

## MODEL ANSWERS OF QUIZ I

1. (3 points) Let  $X_1, X_2, X_3, X_4 \stackrel{i.i.d.}{\sim} \text{Bernoulli}(p)$ , where  $0 < p < 1$  is unknown parameter. Denote

$$U = X_1(X_3 + X_4) + X_2.$$

Show that  $U$  is not a sufficient statistic for  $p$ .

**Solution:** Let us calculate  $P((X_1, X_2, X_3, X_4) = (0, 0, 0, 0) | U = 0)$ . Note that

$$\{U = 0\} = \{(X_1, X_2, X_3, X_4) \in \{(0, 1, 1, 0), (0, 0, 1, 0), (0, 1, 0, 0), (1, 0, 0, 0), (0, 0, 0, 0)\}\}.$$

Therefore,

$$P((X_1, X_2, X_3, X_4) = (0, 0, 0, 0) | U = 0) = \frac{(1-p)^4}{p^2(1-p)^2 + 3p(1-p)^3 + (1-p)^4}.$$

Hence, for  $p = \frac{1}{2}$ ,  $P((X_1, X_2, X_3, X_4) = (0, 0, 0, 0) | U = 0) = \frac{1}{5}$ .

For  $p = \frac{1}{3}$ ,  $P((X_1, X_2, X_3, X_4) = (0, 0, 0, 0) | U = 0) = \frac{4}{11}$ .

Therefore,  $P((X_1, X_2, X_3, X_4) = (0, 0, 0, 0) | U = 0)$  depends on  $p$ , and hence  $U$  is not a sufficient statistics for  $p$ . [3 points]

2. Let  $X_1, X_2, X_3, X_4, X_5 \stackrel{i.i.d.}{\sim} N(2\theta, 25\theta^2)$ , where  $\theta > 0$  is an unknown parameter.

- (a) (3 points) Derive the minimal sufficient statistics.

**Solution:** The joint probability density function of  $X_1, X_2, \dots, X_5$  is

$$f(x_1, x_2, x_3, x_4, x_5) = \left(\frac{1}{5\theta\sqrt{2\pi}}\right)^5 \exp\left[-\frac{1}{50\theta^2} \sum_{i=1}^5 (x_i - 2\theta)^2\right],$$

for all  $(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5$ . [1 point]

Therefore, for two sample points  $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5)$  and  $\mathbf{y} = (y_1, y_2, y_3, y_4, y_5)$ , the ratio of JPDF is

$$h(\mathbf{x}, \mathbf{y}, \theta) = \exp\left[-\frac{1}{50\theta^2} \left\{\left(\sum_{i=1}^5 x_i^2 - \sum_{i=1}^5 y_i^2\right) - 20\theta(\bar{x} - \bar{y})\right\}\right],$$

which does not involve  $\theta$  if and only if  $\sum_{i=1}^5 x_i^2 = \sum_{i=1}^5 y_i^2$  and  $\bar{x} = \bar{y}$ . Thus,  $(\bar{X}, \sum_{i=1}^5 X_i^2)$  is a minimal sufficient statistic. [2 points]

- (b) (4 points) Is the minimal sufficient statistic, that you obtained in part (a), complete? Justify your answer.

**Solution:** Note that  $E(\bar{X}^2) = \text{Var}(\bar{X}) + (E(\bar{X}))^2 = 9\theta^2 \implies E\left(\frac{1}{9}\bar{X}^2\right) = \theta^2$ .

Moreover,  $E\left(\frac{1}{4}\sum_{i=1}^4 (X_i - \bar{X})^2\right) = 25\theta^2 \implies E\left(\frac{25}{4}\sum_{i=1}^4 (X_i - \bar{X})^2\right) = \theta^2$ .

Now, note that  $\frac{1}{9}\bar{X}^2 - \frac{25}{4}\sum_{i=1}^5 (X_i - \bar{X})^2$  is a function of minimal sufficient statistic obtained in the part (a) and  $E\left[\frac{1}{9}\bar{X}^2 - \frac{25}{4}\sum_{i=1}^5 (X_i - \bar{X})^2\right] = 0$ . [2 points]

As  $\bar{X}$  and  $\sum_{i=1}^5 (X_i - \bar{X})^2$  are two independent continuous random variables,  $\frac{1}{9}\bar{X}^2 - \frac{25}{4}\sum_{i=1}^5 (X_i - \bar{X})^2$  is also continuous random variable. Therefore,

$$P\left(\frac{1}{9}\bar{X}^2 - \frac{25}{4}\sum_{i=1}^5 (X_i - \bar{X})^2 = 0\right) = 0$$

and hence the minimal sufficient statistic as obtained in part (a) is not complete. [2 points]

3. (5 points) Let  $X_1, X_2, \dots, X_n$  be a random sample of size  $n (\geq 2)$  from a population having probability density function

$$f(x, \theta) = \begin{cases} \frac{1}{\theta}e^{-\frac{x}{\theta}} & \text{if } x > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where  $\theta > 0$  is unknown. Find the conditional expectation of  $X_{(1)} = \min\{X_1, X_2, \dots, X_n\}$  given  $T = \sum_{i=1}^n X_i$ . You may assume that the conditional expectation exists.

**Solution:** Note that the concerned distribution belongs to a one-parameter exponential family

$$f(x, \theta) = a(\theta)g(x) \exp[b(\theta)R(x)],$$

where  $a(\theta) = \theta^{-1}$ ,  $b(\theta) = -\frac{1}{\theta}$ ,  $R(x) = x$ , and

$$g(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Therefore,  $T = \sum_{i=1}^n X_i$  is complete and sufficient statistic. [1 point]

As the concerned distribution belongs to scale family of distributions, the distribution of  $\frac{X_i}{\theta}$  does not depend on  $\theta$ . Therefore,

$$\frac{X_{(1)}}{T} = \frac{\min\left\{\frac{X_1}{\theta}, \frac{X_2}{\theta}, \dots, \frac{X_n}{\theta}\right\}}{\frac{X_1}{\theta} + \frac{X_2}{\theta} + \dots + \frac{X_n}{\theta}}$$

is ancillary statistic. [1 point]

Hence, using Basu's theorem,  $\frac{X_{(1)}}{T}$  and  $T$  are independent. [1 point]

Now,

$$E(X_{(1)}|T) = E\left(\frac{X_{(1)}}{T} \cdot T \mid T\right) = TE\left(\frac{X_{(1)}}{T}\right).$$

Also,

$$E(X_{(1)}) = E\left(\frac{X_{(1)}}{T} \cdot T\right) = E\left(\frac{X_{(1)}}{T}\right)E(T) \implies E\left(\frac{X_{(1)}}{T}\right) = \frac{E(X_{(1)})}{E(T)} = \frac{1}{n^2},$$

as  $E(T) = \sum_{i=1}^n E(X_i) = n\theta$ , and  $E(X_{(1)}) = \frac{\theta}{n}$ . [2 points]

Therefore,

$$E(X_{(1)}|T) = \frac{T}{n^2}.$$