На правах рукописи

Рогожников Андрей Павлович

ИССЛЕДОВАНИЕ СВОЙСТВ НЕКОТОРЫХ КРИТЕРИЕВ ПРОВЕРКИ СТАТИСТИЧЕСКИХ ГИПОТЕЗ И ОБЕСПЕЧЕНИЕ КОРРЕКТНОСТИ ИХ ПРИМЕНЕНИЯ МЕТОДАМИ КОМПЬЮТЕРНОГО МОДЕЛИРОВАНИЯ

Специальность 05.13.17 — Теоретические основы информатики

Автореферат
диссертации на соискание ученой степени
кандидата технических наук

Новосибирск — 2012

Работа выполнена в Федеральном государственном бюджетном
образовательном учреждении высшего профессионального образования
«Новосибирский государственный технический университет»

Научный руководитель: доктор технических наук, профессор
Лемешко Борис Юрьевич

Официальные оппоненты: Хабаров Валерий Иванович
доктор технических наук, профессор,
Федеральное государственное бюджетное
образовательное учреждение высшего
профессионального образования
«Сибирский государственный университет
путей сообщения», заведующий кафедрой
«Информационные технологии на
транспорте»

Фаддеенков Андрей Владимирович
кандидат технических наук, доцент,
Федеральное государственное бюджетное
образовательное учреждение высшего
профессионального образования
«Новосибирский государственный
технический университет», доцент кафедры
теории рынка

Ведущая организация: Федеральное государственное бюджетное учреждение науки Институт вычислительной математики и математической геофизики Сибирского отделения Российской академии наук, г. Новосибирск

Защита состоится «13» декабря 2012 г. в 14-00 часов на заседании диссертационного совета Д 212.173.06 при Федеральном государственном бюджетном образовательном учреждении высшего профессионального образования «Новосибирский государственный технический университет» по адресу 630092, г. Новосибирск, пр-т К. Маркса, 20.

С диссертацией можно ознакомиться в библиотеке Новосибирского государственного технического университета.

Автореферат разослан «   » ноября 2012 г.

Ученый секретарь
диссертационного совета Чуб Чубич Владимир Михайлович

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Корректность применения множества статистических моделей и методов зависит от справедливости определенных предположений. При этом предположение о принадлежности наблюдений или ошибок измерений нормальному закону является наиболее частым.

В эконометрических моделях принятие нормальной модели без формальной проверки может влиять на точность оценок и выводов, формируемых в результате анализа. Среди разнообразных примеров обращения к модели нормального закона можно увидеть предположение об условной нормальности данных в модели отбора, применяемой к спросу на жилье, или предположение о нормальности при анализе данных фондового рынка, для которых типичным оказывается наличие тяжелых хвостов.

Проверка гипотез о принадлежности ошибок измерений нормальному закону типична в задачах статистического управления качеством.

Как правило, гипотеза о нормальности проверяется относительно вектора ошибок моделей регрессионного анализа, применяемых к временным рядам, к пробит-регрессии и к другим типам временных рядов.

Предположение о нормальности распространено в задачах анализа результатов медицинских экспериментов. Допустимость этого предположения должна проверяться, например, в случае изменчивости данных об экспрессии генов или в ходе клинических испытаний при проверке эффективности новых методов лечения.

В силу отсутствия исчерпывающей информации о предпочтительности тех или иных критериев проверки гипотез о нормальности к их исследованию в последнее время обращаются различные авторы (Dong and Giles, 2007, Doornik and Hansen, 2008, Voinov and Voinov, 2010, Scott and Stewart, 2011).

Показательный закон является второй наиболее популярной моделью, используемой в задачах статистическом анализа, особенно в задачах анализа времени жизни и в теории надежности. Гипотеза о том, что базовый закон модели надежности является показательным, эквивалентна гипотезе о постоянной интенсивности отказов. В пуассоновских потоках интервалы времени между наступлениями событий подчиняются показательному распределению. Среди процессов, генерирующих пуассоновские потоки, можно указать испускание радиоактивных частиц, землетрясения, отказы оборудования и т.п. Относительно множества критериев показательности также нет однозначной информации о предпочтительности конкретных критериев.

Цель и задачи исследований. Основная цель диссертационной работы заключалась в исследовании свойств и сравнительном анализе множества статистических критериев, предназначенных для проверки гипотез о принадлежности данных нормальному или показательному закону, дающих основание для выбора наиболее предпочтительного критерия в конкретной ситуации, в разработке программного обеспечения, позволяющего исследовать и корректно применять соответствующие статистические критерии.

В соответствии с поставленной целью решались следующие задачи:

- создание программного обеспечения, позволяющего осуществлять проверку гипотез по критериям, рассматриваемым в работе, моделировать распределения статистик критериев, вычислять оценки мощности критериев по отношению к различным конкурирующим гипотезам;
- исследование распределений статистик критериев нормальности Фросини, Хегази-Грина, Гири, Дэвида–Хартли–Пирсона, Шпигельхальтера;
- вычисление методами статистического моделирования оценок мощности критериев нормальности Фросини, Хегази–Грина, Гири, Дэвида–Хартли–Пирсона, Шпигельхальтера, Шапиро–Уилка, Ройстона, Эппса–Палли, Д'Агостино, критериев согласия (Колмогорова, Андерсона–Дарлинга, Крамера–Мизеса–Смирнова, χ^2 Пирсона и χ^2 Никулина — при проверке сложной гипотезы о нормальности) по отношению к близким конкурирующим гипотезам;
- сравнительный анализ мощности перечисленных критериев нормальности;
- исследование распределений статистик критериев показательности Гнеденко, Харриса, Холландера–Прошана, Гини, Колмогорова, Крамера–Мизеса–Смирнова, Андерсона–Дарлинга, Дешпанде, Кокса–Оукса, Большева, Клара, Барингхауса–Хензе, Хензе, Хензе–Мейнтаниса и Эппса–Палли при справедливости проверяемой гипотезы;
- вычисление оценок мощности критериев показательности по отношению к конкурирующим законам с различной формой функции интенсивности отказов, сравнительный анализ мощности критериев.

Методы исследования. Для решения поставленных задач использовался аппарат теории вероятностей, математической статистики, статистического моделирования, математического программирования.

Научная новизна диссертационной работы заключается:

- в результатах сравнительного анализа мощности критериев нормальности;
- в построенных таблицах процентных точек, расширяющих возможности применения критериев Фросини, Хегази–Грина, Гири, Дэвида–Хартли–Пирсона, Шпигельхальтера;
- в результатах сравнительного анализа мощности критериев показательности;
- в выявленных отклонениях распределений статистик критериев показательности от теоретических;
- в построенных моделях распределений для статистик критериев типа Колмогорова, Андерсона–Дарлинга, Крамера–Мизеса–Смирнова, основанных на эмпирической функции распределения, и критериев типа Колмогорова и Крамера–Мизеса–Смирнова, основанных на функции среднего остаточного времени безотказной работы;
- в рекомендациях по применению критериев показательности при ограниченных объемах выборок.

Основные положения, выносимые на защиту. На защиту выносятся:

1. Результаты исследования распределений статистик, мощности и сравнительного анализа критериев нормальности.
2. Рекомендации по применению критериев нормальности.
3. Результаты исследования распределений статистик, мощности и сравнительного анализа критериев показательности.
4. Рекомендации по применению критериев показательности.
5. Подход, обеспечивающий корректность применения статистических критериев в условиях нарушения стандартных предположений за счет построения распределений статистик, соответствующих справедливости проверяемой гипотезы, в результате компьютерного моделирования в ходе проводимого статистического анализа.

Обоснованность и достоверность научных положений, выводов и рекомендаций обеспечивается:

- корректным применением математического аппарата и методов статистического моделирования для исследования свойств и распределений статистик критериев;
- совпадением результатов статистического моделирования с известными теоретическими результатами.

Личный творческий вклад автора заключается в проведении исследований, обосновывающих основные положения, выносимые на защиту: в разработке программного обеспечения, в проведении статистического моделирования распределений статистик, в вычислении мощности критериев относительно конкретных альтернатив, в построении моделей распределений статистик и вычислении таблиц процентных точек.

Практическая ценность и реализация результатов. Полученные модели распределений статистик критериев нормальности и показательности позволяют корректно применять их при ограниченных объемах выборок. Результаты сравнительного анализа мощности позволяют обоснованно выбирать критерии для проверки гипотез о принадлежности выборок нормальному или показательному закону, в том числе, при наличии конкурирующих гипотез определенного вида. Результаты исследований и средства моделирования включены в программную систему «Интервальная статистика» ISW и используются в научных исследованиях и учебном процессе.

Апробация работы. Основные результаты работы были представлены на Всероссийской научной конференции молодых ученых «Наука. Технологии. Инновации» (Новосибирск, 2007, 2008), Международной научно-технической конференции «Актуальные проблемы электроники и приборостроения» (Новосибирск, 2008, 2012), VI International Conference Mathematical Methods in Reliability (Moscow, 2009), The Third International Conference on Accelerated Life Testing, Reliability based Analysis and Design (Clermont-Ferrand, France, 2010), Российской НТК «Информатика и проблемы телекоммуникаций» (Новосибирск, 2011), The International Workshop «Applied Methods of Statistical Analysis. Simulations and Statistical Inference» (AMSA’2011, Novosibirsk, 2011), Российской

НТК «Обработка информационных сигналов и математическое моделирование» (Новосибирск, 2012), Всероссийском, с международным участием, научном симпозиуме «НЕПАРАМЕТРИКА — XIV» (Томск, 2012).

