DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF MASSACHUSETTS
BASIC EXAM - STATISTICS
January 27, 2006

Do all five problems. Sixty points are needed to pass at the Master's level and seventy-five at the Ph.D. level.

1. (16 points) For one observation Y from a normal distribution with variance one and mean 0 or 2, consider $H_0 : \mu = 0$ and $H_A : \mu = 2$. Suppose first that we observe only Y.
   (a) Construct a size α likelihood ratio test. Give explicitly the rejection region in terms of Y.
   (b) Find the power for your test in the previous part.
   (c) Is the test unbiased? Explain.
   (d) Is the likelihood ratio test UMP? Explain what result you are applying.

2. (20 points) Children are given an intellegence test. Given a child with true IQ equal to μ, this child's test score Y is a normal random variable with variance 100 and mean μ. Suppose the mean μ is viewed as having a normal distribution with mean 100 and variance 225 (this could be from acutally sampling an individual at random from a population or in the Bayesian perspecive, which can be viewed as a prior distribution for μ.)
   (a) Incorporating the "randomness" in μ and that of Y given μ, the marginal distribution of Y is known to be normal. Find $E(Y)$ and $\text{Var}(Y)$.
   (b) What is the posterior distribution of μ given $Y = y$.
   (c) Suppose a child scores 115 on a test:
      i. Use the previous part to give the posterior distribution of his or her true IQ.
      ii. Find 95% Bayesian confidence interval for his or her true IQ.
      iii. Find the posterior probability that his or her true IQ is less than 100.

3. (20 points) The assessment of the proportion of defective units in a lot of units is an important problem. Suppose you take a random sample of n units from a lot large enough to treat $X_1, \ldots, X_n$ as i.i.d. Bernoulli (p), where $X_i = 1$ if unit i in the sample is defective and is 0 otherwise. Hence, p is the probability of getting a defective unit or equivalently, the proportion of defective units in the population. Let

$$\hat{p} = \sum_{i=1}^{n} X_i / n$$

   (a) Give a complete sufficient statistic for p. *State precisely what result you are applying to give this.*

(b) We usually use $\hat{p}$ to estimate p and use $\hat{p}(1-\hat{p})/n$ to estimate the $V(\hat{p}) = \sigma^2_{\hat{p}}$ (the variance of $\hat{p}$). Show that $\hat{p}(1-\hat{p})/n$ is a biased estimator of $\sigma^2_{\hat{p}}$.

(c) Find the UMVUE of $\sigma^2_{\hat{p}}$. You must justify your answer.

4. (20 points) Let $X_1, \cdots, X_{10}$ be iid random samples taken from $N(\mu_1, \sigma_1^2)$ and $Y_1, \cdots, Y_{12}$ be iid random samples taken from $N(\mu_2, \sigma_2^2)$. Define $\bar{X}$ and S_1^2 to be the sample mean and variance, respectively of $X_1, \ldots, X_{10}$ respectively and define $\bar{Y}$ and S_2^2 similarly.

   (a) State the distribution of each of $\bar{X}$, S_1^2, $\bar{Y}$ and S_2^2.

   (b) In each of the following questions, a pair of null hypothesis and alternative hypothesis is specified, where in each it is stated which parameters are known or unknown. In each case, provide a test for the stated hypotheses. Specify the rejection region (with numeric critical value) for a 0.05-level test. Each will involve a well known distribution. (Tables are provided for your reference.)

      i. Suppose here $\sigma_1^2 = 5$ and μ_1 is unknown.

      $$H_0: \ \mu_1 \leq 2 \text{ vs. } H_1: \ \mu_1 > 2$$

      ii. Here all μ_i and σ_i^2 $(i = 1, 2)$ are unknown, but $\sigma_1^2 = \sigma_2^2$.

      $$H_0: \ \mu_1 = \mu_2 \text{ vs. } H_1: \ \mu_1 \neq \mu_2$$

      iii. Here μ_1, σ_1^2 and σ_2^2 are unknown.

         A. $H_0: \ \mu_1 = 2$ vs. $H_1: \ \mu_1 \neq 2$.
         B. $H_0: \ \sigma_1^2 \leq 5$ vs. $H_1: \ \sigma_1^2 > 5$
         C. $H_0: \ \sigma_1^2 \leq \sigma_2^2$ vs. $H_1: \ \sigma_1^2 > \sigma_2^2$
         D. $H_0: \ \sigma_1^2 = \sigma_2^2$ vs. $H_1: \ \sigma_1^2 \neq \sigma_2^2$

5. (24 points) Consider a random sample $Y_1, \ldots, Y_n$ from exponential distribution $f(y) = (1/\beta)e^{-y/\beta}$ for $y > 0$ $(\beta > 0)$, a distribution with mean $\mu = \beta$ and variance $\sigma^2 = \beta^2$.

   (a) Find the MLE of β. Show your derivation and be sure to justify that you have maximized the likelihood.

   (b) Derive an **exact** 95% confidence interval for μ. (Hint: The variable $2\sum_{i=1}^n Y_i/\beta$ follows a well-known distribution. If you can't name this distribution, trade the points by using a letter to represent the distribution and define its quantiles. So that you can continue to do the following problems.)

   (c) Derive an **exact** 95% confidence interval for σ^2.

   (d) Find the Cramer-Rao Lower bound for unbiased estimators of σ^2.

(e) An interesting property of the exponential distribution is that it is memoryless; that is $P(Y > s|Y > t) = P(Y > s - t) = e^{-(s-t)/\beta}$. Call this quantity $g(\beta)$ where $s > t$ and s and t are fixed.

i. Give the MLE of $g(\beta)$ and then give the approximate large sample distribution of this MLE.

ii. Find an approximate 95% large sample confidence interval for $g(\beta)$ by using the approximate distribution in the previous part, plus whatever other developments (explain them) that are needed.