Homework#2

Name: 范又方 Fan, You-fang

Student number: 106225022

Exercise 7.2 [P355]

Let $X_1,\ldots,X_n$ be a random sample from a gamma(α , β) population.

(a) Find the MLE of β, assuming α is known.

(b) If α and β are both unknown, there is no explicit formula for the MLEs of α and β, but the maximum can be found numerically. The result in part (a) can be used to reduce the problem to the maximization of a univariate function. Find the MLEs for α and β for the data in Exercise 7.10(c).

Sol :

(a)The likelihood function is

$$\mathrm{L}(\beta|\mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} {x_i}^{\alpha-1} \exp(-\frac{x_i}{\beta})$$

$$= (\tfrac{1}{\Gamma(\alpha)})^n \tfrac{1}{\beta^{n\alpha}} (\textstyle\prod_{i=1}^{n} x_i)^{\alpha-1} \exp(-\tfrac{\sum_{i=1}^{n} x_i}{\beta})$$

$$\log\mathrm{L}(\beta|\mathbf{x}) = -log\Gamma(\alpha)^n - log\beta^{n\alpha} + (\alpha-1)log\prod_{i=1}^{n} x_i - \frac{\sum_{i=1}^{n} x_i}{\beta}$$

$$= -nlog\Gamma(\alpha) - n\alpha log\beta + (\alpha-1)\textstyle\sum_{i=1}^{n} logx_i - \tfrac{\sum_{i=1}^{n} x_i}{\beta}$$

$$\frac{\partial logL(\beta|\mathbf{x})}{\partial\beta} = -\frac{n\alpha}{\beta} + \frac{\sum_{i=1}^{n} x_i}{\beta^2} = 0 \Rightarrow \hat{\beta} = \frac{\bar{x}}{\alpha}$$

Setting the partial derivatives equal to 0 and solving the solution $\hat{\beta} = \frac{\bar{x}}{\alpha}$.

To check this is a maximum,

$$\left.\frac{\partial^2 log\mathrm{L}(\beta|\mathbf{x})}{\partial\beta^2}\right|_{\beta=\hat{\beta}} = \left.\frac{n\alpha}{\beta^2} - \frac{2\sum_{i=1}^{n} x_i}{\beta^3}\right|_{\beta=\hat{\beta}} = \frac{n\alpha}{(\frac{\bar{x}}{\alpha})^2} - \frac{2n\bar{x}}{\left(\frac{\bar{x}}{\alpha}\right)^3} = -\frac{n\alpha^3}{\bar{x}^2} < 0.$$

Because $\frac{\bar{x}}{\alpha}$ is the only extreme point and it is a global maximum. Therefore, $\frac{\bar{X}}{\alpha}$ is the MLE of β.

(b) The likelihood function is

$$L(\alpha, \beta|\mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} {x_i}^{\alpha-1} \exp(-\frac{x_i}{\beta})$$

$$= (\frac{1}{\Gamma(\alpha)})^n \frac{1}{\beta^{n\alpha}} (\prod_{i=1}^{n} x_i)^{\alpha-1} \exp(-\frac{\sum_{i=1}^{n} x_i}{\beta})$$

$$\log L(\alpha, \beta|\mathbf{x}) = -log\Gamma(\alpha)^n - log\beta^{n\alpha} + (\alpha-1)log\prod_{i=1}^{n} x_i - \frac{\sum_{i=1}^{n} x_i}{\beta}$$

$$= -nlog\Gamma(\alpha) - n\alpha log\beta + (\alpha-1)\sum_{i=1}^{n} logx_i - \frac{\sum_{i=1}^{n} x_i}{\beta}$$

$$\frac{\partial logL(\alpha, \beta|\mathbf{x})}{\partial\beta} = -\frac{n\alpha}{\beta} + \frac{\sum_{i=1}^{n} x_i}{\beta^2} = 0 \Rightarrow \hat{\beta} = \frac{\bar{x}}{\alpha}$$

$$\frac{\partial logL(\alpha, \beta|\mathbf{x})}{\partial\alpha} = -n(\frac{d}{d\alpha}log\Gamma(\alpha)) - nlog\beta + \sum_{i=1}^{n} logx_i$$