Публикации. По теме диссертации опубликовано 14 печатных работ, в том числе 2 статьи в журналах, входящих в перечень изданий, рекомендуемых ВАК РФ, 1 статья в рецензируемом издании серии «Statistics for Industry and Technology», выпускаемом издательством Birkhäuser, Boston (Springer), 1 статья в сборнике научных трудов, 9 работ в сборниках трудов конференций, 1 свидетельство о государственной регистрации программы для ЭВМ.

Структура работы. Диссертация состоит из введения, 5 глав основного содержания, заключения, списка литературы и приложений. Диссертация изложена на 122 страницах основного текста, включая 34 таблицы, 33 рисунка и список литературы из 102 наименований.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

В первой главе делается постановка задач исследования.

При проведении экспериментальных исследований нельзя исключить возможное наличие факторов, приводящих к систематическим ошибкам, смещенности оценок параметров, коррелированности результатов измерений, появлению тренда в той или иной форме. Это приводит к появлению вопросов, связанных с точностью измерений и корректностью статистических выводов. Поэтому при статистическом анализе результатов измерений обычно первым из проверяемых предположений является гипотеза о принадлежности ошибок измерений нормальному закону. Если данная гипотеза не отвергается, то дальнейший анализ упрощается за счет использования классических результатов.

В то же время среди специалистов в области статистического анализа до сих пор нет однозначного мнения относительно преимуществ какого-либо критерия или группы критериев, применяемых для проверки отклонения эмпирического распределения от нормального закона.

С целью изучения свойств и определения предпочтительности применения того или иного критерия на основании сравнительного анализа мощности по отношению к близким конкурирующим гипотезам в настоящей работе рассматривается широкий набор критериев нормальности.

В анализе надежности и выживаемости показательный закон является важной моделью и характеризуется «отсутствием памяти», то есть свойством $P\{X > x+y \mid X > y\} = P\{X > x\}$. В терминах надежности и выживаемости это эквивалентно постоянству функции интенсивности $h(t) = f(t)/(1-F(t))$. В системах массового обслуживания рассматривают пуассоновские потоки, в которых интервалы времени между событиями подчиняются показательному распределению.

В задачах выживаемости и надежности постоянство интенсивности отказов противопоставляют другим формам функции интенсивности, характерным для тех или иных процессов старения, износа, выздоровления, деградации. Поэтому различные авторы исследуют мощность критериев показательности по отношению к близким конкурирующим гипотезам с различной функцией интенсивности отказов. Подобный подход используется и в настоящей работе.

Цель исследований, связанных с критериями показательности, заключалась в анализе существующих критериев и в получении знаний о предпочтительности того или иного критерия в зависимости от альтернативы с той или иной формой функции интенсивности.

Корректность применения статистических методов базируется на справедливости определенных предположений, которые в реальных ситуациях нередко нарушаются. Например, применение асимптотических результатов оказывается неправомерным вследствие ограниченности объемов выборок, использование параметрических критериев может быть некорректным из-за нарушения предположения о нормальности, применение непараметрических критериев согласия при проверке сложных гипотез ограничивается отсутствием знаний о распределениях статистик.

Экспоненциальный рост производительности компьютеров, дальнейшее развитие в сторону увеличения количества вычислительных устройств на одном чипе и возможность организации распределенных вычислений подталкивает к устранению затруднений, связанных с применением критериев проверки гипотез в нестандартных условиях, за счет поиска требуемых закономерностей (распределений статистик) методами статистического моделирования в интерактивном режиме в условиях, имитирующих реальную ситуацию получения наблюдений. В части диссертационной работы, посвященной программной поддержке исследований, решалась задача корректного применения статистических критериев в условиях, когда действительное распределение их статистик приходится получать в ходе проводимого статистического анализа.

Во второй главе проводится исследование особенностей и мощности некоторых критериев нормальности.

Проверке гипотезы о нормальности посвящен ГОСТ Р ИСО 5479-2002, включающий плохо обоснованный и очень ограниченный список критериев, что послужило предметом специального исследования[1]. В данной главе исследованы распределения статистик, уточнены свойства и проведен анализ мощности некоторых критериев нормальности, не получивших ранее широкого освещения.

Статистика критерия Фросини имеет вид

$$B_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left| \Phi\left(z_i\right) - \frac{i-0.5}{n} \right|,$$

---

[1] Лемешко Б.Ю., Лемешко С.Б. Сравнительный анализ критериев проверки отклонения распределения от нормального закона // Метрология. 2005. № 2. С. 3-24.

где $z_i = \left(x_{(i)} - \bar{x}\right) \big/ s, \bar{x} = n^{-1} \sum_{i=1}^{n} x_i, s^2 = n^{-1} \sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2, \Phi(z)$ — функция распределения стандартного нормального закона. Критерий имеет правостороннюю критическую область.

Два критерия Хегази-Грина опираются на статистики

$$T_1 = \frac{1}{n} \sum_{i=1}^{n} \left| z_i - \eta_i \right| \qquad \text{и} \qquad T_2 = \frac{1}{n} \sum_{i=1}^{n} \left\{ z_i - \eta_i \right\}^2,$$

где $z_i = \left(x_{(i)} - \bar{x}\right) \big/ s, \quad \bar{x} = n^{-1} \sum_{i=1}^{n} x_i, \quad s^2 = (n-1)^{-1} \sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2,$ $\eta_i = \Phi^{-1}\left(i/(n+1)\right)$. Критерий имеет правостороннюю критическую область.

В критерии Гири используется статистика

$$d = \frac{1}{ns} \sum_{i=1}^{n} \left| x_i - \bar{x} \right|, \tag{1}$$

где $\bar{x} = n^{-1} \sum_{i=1}^{n} x_i$; $s^2 = n^{-1} \sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2$. Критерий двусторонний.

В критерии Дэвида–Хартли–Пирсона используется статистика

$$U = \frac{R}{s}, \tag{2}$$

где $R = (x_{(n)} – x_{(1)})$; $s^2 = (n-1)^{-1} \sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2$. Критерий двусторонний.

Статистика критерия Шпигельхальтера базируется на комбинации статистик критериев Гири и Дэвида–Хартли–Пирсона и имеет вид

$$T' = \left\{ \left(C_n U\right)^{-(n-1)} + g^{-(n-1)} \right\}^{1/(n-1)} \tag{3}$$

где $C_n = (2n)^{-1} (n!)^{1/(n-1)}$, U — статистика (2); $g = d \big/ \sqrt{(n-1)/n}$, d — статистика (1). Критерий имеет правостороннюю критическую область.

Показано, что вопреки утверждениям автора, распределения статистики критерия Гири при справедливости проверяемой гипотезы асимметричны и плохо аппроксимируются нормальным законом.

Показано, что общим недостатком рассмотренных критериев является зависимость распределений статистик от объема выборки. Если у критерия Фросини распределение статистики достаточно быстро сходится к некоторому предельному и существенно не меняется при $n \geq 100$, то у остальных критериев зависимость от n более сильная (рис. 1).

В ходе исследований построены таблицы процентных точек для более широкого спектра значений n, что расширяет возможности применения данных критериев.

В данном разделе при анализе мощности критериев рассмотрены те же конкурирующие законы, что и в предшествующей работе[1]:

– H_0: нормальный закон, $f(x) = \left(\theta_1 \sqrt{2\pi}\right)^{-1} e^{-\left(x-\theta_0\right)^2 \big/ 2\theta_1^2}$, $(\theta_0, \theta_1)^T = (0,1)^T$;

– H_1: распределение семейства с плотностью

$$f(x)=\frac{\theta_2}{2\theta_1\Gamma(1/\theta_2)}\exp\left\{-\left(|x-\theta_0|/\theta_1\right)^{\theta_2}\right\} \tag{4}$$

и параметром формы $\theta_2=4$ и $(\theta_0,\theta_1)^T=(0,1)^T$.

