

**22S:008 Final Exam, May 12, 1999 Name:** \_\_\_\_\_

**Please circle one.** Section: 11:30, 12:30, or 2:30

Please enter all of your answers on these exam pages. Circle the correct answers for the true-false and multiple choice questions. There are 40 questions. The Defective Question Report will be supplied separately.

1. (2 pts.) A **statistic** is a numerical characteristic of a sample.  
A) True B) False
2. (2 pts.) **Special causes** are individually small, unobserved influences on process results.  
A) True B) False
3. (2 pts.) A Binomial distribution has its **smallest variability** when  $\pi = 0.5$ .  
A) True B) False
4. (2 pts.) The results produced by a constant-cause system vary, and may vary over a wide band or a narrow band.  
A) True B) False
5. (2 pts.) A **Type I error** is committed when you fail to reject a false null hypothesis.  
A) True B) False
6. (2 pts.) Least squares chooses values for the regression coefficients,  $b_0, b_1, \dots, b_k$ , that minimize the sum of the squared deviations between the observed response values and the fitted regression equation.  
A) True B) False
7. (2 pts.) The units (dollars, pounds, inches, etc.) for the residual standard deviation are the same as the units for the response variable  $y$ .  
A) True B) False
8. (2 pts.) A probability survey uses double blinding to prevent bias in survey results.  
A) True B) False
9. (2 pts.) In a designed experiment, treatments are randomly assigned to the study units (people or things). The reason for the randomization is to prevent bias due to other factors which are not controlled in the experiment.  
A) True B) False
10. (2 pts.) A histogram with too many class intervals may fail to show the general shape of a distribution.  
A) True B) False
11. (2 pts.) Tampering with a process that is in statistical control will usually reduce the variability in the process.  
A) True B) False
12. (2 pts.) A market researcher computed a confidence interval for a population proportion using a 95% confidence level. Her boss decided that she wanted a 99% confidence level instead. True or false: The new interval with 99% confidence level will be **wider** than the original one with a 95% confidence level.  
A) True B) False

13. (2 pts.) The ***p*-value** of a statistical test is the chance of getting an outcome at least as extreme as the outcome actually observed. The *p*-value is computed under the assumption that the null hypothesis is true.
- A) True    B) False
14. (2 pts.) Which of the following statistics are sensitive to outliers?
- I. the correlation coefficient  
II. the standard deviation  
III. the mean
- A) I only  
B) II only  
C) III only  
D) I, II, and III  
E) None of A, B, C, or D is the correct answer.
15. (2 pts.) Suppose *y* is a response variable representing the market value of a house, *x* is a continuous predictor variable representing the size of the house in thousands of square feet, and *z* is a binary indicator variable indicating whether or not the house is located in a prime location (*z* = 1 if prime location, *z* = 0 if not). Consider the regression model with fitted equation:  $\hat{y} = b_0 + b_1x + b_2z$ . Which of the following represents the predicted market value for a three thousand square foot house located in a prime location?
- A)  $3b_1$   
B)  $b_0 + 3b_1$   
C)  $3b_1 + b_2$   
D)  $b_0 + b_1 + b_2$   
E)  $b_0 + 3b_1 + b_2$
16. (2 pts.) When fitting a quadratic curve to 23 data points, the number of degrees of freedom for the residual standard deviation is
- A) 20  
B) 21  
C) 22  
D) 23  
E) None of the above.
17. (2 pts.) Consider a *t*-distribution with 20 degrees of freedom. Which of the following is the value of the 95th percentile, that is, the *t* value below which 95% of the total area lies or  $t_{.95}$ ?
- A) 1.325  
B) 1.725  
C) 1.729  
D) 2.086  
E) 2.093

18. (2 pts.) Consider 18 Bernoulli trials with success probability  $\pi = 0.2$ . What is the chance of observing **at most one** success?