Let $\psi(\alpha) = \frac{d}{d\alpha}log\Gamma(\alpha)$ *be the digamma function and* $\psi'(\alpha)$ *be the trigamma function.*

$$\frac{\partial logL(\alpha, \beta|\mathbf{x})}{\partial\alpha} = -n\psi(\alpha) - nlog\beta + \sum_{i=1}^{n} logx_i$$

$$= \psi(\alpha) + log\beta - \frac{1}{n}\sum_{i=1}^{n} logx_i = 0$$

To found the maximum likelihood estimators is reduced to a one − dimensional problem, changing the α *into* β. $\hat{\beta} = \frac{\bar{x}}{\alpha} \Rightarrow \alpha = \frac{\bar{x}}{\hat{\beta}}$

$$\frac{\partial logL(\alpha, \beta|\mathbf{x})}{\partial\alpha} = \psi\left(\frac{\bar{x}}{\beta}\right) + log\beta - \frac{1}{n}\sum_{i=1}^{n} logx_i = 0$$

In R,

```
y=read.csv("C:/Users/youfang/Desktop/azxc.csv")
y=y$X
qq <- function(beta){
  digamma(mean(y)/beta)-mean(log(y))+log(beta)
}
#to find the root of beta
b <- uniroot(qq,lower=0.0001,upper=100000)
beta.hat<-b$root
beta.hat
#bring the value of beta in alpha
alpha.hat<-mean(y)/beta.hat
alpha.hat
```

```
> beta.hat
[1] 0.04494341
> alpha.hat<-mean(y)/beta.hat
> alpha.hat
[1] 514.2975
```

Exercise 7.6 [P355]

Let $X_1,\ldots,X_n$ be a random sample from the pdf

$$f(x|\theta) = \theta x^{-2}\,, 0 < \theta \le x < \infty.$$

(a) What is a sufficient statistic for θ ? Is complete ?

(b) Find the MLE of θ.

Draw a figure of the likelihood function to explain your answer.(with R)

(c) Find the method of moments estimator of θ.

Sol :

(a) The joint pdf of the sample **X** is

$$f(\mathbf{x}|\theta) = \prod_{i=1}^{n} \theta x_i^{-2} \quad , \qquad 0 < \theta \le x < \infty$$

Now $< \theta \le x < \infty$, using the indicator function, we can see the inequality is $0 < \theta \le x_{(1)}$

So we have

$$f(\mathbf{x}|\theta) = \prod_{i=1}^{n} \theta x_i^{-2} I_{(\theta\,,\infty\,)}(x_i)$$

$$= \theta^n (\prod_{i=1}^{n} x_i^{-2}) I_{(0\,,\,x_{(1)})}(\theta)$$

Since $f(x|\theta) = h(x)g(T(\mathbf{X})|\theta)$, where $h(x) = \prod_{i=1}^{n} x_i^{-2}$ and $g(T(\mathbf{X})|\theta) = \theta^n I_{(0\,,x_{(1)}\,)}(\theta)$, $T(\mathbf{X}) = X_{(1)}$

Therefore, by the Factorization Theorem, $X_{(1)}$ is a sufficient statistic for θ.

Let T $= X_{(1)}$, the pdf $f_T(t) = \frac{n!}{(n-1)!}\theta t^{-2}(\frac{\theta}{t})^{n-1} = \frac{n\theta^n}{t^{n+1}}$, $\forall\, \theta \le x_{(1)} < \infty$

$$E_\theta\{g(T)\} = 0\,, \qquad \forall\, \theta \le x_{(1)} < \infty$$

$$\Rightarrow \int_\theta^\infty g(t)\frac{n\theta^n}{t^{n+1}}dt = 0$$

$$\Rightarrow \frac{d}{d\theta}\int_\theta^\infty g(t)\frac{n\theta^n}{t^{n+1}}dt = \frac{d}{d\theta}0$$

$$\Rightarrow g(\theta)\frac{n\theta^n}{\theta^{n+1}}=0$$

$$\Rightarrow g(\theta) = 0\,, \forall\, \theta \le x_{(1)} < \infty$$

That T $= X_{(1)}$ is a complete statistic.

