

Please circle one. Section: 10:30 or 11:30

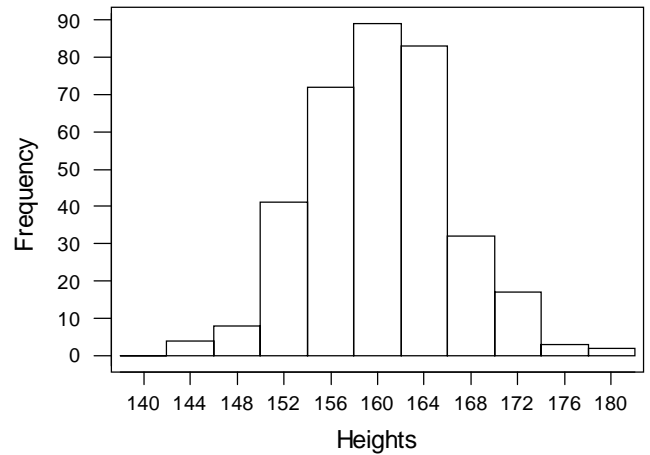
Please enter all of your answers on these exam pages. Circle the correct answers for the true-false and multiple choice questions. There are 29 questions plus one bonus question. The Defective Question Report is supplied separately.

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 temporary employee would be considered a special cause of variation.
A) True B) False
6. (2 pts.) A **Type I error** is failing to reject a false null hypothesis.
A) True B) False
7. (2 pts.) Least squares chooses values for the regression coefficients, b_0, b_1, \dots, b_k , that minimize the sum of the squared vertical deviations from the response values to the fitted regression equation.
A) True B) False
8. (2 pts.) The units (dollars, pounds,...) for the residual standard deviation are the same as the units for the response variable.
A) True B) False
9. A probability survey uses double blinding to prevent bias in survey results.
A) True B) False
10. (2 pts.) In fitting a straight line by least squares, the units (dollars, pounds,...) of the slope coefficient, b_1 , are the same as the units for the response variable.
A) True B) False
11. (2 pts.) With regard to process improvement: The Pareto Principle states that only a few causes are responsible for most of the variation.
A) True B) False
12. (2 pts.) A histogram with too many class intervals may fail to show the general shape of a distribution.
A) True B) False

13. (2 pts.) Tampering with a process that is in statistical control will usually reduce the variability in the process.
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. What 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 4 Bernoulli trials with success probability $\pi = 0.2$. What is the chance of observing **at most one** success?
- A) $4(0.2)^1(0.8)^3$
 - B) $(0.8)^4$
 - C) $1 - (0.8)^4$
 - D) $1 - 4(0.2)^1(0.8)^3$
 - E) $(0.8)^4 + 4(0.2)^1(0.8)^3$
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.) 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. What is 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) $100.5 + 2.306 \times (0.23 / (\sqrt{9}))$
21. (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

22. (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



23. (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.

24. (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. Based on a simple random sample, what sample size is required to achieve this accuracy? (Use the conservative approach. Show your work.)

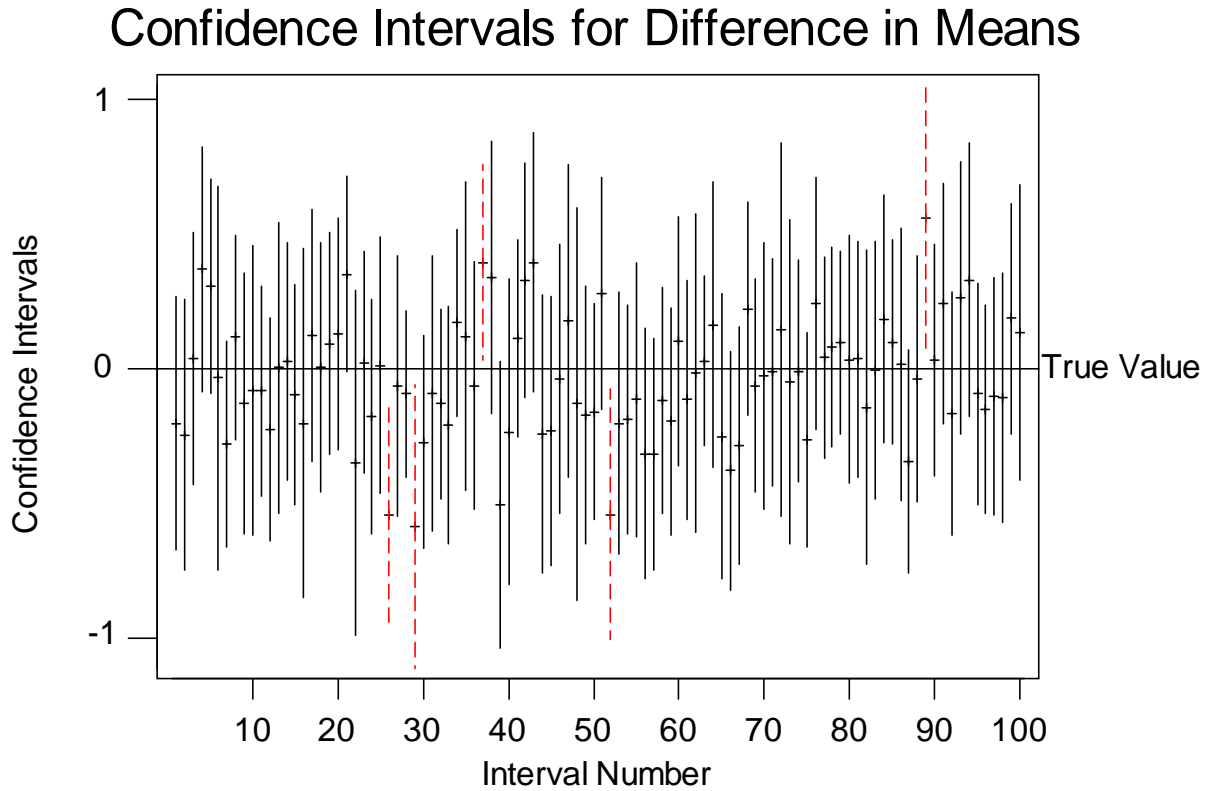
25. You conduct a survey of your 100,000 employees using a simple random sample of size 500. Let π be the fraction of your employees who own stock in your major competitor. You find that 100 of the employees in the sample own stock in your major competitor.

