

Мирошниченко Дарья Вячеславовна

**СТАТИСТИЧЕСКИЙ АНАЛИЗ РЕЗУЛЬТАТОВ  
КОМПЬЮТЕРНОГО ТЕСТИРОВАНИЯ: МЕТОДЫ И  
ИНСТРУМЕНТЫ**

*В современном мире возникает проблема, которая заключается в том, что необходимо обработать результаты тестирования, проведенного в онлайн-формате. Для этого необходимо выбрать подходящий метод и инструменты, чтобы корректно провести анализ данных. Целью работы является рассмотрение этих методов и инструментов, их сравнение, недостатки и достоинства. В работе рассмотрены дисперсионный, корреляционный и кластерный анализы.*

*Статистический анализ, аналитический обзор, анализ дисперсии, корреляционный анализ, кластерный анализ.*

Miroshnichenko Daria Vyacheslavovna

## **STATISTICAL ANALYSIS OF COMPUTER TESTING RESULTS: METHODS AND TOOLS**

*In the modern world, a problem arises, which is that it is necessary to process the results of testing conducted in an online format. To do this, you need to choose the appropriate method and tools to correctly analyze the data. The purpose of the work is to consider these methods and tools, compare them, drawbacks and advantages. The paper considers variance, correlation and cluster analyses.*

*Statistical analysis, analytical review, variance analysis, correlation analysis, cluster analysis.*

### **Введение**

В современном мире набирает популярность проверка обучающихся с помощью компьютерного тестирования для оценки имеющихся знаний и навыков, после прохождения курса, предмета или дисциплины.

На следующем этапе возникает проблема с интерпретацией этих результатов для дальнейшей работы и корректировки материалов курса. Для этого необходимо производить их с помощью статистического анализа, для того чтобы получить объективную и качественную картину ситуации, а также составления дальнейших корректирующих мероприятий [1].

Цель данной статьи рассмотреть основные методы и инструменты статистического анализа для внедрения его в анализ результатов прохождением обучающимися электронного курса.

### **Основная часть**

Рассмотрим основные методы и инструменты статистического анализа результатов компьютерного анализа, которые чаще всего используются для того, чтобы выбрать наиболее подходящий вариант, который будет самым релевантным.

Первый метод – это анализ дисперсии (ANOVA, от англ. Analysis of Variance), который необходимо использовать, когда оцениваются различия между средними значениями результатов тестирования в различных группах обучающихся. Данный способ применяется, когда ключевым критерием

являются существенные различия в результатах и факторы, которые могут повлиять на них.

Дисперсионный анализ построен на понятиях о двух ошибках. Первая – случайная ошибка, которая включает в себя разницу между выборочным значением каждого наблюдения, рассматриваемую как случайный фактор, вторая – системная, означающая разницу между разными факторами (зачастую они разные), которая формируется ошибки [2].

Сравнение данных ошибок основано на коэффициенте дисперсии, если разница между ними существенная. Данный метод предназначен для определения соответствия корреляции переменных по обнаруженным групповым различиям.

Преимущества использования анализа дисперсии:

- 1) производится статистический анализ на основе различий в средних значений выборок;
- 2) оценка различий между обучающимися, и определить ее значимость статистической значимости;
- 3) учет вариативности результатов в рамках выбранной группы;
- 4) имеется возможность оценки с учетом влияния внешних факторов.

Недостатки использования анализа дисперсии:

- 1) необходимы высокие знания в области математики для корректной работы с данными;
- 2) при выборе некорректной модели анализа, либо недостатке имеющихся данных, возникает риск, получения неправильных результатов.

Вторым методом является кластерный анализ, который заключается в выделении схожих элементов данных и дальнейшей идентификации на основе полученной информации, благодаря чему программу курса легко адаптировать и тестировать.

Первым этапом является сбор данных с результатами прохождения тестирования, далее происходит поиск схожих паттернов и группирует данные на основе сформированных кластеров. В результате получают кластеры, которые состоят из результатов с похожими значениями [3].

Метрика для оценки сходства / различия обычно основана на таких показателях расстояния, как евклидово или манхэттенское расстояние. Точки данных расположены ближе в одном кластере по сравнению с точками между кластерами.

Помечая связанные группы идентификаторами кластеров, кластеризация облегчает интерпретацию и оценку распределения данных в различных

областях. Она имеет обширные приложения для сегментации рынка, распознавания образов, анализа изображений и поиска информации.

