STATS 300A Theory of Statistics
Stanford University, Fall 2015

Problem Set 2

Due: Thursday, October 8, 2015 **Reading:** K 3.4-3.6, 4.1-4.2

Instructions:

- You may appeal to any result proved in class or proved in the course textbooks.
- Any request to find a statistic with certain properties requires proof that those properties are satisfied.

Problem 1. Suppose that $X_1, \dots, X_n$ are i.i.d. Poisson random variables with mean λ and that we aim to estimate $g(\lambda) = \exp(-\lambda) = \mathbb{P}_\lambda [X = 0]$.

(a) Show that $S_1 = \mathbb{I}(X_1 = 0)$ and $S_2 = \frac{1}{n} \sum_{i=1}^n \mathbb{I}(X_i = 0)$ are unbiased estimates of $g(\lambda)$.

(b) Show that $T = \sum_{i=1}^n X_i$ is sufficient.

(c) Compute the Rao-Blackwellized estimators $S_i^* = \mathbb{E}[S_i \mid T]$ for $i \in \{1, 2\}$. How does your answer relate to completeness?

Problem 2. An exponential random variable with failure rate parameter $\lambda > 0$ has density

$$f(t; \lambda) = \begin{cases} \lambda \exp(-\lambda t) & \text{for } t > 0 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the distribution of $Y = \lambda X$ when $X \sim \text{Exp}(\lambda)$.

(b) Let $X_1, \dots, X_n$ be an i.i.d. sample from $\text{Exp}(\lambda)$. Show that

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{and} \quad \frac{\sum_{i=1}^n X_i^2}{\bar{X}^2}$$

are independent.

(c) Show that $\bar{X}$ and $\frac{X_{(1)}}{X_{(n)}}$ are independent.

Problem 3. Consider an i.i.d. sample $(X_1, X_2, \dots, X_n)$ with $X_i \sim \text{Unif}[\theta - a, \theta + a]$ for $a \in \mathbb{R}$ known.

(a) Show that $T = (X_{(1)}, X_{(n)})$ is sufficient.

(b) Suppose that we wish to estimate θ under the quadratic loss $L(\theta, d) = (\theta - d)^2$. While the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ appears to be a reasonable estimate, it turns out to be inadmissible when $n > 2$. Show this by deriving an estimator with strictly better MSE than that of $\bar{X}$. (Your estimator should have a simple form.)

Problem 4. Nowadays, unbiasedness is often used to develop estimates of the risk of an estimator. Suppose that $X \in \mathbb{R}^p$ has a multivariate normal distribution with mean vector μ and known covariance $\sigma^2 I$, and let δ be a known differentiable estimator of μ satisfying $\mathbb{E}\left[\|\delta(X)\|_2^2\right] < \infty$ and $\mathbb{E}\left[\left|\frac{\partial}{\partial x_i}\delta_i(X)\right|\right] < \infty$ for $i \in \{1, \ldots, p\}$.

(a) Derive the UMVUE of the squared-error risk $\mathbb{E}\left[\|\mu - \delta(X)\|_2^2\right]$. You may find Problem 2 of Homework 1 useful.

(b) Although the true risk is nonnegative, the UMVUE can take on negative values. How can this be addressed by introducing bias?

Problem 5. Let X and Y be i.i.d. exponential random variables with mean $1/\theta$ or equivalently with density

$$f(t; \theta) = \begin{cases} \theta \exp(-\theta t) & \text{for } t > 0 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the UMVUE for the mean $1/\theta$.

(b) Show that the geometric mean $(XY)^{1/2}$ has smaller mean squared error than the UMVUE for all θ.

(c) Find an estimator which has even smaller mean squared error for all θ.

Problem 6. How much time did you spend on this problem set?