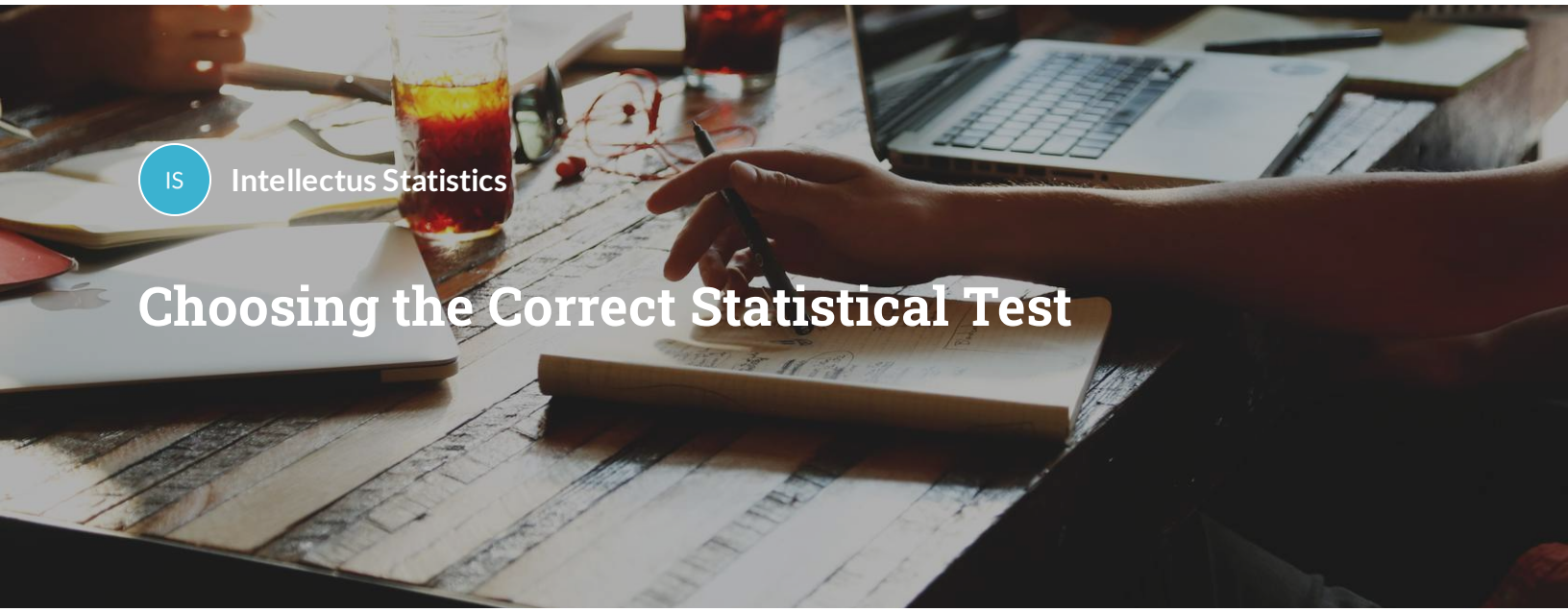


IS Intellectus Statistics

Choosing the Correct Statistical Test

Intellectus Statistics

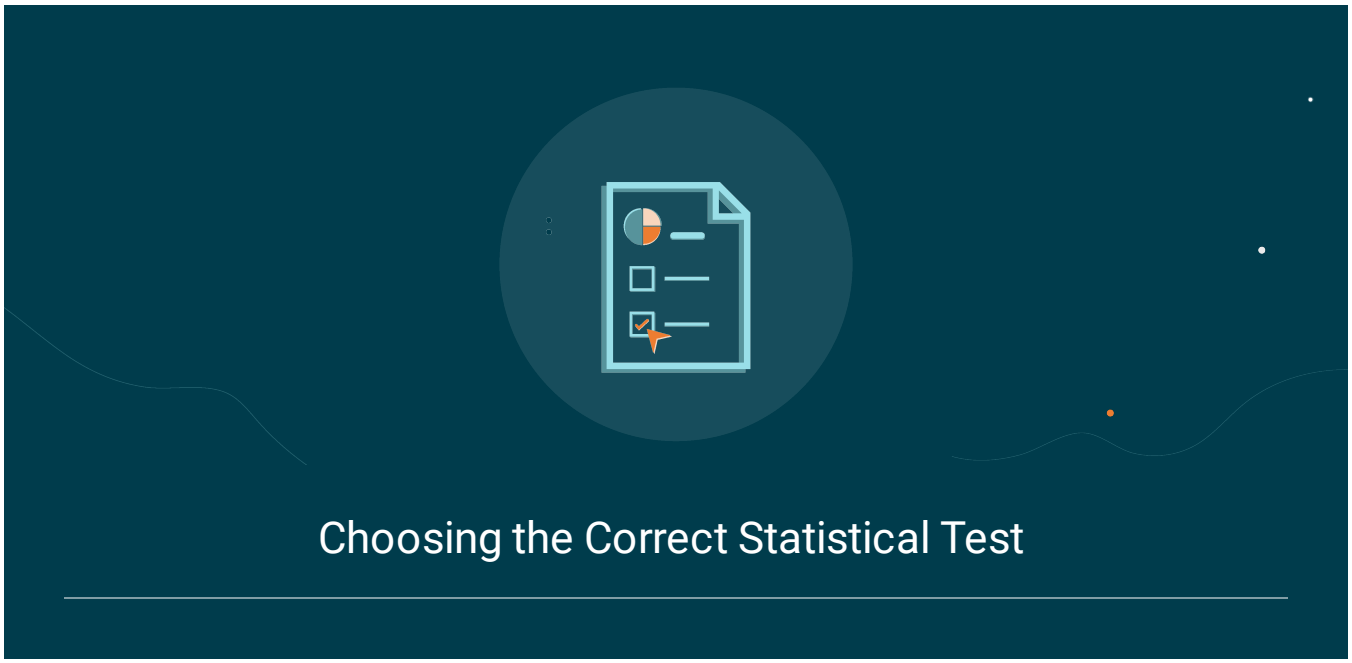
CHOOSING THE CORRECT STATISTICAL TEST

REVIEW

Lesson 1 of 6

Introduction

IS Intellectus Statistics



03:25

Narration.

Learning Objectives

1. Determine the most appropriate statistical test to use based on the research objective and nature of the data
2. Distinguish between independent samples and paired samples (or repeated measures) tests
3. Distinguish between parametric and non-parametric tests

CONTINUE

Types of Variables

Before completing this lesson, it is important to understand the difference between dependent variables, independent variables, and covariates.