– H_2: распределение (4) с параметром формы $\theta_2=1$ (распределение Лапласа) и $(\theta_0,\theta_1)^T=(0,1)^T$;

– H_3: логистическое распределение с плотностью

$$f(x)=\frac{\pi}{\theta_1\sqrt{3}}\exp\left\{-\frac{\pi(x-\theta_0)}{\theta_1\sqrt{3}}\right\}\Bigg/\left[1+\exp\left\{-\frac{\pi(x-\theta_0)}{\theta_1\sqrt{3}}\right\}\right]^2,\ (\theta_0,\theta_1)^T=(0,1)^T.$$

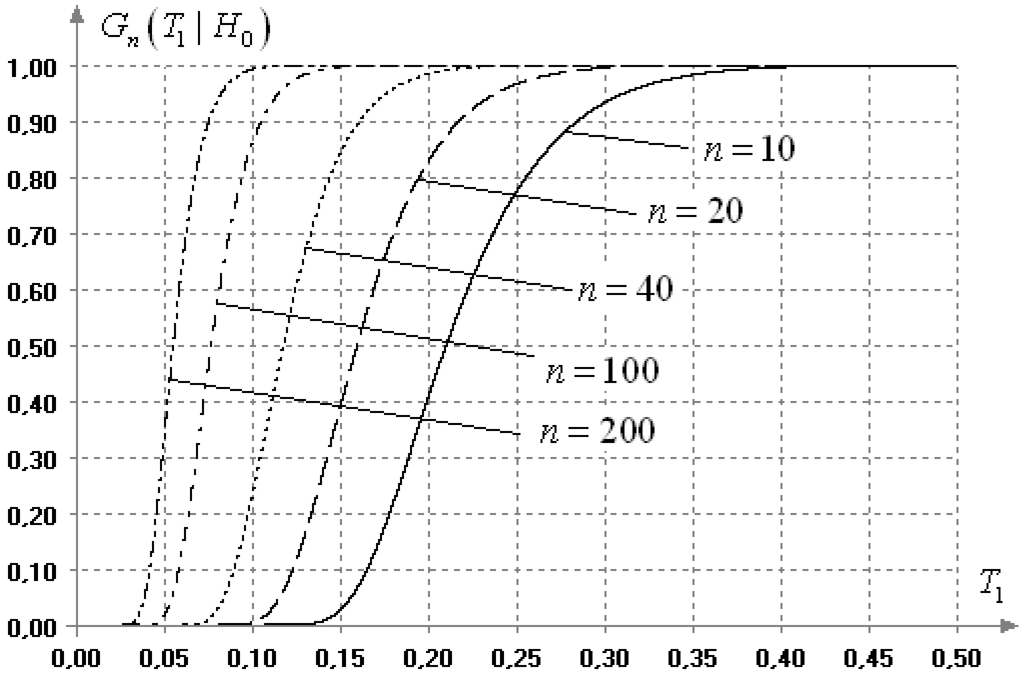


Рис. 1. Зависимость распределения $G(T_1\,|\,H_0)$ статистики T_1 от объема выборки

В результате исследования распределений статистик критериев показано, что критерии Хегази–Грина со статистиками T_1 и T_2 при малых объемах выборок оказываются смещенными относительно гипотезы H_1 (так же, как и критерии Шапиро–Уилка и Эппса–Палли), критерий Шпигельхальтера, показывающий, как правило, хорошую мощность, не способен (даже с возрастанием объемов выборок) отличить H_0 от H_1.

На основании проведенного моделирования и последующего сравнительного анализа мощности рассмотренные критерии можно проранжировать следующим образом:

Гири $\succ$ Шпигельхальтера $\succ$ Хегази–Грина (T_2) $\succ$ Хегази–Грина (T_1) $\succ$ Фросини $\succ$ Дэвида–Хартли–Пирсона.

Однако при выборе и использовании критериев необходимо учитывать отмеченные выше существенные недостатки критериев Шпигельхальтера и Хегази–Грина (рис. 2) и замечание по поводу распределений критерия Гири.

В приведенном ряду предпочтительности критерий Эппса–Палли, включенный в ГОСТ Р ИСО 5479-2002, должен занять место за критерием Хегази–Грина со статистикой T_1, а критерий Шапиро–Уилка — сразу после критерия Дэвида–Хартли–Пирсона.

Отметим здесь же, что по отношению к наиболее близкой конкурирующей гипотезе H_3 критерии согласия Андерсона–Дарлинга и типа χ^2 Никулина уступают по мощности критериям Д'Агостино со статистикой z_2, Хегази–Грина, Шпигельхальтера и Гири, но превосходят остальные критерии проверки отклонения от нормального закона (показано в главе 3).

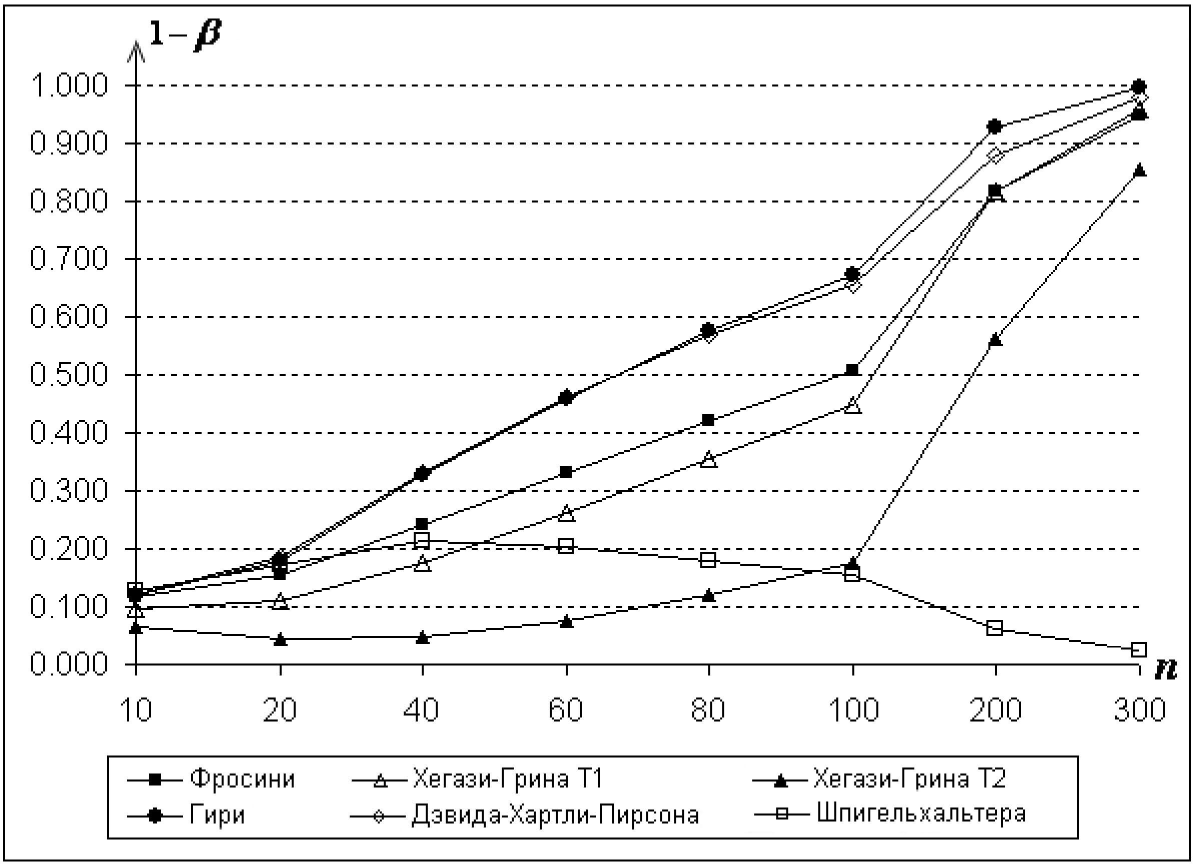


Рис. 2. Мощность критериев по отношению к конкурирующей гипотезе H_1 в зависимости от n при $\alpha = 0.1$

В третьей главе мощность критериев нормальности исследуется на примере наблюдений в классических экспериментах. Результаты классических экспериментов, связанных с измерениями физических констант, вызывают интерес не только в силу их исторической ценности, но и как примеры высокого уровня организации и проведения соответствующих измерений. В данной главе рассматриваются результаты четырех экспериментов: Кавендиша — по измерению средней плотности Земли, Милликена — по определению заряда электрона, Майкельсона и Ньюкомба — по измерению скорости света (табл. 1–4). Принято считать, что погрешности измерений в этих экспериментах подчиняются нормальному закону. Важным является и то обстоятельство, что наблюдения фиксировались и приводятся в первоисточниках полностью, без пропусков и без цензурирования.

К описанию ошибок измерений подобных результатов кроме нормального закона можно пытаться использовать и другие модели. В рамках настоящего исследования рассмотрены возможные конкурирующие законы и исследована мощность множества критериев относительно них.

Кроме описанных выше критериев Фросини, Хегази–Грина, Гири, Дэвида–Хартли–Пирсона и Шпигельхальтера были рассмотрены специализированные критерии Шапиро–Уилка, Ройстона, Эппса–Палли, Д'Агостино; непараметрические критерии согласия Колмогорова, Андерсона–Дарлинга и Крамера–Мизеса–Смирнова; критерии согласия типа χ^2 Пирсона и Никулина (Никулина–Рао–Робсона).

Весь перечень критериев был использован для проверки гипотезы о принадлежности каждого из рассматриваемых наборов данных нормальному закону. Полученные в результате проверки достигнутые уровни значимости показали, что нет оснований для отклонения проверяемой гипотезы.

Таблица 1

Измерения средней плотности Земли, полученные Кавендишем, г/см^3.

5.50	5.55	5.57	5.34	5.42	5.30
5.61	5.36	5.53	5.79	5.47	5.75
4.88	5.29	5.62	5.10	5.63	5.68
5.07	5.58	5.29	5.27	5.34	5.85
5.26	5.65	5.44	5.39	5.46	

Таблица 2

Измерения заряда электрона, полученные Милликеном,
Фр =1/2 997 924 580 Кл.

4.781	4.782	4.764	4.768	4.779	4.761	4.790	4.749
4.795	4.767	4.774	4.801	4.788	4.792	4.747	4.781
4.769	4.764	4.778	4.785	4.772	4.758	4.769	
4.792	4.776	4.791	4.783	4.791	4.764	4.806	
4.779	4.771	4.777	4.808	4.788	4.810	4.779	
4.775	4.789	4.765	4.771	4.783	4.799	4.785	
4.772	4.772	4.785	4.809	4.740	4.799	4.790	
4.791	4.789	4.805	4.790	4.775	4.797	4.777	

Таблица 3

Измерения Майкельсона (+299 000, км/ч).

850	980	1000	880	810	880	970	850	800	890	890	820	950
740	880	1000	800	880	880	950	840	770	860	840	850	800
900	1000	960	850	880	880	880	840	760	880	780	870	810
1070	980	960	880	830	860	910	840	740	720	810	870	870
930	930	960	900	800	720	850	890	750	840	760	810	
850	650	940	840	790	720	870	810	760	850	810	740	
950	760	960	830	760	620	840	810	910	850	790	810	
980	810	940	790	800	860	840	820	920	780	810	940	

С целью выбора моделей конкурирущих законов, к каждому из наборов данных была применена следующая процедура. Из множества законов распределения, реализованных в программной системе ISW, были отобраны те, относительно которых проверяемая гипотезы о принадлежности к ним соответствующих наборов, осуществляемая с использованием критериев согласия Колмогорова, Андерсона–Дарлинга, Крамера–Мизеса–Смирнова и Никулина, не отклонялась. До трех моделей законов, относительно которых при проверке гипотез были получены наиболее высокие достигаемые уровни значимости, были выбраны в качестве конкурирующих законов. Мощность рассматриваемых критериев исследовалась относительно этих конкурирующих законов.

Таблица 4

Измерения Ньюкомба (×10^{-3} + 24.8, с×10^{-6})

28	40	23	36	28	36	26	24	16
26	-2	29	28	26	23	33	25	23
33	29	31	25	30	27	26	32	
24	22	19	21	32	27	32	25	
34	24	24	28	36	28	32	29	
-44	21	20	29	26	27	24	27	
27	25	36	37	30	31	39	28	
16	30	32	25	22	27	28	29	

Для наблюдений Кавендиша с объемом выборки $n = 29$ наилучшими моделями оказались: нормальный закон с параметрами $\theta_0 = 5.4479$ и $\theta_1 = 0.2171$; распределение Лапласа с параметрами $\theta_0 = 5.46$ и $\theta_1 = 0.1737$; логарифмически нормальный закон $f(x) = \left(x\theta_1\sqrt{2\pi}\right)^{-1} \exp\left\{-\left(\ln x - \theta_0\right)^2 / 2\theta_1^2\right\}$ с $\theta_0 = 1.6944$ и $\theta_1 = 0.0403$; логистический закон с $\theta_0 = 5.456$ и $\theta_1 = 9.8961$.

Для данных Милликена с объемом выборки $n = 58$ среди отобранных моделей оказались: нормальный закон с $\theta_0 = 4.7808, \theta_1 = 0.0152$, распределение Лапласа $(\theta_0 = 4.7813, \theta_1 = 0.0125)$ и распределение семейства (4) с параметрами $\theta_0 = 4.7808$, $\theta_1 = 0.0214$, $\theta_2 = 1.9903$.

Для наблюдений Майкельсона с объемом выборки $n = 100$ наилучшие модели: нормальный закон $(\theta_0 = 852.3992, \theta_1 = 78.6145)$, логистический $(\theta_0 = 851.4711, \theta_1 = 80.4661)$ и логарифмически нормальный $(\theta_0 = 6.7437$, $\theta_1 = 0.0935)$.

Для редуцированной выборки Ньюкомба (без выбросов) лучшие модели: нормальное распределение с параметрами $(\theta_0 = 27.75, \theta_1 = 5.0436)$, распределение Лапласа $(\theta_0 = 27.75, \theta_1 = 3.0963)$ и логарифмически нормальное распределение $(\theta_0 = 3.3061, \theta_1 = 0.1876)$.

В результате сравнительного анализа мощности, показанной критериями относительно рассматриваемых конкурирующих законов при объемах выборок, соответствующих экспериментам, специальные критерии нормальности можно упорядочить по мощности следующим образом:
относительно распределения Лапласа —

Шпигельхальтера $\succ$ *Гири* $\sim$ *Хегази–Грина* $T_2 \succ$ *Хегази–Грина* $T_1 \succ$ *Фросини* $\sim$ *Эппса–Палли* $\sim$ *Ройстона* $\succ$ *Шапиро–Уилка* $\succ$ *Дэвида–Хартли–Пирсона* $\succ$ *Д’Агостино;*

относительно логистического распределения —

Шпигельхальтера $\sim$ *Хегази–Грина* $T_2 \succ$ *Гири* $\sim$ *Хегази–Грина* $T_1 \succ$ *Ройстона* $\succ$ *Эппса–Палли* $\succ$ *Шапиро–Уилка* $\succ$ *Дэвида–Хартли–Пирсона* $\succ$ *Д'Агостино* $\succ$ *Фросини;*

относительно логарифмически нормального закона —

Ройстона $\succ$ *Эппса–Палли* $\sim$ *Шапиро–Уилка* $\succ$ *Хегази–Грина* $T_2 \succ$ *Хегази–Грина* $T_1 \succ$ *Фросини* $\succ$ *Шпигельхальтера* $\succ$ *Д'Агостино* $\succ$ *Гири* $\succ$ *Дэвида–Хартли–Пирсона.*

Полученные оценки мощности показывают, что к наиболее предпочтительным следует отнести критерии Шпигельхальтера, Хегази–Грина T_1 и T_2, Эппса–Палли, Ройстона, Шапиро–Уилка, Гири. При этом следует учитывать, что критерии Шпигельхальтера, Хегази–Грина T_1 и T_2, Эппса–Палли, Ройстона и Шапиро–Уилка, имея преимущество в мощности, обладают серьезным недостатком: критерии смещены относительно некоторых конкурирующих гипотез, в частности, по отношению к распределению семейства (4) с параметром формы $\theta_2 = 4$.

Непараметрические критерии согласия и критерии типа χ^2, имея меньшую мощность, не обладают подобным недостатком.

В четвертой главе проводится исследование критериев показательности. Пусть $Exp(\theta)$ — показательное распределение с плотностью $f(x) = \exp(-x/\theta)/\theta,\ x \geq 0$. Будем рассматривать критерии согласия с классом показательных распределений $\{Exp(\theta) : \theta > 0\}$. При заданной выборке независимых наблюдений $X_1, ..., X_n$ неотрицательной случайной величины X с плотностью $f(x)$ и функцией распределения $F(x)$ проверяется сложная гипотеза H_0 о принадлежности X классу $\{Exp(\theta) : \theta > 0\}$ против общих альтернатив.

Ранее предпринимались попытки провести сравнительный анализ некоторых наборов критериев показательности. Среди таких работ следует выделить обзоры Эшера[2] и Хензе и Мейнтаниса[3], из которых почерпнута существенная часть критериев, рассмотренных в настоящей работе.

***F*-критерий Гнеденко** предназначен для проверки гипотезы показательности против конкурирующей гипотезы H_1: распределение имеет монотонную интенсивность отказов. В данном критерии наблюдения разбиваются на две группы, первая из которых содержит *R* наименьших наблюдений, а вторая — остальные *n-R* наблюдений. Статистика критерия имеет вид

---

[2] Ascher S. A survey of tests for exponentiality // Communications in Statistics - Theory and Methods. 1990. Vol. 19. No. 5. pp. 1811-1825.

[3] Henze N., Meintanis S.G. Recent and classical tests for exponentiality: a partial review with comparisons // Metrika. 2005. Vol. 61. pp. 29-45.

$$Q_R = \left(\sum_{j=1}^{R} D_j / R \right) \Bigg/ \sum_{j=R+1}^{n} D_j / (n-R),$$

где $D_j = (n-j+1)\left(X_{(j)} - X_{(j-1)}\right)$, $1 \le j \le n$, $X_{(0)} \equiv 0$, $X_{(j)}$ — соответствующая порядковая статистика. Если H_0 верна, статистика Q_R подчиняется F-распределению Фишера с $2R$ и $2(n-R)$ степенями свободы. Гипотеза H_0 отклоняется при малых и больших значениях Q_R.

Критерий Харриса представляет собой модификацию F-критерия Гнеденко для проверки показательности при наличии альтернатив с выпуклой интенсивностью отказов. Статистика критерия

$$Q'_R = \left[\left(\sum_{j=1}^{R} D_j + \sum_{j=n-R+1}^{n} D_j \right) \Bigg/ 2R \right] \Bigg/ \sum_{j=R+1}^{n-R} D_j / (n-2R)$$

при справедливости H_0 подчиняется F-распределению с $4R$ и $2(n\text{-}2R)$ степенями свободы. Гипотеза отклоняется при малых и больших значениях Q'_R.

Статистика **критерия Холландера–Прошана**:

$$T = \sum_{i>j>k} \mathbf{1}\left\{ X_{(i)} > X_{(j)} + X_{(k)} \right\}.$$

Критерий двусторонний, авторы критерия приводят таблицы приближенных критических значений и следующую нормальную аппроксимацию:

$$T^* = \left(T - \mathrm{E}[T \mid H_0]\right) \Big/ \left(\mathrm{D}[T \mid H_0]\right)^{1/2},$$

где $\mathrm{E}(T \mid H_0) = n(n-1)(n-2)/8$ и $\mathrm{D}[T \mid H_0] = 1.5n(n-1)(n-2) \times$ $\times \left[2(n-3)(n-4)/2592 + 7(n-3)/432 + 1/48 \right]$.

Критерий Гини является двусторонним и основан на статистике

$$G_n = \sum_{j,k=1}^{n} \left| Y_j - Y_k \right| \Big/ 2n(n-1),$$

где $Y_j = X_j / \hat{\theta}_n$, $\hat{\theta}_n$ — оценка максимального правдоподобия параметра θ.

В **критерии согласия Колмогорова** в качестве меры отличия эмпирического распределения от показательного закона используется величина

$$D_n = \sup_{x \ge 0} \left| F_n(x) - \left(1 - \exp\left(-x/\hat{\theta}\right)\right) \right| = \max\left\{ \max_{i \le j \le n} \left[\frac{j}{n} - Z_{(j)} \right], \max_{i \le j \le n} \left[Z_{(j)} - \frac{j-1}{n} \right] \right\},$$

где $Z_j = 1 - \exp\left(-Y_j\right)$. Для уменьшения зависимости распределения статистики от объема выборки следует использовать статистику с поправкой Большева:

$$K_n = (6n \cdot D_n + 1) \big/ 6\sqrt{n}.$$

В **критерии согласия Крамера–Мизеса–Смирнова** при проверке гипотезы о принадлежности выборки показательному закону используется статистика

$$CMS_n = \frac{1}{12n} + \sum_{j=1}^{n} \left(Z_{(j)} - \frac{2j-1}{2n} \right)^2 .$$

Статистика **критерия согласия Андерсона–Дарлинга** при проверке гипотезы о принадлежности выборки показательному закону имеет вид

$$AD_n = -n - 2\sum\nolimits_{j=1}^{n} \left\{ \frac{2j-1}{2n} \ln Z_j + \left(1 + \frac{2j-1}{2n} \right) \ln\left(1 - Z_j\right) \right\}.$$

Статистика **Барингхауса–Хензе критерия типа Колмогорова** имеет вид

$$\bar{K}_n = \sqrt{n} \sup_{t \ge 0} \left| \frac{1}{n} \sum_{j=1}^{n} \min\left(Y_j, t\right) - \frac{1}{n} \sum_{j=1}^{n} \mathbf{1}\left\{Y_j \le t\right\} \right| = \sqrt{n} \max\left(K_n^+, K_n^-\right),$$

где

$$K_n^+ = \max_{j=0,1,\dots,n-1} \left[n^{-1}\left(Y_{(1)} + \dots + Y_{(j)}\right) + Y_{(j+1)}\left(1 - j/n\right) - j/n \right],$$

$$K_n^- = \max_{j=0,1,\dots,n-1} \left[j/n - n^{-1}\left(Y_{(1)} + \dots + Y_{(j)}\right) - Y_{(j)}\left(1 - j/n\right) \right].$$

Как и в классическом критерии Колмогорова, в данном случае целесообразно использовать статистику с поправкой Большева

$$K_n^* = \left(6n \cdot \bar{K}_n \big/ \sqrt{n} + 1\right) \big/ 6\sqrt{n} .$$

Статистика **Барингхауса–Хензе критерия типа Крамера–Мизеса–Смирнова** определяется выражением

$$CMS_n^* = \frac{1}{n} \sum_{j,k=1}^{n} \left[2 - 3e^{-\min\left(Y_j, Y_k\right)} - 2\min\left(Y_j, Y_k\right)\left(e^{-Y_j} + e^{-Y_k}\right) + 2e^{-\max\left(Y_j, Y_k\right)} \right].$$

Критерий Дешпанде был предложен для проверки показательности против конкурирующих распределений с возрастающей интенсивностью отказов. Статистика критерия вычисляется в соответствии с соотношением

$$J_b = n\left(n-1\right)^{-1} \sum \mathbf{1}\left\{X_j > bX_k\right\},$$

где суммирование ведется по всем $1 \le j,\ k \le n$ таким, что $j \ne k$.

Критерий Кокса–Оукса. Проверяемая гипотеза отклоняется при малых и больших значениях статистики

$$CO_n = n + \sum_{j=1}^{n} \left(1 - Y_j\right) \ln Y_j .$$

Критерий Большева. Данный критерий предназначен для проверки гипотезы о принадлежности совокупности малых выборок показательным законам распределения вероятностей, возможно, с разными параметрами.

Критерий Клара основан на интегрированной функции распределения и отвергает гипотезу о показательности при больших значениях статистики

$$KL_{n,a}=\frac{2(3a+2)n}{(2+a)(1+a)^2}-2a^3\sum_{j=1}^{n}\frac{\exp\left(-(1+a)Y_j\right)}{(1+a)^2}-\frac{2}{n}\sum_{j=1}^{n}\exp\left(-aY_j\right)+$$
$$+\frac{2}{n}\sum_{j<k}\left[a\left(Y_{(k)}-Y_{(j)}\right)-2\right]\exp\left(-aY_{(j)}\right).$$

Автор предлагает использовать критерий $KL_n^{1,10}$, отвергающий гипотезу о показательности, если ее отвергает хотя бы один из критериев $KL_{n,1}$ и $KL_{n,10}$.

В **критерии Барингхауса–Хензе** используется факт, что $\psi(t)=E\left[\exp(-tX)\right]=\lambda/(t+\lambda)$ ($X\sim Exp$) удовлетворяет дифференциальному уравнению $(\lambda+t)\psi'(t)+\psi(t)=0$, $t\in R$. Выбирая константу $a>0$, можно получить критерий, отклоняющий гипотезу показательности для больших значений статистики

$$BH_{n,a}=n^{-1}\sum_{j,k=1}^{n}\left[\frac{\left(1-Y_j\right)\left(1-Y_k\right)}{Y_j+Y_k+a}-\frac{Y_j+Y_k}{\left(Y_j+Y_k+a\right)^2}+\frac{2Y_jY_k}{\left(Y_j+Y_k+a\right)^2}+\frac{2Y_jY_k}{\left(Y_j+Y_k+a\right)^3}\right].$$

В **критерии Хензе** гипотеза о показательности отклоняется при больших значениях статистики

$$HE_{n,a}=\frac{1}{n}\sum_{j,k=1}^{n}\frac{1}{Y_j+Y_k+a}-\sum_{j=1}^{n}\exp\left(Y_j+a\right)E_1\left(Y_j+a\right)+n\left(1-a\exp(a)E_1(a)\right),$$

где $E_1(z)=\int_z^\infty t^{-1}\exp(-t)dt$ — показательный интеграл, $a>0$ — константа.

В L**-критерии Хензе–Мейнтаниса** используется статистика

$$L_{n,a}=n\int_0^\infty\left(\psi_n(t)-\frac{1}{1+t}\right)^2(1+t)^2e^{-at}dt=\frac{1}{n}\sum_{j,k=1}^{n}\frac{1+\left(Y_j+Y_k+a+1\right)^2}{\left(Y_j+Y_k+a\right)^3}-2\sum_{j=1}^{n}\frac{1+Y_j+a}{\left(Y_j+a\right)^2}+\frac{n}{a}.$$

Проверяемая гипотеза о принадлежности выборки показательному закону распределения отклоняется при больших значениях статистики.

Критерий показательности Эппса–Палли опирается на эмпирическую характеристическую функцию. Распределение статистики критерия

$$EP_n=(48n)^{1/2}\left[\frac{1}{2}\sum_{j=1}^{n}\exp\left(-Y_j\right)-1/2\right]$$

при $n\to\infty$ описывается стандартным нормальным законом. Проверяемая гипотеза отклоняется при больших значениях $|EP_n|$.

***W*-критерии Хензе–Мейнтаниса**, также опираются на эмпирическую характеристическую функцию. Проверяемая гипотеза отклоняется при больших значениях статистик

$$W_{n,a}^{(1)}=\frac{a}{2n}\sum_{j,k=1}^{n}\left[\frac{1}{a^2+\left(Y_{jk}^{-}\right)^2}-\frac{1}{a^2+\left(Y_{jk}^{+}\right)^2}-\frac{4Y_{jk}^{+}}{\left(a^2+\left(Y_{jk}^{+}\right)^2\right)^2}+\frac{2a^2-6\left(Y_{jk}^{-}\right)^2}{\left(a^2+\left(Y_{jk}^{-}\right)^2\right)^3}+\frac{2a^2-6\left(Y_{jk}^{+}\right)^2}{\left(a^2+\left(Y_{jk}^{+}\right)^2\right)^3}\right],$$

$$W_{n,a}^{(2)}=\frac{\sqrt{\pi}}{4n\sqrt{a}}\sum_{j,k=1}^{n}\left[\left(1+\frac{2a-\left(Y_{jk}^{-}\right)^2}{4a^2}\right)\exp\left(-\frac{\left(Y_{jk}^{-}\right)^2}{4a}\right)+\left(\frac{2a-\left(Y_{jk}^{+}\right)^2}{4a^2}-\frac{Y_{jk}^{+}}{a}-1\right)\exp\left(-\frac{\left(Y_{jk}^{+}\right)^2}{4a}\right)\right],$$

где $Y_{jk}^{+}=Y_j+Y_k$, $Y_{jk}^{-}=Y_j-Y_k$.