(b) The likelihood function is

$$L(\theta|\mathbf{x}) = \prod_{i=1}^{n} \theta {x_i}^{-2}, \qquad 0 < \theta \leq x < \infty$$

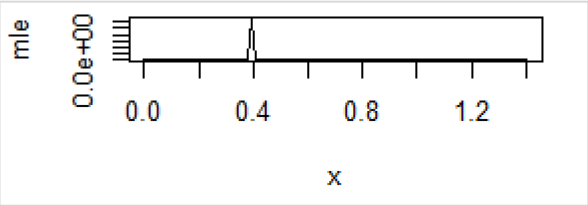
Using the indicator function,

$$L(\theta|\mathbf{x}) = \theta^n (\prod_{i=1}^{n} {x_i}^{-2}) I_{(0\,,\,x_{(1)})}(\theta)$$

Because $L(\theta|\mathbf{x})$ is an increasing function of θ in $\theta \in (0\,,\,x_{(1)})$, so we want to maximize $L(\theta|\mathbf{x})$. But due to the indicator function $L(\theta|\mathbf{x}) = 0$, if $\theta > x_{(1)}$. Therefore, $\hat{\theta} = X_{(1)}$.

In R,

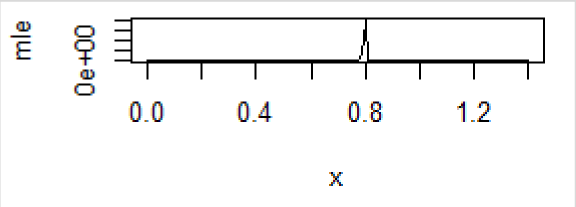
```
par(mfrow=c(2,2))
n=100
x=seq(0.4,10)
x
mle<-function(theta){
  theta^n *prod(x^(-2))*{theta > 0 & theta < min(x)}
}
plot(mle,xlim=c(0,1.4))
```



Suppose that $X_{(1)} = 0.4$, we can see the figure that maximum is at 0.4

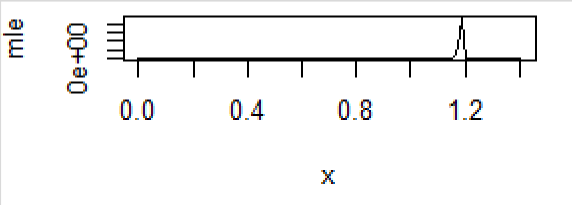
Following the second and the third figures, we have the same conclusion.

```
n=100
x=seq(0.8,10)
mle<-function(theta){
  theta^n *prod(x^(-2))*{theta > 0 & theta < min(x)}
}
plot(mle,xlim=c(0,1.4))
```



Suppose that $X_{(1)} = 0.8$, the figure that maximum is at 0.8

```
n=100
x=seq(1.2,10)
mle<-function(theta){
  theta^n *prod(x^(-2))*{theta > 0 & theta < min(x)}
}
plot(mle,xlim=c(0,1.4))
```



Suppose that $X_{(1)} = 1.2$, the figure that maximum is at 1.2

Therefore, $\hat{\theta} = X_{(1)}$ is the MLE of θ.

(c)

$$\mathrm{E(X)} = \int_{\theta}^{\infty} x\theta x^{-2} dx = \theta \int_{\theta}^{\infty} x^{-1} dx = \theta \log x \Big|_{\theta}^{\infty} = \infty.$$

Therefore, the method of moments estimator of θ does not exist.

Exercise 7.8 [P356]

One observation, **X**, is taken from a n(0,σ^2) population.

(a) Find an unbiased estimator of σ^2.

(b) Find the MLE of σ.

(c) Discuss how the method of moments estimator of σ might be found.

Sol :

(a) $\mathrm{E(X)} = 0$, the 1th moment is independent of σ^2, we need to use 2nd moment.

$\mathrm{E}(X^2) = \mathrm{Var(X)} + (E(X))^2 = \sigma^2 + 0 = \sigma^2$

Therefore, X^2 is an unbiased estimator of σ^2.

