Homework 1 pt. 2

Michael Pena

2024-01-27

Problem 1

(a).

recall the observed information is: $-\nabla^2 \ell(\mu, \sigma)$

$\ell_\mu = \frac{\sum_{i=1}^n (x_i - \mu)}{\sigma^2}$

$\ell_\sigma = \frac{-1}{n} + \frac{\sum_{i=1}^n (x_i - \mu)^2}{\sigma^3}$

$\ell_{\mu\mu} = \frac{-n}{\sigma^2}$

$\ell_{\mu\sigma} = \ell_{\sigma\mu} = \frac{-2\sum_{i=1}^n (x_i - \mu)}{\sigma^3}$

$\ell_{\sigma\sigma} = \frac{n}{\sigma^2} - \frac{3\sum_{i=1}^n (x_i - \mu)^2}{\sigma^4}$

thus…

$$-\nabla^2 \ell(\mu, \sigma) = \begin{bmatrix} \frac{n}{\sigma^2} & \frac{2\sum_{i=1}^n (x_i - \mu)}{\sigma^3} \\ \frac{2\sum_{i=1}^n (x_i - \mu)}{\sigma^3} & \frac{-n}{\sigma^2} + \frac{3\sum_{i=1}^n (x_i - \mu)^2}{\sigma^4} \end{bmatrix}$$

(b).

let's express the Fisher Info. (Fisher Info: $E(-\nabla^2 \ell(\mu, \sigma))$)

$E(\frac{n}{\sigma^2}) = \frac{n}{\sigma^2}$

$E[\frac{2\sum_{i=1}^n (x_i - \mu)}{\sigma^3}] = \frac{-2}{\sigma^3}\sum_{i=1}^n E[x_i] - \mu = \frac{-2}{\sigma^3}\sum_{i=1}^n \mu - \mu = 0$

$E[\frac{-n}{\sigma^2} + \frac{3\sum_{i=1}^n (x_i - \mu)^2}{\sigma^4}] = \frac{-n}{\sigma^2} + \frac{3}{\sigma^4}\sum_{i=1}^n E[x_i^2] - 2\mu E[x_i] + \mu^2 = \frac{-n}{\sigma^2} + \frac{3}{\sigma^4}\sum_{i=1}^n \sigma^2 + \mu^2 - 2\mu^2 + \mu^2 = \frac{-n}{\sigma^2} + \frac{3n\sigma^2}{\sigma^4} = \frac{2n}{\sigma^2}$

$$E[-\nabla^2 \ell(\mu, \sigma)] = \begin{bmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{2n}{\sigma^2} \end{bmatrix}$$

(c).

let $\vec{\theta} = (\theta_1, \theta_2)$

let

$$g(\vec{\theta}) = \begin{bmatrix} \theta_1 \\ \theta_2^2 \end{bmatrix}$$

thus

$$J(\theta) = \begin{bmatrix} (\theta_1)_{\theta_1} & (\theta_1)_{\theta_2} \\ (\theta_2^2)_{\theta_1} & (\theta_2^2)_{\theta_2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2\theta_2 \end{bmatrix}$$

in our case

$$J(\mu, \sigma) = \begin{bmatrix} 1 & 0 \\ 0 & 2\sigma \end{bmatrix}$$

$$I^{-1}(\vec{\theta}) = \begin{bmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{2n}{\sigma^2} \end{bmatrix}^{-1} = \frac{\sigma^4}{2n^2} \begin{bmatrix} \frac{2n}{\sigma^2} & 0 \\ 0 & \frac{n}{\sigma^2} \end{bmatrix} = \begin{bmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{\sigma^2}{2n} \end{bmatrix}$$

thus the Fisher information for $\ell(\mu, \sigma^2)$ is....

$$[J(\vec{\theta})I^{-1}(\vec{\theta})J^T(\vec{\theta})]^{-1} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 2\sigma \end{bmatrix} \begin{bmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{\sigma^2}{2n} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2\sigma \end{bmatrix} \right)^{-1} = \left(\begin{bmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{\sigma^3}{n} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2\sigma \end{bmatrix} \right)^{-1} = \begin{bmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{n}{2\sigma^4} \end{bmatrix}$$

(d).

$$I^{-1}(\mu, \sigma) = \begin{bmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{\sigma^2}{2n} \end{bmatrix}$$

above shows us that

$$SE(\hat{\theta}_1 = \mu) = \sigma/\sqrt{n}$$

and that

$$SE(\hat{\theta}_2 = \sigma) = \frac{\sigma}{\sqrt{2n}}$$

$$I(\mu, \sigma^2) = \begin{bmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{bmatrix}$$

above shows us that $SE(\theta_2^* = \sigma^2) = \sigma^2 \sqrt{\frac{2}{n}}$

Problem 2

(a).

