

Statistics Competency Exam Spring 2019

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

- Which of these questions from a class survey produced variables that are quantitative?
 - What do you think is the ideal number of children for a family to have?
 - Do you favor or oppose the death penalty for persons convicted of murder (Favor, Oppose)?
 - How many pairs of shoes do you own?
 - What is your opinion about a married person having sexual relations with someone other than their marriage partner (Always wrong, Almost always wrong, Wrong only sometimes, Not wrong at all)?
 - Usually, how many hours sleep do you get at night?

(A) i, ii (B) iii, iv, v (C) ii, iv (D) iii (E) i, iii, v
- If you draw an M&M candy at random from a bag of the candies, the candy you draw will have one of six colors. The probability of drawing each color depends on the proportion of each color among all candies made. Here are the probabilities of each color for a randomly chosen plain M&M:

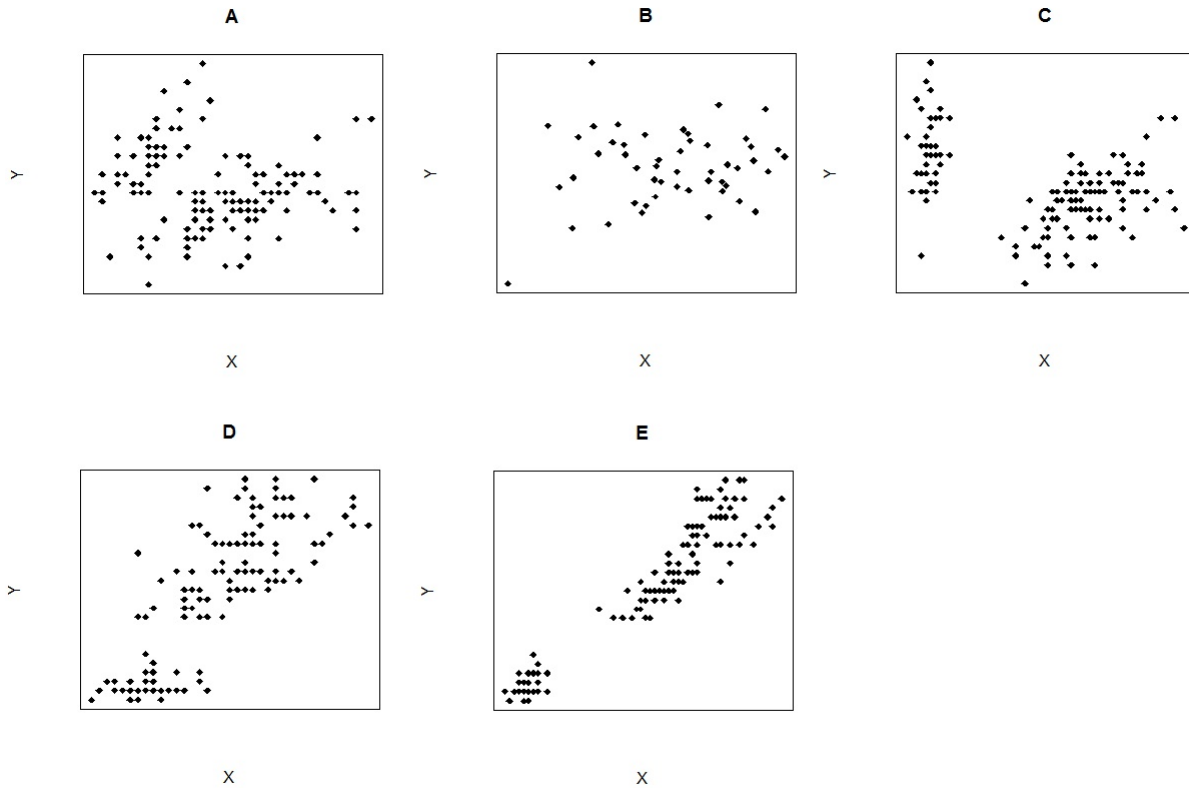
Color	Brown	Yellow	Green	Orange	Blue	Red
Probability	0.13	0.14	0.16	0.20	?	0.13

What is the probability of drawing a blue M&M candy?

