Probability and Statistics Preliminary
August 2006

Work all 7 problems. Begin each problem on a new page, using one side of the sheet. Throughout "p.d.f." means "probability density function" when it is associated with an absolutely continuous random variable, or "probability distribution function" when associated with a discrete random variable. A table of standard normal cumulative probabilities is attached. Fully justify your answers.

1) A distracted job applicant placed N application letters at random into N envelopes. What is the probability that every letter ends up in a wrong envelope? What is the limit of this probability as $N \to \infty$? (Hint: You may start with the situations of $N = 2, 3$.)

2) a) Show that if A has a uniform distribution on $(0, 1)$, then $M = -\ln A$ has an exponential distribution.
   b) Assume random variables A, B, C are independent and uniformly distributed on $(0, 1)$. What is the probability that the equation $Ax^2 + Bx + C = 0$ has no real roots?

3) Let $P(\cdot)$ and $Q(\cdot)$ be probability measures defined on the Borel σ-field $\mathcal{B}$, and let λ be a convex combination of $P(\cdot)$ and $Q(\cdot)$. Show that λ is also a probability measure on $\mathcal{B}$, and use this fact to show that there are random variables on $([0, 1], \mathcal{B}, \lambda))$ that are neither discrete nor absolutely continuous.

4) A small working family consists of a mother, a father, and a child. Assume the monthly salary for each of the three family members is a random variable with the common mean μ and variance σ^2, and the salaries of family members are independent. For n randomly selected small working families, let $\overline{P}$ be the sample mean salaries of the parents and $\overline{C}$ be the sample mean salaries of their children. Give an approximate value of n so that the probability that $|\overline{C} - \overline{P}| < \sigma/5$ is 0.99.

5) Let $U_i, i = 1, 2, \ldots,$ be i.i.d. r.v.'s from $U(0, 1)$, and let X have the distribution:

$$P(X = x) = \frac{1}{2^x}, \qquad x = 1, 2, \ldots.$$

Find the distribution of $Z = \max\{U_1, \ldots, U_X\}$, noting that the distribution of $Z|X = x$ is that of the highest order statistics from a sample of size x.

6) Let $X_1, \ldots, X_n$ be a random sample from the exponential p.d.f.

$$f_\theta = \begin{cases} \theta e^{\theta x}, & x > 0, \\ 0, & \text{elsewhere}, \end{cases}$$

where $\theta > 0$. The null hypothesis $H_0 : \theta = 1$ is to be tested against the alternative $H_1 : \theta \neq 1$.

   a) What is the likelihood ratio statistic λ for testing this problem?
   b) Show that the rejection region for the likelihood ratio test at level α is given by $\overline{X}_n \leq c_1$ or $\overline{X}_n \geq c_2$ for some numbers c_1 and c_2.
   c) Assume two kinds of type one errors in part b) are selected to be equal. What tables are needed for the standard values $\alpha = 0.01, 0.05, 0.1$ to calculate the maunder c_1 and c_2 for various sample size $n = 1, 2, \ldots, 29$. Explain your reasoning.
   d) What popular tables could also be used in part c) for large sample sizes, say for $n \geq 30$?
   e) What is the limiting distribution of $-2 \ln \lambda$ under the null hypothesis in part a)?

7) Let $X_1, \ldots, X_n$ be a random sample of size n from a distribution with parameter $\theta > 0$ whose p.d.f. $f_\theta(x)$ is positive only for $0 < x < 1$ and is proportional to $x^{\theta-1}(1 - x)$.

   a) Show that the $f_\theta(x)$, $\theta > 0$ is a member of the exponential family of distributions.
   b) Find the most powerful level α test for testing $H_0 : \theta = 1$ vs. $H_1 : \theta = 2$.
   c) Derive the Uniformly Most Powerful Test at level α for testing the null hypothesis $H_0 : \theta \leq 1$ against the alternative $H_1 : \theta > 1$, if such a test exists.
   d) Find a complete and sufficient statistic for θ, if there is one.
   e) Find the Cramér Rao lower bound for θ.