

#	Имя	Тип данных	Длина/Знач...	Беззнаковое	Разрешить NULL	Zerofill	По умолчанию
1	schc_id	BIGINT	20	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	AUTO_INCREMENT
2	schc_date	DATE		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Нет значения по умолчанию
3	schc_group	VARCHAR	8	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Нет значения по умолчанию
4	schc_num	TINYINT	3	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Нет значения по умолчанию
5	schc_lesson	VARCHAR	256	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Нет значения по умолчанию
6	schc_url	VARCHAR	256	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	"
7	schc_adddatetime	DATETIME		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	current_timestamp()
8	schc_adduserid	BIGINT	20	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	'0'
9	schc_is_active	TINYINT	1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	'1'

Рис. 4. Структура таблицы Schedule_changes

Программа во время своей работы в цикле проверяет пользователей, ожидающих расписания, когда время отправки достигнуто происходит формирование расписания для каждого пользователя и его отправка. После этого пользователю назначается время следующий отправки сообщения, в зависимости от его настроек.

Бот был протестирован на Windows 10, Ubuntu 20.04, Ubuntu 20.10 были применены некоторые правки для корректной работы на разных ОС. Все выявленные ошибки были исправлены.

Результатом проделанной работы является:

1. В настоящее время бот создан, эксплуатируется.
2. Спроектирована база данных, структура была записана в приложение. Поддержка СУБД с синтаксисом MySQL.
3. Создание кроссплатформенного приложения портативного чат бота на платформе .NET с возможностью запуска на Windows, Linux и macOS.

Разработанный чат бот позволит удобнее ориентироваться в расписании и изменениях, не тратить значительно меньше времени на проверку расписания.

Библиографический список

1. Microsoft Docs. Руководства по .NET | Microsoft Docs // [Электронный ресурс]. URL: <https://docs.microsoft.com/ru-ru/dotnet/core/tutorials/> (дата обращения: 16.03.2021).
2. Документация для разработчиков VK API // [Электронный ресурс]. URL: <https://vk.com/dev/manuals> (дата обращения: 16.03.2021).
3. VKNET ВКонтакте API для .NET (C#) // [Электронный ресурс]. URL: <https://vknet.github.io/vk> (дата обращения: 16.03.2021).
4. Wit — HTTP API// [Электронный ресурс]. URL: <https://wit.ai/docs/http/20170307> (дата обращения: 16.03.2021).
5. ClosedXML is a .NET library. GitHub // [Электронный ресурс]. URL: <https://github.com/ClosedXML/ClosedXML> (дата обращения: 16.03.2021).

УДК 519.233.8

Калиничева Д.И., Радзиевская В.Г.

ИССЛЕДОВАНИЕ НЕКОТОРЫХ ПОСТ-ТЕСТОВ МНОЖЕСТВЕННЫХ СРАВНЕНИЙ И ИХ РАБОТОСПОСОБНОСТИ ПРИ НЕГАУССОВОСТИ ВЫБОРОК МЕТОДОМ МОНТЕ-КАРЛО

INVESTIGATIONS OF SOME POST-TESTS OF MULTIPLE COMPARISONS AND THEIR PERFORMANCE WITH NON-GAUSSIAN SAMPLES BY THE METHOD OF MONTE-CARLO

*Филиал «Протвино» государственного университета «Дубна»
Секция «Естественные и инженерные науки»*

Авторы: Калиничева Диана Ильинична, студентка 4 курса направления «Физика» филиала «Протвино» государственного университета «Дубна»; Радзиевская Вячеслава

Геннадиевна, студентка 4 курса направления «Физика» филиала «Протвино» государственного университета «Дубна».

Научный руководитель: Масликов Александр Альбертович, кандидат физико-математических наук, доцент кафедры технической физики филиала «Протвино» государственного университета «Дубна».

Author: Kalinicheva Diana, 4d year student of the direction "Physics" of the branch "Protvino" state University "Dubna". Radzievskaya Vyacheslava, 4d year student of the direction "Physics" of the branch "Protvino" state University "Dubna".

Scientific adviser: Maslikov Alexander Albertovich, candidate of physical and mathematical Sciences, associate professor of the department of technical physics of the branch "Protvino" state University "Dubna".

Аннотация

Методом Монте-Карло симулируется использование тестов, выполняемых при множественных сравнениях на выборках малого объёма из искаженных (по сравнению с нормальным) распределений. Вычисляются эмпирические эффективные вероятности ошибок 1-го рода для дисперсионного анализа, пост-тестов множественных сравнений Тьюки и процедуры Бенжамини-Хохберга.

