PS2

Samuel Olson

Problem 1

7.11, Casella & Berger

Let $X_1, \ldots, X_n$ be iid with pdf

$$f(x|\theta) = \theta x^{\theta-1}, \quad 0 \leq x \leq 1, \quad 0 < \theta < \infty.$$

Hint: In part (a), you can assume each observation lies in $X_i \in (0,1)$ for finding the MLE (since there is zero probability of "some $X_i = 0$ or 1 for $i = 1, \ldots, n$"). To find the variance in part (a), you should be able to show that $Y_i = -\log(X_i)$ has an exponential distribution with scale parameter $\beta = 1/\theta > 0$ so that

$$W = \sum_{i=1}^{n} Y_i$$

has a gamma ($\alpha = n, \beta$) distribution; then, you can compute the variance by finding moments $E_\theta(W^{-1})$ and $E_\theta(W^{-2})$.

a)

Find the MLE of θ, and show that its variance $\to 0$ as $n \to \infty$.

b)

Find the method of moments estimator of θ.

Problem 2

7.12(a), Casella & Berger

Let $X_1, \ldots, X_n$ be a random sample from a population with pmf

$$P_\theta(X = x) = \theta^x (1-\theta)^{1-x}, \quad x = 0 \text{ or } 1, \quad 0 \le \theta \le \frac{1}{2}.$$

Hint: Note that the parameter space is $\Theta \equiv [0, 1/2]$. In maximizing the likelihood, it might be clearest to consider three data cases:

1. $\sum_{i=1}^n X_i = 0$;
2. $\sum_{i=1}^n X_i = n$; or
3. $0 < \sum_{i=1}^n X_i < n$.

In the last case, the derivative of log-likelihood $L(\theta)$ indicates that $L(\theta)$ is increasing on $(0, \bar{X}_n)$ and decreasing on $(\bar{X}_n, 1)$.

a)

Find the method of moments estimator and MLE of θ.

Problem 3

7.14, Casella & Berger

Let X and Y be independent exponential random variables, with

$$f(x|\lambda) = \frac{1}{\lambda} e^{-x/\lambda}, \quad x > 0, \quad f(y|\mu) = \frac{1}{\mu} e^{-y/\mu}, \quad y > 0.$$

We observe Z and W with

$$Z = \min(X, Y) \quad \text{and} \quad W = \begin{cases} 1 & \text{if } Z = X \\ 0 & \text{if } Z = Y. \end{cases}$$

In Exercise 4.26, the joint distribution of Z and W was obtained. Now assume that $(Z_i, W_i), i = 1, \dots, n$, are n iid observations. Find the MLEs of λ and μ.

Hint: You may use that the joint density of (Z, W) is

$$f(z, w|\lambda, \mu) = \frac{dF(z, w)}{dz} = \begin{cases} \mu^{-1} e^{-z(\lambda + \mu^{-1})}, & z > 0, w = 0 \\ \lambda^{-1} e^{-z(\lambda + \mu^{-1})}, & z > 0, w = 1 \end{cases}$$

where

$$F(z, w|\lambda, \mu) = P(Z \leq z, W = w|\lambda, \mu).$$

Then, based on a random sample $(Z_i, W_i), i = 1, \dots, n$ of pairs, this problem involves using calculus with two variables to find the MLE.

Problem 4

7.49, Casella & Berger

Let $X_1, \dots, X_n$ be iid exponential(λ).

a)

Find an unbiased estimator of λ based only on $Y = \min\{X_1, \dots, X_n\}$.

b)

Find a better estimator than the one in part (a). Prove that it is better.

c)

The following data are high-stress failure times (in hours) of Kevlar/epoxy spherical vessels used in a sustained pressure environment on the space shuttle:

$$50.1, \quad 70.1, \quad 137.0, \quad 166.9, \quad 170.5, \quad 152.8, \quad 80.5, \quad 123.5, \quad 112.6, \quad 148.5, \quad 160.0, \quad 125.4.$$

Failure times are often modeled with the exponential distribution. Estimate the mean failure time using the estimators from parts (a) and (b).

Problem 5

Suppose someone collects a random sample $X_1, X_2, \ldots, X_n$ from an exponential $\beta = 1/\theta$ distribution with pdf

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

and a parameter $\theta > 0$. However, due to a recording mistake, only truncated integer data $Y_1, Y_2, \ldots, Y_n$ are available for analysis, where Y_i represents the integer part of X_i after dropping all digits after the decimal place in X_i's representation. (For example, if $x_1 = 4.9854$ in reality, we would have only $y_1 = 4$ available.) Then, $Y_1, \ldots, Y_n$ represent a random sample of iid (discrete) random variables with pmf

$$f(y|\theta) = P_\theta(Y_i = y) = e^{-\theta y} - e^{-\theta(1+y)}, \quad y = 0, 1, 2, 3, \ldots.$$

a)

Show that the likelihood equals

$$L(\theta) = \left[e^{-\theta \bar{Y}_n} (1 - e^{-\theta}) \right]^n,$$

where $\bar{Y}_n$ is the sample average.

b)

If $Y_n = \sum_{i=1}^n Y_i/n = 0$, show that an MLE for θ does not exist on the parameter space $(0, \infty)$.

(Recall: Y_i is discrete and this corresponds to a pathological MLE case mentioned in class: $Y_1 = \cdots = Y_n = 0$. This event can happen but typically with small probability for large n.)

c)

If $0 < \bar{Y}_n$, show that the MLE $\hat{\theta}$ is

$$\hat{\theta} = \log(\bar{Y}_n^{-1} + 1).$$