(A) 0.14 (B) 0.24 (C) 0.18 (D) 0.20 (E) 0.19
- An IQ test is designed to have scores that have a standard deviation of $\sigma = 10$. A simple random sample of students at a large university will be given the test in order to construct a 95% confidence interval for the mean IQ of all students at the university. How many students must be tested so that the margin of error will be equal to 2 points?

(A) 45 (B) 150 (C) 97 (D) 16 (E) 10
- Several factors are involved in the creation of a confidence interval. Among them are the sample size, the level of confidence, and the margin of error. Which statement is true?
 - For a specified confidence level, larger samples provide smaller margins of error.
 - For a given sample size, higher confidence means a smaller margin of error.
 - For a fixed margin of error, smaller samples provide greater confidence.
 - For a specified confidence level, halving the margin of error means halving the sample size.
 - All of the above.

5. Which plot corresponds to the strongest correlation between Y and X ?



6. Here are the numbers of home runs that Babe Ruth hit in his 15 years with the New York Yankees:

The decimal point is 1 digit(s) to the right of the |

```

2 | 25
3 | 45
4 | 1166679
5 | 449
6 | 0

```

The distribution is best described as

(A) skewed right (B) skewed left (C) roughly symmetric (D) bimodal (E) multimodal

7. A medical experiment compared zinc supplements with a placebo for reducing the duration of colds. Let μ denote the mean decrease, in days, in the duration of a cold. A decrease of $\mu = 1$ is a practically important decrease. A statistical test is more likely to find a significant decrease in the mean duration of a cold if

- (A) it is based on a very large random sample.
- (B) it is based on a very small random sample.
- (C) The size of the sample doesn't have any effect on the significance of the test.
- (D) it is of practical significance.
- (E) the p-value is large.

Use the following for the next 2 questions. The National Center for Health Statistics reports that the mean systolic blood pressure for males 35 to 44 years of age is 128. The medical director of a large company looks at the medical records of 72 executives in this age group and finds that the mean systolic blood pressure is 126.1 and the standard deviation is 15.2. Is this evidence that the company's executives have a different mean blood pressure from the general population?

8. The hypotheses of interest are
- (A) $H_0 : \mu = 126.1$ vs. $H_a : \mu < 126.1$
 (B) $H_0 : \mu = 128$ vs. $H_a : \mu \neq 128$
 (C) $H_0 : \bar{x} = 126.1$ vs. $H_a : \bar{x} \neq 126.1$
 (D) $H_0 : \mu < 128$ vs. $H_a : \mu = 128$
 (E) $H_0 : \mu = 128$ vs. $H_a : \mu < 128$

9. The test statistic value is

- (A) 1.061
 (B) -0.125
 (C) -9
 (D) -1.061
 (E) 0.125

Use the following for the next 4 questions. How do women and men compare in the pursuit of academic degrees? The table below presents counts (in thousands) from the *Statistical Abstract* on degrees earned in 2005 categorized by the level of the degree and gender of the recipient.

