Homework 2

Yubing Li

00808366

1. Nursing home Utilization:

(a)

> nursing= read.csv("WiscNursingHome.csv")

> nurs=nursing[which(nursing$CRYEAR==2001),]

> TPY<-nurs$TPY

> LOGTPY<-log(TPY)

> NUMBED<-nurs$NUMBED

> LOGNUMBED<-log(NUMBED)

> summary(LOGTPY)

Min. 1st Qu. Median Mean 3rd Qu. Max.

2.511 4.041 4.396 4.368 4.700 6.088

> summary(LOGNUMBED)

Min. 1st Qu. Median Mean 3rd Qu. Max.

2.890 4.094 4.500 4.457 4.779 6.125

For natural log of total patients years in 2001, with “patients” as units:

The minimum is 2.511.

The 1st quantile is 4.041.

The median is 4.396.

The 3rd quantile is 4.700.

The maximum is 6.088.

The mean is 4.368.

For natural log of the number of beds in 2001, with “beds” as units:

The minimum is 2.890.

The 1st quantile is 4.094.

The median is 4.500.

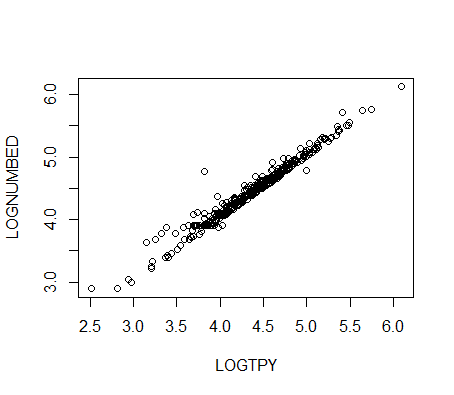
The 3rd quantile is 4.779.

The maximum is 6.125.

The mean is 4.457.

Scatter Plot:

In the scatter plot, the LOGTPY is the explanatory variable and LOGNUMBED is the response variable.



> plot(LOGNUMBED ~ LOGTPY)

> lm.out<-lm(LOGNUMBED ~ LOGTPY)

> summary(lm.out)

Call:

lm(formula = LOGNUMBED ~ LOGTPY)

Residuals:

Min 1Q Median 3Q Max

-0.26255 -0.04970 -0.01103 0.03080 0.83999

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.315492 0.041397 7.621 2.34e-13 \*\*\*

LOGTPY 0.948146 0.009413 100.730 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.09041 on 353 degrees of freedom

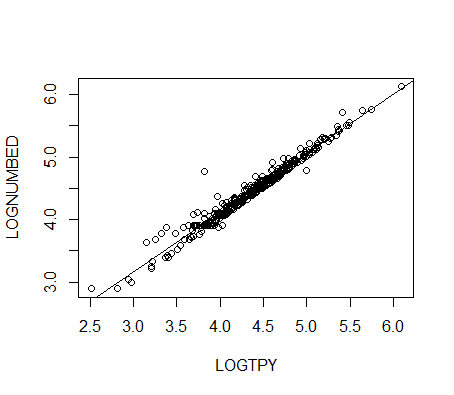
Multiple R-squared: 0.9664, Adjusted R-squared: 0.9663

F-statistic: 1.015e+04 on 1 and 353 DF, p-value: < 2.2e-16

Correlation coefficient for these two variables is:

R2 = 0.9664 r = +√R2 = 0.9831

The r value is very close to 1, which indicates a strong linear association between LOGTPY and LOGNUMBED in 2001. It also indicates that 96.64% of the variability in the LOGNUMBED is accounted for by variation in LOGTPY. The scatter plot also suggests that the two variables have strong positive linear association. In general, the association between LOGTPY and LOGNUMBED is strong, positive linear and there’s no unusual features.

(b) > abline(lm.out)