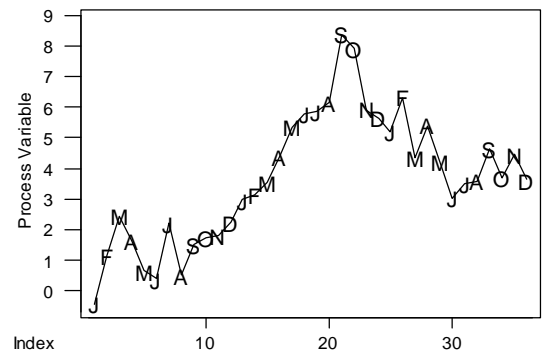
- A)  $18(0.2)^1(0.8)^{17}$
- B)  $(0.8)^{18}$
- C)  $1 - (0.8)^{18}$
- D)  $1 - 18(0.2)^1(0.8)^{17}$
- E)  $(0.8)^{18} + 18(0.2)^1(0.8)^{17}$

19. (2 pts.) A class has asked their instructor to “grade on the curve.” With this system the instructor is required to give preselected percentages of the various possible grades. In particular, the lowest 10% of the class **must receive F**’s. If exam scores are normally distributed with mean 81.4 and standard deviation 5, what exam score corresponds to the highest F grade?

- A) 55
- B) 60
- C) 65
- D) 70
- E) 75

20. (2 pts.) A monthly process is observed for three years. It’s sequence plot is displayed at the right. In the plot J=January, F=February, etc. True or false? This series is more **random** than **meandering**.

- A) True    B) False

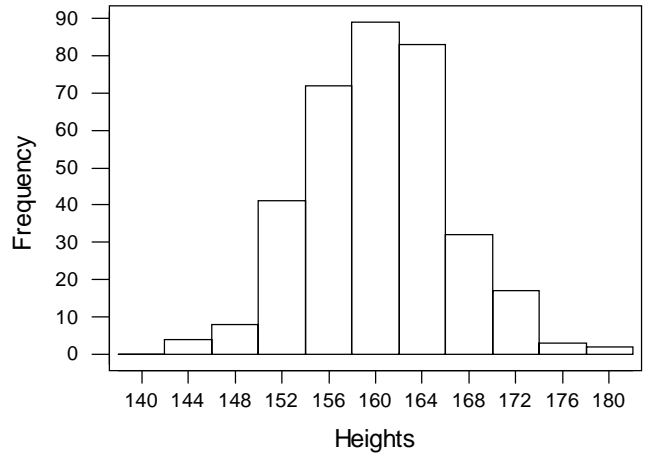


21. (2 pts.) We have observed 9 values in a random sample from a normal population. We obtained a sample mean of 100.5 and sample standard deviation of 0.23. We need to estimate the mean of the population. Which of the following will give the value of the upper endpoint of a 95% confidence interval for the population mean?

- A)  $100.5 + 2.306 \times (0.23)$
- B)  $100.5 + 1.860 \times (0.23)$
- C)  $100.5 + 1.860 \times (0.23 / (\sqrt{9}))$
- D)  $100.5 + 2 \times (0.23 / (\sqrt{8}))$
- E)  $100.5 + 2.306 \times (0.23 / (\sqrt{9}))$

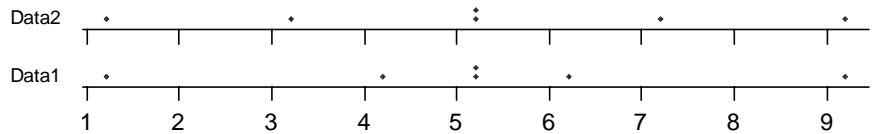
22. (2 pts.) Suppose that 30 subgroups (samples) each of size 50 are selected from a Bernoulli process. If there are 150 defectives overall, what is the value of the upper control limit (UCL) for the  $p$ -chart for fraction defective? (Rounded to the nearest hundredth.)
- A) 0.10  
 B) 0.13  
 C) 0.16  
 D) 0.23  
 E) 0.26

23. (2 pts.) The histogram at the right displays the distribution of heights (in cm.) of 351 women. Which of the following values is closest to the standard deviation in this distribution?



- A) 1  
 B) 3  
 C) 6  
 D) 9  
 E) 12

24. (2 pts.) Dotplots for two datasets are shown at the right. Do not calculate anything!



True or false? Dataset 1 has a larger mean than dataset2.

- A) True B) False

25. (2 pts.) Refer to the previous question. True or false? Dataset 1 has a larger median than dataset 2.