Dependent variable

A dependent variable is an outcome that the researcher is interested in explaining or changing. They are referred to as “dependent” because the researcher proposes that the value of a dependent variable will depend on the values of other variables.

Independent variable

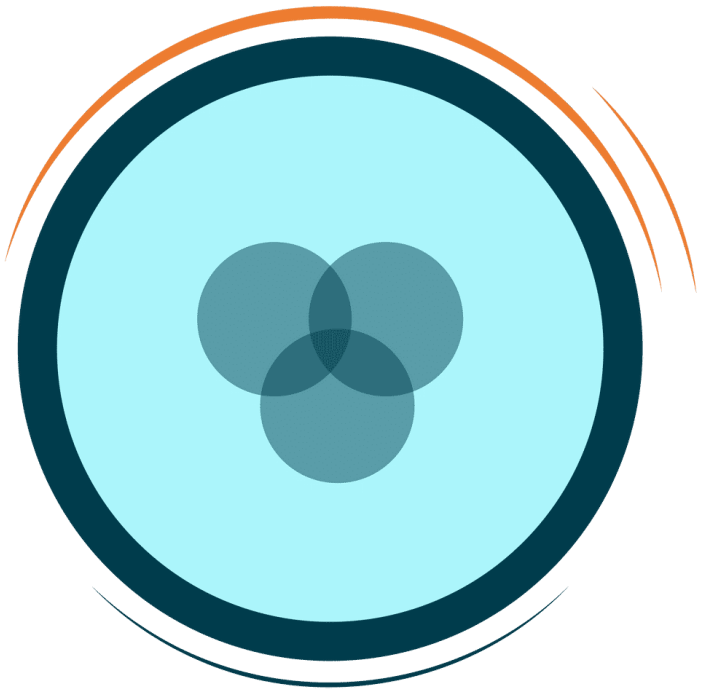
An independent variable is a factor, characteristic, or intervention that the researcher believes may explain, affect, or change the values of the dependent variable. They are referred to as “independent” because the researcher proposes that the value of an independent variable will not depend on the values of any of the other variables being studied.

