MATH-650 Assignment 6

Saket Choudhary (USCID: 2170058637) (skchoudh@usc.edu)

09/15/2015

# Problem 19

pH.data <- read.csv('case0702.csv', header=T)
logT <- log(pH.data$Time)
pH.data$logT <- logT
n <- nrow(pH.data)

Simple linear regression model for $log(pH)$ is given by:

$$pH=β\_{0}+β\_{1}log(T)$$

### Part (a)

fit <- lm(pH~logT, data=pH.data)
s <- summary(fit)
b0 <- fit$coefficients[1]
b1 <- fit$coefficients[2]
se0 <- s$coefficients[3]
se1 <- s$coefficients[4]
t0 <- s$coefficients[5]
t1 <- s$coefficients[6]
p0 <- s$coefficients[7]
p1 <- s$coefficients[8]

Thus, $β\_{0}=6.983626(S.E=0.048532,p−value=6.0839895×10^{−15})$ and $β\_{1}=−0.7256578(S.E=0.0344263,p−value=2.6951582×10^{−8})$

### Part (b)

Xbar <- mean(log(pH.data$Time))#1.190
sx2 <- var(log(pH.data$Time))#0.6344
mu <- b0 +b1\*log(5)

Thus, $^\{Y|log(5)\}=5.8157249$

### Part (c)

sigmahat <- 0.08226
se = sigmahat \*sqrt(1/n+(log(5)-Xbar)^2/(n-1\*sx2))

$SE[^\{Y|log(5)\}]=0.0283496$

# Problem 25

### Part (a)

$$\begin{matrix}SS(β\_{0},β\_{1})&=\sum(\_{i=1}^{N}Y\_{i}−β\_{0}−β\_{1}X\_{i})^{2}\\\frac{∂SS}{∂β\_{0}}&=−2\sum(\_{i=1}^{N}Y\_{i}−β\_{0}−β\_{1}X\_{i})\\\frac{∂SS}{∂β\_{1}}&=−2\sum(\_{i=1}^{N}Y\_{i}−β\_{0}−β\_{1}X\_{i})X\_{i}\end{matrix}$$

$$\begin{matrix}\frac{∂SS}{∂β\_{0}}&=−2\sum(\_{i=1}^{N}Y\_{i}−β\_{0}−β\_{1}X\_{i})=0\\\sumY\_{i}\_{i=1}^{N}&=\sum(\_{i=1}^{N}β\_{0}+β\_{1}X\_{i})\\\sumY\_{i}\_{i=1}^{N}&=β\_{0}n+β\_{1}\sumX\_{i}\_{i=1}^{N}\end{matrix}$$

$$\begin{matrix}\frac{∂SS}{∂β\_{1}}&=−2\sum(\_{i=1}^{N}Y\_{i}−β\_{0}−β\_{1}X\_{i})X\_{i}=0\\\sumY\_{i}\_{i=1}^{N}X\_{i}&=\sum(\_{i=1}^{N}β\_{0}+β\_{1}X\_{i})X\_{i}\\\sumY\_{i}\_{i=1}^{N}X\_{i}&=β\_{0}\sumX\_{i}\_{i=1}^{N}+\sumβ\_{1}\_{i=1}^{N}X\_{i}^{2}\end{matrix}$$

### Part (b)

$$\begin{matrix}^&=\overline{Y}−^\overline{X}\end{matrix}$$

And,

$$\begin{matrix}^&=\frac{\sum(\_{i=1}^{N}X\_{i}−\overline{X})(Y\_{i}−\overline{Y})}{\sum(\_{i=1}^{N}X\_{i}−\overline{X})^{2}}\\&=\frac{\sum(\_{i=1}^{N}X\_{i}Y\_{i}−\overline{X}Y\_{i}+\overline{X}\overline{Y}−X\_{i}\overline{Y})}{\sum(\_{i=1}^{N}X\_{i}^{2}−2\overline{X}X\_{i}+\overline{X}^{2})}\\&=\frac{\sumX\_{i}\_{i=1}^{N}Y\_{i}−n\overline{X}\overline{Y}}{\sumX\_{i}^{2}\_{i=1}^{N}−n\overline{X}^{2}}\end{matrix}$$