- A) True B) False

26. (2 pts.) Refer to the previous question. True or false? Dataset 2 has a larger standard deviation than dataset 1.

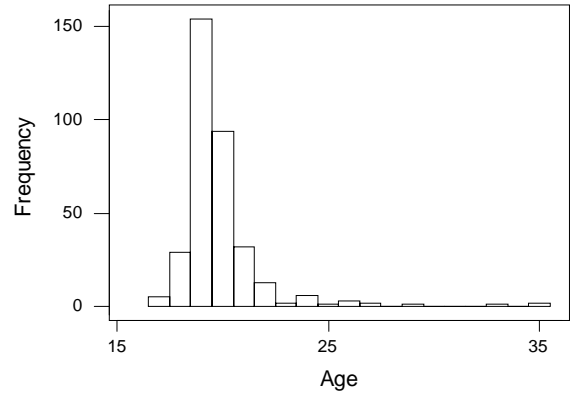
- A) True B) False

27. (2 pts.) A taste test of the new Pepsi One was performed in the following way. Two glasses were tasted. One contained Diet Pepsi and the other contained the new one-calorie Pepsi One. The glasses looked indistinguishable to the tasters. Twenty-five subjects were asked to identify which glass contained Pepsi One. Let  $\pi$  denote the chance that a randomly chosen person can identify the Pepsi One. Which one of the following is the most appropriate *null* hypothesis?

- A)  $\pi = 1/3$   
 B)  $\pi = 1/2$   
 C)  $\pi > 1/3$   
 D)  $\pi > 1/2$   
 E)  $\pi = 1/25$

28. (2 pts.) Here is a histogram of the distribution of ages in a large class of 345 students. Age is recorded as age at last birthday. About how many 19-year-olds were in this class?

- A) less than 30
- B) 30
- C) 40
- D) 90
- E) 150



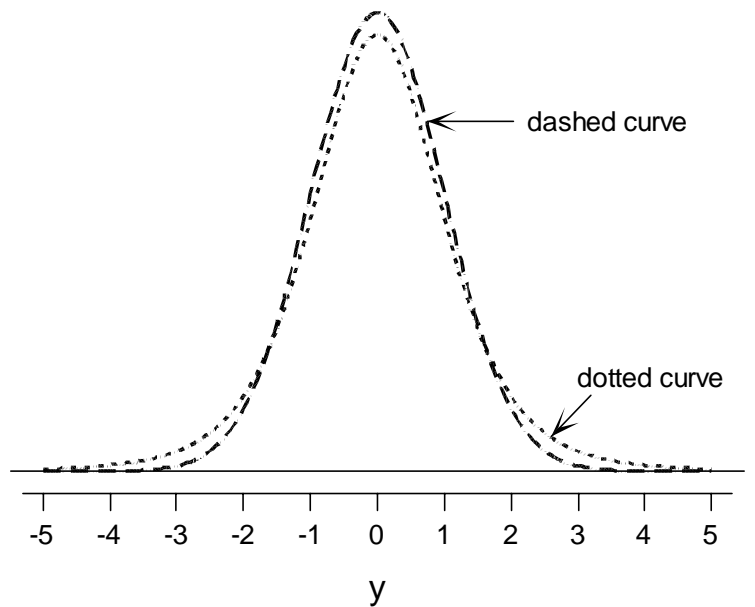
29. (2 pts.) Refer to the previous question. How young were the youngest students?

- A) 16
- B) 17
- C) 18
- D) 19
- E) None of the above.

