| | |
|---|---|
| **Started on** | Wednesday, June 10, 2020, 3:13 PM |
| **State** | Finished |
| **Completed on** | Wednesday, June 10, 2020, 4:23 PM |
| **Time taken** | 1 hour 9 mins |
| **Grade** | **88.00** out of 100.00 |

Information

There are 32 questions. The first 20 are worth 2 points each. The last 12 are worth 5 points each. There will be some True/False, Multiple Choice, and Short Answer questions.

Question **1**

Correct

2.00 points out of 2.00

A stratified random sample

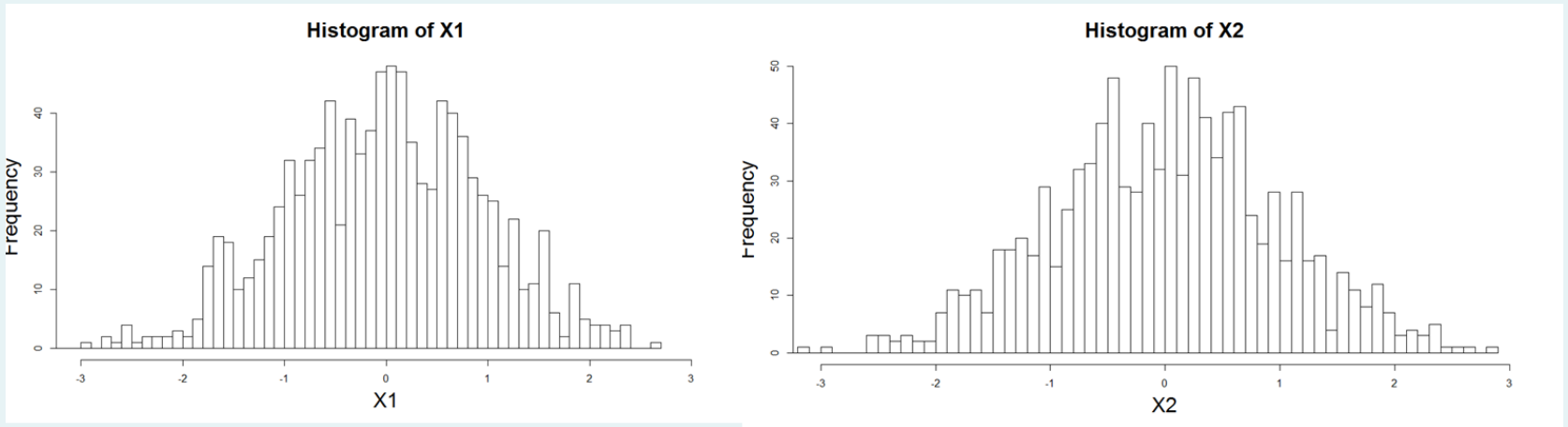
Select one:

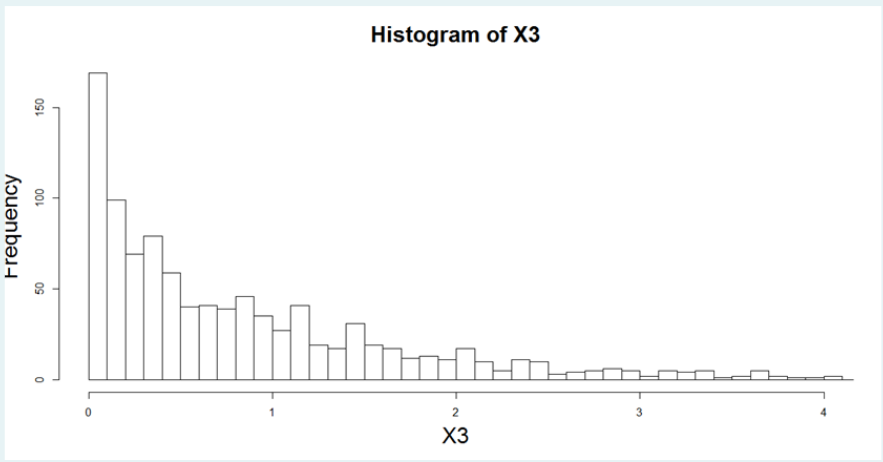
- a. Chooses samples from each "strata." ✔
- b. Often reflects the opinions of those with strong opinions.
- c. Is an example of volunteer response sampling.
- d. Will always produce skewed samples.