Abstract

The Monte Carlo method simulates the use of tests performed in multiple comparisons on small-volume samples from distorted (compared to normal) distributions. Empirical effective probabilities of errors of the 1st kind are calculated for the analysis of variance, Tukey's post-tests of multiple comparisons and the Benjamini-Hochberg procedure.

Ключевые слова: Математическая статистика, множественные сравнения, дисперсионный анализ, апостериорные тесты.

Keywords: Mathematical statistics, multiple comparisons, ANOVA, post-hoc tests.

Условием надёжной работы большинства классических статистических тестов является нормальное (Гауссово) распределение исследуемых генеральных совокупностей и их однородность (гомогенность) по дисперсиям [1]. В то же время во многих отраслях знания (например, в медицине, в фармакологии, в токсикологии) при математико-статистических исследованиях зачастую приходится иметь дело с выборками небольшого объема (5—10 элементов), для которых проблематично гарантированно установить нормальность и/или однородность дисперсий. Поэтому представляется интересным изучить поведение традиционных статистических критериев на ненормальных распределениях и выборках малого объема. Наши исследования и расчёты лежат в русле направления, развиваемого школой профессора Лемешко Б.Ю. [2–4]. Традиционные аналитические методы исследования эффективны только при весьма жёстких ограничениях на соответствующие генеральные совокупности. В то же время для решения этого вопроса с успехом можно применять компьютерные симуляции и метод Монте-Карло.

В качестве инструмента мы использовали программы, написанные в пакете Wolfram Mathematica (WM). Идея 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 мы и использовали. 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 огромного числа случайных выборок из большого набора библиотечных распределений. Мы вполне можем использовать количество генераций порядка 10^5 — 10^6 на каждый набор параметров.

В данном исследовании мы остановили своё внимание на пост-тестах множественных сравнений, т.е. тестах, которые выполняются после того как, например, дисперсионный анализ (ANOVA) определил наличие каких-то различий [5, 6]. Назначение этой группы тестов – определить, какие именно группы статистически значимо отличаются между собой. Разработка этой группы тестов обусловлена катастрофическим падением мощности парного t-сравнения при увеличении множественности сравнений (так называемый эффект поправки Бонферони).

Мы приведём результаты по исследованию апостериорного теста множественных сравнений Тьюки (Tukey) [7] (они остаются справедливыми и для его модификации Ньюмена-Кейлса (Newman-Keuls)), а также по пост-тесту процедуре Бенжамини-Хохберга (Benjamini-Hochberg) [8]. В основе этих тестов лежит t-критерий Стьюдента сравнения средних, опирающийся на известную t-статистику [5, 6].

t-распределение Стьюдента - это непрерывное одномерное распределение с одним параметром - количеством степеней свободы. Пусть Y_0, Y_1, \dots, Y_n — независимые стандартные нормальные случайные величины (матожидание=0, дисперсия=1). Тогда распределение случайной величины t , где

$$t = \frac{Y_0}{\sqrt{\frac{1}{n} \sum_{i=1}^n Y_i^2}}$$
 называется распределением Стьюдента с n степенями свободы ($df=n$).

Плотность вероятности этого распределения $f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi} \Gamma\left(\frac{n}{2}\right) \left(1 + \frac{x^2}{n}\right)^{\frac{n+1}{2}}}$, где $\Gamma(\dots)$ - гамма

функция Эйлера. Отметим, что эта формула допускает и дробное число степеней свободы.

t-критерий может применяться в двух модификациях: 1) в варианте статистически одинаковых дисперсий групп; 2) в варианте статистически различных дисперсий. В первом варианте число степеней свободы в t-статистике $df = n + m - 2$, где n и m – объёмы сравниваемых выборок. Во втором варианте число степеней свободы становится дробным числом и вычисляется по весьма сложной формуле. Для процедуры Бенжамини-Хохберга мы тестируем оба варианта t-критерия.

