lecture 4.10 example 2

preliminaries

```
#clear workspace
rm(list=ls())
#set random seed
set.seed(123)

# import libraries
library(MASS)
library(invgamma)

# load data
data = read.csv("~/Desktop/wages1.csv")

# build variables
wage = data$wage
school = data$school
exper = data$exper
gender = rep(0, length(wage))
gender[data$sex=="female"]=1
```

fit interaction model using OLS (pooling all data) and use fits to build posteriors

```
sC = school - mean(school)
eC = exper - mean(exper)
m = lm(log(wage) ~ sC + gender + gender * sC  )
```

```
#posterior parameters
betaHat = coef(m)
sigmaHat = vcov(m)
shape = (length(wage) - 3) / 2
rate = sum(resid(m)^2)/2
```

Compute gender wage differences for different levels of educational differences

```
deltaSchool = -6:6
nSimsPer = 10000

yMale = rep(-1,nSimsPer)
yFemale = array(rep(-1, length(deltaSchool) * nSimsPer),
                dim = c(length(deltaSchool), nSimsPer))


for (t in 1:nSimsPer) {
  # sample parameters from posteriors
  beta = mvrnorm(n=1, mu=betaHat, sigmaHat,tol = 1e-10)
  sigma2 = rinvgamma(1,shape,rate)
```

```r
  sigma = sqrt(sigma2)
  # simulate Male data
  yMale[t] = rnorm(1, beta[1],sigma)
  # simulate Female data for different extra schooling levels
  for (n in 1:length(deltaSchool)) {
    yFemale[n,t] = rnorm(1, beta[1] + beta[3] + (beta[2]+beta[4]) * deltaSchool[n] ,sigma)
  }
}

probComp = rep(-1,length(deltaSchool))
for (t in 1:length(deltaSchool)) {
  probComp[t] = sum(yMale < yFemale[t,]) / nSimsPer
}

plot(deltaSchool,probComp,
     pch = 19, col="red",
     xlab = "education level for female employees",
     ylab = "prob larger wage than average education male")
abline(h=0.5, lty=2)
abline(v=2, lty=2)
```