Name: _____

Practice FINAL

Dr. Nancy Pfenning
Statistics 1000
Spring 2007

The Final will be a closed book exam. You are allowed to bring needed tables, a calculator, and 2 two-sided sheets of notes.

1. (15 pts.) Select the most appropriate display for each of the following.

   (a) Workers' income and whether they are black, white, Asian, or Hispanic:
   (i) bar graph (ii) histogram (iii) two-way table (iv) scatterplot
   (v) back-to-back stemplot (vi) side-by-side boxplots

   (b) Rent charged and distance from the university for a sample of Oakland apartments:
   (i) bar graph (ii) histogram (iii) two-way table (iv) scatterplot
   (v) back-to-back stemplot (vi) side-by-side boxplots

   (c) Workers' marital status and whether they smoke or not:
   (i) histogram (ii) two-way table (iii) scatterplot (iv) back-to-back stemplot
   (v) side-by-side boxplots

2. (5 pts.) Which one of the following is a matched pairs study? (i) Measure level of depression for a random sample of internet users and for a random sample of non-users (ii) Measure level of depression for a random sample of non-internet users; provide them with internet use for a year and then measure their level of depression.

3. (5 pts.) In general, which is more likely to contain the unknown population mean?
   (i) a 90% confidence interval (ii) a 99% confidence interval (iii) both the same

4. (35 pts.) Select the most appropriate statistical test for each of the following.

   (a) We survey a random sample of Oakland households to test if "nuclear families" (father, mother, offspring) are in a minority there:
   (i) z test about a proportion (ii) z test about a mean with one-sided alternative (iii) z test about a mean with two-sided alternative (iv) t test about a mean with one-sided alternative (v) t test about a mean with two-sided alternative (vi) two-sample t test with one-sided alternative (vii) two-sample t test with two-sided alternative (viii) chi square test (ix) ANOVA (x) inference for regression

(b) We look at the mean and standard deviation of a small random sample of Oakland one-bedroom apartments to decide if the overall mean rent is more than 500 dollars:

(i) z test about a proportion (ii) z test about a mean with one-sided alternative (iii) z test about a mean with two-sided alternative (iv) t test about a mean with one-sided alternative (v) t test about a mean with two-sided alternative (vi) two-sample t test with one-sided alternative (vii) two-sample t test with two-sided alternative (viii) chi square test (ix) ANOVA (x) inference for regression

(c) We want to test if the mean rent of all Oakland one-bedroom apartments is equal to the mean rent of all Shadyside one-bedroom apartments:

(i) z test about a proportion (ii) z test about a mean with one-sided alternative (iii) z test about a mean with two-sided alternative (iv) t test about a mean with one-sided alternative (v) t test about a mean with two-sided alternative (vi) two-sample t test with one-sided alternative (vii) two-sample t test with two-sided alternative (viii) chi square test (ix) ANOVA (x) inference for regression

(d) We test the IQ of a large random sample of children of women smokers to see if their mean IQ is significantly lower than the national mean of 100. We assume population standard deviation to be 16:

(i) z test about a proportion (ii) z test about a mean with one-sided alternative (iii) z test about a mean with two-sided alternative (iv) t test about a mean with one-sided alternative (v) t test about a mean with two-sided alternative (vi) two-sample t test with one-sided alternative (vii) two-sample t test with two-sided alternative (viii) chi square test (ix) ANOVA (x) inference for regression

(e) We examine a random sample of Oakland apartments to see if overall there is a relationship between rent charged and size (in square feet):

(i) z test about a proportion (ii) z test about a mean with one-sided alternative (iii) z test about a mean with two-sided alternative (iv) t test about a mean with one-sided alternative (v) t test about a mean with two-sided alternative (vi) two-sample t test with one-sided alternative (vii) two-sample t test with two-sided alternative (viii) chi square test (ix) ANOVA (x) inference for regression

(f) We take random samples of married, single, and divorced workers to determine if mean earnings differ among these groups:

(i) z test about a proportion (ii) z test about a mean with one-sided alternative (iii) z test about a mean with two-sided alternative (iv) t test about a mean with one-sided alternative (v) t test about a mean with two-sided alternative (vi) two-sample t test with one-sided alternative (vii) two-sample t test with two-sided alternative (viii) chi square test (ix) ANOVA (x) inference for regression

(g) We want to test for a relationship between race and employment status (employed or not):

(i) z test about a proportion (ii) z test about a mean with one-sided alternative (iii) z test about a mean with two-sided alternative (iv) t test about a mean with

one-sided alternative (v) t test about a mean with two-sided alternative (vi) two-sample t test with one-sided alternative (vii) two-sample t test with two-sided alternative (viii) chi square test (ix) ANOVA (x) inference for regression