Для одних критериев распределения статистик получены при построении, для некоторых других авторами предложены нормализующие преобразования. На практике для выборок конечного объема такие асимптотические результаты могут оказаться неприемлемыми. Необходимо знать, начиная с каких объемов выборок можно использовать асимптотическое распределение для статистики соответствующего критерия.

Проведенные исследования показали, что при использовании критериев Гнеденко, Харриса, Гини, Колмогорова, Крамера–Мизеса–Смирнова, Андерсона–Дарлинга и Большева применение предельных (теоретических) распределений статистик остается корректным и в условиях ограниченных объемов выборок, что позволяет при проверке гипотезы определить точный достигнутый уровень значимости.

Для критериев Барингхауса–Хензе со статистиками K^* и CMS^* построены модели распределений статистик, применимые при $n \ge 20$.

Показано, что нормальную аппроксимацию распределения статистики критерия Холландера–Прошана можно с определенной погрешностью использовать при $n \ge 100$; при $n \ge 600$ применение нормальной аппроксимации не приводит к существенным ошибкам.

Показано, что в критериях Дешпанде (со статистикой $J_{0.5}$), Эппса–Палли и Кокса–Оукса применение нормальной аппроксимации не приводит к существенным ошибкам при $n \ge 100$.

Для остальных критериев (Клара, Барингхауса–Хензе со статистикой BH, Хензе, L-критерия Хензе–Мейнтаниса, W-критериев Хензе–Мейнтаниса) следует использовать соответствующие таблицы процентных точек.

Мощность критериев сравнивалась на объемах выборок n=20 и n=50. Эмпирические распределения статистик критериев, соответствующие проверяемой и конкурирующим гипотезам, для получения приемлемой точности строились по 1 660 000 испытаниям. В качестве конкурирующих гипотез рассматривались распределения, принадлежащие к трем классам: с возрастающими, убывающими и немонотонными интенсивностями отказов:

- Вейбулла $W(\theta)$ с плотностью $f(x)=\theta x^{\theta-1}\exp\left(-x^{\theta}\right)$;
- гамма-распределение $\Gamma(\theta)$ — $f(x)=\Gamma(\theta)^{-1}x^{\theta-1}\exp(-x)$;

- бета-распределение $\mathrm{B}(\theta_0,\theta_1)$ — $f(x)=\mathrm{B}(\theta_0,\theta_1)^{-1}x^{\theta_0-1}(1-x)^{\theta_1-1}$;
- равномерное $U(0,1)$ на интервале $[0,1]$;
- логнормальное $LN(\theta)$ — $f(x)=\left(\theta x\sqrt{2\pi}\right)^{-1}\exp\left(-(\ln x)^2/2\theta^2\right)$;
- полунормальное HN — $f(x)=(2/\pi)^{1/2}\exp\left(-x^2/2\right)$.

К распределениям с возрастающими интенсивностями отказов относятся $W(\theta)$ и $\Gamma(\theta)$ при $\theta>1$, $U(0,1)$, HN, $\mathrm{B}(1,2)$, $\mathrm{B}(2,1)$; с убывающими — $W(\theta)$ и $\Gamma(\theta)$ при $\theta<1$; с немонотонными — LN, $\mathrm{B}(0.5,1)$.

Критерий Барингхауса–Хензе со статистикой BH и критерий Хензе ведут себя схожим образом. Показано, что выбор параметра $a=0.5$ обеспечивает более высокую мощность этих критериев по сравнению с другими значениями a. Показано, что в L-критерии Хензе–Мейнтаниса в общем случае имеет смысл выбирать статистику L_1, в W-критериях Хензе–Мейнтаниса — статистику $W_1^{(1)}$, в критерии Клара — статистику $KL^{1,10}$.

Показано, что для достижения наибольшей мощности в критерии Гнеденко следует полагать R равным $[0.3n]$ из рассмотренных возможных целых значений $[0.1n]$, $[0.2n]$, …, $[0.9n]$, а в критерии Харриса — равным $[0.1n]$ из $[0.05n]$, $[0.1n]$, …, $[0.45n]$.

В случае конкурирующих законов с возрастающей интенсивностью отказов необходимо отметить следующие недостатки исследуемых критериев. При $n=20$ критерий Большева смещен относительно конкурирующих законов $W(1.2)$, $\Gamma(1.5)$, полунормального HN и $\mathrm{B}(1,2)$ (т.е. мощность критерия оказывается меньше заданного уровня значимости); критерий $L_{0.1}$ смещен относительно тех же законов и W(1.4); критерий $W_{2.5}^{(2)}$ оказывается смещенным относительно $W(1.2)$. Относительно конкурирующих законов с убывающей интенсивностью отказов примечательно низкую мощность показывает критерий Харриса со статистикой $Q'_{0.1}$.

Показано, что в случае конкурирующих законов с немонотонными интенсивностями отказов критерии со статистиками $W_{2.5}^{(2)}$ Хензе–Мейнтаниса и BH_5 Барингхауса–Хензе смещены относительно альтернативы $\mathrm{B}(0.5,1)$, критерий $L_{0.1}$ Хензе–Мейнтаниса — относительно $LN(1)$ и $LN(0.8)$.

Исследования показали, что по отношению к конкурирующим законам с возрастающими и убывающими интенсивностями отказов стабильно высокую мощность демонстрируют критерии Кокса–Оукса, Андерсона–Дарлинга, Хензе–Мейнтаниса (L_1 и $W_1^{(1)}$) и Барингхауса–Хензе $(BH_{0.5})$.

По отношению к альтернативам с немонотонными интенсивностями отказов высокой мощностью обладают критерии Харриса $(Q'_{0.1})$ и Андерсона–Дарлинга.

Показано, что при малых объемах выборок (или, если не указана конкретная альтернатива) вследствие возможной смещенности нежелательно применение критериев Харриса $\left(Q'_{0.1}\right)$, Большева (при малом суммарном объеме выборок), Хензе–Мейнтаниса ($L_{0.1}$ и $W_{2.5}^{(2)}$), Барингхауса–Хензе$\left(BH_{0.5}\right)$.

Чтобы выбрать наиболее мощный критерий показательности при наличии заданной альтернативы, выходящей за рамки рассмотренных в данной работе конкурирующих гипотез, необходимо провести исследование мощности критериев по аналогичной методике. При этом следует учесть знание об интенсивности отказов, характеризующей данную альтернативу.

Показано, что критерий показательности Большева (B) имеет достаточно высокую мощность против конкурирующих законов с убывающими интенсивностями отказов, а в случае иных альтернатив он уступает другим рассмотренным критериям. При этом следует иметь в виду, что основным достоинством критерия Большева является подход, позволяющий проверять гипотезу о показательности по совокупности малых выборок.

Пятая глава посвящена программному обеспечению проведения исследований.

Корректность применения статистических методов и формируемых выводов требует выполнения «стандартных» предположений, которые нередко нарушаются на практике. Отсутствие аналитических результатов, позволяющих решать соответствующие задачи в условиях нарушения стандартных предположений, может быть компенсировано построением требуемых закономерностей в результате статистического моделирования в условиях, имитирующих реальную ситуацию получения наблюдений.

В качестве распространенной задачи, выигрывающей от применения распределенных вычислений, в работе рассматривается проверка сложных гипотез с использованием классических критериев согласия (Колмогорова, Андерсона–Дарлинга, Крамера–Мизеса–Смирнова). Распределения статистик этих критериев при проверке сложных гипотез зависят от объема выборок (при малых n), вида наблюдаемого закона, используемого метода оценивания, набора оцениваемых параметров, а иногда, что наиболее неприятно, — от значений параметров (не обязательно из числа тех, которые подвергаются оцениванию). Последнее характерно, например, для гамма-распределения, двустороннего экспоненциального, обратного гауссовского, обобщенного распределения Вейбулла и семейств бета-распределений.

