Math 10
Spring 99
Practice Final Exam

Disclaimer: This set of problems is meant neither to indicate th e length nor composition of the actual exam. Many are taken from earlier exams or practice exams. However, they may give an indication of the type of problems which will appear on the exam.

On problems 1-17, circle the letter of the most appropriate response.

1. You read in an article by a statistician that the SAT scores in high school explain only 9% of the variation in students' later grades in college. From this you conclude that the correlation between SAT scores and college grades is
   (a) 0.0081
   (b) 0.09
   (c) .81
   (d) 0.3
   (e) 0.03

2. A study of two methods of teaching reading is carried out in a large elementary school. Third graders at this school are divided into two groups at random. One group is taught by Method A; the other by Method B. At the end of the shcool year, all of the children are given a standard test of reading ability. The response variable in this study is
   (a) the reading test scores
   (b) the teaching method used
   (c) the third graders
   (d) the two groups of students
   (e) the differences in test scores between students in the two groups

3. Use the information in Problem 2. The design of the study is
   (a) a completely randomized design
   (b) a randomized block experiment
   (c) a simple random sample
   (d) a stratified random sample
   (e) not an experiment because there are no controls

4. Use the information in problem 2. To see if the study gives evidence of a difference in the effectiveness of the two teaching methods, you would test hypotheses of the form
   (a) $H_0 : \mu_1 = \mu_2$ against $H_a : \mu_1 > \mu_2$
   (b) $H_0 : \mu_1 = \mu_2$ against $H_a : \mu_1 < \mu_2$
   (c) $H_0 : \mu = 0$ against $H_a : \mu \neq 0$
   (d) $H_0 : \mu_1 = \mu_2$ against $H_a : \mu_1 \neq \mu_2$
   (e) $H_0 : \mu = 0$ against $H_a : \mu > 0$

5. Use the information in problem 2. Of the tests of significance you have learned, the most appropriate for analyzing the results of this study is
   (a) two-sample t test
   (b) matched pairs t test
   (c) single sample t test
   (d) two sample t test for proportions
   (e) none of the above because a z test is more appropriate

6. A news report says that a national opinion poll of 1500 randomly selected adults found that 43% thought that they would be worse off economically during the next year. The news report went on to say that the margin of error in the poll result was $\pm 3\%$ with 95% confidence. This margin of error does *not* include errors due to
   (a) the fact that the poll dialed telephone numbers at random and so missed all people without phones
   (b) the fact that the poll could not contact some of the people whose numbers were chosen
   (c) chance variation in the random selection
   (d) both (a) and (b)
   (e) all of (a), (b), and (c)

7. Use the information from problem 6. A 90% confidence interval based on the poll results would have a margin of error
   (a) less than ± 3 percentage points.
   (b) more than ± 3 percentage points.
   (c) about equal to ± 3 percentage points
   (d) that could be any of (a) , (b), and (c)
   (e) that can't be described by any of the choices above

8. Use the information from problem 6. If the poll had interviewed 1000 persons rather than 1500 (and still found 43% believing they would be worse off), then for 95% confidence, the poll would have a margin of error
   (a) less than ± 3 percentage points.
   (b) more than ± 3 percentage points.
   (c) about equal to ± 3 percentage points
   (d) that could be any of (a) , (b), and (c)
   (e) that can't be described by any of the choices above

9. Use the information from problem 6. If the poll had obtained the outcome 43% by a similar random sampling method from all adults in new York State (population about 20 million) rather than from adults in the US (population about 275 million) then for 95% confidence for the opinion in the state of New York, the poll would have a margin of error
   (a) less than ± 3 percentage points.
   (b) more than ± 3 percentage points.
   (c) about equal to ± 3 percentage points
   (d) that could be any of (a) , (b), and (c)
   (e) that can't be described by any of the choices above

10. A study of blood types measures the blood type (A, B, AB, or O) of 10,000 people. The study includes people living in Florida, Minnesota, and California. The chi-square test is used to see if there are differences among the distributions of blood types in the three states. The degrees of freedom for this test are
   (a) 9999
   (b) 12
   (c) 6
   (d) 9998
   (e) 2

11. In a test of anti-fungal activity of a chemical compound, fungus is grown in petri dishes with different concentrations of the compound and the diameter of the fungus colonies is measured after one day. There are 20 dishes, two at each of 10 concentrations. A plot of diameter against concentration seems to show a straight line pattern. Least squares regression is used to analyze the data. What distribution is used in the test of the hypothesis that concentration has no effect on diameter?
   (a) t(18)
   (b) t(9)
   (c) t(19)
   (d) t(2)
   (e) t(8)

12. A certain binomial distribution consists of three independent trials, each with probability $\frac{1}{3}$ of success. The mean of this binomial distribution is
   (a) 1
   (b) 1/3
   (c) 3
   (d) 1.5
   (e) not possible to compute from the given information

13. Refer to the previous problem. The standard deviation for this probability distribution is
    (a) $\sqrt{2}/3$
    (b) $\sqrt{3}/2$
    (c) $\sqrt{6}/3$
    (d) $\sqrt{6}/2$
    (e) $\sqrt{6}/9$

14. Refer to problem 13. The probability of exactly one success is
    (a) 1/27
    (b) 4/27
    (c) 1/9
    (d) 4/9
    (e) 1/3

15. For a normal distribution
    (a) The interquartile range is larger than the standard deviation
    (b) the standard deviation is larger than the interquartile range
    (c) the mean equals the median
    (d) a and c
    (e) b and c

16. A company has ten employees, including the owner. The sum of the salaries of the 9 non-owners is less than the salary of the owner. How many employees earn less than the mean salary for the company and why?