$$p_1 = \frac{1}{4}, p_2 = \frac{1}{4}, p_3 = \frac{1}{2}$$

$$I(\vec{\pi}) = n \begin{bmatrix} \frac{1}{\pi_1} + \frac{1}{\pi_k} & \frac{1}{\pi_k} \\ \frac{1}{\pi_k} & \frac{1}{\pi_2} + \frac{1}{\pi_k} \end{bmatrix} \Rightarrow I(\vec{p}) = 200 \cdot \begin{bmatrix} 4+2 & 2 \\ 2 & 4+2 \end{bmatrix} = \begin{bmatrix} 1200 & 400 \\ 400 & 1200 \end{bmatrix}$$

```r
# calculating inverse matrix
fish <- matrix(c(1200, 400, 400, 1200), ncol = 2, nrow = 2)
solve(fish)
```

```
##            [,1]       [,2]
## [1,]  0.0009375 -0.0003125
## [2,] -0.0003125  0.0009375
```

$$Var(p_1) = 0.0009375, Var(p_2) = 0.0009375$$

$Var(p_3) = Var(1-p_1-p_2) = Var(p_1)+Var(p_2)+2Cov(p_1, p_2) = 0.0009375+0.0009375+2(-0.0003125) = 0.00125$

thus $SE(p_1) = SE(p_2) = 0.03061862$ and $SE(p_3) = 0.03535534$

(b).

```
set.seed(534)
# building the matrix function
f <- function(n_sim, n, p1, p2) {
    # set p3
    p3 <- 1 - p2 - p1
    # generate a n_sim x 3 zero-matrix
    T <- matrix(0, nrow = n_sim, ncol = 3)
    # run a for loop
    for (i in 1:n_sim) {
        sample <- sample(c("t1", "t2", "t3"), size = n, replace = TRUE, prob = c(p1,
            p2, p3))  # this takes the sample

        T[i, ] <- c(length(sample[sample == "t1"]), length(sample[sample == "t2"]),
            length(sample[sample == "t3"]))
    }
    T
}

# run function
T <- f(1000, 200, 0.25, 0.25)
head(T, n = 5)
```

```
##      [,1] [,2] [,3]
## [1,]   53   49   98
## [2,]   55   41  104
## [3,]   54   55   91
## [4,]   42   56  102
## [5,]   51   59   90
```

```
# make the M matrix
mle_func <- function(T) {
    M <- matrix(0, nrow = length(T[, 1]), ncol = length(T[1, ]))
    for (i in 1:length(T[, 1])) {
        denom <- sum(T[i, ])
        for (j in 1:length(T[1, ])) {
            M[i, j] <- T[i, j]/denom
        }
    }
    return(M)
}
M <- mle_func(T)
head(M, n = 5)
```

```
##       [,1]  [,2]  [,3]
```

```
## [1,] 0.265 0.245 0.490
## [2,] 0.275 0.205 0.520
## [3,] 0.270 0.275 0.455
## [4,] 0.210 0.280 0.510
## [5,] 0.255 0.295 0.450
```

(d).

```r
# obtaining covariance matrix
covM <- cov(M)

# making the greybill matrix
GB <- 200^(-1) * matrix(c(0.25 * 0.75, -0.25 * 0.25, -0.25 * 0.5, -0.25 * 0.25, 0.25 *
    0.75, -0.25 * 0.5, -0.25 * 0.5, -0.25 * 0.5, 0.5 * 0.5), ncol = 3, nrow = 3)
# render
covM
```

```
##               [,1]          [,2]          [,3]
## [1,]  0.0009612857 -0.0002981115 -0.0006631742
## [2,] -0.0002981115  0.0008951239 -0.0005970124
## [3,] -0.0006631742 -0.0005970124  0.0012601866
```

```r
GB
```

