Math 214

REVIEW SHEET EXAM #3

Exam: Wednesday May 9, 2007

Coverage: Exam 3 will cover the following material:

- Estimating Population Parameters: Chapter 6
  1. What is a sampling distribution? What is a parent distribution?
  2. What is a point estimator?
  3. What is an unbiased estimator?
  4. How to chose estimator for the Center of a Distribution? (Mean or Median?)
- Confidence Intervals: Chapter 7: sections 1-3
  1. CI for population proportion
  2. CI for population mean: σ known.
  3. CI for population mean: σ unknown.
- Hypothesis Testing: Chapter 8: sections 1-3
  1. Null and Alternative hypotheses (H_0,H_a).
  2. p-values.
  3. One and two tailed tests.

  You should understand the following facts/formulas and know how and when to apply them:

  The central limit theorem states that if $x_1, x_2, \ldots, x_n$ is a random sample from a population with mean μ and standard deviation σ, then the sample mean is approximately normal with mean μ and standard deviation $\sigma/\sqrt{n}$.

  Under assumptions, $(1-\alpha)100\%$ CIs for the population proportion π based on $\widehat{p}$ and the population mean μ based on $\bar{y}$ are respectively

$$\pi = \widehat{p} \pm z^* \sqrt{\widehat{p}(1-\widehat{p})/n} \quad \text{and} \quad \mu = \bar{y} \pm z^* \sigma/\sqrt{n}, \quad \mu = \bar{y} \pm t^* s/\sqrt{n}.$$

Confidence Intervals:

1. A random sample of 21 homes in a region found that an average of 160 gallons of heating oil were consumed each January. Assume that the amount of oil conusmed is approximately normally distributed. Answer the following:
   (a) If the standard deviation of the sample is 31 gallons, compute a 95% confidence interval based on $\bar{y}$ for the mean amount of heating oil used per household in this region.
   (b) Assume the standard deviation of heating consumption in all homes is 32 gallons. Compute a 95% confidence interval based on $\bar{y}$ for the mean amount of heating oil used per household in this region.
   (c) If the sample size is increased to 100, and the data now shows $\bar{y} = 164$, with sample standard deviation of 33, compute 95% confidence intervals for mean amount of heating oil used per household in the region.
2. Are Pot Penalties too severe? A sample of 2,400 college students is taken. 960 believe that the penalties for possession of small amounts of marijuana are too severe.
   (a) Find a 90% confidence interval for the proportion of all college students who think the penalties are too high.
   (b) Find a 99% confidence interval for the proportion of all college students who think the penalties are too high.
   (c) How large a sample is needed if the margin of error in the poll is to be no greater than 2%?
3. According to a new report from Israel, taking aspirin up to the day of a coronary bypass grafting, as opposed to stopping one week prior to surgery, seems to speed lung function recovery after the surgery without increasing the risk of bleeding significantly.

   Suppose the data summarizing the study are

Group	n	ybar	s
took aspirin	14	4	3
stopped aspirin	18	9	2.5

   The variable x records the time spent on a ventilator after surgery.

   Is there statistically significant evidence level that the mean time spent on the ventilator has decreased? One way of answering this question is to construct 95% C.I. for estimates of the population mean times from each sample. Do this, assuming each sample is approximately normally distributed. Do the estimates overlap? What happens if we changed the C.I. to 80%?

4. The t-test, and the t-distribution, actually evolved from Gossett's diligent work on a data set of Criminals in England. He had 3,000 measurements of prisoner's height. He broke these down into 750 sets of 4 measurements. From these he computed the t-score of each set and plotted the sampling distribution of this statistic.

   Suppose the first set of four heights (in inches) was:

   ```
   70, 65, 63, 68
   ```

   From these values, find a 95% confidence interval for the average height of the 3,000 prisoners.

Hypothesis Testing:

The steps you will need to know to perform any of these significance tests are
(a) Specify the null and alternative hypotheses: H_0 , H_a .
(b) Collect the data.
(c) Compute the observed value of the appropriate test statistic
(d) Either use the table to find the critical value(s) for the problem or the p-value if using the computer. The critical value(s) make use of α and the alternative hypothesis.
(e) Finally, compare the observed value with the critical value(s) or the p-value with α to determine if the difference is statistically significant.

Some sample problems follow:

1. The English mathematican John Kerrich tossed a coin 10,000 times and obtained 5,060 heads. Is this significant evidence (at the 5% level) that the probability that Kerrich's coin comes up heads is different from 0.5?
2. Land's Beginning is considering buying a mailing list of 100,000 people, but only if they are reasonably confident that 5% (or more) will respond to direct mailing. They are given a random sample of 500 from the list and find that only 4% responded to a mailing. Is this difference from 5% statistically significant to make LB not want to buy these names?
3. How accurate are home radon detectors? A company tested 6 in an environment with an accurately measured amount of 115. The 6 cheap detectors found

   ```
   91.9 122.3 105.4 95.0 99.6 120.9
   ```

   with summary

   ```
   xbar        sd      n
   105.85000 13.02839 6.00000
   ```

Is there evidence to indicate that the home measurements are lower than they should be?

4. Does removing a annual credit card fee change consumer purchases? A bank investigates by waiving the fee on a random sample of 500 of its customers and tracks differences in the difference in the amount they spent. The mean increase was \$565 with s = \$267. Is there significant evidence at the 5% level that the mean amount increased?