Covariate or control variable

A covariate or control variable is a factor or characteristic that is not of primary interest, but the researcher believes it may affect the values of the dependent variable and/or covary with the independent variables.

For more information on distinguishing between these types of variables, see the Types of Variables lesson.

CONTINUE



Nominal measures represent categories that CANNOT be ranked or ordered.

- Examples: gender, race/ethnicity, location, region, nationality, occupation

Ordinal measures represent categories that can be ranked or ordered (e.g., from lowest to highest).

- Examples: education (high school, college, graduate school), letter grade (A, B, C, D, F), Likert items (strongly disagree to strongly agree)

Scale measures represent continuous numerical values. Subtypes of scale measures include ratio and interval.

- **Ratio**: Numeric values with a meaningful (non-arbitrary) zero
    - Examples: age (in years), income, height, weight, counts

- **Interval**: Numeric values with no meaningful (arbitrary) zero
    - Examples: temperature (in C/F), dates/years, scale scores and composite measurements

CONTINUE

How to Choose the Correct Statistical Test

The most appropriate statistical test to perform will depend on two main factors: the research **objective** and the **nature of the data.**

The objective of the research is typically the research question you are trying to answer and/or the hypothesis you are testing.

The nature of the research question or hypothesis may determine the general analysis approach to use. Analyses may:

- Summarize or describe the data
- Determine differences between variables
- Determine relationships between variables
- Predict outcomes

Most research questions can be answered through one of the above general analysis approaches.

The **nature of the data** may determine the specific statistical test within an analysis approach to use. Considerations about the nature of the data include:

- The **levels of measurement** of the variables
- Whether the data reflect **independent measurements or paired/repeated measurements**
- If there are **covariate/control variables** that need to be accounted for

CONTINUE

Research Objectives

IS Intellectus Statistics

03:29

Narration.

Summarizing or Describing Data

When the objective of the research is to determine the range, variability, magnitude, frequency, or prevalence of variables, then an analysis that **summarizes or describes the data** is appropriate.

If the objective of the research is to summarize or describe data, there may be a research question, but often there are no null and alternative hypotheses. Examples of research questions about summarizing/describing data:

- What proportion of applicants for entry-level software engineering positions are women?
- What is the prevalence of generalized anxiety disorder among graduate students in the United States?

Descriptive statistics is the appropriate analysis approach for this objective.

Determining Differences Between Variables

An analysis that **determines differences between variables** is appropriate when the researcher is interested in:

- Differences between groups of participants (e.g., females vs. males) on a variable
- Differences between repeated measurements of a variable (e.g., pretest vs. posttest)
- Differences between conditions in an experiment (e.g., treatment group vs. control group)


English
My name is
NIKITA

If the objective of the research is to determine differences between variables, there may be a research question and/or null and alternative hypotheses.

The research question/hypothesis will typically be about a **difference, change, increase, or decrease in a variable.**

Examples of research questions about differences between variables:

- Is there a difference in transformational leadership between CEOs of publicly traded companies and CEOs of privately owned companies?
- Does participation in an after-school reading program result in higher end-of-year English test scores?

A **test of differences** is the appropriate analysis approach for this objective.

Determining Relationships Between Variables

An analysis that **determines relationships between variables** is appropriate when the researcher is interested in the association or correlation between two variables.

If the objective of the research is to determine relationships between variables, there may be a research question and/or null and alternative hypotheses.

The research question/hypothesis will typically be about a **relationship, association, or correlation between variables.**

Examples of research questions about relationships between variables:

- Is there a relationship between self-compassion and burnout among workers in assisted-living facilities?

- Is there a positive correlation between general intelligence and emotional intelligence?

A **test of association or correlation** is the appropriate analysis approach for this objective.

Predicting Outcomes

An analysis that **predicts outcomes** is appropriate when the researcher wants to use one or more independent variables to explain or predict a dependent variable.

