Program in Applied Mathematics

PROBABILITY AND STATISTICS PRELIMINARY EXAMINATION
August 21, 1997

Notice: Do four of the following five problems. Place an X on the line opposite the number of the problem that you are NOT submitting for grading. Please do not write your name anywhere on this exam. You will be identified only by your student number, given below. Show all relevant work.

Student Number

1. For a random sample $X_1, X_2, \ldots, X_n$ from a continuous distribution with density function $f(x)$ and cumulative distribution function $F(x)$, let $U = \min(X_1, X_2, \ldots, X_n)$ and let $V = \max(X_1, X_2, \ldots, X_n)$.

   (a) Find the density function for $U$.

   (b) Find the density function for $V$.

   (c) Find the joint probability density function for $U$ and $V$. State where it is 0 and where it is what you claim it is.

   (d) Let $R = V - U$. Find the density for $R$. State where it is 0 and where it is what you claim it is. Your final answer will be in terms of an integral. In general, the density for $R$ can not be simplified any more than that.

2. $X$ and $Y$ are independent random variables with a common exponential density, that is, with $f_X(x) = \theta e^{-\theta x}$ and $f_Y(y) = \theta e^{-\theta y}$ for $x, y$ and $\theta$ positive. The two densities are 0 otherwise. Set $W = X + Y$. In the following, you are asked to find $f_W(w)$ in three different ways.

   (a) Find the density, $f_W(w)$, for the random variable $W$ by writing down and then evaluating the appropriate convolution integral.

   (b) Use moment generating functions to find the density, $f_W(w)$, for $W$.

   (c) Now consider the situation where at time $t$, the random variable $N(t)$ represents the number of events that have occurred up until time $t$ with $N(0) = 0$. $N(t + s) - N(s)$ has a Poisson distribution with parameter $\theta t$ independent of $s$ (i.e. $P\{N(t + s) - N(s) = k\} = e^{-\theta t} (\theta t)^k / k!$ for $k = 0, 1, 2, \ldots$) and the number of occurrences in disjoint time intervals are independent. Let $X$ denote the time until the first occurrence and $Y$ the time between the first and second occurrences. Argue that $X$ and $Y$ are independent and identically distributed with an exponential density and use the properties of the Poisson distribution to find the density for $W = X + Y$. 
3. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed with density function
$f(x, \theta) = \theta x^{\theta-1}$ for $\theta > 0$ and $0 < x < 1$. Set $T(X) = -\frac{1}{n} \sum_{j=1}^{n} \ln(X_j)$.

(a) Show that $T(X)$ is an unbiased estimator of $1/\theta$.
(b) Show that $T(X)$ is MVUE for $1/\theta$, that is, show that $T(X)$ is a minimum variance unbiased estimator of $1/\theta$. You can do this in one of three ways. You can use the Cramer-Rao inequality, or you can use Rao-Blackwell and completeness, or you can use the notion of minimal sufficiency. Present only one of these methods. If you give more than one of these methods, your work will not be graded.

4. Let $Y_1, Y_2, \ldots, Y_5$ be a random sample of size 5 from a normal population with a mean of 0 and a variance of 1 and let $\bar{Y} = (1/5) \sum_{i=1}^{5} Y_i$. Let $Y_6$ be another independent observation from the same population. In justifying your answers to the following questions, it is sufficient to quote standard definitions and theorems.

(a) What is the distribution of $W = \sum_{i=1}^{5} Y_i^2$? Why?
(b) What is the distribution of $U = \sum_{i=1}^{5} (Y_i - \bar{Y})^2$? Why?
(c) What is the distribution of $\sum_{i=1}^{5} (Y_i - \bar{Y})^2 + Y_6^2$? Why?
(d) What is the distribution of $\sqrt{5}Y_6/\sqrt{W}$? Why?
(e) What is the distribution of $2(5\bar{Y}^2 + Y_6^2)/U$? Why?

5. $n$ independent random variables are to be observed. Each of the random variables can take one of three values and the random variables are identically distributed. Denote the probabilities for the three outcomes by $p_1$, $p_2$ and $p_3$ so that $p_1 + p_2 + p_3 = 1$. Let $X$ denote the number of times that the first value is observed, let $Y$ denote the number of times that the second value is observed and let $Z = n - X - Y$ denote the number of times the third value is observed.

(a) Write down the likelihood function as a function of $p_1$ and $p_2$.
(b) If $p_1 = p$ and $p_2 = 2p$ so that $p_3 = 1 - 3p$, find the maximum likelihood estimator for $p$ as a function of $X$ and $Y$.
(c) Find the likelihood ratio statistic, as a function of $X$ and $Y$, for testing $H_0 : p_1 = p, p_2 = 2p, p_3 = 1 - 3p$ for some $p$ between 0 and 1 against the alternative hypothesis that there are no constraints on the $p_j$ other than $p_1 + p_2 + p_3 = 1$.
(d) Describe in detail how you would carry out the test in part (c), specifically, describe the rejection region for the test.