The correct answer is: Chooses samples from each "strata."

Information

Consider the following histograms of population variables labeled X1, X2, and X3. Use these to answer the next two questions.




Question 2

Correct

2.00 points out of 2.00

Which population distribution would the mean and median be furthest apart if all of the variables were put on the same scale?

Select one:

- a. X1
- b. X2
- c. X3 ✔
- d. It is impossible to tell.

The correct answer is: X3

Question 3

Incorrect

0.00 points out of 2.00

If you repeatedly took large samples, calculated their sample means, and graphed the resulting sampling distributions from each of these populations separately, then which of these sampling distributions is Normally distributed?

Select one:

- a. X1 and X2 only. ✖
- b. All of these populations (X1, X2, and X3).
- c. None of these because sample means never follow a Normal distribution.
- d. Need more information.

The correct answer is: All of these populations (X1, X2, and X3).

Question 4

Correct

2.00 points out of 2.00

Which statistic is resistant to outliers?

Select one:

- a. Variance
- b. Mean
- c. Correlation Coefficient
- d. None of the above. ✔

The correct answer is: None of the above.

Question 5

Correct

2.00 points out of 2.00

A standard deviation is bounded between -1 and 1 only.

Select one:

- True
- False ✔

The correct answer is 'False'.

Question 6

Incorrect

0.00 points out of 2.00

The sampling distribution of a statistic describes how possible values of the statistic vary

Select one:

- a. Within the selected sample.
- b. Among different possible samples.
- c. Within the population of interest.
- d. Between the sample and the population. ✖

The correct answer is: Among different possible samples.

Question 7

Correct

2.00 points out of 2.00

We create a 95% confidence interval for the mean. What does this indicate?

Select one:

- a. There is a 5% probability that the procedure will produce an interval that covers the sample mean.
- b. There is a 95% probability that the procedure will produce an interval that covers the sample mean.
- c. There is a 95% probability that the procedure will produce an interval that covers the population mean. ✔
- d. There is a 95% probability that the procedure will produce an interval that covers the standard error of the mean.

The correct answer is: There is a 95% probability that the procedure will produce an interval that covers the population mean.

Question 8

Correct

2.00 points out of 2.00

Confidence intervals values are different from sample to sample.

Select one:

- True ✔
- False

The correct answer is 'True'.

Question 9

Correct

2.00 points out of 2.00

A manager for a large food company was testing whether a machine that bags potato chips was producing bags that were either too full or too empty. To test this claim the manager conducted a two-sided hypothesis test for mean weight of bags off the line. The manager sampled 140 bags and concluded to not reject the null hypothesis and that the machine was working properly. While the machine went through regular maintenance, it was discovered it was working incorrectly and the manager had made the wrong decision. Which of the following is true?

Select one:

- a. The manager must have calculated something incorrectly because hypothesis tests are always correct if no calculation error is made.
- b. The manager committed a Type I error.
- c. The manager committed a Type II error. ✔
- d. The manager committed both a Type I and Type II error.

The correct answer is: The manager committed a Type II error.

Question 10

Incorrect

0.00 points out of 2.00

The p-value is calculated assuming that the alternative hypothesis is true.

Select one:

- True ✖
- False

The correct answer is 'False'.

Question 11

Correct

2.00 points out of 2.00

Which of the following is a characteristic of the F-distribution?

Select one:

- a. May take positive or negative values.
- b. Right-skewed. ✔
- c. Symmetric.
- d. Only has one set of degrees of freedom.

The correct answer is: Right-skewed.

Question 12

Incorrect

0.00 points out of 2.00

In simple linear regression, the sign of the slope coefficient must be the same as the sign of the correlation coefficient between independent and dependent variables.

Select one:

- True
- False ✖

The correct answer is 'True'.

Question 13

Complete

2.00 points out of 2.00

State two assumptions to the multiple linear regression model.