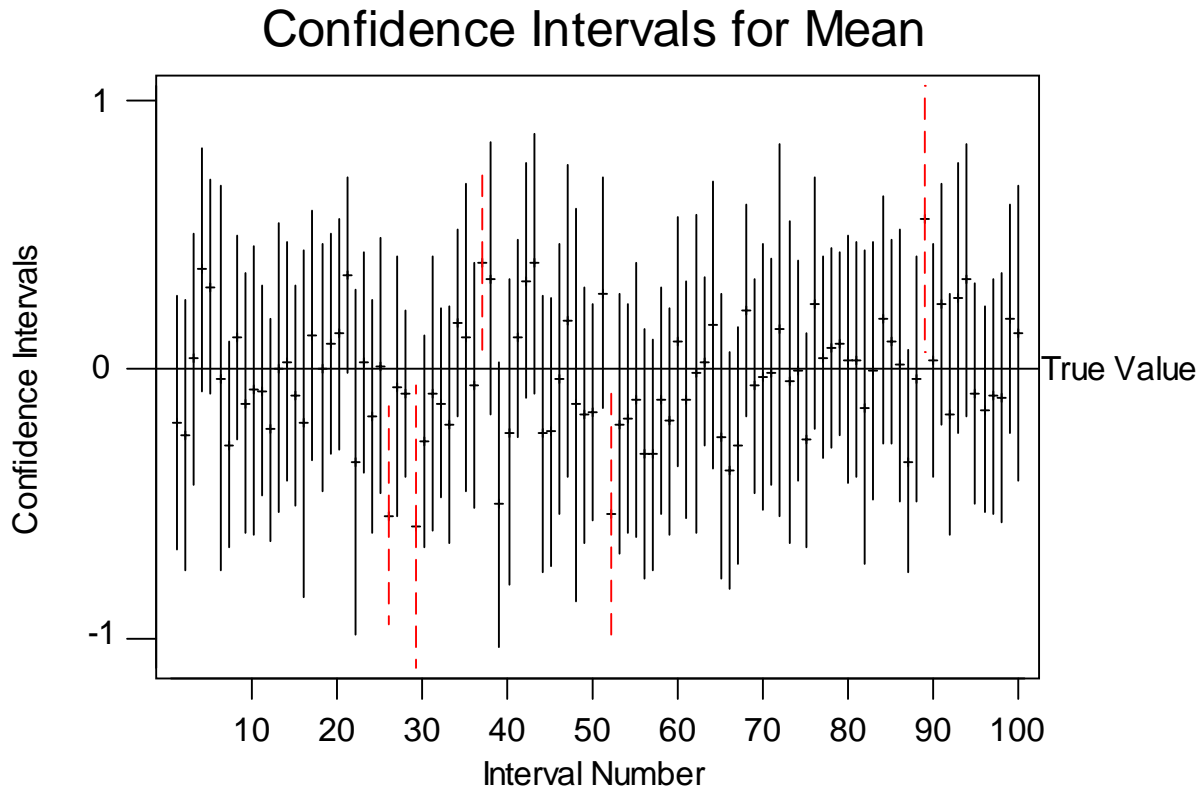
30. (2 pts.) Two theoretical density curves are shown to the right. One is a standard normal curve and the other is a  $t$  distribution with 5 degrees of freedom.

True or false? The dashed curve is the standard normal density.

- A) True
- B) False



31. (2 pts.) George invented a new way to form a confidence interval for a population mean but his theory is a little shaky. To test his method he used simulation to obtain 100 confidence intervals each based on samples of size 20. In his simulation the true population mean was 0. The 100 confidence intervals are displayed below. What *confidence level* do his intervals have (assuming this simulation represents “the long run”)?



- A) 5%  
 B) 10%  
 C) 90%  
 D) 95%  
 E) 99%
32. (5 pts.) The time it takes for an automatic car wash to wash a car varies according to a number of factors—is it a basic wash, a deluxe wash and wax, and so forth. Suppose that the distribution of individual wash times has a mean of 5 minutes and a standard deviation of 20 seconds. You are in line behind 9 cars waiting for a car wash. The first car begins its wash. What are the chances that you will have to wait more than 40 minutes before starting your car wash? Show your work.

Ans \_\_\_\_\_

33. (3 pts.) A model of the form  $y = b_0 + b_1x + b_2z$  has been used to model  $y$ , city gallons per 100 miles, where  $x$  is the weight of the car in pounds and  $z$  is an indicator variable for transmission type (1 = automatic, 0 = manual). The least squares estimates are  $b_0 = -1.3$ ,  $b_1 = 0.0025$ , and  $b_2 = 0.7$ . What does the model predict for city gallons per 100 miles for a car that weighs 2500 pounds and has a manual transmission? (Show your work.)

Ans \_\_\_\_\_

34. (2 pts.) Refer to the previous question. If we put 200 pounds of sand in the trunk of a car of any weight, how many **additional** gallons of gas does this model predict it will take to drive that car 100 miles in city driving? (Show your work.)