Отклонения распределений генеральных совокупностей от нормального Гауссового распределения можно описывать параметрами: эксцесс (коэффициент островершинности E) и асимметрия (A). Для нормального распределения $E = 3$, а $A=0$. На первом шаге в качестве модели искажённого нормального распределения мы взяли распределение χ^2 с числом степеней свободы $df = 4$. Это распределение является частным случаем Гамма распределения и описывается функцией плотности вероятности:

$$f_{\chi^2(n)} = \frac{\left(\frac{1}{2}\right)^{n/2}}{\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}},$$

Параметры данного распределения: матожидание $\mu = n$, дисперсия $\sigma^2 = 2n$, коэффициент асимметрии $\sqrt{8/n}$, эксцесс $3+12/n$, где n – число степеней свободы. В нашем случае χ^2 распределения с $df = 4$ имеем: $\mu = 4$, $\sigma^2 = 8$, коэффициент асимметрии $A = \sqrt{2}$, эксцесс $E = 6$.

Исследования методом Монте-Карло дисперсионного анализа и пост-теста Тьюки на нормальных выборках подтверждают правильность общего подхода. Т.е. дают эмпирическую вероятность ошибки 1-го рода 0.05 при заложенном уровне значимости 5%. При этом наблюдаются ситуации, когда ANOVA указывает на наличие различий, а тест Тьюки не в состоянии их определить. Это следствие меньшей мощности пост-тестов по сравнению с дисперсионным анализом.

Далее, мы исказили исследуемое распределение, взяв χ^2 с числом степеней свободы $df = 4$.

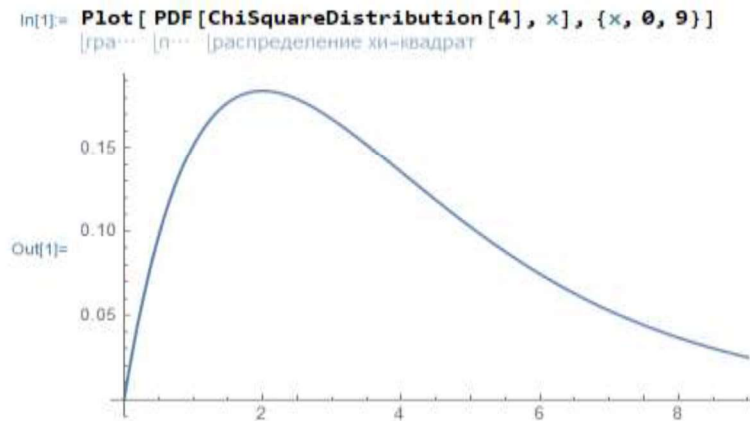


Рис. 1. Распределение хи-квадрат

Обсчитали это распределение, меняя объем выборки, методом Монте-Карло с числом испытаний 250000. Число сравниваемых выборок взяли 4, и, кроме того, исследовали случай, когда асимметрия 2-х из 4-х выборок направлена в отрицательную сторону. Далее применялся дисперсионный анализ (ANOVA) и пост тест Тьюки (эти же результаты справедливы и для пост-теста Ньюмена-Кейлса (НК)). Результаты этого модельного исследования представлены в Таблице 1. Мы видим, что искажение нормальности распределения не сказывается катастрофическим образом на достоверности ANOVA и пост-тестов. Более того, при одинаковой направленности асимметрии наблюдается даже уменьшение вероятности ошибки 1-го рода. Только в случае разнонаправленных асимметрий эта вероятность незначительно возрастает.

Таблица 1. Вероятности ошибок 1-го рода для ANOVA и пост-теста Тьюки (для χ^2).
I – сонаправленные асимметрии, II – разнонаправленные асимметрии.

	Объем	5	6	7	8	9	10
I	ANOVA	0,0440	0.0451	0.0460	0,0458	0,0444	0,0484
	Tukey (НК)	0,0436	0.0432	0.0447	0,0443	0,0436	0,0465
II	ANOVA	0.0582	0.0576	0.0575	0.0560	0.0560	0.0548
	Tukey (НК)	0.0577	0.0569	0.0566	0.0555	0.0557	0.0547

Теперь для примера аналогично изучим ограниченное слева распределение Джонсона: $Y = \sigma \exp((X - \gamma)/\delta) + \mu$. Где X – случайная величина, распределенная по стандартному нормальному закону. Интегральная функция распределения Джонсона выражается формулой $J(Y) = \Phi[\sigma \ln((Y - \gamma)/\delta) + \mu]$, где Φ – интегральная функция стандартного нормального распределения.

