

Statistics Competency Exam Fall 2021

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. What is the fastest you have ever driven a car? (in mph)
 - ii. What is your age? (in years)
 - iii. What is your religious identification? (Christian Religion, Non-Christian Religion, None)
 - iv. Do you believe vaccines are safe? (Yes, No)
 - v. Have you, yourself, ever smoked cigarettes in the past week? (Yes, No)
- (A) i, ii (B) i (C) ii, iii (D) iii, iv, v (E) iv, v
2. Suppose the table below represents the probability distribution for the number of siblings that students in a small statistics class has.

Number of Siblings	0	1	2	3	4	5	6	7
Probability	0.20	0.25	0.30	0.1	0.05	?	0.01	0.00

What is the probability that a randomly selected student has 5 siblings?

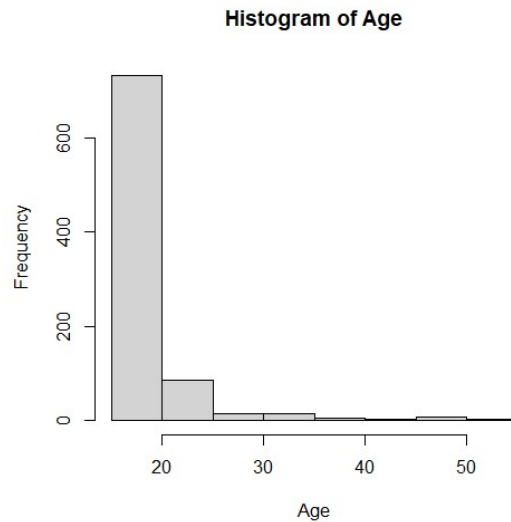
- (A) 0.15 (B) 0.09 (C) 0.01 (D) 0.91 (E) Not enough information provided.
3. According to healthline.com, as of 2018, the average height for American men aged 20 years and older is 69.1 inches. Assume that heights for American men are normally distributed with a mean of 69.1 inches and a standard deviation of 3.3 inches. Approximately what percent of American men, aged 20 years and older, are taller than 72.4 inches?
- (A) 68% (B) 97.5% (C) 34% (D) 16% (E) 2.5%
4. The number of pets owned by students in a small class is as follows:

2 0 1 3 3 4 1 0

The five-number summary (calculated by hand) would be

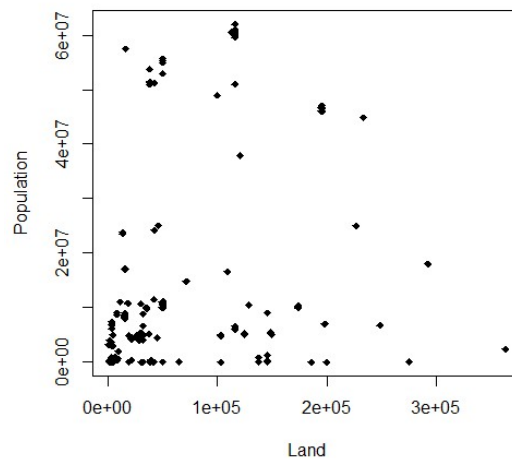
- (A) 0, 0.5, 1.5, 3, 4
- (B) 2, 0.5, 3, 2.5, 0
- (C) 0, 1, 1.5, 3, 4
- (D) Median = 1.5, Mean = 1.75, Range=4, Min =0, Max = 4
- (E) N=8, IQR = 2.5, Mean = 1.75, Range = 4, Median = 1.5

Use the following for the next question. A class survey asked “What is your age? (in years)” The figure below represents the responses of 867 students.



5. Which of the following best describes this distribution? This distribution is
- (A) right-skewed, with the center around 20 years and spread from 15 to 55 years, with possible outliers.
 - (B) left-skewed, with the center around 20 years and spread from 15 to 55 years, with possible outliers.
 - (C) symmetric around 20 years, with very little variation.
 - (D) multi-peaked with the bulk of data from 15 to 20.
 - (E) evenly-spaced from 15 to 55.
6. A student working in ETSU’s Treehouse wanted to see how often people bought something to eat or drink. She decided she would send out a survey to all ETSU faculty, staff and students, asking, “How many times have you bought something at the ETSU Treehouse this week?” She had a total of 150 people respond to her survey. This is an example of
- (A) an experiment.
 - (B) a systematic sample.
 - (C) a matched-pairs design.
 - (D) a simple random sample.
 - (E) a volunteer sample.
7. A researcher performed a hypothesis test to determine if the percentage of U.S. college professors that had at least one child is different than 50%. He conducted his test using a 10% significance level and decided not to reject the null hypothesis. Suppose the researcher constructs a 90% confidence interval on p . For the 90% confidence interval,
- (A) the value of 0.5 would be contained in it.
 - (B) cannot be constructed after the researcher has performed a hypothesis test.
 - (C) as there is insufficient information, cannot say anything about it.
 - (D) the value of 0.5 would not be contained in it.
 - (E) 90% of the sample of professors used in his study had at least one child.
8. A nutritionist compared the pre-diet and post-diet weights of her clients. This is an example of
- (A) a binomial experiment.
 - (B) a placebo-controlled experiment.
 - (C) a matched-pairs design.
 - (D) a prospective study.
 - (E) confounding.

