UDC: 311.213.3:519.233.3
Professional paper

STATISTICAL TESTS FOR DEPENDENT AND INDEPENDENT SOLUTIONS AND THEIR APPLICATION

Rushadije RAMANI[1*], Anduena HAZIRI[1], Sejhan IDRIZI[1], Mirlinda SHAQIRI[1], Lazim KAMBERI[1], Alit IBRAIMI[1]

[1]*University of Tetovo, Faculty of Mathematical and Natural Sciences, Department of Mathematics, Str. Ilinden, n.n., 1200, Tetovo, Republic of North Macedonia*
**Corresponding Author: lazim.kamberi@unite.edu.mk*

Abstract

Statistical tests for dependent and independent solutions are the main ways for testing the fit between a hypothesis and the data in many fields of science. This paper is focused on reviewing these tests, analyzing the theoretical principles that these tests support and their concrete application. The tests that include independent solutions are used when two variables are not correlated, and their goal is to review if there are significant differences between different groups. Two main tests for these kinds of problems are the Chi- square test for independence, a test which assesses the relationship between two categorical variables, and the t- test for independent groups, which is used for comparison of the means of two independent groups. While, tests for dependent variables, analyze related variables, and are useful to review the changes between the same measurements over time, or to review the influence of a variable on another. The t-test for dependent variables and the linear regression model are examples of dependent solutions, which are used to compare measurements before and after this test, or to analyze the relationship between a dependent variable and several independent variables. This paper illustrates the use of these tests on different research studies, including medical research, social research, and economic research, emphasizing the importance of choosing the right test in order to achieve accurate and reliable results. In conclusion, the use of these statistical methods is fundamental to achieving valid conclusions and support in making decisions in various fields of study.

Keywords: Dependent and independent variables, statistical tests, Chi-square test, t-test, application.

1. Introduction

Statistical tests are fundamental tools for analyzing data and drawing valid conclusions in various research and studies. These tests are used to evaluate hypotheses and to check relationships between variables. In this paper, we will examine two main types of statistical tests: independent sample tests and dependent sample tests, as well as their applications in different contexts. Independent sample tests, such as the chi-square test and the independent group t-test, are used to compare different groups that are not related to each other, while dependent sample tests, such as the t-test and linear regression for dependent samples, are applied to analyze relationships between related (dependent) variables. This paper emphasizes the importance of these tests for testing hypotheses and drawing valid conclusions from data, providing concrete examples of applications in fields such as the social sciences, health, and economics. Statistical tests are fundamental tools for improving analysis and supporting decision-making in many fields, guaranteeing accurate and reliable results.
(In politics and social sciences, the terms, independent choice" and, dependent choice" are used to describe how individuals or groups make decisions in different contexts.)

2. Materials and Methods

On this papper, to solve problems that have to do with dependent or independent variables, are used some statistical tests such as the chi-square test and the independent group t-test, who are used to compare different groups that are not related to each other, while for dependent samples, is used the t-test and linear regression for dependent samples, that are applied to analyze relationships between related (dependent) variables.

3. Results and Discussion

By independent choice, we imply a choice that is made without influencing or being influenced by the choices of others. The decision maker makes the decision based on his/her own preferences, interests, or information, without being influenced by external factors.
For example, a person decides to vote for a political party solely based on its program and not because his family or friends vote for the same party.
By dependent choice, we imply a choice that is influenced by the choices or actions of others. The decision may be based on social pressure, media influence, cultural tradition, or information coming from external sources.
For example, a person decides to buy a product because most of his friends have bought and are using it.
In data analysis and statistics, the terms, independent" and, dependent" have a very specific meaning and are very important, especially in statistical modeling, such as in linear regression, correlation analysis, etc.

3.1. Independent "choice": When we talk about "independent choice" in statistics, we are not talking about selection in the sense of personal decisions, but about the way the data were obtained — specifically, about the statistical independence of the data or observations. That is, it is what we do not control or what is assumed not to affect another variable.
What is an independent choice in statistics? An independent choice (selection or observation) means that it is not influenced by and does not affect other choices in the same data set. That is, each observation is independent of the others. In experiments, it is the one that is not manipulated by the researcher. It is the" cause" in a cause-and-effect relationship.
Why is this important? In many statistical analyses (such as regression, t-tests, ANOVA, etc.), independence of choices is an essential condition. If the data are not independent, the results can be inaccurate or misleading.
Some examples of independent choices:
Example: If we are analyzing how study time affects a student's grade: Study time is the independent variable.
Example: Imagine an experiment where:

- 100 different people, randomly selected, are asked about their eating habits.
- Each person answers only for himself.

These answers are independent choices, because one answer does not affect and is not affected by another.
When variables are independent, this means that there is no relationship between them, and one variable is not related to another. Independent choice tests are used to compare two different groups that do not have any direct relationship between them.

3.1.1. The chi-square test for independence: This test is used to verify whether two categorical variables are independent of each other.

For example, it can be used to analyze whether there is a relationship between two categorical variables, such as " gender" and "political preferences". Chi-square test formula:

$$\chi^2 = \sum_{i,j} \frac{\left(x_{ij} - e_{ij}\right)^2}{e_{ij}}$$

Application: Let's suppose we want to analyze whether there is a relationship between gender and political preferences of a population. We use the Chi-square test to check whether the separation between gender and political preferences is random or whether there is a relationship between them.