Plot[PDF[JohnsonDistribution["SL", 0, 0.8, 0, 1], x], {x, 0, 4}]
 график распределения Джонсона

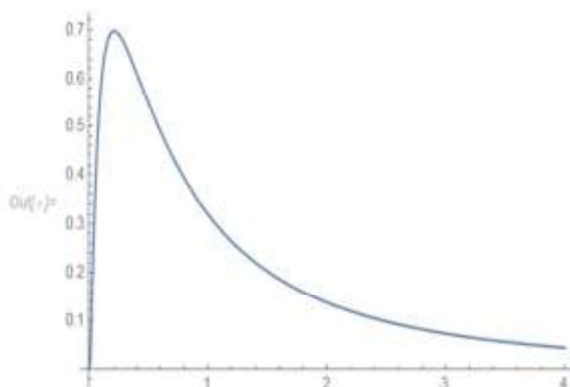


Рис. 2. Распределение Джонсона

При взятых значениях параметров (см. Таблицу 2.) асимметрия $A=13.15$, эксцесс $E=800$. Опять сравнивалось 4 выборки. Результаты представлены в Таблице 2. Для получения каждой вероятности использовалось 10^5 симуляций. Снова наблюдается уменьшение вероятности ошибки 1-го рода при сонаправленных асимметриях и незначительное её увеличение при разнонаправленных.

Таблица 2. Вероятности ошибок 1-го рода для ANOVA и пост-теста Тьюки (для ограниченного слева распределения Джонсона с параметрами: $\gamma=0$, $\delta=0.8$, $\mu=0$, $\sigma=1$).

I – сонаправленные асимметрии, II – разнонаправленные асимметрии.

		Объем	5	6	7	8	9	10
I	ANOVA		0.0285	0.0281	0.0291	0.0306	0.0300	0.0307
	Tukey (NK)		0.0246	0.0242	0.0248	0.0259	0.0262	0.0266
II	ANOVA		0.118	0.132	0.128	0.121	0.117	0.114
	Tukey (NK)		0.115	0.127	0.124	0.117	0.114	0.110

Перейдём к исследованию процедуры Бенжамини-Хохберга. Она основана на контроле частоты ложных обнаружений, который производится с помощью арифметической прогрессии с шагом равным уровню значимости делённому на число пар выборок. Максимально различной паре выборок ставится в соответствие исходный уровень значимости, а минимально различной – минимальный уровень значимости. Приведём результаты тестирования процедуры Бенжамини-Хохберга на нормальных выборках различного объёма. Процедура организована таким образом, что при обнаружении негомогенности дисперсий используется t-сравнение по 2-му варианту, а при однородных дисперсиях по 1-му. Результаты исследований пост-теста процедуры Бенжамини-Хохберга для различного количества и объёма выборок приводим в Таблице 3. Количество симуляций здесь составляет 10^5 . Наблюдаем, что эмпирическая вероятность ошибки 1-го рода укладывается в декларируемый уровень значимости и имеется тенденция к понижению этой вероятности с ростом количества сравниваемых выборок.

Таблица 3. Вероятности ошибок 1-го рода для процедуры Бенжамини-Хохберга на выборках из стандартного нормального распределения.

	Кол-во выборок	4	5	6	7
Кол-во элементов					
6		0.0447	0.0423	0.0419	0.0408
7		0.0448	0.0442	0.0424	0.0416
8		0.0461	0.0436	0.0423	0.0421

Далее мы искажем нормальность распределения в одном и в обоих направлениях и получим для искаженных выборок результаты аналогичные Таблицам 1 и 2. В качестве искаженного будем использовать снова распределение χ^2 с числом степеней свободы $df = 4$ (как в Таблице 1.) Результаты представлены в Таблице 4. Мы указали эмпирические вероятности ошибок 1-го рода для 2-х способов Т-сравнений встроенных в процедуру Б-Х. При I-ом способе делается поправка на возможную негомогенность дисперсий выборок, при II-ом способе дисперсии считаются однородными (что соответствует истинной гипотезе H_0). Естественно, II-ой способ чаще обнаруживает различия. Опять наблюдается отклонение вероятности ошибки 1-го рода в меньшую сторону при сонаправленности асимметрий и в сторону увеличения при разнонаправленности. Причём эти отклонения больше, чем соответствующие отклонения для пост-теста Тьюки (на том же самом распределении).

Таблица 4. Вероятности ошибок 1-го рода для процедуры Бенжамини-Хохберга на выборках из распределения χ^2 с числом степеней свободы $df = 4$.

