

## **Исследование критериев множественных сравнений методом Монте-Карло и их применение к математико-статистическому анализу относительно малых выборок**

**Руководитель темы Масликов Александр Альбертович, кандидат физико-математических наук, доцент кафедры общеобразовательных дисциплин, заместитель директора по учебной и научной работе.**

Выполнение множественных сравнений совершенно необходимо при математико-статистической обработке проведенных исследований в самых разных областях науки и техники. При этом ставится задача установить статистически значимые различия между несколькими выборками, либо их отсутствие. Известный t-критерий Стьюдента (Госсета) разработанный для сравнения пар выборок нуждается в модификации при количестве выборок более двух и от этого резко теряет в мощности. Для этого критерия приходится вводить так называемую поправку Бонферони. При проверке статистических гипотез существует риск совершить ошибки двух видов - ошибки I-го рода, когда ошибочно отклоняется верная гипотеза и ошибки II -го рода, когда ошибочно принимается ложная гипотеза. Вероятность ошибки I-го рода называется уровнем значимости  $\alpha$ , стандартом для него является значение 0,05. Поправка Бонферони заключается в том, что эффективно уровень значимости должен быть разделен на количество попарных сравнений, что приводит к значительной потере чувствительности критерия даже к значительным эффектам.

Вероятность ошибки II -го рода обозначают  $\beta$ . Ошибки I и II рода являются конкурирующими, уменьшение вероятности одной влечёт увеличение вероятности другой. Также полезной характеристикой является мощность критерия  $(1 - \beta)$ . Чем больше мощность критерия, тем надёжнее он обнаруживает различия между выборками. Обычно считают приемлемой мощность порядка 0,8. Разумеется мощность зависит не только от объёма выборок (чем больше элементов в выборках, тем критерий оказывается мощнее), но и от величины эффекта, поэтому полезно иметь таблицы (или графики) для разных значений эффектов.

Мощнейшим методом пригодным для множественных сравнений является дисперсионный анализ (в англоязычной литературе ANOVA). Но, во-первых, этот метод определяет наличие различий по всей совокупности сравниваемых выборок, не указывая какие именно выборки различаются и во-вторых, аналитически доказана корректность процедуры дисперсионного анализа, только при условии нормальности распределений из которых производятся выборки

(и гомогенности их дисперсий). Если не установлен факт нормальности, то в качестве альтернативы используют непараметрический ранговый критерий Крускала—Уоллиса с апостериорным тестом Данна, которые обладают значительно меньшей мощностью, поскольку работают не с числовыми значениями измерений, а с номерами из упорядоченного числового ряда.

Апостериорные тесты применяются после основного дисперсионного анализ или его непараметрического аналога и призваны указать – какие конкретно выборки различаются. После дисперсионного анализа в качестве пост-теста обычно используют тест Тьюки или его модификацию Ньюмена-Кейлса. Эти тесты считаются достаточно мощными, однако, как и ANOVA, строго доказаны только для нормальных и гомогенных выборок.

Мы решили экспериментально исследовать вопрос – насколько оправдан отказ от использования традиционных параметрических критериев, в случае, если критерий Шапиро-Уилка и тест Левене указывают на возможную ненормальность и/или негомогенность выборок.

Исследования проводятся экспериментально с помощью компьютерных симуляций, путём извлечения методом Монте-Карло выборок из искажённых ненормальных распределений. Использовались выборки сравнительно небольшого объёма (до 10 элементов), что характерно для некоторых дорогостоящих исследований, например, в фармакологии, социологии, лингвистике. Мы выдвигаем гипотезу о том, что использование традиционных параметрических критериев не приводит к катастрофическим последствиям. Это утверждение подкрепляется вычислением экспериментальных уровней значимости и мощностей, а также сравнением графиков, построенных на основе вычислений.

В качестве инструментария мы использовали программы, написанные в пакете Wolfram Mathematica (WM). Идея заключается в том, чтобы, следуя методу Монте-Карло, генерировать случайные выборки из различных распределений и применять к ним исследуемые тесты. В WM есть возможность подключить генератор псевдослучайных чисел “Mersenne Twister”, который и был использован в работе. Вихрь Мерсенна – это генератор сдвигового регистра обобщенной обратной связи с огромным периодом ( $2^{19937} - 1$ ), гарантирующим высочайшую степень случайности. Пакет WM позволяет формировать компактные программы для

генерации большого числа случайных выборок из широкого набора библиотечных распределений.

