



Quiz, Lesson 1: Introduction to Statistics

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 for each question.



1. For an asymmetric (or skewed) distribution, which of the following statistics is a good measure for the middle of the data?

- a. mean
- b. median
- c. either mean or median

Your answer: **b**

Correct answer: **b**

The median is not affected by outliers and is less affected by the skewness. The mean, on the other hand, averages in any outliers that might be in your data.

Review: [Measures of Location](#)



2. Which of the following code examples correctly calculates descriptive statistics of popcorn yield (**Yield**) for each level of the class variable (**Type**) in the data set **Statdata.Popcorn**, as well as statistics for all levels combined?

The output should include the following statistics: sample size, mean, median, standard deviation, variance, range, and interquartile range.

- a.

```
proc means data=statdata.popcorn
    maxdec=2 fw=10
    n mean median std var
    range qrange;
    class Type;
    var Yield;
run;
```
- b.

```
proc means data=statdata.popcorn
    maxdec=2 fw=10
    printalltypes
    n mean median std var
    range qrange;
    class Yield;
    var Class;
run;
```
- c.

```
proc means data=statdata.popcorn
    maxdec=2 fw=10
    printalltypes
    n mean median std var
    range qrange;
    class Type;
    var Yield;
run;
```
- d.

```
proc means data=statdata.popcorn
    maxdec=2 fw=10
    printalltypes
    n mean median std range
    IQR;
```

```
class Type;  
var Yield;  
run;
```

Your answer: c

Correct answer: c

The PROC MEANS statement must include the option PRINTALLTYPES in order for SAS to display statistics for all requested combinations of class variables – that is, for each level or occurrence of the variable and for all occurrences combined. The statistics specified on the second line must include the keywords N MEAN MEDIAN STD VAR RANGE Q RANGE. The code must specify **Type** as the class variable and **Yield** as the analysis variable.

Review: [The MEANS Procedure](#)



3. Read the following statement about the central limit theorem and choose the answer that contains the correct values for all of the missing fields.

The central limit theorem states that the distribution of sample __ (1) __ is approximately __ (2) __, regardless of the distribution of the population data, and this approximation improves as the sample size gets __ (3) __.

- a. means, skewed, larger
- b. variance, equal, smaller
- c. means, normal, larger
- d. proportions, equal, smaller

Your answer: c

Correct answer: c

The central limit theorem states that the distribution of sample means is approximately normal, regardless of the distribution of the population data, and this approximation improves as the sample size gets larger.

Review: [Normality and the Central Limit Theorem](#)



4. Psychologists at a college want to know if students are sleeping more or less than the recommended average of 8 hours a day.

Which of the following code choices correctly tests the null hypothesis?

- a.

```
proc univariate data=statdata.sleep mu0<>8;  
var hours;  
run;
```
- b.

```
proc univariate data=statdata.sleep;  
var hours / mu0=8;  
run;
```
- c.

```
proc univariate data=statdata.sleep;  
var hours / mu0<>8;  
run;
```
- d.

```
proc univariate data=statdata.sleep mu0=8;  
var hours;  
run;
```

Your answer: d

Correct answer: d

You specify the MU0= option as part of the PROC UNIVARIATE statement to indicate the test value of the null hypothesis. The alternative hypothesis is that μ is not equal to 8 hours, but

this does not need to be specified in the PROC UNIVARIATE code.

Review: [Using PROC UNIVARIATE to Generate a *t* Statistic](#)



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 ability of the statistical test to detect a true difference, or the ability to successfully reject a false null hypothesis. The probability of committing a Type I error is α . The probability of failing to reject the null hypothesis when it is actually false is a Type II error.

Review: [Types of Errors and Power](#)



6. Select the choice that lists only continuous variables.

- a. body temperature, number of children, gender, beverage size
- b. age, body temperature, gas mileage, income
- c. number of children, gender, gas mileage, income
- d. gender, gas mileage, beverage size, income

Your answer: b

Correct answer: b

The continuous variables are age, body temperature, gas mileage, and income.

Review: [Types of Variables: Quantitative and Categorical](#)



7. Which of the following code choices creates a histogram for the variable **Speed** from the data set **SpeedTest** with a normal curve overlay and a box with the skewness and kurtosis statistics printed in the northeast corner?

- a.

```
proc univariate data=statdata.speedtest;
    histogram Speed / normal(mu=est sigma=est);
    inset skewness kurtosis;
run;
```
- b.

```
proc univariate data=statdata.speedtest;
    histogram Speed / normal (mean std);
    inset skewness kurtosis / position=ne;
run;
```
- c.

```
proc univariate data=statdata.speedtest;
    histogram Speed / normal(mu=est sigma=est);
    inset skewness kurtosis / position=ne;
run;
```
- d.

```
proc univariate data=statdata.speedtest;
    histogram Speed / normal(skewness kurtosis);
run;
```

Your answer: c

Correct answer: c

In the HISTOGRAM statement, you specify the **Speed** variable and the NORMAL option using estimates of the population mean and the population standard deviation. In the INSET statement, you specify the keywords SKEWNESS and KURTOSIS, as well as the POSITION=NE option.

Review: [The UNIVARIATE Procedure](#)



8. 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 of cereal.
- a. You are 95% confident that the true average weight for a box of cereal is between 15.02 and 15.04 ounces.
 - b. The probability is .95 that the true average weight is between 15.02 and 15.04 ounces.
 - c. In the long run, approximately 95% of the intervals calculated with this procedure will capture the true average weight.

Your answer: b

Correct answer: b

A 95% confidence interval means that you are 95% confident that the interval contains the true population mean. If you sample repeatedly and calculate a confidence interval for each sample mean, 95% of the time your confidence interval will contain the true population mean. A confidence interval is not a probability. When a confidence interval is calculated, the true mean is in the interval or it is not. There is no probability associated with it.

Review: [Confidence Intervals](#)



9. 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)
 - d. none of the above

Your answer: c

Correct answer: c

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

Review: [Normal Distribution](#)



10. The standard error of the mean is
- a. used to calculate confidence intervals of the mean.
 - b. always normally distributed.
 - c. sometimes less than 0.
 - d. none of the above

Your answer: a

Correct answer: a

The standard error of the mean is part of the equation used to calculate a confidence interval of the mean. It is not normally distributed, and it is never less than 0.

Review: [Point Estimators, Variability, and Standard Error](#), [Interval Estimators](#)

Close

Copyright © 2017 SAS Institute Inc., Cary, NC, USA. All rights reserved.