Mathematical Statistics Ⅱ

Homework#2

Chen I Hsuan (102225002)

+2

6/5

1. In example 7.2.17, let $\beta = 2, \tau_i$=1,and n=4.

(1) Check if the Hessian Matrix is negative definite or not when X=(3, 2, 0, 10)

$$\begin{cases} X_i \overset{iid}{\sim} Poisson(\tau_i) \\ Y_i \overset{iid}{\sim} Poisson(\beta\tau_i) \end{cases}, \text{where } i = 1,2,\dots,n$$

$$X_i \perp Y_i$$

$$\theta = (\beta, \tau_1, \dots, \tau_n) \in R^{n+1}$$

$$L(\theta|\mathbb{x},\mathbb{y}) = \prod_{i=1}^{n} \frac{\tau_i^{x_i} e^{-\tau_i}}{x_i!} \times \frac{\beta\tau_i^{y_i} e^{-\beta\tau_i}}{y_i!}$$

$$\log L(\theta|\mathbb{x},\mathbb{y}) = \sum_{i=1}^{n} [-\log x_i! + x_i log\tau_i - \tau_i - log y_i! + y_i log\beta + y_i log\tau_i - \beta\tau_i]$$

$$\begin{cases} \dfrac{\partial \log L(\theta)}{\partial \beta} = \sum_{i=1}^{n} (\dfrac{y_i}{\beta} - \tau_i) = 0 & --① \\ \dfrac{\partial \log L(\theta)}{\partial \tau_i} = \dfrac{x_i}{\tau_i} - 1 + \dfrac{y_i}{\tau_i} - \beta = 0 & --② \end{cases}, \forall i = 1, \dots, n$$

$$②: 1 + \beta = \frac{x_i + y_i}{\tau_i} \Rightarrow \tau_i = \frac{x_i + y_i}{1 + \beta}$$

$$\sum_{i=1}^{n} \tau_i = \frac{\sum_{i=1}^{n} x_i + y_i}{1 + \beta} --③$$

① and ③ give $\Rightarrow \hat\beta = \frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} X_i}$

$$\Rightarrow \hat{\tau_i} = \frac{X_i + Y_i}{1 + \hat\beta}$$

$$\frac{\partial^2 \log L(\theta)}{\partial \beta^2} = \frac{-1}{\beta^2} \sum_{i=1}^{n} y_i$$

$$\frac{\partial^2 \log L(\theta)}{\partial \tau_i \partial \beta} = -1$$

$$\frac{\partial^2 \log L(\theta)}{\partial \tau_i \partial \tau_j} = 0, \forall i \neq j$$

$$\frac{\partial^2 \log L(\theta)}{\partial \tau_i^2} = -\frac{x_i + y_i}{\tau_i^2}$$

+1

$$\therefore \frac{\partial^2 \log L(\theta)}{\partial\theta\,\partial\theta'}\Big|_{\theta=\hat{\theta}} = \begin{bmatrix} \frac{-\sum_{i=1}^{n} y_i}{\hat{\beta}^2} & -1 & \cdots & -1 \\ -1 & -\frac{x_1+y_1}{\widehat{\tau_1^2}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & \cdots & -\frac{x_n+y_n}{\widehat{\tau_n^2}} \end{bmatrix}$$

#code

```
#1(1)
rm(list=ls(all=TRUE))
x=c(3,2,0,10)
y=c(8,2,1,16)
beta=2
tau=1
n=4

beta_hat=sum(y)/sum(x) #1.8
tau_hat=(x+y)/(1+beta_hat)

H=matrix(0,n+1,n+1)
H[1,]=-1
H[,1]=-1
H[1,1]=-sum(y)/beta_hat^2

for( i in 2:(n+1)){

c=-(x+y)/(tau_hat^2)
H[i,i]=c[i-1]
}
H

eigen(H)$values
#Since all eigenvalues are all negative, Hessian is negative definite.
```

#Output

```
> H
          [,1]       [,2]  [,3]  [,4]       [,5]
[1,] -8.333333 -1.0000000 -1.00 -1.00 -1.0000000
[2,] -1.000000 -0.7127273  0.00  0.00  0.0000000
[3,] -1.000000  0.0000000 -1.96  0.00  0.0000000
[4,] -1.000000  0.0000000  0.00 -7.84  0.0000000
[5,] -1.000000  0.0000000  0.00  0.00 -0.3015385
>
> eigen(H)$values
[1] -0.1295155 -0.6132033 -1.8330849 -7.2169366 -9.3548588
> #Since all eigenvalues are all negative,Hessian is negative definite.
```

+1

(2) Generate data by R. Check if the Hessian matrix is always negative definite or not.