(b) The likelihood function is

$$\mathrm{L}(\sigma|\mathrm{x}) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{x^2}{2\sigma^2}\right) = (2\pi)^{-\frac{1}{2}}(\sigma)^{-1} exp\left(-\frac{x^2}{2\sigma^2}\right)$$

$$\log \mathrm{L}(\sigma|\mathrm{x}) = -\frac{1}{2}\log(2\pi) - log\sigma - \frac{x^2}{2\sigma^2}$$

$$\frac{\partial \log \mathrm{L}(\sigma|\mathrm{x})}{\partial\sigma} = -\frac{1}{\sigma} + \frac{x^2}{\sigma^3} = 0 \implies \hat{\sigma} = \sqrt{\mathrm{X}^2} = |\mathrm{X}|$$

$$\left.\frac{\partial^2 \log \mathrm{L}(\sigma|\mathrm{x})}{\partial\sigma^2}\right|_{\sigma=\hat{\sigma}} = \frac{1}{\sigma^2} - \frac{3x^2}{\sigma^4} = \frac{1}{x^2} - \frac{3x^2}{x^4} = -\frac{2}{x^2} < 0$$

$\hat{\sigma} = |X|$ is a local maximum, and that is the only zero of the first derivative. It is a global maximum, too. Therefore, $\hat{\sigma} = |X|$ is the MLE of σ.

(c) Because $\mathrm{E(X)} = 0$ is known. $\mathrm{E}(X^2) = \sigma^2 = \frac{1}{n}\sum_{i=1}^{n} x_i{}^2 \implies \hat{\sigma} = |X|$

Exercise 7.9 [P356]

Let $X_1,\ldots,X_n$ be iid with pdf

$$f(x|\theta) = \frac{1}{\theta} \quad ,0 \le x \le \theta \ ,\theta > 0.$$

Estimate θ using both the method of moments and maximum likelihood. Calculate the means and variances of the two estimators. Which one should be preferred and why?

Sol :

MME :

$$\mathrm{E(X)} = \bar{X} = \frac{\theta}{2} \Longrightarrow \theta = 2\bar{X}$$

The method of moments estimators is $\tilde{\theta} = 2\bar{X}$

Therefore, $\mathrm{E}(\tilde{\theta}) = \mathrm{E}(2\bar{X}) = 2\frac{\theta}{2} = \theta\ ,\ Var(\tilde{\theta}) = Var(2\bar{X}) = 4\frac{1}{n}\frac{\theta^2}{12} = \frac{\theta^2}{3n}$

MLE :

$$\mathrm{L}(\theta|\mathbf{x}) = \prod_{i=1}^{n}\frac{1}{\theta}\ ,\ \ 0 \le x \le \theta\ ,\theta > 0$$

Using the indicator function,

$$\mathrm{L}(\theta|\mathbf{x}) = \prod_{i=1}^{n}\frac{1}{\theta}\, I_{[0\,,\theta]}(x_i) = \frac{1}{\theta^n} I_{[x_{(n)},\infty)}(\theta)$$

Because $\mathrm{L}(\theta|\mathbf{x})$ is an increasing function of θ in $\theta \in [x_{(n),\infty})$. For $\theta \ge x_{(n)}$, $\mathrm{L}(\theta|\mathbf{x})$ is maximized at $\hat{\theta} = X_{(n)}$. Therefore, $\hat{\theta} = X_{(n)}$ is the MLE of θ.

Now we can see that $\frac{X_{(n)}}{\theta} \sim Beta(n, 1)$

The mean and variance,

$$\begin{cases} E\left(\frac{X_{(n)}}{\theta}\right) = \frac{n}{n+1} \\ Var\left(\frac{X_{(n)}}{\theta}\right) = \frac{n}{(n+1)^2(n+2)} \end{cases} \Longrightarrow \begin{cases} E(X_{(n)}) = E(\hat{\theta}) = \frac{n}{n+1}\theta \\ Var(X_{(n)}) = Var(\hat{\theta}) = \frac{n}{(n+1)^2(n+2)}\theta^2 \end{cases}$$

To determine the better estimator, we should compare variances. $\tilde{\theta}$ is an unbiased estimator of θ, and $\hat{\theta}$ is a biased estimator. We can easy to see that

$$\mathrm{Var}(\hat{\theta}) = \frac{n}{(n+1)^2(n+2)}\theta^2 < \frac{\theta^2}{3n} = Var(\tilde{\theta})$$

If n is large, $\hat{\theta}$ is probably preferable to $\tilde{\theta}$.

