### HW#2, Nan Deng

#### (1)

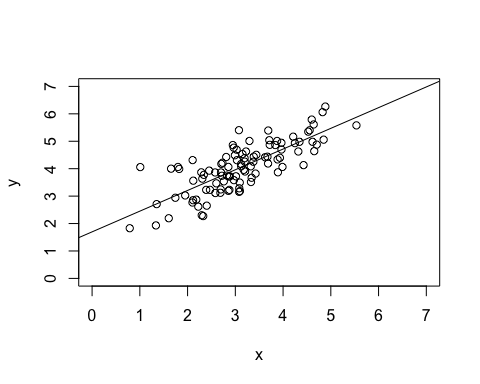
library(MASS)  
mu <- c(3,4)  
sigma <- matrix(c(1.0,0.8,0.8,1.0),nrow=2)  
set.seed(123)  
datam <- data.frame(mvrnorm(100,mu,sigma))  
colnames(datam) <- c("x","y")

#### (a)

library(faraway)  
fit\_line <- lm(y ~ x,data=datam)  
summary(fit\_line)

##   
## Call:  
## lm(formula = y ~ x, data = datam)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.18120 -0.44633 -0.01159 0.36843 1.60020   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.69698 0.19998 8.486 2.31e-13 \*\*\*  
## x 0.75479 0.06144 12.285 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5701 on 98 degrees of freedom  
## Multiple R-squared: 0.6063, Adjusted R-squared: 0.6023   
## F-statistic: 150.9 on 1 and 98 DF, p-value: < 2.2e-16

plot(datam$x,datam$y,xlab="x",ylab="y",xlim=c(0,7),ylim=c(0,7))  
abline(fit\_line)



coefficients(fit\_line)

## (Intercept) x   
## 1.6969772 0.7547914

summary(fit\_line)$sigma

## [1] 0.5701041

summary(fit\_line)$r.squared

## [1] 0.606287

#### (b) α=1.6969772; β=0.7547914; σ=0.5701041; R2 =0.606287

#### (c)

t.test(datam, mu=0, alternative="greater", conf.level=0.95)

##   
## One Sample t-test  
##   
## data: datam  
## t = 49.314, df = 199, p-value < 2.2e-16  
## alternative hypothesis: true mean is greater than 0  
## 95 percent confidence interval:  
## 3.465606 Inf  
## sample estimates:  
## mean of x   
## 3.585767

#### Considering that p-value in this model is less than 2.2e-16 and definitely less than 0.05, H0 should be rejected.

#### (d)

fit\_line1 <- lm(y ~ offset(0.9\*x),data=datam)  
anova(fit\_line,fit\_line1)

## Analysis of Variance Table  
##   
## Model 1: y ~ x  
## Model 2: y ~ offset(0.9 \* x)  
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 98 31.852   
## 2 99 33.667 -1 -1.8154 5.5854 0.02008 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#### Since p-value of this model is around 0.02, which is less than 0.05, H0 should also be rejected.

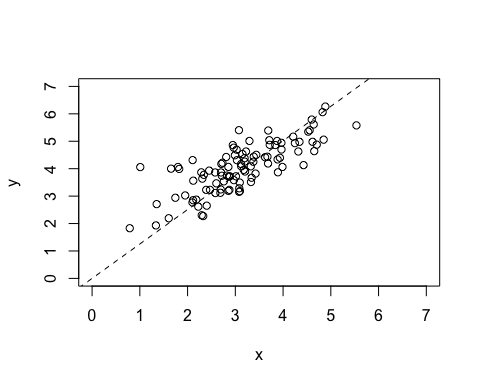
#### (2)

m0 <- lm(y ~ x -1 , data=datam)   
summary(m0)

##   
## Call:  
## lm(formula = y ~ x - 1, data = datam)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.4258 -0.3656 0.1252 0.5823 2.7931   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## x 1.25453 0.02295 54.66 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.7471 on 99 degrees of freedom  
## Multiple R-squared: 0.9679, Adjusted R-squared: 0.9676   
## F-statistic: 2987 on 1 and 99 DF, p-value: < 2.2e-16

#### (a)

plot(datam$x,datam$y,xlab="x",ylab="y",xlim=c(0,7),ylim=c(0,7))  
abline(m0,lty=2)



#### (b)

sum(summary(m0)$resid)

## [1] 13.79104

#### No, the sum of residuals is about 13.8 rather than 0.

#### (c)

sum(datam$x \* residuals(m0))

## [1] -4.010681e-15

sum(fitted.values(m0) \* residuals(m0))

## [1] -4.458239e-15

#### Macintosh HD:Users:CandiceDeng 1:Desktop:2c.JPG

#### (3)

#### (a)

#### Macintosh HD:Users:CandiceDeng 1:Desktop:3a.JPG

In this equation, r and β have the same sign.

#### (b)