3.1.2. The t-test for independence: This test is used to compare the means of two independent groups. For example, it can be used to compare the average expenses between two different citizens.
The t-test formula:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

If we assume equal variances, we use the common variance:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where,

$$s_p^2 = \frac{\left(n_1 - 1\right)s_1^2 + \left(n_2 - 1\right)s_2^2}{n_1 + n_2 - 2}$$

Application: Suppose you want to compare the average spending for two groups of consumers living in two different cities. Use the t-test to assess whether there is a significant difference between the groups.

3.2. Dependent "choice": A dependent variable (or observation) is affected by or related to other variables in the data set. This statistical dependence means that the results of one piece of data have an impact on or are closely related to other pieces of the same data set. So, the dependent variable is the one that is measured and affected by the independent variables. It is also called the outcome variable. It is the effect in the cause-and-effect relationship.
Some examples of dependent choices:
Example: If we measure a person's blood pressure before and after a treatment, these two measurements are:

- These measurements are dependent choices, because they both come from the same person, and the result after the treatment is influenced by the previous one.

Example: If the same people are asked several times (e.g., every day for 1 week), or if we are measuring results from members of the same family, these choices can be dependent, because they are related to each other.
Example: In the same example as above:

- The grade a student gets is a dependent choice, because it is influenced by the time he or she has spent studying.

When do dependent choices occur?

- When measurements are made on the same unit (e.g., the same people, cars, patients, etc.) at different times.
- When there are family, social, or geographical connections between the study subjects.
- In repeated experiments on the same unit.

More formally (precisely, mathematically): Two variables or choices are statistically independent if the occurrence of one does not affect the probability of the occurrence of the other.

Concrete examples:

Situation	**Choices are:**
Measure the blood pressure of 30 people before and after a diet	Dependent
Test 50 students at the beginning and end of the semester	Dependent
Ask 100 different people once each	Independent

Caution during analysis: When we have dependent variables, we cannot use tests for independent data (such as a t-test for independent groups). We must use methods that take into account dependence, such as:

- Paired t-test
- Repeated measures ANOVA

3.2.1. The t-test for dependence: This test is used to compare the means of two related groups. For example, it can be used to compare measurements before and after a medical treatment of a group of individuals.

The t-test formula:

$$t = \frac{\bar{x}}{s_x / \sqrt{n}}$$

where,

$\bar{x}$ – the average of the differences between pairs of data,

s_x – standard deviation of differences,

n – number of couples.

Application: Suppose you want to evaluate whether a medical treatment affects the health of patients. Use the t-test for dependent measures to compare the results before and after the treatment.

3.2.2. Dependent regression test: The regression test is used to analyze the relationship between a dependent variable and one or more independent variables. This test is useful for predicting the values of the dependent variables.
Linear regression formula:

$$Y = a + bX$$

where,

$$b = \frac{\sum XY - \sum X \sum Y}{n\sum X^2 - \left(\sum X\right)^2}, \quad a = \frac{\sum Y - b\sum X}{n}$$

Application: Suppose you want to analyze the impact of temperature and humidity on the production of an agricultural crop. Use linear regression to predict the crop's production depending on weather conditions.

4. Conclusions

Tests for independent and dependent variables are essential for analyzing data and drawing valid conclusions. They can be used to:

- Compare groups: As with t-tests for independent and dependent groups, to assess differences between groups.
- Examine relationships between variables: The Chi-square test can be used to analyze whether two categorical variables are independent.
- Predict different values: Linear regression can be used to predict a dependent value based on independent variables.

Statistical tests for independent and dependent variables are essential tools for any study or research that involves data analysis. Choosing the right test depends on the nature of the data and the purpose of the analysis. By using the right tests, researchers can draw accurate and valid conclusions that can be applied in many fields, such as the social sciences, health, economics, and beyond.

Acknowledgements

I would like to express my sincere gratitude to the members of the research team who contributed to the successful completion of this study.

References

[1] **Andy Field (2013).** *Discovering Statistics Using IBM SPSS Statistics* (4th Ed.). SAGE Publications.
[2] Field, A. (2013). *Discovering statistics using IBM SPSS Statistics* (4th ed.). London: SAGE Publications.
[3] Witte, R. S., & Witte, J. S. (2017). *Statistics* (11th ed.). Hoboken, NJ: Wiley.
[4] **Murat Bejdozi.** *Statistikë Aplikative në Shkencat Shoqërore*. Universiteti i Tiranës, Fakulteti i Shkencave Sociale.
[5] **Robert S. Witte & John S. Witte (2017).** *Statistics* (11th Ed.). Wiley.
[6] Hoxha, E. (2015). *Bazat e statistikës për psikologë dhe pedagogë*. Tiranë: Botime EMAL.
[7] Howell, D. C. (2012). *Statistical Methods for Psychology* (8th Ed.). Wadsworth.
[8] Lakens, D. (2013). "Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs." *Frontiers in Psychology*, 4, 863. https://doi.org/10.3389/fpsyg.2013.00863
[9] Çela, E., & Gëdeshi, I. (2009). *Metoda statistikore për shkencat sociale*. Tiranë: Qendra për Studime Ekonomike dhe Sociale.
[10] **Laerd Statistics – T-test tutorials** https://statistics.laer d.com/