5. (25 pts.) Below are a scatterplot and regression output, where the data consist of average viewer ratings (on a scale of 1 to 10) and gross receipts (in millions of dollars) of 15 recent movies:

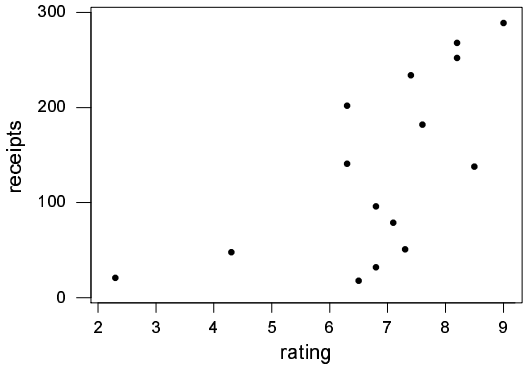
   (a) Considering the appearance of the scatterplot, which of the following is the most reasonable guess for the correlation r?
   (i) -.95 (ii) -.6 (iii) -.1 (iv) +.1 (v) +.6 (vi) +.95

   (b) Use the regression equation to predict the receipts of a movie rated 9 by viewers. UNDERLINE the part of the output that estimates about how far off this prediction would be.

   (c) According to the regression output, is there statistical evidence of a relationship between ratings and receipts? Answer yes or no and CIRCLE THE SPECIFIC PART OF THE OUTPUT THAT YOU USE TO DECIDE.

   (d) Which would be of more interest to the producer of a new movie rated 9 by viewers? (i) a confidence interval for the mean gross receipts of all movies rated 9 by viewers (ii) a prediction interval for the gross receipts of a single movie rated 9 by viewers

   (e) Which of the intervals in (d) would be wider?
   (i) confidence interval (ii) prediction interval (iii) both the same




```
Regression Analysis: receipts versus rating
The regression equation is
receipts = - 113 + 36.5 rating

Predictor        Coef     SE Coef          T        P
Constant      -112.60       84.37      -1.33    0.205
rating          36.45       12.00       3.04    0.010

S = 75.93       R-Sq = 41.5%     R-Sq(adj) = 37.0%
```

6. (40 pts.) This table from 1995 shows layoffs by ethnic group for U.S. Labor Department employees. We want to use a chi-square test to determine if there is convincing evidence of a relationship between race and layoffs.

Observed	Laid Off	Not Laid Off	Total
African American	130	1380	1510
White	90	2810	2900
Total	220	4190	4410

(a) The data were probably obtained from (i) an observational study (ii) an experiment (iii) it could easily have been either (i) or (ii).

(b) State the appropriate null and alternative hypotheses.

(c) In the "*Observed*" table, who has a higher proportion of layoffs? (i) African Americans (ii) whites (iii) both the same.

(d) If there were no relationship between race and layoffs, what counts would be expected in each cell? Display them in the table provided.

Expected	Laid Off	Not Laid Off
African American		
White		

(e) Calculate the value X^2 of the chi-square statistic.

(f) How many degrees of freedom are there?

(g) Use Table A.5 to give a range for the P-value.

(h) Draw your conclusions (circle one of the following):

i. The P-value is small, providing evidence of a relationship.

ii. The P-value is **not** small, providing evidence of a relationship.

iii. The P-value is small, providing **no** evidence of a relationship.

iv. The P-value is **not** small, providing **no** evidence of a relationship.

7. (5 pts.) A biologist observes a sample of 81 cells to have mean dividing time 30.5 minutes and standard deviation 4.9 minutes. Construct a 99% confidence interval for population mean dividing time.

8. (25 pts.) Scientists believe people's ears get larger with age. They measured ear length in hundreds of patients, aged 30 to 93, and concluded that ears grow an average of .01 inches a year.

   (a) Was this an experiment or an observational study?

   (b) What is the explanatory variable? Is it quantitative or categorical?

   (c) What is the response variable? Is it quantitative or categorical?

   (d) According to our notation, .01 is denoted by which of the following?
   $\mu, \sigma, p, \beta_0, \beta_1, \chi^2, n, \bar{x}, s, \hat{p}, b_o, b_1, X^2, \alpha$

   (e) A Chinese physician recalled his mother's childhood nagging: "Stretch your ears daily, child, to ensure long life." In fact, some scientists believe that men with small ears die younger, leaving a population of healthier old people with big ears. What would be the explanatory variable, according to this theory?

