Mathematics 376 – Probability and Statistics II
Review Sheet – Exam II
April 2, 2004

General Information

As announced in the course syllabus, the second midterm exam this semester will be given in class on *Friday, April 16*. The exam will cover the material we have discussed since the last exam, starting from the material on confidence intervals for variances from section 8.9 (see Lab Project 2), and going through the material we covered in Chapter 10 on hypothesis testing, culminating in Lab Project 3. There is a more detailed breakdown of topics given below.

Format

This will be an *open book, open class notes, open problem sets/solutions exam*. But you should still prepare carefully for the exam and understand the key concepts we have talked about. Know how to apply the different techniques we have studied and how to select the most appropriate method when there are several possible choices. If you are flipping through the book to look for *every* formula you need or if you need to find examples similar to the test questions to get started on a problem, you *will not be able to finish the exam in 50 minutes*.

Topics

The topics to be covered (not in chronological order, but according to logical connections):

1) The method of moments and the method of maximum likelihood for deriving estimators. (Some of the other material we discussed in Chapter 9 on consistency of estimators, sufficient statistics, etc. is mainly background theoretical information for your general mathematical education and will not feature in the exam questions.)
2) Hypothesis testing – the general concepts: null hypothesis, alternative hypothesis, test statistic, rejection region, Type I error probability ($= \alpha$, or level of test), Type II error probability ($= \beta$), attained significance level (p-value of a test), interpretation of results.
3) The connection between confidence intervals and rejection/"acceptance" regions for tests.
4) Large sample (Z-) tests and related confidence intervals for means and proportions (Note: some of this overlaps material from Midterm 1!). Questions here might also ask you to design tests with a given α-value to achieve a certain β-value by selecting sample size appropriately.
5) Small sample (t-) tests for means and related confidence intervals.
6) χ^2-tests for variances and related confidence intervals.
7) F-tests for ratios of variances and related confidence intervals.

Comment: As you should be able to tell, the justifications for the methods we have developed here depend heavily on the probability topics we learned last semester, as well as the sampling distribution theory (Z, χ^2, t, F distributions, etc.) and the Central Limit Theorem from Exam 1. If you are feeling "rusty" on any of that, start by reviewing that material.

Review Session

I will be happy to run a pre-exam review session if there is interest. The evening of Wednesday, April 14 would be best, but other times might be possible too.

Suggested Review Problems

From the text: Chapter 9/62, 63, 64, 73 (ignore the Hint – just work it out!), 75, 76 a,b, Chapter 10/3 (this continues problem 2, which we did as an example in class), 103, 105, 106, 109, 111, 112

Sample Exam

Disclaimer: A reasonable 1 hour exam cannot cover every topic we have discussed in this section of the course (in particular every type of estimation and hypothesis testing problem we have seen). You should be prepared for all the possibilities though. The following problems indicate the approximate level of difficulty and cover a possible subset of the topics on the upcoming exam; the actual exam problems may differ substantially from these and might deal with different situations.

Comments: Any general formula we have studied can be used without comment (i.e. you don't need to rederive it in you solution).

I. Let $Y_1, \ldots, Y_n$ be independent samples from the distribution with pdf containing the unknown parameter θ:

$$f(y|\theta) = \begin{cases} \frac{1}{\theta} y^{(1-\theta)/\theta} & \text{if } 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

A) What is the likelihood function $L(y_1, \ldots, y_n|\theta)$?
B) Determine the maximum likelihood estimator for θ.

II. Let X and Y be the forces required to pull two different studs out of an automobile window. Assume that the distributions are both normal with $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$.
A) Assuming that $\sigma_X^2 = \sigma_Y^2$, and you have $n = 10$ X_i measurements and $m = 10$ Y_j measurements, specify a test statistic and rejection region for testing $H_0 : \mu_X = \mu_Y$ against $H_a : \mu_X \neq \mu_Y$, if the test is to have level $\alpha = .05$.
B) The following $n = 10$ measurements of X:

$$111, 120, 139, 136, 138, 149, 143, 145, 111, 123$$

and $m = 10$ measurements of Y:

$$152, 155, 133, 134, 119, 155, 142, 146, 157, 149$$

were obtained. Calculate your test statistic from part A) and state your conclusion.

C) Compute a two-sided 95% confidence interval for $\mu_X - \mu_Y$ and state the relation between your answer here and your answer for part B.

D) What is the approximate p-value of your test in part B?

E) Test whether the assumption $\sigma_X^2 = \sigma_Y^2$ seems valid, using an appropriate test with level .05.

F) Suppose you had $n = 50 = m$ measurements of X and Y instead of $n = 10 = m$. What, if anything, would change in your answer to part A? Explain.

G) Assuming $m = n$ can be chosen arbitrarily, how large would be need to take them to obtain a test as in part F with $\beta = .05$, assuming $\mu_X - \mu_Y = 10$?

III. (Essay) In intuitive terms we might say "if the test statistic turns out to be outside the rejection region, then we accept H_0." Is that precisely correct, though? Should we take the results of the test as indicating H_0 is definitely true in that case?