Problem Set 2

Data Analysis and Interpretation (EE 223)

Instructor: Prof. Prasanna Chaporkar
EE Department, IIT Bombay

1. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with probability density function $f_{\vec{\theta}}(\cdot)$. Find the sufficient statistic in the following cases:

   (a) $f_\theta(x) = e^{\theta - x}$ for $x > \theta$ and 0 otherwise.

   (b) $f_\theta(x) = 1/(2\theta)$ for $x \in [1-\theta, 1+\theta]$ and 0 otherwise.

   (c) $f_\theta(x) = 1/\theta$ for $x \in [0, \theta]$ and 0 otherwise.

   (d) $f_\theta(x) = c(\theta)x^{-\theta}$ for $x \geq 1$ and 0 otherwise. Here, $\theta \geq 2$ and $c(\theta)$ is a normalization constant depending on θ.

2. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with probability density function $f_{\vec{\theta}}(\cdot)$. Define,

$$k(\vec{x}, \vec{y}, \vec{\theta}) = \frac{f_{\vec{\theta}}(\vec{x})}{f_{\vec{\theta}}(\vec{y})}.$$

   A sufficient statistic T is called minimal if the following holds: $k(\vec{x}, \vec{y}, \vec{\theta})$ does not depend on $\vec{\theta}$ if and only if $T(\vec{x}) = T(\vec{y})$. Show that

   (a) if $X_1 \sim G(\mu, 1)$, then $T(\vec{x}) = \sum_{k=1}^n x_k$ is minimal sufficient. Also, $\tilde{T}(\vec{x}) = (\sum_{k=1}^n x_k, \sum_{k=1}^n x_k^2)$ is not minimal.

   (b) Verify that all the sufficient statistics found in Question 1 are minimal.

3. Let $X \sim \text{Poisson}(\lambda)$, the parameter $\lambda \in (0, \infty)$, and $\psi(\lambda) = 1/\lambda$. Find an unbiased estimator for ψ.

4. You need to aid a physicist in estimating the rate at which a radio active material emits gamma particles. It is known that the interval between the two consecutive emissions is an exponential random variable with parameter λ. Moreover, inter-emission periods are independent. You choose to put the radio active material with a photographic plate in a lead container for T time units. At the end of this period, you take out the plate and measure the number of marks on the plate.

   (a) Show that only noting the number of marks on the photographic plate is sufficient to estimate the rate

   (b) Give an unbiased estimator for λ

5. You are tasked with approximating the number of tigers in a tiger reserve. You install sensors near a water body that can uniquely identify a tiger that comes close to the water body. Information from the locals allows you to believe that each tiger visits the water body with probability 0.1 independent of other tigers.

   (a) Show that counting the number of unique tigers that visited the water body is sufficient to estimate the number of tigers

   (b) Provide an unbiased estimator

6. Consider random variables X and Y with joint probability density function $f_{XY}(x, y)$, and marginals $f_X(x)$ and $f_Y(y)$. Let $Z = E[X|Y]$ be a random variable from $(\Re, \mathcal{B}(\Re))$ to $(\Re, \mathcal{B}(\Re))$ defined as follows: for $y \in \Re$, $Z(y) = E[X|Y = y]$. Argue that

   (a)

$$Z = \frac{1}{f_Y(Y)} \int_{-\infty}^{\infty} x f_{XY}(x, Y) dx$$

   (b) Find $E[Z]$.

7. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with probability density function $f_{\vec{\theta}}(\cdot)$. Define $X = X_1$ and $Y = \sum_{k=1}^{n} X_k$.

   (a) For $n = 3$ and let the density function be exponential with parameter $\lambda > 0$. Find $E[X|Y]$. Do explicit calculations.

   (b) Find $E[X|Y]$ for the general density function.

8. Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with probability density function $f_{\vec{\theta}}(\cdot)$. Let T be a sufficient statistic and let δ be any unbiased estimate of the given function ψ. Show that $E_{\theta}[\delta|T]$ is also an unbiased estimate of ψ.

9. Show that the sample variance is an unbiased estimator for the variance.

10. Find unbiased estimate for $\sigma > 0$, where samples are drawn from $G(\sigma, \sigma)$.