EE 503

Homework 13 Solution

Instructor: Christopher Wayne Walker, TA: James Huang

Problem 1. Suppose you are going to conduct a political poll. You wish to have a margin of error of ± 3% and a 95% confidence interval. Determine the number of people you need to poll.

Solution: To estimate $p \in [0,1]$ by conducting a poll with 0.03 margin of error and $1-\alpha = 0.95$, we need

$$n = \frac{z_{\frac{\alpha}{2}}^2 p(1-p)}{(\text{M.O.E.})^2} = \frac{1.96^2 p(1-p)}{0.03^2} = \frac{1.96^2 p(1-p)}{0.03^2} \leq \frac{1.96^2 \cdot 0.25}{0.03^2}.$$

To make sure we have enough samples for any $p \in (0,1)$, we pick n greater than that number, which is 1068.

Problem 2. Suppose you are going to perform a bit error rate (BER) simulation. How many bit errors should you count so that you are 95% confident the true BER is within ± 5% of the BER calculated from you simulation?

Solution: To estimate the BER (p close to 0) with ± 5% margin of error of p and $1-\alpha = 0.95$, we need error counts

$$k = np = \frac{z_{\frac{\alpha}{2}}^2 p(1-p)}{(0.05p)^2} \cdot p = \frac{1.96^2(1-p)}{0.05^2} \approx \frac{1.96^2}{0.05^2}.$$

Pick $k = 1537$.

Problem 3. Suppose you perform a simulation of a communication system and you are processing received bits. When errors occur in these bits the errors are independent. Suppose after processing 10 million bits you observe no errors. How confident are you that the true bit error rate is no higher than 10^{-5}?

Solution: We estimate the error probability $\hat{p} = \frac{0}{10^7} = 0$ with desired M.O.E.

$= 10^{-5}$. By central limit theorem, we know that with probability $1-\alpha$

$$p \in \left[\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}, \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right].$$

Hence, $z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{10^7}} = 10^{-5} \Rightarrow z_{\frac{\alpha}{2}} \approx \frac{10^{-1.5}}{\sqrt{p}}$ since $1-p \approx 1$. The probability that p is outside of the confidence interval is $\alpha = 2 \times \left(1 - \Phi(z_{\frac{\alpha}{2}})\right) = 2 \times \left(1 - \Phi\left(\frac{10^{-1.5}}{\sqrt{p}}\right)\right)$ where $\Phi(\cdot)$ is standard normal cdf. Notice that BER p is very small, thus α is close to 0.

Problem 4. Suppose you gather some data and compute the sample mean and it has a value of 7.7 based on a sample size of 200. Because of the relatively large sample size we can assume the sample mean is normal (or Gaussian).

a. Construct a 95% confidence interval for the true mean of the data if the true variance of the data is known to be 1.2.

b. Construct a 95% confidence interval for the true mean of the data if the variance of the data is unknown but the sample variance is 1.2.

Solution:

a. Assume we know the true variance $\sigma^2 = 1.2$, the 95% confidence interval is

$$\text{C.I.} = \left[\bar{X} - z_{0.025}\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{0.025}\frac{\sigma}{\sqrt{n}}\right] = \left[7.7 - 1.96\sqrt{\frac{1.2}{200}}, 7.7 + 1.96\sqrt{\frac{1.2}{200}}\right].$$

b. If we only have the sample variance $S^2 = 1.2$, the 95% confidence interval is

$$\text{C.I.} = \left[\bar{X} - t_{0.025,n-1}\frac{S}{\sqrt{n}}, \bar{X} + t_{0.025,199}\frac{S}{\sqrt{n}}\right] = \left[7.7 - 1.972\sqrt{\frac{1.2}{200}}, 7.7 + 1.972\sqrt{\frac{1.2}{200}}\right].$$

Problem 5. Let $X = (X_1, X_2, \ldots, X_n)$ be i.i.d. where each $X_i \sim U(\theta_1, \theta_2)$ where θ_1 and θ_2 are unknown. Find the MLE for θ_1 and θ_2.

