Solutions Assignment 1

6

6.3.14(a)

First, note that $Y_n = \sum_{i=1}^n X_i^2/n$ has asymptotically the normal distribution with mean σ^2 and variance $2\sigma^4/n$. Here, we have used the fact that $E(X_i^2) = \sigma^2$ and $E(X_i^4) = \cancel{2}\,3\sigma^4$.

(a) Let $g(x) = 1/x$. Then $g'(x) = -1/x^2$. So, the asymptotic distribution of $g(Y_n)$ is the normal distribution with mean $1/\sigma^2$ and variance $(2\sigma^4/n)/\sigma^8 = 2/[n\sigma^4]$.

Details: By CLT, $\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_i^2 - E[X_1^2]\right) \xrightarrow{d} N(0, Var(X_1^2))$ ✓

Because $E[X_1^2] = \sigma^2$ and $Var(X_1^2) = E[X_1^4] - E[X_1^2]^2 = 3\sigma^4 - (\sigma^2)^2 = 2\sigma^4$ ✓

$\sqrt{n}\left(\frac{1}{n}\sum X_i^2 - \sigma^2\right) \xrightarrow{d} N(0, 2\sigma^4)$

δ-method with $g(x) = 1/x$ ✓ $\quad \sqrt{n}\,\dfrac{\left(\frac{1}{n}\sum X_i^2\right)^{-1} - \sigma^{-2}}{-1/(\sigma^2)^2} \xrightarrow{d} N(0, 2\sigma^4)$ ✓

$\Longrightarrow \left(\frac{1}{n}\sum X_i^2\right)^{-1} - \sigma^{-2} \overset{approx}{\sim} N\left(0, \frac{2\sigma^4}{(\sigma^2)^4 n}\right) = N\left(0, \frac{2}{\sigma^4 n}\right)$

$\Longrightarrow \left(\frac{1}{n}\sum_{i=1}^n X_i\right)^{-1} \overset{approx}{\sim} N\left(\sigma^{-2}, \frac{2}{\sigma^4 n}\right)$ ✓ ✓

NOT GRADED

6.3.15

(a) Clearly, $Y_n \le y$ if and only if $X_i \le y$ for $i = 1, \ldots, n$. Hence,

$$\Pr(Y_n \le y) = \Pr(X_1 \le y)^n = \begin{cases} (y/\theta)^n \text{if } 0 < y < \theta, & \\ 0 & \text{if } y \le 0, \\ 1 & \text{if } y \ge \theta. \end{cases}$$

(b) The c.d.f. of Z_n is, for $z < 0$,

$$\Pr(Z_n \le z) = \Pr(Y_n \le \theta + z/n) = (1 + z/[n\theta])^n. \tag{S.6.9}$$

Since $Z_n \le 0$, the c.d.f. is 1 for $z \ge 0$. According to Theorem 5.3.3, the expression in (S.6.9) converges to $\exp(z/\theta)$.

(c) Let $\alpha(y) = y^2$. Then $\alpha'(y) = 2y$. We have $n(Y_n - \theta)$ converging in distribution to the c.d.f. in part (b). The delta method says that, for $\theta > 0$, $n(Y_n^2 - \theta^2)/[2\theta]$ converges in distribution to the same c.d.f.

NOT GRADED

6.5.2

Because of the property of the Poisson distribution described in Theorem 5.4.4, the random variable X can be thought of as the sum of a large number of i.i.d. random variables, each of which has a Poisson distribution. Hence, the central limit theorem (Lindeberg and Lévy) implies the desired result. It can also be shown that the m.g.f. of X converges to the m.g.f. of the standard normal distribution.

