Mathematics 376 – Probability and Statistics 2
Final Examination
May 6, 2006

Directions

Do all work in the blue exam booklet, and *include all work necessary to justify your answers*. Be sure to read each question carefully before starting to work. There are 200 regular points and 10 Extra Credit points. *General* Note: A "random sample" always consists of *independent* measurements from the indicated distribution.

I. Let $Y_1, \ldots, Y_n$ be a random sample from a distribution with probability density function $f(y|\theta) = \theta y^{\theta-1}$ if $0 < y < 1$ and 0 otherwise. We also assume $\theta > 0$.
A) (15) Find the method of moments estimator for θ.
B) (15) Find the maximum-likelihood estimator for θ.

II. Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ be two random samples from normal distributions with common variance σ^2.
A) (10) Using the fact that the expected value of a χ^2 random variable with ν degrees of freedom is ν, show that $S_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ is an unbiased estimator for σ^2.
B) (10) Show that the pooled estimator S_p^2 using both the X_i and the Y_j is also unbiased for σ^2.

III.
A) (15) Let $Y_1, Y_2, \ldots, Y_{10}$ be a random sample from a normal distribution with mean $\mu = 2$ and variance $\sigma^2 = 81$. Let $U = \frac{1}{81}\sum_{i=1}^{9}(Y_i - 2)^2$. What is the distribution of $V = \frac{Y_{10}-2}{3\sqrt{U}}$?
B) (15) If T has a t-distribution with ν degrees of freedom, what is the distribution of T^2? Explain.

IV. Let p be the proportion of letters mailed in the Netherlands that are delivered the next day.
A) (15) A random sample of $n = 200$ letters are sent out and 142 are delivered the next day. Find an approximate 95% confidence interval for p based on this sample.
B) (15) ("Thought question") Note that part A says "approximate." What is the actual distribution of $Y =$ the number of letters delivered the next day (out of a random sample of size $n = 200$)? Why does the method you used in part A give a reasonable interval estimate for p?

V. A mathematics department wishes to evaluate a new method of teaching calculus with Maple labs. At the end of the course, 15 students who used the labs are given a standardized test. Their average score is 83, with standard deviation 9.
A) (10) Find a 95% confidence interval for the mean test score for students who are taught using the new method.

B) (10) From departmental experience, students who are taught the course without the Maple labs average 79 on the same standardized test, and the standard deviation is also 9 for these students. Is there sufficient evidence to conclude that taking the course with the labs has an effect on students' performance, at the $\alpha = .05$ level?

C) (10) What number n of students who took the course with the labs would have to be tested in order for the department to be "98% sure" that the sample mean test score is no farther than 1 away from the true population mean test score for the students taking the course with the labs. You may assume for this part that n will be significantly > 30.

VI. The fill weights of a random sample of $n_1 = 21$ 6-pound boxes of "Super-Sudsy" laundry soap produced at Plant 1 had a mean of 6.25 pounds and standard deviation $s = .095$ pounds. A similar sample of size $n_2 = 21$ produced at Plant 2 had mean weight 6.12 pounds and standard deviation $s = .065$.

A) (15) Is there sufficient evidence at the $\alpha = .01$ level to conclude that the standard deviations at the two plants are different?

B) (15) Is there sufficient evidence to conclude that the mean fill weights are different? Report the results by giving an estimate of the p-value of your test.

VII. The following table gives measurements of the firmness of pickles stored in low-salt brine as a function of time:

x time (weeks)	y firmness (lb)
1	19.8
4	16.5
14	12.8
32	8.1
52	7.5

A) (15) Find the least-squares estimators for the coefficients in a model $Y = \beta_0 + \beta_1 x + \varepsilon$ for this data set.

B) (15) Is there sufficient evidence to say that $\beta_0 < 21$? Explain, using the p-value of an appropriate test.

Extra Credit (Short Essay – no more than a paragraph, please!) (10) Suppose we needed to choose between the estimators in question I, parts A and B. One criterion for comparing estimators we discussed in class used the relative efficiency $\text{eff}(\widehat{\theta_1}, \widehat{\theta_2}) = V(\widehat{\theta_2})/V(\widehat{\theta_1})$. But it shouldn't be clear how to compute either variance theoretically from your formulas. Propose a method to test which estimator is superior "experimentally," taking into account the possibility that which estimator is superior might depend on the true value of θ.

Have a safe, enjoyable, and productive summer!