

BIOSTAT 602 Biostatistical Inference

Homework 05

Ashton Baker

Thursday February 16, 2017

1. Let X be a discrete random variable with pmf $f_X(x|\theta)$ where $\theta \in \{1, 2, 3\}$ and $x \in \{1, 2, 3, 4, 5, 6\}$.

$$f_X(x|\theta) = \begin{cases} x/21, & \theta = 1 \\ 1/6, & \theta = 2 \\ I(x = 3), & \theta = 3 \end{cases}$$

Find a maximum-likelihood estimator of θ .

Solution. Given an i.i.d. sample X_1, \dots, X_n from this distribution, the likelihood \mathcal{L} of the sample is

$$\begin{aligned} \mathcal{L}(\theta|\mathbf{x}) &= \prod_{i=1}^n f_X(x_i|\theta) \\ &= \begin{cases} \left(\frac{1}{21}\right)^n \prod_{i=1}^n x_i, & \theta = 1 \\ (1/6)^n, & \theta = 2 \\ I(x_1, \dots, x_n = 3), & \theta = 3 \end{cases} \end{aligned}$$

So, the MLE is more of a procedure or algorithm: Determine

$$\max \left\{ 21^{-n} \prod_{i=1}^n x_i, (1/6)^n, I(x_1, \dots, x_n = 3) \right\}$$

and choose $\hat{\theta}$ to be 1, 2, or 3, correspondingly.

2. Let X_1, \dots, X_n be i.i.d. random variables from $U(0, \theta)$ with pdf

$$f_X(x|\theta) = \frac{1}{\theta}, \quad 0 \leq x \leq \theta$$

- (a) Find a method of moments estimator of θ using the lowest-order moments as possible.

Solution. The first moment μ_1 can be calculated

$$\mu_1 = E[X] = \int_0^\theta x \frac{1}{\theta} dx = \frac{\theta}{2}$$

So a method of moments estimator of θ can be determined as the solution to $\hat{\mu}_1 = \frac{\hat{\theta}}{2}$, which is of course

$$\hat{\theta} = 2\hat{\mu}_1 = \frac{2}{n} \sum_{i=1}^n X_i$$

- (b) Calculate the mean and variance of the method of moments estimator.

Solution. The mean of $\hat{\theta}$ is

$$\begin{aligned} E[\hat{\theta}] &= E[2\hat{\mu}_1] \\ &= \frac{2}{n} E\left[\sum_{i=1}^n X_i\right] \\ &= \frac{2}{n} \sum_{i=1}^n E[X_i] \\ &= \frac{2}{n} \left(n \frac{\theta}{2}\right) \\ &= \theta \end{aligned}$$

The variance of the given uniform distribution is $\sigma^2 = \frac{1}{12}\theta^2$, and the variance of the sample mean is given by $\text{Var}(\bar{X}) = \sigma^2/n$, so

$$\text{Var}(\hat{\theta}) = \text{Var}(2\hat{\mu}_1) = 4\text{Var}(\hat{\mu}_1) = \frac{4\sigma^2}{n}$$

- (c) Compare the MLE $\hat{\theta}_{MLE} = X_{(n)}$ with the estimator from (a) in terms of bias and variance. Which estimator is better? Justify your answer.

3. Let X_1, \dots, X_n be a random sample from a double exponential distribution with pdf

$$f_X(x|\mu, \sigma^2) = \frac{1}{2\sigma} \exp[-|x - \mu|/\sigma], \quad x \in \mathbb{R}, \quad \mu \in \mathbb{R}, \quad \sigma > 0$$

Find MLEs of μ and σ . Show all steps. You may use the fact that for a set of real numbers x_1, x_2, \dots, x_n the quantity $\frac{1}{n} \sum_{i=1}^n |x_i - a|$ is minimized when $a = \text{median}(x_1, x_2, \dots, x_n)$

Solution. For the random sample X_1, \dots, X_n , the likelihood \mathcal{L} of the sample is

$$\begin{aligned} \mathcal{L}(\mu, \sigma^2|\mathbf{x}) &= \prod_{i=1}^n f(x_i|\mu, \sigma^2) \\ &= \prod_{i=1}^n \frac{1}{2\sigma} \exp[-|x_i - \mu|/\sigma] \\ &= (2\sigma)^{-n} \exp \left[-\frac{1}{\sigma} \sum_{i=1}^n |x_i - \mu| \right] \end{aligned}$$

and the log-likelihood ℓ is

$$\begin{aligned} \ell(\mu, \sigma^2|\mathbf{x}) &= \log \mathcal{L}(\mu, \sigma^2|\mathbf{x}) \\ &= -n \log(2\sigma) - \frac{1}{\sigma} \sum_{i=1}^n |x_i - \mu| \end{aligned}$$

If σ is constant, then it's pretty clear, given the hint, that ℓ is maximized when $\mu = \text{median}(x_1, x_2, \dots, x_n)$. So that's the MLE for μ . For σ , we need to take the derivative:

$$\frac{\partial \ell}{\partial \sigma} = -n(2\sigma)^{-n} + \sigma^{-2} \sum_{i=1}^n |x_i - \mu|$$

and set it equal to zero:

$$\begin{aligned} 0 &= -n(2\sigma)^{-n} + \sigma^{-2} \sum_{i=1}^n |x_i - \mu| \\ n(2\sigma)^{-n} &= \sigma^{-2} \sum_{i=1}^n |x_i - \mu| \\ \sigma &= \frac{1}{2} \left(\frac{1}{n} \sum_{i=1}^n |x_i - \mu| \right)^{1/(2-n)} \end{aligned}$$

4. Let X_1, \dots, X_n be an i.i.d. random sample from a distribution with the pdf

$$f_X(x|\theta) = \frac{x}{\theta} \exp\left(-\frac{x^2}{2\theta}\right), \quad x > 0 \quad \theta > 0$$

- Find a complete sufficient statistic for θ .
- Find the Cramer-Rao lower bound for the variance of any unbiased estimator of θ .
- Can you find a simple function (constant multiple) of the complete sufficient statistic in part (a) which is unbiased?
- Does the estimator in part (c) attain the Cramer-Rao lower bound obtained in part (b)?

5. Let X_1, \dots, X_n be an i.i.d. random sample from pdf

$$f_X(x|\theta) = \theta x^{\theta-1} I(0 < x < 1)$$

- (a) When $\theta \geq 1$, find the maximum likelihood estimator for θ .

Solution. The likelihood is

$$\begin{aligned}\mathcal{L}(\theta|\mathbf{x}) &= \prod_{i=1}^n \theta x_i^{\theta-1} I(0 < x_i < 1) \\ &= \theta^n \left(\prod_{i=1}^n x_i \right)^{\theta-1} I(0 < x_1, \dots, x_n < 1)\end{aligned}$$

and the log likelihood is

$$\ell(\theta|\mathbf{x}) = n \log(\theta) + (\theta - 1) \sum_{i=1}^n \log(x_i) + \log(I(0 < x_1, \dots, x_n < 1))$$

The whole expression is meaningless if x_1, \dots, x_n are not between 0 and 1, so we will assume this is the case, and simply consider the final term to be zero. Then the MLE for θ can be found by solving $\partial \ell / \partial \theta = 0$. Note that

$$\frac{\partial \ell}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^n \log(x_i)$$

Solving for zero gives

$$\begin{aligned}0 &= \frac{n}{\theta} + \sum_{i=1}^n \log(x_i) \\ -\frac{n}{\theta} &= \sum_{i=1}^n \log(x_i) \\ \theta &= \frac{-n}{\sum_{i=1}^n \log(x_i)},\end{aligned}$$

so $\hat{\theta} = -n / \sum_{i=1}^n \log(x_i)$ is the MLE for θ .

- (b) When $\theta > 1$, find the maximum likelihood estimator for $\tau(\theta) = 1/\theta$.

Solution. First, we reformulate $\ell(\theta|\mathbf{x})$ in terms of τ . This gives

$$\ell(\tau|\mathbf{x}) = -n \log(\tau) + (\tau^{-1} - 1) \sum_{i=1}^n \log(x_i)$$

So the derivative is

$$\frac{\partial \ell}{\partial \tau} = \frac{-n}{\tau} - \frac{1}{\tau^2} \sum_{i=1}^n \log(x_i)$$

and solving for $\partial \ell / \partial \tau$ gives

$$\begin{aligned}0 &= -\frac{n}{\tau} - \frac{1}{\tau^2} \sum_{i=1}^n \log(x_i) \\ 0 &= -n\tau - \sum_{i=1}^n \log(x_i) \\ \tau &= -\frac{1}{n} \sum_{i=1}^n \log(x_i)\end{aligned}$$

So $\hat{\tau} = -\frac{1}{n} \sum_{i=1}^n \log(x_i)$ is the MLE for τ .

- (c) When $\theta > 0$, find the Cramer-Rao lower bound of the variance of unbiased estimators for $\tau(\theta) = 1/\theta$. Does the MLE in (b) attain the bound?