

Quiz: Basic Statistical Concepts

Your Score: 100% Congratulations! Your score of 100% indicates that you've mastered the topics in this lesson. If you'd like, you can review the feedback.

When you're finished, exit the lesson.



1. A sample from a population should be which of the following?

- a. random
- b. representative
- c. biased
- d. all of the above

Your answer: b

Correct answer: b

A sample should be representative. You can attain that by using a random sample. A biased sample is not representative of the population.



2. Which of the following is **not** a characteristic of predictive modeling?

- a. large sample size
- b. many variables
- c. assessed using a holdout or validation data set
- d. answers the question "How is X related to Y?"

Your answer: d

Correct answer: d

Predictive modeling predicts future values of a response variable based on the existing values of predictor variables. You assess the prediction's accuracy using a holdout or validation data set, and the model usually has many variables and a large sample size.



3. The standard error measures the variability associated with the sample mean, \bar{x} .

- a. true
- b. false

Your answer: a
Correct answer: a

The variability associated with the sample mean, \bar{x} , is measured by the standard error.



4. Select the statement below that **incorrectly** interprets a 95% confidence interval (15.02, 15.04) for the population mean, if the sample mean is 15.03 ounces.
- a. Approximately 95% of the intervals calculated with this procedure will capture the true average weight.
 - b. The calculated confidence intervals might or might not contain the true mean.
 - c. The probability is .95 that the true average weight is between 15.02 and 15.04 ounces.

Your answer: c
Correct answer: c

A 95% confidence level means that 95% of a theoretically infinite number of intervals would contain the true population mean, but 5% would not. Therefore, for any given sample, the calculated confidence interval might or might not contain the value of the true population mean.



5. How do you define the term *power*?
- a. the measure of the ability of the statistical hypothesis test to reject the null hypothesis when it is actually false
 - b. the probability of committing a Type I error
 - c. the probability of failing to reject the null hypothesis when it is actually false

Your answer: a
Correct answer: a

Power is the probability that you correctly reject the null hypothesis.



6. The location and spread of a normal distribution depend on the value of which two parameters?
- a. the mean (\bar{x}) and the standard deviation (s)
 - b. the standard deviation (σ) and the variance (σ^2)
 - c. the mean (μ) and the standard deviation (σ)

Your answer: c
Correct answer: c

The location and spread of a normal distribution depend on the value of the mean (μ) and the standard deviation (σ).



7. A bank manager is concerned that the percent of loans processed that contain errors has increased above the acceptable amount of 1%. A significance test is conducted to test his concern (H_0 : loan error rate ≤ 0.01 , H_a : loan error rate > 0.01).

The manager concludes that the rate is indeed above 1%, when in reality it is not. What type of error did the manager make?

- a. Type I
- b. Type II

Your answer: a

Correct answer: a

The null hypothesis is that the percent of loans processed that contain errors is less than or equal to 1%. If the bank manager rejects the null hypothesis when it's actually true, he's made a Type I error. If he fails to reject the null hypothesis and it's actually false, he's made a Type II error.



8. To reject a test with Student's t statistic, the t statistic should be far from 0 and have a small corresponding p -value.

- a. true
- b. false

Your answer: a

Correct answer: a

The Student's t statistic measures how far \bar{x} is from the null hypothesized mean, in standard error units. To reject a test, the t statistic should be much higher or lower than 0 and have a small corresponding p -value.



9. Which PROC TTEST option would you use to change the confidence level in a confidence interval plot?

- a. LEVEL=
- b. INTERVAL=
- c. BOUNDARY=
- d. ALPHA=

Your answer: d

Correct answer: d

The confidence bounds can be changed using the ALPHA= option in the PROC TTEST statement. Set alpha equal to 1 minus confidence. For example, for a 99% confidence interval, you would specify ALPHA=0.01.



10. A CLASS statement is required in a two-sample t test.

- a. true
- b. false

Your answer: a

Correct answer: a

The CLASS statement is required in a two-sample t test. The classification variable must have exactly two levels, because PROC TTEST divides the observations into the two groups using the values of this variable.

Close

Copyright © 2019 SAS Institute Inc., Cary, NC, USA. All rights reserved.