If the objective of the research is to predict outcomes, there may be a research question and/or null and alternative hypotheses.

The research question/hypothesis will typically be about prediction. The research question could also be about the **relationship between multiple independent variables and a single dependent variable.**

Examples of research questions about predicting outcomes:

- To what extent do gender, race, age, and education level predict people's willingness to invest in cryptocurrency?
- Are pay, benefits, and work hours significantly related to employee job satisfaction?

A **regression** is the appropriate analysis approach for this objective.

CONTINUE

Nature of the Data

IS Intellectus Statistics

02:17

Narration.

Considering the Nature of the Data

Once the general analysis approach has been identified based on the research objective, the specific statistical test may be determined based on the nature of the data. Factors to consider are the levels of measurement, whether the data reflect independent or repeated measurements, and if there are covariate/control variables.

Levels of Measurement

The **levels of measurement** of the dependent and independent variables partly determine the specific test to use. Different statistical tests require the variables to be specific levels of measurement. Generally...

- There are different descriptive statistics that are appropriate for scale, ordinal, and nominal variables.
- **Parametric tests** (such as t-tests and ANOVAs) require the dependent variable to be a scale level of measurement. If the dependent variable is ordinal or nominal, **a non-parametric test** should be used. Non-parametric tests can be considered if the data for a parametric test do not meet the normality assumption.
- There are different types of regression that are appropriate for scale, ordinal, and nominal dependent variables.

Independent or Repeated Measurements

The specific test to use also depends on whether the data consists of independent or repeated measurements.

- If each variable is measured only once for each subject/participant, then the data consists of independent measurements. These data are analyzed using independent samples tests.
- If a variable is **measured more than once** for each subject/participant, then the data consists of repeated measurements. This is the case in pretest/posttest studies or longitudinal studies. These data are analyzed using paired samples or repeated measures tests.
- Measurements from different subjects/participants that can be paired or matched in some way (e.g., siblings, spouses) may also be analyzed using paired or repeated measures tests.

Covariate/Control variables

If the research calls for **covariate/control variables** to be analyzed, then the statistical test must permit the inclusion of covariates.

- For tests of differences, analysis of covariance (ANCOVA) and its variants allow the inclusion of covariates.
- For tests of relationships or prediction, partial correlation and regression allow the inclusion of covariates.

CONTINUE

Choosing Tests

IS Intellectus Statistics

04:22

Narration.

Descriptive Statistics

The appropriate descriptive statistics to perform depend on the variable's level of measurement.

- **Scale** variables: mean, standard deviation, median, mode, minimum, maximum, skewness, kurtosis
- **Ordinal** variables: frequencies, median, mode, minimum, maximum
- **Nominal** variables: frequencies, mode

CONTINUE

Tests of Differences

The appropriate test of differences to perform depends on the level of measurement of the variables and if there are repeated measurements.

If the **dependent variable is a scale level of measurement** and **the measurements are independent**, then independent samples parametric tests may be performed. These tests include the following:

- **Independent samples t-test:** used to test for differences between two groups (i.e., a binary independent variable)
- **Analysis of variance (ANOVA):** used to test for differences between two or more groups
- **Multivariate analysis of variance (MANOVA):** used to test for differences between two or more groups on multiple dependent variables
- **Analysis of covariance (ANCOVA or MANCOVA):** variants of ANOVA and MANOVA that permit the inclusion of covariates

If the **dependent variable is a scale level of measurement** and **there are repeated measurements**, then repeated measures parametric tests may be performed. These tests include the following:

- **Paired samples t-test:** used to test for differences between two measurements (e.g., pretest vs. posttest)
- **Repeated Measures ANOVA:** used to test for differences between two or more measurements
- **Repeated Measures MANOVA:** used to test for differences between two or more measurements on multiple dependent variables
- **Repeated Measures ANCOVA or MANCOVA:** variants of repeated measures ANOVA and MANOVA that permit the inclusion of covariates

If the **dependent variable is an ordinal level of measurement,** then non-parametric tests should be performed. These tests include the following:

- **Mann-Whitney Test: Used to test for differences between two groups. The dependent variable may be ordinal or scale. This is the non-parametric version of the independent samples t-test.**
- **Kruskal-Wallis Test:** Used to test for differences between two or more groups. The dependent variable may be ordinal or scale. This is the non-parametric version of ANOVA.
- **Wilcoxon Signed Rank Test:** Used to test for differences between two measurements. The dependent variable may be ordinal or scale. This is the non-parametric version of the paired samples t-test.
- **Friedman Test:** Used to test for differences between two or more measurements. The dependent variable may be ordinal or scale. This is the non-parametric version of repeated measures ANOVA.