Use the following for the next two questions. A class survey asked the following questions: “Pick any country that you would like to visit. For the next two questions, answer the following questions about the country you chose. Land: What is the land area (in square miles)? Population: What is the population?” A scatterplot of the data from these questions is given:



9. A linear model was fit to predict land using the population of the country; the fitted model is $\hat{Population} = 8807652.9 + 178.7Land$. Choose the correct interpretation of the slope.
- On average, for every 1 square mile increase in the country's land area, the population increases by 178.7.
 - For every 1 square mile increase in a country's land area, the population increases by 8807652.9.
 - For every 178.7 square miles increase in a country's land area, the population increases by 8807652.9, on average.
 - When a country is brand new, the population is 178.7, on average.
 - On average, for every 1 person increase in a country's population, the land area increases by 178.7 square miles.
10. A plausible value for the correlation between population and land is:
- 0.9
 - 0.5
 - 0.3
 - 1.0
 - 2.0
11. A regression analysis was performed on an Automobile data set to determine whether we could predict the Horsepower of the engine that manufacturers put in a car by knowing the Weight of the car. The R-squared value was found to be 85.2%. Which of the following is the correct interpretation of this value?
- The correlation between Horsepower and Weight is 0.852.
 - The y-intercept of the linear regression model is 85.2.
 - The slope of the linear regression model is 85.2.
 - 85.2% of the variation in Weight is explained by its linear relationship with Horsepower.
 - 85.2% of the variation in Horsepower is explained by its linear relationship with Weight.
12. An M&M manufacturer believes that certain flavors of M&Ms are more popular than other flavors. She randomly selects 50 sales records, and records the flavor of each M&M pack sold. A hypothesis test is performed in order to determine if the data supports the claim of the M&M manufacturer. The alternate hypothesis would be
- H_a : The flavor "peanut" has a higher preference than the other flavors.
 - H_a : At least one flavor is preferred differently than the other flavors.
 - H_a : The flavor "plain" has a lower preference than the other flavors.
 - H_a : All flavors are equally preferred.
 - None of the above.

Use the following for the next three questions. A question from a class survey asked students, “How much did you spend on your last meal that you ate at an off-campus restaurant? (in US dollars)” On average, is the amount of money spent on a meal off-campus by ETSU students more than \$10? The output from running the appropriate analysis is given:

One Sample t-test

```
data: Meal
t = 1.8274, df = 866, p-value = 0.03399
```

```
95 percent confidence interval:
9.803007 15.519369
```

```
sample estimates:
mean of x
12.66119
```

13. Which of the following are the correct null and alternative hypotheses for the test?
- (A) $H_0 : \bar{x} = 10$ vs $H_a : \bar{x} \neq 10$
 - (B) $H_0 : \mu = 10$ vs $H_a : \mu > 10$
 - (C) $H_0 : \mu = 10$ vs $H_a : \mu \neq 10$
 - (D) $H_0 : \bar{x} = 12.66$ vs $H_a : \bar{x} > 12.66$
 - (E) $H_0 : \bar{x} = 10$ vs $H_a : \bar{x} > 10$
14. Which of the following is the appropriate decision and conclusion for the test?
- (A) Reject H_0 . With 5% significance level, there is sufficient evidence that the true mean amount of money spent on an off-campus meal by all ETSU students is more than \$10.
 - (B) Do not reject H_0 . With 5% significance level, there is not sufficient evidence that the true mean amount of money spent on an off-campus meal by all ETSU students is more than \$10.
 - (C) Reject H_0 . With 5% significance level, there is sufficient evidence that the true mean amount of money spent on an off-campus meal by this sample of ETSU students is not \$10.
 - (D) Reject H_0 . With 5% significance level, there is sufficient evidence that the true mean amount of money spent on an off-campus meal by all ETSU students is more than \$12.66.
 - (E) Do not reject H_0 . With 5% significance level, there is sufficient evidence that the true mean amount of money spent on an off-campus meal by this sample of ETSU students is not \$12.66.
15. Which of the following is the correct interpretation of the confidence interval?
- (A) With 95% confidence, the true mean amount of money spent on an off-campus meal by all ETSU students is between \$9.80 and \$15.52.
 - (B) With 99% confidence, the true mean amount of money spent on an off-campus meal by all ETSU students is between \$9.80 and \$15.52.
 - (C) With 95% confidence, the mean amount of money spent on an off-campus meal by this sample of ETSU students is between \$9.80 and \$15.52.
 - (D) With 99% confidence, the mean amount of money spent on an off-campus meal by this sample of ETSU students is between \$9.80 and \$15.52.
 - (E) With 95% confidence, the true mean amount of money spent on an off-campus meal read by all ETSU students is \$12.66.