Мы подвергли исследованию следующие искажённые распределения: t-распределение Стьюдента,  $\chi^2$ -распределение, распределение Джонсона, гамма-распределение, распределение Вейбулла. Рассматривались распределения как с сонаправленными, так и с разнонаправленными асимметриями. Приведём графики для мощностей некоторых распределений (см. рис 1.) Для получения каждой точки использовалось до  $10^5$  наборов симуляций четвёрок выборок. По горизонтали отложен шаг эффекта в единицах СКО, по вертикали – вероятность обнаружения эффекта. Как видно, все графики расположены весьма близко к графику для нормального распределения, что подтверждает возможность использования традиционных классических параметрических критериев и для искажённых выборок.

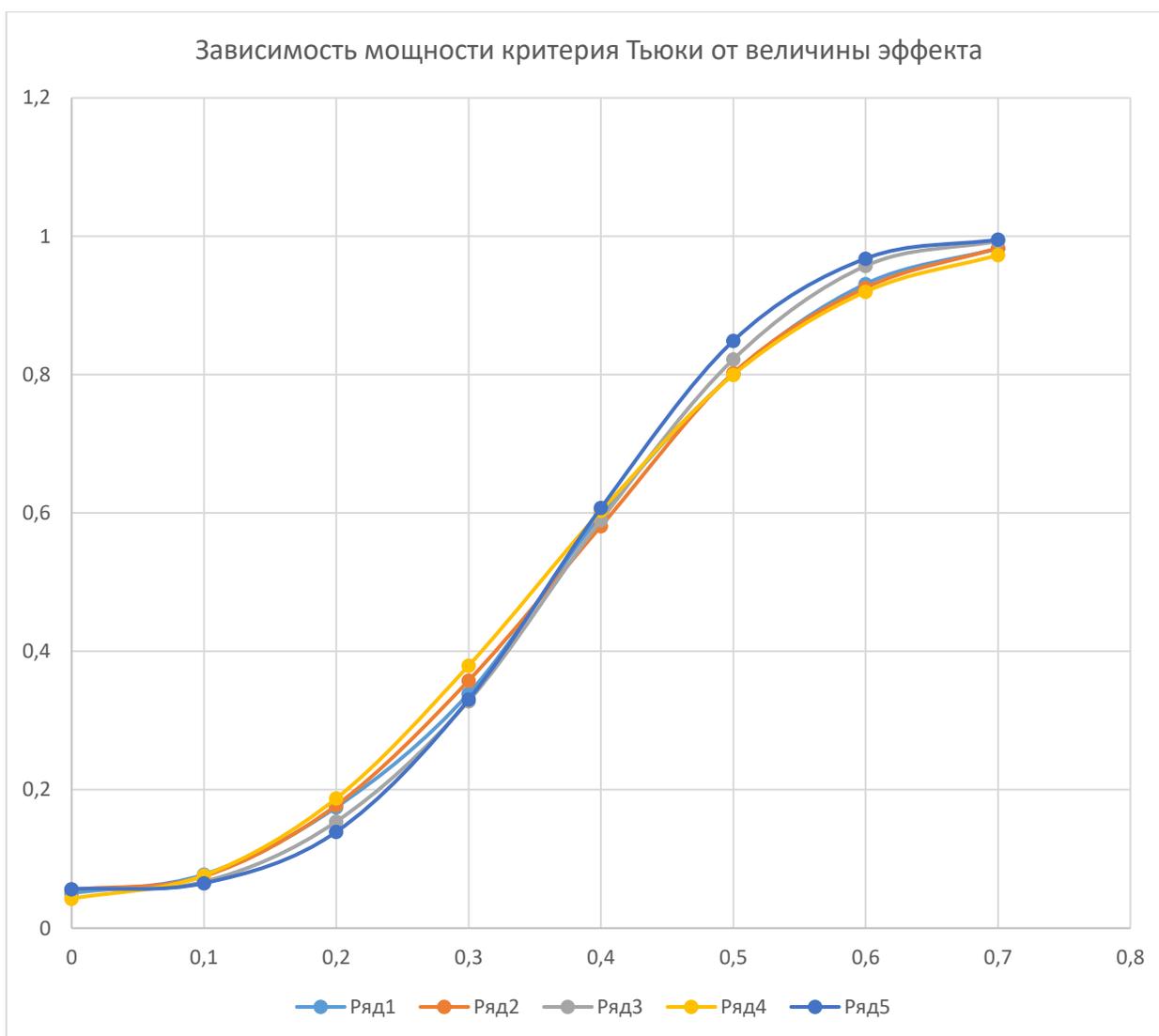


Рис. 1. – Мощность критерия Тьюки (Tukey HSD) для 10-ти элементных выборок.