7

7.5.6

Let $\theta = \sigma^2$. Then the likelihood function is

$$f_n(\boldsymbol{x} \mid \theta) = \frac{1}{(2\pi\theta)^{n/2}} \exp\left\{-\frac{1}{2\theta}\sum_{i=1}^{n}(x_i - \mu)^2\right\}. \checkmark$$

If we let $L(\theta) = \log\ f_n(\boldsymbol{x} \mid \theta)$, then $L(\theta) = -\frac{n}{2}\log(2\pi\theta) - \frac{1}{2\theta}\sum_{i=1}^{n}(x_i-\mu)^2$. ✓ Thus,

$$\frac{\partial}{\partial\theta}L(\theta) = -\frac{n}{2\theta} + \frac{1}{2\theta^2}\sum_{i=1}^{n}(x_i - \mu)^2. \checkmark$$

✓ (for setting $\frac{\partial}{\partial\theta}L(\theta)=0$)

The maximum of $L(\theta)$ will be attained at a value of θ for which this derivative is equal to 0. In this way, we find that

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2. \checkmark$$

Furthermore,

$$\frac{\partial^2}{\partial\theta^2}L(\theta) = \frac{n}{2\theta^2} - \frac{1}{\theta^3}\sum_{i=1}^{n}(x_i-\mu)^2 \checkmark \overset{\text{at } \theta=\hat{\theta}}{=} \frac{n}{2\hat{\theta}^2} - \frac{n}{\hat{\theta}^2} = -\frac{n}{2\hat{\theta}^2} < 0 \checkmark$$

So we conclude that $\hat{\theta}$ is the MLE.

(for showing and not just claiming that it's < 0)

Alternative for [$\cdot$ -]:

$\prod_{i=1}^{n} f(X_i;\theta_1,\theta_2) = \prod_{i=1}^{n} \frac{1}{\theta_2-\theta_1} I_{[\theta_1,\theta_2]}(X_i)$

$= \frac{1}{(\theta_2-\theta_1)^n} \cdot \prod_{i=1}^{n} I_{[\theta_1,\theta_2]}(X_i)$

$\underset{\theta_1 \le X_i \le \theta_2}{=} \frac{1}{(\theta_2-\theta_1)^n} \cdot \prod_{i=1}^{n} I_{(-\infty,X_i]}(\theta_1)\, I_{[X_i,\infty)}(\theta_2)$

means: $\theta_1 \le X_i$, $\theta_2 \ge X_i$

$= \frac{1}{(\theta_2-\theta_1)^n} I_{(-\infty, \min X_i]}(\theta_1)\, I_{[\max X_i, \infty)}(\theta_2)$

6 MLE **7.5.11**

+ 7 MOM The p.d.f. of each observation can be written as follows:

$$f(x \mid \theta_1,\theta_2) = \begin{cases} \frac{1}{\theta_2-\theta_1} & \text{for } \theta_1 \le x \le \theta_2, \\ 0 & \text{otherwise.} \end{cases}$$

[Therefore, the likelihood function is

$$f_n(\boldsymbol{x} \mid \theta_1,\theta_2) = \frac{1}{(\theta_2-\theta_1)^n}$$

if $\theta_1 \le x_i$ for all $i=1,\dots,n$ (i.e., ~~for~~ $\theta_1 \le \min\{x_1,\dots,x_n\}$) and $\theta_2 \ge x_i$ for all $i=1,\dots,n$ (i.e., $\max\{x_1,\dots,x_n\} \le \theta_2$), and $f_n(\boldsymbol{x} \mid \theta_1,\theta_2) = 0$ otherwise.] (*) Hence, $f_n(\boldsymbol{x} \mid \theta_1,\theta_2)$ will be a maximum when $\theta_2 - \theta_1$ is made as small as possible under these constraints on θ_1 and θ_2. Since the smallest possible value of θ_2 is $\max\{x_1,\dots,x_n\}$ and the largest possible value of θ_1 is $\min\{x_1,\dots,x_n\}$, these values are the M.L.E.'s.

