EPI 513-607A Mid-term Examination Oct. 23, 1989

Calculations need not be exact. Questions carry the weights indicated.

PART I - MULTIPLE CHOICE - 3 POINTS EACH

1. The heights of American men aged 18 to 24 are approximately normally distributed with mean 68 inches and standard deviation 2.5 inches. Half of all young men are shorter than
(a) 65.5 inches (b) 68 inches (c) 70.5 inches (d) can't tell, because the median height is not given

2. Use the information in the previous problem. Only about 5% of young men have heights outside the range
(a) 65.5 inches to 70.5 inches (b) 63 inches to 73 inches (c) 60.5 inches to 75.5 inches (d) 58 inches to 78 inches

3. Use the information in Problem 1. What percent of young men are taller than 72 inches?
(a) 94.5% (b) 44.5% (c) 5.5% (d) 2.5%

4. The central limit theorem states that
(a) the sample mean is unbiased
(b) the distribution of the sample mean is normal given sufficiently large n
(c) the binomial distribution is skewed
(d) the sample standard deviation is approximately normal
(e) none of the above

5. A significance test was performed to test the null hypothesis $H_o : \mu = 2$ versus the alternative $H_a : \mu \neq 2$. The test statistic is z = 1.40. The P-value for this test is approximately
(a) .16 (b) .08 (c) .003 (d) .92 (e) .70

6. A significance test gives a P-value of .04. From this we can
(a) reject H_o with $\alpha = .01$ (b) reject H_o with $\alpha = .05$
(c) say that the probability that H_o is false is .04 (d) say that the probability that H_o is true is .04
(e) none of the above

7. You want to compute a 90% confidence interval for the mean of a population with unknown population standard deviation. The sample size is 20. The number of SEM's you would use for this interval is
(a) 1.96 (b) 1.645 (c) 1.729 (d) 0.90 (e) 1.311

8. To use the two-sample t procedure to perform a significance test on the difference between two means, we assume
(a) the population standard deviations are known
(b) the samples from each population are independent
(c) the distributions are exactly normal in each population
(d) the sample sizes are large
(e) all of the above

PART II - PROBLEMS

9. Figure 1.21 displays three density curves with several points marked on each. At which of these points on each curve do the mean and the median fall?

10. Eleanor scores 680 on the mathematics part of the Scholastic Aptitude Test (SAT). The distribution of SAT scores in a reference population is normal with mean 500 and standard deviation 100. Gerald takes the American College Testing (ACT) mathematics test and scores 27, ACT scores are normally distributed with mean 18 and standard deviation 6. Find the standardized scores for both students. Assuming that both tests measure the same kind of ability, who has the higher score?

11. For each of the following situations, indicate whether a binomial distribution is a reasonable probability model for the random variable X. Give your reasons in each case.

(a) You observe the sex of the next 50 children born at a local hospital; X is the number of girls among them.

(b) A couple decides to continue to have children until their first girl is born; X is the total number of children the couple has.

(c) You want to know what percent of married people believe that mothers of young children should not be employed outside the home. You plan to interview 50 people, and for the sake of convenience you decide to interview both the husband and the wife in 25 married couples. The random variable X is the number among the 50 persons interviewed who think mothers should not be employed.

(d) The pool of potential jurors for a murder case contains 100 persons chosen at random from the adult residents of a large city. Each person in the pool is asked whether he or she opposes the death penalty; X is the number who say "Yes"

12. A university that is better known for its basketball program than for its academic strength claims that 80% of its basketball players get degrees. An investigation examines the fate of all 20 players who entered the program over a period of several years that ended 5 years ago. Of these players, 11 graduated and the remaining 9 are no longer in school. If the university's claim is true, the number of players who graduate among the 20 studied should have a _____________ distribution.

(a) Find the probability that exactly 11 players graduate under these assumptions.

(b) Find the probability that 11 or fewer players graduate. Is this probability so small that it casts doubt on the university's claim?

(c) State the formal statistical hypotheses inherent in (1) the university's claim and (2) the investigation's claim.

13. A social psychologist reports that "in our sample, ethnocentrism was significantly higher ($P < 0.05$) among church attenders than among nonattenders." Explain what this means in language understandable to someone who knows no statistics. Do not use the word "significance" in you answer.

14. A government report gives a 99% confidence interval for the 1985 median family income as \$27,735 ± \$357. This result was calculated by advanced methods from the Current Population Survey, a multistage random sample of over 70,000 households.

(a) Would a 95% confidence interval be wider or narrower? Explain you answer.

(b) Would the null hypothesis that the 1985 median family income was \$28,000 be rejected at the 1% significance level in favor of the two-sided alternative? Why?

15. A bank wonders whether omitting the annual credit card fee would increase the amount charged on its credit card. The bank makes this offer to a random sample of 200 of its existing credit card customers. It then compares how much these customers charge this year with the amount that they charged last year. The mean increase is \$332, and the standard deviation is \$108.

(a) What is the appropriate test to assess whether there is significant evidence at the 1% level that the mean amount charged increases under the no-fee offer? State H_o and H_a.

(b) Give a 99% confidence interval for the mean amount of the increase.

(c) The distribution of the amount charged is skewed to the right, but outliers are prevented by the credit limit that the bank enforces on each card. Use of the statistical procedures is justified in this case even thought the population distribution is not normal. Explain why.

(d) Give one methodologic criticism of the design of the study. Briefly describe the design of an experiment to study the effect of the no-fee offer that would avoid this criticism.

16. It is claimed that a new filter for filter-tipped cigarettes leaves less nicotine in the smoke than does the current filter. Because cigarette brands differ in a number of ways, each filter is tested on one cigarette of each of nine brands. The difference between the nicotine content for the current filter and the new filter is recorded. The mean difference is x-bar = 1.20 milligrams (mg), and the standard deviation of the difference is s = 2.40 mg.

(a) What test is appropriate to calculate how significant is the observed difference in means? State H_o and H_a.

(b) Give a 90% confidence interval for the mean amount of additional nicotine removed by the new filter.

17. Do various occupational groups differ in their diets? In a British study of this question, two of the groups compared were 100 drivers and 81 conductors of London double-decker buses. The ticket takers' jobs require more physical activity. The article reporting the study gives the data as "Mean daily consumption (± s.e.)." Some of the study results appear below:

	Drivers	Ticket Takers
Total calories	2821 ± 44	2844 ± 48
Alcohol	.24 ± .06	.39 ± .11

(a) What does "s..e." stand for? Give x-bar and s for one of the four sets of measurements.

(b) What formal statistical test is appropriate to assess if there is significant evidence at the 5% level that ticket takers consume more calories per day than do drivers?

(c) What do you conclude on the basis of a quick "visual" test?

18. A study of rush hour traffic in Montreal records the number X of people in each car crossing the Champlain Bridge. Suppose that this number X has mean 1.5 and standard deviation 0.75 in the population of all cars that cross the bridge during rush hours.

(a) Does the count X have a binomial distribution? Why or why not?

(b) Could the exact distribution of X be normal? Why or why not?

(c) Traffic engineers estimate that the capacity of the bridge is 700 cars per hour. According to the central limit theorem, what is the approximate distribution of the mean number of persons (x-bar) in 700 randomly selected cars crossing the bridge?

(d) The count of people in 700 cars is 700 times x-bar. Use your result from (c) to give an approximate distribution for the count. What is the probability that 700 cars will carry more than 1075 people?