Quiz#2

Q3 Let $X_1, \dots, X_n \sim N(\mu, \sigma^2)$, where μ is restricted to $\mu \le a$ or $\mu \ge b$ for some numbers $a < b$. Assume that σ^2 is known. Hence, the parameter space is $\Theta = (-\infty, a] \cup [b, \infty)$. Obtain the MLE $\hat{\mu}$.

Draw figures to explain 4 cases under $X_1, \dots, X_5 \sim N(\mu, 1)$, $a = 0$, $b = 2$.

Sol :

$$L(\mu|\mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right)$$

$$\log L(\mu|\mathbf{x}) = \left(-\frac{n}{2}\right)\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

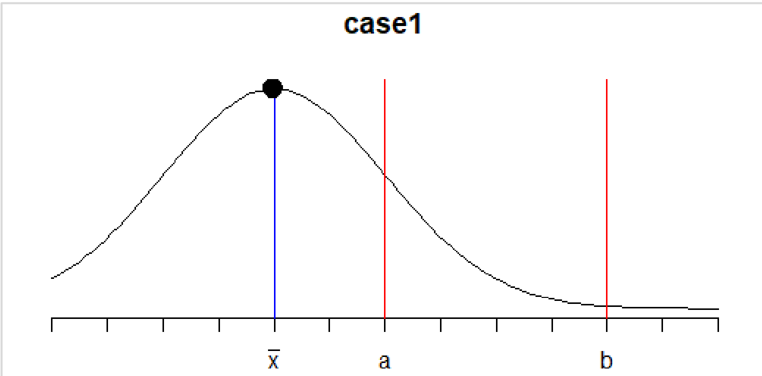
$$\frac{\partial}{\partial\mu}\log L(\mu|\mathbf{x}) = -\frac{1}{2\sigma^2}\sum_{i=1}^{n}\big(2(x_i-\mu)(-1)\big) = 0 \Rightarrow \sum_{i=1}^{n} x_i = n\mu \Rightarrow \hat{\mu} = \bar{x}$$

$$\left.\frac{\partial^2}{\partial\mu^2}\log L(\mu|\mathbf{x})\right|_{\mu=\hat{\mu}} = -\frac{n}{2\sigma^2} < 0$$

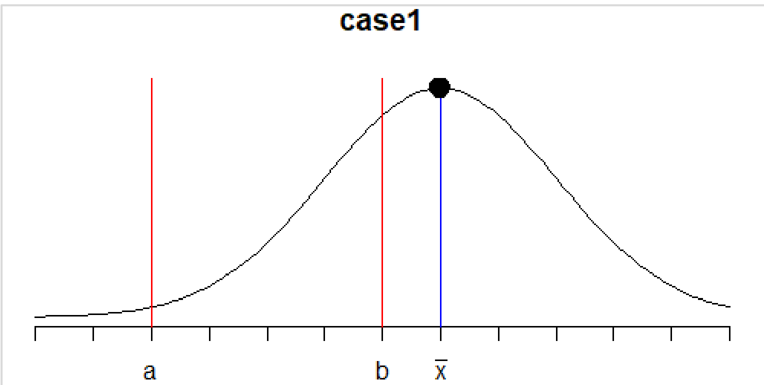
Therefore, $L(\mu|\mathbf{x})$ has a peak.

And we have 4 cases under $X_1, \dots, X_5 \sim N(\mu, 1)$, $a = 0$, $b = 2$.

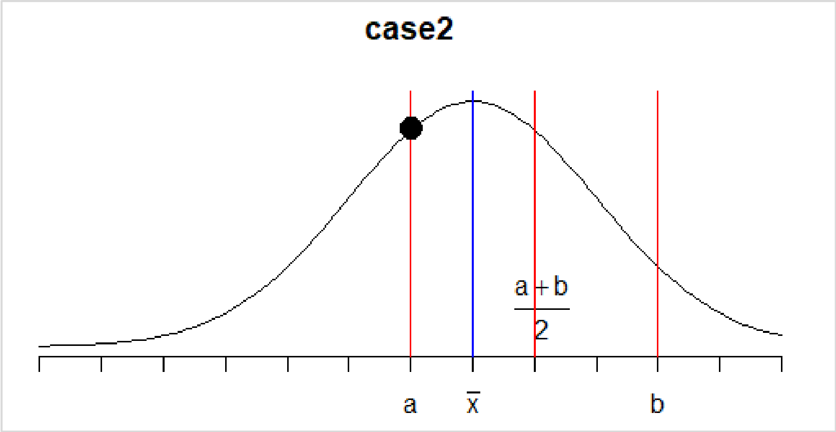
Case 1 : $\hat{\mu} = \bar{X}$, $if\ \bar{X} \le a$