Выборки		Элементы		6		7		8	
		I	II	I	II	I	II		
4	сонаправленность	0.0329	0.0371	0.0313	0.0357	0.0330	0.0380		
	разнонапр-ть	0.0878	0.0909	0.0854	0.0881	0.0816	0.0839		
5	сонаправленность	0.0321	0.0366	0.0312	0.0362	0.0306	0.0355		
6	сонаправленность	0.0303	0.0344	0.0299	0.0347	0.0300	0.0340		
	разнонапр-ть	0.0923	0.0957	0.1015	0.1044	0.0988	0.1017		
7	сонаправленность	0.0313	0.0356	0.0300	0.0349	0.0280	0.0331		

Представляется полезным итерационным методом найти уровни значимости, соответствующие статистически наблюдаемым ошибкам 1-го рода для используемых «ненормальных» распределений. Имеет смысл провести аналогичные исследования для распределений с другими значениями параметров и другими малыми объемами выборок. Также мы планируем аналогичным методом исследовать влияние «ненормальности» на мощность критериев.

Библиографический список

1. Гмурман В.Е. Теория вероятностей и математическая статистика – М., Высшая школа, 2003.
2. Лемешко Б. Ю., Лемешко С. Б., Постовалов С. Н., Чимитова Е. В. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход. Изд-во НГТУ; 2011. 888 с.;
3. Лемешко Б. Ю., Лемешко С. Б. Об устойчивости и мощности критериев проверки однородности средних. Измерительная техника. 2008; (9): 23–28.
4. Корнеев А. А., Кричевец А. Н. Условия применимости критериев Стьюдента и Манна–Уитни. Психологический журнал. 2011; 32 (1): 97–110.

5. С.Гланц. Медико-биологическая статистика. Пер. с англ. – М., Практика, 1998 – 459 с.
6. Лапач С.Н., Чубенко А.В., Бабич П.Н., Статистические методы в медико-биологических исследованиях с использованием Excel. – 2-е изд. перераб. и доп. – К.: МОРИОН, 2001. – 408 с.
7. Тьюки, Джон (1949). "Сравнение индивидуальных средних в Дисперсионном анализе". Биометрия. 5 (2): 99–114.
8. Y. Benjamini, Y. Hochberg. «Controlling the Discovery Rate : a Practical and Powerful Approach to Multiple Testing». Journal of the Royal Statistical Society. Series B (Methodological), Vol.57, No. 1 (1995), pp 289–300.

УДК 20.15.05

Карпов Д.Ю.

**РАЗРАБОТКА КЛИЕНТ-СЕРВЕРНОЙ ПРОГРАММЫ «ПРЕПОДАВАТЕЛЬ -
СТУДЕНТЫ»**
DEVELOPMENT OF THE CLIENT-SERVER PROGRAM «TEACHER - STUDENTS»

*Филиал «Протвино» государственного университета «Дубна»
Секция «Информационные технологии»*

Автор: Карпов Даниил Юрьевич, студент 2-го курса направления «Информатика и вычислительная техника» филиала «Протвино» государственного университета «Дубна».

Научный руководитель: Кульман Татьяна Николаевна, кандидат технических наук, доцент кафедры информационных технологий филиала «Протвино» государственного университета «Дубна».

Author: Karpov Daniil Yurievich, 2nd year student of the direction "Informatics and computer engineering" of the branch "Protvino" state University "Dubna".

Scientific adviser: Kulman Tatiana Nikolaevna, candidate of technical sciences, associate professor of the department information technology of the branch "Protvino" state University "Dubna".

Аннотация

В данной статье рассматривается разработка программы удаленного доступа для образовательных учреждений с дополнительными возможностями для проведения занятий в компьютерных классах и тестирования обучающихся.

Abstract

This article discusses the development of a remote access program for educational institutions with additional capabilities for conducting classes in computer labs and testing students.

Ключевые слова: Удаленный доступ, TCP, средство удаленного администрирования, тестирование, обучение, клиент-сервер.

Keywords: Remote access, TCP, remote administration tool, testing, training, client-server.

Целью работы является создание программы удаленного доступа с дополнительными возможностями для работы в компьютерных классах и тестирования обучающихся.

Актуальность

Данная программа может облегчить жизнь педагогам в образовательных учреждениях. Так же эту программу можно будет использовать для администрирования компьютерных классов, что поможет в работе системным администраторам.

Постановка задачи

Создание клиент-серверной программы для удаленного управления компьютерами с функционалом:

- отправка снимков экрана как от преподавателя студенту, так и на оборот,,
- тестирование обучающихся,