Thus,

$$\begin{matrix}^\sumX\_{i}^{2}\_{i=1}^{N}−n^\overline{X}^{2}&=\sumX\_{i}\_{i=1}^{N}Y\_{i}−n\overline{X}\overline{Y}\\^\sumX\_{i}^{2}\_{i=1}^{N}+n\overline{X}\overline{Y}−n^\overline{X}^{2}&=\sumX\_{i}\_{i=1}^{N}Y\_{i}\\^\sumX\_{i}^{2}\_{i=1}^{N}+n\overline{Y}(\overline{X}−^\overline{X})&=\sumX\_{i}\_{i=1}^{N}Y\_{i}\end{matrix}$$

Substituting (3) in (4), we get

$$\begin{matrix}^\sumX\_{i}^{2}\_{i=1}^{N}+n\overline{X}(\overline{Y}−^\overline{X})&=\sumX\_{i}\_{i=1}^{N}Y\_{i}\\^\sumX\_{i}^{2}\_{i=1}^{N}+n\overline{X}(^)\overline{X}&=\sumX\_{i}\_{i=1}^{N}Y\_{i}\\^\sumX\_{i}^{2}\_{i=1}^{N}+^\sumX\_{i}\_{i=1}^{N}&=\sumX\_{i}\_{i=1}^{N}Y\_{i}\end{matrix}$$

Thus, (5) is same as (2)

Now, From (3)

$$\begin{matrix}^&=\overline{Y}−^\overline{X}\\\overline{Y}&=^+^\overline{X}\\\frac{\sumY\_{i}\_{i=1}^{N}}{N}&=^+^\frac{\sumX\_{i}\_{i=1}^{N}}{N}\\\sumY\_{i}\_{i=1}^{N}&=n^+^\sumX\_{i}\_{i=1}^{N}\end{matrix}$$

Thus, (6) is same as (1)

Consider second order differentials:

$$\begin{matrix}\frac{∂^{2}SS}{∂β\_{0}^{2}}&=−2\*−n=2n\\\frac{∂^{2}SS}{∂β\_{0}∂β\_{1}}&=2\sumX\_{i}\_{i=1}^{N}\\\frac{∂^{2}SS}{∂β\_{1}∂β\_{0}}&=2\sumX\_{i}\_{i=1}^{N}\\\frac{∂^{2}SS}{∂β\_{1}^{2}}&=2\sumX\_{i}^{2}\_{i=1}^{N}\\Hessian(H)&=\left[\begin{matrix}2&2\sumX\_{i}\_{i=1}^{N}\\2\sumX\_{i}\_{i=1}^{N}&2\sumX\_{i}^{2}\_{i=1}^{N}\end{matrix}\right]\\2n>0and, \\det(H)&=4(n\sumX\_{i}^{2}\_{i=1}^{N}−(\sumX\_{i}\_{i=1}^{N})^{2})=4n(n−1)Var(X)>0 always\end{matrix}$$

Thus, the values of $^,^$ indeed guarantee a minima since the Hessian is positive definitee.

### Part (c)

# Problem 13

Intercept = 0.3991 Standard Error = 0.1185 df = 22 $t\_{0.975,22}=2.073$

Upper limit = 0.3991 + 2.073 \* 0.1185 = 0.6447505 Lower limit = 0.3991 - 2.073 \* 0.1185 = 0.1534495

Thus, the 95% CI for Intercept is [0.1534495, 0.6447505]

Also following is the R code:

intercept <- 0.3991
se <- 0.1185
df <- 22
t975 <- qt(0.975,df)
limit.upper <- intercept + t975\*se
limit.lower <- intercept - t975\*se
limit.upper

## [1] 0.644854

limit.lower

## [1] 0.153346