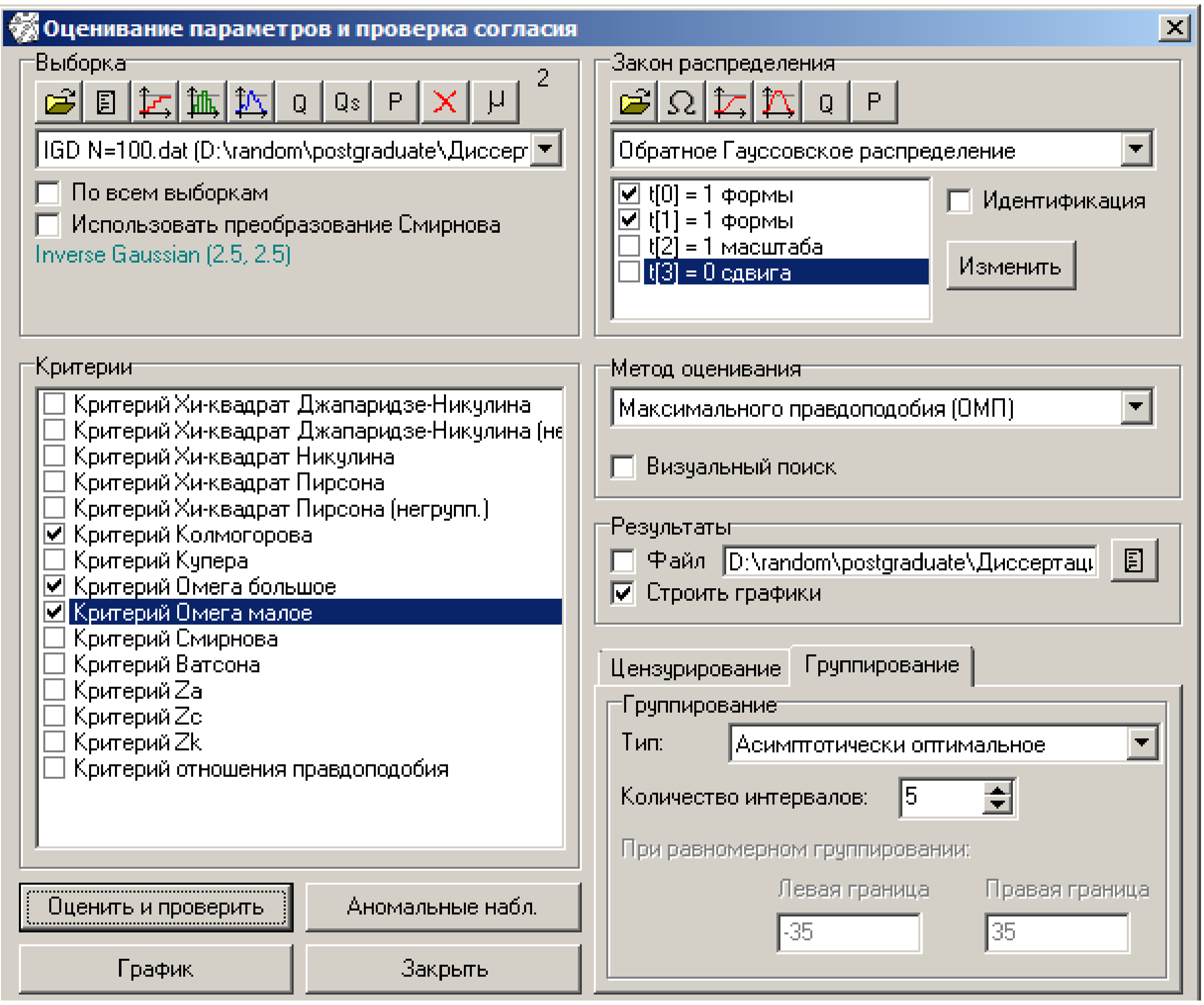
При проверке сложной гипотезы в ситуации, когда распределение статистики критерия, соответствующее справедливости проверяемой гипотезы, неизвестно и должно быть построено в результате компьютерного моделирования в ходе выполнения проверки данной гипотезы (в интерактивном режиме) требуется последовательно выполнить следующие шаги:

1. Вычислить оценки параметров распределения (в случае проверки сложной гипотезы о согласии).
2. Вычислить значение статистики критерия.

3. Смоделировать эмпирическое распределение $G_N\left(S_n \mid H_0\right)$ статистики критерия при том же объеме n, что и в исследуемой выборке, и значениях параметров, равных значениям полученных оценок. Объём N моделируемой выборки статистик определяется требуемой точностью и числом доступных для распараллеливания вычислительных устройств.
4. По полученному эмпирическому распределению статистики определить достигнутый уровень значимости, соответствующий вычисленному значению статистики критерия.

Подобная процедура может быть удобной, если она автоматизирована в программном обеспечении. В данной работе с применением описанного подхода в качестве модуля программной системы ISW реализована проверка гипотез с применением критериев согласия.

После нажатия кнопки «Оценить и проверить» (рис. 3) пользователю будет предложено получить при помощи моделирования достигаемые уровни значимости по тем критериям, для которых в данном случае распределения статистик неизвестны (рис. 4). Здесь распределения неизвестны для всех трех выбранных критериев. Чтобы получить распределения всех критериев, перечисленных в таблице, следует нажать кнопку «Начать».



Оценивание параметров и проверка согласия

Выборка

IGD N=100.dat (D:\random\postgraduate\Диссер

☐ По всем выборкам
☐ Использовать преобразование Смирнова

Inverse Gaussian (2.5, 2.5)

Закон распределения

Обратное Гауссовское распределение

☑ t[0] = 1 формы
☑ t[1] = 1 формы
☐ t[2] = 1 масштаба
☐ t[3] = 0 сдвига

☐ Идентификация

Критерии

☐ Критерий Хи-квадрат Джапаридзе-Никулина
☐ Критерий Хи-квадрат Джапаридзе-Никулина (не
☐ Критерий Хи-квадрат Никулина
☐ Критерий Хи-квадрат Пирсона
☐ Критерий Хи-квадрат Пирсона (негрупп.)
☑ Критерий Колмогорова
☐ Критерий Купера
☑ Критерий Омега большое
☑ Критерий Омега малое
☐ Критерий Смирнова
☐ Критерий Ватсона
☐ Критерий Za
☐ Критерий Zc
☐ Критерий Zk
☐ Критерий отношения правдоподобия

Метод оценивания

Максимального правдоподобия (ОМП)

☐ Визуальный поиск

Результаты

☐ Файл D:\random\postgraduate\Диссертаці
☑ Строить графики

Группирование

Тип: Асимптотически оптимальное

Количество интервалов: 5

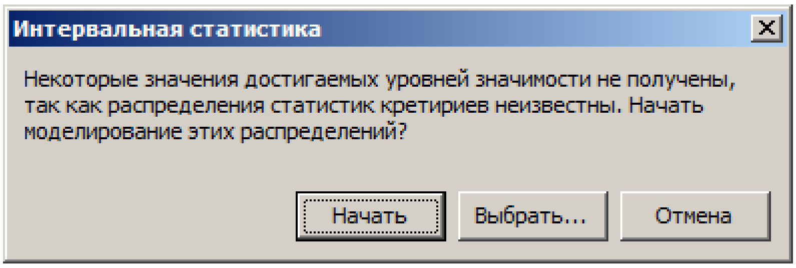
При равномерном группировании:

Левая граница -35

Правая граница 35

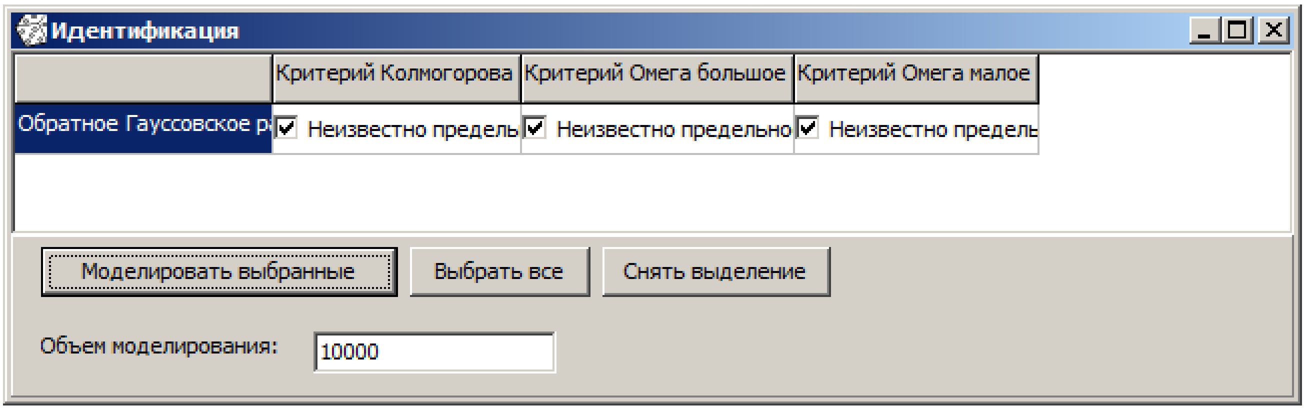
Рис. 3. Экранная форма для проверки гипотез о согласии

Если требуется исключить из процесса моделирования те или иные критерии или изменить объем моделирования, следует нажать кнопку «Выбрать...» и сделать соответствующие изменения. На рис. 5 объем моделирования изменен с 1000 до 10000. После необходимых настроек следует нажать кнопку «Моделировать выбранные». После завершения моделирования в ячейках таблицы будут указаны полученные оценки достигаемых уровней значимости (рис. 6).



Некоторые значения достигаемых уровней значимости не получены, так как распределения статистик кретириев неизвестны. Начать моделирование этих распределений?