9. (20 pts.) For several years, a mathematics placement test has been administered to all incoming freshmen at a certain college. A random sample of 16 students selected over this time period took on the average $\bar{x} = 51.5$ minutes to complete the test, with a standard deviation $s = 3.0$ minutes. Those in charge of administering the test claim the average time required is no more than 50 minutes, but students suspect the average time is in fact more than 50 minutes.

   (a) Test at the $\alpha = .05$ level, making sure to identify the appropriate hypotheses, test statistic, range for the P-value, and conclusion: is there compelling evidence that the mean time exceeds 50 minutes? (yes or no)

   (b) If a test were made against a **two-sided** alternative, then the P-value would lie between _____ and _____.

   (c) Suppose the P-value for the one-sided test is calculated to be 0.03. Which one of the following is correct?

      i. The probability that the test requires on the average 50 minutes is no more than 3%.

      ii. The probability that the test takes longer than 50 minutes is 3%.

      iii. If the average time required to take the test is 50 minutes, then the probability of the sample producing an average time **at least as long as** 51.5 minutes is 3%.

      iv. If the average time required to take the test is 50 minutes, then the probability of the sample producing an average time of 51.5 minutes is **at least** 3%.

   (d) If there is statistical evidence that the test requires more than 50 minutes, then the administration must go through a great deal of trouble to revise the test. Students, on the other hand, could benefit from having them create a shorter test.

      i. Who would be most interested in avoiding a Type I error, administrators or students?

ii. Who would be most interested in avoiding a Type II error, administrators or students?

10. (25 pts.) Researchers at the University of California Earthquake Center want to compare intensities of earthquakes at San Francisco Bay and at the Los Angeles Basin. The data, assumed to be independent random samples from the two areas, represent intensity on the Richter scale of tremors felt during several months in 1994.

    (a) If researchers want to test if mean intensities differ between the two regions, state the appropriate null and alternative hypotheses.

    (b) Use the two-sample t statistic in the output, and Table A.2, to report a range for the P-value.

    (c) Is there statistically significant evidence of a difference?

    (d) Are your conclusions still valid if one of the data sets is very skewed?

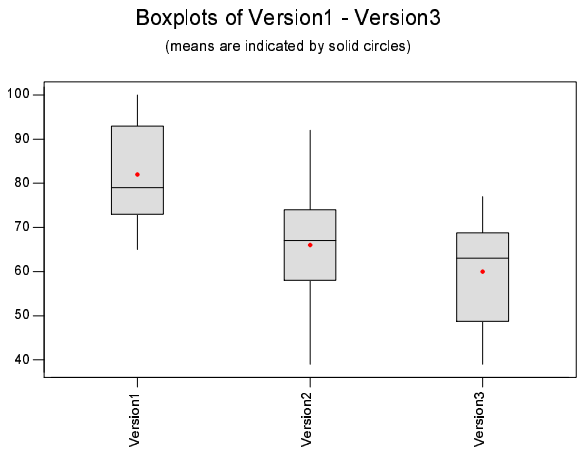
    (e) Circle the part of the output that would help you decide if the Rule of Thumb for use of a pooled procedure is satisfied.

```
Two-sample T for SF vs LA

                N      Mean       StDev     SE Mean
SF             14      2.59        1.02        0.27
LA             16      2.13        1.09        0.27

Difference = mu SF - mu LA
Estimate for difference: 0.454
95% CI for difference: (-0.339, 1.248)
T-Test of difference = 0 (vs not =): T-Value = 1.18
```

11. (10 pts.) Below are boxplots for students' scores on 3 different versions of an exam.

    (a) Which version appears to have the highest mean score?

    (b) Which version appears to have the largest standard deviation?


Boxplots of Version1 - Version3
(means are indicated by solid circles)
100
90
80
70
60
50
40
Version1
Version2
Version3

12. (40 pts.) Three groups of six guinea pigs each were injected, respectively, with small, medium, and large doses of a new tranquilizer, and the following are data for the number of minutes X_i (for $i = 1, 2, 3$) it took them to fall asleep:

$$\begin{array}{rrrr} \text{small:} & n_1 = 6 & \bar{x}_1 = 22.5, & s_1 = 2.168 \\ \text{medium:} & n_2 = 6 & \bar{x}_2 = 20.0, & s_2 = 1.414 \\ \text{large:} & n_3 = 6 & \bar{x}_3 = 13.0, & s_3 = 2.098 \end{array}$$

(a) Is this study an experiment? Circle one of the following:

i. Yes, because more than one dosage level is tested.

ii. Yes, because there is a treatment involved (injecting with various amounts of tranquilizer) for the units.

iii. No, this is an observational study on a random sample of guinea pigs.

iv. No, because there is no control group of guinea pigs which receive no injection.