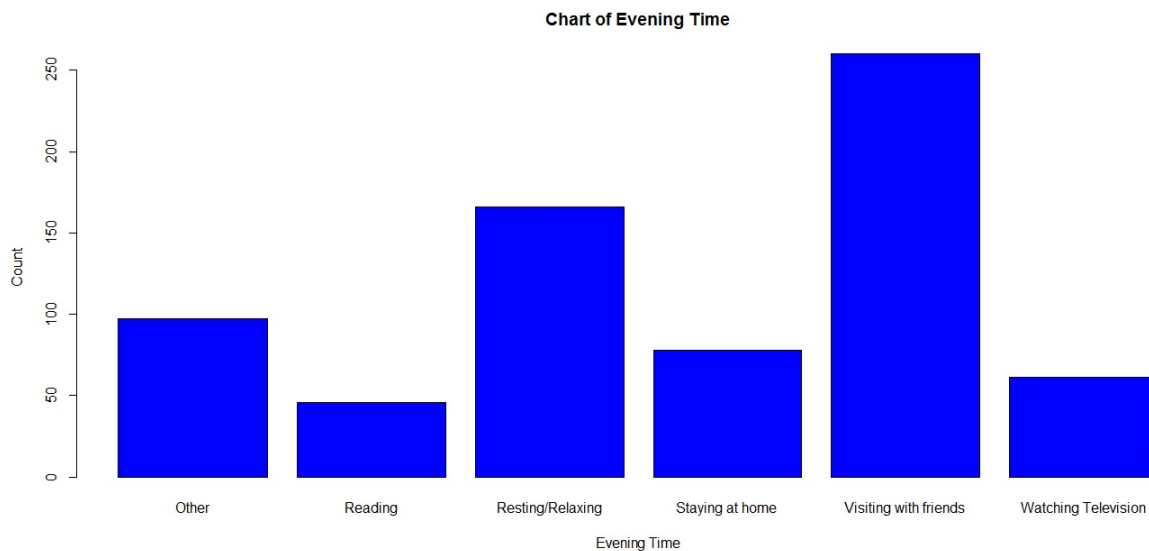
Gender	Bachelor's	Master's	Professional	Doctorate	Total
Female	826	341	43	26	1236
Male	613	234	44	27	918
Total	1439	575	87	53	2154

10. What proportion of degree recipients are Male and earned a Doctorate degree?
 (A) $\frac{971}{2154}$ (B) $\frac{918}{2154}$ (C) $\frac{27}{918}$ (D) $\frac{27}{2154}$ (E) $\frac{53}{2154}$
11. What proportion of the degree recipients earned a Bachelor's degree?
 (A) $\frac{1439}{2154}$ (B) $\frac{826}{1439}$ (C) $\frac{1236}{1439}$ (D) $\frac{1236}{2154}$ (E) $\frac{826}{2154}$
12. Given a degree recipient earned a Master's degree, what proportion are Female?
 (A) $\frac{341}{2154}$ (B) $\frac{575}{1236}$ (C) $\frac{341}{1236}$ (D) $\frac{575}{2154}$ (E) $\frac{341}{575}$
13. We are interested in testing whether gender differs among the level of degree earned. Which of the following statistical methods can be used to address this?
 (A) Chi-square test (B) Two-sample t procedure (C) Paired-t test (D) Z-test (E) Regression analysis

Use the following for the next 2 questions. The weights of legal U.S. quarters have a normal distribution with a mean of 5.65 grams and a standard deviation of 0.05 grams.

14. Vending machines can be adjusted to reject coins above and below certain weights. If a vending machine is adjusted to reject quarters that weigh more than 5.75 grams and less than 5.55 grams, what percentage of legal quarters will be rejected by the machine?
 (A) 95% (B) 2.5% (C) 5% (D) 10% (E) 97.5%
15. In a sample of 25 legal U.S. quarters, what percentage weighs more than 5.70 grams?
 (A) 15.87% (B) 0% (C) 1% (D) 84.13% (E) 70%

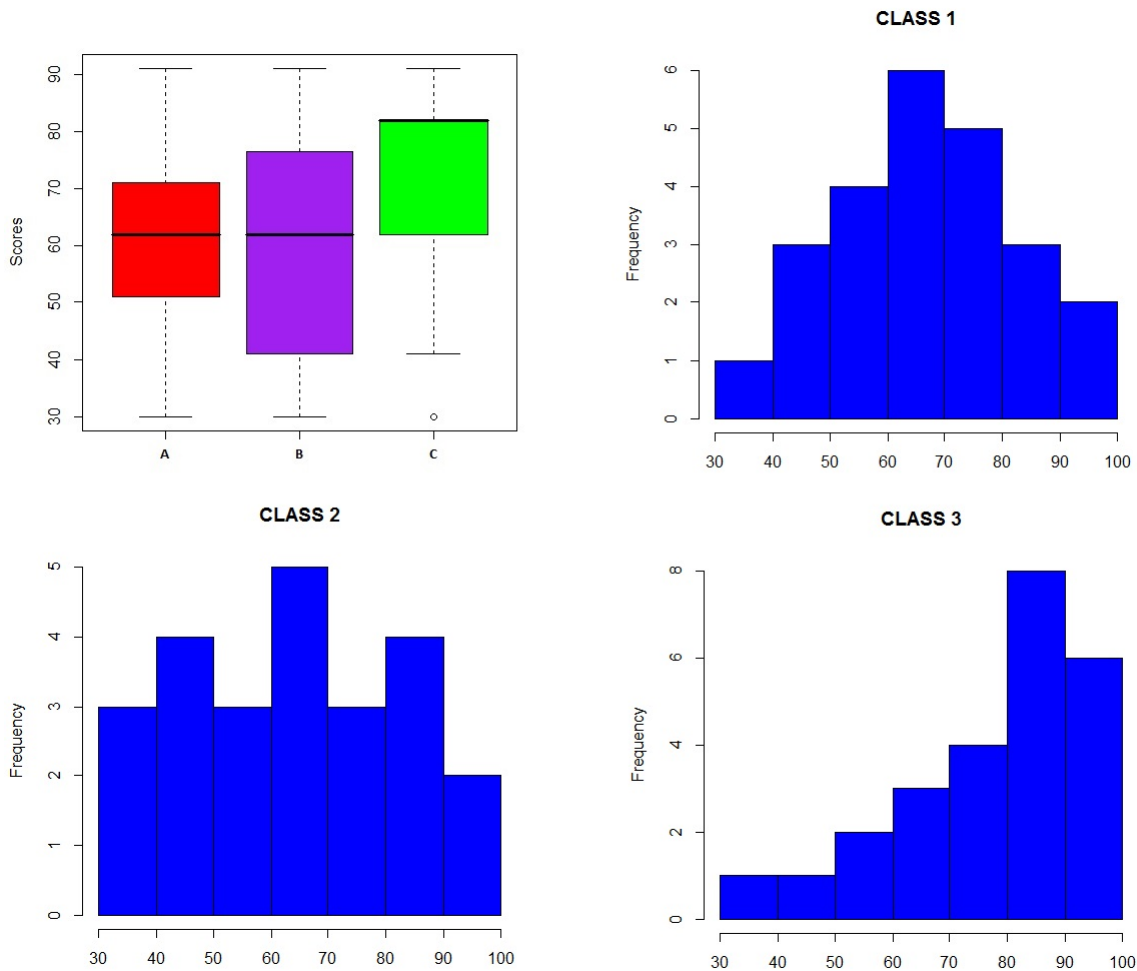
16. Data were collected over a decade from 1021 men and women with a recent history of precancerous colon polyps. Participants were randomly assigned to receive folic acid or a placebo, and the study concluded that those receiving folic acid may actually increase their risk of developing additional precancerous growths. (Source: *JAMA* 2007, 297) This is an example of:
- (A) a sample survey (C) a block design (E) a completely randomized experiment
- (B) an observational study (D) a matched pairs study
17. A sales company that has just recently started taking online orders promises to deliver orders within 3 days. Follow-up calls to randomly selected customers shows that a 95% confidence interval for the proportion of all orders that arrive on time is 0.88 ± 0.06 . Which one is the correct interpretation of this confidence interval?
- (A) We are 95% confident that between 82% and 94% of the orders placed by the sampled customers arrived on time.
- (B) Between 82% and 94% of all orders arrive on time.
- (C) 95% of all random samples of customers show that 88% of orders arrive on time.
- (D) We are 95% confident that between 82% and 94% of all orders will arrive on time.
- (E) On 95% of the days, between 82% and 94% of the orders will arrive on time.
18. In the Spring 2016 Math 1530 Survey, students were asked “What is your favorite way of spending an evening? (Staying at home, Watching TV, Resting/Relaxing, Reading, Visiting with friends, Other)” The following graph displays the information on this variable:



Which of the following describes this distribution?

