

Statistics Competency Exam Fall 2017

Name _____

There are five possible responses to each of the following multiple choice questions. There is only one “BEST” answer. Be sure to read all possible choices before selecting your answer. You may mark on this examination. You can use a calculator but a calculator manual cannot be used.

Stat Comp

CONFIDENCE LEVEL	90%	95%	99%
z^*	1.645	1.96	2.576

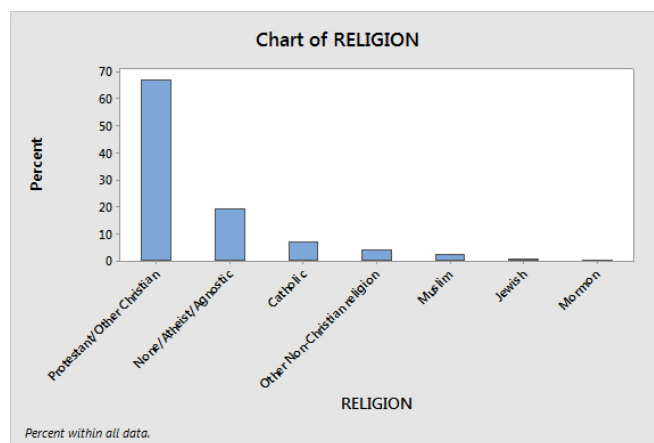
1. Which of these questions from a class survey produced variables that are quantitative?
- i. Are you satisfied with America’s law or policies on guns? (Yes or No)
 - ii. Approximately, how many phone calls do you make per day?
 - iii. What is your ideal number of children?
 - iv. Are you satisfied with America’s security with terrorism? (Yes or No)
 - v. Are you satisfied with America’s state of race relations? (Yes or No)
- (A) i (B) ii (C) iii (D) ii, iii (E) i, iv, v

Use the following for the next 2 questions. The following data (sorted) are the percents of state residents who were born outside the United States.

1.2 1.8 1.9 2.1 2.2 2.7 2.7 2.8 2.9 3.2 3.3 3.6 3.8
3.8 3.9 3.9 4.1 4.2 4.4 4.9 5.1 5.4 5.6 5.6 5.9 6.3
6.6 6.9 7.0 8.1 8.3 9.2 9.7 10.1 10.1 10.3 12.2 12.4 12.6
12.7 12.9 13.8 14.1 15.1 15.9 16.3 18.9 19.1 20.1 21.6 27.2

2. Which of the following best describes the distribution?
- (A) Ignoring the outlier(s), it is close to symmetric. The mean will be close to the median.
 - (B) The distribution is double-peaked and has an outlier.
 - (C) The distribution is strongly skewed left and has several peaks. The center of the distribution is around 10% with a spread of 20%.
 - (D) The distribution is right skewed with a peak which lies between 0% and 5% of residents that are foreign born. There is a possible outlier.
 - (E) Flat, resembles a uniform distribution.
3. The center of the distribution is close to
- (A) 3.8%. (B) 6.3%. (C) 10%. (D) 50%. (E) the modes: 2.7%, 3.8%, 3.9%, 5.6%, 10.1%.

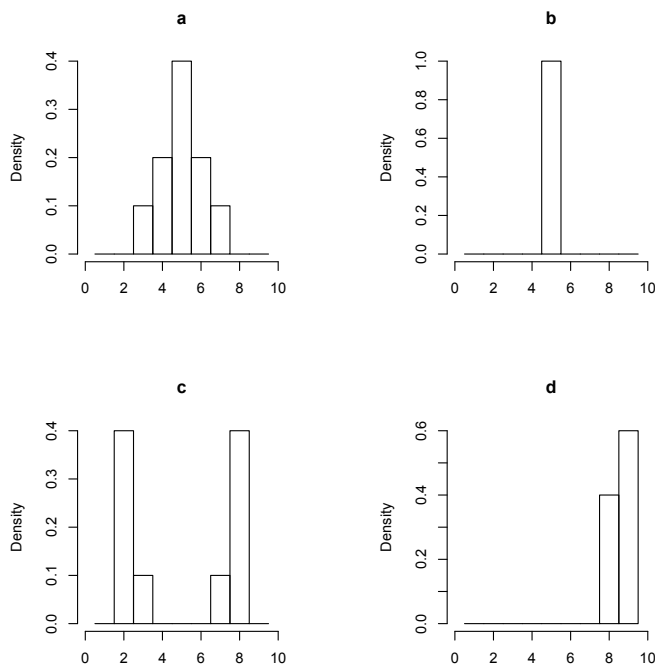
4. The scores of an easy test range from 0 to 100. There are a few scores below 50, some scores ranging from 50 to 80, and many scores above 80. The distribution of test scores will be
- (A) roughly symmetric. (C) correlated to age of the student.
- (B) roughly normal. (D) skewed to the right.
- (E) skewed to the left.
5. The number of deaths among persons aged 15 to 24 years in the United States in 2003 due to the leading causes of death for this age group were: accidents, 14,966; homicide, 5148; suicide, 3921; cancer, 1628; heart disease, 1083; congenital defects, 425. To display these data you should construct a
- (A) histogram. (B) stemplot. (C) bar graph. (D) time plot. (E) 5-number summary.
6. A survey asked "What is your religious identification?" The figure below represents the responses of 775 students.