17. Give a list of four distinct integers between one and ten (inclusive) with the largest possible standard deviation.

18. Give two different lists of four distinct integers between one and ten (inclusive) with the smallest possible standard deviation.

19. When we use a t distribution in computing a confidence interval for the mean of a sample, what is the *primary* reason we use a t distribution instead of a normal distribution? (A one sentence answer is preferred.)

20. A random variable is normally distributed with mean 4 and standard deviation 2. What proportion of the values of this random variable are greater than 7?

21. Suppose we have a result of a t-test or z-test that is significant at the 5% level.
    (a) Is it also significant at the 10% level?

    (b) Is it also significant at the 1% level?

22. A student at a big university reads in the student newspaper that "according to a poll of a simple random sample of 1100 students, we can be better than 95% sure that 55% of the students, plus or minus 3%, favor increasing the hours the student cafeteria is open, even if this means increasing the cost of a meal plan." Explain why, by polling 1100 students, the newspaper can be better than 95% sure that its percentage is correct within a margin of error of plus or minus 3%.

23. A lauguage department has a basic list of 2500 vocabulary words it wants its majors to know. Among the requirements for graduation is taking a test of translation knowledge on vocabulary words and passing it. The department has decided to hold its students to a very high standard on this test. In particular, they want to be quite sure that a student who knows 90% or more of the vocabulary words will pass the test and they also want to be quite sure that a student who knows 80% or fewer of the vocabulary words will not pass the test.
    (a) If the test consists of an independent random sample of 100 vocabulary words for the student to translate, and a score of 85 is the minimum passing score, what is the probability that someone who knows 90% of the words on the basic list will pass the test?

(b) If the test consists of an independent random sample of 100 vocabulary words for the student to translate, and a score of 85 is the minimum passing score, what is the probability that someone who knows 80% of the words on the basic list will fail the test?

(c) How many questions must they have on the test and what must the minimum passing score be for them to be at least 99.9% sure that someone who knows 90% of the vocabulary words will pass the test and at least 99.9% sure that someone who knows 80% of the vocabulary words will not pass the test?

24. In the April 1997 issue, the magazine *Consumer Reports* reported on many statistics, including weight (Wt, in pounds), horsepower (HP), and gas mileage (in miles per gallon, MPG) for 114 passenger vehicles (cars, vans, and sport-utility vehicles). On the next several pages you will see JMP input and output regarding these three variables. Using this information and making any necessary computations, answer the following questions.
    (a) Describe any outliers in the three distributions.

    (b) Write down a regression equation that could be used to predict MPG from weight.

    (c) Which variable is a better predictor of gasoline mileage: weight or horsepower? Why?

(d) In the MPG by wt regression, the slope is very close to zero, (-.005965). Is this slope significantly different from zero? (Explain.)

(e) What is the mean gasoline mileage you would expect in vehicles that have 200 horsepower engines?

(f) What is the variance of the horsepower distribution? Using this information, what is $\sum(x_i - \bar{x})^2$ for the horsepower distribution?

25. A major university classified all students entering PhD programs in a given year by their status 6 years later. The categories used were "completed the degree," "still enrolled," and "dropped out." Here are the data.

Status	Men	Women
completed	423	98
still enrolled	134	33
dropped out	238	98

Assume that these data can be viewed as a random sample of all students at similar universities in recent years.

(a) State an appropriate null hypothesis and alternative hypothesis that addresses the question of sex differences in perseverance in graduate education.

(b) Compute an appropriate statistic and use it to test your hypotheses in the previous part.

(c) What conclusion do you draw from your analysis? Discuss possible factors not mentioned above that might be relevant to interpreting the results of the study.

26. In a study of heart surgery, one issue was the effect of drugs called beta-blockers on the pulse rate of patients during surgery. The available subjects were divided at random into two groups of 30 patients each. One group received a beta-blocker; the other a placebo. The pulse rate of each patient at a critical point during the operation was recorded. The treatment group had mean pulse rate 65.2 and standard deviation 7.8. For the control group, the mean was 70.3 and the standard deviation was 8.3.
    (a) State an appropriate null hypothesis and alternative hypothesis that are relevant to the question of whether beta-blockers reduce the pulse rate during surgery.

    (b) Test your hypotheses with an appropriate test. Do the results show a significant effect of beta-blockers at the 5% level? At the 1% level? If the result is significant, what does it say about the means?

    (c) Give a 95% confidence interval for the difference in mean pulse rates between the control group and the treatment group.