Statistics Prelim Exam

January 2015

Read the following instructions before you begin:

- You may attempt all of 10 problems in this exam. However, you can turn in solutions for **at most** 6 problems. On the outside of your exam booklet, indicate which problem you are turning in.
- Each problem is worth 10 points; 40 points or higher will result in a pass.
- If you think that a problem is misstated, then interpret that problem in such a way as to make it nontrivial.
- If you wish to quote a result that was not in your 5080–5090/6010 texts, then you need to carefully state that result.

Exam problems begin here:

1. Let $X_1, X_2, \ldots, X_n$ be a random sample from the Poisson distribution with parameter $\lambda > 0$. Use the Rao–Blackwell theorem and/or the Lehmann–Scheffé theorem to find a UMVUE of $\mathrm{P}\{X_1 = 0\} = \mathrm{e}^{-\lambda}$.

2. Find the maximum likelihood estimators of θ and α, based on a random sample $X_1, X_2, \ldots, X_n$ from the distribution with density

$$f(x; \theta, \alpha) = \frac{\alpha}{\theta^\alpha} x^{\alpha-1} \qquad 0 \leqslant x \leqslant \theta.$$

Here θ and α are positive.

3. 1000 individuals were classified according to sex and to whether or not they were color-blind, with the following results: male normal 442; female normal 514; male color-blind 38; and female color-blind 6. According to a genetic model, the four relative frequencies should be of the form $p/2$; $p^2/2 + p(1-p)$; $(1-p)/2$; and $(1-p)^2/2$, respectively. Suppose we want to test whether the data fit the model.

   (a) Find the maximum likelihood estimator $\hat{p}$ of p.

   (b) Describe a size $\alpha = 0.10$ test of the null hypothesis that the data fit the model, assuming that $\hat{p} = 0.913$. Specify the critical region as precisely as possible, perhaps in terms of a quantile of a certain χ^2 distribution.

4. A scale has two pans. The measurement given by the scale is the difference between the weights on pan 1 and pan 2 plus a random error. Thus, if a weight μ_1 is put on pan 1 and a weight μ_2 is put on pan 2, then the measurement is $Y = \mu_1 - \mu_2 + \varepsilon$. Suppose that $\mathrm{E}[\varepsilon] = 0$, $\mathrm{Var}(\varepsilon) = \sigma^2$, and in repeated use of the scale observations Y_i are independent.

   Suppose two objects, #1 and #2, have weights β_1 and β_2. Measurements are then taken as follows:

   - Object #1 is put on pan 1 and nothing on pan 2
   - Object #2 is put on pan 2 and nothing on pan 1
   - Object #1 is put on pan 1 and object #2 is put on pan 2
   - Objects #1 and #2 are both put on pan 1 and nothing on pan 2

   (a) Let $\boldsymbol{Y} = (Y_1, Y_2, Y_3, Y_4)'$ be the vector of observations. Formulate this as a linear model.

   (b) Find the least squares estimators of β_1 and β_2.

   (c) Find the variances of the least squares estimators as well as the covariance between them.

5. Let $X_1, \ldots, X_{2n+1}$ be independent lognormal(μ, σ^2). That is, $X_1 = \mathrm{e}^{Y_1}, \ldots, X_{2n+1} = \mathrm{e}^{Y_{2n+1}}$ with $Y_1, \ldots, Y_{2n+1}$ being independent $N(\mu, \sigma^2)$. Use a theorem to show that the sample median $X_{n+1:2n+1}$ is asymptotically normal and find its asymptotic mean and asymptotic variance.

6. Suppose we have two samples of sizes $n_1 = 4$ and $n_2 = 6$. We observe

$$(x_1, x_2, x_3, x_4) = (3.7, 3.3, 4.2, 3.1)$$

and

$$(y_1, y_2, y_3, y_4, y_5, y_6) = (3.6, 5.0, 4.5, 3.9, 5.1, 4.7).$$

The null hypothesis is that $F_X = F_Y$ and the alternative hypothesis is that X is stochastically smaller than Y. Use the Wilcoxon/Mann–Whitney test. Determine the value of the test statistic and evaluate the p-value of the data. Do not use a large-sample approximation.

7. Let $X_1, X_2, \ldots, X_n$ be a random sample from a $N(\theta\,, \sigma^2)$ distribution (σ^2 known), and let the prior distribution of θ be $N(\mu_0\,, \sigma_0^2)$, where μ_0 and σ_0^2 are known.

   (a) Find the posterior distribution of θ.

   (b) Find the Bayes estimator under squared error loss.

8. (a) Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\theta\,, \theta^2)$, where $\theta > 0$. Find a pair of statistics that are jointly sufficient.

   (b) Using some linear combination of $\sum_{i=1}^n X_i^2$ and $(\sum_{i=1}^n X_i)^2$, show that the statistics of part (a) are not jointly complete.

9. Let X_1 and X_2 be independent EXP(1) random variables (density e^{-x}, $x > 0$). Find the joint density of $Y_1 = X_1/X_2$ and $Y_2 = X_1 + X_2$. Find the marginal densities of Y_1 and Y_2 as well.

10. Let $X_1, \ldots, X_n$ be a random sample from UNIF$[0\,, \theta]$, where θ is a positive unknown. Derive the generalized likelihood ratio (GLR) test of $H_0 : \theta = \theta_0$ vs. $H_a : \theta \neq \theta_0$.