Homework 6: Written Part (SOLUTIONS)

STAT 343: Mathematical Statistics

Details

How to Write Up

The written part of this assignment can be either typeset using latex or hand written.

Grading

5% of your grade on this assignment is for turning in something legible and organized

An additional 15% of your grade is for completion. A quick pass will be made to ensure that you've made a reasonable attempt at all problems.

Across both the written part and the R part, in the range of 1 to 3 problems will be graded more carefully for correctness. In grading these problems, an emphasis will be placed on full explanations of your thought process. You don't need to write more than a few sentences for any given problem, but you should write complete sentences! Understanding and explaining the reasons behind what you are doing is at least as important as solving the problems correctly.

Collaboration

You are allowed to work with others on this assignment, but you must complete and submit your own write up. You should not copy large blocks of code or written text from another student.

Sources

You may refer to our text, Wikipedia, and other online sources. All sources you refer to must be cited.

Problem I: Binomial Distribution

We have previously found that if $X \sim \text{Binomial}(n, \theta)$ then the maximum likelihood estimator of θ is $\hat{\theta}^{MLE} = \frac{X}{n}$. We showed that $E(\hat{\theta}^{MLE}) = \theta$, $Var\left(\hat{\theta}^{MLE}\right) = \frac{\theta(1-\theta)}{n}$, and $MSE\left(\hat{\theta}^{MLE}\right) = \frac{\theta(1-\theta)}{n}$.

(a) Is it possible for another estimator $\tilde{\Theta}$ to have lower variance than the MLE? If so, give an example of such an estimator. If not, explain why not with reference to a theorem from class.

Yes, it is possible for another estimator $\tilde{\Theta}$ to have lower variance than the MLE. One example would be $\tilde{\Theta} = \frac{X}{n+1}$. In this case, $Var\left(\tilde{\Theta}\right) = \frac{n\theta(1-\theta)}{(n+1)^2} < \frac{n\theta(1-\theta)}{n^2} = \frac{\theta(1-\theta)}{n} = Var\left(\hat{\theta}^{MLE}\right)$ $\forall n \geq 1$.

(b) Is it possible for another *unbiased* estimator $\tilde{\Theta}$ to have lower variance than the MLE? If so, give an example of such an estimator. If not, explain why not with reference to a theorem from class.

No, it is not possible for another *unbiased* estimator $\tilde{\Theta}$ to have lower variance than the MLE. Since we are considering unbiased estimators for θ, we know that for all unbiased estimators of θ, $\tilde{\Theta}$,

$$Var\left(\tilde{\Theta}\right) \geq \frac{1}{nI_i(\theta)}.$$

Now, we need to find $I_i(\theta)$ to determine if $Var\left(\hat{\Theta}^{MLE}\right) = \frac{1}{nI_i(\theta)}$.

$$\begin{aligned}
I_i(\theta) &= E\left[-\frac{d^2}{d\theta^2}\log\{f_X(x_i|\theta)\}\right] \\
&= E\left[-\frac{d^2}{d\theta^2}\left(X_i \log\theta + (1-X_i)\log(1-\theta)\right)\right] \\
&= E\left[-\frac{d}{d\theta}\left(\frac{X_i}{\theta} - \frac{1-X_i}{1-\theta}\right)\right] \\
&= E\left[-\left(-\frac{X_i}{\theta^2} - \frac{1-X_i}{(1-\theta)^2}\right)\right] \\
&= E\left[\frac{X_i}{\theta^2} + \frac{1-X_i}{(1-\theta)^2}\right] \\
&= E\left[\frac{X_i}{\theta^2}\right] + E\left[\frac{1-X_i}{(1-\theta)^2}\right] \\
&= \frac{\theta}{\theta^2} + \frac{1-\theta}{(1-\theta)^2} \\
&= \frac{1}{\theta} + \frac{1}{1-\theta} \\
&= \frac{1}{\theta(1-\theta)}
\end{aligned}$$

Thus, $Var\left(\hat{\theta}^{MLE}\right) = \frac{\theta(1-\theta)}{n} = \frac{1}{nI_i(\theta)}$, so the MLE attains (is equal to) the Cramer-Rao Lower Bound. This means that it has the smallest variance among all *unbiased* estimators.