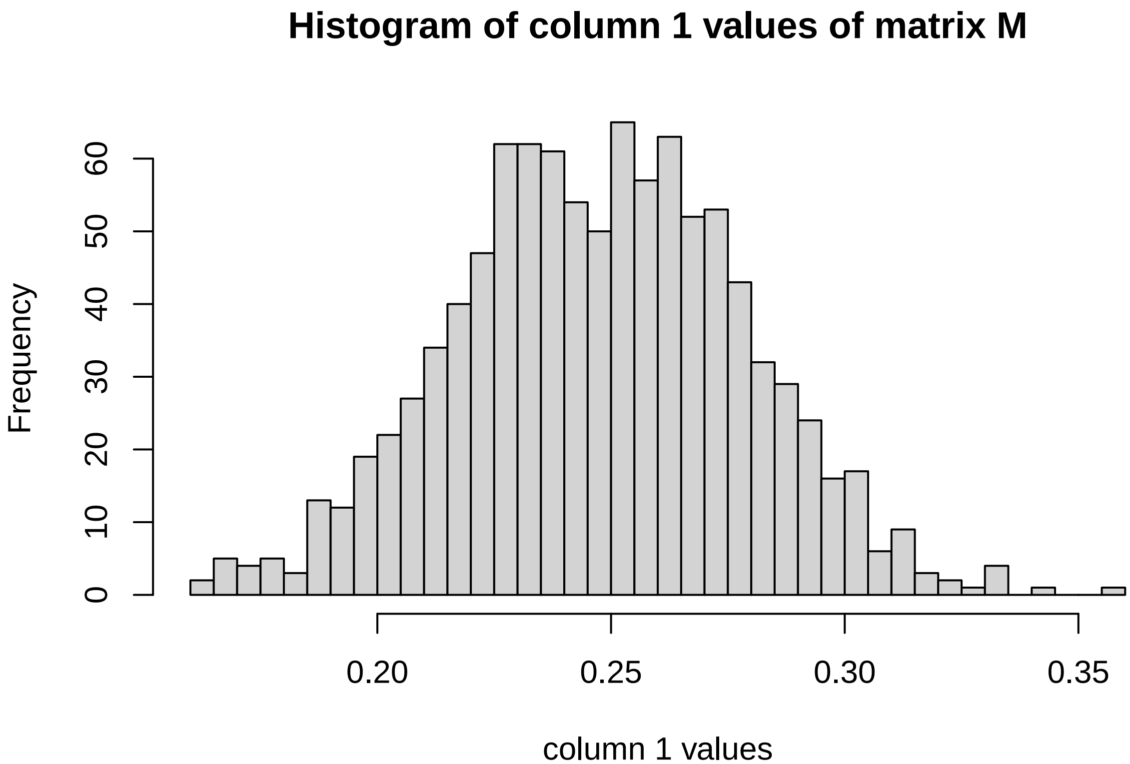
```
##            [,1]       [,2]      [,3]
## [1,]  0.0009375 -0.0003125 -0.000625
## [2,] -0.0003125  0.0009375 -0.000625
## [3,] -0.0006250 -0.0006250  0.001250
```

```r
# make the table
names <- c("var(p1)", "var(p2)", "var(p3)", "cov(p2,p1)", "cov(p3,p2)", "cov(p3,p1)")
theoretical <- c(0.0009375, 0.0009375, 0.00125, -0.0003125, -0.000625, -0.000625)
simulated <- c(0.0009612857, 0.0008951239, -0.0012601866, -0.0002981115, -0.0005970124,
    -0.0006631742)
abs_diff <- abs(theoretical - simulated)

table <- data.frame(names, theoretical, simulated, abs_diff)
table
```

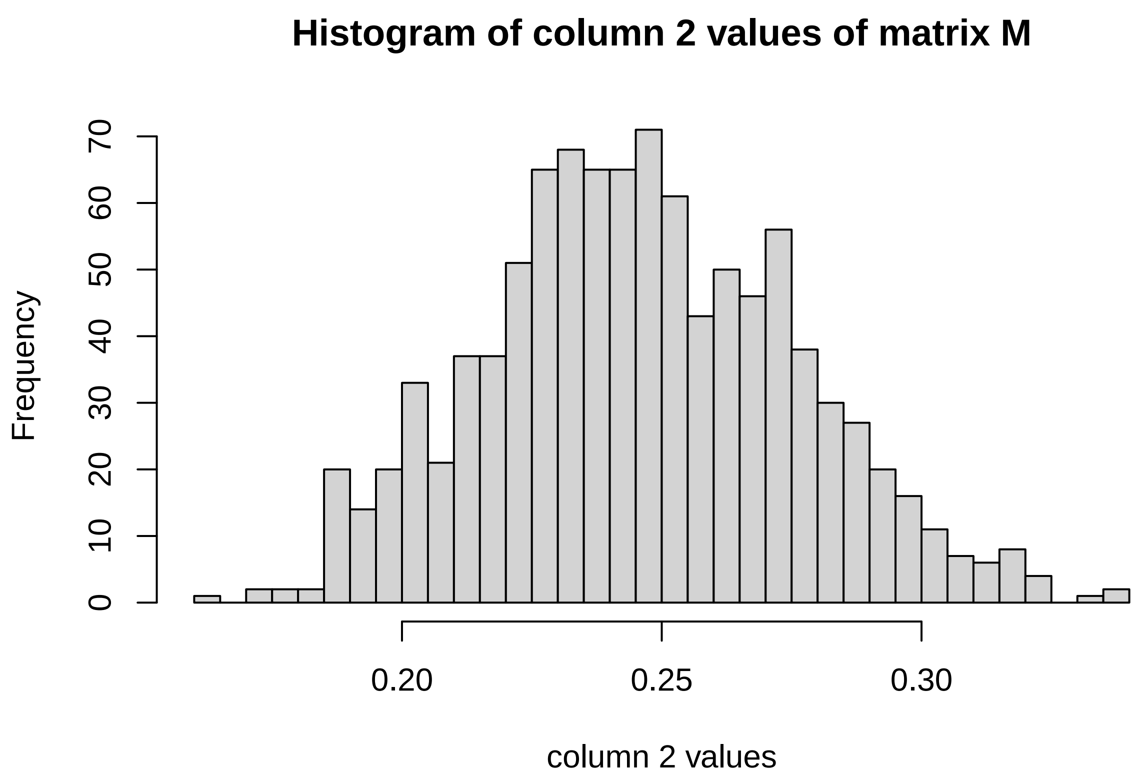
```
##        names theoretical     simulated     abs_diff
## 1    var(p1)   0.0009375  0.0009612857 0.0000237857
## 2    var(p2)   0.0009375  0.0008951239 0.0000423761
## 3    var(p3)   0.0012500 -0.0012601866 0.0025101866
## 4 cov(p2,p1)  -0.0003125 -0.0002981115 0.0000143885
## 5 cov(p3,p2)  -0.0006250 -0.0005970124 0.0000279876
## 6 cov(p3,p1)  -0.0006250 -0.0006631742 0.0000381742
```

```r
# building the histogram
hist(M[, 1], breaks = 30, main = "Histogram of column 1 values of matrix M", xlab = "column 1 values")
```