Note that

$$EX_1 = (\theta_1+\theta_2)/2$$

and

$$EX_1^2 = (\theta_1^2+\theta_2^2+\theta_1\theta_2)/3.$$

Setting $\hat\mu_1 = EX_1$ and $\hat\mu_2 = EX_1^2$ and substituting θ_1 in the second equation by $2\hat\mu_1 - \theta_2$ (the first equation), we obtain that

$$(2\hat\mu_1-\theta_2)^2 + \theta_2^2 + (2\hat\mu_1-\theta_2)\theta_2 = 3\hat\mu_2,$$

which is the same as

$$(\theta_2-\hat\mu_1)^2 = 3(\hat\mu_2-\hat\mu_1^2).$$

Since $\theta_2 > EX_1$, we obtain that

$$\hat\theta_2 = \hat\mu_1 + \sqrt{3(\hat\mu_2-\hat\mu_1^2)} = \bar X + \sqrt{\tfrac{3(n-1)}{n}S^2}$$

and

$$\hat\theta_1 = \hat\mu_1 - \sqrt{3(\hat\mu_2-\hat\mu_1^2)} = \bar X - \sqrt{\tfrac{3(n-1)}{n}S^2}.$$

Alternative: Substituting $\theta_1 = 2\hat\mu_1 - \theta_2$ in the second equation, we obtain

$3\hat\mu_2 = (2\hat\mu_1-\theta_2)^2 + \theta_2^2 + (2\hat\mu_1-\theta_2)\theta_2$

$\iff 3\hat\mu_2 = 4\hat\mu_1^2 - 4\hat\mu_1\theta_2 + \cancel{\theta_2^2} + \theta_2^2 + 2\hat\mu_1\theta_2 - \cancel{\theta_2^2}$

$\iff \theta_2^2 - 2\hat\mu_1\theta_2 + 4\hat\mu_1^2 - 3\hat\mu_2 = 0$

$\iff \theta_2 = \frac{2\hat\mu_1 \pm \sqrt{4\hat\mu_1^2 - 4(4\hat\mu_1^2 - 3\hat\mu_2)}}{2}$

$= \hat\mu_1 \pm \sqrt{3(\hat\mu_2 - \hat\mu_1^2)}$

We take the + solution because $\theta_2 \ge \hat\mu_1$

$\Rightarrow \theta_1 = 2\hat\mu_1 - \theta_2 = \hat\mu_1 - \sqrt{3(\hat\mu_2-\hat\mu_1^2)}$

(*) Only 1 point if the "for all i" part is missing and they directly jump to max x_i and min x_i

NOT GRADED

7.5.12

The likelihood function is

$$f_n(\boldsymbol{x} \mid \theta_1, \dots, \theta_k) = \theta_1^{n_1} \cdots \theta_k^{n_k}.$$

If we let $L(\theta_1, \dots, \theta_k) = \log\ f_n(\boldsymbol{x} \mid \theta_1, \dots, \theta_k)$ and let $\theta_k = 1 - \sum_{i=1}^{k-1} \theta_i$, then

$$\frac{\partial L(\theta_1, \dots, \theta_k)}{\partial \theta_i} = \frac{n_i}{\theta_i} - \frac{n_k}{\theta_k} \qquad \text{for } i = 1, \dots, k-1.$$

If each of these derivatives is set equal to 0, we obtain the relations

$$\frac{\theta_1}{n_1} = \frac{\theta_2}{n_2} = \cdots = \frac{\theta_k}{n_k}.$$

If we let $\theta_i = \alpha n_i$ for $i = 1, \dots, k$, then

$$1 = \sum_{i=1}^{k} \theta_i = \alpha \sum_{i=1}^{k} n_i = \alpha n.$$

Hence $\alpha = 1/n$. It follows that $\hat{\theta}_i = n_i/n$ for $i = 1, \dots, k$.