Solution: Given $\{X_i\}$, we find the MLEs as

$$\begin{aligned}
\hat{\theta}_1 &= \arg\max_{\theta_1} \prod_{i=1}^{n} f_X(X_i; \theta_1, \theta_2) \\
&= \arg\max_{\theta_1} \left(\frac{1}{\theta_2 - \theta_1}\right)^n \prod_{i=1}^{n} \mu(X_i - \theta_1) \\
&= \min_i \{X_i\}. \\
\hat{\theta}_2 &= \arg\max_{\theta_2} \prod_{i=1}^{n} f_X(X_i; \theta_1, \theta_2) \\
&= \arg\max_{\theta_2} \left(\frac{1}{\theta_2 - \theta_1}\right)^n \prod_{i=1}^{n} \mu(\theta_2 - X_i) \\
&= \max_i \{X_i\}.
\end{aligned}$$

Problem 6. Let $X = (X_1, X_2, \ldots, X_n)$ be i.i.d. where each $X_i \sim U(\theta, 2\theta)$ where θ is unknown. Find the MLE for θ.

Solution:

$$\begin{aligned}
\hat{\theta} &= \arg\max_{\theta} \prod_{i=1}^{n} f_X(X_i; \theta) \\
&= \arg\max_{\theta} \left(\frac{1}{2\theta - \theta}\right)^n \prod_{i=1}^{n} \mu(2\theta - X_i)\mu(X_i - \theta) \\
&= \max_i \left\{\frac{X_i}{2}\right\}.
\end{aligned}$$

Notice that we want to find the smallest θ maximizing $\left(\frac{1}{\theta}\right)^n$ while satisfying

$2\theta \geq X_i, \forall i$ and $X_i \geq \theta, \forall i$. The constraints imply

$$\theta \geq \max_i \left\{ \frac{X_i}{2} \right\},$$
$$\theta \leq \min_i \{X_i\}.$$

Hence, we find the smallest available θ which is $\max_i \left\{ \frac{X_i}{2} \right\}$.

Problem 7. Let $X = (X_1, X_2, \ldots, X_n)$ be i.i.d. where each $X_i \sim U(a - \theta, a + \theta)$ where both a and θ are unknown. Find the MLE for a and θ.

Solution:

$$(\hat{a}, \hat{\theta}) = \arg\max_{a,\theta} \left(\frac{1}{2\theta} \right)^n \prod_{i=1}^{n} \mu(a + \theta - X_i)\mu(X_i - a + \theta).$$

Notice that the value of the likelihood function is $\left(\frac{1}{2\theta}\right)^n$ as long as a satisfying $a + \theta \geq X_i$ and $X_i \geq a - \theta$. We want to find the smallest θ maximizing $\left(\frac{1}{2\theta}\right)^n$ while satisfying the constraints

$$\theta \geq \max_i \{X_i\} - a,$$
$$\theta \geq a - \min_i \{X_i\}.$$

Hence,

$$\hat{\theta} = \max \left\{ \max_i \{X_i\} - a, a - \min_i \{X_i\} \right\}.$$

Observe that $\hat{\theta}$ is minimize if we pick

$$a = \frac{\max_i\{X_i\} + \min_i\{X_i\}}{2}.$$

Hence, we have the MLEs as

$$(\hat{a}, \hat{\theta}) = \left(\frac{\max_i\{X_i\} + \min_i\{X_i\}}{2}, \frac{\max_i\{X_i\} - \min_i\{X_i\}}{2} \right).$$

Problem 8. Let $X = (X_1, X_2, \ldots, X_n)$ be i.i.d. where each X_i is Poisson distributed with parameter $\lambda \geq 0$, that is, each X_i has probability mass function

$$P(X_i = x_i) = \frac{e^{-\lambda}\lambda^{x_i}}{x_i!}, \quad i = 0, 1, 2, ...$$

Find the MLE for λ.

