Indian Institute of Technology Guwahati
Statistical Inference (MA682)
Problem Set 02

1. Let $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} Geo(p)$ with common PMF

$$f(x;\, p) = \begin{cases} p(1-p)^x & \text{if } x - 0,\, 1,\, 2,\, \ldots \\ 0 & \text{otherwise,} \end{cases}$$

where $p \in (0,\, 1)$ is an unknown parameter. Find a sufficient statistic for p.

2. Let $X_1, X_2, \ldots, X_m \overset{i.i.d.}{\sim} Poi(\lambda)$, $Y_1, Y_2, \ldots, Y_n \overset{i.i.d.}{\sim} Poi(2\lambda)$, where $\lambda > 0$ is unknown parameter. Also, assume that X's and Y's are independent. Show that $\sum_{i=1}^m X_i + \sum_{i=1}^n Y_i$ is a sufficient statistic for λ.

3. Let $X_1, X_2, X_3, X_4 \overset{i.i.d.}{\sim} Bernoulli(p)$, 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.

4. Suppose that $X_1, X_2, \ldots, X_n$ are *i.i.d.* RVs with common PDF

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

where $\mu \in \mathbb{R}$ and $\sigma > 0$. Show that

(a) $X_{(1)} = \min\{X_1, X_2, \ldots, X_n\}$ is sufficient for μ if σ is known.
(b) $\frac{1}{n}\sum_{i=1}^n (X_i - \mu)$ is sufficient statistic for σ if μ is known.
(c) $\left(X_{(1)},\, \sum_{i=1}^n \left(X_i - X_{(1)}\right)\right)$ is sufficient statistic for $(\mu,\, \sigma)$ if both parameters are unknown.

5. Let $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} N(\theta,\, \theta)$, where $\theta > 0$ is unknown parameter. Derive a sufficient statistic for θ.

6. Let $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} U(-\theta,\, \theta)$, where $\theta > 0$ is unknown parameter. Derive a sufficient statistic for θ.

7. Let $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} Bernoilli(p)$, where $p \in (0,\, 1)$ is an unknown parameter. Evaluate $\mathcal{I}_{\boldsymbol{X}}(p)$, $\mathcal{I}_{\overline{X}}(p)$, and compare. Can we use these Fisher information to claim the sufficiency of $\overline{X}$?

8. Let $X_1, X_2, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu,\, \sigma^2)$, where $\sigma > 0$ is unknown, but $\mu \in \mathbb{R}$ is known parameter. Assume that $n \geq 2$. Let

$$U^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^n \left(X_i - \overline{X}\right)^2.$$

Evaluate $\mathcal{I}_{U^2}(\sigma^2)$ and $\mathcal{I}_{S^2}(\sigma^2)$. Show that $\mathcal{I}_{U^2}(\sigma^2) > \mathcal{I}_{S^2}(\sigma^2)$.

9. Suppose that $X_1, X_2, \ldots, X_n$ are *i.i.d.* RVs with common Rayleigh PDF

$$f(x, \sigma) = \begin{cases} \frac{2}{\sigma} x e^{-\frac{x^2}{\sigma}} & \text{if } x > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where $\sigma > 0$ is unknown parameter. Denote a statistic $T = \sum_{i=1}^{n} X_i^2$. Evaluate $\mathcal{I}_{\boldsymbol{X}}(\sigma)$ and $\mathcal{I}_T(\sigma)$. Are they same? If so, what conclusion can one draw from this?