

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

На правах рукописи

Постовалов Сергей Николаевич

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

05.13.17 – «Теоретические основы информатики»

Диссертация на соискание ученой степени доктора технических наук

Научный консультант: д.т.н., профессор Лемешко Борис Юрьевич

Новосибирск, 2013

Оглавление

ВВЕДЕНИЕ	7
ГЛАВА 1. МЕТОДИКА КОМПЬЮТЕРНОГО МОДЕЛИРОВАНИЯ В ПРИМЕНЕНИИ К ЗАДАЧАМ ПРОВЕРКИ СТАТИСТИЧЕСКИХ ГИПОТЕЗ	17
1.1 Метод Монте-Карло	17
1.2 Статистические гипотезы и критерии их проверки	18
1.3 Вычисление достигаемого уровня значимости	23
1.4 Аппроксимация функции распределения статистики критерия.....	26
1.5 Вычисление критических значений.....	27
1.6 Вычисление мощности критерия	28
1.7 Определение количества повторений.....	29
1.7.1 Определение количества повторений при оценивании вероятности наступления некоторого события.....	29
1.7.2 Погрешность моделирования функции распределения.....	32
1.7.3 Погрешность моделирования процентных точек.....	34
1.7.4 Погрешность моделирования мощности статистического критерия.....	35
1.7.5 Погрешность моделирования мощности статистического критерия при известном виде закона распределения статистики критерия	39
1.8 Исследование скорости сходимости распределения статистики критерия к предельному закону	41
1.8.1 Определение скорости сходимости	42
1.8.2 Алгоритм моделирования закона распределения статистики критерия при конечном объеме выборки.....	43
1.8.3 Аппроксимация расстояния до предельного закона степенной функцией.....	45
1.9 Моделирование псевдослучайных величин.....	46
1.9.1 Генераторы псевдослучайных величин.....	47
1.9.2 Моделирование псевдослучайных величин по непрерывному закону	48
1.10 Выводы	48
ГЛАВА 2. ОПТИМАЛЬНОЕ ПЛАНИРОВАНИЕ СТАТИСТИЧЕСКОГО ЭКСПЕРИМЕНТА ДЛЯ РАЗЛИЧЕНИЯ ДВУХ ПРОСТЫХ ГИПОТЕЗ	51
2.1 Простая гипотеза о виде распределения	52
2.2 Дивергенция Кульбака-Лейблера	53

2.3	Критерий отношения правдоподобия.....	54
2.4	Последовательный критерий отношения правдоподобия.....	56
2.5	Необходимый объем выборки для критерия отношения правдоподобия	60
2.6	Относительная эффективность критерия.....	65
2.7	Оптимальное группирование для различения двух простых гипотез.....	66
2.8	Связь между оптимальным группированием и мощностью критерия χ^2	68
2.9	Оптимальное планирование эксперимента по различению двух гипотез	69
2.10	Необходимый объем выборки и ошибки измерения	74
2.11	Выводы	76
ГЛАВА 3. СОКРАЩЕНИЕ СРЕДНЕГО ОБЪЕМА ВЫБОРКИ В ПОСЛЕДОВАТЕЛЬНЫХ КРИТЕРИЯХ.....		77
3.1	Критерий Вальда	78
3.2	Критерий Айвазяна.....	79
3.3	Критерий Лордена	81
3.4	Оценивание точных критических границ методом Монте-Карло.....	83
3.5	Средний объем выборки до принятия решения по последовательным критериям при использовании оценок точных и приближенных критических границ	90
3.6	Применение последовательных критериев к цензурированным наблюдениям.....	94
3.7	Вычисление критических границ для последовательного t-критерия.....	98
3.8	Вычисление критических границ при проверке сложных гипотез	100
3.9	Выводы	103
ГЛАВА 4. ИССЛЕДОВАНИЕ МОЩНОСТИ КРИТЕРИЕВ СОГЛАСИЯ ПРИ ПРОВЕРКЕ СЛОЖНЫХ ГИПОТЕЗ		105
4.1	Исследуемые критерии	107
4.1.1	Критерий Колмогорова.....	107
4.1.2	Критерии типа ω^2	108
4.2	Методы оценивания	109
4.2.1	Метод максимального правдоподобия.....	109
4.2.2	Методы минимального расстояния	110
4.2.3	Оценивание параметров по порядковым статистикам	110

4.3	Сравнение мощности критериев согласия для пары гипотез «Нормальное распределение против логистического».....	111
4.4	Мощность критериев для пары гипотез «распределение Вейбулла-Гнеденко против гамма-распределения»	116
4.5	Выводы	121
ГЛАВА 5. СРАВНИТЕЛЬНЫЙ АНАЛИЗ МОЩНОСТИ КРИТЕРИЕВ ОДНОРОДНОСТИ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ		124
5.1	Гипотеза однородности.....	125
5.1.1	Критерий Смирнова	125
5.1.2	Критерии типа ω^2	126
5.1.3	Критерий Лемана–Розенблатта	127
5.1.4	Критерий однородности Андерсона-Дарлинга-Петита	128
5.2	Сравнение мощности критериев	131
5.2.1	Мощность критерия Андерсона-Дарлинга-Петита	132
5.2.2	Мощность критерия однородности Смирнова	136
5.2.3	Мощность критерия Лемана-Розенблатта.....	138
5.3	Сравнение мощности критериев однородности по данным типа времени жизни..	140
5.4	Выводы	142
ГЛАВА 6. ИНВАРИАНТНЫЕ КРИТЕРИИ ПРОВЕРКИ ГИПОТЕЗЫ О МНОГОМЕРНОЙ НОРМАЛЬНОСТИ.....		143
6.1	Гипотеза о многомерной нормальности.....	144
6.2	Моделирование многомерной случайной величины	145
6.3	Инвариантные критерии проверки многомерной нормальности	146
6.3.1	Критерии, основанные на вычислении коэффициента асимметрии	147
6.3.2	Критерии, основанные на вычислении коэффициента эксцесса	160
6.3.3	Критерии многомерной нормальности, основанные на полярной декомпозиции вектора наблюдения случайной величины	169
6.4	Сходимость распределений статистик критериев многомерной нормальности к предельному закону	179
6.5	Исследование мощности критериев проверки многомерной нормальности	181
6.5.1	Конкурирующие гипотезы.....	181

6.5.2	Зависимость мощности критериев многомерной нормальности от размерности выборки	183
6.5.3	Анализ мощности инвариантных критериев многомерной нормальности	186
6.6	Выводы	194
ГЛАВА 7. ОПТИМАЛЬНОЕ ПЛАНИРОВАНИЕ ПРОВЕРКИ ГИПОТЕЗЫ ОДНОРОДНОСТИ ПРИ ПРОВЕДЕНИИ ДВУХЭТАПНОГО ПОЛНОГЕНОМНОГО АНАЛИЗА АССОЦИАЦИЙ.....		196
7.1	Полногеномный анализ ассоциаций	196
7.2	Одноэтапный эксперимент по выявлению ассоциации.....	198
7.3	Основная и конкурирующие гипотезы.....	198
7.4	Критерии ассоциаций.....	202
7.4.1	Критерий Хи-квадрат	202
7.4.2	Критерий отношения правдоподобия с ограничением (CLRT).....	202
7.4.3	Критерий тренда Кокрена-Армитеджа.....	203
7.4.4	MERT критерий	205
7.4.5	Критерий максимума (MAX3).....	206
7.4.6	Критерий минимума (MIN2)	207
7.4.7	Критерий выбора генетической модели (GMS)	207
7.4.8	Критерий Кульбака-Лейблера	208
7.5	Сравнительный анализ мощности критериев ассоциаций	209
7.6	Зависимость необходимого объема выборки от дивергенции Кульбака-Лейблера	211
7.7	Относительная эффективность критериев ассоциаций	219
7.8	Оптимальное планирование двухэтапного эксперимента.....	220
7.8.1	Моделирование вероятностей ошибок первого и второго рода	221
7.8.2	Оптимальное планирование двухэтапного эксперимента.....	223
7.8.3	Оптимальное дискретное планирование двухэтапного эксперимента	227
7.9	Выводы	228
ГЛАВА 8. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ СТАТИСТИЧЕСКОГО АНАЛИЗА		230
8.1	Программная система статистического анализа интервальных наблюдений одномерных непрерывных случайных величин "Интервальная статистика" (ISW)	231

8.2	Программное обеспечение статистического моделирования в задачах проведения и обработки измерений «НКЦ ИТР: Статистика 1.0».....	233
8.2.1	Программная платформа 1С:Предприятие 8.2.....	233
8.2.2	Объектная модель программной системы	234
8.2.3	Идентификация закона распределения	239
8.2.4	Проверка статистических гипотез	241
8.2.5	Выявление зависимостей и значимых факторов	242
8.2.6	Поддержка ГОСТов по статистике	243
8.3	Выводы	248
	ЗАКЛЮЧЕНИЕ.....	249
	СПИСОК ЛИТЕРАТУРЫ.....	251
	ПРИЛОЖЕНИЕ 1. АКТ О ВНЕДРЕНИИ В УЧЕБНЫЙ ПРОЦЕСС.....	286
	ПРИЛОЖЕНИЕ 2. СВИДЕТЕЛЬСТВА О РЕГИСТРАЦИИ ПРОГРАММЫ ДЛЯ ЭВМ «ИНТЕРВАЛЬНАЯ СТАТИСТИКА» ISW	287
	ПРИЛОЖЕНИЕ 3. СВИДЕТЕЛЬСТВА О РЕГИСТРАЦИИ ПРОГРАММЫ ДЛЯ ЭВМ «НКЦ ИТР: СТАТИСТИКА 1.0».....	290
	ПРИЛОЖЕНИЕ 4. ГОСТ Р 50.1.033-2001 (ФРАГМЕНТ).....	293
	ПРИЛОЖЕНИЕ 5. ГОСТ Р 50.1.037-2002 (ФРАГМЕНТ).....	296

ВВЕДЕНИЕ

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 удастся 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 проследивать 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. Соболев, 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-Дарлингга;
- в построении точных критических границ последовательных критериев Вальда, Айвазяна и Лордена;
- в результатах сравнительного анализа мощности критерия однородности Андерсона-Дарлингга-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 МАХЗ;

- в оценке относительной эффективности критерия МАХЗ 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 [24] и Р 50.1.037-2002 [25]. Рекомендации введены в действие постановлениями Госстандарта с 01.07.2002.

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. На защиту выносятся:

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 KBM-2011 (Новосибирск), 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.В37.21.0860), 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 А/11/76161).

Основные результаты исследований по теме диссертации опубликованы в 106 печатных работах [5, 32, 37, 38, 43-47, 49-61, 63-72, 74-85, 87, 89, 92-93, 95-99, 101, 107-149, 154, 172, 206-221, 242-246] общим объемом 113 п.л., в том числе 23 статьи в рецензируемых научных журналах (из них 19 статей в журналах, входящих в перечень изданий, рекомендуемых ВАК РФ), монография, рекомендации по стандартизации [24, 25], учебное пособие с главами научного содержания [89], 78 публикаций в сборниках научных работ, трудах и материалах научных конференций. Получены пять свидетельств о государственной регистрации программ для ЭВМ [145, 146, 147, 148, 149].

В опубликованных работах автору принадлежат результаты, изложенные в тексте диссертации.

Структура работы. Диссертация состоит из введения, 8 глав основного содержания, заключения, списка литературы и приложений. Основное содержание представлено на 285 страницах, включая 58 таблиц, 82 рисунков и списка литературы из 281 источника.

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.

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.

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, найденных методом компьютерного моделирования.

В четвертой главе рассмотрены непараметрические критерии 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 пятой главе рассмотрены критерии проверки однородности распределений Смирнова, Лемана-Розенблатта и Андерсона-Дарлинга-Петита. Проведено исследование скорости сходимости распределения статистики критерия Андерсона-Дарлинга-Петита к предельному закону. Выполнено сравнение мощностей критериев однородности. Проведено сравнение мощностей критериев однородности на данных типа времени жизни для альтернатив с пересечениями и без пересечений.

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 мощностей инвариантных критериев многомерной нормальности.

В седьмой главе рассмотрены критерии проверки однородности выборки случаев и контрольной выборки при проведении полногеномного анализа ассоциаций. Проведено сравнение мощностей 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 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.

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 [24] и Р 50.1.037-2002 [25].

Автор выражает глубокую признательность научному консультанту д.т.н. профессору Б.Ю. Лемешко за помощь и ценные советы при выполнении диссертационной работы.

ГЛАВА 1. МЕТОДИКА КОМПЬЮТЕРНОГО МОДЕЛИРОВАНИЯ В ПРИМЕНЕНИИ К ЗАДАЧАМ ПРОВЕРКИ СТАТИСТИЧЕСКИХ ГИПОТЕЗ

1.1 Метод Монте-Карло

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 основан на получении большого числа реализаций стохастического (случайного) процесса.

Идея метода довольно проста, и ее история тесно связана с появлением теории вероятностей в XVII веке. Классическим примером этого метода является алгоритм вычисления числа π , предложенный Бюффеном в 1777 году [19]. Однако по-настоящему доступным этот метод стал только после появления первых компьютеров, т.к. моделирование случайных чисел вручную (например, в алгоритме Бюффона – это подбрасывание иглы) является весьма трудоемким.

Современное название этого метода – «Метод Монте-Карло» появилось в группе ученых (Дж. фон Нейман, С. Улам, Н. Метрополис, Г. Кан, Э. Ферми и др.), работавших над созданием атомных реакторов в г. Лос-Алмос (США) в 1945 году [229]. Поскольку работа над проектом была засекречена, то впервые официальное название этого метода появилось в 1949 году, когда вышла статья Н. Метрополиса и С. Улама «Метод Монте-Карло» [228], в которой были сформулированы основные принципы этого метода.

В частности, они пишут: «Вычисление вероятности успешного раскладывания пасьянса (мы подразумеваем здесь только такие игры, где умение игрока не играет никакой роли) является трудноразрешимой задачей. С другой стороны законы больших чисел и асимптотические теоремы теории

вероятностей не слишком проливают свет даже на качественные вопросы, касающиеся такой вероятности. Очевидно, что практическая процедура может состоять в том, чтобы произвести большое число реализаций заданной игры и затем вычислить относительную долю успехов. Пасьянс в данном примере является только иллюстрацией целого класса комбинаторных задач, встречающихся как в чистой, так и в прикладной науке. Мы можем показать сразу, что оценка никогда не будет ограничена заданными пределами с полной определенностью, но только (если число испытаний большое) – с большой вероятностью. Чтобы установить это, мы должны обратиться к законам больших чисел и другим результатам теории вероятностей» [228].

С помощью метода Монте-Карло решается множество задач в математике, физике, химии, экономике, социологии и других науках [181, 153, 35, 36, 189, 6, 33, 100]. В настоящей работе метод Монте-Карло применяется для решения теоретических задач математической статистики, в частности в применении к задачам проверки статистических гипотез.

В данной главе изложена методика применения метода Монте-Карло к задачам проверки статистических гипотез, основанная на многолетнем опыте компьютерного моделирования при исследованиях критериев проверки статистических гипотез [32, 67, 89, 93, 97].

1.2 Статистические гипотезы и критерии их проверки

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 независимости и одинаковом распределении наблюдений в

выборке; *о виде распределения*, если есть предположения о законе распределения случайной величины (рисунок 1.1).

Проверка статистической гипотезы состоит в том, чтобы сформулировать такое правило, которое позволило бы по результатам проведенных наблюдений принять или отклонить гипотезу. Правило, согласно которому гипотеза принимается или отвергается, называется *критерием* проверки статистической гипотезы.

С проверкой статистических гипотез связывают ошибки двух типов. *Ошибкой первого рода* называют событие, когда верная проверяемая гипотеза отвергается критерием. *Ошибкой второго рода* называют событие, когда неверная проверяемая гипотеза принимается критерием. Вероятности ошибок первого и второго рода обозначают α и β , соответственно. Вероятность ошибки второго рода зависит от выдвигаемой конкурирующей гипотезы. Вероятность отклонения ложной проверяемой гипотезы, т.е. принятия правильного решения в пользу конкурирующей, называется *мощностью*, и она равна $1 - \beta$. Вероятность ошибки первого рода также называют *уровнем значимости* критерия.

Есть небольшая терминологическая тонкость в использовании выражения «гипотеза принимается критерием». Часто, вместо этого говорят «гипотеза не отвергается критерием», понимая, что по выборке большего объема гипотеза может быть отвергнута этим же критерием, или же говорят «нет оснований для отвержения гипотезы по данной выборке». В дальнейшем, для упрощения текста будет использоваться самый краткий вариант.

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, \dots, H_m .

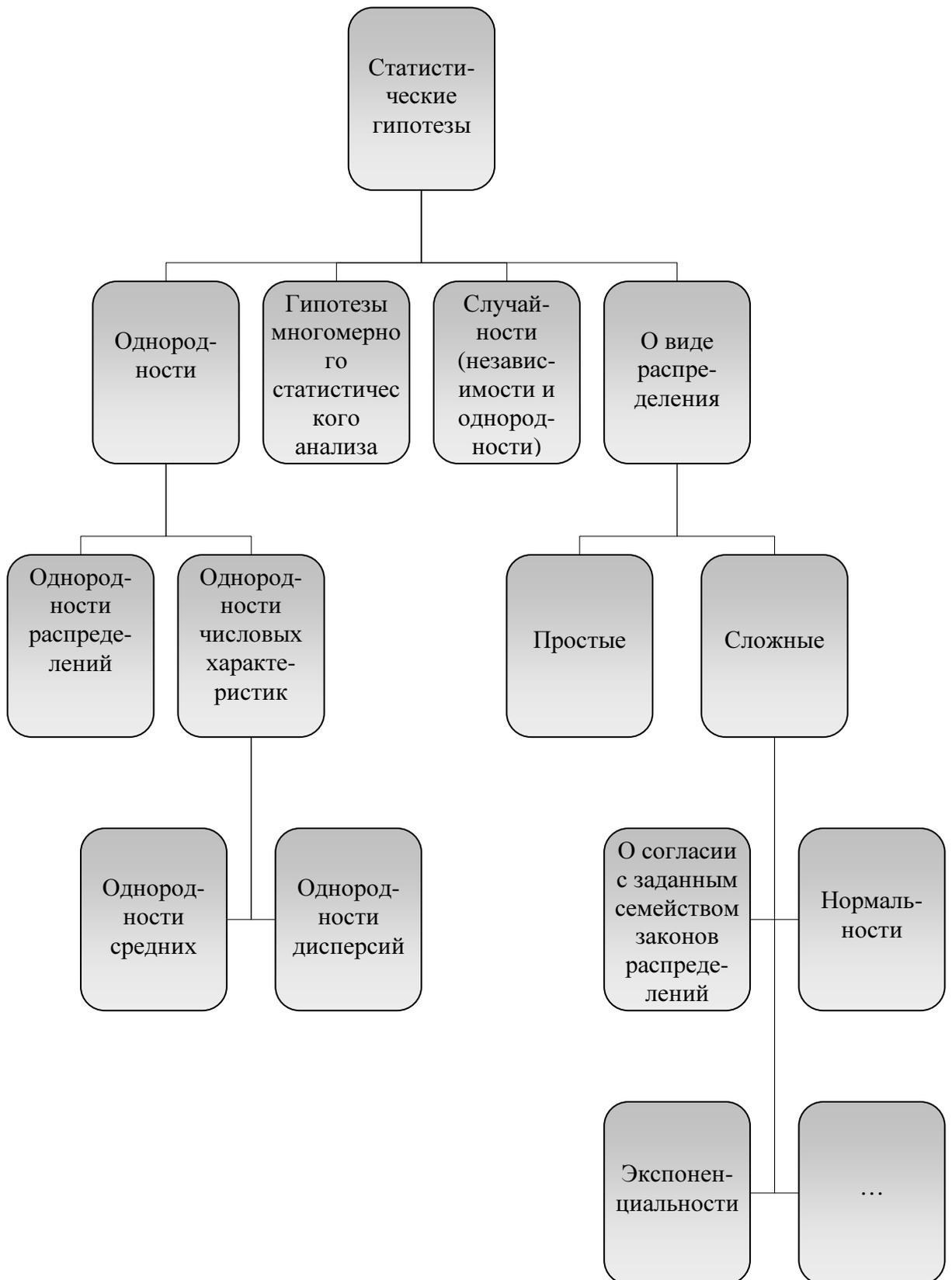


Рисунок 1.1 – Классификация статистических гипотез

Для проверки одной и той же гипотезы, как правило, существует несколько различных статистических критериев T_1, T_2, \dots, 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). Так, например, нельзя применять критерий Колмогорова по группированным данным, или по наблюдениям дискретной случайной величины.
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.

Добиться выполнения последнего принципа на практике не представляется возможным, потому что построить наиболее мощный критерий удастся только в очень редких случаях, например, когда основная и конкурирующая гипотезы являются *простыми* [39]. Чаще всего, для разных конкурирующих гипотез, для разных уровней значимости, для разных объемов выборки, более мощными оказываются разные критерии.

В этой ситуации для выбора оптимального критерия можно применить классическую теорию *принятия решений в условиях неопределенности* [271]. Критерии являются стратегиями, конкурирующие гипотезы – состояниями среды, функция полезности $u(T_i, H_j)$ – это мощность критерия (таблица 1.1). Другим способом определения функции полезности при выборе критерия может быть стоимость проведения эксперимента по различению основной и

конкурирующей гипотез с заданными вероятностями ошибок первого и второго рода.

Таблица 1.1 – Матрица полезности выбора критерия T_i при конкурирующих гипотезах H_j

$T_i \setminus H_j$	H_1	H_2	...	H_m
T_1	$u(T_1, H_1)$	$u(T_1, H_2)$...	$u(T_1, H_m)$
T_2	$u(T_2, H_1)$	$u(T_2, H_2)$...	$u(T_2, H_m)$
...
T_k	$u(T_k, H_1)$	$u(T_k, H_2)$...	$u(T_k, H_m)$

Существуют разные подходы к выбору оптимальной стратегии при принятии решения в условиях неопределенности. В случае, когда нет никакой информации о том, какая конкурирующая гипотеза может быть верна, рациональным выглядит выбор критерия по правилу Вальда [271] (известны также такие названия как «критерий крайнего пессимиста» или «критерий осторожного наблюдателя»):

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

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

Любой критерий проверки статистической гипотезы разбивает выборочное пространство на доверительную область \mathcal{R}_0 и критическую область \mathcal{R}_1 . При попадании выборки в критическую область гипотеза отвергается, а при попадании в доверительную область – принимается. Чаще всего такое разбиение производится с помощью одномерной *статистики* – функции от выборки, поэтому критическая и доверительная область формулируются уже как подмножества множества вещественных чисел.

Доверительная область включает такие значения статистики критерия, при которых гипотеза принимается, а критическая область – значения, при которых гипотеза отвергается. Кроме того, вероятность попадания выборки (статистики критерия) в критическую область, когда гипотеза верна, по определению равна вероятности ошибки первого рода, а вероятность попадания выборки (статистики критерия) в доверительную область, когда гипотеза не верна, равна вероятности ошибки второго рода.

Как правило, встречаются три вида критических областей для статистики критерия:

- правосторонняя критическая область (t_α, ∞) ;
- левосторонняя $(-\infty, t_\alpha)$;
- двусторонняя $(-\infty, t_{\alpha_1}) \cup (t_{\alpha_2}, \infty)$, $\alpha_1 + \alpha_2 = \alpha$.

1.3 Вычисление достигаемого уровня значимости

Достижимый уровень значимости (p-value) определяется как вероятность попадания статистики критерия:

- в область $(S(X_n), \infty)$, если критическая область правосторонняя;
 - в область $(-\infty, S(X_n))$, если критическая область левосторонняя;
- где $S(X_n)$ - вычисленное значение статистики по реализации выборки.

Гипотеза отвергается, если достигаемый уровень значимости оказывается меньше заданной вероятности ошибки первого рода. Достоинство процедуры проверки гипотезы с использованием p -value в том, что не нужно заранее фиксировать уровень значимости и определять критическую область для значений статистики критерия. Кроме того, p -value характеризует “степень уверенности” в принимаемом решении, т.е. чем меньше p -value, тем больше оснований для отвержения основной гипотезы.

Если критическая область двусторонняя, то однозначного способа вычисления достигаемого уровня значимости нет. Например, можно вычислять p -value как $2 \cdot \min(p, 1 - p)$, где $p = P\{S \in (S(X_n), \infty) | H_0\}$.

Достижимый уровень значимости является случайной величиной, определенной на интервале $[0,1]$, на основании которой делается статистический вывод о принятии гипотезы. Чем ближе значение p -value к 1, тем больше оснований для принятия гипотезы, чем ближе значение p -value к 0, тем больше оснований для отвержения гипотезы. Однако следует помнить о следующем важном замечании относительно p -value.

Когда основная гипотеза ложна, то p -value будет стремиться к 0 с ростом объема наблюдаемой выборки. Однако, когда основная гипотеза истинна, p -value не стремится к 1, а распределено равномерно на интервале $[0,1]$.

Вычисление 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). Очень важным для корректного вычисления p -value (с требуемой точностью) является вопрос о правильном выборе числа повторений N , который будет рассмотрен в пункте 1.7.

Алгоритм 1.1. Вычисление достигаемого уровня значимости статистического критерия с правосторонней или левосторонней критической областью методом Монте-Карло.

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

Действия.

1. Вычислить $S = S(X_n)$ – статистику критерия по выборке.
2. Установить $m = 0$.
3. Сгенерировать выборку Y_n при верной гипотезе H_0 .
4. Вычислить значение $S(Y_n)$.
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;
- $\hat{p} = 2 \cdot \min\left(\frac{m}{N}, 1 - \frac{m}{N}\right)$ для двусторонней критической области.

Применение метода Монте-Карло для вычисления достигаемого уровня значимости удобно, если нам нужно проверить гипотезу однократно, по одной выборке. Однако если нужно проверить гипотезу по нескольким выборкам, то алгоритм 1.1 будет неэффективным, т.к. достаточно трудоемкая процедура по моделированию статистики критерия будет повторяться много раз. Чтобы сократить затраты на моделирования возможны следующие варианты.

1. Сохранить выборку статистик $S(Y_n)$ в массиве или во внешнем файле. Тогда p-value можно вычислить по формуле $1 - G_N(S(X_n))$ для

правосторонней критической области или по формуле $G_N(S(X_n))$ для левосторонней критической области, где $G_N(t)$ – эмпирическая функция распределения статистики критерия.

2. Подобрать подходящую вероятностную модель $G(t)$, хорошо аппроксимирующую $G_N(t)$. Тогда p -value можно вычислить по формуле $1 - G(S(X_n))$ для правосторонней критической области или по формуле $G(S(X_n))$ для левосторонней критической области.

1.4 Аппроксимация функции распределения статистики критерия

Для аппроксимации функции распределения статистик используется процедура подбора закона распределения в системе ISW [89, 97], которая включает более 30 стандартных законов и распределений, получаемых с помощью операций над этими стандартными моделями: операций сдвига, масштаба, смеси законов, произведения, усечения, логарифмирования.

После визуального анализа эмпирической функции распределения статистики критерия, делается отбор вероятностных моделей похожих по форме на искомое распределение. Далее для каждой модели выполняется идентификация – оценивание параметров модели и проверка гипотезы о согласии, и выбирается наилучшая модель.

Даже если подходящей модели в системе ISW не оказалось, можно разбить область определения на три области – левый «хвост», середину и правый «хвост», и попробовать аппроксимировать функцию распределения на каждой области по-отдельности разными моделями. Для этого выборка цензурируется (для левого «хвоста» – справа, для правого «хвоста» – слева, а для середины – с обеих сторон), по такой цензурированной выборке подбирается закон распределения по описанной выше процедуре.

По описанной технологии, в частности, были построены аппроксимации законов распределения непараметрических критериев согласия при проверке сложных гипотез [220].

1.5 Вычисление критических значений

Критическое значение (иногда называемое процентной точкой) может быть вычислено по формуле $t_\alpha = G^{-1}(\alpha)$ для левосторонней критической области или $t_\alpha = G^{-1}(1 - \alpha)$ для правосторонней критической области, где $G(t)$ – функция распределения статистики критерия.

Если функция распределения неизвестна, но имеется смоделированная по методу Монте-Карло выборка статистик S_1, S_2, \dots, S_N , упорядоченная по возрастанию, то критическое значение t_α можно найти по формуле для вычисления выборочного квантиля порядка q . Выборочная квантиль порядка q – это член вариационного ряда с номером $[Nq] + 1$, если Nq – не целое число, и полусумма членов вариационного ряда с номерами Nq и $Nq + 1$, если число Nq – целое. Таким образом,

$$t_\alpha = \begin{cases} \frac{1}{2}(S_{[qN]} + S_{[qN]+1}), & qN \in \mathbb{N} \\ S_{[qN]+1}, & qN \notin \mathbb{N} \end{cases} \quad (1.1)$$

где $q = \alpha$ для левосторонней критической области, и $q = 1 - \alpha$ для правосторонней критической области. В данной записи $[x]$ обозначает целую часть от числа x , а \mathbb{N} – это множество натуральных чисел.

Выборочная квантиль является состоятельной оценкой теоретической квантили, если функция распределения случайной величины является непрерывной и строго монотонной в точке теоретической квантили [39].

1.6 Вычисление мощности критерия

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, и оно, естественно, 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 $S(Y_n)$.
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

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

1.7 Определение количества повторений

В рассмотренных алгоритмах 1.1 и 1.2 вычисления достигаемого уровня значимости и мощности имеет параметр – число повторений, от которого зависит, насколько точно мы найдем оценки.

Выбор требуемого количества повторений производится с помощью построения доверительного интервала для искомой вероятности. Половину длины этого доверительного интервала будем называть *абсолютной погрешностью моделирования*.

Оценка точности характеристик и необходимое число реализаций (повторений) в методе Монте-Карло приводится в книге Вентцель Е.С. [8], причем рассматривается простейший случай, когда с помощью метода Монте-Карло оценивается вероятность наступления некоторого события. В пунктах 1.7.2-1.7.4 будет рассмотрена задача определения количества повторений при оценивании функции распределения, процентных точек и мощности статистического критерия.

1.7.1 Определение количества повторений при оценивании вероятности наступления некоторого события

Пусть требуется вычислить вероятность p появления некоторого случайного события A . В каждой из N реализаций процесса количество наступлений события A является случайной величиной ξ , принимающей значение $x_1 = 1$ с вероятностью p , и значение $x_2 = 0$ с вероятностью $1 - p$. Случайная

величина ξ подчиняется распределению Бернулли с вероятностью успеха p и имеет математическое ожидание p и дисперсию $p(1-p)$.

В качестве оценки для искомой вероятности p принимается частота m/N наступлений события A при N реализациях

$$\frac{m}{N} = \frac{1}{N} \sum_{i=1}^N x_i,$$

где x_i – количество наступлений события A в реализации с номером i .

В силу центральной предельной теоремы теории вероятностей частота при достаточно больших N имеет распределение, близкое к нормальному:

$$\sqrt{N} \frac{m/N - M\xi}{\sqrt{D\xi}} = \sqrt{N} \frac{m/N - p}{\sqrt{p(1-p)}} \rightarrow \eta \in N(0,1).$$

Отсюда

$$\mathbb{P} \left\{ \left| \sqrt{N} \frac{m/N - p}{\sqrt{p(1-p)}} \right| < t_\gamma \right\} = \Phi(t_\gamma) - \Phi(-t_\gamma) = 2\Phi(t_\gamma) - 1 = \gamma,$$

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

$$\mathbb{P} \left\{ |m/N - p| < t_\gamma \frac{\sqrt{p(1-p)}}{\sqrt{N}} \right\} = \gamma.$$

Таким образом, погрешность моделирования p с доверительной вероятностью γ равна

$$\varepsilon = t_\gamma \frac{\sqrt{p(1-p)}}{\sqrt{N}}. \quad (1.2)$$

Отсюда количество повторений, для которого погрешность моделирования не превышает ε , равно

$$N = \left[t_{\gamma}^2 \frac{p(1-p)}{\varepsilon^2} \right] + 1, \quad (1.3)$$

где $[\cdot]$ означает целую часть, так как количество повторений должно быть натуральным числом.

Погрешность моделирования зависит от величины p и достигает максимального значения при $p = 0,5$. При $p \rightarrow 0$ и при $p \rightarrow 1$ погрешность стремится к нулю. Для значений p близких к нулю или единице имеет смысл рассматривать *относительную* погрешность моделирования

$$\delta = \frac{\varepsilon}{p(1-p)} \quad (1.4)$$

Величина $\delta_0 = \frac{\varepsilon}{p}$ показывает, во сколько раз погрешность моделирования превосходит вероятность p , близкую к 0. Величина $\delta_1 = \frac{\varepsilon}{1-p}$ показывает, во сколько раз погрешность моделирования превосходит вероятность $1-p$, когда p близко к 1. Формула (1.4) обобщает эти две ситуации.

Тогда формулу (1.3) можно переписать для относительной погрешности в виде

$$N = \left[t_{\gamma}^2 \frac{1}{p(1-p)\delta^2} \right] + 1. \quad (1.5)$$

Для надежного оценивания вероятностей желательно, чтобы относительная погрешность не превышала 5%. В таблице 1.2 приведены рекомендуемые количества повторений для разных значений вероятности p при относительной погрешности 5%.

Таблица 1.2 – Требуемое количество повторений при относительной погрешности моделирования 5%

p	0,15	0,1	0.05	0,01	10^{-3}	10^{-4}	10^{-5}
N	8082	11449	21692	10^5	10^6	10^7	10^8

1.7.2 Погрешность моделирования функции распределения

Пусть требуется определить функцию распределения $F_{\xi}(x)$ некоторой случайной величины ξ . В качестве непараметрической оценки распределения можно использовать эмпирическую функцию распределения:

$$F_N(x) = \begin{cases} 0, & x < x_{(1)} \\ m/N, & x_{(m)} \leq x < x_{(m+1)}, m = 1, 2, \dots, N-1, \\ 1, & x \geq x_{(N)}, \end{cases}$$

где $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(N)}$ – вариационный ряд, построенный по выборке.

Для любого фиксированного значения x функция $F_N(x)$ представляет собой дискретную случайную величину, принимающую значения $0, 1/N, 2/N, \dots, 1$. Математическое ожидание и дисперсия $F_N(x)$ равны:

$$M[F_N(x)] = F(x), \quad D[F_N(x)] = \frac{F(x)(1-F(x))}{N}.$$

Тогда формулы (1.2) и (1.3) для вычисления значения $p = F(x)$ с помощью случайной величины $m/N = F_N(x)$ примут вид

$$\varepsilon = t_{\gamma} \frac{\sqrt{F(x)(1-F(x))}}{\sqrt{N}}, \quad N = \left[t_{\gamma}^2 \frac{F(x)(1-F(x))}{\varepsilon^2} \right] + 1, \quad (1.6)$$

где $[\cdot]$ означает целую часть числа.

Абсолютная погрешность моделирования зависит от величины $F(x)$ и достигает максимума при $F(x)=0,5$. Поэтому можно найти оценку сверху абсолютной погрешности и необходимого количества повторений:

$$\varepsilon \leq t_\gamma \frac{1}{2\sqrt{N}}, N \leq \left[t_\gamma^2 \frac{1}{4\varepsilon^2} \right] + 1. \quad (1.7)$$

Так, например, для доверительной вероятности $\gamma=0,99$ необходимое число повторений равно для $\varepsilon = 0,001$ равно 1 658 725.

Отметим, что на практике при проверке статистических гипотез часто опираются на значения *процентных точек* (квантилей) распределения статистики критерия, то есть на такие значения x , при которых $F(x)=0,85; 0,90; 0,95; 0,99; 0,999$. Абсолютная и относительная погрешности моделирования $F(x)$ при разных количествах повторений приведены в таблицах 1.3 и 1.4, соответственно.

Как и следовало ожидать, погрешности моделирования $F(x)$ в этих точках при достаточно больших количествах повторений малы. Однако какова при этом будет погрешность моделирования самих процентных точек?

Таблица 1.3 – Абсолютная погрешность моделирования $F(x)$

$F(x)$	$F(x)(1-F(x))$	$\varepsilon_N,$ $N = 1000$	$\varepsilon_N,$ $N = 10\ 000$	$\varepsilon_N,$ $N = 100\ 000$
0,850	0,128	0,0291	0,0092	0,0029
0,900	0,090	0,0244	0,0077	0,0024
0,950	0,048	0,0178	0,0056	0,0018
0,990	0,010	0,0081	0,0026	0,0008
0,999	0,001	0,0026	0,0008	0,0003

Таблица 1.4 – Относительная погрешность моделирования $F(x)$

$F(x)$	$F(x)(1-F(x))$	$\delta_N,$ $N = 1000$	$\delta_N,$ $N = 10\ 000$	$\delta_N,$ $N = 100\ 000$
0,850	0,128	23%	7%	2%
0,900	0,090	27%	9%	3%
0,950	0,048	37%	12%	4%
0,990	0,010	82%	26%	8%
0,999	0,001	258%	81%	26%

1.7.3 Погрешность моделирования процентных точек

В качестве оценки процентной точки порядка p используется выборочная квантиль, вычисляемая по формуле (1.1). Известно [39], что данная статистика в асимптотике имеет нормальное распределение с математическим ожиданием и дисперсией

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

где $f(x)$ – функция плотности. Тогда формулы (1.2) и (1.3) для оценивания значения $F^{-1}(p)$ с помощью выборочной квантили $x_{(pN)}$ примут вид

$$\varepsilon = t_\gamma \frac{\sqrt{p(1-p)}}{f(F^{-1}(p))\sqrt{N}}, \quad N = \left[t_\gamma^2 \frac{p(1-p)}{\varepsilon^2 f^2(F^{-1}(p))} \right] + 1. \quad (1.8)$$

Естественно, что на величину погрешности в (1.8) существенный вклад оказывает вид функции распределения случайной величины ξ . Вычислим погрешности при определении процентных точек методом Монте-Карло в предположении, что случайная величина ξ имеет экспоненциальное распределение с параметром масштаба λ :

$$F(x, \lambda) = 1 - \exp(-x/\lambda), \quad x > 0;$$

$$f(x, \lambda) = \frac{1}{\lambda} \exp(-x/\lambda), \quad x > 0.$$

В этом случае

$$f(F^{-1}(p)) = \frac{1-p}{\lambda}.$$

Тогда получаем, что

$$\varepsilon = t_\gamma \frac{\sqrt{\lambda p}}{\sqrt{N(1-p)}}, \quad N = t_\gamma^2 \frac{\lambda p}{(1-p)\varepsilon^2}. \quad (1.9)$$

Абсолютная и относительная погрешности моделирования $F^{-1}(p)$, при $\lambda = 1$, и разных N приведены в таблицах 1.5 и 1.6 соответственно. Как видно из таблиц 1.3 и 1.5, в то время как абсолютная погрешность моделирования функции распределения при заданном N уменьшается с приближением p к единице, погрешность моделирования процентных точек увеличивается.

Если сравнить таблицы 1.4 и 1.6, то видно, что и относительная погрешность моделирования функции распределения, и относительная погрешность моделирования процентных точек растут с приближением p к единице, причем порядок погрешностей примерно одинаков.

Таблица 1.5 – Абсолютная погрешность моделирования $F^{-1}(p)$ для экспоненциального закона распределения

p	$F^{-1}(p)$	$\varepsilon_N, N = 1000$	$\varepsilon_N, N = 10\,000$	$\varepsilon_N, N = 100\,000$
0.850	0,82	0,194	0,061	0,019
0.900	1,00	0,244	0,077	0,024
0.950	1,30	0,355	0,112	0,036
0.990	2,00	0,810	0,256	0,081
0.999	3,00	2,575	0,814	0,257

Таблица 1.6 – Относительная погрешность моделирования $F^{-1}(p)$ для экспоненциального закона распределения

p	$F^{-1}(p)$	$\delta_N, N = 1000$	$\delta_N, N = 10\,000$	$\delta_N, N = 100\,000$
0.850	0,82	24%	7%	2%
0.900	1,00	24%	8%	2%
0.950	1,30	27%	9%	3%
0.990	2,00	41%	13%	4%
0.999	3,00	86%	27%	9%

1.7.4 Погрешность моделирования мощности статистического критерия

При оценивании 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 еще более сложная задача, чем оценка погрешностей при вычислении процентных точек. Поэтому для оценки погрешности вычисления мощности воспользуемся методом Монте-Карло.

Пусть статистика критерия при верной гипотезе H_0 имеет экспоненциальное распределение с параметром масштаба 1, а при верной гипотезе H_1 – экспоненциальное распределение с параметром 4. Критерий имеет правостороннюю критическую область. В данном случае критическое значение можно вычислить по формуле $t_\alpha = -\ln \alpha$, а мощность критерия равна $1 - \beta = \alpha^{1/4}$.

Для $N = 1000, 10000, 100000$ мощность критерия была вычислена 100000 раз по алгоритму 1.2. Далее по эмпирической функции распределения $F_n(t)$ была вычислена половина длины 99% доверительного интервала по формуле

$$\varepsilon_N = \frac{F_n^{-1}(0,995) - F_n^{-1}(0,005)}{2}.$$

Оценки абсолютной погрешности моделирования мощности приведены в таблице 1.7, относительной погрешности – в таблице 1.8. Погрешность моделирования мощности превосходит погрешности, соответствующие погрешностям моделирования вероятностей α и $1 - \beta$ (см. таблицу 1.9), причем погрешность растет при приближении α к нулю.

Следует также отметить, что распределение погрешности моделирования мощности хорошо аппроксимируется нормальным законом (рисунок 1.2). Из-за сильной дискретности распределения при небольшом N гипотеза о согласии с нормальным распределением отвергается (таблица 1.10), однако при $N = 100000$ гипотеза о согласии не отвергается. Таким образом, для оценивания мощности можно выполнять моделирование сериями с последующим вычислением доверительного интервала.

Еще один недостаток использования небольшого N в том, что оценка мощности в этом случае имеет небольшое смещение. При $N = 1000$ смещение составляет примерно 0,001. С ростом N смещение уменьшается.

Поскольку погрешность моделирования мощности сильно зависит от формы распределений статистики критерия при верной основной и конкурирующей гипотезах, то дать общую рекомендацию о выборе числа повторений при оценке мощности невозможно. Пожалуй, лучше всего вычислить мощность критерия для одной пары конкурирующих гипотез несколько раз, затем полученные значения усреднить и оценить погрешность моделирования. При этом можно предполагать, что распределение погрешностей подчиняется нормальному закону.

Количество повторений при моделировании распределения статистики критерия при верной гипотезе H_0 не обязательно должно равняться количеству повторений при моделировании распределения статистики критерия при верной гипотезе H_1 . Так как *большая* погрешность связана с оцениванием процентной точки t_α , то *большее* количество повторений следует делать при моделировании распределения статистики критерия при верной основной гипотезе.

Таблица 1.7 – Абсолютная погрешность моделирования мощности критерия

α	$1 - \beta$	$\varepsilon_N,$ $N = 1000$	$\varepsilon_N,$ $N = 10\ 000$	$\varepsilon_N,$ $N = 100\ 000$
0,15	0,622333	0,0495	0,0157	0,0050
0,10	0,562341	0,0540	0,0168	0,0053
0,05	0,472871	0,0580	0,0185	0,0059
0,01	0,316228	0,0735	0,0235	0,0075
0,001	0,177828	0,1035	0,0368	0,0118

Таблица 1.8 – Относительная погрешность моделирования мощности критерия

α	$1 - \beta$	δ_N , $N = 1000$	δ_N , $N = 10\,000$	δ_N , $N = 100\,000$
0,15	0,622333	8,0%	2,5%	0,8%
0,10	0,562341	9,6%	3,0%	0,9%
0,05	0,472871	12,3%	3,9%	1,2%
0,01	0,316228	23,2%	7,4%	2,4%
0,001	0,177828	58,2%	20,7%	6,6%

Таблица 1.9 – Сравнение абсолютной погрешности моделирования вероятностей ошибок первого и второго рода и мощности критерия

α	$1 - \beta$	ε_N для α , $N = 10\,000$	ε_N для $1 - \beta$, $N = 10\,000$	ε_N , $N = 10\,000$
0,15	0,622333	0,0092	0,0125	0,0157
0,10	0,562341	0,0077	0,0128	0,0168
0,05	0,472871	0,0056	0,0129	0,0185
0,01	0,316228	0,0026	0,0120	0,0235
0,001	0,177828	0,0008	0,0098	0,0368

Таблица 1.10 – Согласие распределения погрешности моделирования мощности с нормальным законом при $\alpha = 0.05$ и разных объемах моделирования

N	Закон распределения	Достижимый уровень значимости		
		Критерий Колмогорова	Критерий Крамера-Мизеса-Смирнова	Критерий Андерсона-Дарлинга
1000	Norm(0.4720,0.0226)	0	0	0
10000	Norm(0.4728,0.0072)	0.002	0.011	0.012
100000	Norm(0.4729,0.0023)	0.067	0.256	0.319

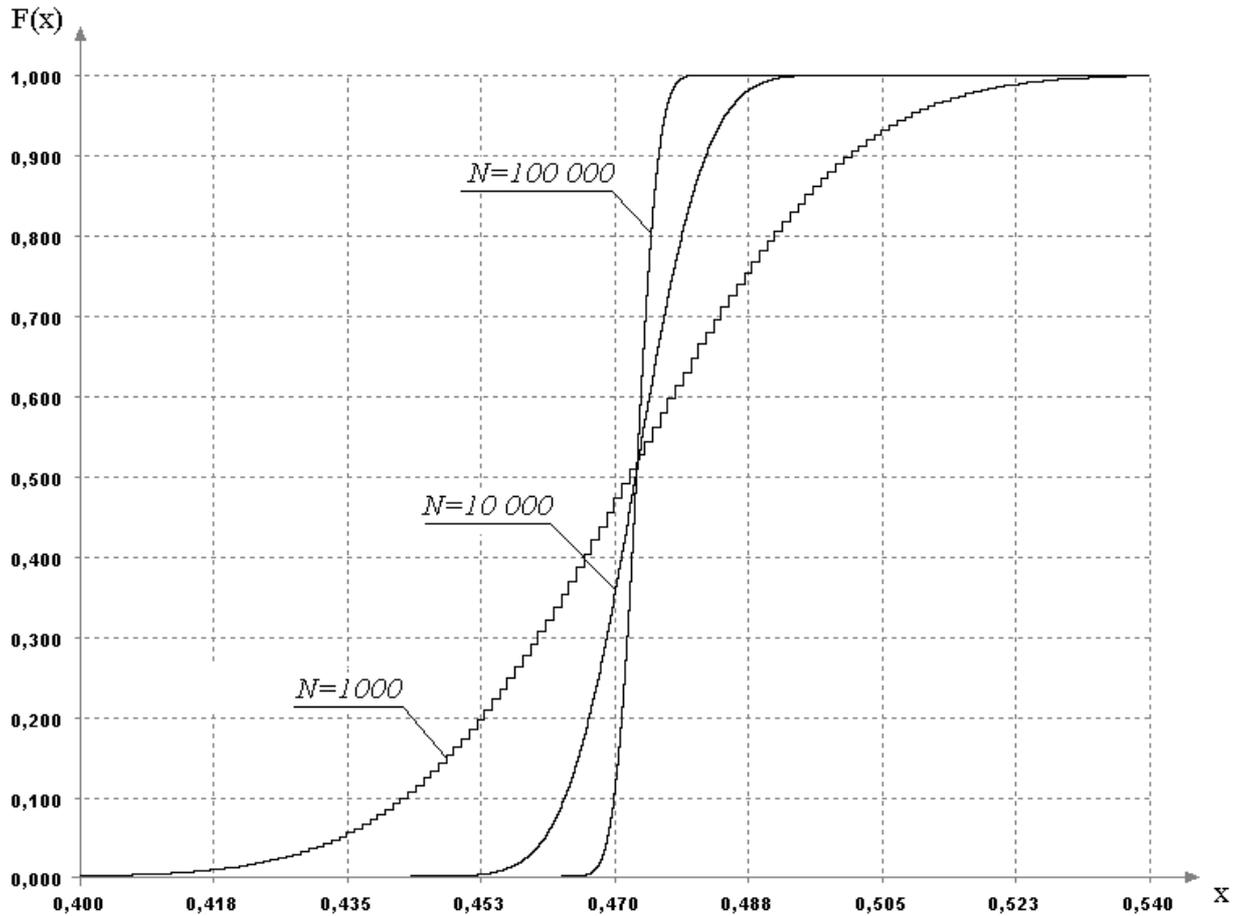


Рисунок 1.2 – Распределения погрешности моделирования мощности

1.7.5 Погрешность моделирования мощности статистического критерия при известном виде закона распределения статистики критерия

Если известен закон распределения, которому подчиняется статистика критерия, но неизвестны параметры этого закона, то можно существенно уменьшить погрешность моделирования. Рассмотрим эту ситуацию на следующем примере.

Пусть статистика критерия при верной гипотезе H_0 имеет экспоненциальное распределение с параметром масштаба λ_0 , а при верной гипотезе H_1 – экспоненциальное распределение с параметром λ_1 . Критерий имеет правостороннюю критическую область. Эффективную оценку параметра масштаба экспоненциального закона можно найти по формуле

$$\hat{\lambda} = \frac{1}{N} \sum_{i=1}^N S_i,$$

где S_i – это значения статистики критерия при моделировании.

Тогда критическое значение можно оценить по формуле $\hat{t}_\alpha = -\hat{\lambda}_0 \ln \alpha$, а мощность критерия оценивается как

$$1 - \hat{\beta} = \alpha^{\frac{\hat{\lambda}_0}{\hat{\lambda}_1}}.$$

Результаты моделирования представлены на рисунке 1.3 для $\alpha=0,05$. Видно, что при оценивании параметров закона распределения статистики критерия дисперсия погрешности моделирования уменьшается. Более существенно погрешность моделирования уменьшается при приближении уровня значимости к нулю (таблица 1.11). Так, например, при $\alpha=0,001$ погрешность моделирования уменьшается на 66%, что эквивалентно сокращению объема моделирования в 10 раз.

Таблица 1.11 – Сравнение погрешности моделирования мощности критерия при известном и неизвестном законе распределения статистики критерия, при $N=1000$

α	$1 - \beta$	ε_N , закон распределения статистики неизвестен	ε_N , закон распределения статистики известен	Сокращение погрешности
0,15	0,622333	0,0495	0,0341	31%
0,10	0,562341	0,0540	0,0374	31%
0,05	0,472871	0,0580	0,0409	30%
0,01	0,316228	0,0735	0,0420	43%
0,001	0,177828	0,1035	0,0354	66%

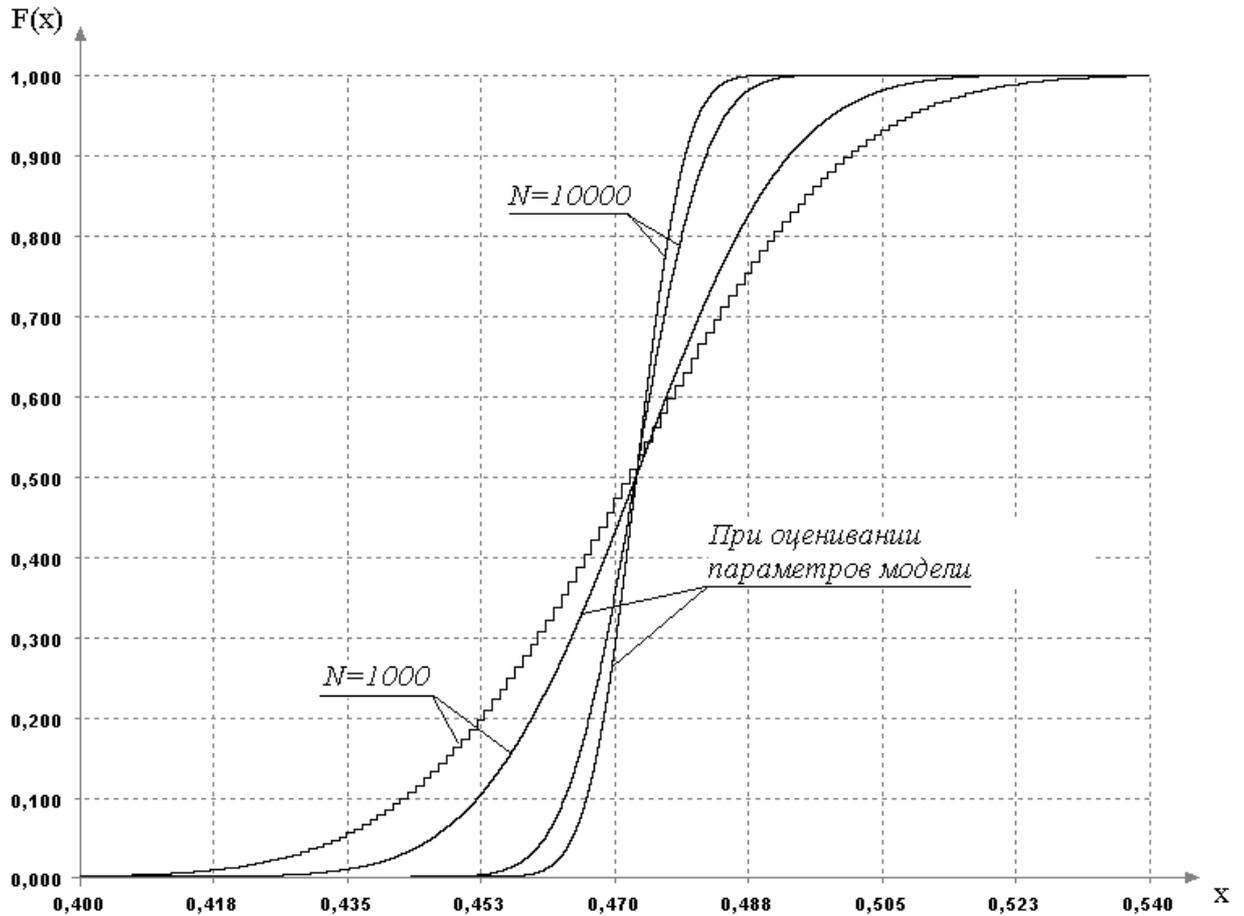


Рисунок 1.3 – Распределения погрешности моделирования мощности при уровне значимости $\alpha=0,05$

Отметим, что сокращению дисперсии погрешности моделирования в некоторых случаях посвящен ряд работ [193, 252, 263]. Подробное описание техники сокращения дисперсии изложено в монографии [253].

1.8 Исследование скорости сходимости распределения статистики критерия к предельному закону

Пусть в случае верной гипотезы H_0 статистика критерия $S(X_n)$ имеет функцию распределения $G_n(x)$, а при $n \rightarrow \infty$ - предельную функцию распределения $G(x)$. Доказать сходимость $G_n(x)$ к $G(x)$ с помощью компьютерного моделирования невозможно, однако можно решить следующие задачи:

- оценить скорость сходимости $G_n(x)$ к $G(x)$;
- определить объем выборки, при котором расстояние до предельного закона не превышает заданного числа τ .

Эти две задачи близки по смыслу, но не эквивалентны. Первая задача представляет теоретический интерес, например, если нужно проверить некоторые теоретические предположения о скорости сходимости. Вторая задача имеет важный смысл для практики, – при каком объеме выборки с приемлемой величиной ошибки вместо “истинного” можно использовать предельный закон распределения статистики.

1.8.1 Определение скорости сходимости

Пусть $\rho(G_n, G)$ - расстояние между двумя функциями распределения $G_n(x)$ и $G(x)$. Например, свойствами расстояния в пространстве законов распределений обладает статистика Колмогорова:

$$D_n = \sup_{|x| < \infty} |G_n(x) - G(x)|.$$

Скоростью сходимости $G_n(x)$ (например, распределения статистики при объеме выборки n) к $G(x)$ (предельному) называется число b , при котором $\rho(G_n, G) = O(n^{-b})$. Чем больше величина b , тем больше скорость сходимости распределения статистики к предельному закону.

Задача определения скорости сходимости распределения статистики к предельному закону является одной из классических задач математической статистики и может быть решена только аналитическими методами. Однако предложенные далее алгоритмы позволяют дать численные оценки скорости сходимости, с одной стороны, и оценить необходимый объем выборки, при котором погрешность, вызванная использованием предельного закона при

вычислении достигаемого уровня значимости, будет несущественной, с другой стороны.

Для оценки скорости сходимости будет использоваться предположение о том, что расстояние между предельным и допредельным законом распределения статистики аппроксимируется как $\rho(n) = an^{-b}$ с неизвестными параметрами a и b .

1.8.2 Алгоритм моделирования закона распределения статистики критерия при конечном объеме выборки

Аналитическое нахождение функции распределения $G_n(x)$, как правило, представляет собой более сложную задачу, чем нахождение предельного закона распределения. Однако достаточно просто можно построить эмпирическую функцию распределения для $G_n(x)$, используя метод Монте-Карло. Для этого нужно сгенерировать выборку значений статистик критерия объемом N : $\{S_1, S_2, \dots, S_N\}$ и построить по ней эмпирическую функцию распределения $G_{n,N}(x)$. В результате вместо расстояния D_n будет вычисляться

$$D_{n,N} = \sup_{|x| < \infty} |G_{n,N}(x) - G(x)|. \quad (1.10)$$

Естественно, что эмпирическое распределение $G_{n,N}(x)$ отличается от $G_n(x)$, но величину отклонения ε_N можно определить по теореме Колмогорова и подобрать такое N , при котором величина погрешности моделирования будет меньше, чем расстояние (рисунок 1.4). Если ошибка моделирования будет больше чем расстояние до предельного закона, то тогда восстановить зависимость расстояния от объема выборки будет очень сложно.

По теореме Колмогорова [39]

$$\lim_{n \rightarrow \infty} P \left\{ \sqrt{N} \sup_{|x| < \infty} |G_{n,N}(x) - G_n(x)| < t \right\} = K(t),$$

где $K(t)$ - распределение Колмогорова. Отсюда можно найти такое N , при котором величина погрешности моделирования $\varepsilon_N = \sup_{|x| < \infty} |G_{n,N}(x) - G_n(x)|$ не превосходит заданного значения с некоторой доверительной вероятностью. Так, например, если мы хотим, чтобы погрешность моделирования ε_N не превышала 0,001 с вероятностью 0,99, то мы должны взять объем выборки статистик, равный

$$N = \left[\left(\frac{K^{-1}(0,99)}{0,001} \right)^2 \right] + 1 \approx \left[\left(\frac{1,62762}{0,001} \right)^2 \right] + 1 = 2\,649\,147.$$

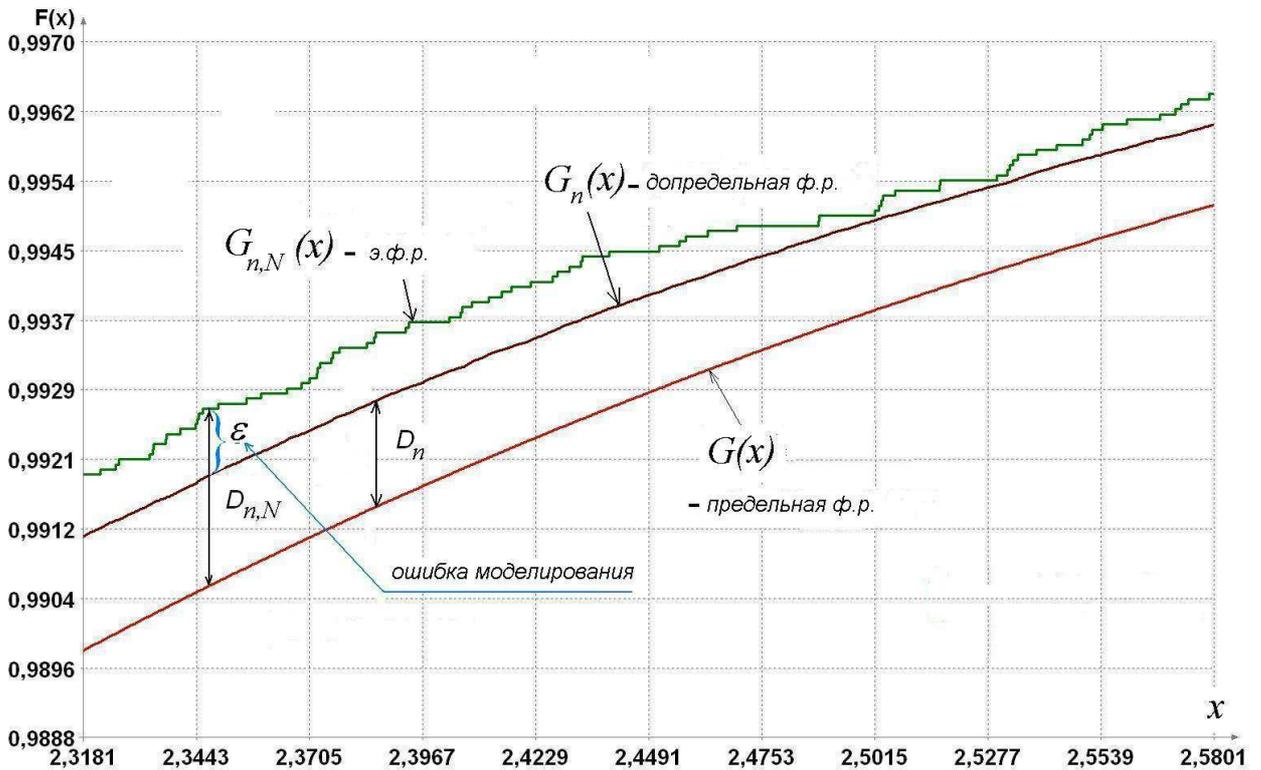


Рисунок 1.4 – Определение функции распределения $G_n(x)$ с помощью эмпирической функции распределения $G_{n,N}(x)$

1.8.3 Аппроксимация расстояния до предельного закона степенной функцией

В результате моделирования должна получиться таблица расстояний следующего вида $r_{n_1}, r_{n_2}, r_{n_3}, r_{n_4}, r_{n_5}, \dots$ (таблица 1.12). Дальнейшее увеличение n , когда $D_{n,N} < 2\varepsilon_N$, уже не имеет смысла, т.к. в этом случае $D_{n,N}$ будет показывать ошибку моделирования, а не расстояние до предельного закона распределения.

Далее по полученным значениям можно оценить параметры степенной регрессии $\rho(n) = an^{-b}$, или линейной регрессии, если рассматривать уравнение $\ln \rho = \ln a - b \ln n$. В случае высокого значения коэффициента детерминации можно считать, что скорость сходимости приближенно равна оценке параметра b . Используя найденное уравнение степенной регрессии, можно решить уравнение $\hat{a}n^{-\hat{b}} = \tau$ и найти объем выборки n , начиная с которого расстояние до предельного закона распределения не превышает заданного значения τ :

$$n = \left(\frac{\tau}{\hat{a}} \right)^{-1/\hat{b}}. \quad (1.10)$$

Таблица 1.12 – Зависимость расстояния от n

n	n_1	n_2	n_3	n_4	n_5	...
$D_{n,N}$	r_{n_1}	r_{n_2}	r_{n_3}	r_{n_4}	r_{n_5}	...

Задачу выбора точек n_i , в которых нужно проводить моделирование функции распределения статистики критерия, можно решить с помощью методов планирования эксперимента [155].

По предложенной автором методике оценивания скорости сходимости производились исследования скорости сходимости критериев экспоненциальности [127], критериев однородности по случайно

цензурированным данным [120, 115, 116], критериев согласия типа хи-квадрат по случайно цензурированным данным [117, 118], критерия однородности Андерсона-Дарлинга-Петита (*глава 5*).

1.9 Моделирование псевдослучайных величин

В основе метода Монте-Карло лежит моделирование псевдослучайных величин. Хороший генератор псевдослучайных величин (ГПСЧ) для метода Монте-Карло должен обладать следующими свойствами:

1. Распределение генерируемых чисел должно подчиняться *равномерному* закону.
2. Последовательность должна иметь длинный *период* – количество элементов последовательности до ее зацикливания должно быть достаточно большим.
3. Количество *различных* псевдослучайных чисел должно быть достаточно велико.

Нарушение равномерности генерируемых чисел будет приводить к ошибочным результатам моделирования по методу Монте-Карло.

Маленький период последовательности будет приводить к повторению выборок, возможно, будет появляться статистическая зависимость между соседними элементами выборки статистик критерия.

Небольшое количество различных псевдослучайных чисел будет приводить к дополнительной *дискретизации* распределений статистик критериев, вызванной моделированием.

Поэтому очень важно выполнение всех трех перечисленных свойств. Кроме того, немаловажным является вопрос *времени* генерирования одного числа, так как это существенно влияет на количество повторений в методе Монте-Карло, которые можно сделать за ограниченное время. Чем быстрее

алгоритм ГПСЧ, тем меньше будет погрешность моделирования при ограничении на время моделирования.

Отметим, что выбор начального значения для инициализации ГПСЧ в методе Монте-Карло не так важен, как в криптографии. Поскольку генерируются последовательности выборок случайной величины, каждая из которых имеет объем n наблюдений, то важно, чтобы n не было кратно *периоду* ГПСЧ, иначе произойдет заикливание – выборки начнут повторяться.

1.9.1 Генераторы псевдослучайных величин

Первым компьютерным ГПСЧ был генератор «Метод середины квадрата» (Middle square method), предложенный Нейманом в 1946 году для реализации метода Монте-Карло на первом компьютере ENIAC [234]. Он оказался неудачным из-за короткого периода последовательности и от него быстро отказались [40].

Одним из первых ГПСЧ, вошедшим в употребление в 60-х годах прошлого столетия, был генератор RANDU [40], основанный на рекуррентной последовательности

$$V_{j+1} = (65539V_j) \bmod 2^{31},$$

где V_0 является нечетным числом. Генератор RANDU является линейным конгруэнтным ГПСЧ, у которого параметры были подобраны очень неудачно. Оказалось, что между любыми тремя последовательными элементами последовательности RANDU имеется линейная зависимость [40].

В 1997 году Мацумото и Нисимурой был предложен ГПСЧ «Mersenne Twister». У него большой период ($2^{19937}-1$), равномерное распределение в 623 измерениях, быстрая генерация случайных чисел [227]. Принцип работы генератора основывается на свойствах простых чисел Мерсенна, которое обеспечивает равномерное распределение ПСЧ в 623 измерениях. Этот ГПСЧ

реализован в большинстве библиотек генераторов псевдослучайных чисел, в частности от компаний Intel и Microsoft [157].

В то же время ГПСЧ «Mersenne Twister» чувствителен к плохой начальной инициализации. В 2006 году была разработана его модификация - ГПСЧ «WELL» (Well Equidistributed Long-period Linear), которая имеет меньшую зависимость от начального приближения, такую же или большую производительность [239, 183].

1.9.2 Моделирование псевдослучайных величин по непрерывному закону

Непрерывная случайная величина ξ , имеющая функцию распределения $F(x)$, может быть смоделирована на основании непрерывной равномерной случайной величины η на интервале $[0,1]$ по методу обратной функции [153]:

$$\xi = F^{-1}(\eta).$$

Если обратная функция не выражается аналитически, то ее достаточно просто найти численно, используя метод Ньютона, который обычно сходится за несколько итераций.

1.10 Выводы

В данной главе описана методика проведения компьютерного моделирования в применении к задачам проверки статистических гипотез [49, 67, 87, 89, 97, 98, 99, 216].

Данная методика применялась автором при исследовании мощности критериев экспоненциальности [139], критериев согласия [92, 138, 95, 96, 112, 142, 37], непараметрических критериев согласия [72, 209, 212, 213, 215, 210, 217, 219], критериев согласия типа Хи-квадрат [80, 5, 134], критериев случайности [107, 108, 110, 114], критериев однородности [126, 143, 144], критериев многомерной нормальности [125, 132, 154]; при исследовании

систем массового обслуживания [109]; при исследовании критериев отбраковки аномальных наблюдений [101]; при исследовании робастности критериев [50].

Разработана методика исследования сходимости распределения статистики критерия к предельному закону, которая применялась при исследовании сходимости распределений статистик к своим предельным законам для критериев экспоненциальности [127], для непараметрических критериев согласия [58], для критериев однородности по случайно цензурированным данным [243, 115, 116], для критериев согласия типа хи-квадрат по группированным [57, 64, 80] и по случайно цензурированным данным [117, 118], для критерия однородности Андерсона-Дарлинга-Петита (глава 5).

Особое внимание уделено вопросам оценки погрешности моделирования функций распределения статистик критериев, достигаемого уровня значимости, критических значений и мощности статистического критерия при использовании метода Монте-Карло. С помощью компьютерного моделирования на рассмотренном примере показано, что знание семейства закона распределения статистики позволяет уменьшить объем моделирования примерно на один порядок.

Следует также отметить, что моделирование статистических закономерностей хорошо распараллеливается, что позволяет эффективно использовать вычислительные ресурсы на параллельных компьютерах [111, 218].

Исследование критериев проверки статистических гипотез, представленных на рисунке 1.1, с помощью *компьютерного моделирования* проводилось в течение последних 20 лет учеными научной школы Б.Ю. Лемешко. В диссертации Е.В. Чимитовой проводились исследования свойств критериев согласия по группированным и цензурированным наблюдения [156]. В диссертации С.С. Помадина проводилось исследование

критериев проверки гипотез корреляционного анализа в условиях, когда ошибки наблюдений отклоняются от нормального распределения [106]. В диссертации В.М. Волковой исследовались распределения статистик критериев дисперсионного анализа в условиях нарушения предположений нормальности [9]. В диссертации С.Б. Лемешко проводились исследования критериев однородности Лемана-Розенблатта и Смирнова, критериев нормальности, критериев отбраковки аномальных наблюдений, а также были построены более точные вероятностные модели распределений статистик критериев Колмогорова, Крамера-Мизеса-Смирнова, Андерсона-Дарлинга при проверке сложных гипотез [94]. В диссертации А.П. Рогожникова проводились исследования свойств критериев нормальности и экспоненциальности [150].

В настоящей работе представлены результаты исследования последовательных критериев Вальда, Лордена и Айвазяна; критериев согласия при проверке простых и сложных гипотез, при использовании разных методов оценивания параметров; критериев однородности; критериев многомерной нормальности; критериев ассоциаций при проведении полногеномного анализа ассоциаций.

Таким образом, результаты автора в совокупности с результатами других ученых научной школы Б.Ю. Лемешко охватывают практически все востребованные на практике задачи проверки статистических гипотез.

ГЛАВА 2. ОПТИМАЛЬНОЕ ПЛАНИРОВАНИЕ СТАТИСТИЧЕСКОГО ЭКСПЕРИМЕНТА ДЛЯ РАЗЛИЧЕНИЯ ДВУХ ПРОСТЫХ ГИПОТЕЗ

Одним из самых простых случаев проверки гипотез является различение двух простых гипотез, когда и основная, и конкурирующая гипотезы полностью определены. Фундаментальный результат в этой ситуации был получен Нейманом и Пирсоном [235], 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 [272]. Вальдом и Вольфовицем было доказано, что SPRT имеет минимальный средний объем выборки среди всех критериев при заданных вероятностях ошибок первого и второго рода [274]. 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 [273], 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 [203].

Экономичность последовательного критерия отношения правдоподобия по сравнению с критерием принятия решения по выборке фиксированного

объема, т.е. выигрыш в среднем числе наблюдений, была оценена Вальдом [273] и уточнена Айвазяном [1].

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 Володин Н.И. [10, 11, 12, 13, 14, 15, 16, 17, 18, 34].

Если при проведении эксперимента планируется группировать наблюдения, то кроме определения оптимального объема можно также определять оптимальные границы интервалов группирования, при которых максимизируется информация о различии между распределениями статистики критерия при справедливости основной и конкурирующей гипотез. В параметрическом случае, когда конкурирующая гипотеза принадлежит тому же семейству распределений, но с другим значением параметра, можно максимизировать информационное количество Фишера [41, 62, 207]. В общем случае, когда основная и конкурирующая гипотезы соответствуют разным семействам законов распределения, можно максимизировать дивергенцию Кульбака-Лейблера [203].

2.1 Простая гипотеза о виде распределения

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, \dots, 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.

2.2 Дивергенция Кульбака-Лейблера

Отношение $z_i = \ln \frac{f_1(x_i)}{f_0(x_i)}$ называется *информацией о различии* между гипотезами H_0 и H_1 , которую несет одно наблюдение x_i [203]. Объяснение именно такой формы информации может быть дано в смысле теоремы Байеса как разность между логарифмами шансов в пользу гипотезы H_1 до и после наблюдения x_i [204, 42]:

$$\ln \frac{f_1(x_i)}{f_0(x_i)} = \ln \frac{P\{H_1 | \xi = x_i\}}{P\{H_0 | \xi = x_i\}} - \ln \frac{P\{H_1\}}{P\{H_0\}}.$$

Среднее значение информации о различии между H_0 и H_1 называется *дивергенцией Кульбака-Лейблера*

$$D(H_0, H_1) = -E_0 z_i = \int_R f_0(x) \ln \frac{f_0(x)}{f_1(x)} dx, \quad (2.1)$$

$$D(H_1, H_0) = E_1 z_i = \int_R f_1(x) \ln \frac{f_1(x)}{f_0(x)} dx. \quad (2.2)$$

Дивергенция Кульбака-Лейблера обладает двумя фундаментальными свойствами [42].

1. Дивергенция Кульбака-Лейблера является неотрицательной, причем $D(H_0, H_1) = 0$ тогда, и только тогда, когда $H_0 = H_1$.
2. Дивергенция Кульбака-Лейблера является асимметричной, т.е. $D(H_0, H_1) \neq D(H_1, H_0)$.

Симметричная дивергенция Кульбака-Лейблера равна [42]

$$\rho(H_0, H_1) = D(H_0, H_1) + D(H_1, H_0) = \int_R (f_1(x) - f_0(x)) f_0(x) \ln \frac{f_1(x)}{f_0(x)} dx. \quad (2.3)$$

Хотя $\rho(H_0, H_1)$ и обладает свойством симметричности, она не является расстоянием в пространстве распределений, т.к. для неё не выполняется неравенство треугольника [42].

Следует отметить, что выражения для $D(H_0, H_1)$ и $\rho(H_0, H_1)$ были известны и использовались до работы [203]. Роль Кульбака и Лейблера состоит в рассмотрении этих выражений с точки зрения теории информации и изучении их свойств.

2.3 Критерий отношения правдоподобия

Критерий отношения правдоподобия строится на основании *статистики отношения правдоподобия*

$$\Lambda_n = \prod_{i=1}^n \frac{f_1(x_i)}{f_0(x_i)}, \quad (2.4)$$

или ее логарифма

$$\lambda_n = \ln \Lambda_n = \ln \prod_{i=1}^n \frac{f_1(x_i)}{f_0(x_i)} = \sum_{i=1}^n \ln \frac{f_1(x_i)}{f_0(x_i)} = \sum_{i=1}^n z_i. \quad (2.5)$$

С точки зрения теории информации статистика отношения правдоподобия λ_n равна информации о различии между гипотезами H_0 и H_1 о

принадлежности выборки соответствующему закону с плотностью $f_j(x)$, содержащейся во всей выборке.

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

Отсюда, в случае непрерывности случайной величины ξ

$$t_\alpha = F_{\lambda_n|H_0}^{-1}(1 - \alpha), \quad (2.6)$$

где $F_{\lambda_n|H_0}^{-1}(y)$ - обратная функция распределения случайной величины λ_n при верной гипотезе H_0 . Мощность критерия отношения правдоподобия определяется соотношением

$$1 - \beta = P\{\lambda_n > t_\alpha | H_1\} = F_{\lambda_n|H_1}(t_\alpha). \quad (2.7)$$

Фундаментальную роль критерия отношения правдоподобия устанавливает следующая теорема [235].

Теорема 2.1 (Лемма Неймана-Пирсона)

Среди всех критериев заданного уровня значимости α , проверяющих две простые гипотезы H_0 и H_1 , критерий отношения правдоподобия является наиболее мощным.

Егучи и Копас [174] определили, что если конкурирующая гипотеза H_1 была выбрана неправильно и на самом деле верна гипотеза H_2 , то критерий отношения правдоподобия уже не будет самым мощным, и потеря мощности равна $D(H_1, H_2)$. Аналогично, при неправильном выборе гипотезы H_0 вместо H_2 , критерий отношения правдоподобия имеет потерю мощности $D(H_0, H_2)$.

2.4 Последовательный критерий отношения правдоподобия

Последовательный критерий отношения правдоподобия (SPRT, критерий Вальда) строится следующим образом. Сначала выбирают критические границы c_0 и c_1 ($c_0 < c_1$). После первого наблюдения x_1 вычисляют значение статистики критерия

$$\lambda_1 = \ln \left(\frac{f_1(x_1)}{f_0(x_1)} \right) = z_1.$$

Если $\lambda_1 < c_0$, то принимают гипотезу H_0 ; если $\lambda_1 > c_1$, принимают H_1 ; если $c_0 \leq \lambda_1 \leq c_1$, то производят второе наблюдение x_2 и также исследуют величину

$$\lambda_2 = \ln \left(\frac{f_1(x_1)f_1(x_2)}{f_0(x_1)f_0(x_2)} \right) = z_1 + z_2.$$

Известно [273], что с вероятностью, равной единице, процесс оканчивается либо выбором H_0 , либо выбором H_1 , т.е. на каком-то шаге n статистика

$$\lambda_n = \sum_{i=1}^n \ln \left(\frac{f_1(x_i)}{f_0(x_i)} \right) = \sum_{i=1}^n z_i \quad (2.8)$$

выйдет за интервал $[c_0, c_1]$.

Величины c_0 и c_1 должны определяться из условия равенства вероятностей ошибок первого и второго рода заданным значениям α и β . Обозначим эту зависимость через $c_0(\alpha, \beta)$ и $c_1(\alpha, \beta)$. В случае проверки простой гипотезы известна следующая теорема [273].

Теорема 2.2. Критические значения c_0 и c_1 критерия Вальда удовлетворяют неравенствам:

$$c_0 \geq \ln\left(\frac{\beta}{1-\alpha}\right), \quad c_1 \leq \ln\left(\frac{1-\beta}{\alpha}\right), \quad (2.9)$$

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

При этом, если границы c_0 и c_1 заменить правыми частями в (2.9),

$$c'_0 = \ln\left(\frac{\beta}{1-\alpha}\right), \quad c'_1 = \ln\left(\frac{1-\beta}{\alpha}\right) \quad (2.10)$$

то вероятности ошибок первого и второго рода будут равны (α', β') , где

$$\alpha' \leq \frac{\alpha}{1-\beta}, \quad \beta' \leq \frac{\beta}{1-\alpha} \quad \text{и} \quad \alpha' + \beta' \leq \alpha + \beta. \quad (2.11)$$

Эта теорема позволяет использовать последовательный критерий отношения правдоподобия с приближенными границами (2.10), которые не зависят от распределений, заданных гипотезами H_0 и H_1 , но при этом дают меньшую суммарную вероятность ошибок первого и второго рода (2.11). Оцениванию точных критических границ с помощью метода Монте-Карло будет посвящена третья глава.

В последовательных критериях количество наблюдений является случайной величиной, поэтому в качестве основной характеристики последовательного критерия рассматривается математическое ожидание объема выборки (average sample number, ASN). Вальдом найдена нижняя граница для SPRT [273], причем эта граница обратно пропорциональна дивергенции Кульбака-Лейблера.

Теорема 2.3. Оценка снизу среднего числа наблюдений для любого последовательного критерия с вероятностями ошибок α и β имеет вид:

$$E_{H_0} n(\alpha, \beta) \geq \frac{\omega(\alpha, \beta)}{D(H_0, H_1)}, \quad (2.12)$$

$$E_{H_1} n(\alpha, \beta) \geq \frac{\omega(\beta, \alpha)}{D(H_1, H_0)}, \quad (2.13)$$

где

$$\omega(x, y) = \left((1-x) \ln \left(\frac{1-x}{y} \right) + x \ln \left(\frac{x}{1-y} \right) \right). \quad (2.14)$$

Нижние границы 2.12 и 2.13 являются неулучшаемыми, но для большинства конкурирующих гипотез средний объем выборки существенно больше.

Айвазяном в 1959 были получены асимптотические (по сближению различаемых гипотез) точные формулы для среднего объема выборки [1].

Теорема 2.4. Если $f_0(x) = f(x, \theta_0)$, $f_1(x) = f(x, \theta_1)$ и $\theta_1 \rightarrow \theta_0$, то при выполнении ряда условий справедливо:

$$E_{H_0} n(\alpha, \beta) = \frac{2\omega(\alpha, \beta)}{\rho(H_0, H_1)}, \quad (2.15)$$

$$E_{H_1} n(\alpha, \beta) = \frac{2\omega(\beta, \alpha)}{\rho(H_0, H_1)}. \quad (2.16)$$

Рассмотрим, насколько применимы формулы (2.15) и (2.16). Отношение правой части (2.15) к (2.12) равно

$$\frac{2\omega(\alpha, \beta)}{\rho(H_0, H_1)} : \frac{\omega(\alpha, \beta)}{D(H_0, H_1)} = \frac{2D(H_0, H_1)}{\rho(H_0, H_1)} = \frac{2D(H_0, H_1)}{D(H_0, H_1) + D(H_1, H_0)}. \quad (2.17)$$

Аналогично, отношение правой части (2.16) к (2.13) равно

$$\frac{2\omega(\beta, \alpha)}{\rho(H_0, H_1)} : \frac{\omega(\beta, \alpha)}{D(H_1, H_0)} = \frac{2D(H_1, H_0)}{\rho(H_0, H_1)} = \frac{2D(H_1, H_0)}{D(H_0, H_1) + D(H_1, H_0)}. \quad (2.18)$$

Анализируя выражения (2.17) и (2.18), можно сделать вывод, что одно из них должно быть больше 1, а другое – меньше 1, в зависимости от того, что больше: $D(H_1, H_0)$ или $D(H_0, H_1)$. Таким образом, (2.15) и (2.16) могут быть верными одновременно тогда и только тогда, когда $D(H_1, H_0) = D(H_0, H_1)$. Поэтому правые части в (2.15) и (2.16) совпадут с нижними границами в (2.13) и (2.14), и, следовательно, теорема 2.4 дает пример (асимптотический), при котором достигается нижняя граница в неравенстве Вальда (2.12)-(2.13).

В 1948 году Вальд и Вольфовиц доказали оптимальность среднего объема выборки в SPRT [274].

Теорема 2.5. Пусть T – последовательный критерий отношения правдоподобия с критическими границами $-\infty < c_0 < 0 < c_1 < +\infty$, с вероятностями ошибок первого и второго рода α и β , $E_{H_0} n(\alpha, \beta)$ – средний объем выборки при верной гипотезе H_0 , $E_{H_1} n(\alpha, \beta)$ – средний объем выборки при верной гипотезе H_1 , а T' – другой критерий с вероятностями ошибок первого и второго рода α' и β' , и конечными средними объемами выборок $E_{H_0} n'(\alpha', \beta') < \infty$ и $E_{H_1} n'(\alpha', \beta') < \infty$.

Тогда, если $\alpha' < \alpha$ и $\beta' < \beta$, то

$$E_{H_0} n(\alpha, \beta) \leq E_{H_0} n'(\alpha', \beta') \text{ и } E_{H_1} n(\alpha, \beta) \leq E_{H_1} n'(\alpha', \beta').$$

В 1976 году Симонс несколько усилил данную теорему [264].

Теорема 2.6. Пусть T – последовательный критерий отношения правдоподобия с критическими границами $-\infty < c_0 < 0 < c_1 < +\infty$, с вероятностями ошибок первого и

второго рода α и β , $E_{H_0} n(\alpha, \beta)$ - средний объем выборки при верной гипотезе H_0 , $E_{H_1} n(\alpha, \beta)$ - средний объем выборки при верной гипотезе H_1 , а T' - другой критерий с вероятностями ошибок первого и второго рода α' и β' , и конечными средними объемами выборок $E_{H_0} n'(\alpha', \beta') < \infty$ и $E_{H_1} n'(\alpha', \beta') < \infty$.

Тогда, если $\frac{\alpha'}{1-\beta'} < \frac{\alpha}{1-\beta}$ и $\frac{\beta'}{1-\alpha'} < \frac{\beta}{1-\alpha}$, то

$$E_{H_0} n(\alpha, \beta) \leq E_{H_0} n'(\alpha', \beta') \text{ и } E_{H_1} n(\alpha, \beta) \leq E_{H_1} n'(\alpha', \beta').$$

Суть теорем 2.5 и 2.6 в том, что критерий с меньшими ошибками первого и второго рода будет требовать в среднем большего или такого же числа наблюдений, что и SPRT.

2.5 Необходимый объем выборки для критерия отношения правдоподобия

Рассмотрим теперь задачу про нахождение необходимого объема выборки для критерия отношения правдоподобия при различении двух простых гипотез.

Распределение логарифма статистики отношения правдоподобия при достаточно большом объеме (что, например, возможно при стремлении вероятностей ошибок первого и второго рода к нулю), согласно центральной предельной теореме, подчиняется нормальному распределению

$$\begin{aligned} F_{\lambda_n}(t) &= P\{\lambda_n < t\} = P\left\{\sum_{i=1}^n z_i < t\right\} = \\ &= P\left\{\sqrt{n} \frac{\frac{1}{n} \sum_{i=1}^n z_i - Ez}{\sqrt{Dz}} < \sqrt{n} \frac{\frac{t}{n} - Ez}{\sqrt{Dz}}\right\} \approx \Phi\left(\sqrt{n} \frac{\frac{t}{n} - Ez}{\sqrt{Dz}}\right), \end{aligned} \quad (2.19)$$

где $\Phi(x)$ – функция распределения стандартного нормального закона.

Из (2.6), (2.7) и (2.19) можно выразить критическое значение t_α :

$$\left\{ \begin{array}{l} 1 - \Phi \left(\frac{t_\alpha - E_{H_0} z}{\sqrt{n} \frac{n}{\sqrt{D_{H_0} z}}} \right) = \alpha \\ 1 - \Phi \left(\frac{t_\alpha - E_{H_1} z}{\sqrt{n} \frac{n}{\sqrt{D_{H_1} z}}} \right) = 1 - \beta \end{array} \right. \Rightarrow \left\{ \begin{array}{l} \frac{t_\alpha}{n} = E_{H_0} z + \frac{\sqrt{D_{H_0} z}}{\sqrt{n}} \Phi^{-1}(1 - \alpha) \\ \frac{t_\alpha}{n} = E_{H_1} z + \frac{\sqrt{D_{H_1} z}}{\sqrt{n}} \Phi^{-1}(\beta) \end{array} \right. \quad (2.20)$$

Приравнявая t_α , из (2.20), можно выразить n :

$$n = \left(\frac{\sqrt{D_{H_0} z} \cdot \Phi^{-1}(1 - \alpha) - \sqrt{D_{H_1} z} \cdot \Phi^{-1}(\beta)}{E_{H_1} z - E_{H_0} z} \right)^2. \quad (2.21)$$

Формулой (2.21) можно пользоваться для оценки необходимого объема выборки при достаточно малых вероятностях ошибок первого и второго рода. Если достаточна довольно грубая оценка необходимого объема выборки, то можно воспользоваться следующими выкладками.

Предполагая, что $\alpha = \beta$, получаем

$$n = \left(\frac{\sqrt{D_{H_0} z} - \sqrt{D_{H_1} z}}{\rho(H_0, H_1)} \right)^2 \cdot (\Phi^{-1}(\alpha))^2. \quad (2.22)$$

Подставляя (2.22) в неравенства Вальда (2.12) и (2.13), получаем

$$\left\{ \begin{array}{l} \left(\frac{\sqrt{D_{H_0} z} - \sqrt{D_{H_1} z}}{\rho(H_0, H_1)} \right)^2 \cdot (\Phi^{-1}(\alpha))^2 \geq \frac{\omega(\alpha, \alpha)}{D(H_0, H_1)}; \\ \left(\frac{\sqrt{D_{H_0} z} - \sqrt{D_{H_1} z}}{\rho(H_0, H_1)} \right)^2 \cdot (\Phi^{-1}(\alpha))^2 \geq \frac{\omega(\alpha, \alpha)}{D(H_1, H_0)}; \end{array} \right. \quad (2.23)$$

Отсюда следует, что

$$\frac{\left(\sqrt{D_{H_0} z} - \sqrt{D_{H_1} z}\right)^2}{\rho(H_0, H_1)} \geq \frac{2\omega(\alpha, \alpha)}{\left(\Phi^{-1}(\alpha)\right)^2}. \quad (2.24)$$

Левая часть неравенства (2.24) есть функционал от распределений, соответствующих гипотезам, а правая часть – функция от вероятности ошибки первого рода. Т.к. неравенство (2.24) должно выполняться при любом α , то мы можем найти максимальное значение правой части (2.24).

Функция $\frac{2\omega(\alpha, \alpha)}{\left(\Phi^{-1}(\alpha)\right)^2}$ является выпуклой, и ее максимум равен $d = 2,5465$

и достигается при $\alpha=0.5$ (рисунок 2.1). Подставляя (2.24) в (2.22), получаем, что

$$n \geq \frac{d \cdot \left(\Phi^{-1}(\alpha)\right)^2}{\rho(H_0, H_1)}. \quad (2.25)$$

Формула (2.25) дает довольно простую по виду нижнюю границу необходимого объема выборки для различения двух простых гипотез при $\alpha = \beta$ на основании нижней границы последовательного критерия. Однако последовательный критерий часто требует в два раза меньше наблюдений, чем наилучший критерий с фиксированным объемом выборки [273, стр. 1].

Айвазяном в 1959 г. найден необходимый асимптотический объем выборки (по сближению гипотез H_0 и H_1) [1].

Теорема 2.7. Если $f_0(x) = f(x, \theta_0)$, $f_1(x) = f(x, \theta_1)$ и $\theta_1 \rightarrow \theta_0$, то при выполнении ряда условий справедливо:

$$n \approx \frac{\left(\Phi^{-1}(1-\alpha) + \Phi^{-1}(\beta)\right)^2}{\rho(H_0, H_1)}. \quad (2.26)$$

Отметим, что (2.26) является правой частью (2.25) при $d = 4$ и $\alpha = \beta$.

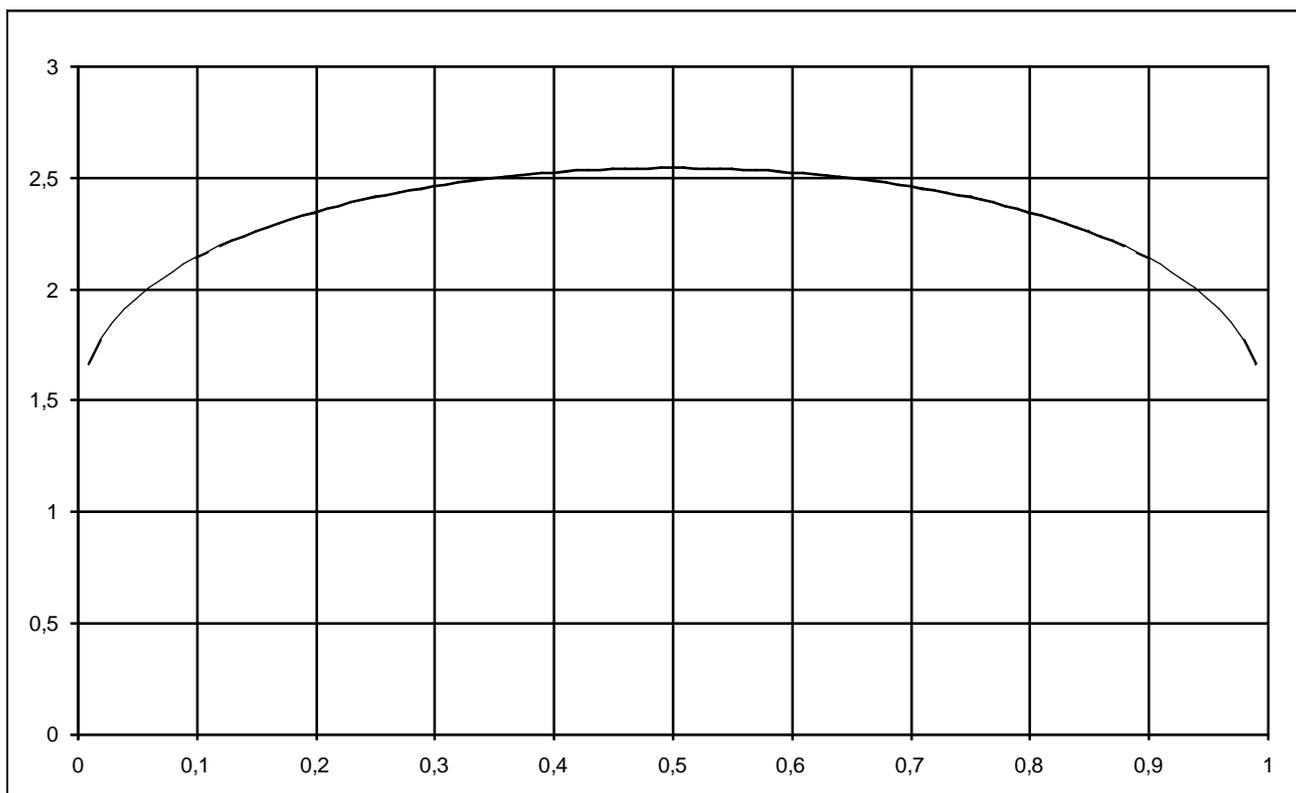


Рисунок 2.1 – Функция правой части (2.24) от α

Володиным получены более точные формулы для вычисления необходимого объема выборки для близких гипотез [10].

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

$$\lambda_n = \sum_{i=1}^n \ln \frac{f_1(x_i)}{f_0(x_i)} = \sum_{i=1}^{n-1} \ln \frac{f_1(x_i)}{f_0(x_i)} + \ln \frac{f_1(x_n)}{f_0(x_n)} = \lambda_{n-1} + z_n. \quad (2.27)$$

Формула (2.27) 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 правдоподобия по выборке объема i , при верных гипотезах H_0 и H_1 , соответственно.

Алгоритм 2.1. Определение необходимого объема выборки (для точечных наблюдений)

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 (1.2), 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 $\lambda_{ij}^0 = \lambda_{(i-1)j}^0 + \ln \frac{f_1(x_j^0)}{f_0(x_j^0)}$ и $\lambda_{ij}^1 = \lambda_{(i-1)j}^1 + \ln \frac{f_1(x_j^1)}{f_0(x_j^1)}, j = \overline{1, N}$
4. Вычислить \hat{t}_α – критическое значение статистики критерия при уровне значимости α по выборке $\{\lambda_{ij}^0, j = \overline{1, N}\}$ по формуле (1.1), где $q = 1 - \alpha$.
5. Вычислить оценку вероятности ошибки второго рода $\hat{\beta}_i$, как количество λ_{ij}^1 , не превосходящих \hat{t}_α , деленное на N .
6. Если $\hat{\beta}_i > \beta - \varepsilon_N$, то увеличить $i := i + 1$ и перейти на шаг 2.

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

В алгоритме 2.1 вычисляется оценка необходимого объема выборки, т.к. значения \hat{t}_α и $\hat{\beta}_i$ имеют статистическую погрешность моделирования, зависящую от числа повторений N . Чтобы уменьшить риск неправильного

определения необходимого объема выборки на шаге 6 учитывается погрешность моделирования ε_N .

2.6 Относительная эффективность критерия

Предположим, что имеется два критерия T_1 и T_2 для проверки гипотезы H_0 против гипотезы H_1 . Тогда *относительная эффективность* первого критерия по отношению ко второму равна

$$RE(T_1, T_2) = \frac{n_2(\alpha, \beta)}{n_1(\alpha, \beta)}, \quad (2.28)$$

где $n_i(\alpha, \beta)$ - это необходимый объем выборки для различения гипотез H_0 и H_1 с вероятностями ошибок первого и второго рода, равными α и β , соответственно.

Поскольку часто выражение (2.28) трудно вычислить аналитически, то рассматривают *асимптотическую относительную эффективность* [248]. Известно три асимптотических подхода – эффективность по Питмену, которая определяется как

$$ARE(T_1, T_2) = \lim_{H_1 \rightarrow H_0} \frac{n_2(\alpha, \beta)}{n_1(\alpha, \beta)}, \quad (2.29)$$

эффективность по Ходжесу-Леману, которая определяется как [186]

$$ARE(T_1, T_2) = \lim_{\beta \rightarrow 0} \frac{n_2(\alpha, \beta)}{n_1(\alpha, \beta)}, \quad (2.30)$$

и эффективность по Бахадуру, которая определяется как [163]

$$ARE(T_1, T_2) = \lim_{\alpha \rightarrow 0} \frac{n_2(\alpha, \beta)}{n_1(\alpha, \beta)}. \quad (2.31)$$

Имеются обобщения асимптотической эффективности для последовательных критериев [166, 167]

Следует отметить, что относительная эффективность критерия $RE(T_1, T_2)$ для заданных α и β может быть вычислена с помощью компьютерного моделирования. В главе 7 таким способом вычислена относительная эффективность критерия ассоциации МАХЗ по отношению к критерию тренда Кокрена-Армитеджа.

2.7 Оптимальное группирование для различения двух простых гипотез

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 [203]. Они рассматривали пример разбиения наблюдений на две группы при проверке полиномиального распределения против равномерного распределения путем максимизации дивергенции Кульбака-Лейблера.

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

$$\max_{t_1 < t_2 < \dots < t_{k-1}} \sum_{i=1}^k P_{0i} \ln \frac{P_{0i}}{P_{1i}}, \quad (2.32)$$

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

Существует взаимосвязь между дивергенцией Кульбака-Лейблера и информационной матрицей Фишера. Пусть $f_0(x) = f(x, \theta)$, $f_1(x) = f(x, \theta + \Delta\theta)$.

Тогда

$$D(H_0, H_1) \approx \frac{1}{2} I(\theta) (\Delta\theta)^2, \quad (2.33)$$

где $I(\theta) = E \left[\frac{\partial \ln f(x, \theta)}{\partial \theta} \right]^2$ - информационное количество Фишера. В монографии [42] приведено выражение (2.33) для общего случая, когда распределение имеет вектор параметров. Таким образом, если конкурирующие гипотезы принадлежат одному параметрическому семейству, то максимизация дивергенции Кульбака-Лейблера эквивалентна максимизации информационного количества Фишера.

В связи с проблемой оценивания параметров по группированным выборкам задача асимптотически оптимального группирования (АОГ) данных рассматривалась в [171, 41], неоднократно она возникала при использовании для оценивания параметров выборочных квантилей. Наиболее полная совокупность таблиц асимптотически оптимального группирования для распределений экспоненциального, полунормального, Рэлея, Максвелла, модуля многомерного нормального вектора, Парето, Эрланга, Лапласа, нормального, логарифмически-нормальных (\ln и \lg), Коши, Вейбулла, распределений минимального и максимального значения, двойного показательного, гамма-распределения представлена в [31, 97].

Использование АОГ увеличивает мощность критериев типа χ^2 , когда альтернатива принадлежит тому же семейству распределений, что и основная гипотеза [62].

2.8 Связь между оптимальным группированием и мощностью критерия χ^2

Несложно показать, что оптимальное группирование (2.32) максимизирует мощность критерия χ^2 , когда основная и конкурирующая гипотеза принадлежат различным семействам распределений.

Пусть $g(x, \lambda) = (1 - \lambda)f_0(x) + \lambda f_1(x)$ - смесь распределений 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 $H_0: \lambda = 0$ против $H_1: \lambda = 1$, где λ - это параметр смеси $g(x, \lambda)$.

Тогда дивергенция Кульбака-Лейблера эквивалентна информационному количеству Фишера о параметре λ , согласно формуле (2.32). Найдем, чему равно информационное количество Фишера о параметре λ по группированной выборке:

$$I_{gr}(\lambda) = \sum_{i=1}^k ((1 - \lambda)P_{0i} + \lambda P_{1i}) \left(\frac{d \ln((1 - \lambda)P_{0i} + \lambda P_{1i})}{d\lambda} \right)^2 = \sum_{i=1}^k \frac{(P_{1i} - P_{0i})^2}{(1 - \lambda)P_{0i} + \lambda P_{1i}}. \quad (2.34)$$

Когда $\lambda \rightarrow 0$, то $I_{gr}(\lambda) \rightarrow \sum_{i=1}^k \frac{(P_{1i} - P_{0i})^2}{P_{0i}} = \chi^2(H_0, H_1)$, а это значит, что

максимизация дивергенции Кульбака-Лейблера (2.32) эквивалентна максимизации расстояния $\chi^2(H_0, H_1)$ между законами, соответствующими основной и конкурирующей гипотезам и эквивалентна максимизации параметра нецентральности, связанного с нецентральным χ^2 -распределением, которому будет подчиняться статистика критерия χ^2 Пирсона при справедливости конкурирующей гипотезы, то есть максимизации мощности этого критерия.

2.9 Оптимальное планирование эксперимента по различению двух гипотез

Таким образом, для того чтобы построить оптимальный план эксперимента по различению двух простых гипотез, необходимо:

1. Определить способ представления наблюдений (точечный или группированный).
2. Если наблюдения будут фиксироваться в группированном виде, то следует найти точки оптимального группирования по (2.32) и количество интервалов группирования, при котором отношение дивергенции Кульбака-Лейблера по группированной выборке к дивергенции Кульбака-Лейблера по полной выборке близко к единице.
3. Для определения необходимого объема выборки по критерию отношения правдоподобия следует воспользоваться алгоритмом 2.1.

Проиллюстрируем рассмотренные в этой главе алгоритмы и методы на двух численных примерах.

В первой паре проверяемой гипотезе H_0 соответствует нормальный закон с плотностью

$$f(x) = \frac{1}{\theta_0 \sqrt{2\pi}} \exp\left\{-\frac{(x - \theta_1)^2}{2\theta_0^2}\right\}, \quad (2.35)$$

и параметрами $\theta_0 = 1$, $\theta_1 = 0$, а конкурирующей гипотезе H_1 - логистический закон с функцией плотности

$$f(x) = \frac{\pi}{\theta_0 \sqrt{3}} \exp\left\{-\frac{\pi(x - \theta_1)}{\theta_0 \sqrt{3}}\right\} / \left[1 + \exp\left\{-\frac{\pi(x - \theta_1)}{\theta_0 \sqrt{3}}\right\}\right]^2 \quad (2.36)$$

и параметрами $\theta_0 = 1$, $\theta_1 = 0$.

Во второй паре H_0 соответствует распределение Вейбулла-Гнеденко с плотностью

$$f(x) = \frac{\theta_0}{\theta_1} \left(\frac{x - \theta_2}{\theta_1} \right)^{\theta_0 - 1} \exp \left\{ - \left(\frac{x - \theta_2}{\theta_1} \right)^{\theta_0} \right\}, x \geq \theta_2 \quad (2.37)$$

и параметрами $\theta_0 = 2, \theta_1 = 2, \theta_2 = 0$, а H_1 - гамма-распределение с плотностью

$$f(x) = \frac{1}{\theta_1 \Gamma(\theta_0)} \left(\frac{x - \theta_2}{\theta_1} \right)^{\theta_0 - 1} e^{-\frac{x - \theta_2}{\theta_1}}, x \geq \theta_2 \quad (2.38)$$

и параметрами $\theta_0 = 3,12154, \theta_1 = 0,557706, \theta_2 = 0$.

Выбор данных пар конкурирующих гипотез обусловлен двумя причинами. Во-первых, эти законы часто используются на практике: в метрологии, в теории массового обслуживания, теории надежности и теории выживаемости. Во-вторых, эти семейства законов распределений очень близки друг к другу, и различие их является тяжелой задачей (рисунки 2.2 и 2.3). Параметры конкурирующих законов распределения найдены минимизацией расстояния между функциями распределения. Этим объясняется большое число знаков после запятой для параметров гамма-распределения.

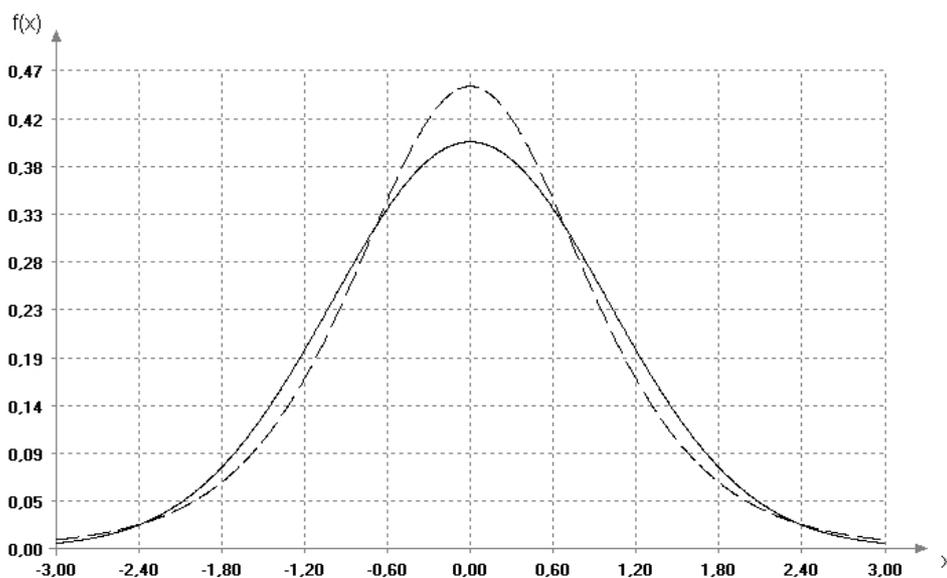


Рисунок 2.2 – Функции плотности нормального (сплошная линия) и логистического распределений (пунктирная линия)

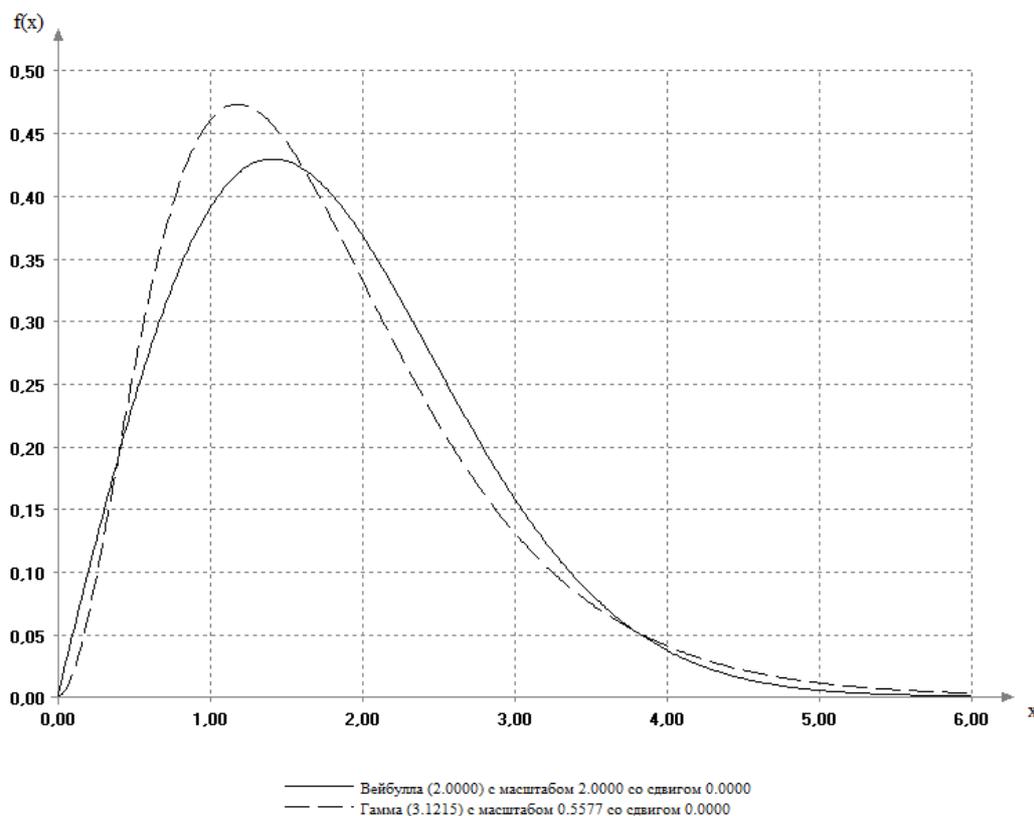


Рисунок 2.3 – Функции плотности распределения Вейбулла-Гнеденко (сплошная линия) и гамма-распределения (пунктирная линия)

В таблице 2.1 приведены результаты расчетов дивергенции Кульбака-Лейблера для двух пар проверяемых гипотез.

Расчет необходимого объема выборки для разных вероятностей приведен в таблице 2.2 для пары альтернатив нормальное против логистического и в таблице 2.3 для пары альтернатив Вейбулла-Гнеденко против гамма-распределения в сравнении с различными теоретическими оценками необходимого объема выборки для проверки гипотезы. Результаты моделирования показывают, что необходимый объем выборки больше, чем асимптотическая оценка (2.26) при сближении гипотез, полученная Айвазяном, но меньше, чем асимптотическая оценка (2.21), полученная при использовании центральной предельной теоремы. Таким образом, алгоритма 2.1 действительно

более точно оценить необходимый объем выборки при различении двух простых гипотез.

На рисунках 2.4 и 2.5 показано, как изменяется вероятность ошибки второго рода при увеличении объема выборки.

Таблица 2.1 – Дивергенция Кульбака-Лейблера

H_0	H_1	$D(H_0, H_1)$	$D(H_1, H_0)$	$\rho(H_0, H_1)$	$\sqrt{D_{H_0} z}$	$\sqrt{D_{H_1} z}$
Нормальное	Логистическое	0.0104965	0.0143623	0.0248588	0.133096629	0.220468
Вейбулла-Гнеденко	Гамма-распределение	0.0166951	0.0192277	0.0359228	0,18178744	0,229058

Таблица 2.2 – Необходимый объем выборки для различения нормального и логистического распределений

α	β	(2.12)	(2.25)	(2.26)	(2.21)	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

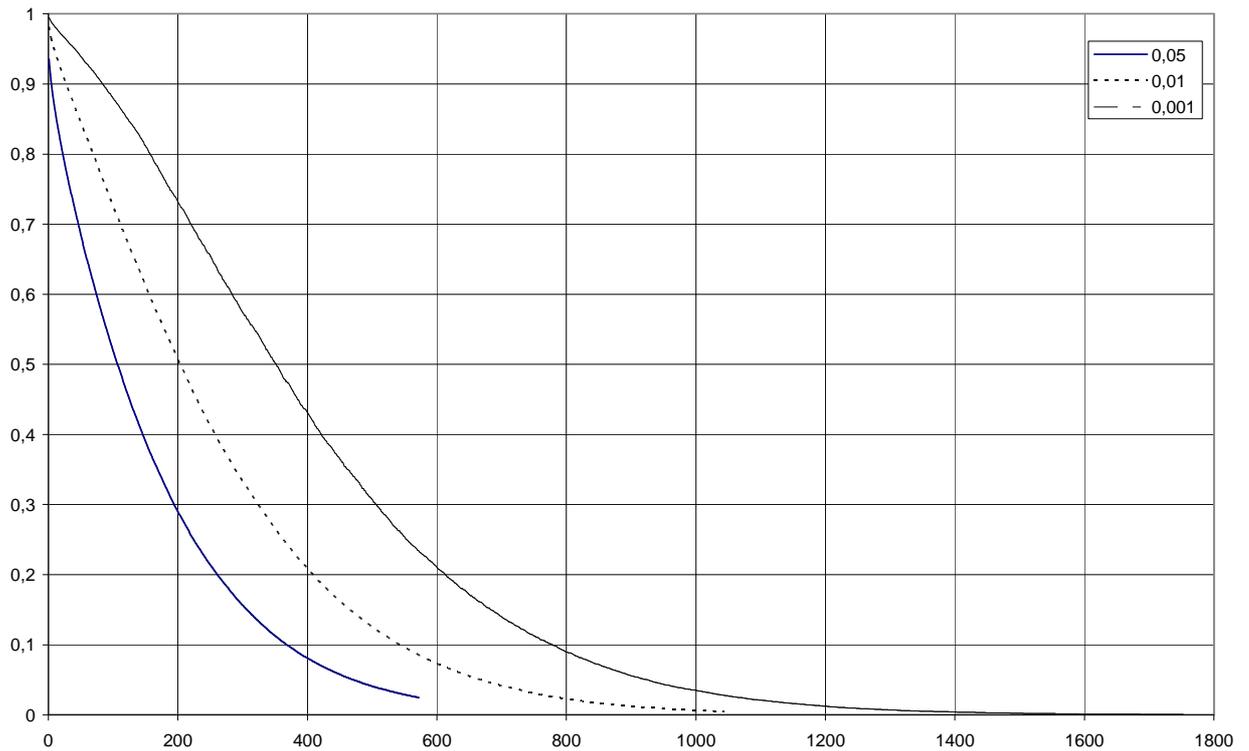


Рисунок 2.4 – Зависимость вероятности ошибки второго рода от объема выборки при проверке гипотезы о нормальном распределении против логистического распределения по критерию отношения правдоподобия

Таблица 2.3 – Необходимый объем выборки для различения распределения Вейбулла-Гнеденко и гамма-распределения

α	β	(2.12)	(2.25)	(2.26)	(2.21)	Алг. 2.1 при $N = 1660000$
0,05	0,05	159	192	302	354	319
0,01	0,01	270	384	603	708	631
10^{-3}	10^{-3}	413	677	1064	1250	1117

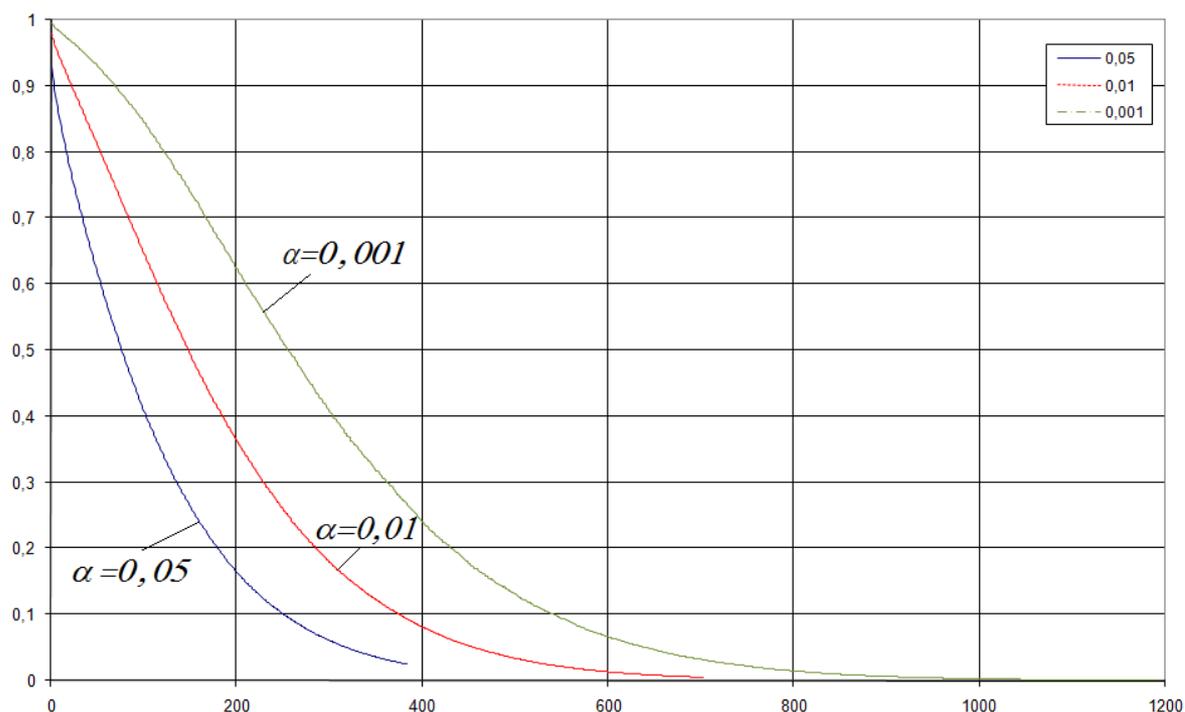


Рисунок 2.5 – Зависимость вероятности ошибки второго рода от объема выборки при проверке гипотезы о распределении Вейбулла против гамма-распределения по критерию отношения правдоподобия

2.10 Необходимый объем выборки и ошибки измерения

Если наблюдения получаются в результате измерений с некоторой погрешностью, то это влияет на проверку гипотезы.

Пусть в результате эксперимента получена точечная выборка $Y_n = \{y_1, y_2, \dots, y_n\}$, однако исследователя фактически интересует выборка $X_n = \{x_1, x_2, \dots, x_n\}$, $y_i = x_i + \Delta_i$, $\Delta_i = \delta_i |y_i| + \varepsilon_i$, где ε_i — абсолютная и δ_i — относительная погрешность измерения. Об относительной и абсолютной погрешности из технического паспорта прибора известно только, что $|\varepsilon_i| \leq \Delta_\varepsilon$, $|\delta_i| \leq \Delta_\delta$.

Таким образом, границы интервала, содержащего точное значение наблюдения, определяются по формулам: $a_i = y_i + \Delta_\delta |y_i| + \Delta_\varepsilon$,

$b_i = y_i - \Delta_\delta |y_i| - \Delta_\varepsilon$. В результате получаем интервальную выборку:
 $\mathbf{X}_n = \{[a_i, b_i], i = 1, \dots, n\}$ [54, 59, 66].

Пусть задана статистика $T: R_n \rightarrow R$. *Нотной* N_T статистики T называется максимально возможное абсолютное отклонение $T(Y_n)$ от $T(X_n)$, т.е. $N_T = N_T(\mathbf{X}_n) = \max_{X_n \in \mathbf{X}_n} |T(Y_n) - T(X_n)|$ [103, 104].

При достаточно малых ошибках Δ_i измерения, с точностью до $o(|\Delta_i|)$ выполняется $T(Y_n) - T(X_n) = \sum_{i=1}^n \frac{\partial T}{\partial x_i} \Delta_i$. Отсюда следует, что

$$N_T(X_n) = \Delta_\varepsilon \sum_{i=1}^n \left| \frac{\partial T}{\partial x_i} \right| + \Delta_\delta \sum_{i=1}^n x_i \left| \frac{\partial T}{\partial x_i} \right|.$$

Таким образом, нотна статистики критерия пропорциональна максимально допустимым погрешностям.

С учетом погрешностей измерений граничное значение t_α можно заменить на $t_\alpha + N_T(Y_n)$ [105]. Такая замена дает гарантию, что вероятность отклонения нулевой гипотезы H_0 , когда она верна, не более α . Исходя из принципа уравнивания погрешностей [102], можно определить *рациональный объем выборки* из условия

$$t_\alpha = N_T(Y_n).$$

Если рациональный объем выборки оказался меньше, чем необходимый объем выборки, то это означает, что различить гипотезы с заданными вероятностями ошибок первого и второго рода при таких погрешностях измерения невозможно.

Аналогичные результаты были получены автором в работе [59] на примере критерия Колмогорова.

2.11 Выводы

В данной главе рассмотрена методика оптимального планирования статистического эксперимента по различению двух простых гипотез, которая включает определение необходимого объема выборки и асимптотически оптимальное группирование.

Нижняя граница необходимого объема выборки для различения двух простых гипотез обратно пропорциональна симметричной дивергенции Кульбака-Лейблера $\rho(H_0, H_1)$, а сам необходимый объем выборки при малых равных вероятностях ошибок первого и второго рода вычисляется по формуле (2.21). Для общего случая предложен алгоритм оценивания необходимого объема выборки методом Монте-Карло.

Асимптотически оптимальное группирование для различения двух простых гипотез основано на максимизации дивергенции Кульбака-Лейблера. Показано, что при асимптотически оптимальном группировании максимизируется мощность критерия χ^2 .

Рассмотренная методика применена на двух численных примерах при различении нормального распределения против логистического и распределения Вейбулла-Гнеденко против гамма-распределения.

ГЛАВА 3. СОКРАЩЕНИЕ СРЕДНЕГО ОБЪЕМА ВЫБОРКИ В ПОСЛЕДОВАТЕЛЬНЫХ КРИТЕРИЯХ

Последовательный критерий отношения правдоподобия, разработанный Вальдом, является оптимальным по величине среднего объема выборки, требуемого для различения двух простых гипотез. Однако его оптимальность нарушается, если ни одна из двух 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» [194] были предложены различные процедуры усечения последовательного критерия отношения правдоподобия. Андерсоном [159] был сформулирован оптимальный обобщенный критерий отношения правдоподобия (GSPRT), который имеет ограниченную область неопределенности и минимальный средний объем выборки. Айвазяном в 1965 г. [2] была найдена линейная аппроксимация GSPRT в виде треугольной области неопределенности. Другое решение проблемы Кифера-Вейсса было предложено Лорденом в 1976 году [222] 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 [169] 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 [26, 27, 28, 29, 30]. В частности, в [26] был предложен 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, 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 $g(x, \theta)$.

3.1 Критерий Вальда

В последовательном критерии Вальда используется статистика критерия отношения правдоподобия (2.5), 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 (2.10). Графически процедура проверки гипотезы представлена на рисунке 3.1.

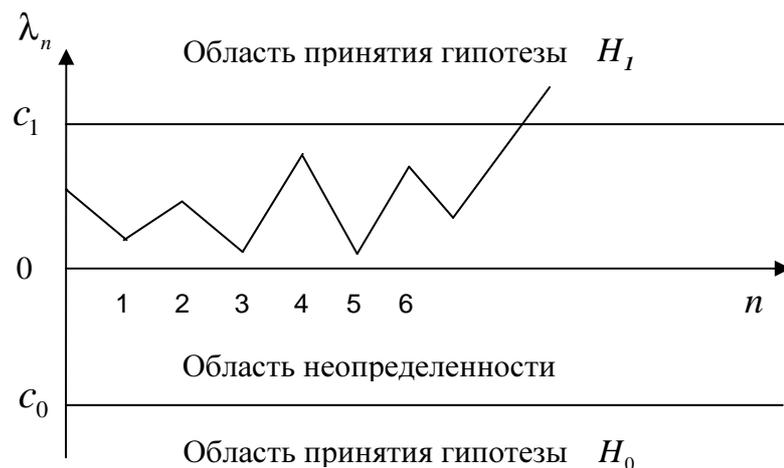


Рисунок 3.1 – Критическая область критерия Вальда

3.2 Критерий Айвазяна

Дж. Кифер и Л. Вейсс впервые рассмотрели возможность усовершенствования критерия Вальда определением границ критерия c_0 , c_1 в зависимости от текущего количества наблюдений n . Эта модификация получила название *обобщенного последовательного критерия отношения правдоподобия* (GSPRT). Он был разработан для проверки простых гипотез в ситуации, когда истинное значение параметра может быть равным θ' , где $\theta' \neq \theta_0$ и $\theta' \neq \theta_1$, т.е. принимает промежуточное относительно H_0 и H_1 значение [194].

Если аппроксимировать криволинейные границы GSPRT касательными, то получится оптимальный обобщенный последовательный критерий, полученный Айвазяном [2, 3].

Области принятия гипотез H_0 ($R_n^{H_0}$) и H_1 ($R_n^{H_1}$) и область неопределенности (R_n^*) задаются соотношениями:

$$R_n^{H_0} = \left\{ \lambda_n \leq -c_0 \left(1 - \frac{n\rho(H_0, H_1)}{8 \ln(1 / \min(\alpha, \beta))} \right) \right\}, \quad (3.1)$$

$$R_n^{H_1} = \left\{ \lambda_n \leq c_1 \left(1 - \frac{n\rho(H_0, H_1)}{8\ln(1/\min(\alpha, \beta))} \right) \right\}, \quad (3.2)$$

$$R_n^* = \left\{ -c_0 \left(1 - \frac{n\rho(H_0, H_1)}{8\ln(1/\min(\alpha, \beta))} \right) < \lambda_n < c_1 \left(1 - \frac{n\rho(H_0, H_1)}{8\ln(1/\min(\alpha, \beta))} \right) \right\}, \quad (3.3)$$

где c_0 и c_1 – положительные числа, равные соответственно $2 \cdot \ln \frac{1}{\beta}$ и $2 \cdot \ln \frac{1}{\alpha}$,

$\rho(H_0, H_1)$ – симметричная дивергенция Кульбака-Лейблера.

Если рассматривать области $R_n^{H_0}$ и $R_n^{H_1}$ в плоскости (n, λ_n) , то в последовательном критерии Вальда их границы задаются прямыми $\lambda_n = \ln \frac{\beta}{1-\alpha}$ и $\lambda_n = \ln \frac{1-\beta}{\alpha}$, параллельными горизонтальной оси, в то время как в критерии Айвазяна границами областей являются две сходящиеся прямые (рисунок 3.2), пересекающиеся в точке

$$n^* = -\frac{8\ln(\min(\alpha, \beta))}{\rho(H_0, H_1)}. \quad (3.4)$$

В таблице 3.1 приведены значения n^* для двух пар конкурирующих гипотез.

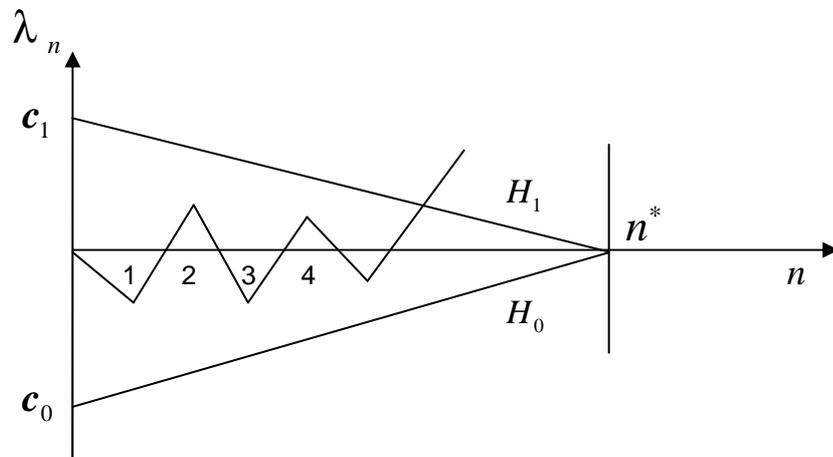


Рисунок 3.2 – Процедура принятия решения по критерию Айвазяна

Таблица 3.1 - Точка пересечения n^* критических границ критерия Айвазяна

Пара конкурирующих гипотез	Минимальная вероятность ошибки			
	0,01	0,05	0,1	0,15
Нормальное (2.35) против логистического (2.36)	1482	964	741	610
Вейбулла-Гнеденко (2.37) против гамма-распределения (2.38)	1054	685	527	434

3.3 Критерий Лордена

Критерий Лордена (или *двойной последовательный критерий отношения вероятностей*, 2-SPRT) представляет собой еще один способ усечения критерия Вальда, и основан на одновременной проверке двух пар гипотез [222].

В критерии Лордена используется вспомогательная гипотеза $H_\nu : f(x) = \nu f_0(x) + (1 - \nu)f_1(x)$ и параметр ν определяется из условия:

$$\rho(H_\nu, H_0) = \rho(H_\nu, H_1), \quad (3.5)$$

где $\rho(H_\nu, H_i)$ – симметричная дивергенция Кульбака-Лейблера (2.3). Для пары гипотез нормальное распределение (2.35) против логистического (2.36) имеем $\nu = 0.55261494$, а для проверки гипотезы о распределении Вейбулла-Гнеденко (2.37) против гамма-распределения (2.38) – $\nu = 0.525252$.

В критерии Лордена, графической интерпретацией которого является рисунок 3.3, одновременно рассматриваются две пары гипотез: H_0 против H_ν и H_ν против H_1 , однако процесс наблюдения продолжается до тех пор, пока не будет отвергнута одна из гипотез H_0 или H_1 . Если на шаге n статистика

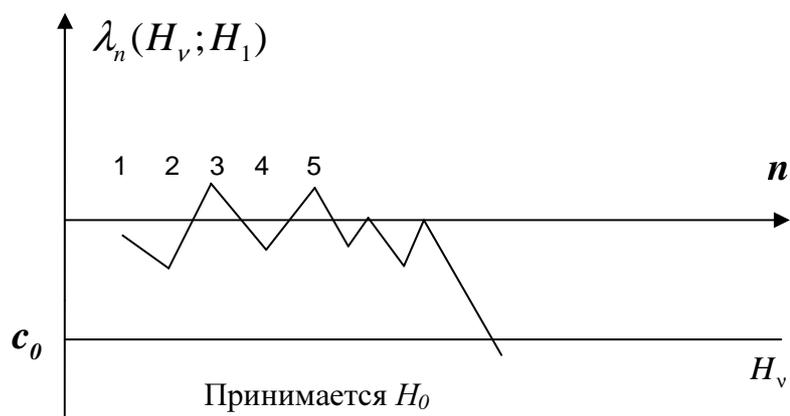
$$\lambda_n(H_\nu; H_1) = \sum_{i=1}^n \ln \frac{f_1(x_i)}{(1 - \nu)f_0(x_i) + \nu f_1(x_i)} < c_0, \quad (3.6)$$

то принимается гипотеза H_0 (рисунок 3.3(a)), если же

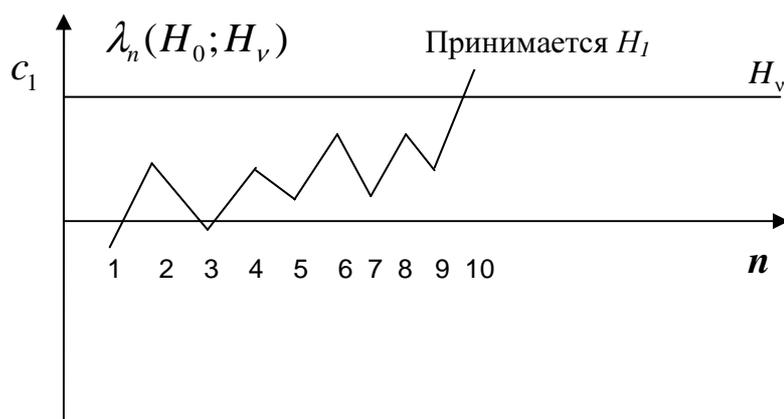
$$\lambda_n(H_0; H_\nu) = \sum_{i=1}^n \ln \frac{(1-\nu)f_0(x_i) + \nu f_1(x_i)}{f_0(x_i)} > c_1, \quad (3.7)$$

то принимается гипотеза H_1 (рисунок 3.3(б)). Приближенные критические границы находятся так же, как в критерии Вальда

$$c_0 \approx \ln \frac{\beta}{1-\alpha}, \quad c_1 \approx \ln \frac{1-\beta}{\alpha}. \quad (3.8)$$



(a)



(б)

Рисунок 3.3 – Процедура принятия решения по критерию Лордена

3.4 Оценивание точных критических границ методом Монте-Карло

Использование приближенных границ в последовательных критериях приводит к увеличению среднего объема выборки, однако, точные критические границы найти с помощью аналитических методов очень сложно – по причине того, что при вычислении вероятности ошибок первого и второго рода

$$1 - \alpha = P\{H_0 | H_0\} = \sum_{i=1}^{\infty} P\left\{\lambda_i < c_0 \bigcap_{j=1}^{i-1} \lambda_j \in [c_0, c_1] | H_0\right\}, \quad (3.9)$$

$$1 - \beta = P\{H_1 | H_1\} = \sum_{i=1}^{\infty} P\left\{\lambda_i > c_1 \bigcap_{j=1}^{i-1} \lambda_j \in [c_0, c_1] | H_1\right\} \quad (3.10)$$

случайные величины λ_i и λ_j 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 работах [113, 122, 123, 129], для чего может быть использован следующий алгоритм.

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 производится 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 $\hat{\beta}(c_0, c_1)$, пример которых показан на рисунках 3.4 и 3.5. Для нахождения точных критических границ c_0, c_1 необходимо решить систему

$$\begin{cases} \hat{\alpha}(c_0, c_1) = \alpha; \\ \hat{\beta}(c_0, c_1) = \beta. \end{cases} \quad (3.11)$$

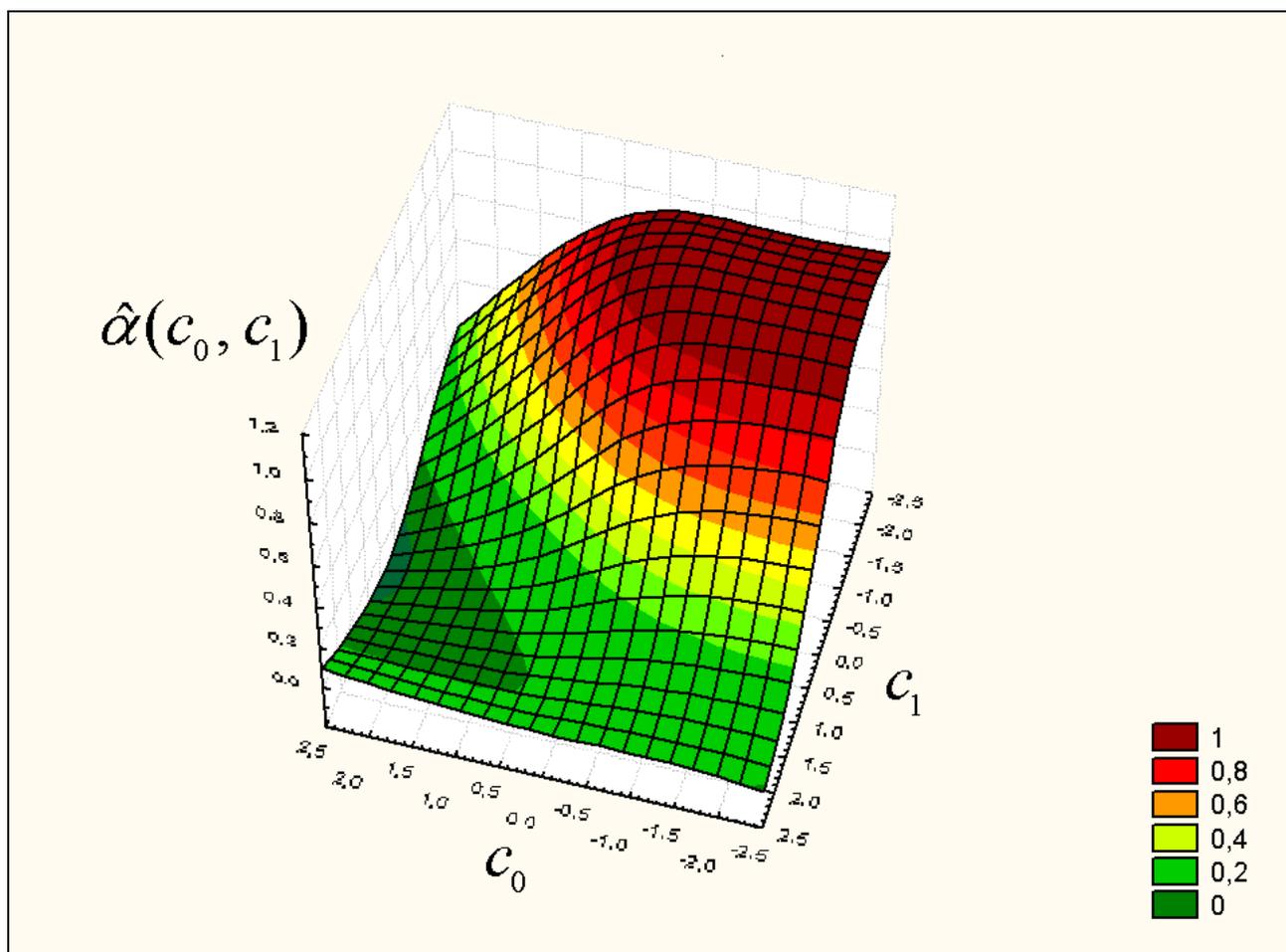


Рисунок 3.4 – Распределение вероятности ошибки первого рода по узлам сетки при проверке гипотезы о нормальном распределении против логистического распределения

Алгоритм 3.1 и решение системы (3.11) позволяют оценить точные критические границы, однако у данного подхода имеется ряд недостатков. Во-первых, моделирование вероятностей ошибок первого и второго рода требует достаточно много времени. Например, в таблице 3.2 приведено время моделирования при $N = 100000$ повторений на компьютере с процессором Intel Core 2 Quad Q6600 с тактовой частотой 2.4 ГГц. Во-вторых, алгоритм 3.1 дает оценки вероятностей ошибок первого и второго рода, но не дает самих критических границ, которые приходится находить, решая систему (3.11).

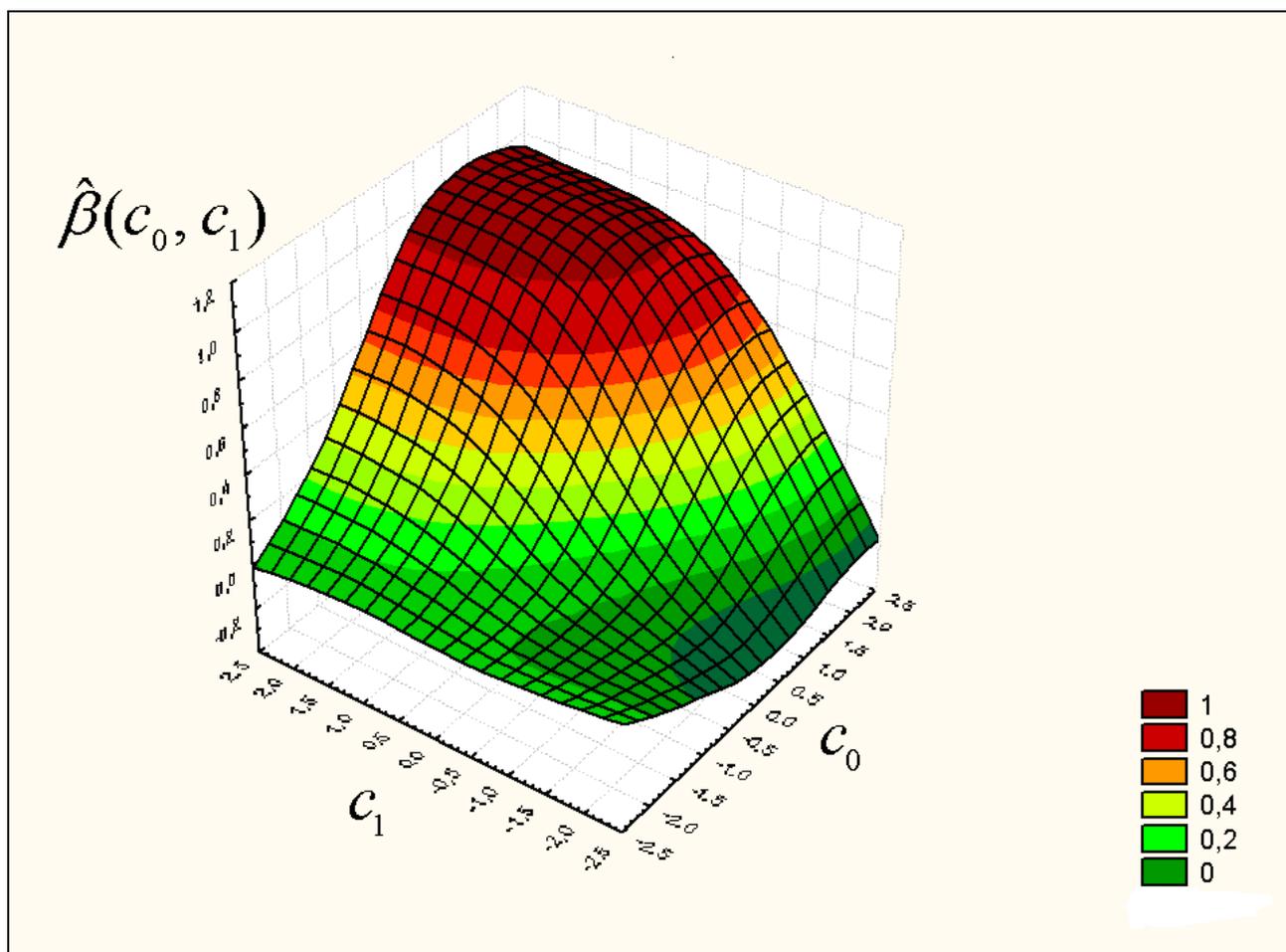


Рисунок 3.5 – Распределение вероятности ошибки второго рода по узлам сетки при проверке гипотезы о нормальном распределении против логистического распределения

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 (3.11), 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.

Таблица 3.2. Затраты времени на моделирование вероятностей ошибки первого и второго рода при проверке нормального распределения против логистического

Гипотеза	Область	Шаг	Время моделирования α , мин.	Время моделирования β , мин.	Общее время моделирования, мин.
Простая	$c_0 \in [-4, 6; 0]$ $c_1 \in [0; 4, 6]$	0,01	675	592	1267

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_0 и β_0 .
2. Строится сетка из $(k \times k)$ узлов ($k = 10$, подбирается эмпирически).
3. Вычисляются оценки $\alpha(c_0, c_1)$ и $\beta(c_0, c_1)$ по алгоритму 3.1.

4. Находится узел (i, j) , для которого достигается

$$d = \min_{i,j} \max \left(\left| \alpha(c_{0i}, c_{1j}) - \alpha_0 \right|, \left| \beta(c_{0i}, c_{1j}) - \beta_0 \right| \right)$$

5. Если $d \leq \varepsilon$ & $\max \left(t_{\gamma} \sqrt{\frac{\alpha_0(1-\alpha_0)}{N_{\alpha}}}, t_{\gamma} \sqrt{\frac{\beta_0(1-\beta_0)}{N_{\beta}}} \right) \leq \varepsilon$, 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, если на границе – то размер остается прежним.

7. Если погрешность моделирования больше ε , то $N_{\alpha} = N_{\alpha} * 2, N_{\beta} = N_{\beta} * 2$.

8. Переход на шаг 2.

Для наглядности построим линии равного уровня вероятностей ошибок первого и второго рода как функции зависимости вероятностей ошибок от границ критерия c_0, c_1 (рисунок 3.6). Точки пересечения этих линий дают искомые критические границы критерия Вальда при проверке гипотезы о нормальном распределении против гипотезы о логистическом законе распределения.

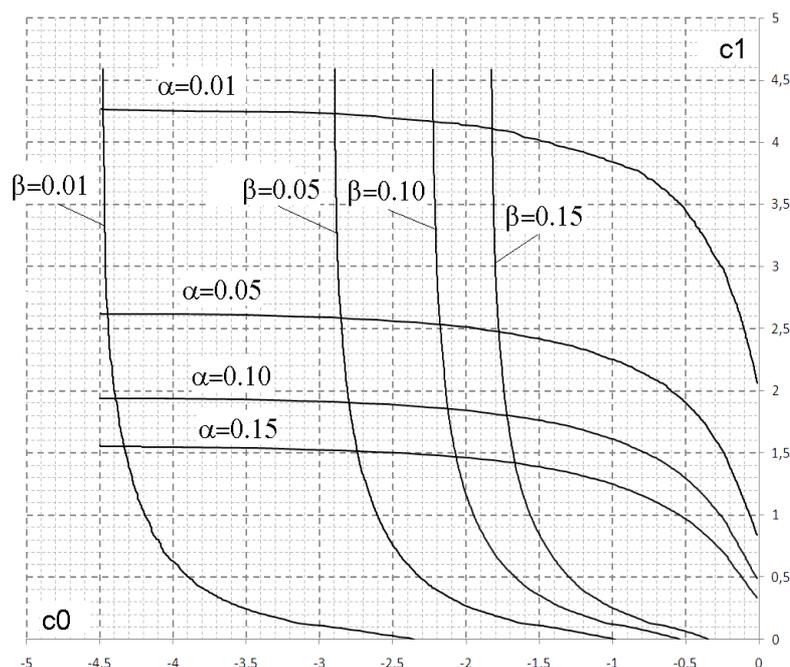


Рисунок 3.6 – Линии равного уровня вероятностей ошибок первого и второго рода, принимающих значения 0,05; 0,01; 0,1; 0,15; при различении гипотез по критерию Вальда

Рассчитанные по алгоритму 3.2 при $\varepsilon = 0,001$ оценки точных критических границ приведены в таблице 3.3. Объемы моделирования и затраченное время для вычисления приведены в таблице 3.4. Суммарное время на расчет 16 критических точек составляет 559 минут, что более чем в два раза меньше, чем время вычислений по алгоритму 3.1, причем сами критические значения посчитаны более точно, например, при $\alpha = \beta = 0,15$ количество повторений в методе Монте-Карло равно 5 120 000.

Аналогичным образом были получены оценки точных критических границ для критерия Лордена (таблица 3.5) и критерия Айвазяна (таблица 3.6).

Таблица 3.3 – Критические границы критерия Вальда при проверке гипотезы о нормальном законе распределения против гипотезы о логистическом распределении

$\alpha \backslash \beta$	0.15	0.10	0.05	0.01
0.15	-1.668; 1.416	-2.065; 1.463	-2.755; 1.514	-4.403; 1.555
0.10	-1.724; 1.801	-2.127; 1.858	-2.814; 1.904	-4.426; 1.944
0.05	-1.777; 2.479	-2.181; 2.533	-2.876; 2.585	-4.489; 2.636
0.01	-1.822; 4.141	-2.223; 4.149	-2.911; 4.184	-4.533; 4.224

Таблица 3.4 – Длительность расчета для нахождения оценок точных критических границ критерия Вальда при проверке гипотезы о нормальном законе распределения против гипотезы о логистическом распределении

$\alpha \backslash \beta$	0.15	0.10	0.05	0.01
0.15	5 120 000 32 мин.	5 120 000 38 мин	5 120 000 48 мин.	5 120 000 70 мин.
0.10	5 120 000 37 мин.	2 560 000 22 мин.	2 560 000 28 мин.	2 560 000 39 мин.
0.05	512 0000 47 мин.	2 560 000 28 мин	1 280 000 17 мин.	1 280 000 23 мин.
0.01	5 120 000 65 мин.	2 560 000 36 мин.	1 280 000 21 мин.	320 000, 7 мин.

Таблица 3.5 – Критические границы критерия Лордена при проверке гипотезы о нормальном законе распределения против логистического распределения

$\alpha \backslash \beta$	0.15	0.10	0.05	0.01
0.15	-1.169; 0.924	-1.463; 0.995	-2.010; 1.112	-3.274; 1.270
0.10	-1.251; 1.207	-1.574; 1.312	-2.104; 1.422	-3.334; 1.596
0.05	-1.359; 1.734	-1.683; 1.835	-2.245; 1.978	-3.598; 2.205
0.01	-1.505; 2.968	-1.859; 3.149	-2.430; 3.267	-3.829; 3.577

Таблица 3.6 - Критические границы критерия Айвазяна при проверке гипотезы о нормальном законе распределения против логистического распределения по полным данным

$\beta \backslash \alpha$	0.15	0.10	0.05	0.01
0.15	-2.105; 1.726 $n^* = 610$	-2.675; 1.762 $n^* = 741$	-3.678; 1.778 $n^* = 964$	-6.148; 1.747 $n^* = 1482$
0.10	-2,134; 2,220 $n^* = 741$	-2,849; 2,366 $n^* = 741$	-3,946; 2,366 $n^* = 964$	-6,484; 2,256 $n^* = 1482$
0.05	-2,149; 3,161 $n^* = 964$	-2,784; 3,283 $n^* = 964$	-4,152; 3,505 $n^* = 964$	-6,875; 2,636 $n^* = 1482$
0.01	-2,080; 5,370 $n^* = 1482$	-2,635; 4,149 $n^* = 1482$	-3,721; 5,790 $n^* = 1482$	-7,529; 6,405 $n^* = 1482$

3.5 Средний объем выборки до принятия решения по последовательным критериям при использовании оценок точных и приближенных критических границ

Сравним между собой функции среднего объема (ASN) выборки до принятия решения по последовательным критериям при использовании приближенных и оцененных точных критических границ, чтобы убедиться в том, что точные границы дают действительно существенный выигрыш по количеству необходимых для принятия решения наблюдений.

Возьмем оценки значений границ c_0, c_1 , при которых вероятность ошибок будет приближенно равна $\alpha = \beta \approx 0.05$ с точностью 0,01. Для достижения этой точности количество серий наблюдений (количество реализаций) зададим равным $N = 100000$.

Графики на рисунках 3.7-3.9, представляют собой эмпирические функции распределения требуемого для принятия решения объема выборки n по критериям Вальда, Айвазяна и Лордена.

Особенностью распределений требуемого для принятия решения объема выборки является их асимметричность – правый хвост распределения сильно

вытянут. При среднем значении объема выборки для критерия Вальда в 267 наблюдений при верной основной гипотезе и в 232 наблюдения при верной конкурирующей гипотезе встречались ситуации, когда необходимый объем выборки достигал 2000 наблюдений до принятия гипотезы. Поэтому усечение критерия, которое получается в критериях Айвазяна и Лордена, имеет смысл даже в классическом случае, когда верна одна из двух гипотез, не говоря уже о вполне возможном на практике варианте, когда неверна ни одна из рассматриваемых гипотез.

Использование оценок точных границ существенно сдвигает распределение требуемого объема выборки влево, причем для критериев Айвазяна и Лордена этот сдвиг более значителен.

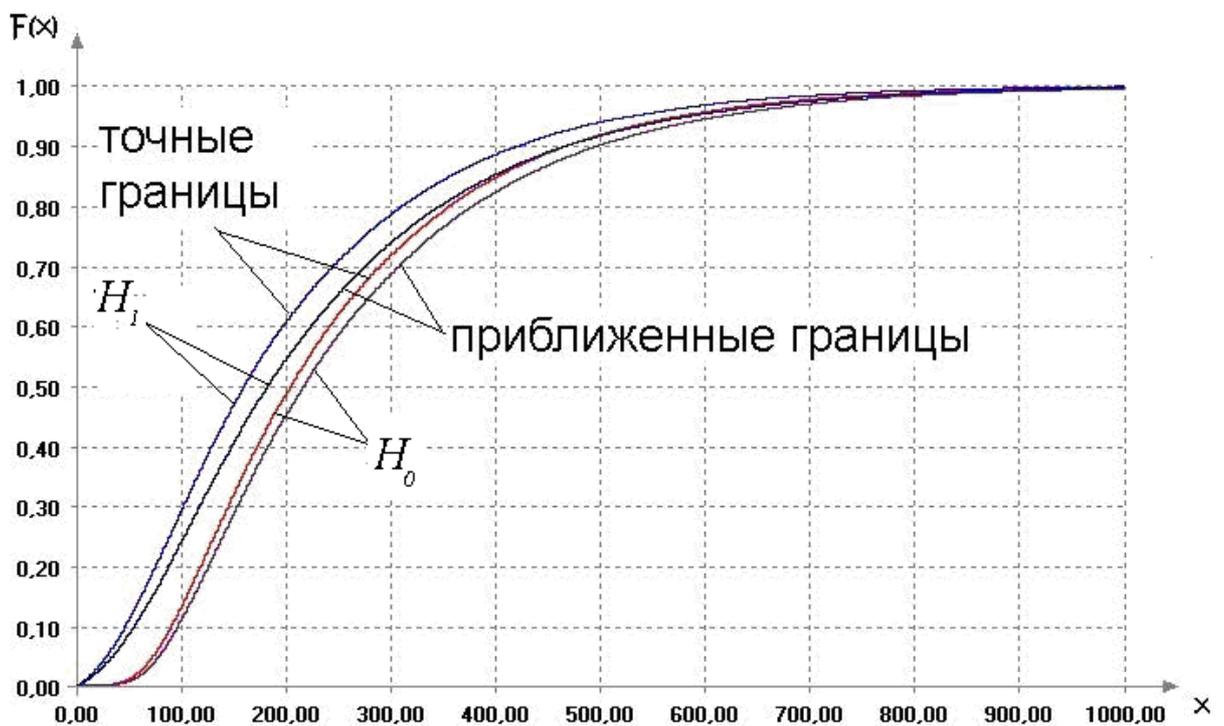


Рисунок 3.7 – Распределение требуемого объема выборки в критерии Вальда при использовании оценок точных и приближенных значений границ и $\alpha = \beta \approx 0.05$

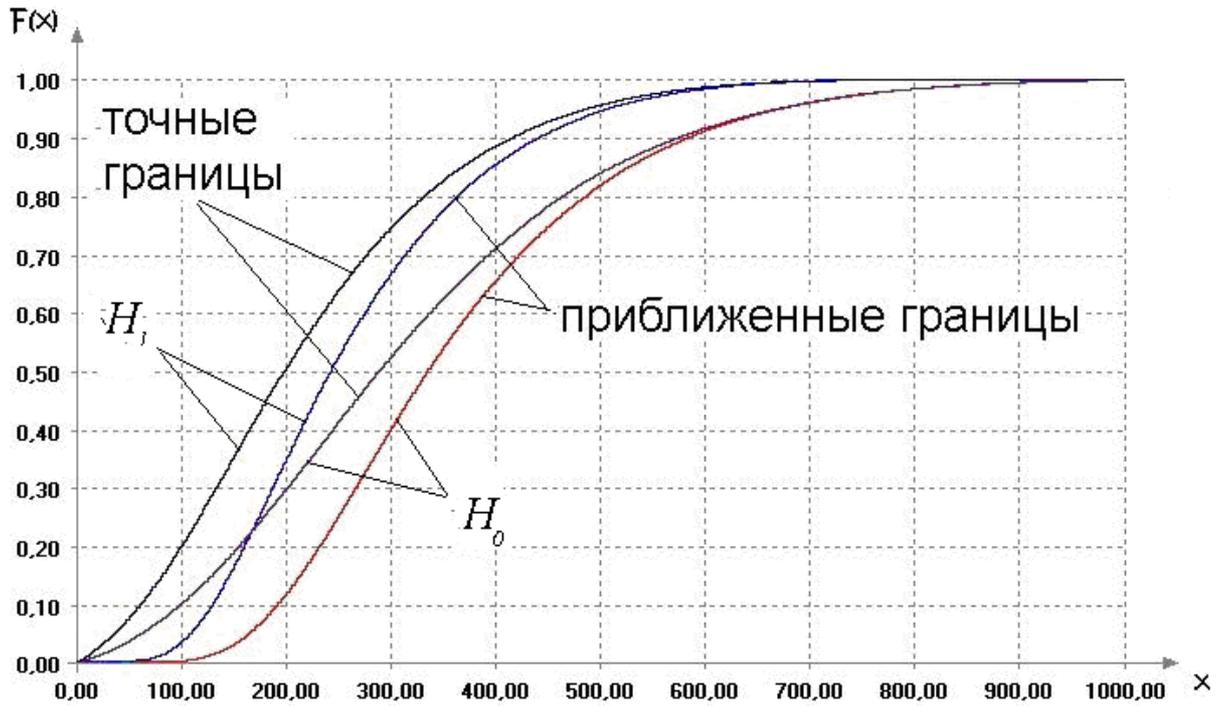


Рисунок 3.8 – Распределение требуемого объема выборки в критерии Лордена при использовании оценок точных и приближенных значений границ и $\alpha = \beta \approx 0.05$

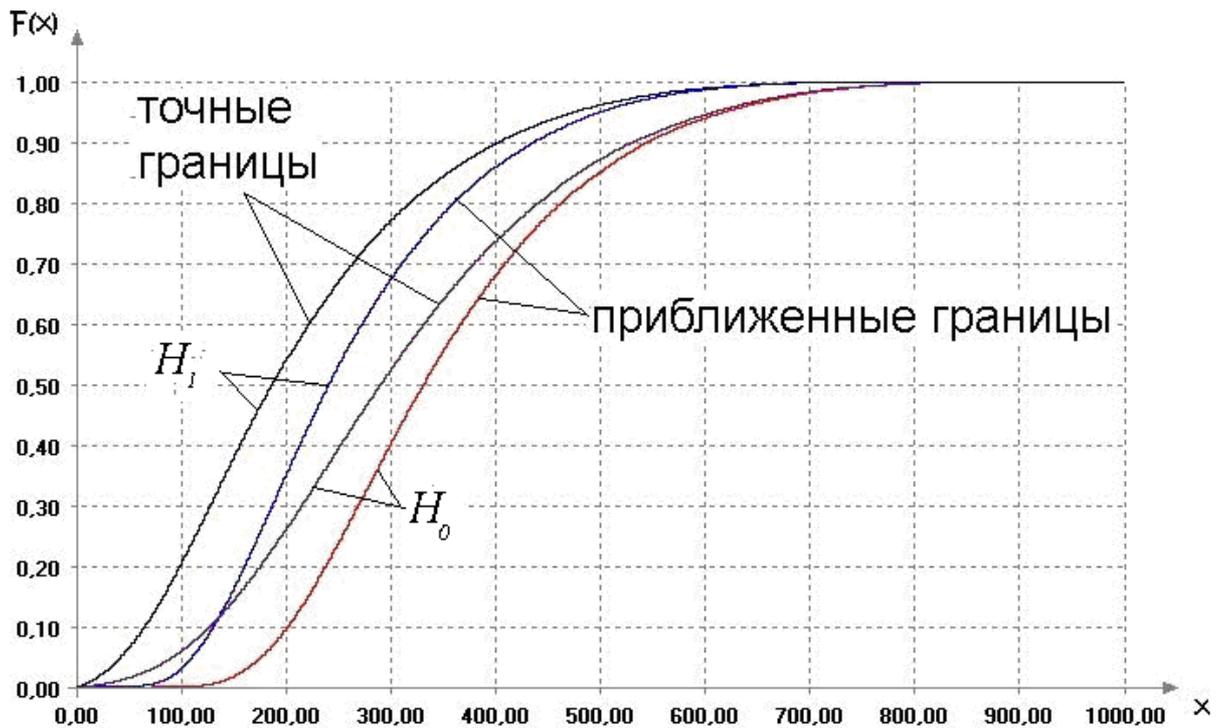


Рисунок 3.9 – Распределение требуемого объема выборки в критерии Айвазяна при использовании оценок точных и приближенных значений границ и $\alpha = \beta \approx 0.05$

В таблице 3.7 приведены данные о среднем требуемом объеме выборки наблюдений до принятия решения по последовательным критериям при справедливости основной и конкурирующей гипотез с использованием приближенных критических границ и оценок точных критических границ критериев при заданных значениях вероятностей ошибок. Сокращение среднего объема выборки приведено в процентах в таблице 3.7 и представлено графически на рисунках 3.10 и 3.11.

Интересно, что использование оценок точных границ существенно приблизило средний требуемый объем выборки усеченных критериев к оптимальному критерию. Например, если при использовании приближенных границ критерий Айвазяна уступал критерию Вальда на 25%, то при использовании оценок точных границ потеря в среднем объеме выборки составляет только 5%.

Таблица 3.7. Оценки среднего требуемого объема выборки наблюдений при проверке гипотез по последовательным критериям с использованием приближенных и оценок точных критических границ ($\alpha = \beta = 0,05$)

Используемые границы	Критерий Вальда		Критерий Лордена		Критерий Айвазяна	
	H_0	H_1	H_0	H_1	H_0	H_1
Приближенные	267	232	367	318	355	312
Оценки точных границ	251	206	268	223	265	214
Сокращение среднего объема выборки (%)	6%	11%	27%	30%	25%	31%

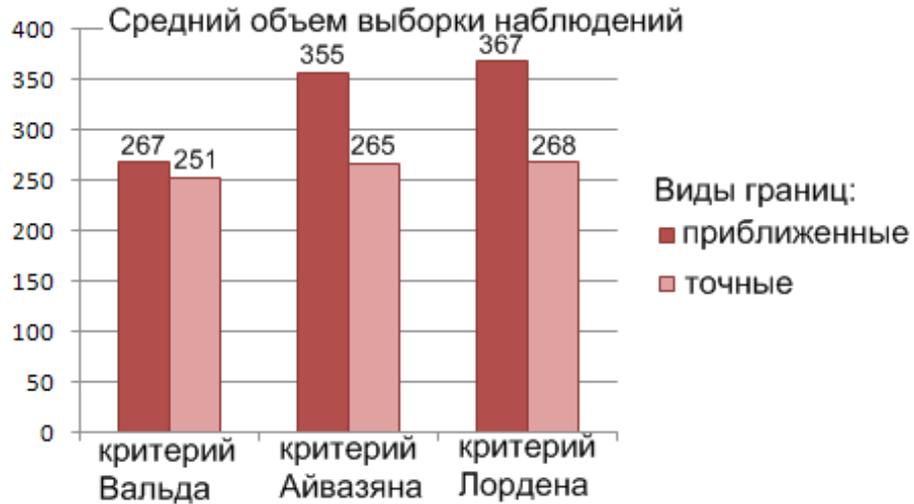


Рисунок 3.10 – Выигрыш по среднему требуемому объему выборки наблюдений до принятия решения при использовании оценок точных границ вместо приближенных значений при справедливости основной гипотезы H_0 ($\alpha = \beta = 0,05$)



Рисунок 3.11 – Выигрыш по среднему требуемому объему выборки наблюдений до принятия решения при использовании оценок точных границ вместо приближенных значений при справедливости альтернативной гипотезы H_1 ($\alpha = \beta = 0,05$)

3.6 Применение последовательных критериев к цензурированным наблюдениям

Цензурированные наблюдения встречаются в медицине, социальных науках, экономике, задачах контроля качества и многих других областях. В

экономике изучается "выживание" новых предприятий или времена "жизни" продуктов. В задачах контроля качества типичным является изучение "выживания" элементов изделий под нагрузкой (анализ времен отказов). Медицина изучает влияние различных видов вакцины на пациентов.

Цензурированные наблюдения нецелесообразно рассматривать как пропущенные, так как они содержат важную информацию, и при их исключении меняется закон распределения наблюдаемой случайной величины. Однако для выборок, содержащих такие наблюдения, необходимо применять отличные от классических методы оценки показателей и критерии проверки статистических гипотез. В частности, применение методов последовательного анализа в том виде, в котором они сформулированы в работах Вальда и его последователей, невозможно без некоторой адаптации. Подобная адаптация впервые встречается в работе Бангдивалы [164].

Использование метода Монте-Карло для определения критических границ по цензурированным данным рассмотрено в работах автора [120, 124, 130]. Используя априорную информацию, можно модифицировать статистики для последовательных критериев по цензурированной выборке следующим образом.

Введем обозначение:

$$q_i(x) = \begin{cases} f_i(x), & x - \text{нецензурировано} \\ 1 - F_i(x), & x - \text{цензурировано} \end{cases} \quad (3.12)$$

где $f_i(x)$ – функция плотности случайной величины ξ при верной гипотезе H_i , $i = 0, 1$; $F_i(x)$ – ее функция распределения.

Тогда статистика критерия отношения правдоподобия (2.4) примет вид:

$$\Lambda_n = \prod_{i=1}^n \frac{q_1(x_i)}{q_0(x_i)}, \quad (3.13)$$

где в правой части находится произведение отношений функций правдоподобия по нецензурированным наблюдениям или отношений функций выживаемости по моментам цензурирования (в зависимости от того, чем является x_i – наблюдением или моментом цензурирования).

Логарифм (3.13) имеет вид

$$\lambda_n = \sum_{i=1}^n \ln \left(\frac{q_1(x_i)}{q_0(x_i)} \right). \quad (3.14)$$

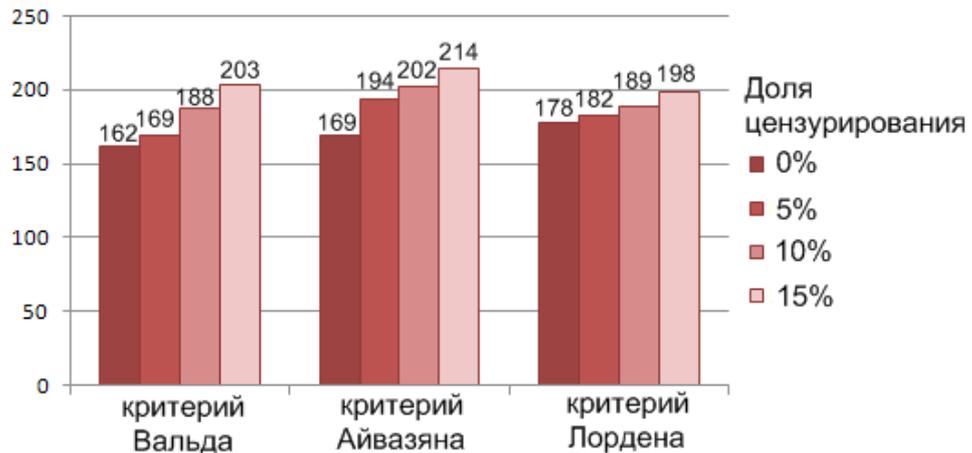
Границы критерия, базирующегося на статистике (3.14), не подчиняются классическим границам критериев Вальда, Лордена и Айвазяна. Численные значения критических границ для заданных простых гипотез при заданном уровне цензурирования можно найти методом Монте-Карло [130].

Характерной особенностью проверки гипотезы по цензурированным данным последовательными критериями является увеличение среднего требуемого объема выборки при увеличении степени цензурирования. В таблице 3.8 приведены средние значения требуемого объема выборки при проверке гипотезы о распределении Вейбулла-Гнеденко против гамма-распределения при различной степени цензурирования (в процентах).

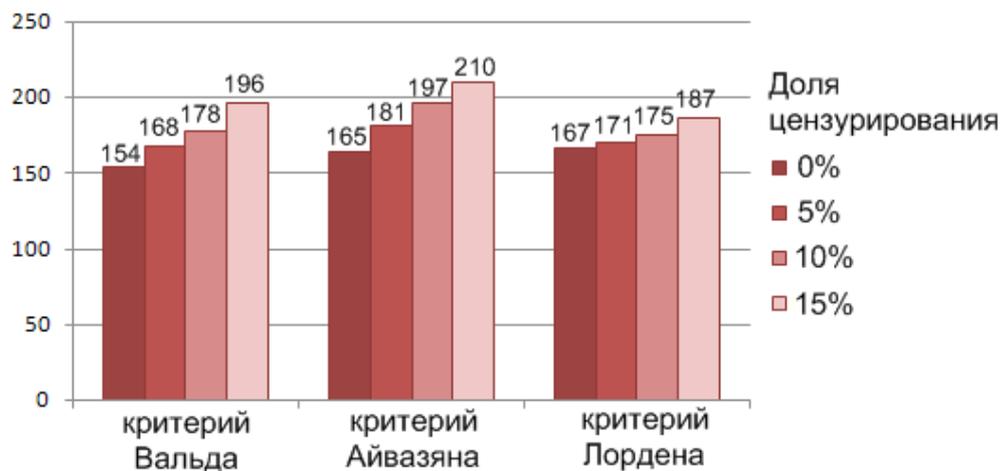
Таблица 3.8 - Оценки среднего значения требуемого объема выборки при проведении процедуры последовательной проверки гипотез по полным и случайно цензурированным данным ($\alpha = \beta = 0.05$)

Критерий \ Справедливая гипотеза	H_0				H_1			
	0%	5%	10%	15%	0%	5%	10%	15%
Процент цензурирования								
Вальда	162	169	188	203	154	168	178	196
Айвазяна	169	194	202	214	165	181	197	210
Лордена	178	182	189	198	167	171	175	187

Увеличение среднего значения требуемого объема выборки при увеличении процента цензурированных наблюдений характерно для всех последовательных критериев, однако, для критерия Лордена это увеличение меньше (рисунок 3.12), что позволяет рекомендовать использование критерия Лордена при проверке гипотез по цензурированным наблюдениям.



(a)



(б)

Рисунок 3.12 – Средний требуемый объем выборки последовательных критериев при проверке по данным с различной степенью цензурирования, вероятность ошибок $\alpha = \beta = 0.05$, (а) – справедлива гипотеза H_0 , (б) – справедлива гипотеза H_1

3.7 Вычисление критических границ для последовательного t-критерия

Пусть рассматривается задача проверки гипотезы о том, что неизвестное среднее значение θ нормального распределения равно некоторой определенной величине θ_0 , в то время как относительно дисперсии σ^2 этого распределения ничего неизвестно, т.е. $H_0: \theta = \theta_0$, а альтернативная ей гипотеза имеет вид: $H_1: \theta \neq \theta_0$.

Вальд в своей монографии [273] несколько модифицировал классическую постановку следующим образом. Он предложил определить такую положительную величину δ , что принятие гипотезы H_0 может рассматриваться как ошибка, имеющая практические последствия, только при $\left| \frac{\theta - \theta_0}{\sigma} \right| \geq \delta$.

В соответствии с этим определяется три области параметрического пространства. Область принятия ω_{np} состоит из всех точек (θ, σ) , для которых $\theta = \theta_0$, т.е. область ω_{np} состоит из всех точек (θ, σ) , где σ может принимать любое положительное значение. Область отклонения $\omega_{отк}$ состоит из всех точек (θ, σ) , для которых $\left| \frac{\theta - \theta_0}{\sigma} \right| \geq \delta$. Область безразличия $\omega_{без}$ состоит из всех точек (θ, σ) , для которых $0 < \left| \frac{\theta - \theta_0}{\sigma} \right| < \delta$ [273].

Функция плотности вероятностей выборки (x_1, \dots, x_n) , соответствующая нормальному распределению со средним значением θ и стандартным отклонением σ , определяется формулой [273]:

$$p_n = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \theta)^2 \right\}.$$

Критерий в общем случае основан на отношении правдоподобия $\frac{P_{1n}}{P_{0n}}$, где P_{0n} – некоторое взвешенное среднее p_n , соответствующее различным точкам (θ, σ) , принадлежащим к области ω_{np} , и P_{1n} – некоторое взвешенное среднее p_n , соответствующее различным точкам (θ, σ) , принадлежащим к области $\omega_{отк}$.

Тогда отношение правдоподобия имеет вид [273]:

$$S_n = \frac{P_{1n}}{P_{0n}} = \frac{\frac{1}{2} \int_0^{\infty} \frac{1}{\sigma^n} \left[\exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \theta_0 - \delta\sigma)^2 \right\} + \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \theta_0 + \delta\sigma)^2 \right\} \right] d\sigma}{\int_0^{\infty} \frac{1}{\sigma^n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \theta_0)^2 \right\} d\sigma} \quad (3.15)$$

Сама проверка проводится по последовательной схеме. Производятся дополнительные наблюдения до тех пор, пока выполняются неравенства $B < S_n < A$; гипотеза H_0 принимается, если на некотором этапе проверки $S_n \leq B$; если же $S_n \geq A$, то гипотеза H_0 отклоняется. Критические границы вычисляются по приближенным формулам [273]:

$$\tilde{A} = \frac{1 - \beta}{\alpha}, \quad (3.16)$$

$$\tilde{B} = \frac{\beta}{1 - \alpha}, \quad (3.17)$$

где α – вероятность ошибки первого рода, а β – вероятность ошибки второго рода.

В рамках проведенных исследований [131] удалось установить зависимость оценок точных критических границ относительно параметров θ_1 (значение параметра θ при верной гипотезе H_1), δ и σ – чем меньше значение $|\theta_0 - \theta_1|$, тем больше длина интервала $[A, B]$. При увеличении δ длина

интервала $[A, B]$ также увеличивается. При увеличении параметра σ длина интервала $[A, B]$ уменьшается.

Таким образом, поскольку значения σ и θ_1 неизвестны, применение метода Монте-Карло для оценивания критических границ A и B теряет смысл. Однако, если известно, что $\sigma < \bar{\sigma}$ и $\theta_1 > \underline{\theta}_1$, то с помощью метода Монте-Карло можно найти критические границы $[\underline{A}, \bar{B}]$ при $\sigma = \bar{\sigma}$ и $\theta_1 = \underline{\theta}_1$, которые будут лучше, чем границы $[\tilde{A}, \tilde{B}]$.

Сравнительный анализ использования приближенных $[\tilde{A}, \tilde{B}]$ и оценок точных критических границ $[\underline{A}, \bar{B}]$ показал довольно большой выигрыш в среднем объеме выборки. В некоторых задачах при использовании оценок точных критических границ $[\underline{A}, \bar{B}]$ для принятия или отклонения гипотезы средний объем выборки оказался на 65-75% меньше, чем при использовании приближенных критических границ $[\tilde{A}, \tilde{B}]$ [131].

3.8 Вычисление критических границ при проверке сложных гипотез

Применение критерия Вальда для проверки сложных гипотез затруднено тем, что теорема об оценках границ (2.10) в данном случае не имеет места [122]. Поэтому перед применением алгоритма нахождения критических границ необходимо найти область, на которой следует осуществлять моделирование. Для этого выполняется предварительное моделирование на заведомо большей области с большим шагом и малым количеством экспериментов [119].

На основании предварительного шага было выявлено, что при проверке *сложной* гипотезы о нормальном распределении с неизвестным математическим ожиданием (2.35) против логистического закона распределения (2.36) функции вероятности ошибок первого и второго рода

$\alpha(c_0, c_1)$ и $\beta(c_0, c_1)$ определены (т.е. отличны от 0 и 1 при заданном объеме моделирования) при $c_0 \in [-30; 2]$ и $c_1 \in [-1; 5]$.

Для проверки зависимости критических границ от истинного значения оцениваемого параметра было рассмотрено два варианта моделирования:

а) при верной основной гипотезе выборка моделировалась по стандартному нормальному закону;

б) при верной основной гипотезе выборка моделировалась по нормальному закону с параметром сдвига $\theta = 0,1$.

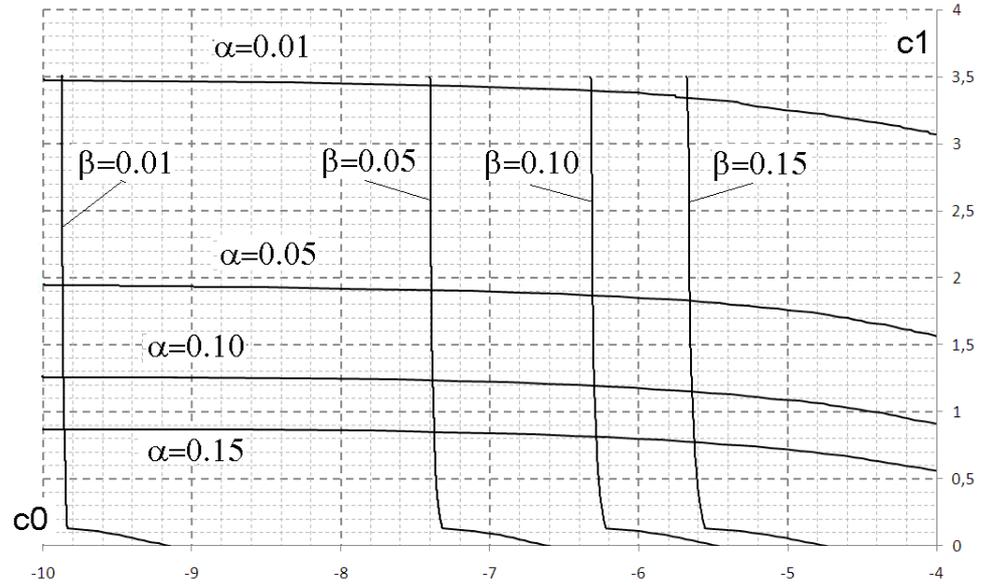
В результате проведенных исследований были получены функции $\alpha(c_0, c_1)$ и $\beta(c_0, c_1)$, на основании которых были построены линии равного уровня (рисунок 3.13).

Хорошо видно, что по сравнению с рисунком 3.6 линии равного уровня для $\alpha(c_0, c_1)$ и $\beta(c_0, c_1)$ сместились вниз и влево, в результате чего значения точек их пересечения стали существенно меньше. Данный результат вполне логичен, так как при оценивании параметров закона, соответствующего гипотезе H_0 , знаменатель отношения правдоподобия увеличивается, а статистика отношения правдоподобия уменьшается.

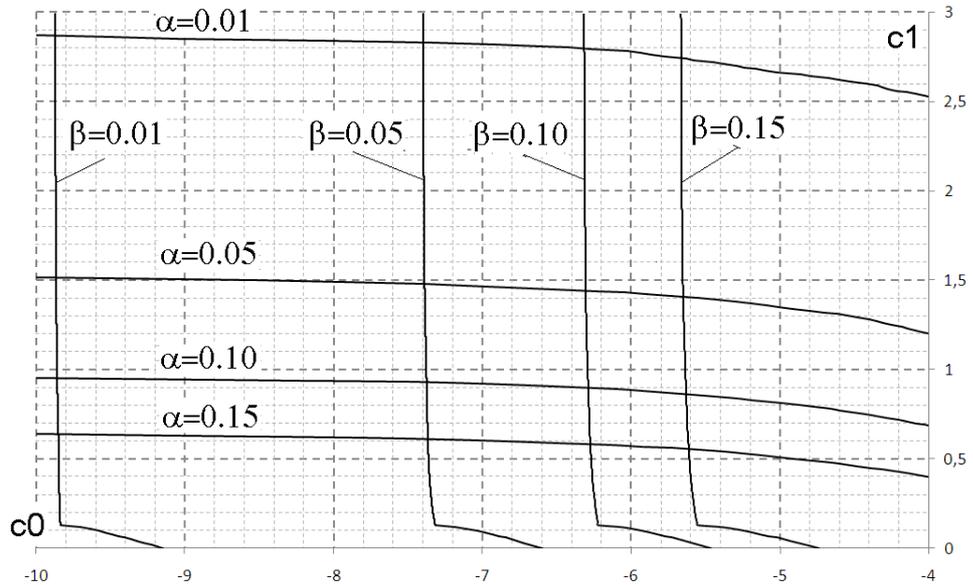
Также была выявлена зависимость критических границ от истинного значения оцениваемого параметра, причем при увеличении расстояния между основной и альтернативной гипотезой критические границы c_0 и c_1 становятся ближе.

Критические значения, полученные в результате пересечения линий равного уровня функций $\alpha(c_0, c_1)$ и $\beta(c_0, c_1)$, при проверке сложной гипотезы о нормальном распределении против простой конкурирующей гипотезы о логистическом законе (при моделировании стандартного нормального закона) приведены в таблице 3.9.

Таким образом, проведенные исследования показали, что применение метода Монте-Карло для оценивания точных границ при проверке сложных гипотез также возможно. Однако, полученные границы зависят от значений параметров, при которых моделируются распределение статистики при основной и конкурирующей гипотезах.



а)



б)

Рис. 3.13. Линии равного уровня функции $\alpha(c_0, c_1)$ и $\beta(c_0, c_1)$ при проверке сложной гипотезы о нормальном законе против простой конкурирующей гипотезы о логистическом законе, при моделировании: а) стандартного нормального распределения; б) нормального распределения со сдвигом 0,1

Таблица 3.9 - Критические значения c_0 , c_1 при проверке сложной гипотезы о нормальном распределении против простой конкурирующей гипотезы о логистическом законе

$\alpha \backslash \beta$	0.15	0.10	0.05	0.01
0.15	-5.63, 0.81	-6.30, 0.85	-7.37, 0.88	-9.93, 0.92
0.10	-5.64, 1.18	-6.31, 1.23	-7.38, 1.27	-9.93, 1.31
0.05	-5.66, 1.86	-6.32, 1.90	-7.39, 1.93	-9.93, 1.99
0.01	-5.67, 3.37	-6.33, 3.43	-7.40, 3.46	-9.94, 3.49

3.9 Выводы

Таким образом, в данном разделе показано, что 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% в случае нормального распределения против логистического и до 35% в случае распределения 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 [244, 245].

Предложен алгоритм оценивания точных критических границ для различения двух простых гипотез, который позволяет осуществить процесс поиска за сравнительно небольшое время. В частности, при проверке гипотезы о нормальном законе распределения против логистического закона

распределения максимальное время вычисления оценки одного критического значения критерия Вальда составляет порядка 70 минут при вероятностях ошибок первого и второго рода $\alpha=0.15$, $\beta=0.01$.

Использование метода Монте-Карло для нахождения критических границ позволяет расширить круг задач, решаемых методами последовательного анализа, проверкой гипотез по *цензурированным* и *группированным* данным.

Как показал рассмотренный пример проверки гипотезы о распределении Вейбулла-Гнеденко против гамма-распределения, средний объем выборки в последовательных критериях по цензурированным данным немного увеличивается по сравнению с теми же критериями по полным данным, причем средний объем выборки растет с ростом степени цензурирования. Наименее чувствительным к цензурированию (в смысле увеличения объема выборки) показал себя критерий Лордена, соответственно, при наличии цензурированных данных можно рекомендовать его к использованию в первую очередь.

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.

ГЛАВА 4. ИССЛЕДОВАНИЕ МОЩНОСТИ КРИТЕРИЕВ СОГЛАСИЯ ПРИ ПРОВЕРКЕ СЛОЖНЫХ ГИПОТЕЗ

Критерии согласия используются при проверке гипотез о виде распределения, когда конкурирующая гипотеза является отрицанием основной гипотезы.

В работах Б.Ю. Лемешко, С.Б. Лемешко, С.Н. Постовалова [92, 95, 96] 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.

Для случая проверки простых гипотез рассмотренные критерии можно расположить по мощности следующим образом:

χ^2 Пирсона (АОГ) \succ Ω^2 Андерсона-Дарлинга \succ ω^2 Мизеса
 \succ Колмогорова.

Такая шкала справедлива при использовании в критерии χ^2 Пирсона АОГ, при котором минимизируются потери в информации Фишера. В случае близких конкурирующих гипотез преимущество в мощности критерия χ^2 Пирсона может быть существенным.

При проверке сложных гипотез порядок предпочтения оказывается существенно иным:

Ω^2 Андерсона-Дарлинга \succ ω^2 Мизеса \succ Y_n^2 (АОГ) \succ χ^2 Пирсона (АОГ)
 \succ Колмогорова.

При очень близких гипотезах может быть:

Ω^2 Андерсона-Дарлинга $\succ Y_n^2$ (АОГ) $\succ \omega^2$ Мизеса $\succ \chi^2$ Пирсона (АОГ)
 \succ Колмогорова.

Указанные выводы носят интегрированный характер. Такое упорядочение не является жёстким. Иногда критерий имеет преимущество по мощности при одних значениях α и объемах выборок n и уступает при других значениях α и n . Мощность критериев типа χ^2 (Пирсона и Рао-Робсона-Никулина) зависит не только от гипотез H_0 , H_1 и объема выборок n , но при заданных гипотезах H_0 и H_1 – от способа группирования и числа интервалов. Число интервалов, при котором мощность критериев для пары альтернатив H_0 и H_1 максимальна, зависит от этих гипотез и от способа группирования. Увеличение числа интервалов не всегда приводит к росту мощности критериев типа χ^2 [88].

При близких гипотезах H_0 и H_1 при использовании критерия χ^2 Пирсона выбор АОГ дает положительный эффект как при простых, так и при сложных гипотезах. Однако это не означает, что применение АОГ всегда гарантирует максимальную мощность данного критерия. При конкретных и не очень близких гипотезах оптимальным может оказаться некоторый другой способ группирования, который может быть найден в результате максимизации мощности критерия.

Вывод о безоговорочно положительном эффекте применения АОГ нельзя распространять на критерий Рао-Робсона-Никулина: при одной и той же паре гипотез H_0 и H_1 при одном числе интервалов k критерий оказывается более мощным при АОГ, при другом k – более мощным при РВГ. Зависимость мощности от способа группирования оказывается более сложной и требует исследования.

В работе [96] 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 Б.Ю. Лемешко и С.Н. Постоваловым в 2001 году [55, 56, 60, 63, 65, 68, 69, 72, 76, 81, 82, 38, 97, 208].

В работах Лемешко, Постовалова и Французова [71, 79, 84, 85, 211] рассматривается применение критериев согласия для проверки адекватности непараметрических моделей при использовании ядерных оценок функции плотности с ядром Епанечникова.

В данной главе приведены результаты исследования мощности непараметрических критериев согласия при тех же альтернативах, что и в работе [96], 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*) [142].

4.1 Исследуемые критерии

4.1.1 Критерий Колмогорова

В критериях типа Колмогорова измеряемое расстояние между эмпирическим $F_n(x)$ и теоретическим $F(x, \theta)$ распределениями имеет вид

$$D_n = \sup_{|x| < \infty} |F_n(x) - F(x, \theta)|, \quad (4.1)$$

где n – объем выборки.

Предпочтительнее в критерии Колмогорова (Колмогорова-Смирнова) использовать статистику с поправкой Большева [7] вида

$$S_k = \frac{6nD_n + 1}{6\sqrt{n}}, \quad (4.2)$$

где

$$D_n = \max(D_n^+, D_n^-), \quad D_n^+ = \max_{1 \leq i \leq n} \left\{ \frac{i}{n} - F(x_i, \theta) \right\}, \quad D_n^- = \max_{1 \leq i \leq n} \left\{ F(x_i, \theta) - \frac{i-1}{n} \right\}, \quad (4.3)$$

и x_1, x_2, \dots, x_n – упорядоченные по возрастанию выборочные значения.

Статистика S_k при справедливости простой проверяемой гипотезы в пределе подчиняется закону Колмогорова, а в случае сложной гипотезы – различным законам, в зависимости от вида распределения и оцениваемых параметров. Статистические модели распределений статистик $G(S_k | H_0)$ для наиболее распространенных семейств законов распределений приведены в [97, 220].

Если для вычисленного по выборке значения статистики S_k^* выполняется неравенство $P\{S > S_k^*\} = 1 - G(S_k^* | H_0) > \alpha$, то нет оснований для отклонения гипотезы H_0 .

4.1.2 Критерии типа ω^2

В критериях типа ω^2 расстояние между гипотетическим и истинным распределениями рассматривают в квадратичной метрике. Статистика критерия [7] выражается соотношением

$$\begin{aligned} \omega_n^2[\psi(F)] &= \int_{-\infty}^{\infty} \left\{ E[F_n(x)] - F(x) \right\}^2 \psi(F(x)) dF(x) = \\ &= \frac{2}{n} \sum_{i=1}^n \left\{ g[F(x_i)] - \frac{2i-1}{2n} f[F(x_i)] \right\} + \int_0^1 (1-t)^2 \psi(t) dt, \end{aligned} \quad (4.4)$$

где

$$f(t) = \int_0^1 \psi(s) ds, \quad g(t) = \int_0^1 s\psi(s) ds.$$

При выборе $\psi(t) \equiv 1$ получается статистика критерия Крамера-Мизеса-Смирнова:

$$S_\omega = n\omega_n^2 = \frac{1}{12n} + \sum_{i=1}^n \left\{ F(x_i, \theta) - \frac{2i-1}{2n} \right\}^2. \quad (4.5)$$

При выборе $\psi(t) \equiv 1/t(1-t)$ получается статистика критерия Андерсона-Дарлинга:

$$S_\Omega = n\Omega_n^2 = -n - 2 \sum_{i=1}^n \left\{ \frac{2i-1}{2n} \ln F(x_i, \theta) + \left(1 - \frac{2i-1}{2n} \right) \ln(1 - F(x_i, \theta)) \right\}. \quad (4.6)$$

Статистики S_ω и S_Ω при простой гипотезе в пределе подчиняется законам $a_1(s)$ и $a_2(s)$, соответственно [7], а в случае сложной гипотезы – различным законам, в зависимости от вида распределения, числа и типа оцениваемых параметров, значений параметров формы, от метода оценивания. Статистические модели распределений статистик $G(S_\omega|H_0)$ и $G(S_\Omega|H_0)$ для наиболее распространенных семейств законов распределений приведены в [97, 220].

4.2 Методы оценивания

4.2.1 Метод максимального правдоподобия

Оценки максимального правдоподобия (ОМП) вычисляются в результате минимизации по θ функции правдоподобия

$$L(\theta) = \prod_{i=1}^n f(x_i, \theta), \quad (4.7)$$

где $f(x, \theta)$ – функция плотности закона распределения. Оценки максимального правдоподобия при достаточно общих условиях являются асимптотически несмещенными и асимптотически эффективными, что объясняет их популярность на практике. Однако ОМП не всегда являются устойчивыми к наличию аномальных наблюдений в выборке.

4.2.2 Методы минимального расстояния

При вычислении MD -оценок (оценок по методу минимального расстояния) по θ минимизируется некоторая мера близости, задающая расстояние $\rho(F(x, \theta), F_n(x))$ между теоретической и эмпирической функциями распределения. MD -оценки находятся в процессе решения задачи

$$\tilde{\theta} = \arg \min_{\theta} \rho(F(x, \theta), F_n(x)). \quad (4.8)$$

В качестве меры близости можно использовать статистики критериев согласия (4.1), (4.5), (4.6), задающие соответственно метод минимального расстояния Колмогорова, методы минимального расстояния ω^2 и Ω^2 .

4.2.3 Оценивание параметров по порядковым статистикам

Для нахождения оценок часто используются линейные комбинации порядковых статистик или выборочных квантилей. Такие оценки называются L -оценками. Они обладают двумя важными для практического применения качествами: простотой вычисления и хорошими свойствами робастности.

При построении L -оценок по выборочным квантилями $z_1 < z_2 < \dots < z_k$ рассматриваемого закона оценки находят в виде:

$$\hat{\theta} = \sum_{i=1}^k \alpha_i z(p_i), \quad z(p) = (x_{(np)} + x_{(np+1)}) / 2, \quad (4.9)$$

где $x_{(i)}$ – i -я порядковая статистика; α_i и p_i – набор коэффициентов и вероятностей, которыми определяется конкретная оценка; n – объем выборки.

Оптимальные несмещенные оценки для параметров сдвига и масштаба по выборочным квантилям получены в [237]. В [151] установлено, что оптимальные оценки параметров сдвига и масштаба являются асимптотически эффективными. В работах Б.Ю. Лемешко, Е.В. Чимитовой и С.Н. Постовалова [73, 75] предложено при построении таких оценок использовать асимптотически оптимальное группирование, а также получены соответствующие коэффициенты для L -оценок.

4.3 Сравнение мощности критериев согласия для пары гипотез «Нормальное распределение против логистического»

Функции распределения нормального и логистического законов, соответствующих конкурирующим гипотезам H_0 и H_1 , очень близки. В этом можно убедиться, сравнив их графики. А если по выборкам, соответствующим нормальному или логистическому закону, находить оценки параметров этих законов, то значения оценок параметров сдвига и масштаба практически совпадают.

В таблицах 4.1-4.3 приведены полученные в результате моделирования методом Монте-Карло значения мощности критериев согласия Колмогорова, Крамера-Мизеса-Смирнова и Андерсона-Дарлингга при использовании различных методов: ММП (колонка «ОМП»), оценивания по порядковым статистикам (колонка «L-оц.»), методов минимального расстояния ω^2 Мизеса (колонка «ММР ω^2 »), Ω^2 Мизеса (колонка «ММР Ω^2 ») и Колмогорова (колонка «ММР Колм.») при объемах выборок $n=100$ и 500 наблюдений. Уровень значимости критерия α (вероятность ошибки первого рода)

варьировался от 0,01 до 0,15. Полужирным шрифтом выделены наибольшие значения мощности в строке при одинаковом объеме выборки.

Объемы моделирования выбраны таким образом, чтобы статистическая погрешность приводимых в таблицах оценок мощности критериев при 95% доверительном интервале не превышала величины 10^{-3} .

В таблице 4.1 мощность определялась при оценивании обоих параметров, в таблице 4.2 – при оценивании только параметра масштаба, а в таблице 4.3 – при оценивании только параметра сдвига.

Сравнивая полученные значения мощности при оценивании обоих параметров нормального закона распределения, можно заметить, что рассмотренные критерии имеют бóльшую мощность при использовании метода максимального правдоподобия. Это хорошо видно на графиках функций распределения статистик критериев, приведенных на рисунках 4.1-4.3: распределения статистик критериев наиболее сильно сдвинуты вправо именно при использовании оценок максимального правдоподобия.

При оценивании только параметра масштаба нормального распределения при $n = 100$ мощность критериев согласия при оценивании по ММП высокая, но в ряде случаев уступает мощности при других методах оценивания. Однако уже при $n = 500$ мощность всех рассмотренных критериев выше при оценивании по ММП, за одним исключением: при $\alpha = 0.01$ мощность критерия Андерсона-Дарлинга выше при оценивании параметра масштаба по ММР Колмогорова.

При оценивании только параметра сдвига наблюдаемая картина меняется. При $n = 500$ выявляется следующая закономерность: мощность непараметрических критериев согласия выше при использовании метода минимального расстояния, соответствующего статистике критерия. Т.е. для

критерия Колмогорова это MMP Колмогорова, для критерия Крамера-Мизеса-Смирнова – MMP ω^2 , а для критерий Андерсона-Дарлинга – MMP Ω^2 .

Следует отметить, что мощность критериев при проверке сложной гипотезы в случае оценивания параметров сдвига и масштаба нормального распределения выше мощности, вычисленной при оценивании тем же методом только параметра сдвига, которая в свою очередь выше мощности при оценивании тем же методом только параметра масштаба.

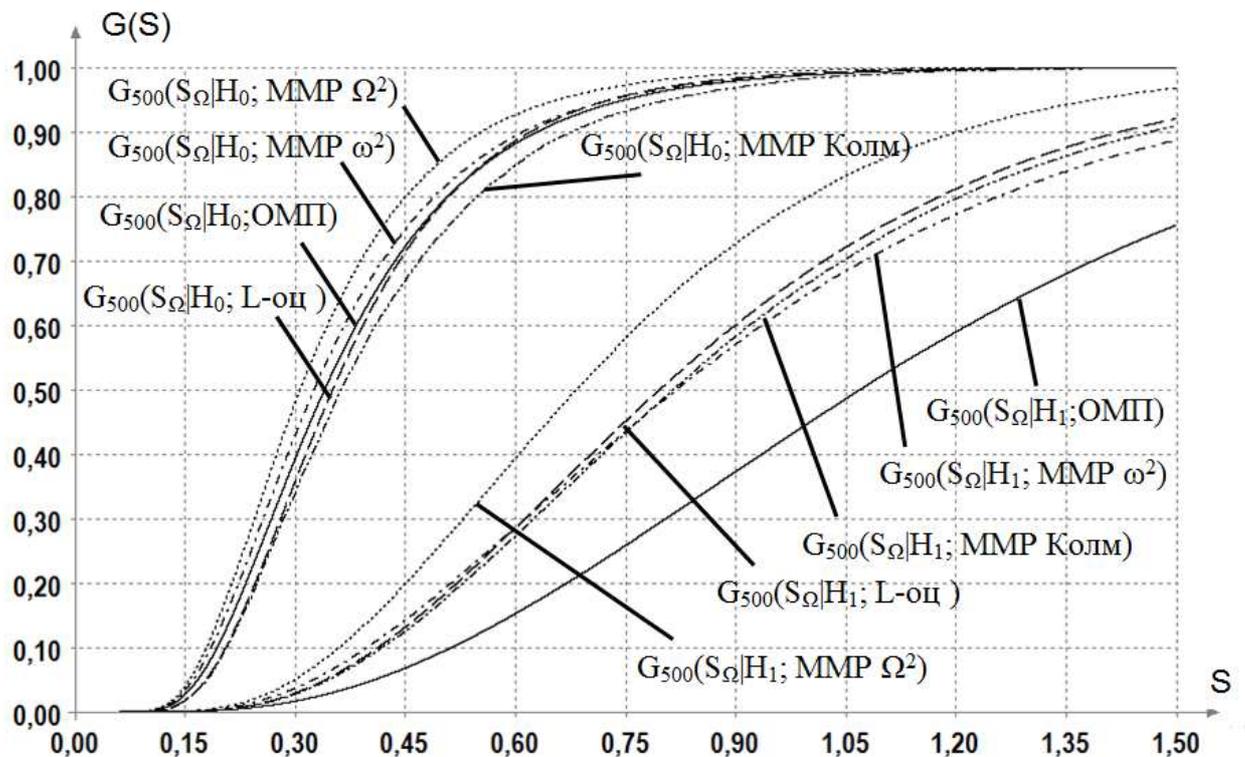


Рисунок 4.1 - Распределения статистики критерия Ω^2 Андерсона — Дарлинга $G(S_\Omega|H_0)$ и $G(S_\Omega|H_1)$ при проверке сложной гипотезы H_0 о согласии с нормальным законом в случае использования разных методов оценивания параметров

Таблица 4.1 - Мощность критериев согласия при оценивании разными методами параметров сдвига и масштаба нормального закона

α	$n=100$					$n=500$				
	ОМП	L-оц.	ММР ω^2	ММР Ω^2	ММР Колм.	ОМП	L-оц.	ММР ω^2	ММР Ω^2	ММР Колм.
	Мощность критерия Ω^2 Андерсона-Дарлинга									
0,150	0,401	0,365	0,375	0,345	0,337	0,835	0,762	0,773	0,745	0,725
0,100	0,325	0,288	0,297	0,266	0,259	0,828	0,691	0,704	0,667	0,643
0,050	0,226	0,190	0,197	0,166	0,166	0,742	0,571	0,582	0,530	0,506
0,025	0,155	0,124	0,129	0,102	0,101	0,649	0,457	0,469	0,405	0,381
0,010	0,093	0,069	0,071	0,052	0,054	0,528	0,328	0,336	0,271	0,253
	Мощность критерия ω^2 Крамера-Мизеса-Смирнова									
0,150	0,367	0,190	0,180	0,238	0,195	0,836	0,314	0,342	0,487	0,360
0,100	0,292	0,132	0,123	0,173	0,137	0,781	0,230	0,251	0,390	0,272
0,050	0,195	0,072	0,063	0,097	0,077	0,678	0,133	0,146	0,263	0,171
0,025	0,130	0,039	0,031	0,054	0,046	0,575	0,075	0,082	0,169	0,106
0,010	0,075	0,017	0,013	0,025	0,023	0,445	0,035	0,038	0,093	0,057
	Мощность критерия Колмогорова									
0,150	0,315	0,179	0,178	0,220	0,183	0,732	0,025	0,316	0,398	0,314
0,100	0,239	0,123	0,122	0,156	0,129	0,647	0,176	0,233	0,303	0,224
0,050	0,145	0,065	0,063	0,085	0,073	0,504	0,094	0,138	0,186	0,138
0,025	0,088	0,034	0,033	0,046	0,043	0,375	0,050	0,078	0,110	0,084
0,01	0,044	0,015	0,015	0,020	0,023	0,242	0,022	0,037	0,056	0,045

Таблица 4.2 - Мощность критериев согласия при оценивании разными методами параметра масштаба нормального закона

α	$n=100$					$n=500$				
	ОМП	L-оц.	ММР ω^2	ММР Ω^2	ММР Колм.	ОМП	L-оц.	ММР ω^2	ММР Ω^2	ММР Колм.
	Мощность критерия Ω^2 Андерсона-Дарлинга									
0,150	0,196	0,198	0,203	0,179	0,199	0,506	0,350	0,376	0,295	0,326
0,100	0,125	0,131	0,134	0,117	0,136	0,356	0,230	0,251	0,190	0,227
0,050	0,057	0,064	0,065	0,057	0,072	0,174	0,110	0,120	0,089	0,124
0,025	0,026	0,032	0,032	0,028	0,038	0,079	0,053	0,057	0,043	0,069
0,010	0,010	0,013	0,012	0,011	0,017	0,027	0,021	0,021	0,016	0,032
	Мощность критерия ω^2 Крамера-Мизеса-Смирнова									
0,150	0,175	0,151	0,152	0,156	0,152	0,403	0,166	0,164	0,185	0,168
0,100	0,111	0,100	0,101	0,103	0,101	0,265	0,110	0,109	0,122	0,113
0,050	0,052	0,049	0,050	0,051	0,051	0,120	0,054	0,054	0,060	0,058
0,025	0,024	0,024	0,025	0,025	0,025	0,053	0,026	0,027	0,029	0,029
0,010	0,009	0,009	0,010	0,010	0,010	0,018	0,010	0,011	0,012	0,012
	Мощность критерия Колмогорова									
0,150	0,194	0,149	0,154	0,171	0,152	0,435	0,170	0,170	0,227	0,165
0,100	0,129	0,098	0,104	0,115	0,101	0,319	0,133	0,114	0,155	0,107
0,050	0,064	0,048	0,052	0,059	0,049	0,181	0,056	0,058	0,081	0,055
0,025	0,032	0,023	0,027	0,030	0,023	0,100	0,026	0,029	0,042	0,026
0,010	0,013	0,009	0,011	0,013	0,009	0,044	0,011	0,012	0,017	0,012

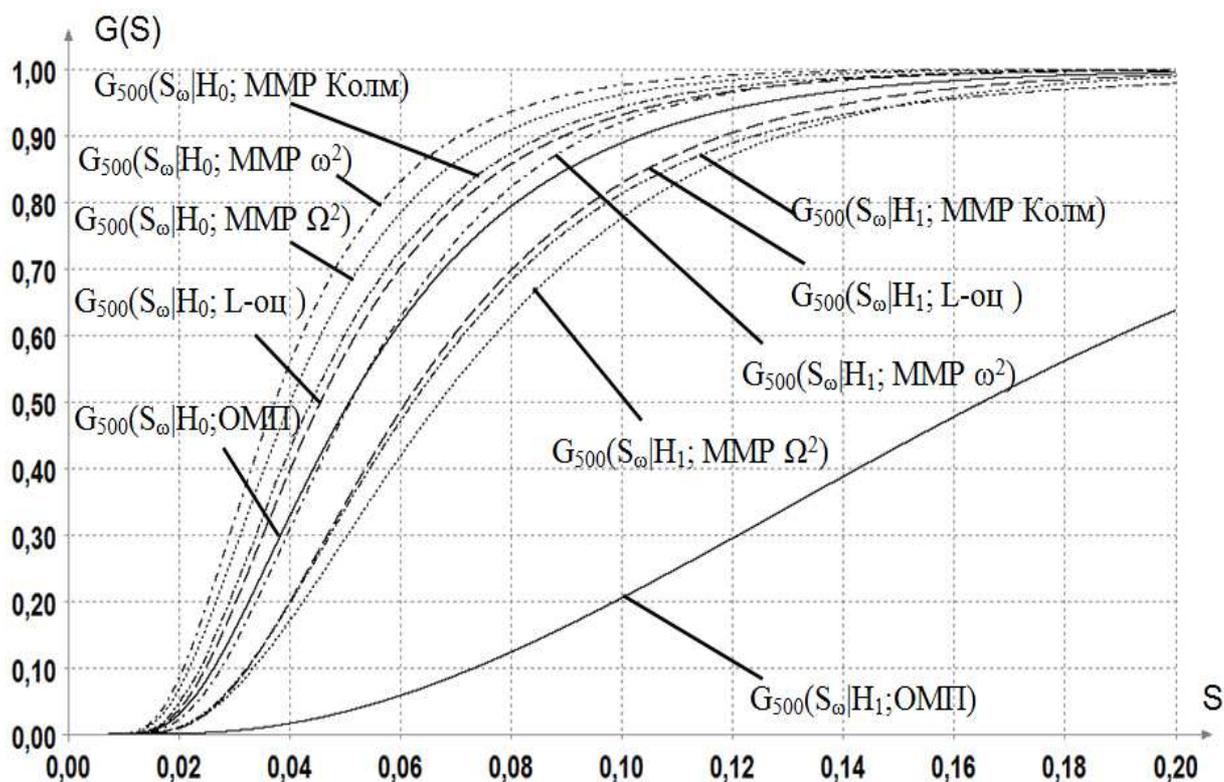


Рисунок 4.2 - Распределения статистики критерия ω^2 Крамера — Мизеса — Смирнова $G(S_\omega | H_0)$ и $G(S_\omega | H_1)$ при проверке сложной гипотезы H_0 о согласии с нормальным законом в случае использования разных методов оценивания параметров

4.4 Мощность критериев для пары гипотез «распределение Вейбулла-Гнеденко против гамма-распределения»

В таблицах 4.4-4.6 приведены полученные значения мощности критериев согласия Колмогорова, Крамера-Мизеса-Смирнова и Андерсона-Дарлинга для второй пары гипотез «распределение Вейбулла-Гнеденко против гамма-распределения». У распределения Вейбулла-Гнеденко теми же методами оценивались параметры масштаба и формы в разных комбинациях: оба параметра (таблица 4.4), только параметр масштаба (таблица 4.5) и только параметр формы (таблица 4.6).

Таблица 4.3 - Мощность критериев согласия при оценивании разными методами параметра сдвига нормального закона

α	$n=100$					$n=500$				
	ОМП	L-оц.	ММП ω^2	ММП Ω^2	ММП Колм.	ОМП	L-оц.	ММП ω^2	ММП Ω^2	ММП Колм.
	Мощность критерия Ω^2 Андерсона-Дарлинга									
0,150	0,349	0,342	0,340	0,345	0,338	0,776	0,777	0,773	0,780	0,772
0,100	0,265	0,259	0,256	0,261	0,256	0,700	0,702	0,698	0,706	0,695
0,050	0,162	0,157	0,155	0,160	0,149	0,570	0,571	0,564	0,571	0,564
0,025	0,097	0,093	0,092	0,095	0,089	0,447	0,449	0,441	0,451	0,441
0,010	0,047	0,045	0,044	0,045	0,044	0,310	0,311	0,307	0,315	0,305
	Мощность критерия ω^2 Крамера-Мизеса-Смирнова									
0,150	0,330	0,314	0,328	0,325	0,319	0,732	0,728	0,747	0,741	0,735
0,100	0,256	0,243	0,258	0,254	0,251	0,662	0,660	0,679	0,674	0,669
0,050	0,164	0,154	0,168	0,163	0,160	0,545	0,543	0,567	0,560	0,552
0,025	0,103	0,097	0,108	0,103	0,105	0,437	0,437	0,461	0,457	0,446
0,010	0,055	0,051	0,058	0,056	0,054	0,317	0,318	0,336	0,332	0,334
	Мощность критерия Колмогорова									
0,150	0,300	0,279	0,298	0,295	0,311	0,663	0,653	0,685	0,678	0,705
0,100	0,225	0,207	0,225	0,225	0,240	0,579	0,569	0,604	0,600	0,639
0,050	0,136	0,122	0,139	0,139	0,150	0,445	0,436	0,479	0,475	0,516
0,025	0,081	0,071	0,085	0,084	0,091	0,331	0,324	0,367	0,367	0,410
0,010	0,040	0,034	0,041	0,042	0,049	0,215	0,210	0,249	0,253	0,304

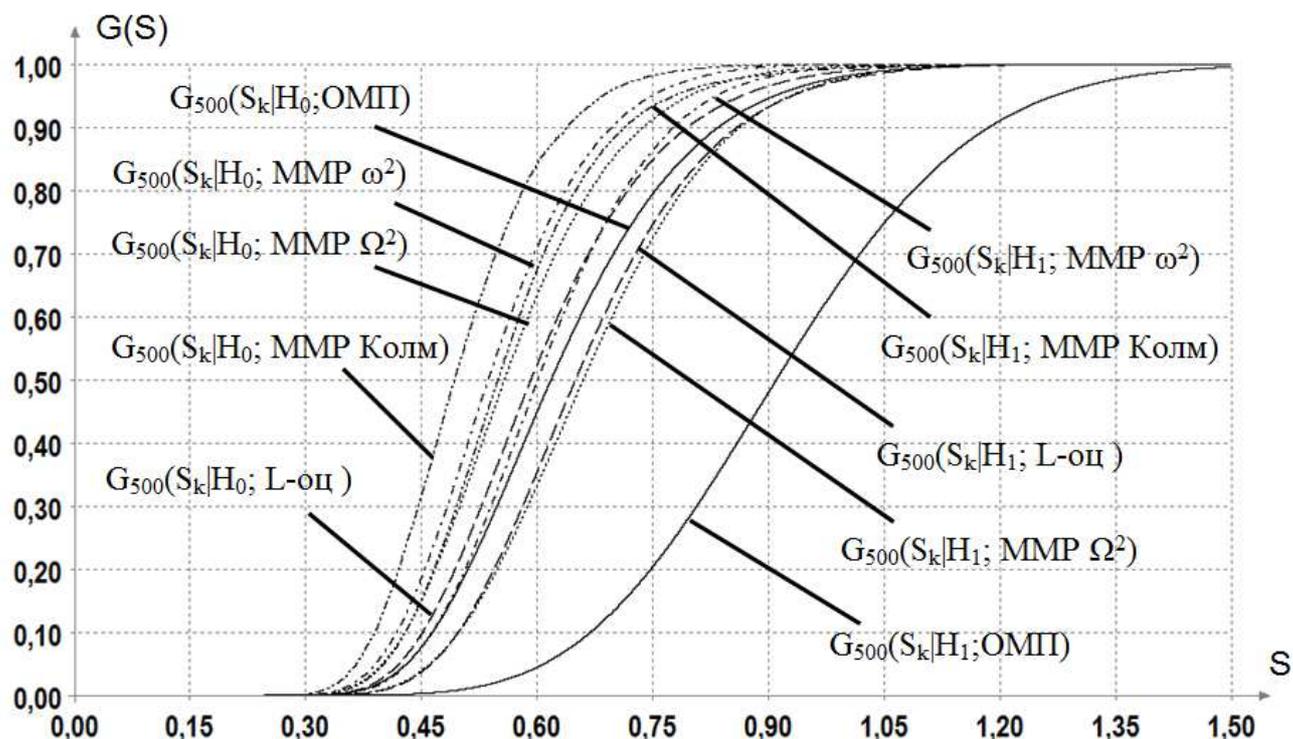


Рис. 4.3. Распределения статистики критерия Колмогорова $G(S_k|H_0)$ и $G(S_k|H_1)$ при проверке сложной гипотезы H_0 о согласии с нормальным законом в случае использования разных методов оценивания параметров

При оценивании обоих параметров наибольшая мощность оказалась выше, также как и для первой пары гипотез, при использовании метода максимального правдоподобия.

Похожая ситуация оказалась и при оценивании только параметра масштаба распределения Вейбулла-Гнеденко: мощность всех критериев выше при использовании ММП уже при $n = 100$.

Однако при оценивании только параметра формы распределения Вейбулла-Гнеденко мощность оказалась выше при использовании L -оценок.

Следует отметить, что мощность критериев проверки сложной гипотезы при оценивании параметров масштаба и формы распределения Вейбулла-Гнеденко выше мощности, вычисленной при оценивании тем же методом только параметра формы, которая в свою очередь выше мощности при оценивании тем же методом только параметра масштаба.

Таблица 4.4 - Мощность критериев согласия при различных методах оценивания параметров масштаба и формы распределения Вейбулла-Гнеденко

α	$n=100$					$n=500$				
	ОМП	L-оц.	ММР ω^2	ММР Ω^2	ММР Колм.	ОМП	L-оц.	ММР ω^2	ММР Ω^2	ММР Колм.
	Мощность критерия Ω^2 Андерсона-Дарлинга									
0,150	0,436	0,339	0,421	0,421	0,379	0,964	0,868	0,934	0,939	0,899
0,100	0,356	0,253	0,340	0,340	0,301	0,945	0,801	0,902	0,910	0,851
0,050	0,247	0,150	0,233	0,232	0,204	0,904	0,667	0,834	0,847	0,752
0,025	0,170	0,087	0,156	0,156	0,140	0,852	0,523	0,753	0,771	0,638
0,010	0,100	0,042	0,089	0,090	0,087	0,768	0,356	0,637	0,657	0,484
	Мощность критерия ω^2 Крамера-Мизеса-Смирнова									
0,150	0,398	0,277	0,294	0,325	0,246	0,929	0,657	0,758	0,798	0,676
0,100	0,318	0,206	0,221	0,251	0,184	0,897	0,566	0,683	0,731	0,575
0,050	0,214	0,124	0,135	0,159	0,114	0,829	0,426	0,556	0,614	0,412
0,025	0,142	0,074	0,082	0,100	0,074	0,752	0,310	0,440	0,501	0,277
0,010	0,082	0,037	0,041	0,054	0,047	0,641	0,198	0,311	0,371	0,146
	Мощность критерия Колмогорова									
0,150	0,344	0,257	0,275	0,282	0,205	0,858	0,557	0,707	0,683	0,574
0,100	0,264	0,189	0,204	0,209	0,150	0,799	0,464	0,623	0,591	0,465
0,050	0,166	0,111	0,121	0,124	0,091	0,688	0,331	0,486	0,446	0,302
0,025	0,102	0,065	0,071	0,072	0,059	0,574	0,232	0,368	0,325	0,188
0,010	0,053	0,032	0,035	0,035	0,035	0,432	0,141	0,244	0,203	0,097

Таблица 4.5 - Мощность критериев согласия при различных методах оценивания параметра масштаба распределения Вейбулла-Гнеденко

α	$n=100$					$n=500$				
	ОМП	L-оц.	ММР ω^2	ММР Ω^2	ММР Колм.	ОМП	L-оц.	ММР ω^2	ММР Ω^2	ММР Колм.
	Мощность критерия Ω^2 Андерсона-Дарлинга									
0,150	0,428	0,339	0,369	0,346	0,344	0,941	0,868	0,898	0,886	0,877
0,100	0,348	0,253	0,285	0,261	0,259	0,910	0,801	0,847	0,828	0,814
0,050	0,242	0,150	0,180	0,156	0,152	0,842	0,667	0,742	0,709	0,686
0,025	0,167	0,087	0,112	0,091	0,089	0,759	0,523	0,624	0,577	0,550
0,010	0,101	0,042	0,058	0,043	0,041	0,638	0,356	0,469	0,411	0,380
	Мощность критерия ω^2 Крамера-Мизеса-Смирнова									
0,150	0,441	0,277	0,217	0,261	0,219	0,936	0,657	0,608	0,705	0,631
0,100	0,366	0,206	0,147	0,186	0,150	0,906	0,566	0,497	0,608	0,520
0,050	0,264	0,124	0,075	0,103	0,076	0,845	0,426	0,333	0,452	0,354
0,025	0,189	0,074	0,037	0,055	0,036	0,772	0,310	0,208	0,318	0,225
0,010	0,120	0,037	0,014	0,024	0,014	0,667	0,198	0,105	0,188	0,114
	Мощность критерия Колмогорова									
0,150	0,408	0,257	0,210	0,243	0,213	0,900	0,557	0,558	0,600	0,586
0,100	0,332	0,189	0,145	0,174	0,145	0,859	0,464	0,453	0,498	0,478
0,050	0,230	0,111	0,076	0,097	0,075	0,776	0,331	0,305	0,348	0,327
0,025	0,159	0,065	0,039	0,053	0,038	0,685	0,232	0,199	0,233	0,211
0,010	0,095	0,032	0,017	0,024	0,015	0,561	0,141	0,109	0,133	0,112

Таблица 4.6 - Мощность критериев согласия при различных методах оценивания параметра формы распределения Вейбулла-Гнеденко

α	$n=100$					$n=500$				
	ОМП	L-оц.	ММР ω^2	ММР Ω^2	ММР Колм.	ОМП	L-оц.	ММР ω^2	ММР Ω^2	ММР Колм.
	Мощность критерия Ω^2 Андерсона-Дарлинга									
0,150	0,277	0,302	0,269	0,277	0,276	0,751	0,781	0,739	0,753	0,729
0,100	0,198	0,222	0,189	0,196	0,202	0,637	0,690	0,629	0,644	0,629
0,050	0,111	0,130	0,101	0,107	0,113	0,454	0,534	0,456	0,471	0,479
0,025	0,063	0,076	0,052	0,057	0,062	0,311	0,396	0,319	0,331	0,360
0,010	0,030	0,037	0,020	0,024	0,024	0,184	0,257	0,191	0,200	0,242
	Мощность критерия ω^2 Крамера-Мизеса-Смирнова									
0,150	0,276	0,295	0,274	0,272	0,266	0,654	0,715	0,683	0,679	0,673
0,100	0,203	0,223	0,200	0,198	0,202	0,558	0,638	0,595	0,591	0,584
0,050	0,121	0,138	0,113	0,111	0,113	0,414	0,508	0,455	0,452	0,454
0,025	0,072	0,084	0,061	0,060	0,057	0,300	0,393	0,336	0,333	0,336
0,010	0,036	0,043	0,026	0,025	0,024	0,191	0,270	0,214	0,213	0,228
	Мощность критерия Колмогорова									
0,150	0,281	0,294	0,263	0,269	0,257	0,637	0,700	0,641	0,658	0,664
0,100	0,211	0,225	0,193	0,199	0,186	0,551	0,627	0,554	0,577	0,582
0,050	0,129	0,141	0,112	0,116	0,106	0,422	0,508	0,420	0,447	0,448
0,025	0,077	0,088	0,064	0,066	0,052	0,317	0,402	0,309	0,336	0,333
0,010	0,039	0,046	0,029	0,030	0,023	0,209	0,282	0,167	0,221	0,210

4.5 Выводы

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* [73, 90] по выборочным квантилям.

Таким образом, можно сделать следующие выводы:

1. Распределения статистик непараметрических критериев согласия при проверке сложных гипотез зависят от вида проверяемой гипотезы и существенно зависят от метода оценивания [72].

2. Метод оценивания параметров оказывает существенное влияние на мощность непараметрических критериев [76].

3. Если подходить к выбору метода оценивания с позиций максимизации мощности используемого критерия, то в общем случае предпочтение следует отдать, как и было показано в [76], методу максимального правдоподобия. К тому же оценки максимального правдоподобия обладают наилучшими статистическими свойствами.

4. 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 [142].

ГЛАВА 5. СРАВНИТЕЛЬНЫЙ АНАЛИЗ МОЩНОСТИ КРИТЕРИЕВ ОДНОРОДНОСТИ ЗАКОНОВ 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 большинстве случаев критерии однородности являются двухвыборочной модификацией критериев согласия. Так критерий однородности Смирнова [152] соответствует критерию согласия Колмогорова, критерий однородности Лемана-Розенблатта [205, 251, 178] – критерию согласия Крамера-Мизеса-Смирнова, критерий однородности Андерсона-Дарлинга-Петита [241] – критерию согласия Андерсона-Дарлинга, критерий согласия Хи-квадрат – критерию однородности Хи-квадрат.

Другой подход к проверке однородности – это критерии, основанные на расстоянии между ядерными оценками функций плотности первой и второй выборки в метрике L_1 и L_2 [158].

Иногда в качестве критериев однородности используются критерии однородности средних или дисперсий, однако следует понимать, что их применение как критериев однородности возможно только при наложении ограничений на вид распределения случайной величины.

В работе Б.Ю. Лемешко, С.Б. Лемешко [91] представлены результаты исследований критериев однородности Смирнова и Лемана-Розенблатта. Эти критерии являются двухвыборочными аналогами критериев согласия Колмогорова и Крамера-Мизеса-Смирнова соответственно. Однако по результатам статьи [96] по сравнительному анализу мощности критериев согласия относительно ряда конкурирующих гипотез из рассмотренных

непараметрических критериев согласия Колмогорова, Крамера-Мизеса-Смирнова и Андерсона-Дарлинга наиболее мощным оказался критерий Андерсона-Дарлинга.

По-видимому, первое исследование критерия Андерсона-Дарлинга как критерия однородности было проведено Петитом в 1976 году [241]. Позднее Шольц и Стефенс обобщили этот критерий на случай k выборок [259]. В данной главе проведены исследования мощности критерия однородности Андерсона-Дарлинга-Петита для ряда фиксированных альтернатив и выполнено его сравнение с другими критериями однородности.

5.1 Гипотеза однородности

Гипотеза однородности формулируется следующим образом: пусть имеются две независимые случайные выборки $X = (x_1, x_2, \dots, x_m)$ и $Y = (y_1, y_2, \dots, y_n)$ объемов m и n соответственно. Выборке X соответствует функция распределения $F(x)$, выборке Y – функция распределения $G(x)$. Проверяемая нулевая гипотеза H_0 имеет вид: $F(x) = G(x)$ против конкурирующей $H_1: F(x) \neq G(x)$. Функции $F(x)$ и $G(x)$ будем считать непрерывными.

5.1.1 Критерий Смирнова

Критерий однородности Смирнова предложен в работе [152]. В критерии Смирнова используется статистика вида:

$$S_C = \sqrt{\frac{mn}{m+n}} D_{mn}, \quad (5.1)$$

где

$$D_{mn}(X, Y) = \sup_{-\infty < x < \infty} |F_m(x) - G_n(x)|, \quad (5.2)$$

$F_m(x)$ и $G_n(x)$ – эмпирические функции распределения, построенные по выборкам X и Y соответственно.

На практике значение $D_{mn}(X, Y)$ рекомендуется вычислять в соответствии с выражением (5.3):

$$D_{mn}(X, Y) = \max(D_{mn}^+, D_{mn}^-), \quad (5.3)$$

где D_{mn}^+ и D_{mn}^- вычисляются по формуле

$$D_{mn}^+ = \max_{1 < i < m} \left\{ \frac{i}{m} - F_n(x_i) \right\}, \quad D_{mn}^- = \max_{1 < i < m} \left\{ F_n(x_i) - \frac{i-1}{m} \right\}, \quad (5.4)$$

где x_i – i -й элемент упорядоченной по возрастанию выборки X .

Критерий Смирнова не зависит от конкретного вида распределений $F(x)$ и $G(x)$, и при стремлении объемов выборок к бесконечности статистика (5.4) сходится к распределению Колмогорова. Однако при малых значениях объемов m и n распределение статистики (5.4) может значительно отклоняться от предельного закона, в связи с чем в работе [91] была предложена модификация критерия Смирнова вида

$$S_{CM} = \sqrt{\frac{mn}{m+n}} \left(D_{m,n} + \frac{m+n}{4.6mn} \right). \quad (5.5)$$

При использовании критерия Смирнова рекомендуется брать объемы выборок m и n , представляющие собой взаимно простые числа [91].

5.1.2 Критерии типа ω^2

В критериях типа ω^2 расстояние между гипотетическим и истинным распределениями рассматривается в квадратичной метрике, в соответствии с выражением:

$$\int_{-\infty}^{\infty} \{E[F_n(x)] - F(x)\}^2 \psi(F(x)) dF(x), \quad (5.6)$$

где $E[\cdot]$ - оператор математического ожидания.

Данный критерий применим для проверки согласия полученного опытного распределения с теоретическим. Если же в данном критерии сделать переход от $F(x)$ функции распределения вероятностей, с которой проверяется согласие наблюдаемой выборки к эмпирической функции распределения $F_m(x)$, то данный критерий можно использовать для проверки гипотезы однородности. В этом случае статистика критерия будет опираться на разность эмпирических функций распределений. Одним из таких критериев является критерий Лемана–Розенблатта.

5.1.3 Критерий Лемана–Розенблатта

Поскольку при проверке однородности в (5.6) имеется два равноправных распределения, то можно рассматривать двухвыборочный аналог, в котором используется статистика вида:

$$T = \frac{mn}{m+n} \int_{-\infty}^{\infty} [G_m(x) - F_n(x)]^2 dH_{m+n}(x), \quad (5.7)$$

где $H_{m+n}(x) = \frac{m}{m+n} G_m(x) + \frac{n}{m+n} F_n(x)$ – эмпирическая функция распределения, построенная по вариационному ряду объединения двух выборок.

Данный критерий был предложен Леманом [205] и исследован Розенблаттом [251]. Как правило, критерий используют со статистикой вида

$$T = \frac{1}{mn(m+n)} \left[n \sum_{i=1}^n (r_i - i)^2 + m \sum_{j=1}^m (s_j - j)^2 \right] - \frac{4mn-1}{6(m+n)}, \quad (5.8)$$

где r_i – порядковый номер (ранг) наблюдения y_i , s_j – порядковый номер (ранг) наблюдения x_j в общем вариационном ряде, построенном по объединенной выборке $X \cup Y$ [7].

В [251] Розенблаттом было показано, что статистика (5.7) в пределе распределена как $a1(t)$, то есть выполняется:

$$\lim_{\substack{m \rightarrow \infty \\ n \rightarrow \infty}} P\{T < t\} = a1(t). \quad (5.9)$$

5.1.4 Критерий однородности Андерсона-Дарлингга-Петита

В критерии Лемана-Розенблатта весовая функция $\psi(t)$ предполагается равной единице. Если в данном критерии положить весовую функцию $\psi(t)$ равной $\frac{1}{t(1-t)}$, то получится двухвыборочный аналог критерия Андерсона-Дарлингга

$$A^2 = \frac{mn}{m+n} \int_{-\infty}^{\infty} \frac{[G_m(x) - F_n(x)]^2}{(1 - H_{m+n}(x))H_{m+n}(x)} dH_{m+n}(x). \quad (5.10)$$

В работе [241] получено простое выражение для статистики (5.10)

$$A^2 = \frac{1}{mn} \sum_{i=1}^{m+n-1} \frac{(M_i(m+n) - mi)^2}{i(m+n-i)}, \quad (5.11)$$

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 (п. 1.8).

Моделировались выборки одинакового объема m и n от 5 до 50, подчиняющиеся:

- а) стандартному нормальному закону распределения;
- б) стандартному экспоненциальному закону.

По каждой выборке вычислялась статистика A^2 . Процедура повторялась N раз (во всех исследованиях $N=1660000$). Расстояние до предельного закона распределения $D_{n,N}$ вычислялась по формуле (1.10). Результаты моделирования приведены в таблице 5.1, а график зависимости $D_{n,N}$ от $n=t$ на рисунке 5.1. Хорошо видно, что расстояние $D_{n,N}$ уменьшается с ростом числа наблюдений в выборках и при $n=t > 15$ описывается степенной зависимостью с высоким коэффициентом детерминации. Скорость сходимости равна $O(n^{-0,93})$ и практически не зависит от закона распределения моделируемых случайных величин, а различия объясняются погрешностью моделирования.

Для того чтобы расстояние (1.10) от распределения статистики A^2 до предельного закона распределения $a_2(t)$ не превышало 0,01 достаточно взять выборки объемом 45 наблюдений.

Таблица 5.1 – Поведение расстояния распределения статистики критерия до предельного закона для разных законом моделирования выборок с использованием статистики Колмогорова

$n = m$	$D_{n,N}$ для выборок из экспоненци- ального закона	$D_{n,N}$ для выборок из нормаль- ного закона	$n = m$	$D_{n,N}$ для выборок из экспонен- циального закона	$D_{n,N}$ для выборок из нормаль- ного закона
5	0,1508	0,1508	28	0,0140	0,0140
6	0,1061	0,1061	29	0,0139	0,0139
7	0,0861	0,0861	30	0,0133	0,0133
8	0,0684	0,0684	31	0,0126	0,0126
9	0,0574	0,0574	32	0,0123	0,0123
10	0,0451	0,0451	33	0,0121	0,0119
11	0,0388	0,0388	34	0,0115	0,0119
12	0,0324	0,0324	35	0,0113	0,0118
13	0,0291	0,0291	36	0,0114	0,0109
14	0,0277	0,0277	37	0,0108	0,0108
15	0,0255	0,0255	38	0,0107	0,0105
16	0,0241	0,0241	39	0,0102	0,0102
17	0,0228	0,0228	40	0,0102	0,0098
18	0,0214	0,0214	41	0,0096	0,0100
19	0,0201	0,0201	42	0,0102	0,0098
20	0,0195	0,0195	43	0,0097	0,0095
21	0,0187	0,0187	44	0,0093	0,0090
22	0,0172	0,0172	45	0,0089	0,0088
23	0,0169	0,0169	46	0,0088	0,0094
24	0,0161	0,0161	47	0,0088	0,0088
25	0,0155	0,0155	48	0,0089	0,0086
26	0,0156	0,0156	49	0,0083	0,0083
27	0,0145	0,0145	50	0,0082	0,0083

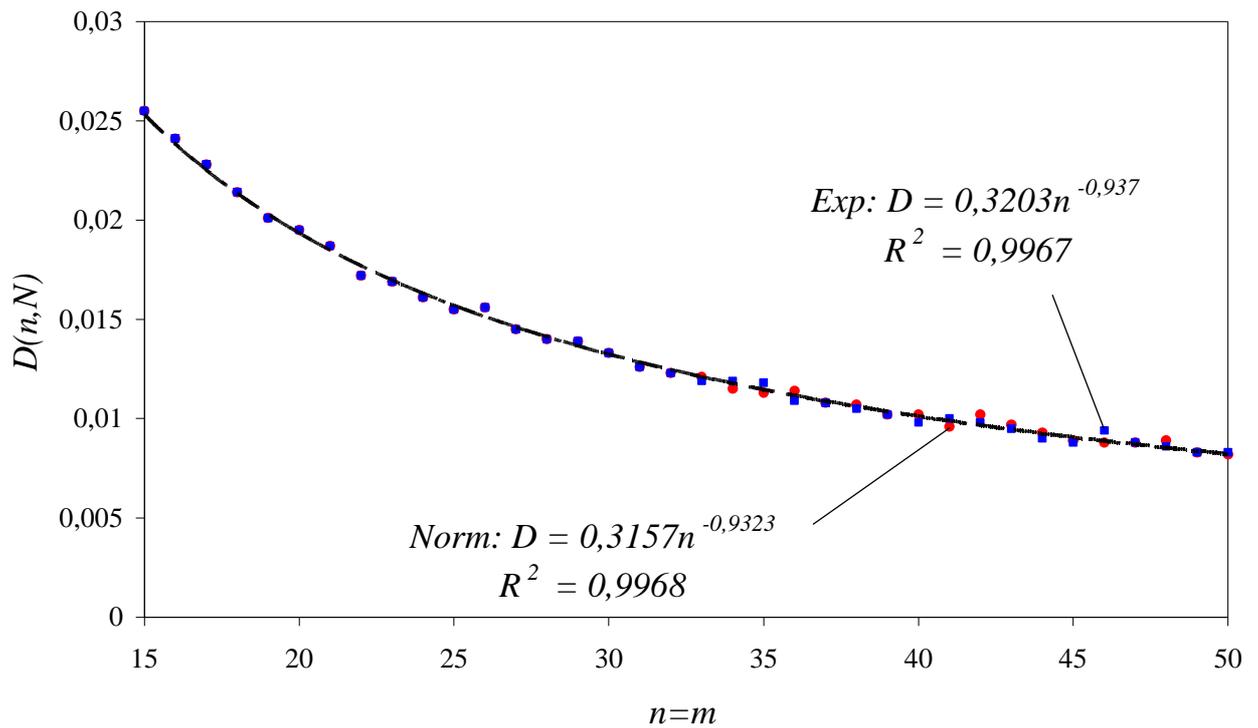


Рисунок 5.1 – Зависимость расстояния от распределения статистики A^2 до предельного закона $a_2(t)$

5.2 Сравнение мощности критериев

В данной работе мощность критериев проверки однородности исследовалась при ряде альтернатив. Для определенности, гипотезе H_0 соответствовала принадлежность выборок одному и тому же стандартному нормальному закону распределения (2.35) с параметрами сдвига $\theta_1 = 0$ и масштаба $\theta_0 = 1$. При всех рассмотренных альтернативах первая выборка всегда соответствовала стандартному нормальному закону, а вторая – некоторому другому. В частности, в случае конкурирующей гипотезы H_1 вторая выборка соответствовала нормальному закону с параметром сдвига $\theta_1 = 0.1$ и параметром масштаба $\theta_0 = 1$. В случае гипотезы H_2 – нормальному закону с параметрами $\theta_1 = 0.5$ и $\theta_0 = 1$. В случае гипотезы H_3 – нормальному закону с параметрами $\theta_1 = 0$ и $\theta_0 = 1.1$. В случае гипотезы H_4 – нормальному закону с параметрами $\theta_1 = 0$ и $\theta_0 = 1.5$. В случае гипотезы H_5 – вторая выборка

соответствовала логистическому закону с плотностью (2.36) и параметрами $\theta_1 = 0$ и $\theta_0 = 1$.

Рассматриваемые критерии исследовались на выборках объёма m и n , ($m=n$), равных 20, 50, 100, 300, 1000 и 2000.

5.2.1 Мощность критерия Андерсона-Дарлинга-Петита

На рисунке 5.2 представлены полученные в результате моделирования распределения статистики $G(A^2 | H_1)$ при справедливости H_1 и m, n равным 20.

На основе полученных результатов можно оценить значения мощности при различных объемах выборок m и n . В таблице 5.2 можно видеть полученные значения мощности критерия относительно конкурирующей гипотезы H_1 .

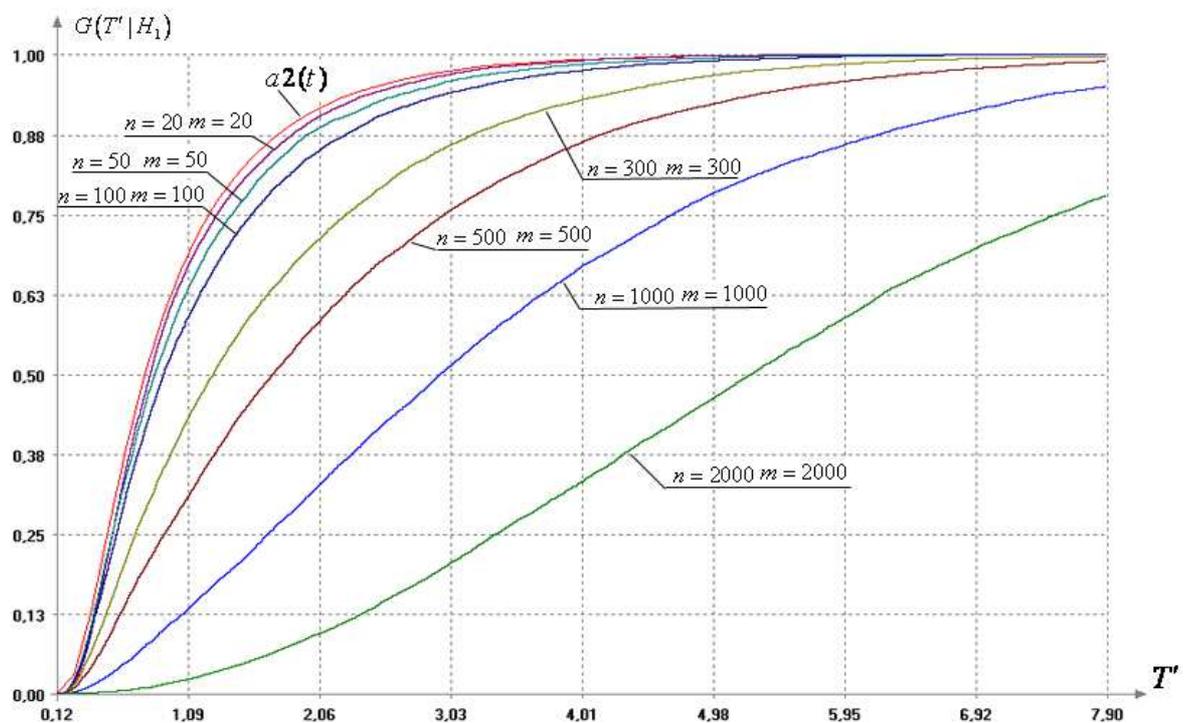


Рисунок 5.2 – Распределения статистики критерия Андерсона-Дарлинга-Петита при справедливости H_1

Таблица 5.2 – Мощность критерия однородности Андерсона-Дарлинга-Петита относительно альтернативы H_1

Уровень значимости α	Значения мощности относительно альтернативы H_1						
	$n = 20$	$n = 50$	$n = 100$	$n = 300$	$n = 500$	$n = 1000$	$n = 2000$
0.1	0.1169	0.1325	0.1707	0.3145	0.4425	0.6991	0.9174
0.05	0.0607	0.0729	0.0961	0.2205	0.3251	0.5909	0.8647
0.025	0.0349	0.0404	0.0586	0.1406	0.2431	0.4847	0.8022

На рисунке 5.3 при различных объемах выборок приведены распределения статистики $G(A^2 | H_2)$.

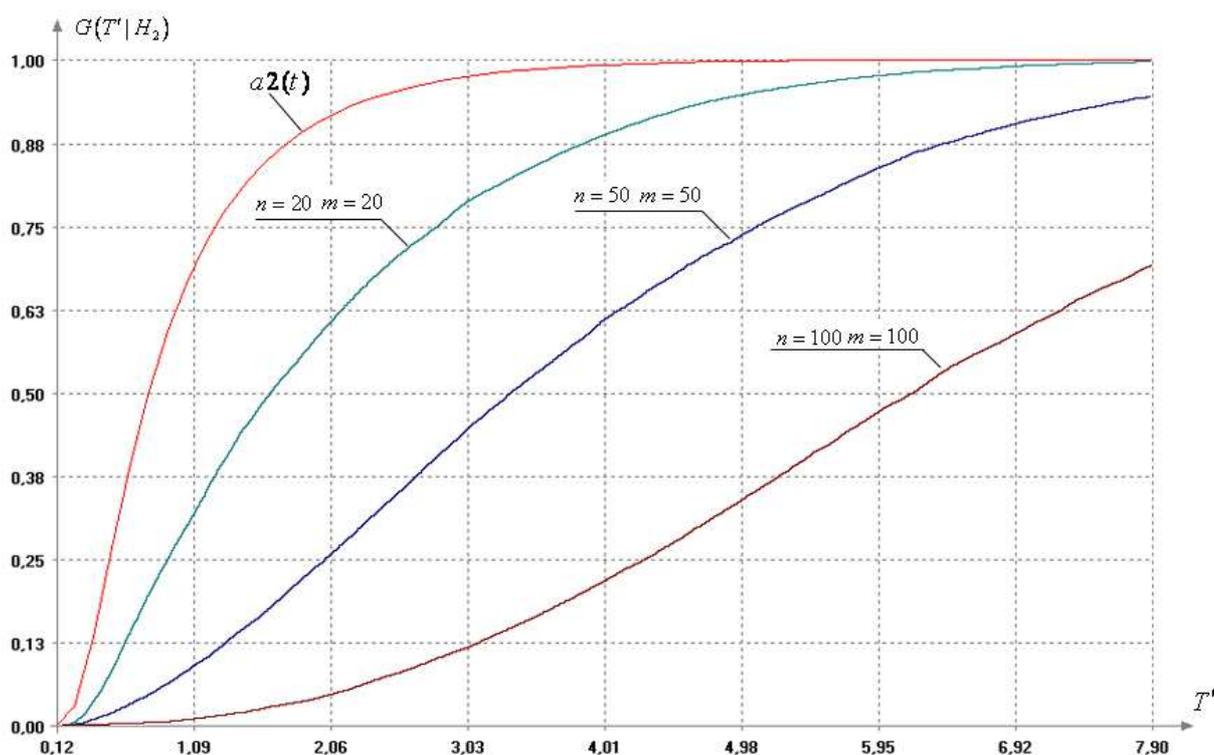


Рисунок 5.3 – Распределения статистики критерия Андерсона-Дарлинга-Петита при справедливости H_2

При объемах выборок m и n , равных 300, 500, 1000 и 2000, распределения $G(A^2 | H_2)$ существенно отклоняются от распределения $a_2(t)$ и на рисунке не приводятся.

Значения мощности критерия относительно конкурирующей гипотезы H_2 представлены в таблице 5.3.

Аналогично, на рисунках 5.4, 5.5, 5.6 приведены распределения статистики $G(A^2 | H_3)$, $G(A^2 | H_4)$, $G(A^2 | H_5)$ при справедливости соответствующих конкурирующих гипотез. Вычисленные значения мощности критерия Андерсона-Дарлинг-Петита представлены соответственно в таблицах 5.4, 5.5, 5.6.

Таблица 5.3 – Мощность критерия Андерсона-Дарлинг-Петита относительно конкурирующей гипотезы H_2

Уровень значимости α	Значения мощности относительно альтернативы H_2						
	$n = 20$	$n = 50$	$n = 100$	$n = 300$	$n = 500$	$n = 1000$	$n = 2000$
0.1	0.4283	0.7669	0.9604	0.9605	1	1	1
0.05	0.3101	0.6566	0.9235	0.9274	1	1	1
0.025	0.2248	0.5431	0.8791	0.8813	1	1	1

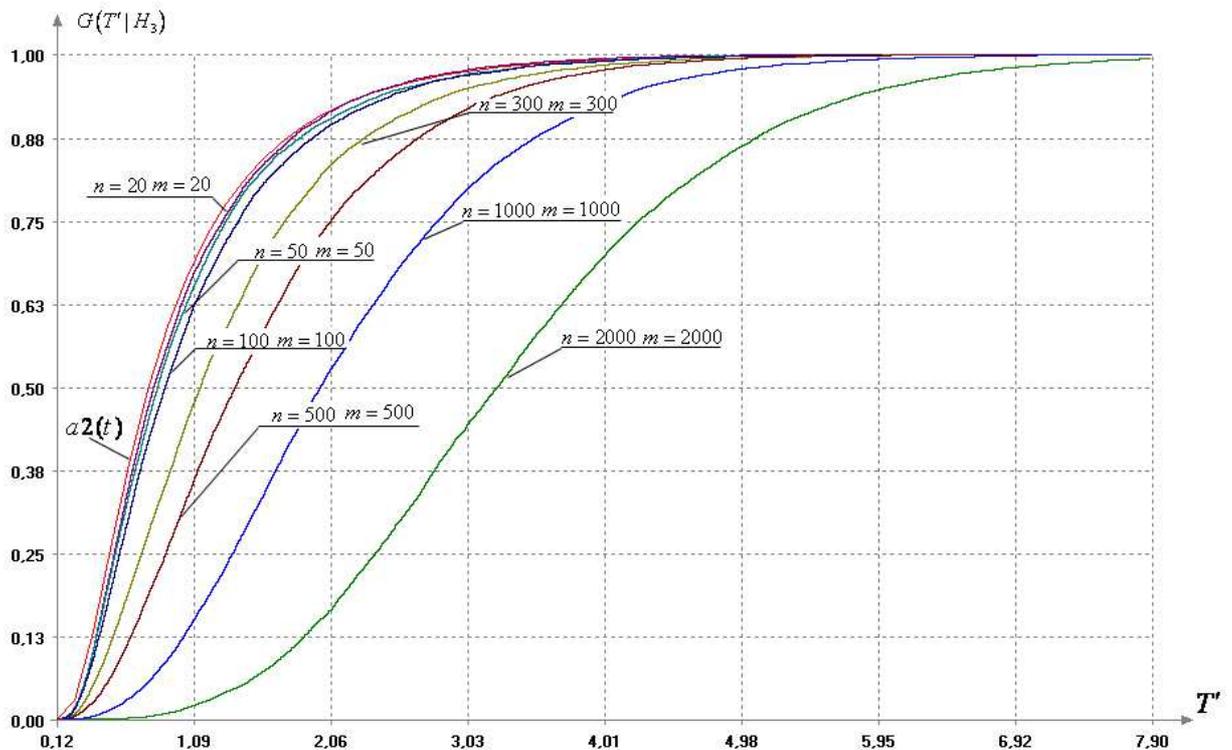


Рисунок 5.4 – Распределения статистики критерия Андерсона-Дарлинг-Петита при справедливости H_3

Таблица 5.4 – Мощность критерия Андерсона-Дарлинга-Петита относительно конкурирующей гипотезы H_3

Уровень значимости α	Значения мощности относительно альтернативы H_3						
	$n = 20$	$n = 50$	$n = 100$	$n = 300$	$n = 500$	$n = 1000$	$n = 2000$
0.1	0.1057	0.1112	0.1258	0.2003	0.2842	0.5161	0.8634
0.05	0.0512	0.0549	0.0605	0.1065	0.1553	0.3425	0.7252
0.025	0.0269	0.0282	0.0311	0.0508	0.0819	0.1993	0.5708

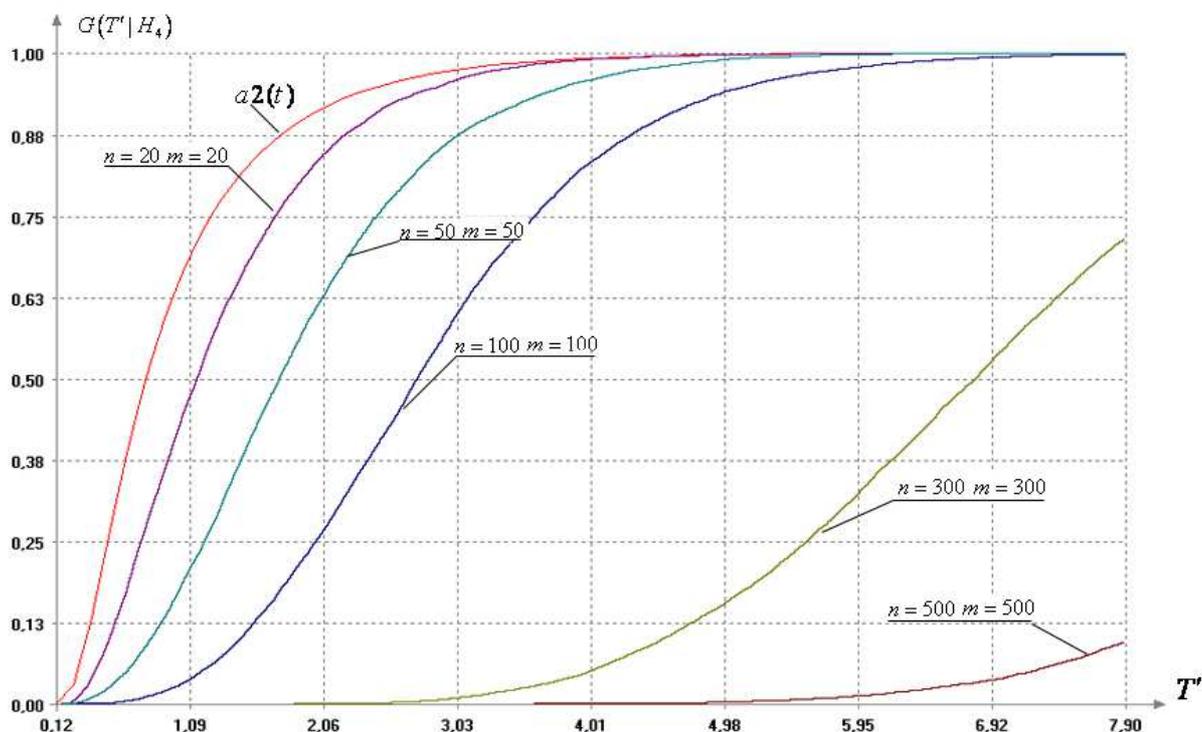


Рисунок 5.5 – Распределения статистики критерия Андерсона-Дарлинга-Петита при справедливости H_4

Таблица 5.5 – Мощность критерия однородности Андерсона-Дарлинга-Петита относительно конкурирующей гипотезы H_4

Уровень значимости α	Значения мощности относительно альтернативы H_4						
	$n = 20$	$n = 50$	$n = 100$	$n = 300$	$n = 500$	$n = 1000$	$n = 2000$
0.1	0.1903	0.4184	0.7714	0.9999	1	1	1
0.05	0.0902	0.2336	0.5832	0.9986	1	1	1
0.025	0.0464	0.1178	0.3868	0.9901	1	1	1

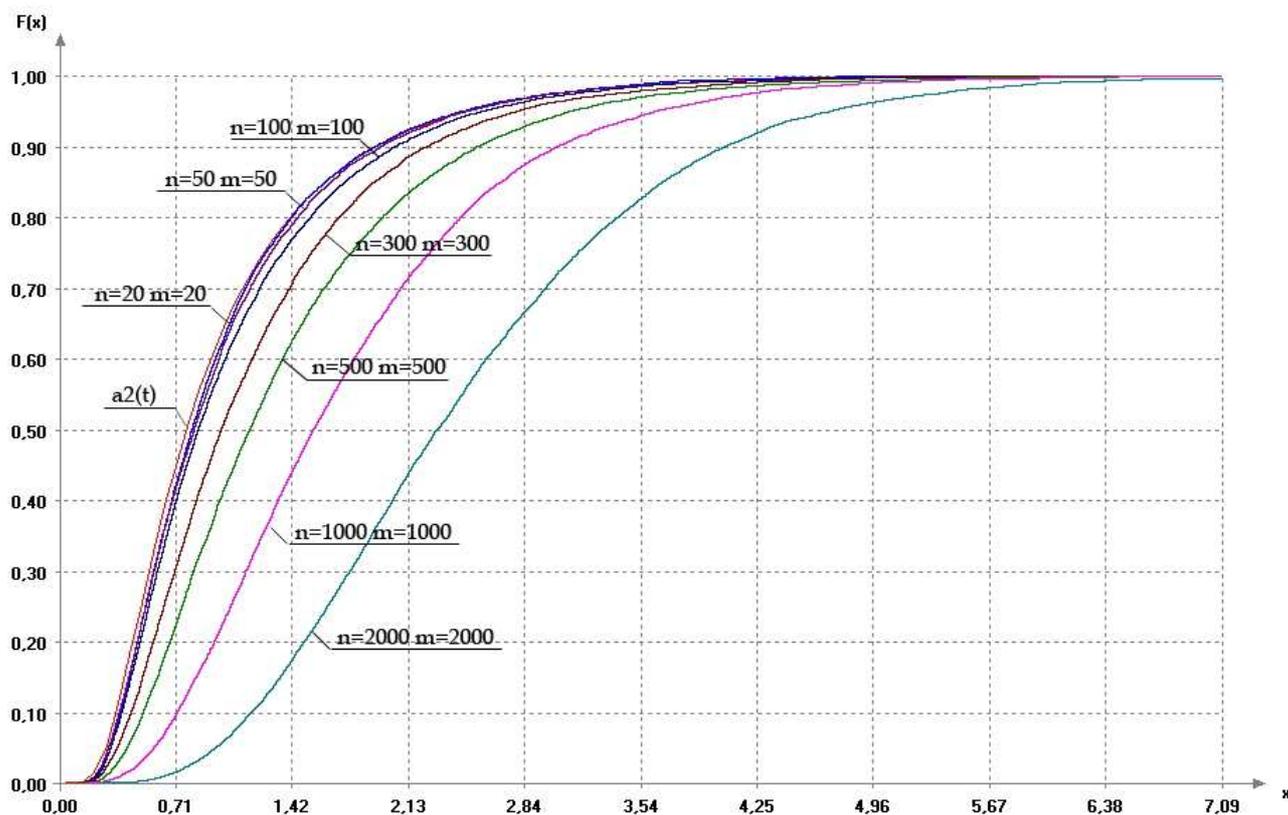


Рисунок 5.6 – Распределения статистики критерия Андерсона-Дарлинга-Петита при справедливости H_5

Таблица 5.6 – Мощность критерия однородности Андерсона-Дарлинга-Петита относительно конкурирующей гипотезы H_5

Уровень значимости α	Значения мощности относительно альтернативы H_5						
	$n = 20$	$n = 50$	$n = 100$	$n = 300$	$n = 500$	$n = 1000$	$n = 2000$
0.1	0.1008	0.1058	0.1199	0.1454	0.2066	0.3482	0.6398
0.05	0.0504	0.0483	0.0592	0.0718	0.1072	0.194	0.4436
0.025	0.0268	0.023	0.0287	0.035	0.0563	0.0992	0.2933

5.2.2 Мощность критерия однородности Смирнова

Критерий Смирнова был исследован в работе [152]. В ходе исследования было установлено, что условное распределение $G(S_C | H_0)$ статистики S_C при справедливости гипотезы H_0 , медленно сходится к $K(s)$ и существенно отличается от него при не очень больших m и n . Проблемы сходимости

распределения статистики критерия однородности Смирнова к предельному и мощность относительно анализируемых конкурирующих гипотез в зависимости от объёмов выборок рассмотрены в [91]. Результаты оценки мощности представлены в таблице 5.7.

Таблица 5.7 – Мощность критерия однородности Смирнова относительно конкурирующих гипотез $H_1 \div H_5$ в зависимости от объемов выборок ($m = n$)

Уровень значимости α	Значения мощности относительно альтернативы H_1						
	$n = 20$	$n = 50$	$n = 100$	$n = 300$	$n = 500$	$n = 1000$	$n = 2000$
0.1	0.0937	0.1480	0.1766	0.2775	0.3806	0.6171	0.8688
0.05	0.0410	0.0569	0.0944	0.1883	0.2682	0.4899	0.7762
0.025	0.0410	0.0344	0.0505	0.1163	0.1829	0.3859	0.6737
Значения мощности относительно альтернативы H_2							
0.1	0.3457	0.7200	0.9332	1	1	1	1
0.05	0.2202	0.5341	0.8722	0.9996	1	1	1
0.025	0.2202	0.4328	0.7842	0.9992	1	1	1
Значения мощности относительно альтернативы H_3							
0.1	0.0884	0.1229	0.1257	0.1466	0.1856	0.2967	0.5508
0.05	0.0352	0.0458	0.0630	0.0789	0.1024	0.1677	0.3520
0.025	0.0352	0.0257	0.0280	0.0410	0.0518	0.0967	0.2098
Значения мощности относительно альтернативы H_4							
0.1	0.1396	0.2986	0.5213	0.9609	0.9989	1	1
0.05	0.0570	0.1268	0.3161	0.8977	0.9952	1	1
0.025	0.0570	0.0763	0.1689	0.7738	0.9786	1	1
Значения мощности относительно альтернативы H_5							
0.1	0.0836	0.1069	0.1308	0.1568	0.1976	0.3191	0.5639
0.05	0.0341	0.0455	0.0673	0.0891	0.1158	0.1879	0.3754
0.025	0.0341	0.0258	0.0316	0.0471	0.0618	0.1119	0.2390

5.2.3 Мощность критерия Лемана-Розенблатта

В работе [91] было показано, что распределение статистики Лемана – Розенблатта уже при $m=100$ и $n=100$ при справедливости гипотезы H_0 практически совпадает с $a1(t)$. Там же была получена мощность критерия относительно альтернатив $H_1 \div H_5$. Результаты представлены в таблице 5.8.

Таблица 5.8 – Мощность критерия однородности Лемана – Розенблатта относительно конкурирующих гипотез $H_1 \div H_5$ в зависимости от объемов выборок ($m = n$)

Уровень значимости α	Значения мощности относительно альтернативы H_1						
	$n = 20$	$n = 50$	$n = 100$	$n = 300$	$n = 500$	$n = 1000$	$n = 2000$
0.1	0.1241	0.1382	0.1727	0.3125	0.4369	0.6874	0.9114
0.05	0.0615	0.0770	0.0999	0.2078	0.3211	0.5703	0.8469
0.025	0.0324	0.0410	0.0590	0.1333	0.2288	0.4589	0.7681
Значения мощности относительно альтернативы H_2							
0.1	0.4321	0.7628	0.9549	1	1	1	1
0.05	0.3121	0.6473	0.9154	1	1	1	1
0.025	0.2120	0.5355	0.8661	0.9998	1	1	1
Значения мощности относительно альтернативы H_3							
0.1	0.1096	0.1107	0.1147	0.1459	0.1898	0.3265	0.6237
0.05	0.0508	0.0567	0.0563	0.0691	0.0945	0.1675	0.3986
0.025	0.0252	0.0291	0.0283	0.0334	0.0442	0.0805	0.2259
Значения мощности относительно альтернативы H_4							
0.1	0.1655	0.2875	0.5513	0.9875	0.9999	1	1
0.05	0.0801	0.1437	0.3199	0.9470	0.9993	1	1
0.025	0.0361	0.0727	0.1687	0.8587	0.9952	1	1
Значения мощности относительно альтернативы H_5							
0.1	0.1087	0.1069	0.1135	0.1422	0.1826	0.2978	0.5463
0.05	0.0511	0.0549	0.0581	0.0668	0.0910	0.1450	0.3390
0.025	0.0241	0.0276	0.0290	0.0332	0.0431	0.0712	0.1822

Сравнивая мощность критериев относительно рассмотренных конкурирующих гипотез, можно заметить, что мощности критериев Смирнова, Лемана-Розенблатта и Андерсона-Дарлинга-Петита при близких альтернативах и при небольших объёмах выборок невысока.

Относительно конкурирующей гипотезы H_1 при объёмах выборок более 300 мощность критерия Андерсона-Дарлинга-Петита оказывается чуть выше мощности критериев Смирнова и Лемана-Розенблатта.

Относительно конкурирующей гипотезы H_2 мощность критерия Смирнова достигает 1 при объёме выборок равных 300 при уровне значимости 0.1, в то время как мощность критерия Андерсона-Дарлинга-Петита достигает 1 при объёме выборок равных 500, а при объёме выборок равном 300 составляет около 0.92.

При сдвиге масштаба на 0.1σ (конкурирующая гипотеза H_3), мощность критерия Андерсона-Дарлинга-Петита, начиная с объёма выборок равного 100, выше мощности критериев Смирнова и Лемана-Розенблатта. И при объёме равном 2000 при уровне значимости 0,1 составляет 0,86, в то время как, мощность критерия Лемана-Розенблатта (при тех же условиях) составляет 0,62, а мощность критерия Смирнова – 0,55. При более существенных отклонениях масштаба мощность критерия Андерсона-Дарлинга-Петита оказывается выше мощности критериев Смирнова и Лемана-Розенблатта при объёме выборок равном 500.

Если посмотреть на мощность критериев относительно очень близкой конкурирующей гипотезы H_5 , то можно заметить, что мощность критерия Смирнова и Лемана-Розенблатта при объёме равном 2000 и уровне значимости 0,05 составляет 0,34 и 0,37 соответственно. Мощность же критерия Андерсона-Дарлинга-Петита при этих условиях равна 0,44.

5.3 Сравнение мощности критериев однородности по данным типа времени жизни

В работах Постовалова и Филоненко [143,242] проведено исследование различных критериев однородности по данным типа времени жизни. Были рассмотрены критерии проверки гипотезы однородности Смирнова, Лемана-Розенבלата, Андерсона-Дарлинга-Петита, Вилкоксона, логарифмический ранговый критерий, Q-критерий [226], критерий Багдонавичуса-Никулина [162] при близких конкурирующих гипотезах.

Конкурирующая гипотеза H_1 предполагает, что первая выборка подчинена распределению Вейбулла-Гнеденко с параметрами формы и масштаба равными двум, а вторая выборка – гамма-распределению с параметром формы 3,12154 и параметром масштаба 0,557706. Гипотеза H_2 предполагает, что первая выборка подчинена распределению Вейбулла-Гнеденко с параметрами формы и масштаба равными двум, а вторая выборка – логнормальному закону с параметрами $\mu = 0,4096; \sigma = 0,6179$. Для гипотезы H_3 обе выборки подчиняются экспоненциальному закону, но с разным параметром масштаба: в первой выборке – 1, а во второй – 1,1.

Результаты вычисления мощности приведены в таблице 5.9. Полужирным шрифтом выделены ячейки, в которых мощность является максимальной по столбцу. По полученным результатам можно сделать вывод, что для альтернативы с пересечениями функций распределения (против конкурирующих гипотез H_1 и H_2) критерий Багдонавичуса-Никулина имеет наибольшую мощность, а для альтернативы без пересечений (против конкурирующей гипотезы H_3) – логранговый критерий.

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

Таблица 5.9 – Мощность критериев однородности относительно конкурирующих гипотез $H_1 \div H_3$ в зависимости от объемов выборок ($m = n$)

Критерий	$n = 20$	$n = 50$	$n = 200$	$n = 500$
Значения мощности относительно альтернативы H_1				
Вилкоксона	0,055	0,066	0,109	0,207
Логранговый	0,054	0,051	0,052	0,055
Q-критерий	0,054	0,053	0,058	0,072
Андерсона-Дарлинга-Петита	0,056	0,069	0,125	0,256
Лемана-Розенблатта	0,058	0,070	0,140	0,275
Смирнова	0,059	0,071	0,142	0,288
Багдонавичуса-Никулина	0,056	0,075	0,194	0,452
Значения мощности относительно альтернативы H_2				
Вилкоксона	0,060	0,068	0,121	0,226
Логранговый	0,051	0,048	0,065	0,110
Q-критерий	0,049	0,044	0,053	0,091
Андерсона-Дарлинга-Петита	0,064	0,081	0,270	0,781
Лемана-Розенблатта	0,063	0,091	0,234	0,554
Смирнова	0,071	0,096	0,250	0,543
Багдонавичуса-Никулина	0,078	0,170	0,729	0,992
Значения мощности относительно альтернативы H_3				
Вилкоксона	0,056	0,070	0,129	0,257
Логранговый	0,059	0,076	0,158	0,327
Q-критерий	0,061	0,073	0,147	0,300
Андерсона-Дарлинга-Петита	0,058	0,067	0,132	0,264
Лемана-Розенблатта	0,060	0,067	0,127	0,251
Смирнова	0,059	0,067	0,115	0,216
Багдонавичуса-Никулина	0,055	0,066	0,121	0,248

5.4 Выводы

В результате проведенных исследований можно сделать следующие выводы.

1. Распределение статистики критерия Андерсона-Дарлинга-Петита достаточно быстро, со скоростью примерно $O(n^{-0,93})$, сходится к предельному распределению $a_2(t)$. Для практических целей, когда погрешность в определении достигаемого уровня значимости критерия не превосходит 0,01 достаточно иметь выборки объемом 45 наблюдений.

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

3. 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.

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

Таким образом, проведенные исследования подтвердили предположение о высокой мощности критерия однородности Андерсона-Дарлинга-Петита в большинстве случаев, что позволяет рекомендовать его широкое применение на практике.

ГЛАВА 6. ИНВАРИАНТНЫЕ КРИТЕРИИ ПРОВЕРКИ ГИПОТЕЗЫ О МНОГОМЕРНОЙ НОРМАЛЬНОСТИ

Предположение о нормальности многомерной случайной величины часто используется при нахождении распределений статистик критериев в многомерном статистическом анализе. Например, многомерная нормальность предполагается при проведении дисперсионного анализа (MANOVA), дискриминантного анализа и многомерной регрессии. Исследования, проведенные Хопкинсом (Hopkins) и Клаем (Clay) [187], Мардия (Mardia) [225], Коновером (Conover) и Иманом (Iman) [170] показали важность предположений о многомерной нормальности, неробастность многих процедур и влияние на них нарушения предположения о нормальности данных [177].

Для проверки многомерной нормальности разработаны специальные статистические критерии. В данной главе будут рассмотрены критерии, которые являются подмножеством класса инвариантных критериев проверки многомерной нормальности, приведенные в обзоре Хенце (Henze) в 2002 году [185].

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. Так еще в 1970 году Мардия предложил критерии проверки многомерной выборки на нормальность. Вычисление статистик этих критериев основывалось на выборочных коэффициентах асимметрии и эксцесса [223]. Позже Козиол (Koziol) занимался исследованием обоих критериев [199], [200]. Мори (Mori), Рохатги (Rohatgi) и Шекли (Szekely) предложили свою модификацию критерия, основанного на вычислении коэффициента асимметрии [231]. В 1984 Шривастава (Srivastava)

предложил использовать разложение ковариационной матрицы на собственные вектора для проверки выборки на коэффициенты асимметрии и эксцесса [261]. С начала 90-х годов Сео (Seo) и Арига (Ariga) [182], [260], [261] занимаются исследованием преобразований группы критериев, основанных на вычислении эксцесса Мардия и Шриваставы.

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 Хи-квадрат распределению с d степенями свободы.

Критерии нормальности, основанные на полярной декомпозиции, исследовались в работах [197, 198, 206, 247, 240, 250, 230, 232].

6.1 Гипотеза о многомерной нормальности

Пусть X_1, X_2, \dots, X_n последовательность независимых случайных d -мерных векторов X и $N_d(\mu, \Sigma)$ – d -мерное нормальное распределение с вектором математического ожидания $M = [M_1, \dots, M_d]$ и ковариационной матрицей Σ

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1d} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2d} \\ \dots & \dots & \dots & \dots \\ \sigma_{d1} & \sigma_{d2} & \dots & \sigma_{dd} \end{bmatrix}. \quad (6.1)$$

Функция плотности многомерного нормального закона имеет вид

$$f(X) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(X-M)^T \Sigma^{-1} (X-M)} \quad (6.2).$$

Обозначим N_d – как класс всех невырожденных d -мерных нормальных распределений. Задача проверки многомерной нормальности состоит в проверке гипотезы:

$$H_0 : P^X \in N_d, \quad (6.3)$$

где P^X – распределение вектора X .

6.2 Моделирование многомерной случайной величины

Для исследования свойств критериев методом Монте-Карло необходимо моделировать многомерные псевдослучайные величины. В данной работе использовался алгоритм 6.1, описанный в [36].

Алгоритм 6.1. Моделирование многомерной случайной нормальной выборки [36].

Входные данные. M – вектор математического ожидания и Σ – ковариационная матрица размерности d , n – объем выборки.

1. Моделируется $\{Z_i\}$, $i = 1, \dots, d$, – совокупность одномерных случайных величин, подчиненных стандартному нормальному закону.
2. Вектор X распределен по многомерному нормальному закону с математическим ожиданием M и ковариационной матрицей $\Sigma = \|\sigma_{ij}\|_d^1$:

$$X = AZ + M, \quad (6.4)$$

где коэффициенты a_{ij} нижней треугольной матрицы A вычисляются по формуле:

$$a_{ij} = \frac{\sigma_{ij} - \sum_{k=1}^{j-1} a_{ik} a_{jk}}{\sqrt{\sigma_{jj} - \sum_{k=1}^{j-1} a_{jk}^2}}, \quad 1 \leq j \leq i \leq d. \quad (6.5)$$

Описанный алгоритм можно применять и для моделирования многомерной случайной величины, распределение которой отлично от нормального. При этом совокупность $\{Z_i\}$ моделируется не по нормальному закону, а в соответствии с некоторым одномерным распределением (отличным от нормального) [86].

Математическое ожидание полученного многомерного распределения будет равно вектору M , а ковариационная матрица будет равна заданной матрице Σ [86].

6.3 Инвариантные критерии проверки многомерной нормальности

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 $T_n = T_n(X_1, \dots, X_n)$ инвариантного критерия проверки гипотезы H_0 должна удовлетворять условию инвариантности:

$$T_n(AX_1 + b, \dots, AX_n + b) = T_n(X_1, \dots, X_n), \quad (6.6)$$

для произвольного вектора $b \in R^d$ и невырожденной матрицы $A \in R^{d \times d}$.

Все критерии проверки многомерной нормальности, рассматриваемые в данной главе, инвариантны относительно аффинных преобразований. Это означает, что результат проверки гипотезы H_0 не зависит от параметров распределения (вектора математических ожиданий и ковариационной матрицы) исходной многомерной выборки (принадлежащей нормальному или некоторому другому закону).

Пусть X_1, X_2, \dots, X_n последовательность независимых случайных d -мерных векторов X , где $d \geq 1$. Распределение вектора X обозначим как P^X .

6.3.1 Критерии, основанные на вычислении коэффициента асимметрии

В критериях, основанных на выборочном коэффициенте асимметрии, вместо гипотезы (6.2) проверяется гипотеза о равенстве коэффициента асимметрии нулю. Критерий асимметрии позволяет отклонить гипотезу о нормальности, но недостаточен для принятия гипотезы о согласии с нормальным распределением.

В одномерном случае выборочный коэффициент асимметрии вычисляется по формуле

$$\sqrt{b_{n,1}^{(1)}} = \frac{1}{n} \sum_{j=1}^n \left(\frac{X_j - \bar{X}_n}{\sigma_n} \right)^3, \quad (6.7)$$

где $\sigma_n^2 = \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X}_n)^2$ – выборочная дисперсия. На вычисление этого коэффициента опирается часть критериев проверки многомерной нормальности.

6.3.1.1 Критерий проверки на асимметрию Мардиа

Обобщение формулы (6.7) для многомерного случая выражается в вычислении невырожденной инвариантной асимметрии по формуле

$$b_M^{(1)} = \frac{1}{n^2} \sum_{j,k=1}^n D_{n,jk}^3, \quad (6.8)$$

где

$$\begin{aligned}
D_{n,jk} &= Y_{n,j}^T Y_{n,k}, \\
Y_{n,j} &= S_n^{-1/2} (X_j - \bar{X}_n), \\
S_n &= \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X}_n)(X_j - \bar{X}_n)^T, \\
\bar{X}_n &= \frac{1}{n} \sum_{j=1}^n X_j.
\end{aligned} \tag{6.9}$$

Эта статистика была предложена Мардия в 1970-х годах [224, 225, 223]. Заметим, что статистика $b_M^{(1)}$ – это квадрат выборочной асимметрии для одномерных данных. В работе [225] Мардия советует использовать статистику $b_M^{(1)}$ (как верхнюю грань отклонения) для проверки многомерной нормальности, говоря, что асимметрия и эксцесс являются естественной мерой измерения отклонения от (многомерной) нормальности.

Очевидно, что (6.9) – это состоятельная оценка инвариантного функционала $\beta_d^{(1)} = E(X^T X)^3$. Последнее предположение справедливо при ограничении $E \|X\|^6 < \infty$, где $\|\cdot\|$ означает Евклидову норму в пространстве R^d . Если X является сферически-симметричным (т.е. распределение X инвариантно относительно ортогональных преобразований) и $E \|X\|^6 < \infty$, то

$$nb_M^{(1)} \xrightarrow{D} \alpha_1 \chi_d^2 + \alpha_2 \chi_{d(d-1)(d+4)/6}^2$$

где

$$\alpha_1 = \frac{3}{d} \left[\frac{E \|X\|^6}{d+2} - 2E \|X\|^4 + d(d+2) \right], \quad \alpha_2 = \frac{6E \|X\|^6}{d(d+2)(d+4)} \tag{6.10}$$

и χ_d^2 , $\chi_{d(d-1)(d+4)/6}^2$ – распределение χ^2 с d и $d(d-1)(d+4)/6$ степенями свободы соответственно [165].

Если $X \sim N_d(0, I_d)$, то коэффициенты из формул (6.10) $\alpha_1 = \alpha_2 = 6$ и $nb_M^{(1)} \xrightarrow{D} 6\chi_{d(d+1)(d+2)/6}^2$, при верной гипотезе H_0 [225].

Рассмотрим статистику (6.8) критерия асимметрии Мардя. Т.к. распределение статистики зависит от размерности выборки и от ее объема, то будем анализировать сходимость к предельному распределению не самой статистики $b_M^{(1)}$, а величины $\frac{n}{6}b_M^{(1)}$:

$$\frac{n}{6}b_M^{(1)} \rightarrow \chi_{d(d+1)(d+2)/6}^2 \quad (6.11)$$

На рисунках 6.1, 6.2 и 6.3 представлены функции распределения статистики критерия Мардя, основанного на коэффициенте асимметрии, умноженного на величину $\frac{n}{6}$ для размерностей $d=1$, $d=3$, $d=6$ соответственно.

По графикам видно, что статистика $\frac{n}{6}b_M^{(1)}$ сходится к предельному Хи-квадрат распределению. При объеме выборки $n > 200$ в одномерном случае и при $n > 400$ при $d=3$ и $d=6$ распределение статистики почти не отличается от предельного закона. Для того, чтобы определить более точные объемы выборки, при которых статистика критерия отклоняется от предельного закона не более чем на 0,01, вычислим расстояние Колмогорова D_n при разных объемах выборки. Полученные значения аппроксимируем уравнением степенной регрессии.

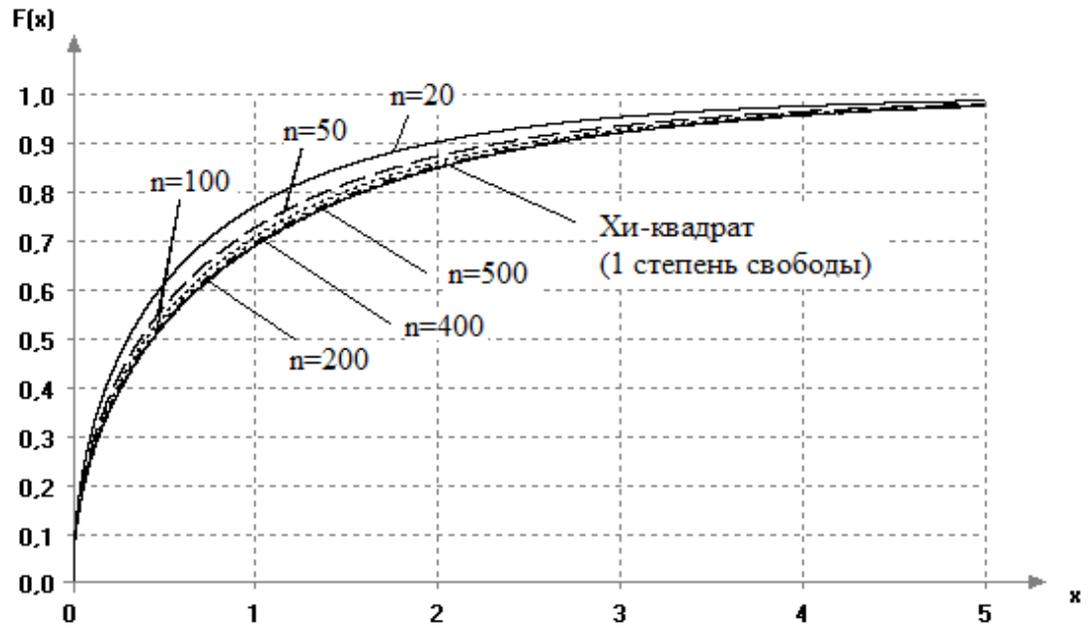


Рисунок 6.1 – Распределение статистики $\frac{n}{6}b_M^{(1)}$ при $d=1$

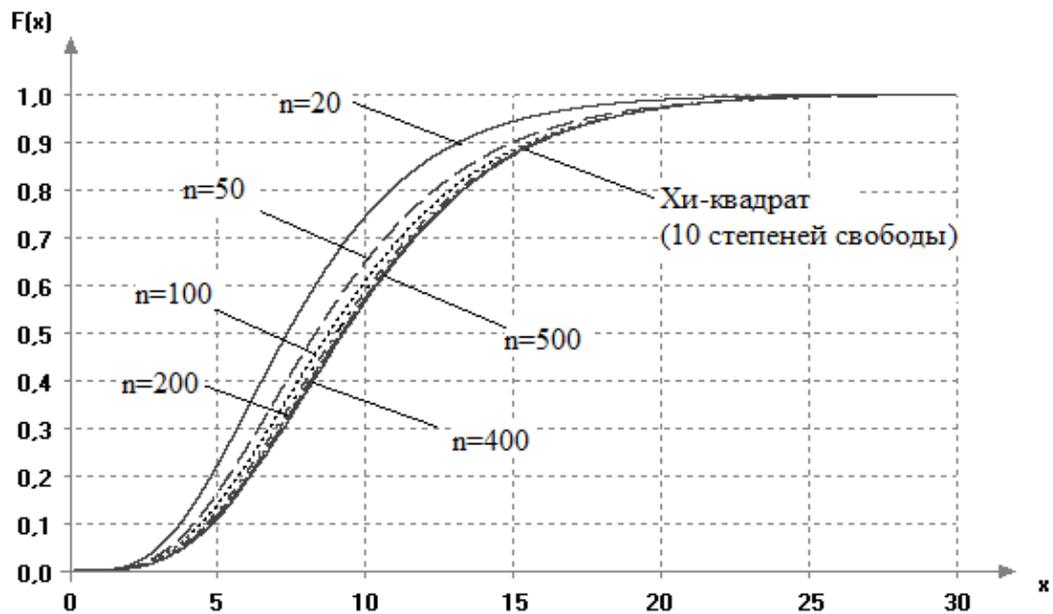


Рисунок 6.2 – Распределение статистики $\frac{n}{6}b_M^{(1)}$ при $d=3$

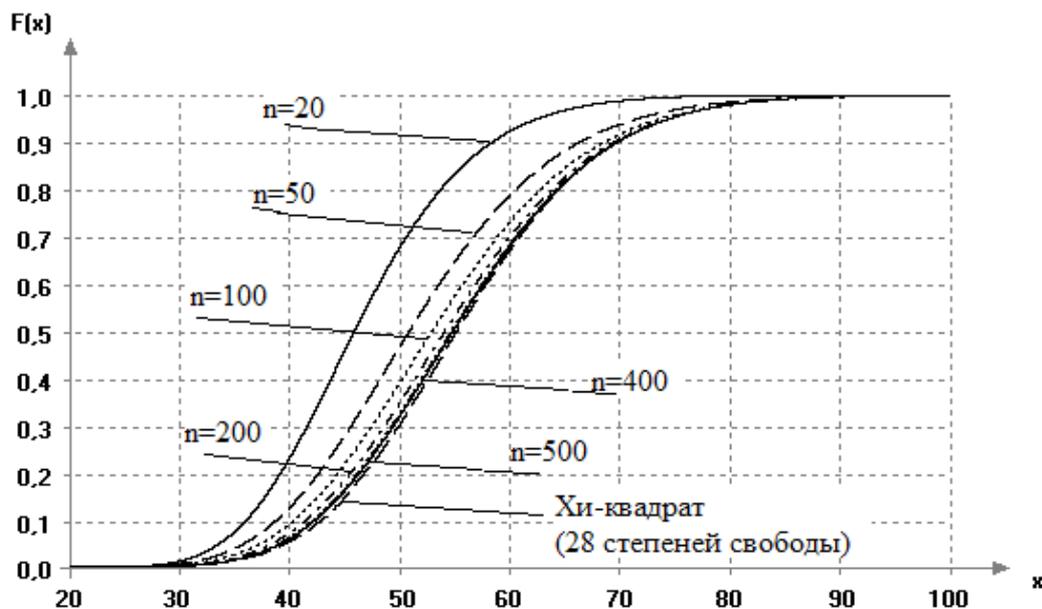


Рисунок 6.3 – Распределение статистики $\frac{n}{6}b_M^{(1)}$ при $d=6$

На рисунке 6.4 представлена зависимость расстояния D_n от объема выборки n для критерия асимметрии Мардя при различных размерностях выборки. На этом рисунке и далее, пунктирной линией обозначен уровень, при котором расстояние D_n равно 0,01.

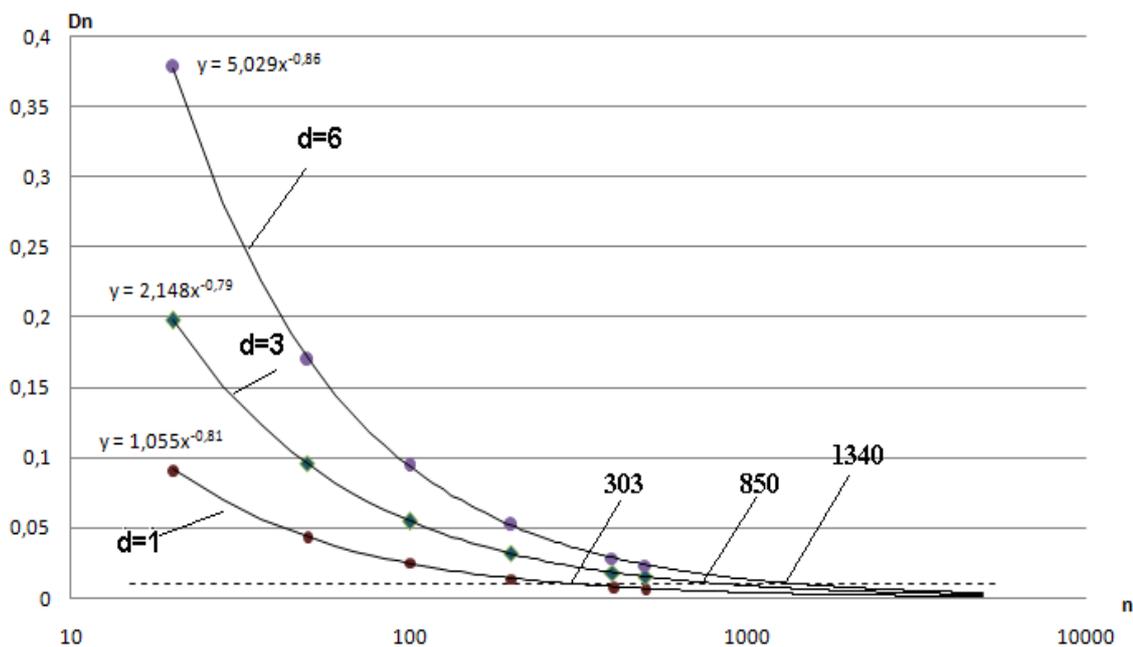


Рисунок 6.4 – Расстояние до предельного закона распределений статистик критерия Мардя

Видно, что с ростом размерности многомерной выборки скорость сходимости к предельному закону уменьшается. "Пороговым" значением объема выборки n , при котором расстояние до предельного закона меньше или равно 0.01, для одномерной выборки является 303 наблюдения, для выборки размерности $d=3$ уже 850 наблюдений, и 1340 наблюдений при $d=6$. Очевидно, что использовать Хи-квадрат распределение, как предельный закон для рассматриваемого критерия можно только при больших объемах выборки.

6.3.1.2 Критерий проверки на асимметрию Мори

Мори, Рохатги и Шекли [231] изучали статистику

$$\tilde{\beta}_d^{(1)} = \| E(\| X \|^2 X) \|^2 \quad (6.12)$$

в качестве выборочной оценки многомерной асимметрии.

Так же как и $\beta_d^{(1)}$, вычисление $\tilde{\beta}_d^{(1)}$ сводится к квадрату выборочной асимметрии в одномерном случае. Как и в случае критерия со статистикой $b_M^{(1)}$, естественно отклонять гипотезу многомерной нормальности при больших значениях инвариантной выборочной статистики

$$\tilde{b}_{n,d}^{(1)} = \frac{1}{n^2} \sum_{j,k=1}^n D_{n,jk} D_{n,jj} D_{n,kk} \quad , \quad (6.13)$$

где $D_{n,jk}$ вычисляются по (6.9).

В работе [184] было получено предельное распределение статистики $\tilde{b}_{n,d}^{(1)}$ и показаны недостатки критерия, основанного на вычислении статистики $\tilde{b}_{n,d}^{(1)}$.

Если $E \| X \|^6 < \infty$ и распределение X сферически-симметричное, то

$$n\tilde{b}_{n,d}^{(1)} \xrightarrow{D} \lambda \chi_d^2,$$

где $\lambda = (d + 2)\alpha_1 / 3$, а коэффициент α_1 вычисляется по формуле (6.10). С другой стороны, $\sqrt{n}(\tilde{b}_{n,d}^{(1)} - \tilde{\beta}_d^{(1)})$ подчиняется в пределе нормальному распределению, если $\tilde{\beta}_d^{(1)} > 0$ [184, 195].

Критерий проверки многомерной выборки на асимметрию Мори является модификацией статистики критерия Мардия. Статистика критерия Мори вычисленная по формуле (6.13) зависит от объема и размерности выборки. Предельным распределением статистики является распределение Хи-квадрат с числом степеней свободы, равным размерности выборки:

$$\frac{n}{2(d+2)} b_{Mori}^{(1)} \rightarrow \chi_d^2. \quad (6.14)$$

Рисунок 6.5 иллюстрирует скорость сходимости распределения статистики критерия Мори к предельному Хи-квадрат распределению. По графику видно, что распределение статистики критерия Мори при $d > 1$ сходится к предельному распределению быстрее, чем распределение статистики критерия асимметрии Мардия. Для $d = 3$ расстояние до предельного Хи-квадрат распределения равно 0,01 уже при объеме выборки около 560 наблюдений. Это в полтора раза меньше, чем в случае критерия асимметрии Мардия.

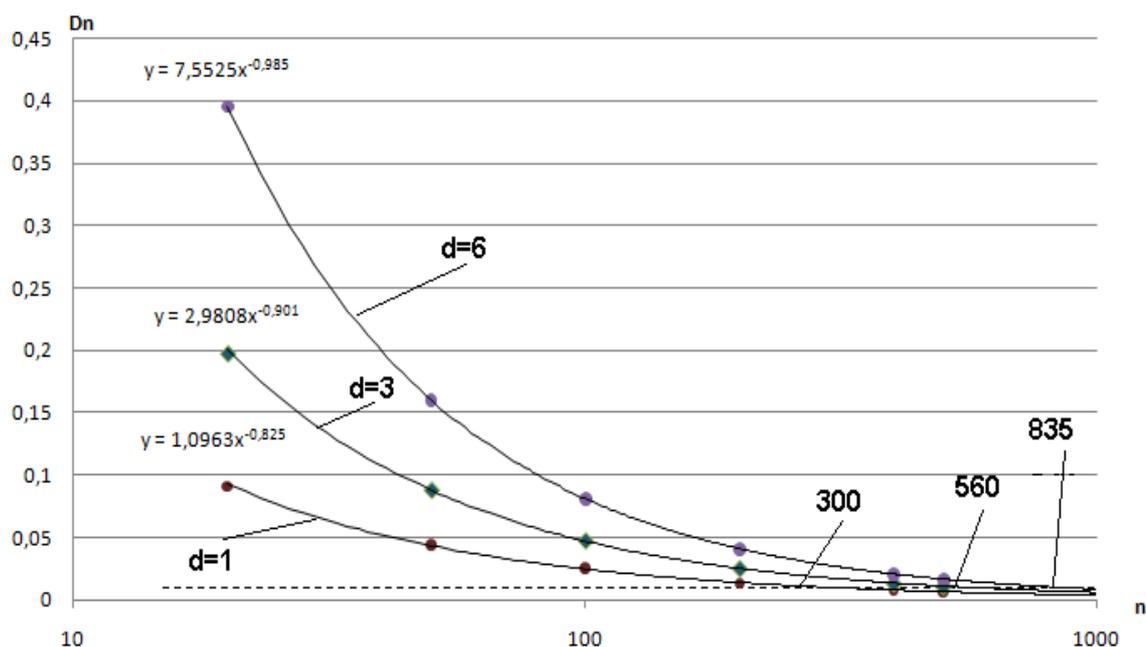


Рисунок 6.5 – Расстояние до предельного закона распределений статистик критерия Мори

Аналогичный результат получился и при других размерностях $d > 1$ многомерной выборки. Для одномерной выборки статистики критериев асимметрии Мардя и Мори совпадают, а значит и объемы выборок, при которых расстояние D_n одинаково мало, так же должны быть одинаковыми. На рисунках 6.4 и 6.5 эти значения различаются на уровне статистической погрешности моделирования.

При малых объемах выборок распределение статистики $\frac{n}{2(d+2)} b_{Mori}^{(1)}$ значительно удалено от предельного Хи-квадрат распределения. Но с ростом числа наблюдений расстояние до предельного закона быстро сокращается, что хорошо видно на графиках функций распределения статистик критерия Мори, приведенных на рисунках 6.6, 6.7, 6.8.

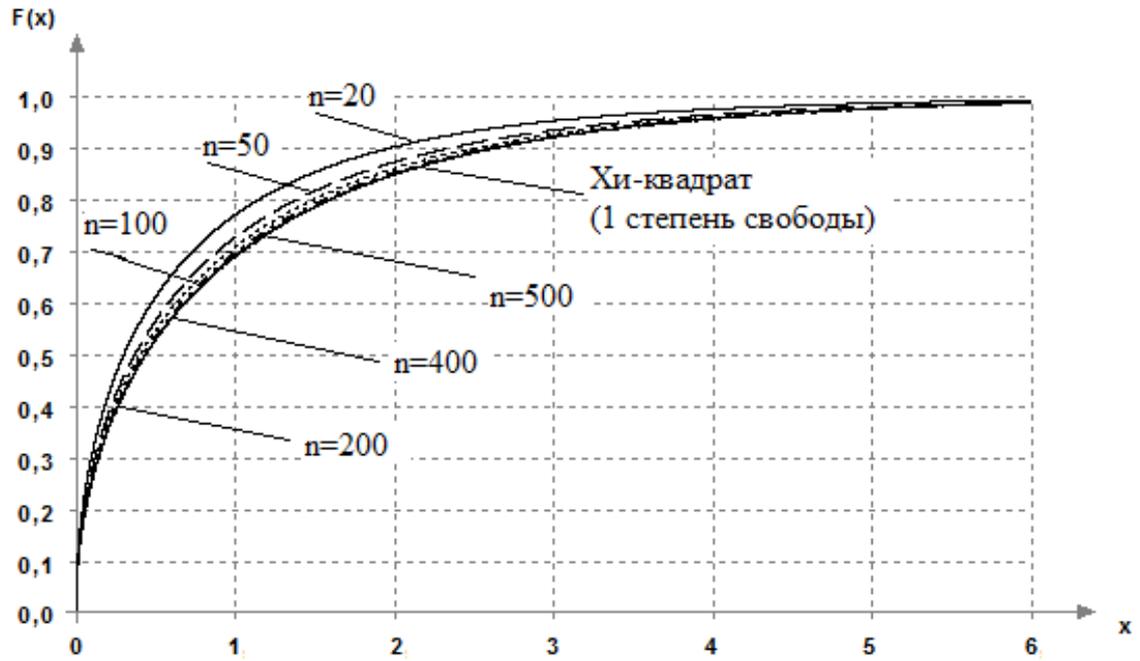


Рисунок 6.6 – Распределение статистики $\frac{n}{2(d+2)} b_{Mori}^{(1)}$ при $d=1$

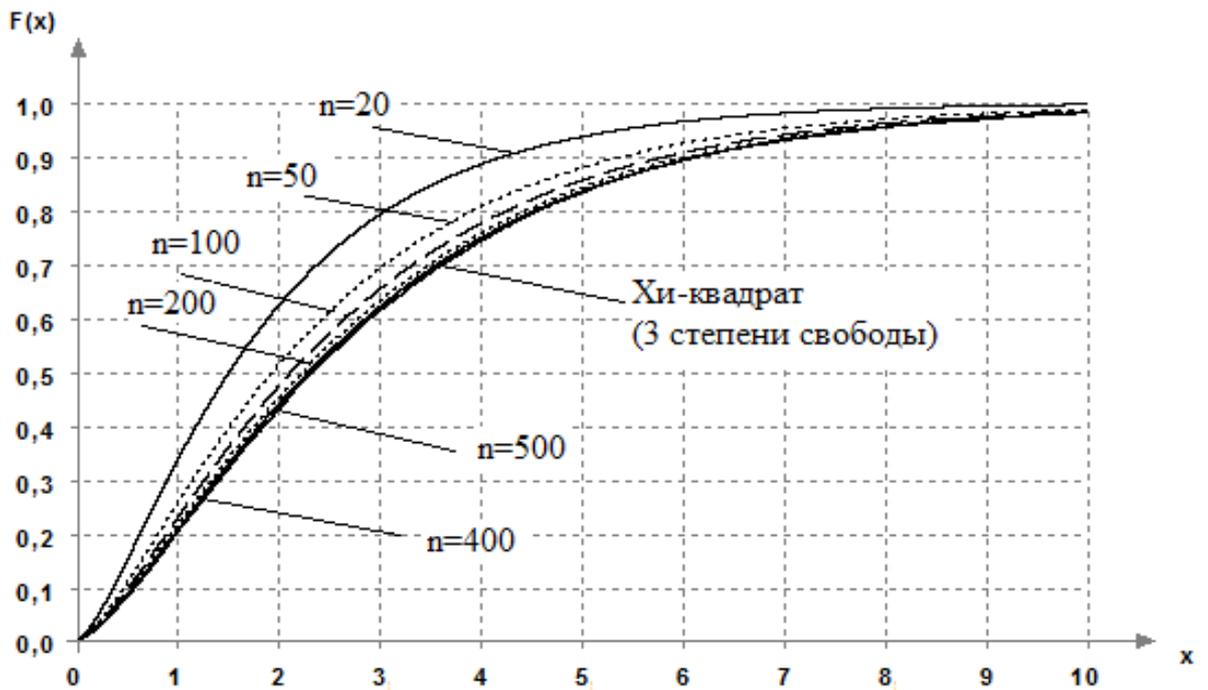


Рисунок 6.7 – Распределение статистики $\frac{n}{2(d+2)} b_{Mori}^{(1)}$ при $d=3$

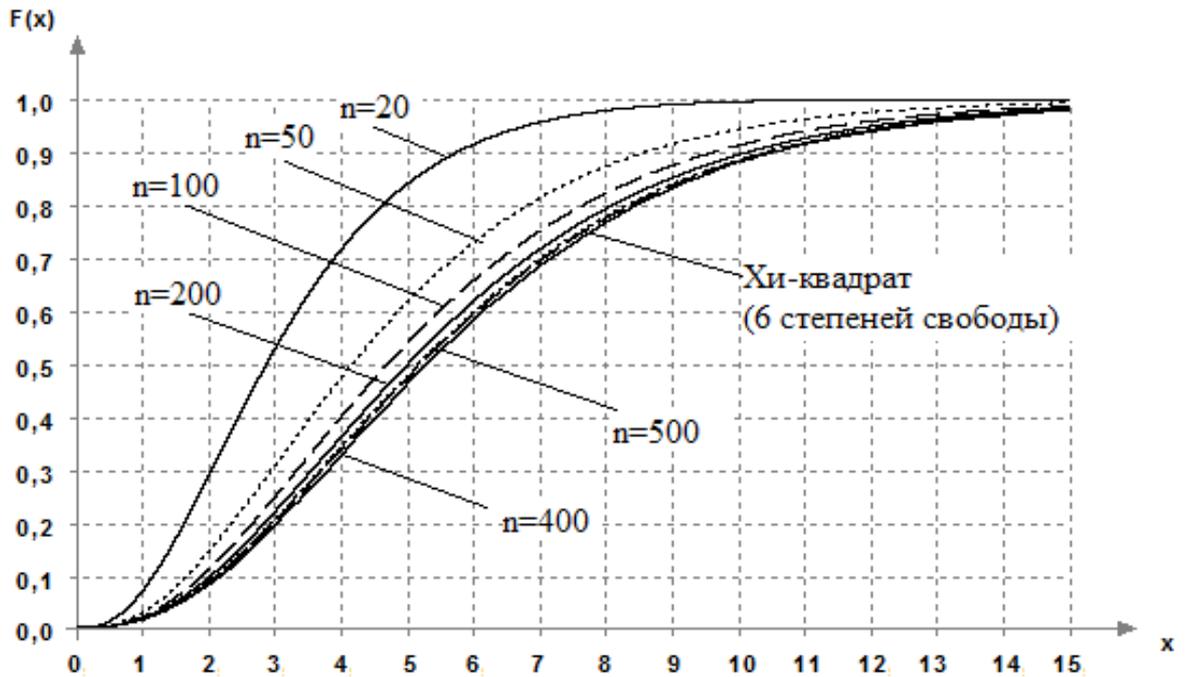


Рисунок 6.8 – Распределение статистики $\frac{n}{2(d+2)}b_{Mori}^{(1)}$ при $d=6$

6.3.1.3 Критерий проверки на асимметрию Шриваставы

Шривастава в 1984 году [268] предложил свою модификацию критерия многомерной нормальности, основанную на выборочном коэффициенте асимметрии.

Пусть

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{и} \quad S = n^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)(X_i - \bar{X}_n)^T = HD_\omega H^T \quad (6.15)$$

– это вектор выборочного среднего и выборочная ковариационная матрица соответственно, где $H = (h_1, h_2, \dots, h_d)$ является ортогональной матрицей, составленной из собственных векторов матрицы S , а $D_\omega = \text{diag}(\omega_1, \omega_2, \dots, \omega_d)$ и $\omega_1, \omega_2, \dots, \omega_d$ – это собственные значения матрицы S . Статистика критерия Шриваставы проверки многомерной нормальности вычисляется по формуле

$$b_s^{(1)} = \frac{1}{d} \sum_{i=1}^d \left\{ \frac{1}{\omega_i^{3/2}} \sum_{j=1}^n \frac{(y_{ij} - \bar{y}_i)^3}{n} \right\}^2, \quad (6.16)$$

где

$$y_{ij} = h_i' x_j \quad (i=1,2,\dots,d; j=1,2,\dots,n), \quad \bar{y}_i = \frac{1}{n} \sum_{j=1}^n y_{ij}.$$

В работе [268] Шривастава показал, что при больших значениях n статистика критерия (6.16) сходится к предельному Хи-квадрат распределению:

$$\frac{nd}{6} b_s^{(1)} \rightarrow \chi_d^2. \quad (6.17)$$

В работах [238, 176], представлена следующая модификация статистики $b_s^{(1)}$:

$$\frac{(n+1)(n+3)d}{6(n-2)} b_s^{(1)} \rightarrow \chi_d^2. \quad (6.18)$$

Скорость сходимости к предельному закону модифицированной статистики асимметрии Шриваставы ниже, чем у критериев Мори и Мардия, однако расстояние до предельного закона в 0,01 достигается в критерии Шриваставы при меньших объемах выборки. По рисунку 6.9 можно определить, что уже при объеме выборки $n=100$ расстояние до предельного закона мало для любого из рассмотренных значений d . А уровень в 0,01 расстояния D_n для $d=1, 3, 6$ достигается при $n=110, 200$ и 375 соответственно. Эти объемы n существенно меньше, чем в случае других рассматриваемых критериев, что делает критерий Шриваставы (в этом смысле) более предпочтительным для использования.

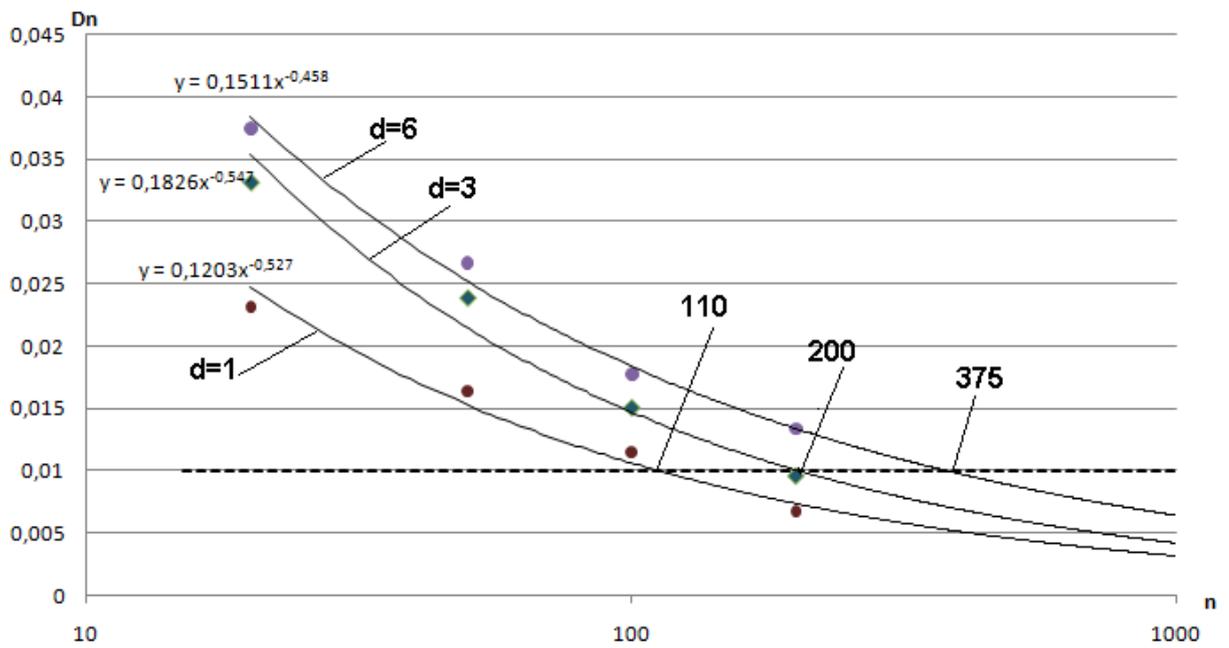


Рисунок 6.9 – Расстояние до предельного закона распределений статистик критерия Шриваставы

Графики функций распределения статистик критерия, приведенные на рисунках 6.10, 6.11 и 6.12, так же показывают быструю сходимость распределений статистик к предельному закону при рассмотренных объемах выборок $n = 20, 50, 100, 200, 400, 500$.

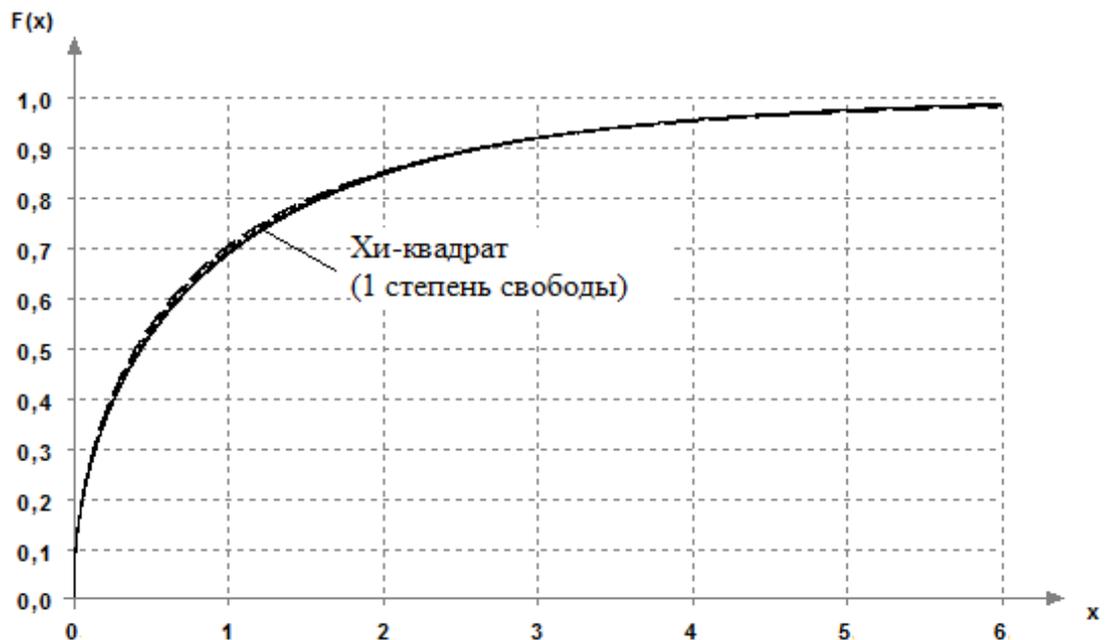


Рисунок 6.10 – Распределение статистики $\frac{(n+1)(n+3)}{6(d-2)}b_s^{(1)}$ при $d=1$

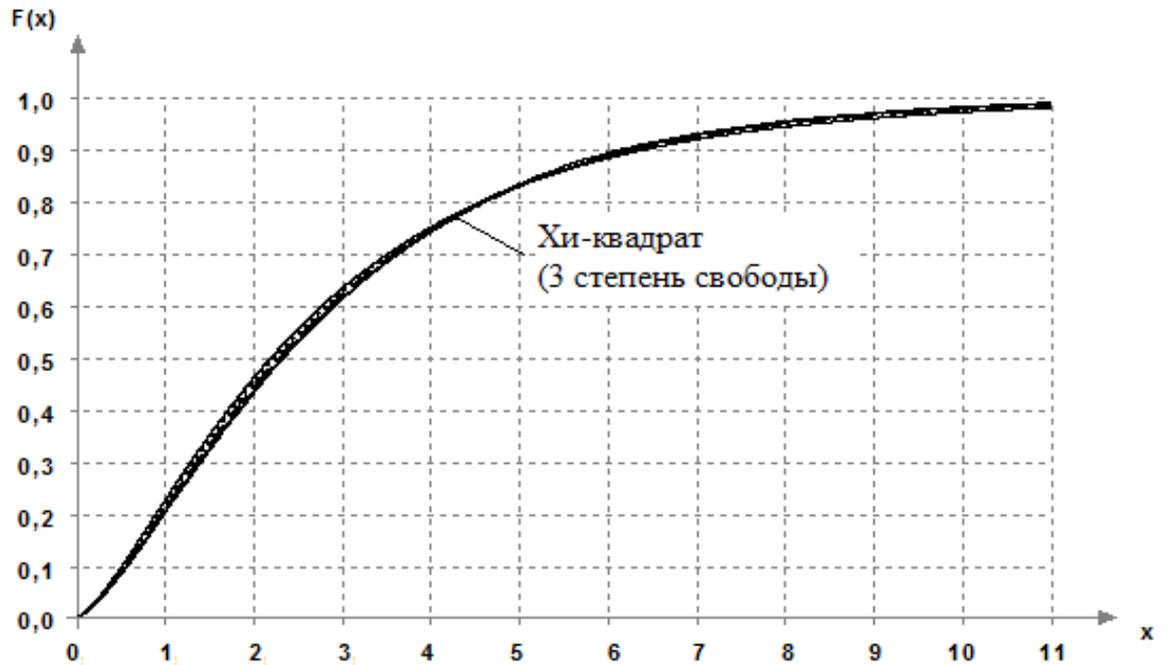


Рисунок 6.11 – Распределение статистики $\frac{(n+1)(n+3)}{6(d-2)}b_s^{(1)}$ при $d=3$

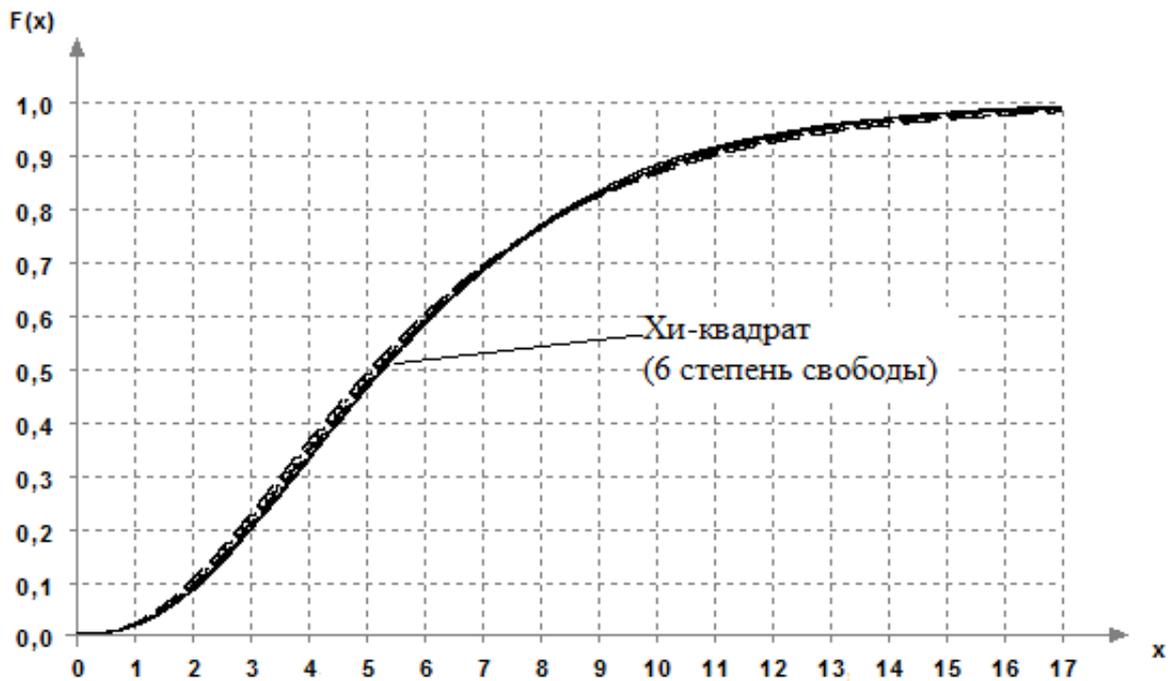


Рисунок 6.12 – Распределение статистики $\frac{(n+1)(n+3)}{6(d-2)}b_s^{(1)}$ при $d=6$

6.3.2 Критерии, основанные на вычислении коэффициента эксцесса

В критериях, основанных на выборочном коэффициенте эксцесса, вообще говоря, вместо гипотезы (6.2) проверяется гипотеза о том, что коэффициент эксцесса равен $d(d+2)$, где d – размерность многомерной случайной величины. Критерий эксцесса позволяет отклонить гипотезу о нормальности, но этого недостаточно для принятия гипотезы о согласии с нормальным распределением.

Рассмотрим критерии проверки многомерной нормальности, основанные на обобщении одномерного выборочного эксцесса

$$b_{n,1}^{(2)} = \frac{1}{n} \sum_{j=1}^n \left(\frac{X_j - \bar{X}_n}{\sigma_n} \right)^4, \quad (6.19)$$

который является оценкой центрального четвертого момента исходного распределения. Как обозначалось и выше, $\sigma_n^2 = \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X}_n)^2$ – это выборочная дисперсия.

6.3.2.1 Критерий проверки на эксцесс Мардия

В работе [225] предложена статистика для вычисления многомерного эксцесса

$$b_M^{(2)} = \frac{1}{n} \sum_{j=1}^n D_{n,jj}^2 = \frac{1}{n} \sum_{j=1}^n \|Y_{n,j}\|^4. \quad (6.20)$$

Критерии многомерной нормальности, основанные на вычислении статистики $b_M^{(2)}$, отклоняют гипотезу H_0 при больших и малых значениях статистики.

Если выборка величины X распределена нормально [225], то

$$\sqrt{n} (b_M^{(2)} - d(d+2)) \xrightarrow{D} N(0, 8d(d+2)), \quad (6.21)$$

или

$$\sqrt{\frac{n}{8d(d+2)}}(b_M^{(2)} - d(d+2)) \xrightarrow{D} N(0,1). \quad (6.22)$$

Теоретические свойства критерия проверки на эксцесс Мардия для многомерной нормальности изучались в [184] и [199].

Используя представление

$$\sqrt{n}(b_M^{(2)} - d(d+2)) = \int_0^\infty \xi^2 d\vartheta_n(\xi), \quad (6.23)$$

где $\vartheta_n(\cdot)$ является некоторым эмпирическим процессом $D_{n,1}, \dots, D_{n,n}$, в работе [199] доказывается, что предельное распределение $\sqrt{n}(b_M^{(2)} - d(d+2))$ для близких альтернатив, относительно нормального распределения, есть нормальное распределение с дисперсией $8d(d+2)$.

Для фиксированных (стандартизированных) альтернативных распределений удовлетворяющих $E \|X\|^8 < \infty$, верно

$$\sqrt{n}(b_M^{(2)} - E \|X\|^4) \xrightarrow{D} N(0, \sigma^2), \quad (6.24)$$

где σ^2 зависит от распределения X [184]. Если распределение P^X сферически-симметричное, то $\sigma^2 = r_8 - r_4^2 + 4d^{-1}r_4(d^{-1}r_4^2 - r_6)$, где, $r_i = E \|X\|^i$.

6.3.2.2 Критерий проверки на эксцесс Козиола

Статистика критерия проверки на эксцесс, рассматриваемая в [200], имеет вид:

$$\tilde{b}_{n,d}^{(2)} = \frac{1}{n^2} \sum_{j=1}^n D_{n,jk}^4 = \frac{1}{n^2} \sum_{j,k=1}^n (Y_{n,j}^T Y_{n,k})^4. \quad (6.25)$$

Заметим, статистика $\tilde{b}_{n,d}^{(2)}$ оценивает по выборке величину $\tilde{\beta}_d^{(2)} = E(X_1^T X_2)^4$, и в одномерном случае является квадратом четвертого момента (стандартизированного) распределения.