If we don't have any condition for the Hessian matrix, the Hessian matrix will not converge.

Thus, if we take some conditions for the Hessian matrix ($\forall x_i \neq 0$ *or* $\forall y_i \neq 0$, $\sum x_i \neq 0$ and $\sum y_i \neq 0$), the Hessian matrix will exist and be always negative definite.

#code

```
#1(2)
rm(list=ls(all=TRUE))
aa=c()
x=c()
y=c()

for(j in 1:1000){
beta=2
tau=1
n=4

for(k in 1:n){
repeat{
x[k]=rpois(1,tau)
y[k]=rpois(1,beta*tau)
if(x[k]!=0||y[k]!=0 & sum(y)!=0& sum(x)!=0 ){break}
#condition(1)x and y cannot be 0 simultaneously.
#condition(2)sum(x)and sum(y)are not 0
}
}
beta_hat=sum(y)/sum(x) #1.8
tau_hat=(x+y)/(1+beta_hat)

H=matrix(0,n+1,n+1)
H[1,]=-1
H[,1]=-1
H[1,1]=-sum(y)/beta_hat^2

for( i in 2:(n+1)){

c=-(x+y)/(tau_hat^2)
H[i,i]=c[i-1]
}

aa[j]=length((eigen(H)$values)[eigen(H)$values>0])
}
sum(aa)
```

Not

#Output	
> sum(aa) [1] 0	➤ We check Hessian matrix by R for 1000 times, and the result Indicates Hessian matrix is always negative definite.

+1

(3) Generate data by R under β=1, 2, 5, 10 , τ_i=1, and n=1000. Check if $\hat{\beta}$ is close to the true value.

#code

```
#1(3)
rm(list=ls(all=TRUE))
beta_hat=c()
x=c()
y=c()

for(i in 1:4){
beta=c(1,2,5,10)
tau=1
x=rpois(1000,tau)
y=rpois(1000,beta[i]*tau)

beta_hat[i]=sum(y)/sum(x)
}
beta_hat
beta
```

#Output

```
> beta_hat
[1]  1.039837  2.120763  5.127638 10.597294
> beta
[1]  1  2  5 10
```

➤ The result shows that $\hat{\beta}$ is close to the true value.

+1 (4) For incomplete data X=(*, 2, 0, 10) and Y=(8, 2, 1, 16), use the EM algorithm to find the MLE. How many repetitions are required to get convergence? [start from $\beta^{(1)}$=1, $\tau_i^{(1)}$=1, i=1, 2, 3, 4]

+2 (5) Compare the above result with the MLE derived from the derivatives.

#code

```
#1(4)(5)
rm(list=ls(all=TRUE))
x=c(NA,2,0,10)
y=c(8,2,1,16)

##initial value
beta=1
tau=rep(1,4)


##MLE derived from the derivatives
tau_hat=c()

beta_hat=sum(y[2:4])/sum(x[2:4])
tau_hat[1]=y[1]/beta_hat

for(i in 2:4){
tau_hat[i]=(x[i]+y[i])/(1+beta_hat)
}

##
m=100
r=1

##EM method
repeat{
beta_rplus1=sum(y)/(tau[1]+sum(x[2:4]))
tau[1]=(tau[1]+y[1])/(1+beta_rplus1)

for(i in 2:4){
tau[i]=(x[i]+y[i])/(1+beta_rplus1)
}

r=r+1
if(abs(beta_rplus1-beta_hat)<0.001 && abs(tau-tau_hat)<0.001 ){break}
}

EM=c("beta_hat_EM","tau1_hat_EM","tau2_hat_EM","tau3_hat_EM","tau4_hat_EM
")
value1=c(beta_rplus1,tau)

Derivatives=c("beta_hat","tau1_hat","tau2_hat","tau3_hat","tau4_hat")
value2=c(beta_hat,tau_hat)
```

```
cbind(EM, value1, Derivatives, value2)
cbind("repetitions", r)
```