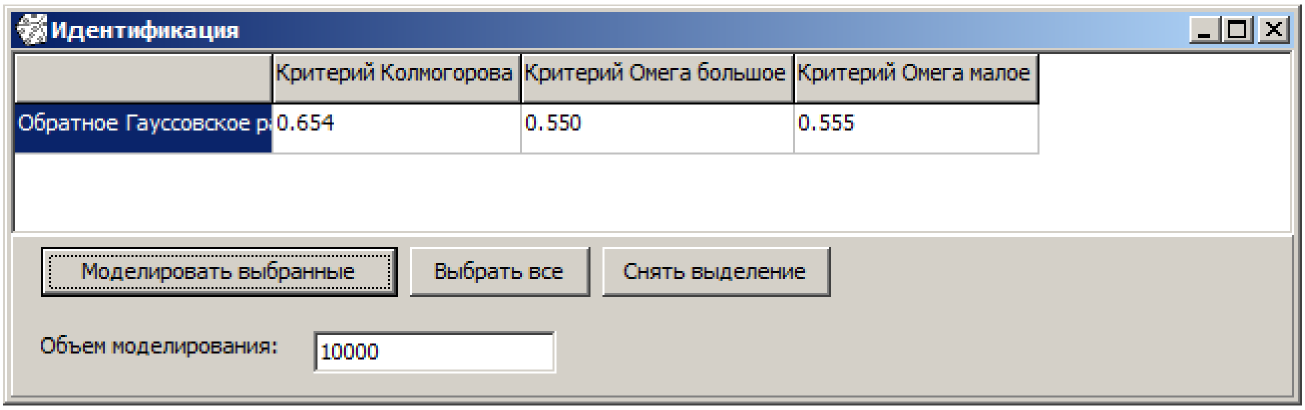
Рис. 4. Диалоговое окно, предлагающее получить достигаемые уровни значимости при помощи моделирования



	Критерий Колмогорова	Критерий Омега большое	Критерий Омега малое
Обратное Гауссовское р	☑ Неизвестно предель	☑ Неизвестно предельно	☑ Неизвестно предель

Объем моделирования: 10000

Рис. 5. Выбор критериев и объема моделирования



	Критерий Колмогорова	Критерий Омега большое	Критерий Омега малое
Обратное Гауссовское р	0.654	0.550	0.555

Объем моделирования: 10000

Рис. 6. Оценки достигаемых уровней значимости, полученные в результате моделирования.

ЗАКЛЮЧЕНИЕ

В соответствии с целями исследований получены следующие результаты.

1. Исследованы распределения статистик критериев нормальности Фросини, Хегази–Грина, Гири, Дэвида–Хартли–Пирсона, Шпигельхальтера, Шапиро–Уилка, Ройстона, Эппса–Палли, Д'Агостино, критериев согласия Колмогорова,

Андерсона–Дарлинга, Крамера–Мизеса–Смирнова, χ^2 Пирсона и χ^2 Никулина (при проверке сложной гипотезы о нормальности с вычислением ОМП параметров) при справедливости проверяемой гипотезы.

2. Методами статистического моделирования исследована мощность перечисленных критериев по отношению к близким конкурирующим гипотезам. Проведен сравнительный анализ критериев и сделаны рекомендации по применению критериев при наличии альтернатив того или иного вида. Показаны достоинства и недостатки отдельных критериев. Расширены таблицы процентных точек.

3. Исследованы распределения статистик критериев показательности Гнеденко, Харриса, Холландера–Прошана, Гини, Колмогорова, Крамера–Мизеса–Смирнова, Андерсона–Дарлинга, Дешпанде, Кокса–Оукса, Большева, Клара, Барингхауса–Хензе, Хензе, Хензе–Мейнтаниса и Эппса–Палли при справедливости проверяемой гипотезы.

4. Методами статистического моделирования получены оценки мощности перечисленных критериев показательности по отношению к конкурирующим гипотезам с различной формой функции интенсивности отказов. Проведен сравнительный анализ мощности критериев, сделаны рекомендации по применению критериев при наличии альтернатив того или иного вида. Показаны достоинства и недостатки отдельных критериев. Расширены таблицы процентных точек.

5. Разработано программное обеспечение для проверки гипотез с использованием перечисленных критериев нормальности и показательности, для статистического моделирования распределений их статистик и вычисления оценок мощности критериев по отношению к различным конкурирующим гипотезам.

6. Разработано программное обеспечение, позволяющее применять критерии согласия в тех случаях, когда распределения статистик, соответствующие справедливости проверяемой гипотезе, неизвестны и в каждом конкретном случае находятся с использованием интерактивного моделирования и распределенных вычислений в ходе проверки гипотезы.

7. Разработанное программное обеспечение для моделирования, проверки статистических гипотез и исследования статистических закономерностей, построенные таблицы процентных точек встроены в программную систему «Интервальная статистика» ISW, развиваемую на кафедре прикладной математики. Разработанное программное обеспечение используется при проведении научных исследований и в учебном процессе факультета прикладной математики и информатики при проведении лабораторных работ по курсу «Компьютерные технологии анализа данных и исследования статистических закономерностей» для студентов, обучающихся по направлению подготовки 010400 — прикладная математика и информатика, что подтверждается актом о внедрении.

Разработанное программное обеспечение зарегистрировано в Федеральной службе по интеллектуальной собственности, свидетельство о государственной регистрации программы для ЭВМ № 2012613664.

СПИСОК ПУБЛИКАЦИЙ

1. Рогожников А.П., Лемешко Б.Ю. Исследование методами статистического моделирования свойств некоторых критериев // Материалы всерос. научной конф. молодых ученых «Наука. Технологии. Инновации». Новосибирск. 2007. Т. 1. С. 86–88.
2. Рогожников А.П., Лемешко Б.Ю. Исследование свойств некоторых критериев проверки нормальности // Материалы всерос. научной конф. молодых ученых «Наука. Технологии. Инновации». Новосибирск. 2008. Т. 1. С. 31–33.
3. Рогожников А.П., Лемешко Б.Ю. Исследование методами статистического моделирования свойств некоторых критериев нормальности // Материалы IX междунар. конф. «Актуальные проблемы электронного приборостроения» АПЭП-2008. Новосибирск. 2008. Т. 6. С. 61–64.
4. Лемешко Б.Ю., Рогожников А.П. Исследование особенностей и мощности некоторых критериев нормальности // Метрология. 2009. № 4. С. 3–24.
5. Lemeshko B.Y., Rogozhnikov A.P. Simulation in Comparative Analysis of Several Tests for Normality // MMR 2009 - Mathematical Methods in Reliability. Theory. Methods. Applications. VI International Conference. Extended Abstracts. Moscow. 2009. pp. 403–407. [Моделирование в сравнительном анализе нескольких критериев нормальности]
6. Chimitova E.V., Lemeshko S.B., Lemeshko B.Y., Postovalov S.N., and Rogozhnikov A.P. Distributed computing system for simulation of classical test statistic distributions under nonstandard conditions // Proceedings Third International Conference on Accelerated Life Testing, Reliability-based Analysis and Design. Clermont-Ferrand. 2010. pp. 107–109. [Распределенная вычислительная система для моделирования распределений статистик классических критериев в нестандартных условиях]
7. Рогожников А.П., Лемешко Б.Ю. Исследование критериев отклонения эмпирического распределения вероятностей от нормального закона методами статистического моделирования // Материалы Российской НТК «Информатика и проблемы телекоммуникаций». Новосибирск. 2011. Т. 1. С. 109–112.
8. Lemeshko B.Y., Lemeshko S.B., and Rogozhnikov A.P. Real-Time Studying of Statistic Distributions of Non-Parametric Goodness-of-Fit Tests when Testing Complex Hypotheses // Proceedings of the International Workshop «Applied Methods of Statistical Analysis. Simulations and Statistical Inference» — AMSA’2011. Novosibirsk. 2011. pp. 19–27. [Изучение распределений статистик непараметрических критериев согласия в реальном времени при проверке сложных гипотез]
9. Lemeshko B.Y., Lemeshko S.B., Chimitova E.V., Postovalov S.N., and Rogozhnikov A.P. Software System for Simulation and Research of Probabilistic

Regularities and Statistical Data Analysis in Reliability and Quality Control // In: Mathematical and Statistical Models and Methods in Reliability. Applications to Medicine, Finance, and Quality Control / Ed. by Rykov V., Balakrishnan N., and Nikulin M. Boston: Birkhäuser, 2011. pp. 417–432. [Программная система для моделирования и исследования вероятностных закономерностей и статистического анализа данных в теории надежности и контроле качества]

10. Лемешко Б.Ю., Рогожников А.П. Исследование мощности критерия показательности Большева // Сборник научных трудов НГТУ. 2012. № 1(67). С. 107–114.
11. Рогожников А.П., Лемешко Б.Ю. Исследование критериев отклонения эмпирического распределения вероятностей от нормального закона методами статистического моделирования // Материалы Российской научно-технической конференции «Обработка информационных сигналов и математическое моделирование». Новосибирск. 2012. С. 66–69.
12. Лемешко Б.Ю., Рогожников А.П. О нормальности погрешностей измерений в классических экспериментах и мощности критериев, применяемых для проверки отклонения от нормального закона // Метрология. 2012. № 5. С. 3–26.

13 Рогожников А.П., Лемешко Б.Ю. Обзор критериев показательности // Материалы XI международной конференции «Актуальные проблемы электронного приборостроения» АПЭП-2012. Новосибирск. 2012. Т.6. С.47–55.

14 Лемешко Б.Ю., Постовалов С.Н., Лемешко С.Б., Чимитова Е.В., Рогожников А.П., Щеглов А.Е., Горбунова А.А. Статистистический анализ интервальных наблюдений одномерных непрерывных случайных величин “Интервальная статистика 5.0”. Свидетельство о государственной регистрации программы для ЭВМ № 2012613664 от 19 апреля 2012 г.

Подписано в печать 09.11.2012 г. Формат 60 × 84 × 1/16
Бумага офсетная. Тираж 100 экз. Печ. л. 1.5.
Заказ № 1545

Отпечатано в типографии
Новосибирского государственного технического университета
630092, г. Новосибирск, пр-т К. Маркса, 20