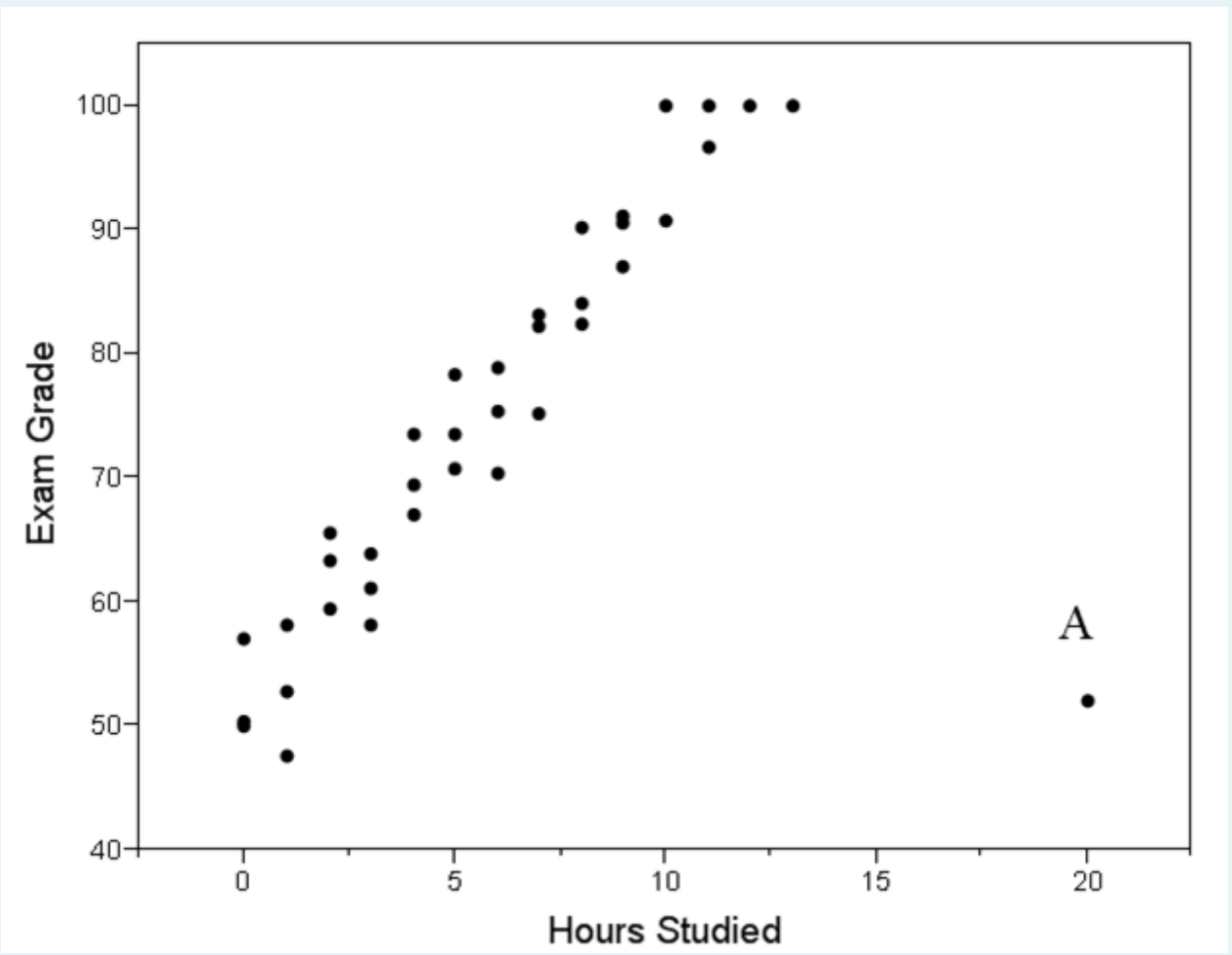
The probability distribution of errors is normal (Normality of Errors).

Errors associated with any two different observations are independent of each other (Independence of Errors).

Comment:

Information

An instructor of a college literature course records the grades of the students on the final exam for the course along with the students self-reported time spent on studying for the exam. Use the following plot to answer question 12.



