Stat511(Section001): Homework #4

Due on Feb. 9, 2021 at 5:00pm

Instructor:Xuming He

Tiejin Chen **tiejin@umich.edu**

Problem 1

Since $f_\theta(x)$ is the pdf of X. Thus, if we take intergation of it, we can get:

$$\int a(\theta)h(x)exp(T(x)^T\theta)dx = 1$$

If we see $\int a(\theta)h(x)exp(T(x)^T\theta)dx$ as a function of θ. Then it is a constant function. And we can get:

$$\frac{d}{d\theta}\int a(\theta)h(x)exp(T(x)^T\theta)dx = 0$$

In other hand, we can see $\int a(\theta)h(x)exp(T(x)^T\theta)dx = a(\theta)\int h(x)exp(T(x)^T\theta)dx$ which is the multiply by two function of θ. Thus:

$$\frac{d}{d\theta}\int a(\theta)h(x)exp(T(x)^T\theta)dx = \frac{da(\theta)}{d\theta}\int h(x)exp(T(x)^T\theta)dx + a(\theta)\int h(x)exp(T(x)^T\theta)T(x)dx = 0$$

And we find that$a\ (\theta)\int h(x)exp(T(x)^T\theta)T(x)dx$ is $E_\theta(T(x))$ since $T(x)$ is nature parameter.
And $\int a(\theta)h(x)exp(T(x)^T\theta)dx = \frac{1}{a(\theta)}$. Thus we have:

$$\frac{da(\theta)}{d\theta}\frac{1}{a(\theta)} + E_\theta(T(x)) = 0 \rightarrow E_\theta(T(x)) = -\frac{dln(a(\theta))}{d\theta}$$

Problem 2

(a)

We can rewrite the pdf of pdf of x as:

$$f_\theta(X) = exp(-X+\theta)I(X>\theta)$$

Hence, we can get:

$$f_\theta(X_1,..,X_n) = exp(n\theta - \sum_{i=1}^{n}X_i)\prod_{i=1}^{n}I(X_i>\theta) = exp(n\theta - \sum_{i=1}^{n}X_i)I(X_{(1)}>\theta)$$

Thus, we have $X_{(1)}$ is the sufficient statistics for θ by factorization therom. Now we need to prove it is complete. First, we need to compute pdf of $X(1)$. Since it is first order statistics, we have:

$$f_{x_{(1)}}(x) = nexp(-x+\theta)(1-F(x))^{n-1}, x>\theta$$

where $F(x)$ is the cdf of x.And we all know that since x must greater than $\theta, F(\theta) = 0$. For any function g we let:

$$\int_\theta^\infty ng(t)exp(-t+\theta)(1-F(t))^{n-1}dt = 0$$

For all θ, Then, We take derivation of it to get:

$$-nexp(-\theta+\theta)g(\theta)(1-F(\theta))^{n-1} = 0 \rightarrow g(\theta) = 0$$

Problem 2 continued on next page...

for all θ. Hence we can get function $g(x)$ is zero function. Thus we prove that $X_{(1)}$ is also complete. And we get it is a complete sufficient statistics.
(b)

Intuitively, we think it is independent. To prove this, we only need to prove $X_{(n)} - X_{(1)}$ is an ancillary statistics. Assumed $Y = X - \theta$, then we have:

$$f_\theta(Y) = exp(-Y + 0)$$

Thus Y is also one of family of distribution. Then we have:

$$Y_{(n)} - Y_{(1)} = X_{(n)} - \theta - (X_{(1)} - \theta) = X_{(n)} - X_{(1)}$$

Thus, $X_{(n)} - X_{(1)}$ does not depends on θ and it is an ancillary statistics. Thus we get $X_{(1)}$ and $X_{(n)} - X_{(1)}$ are independent.

Problem 3

(a)

We know mean is λ, thus pdf of x is:

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

Then we have:

$$f(x_1, \ldots, x_n) = \prod_{\iota=1}^{n} \frac{e^{-\lambda}\lambda^{x_\iota}}{x_i!} = \frac{e^{-n\lambda}\lambda^{\Sigma_{\iota=1}^{n} x_x}}{\prod_{i=1}^{n} x_i!}$$

We can rewrite $\lambda^{\Sigma_{i=1}^{n} x_i}$ as $exp\ (\sum_{i=1}^{n} x_i log(\lambda))$. Thus we can get:

$$f(x_1, ..., x_n) = \frac{e^{(-n\lambda)}}{\prod_{i=1}^{n} x_i!} exp(\sum_{i=1}^{n} x_i log(\lambda))$$

which is an exponential family of distribution and we have $T(x) = \sum_{i=1}^{n} x_i$ is a sufficient statistics. And we have $b(\lambda) = log(\lambda)$. Hence it is a full rank expoential family of distribution. Thus $T(x)$ is complete. We also know $T(x) = n\overline{X}$. Hence $\overline{X}$ is also a complete and sufficent statistics.
(b)

We know for $\overline{X}$, it is unbias estimator of λ, Thus the MSE of $\overline{X}$ only depends on $var(\overline{X})$. And we know:

$$var(\overline{X}) = \frac{\lambda}{n} = \frac{\lambda}{3}$$

when $n = 3$ since the variance of x is also λ. We also know that the expectation of sample variance is $var(X) = \lambda$, hence it is also unbias estimator. Then, we know:

$$Var(S^2) = \frac{1}{n}(\theta_4 - \frac{n-3}{n-1}\theta_2^2) = \frac{\theta_4}{3}$$

when $n = 3$, where θ_4 is fourth centered moment, θ_2 is second centered moment. We prove this in homework of stat510 last semester. Now let us consider$E\ [(x - \lambda)^4]$:

Problem 3 continued on next page. . .

$$E[(x-\lambda)^4] = \sigma^4 \times Kurtosis(X) = \lambda^2 \times (\lambda^{-1}+3) = \lambda + 3\lambda^2$$

Thus the MSE of S^2 is $\dfrac{\lambda}{3} + \lambda^2$ and the MSE of $\overline{X}$ is $\dfrac{\lambda}{3}$. Hence we conclude sample mean $\overline{X}$ is better under MSE.

Problem 4

(a)

We know the expectation of X will be zero. Hence we will consider from second moment.We have:

$$E(X^2) = \int_{-\theta}^{\theta} \frac{1}{2\theta} x^2 dx = \frac{\theta^2}{3}$$

Thus we have:

$$T(x) = \sqrt{\frac{3\sum_{i=1}^{n} X_i^2}{n}}$$

is a method of moments estimator.

(b)

We know:

$$E(T(X)^2) = 3E(X_i^2) = \theta^2$$

And the variance of$T\ (x)$ should not be zero. Hence:

$$Var(T(X)) = E(T(X)^2) - (E(T(X)))^2 \neq 0 \rightarrow E(T(X)) \neq \sqrt{E(T(X)^2)} \rightarrow E(T(X)) \neq \theta$$

Thus it is a biased estimator.

(c)

We have:

$$f(x_1, \ldots, x_n) = \frac{1}{(2\theta)^n} \prod_{i=1}^{n} I(-\theta < x_i < \theta) = \frac{1}{2^n \theta^n} I(\max(|x_i|) < \theta)$$

This can be also wirren as:

$$f(x_1, \ldots, x_n) = \frac{1}{2^n \theta^n} I(\theta > \max(|x_i|))$$

To make this function as large as possible, we need to make θ as small as possible. And it should make $I(\theta > \max(|x_i|) = 1$. Thus the MLE estimator is $max\ (|x_i|)$.