UC Berkeley
Department of Statistics

STAT 210A: Introduction to Mathematical Statistics

Problem Set 1
Fall 2014

Issued: Tuesday, September 2 **Due:** Tuesday, September 9

Problem 1.1
Suppose $X_1, \dots, X_n$ are independent random variables and that for $i = 1, \dots, n$, X_i has a Poisson distribution with mean $\lambda_i = \exp(\alpha + \beta t_i)$, where $t_1, \dots, t_n$ are observed constants and α and β are unknown parameters. Show that the joint distributions for $X_1, \dots, X_n$ form a two-parameter exponential family and identify the statistics T_1 and T_2.

Problem 1.2
Let $\{p_\theta(x) : \theta \in \Omega\}$ be an exponential family of densities with respect to some measure μ, where

$$p_\theta(x) = h(x) \exp\left[\sum_{i=1}^{s} \eta_i(\theta) T_i(x) - B(\theta)\right].$$

In some situations, a potential observation X with density $p_\theta(x)$ can only be observed if it happens to lie in some region S. For regularity, assume that $P_\theta(X \in S) > 0$. In this case, the appropriate distribution for the observed variable Y is given by

$$P_\theta(Y \in B) = P_\theta(X \in B \mid X \in S).$$

This distribution for Y is called the *truncation* of the distribution for X to the set S.

1. Show that Y has a density with respect to μ, giving a formula for its density q_θ.
2. Show that the densities $\{q_\theta(x)\}, \theta \in \Omega$, form an exponential family.

Problem 1.3
Consider an i.i.d. sample $\{X_1, \dots, X_n\}$ from the uniform distribution on $[0, \theta]$, and the estimator $M_n = \max\{X_1, \dots, X_n\}$.

(a) Prove that $M_n \xrightarrow{p} \theta$ as $n \to +\infty$.

(b) Compute the bias and variance of the estimator M_n as a function of θ.

(c) Compute the risk of M_n under quadratic loss.

Problem 1.4
Find the natural parameter space Ξ and densities p_η for a canonical one-parameter exponential family with μ counting measure on $\{1, 2, \dots\}$, $h(x) = x^2$ and $T(x) = -x$. Also, determine the mean and variance for a random variable X with this density. Hint: Consider what Theorem 2.4 in Keener has to say about the derivatives of $\sum_{i=1}^{\infty} e^{-\eta x}$.

Problem 1.5

A parameterization $\theta \mapsto \mathbb{P}_\theta$ is identifiable if $\theta_1 \neq \theta_2$ implies that $\mathbb{P}_{\theta_1} \neq \mathbb{P}_{\theta_2}$. Which of the following parameterizations are identifiable?

(a) Suppose that $X_1, \ldots, X_p$ are independent with $X_i \sim \mathcal{N}(\alpha_i + \nu, \sigma^2)$. Set

$$\theta = (\alpha_1, \ldots, \alpha_p, \nu, \sigma^2)$$

and consider the family of joint distributions P_θ over $(X_1, \ldots, X_p)$.

(b) Suppose that X is a Bernoulli variable, and we parameterize its probability mass function with $\theta = (\theta_0, \theta_1)$ and $p(x; \theta) \propto \exp(\theta_0 x + \theta_1(1 - x))$.

(c) Autoregressive process: consider the collection of RVs $(X_1, \ldots, X_p)$ given by $X_1 \sim \mathcal{N}(0, \sigma^2)$, and $X_{i+1} = \alpha X_i + \sqrt{1 - \alpha^2} W_i$ where $W_i \sim \mathcal{N}(0, \sigma^2)$, independent of the X_i, and $\alpha \in [0, 1]$. Let $\theta = (\alpha, \sigma^2)$ and let P_θ index the joint distribution over $(X_1, \ldots, X_p)$.

Problem 1.6

Suppose that $(X_1, \ldots, X_n)$ are i.i.d. Poisson random variables with parameter θ. Show that $T = \sum_{i=1}^n X_i$ is sufficient in two ways:

(a) first use direct methods: compute the conditional distribution given $T = t$.

(b) apply the factorization theorem.