Math 281 **Qualifying Exam.** **May 22, 2002**

(1) Consider the distribution that has a p.d.f.

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \lambda > 0, \ x > 0 \\ 0 & \text{otherwise} \end{cases} .$$

a. Compute the characteristic function for this distribution.

Also, compute the mean μ and variance σ^2.

b. Now let $X_1, X_2, \ldots, X_n$ be i.i.d. from this distribution, and use a characteristic function argument in order to prove that

$$\sqrt{n}\ (\overline{X}_n - \mu)/\sigma \xrightarrow{L} Z \sim N(0,1) .$$

c. Give a precise statements (no proof required!) of the Helly-Bray and continuity theorems. Then write an account of how these results are used to justify the technique that was used in part b. (above).

(2) Let $X_1, X_2, \ldots, X_n$ be i.i.d. from a distribution that has a finite 4^{th} moment.

a. Derive the 2-dimensional limiting distribution of

$$\overline{X}_n = \frac{X_1 + X_2 + \ldots + X_n}{n} \qquad \text{and} \qquad S_n^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n} .$$

b. Use the result in part a. to write out an (approximate) 95% confidence ellipsoid for the pair (μ, σ^2) .

c. The limiting distribution of S_n^2, in particular, is *not* distribution free. Explain what this means along with its practical importance.

d. Finally, suppose the underlying distribution is Poisson, i.e.,

$$f_X(x;\lambda) = \begin{cases} \dfrac{\lambda^x e^{-\lambda}}{x!} & \lambda > 0, \ x = 0,1,2, \ldots \\ 0 & \text{otherwise} \end{cases}$$

Thus the mean and variance are both equal to λ; and, in this case, $\overline{X}_n$ and S_n^2 are both unbiased (for λ). From the asymptotic point-of-view, which estimator is preferred and why?

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 θ.
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 θ?
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 θ were known to be positive, would the sample mean be admissible? minimax?

4. Compute the Fisher information for θ.

5. Derive the Bayes estimator of θ, 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 θ.

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

4. Two naive estimators of θ are given by

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

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.