#Output

```
> cbind(EM,value1,Derivatives,value2)
     EM             value1              Derivatives value2
[1,] "beta_hat_EM" "1.58345132298443"  "beta_hat"  "1.58333333333333"
[2,] "tau1_hat_EM" "5.05190897121778"  "tau1_hat"  "5.05263157894737"
[3,] "tau2_hat_EM" "1.54831637988022"  "tau2_hat"  "1.54838709677419"
[4,] "tau3_hat_EM" "0.387079094970054" "tau3_hat"  "0.387096774193548"
[5,] "tau4_hat_EM" "10.0640564692214"  "tau4_hat"  "10.0645161290323"
> cbind("repetitions",r)
                   r
[1,] "repetitions" "16"
```

When MLE by using EM method get convergence, the value will be close to MLE from the derivatives. Here with error <0.001.

And I also get the repetitions=16.

EM method to find MLE

$$logL(\theta|\mathbb{x},\mathbb{y}) = \sum_{i=1}^{n}[-\log x_i! + x_i log\tau_i - \tau_i - logy_i! + y_i log\beta + y_i log\tau_i - \beta\tau_i]$$

$$E\left[logL(\theta|\mathbb{x},\mathbb{y})\middle|\theta^{(r)},\mathbb{y}\right]$$

$$= E\left[\textstyle\sum_{i=1}^{n}[-\log x_1! + x_1 log\tau_1 - \tau_1\middle|\theta^{(r)}\right]$$

$$+\sum_{i=2}^{n}[-\log x_i! + x_i log\tau_i - \tau_i] + \sum_{i=1}^{n}[-logy_i! + y_i log\beta + y_i log\tau_i - \beta\tau_i]$$

$\boxed{\text{E}-\text{step}}$

$$E\left[\textstyle\sum_{i=1}^{n}[-\log x_1! + x_1 log\tau_1 - \tau_1\middle|\theta^{(r)}\right] = const. + E\left[X_1\middle|\theta^{(r)}\right]log\tau_1 - \tau_1$$
$$= const. + {\tau_1}^{(r)} log\tau_1 - \tau_1$$

Let A$= const. + {\tau_1}^{(r)} log\tau_1 - \tau_1$

$$Q(\theta) = E\left[logL(\theta|\mathbb{x},\mathbb{y})\middle|\theta^{(r)},\mathbb{y}\right]$$
$$= A + \sum_{i=2}^{n}[-\log x_i! + x_i log\tau_i - \tau_i] + \sum_{i=1}^{n}[-logy_i! + y_i log\beta + y_i log\tau_i - \beta\tau_i]$$

$\boxed{M-step}$

$$\frac{\partial Q(\theta)}{\partial\beta} = \frac{1}{\beta}\sum_{i=1}^{n} y_i - \sum_{i=1}^{n}\tau_i = 0 \quad --①$$

$$\frac{\partial Q(\theta)}{\partial\tau_1} = \frac{{\tau_1}^{(r)}}{\tau_1} - 1 + \frac{Y_1}{\tau_1} - \beta = 0 \quad --②$$

$$\frac{\partial Q(\theta)}{\partial\tau_i} = \frac{X_i}{\tau_1} - 1 + \frac{Y_i}{\tau_i} - \beta = 0 \quad --③$$

$$② \Rightarrow {\tau_1}^{(r)} + Y_1 = \tau_1(1+\beta) \Rightarrow \tau_1 = \frac{{\tau_1}^{(r)} + Y_1}{1+\beta}$$

+1

$$③ \Rightarrow X_i + Y_1 = \tau_i(1+\beta) \Rightarrow \tau_i = \frac{X_i + Y_i}{1+\beta}$$

Using ①, ② *and* ③

$$\frac{\sum_{i=1}^{n} Y_i}{\beta} - \frac{{\tau_1}^{(r)} + Y_1}{1+\beta} - \frac{\sum_{i=2}^{n}(X_i + Y_i)}{1+\beta} = 0$$

$$\Rightarrow \frac{\sum_{i=1}^{n} y_i}{\beta} = \frac{{\tau_1}^{(r)} + Y_1 + \sum_{i=2}^{n}(x_i + Y_i)}{1+\beta}$$

$$\Rightarrow (1+\beta)\sum_{i=1}^{n} Y_i = \beta\left({\tau_1}^{(r)} + \sum_{i=1}^{n} Y_i + \sum_{i=2}^{n} X_i\right)$$

