Statistical analysis

(Basic, classified data, unclassified data)

LEC:1

OLA ALI

Statistical analysis means investigating trends, patterns, and relationships using quantitative data.

It is an important research tool used by scientists, governments, businesses, and other organizations.

To draw valid conclusions, statistical analysis requires careful planning from the very start of the research process.

You need to specify your hypotheses and make decisions about your research design, sample size, and sampling procedure.

After collecting data from your sample, you can organize and summarize the data using descriptive statistics.

Then, you can use inferential statistics to formally test hypotheses and make estimates about the population. Finally, you can interpret and generalize your findings.

This article is a practical introduction to statistical analysis for students and researchers.

We'll walk you through the steps using two research examples.

The first investigates a potential cause-and-effect relationship, while the second investigates a potential correlation between variables.

Example: Causal research questionCan meditation improve exam performance in teenagers?

Example: Correlational research questionIs there a relationship between parental income and college grade point average (GPA)

Table of contents

To collect valid data for statistical analysis, you first need to specify your hypotheses and plan out your research design.

Writing statistical hypotheses

The goal of research is often to investigate a relationship between variables within a population.

You start with a prediction, and use statistical analysis to test that prediction.

A statistical hypothesis is a formal way of writing a prediction about a population.

Every research prediction is rephrased into null and alternative hypotheses that can be tested using sample data.

While the null hypothesis always predicts no effect or no relationship between variables, the alternative hypothesis states your research prediction of an effect or relationship.

Example: Statistical hypotheses to test an effect

Null hypothesis: A 5-minute meditation exercise will have no effect on math test scores in teenagers.

Alternative hypothesis: A 5-minute meditation exercise will improve math test scores in teenagers.

Example: Statistical hypotheses to test a correlation

Null hypothesis: Parental income and GPA have no relationship with each other in college students.

Alternative hypothesis: Parental income and GPA are positively correlated in college students.

Planning your research design

A research design is your overall strategy for data collection and analysis. It determines the statistical tests you can use to test your hypothesis later on.

First, decide whether your research will use a descriptive, correlational, or experimental design. Experiments directly influence variables, whereas descriptive and correlational studies only measure variables.

In an experimental design, you can assess a cause-and-effect relationship (e.g., the effect of meditation on test scores) using statistical tests of comparison or regression.

In a correlational design, you can explore relationships between variables (e.g., parental income and GPA) without any assumption of causality using correlation coefficients and significance tests.

In a descriptive design, you can study the characteristics of a population or phenomenon (e.g., the prevalence of anxiety in U.S. college students) using statistical tests to draw inferences from sample data.

Your research design also concerns whether you'll compare participants at the group level or individual level, or both.

In a between-subjects design, you compare the group-level outcomes of participants who have been exposed to different treatments (e.g., those who performed a meditation exercise vs those who didn't).

In a within-subjects design, you compare repeated measures from participants who have participated in all treatments of a study (e.g., scores from before and after performing a meditation exercise).

xample: Experimental research designYou design a within-subjects experiment to study whether a 5-minute meditation exercise can improve math test scores. Your study takes repeated measures from one group of participants.

First, you'll take baseline test scores from participants. Then, your participants will undergo a 5-minute meditation exercise. Finally, you'll record participants' scores from a second math test.

In this experiment, the independent variable is the 5-minute meditation exercise, and the dependent variable is the change in math test scores from before and after the intervention.

Measuring variables

When planning a research design, you should operationalize your variables and decide exactly how you will measure them.

For statistical analysis, it's important to consider the level of measurement of your variables, which tells you what kind of data they contain:

Categorical data represents groupings. These may be nominal (e.g., gender) or ordinal (e.g. level of language ability).

Quantitative data represents amounts. These may be on an interval scale (e.g. test score) or a ratio scale (e.g. age).

Many variables can be measured at different levels of precision. For example, age data can be quantitative (8 years old) or categorical (young). If a variable is coded numerically (e.g., level of agreement from 1–5), it doesn't automatically mean that it's quantitative instead of categorical.

Identifying the measurement level is important for choosing appropriate statistics and hypothesis tests. For example, you can calculate a mean score with quantitative data, but not with categorical data.

In a research study, along with measures of your variables of interest, you'll often collect data on relevant participant characteristics.