# QUALIFYING EXAM IN THEORETICAL STATISTICS

5/31/01

## Question I.

1. Define sufficiency and state the Neyman factorization criterion for it.

$(X_i, Y_i) \sim$ iid $f$, $i = 1, \dots, n$, where $f$ is the uniform density on the triangle with vertices at $(0,0)$, $(0,\theta)$, $(\theta,0)$ $(\theta > 0)$.

2. Are $X_1$ and $Y_1$ independent? uncorrelated?
3. Show that $T = \max_i (X_i + Y_i)$ is sufficient for $\theta$.
4. Show that $X_1 + Y_1$ has a triangular density (rising on $[0, 2\theta]$), and deduce the distribution of $T$.
5. What is the MLE $\hat{\theta}$ of $\theta$?
6. By what quick argument do you know $\hat{\theta}$ is biased? Is the bias removable?
7. Show that $\hat{\theta}$ has the same distribution as if it were the MLE on $n$ iid data pairs on the square $[0,\theta] \times [0,\theta]$. Avoid messy calculation where possible.
8. What is the order of consistency of $\hat{\theta}$: $\sqrt{n}$?; better?; worse? comment, skipping detailed calculation.

## Question II.

1. Define admissibility and minimaxity of an estimation procedure for a statistical problem.

$X_i \sim$ iid $\mathcal{N}(\theta, 1)$; $\theta \in \mathbb{R}$.

2. Is the sample mean $\bar{X}$ admissible? minimax? (brief comments). What about in higher dimensions?

3. If $\theta$ were known to be positive, would the sample mean be admissible? minimax?

4. Compute the Fisher information for $\theta$.

5. Derive the Bayes estimator of $\theta$, where $\theta \sim N(0,1)$ a priori. Is it admissible? minimax?

6. Perform an efficiency comparison of the Bayes estimator against $\bar{X}$.

## Question III.

Let $X_1, X_2, \ldots, X_n, \ldots$ be i.i.d. from a $N(\theta, \theta)$ distribution in which $\theta > 0$ is unknown.

1. In general terms, describe an application in which this particular model might be of interest.

2. Compute the maximum likelihood estimator (MLE) of $\theta$.

3. Obtain the asymptotic distribution of the MLE. Then write down a (large sample) 95% confidence interval for $\theta$.

4. Two naive estimators of $\theta$ are given by

$$\hat{\theta}_1 = \bar{X} \qquad \text{and} \qquad \hat{\theta}_2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}.$$

   Compare each of these estimators with the MLE by using the techniques of large sample theory.

5. Is $\bar{X}$ complete? sufficient? Find a statistic that is both complete and sufficient.