Problem II. Exponential Distribution

Let $X_1, ..., X_n$ be an i.i.d. sample from an exponential distribution with the density function

$$f(x|\tau) = \frac{1}{\tau}e^{-x/\tau},\ 0 \le x < \infty.$$

It can be shown that the MLE of τ is $\hat{\tau}^{MLE} = \bar{X}$.

(a) Show that exact sampling distribution of the MLE is Gamma$(n, \tau/n)$? *Hint: You can use Moment Generating Functions to show this.*

First, rewrite $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

Because we know the distribution for X_i, we can look up (or derive) the moment generating function: $M_X(t) = [1 - \tau t]^{-1}$ for our parameterization.

Since $\sum_{i=1}^{n} X_i$ is the sum of i.i.d. random variables, $M_{\sum_{i=1}^{n} X_i}(t) = [M_X(t)]^n$.

Also, since $\frac{1}{n}$ is a constant,

$$\begin{aligned} M_{\bar{X}}(t) &= M_{\frac{1}{n}\sum_{i=1}^{n} X_i}(t) \\ &= \left[M_{\frac{1}{n}X}(t)\right]^n \\ &= \left[M_X\left(\frac{1}{n}t\right)\right]^n \\ &= \left[\left(1 - \tau\left(\frac{1}{n}t\right)\right)^n\right]^{-1} \\ &= \left[1 - \left(\frac{\tau}{n}\right)t\right]^{-n} \end{aligned}$$

This is the moment generating function for a $Gamma\,(n, \tau/n)$ distribution, so this is the exact sampling distribution of the MLE.

(b) Use the Central Limit Theorem to find a normal approximation of the sampling distribution.

By the CLT,

$$\sqrt{n}\left(\hat{\tau}^{MLE} - \tau\right) \xrightarrow{d} Normal\left(0, \frac{1}{I_i(\tau)}\right)$$

or alternatively,

$$\hat{\tau}^{MLE} \overset{\text{approx.}}{\sim} Normal\left(\tau, \frac{1}{nI_i(\tau)}\right)$$

where $I_i(\tau) = \frac{1}{\tau^2}$. (Be sure to indicate what $I_i(\tau)$ is specifically in this context.)

(c) Show that the MLE is unbiased, and find its exact variance. (*Hint: The sum of the X_i follows a gamma distribution.*)

WTS: $E\left(\hat{\tau}^{MLE}\right) - \tau = 0$.

$E\left(\hat{\tau}^{MLE}\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} \times n\tau = \tau$. Therefore, $E\left(\hat{\tau}^{MLE}\right) - \tau = 0$ so $\hat{\tau}^{MLE}$ is an unbiased estimator for τ.

$Var\left(\hat{\tau}^{MLE}\right) = n\left(\frac{\tau}{n}\right)^2 = \frac{\tau^2}{n}$, using the exact sampling distribution result from (a).

(d) Is there any other unbiased estimate with smaller variance?

We already have shown that $\hat{\tau}^{MLE}$ is unbiased, and we found the variance in (c). If $Var\left(\hat{\tau}^{MLE}\right) = \frac{1}{nI_i(\tau)}$, then it attains the Cramer Rao Lower Bound, and there is no unbiased estimator with smaller variance.

With a few more steps than are shown here, we can show

$$\begin{aligned}
nI_i(\tau) &= I(\tau) \\
&= E\left[-\frac{d^2}{d\tau^2}\ell(\tau|x_1, ..., x_n)\right] \\
&= -E\left[\frac{n}{\tau^2} - \frac{2}{\tau^3}\sum_{i=1}^{n} X_i\right] \\
&= -\frac{n}{\tau^2} + \frac{2}{\tau^3}\sum_{i=1}^{n} E(X_i) \\
&= -\frac{n}{\tau^2} + \frac{2n}{\tau^3}\tau \\
&= \frac{n}{\tau^2}
\end{aligned}$$

So, $Var\left(\hat{\tau}^{MLE}\right) = \frac{1}{nI_i(\tau)}$. Therefore, there is no unbiased estimator for τ with smaller variance.

