Department of Mathematics
University of Toledo

Master of Science Degree
Comprehensive Examination
Probability and Statistical Theory

July 2, 2002

Instructions:
Do all four problems
Show all of your computations in your blue book
Prove all of your assertions or quote appropriate theorems
Books, notes, and calculators may be used
Problems are equally weighted
This is a three hour test

1. A quality control plan for an assembly line involves sampling $n = 10$ finished items per day and counting Y, the number of defective items. If θ denotes the probability of observing a defective item, then Y has a binomial distribution, when the number of items produced by the line is large. In other words, we can assume that $Y \sim B(n, \theta)$ with $n = 10$. However, θ varies from day to day and is assumed to have a uniform distribution on the interval from 0 to 1/4.

(a) Find the expected value of Y for any given day.

(b) Find the standard deviation of Y.

(c) Find the correlation coefficient between Y and θ.

2. Let

$$
\begin{aligned}
Y_1 &= \beta_1 + \epsilon_1 \\
Y_2 &= 2\beta_1 - \beta_2 + \epsilon_2 \\
Y_3 &= \beta_1 + 2\beta_2 + +\epsilon_3,
\end{aligned}
$$

where ϵ_1, ϵ_2, and ϵ_3 are uncorrelated random variables with $E(\epsilon_i) = 0$ and $\text{Var}(\epsilon_i) = \sigma^2$ for $i = 1, 2, 3$.

(a) Find the least squares estimates $(\hat{\beta}_1, \hat{\beta}_2)$ of (β_1, β_2).

(b) Find $\text{Var}(\hat{\beta}_1)$ and $\text{Var}(\hat{\beta}_2)$.

(c) Find $\text{Cov}(\hat{\beta}_1, \hat{\beta}_2)$.

(d) Suppose further that $\sigma^2 = 1$ and $\epsilon_1, \epsilon_2, \epsilon_3$ are iid $N(0, 1)$ random variables. Find a complete and sufficient statistic for (β_1, β_2).

3. A random sample $X_1, \ldots, X_n$ is drawn from the distribution with the following pdf:

$$f(x; \alpha) = 2x / \alpha^2 \text{ for } 0 < x < \alpha.$$

   a. Find a sufficient statistic for α.
   b. Find the maximum likelihood estimator of α.
   c. Show that the maximum likelihood estimator of α is biased, but asymptotically unbiased.
   d. Find its CDF and show that it is consistent.
   e. Consider the test of $H_0: \alpha = 2$ versus the one-sided alternative $H_1: \alpha < 2$. Find the likelihood ratio test for these hypotheses.
   f. If the observed data were .8, 1.2, 1.2, 1.4, what would be your conclusion? Use level of significance equal to .02.

4. In this problem we will consider the ***Sign test*** and its power for two different distributions. We will see that the test is independent of the underlying distribution, but the power is not. Let m_X and m_Y denote the medians of the two distributions, and let μ_X and μ_Y denote the means of X and Y. We will test H_0: $m_X = m_Y$ versus the one tailed alternative, H_1: $m_X > m_Y$.

   a. For n=2, that is, with data $\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}$ and $\begin{pmatrix} X_2 \\ Y_2 \end{pmatrix}$, find the critical region for the test based on the Sign test statistic that gives the smallest $\alpha > 0$. What is the level of significance for this test?
   b. Let $\Phi(x)$ denote the CDF for the standard normal distribution. Say that our data $(X,Y)'$ comes from a bivariate normal distribution with mean $(\mu_X, \mu_Y)'$ and covariance matrix $\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$. In terms of Φ, find the power of the test from part a as a function of these five parameters. Show that the power is an increasing function of $\Delta = \mu_X - \mu_Y$.
   c. Now say that our data comes from a distribution defined on the unit square $[0,1]\times[0,1]$ with $f(x) = c(1-\alpha x)$, with $\alpha < 1$. Find the constant c and the means μ_X and μ_Y and the medians m_X and m_Y.
   d. Find the power of the test from part a for this distribution. Show that it is an increasing function of either $\Delta_\mu = \mu_X - \mu_Y$ or $\Delta_m = m_X - m_Y$.