Survmeth 895 Homework 3

David (Daiwei) Zhang

February 1, 2017

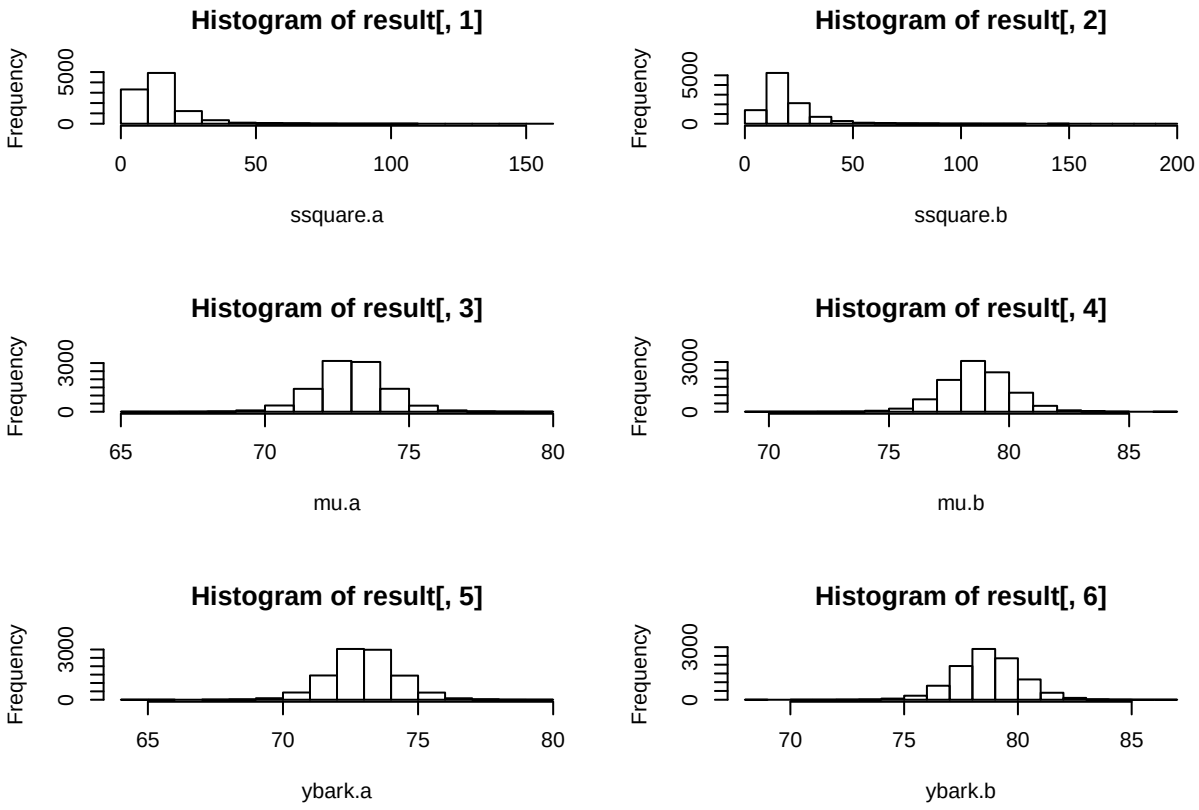
Problem 3

The posterior distribution of the population mean is

$$\begin{aligned}\bar{Y}|y_{\text{inc}} &\sim (\bar{Y}|\mu,\sigma^2,y_{\text{inc}})(\mu|\sigma^2,y_{\text{inc}})(\sigma^2|y_{\text{inc}}) \\ &\sim [(n-1)s^2\chi^2_{n-1}][N(\bar{y}_{\text{inc}},\sigma^2/n)][N(\mu,\sigma^2/k)]\end{aligned}$$

We can use simulation to get more concrete results.

```
a = c(60, 72, 78, 92, 78, 57, 62, 72, 88, 71)
b = c(68, 90, 88, 88, 68, 72, 72, 92, 50, 99)
y = cbind(a, b)
sampsize = length(a)
ybar = c(mean(a), mean(b))
ssquare = c(sd(a), sd(b))
k = 120
nsimul = 10000
result = matrix(0,nsimul,8)
for (i in 1:nsimul){
  tmp=rnorm(sampsize-1)
  chisq=sum(tmp*tmp)
  sigmasq=(sampsize-1)*ssquare/chisq;
  mu=ybar+sqrt(sigmasq/
                 sampsize)*rnorm(1)
  ybark=mu+sqrt(sigmasq/k)*rnorm(1)
  yind=mu+sqrt(sigmasq)*rnorm(1)
  result[i,] = c(sigmasq, mu, ybark, yind)
}
par(mfrow=c(3,2))
hist(result[,1], xlab="ssquare.a")
hist(result[,2], xlab="ssquare.b")
hist(result[,3], xlab="mu.a")
hist(result[,4], xlab="mu.b")
hist(result[,5], xlab="ybark.a")
hist(result[,6], xlab="ybark.b")
```



(a)

We calculate the 95% credibility intervals for the population mean for firm A and firm B

```
## Firm A
quantile(result[,5], c(0.025, 0.975))
```

```
##     2.5%    97.5%
## 70.49443 75.51137
```

```
## Firm B
quantile(result[,6], c(0.025, 0.975))
```

```
##     2.5%    97.5%
## 75.83457 81.57206
```

(b)

The proficient rate is

```
## Firm A
sum(result[,7] >= 75) / nsimul
```

```
## [1] 0.2971
```

```
## Firm B
sum(result[,8] >= 75) / nsimul
```

```
## [1] 0.8009
```

(c)

Based on the credibility intervals and proficiency rates, Firm B provides significantly better training than Firm A.