Use the following for the next three questions. Snee (1974) reports the following table corresponding to hair and eye color for 592 students:

	Hair Color				
Eye Color	Black	Brown	Red	Blond	Total
Brown	68	119	26	7	220
Blue	20	84	17	94	215
Hazel	15	54	14	10	93
Green	5	29	14	16	64
Total	108	286	71	127	592

16. Find the probability that a person will have green eyes given that they have brown hair.
 (A) $\frac{29}{592}$ (B) $\frac{64}{592}$ (C) $\frac{286}{592}$ (D) $\frac{64}{286}$ (E) $\frac{29}{286}$
17. Are the events having green eyes and having brown hair independent?
 (A) Yes because $P(\text{Green eyes}|\text{Brown hair}) = P(\text{Green eyes})$.
 (B) No because $P(\text{Green eyes}|\text{Brown hair}) = P(\text{Green eyes})$.
 (C) Yes because $P(\text{Green eyes}|\text{Brown hair}) \neq P(\text{Green eyes})$.
 (D) No because $P(\text{Green eyes}|\text{Brown hair}) \neq P(\text{Green eyes})$.
 (E) Not enough information given to answer.
18. We are interested in testing whether eye color differs among the different hair colors. Which of the following statistical methods can be used to address this?
 (A) Regression analysis (B) Chi-square test (C) Two-sample t procedure (D) Z-test (E) Paired-t test

Use the following for the next two questions. According to the US Health Department (in 2017), 90.2% of American households have a wireless phone, 45.9% have a landline, and 39.4% had both a wireless phone and a landline.

19. For a randomly selected American household, what is the probability that the home has either a landline or wireless phone, but not both?
 (A) 0.508 (B) 0.394 (C) 0.065 (D) 0.573 (E) 0.033
20. For a randomly selected American household, what is the probability that the home doesn't have a landline, nor a wireless phone?
 (A) 0.508 (B) 0.394 (C) 0.065 (D) 0.573 (E) 0.033

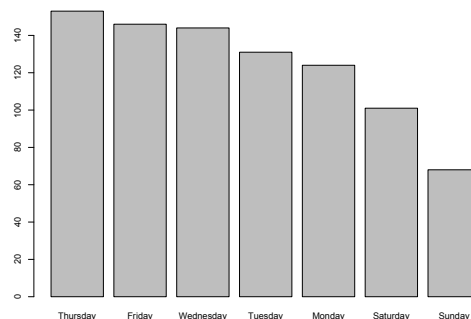
Use the following for the next two questions. According to one source, IQ test scores are normally distributed with a mean of 100 and a standard deviation of 15.

21. How high would a person have to score to be at the 95th percentile?
 (A) 119.35 (B) 114.25 (C) 130 (D) 124.675 (E) 115
22. One test center administered 25 IQ tests in one day. What's the probability that the mean test score for these 25 tests is greater than 130?
 (A) 0.97725 (B) 0 (C) 0.02275 (D) 2 (E) 1
23. Lee is rolling a fair, six-sided die a total of 25 times and counting the number of times the die lands with a "4" face up. This is an example of
 (A) a matched-pairs design.
 (B) a double blind experiment.
 (C) a randomized block design.
 (D) an observational study.
 (E) a binomial experiment.

Use the following for the next question. A class survey asked “Do you agree, disagree, or neither agree nor disagree with the following statement? Vaccines are safe. (Agree, Disagree, Neither agree nor disagree)” A total of 867 students responded to this survey question. 81.1% of the sample responded “Agree”.