Question 14

Correct

2.00 points out of 2.00

There is an observation in this data (marked with the letter "A"). If we removed this observation the following would happen to the correlation coefficient:

Select one:

- a. Get closer to zero.
- b. Get closer to one. ✔
- c. Stay the same.
- d. Need more information.

The correct answer is: Get closer to one.

Question 15

Incorrect

0.00 points out of 2.00

A data analyst is trying to interpret the following results of a multiple linear regression with an $R^2$ = 0.725.

| Parameter | Estimate | Std. Error | DF | T-statistic | P-value |
|---|---|---|---|---|---|
| Intercept | 78.21 | 33.53 | 68 | 2.333 | 0.0266 |
| X1 | -5.81 | 1.32 | 68 | -4.402 | < 0.0001 |
| X2 | 3.64 | 0.47 | 68 | 7.745 | < 0.0001 |

What is the value of the correlation coefficient, r, between the response variable and the first independent variable X1?

Select one:

- a. 0.851
- b. 0.725
- c. -0.851 ✖
- d. Not enough information to tell.

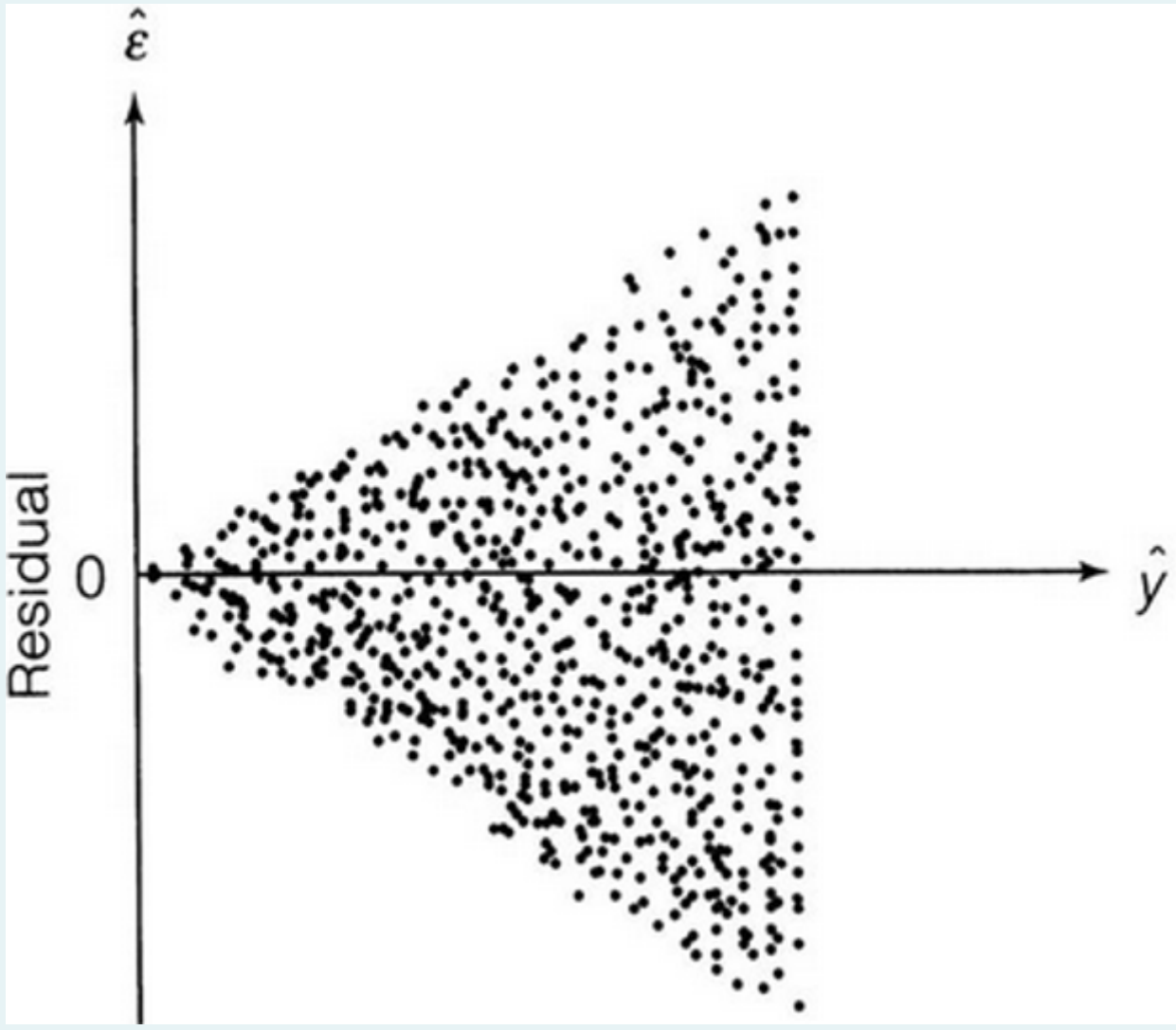
The correct answer is: Not enough information to tell.

