Solutions

STAT 417 Fall 2016 Dr. DasGupta

First Midterm

NOTE a. Please explain your work legibly and clearly to get credit.

b. You may get a maximum of 4 free bonus points. Answer as much as you can.

1. A manufacturing process produces fibers of variable length. The pdf of the length of a fiber is

$$f(x|\theta) = \theta^{-2} x e^{-\frac{x}{\theta}}, x > 0, \theta > 0.$$

Your data are the lengths $X_1, X_2, \cdots, X_n$ of n independently sampled fibers.

(a) Find with proof a sufficient statistic.
(b) Find the MLE of θ.
(c) Prove or disprove that the MLE of θ is unbiased.
6+6+4 = 16 points

a $$l(\theta) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \theta^{-2} x_i e^{-\frac{x_i}{\theta}} = e^{-\frac{\sum x_i}{\theta}} \theta^{-2n} \prod_{i=1}^{n} x_i$$

$\therefore$ by the Fisher–Neyman factorization theorem, $\sum_{i=1}^{n} X_i$ is a sufficient statistic.

b $$\log l(\theta) = -\frac{\sum x_i}{\theta} - 2n \log\theta + \sum \log x_i$$

$$\therefore \frac{\partial}{\partial\theta} \log l(\theta) = \frac{\sum x_i}{\theta^2} - \frac{2n}{\theta} = \frac{\sum x_i - 2n\theta}{\theta^2}$$

< 0 if $\theta > \frac{\sum x_i}{2n}$,

> 0 if $\theta < \frac{\sum x_i}{2n}$.

$\therefore$ $\log l(\theta)$ (and $l(\theta)$) is maximized at $\theta = \frac{\sum x_i}{2n}$.

c $$E_\theta\left(\frac{\sum X_i}{2n}\right) = \frac{1}{2n} E_\theta(\sum X_i) = \frac{1}{2n} \sum E_\theta(X_i)$$

$$= \frac{1}{2n} n E_\theta(X_1) = \frac{1}{2n} n (2\theta) = \theta.$$

$\therefore$ the mle of θ is unbiased.

2. Let $X_1, X_2, \cdots, X_n \overset{iid}{\sim} N(\theta, (1-\theta)^2), 0 < \theta < 1$.

a. Derive the Fisher information function.

b. Find the CRLB(Cramér-Rao lower bound) for the variance of an unbiased estimate of θ.

c. Give an explicit unbiased estimate of θ.

d. Does your estimate of part c) attain the CRLB? Prove your answer.

5+2+3+4 = 14 points

a

$$f(x|\theta) = \frac{1}{\sqrt{2\pi}\,(1-\theta)} e^{-\frac{(\theta-x)^2}{2(1-\theta)^2}}$$

$$\log f(x|\theta) = -\frac{1}{2}\log 2\pi - \log(1-\theta) - \frac{(\theta-x)^2}{2(1-\theta)^2}$$

$$\frac{\partial}{\partial\theta}\log f(x|\theta) = \frac{1}{1-\theta} - \frac{x-\theta}{(1-\theta)^2} \cdot - \frac{(x-\theta)^2}{(1-\theta)^3}$$

$$\left[\frac{\partial}{\partial\theta}\log f(x|\theta)\right]^2 = \frac{1}{(1-\theta)^2} + \frac{(x-\theta)^2}{(1-\theta)^4} + \frac{(x-\theta)^4}{(1-\theta)^6}$$

$$- \frac{2(x-\theta)}{(1-\theta)^3} - \frac{2(x-\theta)^2}{(1-\theta)^4} + \frac{2(x-\theta)^3}{(1-\theta)^5}.$$

$$\therefore E_\theta\left[\frac{\partial}{\partial\theta}\log f(x|\theta)\right]^2 = \frac{1}{(1-\theta)^2} + \frac{(1-\theta)^2}{(1-\theta)^4} + \frac{3(1-\theta)^4}{(1-\theta)^6}$$

$$- 0 - \frac{2(1-\theta)^2}{(1-\theta)^4} + 0$$

$$= \frac{3}{(1-\theta)^2}$$

$$= I(\theta).$$

3. Suppose $n = 4$ observations with values 0.2, .5, 1.0, 2.0 are available from an $Exp(\lambda)$ density.

a. Give an approximate plot of the likelihood function. The plot should be correct about the main features, but *NEED NOT* be drawn to exact scale.

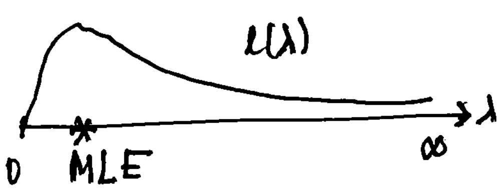
b. Find the numerical value of the MLE of $1/\lambda$.

c. Find the mean squared error of the MLE of $1/\lambda$, and plot it.

4+4+3 = 11 points

$f(x|\lambda) = \lambda e^{-\lambda x}$.

a $\quad l(\lambda) = \prod_{i=1}^{4} f(x_i|\lambda) = \lambda^4 e^{-\lambda(.2+.5+1+2)} = \lambda^4 e^{-3.7\lambda}$, $\lambda > 0$.



b $\quad \log l(\lambda) = 4\log\lambda - 3.7\lambda$

$$\frac{d}{d\lambda}\log l(\lambda) = \frac{4}{\lambda} - 3.7 = \frac{4 - 3.7\lambda}{\lambda}$$

< 0 for $\lambda > \frac{4}{3.7}$

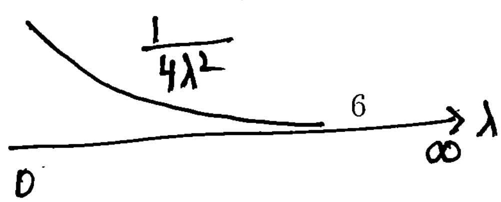
> 0 for $\lambda < \frac{4}{3.7}$.

$\therefore$ the mle of λ is $\frac{4}{3.7}$, and so the mle of $\frac{1}{\lambda}$ is $\frac{3.7}{4}$

$= .925. = \bar{X}$

c $\quad MSE = E(MLE - \frac{1}{\lambda})^2 = E(\bar{X} - \frac{1}{\lambda})^2 = Var(\bar{X})$

$= \frac{\sigma^2}{n} = \frac{1}{4\lambda^2}$.



4. True or False?

(No proofs required; but your answers must be visually unambiguous)

a) The Fisher information function has to be a decreasing function of the underlying parameter. F

b) If unbiased estimates for some parameter exist, then a best unbiased estimate exists. F

c) The best unbiased estimate attains the Cramér-Rao lower bound. F

d) A lower dimensional sufficient statistic exists only in the Exponential family of distributions. F

e) UMVUEs are always asymptotically unbiased. T

f) The likelihood function is a random variable. T

g) The factorization theorem is due to Fisher, Rao, and Neyman. F

$7 \times 3 = 21$ points