$$\Rightarrow \beta^{(r+1)} = \frac{\sum_{i=1}^{n} Y_i}{{\tau_1}^{(r)} + \sum_{i=2}^{n} X_i}$$

$$\therefore \theta^{(r+1)} = \begin{cases} \beta^{(r+1)} = \dfrac{\sum_{i=1}^{n} y_i}{{\tau_1}^{(r)} + \sum_{i=2}^{n} x_i} \\ {\tau_1}^{(r+1)} = \dfrac{{\tau_1}^{(r)} + y_i}{1+\beta^{(r+1)}} \\ {\tau_i}^{(r+1)} = \dfrac{x_i + y_i}{1+\beta^{(r+1)}} \end{cases} \quad , i = 2, \dots, n$$

$$\text{By the property of EM alogrithm,} \begin{cases} \lim_{r\to\infty} \beta^{(r)} = \hat{\beta} \\ \lim_{r\to\infty} {\tau_i}^{(r)} = \hat{\tau} \end{cases}$$

MLE derived from the derivatives

$$\begin{cases} \mathrm{X_i} \overset{iid}{\sim} Poisson(\tau_i) \\ \mathrm{Y_i} \overset{iid}{\sim} Poisson(\beta\tau_i) \end{cases} \text{, where } \mathrm{i} = 1,2,\dots,\mathrm{n}$$

$\mathrm{X_i} \perp Y_i$

$\theta = (\beta, \tau_1, \dots, \tau_n) \in R^{n+1}$

$$\mathrm{L}(\theta|\mathbb{x},\mathbb{y}) = \prod_{i=1}^{n} \frac{\tau_i^{x_i} e^{-\tau_i}}{x_i!} \times \frac{\beta\tau_i^{y_i} e^{-\beta\tau_i}}{y_i!}$$

Incomplete data:

$\mathrm{X} = X_1$ is missing

$\mathbb{y} = (y_1, (x_2, y_2), \cdots (x_n, y_n))$

$$L(\theta|\mathbb{y}) = \sum_{x_1=0}^{\infty} L(\theta|X_1,\mathbb{y}) = \sum_{x_1=0}^{\infty} \prod_{i=2}^{n} \frac{\tau_i^{x_i} e^{-\tau_i}}{x_i!} \times \prod_{i=2}^{n} \frac{\beta\tau_i^{y_i} e^{-\beta\tau_i}}{y_i!} \times \frac{\tau_1^{x_1} e^{-\tau_1}}{x_1!}$$

$$= \prod_{i=2}^{n} \frac{\tau_i^{x_i} e^{-\tau_i}}{x_i!} \times \prod_{i=2}^{n} \frac{\beta\tau_i^{y_i} e^{-\beta\tau_i}}{y_i!}$$

$$\log \mathrm{L}(\theta|\mathbb{y}) = \sum_{i=2}^{n}[-\log x_i! + x_i log\tau_i - \tau_i] + \sum_{i=1}^{n} -log y_i! + y_i log\beta + y_i log\tau_i - \beta\tau_i]$$

$$\begin{cases} \dfrac{\partial \log L(\theta|\mathbb{y})}{\partial \beta} = \dfrac{1}{\beta}\sum_{i=2}^{n} y_i - \sum_{i=2}^{n} \tau_i = 0 & --① \\ \dfrac{\partial \log L(\theta|\mathbb{y})}{\partial \tau_1} = \dfrac{y_1}{\tau_1} - \beta = 0 & --② \\ \dfrac{\partial \log L(\theta|\mathbb{y})}{\partial \tau_i} = \dfrac{x_i}{\tau_i} - 1 + \dfrac{y_i}{\tau_i} - \beta = 0 & --③ \end{cases} , \forall i = 2, \dots, n$$

②: $y_1 = \beta\tau_1 \Rightarrow \hat{\tau}_1 = \dfrac{Y_1}{\hat{\beta}}$

③ : $x_i + y_i = (1+\beta)\,\tau_i \Rightarrow \ \hat{\tau}_i = \frac{X_i + Y_i}{1+\hat{\beta}}$, i=2, 3⋯, n

① and②and ③ give $\Rightarrow \hat{\beta} = \frac{\sum_{i=2}^{n} Y_i}{\sum_{i=2}^{n} X_i}$