(e) Find the form of an approximate $(1-\alpha)*100\%$ confidence interval for τ.

Pivotal quantity: $\frac{\hat{\tau}^{MLE}-\tau}{SE(\hat{\tau}^{MLE})} = \frac{\hat{\tau}^{MLE}-\tau}{\sqrt{\tau^2/n}} \sim Normal(0,1)$, which we get by centering and scaling the result from part (b).

$$\begin{aligned}
1-\alpha &= P\left(z_{\alpha/2} \le \frac{\bar{X}-\tau}{\tau/\sqrt{n}} \le z_{1-\alpha/2}\right); \; z_q = q^{th}\text{quantile from a standard normal distribution} \\
&= P\left(z_{\alpha/2} \le \sqrt{n}\left(\bar{X}/\tau - 1\right) \le z_{1-\alpha/2}\right) \\
&= P\left(\frac{z_{\alpha/2}}{\sqrt{n}} \le \left(\bar{X}/\tau - 1\right) \le \frac{z_{1-\alpha/2}}{\sqrt{n}}\right) \\
&= P\left(1 + \frac{z_{\alpha/2}}{\sqrt{n}} \le \frac{\bar{X}}{\tau} \le 1 + \frac{z_{1-\alpha/2}}{\sqrt{n}}\right) \\
&= P\left(\frac{1}{\bar{X}}\left(1 + \frac{z_{\alpha/2}}{\sqrt{n}}\right) \le \frac{1}{\tau} \le \frac{1}{\bar{X}}\left(1 + \frac{z_{1-\alpha/2}}{\sqrt{n}}\right)\right) \\
&= P\left(\bar{X}\left(1 + \frac{z_{1-\alpha/2}}{\sqrt{n}}\right)^{-1} \le \tau \le \bar{X}\left(1 + \frac{z_{\alpha/2}}{\sqrt{n}}\right)^{-1}\right) \\
&= P\left(\bar{X}\left(1 + \frac{z_{1-\alpha/2}}{\sqrt{n}}\right)^{-1} \le \tau \le \bar{X}\left(1 - \frac{z_{1-\alpha/2}}{\sqrt{n}}\right)^{-1}\right)
\end{aligned}$$

(f) Find the form of an exact confidence interval for τ.

Step 1: Find pivotal quantity.

First thought - rescale $\hat{\tau}^{MLE}$ by $\frac{n}{\tau}$. Then,

$$M_{\frac{n}{\tau}\bar{X}}(t) = M_{\bar{X}}\left(\frac{n}{\tau}t\right) = [1-t]^{-n}$$

This is the MGF for a $Gamma(n,1)$ random variable, and since the distribution does not depend on τ, and $\frac{n}{\tau}\bar{X}$ contains the unknown parameter τ, this is a viable pivotal quantity.

Another alternative is to rescale $\hat{\tau}^{MLE}$ by $\frac{2n}{\tau}$. Then,

$$M_{\frac{n}{\tau}\bar{X}}(t) = M_{\bar{X}}\left(\frac{2n}{\tau}t\right) = [1-2t]^{-n} = [1-2t]^{-2n/2}$$

which is the MGF for a χ^2_{2n} random variable. This may be a more typical choice for a pivotal quantity, but either will work. I am going to proceed with the second one.

Step 2:

Set up the interval: let $\chi(q)$ denote the q^{th} quantile of a χ^2_{2n} distribution.

$$\begin{aligned}
1-\alpha &= P\left(\chi(\alpha/2) \leq \frac{2n}{\tau}\bar{X} \leq \chi(1-\alpha/2)\right) \\
&= P\left(\frac{\chi(\alpha/2)}{2n\bar{X}} \leq \frac{1}{\tau} \leq \frac{\chi(1-\alpha/2)}{2n\bar{X}}\right) \\
&= P\left(\frac{2n\bar{X}}{\chi(\alpha/2)} \geq \tau \geq \frac{2n\bar{X}}{\chi(1-\alpha/2)}\right) \\
&= P\left(\frac{2n\bar{X}}{\chi(1-\alpha/2)} \leq \tau \leq \frac{2n\bar{X}}{\chi(\alpha/2)}\right)
\end{aligned}$$