If someone mentions Bernoulli/binomial experiment, also give the point

5

7.6.4

The probability that a given lamp will fail in a period of T hours is $p = 1 - \exp(-\beta T)$, and the probability that exactly x lamps will fail is $\binom{n}{x} p^x (1-p)^{n-x}$. It was shown in Example 7.5.4 (class) that $\hat{p} = x/n$. Since $\beta = -\log(1-p)/T$, it follows that $\hat{\beta} = -\log(1 - x/n)/T$.

by the invariance principle of MLEs

NOT GRADED

7.6.6

The distribution of $Z = (X - \mu)/\sigma$ will be a standard normal distribution. Therefore,

$$0.95 = \Pr(X < \theta) = \Pr\left(Z < \frac{\theta - \mu}{\sigma}\right) = \Phi\left(\frac{\theta - \mu}{\sigma}\right).$$

Hence, from a table of the values of Φ it is found that $(\theta - \mu)/\sigma = 1.645$. Since $\theta = \mu + 1.645\sigma$, it follows that $\hat{\theta} = \hat{\mu} + 1.645\hat{\sigma}$. By example 6.5.4, we have

$$\hat{\mu} = \overline{X}_n \qquad \text{and} \qquad \hat{\sigma} = \left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2\right]^{1/2}.$$

2

7.6.8

✓ ✓

Let $\theta = \Gamma'(\alpha)/\Gamma(\alpha)$. Then $\hat{\theta} = \Gamma'(\hat{\alpha})/\Gamma(\hat{\alpha})$. It follows from Eq. (7.6.5) that $\hat{\theta} = \sum_{i=1}^{n}(\log\ X_i)/n$.

NOT GRADED

7.6.10

The likelihood function is

$$f_n(\boldsymbol{x} \mid \alpha, \beta) = \left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right]^n \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} \left[\prod_{i=1}^{n}(1 - x_i)\right]^{\beta-1}.$$

If we let $L(\alpha, \beta) = \log\ f_n(\boldsymbol{x} \mid \alpha, \beta)$, then

$$\begin{aligned} L(\alpha, \beta) &= n\ \log\ \Gamma(\alpha+\beta) - n\ \log\ \Gamma(\alpha) - n\ \log\ \Gamma(\beta) \\ &\quad +(\alpha - 1)\sum_{i=1}^{n}\log\ x_i + (\beta - 1)\sum_{i=1}^{n}\log(1 - x_i). \end{aligned}$$

Hence,

$$\frac{\partial L(\alpha, \beta)}{\partial \alpha} = n\frac{\Gamma'(\alpha+\beta)}{\Gamma(\alpha+\beta)} - n\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_{i=1}^{n}\log\ x_i$$

and

$$\frac{\partial L(\alpha,\beta)}{\partial\beta} = n\frac{\Gamma'(\alpha+\beta)}{\Gamma(\alpha+\beta)} - n\frac{\Gamma'(\beta)}{\Gamma(\beta)} + \sum_{i=1}^{n}\log(1-x_i).$$

The estimates $\hat{\alpha}$ and $\hat{\beta}$ must satisfy the equations $\partial L(\alpha,\beta)/\partial\alpha = 0$ and $\partial L(\alpha,\beta)/\partial\beta = 0$. Therefore, $\hat{\alpha}$ and $\hat{\beta}$ must also satisfy the equation $\partial L(\alpha,\beta)/\partial\alpha = \partial L(\alpha,\beta)/\partial\beta$. This equation reduces to the one given in the exercise.

NOT GRADED

7.6.12

We know that $\hat{\beta} = 1/\overline{X}_n$. Also, since the mean of the exponential distribution is $\mu = 1/\beta$, it follows from the law of large numbers that $\overline{X}_n \xrightarrow{p} 1/\beta$. Hence, $\hat{\beta} \xrightarrow{p} \beta$.

3

7.6.22

✓ ✓

The mean of X_i is $\theta/2$, so the method of moments estimator is $2\overline{X}_n$. The M.L.E. is the maximum of the X_i values. ✓

Total: 36 points