- (A) The distribution is skewed left and the range is from “Other” to “Watching Television.”
- (B) The distribution is multimodal and skewed left.
- (C) The distribution is evenly spaced from “Other” to “Watching Television.”
- (D) “Visiting with friends” is the favorite way of spending an evening while “Reading” is the least favorite way of spending an evening.
- (E) The distribution is irregular and centered around “Staying at home.”

19. After hearing of the national result that 40% of students engage in binge drinking (5 or more drinks for a man, 4 or more drinks for a woman over the course of about 2 hours), a professor surveyed a random sample of 265 students at her college and found that 113 of them admitted to binge drinking in the past week. Should she be surprised by this result?
- (A) Probably, the proportion at her college is almost 3 standard deviations above the national result.
 - (B) Probably not, the proportion at her college is less than 1 standard deviation above the national result.
 - (C) Probably, the proportion at her college is slightly higher than the national result.
 - (D) Probably not, the proportion at her college is less than 1 standard deviation below the national result.
 - (E) Probably not, what students do in their free time is of no concern to her.
20. Three Statistics classes all took the same test. Histograms and boxplots of the scores for each class are shown below. Match each class (CLASS 1, CLASS 2, or CLASS 3) with the corresponding boxplot (A (first boxplot), B (middle boxplot) or C (third boxplot)).



- (A) CLASS 1 is A, CLASS 2 is C, CLASS 3 is B
- (B) CLASS 1 is C, CLASS 2 is A, CLASS 3 is B
- (C) CLASS 1 is A, CLASS 2 is B, CLASS 3 is C
- (D) CLASS 1 is B, CLASS 2 is C, CLASS 3 is A
- (E) CLASS 1 is B, CLASS 2 is A, CLASS 3 is C

21. If we assume the conditions for correlation are met, which of the following are true?

- (A) A correlation of -0.97 indicates a strong, negative linear association.
- (B) Standardizing the variables will make the correlation 0.
- (C) A correlation of 0.03 indicates a strong, positive linear association.
- (D) The units of the correlation are the same as the units of y .
- (E) Multiplying every value of x by 3 will triple the correlation.

22. A lurking variable is

- (A) one whose effect on a response variable cannot be distinguished from other variables in the experiment.
- (B) one whose effect on a response variable can be distinguished from other variables in the experiment.
- (C) one that is changed in an experiment.
- (D) one that occurs in one of two possible states, labeled success and failure.
- (E) an extraneous one which causes an explanatory and response variable to appear to be related in a way that they are not.

Use the following for the next 2 questions. National research indicates that full-time (12 credit hours or more) students spend, on the average, \$500 on textbooks per semester at 4-year public institutions. The Fall 2015 MATH1530 survey asked students “Approximately how much did you spend on books for school this semester?” Is there evidence that the mean textbook cost for ETSU students this semester differs from \$500? Use the output below to answer the next two questions.

One-Sample T: BOOKS

Test of $\mu = 500$ vs not equal 500

Variable	N	Mean	StDev	SE Mean	95% Lower Bound	95% Upper Bound	T	P
BOOKS	759	410.1	182.59	6.63	397.1	423.1	-13.6	0

23. Which of the following is the best conclusion?

- (A) There is strong evidence ($P = 0$) that, on average, the cost of textbooks for full-time undergraduate students at ETSU is different from \$500.
- (B) No. There is weak evidence ($P = 0$) that, on average, the cost of textbooks for full-time undergraduate students at ETSU is different from \$500.
- (C) No. We fail to reject H_0 since the P is close to 0.
- (D) The probability that H_0 is true is 1 since the P -value is 0.
- (E) The data are not statistically significant since the P -value is small.

24. Which of the following is the best interpretation of the above confidence interval?

- (A) With 95% confidence, the mean cost of textbooks for full-time undergraduate students at ETSU should be within the calculated interval.
- (B) There is a 95% probability that the cost of textbooks for this sample of students at ETSU should be within the calculated interval.
- (C) The cost of textbooks for full-time undergraduate students at ETSU is \$410.10 and it is within the 95% confidence interval.
- (D) 95% of all textbooks for full-time undergraduate students at ETSU cost \$410.10 and it is within the calculated confidence interval.
- (E) With 95% confidence, all the sample means will be within the calculated confidence interval.

25. Major League Baseball tests players to see whether they are using performance enhancing drugs. Officials select a team at random, and they test all 40 players on the team. What kind of sample is this?

- (A) Simple random sample
- (B) Multistage sample
- (C) Volunteer sample
- (D) Systematic sample
- (E) Cluster sample

26. The table below lists the number of car thefts in a large city over a 14-day period:

51 76 57 63 55 64 54 58 79 65 77 59 78 53

The five-number summary is

- (A) 51, 55, 61, 76, 79
 - (B) 51, 55, 63.5, 76, 79
 - (C) Median = 61, Mean = 63.5, Standard Deviation = 10, Min = 51, Max = 79
 - (D) $N=14$, IQR = 17.75, Standard Deviation = 10, Range = 28, Median = 61
 - (E) 51, 63, 56, 77, 53
27. A vice president at a bank states that customers obtain service within an average time of 4.5 minutes of arriving at the bank's drive-up service window. However a long-time customer of the bank wants to test the hypothesis that the average time to obtain service exceeds 4.5 minutes, using a significance level of $\alpha = 0.10$. The customer collects data from 25 randomly selected customers, and calculates the average time to obtain service. The sample mean time is 5.0 minutes and the sample standard deviation is 1.5 minutes. If the researcher decides to reject the null hypothesis, the researcher would conclude that
- (A) at the 1% significance level, the average waiting time for these 25 bank customers to receive service does not exceed 4.5 minutes
 - (B) at the 10% significance level, the average waiting time for these 25 bank customers to receive service does exceed 4.5 minutes
 - (C) at the 10% significance level, the average waiting time for all of the bank customers to receive service does not exceed 4.5 minutes
 - (D) at the 10% significance level, the average waiting time for all of the bank customers to receive service does exceed 4.5 minutes
 - (E) at the 10% significance level, the total waiting time for these 25 bank customers to receive service does not exceed 4.5 minutes
28. Which of the following statements are false?
- (A) It is better to use an alpha level of 0.05 than an alpha level of 0.01.
 - (B) If we use an alpha level of 0.01, then a P-value of 0.001 is statistically significant.
 - (C) If we use an alpha level of 0.01, then we reject the null hypothesis if the P-value is 0.001.
 - (D) If the P-value is 0.01, we reject the null hypothesis for any alpha level greater than 0.01.
 - (E) Using an alpha level of 0.05, a P-value of 0.025 results in rejecting the null hypothesis.
29. Many states require that cars be tested annually to be sure that exhaust emissions not exceed specified limits for certain pollutants. Suppose that state regulators double-check a random sample of cars that a repair shop has certified as okay. The state regulators will revoke the shop's license if they find significant evidence that the shop is certifying cars that do not meet the standards. In this context, what is a Type I error?
- (A) The shop is certified as meeting standards when it is not.
 - (B) It is decided that the shop is not meeting standards when it is.
 - (C) It is the significance level that the state regulators use for the test.
 - (D) The shop is certified as meeting standards when it is.
 - (E) It is decided that the shop is not meeting standards when it is not.
30. Suppose we have data on the length of childrens' lullabies (in seconds). Which type of graph would be appropriate to display the data?
- (A) Bar chart
 - (B) Side by side bar charts
 - (C) Pie chart
 - (D) Scatterplot
 - (E) Histogram