Solution: We find the MLE of λ with the likelihood function $L(\lambda) = \prod_{i=1}^{n} p_X(X_i; \lambda)$.

$$\begin{aligned}
\hat{\lambda} &= \arg\max_{\lambda} \prod_{i=1}^{n} p_X(X_i; \lambda) \\
&= \arg\max_{\lambda} \log\left(\prod_{i=1}^{n} p_X(X_i; \lambda)\right) \\
&= \arg\max_{\lambda} -n\lambda + \log\lambda \sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \log(X_i!).
\end{aligned}$$

Observe that the log-likelihood function is a concave function of λ, we can take derivative w.r.t. λ and find the maximum by setting the derivative to 0, i.e.

$$\begin{aligned}
&\frac{\partial}{\partial\lambda} \log L(\lambda)\Big|_{\lambda=\hat{\lambda}} = 0 \\
\Rightarrow & -n + \frac{1}{\lambda}\sum_{i=1}^{n} X_i\Big|_{\lambda=\hat{\lambda}} = 0 \\
\Rightarrow & \hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} X_i,
\end{aligned}$$

which is the sample mean of the iid Poisson random samples.

Problem 9. Let $X = (X_1, X_2, \ldots, X_n)$ be i.i.d. where each X_i is Poisson distributed with parameter $\lambda \geq 0$, that is, each X_i has probability mass function

$$P(X_i = x_i) = \frac{e^{-\lambda}\lambda^{x_i}}{x_i!}, \quad i = 0, 1, 2, ...$$

Find the Cramer-Rao Lower Bound for an unbiased estimator of λ. Note that the Poisson distribution is a member of the exponential family. Does the MLE for λ found in Problem 8 achieve this lower bound?

Solution:

Cramer-Rao Lower Bound tells us that for any unbiased estimator $\hat{\lambda}(X_1, X_2, \cdots, X_n)$ constructed from iid Poisson samples, its variance is lower bounded by

$$\mathrm{var}(\hat{\theta}(\mathbf{X})) \geq \frac{1}{-n\mathbf{E}\left[\frac{\partial^2}{\partial\lambda^2}\log p_X(X_i;\lambda)\right]}.$$

where

$$\begin{aligned}\frac{\partial}{\partial\lambda}\log p_X(X_i;\lambda) &= -1 + \frac{X_i}{\lambda},\\ \frac{\partial^2}{\partial\lambda^2}\log p_X(X_i;\lambda) &= \frac{-X_i}{\lambda^2},\\ \mathbf{E}\left[\frac{\partial^2}{\partial\lambda^2}\log p_X(X_i;\lambda)\right] &= \frac{-1}{\lambda}.\end{aligned}$$

Therefore,

$$\mathrm{var}(\hat{\theta}(\mathbf{X})) \geq \frac{\lambda}{n}.$$

The MLE we found has variance

$$\begin{aligned}\mathrm{var}(\hat{\lambda}) &= \mathrm{var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)\\ &= \frac{\lambda}{n},\end{aligned}$$

which meets CRLB.

Problem 10. Let $X = (X_1, X_2, \ldots, X_n)$ be i.i.d. where each X_i has pdf $f(x|\theta) = \frac{1}{\theta}$, $0 < x < \theta$, and zero elsewhere.

a. Show the Cramer-Rao Lower Bound for the variance of any estimator, W, of θ satisfies

$$Var_\theta W \geq \frac{\theta^2}{n}.$$

b. Let $Y = X_{(n)}$, that is, Y is the maximum value observed. Let

$$Z = \frac{n+1}{n}Y.$$

Show Z is an unbiased estimator of θ.

c. Show

$$Var_\theta Z = \frac{\theta^2}{n(n+2)}.$$

d. The variance of Z is smaller than the Cramer-Rao Lower Bound. Explain, mathematically, why this does not violate the Cramer-Rao Inequality. Note: for this part you may need to use