CONTINUE

Tests of Association/Correlation

The appropriate correlation or test of association to perform depends on the level of measurement of the variables.

- **Pearson correlation:** used to test for an association between two scale level variables
- **Point biserial correlation:** used to test for an association between one binary variable and one scale level variable
- **Spearman correlation:** used to test for an association between two variables that are either ordinal or scale. This is the non-parametric version of the Pearson correlation.
- **Partial correlation:** used to test for an association between two scale-level variables while controlling for covariates.

- **Chi-Square Test of Independence:** used to test for an association between two nominal or ordinal variables. This is a non-parametric test.

CONTINUE

Tests of Prediction/Regressions

The appropriate regression to perform depends on the level of measurement of the dependent (criterion) variable. Note that in a regression analysis, the independent (predictor) variables can be any level of measurement. Also, covariates may be included in all types of regression analyses.

- **Linear regression:** used to predict a scale level dependent variable
- **Ordinal logistic regression:** used to predict an ordinal level dependent variable
- **Binary logistic regression:** used to predict a binary (nominal level) dependent variable. The dependent variable must have only two possible values.
- **Multinomial logistic regression:** used to predict a nominal level dependent variable. The dependent variable may have more than two possible values.

CONTINUE

Knowledge Check

IS Intellectus Statistics

Question

01/10

Multiple Choice

What are the two main factors that determine the appropriate statistical test to use?

- ◯ The objective of the research and the size of the population
- ◯ The objective of the research and the nature of the data
- ◯ The nature of the data and the size of the population
- ◯ The size of the population and the number of variables

Question

02/10

Multiple Choice

Which of the following is an aspect of the data that a researcher should consider in choosing a statistical test?

- ◯ Whether or not the data contain repeated measurements
- ◯ Whether there are covariates in the data that need to be controlled for
- ◯ The levels of measurement of the variables
- ◯ All of the above should be considered in choosing a statistical test

Question

03/10

Multiple Choice

Which of the following statements is true?

- The same types of descriptive statistics may be used for all levels of measurement.
- A parametric test should only be used when the dependent variable is ordinal or nominal.
- Independent samples tests are used when each variable is measured only once for each subject or participant.
- All of the above statements are true.

Question

04/10

Multiple Choice

Choose the most appropriate statistical test for the following scenario:

A teacher wants to determine if there is a difference between two classrooms in their SAT scores. Assume that SAT score is a scale level of measurement and that each student takes the SAT only once.

- ◯ Independent samples t-test
- ◯ Paired samples t-test
- ◯ ANCOVA
- ◯ Binary logistic regression

Question

05/10

Multiple Choice

Choose the most appropriate statistical test for the following scenario:

A teacher wants to determine if there is a difference between two classrooms in their SAT scores while controlling for gender and GPA. Assume that SAT score is a scale level of measurement, and that each student takes the SAT only once.

- ◯ Independent samples t-test
- ◯ Paired samples t-test
- ◯ ANCOVA
- ◯ Binary logistic regression

Question

06/10

Multiple Choice

Choose the most appropriate statistical test for the following scenario:

A teacher wants to determine if there is an improvement in her students' SAT scores the second time they take the test. Assume that SAT score is a scale level of measurement, and that each student takes the SAT twice.

- ◯ Independent samples t-test
- ◯ Paired samples t-test
- ◯ ANCOVA
- ◯ Binary logistic regression

Question

07/10

Multiple Choice

Choose the most appropriate statistical test for the following scenario:

A researcher wants to determine if there is an association between education level and starting salary among software engineers. Assume that education level is an ordinal level of measurement, and that starting salary is a scale level of measurement.