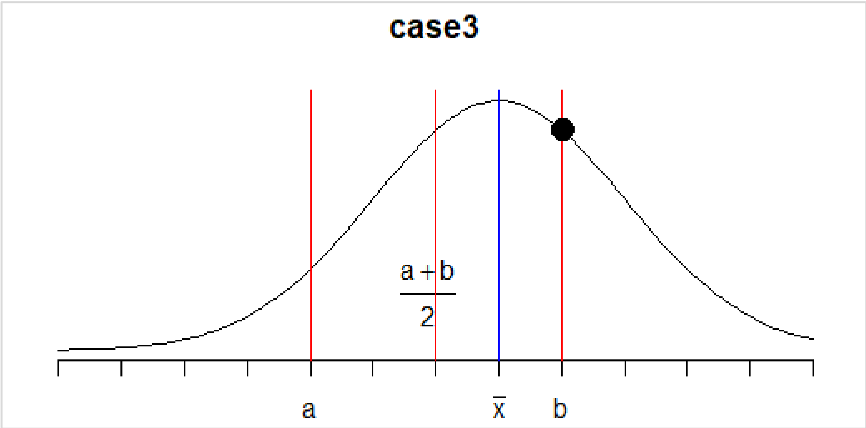
Case 1 : $\hat{\mu} = \bar{X}$, $if\ b \le \bar{X}$



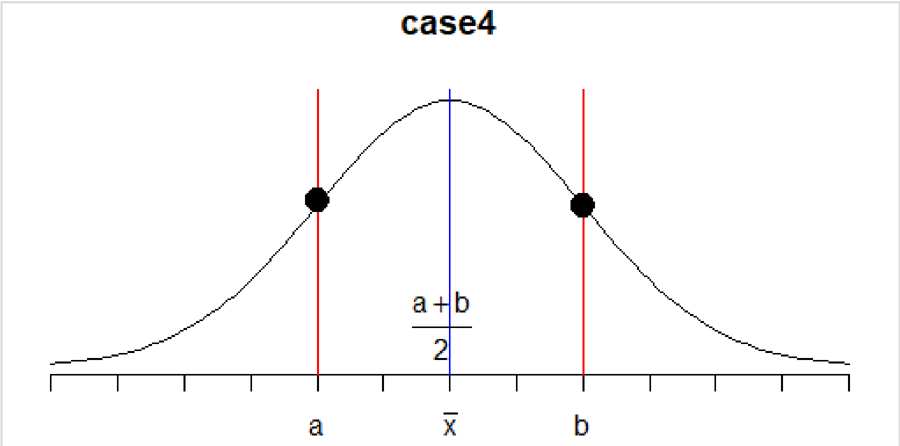
Case 2 : $\hat{\mu} = \mathrm{a}, \ \ if \ \ a < \bar{X} < \frac{a+b}{2}$



Case 3 : $\hat{\mu} = \mathrm{b}, \ \ if \ \ \frac{a+b}{2} < \bar{X} < b$



Case 4 : $\hat{\mu} = a \ or \ b, \ \ if \ \ \bar{X} = \frac{a+b}{2}$



(This case happens with probability 0.)

Q4 Let $X_1, \dots, X_n \sim Gamma(\alpha, \beta)$ as in HW#1. Let $\psi(\alpha) = \frac{d}{d\alpha} log\Gamma(\alpha)$ be the digamma function and $\psi'(\alpha)$ be the trigamma function.

(1) Write down the score functions using the sufficient statistic (T_1, T_2).

(2) Write down the Hessian matrix $H(\alpha, \beta)$.

(3) Let $(\hat{\alpha}, \hat{\beta})$ be the solution to $S_1(\alpha, \beta) = S_2(\alpha, \beta) = 0$. Write down $H(\hat{\alpha}, \hat{\beta})$ in terms of $(\hat{\alpha}, \hat{\beta})$.