Ans \_\_\_\_\_

35. (5 pts.) You observe 100 items from a Bernoulli process with  $\pi = 0.1$ . **Approximate** the probability of observing **21 or more** successes. (You do **not** need to use the continuity correction. Show your work.)

Ans \_\_\_\_\_

36. (5 pts.) You want to estimate the percentage of young adults in the Midwest that would buy your new, innovative product. You decide that you want your estimate to have a margin of sampling error of plus or minus **one** percentage point with 95% confidence. Based on a simple random sample, what sample size is required to achieve this accuracy? (Use the conservative approach. Show your work.)

Ans \_\_\_\_\_

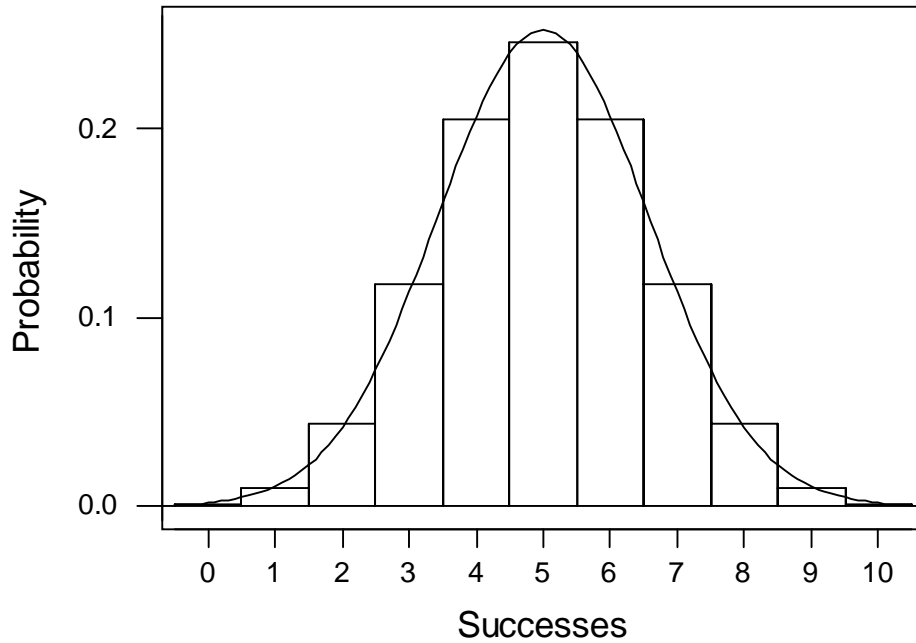
37. (5 pts.) You conduct a survey of your 100,000 employees using a simple random sample of size 500. Let  $\pi$  be the fraction of your employees who own stock in your major competitor. You find that 100 of the employees in the sample own stock in your major competitor. Find a 95% confidence interval for the fraction of all of your employees who own stock in your major competitor? Show your work.

Ans \_\_\_\_\_

38. (5 pts.) A nickel was spun on its edge 100 times. It fell heads up 65 times. Do these results give statistically significant evidence against the null hypothesis that the chance of heads is  $1/2$ ? (Use a **two-sided** test with a significance level of 5% and use a normal approximation to the appropriate sampling distribution. Show your work.)

Your decision or conclusion concerning  $H_0$ : \_\_\_\_\_

39. (2 pts.) The graph below displays the Binomial distribution with 10 trials and success probability  $1/2$ . The corresponding normal curve is also shown. Shade the area that gives the **normal approximation** for the chance of 8 or more successes *when using the continuity correction*. (Think before you shade. Don't mess up the graph. You only get one chance.)



40. (2 pts.) Refer to the previous question. Cross-hatch the area that gives the chance of getting exactly 2 successes.

# Defective Question Report

Name: \_\_\_\_\_

Section: \_\_\_\_\_

If you believe that a test question is defective in some way, please list your complaint here. All complaints will be considered in our interpretation of the test results.

Question number: \_\_\_\_\_

Your complaint:

Question number: \_\_\_\_\_

Your complaint:

Question number: \_\_\_\_\_

Your complaint: