На правах рукописи

Постовалов Сергей Николаевич

ПРИМЕНЕНИЕ КОМПЬЮТЕРНОГО МОДЕЛИРОВАНИЯ ДЛЯ РАСШИРЕНИЯ ПРИКЛАДНЫХ ВОЗМОЖНОСТЕЙ КЛАССИЧЕСКИХ МЕТОДОВ ПРОВЕРКИ СТАТИСТИЧЕСКИХ ГИПОТЕЗ

05.13.17 – «Теоретические основы информатики»

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
доктора технических наук

Новосибирск - 2014

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

Научный консультант:	доктор технических наук, профессор Лемешко Борис Юрьевич
Официальные оппоненты:	Антонов Александр Владимирович, доктор технических наук, профессор, Обнинский институт атомной энергетики – филиал федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский ядерный университет «МИФИ», декан факультета кибернетики
	Рябко Борис Яковлевич, доктор технических наук, профессор, Федеральное государственное образовательное бюджетное учреждение высшего профессионального образования «Сибирский государственный университет телекоммуникаций и информатики», ректор
	Тарасенко Феликс Петрович, доктор технических наук, профессор, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Национальный исследовательский Томский государственный университет», профессор кафедры теоретической кибернетики
Ведущая организация:	Федеральное государственное бюджетное учреждение науки «Институт вычислительной математики и математической геофизики Сибирского отделения Российской академии наук».

Защита состоится «05» июня 2014 г. в 14^{00} часов на заседании диссертационного совета Д 212.173.06 при Федеральном государственном бюджетном образовательном учреждении высшего профессонального образования «Новосибирский государственный технический университет» по адресу: 630073, Новосибирск, пр. К.Маркса, 20.

С диссертацией можно ознакомиться в библиотеке Новосибирского государственного технического университета.

Автореферат разослан «___» марта 2014 г.

Ученый секретарь

диссертационного совета Чубич Владимир Михайлович

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Использование методов статистического анализа на практике всегда связано с применением вычислительных и, как правило, сложных алгоритмов. Однако численные методы, к сожалению, редко используются для получения новых фундаментальных знаний в самой математической статистике.

Методика компьютерного моделирования, основанная на методе Монте-Карло, позволяет более эффективно решать классические задачи статистического анализа. Данная методика дополняет аналитические методы, обеспечивая нахождение приближенного решения в тех случаях, когда этого не удается сделать аналитическими методами. Численное моделирование на компьютере дает наиболее реальный, надежный и относительно простой аппарат для исследования законов распределений различных статистик, для исследования их изменчивости в зависимости от различных факторов. На основании результатов моделирования можно прослеживать изменения закономерностей с ростом объемов выборок и изменением размерности данных. Методика позволяет на основе результатов имитационного моделирования строить модели распределений любой исследуемой статистики в конкретной ситуации.

Появление метода Монте-Карло совпало по времени с появлением первых электронных вычислительных машин (ЭВМ). Именно рост мощности современных ЭВМ сделал возможным применение компьютерного моделирования не только для исследования фундаментальных закономерностей, но и для исследования в интерактивном режиме (в ходе проводимого статистического анализа) закономерностей, имеющих место в реальных (нестандартных) условиях приложений, с последующим использованием полученных результатов (вместо асимптотических, часто существенно отличающихся от имеющих место) в процессе принятия решения.

Степень разработанности. Метод Монте-Карло предложен в 1945 году в процессе работы группы американских физиков и математиков (Дж. фон Нейман, С. Улам, Н. Метрополис, Г.Кан, Э. Ферми и др.) над созданием атомного реактора. Значительный вклад в развитие метода Монте-Карло внесли С. М. Ермаков, Г. А. Михайлов, И. М. Соболь, G. S. Fishman, C. P. Robert, G. Casella.

Обширные исследования статистических критериев с помощью компьютерного моделирования проводились Б.Ю. Лемешко, С.Б. Лемешко, Е.В. Чимитовой, С.С. Помадиным, В.М. Волковой, А.П. Рогожниковым.

Объект исследования. Объектом исследования диссертационной работы являются критерии проверки статистических гипотез. Проверка статистических гипотез является одной из важнейших задач как математической, так и прикладной статистики. При проверке статистических гипотез на практике любой исследователь сталкивается со следующими вопросами.

Во-первых, как выбрать статистический критерий, и какой критерий наиболее предпочтителен? Во-вторых, насколько точными (корректными) являются статистические выводы при проверке гипотезы по применяемому критерию? В-третьих, как уменьшить затраты на проведение экспериментов, необходимых для проверки статистической гипотезы?

Традиционно при ответе на эти вопросы и при решении соответствующих проблем использовались аналитические методы. В то же время для разрешения множества проблем с успехом можно применять компьютерное моделирование.

Цели и задачи. Основной целью диссертации является *развитие аппарата прикладной математической статистики, предназначенного для решения задач проверки статистических гипотез, за счет интенсивного использования методов компьютерного моделирования для исследования вероятностных и статистических закономерностей*.

Для достижения этой цели решаются следующие задачи.

1. Построение более точных аппроксимаций законов распределений статистик критериев при конечных объемах выборок.
2. Сравнительный анализ мощности критериев и решение задачи выбора наиболее предпочтительного критерия при разных конкурирующих гипотезах.
3. Построение точных критических границ в последовательных критериях проверки статистических гипотез.
4. Оптимальное планирование эксперимента для различения двух статистических гипотез с заданными вероятностями ошибок I и II рода.

При этом в диссертации рассматриваются, главным образом, гипотезы о виде распределения (простые и сложные гипотезы, с оцениванием параметров законов распределений) и гипотезы однородности распределений.

Научная новизна диссертационной работы заключается:

- в выявлении зависимости мощности непараметрических критериев согласия при проверке сложных гипотез от метода оценивания параметров; в сравнительном анализе мощности критериев Колмогорова, Крамера-Мизеса-Смирнова, Андерсона-Дарлинга;

- в построении точных критических границ последовательных критериев Вальда, Айвазяна и Лордена;
- в результатах сравнительного анализа мощности критерия однородности Андерсона-Дарлинга-Петита с другими критериями однородности относительно ряда близких альтернатив;
- в результатах сравнительного анализа мощности инвариантных критериев многомерной нормальности для ряда близких альтернатив;
- в выявлении зависимости оптимального объема выборки и стоимости проведения эксперимента от симметричной дивергенции Кульбака-Лейблера между распределениями в выборке случаев и в контрольной выборке при проведении полногеномного анализа ассоциаций по критерию MAX3;
- в оценке относительной эффективности критерия MAX3 по сравнению с критерием тренда Кокрена-Армитеджа при оптимальном наборе коэффициентов.

Теоретическая и практическая значимость работы. В диссертационной работе численные методы и статистическое моделирование направлены на изучение закономерностей самой математической статистики, на уточнение условий, в которых корректно применение конкретных теоретических результатов математической статистики, на исследование постановок, появившихся в последнее время в связи с потребностями практики. Теоретическая значимость работы заключается в том, что полученные результаты развивают аппарат прикладной математической статистики.

Практическая значимость заключается в расширении сферы корректного применения ряда статистических критериев в приложениях, в повышении точности статистических выводов при проверке статистических гипотез, в случае применения последовательных критериев (за счет использования более точных критических границ) в сокращении средних объемов выборок, требуемых для принятия решения (следовательно, в сокращении стоимости проведения экспериментов).

Результаты исследований и средства моделирования включены в программные системы «Интервальная статистика» ISW и конфигурации «НКЦ ИТР:Статистика 1.0», разработанной для платформы «1С:Предприятие 8.2».

На основе результатов исследований свойств критериев согласия при проверке простых и сложных гипотез разработаны рекомендации по стандартизации Госстандарта РФ по правилам применения критериев согласия:

- Р 50.1.033-2001 (Рекомендации по стандартизации. Прикладная статистика.

Правила проверки согласия опытного распределения с теоретическим. Часть I. Критерии типа хи-квадрат. - М.: Изд-во стандартов. – 2002. - 87 с.);

- Р 50.1.037-2002 (Рекомендации по стандартизации. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Часть II. Непараметрические критерии. - М.: Изд-во стандартов. 2002. - 64 с.).

Методология и методы исследования. Для решения поставленных задач использовался аппарат теории вероятностей, математической статистики, статистического моделирования, математического программирования, теории принятия решений в условиях неопределенности.

Положения, выносимые на защиту:

1. Алгоритм определения точных критических границ для последовательных критериев Вальда, Айвазяна и Лордена;
2. Результаты исследования распределений статистик и сравнительного анализа мощности критериев согласия Колмогорова, Крамера-Мизеса-Смирнова и Андерсона-Дарлинга при проверке сложных гипотез о виде распределения и различных методах оценивания параметров.
3. Результаты исследования распределений статистики и мощности критерия однородности Андерсона-Дарлинга-Петита, сравнительного анализа мощности с другими критериями однородности.
4. Результаты исследования распределений статистик и сравнительного анализа мощности критериев многомерной нормальности.
5. Результаты исследования распределений статистик и сравнительного анализа мощности критериев ассоциаций.
6. Алгоритмы оптимального планирования экспериментов при проведении полногеномного анализа ассоциаций.

Соответствие диссертации паспорту научной специальности. Содержание диссертации соответствует п.5 «Разработка и исследование моделей и алгоритмов анализа данных, обнаружения закономерностей в данных и их извлечениях, разработка и исследование методов и алгоритмов анализа текстов, устной речи и изображений» паспорта специальности 05.13.17 – «Теоретические основы информатики» (в области технических наук).

Степень достоверности и апробация результатов. Результаты исследований докладывались на Российской научно-технической конференции «Информатика и проблемы телекоммуникаций" (Новосибирск, 1996, 2004, 2005, 2006, 2010, 2011); Международной научно-технической конференции «Информатика и проблемы телекоммуникаций» (Новосибирск, 1995, 1997, 1998, 1999, 2001, 2002); Российской научно-технической конференции "Обработка информационных сигналов и математическое моделирование" (Новосибирск, 2012, 2013); Международном симпозиуме по непараметрическим и робастным методам в кибернетике (Красноярск, 1995, Железногорск, 1997); Международных конференциях «Актуальные проблемы электронного приборостроения" (АПЭП) (Новосибирск, 1996, 1998, 2000, 2004, 2006, 2008, 2010, 2012); Международной научно-технической конференции «Микропроцессорные системы автоматики" (Новосибирск, 1996); Сибирском Конгрессе по Прикладной и Индустриальной Математике (ИНПРИМ) (Новосибирск, 1996); Межреспубликанском совещании по интервальной математике (Новосибирск, 1996); Международной конференции «Информационные технологии в моделировании и управлении» (Санкт-Петербург, 1996); Международной научно-методической конференции «Новые информационные технологии в университетском образовании» (Новосибирск, 1997); Международной научной конференции «Всесибирские чтения по математике и механике" (Томск, 1997); Международном совещании по интервальной математике (Красноярск, 1997); Международной конференции “Korea-Russia International Symposium of Science and Technology” (KORUS) (Ulsan, 1997, 2003; Novosibirsk, 1999; Tomsk, 2004); Международной конференции “Computer Data Analysis and Modeling: Robustness and Computer Intensive Methods” (CDAM) (Минск, 2004); Международной конференции “Mathematical Methods in Reliability. Theory. Methods. Applications” (MMR) (Москва, 2009); Международной конференции «Accelerated Life Testing, Reliability-based Analysis and Design» (ALT) (Clermont-Ferrand, France, 2010); Международной научно-практической конференции "Новые информационные технологии в образовании" (Москва, 2011, 2013); Международной конференции «Applied Stochastic Models and Data Analysis» (ASMDA) (Крит, Греция, 2007, Рим, 2011); Международной конференции “Applied Methods of Statistical Analysis. Simulations and Statistical Inference” (AMSA) (Новосибирск, 2011, 2013), Всероссийской конференции по вычислительной математике (КВМ) (Новосибирск, 2011), на семинаре "Многомерный статистический анализ и вероятностное моделирование реальных процессов" в ЦЭМИ РАН (Москва, 2012).

Работа выполнена при поддержке федеральной целевой научно-технической программы «Исследования и разработки по приоритетным направлениям развития науки и техники» на 2002-2006 годы (проекты № РИ-19.0/002/091, 2006-РИ-19.0/001/119), федеральной целевой программы «Научные и научно-педагогические кадры инновационной России 2009-2013 гг» (проекты НК-

421П, НК-15П/15, ГК № 02.740.11.5187, соглашение № 14.B37.21.0860), аналитической ведомственной целевой программы «Развитие научного потенциала высшей школы» (проект № 2.1.2/11855), грантов РФФИ (№ № 00-01-00913а, 06-01-00059а, 09-01-00056а), Министерства образования и науки РФ в рамках госзадания (проект 8.1274.2011). Результаты главы 7 получены во время научной стажировки в институте медицинской биометрии и статистики (г. Любек, Германия) при поддержке DAAD (грант A/11/76161).

Публикации. Основные результаты исследований по теме диссертации опубликованы в 106 печатных работах общим объемом 113 п.л., в том числе монография [1], 23 статьи в рецензируемых научных журналах (из них 19 статей в журналах, входящих в перечень изданий, рекомендуемых ВАК РФ) [2-24], учебные пособия с главами научного содержания [25-27], 78 публикаций в сборниках научных работ, трудах и материалах научных конференций. Получено пять свидетельств о государственной регистрации программ для ЭВМ [35-39].

Структура и объем работы. Диссертация состоит из введения, 8 глав основного содержания, заключения, списка литературы и приложений. Общий объем работы составляет 285 страниц, включая 58 таблиц, 82 рисунка и список литературы из 281 источника.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

В первой главе рассмотрена методика компьютерного моделирования в применении к задачам проверки статистических гипотез. Сформулированы алгоритмы моделирования достигаемого уровня значимости и мощности статистического критерия. Рассмотрены вопросы определения необходимого числа повторений метода Монте-Карло для получения заданной погрешности. Предложена методика определения скорости сходимости распределения статистики к предельному закону.

Компьютерное моделирование уже много лет успешно применяется при изучении сложных объектов во многих областях науки и техники. Одним из методов компьютерного моделирования является *метод Монте-Карло* (или *метод статистических испытаний*), который базируется на получении большого числа реализаций стохастического (случайного) процесса.

С помощью метода Монте-Карло решается множество задач в математике, физике, химии, экономике, социологии и других науках. В настоящей работе метод Монте-Карло применяется для решения теоретических задач математической статистики, в частности в применении к задачам проверки статистических гипотез.

Статистической гипотезой называется любое утверждение о виде или свойствах распределения наблюдаемых в эксперименте случайных величин. Обычно статистические гипотезы делят на следующие виды: *однородности*, если имеется две или более выборок случайных величин; *независимости*, если имеется выборка многомерной случайной величины; *случайности*, если есть предположения о наличии в последовательности наблюдений систематических изменений; *о виде распределения*, если есть предположения о законе распределения случайной величины.

Гипотезу, которую мы проверяем, будем называть *основной* или *нулевой* гипотезой, и будем всегда обозначать H_0. Альтернативные или *конкурирующие* гипотезы будем обозначать $H_1,\ H_2,\ \ldots,\ H_m$.

Для проверки одной и той же гипотезы, как правило, существует несколько различных статистических критериев $T_1, T_2, ..., T_k$. Выбор подходящего статистического критерия, вообще говоря, не является тривиальной задачей. Можно сформулировать следующие принципы подбора статистических критериев.

1. Должны выполняться “стандартные” предположения, обуславливающие возможность применения рассматриваемого критерия (например, о виде распределения случайной величины и о наблюдаемых данных).
2. Критерий должен быть *состоятельным*, т.е. его мощность должна стремиться к единице с ростом объема выборки.
3. Критерий должен быть *несмещенным*, т.е. мощность должна быть больше, чем вероятность ошибки первого рода.
4. Критерий должен обладать наибольшей мощностью при заданном объеме выборки и заданном уровне значимости критерия.

Реализация последнего принципа на практике, как правило, оказывается невозможной, потому что построить наиболее мощный критерий удается только в очень редких случаях, например, когда основная и конкурирующая гипотезы являются *простыми*. Чаще всего, для разных конкурирующих гипотез, для разных уровней значимости, для разных объемов выборки, более мощными оказываются различные критерии.

В такой ситуации для выбора оптимального критерия можно применить классическую теорию *принятия решений в условиях неопределенности*. В этом случае критерии отождествляются со стратегиями, конкурирующие гипотезы –

с состояниями среды, функция полезности $u\left(T_i, H_j\right)$ – это мощность критерия. Другим способом задания функции полезности при выборе критерия может быть стоимость проведения эксперимента по различению основной и конкурирующей гипотез с заданными вероятностями ошибок I и II рода.

Существуют разные подходы к выбору оптимальной стратегии при принятии решения в условиях неопределенности. Когда нет никакой информации о том, какая конкурирующая гипотеза может быть верна, рациональным выглядит выбор критерия по правилу Вальда:

$$T^* = \arg\max_{T_i} \min_{H_j} u\left(T_i, H_j\right).$$

Критерий, выбранный по правилу Вальда, максимизирует полезность против самой «неудобной» конкурирующей гипотезы.

Вычисление достигаемого уровня значимости (p-value) относительно просто, когда известно теоретическое распределение статистики критерия при справедливости основной гипотезы. Однако возможны ситуации, когда:

- неизвестен аналитический вид закона распределения статистики;
- известен только асимптотический закон распределения статистики;
- закон распределения статистики меняется от объема выборки, от метода оценивания параметров, от доли цензурирования и т.п.

В таких ситуациях для вычисления p-value можно эффективно применить метод Монте-Карло (алгоритм 1.1).

Алгоритм 1.1. Вычисление достигаемого уровня значимости статистического критерия методом Монте-Карло.

Входные данные: гипотеза H_0, выборка X_n, количество повторений N, функция вычисления статистики критерия $S\left(X_n\right)$.

Действия.

1. Вычислить $S = S\left(X_n\right)$ – статистику критерия по выборке.
2. Установить $m = 0$.
3. Сгенерировать выборку Y_n при верной гипотезе H_0.
4. Вычислить значение $S\left(Y_n\right)$.
5. Если $S(Y_n) > S(X_n)$, то $m = m + 1$.
6. Повторять шаги 3-5 N раз.

Выходные данные: оценка достигаемого уровня значимости (p-value) равна

- $\hat{p} = \frac{m}{N}$ для правосторонней критической области;
- $\hat{p} = 1 - \frac{m}{N}$ для левосторонней критической области;
- $\hat{p} = 2\min\left(\frac{m}{N}, 1 - \frac{m}{N}\right)$ для двусторонней критической области.

Для вычисления мощности статистического критерия необходимо знать распределения статистики критерия при основной и конкурирующей гипотезах. Если распределение статистики критерия при верной основной гипотезе часто можно получить при достаточно общих предположениях, то распределение статистики критерия при верной конкурирующей гипотезе найти аналитическими методами очень сложно. Естественно, что это распределение будет зависеть от множества факторов: от вида конкурирующей гипотезы, от объема выборки, от метода оценивания параметров (если он используется при вычислении статистики), от процента цензурирования (для цензурированных выборок) и т.п. Поэтому основным методом проведения исследований, связанных со сравнительным анализом мощности статистических критериев, является метод Монте-Карло.

Для вычисления мощности критерия по методу Монте-Карло можно воспользоваться алгоритмом 1.2.

Алгоритм 1.2. Вычисление мощности статистического критерия методом Монте-Карло.

Входные данные: гипотеза H_1, критическая область W_α с уровнем значимости α, количество повторений *N*, функция вычисления статистики $S(Y_n)$ по выборке Y_n.

Действия.

1. Установить $m = 0$.
2. Сгенерировать выборку Y_n при верной гипотезе H_1.
3. Вычислить значение $S(Y_\text{n})$.
4. Если критическая область правосторонняя и $S(Y_n) \in W_\alpha$, то $m = m + 1$.
5. Повторять шаги 2-4 *N* раз.

Выходные данные: оценка мощности равна

$$1 - \hat{\beta} = \frac{m}{N}.$$

Оценка точности характеристик и необходимое число реализаций (повторений) в методе Монте-Карло приводится в книге Вентцель Е.С. (Вентцель Е.С. Исследование операций. М.: Советское радио, 1972. – 552 с.), причем рассматривается простейший случай, когда с помощью метода Монте-Карло оценивается вероятность наступления некоторого события. В параграфе 1.7 диссертации рассмотрена задача определения количества повторений при оценивании функции распределения, процентных точек (квантилей распределения статистики критерия) и мощности статистического критерия.

Погрешность оценивания достигаемого уровня значимости p с доверительной вероятностью γ равна

$$\varepsilon = t_\gamma \frac{\sqrt{p(1-p)}}{\sqrt{N}}, \tag{1}$$

где $t_\gamma = \Phi^{-1}\left(\frac{\gamma+1}{2}\right)$ – квантиль стандартного нормального распределения.

Отсюда количество повторений, при котором погрешность моделирования не превышает ε, равно

$$N = \left[t_\gamma^2 \frac{p(1-p)}{\varepsilon^2} \right] + 1, \tag{2}$$

где $[\cdot]$ означает целую часть.

В качестве оценки процентной точки порядка p используется выборочная квантиль, вычисляемая по формуле

$$t_\alpha = \begin{cases} \frac{1}{2}\left(S_{[qN]} + S_{[qN]+1}\right), & qN \in \mathbb{N}; \\ S_{[qN]+1}, & qN \notin \mathbb{N}. \end{cases} \tag{3}$$

Известно, что данная статистика в асимптотике имеет нормальное распределение с математическим ожиданием и дисперсией

$$M\left[x_{([pN])}\right] = F^{-1}(p), \; D\left[x_{([pN])}\right] = \frac{p(1-p)}{N \cdot f^2(F^{-1}(p))},$$

где $f(x)$ – функция плотности. Тогда формулы (1) и (2) для оценивания значения $F^{-1}(p)$ с помощью выборочной квантили $x_{([pN])}$ примут вид

$$\varepsilon = t_\gamma \frac{\sqrt{p(1-p)}}{f(F^{-1}(p))\sqrt{N}}, \quad N = \left[t_\gamma^2 \frac{p(1-p)}{\varepsilon^2 f^2(F^{-1}(p))} \right] + 1. \tag{4}$$

На оценку мощности методом Монте-Карло накладываются две погрешности – первая при определении процентной точки t_α и вторая – при вычислении вероятности ошибки второго рода. Оценка погрешностей в этом случае оказывается более сложной задачей, чем оценка погрешностей при вычислении процентных точек. С помощью метода Монте-Карло на ряде примеров было показано, что распределение погрешности подчиняется нормальному закону, параметры которого зависят от распределений статистики, поэтому для оценивания мощности предложено выполнять моделирование сериями с последующим вычислением доверительного интервала.

В первой главе предложен также алгоритм оценивиания скорости сходимости распределения статистики к предельному закону.

Во второй главе рассмотрены критерии различения двух простых гипотез, как с фиксированным объемом выборки, так и для случая, когда объем определяется в процессе выполнения наблюдений. Рассмотрена симметричная дивергенция Кульбака-Лейблера и ее связь с процедурой оптимального планирования эксперимента по различению двух простых гипотез. Предложен численный алгоритм для определения оптимального объема выборки для критерия отношения правдоподобия.

Различение двух простых гипотез, когда и основная, и конкурирующая гипотезы полностью определены, является самым простым случаем проверки гипотез. Фундаментальный результат в этой ситуации был получен в 1933 году Нейманом и Пирсоном, которые установили, что наиболее мощным критерием при различении двух простых гипотез является критерий отношения правдоподобия.

Позднее Вальдом был предложен последовательный критерий отношения правдоподобия (SPRT), в котором объем выборки не фиксируется заранее, а определяется в процессе наблюдения. Вальдом и Вольфовицем в 1948 году было доказано, что SPRT имеет минимальный средний объем выборки среди всех критериев при заданных вероятностях ошибок первого и второго рода. Критерий с фиксированным объемом выборки можно рассматривать как частный случай последовательного критерия, поэтому свойство оптимальности SPRT распространяется как на последовательные критерии, так и на критерии с фиксированным объемом выборки.

Вальд также нашел нижнюю границу среднего объема выборки любого последовательного критерия, как функцию от вероятностей ошибок первого и второго рода, деленную на меру различия функций распределения основной и конкурирующей гипотез, позднее названную *дивергенцией Кульбака-Лейблера*. Дивергенция Кульбака-Лейблера есть средняя информация, которую несет одно наблюдение для различения основной гипотезы от конкурирующей.

Экономичность последовательного критерия отношения правдоподобия по сравнению с критерием принятия решения по выборке фиксированного объема, т.е. выигрыш в среднем числе наблюдений, была оценена А. Вальдом и уточнена С.А. Айвазяном.

Несмотря на то, что последовательная процедура различения двух простых гипотез является самой «экономичной» по числу наблюдений, условия проведения эксперимента не всегда позволяют выполнять последовательную процедуру проверки гипотезы. В связи с этим возникает задача *оптимального планирования статистического эксперимента*, которая заключается в выборе критической области и объема выборки таким образом, чтобы вероятности ошибок первого и второго рода были меньше или равны заданным значениям.

Значительный вклад в определение необходимого объема выборки для разных статистических критериев внес Н.И. Володин.

Пусть относительно вида распределения случайной величины ξ сделаны два предположения: гипотеза H_0 о том, что распределение имеет плотность $f_0(x)$, и гипотеза H_1 о том, что распределение имеет плотность $f_1(x)$. Если случайная величина дискретная, то $f(x)$ обозначает вероятность события $\xi = x$. Так как гипотезы полностью определены, то это простые гипотезы.

Для различения двух гипотез H_0 и H_1 проводятся наблюдения случайной величины ξ: $x_1, x_2, \ldots, x_n$. Момент прекращения наблюдения (n) может быть выбран заранее, тогда это процедура проверки статистической гипотезы с фиксированным объемом выборки, или определяться в процессе наблюдения – при последовательной процедуре проверки гипотезы.

Отношение $z_i = \ln \dfrac{f_1(x_i)}{f_0(x_i)}$ называется *информацией о различии* между гипотезами H_0 и H_1, которую несет одно наблюдение x_i. Среднее значение информации о различии между H_0 и H_1 называется *дивергенцией Кульбака-Лейблера*

$$D(H_0, H_1) = -E_0 z_i = \int_R f_0(x) \ln \frac{f_0(x)}{f_1(x)} dx, \quad D(H_1, H_0) = E_1 z_i = \int_R f_1(x) \ln \frac{f_1(x)}{f_0(x)} dx. \quad (5)$$

Симметричная дивергенция Кульбака-Лейблера определяется выражением

$$\rho(H_0,H_1)=D(H_0,H_1)+D(H_1,H_0)=\int_R \left(f_1(x)-f_0(x)\right) f_0(x)\ln\frac{f_1(x)}{f_0(x)}dx\,. \quad (6)$$

Критерий отношения правдоподобия строится на основании *статистики отношения правдоподобия*

$$\Lambda_n=\prod_{i=1}^{n}\frac{f_1(x_i)}{f_0(x_i)}, \quad (7)$$

или ее логарифма

$$\lambda_n=\ln\Lambda_n=\ln\prod_{i=1}^{n}\frac{f_1(x_i)}{f_0(x_i)}=\sum_{i=1}^{n}\ln\frac{f_1(x_i)}{f_0(x_i)}=\sum_{i=1}^{n}z_i\,. \quad (8)$$

С точки зрения теории информации статистика отношения правдоподобия λ_n равна информации, содержащейся во всей выборке, о различии между гипотезами H_0 и H_1.

Гипотеза H_0 отвергается по критерию отношения правдоподобия, если $\lambda_n>t_\alpha$, где критическое значение t_α определяется из условия $P\left\{\lambda_n>t_\alpha \middle| H_0\right\}=\alpha$.

Все рассмотренные способы нахождения необходимого объема выборки – являются асимптотическими – требуется либо стремление вероятностей ошибок первого и второго рода к нулю, либо сближение основной и конкурирующей гипотез. В то же время, можно воспользоваться методом Монте-Карло для определения необходимого объема выборки, используя следующее свойство логарифма статистики отношения правдоподобия

$$\lambda_n=\sum_{i=1}^{n}\ln\frac{f_1(x_i)}{f_0(x_i)}=\sum_{i=1}^{n-1}\ln\frac{f_1(x_i)}{f_0(x_i)}+\ln\frac{f_1(x_n)}{f_0(x_n)}=\lambda_{n-1}+z_n\,. \quad (9)$$

Формула (9) означает, что очередное значение λ_n можно получить, зная предыдущее λ_{n-1}. В результате получается алгоритм 2.1, в котором λ_{ij}^0 и λ_{ij}^1- j-я реализация статистики отношения правдопоподбия по выборке объема i, при верных гипотезах H_0 и H_1, соответственно.

Алгоритм 2.1. Определение необходимого объема выборки для критерия отношения правдоподобия по точечным наблюдениям.

Входные данные: гипотезы H_0 и H_1, вероятности ошибок первого и второго рода α и β, количество повторений N.

Действия.

1. Задать начальные значения: $i=1, \lambda_{0j}^0=0$ и $\lambda_{0j}^1=0$, $j=\overline{1,N}$. Вычислить ε_N по формуле (1), где $p=\beta$, $\gamma=0.99$.

2. Смоделировать случайные величины x_j^0 и x_j^1, подчиняющиеся законам распределения в соответствии с гипотезами H_0 и H_1, $j=\overline{1,N}$.

3. Вычислить $\lambda_{ij}^0 = \lambda_{(i-1)j}^0 + \ln\dfrac{f_1(x_j^0)}{f_0(x_j^0)}$ и $\lambda_{ij}^1 = \lambda_{(i-1)j}^1 + \ln\dfrac{f_1(x_j^1)}{f_0(x_j^1)}$, $j=\overline{1,N}$

4. Вычислить $\hat{t}_\alpha$ – критическое значение статистики критерия при уровне значимости α по выборке $\left\{\lambda_{ij}^0,\ j=\overline{1,N}\right\}$ по формуле (3), где $q = 1-\alpha$.

5. Вычислить оценку вероятности ошибки второго рода $\hat{\beta}_i$, как количество λ_{ij}^1, не превосходящих $\hat{t}_\alpha$, деленное на *N*.

6. Если $\hat{\beta}_i > \beta - \varepsilon_N$, то увеличить индекс $i := i+1$ и перейти на шаг 2.

Выходные данные: оценка необходимого объема выборки $n=i$.

Пример использования алгоритма 2.1 приведен в таблице 1 в сравнении с различными теоретическими оценками необходимого объема выборки для проверки гипотезы о нормальном распределении против логистичесого распределения с заданными вероятностями ошибок первого и второго рода. Результаты моделирования показывают, что необходимый объем выборки больше, чем асимптотическая оценка при сближении гипотез, полученная Айвазяном, но меньше, чем асимптотическая оценка, полученная при использовании центральной предельной теоремы.

Таблица 1 – Необходимый объем выборки для различения нормального и логистического распределений

α	β	Оценка снизу по Вальду	Оценка снизу	Оценка Айвазяна	Асимптотическая оценка	Алг. 2.1 при $N=1660000$
0,05	0,05	253	278	436	548	472
0,01	0,01	429	555	871	1095	935
10^{-3}	10^{-3}	657	978	1537	1932	1640

В параграфах 2.7-2.8 рассмотрена задача оптимального группирования при различении двух простых гипотез по критерию отношения правдоподобия. Очевидно, что группирование наблюдений приводит к потере информации, и эти потери зависят от выбора способа группирования данных. На практике обычно строят интервалы равной длины или, в лучшем случае, интервалы равной вероятности. Однако еще в 1951 Кульбак и Лейблер предлагали выполнять группирование, минимизируя потери информации о различии между законами, соответствующими основной и конкурирующей гипотезам.

В общем виде задача оптимального группирования имеет следующий вид. Пусть область определения случайной величины ξ разбита на k интервалов граничными точками $t_1 < t_2 < ... < t_{k-1}$. Нижняя граница области определения равна t_0, верхняя граница равна t_k. Требуется найти такие точки $t_1 < t_2 < ... < t_{k-1}$, при которых достигается максимум дивергенции Кульбака-Лейблера по группированной выборке:

$$\max_{t_1<t_2<...<t_{k-1}} \sum_{i=1}^{k} P_{0i} \ln \frac{P_{0i}}{P_{1i}}, \tag{10}$$

где $P_{0i} = \int\limits_{t_{i-1}}^{t_i} f_0(x)dx$ и $P_{1i} = \int\limits_{t_{i-1}}^{t_i} f_1(x)dx$ - вероятности попадания в интервал при основной и конкурирующей гипотезах соответственно.

Существует взаимосвязь между дивергенцией Кульбака-Лейблера и информационной матрицей Фишера (Кульбак С. Теория информации и статистика. М.: Наука, 1967. - 408 с). Пусть $f_0(x) = f(x,\theta)$, $f_1(x) = f(x,\theta+\Delta\theta)$. Тогда

$$D(H_0,H_1) \approx \frac{1}{2} I(\theta)(\Delta\theta)^2, \tag{12}$$

где $I(\theta) = E\left[\frac{\partial \ln f(x,\theta)}{\partial\theta}\right]^2$ - информационное количество Фишера. Таким образом, если конкурирующие гипотезы принадлежат одному параметрическому семейству, то максимизация дивергенции Кульбака-Лейблера эквивалентна максимизации информационного количества Фишера.

Автором покзано, что оптимальное группирование (10) максимизирует мощность критерия χ^2, когда основная и конкурирующая гипотеза принадлежат различным семействам распределений.

В третьей главе рассмотрены последовательные критерии различения двух гипотез о виде распределения Вальда, Айвазяна и Лордена. Предложен алгоритм для оценивания точных критических границ с помощью компьютерного моделирования. Проведено численное сравнение среднего объема выборки для последовательных критериев с использованием приближенных границ и оценок точных границ.

В последовательном критерии Вальда используется статистика критерия отношения правдоподобия (8), где объем выборки увеличивается (процедура проверки продолжается) до момента наступления одного из событий: $\lambda_n > c_1$, тогда принимается гипотеза H_1, или $\lambda_n < c_0$, тогда принимается гипотеза H_0. При этом границы c_0, c_1 определяются по приближенным формулам.

Последовательный критерий отношения правдоподобия является оптимальным по величине среднего объема выборки, требуемого для различения двух простых гипотез. Однако его оптимальность нарушается, если ни одна из заданных гипотез не верна. В этом случае средний объем выборки увеличивается, вплоть до того, что процедура различения двух простых гипотез может стать бесконечной.

Для решения данной проблемы, получившей название «проблема Кифера-Вейсса», были предложены различные процедуры усечения последовательного критерия отношения правдоподобия. Андерсоном был сформулирован оптимальный обобщенный критерий отношения правдоподобия (GSPRT), который имеет ограниченную область неопределенности и минимальный средний объем выборки. С.А. Айвазяном в 1965 году была найдена линейная аппроксимация GSPRT в виде треугольной области неопределенности. Другое решение проблемы Кифера-Вейсса было предложено Лорденом в 1976 году в виде одновременной проверки двух пар гипотез (2SPRT).

Последовательные критерии Вальда, Айвазяна и Лордена используют приближенные критических границы, применение которых приводит к увеличению среднего объема выборки. В 1977 году Каннер с помощью метода Монте-Карло нашел оценки точных критических границ при различении двух биномиальных распределений.

Обширные исследования последовательных критериев были проведены С.Я. Гродзенским, И.С. Гродзенской и Я.С. Гродзенским. В частности, И.С. Гродзенская разработала последовательный критерий с параболическими границами, параметры которых определяются при помощи моделирования.

Следует отметить, что во всех рассмотренных работах основная и конкурирующая гипотеза принадлежат одному и тому же семейству распределений. Однако не существует никаких препятствий использовать эти критерии для распределений из разных семейств. В частности, в данной главе все результаты получены при рассмотрении двух пар простых конкурирующих гипотез: нормального распределения против логистического и распределения Вейбулла-Гнеденко против гамма-распределения.

Переход от задачи различения двух гипотез из разных семейств распределений к задаче проверки параметрической гипотезы из одного семейства решается таким же образом, как в п. 2.7 диссертации, путем рассмотрения смеси распределений $g(x,\theta)=(1-\theta)f_0(x)+\theta f_1(x)$ при основной и конкурирующей гипотезах. Тогда проверка гипотезы $H_0: f(x)=f_0(x)$ против $H_1: f(x)=f_1(x)$ эквивалентна проверке параметрической гипотезы $H_0:\theta=0$ против $H_1:\theta=1$, где θ – параметр смеси $g(x,\theta)$.

Использование приближенных границ вместо точных в последовательных критериях приводит к увеличению среднего объема выборки. Найти точные критические границы с помощью аналитических методов очень сложно вследствие того, что при вычислении вероятности ошибок первого и второго рода случайные величины λ_i и λ_j (статистики отношения правдоподобия на i-м и на j-м шаге) являются зависимыми. Однако это не является препятствием для оценивания вероятностей ошибок первого и второго рода для последовательных критериев с помощью метода Монте-Карло, как было показано автором в работе [24], для чего может быть использован алгоритм 3.1.

Алгоритм 3.1 Алгоритм оценивания вероятности ошибки первого рода для последовательных критериев методом Монте-Карло

1. Выбирается область моделирования на основании приближенных границ последовательного критерия.
2. Строится сетка с маленьким шагом на выбранной области. Для каждого узла сетки счетчик числа ошибок $m(c_0, c_1)$ устанавливается равным нулю.
3. Моделируется случайная величина в соответствии с гипотезой H_0.
4. Вычисляется статистика критерия.
5. Проверяется условие выхода для каждого узла сетки (c_0, c_1). Если в узле (c_0, c_1) произошла ошибка первого рода, т.е. была принята гипотеза H_1, то счетчик $m(c_0, c_1)$ увеличивается на единицу.
6. Если для какой-то точки сетки условие выхода не выполнено, то следует перейти на шаг 3.
7. Шаги 3-6 повторяются N раз.
8. Для каждой точки сетки оценка вероятности ошибки первого рода вычисляется по формуле:

$$\hat{\alpha}(c_0, c_1)=\frac{m(c_0, c_1)}{N},$$

где $m(c_0, c_1)$ – количество ошибок первого рода в узле (c_0, c_1).

Для вычисления вероятности ошибки второго рода используется аналогичный алгоритм, только на *шаге 3* моделирование случайной величины проводится по закону распределения, соответствующему конкурирующей гипотезе H_1, а на *шаге 5* накапливается количество ошибок второго рода, когда принимается гипотеза H_0.

В результате моделирования получаем две функции $\hat{\alpha}(c_0, c_1)$ и $\hat{\beta}(c_0, c_1)$. Для нахождения точных критических границ c_0, c_1 необходимо решить систему

$$\begin{cases} \hat{\alpha}(c_0, c_1) = \alpha; \\ \hat{\beta}(c_0, c_1) = \beta. \end{cases} \tag{12}$$

Алгоритм 3.1 и решение системы (12) позволяют оценить точные критические границы, однако у данного подхода имеется ряд недостатков. Во-первых, оценка вероятностей ошибок первого и второго рода по результатам моделирования требует достаточно много времени. Во-вторых, алгоритм 3.1 дает оценки вероятностей ошибок первого и второго рода, но не дает самих критических границ, которые приходится находить, решая систему (12).

Причина излишней трудоемкости алгоритма 3.1 заключается в избыточном количестве узлов, в которых вычисляются оценки вероятностей ошибок первого и второго рода. На самом деле область моделирования, заданная приближенными границами, слишком большая. Если требуется найти одну точку (c_0, c_1), в которой выполняются равенства (12), то можно локализовать ее окрестность при относительно небольшом числе повторений. По этому принципу построен адаптивный *алгоритм 3.2* нахождения критических границ, когда сетка сгущается в области, близкой к заданным вероятностям ошибок первого и второго рода.

Алгоритм 3.2. Алгоритм оценивания точных критических границ последовательных критериев

1. Выбирается область, которая содержит критические границы c_0, c_1 для заданных вероятностей α и β.
2. Строится сетка из ($k \times k$) узлов.
3. Вычисляются оценки $\hat{\alpha}(c_0, c_1)$ и $\hat{\beta}(c_0, c_1)$ *по алгоритму 3.1*.
4. Находится узел (i, j), для которого достигается $d = \min\limits_{i,j} \max\left\{\left|\hat{\alpha}(c_{0i}, c_{1j}) - \alpha\right|, \left|\hat{\beta}(c_{0i}, c_{1j}) - \beta\right|\right\}$

5. Если $d \le \varepsilon$ & $\max\left\{t_\gamma\sqrt{\frac{\alpha(1-\alpha)}{N_\alpha}}, t_\gamma\sqrt{\frac{\beta(1-\beta)}{N_\beta}}\right\} \le \varepsilon$, где N_α – количество повторов по методу Монте-Карло при справедливости основной гипотезы H_0, N_β – количество повторов при справедливости конкурирующей гипотезы H_1, то (c_{0i}, c_{1j}) – искомые критические границы, выход.
6. Строится новая область с центром в найденном узле (i, j). Размер области уменьшается, если найденный узел находится внутри области, иначе размер остается прежним.
7. Если погрешность моделирования больше ε, то $N_\alpha := 2 \cdot N_\alpha$, $N_\beta := 2 \cdot N_\beta$.
8. Переход на шаг 2.

Использование оценок точных критических границ для последовательных критериев проверки гипотез позволяет существенно снизить средний объем выборки наблюдений, требуемый для принятия решения о результатах проверки. На рассмотренных примерах (для вероятностей ошибок первого и второго рода 0.05) показано, что более существенный выигрыш в требуемом для принятия решения объеме выборки имеет место для критерия Лордена (до 30% если рассматриваетсяе нормальное распределение против логистического, и до 35% если рассматривается распределение Вейбулла-Гнеденко против гамма-распределения) и для критерия Айвазяна (до 31% и до 37% соответственно). В то же время, выигрыш в случае критерия Вальда составляет 11% при проверке гипотезы о нормальном распределении против логистического закона и 23% при проверке гипотезы о распределении Вейбулла-Гнеденко против гамма-распределения. Очевидно, что преимущество в требуемом объеме выборок от использования оценок точных критических границ меняется в зависимости от значений вероятностей ошибок первого и второго рода и вида проверяемых гипотез [18].

Предложенный автором алгоритм оценивания точных критических границ для различения двух простых гипотез позволяет осуществить процесс поиска за сравнительно небольшое время. Использование метода Монте-Карло для нахождения критических границ позволяет расширить круг задач, решаемых методами последовательного анализа, возможностью проверки гипотез по *цензурированным* и *группированным* данным.

Использование метода Монте-Карло для нахождения критических границ возможно также и при проверке сложных гипотез, однако в этом случае при постановке задачи о проверке гипотез должны быть заданы дополнительные ограничения на возможные значения неизвестных параметров [18].

В четвертой главе рассмотрены непараметрические критерии согласия Колмогорова, Крамера-Мизеса-Смирнова и Андерсона-Дарлинга и их обобщение для проверки сложных гипотез. Показано, что распределения статистик и мощность критериев зависят от используемого метода оценивания

В работах [9,10,12,14,15,16] представлены результаты исследований мощности ряда критериев согласия в случае проверки простых и сложных гипотез. Для двух рассматриваемых пар конкурирующих гипотез были проанализированы мощности критериев χ^2 Пирсона (при разных способах группирования и числе интервалов), Колмогорова, Смирнова, Крамера-Мизеса-Смирнова, Андерсона-Дарлинга и сделаны выводы о предпочтительности тех или иных критериев в разных условиях.

В работе [14] при проверке сложных гипотез с использованием всех исследуемых критериев согласия для оценивания неизвестных параметров применялся *метод максимального правдоподобия* (ММП). Метод оценивания параметров закона распределения влияет на законы распределения статистик непараметрических критериев согласия при проверке сложной гипотезы о виде распределения и, следовательно, влияет на мощность критериев согласия, что было показано в 2001 году [11].

В работах [13] рассматривается применение критериев согласия для проверки адекватности непараметрических моделей при использовании ядерных оценок функции плотности с ядром Епанечникова.

В работе [19] получены результаты исследования мощности непараметрических критериев согласия при тех же альтернативах, что и в работе [14], но с применением методов оценивания, отличных от ММП, а именно: метода минимального расстояния Колмогорова, метода минимального расстояния ω^2 Мизеса, метода минимального расстояния Ω^2 Мизеса (*MD*-оценки) и метода оценивания по порядковым статистикам (*L*-оценки).

Сравнивая полученные значения мощности при оценивании обоих параметров нормального закона распределения, можно заметить, что рассмотренные критерии имеют бо́льшую мощность при использовании метода максимального правдоподобия. Это хорошо иллюстрирует рисунок 1, на котором приведены распределения статистики критерия Андерсона-Дарлинга при проверке сложной гипотезы H_0 о согласии с нормальным распределением против гипотезы H_1, соответствующей логистическому закону. Распределения статистик критериев, соответствующие справедливости H_1, наиболее сильно сдвинуты вправо именно при использовании оценок максимального правдоподобия.

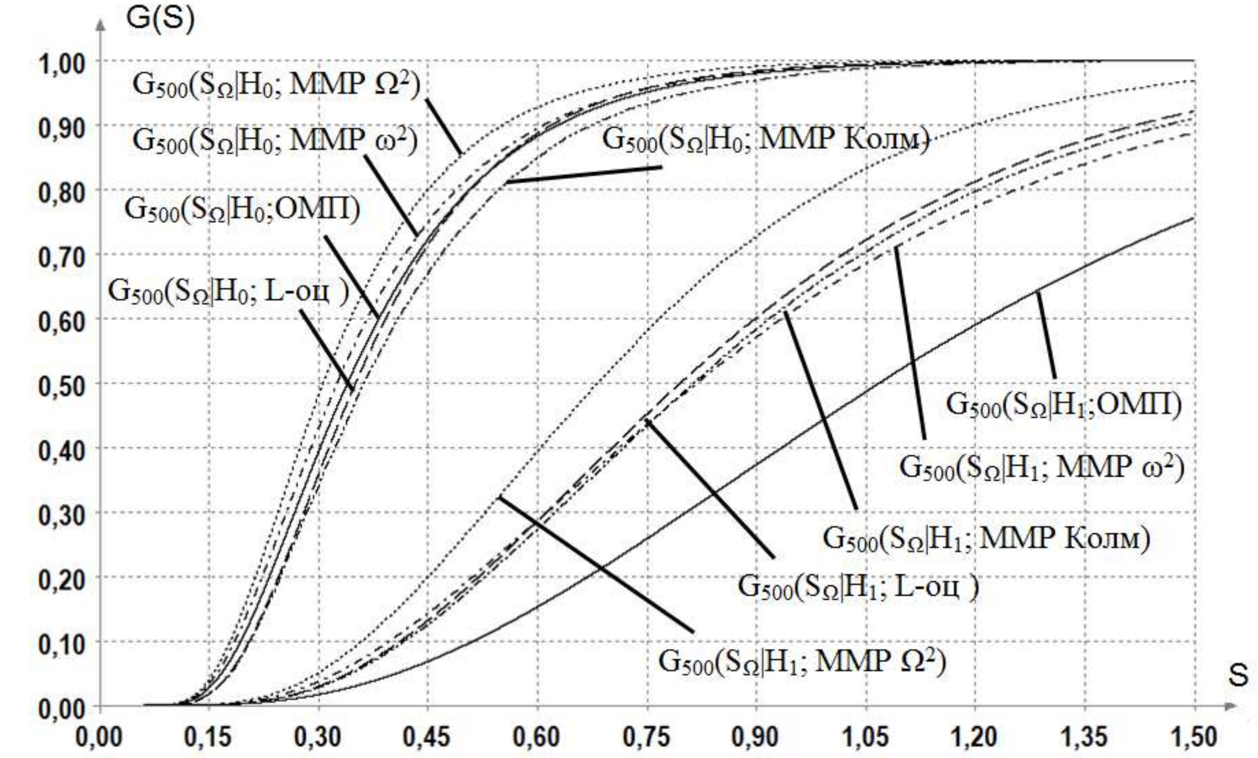

Рисунок 1 - Распределения статистики критерия Ω^2 Андерсона — Дарлинга $G(S_\Omega|H_0)$ и $G(S_\Omega|H_1)$ при проверке сложной гипотезы H_0 о согласии с нормальным законом в случае использования разных методов оценивания параметров при объеме выборки 500 наблюдений

Анализируя результаты оценки мощности критериев относительно двух пар конкурирующих гипотез (нормальное распределение против логистического и распределение Вейбулла-Гнеденко против гамма-распределения), можно отметить следующее.

В тех случаях, когда при проверке сложной гипотезы оцениваются оба параметра или только параметр масштаба, рассмотренные критерии согласия обладают большей мощностью, как правило, при использовании метода максимального правдоподобия.

Однако при сложной гипотезе, в процессе проверки которой оценивается *только параметр сдвига* нормального распределения, критерии оказываются более мощными при использовании оценок, получаемых в результате минимизации статистики применяемого критерия согласия.

В то же время при проверке сложной гипотезы относительно распределения Вейбулла-Гнеденко с оцениванием *только параметра формы* закона критерии показали более высокую мощность в случае использования оптимальных *L-оценок* по выборочным квантилям.

Таким образом, распределения статистик непараметрических критериев согласия при проверке сложных гипотез зависят от вида проверяемой гипотезы и существенно зависят от метода оценивания [6]. Метод оценивания параметров оказывает существенное влияние на мощность непараметрических критериев [11]. Если подходить к выбору метода оценивания с позиций максимизации мощности используемого критерия, то в общем случае предпочтение следует отдать, как и было показано в [11], методу максимального правдоподобия. К тому же оценки максимального правдоподобия обладают наилучшими статистическими свойствами.

В то же время, как показывают результаты данных исследований, в “частных” случаях сложных проверяемых гипотез (при оценивании части параметров), преимущество в мощности (конечно, это зависит от конкурирующей гипотезы) может быть за другим методом оценивания. Вообще говоря, это открывает некоторые перспективы. Если нас интересует задача построения критерия максимальной мощности, можно (в зависимости от вида сложной гипотезы) для заданной пары конкурирующих законов выбрать тот метод оценивания, при котором мощность соответствующего критерия будет наиболее высокой [19].

В пятой главе рассмотрены критерии проверки однородности распределений Смирнова, Лемана-Розенблатта и Андерсона-Дарлинга-Петита, исследована мощность последнего по сравнению с двумя первыми. Проведено сравнение мощности ряда критериев однородности на данных типа времени жизни для альтернатив с пересечениями и без пересечений функций распределения.

Гипотеза однородности – это предположение о том, что две (и более) выборки взяты из одной и той же генеральной совокупности. Для данной гипотезы неважно, какое именно распределение имеет генеральная совокупность, поэтому «идеальный» критерий однородности не должен зависеть от вида закона распределения случайной величины.

В работе (Лемешко Б.Ю., Лемешко С.Б. О сходимости распределений статистик и мощности критериев однородности Смирнова и Лемана-Розенблатта // Измерительная техника. 2005. № 12. – С.9-14.) представлены результаты исследований критериев однородности Смирнова и Лемана-Розенблатта. Эти критерии являются двухвыборочными аналогами критериев согласия Колмогорова и Крамера-Мизеса-Смирнова соответственно. В работах [14-16] по сравнительному анализу мощности критериев согласия относительно ряда конкурирующих гипотез было показано, что из рассмотренных непараметрических критериев согласия Колмогорова, Крамера-Мизеса-Смирнова и Андерсона-Дарлинга наиболее мощным, как правило, оказывается критерий Андерсона-Дарлинга.

По-видимому, первое исследование критерия Андерсона-Дарлинга как критерия однородности было проведено Петитом в 1976 году. Позднее Шольц и Стефенс обобщили этот критерий на случай k выборок. В данной главе проведены исследования мощности критерия однородности Андерсона-Дарлинга-Петита для ряда фиксированных альтернатив и выполнено его сравнение с другими критериями однородности.

Двухвыборочный аналог критерия Андерсона-Дарлинга имеет вид

$$A^2 = \frac{mn}{m+n} \int_{-\infty}^{\infty} \frac{[G_m(x) - F_n(x)]^2}{(1 - H_{m+n}(x)) H_{m+n}(x)} dH_{m+n}(x). \tag{13}$$

Петитом в 1976 году получено простое выражение для статистики (13)

$$A^2 = \frac{1}{mn} \sum_{i=1}^{m+n-1} \frac{\left(M_i(m+n) - mi\right)^2}{i(m+n-i)}, \tag{14}$$

где M_i – число элементов из первой выборки, меньших или равных i-му элементу вариационного ряда объединенной выборки.

Для исследования того, насколько значительно распределение статистики A^2 при ограниченных объемах сравниваемых выборок отличается от предельного закона $a2(t)$ (насколько быстро распределение статистики A^2 сходится к предельному закону), использовалась методика исследования скорости сходимости распределения статистики к предельному, описанная *в главе 1* диссертации (п. 1.8).

Моделировались выборки одинакового объёма m и n от 5 до 50. По каждой выборке вычислялась статистика A^2. Процедура повторялась N раз (N=1660000). Расстояние до предельного закона распределения вычислялась по формуле

$$D_{n,N} = \sup_{|x|<\infty} \left| G_{n,N}(x) - G(x) \right|, \tag{15}$$

где $G_{n,N}(x)$ – эмпирическая функция распределения статистики при конечном n, а $G(x)$ – предельное распределение статистики.

Расстояние $D_{n,N}$ уменьшается с ростом числа наблюдений в выборках и при $n = m > 15$ описывается степенной зависимостью с высоким коэффициентом детерминации, а скорость сходимости примерно равна $O\left(n^{-0,93}\right)$. Для того чтобы расстояние (15) от распределения статистики A^2 до предельного закона распределения $a2(t)$ не превышало 0,01 достаточно взять выборки объемом 45 наблюдений.

В работе [32] проведено исследование различных критериев однородности по данным типа времени жизни. Были рассмотрены критерии проверки гипотезы однородности Смирнова, Лемана-Розенблата, Андерсона-Дарлинга-Петита, Вилкоксона, логарифмический ранговый критерий, Q-критерий, критерий Багдонавичуса-Никулина при близких конкурирующих гипотезах.

Результаты исследований позволяют сделать следующие выводы.

1.Мощность критерия Андерсона-Дарлинга-Петита, как правило, превосходит мощность критериев Смирнова и Лемана-Розенблатта, особенно при близких альтернативах.

2. При проверке гипотезы однородности по данным типа времени жизни для альтернативы с пересечениями функций распределения наибольшую мощность имеет критерий Багдонавичуса-Никулина, а для альтернативы без пересечений – логранговый критерий.

3. Если вид конкурирующей гипотезы неизвестен, то наиболее предпочтительными по правилу Вальда оказываются критерии Лемана-Розенблатта, Андерсона-Дарлинга-Петита и Багдонавичуса-Никулина.

В шестой главе рассмотрены инвариантные критерии проверки гипотезы о многомерной нормальности. Исследованы вопросы сходимости распределений статистик критериев многомерной нормальности к предельному закону. Найдены аппроксимации законов распределения статистик критериев радиуса. Проведено сравнение мощности инвариантных критериев многомерной нормальности.

Выполнение предположения о нормальности многомерной случайной величины является необходимым условием корректного применения большинства классических критериев многомерного статистического анализа. Например, многомерная нормальность предполагается при проведении дисперсионного анализа (MANOVA), дискриминантного анализа и многомерной регрессии. Исследования, проведенные Хопкинсом и Клаем, Мардия, Коновером и Иманом показали важность выполнения предположения о многомерной нормальности, неробастность многих процедур и влияние на корректность результатов анализа нарушения предположения о нормальности данных.

Для проверки многомерной нормальности разработаны специальные статистические критерии. В данной главе рассмотрены критерии, которые являются подмножеством класса инвариантных критериев проверки многомерной нормальности, приведенные в обзоре Хенце в 2002 году.

В многомерном случае нормальное распределение является симметричным, поэтому его коэффициент многомерной асимметрии равен нулю, а коэффициент эксцесса зависит от размерности d случайной величины и равен $d(d+2)$. Критерии, основанные на вычислении коэффициентов асимметрии и эксцесса, исследовались многими авторами. В 1970 году Мардия предложил критерии проверки гипотезы о принадлежности многомерной выборки нормальному закону, в которых статистики критериев основывались на выборочных коэффициентах асимметрии и эксцесса. Мори, Рохатги и Шекли предложили свою модификацию критерия, основанного на вычислении коэффициента асимметрии. В 1984 Шривастава предложил использовать разложение ковариационной матрицы на собственные вектора в критериях проверки многомерной нормальности

Другой подход к проверке многомерной нормальности основан на *полярной декомпозиции* – представлении нормального вектора в виде произведения угловой части и радиуса. При этом угловая часть имеет равномерное распределение на поверхности единичной сферы, а квадрат радиуса подчиняется χ_d^2-распределению.

Критерии нормальности, основанные на полярной декомпозиции, исследовались в работах автора [29, 30].

Пусть $X_1, X_2, \ldots, X_n$ последовательность независимых случайных d-мерных векторов X и $N_d(\mu,\Sigma)$ – d-мерное нормальное распределение с вектором математического ожидания $M=[M_1,\ldots,M_d]$ и ковариационной матрицей Σ. Функция плотности многомерного нормального закона имеет вид

$$f(X)=\frac{1}{(2\pi)^{n/2}\left|\Sigma\right|^{1/2}}e^{-\frac{1}{2}(X-M)^{\mathrm{T}}\Sigma^{-1}(X-M)}. \qquad (16)$$

Обозначим через N_d – класс всех невырожденных d-мерных нормальных распределений. Задача проверки многомерной нормальности состоит в проверке гипотезы: $H_0: P^X \in N_d$, где P^X – распределение вектора X.

Среди множества критериев проверки многомерной нормальности выделяются *инвариантные критерии*. Критерий называется *инвариантным* относительно аффинных преобразований, если выполняется следующее условие: по выборкам $\{X_1, X_2, \dots X_n\}$ и $\{AX_1+b, AX_2+b, \dots AX_n+b\}$ они должны с одинаковой вероятностью отклонять или принимать гипотезу H_0. Следовательно, любая статистика $T_n = T_n(X_1,...,X_n)$ инвариантного критерия проверки гипотезы H_0 должна удовлетворять условию инвариантности: $T_n(AX_1+b,...,AX_n+b) = T_n(X_1,...,X_n)$, для произвольного вектора $b \in R^d$ и невырожденной матрицы $A \in R^{d\times d}$.

В основе критериев типа «радиуса» лежит статистика:

$$R_j^2 = D_{n,jj} = \left\|Y_{n,j}\right\|^2 = \left(X_j - \bar{X}_n\right)^T S_n^{-1}\left(X_j - \bar{X}_n\right), \tag{17}$$

где $\bar{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$, $S_n = \frac{1}{n}\sum_{j=1}^{n}\left(X_j - \bar{X}_n\right)\left(X_j - \bar{X}_n\right)^T$, которая асимптотически распределена по χ_d^2-распределению.

Пусть $G_d(t)$ – функция χ_d^2-распределения. Тогда проверка гипотезы о нормальности эквивалентна проверке гипотезы о том, что статистика (17) подчинена χ_d^2-распределению. Это можно сделать с помощью критериев согласия, описанных в главе 4 диссертации. В 1982 году Козиол предложил использовать статистику Крамера-Мизеса, в 1987 году Полсон, Рухон и Сулло предложили использовать критерий Андерсона-Дарлинга.

Гипотеза о том, что статистика (17) подчинена χ_d^2-распределению, является простой. Но так как $D_{n,jj}$ являются зависимыми друг от друга величинами, то распределения статистик критериев согласия при справедливости проверяемой гипотезы в данном случае отличаются от предельных законов, имеющих место при проверке простых гипотез о выборках независимых одинаково распределенных случайных величин.

В таких ситуациях распределения статистик критериев Колмогорова, Андерсона-Дарлинга и Крамера-Мизеса зависят от размерности d и хорошо аппроксимируются бета-распределениями III-го рода. При построении моделей распределений статистик критериев аппроксимация проводилась по выборкам статистик критериев радиуса при $n = 1000$ (объем выборки многомерной нормальной случайной величины), $N = 1000000$ (количество выборок). С ростом размерности d различие между законами распределения статистик уменьшается так, что графики законов распределений при $d = 7$ и $d = 8$ становятся практически неразличимы. Поэтому для практических целей при размерностях больших 8 можно пользоваться аппроксимацией для $d = 8$.

В данной работе при исследовании мощности в качестве конкурирующих рассматриваются гипотезы, предполагающие принадлежность выборки многомерному распределению, построенному на базе одномерных законов, отличных от нормального: H_1 – на основании распределения Коши; H_2 – на основаниии распределения Лапласа; H_3 – на основании логистического распределения; H_4 – на основании двустороннего экспоненциального распределения с параметром формы 4; H_5 – на основании равномерного распределения; H_6 – на основании распределения наибольшего значения. Моделирование многомерного распределения основано на известном алгоритме (Ермаков С.М., Михайлов Г.А. Статистическое моделирование. М.: ФИЗМАТЛИТ, 1982. - 296 с.), но вместо выборок из одномерного нормального распределения используются выборки из другого одномерного (ненормального) закона.

Результаты исследования мощности критериев показали следующее.

Наибольшей мощностью относительно конкурирующих гипотез H_1 (распределение Коши), H_2 (распределение Лапласа) и H_3 (логистическое распределение) обладает критерий эксцесса Мардия. Немного уступают ему критерии эксцесса Козиола и Шриваставы.

Наибольшей мощностью относительно гипотез H_4 (двустороннее экспоненциальное распределение с параметром формы 4), H_5 (равномерное распределение) обладает “критерий радиуса” Крамера-Мизеса-Смирнова. Немного уступает ему “критерий радиуса” Андерсона-Дарлинга.

Наибольшей мощностью относительно конкурирующей гипотезы H_6 (распределение наибольшего значения) обладают критерии асимметрии Мардия и Мори. Немного меньше мощность у критерия асимметрии Шриваставы.

Критерии асимметрии, как и следовало ожидать, оказались наилучшими для случая несимметричной альтернативы, однако относительно конкурирующих законов с легкими хвостами (без асимметрии) они имеют нулевую мощность, т.е. являются смещенными и несостоятельными. *Критерии эксцесса* оказались очень хороши относительно конкурирующих законов с более тяжелыми хвостами (Коши, Лапласа), а также против логистического распределения. Однако, как и критерии асимметрии, критерии эксцесса являются смещенными и несостоятельными против альтернативы с легкими хвостами. *Критерий “уг-*

лов” показал невысокую мощность практически на всех альтернативах, и также как критерии асимметрии и эксцесса оказался смещенным при альтернативах с легкими хвостами. *Критерии “радиуса”* оказались наилучшими против альтернатив с легкими хвостами, уступая при этом критериям асимметрии и эксцесса в случае остальных альтернатив. Недостатком критериев данного типа является быстрое падение мощности при увеличении размерности d при небольших объемах выборок.

Применение правила Вальда к результатам исследования критериев позволяет указать в качестве наиболее предпочтительного для применения на практике критерий “радиуса” Андерсона-Дарлинга. Для всех рассмотренных конкурирующих гипотез этот критерий показал наибольшую мощность в наихудших условиях.

В седьмой главе рассмотрены критерии проверки однородности выборки случаев и контрольной выборки при проведении полногеномного анализа ассоциаций. Проведено сравнение мощности критериев ассоциаций. Рассмотрена задача оптимального планирования эксперимента для проверки гипотезы однородности. Показана связь между симметричной дивергенцией Кульбака-Лейблера и процедурой оптимального планирования эксперимента при проверке гипотезы однородности при полногеномном анализе ассоциаций.

В генетической эпидемиологии *полногеномный анализ ассоциаций* (genome-wide association study, GWAS) – это изучение общей генетической информации у разных индивидуумов с целью обнаружения того, какие генетические варианты ассоциированы с заболеванием. Обычно в GWAS исследуют связь между однонуклеотидным полиморфизмом (single nucleotide polymorphism, SNP) и основными заболеваниями.

Предположим, что имеется три генотипа AA, Aa, и aa для однонуклеотидного полиморфизма с двумя аллелями, A и a. Пусть имеется r объектов в выборке случаев (*cases*) и s объектов в контрольной выборке (*controls*). Среди r объектов в выборке случаев r_0, r_1 и r_2 человек имеют генотипы AA, Aa и aa соответственно. Среди s объектов в контрольной выборке s_0, s_1 и s_2 человек имеют генотип AA, Aa и aa соответственно.

Основной гипотезой H_0 является предположение, что нет никакой связи между заболеванием и SNP. Имеется три группы конкурирующих гипотез, основанных на *модели наследования* (MOI – mode of inheritance): H_r – рецессивной, H_d – доминантной и H_a – аддитивной модели наследования.

Пусть $p_a = P(a)$ – частота аллели a в популяции, p_i = P{число аллелей a равно i |объект из выборки случаев} и q_i = P{число аллелей a равно i |объект из контрольной выборки}; f_i = P{объект в выборке случаев | число аллелей a равно i} – *пенетрантность*, K=P{попадание в выборку случаев} – распространенность заболевания в популяции.

Конкретная конкурирующая гипотеза задается MOI, частотой p_a, и отношением шансов $\psi_i = \dfrac{f_i(1-f_0)}{f_0(1-f_i)} = \dfrac{p_i \cdot q_0}{p_0 \cdot q_i}$, $i = 1,2$, или относительным генетическим риском $\gamma_i = \dfrac{f_i}{f_0}$, $i = 1,2$.

Разницу между распределениями в выборке случаев и контрольной выборке ($P = \{p_0, p_1, p_2\}$ и $Q = \{q_0, q_1, q_2\}$ соответственно) можно измерить с помощью дивергенции Кульбака-Лейблера (которая была рассмотрена *в главе 2* диссертации, п.2.3):

$$D_{KL} = \sum_{i=0}^{2} p_i \ln \frac{p_i}{q_i}. \tag{18}$$

Если гипотеза H_0 верна, то $D_{KL} = 0$. Дивергенция Кульбака-Лейблера связана с p_a, K, ψ_1 и ψ_2 как [34]

$$D_{KL} = \ln\frac{1-K}{\lambda} + \frac{2p_a(1-p_a)\psi_1}{\lambda + K\psi_1}\ln\psi_1 + \frac{p_a^2 \cdot \psi_2}{\lambda + K\psi_2}\ln\psi_2, \tag{19}$$

где λ может быть найдено из уравнения [34]

$$\frac{\left(1-p_a\right)^2}{\lambda + K} + \frac{2p_a(1-p_a)\psi_1}{\lambda + K\psi_1} + \frac{p_a^2 \cdot \psi_2}{\lambda + K\psi_2} = 1. \tag{20}$$

Дивергенция Кульбака-Лейблера связана с p_a, K, γ_1 и γ_2 как [34]

$$D_{KL} = \tau\left[\left(1-p_a\right)^2 \ln\frac{1-K}{\tau - K} + 2p_a\left(1-p_a\right)\gamma_1 \ln\frac{\gamma_1(1-K)}{\tau - \gamma_1 K} + p_a^2 \cdot \gamma_2 \ln\frac{\gamma_2(1-K)}{\tau - \gamma_2 K}\right], \tag{21}$$

где $\tau = (1-p_a)^2 + 2p_a(1-p_a)\gamma_1 + p_a^2 \cdot \gamma_2$.

Сравнение мощности критериев ассоциаций проведено для трех конкурирующих гипотез с разными видами наследуемости. Показано, что мощность критерия основанного на эмпирической дивергенции Кульбака-Лейблера больше, чем мощность критерия χ^2, но меньше, чем мощность критериев MAX3, GMS и CLRT, которые наиболее предпочтительны при использовании на практике [33].

В результате исследований выявлена статистическая зависимость необходимого объема выборки от симметричной дивергенции Кульбака-Лейблера между законами распределения в выборке случаев и в контрольной выборках для заданной альтернативной гипотезы (модели наследования).

Различные критерии ассоциаций можно сравнить на основании объемов выборок, требуемых для принятия решения с заданными вероятностями ошибок первого и второго рода. Необходимый объем выборки для критерия тренда Кокрена-Армитеджа пропорционален величине

$$\vartheta(\alpha,\beta) = \left(\Phi^{-1}\left(1-\frac{\alpha}{2}\right) + \Phi^{-1}\left(1-\beta\right)\right)^2, \tag{22}$$

где α и β - вероятности ошибок первого и второго рода, соответственно.

Тогда для критерия тренда Кокрена-Армитеджа зависимость необходимого объема выборки от вероятностей ошибок первого и второго рода и симметричной дивергенции Кульбака-Лейблера между распределениями в выборке случаев и контрольной выборке имеет вид [34]

$$n_{CATT} = \frac{4,035 \cdot \vartheta(\alpha,\beta)}{\rho_{KL}}, \tag{23}$$

а для критерия MAX3 соответственно [34]

$$n_{MAX\,3} = \frac{4,238 \cdot \vartheta(\alpha,\beta)}{\rho_{KL}}. \tag{24}$$

В этом случае относительная эффективность критерия MAX3 по отношению к критерию тренда Кокрена-Армитеджа с оптимальным набором коэффициентов, равна 0,95, т.е. критерий MAX3 требует в среднем на 5% больше наблюдений, чем оптимальный критерий [34].

Целью *оптимального планирования эксперимента* является нахождение таких планов, при которых затраты на проведение эксперимента минимальны. Для одноэтапного эксперимента это эквивалентно минимизации объема выборки при заданных вероятностях ошибок первого и второго рода, к чему и сводится оптимальное планирование.

Оптимизация для двухэтапного эксперимента состоит в выборе пропорции между размерами выборок на первом и втором этапах, а также критических границ таким образом, чтобы минимизировать совокупные затраты при заданных вероятностях ошибок первого и второго рода.

Совокупные затраты двухэтапного эксперимента вычисляются как

$$C = n\left(C_R + \pi m C_{G1} + (1-\pi)\left((m-d)\alpha_1 + d(1-\beta_1)\right)C_{G2}\right), \tag{25}$$

где C_R – затраты на фенотипирование одного человека, C_{G1} – затраты на генотипирование одного маркера на первом этапе и C_{G2} – на втором этапе, m – количество маркеров, d – число ассоциированных с заболеванием маркеров, n – общее число людей, протестированных на обоих этапах, π – пропорция между числом человек, протестированных на первом этапе, и общим количеством людей, α_1 – вероятность ошибки первого рода на первом этапе, β_1 – вероятность ошибки второго рода на первом этапе.

Показано, что зависимость затрат (25) на проведение эксперимента обратно пропорционально симметричной дивергенции Кульбака-Лейблера между распределениями в выборке случаев и контрольной выборке (рисунок 2).

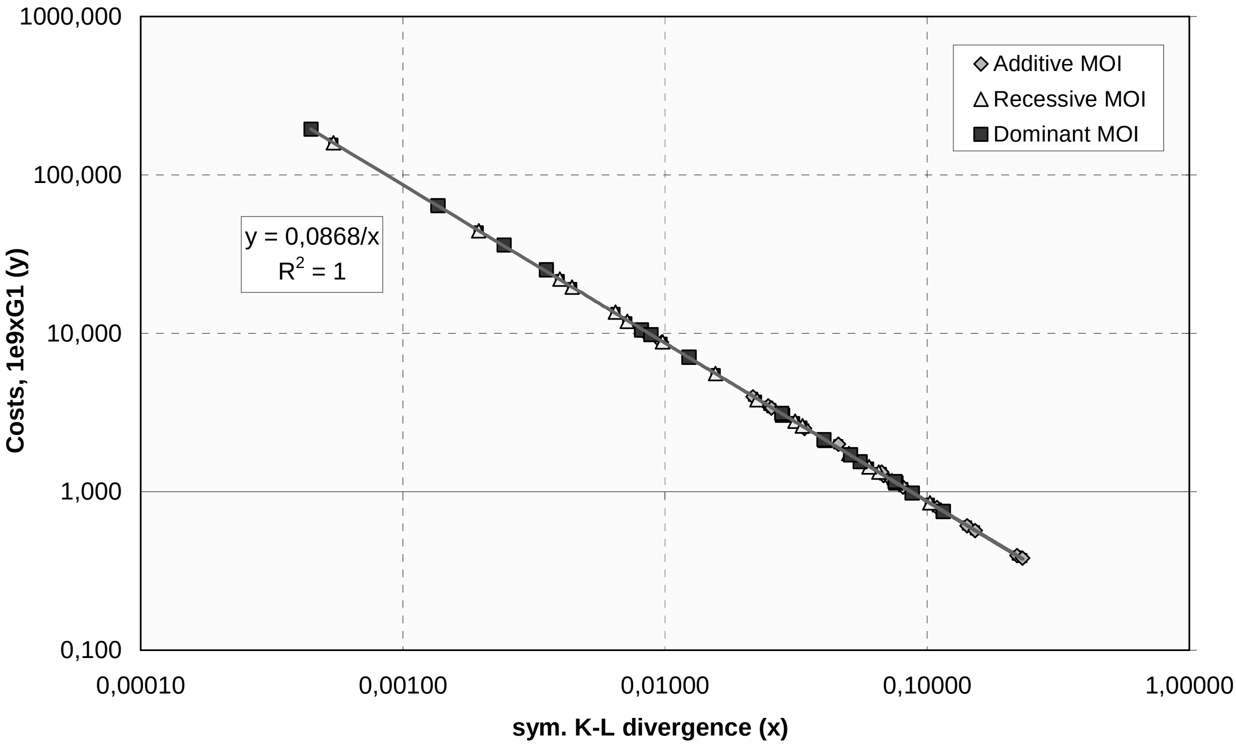

Рисунок 2 – Зависимость затрат от ρ_{KL} для критерия MAX3

В восьмой главе рассмотрено программное обеспечение статистического анализа данных, которое использовалось при получении представленных в диссертации результатов, а также программное обеспечение, в котором были реализованы полученные в диссертации результаты.

Современный статистический анализ невозможен без специализированного программного обеспечения. Большинство результатов, полученных в данной работе, связаны с разработанным программным обеспечением для статистического анализа и компьютерного моделирования статистических закономерностей.

Автор диссертационной работы занимался разработкой системы статистического анализа интервальных наблюдений одномерных непрерывных случайных величин "Интервальная статистика" (ISW) с 1996 года [2,3,28]. В результате была разработана объектная модель системы, реализованы процедуры ввода/вывода выборок, оценивания параметров, проверки гипотезы о согласии, моделирования выборок, распределений оценок и распределений статистик непараметрических критериев. На программные системы ISW версий 4.0, 5.0 и 5.1 получены свидетельства о государственной регистрации [37-39].

Программная система «НКЦ ИТР: Статистика 1.0» разрабатывалась по заказу ООО «Научно-консультационный центр инженерно-технических решений» и была поддержана грантом Фонда содействия развитию малых форм предприятий в научно-технической сфере в 2011 году (проект №13653). Автор диссертационной работы занимался общим руководством проекта, разработал объектную модель, интерфейс системы, модули визуализации [31]. На систему «НКЦ ИТР: Статистика 1.0» имеются свидетельства о государственной регистрации программы ЭВМ [35, 36], а также сертификат фирмы «1С» о совместимости с системой 1С:Предприятие 8.2.

Программная платформа 1С:Предприятие 8.2 широко распространена на территории Российской Федерации и ряда стран СНГ. Особенностью данной платформы является масштабируемость разрабатываемых на ее базе конфигураций: приложение может работать как на локальном компьютере, так и в связке с мощными системами управления базами данных, такими как MS SQL Server и ORACLE. Платформа позволяет работать с базой данных в различных режимах: толстый клиент, тонкий клиент и Web-клиент.

Разработанные системы базируются на фундаментальных результатах, представленных в монографии [1]. Программное обеспечение может быть использовано в промышленности в отделах по контролю качества и метрологических лабораториях; в медицинских, научных и образовательных учреждениях.

В приложении 1 приведен акт о внедрении результатов диссертационной работы в учебный процесс Новосибирского государственного технического университета. **В приложениях 2 и 3** приведены копии свидетельств о регистрации программ для ЭВМ. **В приложениях 4 и 5** приведены фрагменты рекомендаций по стандартизации Госстандарта РФ по правилам применения критериев согласия Р 50.1.033-2001 и Р 50.1.037-2002.

ЗАКЛЮЧЕНИЕ

Таким образом, в диссертации на базе развиваемого программного обеспечения, ориентированного на интенсивное использование компьютерных методов исследования статистических закономерностей, получил дальнейшее развитие классический аппарат прикладной математической статистики. Основные результаты, полученные в диссертационной работе и расширяющие возможности методов прикладной математической статистики, состоят в следующем.

1. Разработан алгоритм оптимального планирования эксперимента для различения двух простых гипотез по критерию отношения правдоподобия.
2. Разработан алгоритм оценивания точных критических границ для последовательных критериев Вальда, Айвазяна и Лордена. Показано, что использование точных критических границ может сократить средний объем наблюдений примерно на 6-11% для критерия Вальда и на 25-30% для критериев Айвазяна и Лордена.
3. Показано, что распределения статистик непараметрических критериев согласия и их мощность при проверке сложных гипотез зависят от вида проверяемой гипотезы и существенно зависят от метода оценивания.
4. Показано, что распределение статистики критерия однородности Андерсона-Дарлинга-Петита сходится к предельному распределению со скоростью примерно $O\left(n^{-0,93}\right)$. Для практических целей можно использовать распределение уже при выборках объемом 45 наблюдений.
5. Показано, что наиболее предпочтительными по правилу Вальда при проверке гипотезы однородности распределений являются критерии Лемана-Розенблатта, Андерсона-Дарлинга-Петита и Багдонавичуса-Никулина.
6. Построены модели, аппроксимирующие законы распределения статистик критериев “радиуса” Крамера-Мизеса-Смирнова, Андерсона-Дарлинга и Колмогорова при проверке гипотезы о многомерной нормальности.
7. На основании правила Вальда принятия решения в условиях неопределенности сделана рекомендация о предпочтительном использовании для проверки многомерной нормальности критерия “радиуса” Андерсона-Дарлинга.
8. В результате исследования критерия тренда Кокрена-Армитеджа с использованием оптимального набора коэффициентов и критерия MAX3 показано, что оптимальный объем выборки обратно пропорционально зависит от симметричной дивергенции Кульбака-Лейблера между законами распределения в выборке случаев и в контрольной выборке для любой мо-

дели наследования. Показано, что критерий MAX3 требует в среднем на 5% больше наблюдений, чем оптимальный критерий.

9. При участии автора разработано программное обеспечение для проведения статистического анализа, которое использовалось для получения новых результатов. Разработанные системы, в свою очередь, базируются на полученных фундаментальных результатах исследований.
10. На основе результатов исследований свойств критериев согласия при проверке простых и сложных гипотез разработаны рекомендации по стандартизации Госстандарта РФ по правилам применения критериев согласия Р 50.1.033-2001 и Р 50.1.037-2002.

Основные публикации автора по теме диссертации

Монография

1. Постовалов С.Н. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход: монография / Б.Ю. Лемешко, С.Б. Лемешко, С.Н. Постовалов, Е.В. Чимитова. – Новосибирск : Изд-во НГТУ, 2011. – 888 с. (серия "Монографии НГТУ").

Журналы из Перечня ВАК ведущих рецензируемых научных изданий для опубликования основных результатов диссертаций на соискание учёной степени доктора и кандидата наук:

2. Постовалов С.Н. Статистический анализ одномерных наблюдений по частично группированным данным / Лемешко Б.Ю., Постовалов С.Н.// Изв. вузов. Физика. - Томск, 1995. - № 9. - С. 39-45.
3. Постовалов С.Н. Вопросы обработки выборок одномерных случайных величин / Лемешко Б.Ю., Постовалов С.Н. // Научный вестник НГТУ. - Новосибирск, 1996. - № 2. - С. 3-24.
4. Постовалов С.Н. О решении задач статистического анализа интервальных наблюдений / Лемешко Б.Ю., Постовалов С.Н. // Вычислительные технологии. - 1997. - Т.2. - № 1. - С. 28-36.
5. Постовалов С.Н. О зависимости предельных распределений статистик Хи-квадрат Пирсона и отношения правдоподобия от способа группирования данных / Лемешко Б.Ю., Постовалов С.Н. // Заводская лаборатория. 1998. Т. 64. - № 5. - С. 56-63.
6. Постовалов С.Н. О распределениях статистик непараметрических критериев согласия при оценивании по выборкам параметров наблюдаемых законов / Лемешко Б.Ю., Постовалов С.Н. // Заводская лаборатория. 1998. Т. 64. - № 3. - С. 61-72.

7. Постовалов С.Н. Об оценивании параметров распределений по интервальным наблюдениям / Лемешко Б.Ю., Постовалов С.Н. // Вычислительные технологии. 1998. Т.3. - № 2. - С. 31-38.
8. Постовалов С.Н. К оцениванию параметров надежности по цензурированным выборкам / Лемешко Б.Ю., Гильдебрант С.Я., Постовалов С.Н. // Заводская лаборатория. Диагностика материалов. 2001. Т. 67. - № 1. - С. 52-64.
9. Постовалов С.Н. Проверка сложных гипотез о согласии с распределениями Джонсона на основе использования непараметрических критериев / Зайцева Е.А., Лемешко Б.Ю., Постовалов С.Н. // Вестник СибГАУ, 2002. Вып.3. - Красноярск: Издательство СибГАУ. - С.70-77.
10. Постовалов С.Н. Применение непараметрических критериев согласия при проверке сложных гипотез / Лемешко Б.Ю., Постовалов С.Н. // Автометрия, 2001. - № 2. - С. 88-102.
11. Постовалов С.Н. О зависимости распределений статистик непараметрических критериев и их мощности от метода оценивания параметров / Лемешко Б.Ю., Постовалов С.Н. // Заводская лаборатория. Диагностика материалов, 2001. - Т.67. - № 7. - С. 62-71.
12. Постовалов С.Н. О распределениях статистики и мощности критерия типа Хи-квадрат Никулина / Лемешко Б.Ю., Постовалов С.Н., Чимитова Е.В. // Заводская лаборатория. Диагностика материалов, 2001. - Т. 67. - № 3. - С. 52-58.
13. Постовалов С.Н., Французов А.В. К применению непараметрических критериев согласия для проверки адекватности непараметрических моделей / Лемешко Б.Ю., Постовалов С.Н., Французов А.В. // Автометрия, 2002. - № 2. - С.3-14.
14. Постовалов С.Н. Сравнительный анализ мощности критериев согласия при близких альтернативах. II. Проверка сложных гипотез / Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н. // Сибирский журнал индустриальной математики, 2008. - Т.11. - № 4(36). - С.78-93.
15. Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н. Сравнительный анализ мощности критериев согласия при близких конкурирующих гипотезах. I. Проверка простых гипотез // Сибирский журнал индустриальной математики, 2008. - Т.11. - № 2(34). - С.96-111.
16. Постовалов С.Н. Мощность критериев согласия при близких альтернативах / Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н. // Измерительная техника, 2007. - № 2. – С.22-27.
17. Postovalov S.N. Statistic Distribution Models for Some Nonparametric Goodness-of-Fit Tests in Testing Composite Hypotheses / Lemeshko B.Yu., Lemeshko S.B. and Postovalov S.N. // Communications in Statistics - Theory and Methods, 2010. - Vol. 39. - No. 3. – P. 460-471.

18. Постовалов С. Н. Проверка простых и сложных гипотез с использованием последовательных критериев Лордена и Айвазяна / С. Н. Постовалов, М. Р. Шахмаметова // Научный вестник НГТУ, 2011. - № 3 (44). - С. 17-28.
19. Постовалов С. Н. Сравнительный анализ мощности критериев согласия при проверке сложных гипотез в зависимости от метода оценивания / С. Н. Постовалов, Е. А. Наумова // Сибирский журнал индустриальной математики. - 2013. - том XVI, №1 (53). - С. 84-94
20. Постовалов С. Н. Компьютерное моделирование и исследование вероятностных закономерностей / Б. Ю. Лемешко, А. А. Горбунова, С. Б. Лемешко, С. Н. Постовалов, А. П. Рогожников, Е. В. Чимитова // Вестник Томского государственного университета. Управление, вычислительная техника и информатика, 2013. – №1(22). – С. 74-85.

Публикации в рецензируемых изданиях:

21. Постовалов С.Н. Прикладные аспекты использования критериев согласия в случае проверки сложных гипотез / Лемешко Б.Ю., Постовалов С.Н. // Надежность и контроль качества, 1997. - № 11. - С. 3-17.
22. Постовалов С.Н. О правилах проверки согласия опытного распределения с теоретическим / Лемешко Б.Ю., Постовалов С.Н. // Методы менеджмента качества. Надежность и контроль качества, 1999. - № 11. - С. 34-43.
23. Постовалов С.Н. Непараметрические критерии при проверке сложных гипотез о согласии с распределениями Джонсона / Лемешко Б.Ю., Постовалов С.Н. // Доклады СО АН ВШ, 2002. - № 1(5). - С.65-74.
24. Постовалов С.Н. Проверка простых и сложных гипотез с использованием последовательного критерия Вальда / Постовалов С.Н. // ДОКЛАДЫ АН ВШ РФ, 2011. - № 2(17). - С.140-150.
25. Постовалов С.Н. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Методические рекомендации. Часть I. Критерии типа Хи-квадрат. / Денисов В.И., Лемешко Б.Ю., Постовалов С.Н.- Новосибирск: Изд-во НГТУ, 1998. - 126 с.
26. Постовалов С.Н. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Методические рекомендации. Часть II. Непараметрические критерии. / Лемешко Б.Ю., Постовалов С.Н. - Новосибирск: Изд-во НГТУ, 1999. - 85 с.
27. Постовалов С.Н Компьютерные технологии анализа данных и исследования статистических закономерностей. Учебное пособие / Лемешко Б.Ю., Постовалов С.Н. – Новосибирск: Изд-во НГТУ, 2004. - 119 с.
28. Software System for Simulation and Research of Probabilistic Regularities and Statistical Data Analysis in Reliability and Quality Control / B.Yu. Lemeshko, S.B. Lemeshko, E.V. Chimitova, S.N. Postovalov, A.P. Rogozhnikov // In: Mathematical and

Statistical Models and Methods in Reliability. Applications to Medicine, Finance, and Quality Control / Editors: V. Rykov, N. Balakrishnan, M. Nikulin / Series “Statistics for Industry and Technology” / Birkhäuser, Boston. - 2011. – P. 417-432.

Материалы конференций:

29. Постовалов С.Н. Исследование критериев проверки многомерной нормальности / С. В. Чебукина, С. Н. Постовалов // Рос. науч.-техн. конф. "Информатика и проблемы телекоммуникаций", Новосибирск, 21 – 22 апр., 2011.: Материалы конф. – Новосибирск: Изд-во СибГУТИ, 2011. – Т. 1. – С.121-124
30. Постовалов С. Н. Критерии проверки многомерной нормальности, основанные на углах и радиусе / С. Н. Постовалов, М. В. Каньшин // Рос. науч.-техн. конф. "Обработка информационных сигналов и математическое моделирование", Новосибирск, 26 – 27 апр., 2012.: Материалы конф. – Новосибирск: Изд-во СибГУТИ, 2012. – С.34-36
31. Программная система статистического моделирования в задачах проведения и обработки измерений на платформе 1С:Предприятие 8.2 / С. Н. Постовалов, Д. Г. Демин, М. В. Каньшин, Б. И. Окурин, А. В. Поздеева, Т. В. Тимошенко, П. А. Филоненко, М. В. Шиловский // Материалы XI международной конференции "Актуальные проблемы электронного приборостроения" АПЭП-2012, Новосибирск, 2-4 октября 2012г. - Новосибирск: Изд-во НГТУ, 2012. - Т. 6. - С. 41-47
32. Постовалов С. Н. Сравнение мощности критериев однородности для альтернатив с пересечениями и без пересечений / С. Н. Постовалов, П. А. Филоненко // Материалы Российской научно-технической конференции "Обработка информационных сигналов и математическое моделирование", Новосибирск, 23–24 мая 2013 г. – Новосибирск : СибГУТИ, 2013. – С. 91–94.
33. Постовалов С. Н. Сравнение мощности критериев однородности при проведении полногеномного анализа ассоциаций / С. Н. Постовалов // Материалы Российской научно-технической конференции "Обработка информационных сигналов и математическое моделирование", Новосибирск, 23–24 мая 2013 г. – Новосибирск: СибГУТИ, 2013. – С. 88–91.
34. Postovalov S. Optimal Discrete Two-Stage Study Design for Genome-Wide Association Studies / S. Postovalov, A. Ziegler, E. Konomanina // Applied Methods of Statistical Analysis. Simulations and Statistical Inference (AMSA 2013) International Conference. Proceedings. – 2013. Novosibirsk, Russia – P. 238-249

Свидетельства о государственной регистрации программы для ЭВМ:

35. Пат. 2009613046. Набор классов для статистического анализа интервальных наблюдений одномерных непрерывных случайных величин / Постовалов С.Н. // Свидетельство о государственной регистрации программы на ЭВМ № 2009613046 от 16.06.09

36. Пат. 2011618443. Программное обеспечение статистического моделирования в задачах проведения и обработки измерений «НКЦ ИТР: Статистика 1.0» / Постовалов С.Н., Макеева П.М., Постовалова А.Ю., Шиловский М.В., Чебукина С.В., Каньшин М.В., Макеев А.Г., Косарева Л.А. // Свидетельство о государственной регистрации программы на ЭВМ № 2011618443 от 10.01.12
37. Пат. 2012613663. Статистический анализ интервальных наблюдений одномерных непрерывных случайных величин "Интервальная статистика 4.0" / Лемешко Б.Ю., Постовалов С.Н. // Свидетельство о государственной регистрации программы на ЭВМ № 2012613663 от 19.04.12
38. Пат. 2012613664. Статистический анализ интервальных наблюдений одномерных непрерывных случайных величин "Интервальная статистика 5.0" / Лемешко Б.Ю., Постовалов С.Н., Лемешко С.Б., Чимитова Е.В., Рогожников А.П., Щеглов А.Е., Горбунова А.А. // Свидетельство о государственной регистрации программы на ЭВМ № 2012613664 от 19.04.12
39. Пат. 2013615968. Статистический анализ интервальных наблюдений одномерных непрерывных случайных величин «Интервальная статистика 5.1» / Лемешко Б.Ю., Постовалов С.Н., Лемешко С.Б., Чимитова Е.В., Рогожников А.П., Горбунова А.А. // Свидетельство о государственной регистрации программы на ЭВМ № 2013615968 от 25.06.13

Отпечатано в типографии Новосибирского
государственного технического университета
6300073, Новосибирск, пр. К. Маркса, 20
Тел./факс (383) 346-08-57
Формат 60х 84/16 Объем 2,5 п.л. Тираж 120 экз.
Заказ№ ____ Подписано в печать « » _________ 2014 г.