Statistical Inference Assignment 4

Junhao Yuan (20307130129)

October 21, 2022

PROBLEM 1.

Let $X_1, \dots, X_n \overset{iid}{\sim} \mathcal{N}(\theta, a\theta^2)$, where a is a known positive constant and $\theta > 0$.

(a) Show that the parameter space does not contain a two-dimensional open set.

(b) Show that the statistics $T = (\bar{X}, S^2)$ is a sufficient statistics for θ, but the family of distributions is not complete.

SOLUTION.

(a) The parameter space $\Theta = \{(\theta, a\theta^2) : \theta > 0, a > 0\}$ is a parabola in a two-dimensional space, which does not contains a two-dimensional open set.

(b) The pdf of the sample is

$$f(\mathbf{X}, \theta) = (2a\pi\theta^2)^{-n/2} \exp\left(-\frac{nS^2 + n(\bar{x} - \theta)^2}{2a\theta^2}\right).$$

Hence, by the factorization theorem, $T = (\bar{X}, S^2)$ is a sufficient statistics for θ.

To prove that the family is not complete, we just need to find a function $g(\bar{X}, S^2)$ satisfies $\mathbb{E}(g(\bar{X}, S^2)) = 0$, but $g(\bar{X}, S^2)$ don't have to be zero identically. We have already known that $\frac{nS^2}{a\theta^2} \sim \chi^2_{n-1}$. If we define $Y = \frac{nS^2}{a\theta^2}$, then we have

$$\begin{aligned} \mathbb{E}(\sqrt{Y}) &= \int_0^\infty y^{\frac{1}{2}} \frac{y^{\frac{n-1}{2}-1} e^{-\frac{y}{2}}}{2^{\frac{n-1}{2}} \Gamma(\frac{n-1}{2})} dx \\ &= \frac{\sqrt{2}\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})}. \end{aligned}$$

Hence,

$$\mathbb{E}(S) = \sqrt{\frac{a}{n}} \theta \mathbb{E}(\sqrt{Y}) = \sqrt{\frac{2a}{n}} \frac{\Gamma(\frac{n}{2})\theta}{\Gamma(\frac{n-1}{2})}.$$

Therefore, if we define

$$g(\bar{X}, S^2) = \bar{X} - \sqrt{\frac{n}{2a}} \frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n}{2})} S,$$

then we have $\mathbb{E}(g(\bar{X}, S^2)) = 0$, but $g(\bar{X}, S^2)$ doesn't have to be zero identically, which means this family of distribution is not complete.

PROBLEM 2.

Let $X_1, \ldots, X_n$ be a random sample from the following population:

$$f(x, \theta) = \theta x^{\theta - 1}, \qquad 0 < x < 1, \qquad \theta > 0. \tag{1}$$

(a) Is $\sum_{i=1}^{n} X_i$ sufficient for θ?

(b) Find a complete sufficient statistics for θ.

SOLUTION.

(a) The pdf of the sample is

$$\begin{aligned} f(\mathbf{X}|\theta) &= \prod_{i=1}^{n} \theta x_i^{\theta-1} \\ &= \theta^n \cdot \exp\left(\theta\left(\sum_{i=1}^{n} \ln(x_i)\right)\right)\left(\prod_{i=1}^{n} x_i\right)^{-1}, \end{aligned}$$

which belongs to the exponential family and $n \geq 1$. Therefore,

$$T(\mathbf{X}) = \sum_{i=1}^{n} \ln(X_i),$$

is a sufficient statistics for θ. Now we prove that $T(\mathbf{X})$ is also the minimal sufficient statistics. Since

$$\begin{aligned} \frac{f(\mathbf{X}|\theta)}{f(\mathbf{Y}|\theta)} &= \left(\frac{\prod_{i=1}^{n} x_i}{\prod_{j=1}^{n} y_j}\right)^{\theta-1} \\ &= e^{(\theta-1)(T(\mathbf{X})-T(\mathbf{Y}))} \end{aligned}$$

is a constant function of θ if and only if $T(\mathbf{X}) = T(\mathbf{Y})$, $T(\mathbf{X})$ is the minimal sufficient statistics for θ.

However, $T(\mathbf{X}) = T(\mathbf{Y})$ does not implies that $\sum_{i=1} x_i = \sum_{j=1}^{n} y_j$, which means $\sum_{i=1}^{n} x_i$ is not sufficient.

(b) Since the natural parameter space $\Theta = (0, \infty)$ contains an open set, $T(\mathbf{X})$ is also a complete statistics for θ.

PROBLEM 3.

Suppose $X_1, \ldots, X_n$ are independently sampled from the following pmf

$$P(X = k) = -\frac{1}{\ln(1-p)} \frac{p^k}{k}, \qquad 0 < p < 1, k = 1, 2, \ldots \tag{2}$$

Use the method of moment to find an estimator for p.

SOLUTION.

Since

$$\mathbb{E}(X) = \sum_{i=1}^{\infty} -\frac{p^k}{\ln(1-p)} = -\frac{1}{\ln(1-p)}\frac{p}{1-p},$$
$$\mathbb{E}(X^2) = \sum_{i=1}^{\infty} -\frac{kp^k}{\ln(1-p)} = -\frac{1}{\ln(1-p)}\frac{p}{(1-p)^2},$$

we have

$$p = 1 - \frac{\mathbb{E}(X)}{\mathbb{E}(X^2)}.$$

Thus, by the method of moment, we have

$$\hat{p} = 1 - \frac{m_1}{m_2},$$

where

$$m_1 = \frac{\sum_{i=1}^n X_i}{n}, \qquad m_2 = \frac{\sum_{i=1}^n X_i^2}{n}.$$

Problem 4.

Suppose $X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$. Find a method of moment estimator for $P(X > 1)$.

Solution.

Let $Y = \frac{X-\mu}{\sigma}$, then $P(X > 1) = P(Y > \frac{1-\mu}{\sigma}) = 1 - \Phi\left(\frac{1-\mu}{\sigma}\right)$, where $\Phi(x)$ is the cdf of the standard normal distribution. By the method of moment, we have $\hat{\mu} = \bar{X}$ and $\hat{\sigma}^2 = S^2 = \sum_{i=1}^n (X_i - \bar{X})^2/n$. Therefore, a method of moment estimator of $P(X > 1)$ is

$$1 - \Phi\left(\frac{1-\bar{X}}{S}\right).$$