- ◯ Partial correlation
- ◯ Pearson correlation
- ◯ Spearman correlation
- ◯ Chi-square test of independence

Question

08/10

Multiple Choice

Choose the most appropriate statistical test for the following scenario:

A researcher wants to determine if there is an association between education level and job department among employees at software development companies. Assume that education level is an ordinal level of measurement, and that job department is a nominal level of measurement.

- ◯ Partial correlation
- ◯ Pearson correlation
- ◯ Spearman correlation
- ◯ Chi-square test of independence

Question

09/10

Multiple Choice

Choose the most appropriate statistical test for the following scenario:

A meteorologist wants to predict whether or not it will rain based on daily temperature, humidity, and atmospheric pressure levels. Assume that the dependent variable is a nominal level of measurement with two categories (rain or no rain), and that temperature, humidity, and pressure are all scale levels of measurement.

- ◯ Linear regression
- ◯ Ordinal logistic regression
- ◯ Binary logistic regression
- ◯ Pearson correlation

Question

10/10

Multiple Choice

Choose the most appropriate statistical test for the following scenario:

A meteorologist wants to predict the amount of rainfall (in inches) based on daily temperature, humidity, and atmospheric pressure levels. Assume that rainfall, temperature, humidity, and pressure are all scale levels of measurement.

- ◯ Linear regression
- ◯ Ordinal logistic regression
- ◯ Binary logistic regression
- ◯ Pearson correlation

Summary

IS Intellectus Statistics

Types of Variables

Concept	Description	Additional Information
Dependent Variable	A dependent variable is an outcome that the researcher is interested in explaining or changing.	They may be referred to as criterion variables in predictive research. The appropriate statistical test may depend on the level of measurement of the dependent variable.
Independent Variable	An independent variable is a factor, characteristic, or intervention that the researcher believes may explain, affect, or change the values of the dependent variable.	They may be referred to as predictor variables in predictive research.
Covariate or Control Variable	A covariate or control variable is a factor or characteristic that is not of primary interest, but the researcher believes it may affect the values of	Some statistical tests do (or do not) permit the inclusion of covariates.

Concept	Description	Additional Information
	the dependent variable and/or covary with the independent variables.	

CONTINUE

Levels of Measurement

Concept	Description	Additional Information
Scale	Represent continuous numerical values.	Sublevels include ratio and interval. Appropriate descriptive statistics include mean, standard deviation, median, mode, minimum, maximum, skewness, and kurtosis.
Ordinal	Represent categories that can be ranked or ordered (e.g., from lowest to highest).	Appropriate descriptive statistics include frequencies, median, mode, minimum, and maximum.
Nominal	Represent categories that cannot be ranked or ordered.	Appropriate descriptive statistics include frequencies and mode.

CONTINUE

Research Objectives

Concept	Description	Additional Information
Descriptive statistics	A type of analysis used to summarize or describe the data.	Used when the objective of the research is to determine the range, variability, magnitude, frequency, or prevalence of variables.
Test of differences	A type of analysis used to determine differences between variables.	Used to test differences between groups of participants, between repeated measurements, or between conditions in an experiment. The research question/hypothesis will typically be about a difference, change, increase, or decrease in a variable.
Test of association/correlation	A type of analysis used to determine relationships between variables.	Used to address a research question/hypothesis about a relationship, association, or correlation between variables.
Test of prediction/regression	A type of analysis used to predict outcomes.	Used to address a research question/hypothesis about prediction. The research question could also be about the relationship

Concept	Description	Additional Information
		between multiple independent variables and a single dependent variable.

CONTINUE

Types of Tests and Measurements

Concept	Description	Additional Information
Parametric test	A type of test that makes underlying assumptions about the distribution of data in the population.	Generally may be used when the dependent variable is a scale level of measurement.
Non-parametric test	A type of test that does not make underlying assumptions about the distribution of data in the population.	Generally should be used when the dependent variable is an ordinal or nominal level of measurement.
Independent measurements	Refer to data in which each variable is measured only once for each subject/participant.	These data are analyzed using independent samples tests.
Repeated measurements	A type of analysis used to predict outcomes. Refer to data in which a	Applies to pretest/posttest studies or longitudinal studies. These data are

Concept	Description	Additional Information
	variable is measured more than once for each subject/participant.	analyzed using paired samples or repeated measures tests.