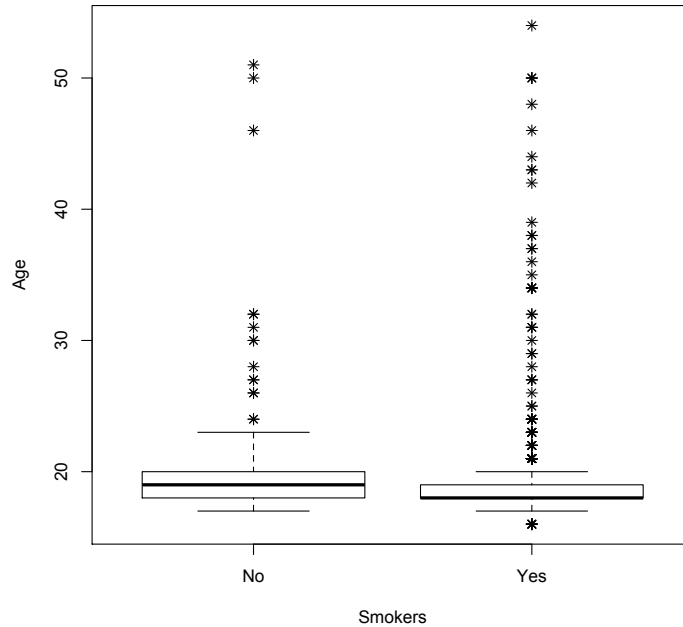
24. Calculate a 95% confidence interval for the proportion of all ETSU students who agree that vaccines are safe.
- (A) 0.811 ± 0.0599
(B) 0.811 ± 0.0306
(C) 0.811 ± 0.0261
(D) 0.811 ± 0.0133
(E) 0.811 ± 0.0219
25. A normal distribution has
- (A) mean $\mu = 1$ and standard deviation $\sigma = 0$.
(B) mean μ and standard deviation σ , where $\sigma > 0$.
(C) mean $\mu = 0$ and standard deviation $\sigma = 1$.
(D) mean μ and standard deviation $\sigma = 1$.
(E) mean $\mu = 1$ and standard deviation $\sigma = 1$.
26. Suppose it is known that 2.5% of all credit transactions in a certain region are fraudulent. In a random sample of 50 transactions, the probability that less than 2 transactions are fraudulent is
- (A) 0.870622 (B) 0.025 (C) 0.9268594 (D) 0.6435113 (E) 0.0001854

Use the following for the next question. A class survey asked “Go to this website <https://www.mathsisfun.com/games/dayofweek.html>. Enter your birthday. What day of the week were you born on? (Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, Saturday)” A total of 867 students responded to this survey.



27. Which of the following is an accurate description for the data displayed in the figure given above?
- (A) The distribution is skewed right.
(B) The distribution is bimodal.
(C) Sunday had the least amount of births.
(D) The mean of the distribution is Wednesday.
(E) The distribution is uniform.

Use the following for the next two questions. A class survey asked “What is your age? (in years)” and “Do you feel it is okay for insurance companies to charge higher rates to smokers? (Yes, No)” A total of 867 students responded to this survey.



Age	n	Min.	Q1	Mean	Q3	Max.
Smokers=Yes	613	16.00	18.00	19.59	19.00	54.00
Smokers=No	254	17.00	18.00	19.62	20.00	51.00

28. According to the figure above, which of the following statements are false?
- (A) The median Age is the same for both groups.
 - (B) The group that had the oldest respondent was the group that replied “Yes” to the Smokers question.
 - (C) There are more outliers for Age in the “Yes” group.
 - (D) The IQR of Age of the respondents that replied “Yes” is smaller than that of the respondents that replied “No” to the Smokers question.
 - (E) The Age of the respondents that replied “Yes” to the Smokers question has the largest range.
29. Using the IQR criterion, Age values larger than ____ for the group that replied “Yes” to the Smokers question would be considered upper outliers.
- (A) 20 (B) 23 (C) 16.5 (D) 20.5 (E) 22
30. Suppose a researcher performs a test of $H_0 : p = 0.75$ vs. $H_a : p > 0.75$. Which of the following statements is correct?
- (A) A large p-value indicates that we should reject the null hypothesis.
 - (B) An extremely small p-value indicates that the actual data differs significantly from that which would be expected if the null hypothesis were true.
 - (C) The p-value measures the probability that the hypothesis is true.
 - (D) The larger the p-value, the stronger the evidence against the null hypothesis.
 - (E) None of the above.