Sol :

(1)

$$f(x|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} exp(-\frac{x}{\beta})$$

$$L(\alpha, \beta|\mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} {x_i}^{\alpha-1} \exp(-\frac{x_i}{\beta})$$

$$= (\tfrac{1}{\Gamma(\alpha)})^n \tfrac{1}{\beta^{n\alpha}} (\textstyle\prod_{i=1}^{n} x_i)^{\alpha-1} \exp(-\tfrac{\sum_{i=1}^{n} x_i}{\beta})$$

$$\log L(\alpha, \beta|\mathbf{x}) = -log\Gamma(\alpha)^n - log\beta^{n\alpha} + (\alpha-1) log \prod_{i=1}^{n} x_i - \frac{\sum_{i=1}^{n} x_i}{\beta}$$

$$= -nlog\Gamma(\alpha) - n\alpha log\beta + (\alpha-1) log \textstyle\prod_{i=1}^{n} x_i - \tfrac{\sum_{i=1}^{n} x_i}{\beta}$$

$$\frac{\partial logL(\alpha, \beta|\mathbf{x})}{\partial \alpha} = -n(\frac{d}{d\alpha} log\Gamma(\alpha)) - nlog\beta + log \prod_{i=1}^{n} x_i$$

$$= -n\,\psi(\alpha) - nlog\beta + log \prod_{i=1}^{n} x_i$$

$$\frac{\partial logL(\alpha, \beta|\mathbf{x})}{\partial \beta} = -\frac{n\alpha}{\beta} + \frac{\sum_{i=1}^{n} x_i}{\beta^2}$$

Since $f(x|\alpha, \beta) = h(x) g(T(\mathbf{X})|\alpha, \beta)$, where $h(x) = 1$ and $g(T(\mathbf{X})|\alpha, \beta) = \left(\frac{1}{\Gamma(\alpha)\beta^{\alpha}}\right)^n (\prod_{i=1}^{n} x_i)^{\alpha-1} \exp(-\frac{\sum_{i=1}^{n} x_i}{\beta})$, $T_1(\mathbf{X}) = \prod_{i=1}^{n} X_i$, $T_2(\mathbf{X}) = \sum_{i=1}^{n} X_i$

The score functions are

$$\begin{cases} S_1(\alpha, \beta) = \dfrac{\partial logL(\alpha, \beta|\mathbf{x})}{\partial \alpha} = -n\,\psi(\alpha) - nlog\beta + log\, T_1(X) \\ S_2(\alpha, \beta) = \dfrac{\partial logL(\alpha, \beta|\mathbf{x})}{\partial \beta} = -\dfrac{n\alpha}{\beta} + \dfrac{T_2(X)}{\beta^2} \end{cases}$$

(2) $H(\alpha, \beta) = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 logL(\alpha,\beta|\mathbf{x})}{\partial \alpha^2} & \frac{\partial^2 logL(\alpha,\beta|\mathbf{x})}{\partial \alpha \partial \beta} \\ \frac{\partial^2 logL(\alpha,\beta|\mathbf{x})}{\partial \alpha \partial \beta} & \frac{\partial^2 logL(\alpha,\beta|\mathbf{x})}{\partial \beta^2} \end{bmatrix}$

$$= \begin{bmatrix} (-n)\psi'(\alpha) & -\frac{n}{\beta} \\ -\frac{n}{\beta} & \frac{n\alpha}{\beta^2} - \frac{2T_2(X)}{\beta^3} \end{bmatrix}$$

(3) Because $S_2(\alpha, \beta) = -\frac{n\alpha}{\beta} + \frac{T_2(\mathrm{X})}{\beta^2} = 0 \quad \Rightarrow \hat{\beta} = \frac{T_2(X)}{n\alpha} \quad \Rightarrow \alpha = \frac{T_2(X)}{n\hat{\beta}}$

Therefore, $H(\hat{\alpha}, \hat{\beta}) = \begin{bmatrix} -n\psi'\left(\frac{T_2(X)}{n\hat{\beta}}\right) & -\frac{n}{\hat{\beta}} \\ -\frac{n}{\hat{\beta}} & -\frac{T_2(\mathrm{X})}{\beta^3} \end{bmatrix}$