Which of following best describes this distribution?

- (A) Approximately 2 times the number of students identify as Protestant/Other Christian than all other religious classes combined.
- (B) The distribution is strongly skewed right.
- (C) The distribution is strongly skewed left.
- (D) The distribution is evenly spaced from Protestant/Other Christian to Mormon.
- (E) The distribution is skewed right, the center is Other/Non-Christian religion and the range is from Protest/Other Christian to Mormon.
7. If you were told a population has a mean of 25 and a variance of 0, what must you conclude?
- (A) The data are probably bell-shaped. (C) 95% of the data values lie between $25 \pm 2\sqrt{0}$.
- (B) There is very little variation in the data. (D) Not sure without looking at a graph and the five-number summary.
- (E) All the values in the population are 25.

8. Which of the following histograms has the largest standard deviation? Each histogram ranges from 0 to 1 and is fairly symmetric around 0.5.



- (A) Histogram A (B) Histogram B (C) Histogram C (D) Histogram D (E) No way to tell.

Use the following for the next 2 questions. Suppose that the blood alcohol count (BAC) of students who drink 5 beers varies from student to student according to a normal distribution with mean 0.08 and standard deviation 0.01.

9. What percent of students who drink 5 beers have a BAC above 0.08 (the legal limit for driving in all states)?
 (A) 2.5% (B) 5% (C) 16% (D) 32% (E) 50%
10. The middle 95% of students who drink 5 beers have a BAC between
 (A) 0.06 and 0.10. (B) 0.07 and 0.09. (C) 0.05 and 0.11. (D) 0.04 and 0.12. (E) 0% and 8%.
11. The Community Intervention Trial for Smoking Cessation asked whether a community-wide advertising campaign would reduce smoking. The researchers located 11 pairs of communities, each pair similar in location, size, economic status, and so on. One community in each pair participated in the advertising campaign and the other did not. To decide which community in each pair should get the advertising campaign the researchers tossed a coin. This is
 (A) a completely randomized experiment. (C) a stratified random sample.
 (B) a matched pairs experiment. (D) an observational study.
 (E) an uncontrolled experiment.
12. Make _____ to display the relationship between two quantitative variables measured on the same subjects.
 (A) side-by-side boxplots (C) the five-number summary
 (B) back-to-back stemplots (D) a scatterplot
 (E) a two sample t-test

13. Does regular exercise reduce the risk of a heart attack? Here are two ways to study this question.
- A researcher finds 2000 men over 40 who exercise regularly and have not had heart attacks. She matches each with a similar man who does not exercise regularly, and she follows both groups for 5 years.
 - Another researcher finds 4000 men over 40 who have not had heart attacks and are willing to participate in a study. She assigns 2000 of the men to a regular program of supervised exercise. The other 2000 continue their usual habits. The researcher follows both groups for 5 years.

Which of the following is correct?

- The first design is an observational study, and the second is an experiment.
 - Both designs are experiments.
 - The first design is an experiment, and the second is an observational study.
 - Both designs are observational studies.
 - The first design is an SRS and the second is a stratified random sample.
- Use the following for the next 2 questions.** Can changing diet reduce high blood pressure? Vegetarian diets and low-salt diets are both promising. There are 240 men with high blood pressure to serve as subjects and they are to be assigned at random to four diets: (1) normal diet with unrestricted salt; (2) vegetarian with unrestricted salt; (3) normal with restricted salt; and (4) vegetarian with restricted salt. The study ran for 8 weeks.
14. An important response variable in this experiment is
- control, randomization and replication.
 - the randomization of the subjects to the treatments.
 - the length of the study.
 - the change in blood pressure after 8 weeks on the assigned diet.
 - normal/vegetarian diet or unrestricted/restricted salt.
15. It turned out that the group of men that were on the vegetarian diet with restricted salt had reduced their high blood pressure the most. The researchers said the results were statistically significant at the 0.05 significance level. What does “statistically significant” mean in this context?
- Perhaps the results are attributable to some confounding variable (e.g., exercise).
 - There is evidence at the 0.05 significance level that there is a difference in the mean reduction of blood pressure between the diets.
 - They mean that the change in blood pressure for the men on the vegetarian diet with restricted salt was larger than what they would expect to occur by chance alone.
 - The results are of practical importance and this can be stated with 95% confidence.
 - The chance that the null hypothesis is true is very small.
16. Researchers asked mothers how much soda (in ounces) their kids drank in a typical day. They also asked these mothers to rate how aggressive their kids were on a scale of 1 to 10, with larger values corresponding to a greater degree of aggression. The correlation between amount of soda consumed and aggression rating was found to be $r = 0.3$. If the researchers had measured amount of soda consumed in liters instead of ounces, what would be the correlation? (There are about 34 ounces in a liter.)
- 0.3
 - $0.3/34 = 0.009$
 - $(0.3)(34) = 10.2$
 - 0.3 ounces per point on scale
 - Cannot be determined without converting ounces to liters.

