278 0.703 0.80787 0.3182 \*  
## 7 0.0440 -3.40e-02 0.0495 1.525 0.00137 0.1727   
## 8 -1.4889 1.29e+00 -1.5278 0.703 0.80787 0.3182 \*  
## 9 -0.0202 3.40e-02 0.0495 1.525 0.00137 0.1727   
## 10 0.1901 -1.30e-01 0.2423 1.315 0.03145 0.1273   
## 11 -0.3591 2.98e-01 -0.3798 1.480 0.07666 0.2364

influence.measures(IIIlm)

## Influence measures of  
## lm(formula = y ~ x, data = IIIdf) :  
##   
## dfb.1\_ dfb.x dffit cov.r cook.d hat inf  
## 1 -4.63e-03 -4.41e-02 -0.1464 1.34e+00 0.011765 0.1000   
## 2 -3.71e-02 1.86e-02 -0.0618 1.39e+00 0.002141 0.1000   
## 3 -3.58e+02 5.25e+02 669.5875 5.06e-11 1.392849 0.2364 \*  
## 4 -3.29e-02 -3.37e-18 -0.0992 1.36e+00 0.005473 0.0909   
## 5 4.92e-02 -1.17e-01 -0.2193 1.34e+00 0.025984 0.1273   
## 6 4.90e-01 -6.67e-01 -0.7897 1.36e+00 0.300571 0.3182   
## 7 2.70e-02 -2.09e-02 0.0303 1.53e+00 0.000518 0.1727   
## 8 2.41e-01 -2.09e-01 0.2472 1.80e+00 0.033817 0.3182 \*  
## 9 1.37e-01 -2.31e-01 -0.3362 1.34e+00 0.059536 0.1727   
## 10 -1.97e-02 1.34e-02 -0.0251 1.45e+00 0.000355 0.1273   
## 11 1.05e-01 -8.74e-02 0.1114 1.64e+00 0.006948 0.2364

influence.measures(IVlm)

## Influence measures of  
## lm(formula = y ~ x, data = IVdf) :  
##   
## dfb.1\_ dfb.x dffit cov.r cook.d hat inf  
## 1 -0.06827 0.03428 -0.1137 1.366 7.17e-03 0.1   
## 2 -0.21353 0.10721 -0.3556 1.078 6.23e-02 0.1   
## 3 0.11654 -0.05851 0.1941 1.294 2.03e-02 0.1   
## 4 0.34734 -0.17439 0.5784 0.742 1.37e-01 0.1   
## 5 0.26029 -0.13069 0.4334 0.958 8.72e-02 0.1   
## 6 0.00628 -0.00315 0.0105 1.406 6.15e-05 0.1   
## 7 -0.32505 0.16320 -0.5413 0.795 1.24e-01 0.1   
## 8 0.00000 0.00000 NaN NaN NaN 1.0 \*  
## 9 -0.25432 0.12769 -0.4235 0.974 8.39e-02 0.1   
## 10 0.15149 -0.07606 0.2523 1.225 3.34e-02 0.1   
## 11 -0.01788 0.00898 -0.0298 1.403 4.98e-04 0.1

##Lara's code  
library(lmtest)

## Warning: package 'lmtest' was built under R version 3.1.1

## Loading required package: zoo

## Warning: package 'zoo' was built under R version 3.1.1

##   
## Attaching package: 'zoo'  
##   
## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

dwtest(Ilm)

##   
## Durbin-Watson test  
##   
## data: Ilm  
## DW = 3.212, p-value = 0.988  
## alternative hypothesis: true autocorrelation is greater than 0

dwtest(IIlm)

##   
## Durbin-Watson test  
##   
## data: IIlm  
## DW = 2.188, p-value = 0.6299  
## alternative hypothesis: true autocorrelation is greater than 0

dwtest(IIIlm)

##   
## Durbin-Watson test  
##   
## data: IIIlm  
## DW = 2.144, p-value = 0.6016  
## alternative hypothesis: true autocorrelation is greater than 0

dwtest(IVlm)

##   
## Durbin-Watson test  
##   
## data: IVlm  
## DW = 1.662, p-value = 0.2892  
## alternative hypothesis: true autocorrelation is greater than 0