Sociology 592 - Research Statistics I
Exam 1
September 24, 2004

Where appropriate, show your work - partial credit may be given. (On the other hand, don't waste a lot of time on excess verbiage.) Do not spend too much time on any one problem. It is legitimate (and probably essential) to refer to results that have previously been proven in class or homework, without re-proving them - for example, you wouldn't need to prove that $P(-1.96 \leq Z \leq 1.96) = .95$, since we have already shown that in class. Likewise, you are free to refer to anything that was demonstrated in the homework or handouts.

---

1. (4 points each, 20 points total). Indicate whether the following statements are true or false. If you think the statement is false, indicate how the statement could be corrected. For false statements, do not just say that you could substitute not equals for equals. For example, the statement $P(Z \leq 0) = .7$ is false. To make it correct, don't just say $P(Z \leq 0) <> .7$, instead say $P(Z \leq 0) = .5$ or $P(Z \leq .525) = .7$.

A. If X has a binomial distribution, then if N is large $\overline{X}$ will also have a binomial distribution.

B. $P(Z \geq 1.96) = .05$

C. When two events are mutually exclusive, the occurrence of one event in no way affects the occurrence of the other.

D. $V(8 + X) = 64 + V(X)$

E. When N is small, the T distribution and the N(0, 1) distribution are almost identical.

---

2. (10 points each, 30 points total) Answer <u>three</u> of the following. The answers to most of these are fairly straightforward, so do not spend a great deal of time on any one problem. NOTE: I will give up to 5 points extra credit for each additional problem you do correctly.

A. $\overline{X} = 30$, N = 25. Determine the 95% confidence interval when

a. $\hat{\sigma} = 15$

b. $\sigma = 10$

B. Here is a random sample of 7 scores from a previous cohort's first exam in statistics. Compute the sample mean and the sample standard deviation.

Score
76
86
90
91
93
99
105

C. A school district is concerned about the practice of "social promotion," i.e. the practice of moving students on to the next grade whether they are ready or not. It therefore requires that all 8th graders take a standardized test that measures how much they have learned in grade school. All students scoring in the bottom 10% of the test will be held back and made to take 8th grade again. If test scores ~ N(90, 10), how high does your score have to be to avoid getting held back a year?

D. It is election day, 2004. Once again, the American Presidential race is too close to call – and, just like in 2000, it appears that the state of Florida could play a critical role in determining the next President.

John Kerry estimates that if he wins Florida, there is an 80% chance that he will win the Presidency. But, if he loses Florida, he has only a 24% chance of becoming president. Finally, he estimates that he and George Bush both have a 50% chance of carrying the state.

What is the probability that John Kerry will be elected President? What is the probability that he will lose both the State of Florida and the Presidency?

E. The population variance of x is known to be 4. Stata's `ztesti` command produces the following output:

```
------------------------------------------------------------------------------
Variable |      Mean    Std. Err.       z     P>|z|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
       x |         5           1         5   0.0000       3.040036    6.959964
------------------------------------------------------------------------------
```

What was the sample size? How big would the sample need to be if you wanted the true standard error of the mean to be only 0.1?

3. (25 points) A large company is being forced to lay off several thousand of its employees. It has agreed to hire an outside firm that will help its ex-employees find new jobs. Two employment placement services are being considered for the task. Each has provided information on the last 1,000 individuals who used their services. A review of the two companies reveals that

- Company A placed 840 of its clients in new jobs in three months or less. Company B, on the other hand, placed only 600 of its clients that quickly.
- 70% of Company A's clients previously had jobs in management, while the rest were manual laborers. For Company B, only 40% of its clients had been in management.
- For company A, 630 of those placed within 3 months were previously in management. For Company B, 300 of those placed within 3 months were previously managers.

a. (10 pts) Complete the following table

	Company A			Company B		
	Placed within 3 months	Not Placed within 3 months	Σ	Placed within 3 months	Not Placed within 3 months	Σ
Management						
Labor						
Σ			1000			1000

b. (10 pts) Company A was much more successful at finding jobs for its clients within 3 months. However, Company A also had many more clients who were previously in management, and those who were in management were much more likely to get new jobs quickly. Suppose that Company B had had Company A's clients (i.e. it had 700 managers and 300 laborers). How many of those individuals would have been placed in new jobs within 3 months, assuming that Company B's placement rates for managers and for laborers each stayed the same?

c. (5 pts) Based on these results, would you recommend that Company A or Company B be hired to provide placement services? Explain your reasoning.

4. (25 points) Supporters of Republican Congressman Chris Chocola claim that at least 60% of the voters in the Congressional District support his re-election. Skeptical Democrats immediately dismiss these claims. A random sample of 200 registered voters is drawn, 107 of whom say they will vote for Chocola. Using the .05 level of significance, test whether the Chocola supporters' claim is supported. Be sure to indicate:

(a) The null and alternative hypotheses - and whether a one-tailed or two-tailed test is called for.

(b) The appropriate test statistic

(c) The critical region

(d) The computed value of the test statistic

(e) Your decision - should the null hypothesis be rejected or not be rejected? Why?

(f) Would your decision change if you used the .01 level of significance instead? Why or why not?