Соответствие цветов графиков распределениям следующее: голубой – нормальное, оранжевый - Вейбулла, серый – Вейбулла с инверсией, жёлтый – гамма, синий – гамма с инверсией.

Кроме теоретических исследований с помощью метода Монте-Карло, также проводились математико-статистические расчёты для Центра Токсикологии (г.Большевик) и для лингвистов из ФГАОУ ВО «Российский университет дружбы народов имени Патриса Лумумбы», результаты были опубликованы в журналах «Мир Науки», «Известия ВГУ» (из списков ВАК, РИНЦ) и в сборниках университета «Дубна». 3 публикации сделаны в соавторстве с выпускниками филиала «Протвино».

За период 2020—2024 гг по описанным выше материалам защитили бакалаврские квалификационные работы 18 выпускников филиала «Протвино». Доклад студентки 3-го курса филиала по теме «Изучение поведения мощности стандартных критериев множественных сравнений на примере некоторых негауссовых выборок» представлен на конференцию "Молодёжь и инноватика" 6 февраля, 2025 (г. Серпухов).

1. Максимова О.Б., Масликов А.А., «Гендерные стереотипы как разновидность социальных стереотипов: функции в социальной коммуникации», Мир науки. Социология, филология, культурология. – 2021. – №3, Том 12. – URL: <https://sfk-mn.ru/PDF/01SCSK321.pdf> ЭЛ № ФС77-67275 от 21.09.2016, Роскомнадзор (ВАК, РИНЦ).
2. Алдобаев В.Н., Артемьева А.Д., Масликов А.А., «Исследование поведения классических критериев множественных сравнений, на ненормальных неоднородных распределениях, методом Монте-Карло», Вестник ВГУ. Серия: Системный анализ и информационные технологии. – 2021. – №3. – С. 72-80 (ВАК, ядро РИНЦ).
3. Алдобаев В.М., Артемьева А.Д., Масликов А.А., Саморокова В.И., «Изучение поведения критериев множественных сравнений на выборках из ненормальных и негомогенных распределений методом Монте-Карло (тезисы доклада)», Задачи и методы нейтронных исследований конденсированных сред: Вторая Всероссийская научно-практическая конференция (с международным участием) (Дубна, 21-23 ноября 2022 г.): сборник тезисов докладов / под. общ. ред. В.И. Боднарчука. – Дубна: Гос. ун-т «Дубна», 2022. – С. 23-24.
4. Максимова О.Б., Масликов А.А., «Лексическая интерференция и позитивный перенос при изучении английского языка в сфере профессиональной коммуникации как второго иностранного», Мир науки. Педагогика и психология. – 2022. – №1. – URL: <https://mir-nauki.com/PDF/41PDMN122.pdf> ЭЛ № ФС77-74663 от 24.12.2018, Роскомнадзор (ВАК, РИНЦ).
5. Максимова О.Б., Масликов А.А., «Межъязыковая интерференция у студентов мультилингов, изучающих английский язык как второй иностранный язык», Мир науки. Педагогика и психология. – 2022. – №5. – URL: <https://mir-nauki.com/PDF/32PDMN522.pdf> ЭЛ № ФС77-74663 от 24.12.2018, Роскомнадзор (ВАК, РИНЦ).
6. Максимова О.Б., Масликов А.А., «Межъязыковая интерференция при изучении английского языка в сфере профессиональной коммуникации как второго иностранного: пилотные исследования и математико-статистическая обработка результатов», Мир науки. Педагогика и психология. — 2023. — Т 11. — №2. — URL: <https://mir-nauki.com/PDF/46PDMN223.pdf> ЭЛ № ФС77-74663 от 24.12.2018, Роскомнадзор (ВАК, РИНЦ).
7. Алдобаев В.Н., Масликов А.А., Скворцова М.С., «Исследование мощности критериев множественных сравнений, применённых к выборкам с негауссовым распределением элементов», Сборник «Государственный университет «Дубна». 30 лет в науке», раздел «Естественные и инженерные науки», - 2024 - С. 233-240.