(b) Is it safe to assume that population standard deviations are equal (as required in an analysis of variance)? Answer yes or no, and circle the values above that you use to decide.

(c) State the appropriate null and alternative hypotheses.

(d) Complete the ANOVA table below and test the null hypothesis of ANOVA at the 5% level of significance. Do you conclude that the mean time required to fall asleep is the same for all three dosage levels? (Answer yes or no.)

Source	DF	Sum of Squares	Mean Square	F	P-value
Group		291.0			
Error		55.5		XXXXXXX	XXXXXXX
Total	XX	XXXXX	XXXXXXX	XXXXXXX	XXXXXXX

(e) Suppose the 18 guinea pigs used in this study have already been randomly selected and must now be assigned randomly to treatments. If they are given number labels "1" through "18", use the following lines of random digits to select which guinea pigs should go into the first treatment group: [Your answer should be a list of 6 numbers.]

14459 26056 31424 80371 65103 62253 50490 61181

38167 98532 62183 70632 23417 26185 41448 75532

(f) Which is the best way to assign guinea pigs to treatments?

i. use random digits as above

ii. give small doses to the 6 least active guinea pigs, and large doses to the 6 most active

iii. give small doses to the 6 most active guinea pigs, and large doses to the 6 least active

13. (30 pts.) In a test for Extra Sensory Perception, a subject chooses each time from 4 images the one that is supposed to be "sent" to him telepathically. Out of 60 questions, he gets 18 correct.

    (a) If he is really only guessing, then the **sample count** X of correct answers is a binomial random variable with $n = 60$ and $p = .25$. Its mean is _____and standard deviation is _____.

    (b) Using a normal approximation, the probability of scoring at least 18 is _____.

    (c) To test if the subject has ESP, we can test if his proportion of correct answers is significantly higher than it would be if he were only guessing. Under the null hypothesis of random guessing, the overall proportion of correct answers is $p = .25$. For a sample of 60 questions, sample proportion $\hat{p}$ would have mean _____and standard deviation _____.

    (d) The p-value is the probability of observing a sample proportion $\hat{p}$ at least as high as 18 out of 60. Use Table A to find the p-value.

    (e) Is there statistical evidence that the subject has ESP?

    (f) Construct a 90% confidence interval for his long-run proportion of correct answers, based on the sample proportion correct.

14. (20 pts.) Recently, warning labels covering 30% of each cigarette pack were made mandatory in all European Union countries. Although the U.S. pioneered tobacco warning labels in 1965, it has not upgraded its cautions since 1984; they cover less than 20% of a pack, and are fairly unobtrusive. In Canada, warnings on cigarette labels cover 50% of the front and back of each pack, and include full-color pictures of organs ravaged by smoking. Researchers would like to know the impact of the various labels on people's smoking habits.

    (a) What would be the most obvious pitfall in a study which would compare percentages smoking in Europe, the U.S., and Canada in order to determine the impact of the 3 types of label? (i) confounding variables (ii) placebo effect (iii) volunteer bias (iv) response bias (v) non-response bias

    (b) Suppose smokers in Europe, the U.S., and Canada were asked to rate the impact of the labels on their desire to smoke, from 0 = "no impact" to 10 = "they convinced me to quit". Would this be an experiment or an observational study? To test for a significant difference among impacts depending on whether a smoker was exposed to the labels in Europe, the U.S., or Canada, researchers should use (i) z test about a proportion (ii) paired t (iii) two-sample t (iv) ANOVA (v) regression (vi) chi-squared

    (c) In a random sample of 400 Canadian smokers, 44% said the new warnings increased their desire to quit. Construct a 95% confidence interval based on these results. which one of the following is a correct interpretation of your interval?

        i. 95% is the probability that the confidence interval produced from this sample contains the proportion of all Canadian smokers whose desire to quit would be increased by the warnings.

ii. 95% is the probability that the proportion of all Canadian smokers whose desire to quit would be increased by the warnings falls in the interval.

iii. 95% is the probability that the confidence interval produced from this sample contains the proportion of all smokers whose desire to quit would be increased by the warnings.

iv. 95% is the probability that the proportion of all smokers whose desire to quit would be increased by the warnings falls in the interval.

(d) Based on your confidence interval in (c), what would be the outcome of a test of $H_o : p = .5$ vs $H_a : p < .5$ where p refers to the population proportion whose desire to quit would be increased by the warnings?