Question 16

Correct

2.00 points out of 2.00

What assumption appears to be broken with this residual plot?




Select one:

- a. Perfect Multicollinearity
- b. Equal Variance of Errors ✔
- c. Normality of Errors
- d. Independence of Errors

Your answer is correct.

The correct answer is: Equal Variance of Errors

Question 17

Correct

2.00 points out of 2.00

Two sample hypothesis tests for means need the two populations to have equal variances or the analysis cannot be done.

Select one:

- True
- False ✔

The correct answer is 'False'.

Question 18

Complete

2.00 points out of 2.00

Name two assumptions for a One-way ANOVA.

Normally distributed categories and equality of variance between all categories.

Comment:

Question 19

Correct

2.00 points out of 2.00

Why do we use blocking in an ANOVA?

Select one:

- a. To isolate the effect of a certain variable by accounting for "nuisance" factors. ✔
- b. The main focus of the analysis is on the blocking variable, not the target variable.
- c. Blocking isn't done in an ANOVA, it is for regression.

Your answer is correct.

The correct answer is: To isolate the effect of a certain variable by accounting for "nuisance" factors.

Question 20

Correct

2.00 points out of 2.00

Which Chi-square test is designed to test an association between ordinal categorical variables?

Select one:

- a. Pearson Chi-square test.
- b. Mantel-Haenszel Chi-square test. ✔
- c. Likelihood Ratio Chi-square test.
- d. Cramer's V

The correct answer is: Mantel-Haenszel Chi-square test.

Information

A fast food restaurant is interested in answering some questions about different food options that were ordered by their customers. The restaurant calculated that 40% of all of the customers ordered hamburgers, 67% ordered a side of French fries, and 29% of all customers had both hamburgers and French fries. Use this information to answer the next two questions.

Question 21

Correct

5.00 points out of 5.00

What is the probability that any one random customer orders either a hamburger OR a side of French fries?

Answer: .78 ✔

The correct answer is: 0.78

Question 22

Correct

5.00 points out of 5.00

What is the probability that a customer orders French fries given that they already ordered a hamburger?

Answer: .725 ✔

The correct answer is: 0.725

Question 23

Complete

5.00 points out of 5.00

A large law firm's board of directors is brain-storming ideas for a commercial. One of the members of the board remembered that the success rate for all of the firm's cases was 68% and mentions that it should be used in the commercial. What is the probability that you randomly select 159 cases from this firm and more than 115 of them were successful?

0.121

Comment:

Information

A large retailer wants to determine the average sales (in millions of dollars) across all of their stores nationwide. However, it is too costly and time consuming to contact every store. The manager in charge of the task wants to use a sample size of 250 stores and does not know if the sales of the stores nationwide follow a Normal distribution. Use this information to answer the next three questions.

Question 24

Complete

5.00 points out of 5.00

Does the sample meet all of the needed assumptions to conduct a confidence interval? Explain.

The sample may meet the assumptions as long as np >= 5 and n(1-p)>=5.

Comment:

Question 25

Complete

5.00 points out of 5.00

An analyst working on the project thinks the sample of 250 stores is too big for the margin of error of $0.75 million specified by the vice president wanting to know the information. The analyst knows the standard deviation of previous studies of this type was $4.5 million. Calculate the estimated sample size needed for the desired margin of error and a 95% confidence interval.

139

Comment:

Question 26

Complete

5.00 points out of 5.00

The manager doesn't listen to the suggestion by the analyst and collects a sample of 250 stores with a sample average sales of $2.37 million with a sample standard deviation of $4.72 million. Calculate a 95% confidence interval for the true average sales of all stores.

(1.77, 2.96)

Comment:

Information

An analyst working for a hospital is trying to develop a model that predicts the average cost in dollars of a stay at the hospital for incoming patients. The analyst comes up with the following model from a sample of 68 patients. Use the following information to answer the following questions.

$x_1$ = Age of patients in years
$x_2$ = 1 if patient is male, 0 if patient is female
$x_3$ = 1 if patient has insurance, 0 if patient does not have insurance
$x_4$ = Number of visits to the hospital over the past year

y-hat = 647.12 + 21.57$x_1$ - 582.36$x_2$ - 1137.90$x_3$ + 185.22$x_4$

SSE = 418741, TSS = 948796

Question 27

Complete

4.00 points out of 5.00

Interpret the coefficient for $x_3$ in terms of the problem.

If a patient has insurance, the cost of staying at a hospital is expected to decrease by $1137.90.

Comment:
All else equal

Question 28

Complete

4.00 points out of 5.00

Interpret the coefficient for $x_4$ in terms of the problem.

The cost of staying at a hospital is expected to increase on average by $185.22 per visit to the hospital that year.

Comment:
All else equal

Question 29

Complete

5.00 points out of 5.00

Calculate the $R^2$ and adjusted-$R^2$ values for this model.

R^2 = 0.55866

R^2A = 0.5306

Comment:

Information

Use the following additional information to answer the following questions.

$R^2_3 = 0.24$

$\Sigma(x_{3,i}$- x-bar$_3)^2 = 196$

Question 30

Complete

5.00 points out of 5.00

Calculate the VIF for the variable $x_3$. Is there a problem of multicollinearity? Explain.

VIF = 1.31 Therefore, multicollinearity is not a problem because 1.31 < 5.

Comment:

Information

A researcher is conducting a study at a large university on the east cost. The researcher is trying to determine whether income bracket of parents of incoming freshman (separated into 4 group: lower class, middle class, upper-middle class, upper class) helps determine the overall GPA at the end of freshman year. The researcher samples 10 people from each income group to collect a total sample of 40. Assume all of the needed assumptions are met. Use the following information to answer following question.

| Source | DF | Sum of Squares | Mean Square | F-Ratio | P-value |
|---|---|---|---|---|---|
| Between | | | | | < 0.0001 |
| Within | | 38.2 | 1.06 | XXXXXXXXX | XXXXXXXXX |
| Total | 39 | 107.3 | XXXXXXXXX | XXXXXXXXX | XXXXXXXXX |
| XXXXXXXXX | XXXXXXXXX | XXXXXXXXX | XXXXXXXXX | XXXXXXXXX | XXXXXXXXX |

Question 31

Complete

5.00 points out of 5.00

What are the 5 missing values in the ANOVA table?

Between DF = 3

Within DF = 36

Between Sum of Sq = 69.1

Between Mean Sq = 23.03

Between F Ratio = 21.73

Comment:

Information

Below is data collected on the habits of both loyalty and non-loyalty customers at a store with regards to buying a new product. Use this table to answer the following questions.

| Customer Type | Bought New Product | Didn't Buy New Product | Total |
|---|---|---|---|
| Loyalty | 10 | 30 | 40 |
| Non-Loyalty | 5 | 25 | 30 |
| **Total** | **15** | **55** | **70** |

Question 32

Complete

5.00 points out of 5.00

Calculate the odds a loyalty customer buys the product, the odds a non-loyalty customer buys the product, and the odds ratio between them (loyalty vs. non-loyalty customer buying the product).

Odds a loyalty customer buys the product = .3333

Odds a non loyalty customer buys the product = 0.20

odds ratio = 1.665

Comment:

◄ Slide Deck 9 - Practice Problems

Jump to...

Final Statistics Assessment - Calculation Answers ►