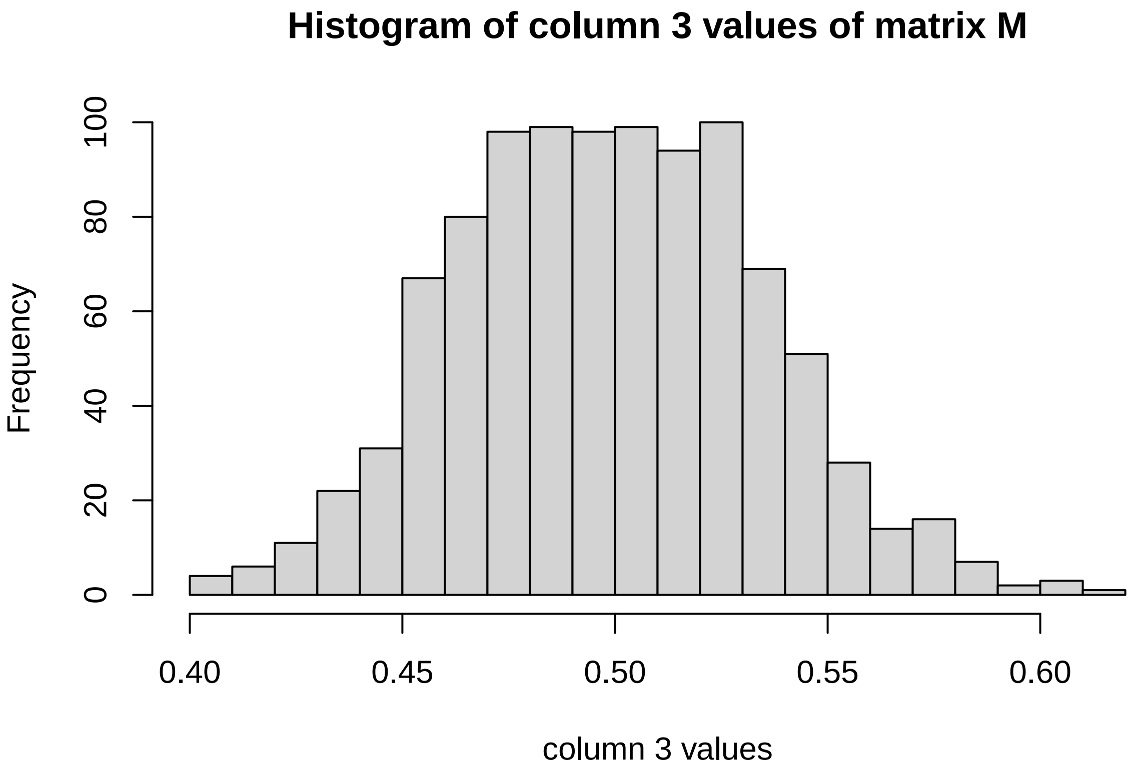


(e).

```
hist(M[, 2], breaks = 30, main = "Histogram of column 2 values of matrix M", xlab = "column 2 values")
```




```
hist(M[, 3], breaks = 30, main = "Histogram of column 3 values of matrix M", xlab = "column 3 values")
```



We observe that these histograms almost follows a normal distribution. This makes sense because cause in our theorem in class $\sqrt{n}(\hat{\theta}_n - \theta^*)$ converges in distribution to $N(\vec{0}, I^{-1}(\theta^*))$.