Leibnitz's Rule: If $f(x,\theta), a(\theta)$ and $b(\theta)$ are differentiable with respect to θ then

$$\begin{aligned}\frac{d}{d\theta}\int_{a(\theta)}^{b(\theta)} f(x,\theta)dx &= f(b(\theta),\theta)\frac{d}{d\theta}b(\theta) - f(a(\theta),\theta)\frac{d}{d\theta}a(\theta) \\ &+ \int_{a(\theta)}^{b(\theta)} \frac{\partial}{\theta} f(x,\theta)dx.\end{aligned}$$

Solution:

a. If we didn't check that uniform distribution $f_X(x;\theta) = \frac{1}{\theta}, \forall x \in (0,\theta)$ does not satisfy the regularity conditions before applying CRLB, it may lead us wrongly that for all unbiased estimator $W(X_1, X_2, \cdots, X_n)$,

$$\begin{aligned}\mathrm{var}(W) &\geq \frac{1}{n\mathbf{E}\left[\left(\frac{\partial}{\partial\theta}\log f_X(X;\theta)\right)^2\right]} \\ &= \frac{1}{n\mathbf{E}\left[\left(\frac{\partial}{\partial\theta}(-\log\theta)\right)^2\right]} \\ &= \frac{\theta^2}{n}.\end{aligned}$$

b. Let $Y = \max_i\{X_i\}$ where $F_Y(y) = \prod_i F_{X_i}(y)$ and $F_{X_i}(y) = y/\theta$ if

$y \in (0, \theta)$, we have

$$\begin{aligned}
\mathbf{E}[Z] &= \frac{n+1}{n}\mathbf{E}[Y] \\
&= \frac{n+1}{n}\int_0^\infty (1 - F_Y(y))dy \\
&= \frac{n+1}{n}\int_0^\theta \left(1 - \left(\frac{y}{\theta}\right)^n\right) dy \\
&= \frac{n+1}{n} \cdot \left(y - \frac{y^{n+1}}{(n+1)\theta^n}\right)_{y=0}^\theta \\
&= \theta.
\end{aligned}$$

Hence, Z is an unbiased estimator of θ.

c. Since $F_Y(y) = \left(\frac{y}{\theta}\right)^n$ if $y \in (0, \theta)$, it has pdf $f_Y(y) = \frac{ny^{n-1}}{\theta^n}$ if $y \in (0, \theta)$.

$$\begin{aligned}
\mathbf{E}[Y^2] &= \int_0^\theta y^2 f_Y(y)\, dy \\
&= \int_0^\theta \frac{ny^{n+1}}{\theta^n}\, dy \\
&= \frac{n}{n+2}\theta^2. \\
\mathrm{var}(Y) &= \frac{n}{n+2}\theta^2 - \frac{n^2}{(n+1)^2}\theta^2 \\
&= \frac{n}{(n+2)(n+1)^2}\theta^2. \\
\mathrm{var}(Z) &= \frac{(n+1)^2}{n^2}\mathrm{var}(Y) \\
&= \frac{\theta^2}{n(n+2)}.
\end{aligned}$$

d. The uniform distribution $f_X(x; \theta) = \frac{1}{\theta}, \forall x \in (0, \theta)$ does not satisfy the regularity condition

$$\mathbf{E}\left[\frac{\partial}{\partial\theta}\log f_X(X; \theta)\right] = 0,$$

or equivalently,

$$\begin{aligned}
\int \frac{\partial}{\partial \theta} \log f_X(x;\theta) \cdot f_X(x;\theta) dx &= \int \frac{\frac{\partial}{\partial \theta} f_X(x;\theta)}{f_X(x;\theta)} \cdot f_X(x;\theta) dx \\
&= \int \frac{\partial}{\partial \theta} f_X(x;\theta) dx \\
&= \frac{\partial}{\partial \theta} \int f_X(x;\theta) dx \\
&= \frac{\partial}{\partial \theta}(1) = 0,
\end{aligned}$$

i.e., the integration and differentiation are interchangeable.

Since $\mathbf{E}\left[\frac{\partial}{\partial \theta} \log f_X(X;\theta)\right] \neq 0$, the Cramer-Rao lower bound is not applicable. It is possible that there is an unbiased estimator with variance less than the right-hand side of CRLB.