a) (5 pts.) Find a 95% confidence interval for the fraction of all of your employees who own stock in your major competitor? Show your work.

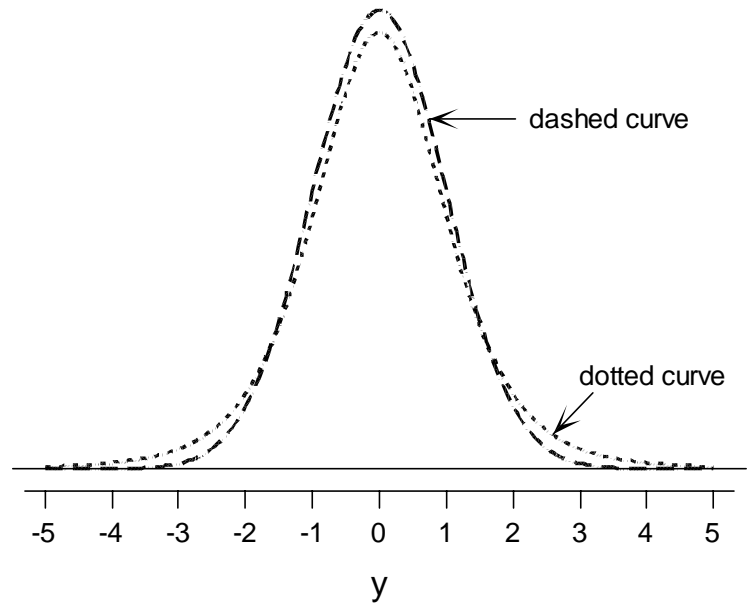
b) (2 pts.) Translate this confidence interval on π into an interval on the *number* of your employees who own stock in your competitor. Show your work.

26. (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 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.)

27. (5 pts.) George invented a new way to form confidence intervals for the difference in population means 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 difference in means was 0. The 100 confidence intervals are displayed below. What *confidence level* do his intervals have (assuming this simulation represents “the long run”)? Explain your reasoning.



28. (5 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. Which is which? Explain your reasoning.



29. In class we used a paddle with 52 holes to sample from a container containing many thousands of beads. Some of the beads were white, some yellow, and some red. Red beads are considered defective and cannot be sold. We counted the number of red beads made by each “worker” each “day.” Five workers each made 4 paddles of beads.
- a) (5 pts.) Explain why each of the BINS assumptions either is or is not reasonably well satisfied in this example.
- b) (5 pts.) In our red bead activity the 52 hole paddle was used 20 times and a total of 86 red beads were “manufactured.” Give a 95% confidence interval for π , the chance of producing a red bead on a single trial, that is, in a single hole. Show your work.

30. **(10 points extra credit. No partial credit for BS!)** We had a sampling activity that led to choosing a random sample of 5 “properties” from a universe of 100 properties and finding the average assessed value in the sample. Suppose instead that we had obtained random samples (*without replacement*) of size 99. Compare the amount of uncertainty in averages based on samples of size 99 with the amount of uncertainty in averages based on samples of size 5. In general, what can you say about the sampling distribution for averages based on samples of size 99? (Think carefully! We are sampling *without* replacement! How many different samples are there?)

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: