UC Berkeley
Department of Statistics

STAT 210A: Introduction to Mathematical Statistics

Problem Set 2
Fall 2014

Issued: Thursday, September 11 **Due:** Thursday, September 18 (beginning of class)

Problem 2.1
Suppose that $(X_i, Y_i),\ i = 1, \ldots, n$ are sampled i.i.d. from the two-dimensional normal distribution

$$\begin{bmatrix} X & Y \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \theta \\ \theta & 1 \end{bmatrix} \right).$$

with $\theta \in \Omega = (-1, 1)$.

(a) Find a two-dimensional minimal sufficient statistic.

(b) Prove that the minimal sufficient statistic found in (a) is not complete.

(c) Prove that $Z_1 = \sum_{i=1}^n X_i^2$ and $Z_2 = \sum_{i=1}^n Y_i^2$ are both ancillary, but that (Z_1, Z_2) is not ancillary.

Problem 2.2
Let $X_1, \ldots, X_n$ be i.i.d. from a uniform distribution on $(-\theta, \theta)$, where $\theta > 0$ is an unknown parameter.

(a) Find a minimal sufficient statistic T.

(b) Let $\bar{X} = \frac{X_1 + \ldots + X_n}{n}$ and define

$$V = \frac{\bar{X}}{X_{(n)} - X_{(1)}},$$

where $X_{(n)} = \max_i\{X_i\}$ and $X_{(1)} = \min_i\{X_i\}$. Show that T and V are independent.

Problem 2.3
Consider the family of variates $(X_1, X_2, \ldots, X_n)$ with $X_i \sim \text{Uni}[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$.

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

(b) Suppose that we wish to estimate θ under the quadratic loss $L(\theta, a) = (\theta - a)^2$. The sample mean $\bar{X}$ might seem to be one reasonable estimate of θ, but turns out to be inadmissible. Indeed, show that the estimator

$$\delta(X_1, \ldots, X_n) \ := \ \mathbb{E}\left[\bar{X} \mid \min_i\{X_i\}, \max_i\{X_i\} \right]$$

has a strictly better MSE than the MSE of $\bar{X}$, and moreover that $\delta(X_1, \ldots, X_n) = \frac{\min_i\{X_i\} + \max_i\{X_i\}}{2}$.

Problem 2.4
Suppose that $X_1, \ldots, X_n$ are an i.i.d. sample from a *location-scale family* specified by the distribution function $F((x-a)/b)$. (Here F is a known cumulative distribution function; the real numbers a and b are the location and scale parameters respectively.)

(a) If b is known, show that the differences $(X_1 - X_i)/b$ for $i = 2, \ldots, n$ are ancillary.

(b) If a is known, show that the ratios $(X_1 - a)/(X_i - a)$ for $i = 2, \ldots, n$ are ancillary.

(c) If neither a nor b are known, show that the ratios $(X_1 - X_i)/(X_2 - X_i)$ for $i = 3, \ldots, n$ are ancillary.

Problem 2.5
Let $X_1, \ldots, X_n$ be i.i.d. Poisson random variables with mean λ, and suppose that we wish 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 both unbiased for estimating $g(\lambda)$.

(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 $i = 1, 2$. What does your answer have to do with completeness?

Problem 2.6
Let X be a single observation from a Poisson distribution with mean λ. Determine the UMVU estimator for

$$e^{-2\lambda} = [P_\lambda(X = 0)]^2 .$$