Кластерный анализ в интеллектуальном анализе данных - это метод, используемый для группировки набора объектов таким образом, чтобы объекты в одной группе (или кластере) были более похожи друг на друга, чем объекты в других группах. Он обычно используется для обнаружения структур в немаркированных данных, классификации данных по различным группам и определения взаимосвязей между точками данных.

Преимуществами кластерного анализа является:

3) возможность выявить характерные паттерны при ответе на вопросы студентами;

4) данный способ позволит сформировать индивидуальные планы обучения и план корректирующих мероприятий;

5) позволяет выявить скрытые закономерности при ответе на вопросы тестирования.

Недостатками кластерного анализа является:

1) необходим большой объем данных;

2) имеется вероятность некорректного результата.

Последним рассмотрим корреляционный метод, который заключается в нахождении взаимосвязи с внешними факторами (уровень подготовки, погодные условия, фактор стресса и т.п.). Он необходим, когда в обязательном порядке необходимо найти и учесть данную взаимосвязь.

Корреляционный анализ - это статистический метод, который используется для измерения силы взаимосвязи между двумя или более переменными. Корреляционный анализ является важным инструментом для понимания взаимосвязи между переменными в наборе данных. Он может помочь исследователям выявить связи между переменными и определить, являются ли эти ассоциации статистически значимыми. Корреляционный анализ широко используется во многих областях, включая социальные науки, финансы и здравоохранение [4].

Важно подчеркнуть, что корреляция устанавливает статистическую взаимосвязь, но не доказывает причинно-следственную связь. Эти две переменные статистически связаны в некоторых географических регионах, но на самом деле причиной такой взаимосвязи может быть их взаимная связь с другой "неизвестной" переменной - например, с недавними достижениями в технологии. Это означает, что для определения фактической причинно-следственной связи может потребоваться дальнейший анализ.

Преимуществами кластерного анализа является:

- 1) имеется возможность выявить взаимосвязь результатов тестирования с внешними факторами;
- 2) по сравнению с другими методами легче интерпретировать полученные результаты.

Недостатками кластерного анализа является:

- 1) данный способ не позволяет исключить в дальнейшем проблемы, так как только позволяет их выявить;
- 2) возможность наличия ложных связей.

### **Выводы**

Подводя итоги, рассмотрели самые популярные методы статистического анализа результатов компьютерного тестирования и пришли к выводу, что самым актуальным и подходящим является анализ дисперсии. Данный способ позволит выявить закономерности результатов, оптимизировать процесс образования и упростить работу с анализом полученных данных для дальнейшего построения корректирующих мероприятий.

### **БИБЛИОГРАФИЧЕСКИЙ СПИСОК**

1. Статистический анализ результатов компьютерного тестирования URL: <https://cyberleninka.ru/article/n/statisticheskiy-analiz-rezultatov-kompyuternogo-testirovaniya> (дата обращения: 20.04.2024).
2. Дисперсионный анализ (ANOVA) URL: <https://habr.com/ru/companies/otus/articles/734258/> (дата обращения: 25.05.2024).
4. Корреляционный анализ URL: [https://e.vyatsu.ru/pluginfile.php/462616/mod\\_resource/content/3/Теоретический%20материал\\_корреляционный%20анализ.pdf](https://e.vyatsu.ru/pluginfile.php/462616/mod_resource/content/3/Теоретический%20материал_корреляционный%20анализ.pdf) (дата обращения: 25.05.2024).
5. Что такое кластерный анализ и как аналитики с его помощью группируют данные URL: <https://practicum.yandex.ru/blog/chto-takoe-klasterizaciya-i-klasternyi-analiz/> (дата обращения: 27.05.2024).

**Мирошниченко Дарья Вячеславовна**, студентка Донского государственного технического университета, Россия, город Ростов-на-Дону, площадь Гагарина 1, 344018, телефон: +7(988)552-45-61, email: [my\\_sun\\_md@mail.ru](mailto:my_sun_md@mail.ru).

**Miroshnichenko Daria Vyacheslavovna**, student of the Don State Technical University, Russia, Rostov-on-Don, Gagarin Square 1, 344018, phone: +7(988)552-45-61, email: [my\\_sun\\_md@mail.ru](mailto:my_sun_md@mail.ru).