CONTINUE

Test for Differences

If the dependent variable is a scale level of measurement and the measurements are independent.

Concept	Description	Additional Information
Independent samples t-test	A parametric test used to test for differences between two groups on a scale level dependent variable.	The dependent variable must be scale and the data must consist of independent measurements.
Analysis of variance (ANOVA)	A parametric test used to test for differences between two or more groups on a scale level dependent variable.	The dependent variable must be scale and the data must consist of independent measurements.
Multivariate analysis of variance (MANOVA)	A parametric test used to test for differences between two or more groups on multiple scale level dependent variables.	The dependent variables must be scale and the data must consist of independent measurements.

Concept	Description	Additional Information
Analysis of covariance (ANCOVA and MANCOVA)	Variants of ANOVA and MANOVA that permit the inclusion of covariates.	The dependent variable(s) must be scale and the data must consist of independent measurements.

CONTINUE

Test for Differences

If the dependent variable is a scale level of measurement and there are repeated measurements.

Concept	Description	Additional Information
Paired samples t-test	A parametric test used to test for differences between two measurements of a scale level dependent variable.	The dependent variable must be scale and the data must consist of repeated measurements
Repeated measures ANOVA	A parametric test used to test for differences between two or more measurements of a scale level dependent variable.	The dependent variable must be scale and the data must consist of repeated measurements.
Repeated measures MANOVA	A parametric test used to test for differences between two or more	The dependent variables must be scale and the data

Concept	Description	Additional Information
	measurements of multiple scale level dependent variables.	must consist of repeated measurements.
Repeated measures ANCOVA and MANCOVA	Variants of repeated measures ANOVA and MANOVA that permit the inclusion of covariates.	The dependent variable(s) must be scale and the data must consist of repeated measurements.

CONTINUE

Test for Differences

If the dependent variable is an ordinal level of measurement, the non-parametric tests should be used.

Concept	Description	Additional Information
Mann-Whitney test	A non-parametric test used to test for differences between two groups on an ordinal or scale dependent variable.	This is the non-parametric version of the independent samples t-test.
Kruskal-Wallis test	A non-parametric test used to test for differences between two or more groups on an ordinal or scale dependent variable.	This is the non-parametric version of ANOVA.

Concept	Description	Additional Information
Wilcoxon Signed Rank test	A non-parametric test used to test for differences between two measurements of an ordinal or scale dependent variable.	This is the non-parametric version of the paired samples *t*-test.
Friedman test	A non-parametric test used to test for differences between two or more measurements of an ordinal or scale dependent variable.	This is the non-parametric version of repeated measures ANOVA.

CONTINUE

Tests of Association/Correlation

Concept	Description	Additional Information
Pearson correlation	Used to test for an association between two scale level variables.	This is a parametric test of association.

Concept	Description	Additional Information
Point biserial correlation	Used to test for an association between one binary variable and one scale level variable.	This is a parametric test of association.
Spearman correlation	Used to test for an association between two variables that are either ordinal or scale.	This is the non-parametric version of the Pearson correlation.
Partial correlation	Used to test for an association between two scale level variables while controlling for covariates.	This is a correlation that permits the inclusion of covariates.
Chi-square test of independence	A non-parametric test used to test for associations between categorical variables.	The variables being tested must be categorical (ordinal or nominal).

CONTINUE

Tests of Prediction/Regression

Concept	Description	Additional Information
Linear regression	Used to predict a scale level dependent variable.	The independent (predictor) variable(s) may be any level

Concept	Description	Additional Information
		of measurement.
Ordinal logistic regression	Used to predict an ordinal level dependent variable.	The independent (predictor) variable(s) may be any level of measurement
Binary logistic regression	Used to predict a binary (nominal level) dependent variable.	The dependent variable must have only two possible values. The independent (predictor) variable(s) may be any level of measurement.
Multinomial logistic regression	Used to predict a nominal level dependent variable.	The dependent variable may have more than two possible values. The independent (predictor) variable(s) may be any level of measurement.

CONTINUE

Congratulations!

You have completed this lesson.