17. An advertising firm, interested in determining how much to emphasize television advertising in a certain county decides to conduct a sample survey to estimate the average number of hours each week that households within that county watch television. The county has two towns, A and B, and a rural area C. Town A is built around a factory and most households contain factory workers with school-aged children. Town B contains mainly retirees and the rural area C are mainly farmers. There are 155 households in town A, 62 in town B and 93 in the rural area, C. The firm decides to select 20 households from Town A, 8 households from Town B and 12 households from the rural area. This is an example of
- (A) an observational sample. (C) a simple random sample.
(B) an experiment. (D) a multistage random sample.
(E) a stratified random sample.
18. A person is testing whether a coin that a magician uses is biased. After analyzing the results from his coin flipping, the P -value ends up being 0.20, so he concludes that there is no evidence that the coin is biased. Based on this information, which of these is a possible 95% confidence interval on the population proportion of times heads comes up?
- (A) (.32, .46) (B) (.76, .98) (C) (.43, .55) (D) (.14, .45) (E) Any interval is possible.
19. Studies have shown that walnuts can reduce blood cholesterol. Walnuts are rich in polyunsaturated fatty acids, and they also help keep blood vessels healthy and elastic. Almonds appear to have a similar effect, resulting in a marked improvement within just four weeks. A statistical test is more likely to find a significant decrease in blood cholesterol if
- (A) it is based on a very small random sample. (C) the p -value is large.
(B) it is based on a very large random sample. (D) the test of hypotheses is not rejected.
(E) the size of the sample doesn't have any effect on the significance of the test.
20. There is a very promising new cancer drug that only has a very small amount to test for further study. In an early clinical trial, the research found that there was a dramatic reduction in cancer growth rate on three subjects. Although the cancer growth rate dramatically decreased, the results are not statistically significant. The explanation is
- (A) the placebo effect is present, which limits statistical significance.
(B) the calculation was in error. The researchers forgot to include the P -value.
(C) that although the cancer growth rate has decreased, the test statistic was large in magnitude.
(D) that the sample size is small and it isn't surprising to find the results not statistically significant.
(E) the results are not of practical significance.
21. Which of the following questions does a test of significance answer?
- (A) Is the observed effect important? (C) Is the null hypothesis true?
(B) Is the sample or experiment properly designed? (D) Is the observed effect due to chance?
(E) Is it based on a very small random sample?

Use the following for the next 2 questions. How much sleep do we need? Several sleep studies have found that seven hours is the optimal amount of sleep. A survey asked students "Usually, how many hours sleep do you get a night?" Assume that the 1216 students that responded to this question are an SRS of all students. We suspect that college students sleep less than seven hours. To verify the suspicion, we test the hypotheses

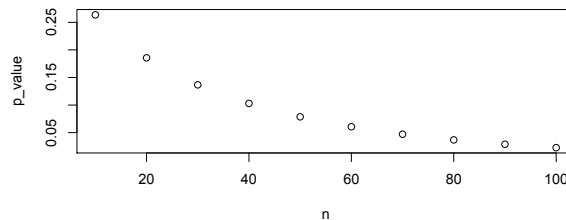
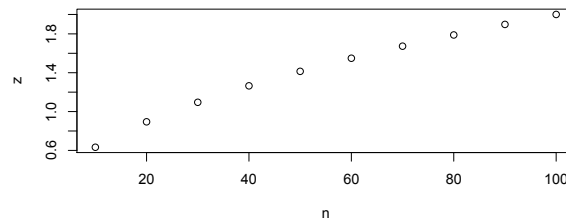
$$H_0 : \mu = 7.0 \text{ vs. } H_a : \mu < 7.0$$

We found the following: $\bar{x} = 6.58$ hours, $s = 1.33$ hours, $t = -176$, p -value ≈ 0 . Hence, we reject H_0 , and conclude that there is strong evidence that, on average, college students sleep less than seven hours per night.

22. The average number of hours sleep per night was 6.58 hours for this sample of 1216 students. The number 6.58 is a
- (A) statistic. (B) parameter. (C) population mean. (D) significant number. (E) practical number.

23. Which of the following would be most helpful in assessing the practical significance of the results?
- (A) Test the hypotheses again, using a smaller significance level.
 - (B) Take a larger sample and retest the hypotheses.
 - (C) Plot the data to display the effect we are seeking, and use a confidence interval to estimate the actual value of μ .
 - (D) Take a random sample of non-college students and retest the hypotheses.
 - (E) Compute the probability of how many students slept less than 7.0 hours.
24. After surveying students at ETSU, a 95% confidence interval for the mean cost of books for one term was computed to be (\$380, \$406). Which of the following is the correct interpretation of the computed interval?
- (A) 95% of all students pay between \$380 and \$406 for books.
 - (B) With 95% confidence, the average amount all students pay is between \$380 and \$406.
 - (C) 95% of the sampled students paid between \$380 and \$406.
 - (D) We are 95% sure that students in this sample averaged between \$380 and \$406 for books.
 - (E) 95% of the samples of students will have average books costs between \$380 and \$406.
25. Inference is _____.
- (A) the process of deducing sample values for populations.
 - (B) the process of selecting a sample.
 - (C) valid for voluntary response samples.
 - (D) drawing conclusions about the sample mean and the sample standard deviation.
 - (E) the process of drawing conclusions about the population from sample information.
26. An analysis of published medical studies about heart attacks noted that in the studies having randomization and strong control for bias, the new therapy provided improved treatment 9% of the time. In studies without randomization or other controls for bias, the new therapy provided improved treatment 58% of the time. How should we interpret these reported findings?
- (A) Some articles that suggest a particular food, drug, or environmental agent is harmful or beneficial should be viewed skeptically, unless we learn more about the statistical design and analysis for the study.
 - (B) The randomized studies were poorly conducted, or they would have found the new treatment to be better much more than 9% of the time.
 - (C) These results suggest it is better not to use randomization in medical studies, because it is harder to show that new ideas are beneficial.
 - (D) In this case, it is better to use the uncontrolled study results instead of the controlled experiment results.
 - (E) The findings of these studies suggest that the answer will most likely lie in the middle.
27. A recent survey of students found that 70% of the students in the sample said that they are not satisfied with the state of race relations in the U.S. The poll's margin of error for 95% confidence was 3%. This means that
- (A) we can be sure that the percent of all students who are not satisfied with the state of race relations in the U.S. is between 67% and 73%.
 - (B) another survey of students using the same method would produce results that will lie between 67% and 70%.
 - (C) there is a 95% chance that 70% will be within $\pm 3\%$.
 - (D) the survey wasn't conducted correctly 3% of the time and this can be stated with 95% probability.
 - (E) the survey used a method that gets an answer within 3% of the truth about the population 95% of the time.

28. In Pick 3 lottery game, the player chooses 3 numbers from 0 to 9. For the straight bet, this is simply matching all 3 numbers in the correct order. The probability of getting each number correct is $1/10$ so the probability of matching all 3 is $(1/10)^3 = 1/1000$. This means that
- (A) if you play the Pick 3 lottery 1000 times, exactly 1 of them will be a perfect match of all 3 numbers.
 - (B) if you play the Pick 3 lottery thousands of times, the fraction of them that would have a perfect match with your 3 numbers will be very close to $1/1000$.
 - (C) since there are only two possibilities (win or lose) your chances of winning will be very close to 50 – 50 in the long run.
 - (D) if you played the Pick 3 lottery 999 times and lost every time, the next time that you play you should have a perfect match of all 3 numbers.
 - (E) if you played the Pick 3 lottery and lost then your chances of winning will increase the next time you play.
29. You are testing the null hypothesis that $\mu = 0$ versus the alternative $\mu > 0$ using $\alpha = 0.05$. Suppose that $\bar{x} = 1$ and $\sigma = 5$. The following plots represent the z -test statistic values versus the sample size (n) and the P -values versus the sample size (n). Which of the following summarizes the effect of the sample size on significance testing?



- (A) The plots show that as sample size increases the results become very important.
 - (B) The plots show that the null hypothesis is true as sample size increases.
 - (C) The plots indicate that as sample size increases the test of hypotheses will become statistically significant.
 - (D) The plots are correlated with each other.
 - (E) The plots indicate there is a problem with the test of hypotheses since one is increasing and the other is decreasing. Both plots should be increasing since n is getting larger.
30. A researcher investigating whether joggers are less likely to get colds than people who do not jog found a P -value of 3%. This means that
- (A) joggers get 3% fewer colds than non-joggers.
 - (B) the differences observed would occur only 3% of the time if jogging had no effect on getting colds.
 - (C) there's a 3% chance that joggers get fewer colds.
 - (D) 3% of joggers get colds.
 - (E) there's a 3% chance that joggers don't get fewer colds.