В [184] показано, что если $E \| X \|^8 < \infty$, то $\sqrt{n}(\tilde{b}_{n,d}^{(2)} - \tilde{\beta}_d^2) \xrightarrow{D} N(0, \tilde{\sigma}^2)$, где $\tilde{\sigma}^2$ зависит от некоторых смешанных выборочных моментов X высшего порядка. Если P^X сферически-симметричное, то

$$\tilde{\sigma}^2 = \frac{36r_4^2}{d^2(d+2)^2} (r_8 - r_4^2 + 4d^{-1}r_4(d^{-1}r_4^2 - r_6)), \quad (6.26)$$

где $r_i = E \| X \|^i$ - момент i -го порядка. В частности, $\sqrt{n}(\tilde{b}_{n,d}^{(2)} - 3d(d+2)) \xrightarrow{D} N(0, 288d(d+2))$ при справедливой гипотезе H_0 . В [184] также доказывается, что в асимптотике, корреляция между $b_M^{(2)}$ и $\tilde{b}_{n,d}^{(2)}$ равна 1, если P^X сферически-симметричное распределение.

Предельным распределением статистик критериев проверки на коэффициент эксцесса является нормальный закон распределения с дисперсией, зависящей от объема выборки n и размерности d . На математическое ожидание предельного закона влияет объем многомерной выборки. Для статистики критерия (6.25) параметр сдвига предельного закона равен $3d(d+2)$. Учитывая приведенные взаимосвязи, статистика (6.27) имеет стандартное нормальное распределение:

$$\frac{\sqrt{n}}{\sqrt{288d(d+2)}} (b_K^{(2)} - 3d(d+2)) \rightarrow N(0,1). \quad (6.27)$$

На рисунке 6.13 и 6.14 приведена зависимость расстояния распределения статистики критерия (6.25) до предельного закона.

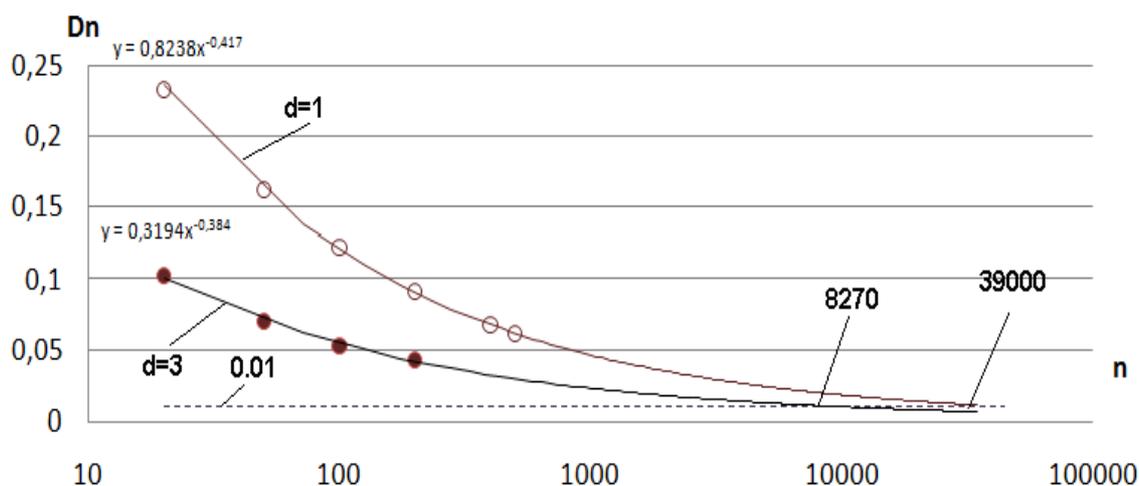


Рисунок 6.13 – Расстояние до предельного закона распределений статистик критерия (6.25) ($d = 1, 3$)

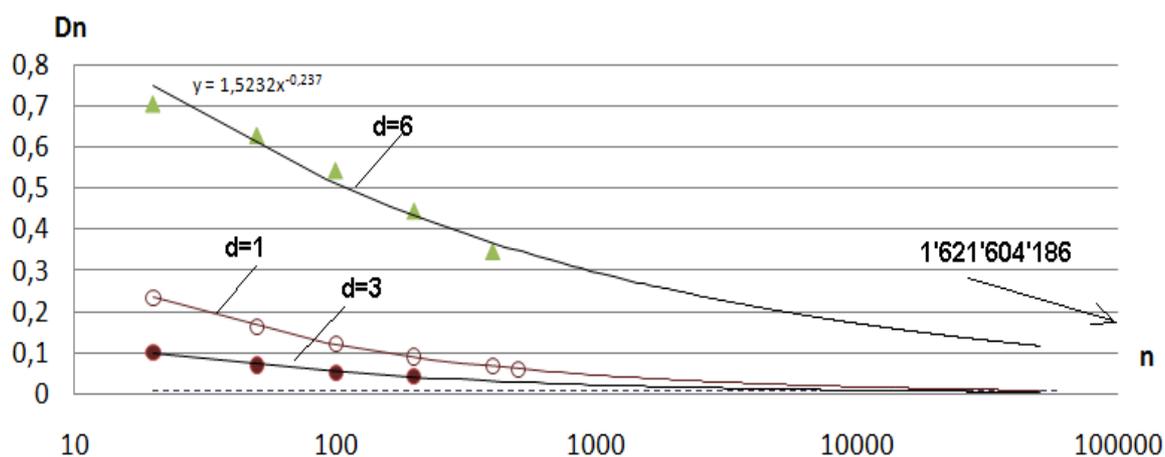


Рисунок 6.14 – Расстояние до предельного закона распределений статистик критерия (6.25) ($d = 1, 3, 6$)

Графики на рисунках 6.13 и 6.14 показывают, что при увеличении размерности многомерной нормально распределенной выборки график расстояния до предельного закона сначала оказывается ниже ($d = 3$), затем выше ($d = 6$), чем при ($d = 1$). И при больших размерностях ($d = 6$) границу в 0.01 функция расстояния достигает при объеме выборки порядка 10^9 . Однако последнее утверждение получено с помощью аппроксимации и нахождения коэффициентов уравнений регрессии, что увеличивает погрешность вычислений. Поэтому нельзя говорить, что лишь при $n = 10^9$ расстояние до

предельного закона сократится до 0,01. Но все же, объем выборки слишком велик для корректного применения в качестве предельного нормального стандартного закона.

6.3.2.3 Критерий проверки на эксцесс Шриваставы

Критерий Шриваставы проверки на эксцесс основан на разложении (6.15) ковариационной матрицы S . В работе [268] выборочный многомерный эксцесс определяется как

$$b_S^{(2)} = \frac{1}{nd} \sum_{i=1}^d \frac{1}{\omega_i^2} \sum_{j=1}^n (y_{ij} - \bar{y}_i)^4, \quad (6.28)$$

где

$$y_{ij} = h_i' x_j (i = 1, 2, \dots, d; \quad j = 1, 2, \dots, n), \quad \bar{y}_i = \frac{1}{n} \sum_{j=1}^n y_{ij}.$$

Предельным распределением статистики (6.28) является нормальное с параметрами, зависящими от объема выборки и размерности случайной величины:

$$b_S^{(2)} \rightarrow N\left(3, \frac{24}{nd}\right). \quad (6.29)$$

6.3.2.4 Модификации статистик критериев проверки на эксцесс Мардия и Шриваставы

В [225] и [268] получены асимптотические распределения статистик $b_M^{(2)}$ и $b_S^{(2)}$. При неизвестной ковариационной матрице Σ статистика критерия Мардия имеет вид

$$z_M = \sqrt{n} \frac{b_M^{(2)} - d(d+2)}{\sqrt{8d(d+2)}}, \quad (6.30)$$

а статистика критерия Шриваставы (6.31)

$$z_S = \sqrt{nd} \frac{b_S^{(2)} - 3}{\sqrt{24}}. \quad (6.31)$$

Распределения статистик Z_M и Z_S сходятся к стандартному нормальному распределению с ростом объема выборки. Однако сходимость к предельному закону статистик (6.30) и (6.31) достаточно медленная, в связи с чем в работе [182] получены модификации (6.32) и (6.33) статистик критериев, ускоряющие сходимость к предельному закону

$$z_M^* = \frac{b_M^{(2)} - d(d+2)(n-1)(n+1)^{-1}}{\sqrt{8d(d+2)(n-3)(n-d-1)(n-d+1)(n+1)^{-2}(n+3)^{-1}(n+5)^{-1}}}, \quad (6.32)$$

$$z_S^* = \left(b_S^{(2)} - (3 - 6n^{-1} + 3n^{-2}) \right) \left(\frac{24}{nd} - \frac{360}{n^2d} - \frac{2976}{n^3d} \right)^{-1/2}. \quad (6.33)$$

Как и в случае статистик Z_M и Z_S , предельным распределением вычисленных по формулам (6.32) и (6.33) статистик критериев Мардия и Шриваставы является стандартный нормальный закон.

Следующие статистики получены при использовании нормализующего преобразования для $b_M^{(2)}$ и $b_S^{(2)}$. Метод, нормализующий некоторые статистики многомерного анализа, обсуждался в [199] и [182]. Нормализующее преобразование критерия Мардия основано на вычислении коэффициента эксцесса, если ковариационная матрица неизвестна:

$$z_{MNT} = \frac{\sqrt{n} \left\{ p \cdot \exp[b_M^{(2)} / p] - p \cdot \exp[d(d+2) / p] - c / n \right\}}{\sqrt{8d(d+2) \exp[d(d+2) / p]}}, \quad (6.34)$$

где

$$p = -\frac{3d(d+2)}{d+8} \quad c = -\frac{2}{3}(3d^2 + 8d + 16) \exp[-(d+8)/3].$$

Нормализующее преобразование критерия Шриваставы [261] имеет вид:

$$z_{SNT} = \frac{\sqrt{nd} \{-\exp(-b_s^{(2)}) + e^{-3} - \tilde{c}n^{-1}\}}{\sqrt{24}e^{-3}}, \quad (6.35)$$

где

$$\tilde{c} = -6e^{-3}(1 + 2/d).$$

Распределения последних статистик хорошо описывается стандартным нормальным законом.

По рисункам 6.16 и 6.17 видно, что распределение статистики Z_{Mnt} быстрее других сходится к предельному нормальному закону. Ниже скорость сходимости у статистики Z_M .

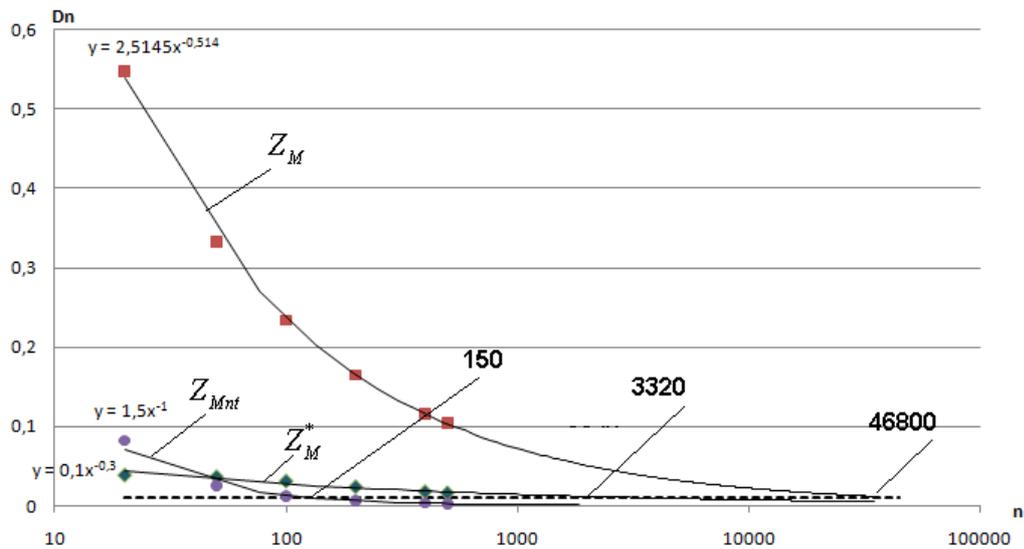


Рисунок 6.15 – Расстояние до предельного закона распределений статистик критерия эксцесса Мардия ($d = 6$)

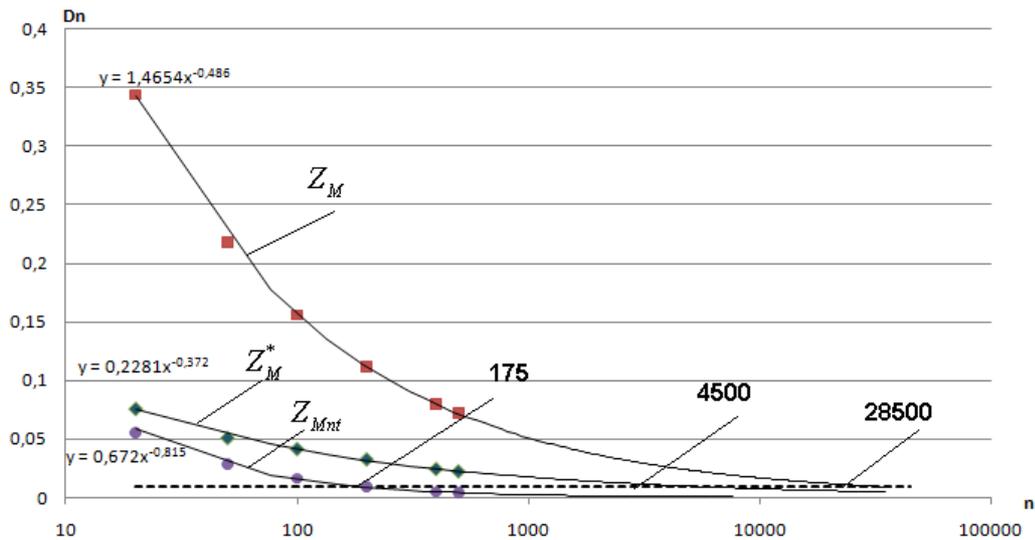


Рисунок 6.16 – Расстояние до предельного закона распределений статистик критерия эксцесса Мардия ($d = 3$)

Для проверки нормальности многомерной выборки при количестве наблюдений $n \sim 100$ можно использовать критерии Мардия со статистиками Z_{Mnt} и Z_M^* , и в качестве предельного закона брать стандартное нормальное распределение, поскольку расстояние D_n до предельного закона около 0,01.

Так же как и критерий проверки на эксцесс Мардия, критерий Шриваставы имеет три модификации (6.31), (6.33) и (6.35). Распределения статистик этих модификаций сходятся к стандартному нормальному закону.

Среди статистик Z_{Snt} , Z_S^* и Z_S быстрее сходится к предельному стандартному нормальному закону распределение статистики Z_{Snt} , медленней – Z_S .

На рисунках 6.17 и 6.18 представлены графики, показывающие зависимость от объема выборки расстояния функции распределения статистики критерия эксцесса Шриваставы от стандартного нормального закона. На рисунках 6.15 и 6.16 показаны аналогичные зависимости для статистик критерия эксцесса Мардия. Статистики Z_{Snt} и Z_{Mnt} , полученные с помощью

нормализующих преобразований, уже при объеме выборки $n=100$ хорошо аппроксимируются стандартным нормальным законом. Распределения модифицированных статистик Z_S , Z_S^* и Z_M, Z_M^* приходят к предельному нормальному закону при очень больших объемах наблюдений, которые на практике встречаются крайне редко.

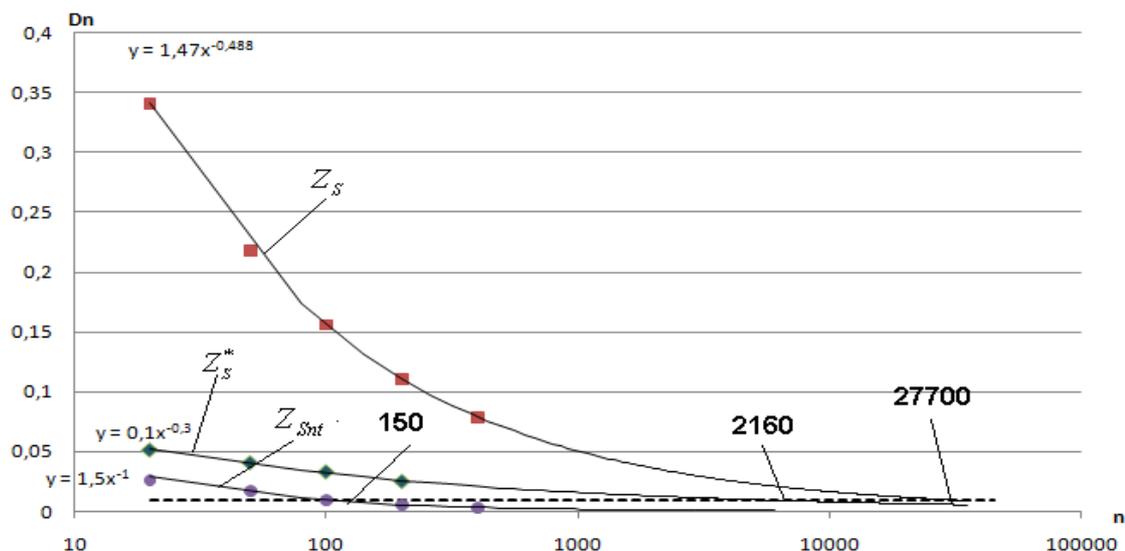


Рисунок 6.17 – Расстояние до предельного закона распределений статистик критерия эксцесса Шриваставы ($d = 6$)

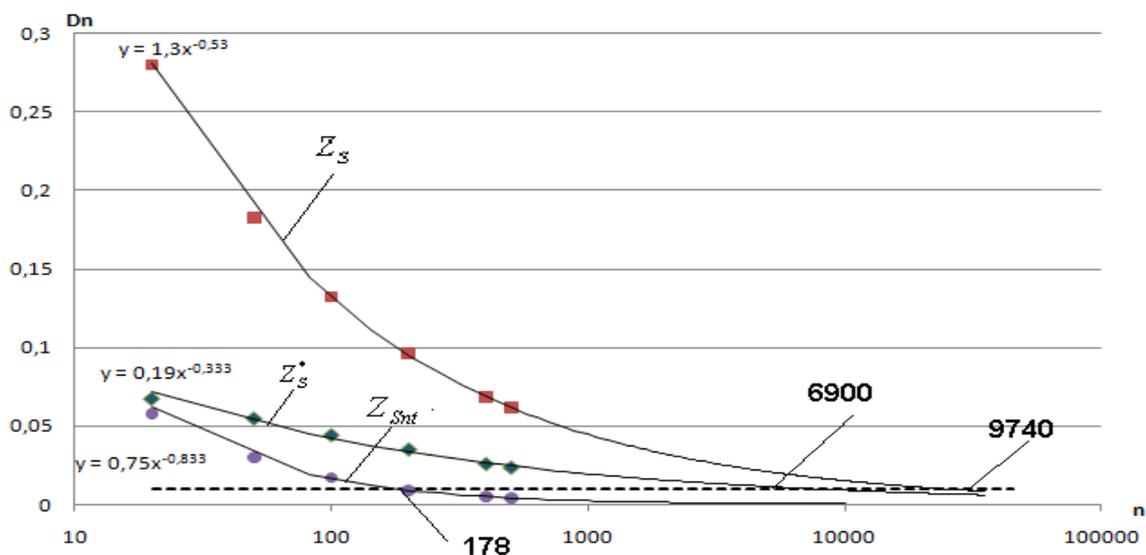


Рисунок 6.18 – Расстояние до предельного закона распределений статистик критерия эксцесса Шриваставы ($d = 3$)

6.3.3 Критерии многомерной нормальности, основанные на полярной декомпозиции вектора наблюдения случайной величины

Вектор наблюдений X можно представить в виде $X = U \cdot R$, где $U = X / \|X\|$ угловая часть вектора X , а $R = \|X\|$ радиальная часть X . При этом нормальное распределение для d -мерного нормального распределения выполняются следующие свойства [185]:

- U и R независимы;
- U имеет равномерное распределение на поверхности единичной d -сферы $S^{d-1} = \{u \in R^d : \|u\| = 1\}$;
- квадрат радиуса R^2 подчиняется распределению χ_d^2 .

6.3.4.1 Критерий, основанный на угловой части X

Данная процедура была впервые изучена Рэлеем (Rayleigh) в 1919 году для проверки однородности S^2 [249], а в 1983 предложена Козиолом [198] в качестве критерия многомерной нормальности.

Статистика критерия имеет вид

$$U_n = \left\| \frac{1}{n} \sum_{j=1}^n U_{n,j} \right\| = \frac{1}{n} \sum_{j,k=1}^n \frac{Y_{n,j}^T Y_{n,k}}{\left(Y_{n,j}^T Y_{n,j} \right)^{1/2} \left(Y_{n,k}^T Y_{n,k} \right)^{1/2}} \quad (6.36)$$

Гипотеза о нормальности отвергается, при больших значениях U_n .

Козиол установил, что при $n \rightarrow \infty$ [198]

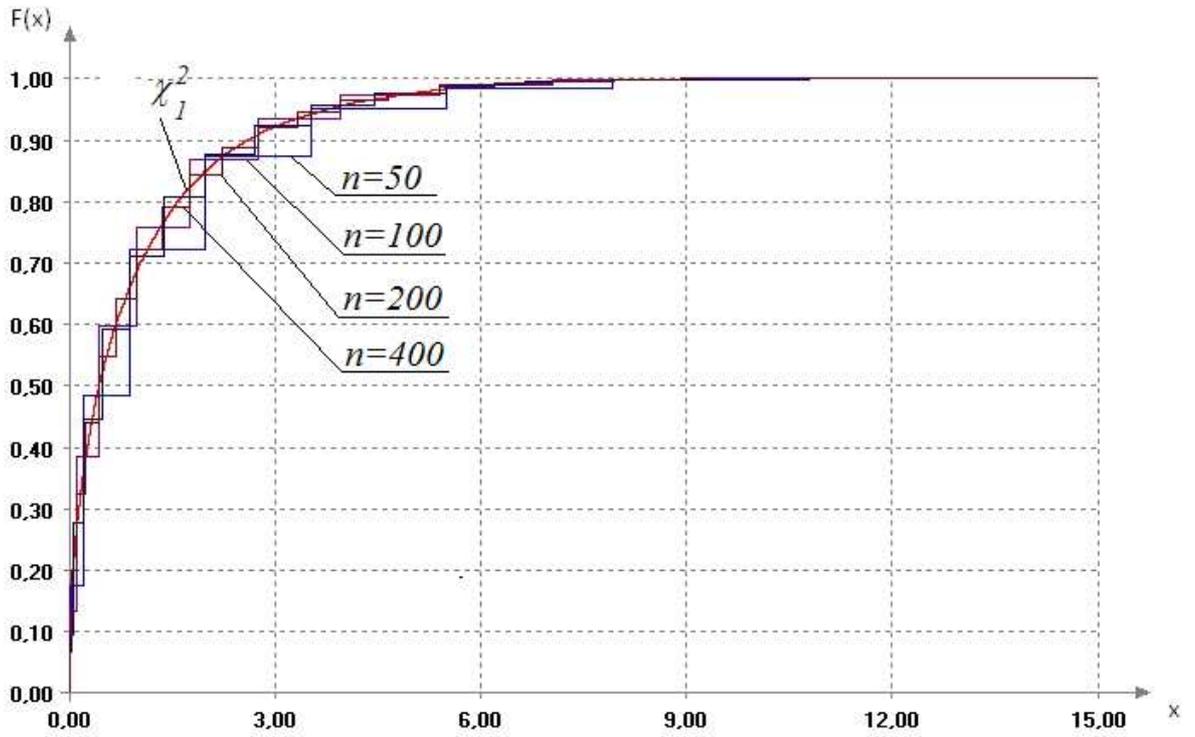
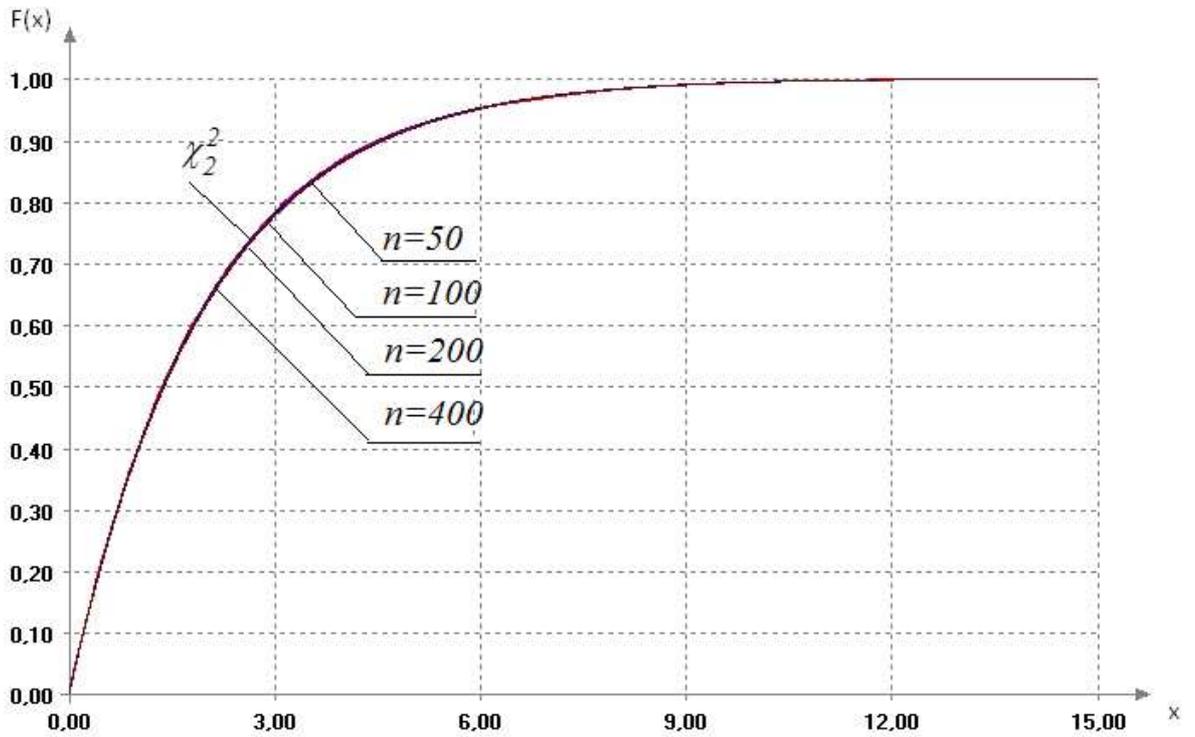
$$nU_n \rightarrow \left[1 - \frac{2}{d} \left(\frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)} \right)^2 \right] \chi_d^2. \quad (6.37)$$

Распределение статистики зависит от размерности величины и объема выборки. Поэтому будем удобнее анализировать сходимость к предельному распределению не самой статистики U_n , а величины

$$\tilde{U}_n = \frac{n}{1 - \frac{2}{d} \left(\frac{\Gamma((d+1)/2)}{\Gamma(d/2)} \right)^2} U_n \rightarrow \chi_d^2. \quad (6.38)$$

На рисунках 6.19-6.22 представлены функции распределения статистики (6.38) для размерностей $d=1,2,4$. По рисункам видно, что статистика (6.38) сходится к предельному Хи-квадрат распределению. Для того, что бы определить более точные объемы выборки, при которых статистика критерия отклоняется от предельного закона не более чем на 0,01, вычислим расстояние Колмогорова D_n при разных объемах выборки.

Полученные значения аппроксимируем уравнением степенной регрессии. На рисунке 6.23 и в таблице 6.1 представлена зависимость расстояния D_n от объема выборки n для критерия, основанного на углах, при разных n . Пунктирной линией обозначен уровень, при котором расстояние D_n равно 0,01.

Рисунок 6.19 – Распределение статистики (6.38) для $d=1$ Рисунок 6.20 – Распределение статистики (6.38) при $d=2$

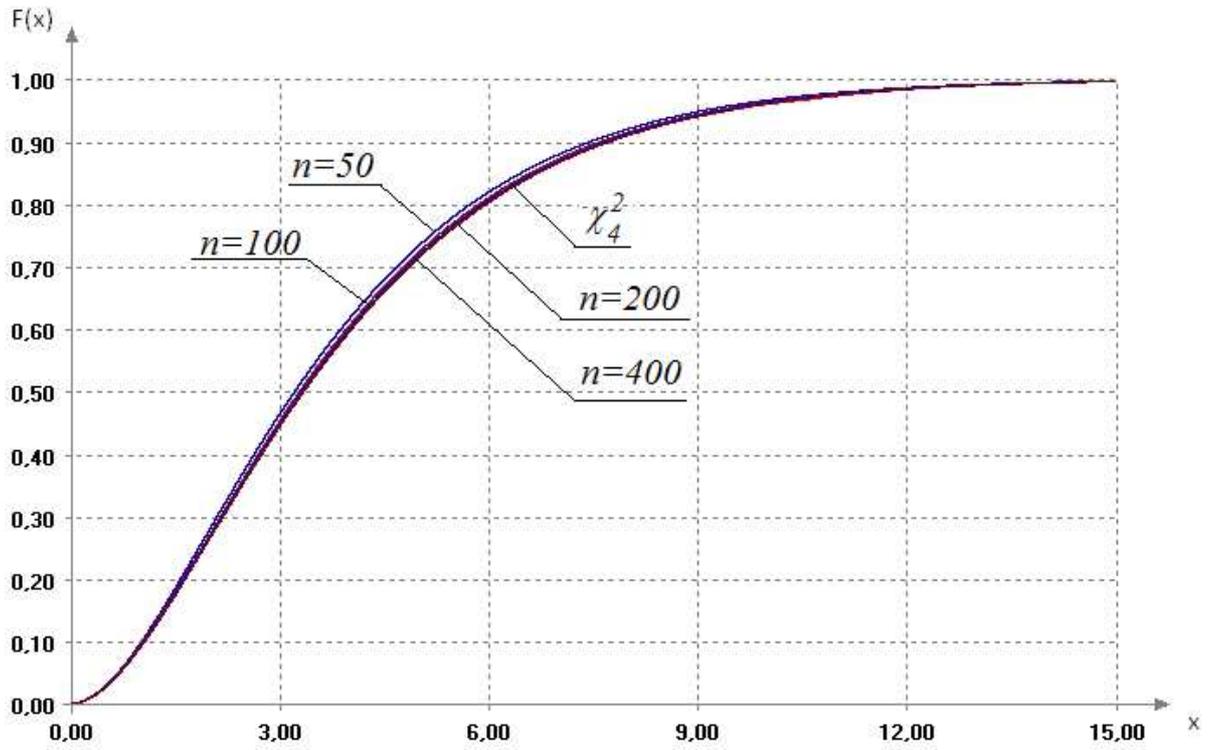


Рисунок 6.21 – Распределение статистики (6.38) при $d=4$

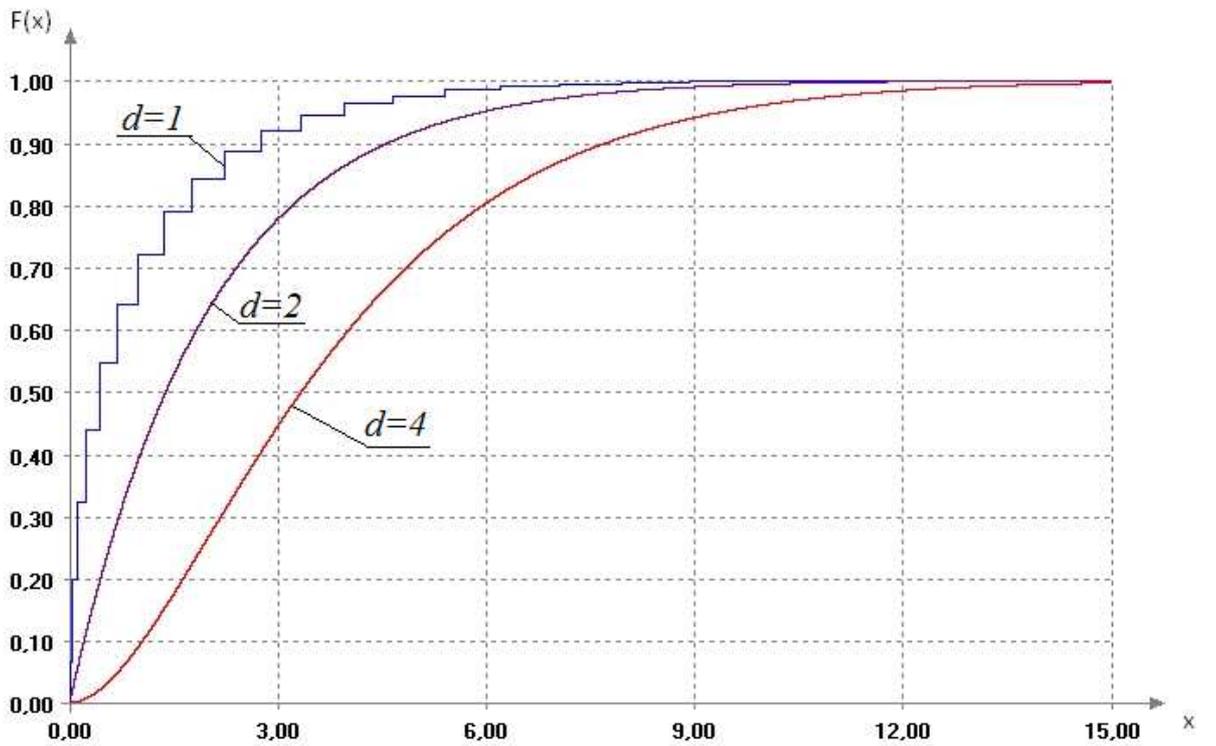


Рисунок 6.22 – Распределение статистики (6.38) при различных d , для $n = 400$

Таблица 6.1 – Расстояние $D_{n,N}$ до предельного закона распределений статистик критерия, основанного на углах

	$n = 50$	$n = 100$	$n = 200$	$n = 400$
$d = 1$	0,186684	0,132355	0,093564	0,066279
$d = 2$	0,003514	0,001541	0,000954	0,000716
$d = 4$	0,023007	0,011214	0,005359	0,002359

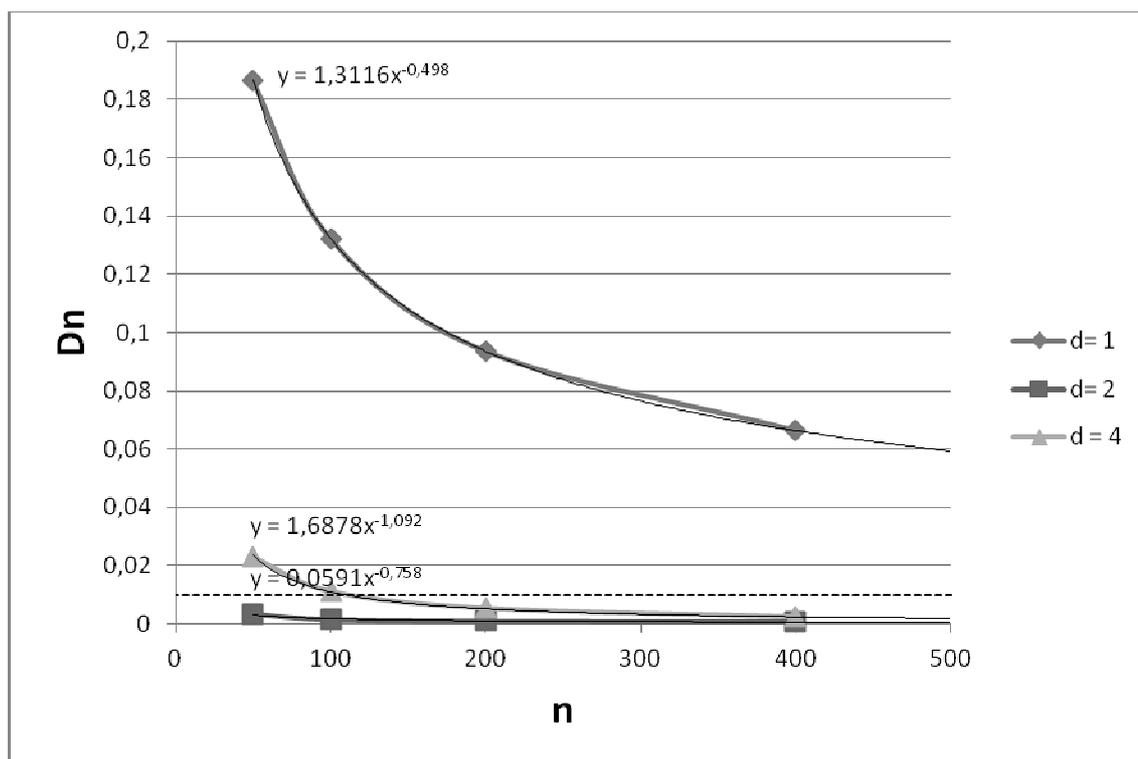


Рисунок 6.23 – Расстояние до предельного закона распределений статистик критерия основанного на углах

Для одномерной выборки наблюдается сильная "ступенчатость" слабо зависящая от n , что дает более медленную сходимость к предельному закону. При $d > 1$, видно, что с ростом размерности многомерной выборки скорость сходимости к предельному закону уменьшается. "Пороговым" значением объема выборки n , при котором расстояние до предельного закона меньше или равно 0,01, является 20972, 10 и 110 наблюдений, для $d = 1$, $d = 2$ и $d = 4$ соответственно.

Очевидно, что использовать Хи-квадрат распределение, как предельный закон для критерия, основанного на углах, можно только при $d > 1$, для $d = 1$ требуется слишком большой объем выборки.

6.3.4.2 Критерии, основанные на радиальной части X

В основе данных критериев лежит квадрат радиуса:

$$R_j^2 = D_{n,jj} = \|Y_{n,j}\|^2 = (X_j - \bar{X}_n)^T S_n^{-1} (X_j - \bar{X}_n). \quad (6.39)$$

Упорядочим значения R_j^2 и обозначим их как $D_{(j)}$, $j = 1, 2, \dots, n$:

$$D_{(1)} \leq D_{(2)} \leq \dots \leq D_{(n)}. \quad (6.40)$$

Пусть G_d - функция χ_d^2 -распределения. Тогда проверка гипотезы о нормальности эквивалентна проверке гипотезы о том, что выборка (6.40) подчинена χ_d^2 -распределению. Это можно сделать с помощью критериев согласия, описанных в главе 4. В 1982 Козиол предложил использовать статистику Крамера-Мизеса [197], в 1987 Полсон, Рухон и Сулло предложили использовать критерий Андерсона-Дарлинга [240].

Гипотеза о том, что выборка (6.40) подчинена χ_d^2 -распределению, является простой. Но так как $D_{n,jj}$ являются зависимыми друг от друга величинами [185], 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.

Распределения статистик критериев Колмогорова, Андерсона-Дарлинга и Крамера-Мизеса зависят от размерности d (см. рисунки 6.23-6.25) и хорошо аппроксимируются бета-распределением III-го рода (таблицы 6.2-6.4). Аппроксимация проводилась по выборкам статистик критериев радиуса при $n = 1000$ (объем выборки многомерной нормальной случайной величины),

$N = 1000000$ (количество выборок). С ростом размерности d различия между законами распределения статистик уменьшается так, что графики законов распределений на рисунках 6.23-25 при $d = 7$ и $d = 8$ практически неразличимы. Поэтому для практических целей при размерностях больших 8 можно пользоваться аппроксимацией для $d = 8$.

Бета-распределение III-го рода является обобщением бета-распределения с функцией распределения (6.41) и функцией плотности (6.42).

$$F(x; \bar{\theta}) = \frac{1}{B(\theta_0, \theta_1)} B\left(\frac{\theta_2 \frac{x - \theta_4}{\theta_3}}{1 + (\theta_2 - 1) \frac{x - \theta_4}{\theta_3}}, \theta_0, \theta_1\right), x \in [\theta_4, \theta_4 + \theta_3] \quad (6.41)$$

$$f(x; \bar{\theta}) = \begin{cases} \frac{\theta_2^{\theta_0}}{\theta_3 B(\theta_0, \theta_1)} \frac{\left(\frac{x - \theta_4}{\theta_3}\right)^{\theta_0 - 1} \left(1 - \frac{x - \theta_4}{\theta_3}\right)^{\theta_1 - 1}}{\left[1 + (\theta_2 - 1) \frac{x - \theta_4}{\theta_3}\right]^{\theta_0 + \theta_1}}, & x \in [\theta_4, \theta_4 + \theta_3]; \\ 0, & x \notin [\theta_4, \theta_4 + \theta_3]. \end{cases} \quad (6.42)$$

Так как бета-распределение III-го рода в качестве области определения имеет отрезок $[\theta_4, \theta_4 + \theta_3]$, то параметры θ_3 и θ_4 выбирались вручную: левая граница $\theta_4 = 0$, а правая граница получена, исходя из условия, чтобы правая граница была заведомо больше максимальной порядковой статистики из смоделированных выборок. Расстояние между эмпирическими распределениями и аппроксимирующими законами вычислено по расстоянию Колмогорова $D_{n,N}$ и показано в последних колонках таблиц 6.2-6.4. Погрешность аппроксимации принимает значения от 0,006 до 0,009, и, таким образом, найденные аппроксимирующие законы могут использоваться в практических целях.

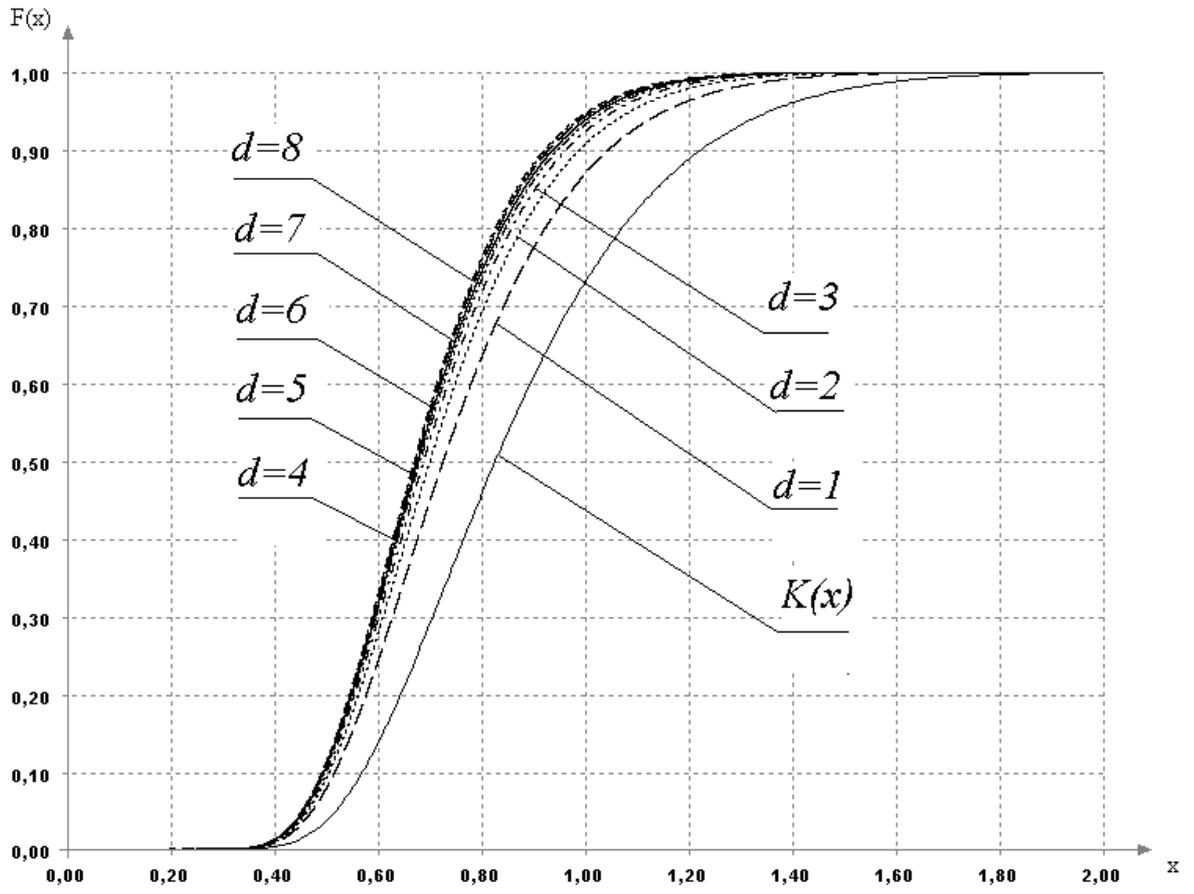


Рисунок 6.24 – Распределение статистики критерия радиуса Колмогорова при $n=1000$ и $d=1,2,\dots,8$; $K(x)$ - распределение Колмогорова

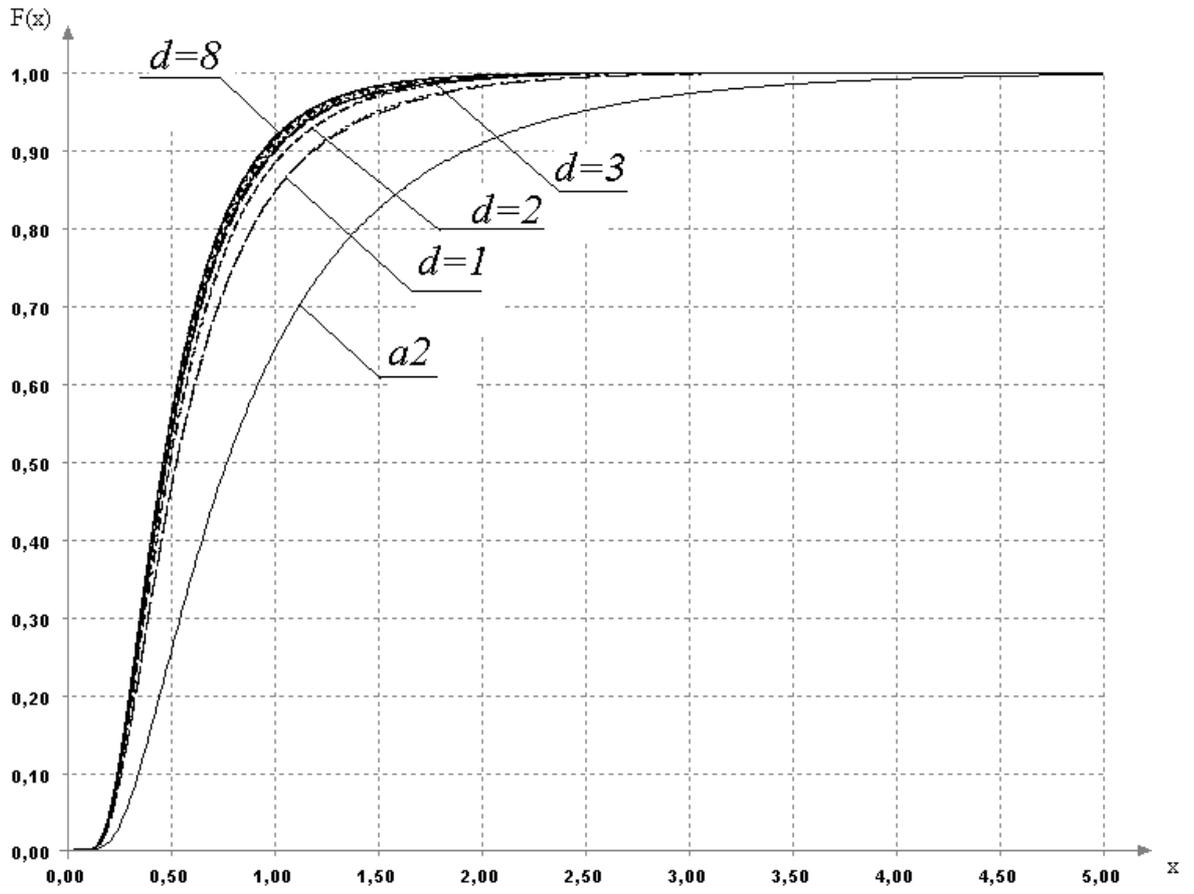


Рисунок 6.25 – Распределение статистики критерия радиуса Андерсона-Дарлинга при $n = 1000, d = 1, 2, \dots, 8$; $a2$ - распределение статистики Андерсона-Дарлинга

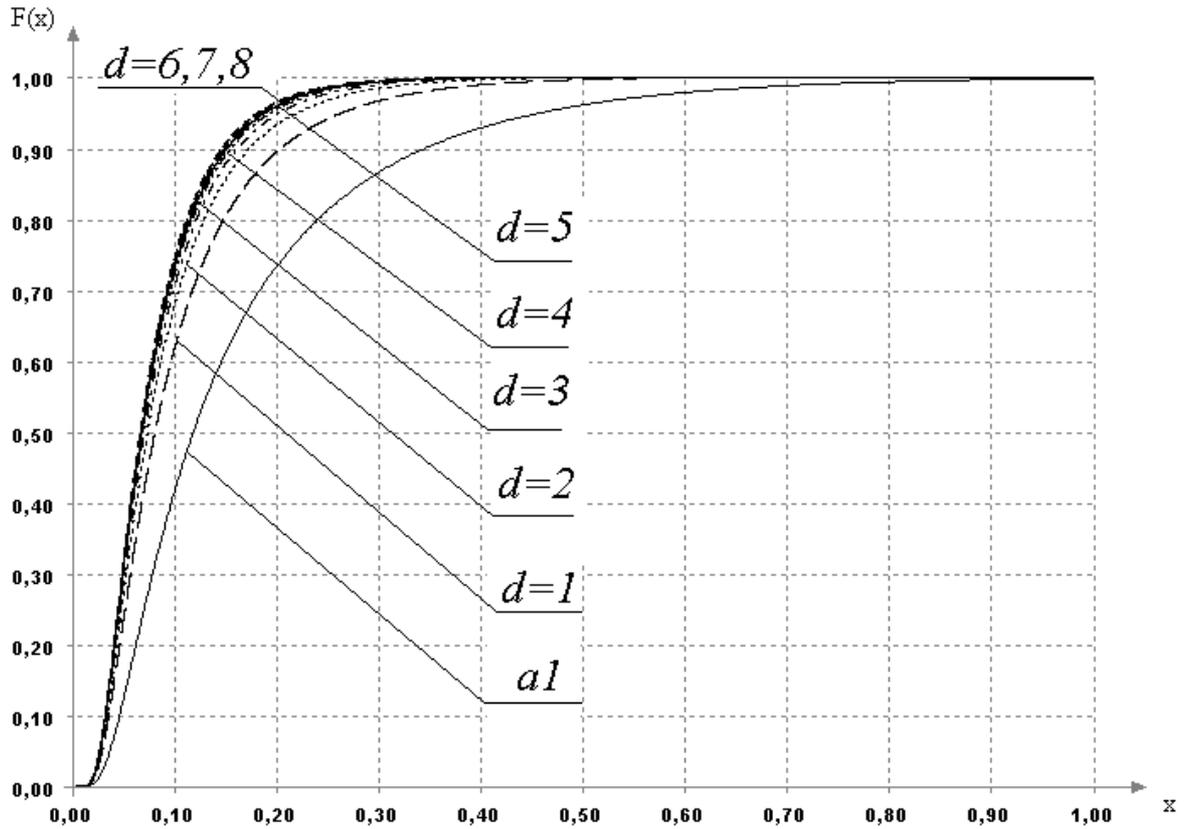


Рисунок 6.26 – Распределение статистики радиуса Крамера-Мизеса-Смирнова при $n = 100$, $d = 1, 2, \dots, 8$; $a1$ - распределение статистики Крамера-Мизеса-Смирнова

Таблица 6.2 – Аппроксимации предельных распределений критерия радиуса Колмогорова

d	Аппроксимирующий закон распределения	$D_{n,N}$
1	Бе-III (53.6236, 15.3955, 45.0496, 10.0000, 0.0000)	0,009
2	Бе-III (61.3277, 16.7835, 49.1481, 10.0000, 0.0000)	0,007
3	Бе-III (65.7375, 17.6448, 51.0536, 10.0000, 0.0000)	0,007
4	Бе-III (68.9325, 18.1012, 52.7890, 10.0000, 0.0000)	0,006
5	Бе-III (69.5879, 18.5017, 52.4714, 10.0000, 0.0000)	0,007
6	Бе-III (68.1034, 18.9962, 50.2919, 10.0000, 0.0000)	0,006
7	Бе-III (68.1446, 19.2711, 49.7699, 10.0000, 0.0000)	0,006
8	Бе-III (51.2230, 21.5792, 33.3143, 10.0000, 0.0000)	0,007

Таблица 6.3 – Аппроксимации предельных распределений критерия радиуса Андерсона-Дарлингга

d	Аппроксимирующий закон распределения	$D_{n,N}$
1	Be-III (14.0000, 3.6001, 74.3281, 10.0000, 0.0000)	0,008
2	Be-III (14.0000, 4.0891, 69.4148, 10.0000, 0.0000)	0,007
3	Be-III (14.0000, 4.3370, 67.1246, 10.0000, 0.0000)	0,007
4	Be-III (14.0000, 4.4716, 66.0749, 10.0000, 0.0000)	0,007
5	Be-III (14.0000, 4.5555, 65.3704, 10.0000, 0.0000)	0,007
6	Be-III (14.0000, 4.6272, 64.8145, 10.0000, 0.0000)	0,007
7	Be-III (14.0000, 4.6756, 64.3915, 10.0000, 0.0000)	0,007
8	Be-III (14.0000, 4.6935, 64.3243, 10.0000, 0.0000)	0,007

Таблица 6.4 – Аппроксимации предельных распределений критерия радиуса Крамера-Мизеса-Смирнова

d	Аппроксимирующий закон распределения	$D_{n,N}$
1	Be-III (7.9674, 3.1108, 64.5654, 2.0000, 0.0000)	0,009
2	Be-III (9.4144, 3.5315, 73.9401, 2.0000, 0.0000)	0,009
3	Be-III (9.5174, 3.7138, 73.9973, 2.0000, 0.0000)	0,007
4	Be-III (9.8276, 3.8724, 75.0000, 2.0000, 0.0000)	0,009
5	Be-III (9.7771, 3.9000, 75.0000, 2.0000, 0.0000)	0,007
6	Be-III (9.6594, 3.9050, 75.0000, 2.0000, 0.0000)	0,006
7	Be-III (9.6528, 3.9400, 75.0000, 2.0000, 0.0000)	0,006
8	Be-III (9.6460, 3.9592, 75.0000, 2.0000, 0.0000)	0,006

6.4 Сходимость распределений статистик критериев многомерной нормальности к предельному закону

Все рассмотренные в данной главе критерии имеют предельные распределения. Однако скорость сходимости статистик критериев к предельному закону ранее практически не исследовалась, за исключением технического отчета Hara и Seo [182] за февраль 2011, в котором с помощью метода Монте-Карло исследовалось изменение выборочных характеристик статистик с ростом объема выборки. Информации о том, какой из рассматриваемых критериев является более мощным, в современной литературе найти не удалось.

По результатам настоящих исследований видно, что статистики критериев асимметрии и эксцесса сходятся к своим предельным распределениям. С увеличением размерности при небольших объемах выборок ($n = 20, n = 50$) расстояние от распределения статистики до предельного закона увеличивается.

Статистика критерия проверки на коэффициент асимметрии Шриваставы обладает высокой скоростью сходимости к предельному закону: уже при объеме выборки $n = 20$ расстояние D_n до предельного Хи-квадрат распределения оказывается меньше 0,05. Медленнее других статистик критериев проверки на асимметрию, сходится к предельному распределению статистики критерия Мардия.

Распределения статистик критериев проверки многомерной выборки на эксцесс Козиола, Мардия и Шриваставы плохо сходятся к предельным распределениям. Модификации критериев Мардия (Z_{Mnt}, Z_M^*) и Шриваставы (Z_{Snt}, Z_S^*) увеличивают скорость сходимости распределений статистик к предельному закону: при объеме выборки $n = 20$ расстояние до предельного закона $D_n \approx 0.05$.

Результаты исследований показали, что распределение статистики критерия, основанного на угловой части X , также с ростом n сходится к предельному закону. При увеличении размерности d , расстояние до предельного закона (при том же n) увеличивается.

Для одномерной выборки распределения статистики критерия, основанного на угловой части вектора наблюдений X , отмечена сильная "ступенчатость", слабо зависящая от n , что приводит к более медленной сходимости к предельному закону. С ростом размерности $d > 1$ многомерной выборки скорость сходимости к предельному закону уменьшается. "Пороговым" значением объема выборки n , при котором расстояние до

предельного закона меньше или равно 0,01, является 20972, 10 и 110 наблюдений, для $d=1$, $d=2$ и $d=4$ соответственно. Очевидно, что использовать Хи-квадрат распределение в качестве предельного закона для рассматриваемого критерия можно только при $d > 1$, так как для $d=1$ требуются большие объемы выборки.

Показано, что распределения статистик критериев согласия Андерсона-Дарлинга, Крамера-Мизеса и Колмогорова, применяемых для анализа многомерных статистик, формируемых при проверке многомерной нормальности и базирующихся на анализе радиальной части X , хорошо аппроксимируются бета-распределениями 3-го рода.

6.5 Исследование мощности критериев проверки многомерной нормальности

В данном пункте рассмотрена мощность инвариантных критериев многомерной нормальности.

6.5.1 Конкурирующие гипотезы

Альтернативная гипотеза имеет вид:

$$H_i : X \notin N_d. \quad (6.3)$$

В данной работе при исследовании мощности в качестве конкурирующих рассматриваются гипотезы, предполагающие принадлежность выборки многомерному распределению, построенному на базе одномерных законов, отличных от нормального.

Соответствующее многомерное распределение получается из одномерного с помощью специального преобразования, описанного в п. 6.2.

Определим 6 конкурирующих гипотез следующим образом.

H_1 : многомерное распределение, основанное на одномерном распределении Коши;

H_2 : многомерное распределение, основанное на одномерном распределении Лапласа;

H_3 : многомерное распределение, основанное на одномерном логистическом распределении;

H_4 : многомерное распределение, основанное на одномерном двустороннем экспоненциальном распределении (его еще называют *обобщенное нормальное распределение*) с параметром формы 4;

H_5 : многомерное распределение, основанное на одномерном равномерном распределении;

H_6 : многомерное распределение, основанное на одномерном распределении наибольшего значения.

На рисунке 6.27 представлены функции распределения, соответствующие проверяемой гипотезе H_0 и конкурирующим гипотезам H_1, \dots, H_6 в одномерном случае. Видно, что конкурирующие распределения отличаются от нормального тяжестью хвостов (в случае гипотез H_1 и H_2 – имеют более тяжелые хвосты, при H_4 – более легкие хвосты, при H_5 – хвосты отсутствуют) и формой (у закона, соответствующего гипотезе H_3 , коэффициент эксцесса равен $6/5$; у соответствующего гипотезе H_6 –коэффициент эксцесса $12/5$, а коэффициент асимметрии примерно 1,14).

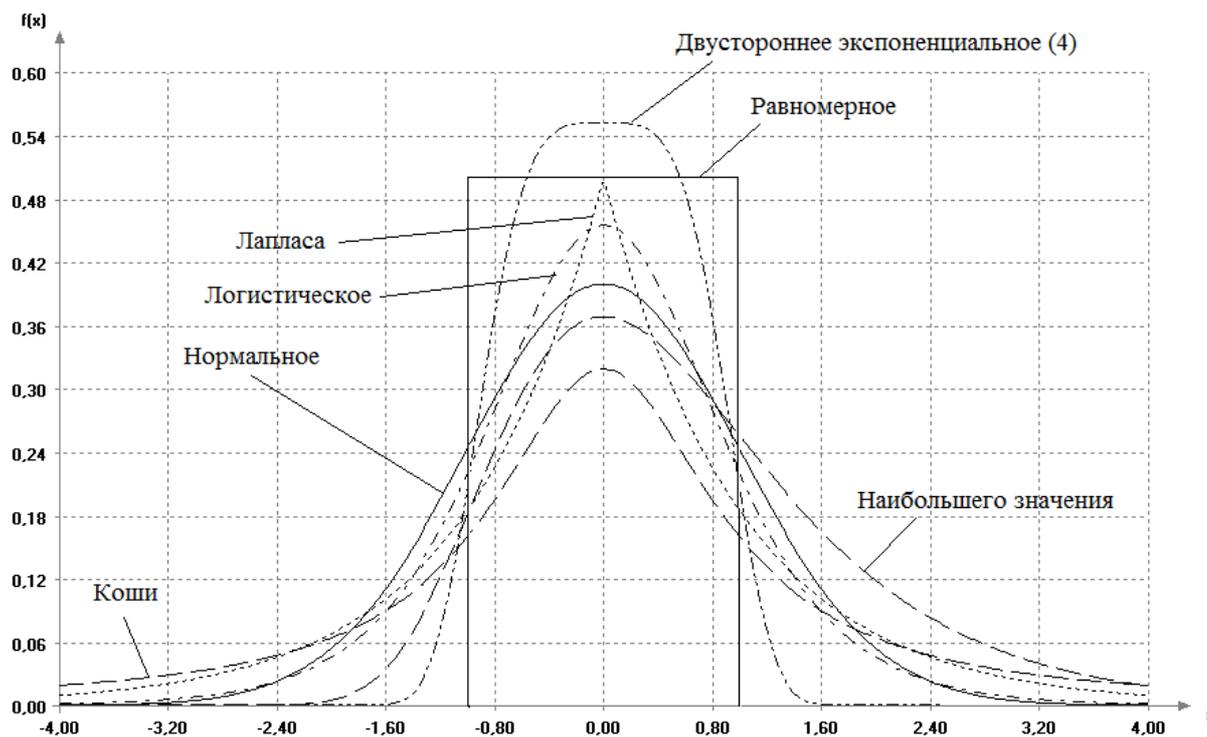


Рисунок 6.27 – Одномерные функции распределения конкурирующих гипотез $H_0 \div H_6$

6.5.2 Зависимость мощности критериев многомерной нормальности от размерности выборки

Результаты исследования мощности приведены в таблицах 6.5-6.9 для размерностей d от 1 до 5. Полужирным шрифтом выделены максимальные значения мощности разных критериев против одной альтернативы. В последней колонке приведено минимальное значение мощности одного критерия против разных альтернатив.

Оценки мощности находились методом Монте-Карло при числе повторений в 1 000 000 раз.

Анализ результатов исследования мощности показывает, что размерность d влияет на мощность критериев многомерной нормальности, которая, как правило, растет с увеличением размерности, а затем уменьшается, как показано на рисунках 6.28-6.29 для объемов выборки $n = 20, 50, 100$ и 200 для мощности

относительно конкурирующей гипотезы H_3 (логистическое распределение) и вероятности ошибки первого рода 0,05.

Хотя мощность критериев зависит от размерности выборки, однако набор критериев, обладающих наибольшей мощностью (из рассматриваемых критериев), практически не меняется при изменении размерности d . Исключением является одномерный случай ($d = 1$), когда совпадают мощности критериев асимметрии Мардия, Мори и Шриваставы, а также совпадают мощности критериев эксцесса Мардия, Козиола и Шриваставы.

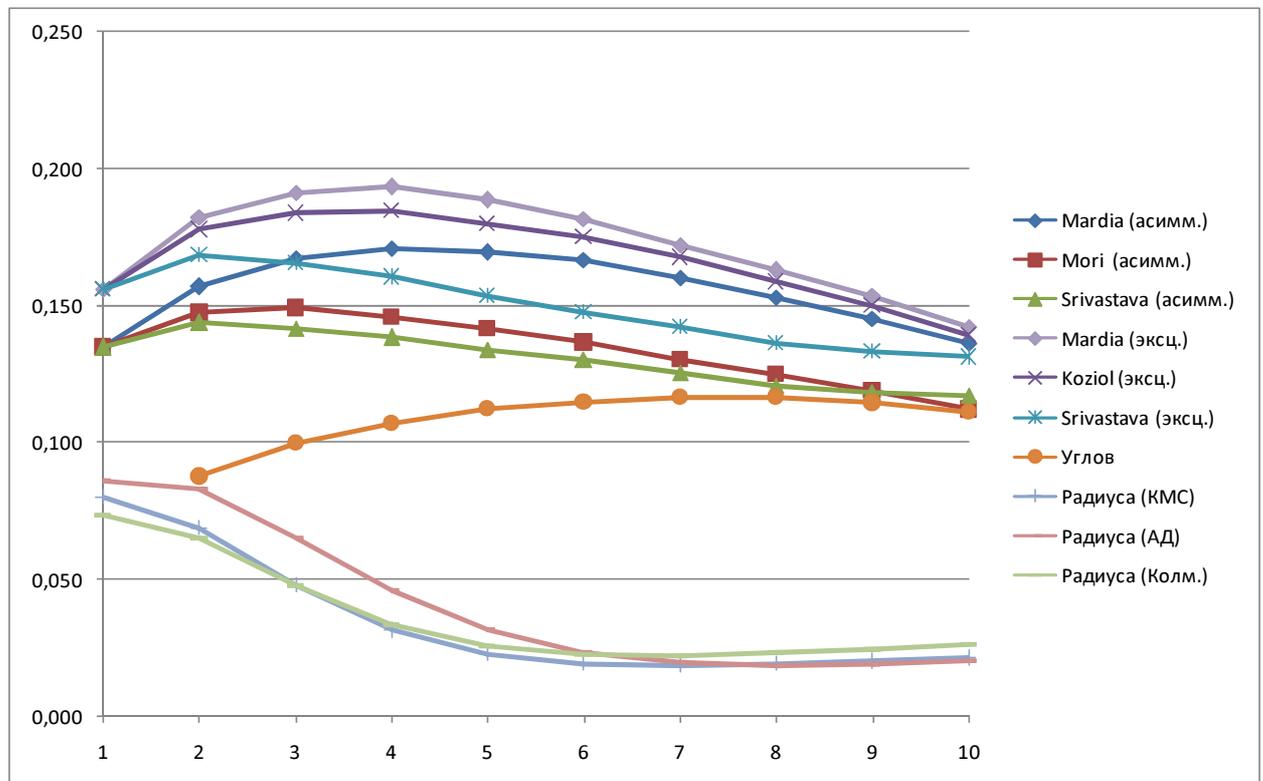


Рисунок 6.28 – Зависимость мощности от размерности d при альтернативе H_3 (логистическое распределение) и объеме $n = 20$

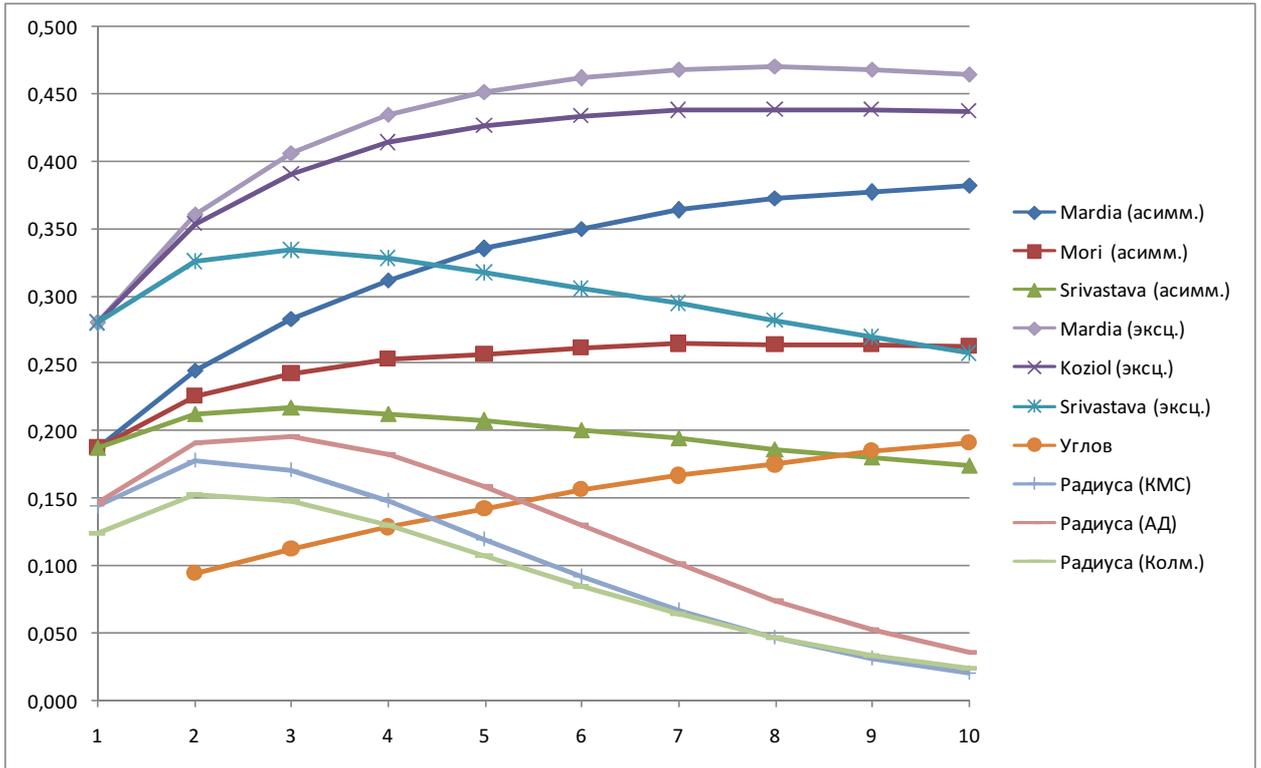


Рисунок 6.29 – Зависимость мощности от размерности d при альтернативе H_3 (логистическое распределение) и объеме $n = 50$

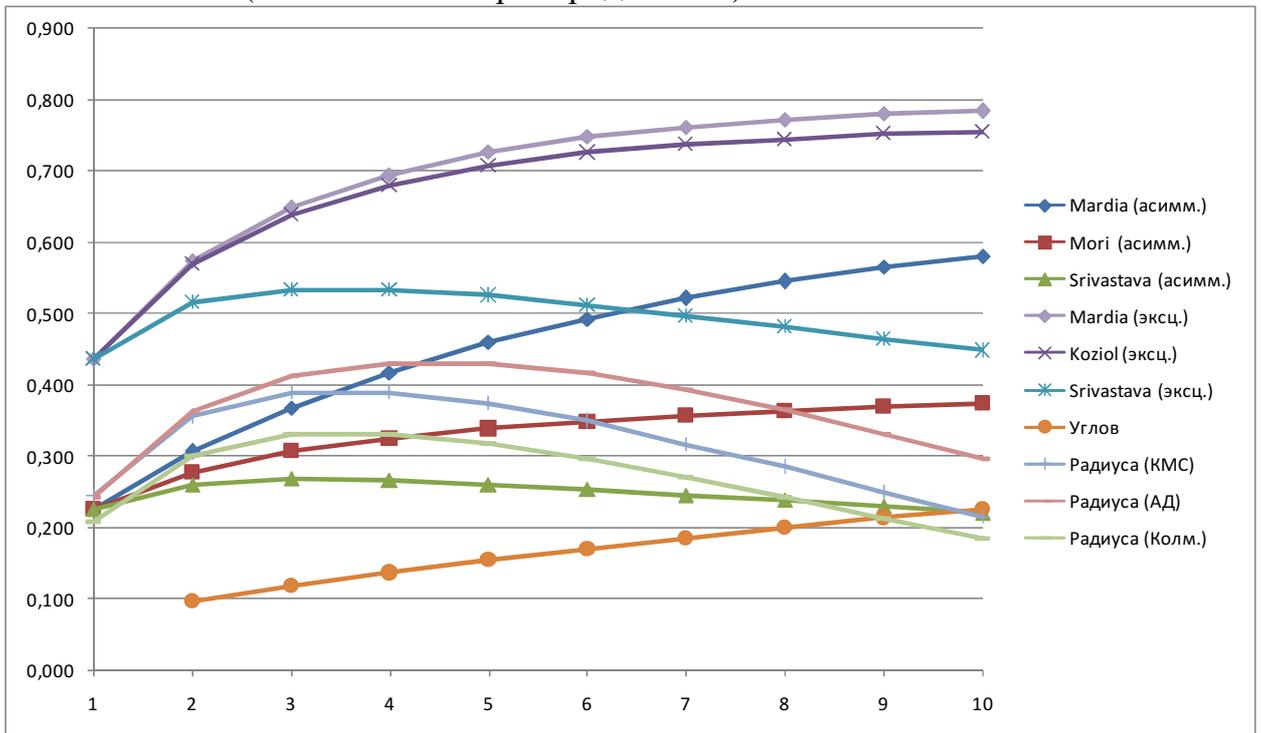


Рисунок 6.30 – Зависимость мощности от размерности d при альтернативе H_3 (логистическое распределение) и объеме $n = 100$

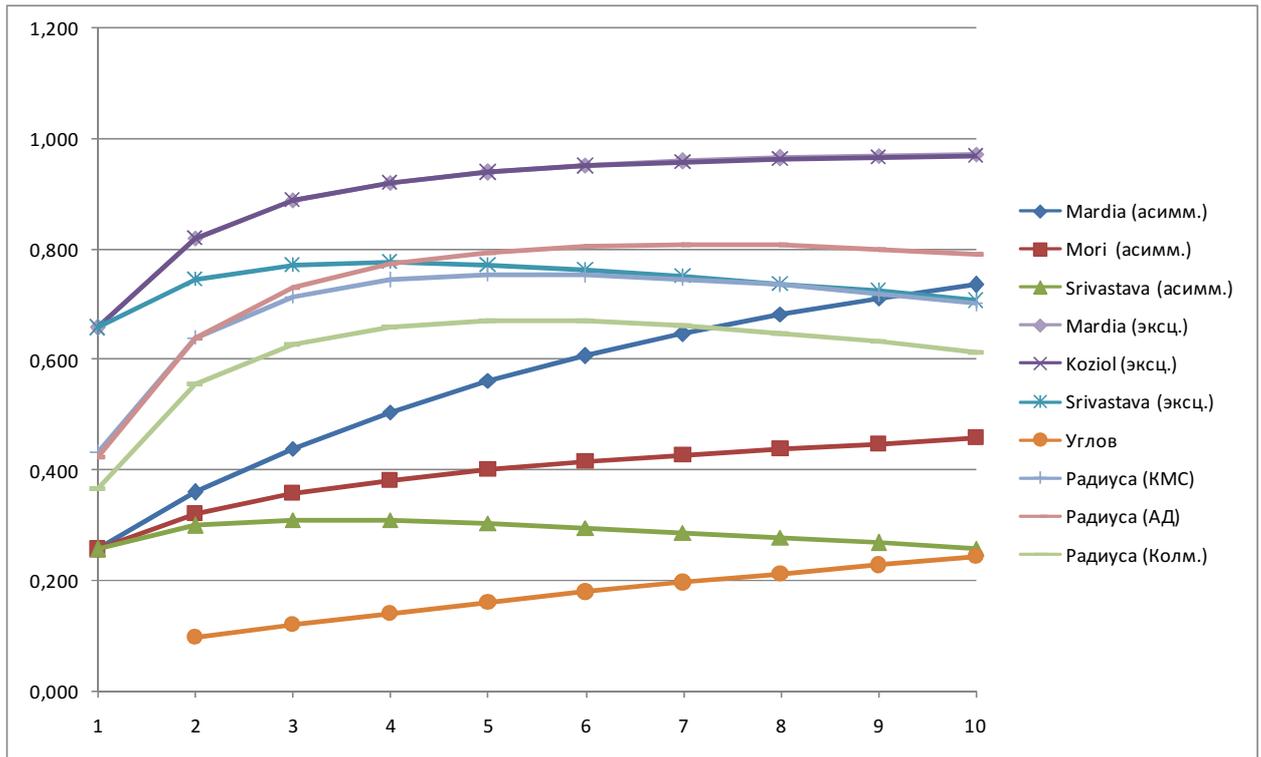


Рисунок 6.31 – Зависимость мощности от размерности d при альтернативе H_3 (логистическое распределение) и объеме $n = 200$

6.5.3 Анализ мощности инвариантных критериев многомерной нормальности

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-Дарлингга.

Наибольшей мощностью относительно конкурирующей гипотезы H_6 (распределение наибольшего значения) обладают критерии асимметрии Мардия и Мори. Немного меньшую мощность имеет критерий асимметрии Шриваставы.

Рассмотрим достоинства и недостатки каждой группы критериев. 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 оказались смещенными и несостоятельными против альтернативы с легкими хвостами.

Критерий “углов” 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 выборки.

Какой же из критериев лучше использовать на практике? С точки зрения робастности, используя правило Вальда, оптимальным является использование критерия “радиуса” Андерсона-Дарлинга. Для всех рассмотренных конкурирующих гипотез этот критерий показал наибольшую мощность в наихудших условиях.

Таблица 6.5 – Мощность критериев многомерной нормальности, $d = 1$

Критерии	Альтернативы							Ми- ни- мум
	Нор- маль- ное	Ко- ши	Лап- ласа	Лог- исти- чес- кое	Дв. эксп. (4)	Рав- но- мер- ное	Наи- боль- шего зна- чения	
$n = 20$								
Mardia (асимм.)	0,050	0,773	0,257	0,135	0,010	0,006	0,323	0,006
Mori (асимм.)	0,050	0,773	0,257	0,135	0,010	0,006	0,323	0,006
Srivastava (асимм.)	0,050	0,773	0,257	0,135	0,010	0,006	0,323	0,006
Mardia (эксц.)	0,050	0,876	0,328	0,156	0,004	0,001	0,217	0,001
Koziol (эксц.)	0,050	0,876	0,328	0,156	0,004	0,001	0,217	0,001
Srivastava (эксц.)	0,050	0,876	0,328	0,156	0,004	0,001	0,217	0,001
Углов	0,057	0,691	0,166	0,080	0,045	0,046	0,191	0,045
Радиуса (КМС)	0,050	0,841	0,225	0,080	0,118	0,259	0,085	0,080
Радиуса (АД)	0,050	0,840	0,239	0,086	0,098	0,224	0,087	0,086
Радиуса (Колм.)	0,050	0,820	0,198	0,074	0,105	0,213	0,077	0,074
$n = 50$								
Mardia (асимм.)	0,050	0,907	0,353	0,188	0,004	0,002	0,702	0,002
Mori (асимм.)	0,050	0,907	0,353	0,188	0,004	0,002	0,702	0,002
Srivastava (асимм.)	0,050	0,907	0,353	0,188	0,004	0,002	0,702	0,002
Mardia (эксц.)	0,050	0,997	0,610	0,280	0,000	0,000	0,403	0,000
Koziol (эксц.)	0,050	0,997	0,610	0,280	0,000	0,000	0,403	0,000
Srivastava (эксц.)	0,050	0,997	0,610	0,280	0,000	0,000	0,403	0,000
Углов	0,097	0,843	0,257	0,128	0,082	0,083	0,479	0,082
Радиуса (КМС)	0,050	0,997	0,571	0,144	0,264	0,657	0,134	0,134
Радиуса (АД)	0,050	0,997	0,575	0,147	0,233	0,641	0,142	0,142
Радиуса (Колм.)	0,050	0,996	0,508	0,125	0,221	0,543	0,114	0,114
$n = 100$								
Mardia (асимм.)	0,050	0,954	0,412	0,227	0,003	0,001	0,948	0,001
Mori (асимм.)	0,050	0,954	0,412	0,227	0,003	0,001	0,948	0,001
Srivastava (асимм.)	0,050	0,954	0,412	0,227	0,003	0,001	0,948	0,001
Mardia (эксц.)	0,050	1,000	0,852	0,437	0,000	0,000	0,605	0,000
Koziol (эксц.)	0,050	1,000	0,852	0,437	0,000	0,000	0,605	0,000
Srivastava (эксц.)	0,050	1,000	0,852	0,437	0,000	0,000	0,605	0,000
Углов	0,067	0,882	0,226	0,094	0,055	0,055	0,674	0,055

Продолжение таблицы 6.5

Критерии	Альтернативы							Ми- ни- мум
	Нор- маль- ное	Ко- ши	Лап- ласа	Лог- исти- чес- кое	Дв. эксп. (4)	Рав- но- мер- ное	Наи- боль- шего зна- чения	
Радиуса (КМС)	0,050	1,000	0,872	0,246	0,510	0,953	0,201	0,201
Радиуса (АД)	0,050	1,000	0,872	0,244	0,476	0,960	0,223	0,223
Радиуса (Колм.)	0,050	1,000	0,818	0,208	0,424	0,889	0,169	0,169

Таблица 6.6 – Мощность критериев многомерной нормальности, $d = 2$

Критерии	Альтернативы							Ми- ни- мум
	Нор- маль- ное	Ко- ши	Лап- ласа	Лог- исти- чес- кое	Дв. эксп. (4)	Рав- но- мер- ное	Наи- боль- шего зна- чения	
$n = 20$								
Mardia (асимм.)	0,050	0,928	0,331	0,157	0,008	0,003	0,357	0,003
Mori (асимм.)	0,050	0,914	0,303	0,147	0,011	0,005	0,366	0,005
Srivastava (асимм.)	0,050	0,916	0,294	0,144	0,014	0,007	0,347	0,007
Mardia (эксц.)	0,050	0,971	0,419	0,182	0,004	0,001	0,262	0,001
Koziol (эксц.)	0,050	0,969	0,400	0,178	0,005	0,001	0,268	0,001
Srivastava (эксц.)	0,050	0,968	0,384	0,168	0,007	0,002	0,245	0,002
Углов	0,050	0,877	0,209	0,088	0,029	0,024	0,237	0,024
Радиуса (КМС)	0,050	0,940	0,225	0,069	0,175	0,405	0,087	0,069
Радиуса (АД)	0,050	0,949	0,274	0,083	0,151	0,359	0,091	0,083
Радиуса (Колм.)	0,050	0,923	0,198	0,065	0,151	0,338	0,080	0,065
$n = 50$								
Mardia (асимм.)	0,050	0,990	0,496	0,245	0,003	0,000	0,813	0,000
Mori (асимм.)	0,050	0,986	0,449	0,226	0,004	0,001	0,802	0,001
Srivastava (асимм.)	0,050	0,987	0,428	0,213	0,006	0,002	0,730	0,002
Mardia (эксц.)	0,050	1,000	0,769	0,361	0,000	0,000	0,523	0,000
Koziol (эксц.)	0,050	1,000	0,757	0,354	0,000	0,000	0,527	0,000
Srivastava (эксц.)	0,050	1,000	0,713	0,326	0,001	0,000	0,483	0,000
Углов	0,050	0,962	0,253	0,095	0,028	0,023	0,569	0,023

Продолжение таблицы 6.6

Критерии	Альтернативы							Ми- ни- мум
	Нор- маль- ное	Ко- ши	Лап- ласа	Лог- исти- чес- кое	Дв. эксп. (4)	Рав- но- мер- ное	Наи- боль- шего зна- чения	
Радиуса (КМС)	0,050	1,000	0,677	0,178	0,415	0,880	0,219	0,178
Радиуса (АД)	0,050	1,000	0,716	0,191	0,382	0,858	0,223	0,191
Радиуса (Колм.)	0,050	1,000	0,607	0,154	0,341	0,789	0,187	0,154
$n = 100$								
Mardia (асимм.)	0,050	0,998	0,591	0,307	0,002	0,000	0,990	0,000
Mori (асимм.)	0,050	0,996	0,531	0,278	0,002	0,000	0,985	0,000
Srivastava (асимм.)	0,050	0,997	0,508	0,261	0,004	0,001	0,932	0,001
Mardia (эксц.)	0,050	1,000	0,958	0,575	0,000	0,000	0,765	0,000
Koziol (эксц.)	0,050	1,000	0,957	0,570	0,000	0,000	0,769	0,000
Srivastava (эксц.)	0,050	1,000	0,914	0,516	0,000	0,000	0,715	0,000
Углов	0,050	0,983	0,271	0,097	0,027	0,022	0,877	0,022
Радиуса (КМС)	0,050	1,000	0,948	0,356	0,736	0,998	0,423	0,356
Радиуса (АД)	0,050	1,000	0,959	0,363	0,712	0,997	0,428	0,363
Радиуса (Колм.)	0,050	1,000	0,914	0,301	0,628	0,987	0,363	0,301

Таблица 6.7 – Мощность критериев многомерной нормальности, $d = 3$

Критерии	Альтернативы							Ми- ни- мум
	Нор- маль- ное	Ко- ши	Лап- ласа	Лог- исти- чес- кое	Дв. эксп. (4)	Рав- но- мер- ное	Наи- боль- шего зна- чения	
$n = 20$								
Mardia (асимм.)	0,050	0,975	0,377	0,167	0,008	0,003	0,367	0,003
Mori (асимм.)	0,050	0,960	0,321	0,149	0,012	0,006	0,374	0,006
Srivastava (асимм.)	0,050	0,962	0,300	0,142	0,017	0,011	0,341	0,011
Mardia (эксц.)	0,050	0,991	0,465	0,191	0,004	0,001	0,282	0,001
Koziol (эксц.)	0,050	0,990	0,433	0,184	0,006	0,002	0,291	0,002
Srivastava (эксц.)	0,050	0,988	0,391	0,165	0,010	0,004	0,251	0,004
Углов	0,050	0,945	0,244	0,100	0,022	0,015	0,262	0,015

Продолжение таблицы 6.7

Критерии	Альтернативы							Ми- ни- мум
	Нор- маль- ное	Ко- ши	Лап- ласа	Лог- исти- чес- кое	Дв. эксп. (4)	Рав- но- мер- ное	Наи- боль- шего зна- чения	
Радиуса (КМС)	0,050	0,956	0,165	0,048	0,189	0,413	0,067	0,048
Радиуса (АД)	0,050	0,972	0,239	0,065	0,174	0,389	0,075	0,065
Радиуса (Колм.)	0,050	0,941	0,147	0,048	0,162	0,344	0,065	0,048
<i>n</i> = 50								
Mardia (асимм.)	0,050	0,999	0,596	0,283	0,002	0,000	0,861	0,000
Mori (асимм.)	0,050	0,997	0,502	0,243	0,004	0,001	0,838	0,001
Srivastava (асимм.)	0,050	0,998	0,457	0,217	0,008	0,003	0,728	0,003
Mardia (эксц.)	0,050	1,000	0,841	0,407	0,000	0,000	0,593	0,000
Koziol (эксц.)	0,050	1,000	0,821	0,391	0,000	0,000	0,597	0,000
Srivastava (эксц.)	0,050	1,000	0,743	0,334	0,001	0,000	0,516	0,000
Углов	0,050	0,992	0,307	0,112	0,019	0,012	0,669	0,012
Радиуса (КМС)	0,050	1,000	0,679	0,171	0,466	0,914	0,264	0,171
Радиуса (АД)	0,050	1,000	0,748	0,196	0,448	0,904	0,268	0,196
Радиуса (Колм.)	0,050	1,000	0,609	0,148	0,381	0,831	0,229	0,148
<i>n</i> = 100								
Mardia (асимм.)	0,050	1,000	0,715	0,368	0,001	0,000	0,997	0,000
Mori (асимм.)	0,050	1,000	0,600	0,307	0,002	0,000	0,993	0,000
Srivastava (асимм.)	0,050	1,000	0,548	0,270	0,005	0,001	0,935	0,001
Mardia (эксц.)	0,050	1,000	0,983	0,649	0,000	0,000	0,844	0,000
Koziol (эксц.)	0,050	1,000	0,982	0,639	0,000	0,000	0,848	0,000
Srivastava (эксц.)	0,050	1,000	0,935	0,534	0,000	0,000	0,759	0,000
Углов	0,050	0,998	0,331	0,119	0,018	0,011	0,948	0,011
Радиуса (КМС)	0,050	1,000	0,961	0,390	0,795	0,999	0,560	0,390
Радиуса (АД)	0,050	1,000	0,976	0,414	0,789	0,999	0,556	0,414
Радиуса (Колм.)	0,050	1,000	0,932	0,331	0,685	0,994	0,491	0,331

Таблица 6.8 – Мощность критериев многомерной нормальности, $d = 4$

Критерии	Альтернативы							Ми- ни- мум
	Нор- маль- ное	Ко- ши	Лап- ласа	Лог- исти- чес- кое	Дв. эксп. (4)	Рав- но- мер- ное	Наи- боль- шего зна- чения	
$n = 20$								
Mardia (асимм.)	0,050	0,990	0,404	0,171	0,008	0,003	0,361	0,003
Mori (асимм.)	0,050	0,979	0,326	0,146	0,013	0,006	0,368	0,006
Srivastava (асимм.)	0,050	0,980	0,296	0,138	0,020	0,014	0,331	0,014
Mardia (эксц.)	0,050	0,997	0,487	0,193	0,004	0,001	0,286	0,001
Koziol (эксц.)	0,050	0,996	0,449	0,185	0,007	0,002	0,301	0,002
Srivastava (эксц.)	0,050	0,995	0,384	0,160	0,013	0,006	0,251	0,006
Углов	0,050	0,972	0,267	0,107	0,018	0,010	0,267	0,010
Радиуса (КМС)	0,050	0,953	0,101	0,032	0,185	0,380	0,044	0,032
Радиуса (АД)	0,050	0,977	0,179	0,046	0,181	0,379	0,054	0,046
Радиуса (Колм.)	0,050	0,937	0,091	0,034	0,158	0,315	0,045	0,034
$n = 50$								
Mardia (асимм.)	0,050	1,000	0,664	0,312	0,002	0,000	0,881	0,000
Mori (асимм.)	0,050	0,999	0,532	0,253	0,003	0,001	0,851	0,001
Srivastava (асимм.)	0,050	1,000	0,461	0,213	0,009	0,004	0,711	0,004
Mardia (эксц.)	0,050	1,000	0,876	0,435	0,000	0,000	0,638	0,000
Koziol (эксц.)	0,050	1,000	0,853	0,414	0,000	0,000	0,641	0,000
Srivastava (эксц.)	0,050	1,000	0,749	0,328	0,002	0,000	0,526	0,000
Углов	0,050	0,998	0,349	0,129	0,014	0,007	0,721	0,007
Радиуса (КМС)	0,050	1,000	0,647	0,148	0,480	0,915	0,275	0,148
Радиуса (АД)	0,050	1,000	0,744	0,183	0,480	0,916	0,284	0,183
Радиуса (Колм.)	0,050	1,000	0,578	0,131	0,392	0,830	0,241	0,131
$n = 100$								
Mardia (асимм.)	0,050	1,000	0,799	0,418	0,001	0,000	0,999	0,000
Mori (асимм.)	0,050	1,000	0,644	0,325	0,002	0,000	0,995	0,000
Srivastava (асимм.)	0,050	1,000	0,564	0,267	0,006	0,002	0,931	0,002
Mardia (эксц.)	0,050	1,000	0,991	0,695	0,000	0,000	0,887	0,000
Koziol (эксц.)	0,050	1,000	0,991	0,680	0,000	0,000	0,890	0,000
Srivastava (эксц.)	0,050	1,000	0,942	0,534	0,000	0,000	0,779	0,000
Углов	0,050	1,000	0,380	0,137	0,013	0,006	0,971	0,006

Продолжение таблицы 6.8

Критерии	Альтернативы							Ми- ни- мум
	Нор- маль- ное	Ко- ши	Лап- ласа	Лог- исти- чес- кое	Дв. эксп. (4)	Рав- но- мер- ное	Наи- боль- шего зна- чения	
Радиуса (КМС)	0,050	1,000	0,963	0,390	0,813	0,999	0,632	0,390
Радиуса (АД)	0,050	1,000	0,981	0,431	0,821	0,999	0,630	0,431
Радиуса (Колм.)	0,050	1,000	0,934	0,330	0,702	0,995	0,561	0,330

Таблица 6.9 – Мощность критериев многомерной нормальности, $d = 5$

Критерии	Альтернативы							Ми- ни- мум
	Нор- маль- ное	Ко- ши	Лап- ласа	Лог- исти- чес- кое	Дв. эксп. (4)	Рав- но- мер- ное	Наи- боль- шего зна- чения	
$n = 20$								
Mardia (асимм.)	0,050	0,996	0,417	0,169	0,008	0,003	0,347	0,003
Mori (асимм.)	0,050	0,988	0,322	0,142	0,014	0,007	0,353	0,007
Srivastava (асимм.)	0,050	0,989	0,288	0,134	0,023	0,017	0,319	0,017
Mardia (эксц.)	0,050	0,998	0,488	0,188	0,005	0,001	0,280	0,001
Koziol (эксц.)	0,050	0,998	0,450	0,180	0,008	0,002	0,298	0,002
Srivastava (эксц.)	0,050	0,998	0,370	0,153	0,015	0,008	0,247	0,008
Углов	0,050	0,984	0,282	0,112	0,016	0,008	0,263	0,008
Радиуса (КМС)	0,050	0,935	0,052	0,023	0,174	0,335	0,027	0,023
Радиуса (АД)	0,050	0,974	0,116	0,032	0,176	0,346	0,034	0,032
Радиуса (Колм.)	0,050	0,916	0,049	0,026	0,148	0,274	0,029	0,026
$n = 50$								
Mardia (асимм.)	0,050	1,000	0,719	0,335	0,001	0,000	0,893	0,000
Mori (асимм.)	0,050	1,000	0,556	0,257	0,004	0,000	0,855	0,000
Srivastava (асимм.)	0,050	1,000	0,461	0,207	0,011	0,005	0,691	0,005
Mardia (эксц.)	0,050	1,000	0,897	0,452	0,000	0,000	0,668	0,000
Koziol (эксц.)	0,050	1,000	0,871	0,426	0,001	0,000	0,670	0,000
Srivastava (эксц.)	0,050	1,000	0,744	0,317	0,003	0,000	0,527	0,000
Углов	0,050	0,999	0,388	0,143	0,011	0,004	0,745	0,004

Продолжение таблицы 6.9

Критерии	Альтернативы							Ми- ни- мум
	Нор- маль- ное	Ко- ши	Лап- ласа	Лог- исти- чес- кое	Дв. эксп. (4)	Рав- но- мер- ное	Наи- боль- шего зна- чения	
Радиуса (КМС)	0,050	1,000	0,597	0,119	0,480	0,906	0,261	0,119
Радиуса (АД)	0,050	1,000	0,719	0,158	0,493	0,916	0,277	0,158
Радиуса (Колм.)	0,050	1,000	0,528	0,108	0,387	0,813	0,231	0,108
$n = 100$								
Mardia (асимм.)	0,050	1,000	0,856	0,460	0,001	0,000	0,999	0,000
Mori (асимм.)	0,050	1,000	0,677	0,339	0,002	0,000	0,996	0,000
Srivastava (асимм.)	0,050	1,000	0,567	0,261	0,007	0,002	0,922	0,002
Mardia (эксц.)	0,050	1,000	0,995	0,727	0,000	0,000	0,914	0,000
Koziol (эксц.)	0,050	1,000	0,994	0,708	0,000	0,000	0,916	0,000
Srivastava (эксц.)	0,050	1,000	0,945	0,526	0,000	0,000	0,788	0,000
Углов	0,050	1,000	0,426	0,155	0,009	0,003	0,981	0,003
Радиуса (КМС)	0,050	1,000	0,961	0,374	0,818	0,999	0,668	0,374
Радиуса (АД)	0,050	1,000	0,982	0,429	0,837	1,000	0,671	0,429
Радиуса (Колм.)	0,050	1,000	0,930	0,318	0,706	0,994	0,597	0,318

6.6 Выводы

В данной главе рассмотрены инвариантные критерии многомерной нормальности: асимметрии (Мардия, Мори, Шриваставы), эксцесса (Мардия, Козиола, Шриваставы), критерии, основанные на полярной декомпозиции (критерий “углов” и критерии “радиуса”).

Проведено исследование скорости сходимости распределений статистик критериев к предельным законам распределения.

Впервые построены модели, аппроксимирующие законы распределения статистик критериев “радиуса” Крамера-Мизеса-Смирнова, Андерсона-Дарлинга и Колмогорова [132].

Проведен сравнительный анализ мощности критериев относительно шести конкурирующих гипотез, представляющих собой многомерные законы распределения, построенные на основании соответствующих одномерных законов (Коши, Лапласа, логистического, обобщенного нормального с параметром формы 4, равномерного и наибольшего значения) в соответствии с рассмотренной процедурой преобразования (моделирования) [125, 132].

Проведено исследование зависимости мощности критериев многомерной нормальности от размерности случайных величин. Показано, что мощность критериев, чаще всего, сначала увеличивается при увеличении размерности d , а затем уменьшается.

Поскольку относительно разных конкурирующих гипотез более мощными оказались критерии разных типов, то на основании правила Вальда принятия решения в условиях неопределенности сделана рекомендация о предпочтительном использовании критерия “радиуса” Андерсона-Дарлинга для проверки гипотезы о многомерной нормальности.

ГЛАВА 7. ОПТИМАЛЬНОЕ ПЛАНИРОВАНИЕ ПРОВЕРКИ ГИПОТЕЗЫ ОДНОРОДНОСТИ ПРИ ПРОВЕДЕНИИ ДВУХЭТАПНОГО ПОЛНОГЕНОМНОГО АНАЛИЗА АССОЦИАЦИЙ

Двухэтапный план эксперимента с использованием контрольной выборки и выборки случаев был впервые использован для эпидемиологических исследований Уайтом [277] и был развит Бреслоу [168]. Элстон предложил минимизировать затраты в двухэтапном планировании [175] для анализа сцепленного наследования. Позднее этот подход был перенесен для анализа ассоциаций [256, 257, 258].

Оптимальное планирование заключается в определении необходимых объемов выборок на первом и втором этапах, и выборе критических значений таким образом, чтобы общая стоимость эксперимента была минимальной при заданных значениях вероятностей ошибок первого и второго рода [254; 262; 276; 201, 202; 233; 265; 236].

7.1 Полногеномный анализ ассоциаций

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

В GWAS обычно сравнивают ДНК двух групп участников: первая группа с заболеваниями (группа случаев, cases), а вторая – без (контрольная группа, controls). У каждого человека производится сканирование ДНК (*генотипирование*), из которых формируются массивы SNP. Если один тип варианта (одна *аллель*) встречается чаще в группе случаев, то говорят, что SNP

ассоциирован с болезнью. Такими SNP отмечают область генома человека, повышающую риск заболевания.

Два произвольно взятых человека имеют отличия в геноме примерно в 3 миллионах нуклеотидных оснований, в то время как общее количество нуклеотидных оснований в геноме человека 3 миллиарда. В полногеномных исследованиях анализируют от 100 тысяч до миллиона SNP.

Первый успешный полногеномный анализ ассоциаций был проведен в 2005 году при выявлении генетических причин *макулодистрофии* (macular degeneration) – заболевания, поражающего сетчатку глаза [196].

На основании баз данных, в которых собираются результаты GWAS [190] можно сказать, что к 2011 году было проведено более 1200 полногеномных экспериментов, и было найдено почти 4000 SNP, ассоциированных с 200 заболеваниями.

В *полногеномном анализе ассоциаций* исследователи часто рассматривают три типа альтернативных гипотез: при рецессивной, аддитивной и доминантной модели наследования, с различными заданными частотами аллелей, относительным генетическим риском (GRR, genetic relative risk) или отношением шансов (OR, odds ratio).

По характеру проявления признака в гетерозиготном организме наследование разделяют на доминантное и рецессивное. При доминантном наследовании заболевание проявляется, если хотя бы одна из гомологичных хромосом несет патологический аллель, при рецессивном — только в случае, когда обе гомологичные хромосомы несут патологический аллель [4]. Аддитивное наследование является комбинацией доминантного и рецессивного наследования.

7.2 Одноэтапный эксперимент по выявлению ассоциации

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 (таблица 7.1).

Таблица 7.1 – Результаты наблюдений за группой больных (группа случаев) и группой здоровых (контрольная группа)

Генотип	AA	Aa	aa	Всего
Выборка случаев (<i>cases</i>)	r_0	r_1	r_2	R
Контрольная выборка (<i>controls</i>)	s_0	s_1	s_2	S
Всего	n_0	n_1	n_2	N

7.3 Основная и конкурирующие гипотезы

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. Далее в таблицах рецессивная модель наследования будет обозначаться как « R », доминантная – « D », а аддитивная – « A ».

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

Конкретная конкурирующая гипотеза задается MOI , частотой p_a , и отношением шансов (OR)

$$\psi_i = \frac{f_i(1-f_0)}{f_0(1-f_i)} = \frac{p_i \cdot q_0}{p_0 \cdot q_i}, i=1,2, \quad (7.1)$$

или относительным генетическим риском (GRR)

$$\gamma_i = \frac{f_i}{f_0}, i=1,2. \quad (7.2)$$

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, п.2.3):

$$D_{KL} = \sum_{i=0}^2 p_i \ln \frac{p_i}{q_i}. \quad (7.3)$$

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

$$D_{KL} = \ln \frac{1-K}{\lambda} + \frac{2p_a(1-p_a)\psi_1}{\lambda + K\psi_1} \ln \psi_1 + \frac{p_a^2 \cdot \psi_2}{\lambda + K\psi_2} \ln \psi_2, \quad (7.4)$$

где λ может быть найдено из уравнения

$$\frac{(1-p_a)^2}{\lambda + K} + \frac{2p_a(1-p_a)\psi_1}{\lambda + K\psi_1} + \frac{p_a^2 \cdot \psi_2}{\lambda + K\psi_2} = 1. \quad (7.5)$$

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

$$D_{KL} = \tau \left[(1-p_a)^2 \ln \frac{1-K}{\tau-K} + 2p_a(1-p_a)\gamma_1 \ln \frac{\gamma_1(1-K)}{\tau-\gamma_1 K} + p_a^2 \cdot \gamma_2 \ln \frac{\gamma_2(1-K)}{\tau-\gamma_2 K} \right], \quad (7.6)$$

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

Поскольку дивергенция Кульбака-Лейблера D_{KL} является асимметричной, а распределения в выборке случаев и контрольной выборке неизвестны, то предпочтительнее использовать симметричную дивергенцию Кульбака-Лейблера ρ_{KL}

$$\rho_{KL}(P, Q) = D_{KL}(P, Q) + D_{KL}(Q, P) = \left(\sum_{i=0}^2 p_i \ln \frac{p_i}{q_i} + \sum_{i=0}^2 q_i \ln \frac{q_i}{p_i} \right) = \sum_{i=0}^2 (p_i - q_i) \ln \frac{p_i}{q_i}. \quad (7.7)$$

Симметричная дивергенция Кульбака-Лейблера связана с p_a , K , γ_1 и γ_2 как

$$\rho_{KL} = \frac{(1 - p_a)^2 (1 - \tau) \ln \frac{1 - K}{\tau - K} + 2p_a (1 - p_a) (\gamma_1 - \tau) \ln \frac{\gamma_1 (1 - K)}{\tau - \gamma_1 K} + p_a^2 \cdot (\gamma_2 - \tau) \ln \frac{\gamma_2 (1 - K)}{\tau - \gamma_2 K}}{\tau (1 - K)}.$$

Симметричная дивергенция Кульбака-Лейблера ρ_{KL} был рассчитана для разных генетических моделей по формулам (7.7) и (7.4) при распространенности заболевания в популяции $K = 0,05$. Результаты приведены в таблице 7.2. Для нахождения параметра λ уравнение (7.5) решалось численно с помощью метода Ньютона.

Таблица 7.2 - Симметричная дивергенция Кульбака-Лейблера ρ_{KL} для разных генетических моделей при $K = 0,05$.

MOI	p_a	ψ_1	ψ_2	ρ_{KL}
R	0,1	1,00	1,25	0,00054
R	0,1	1,00	1,50	0,00195
R	0,1	1,00	1,75	0,00398
R	0,1	1,00	2,00	0,00648
R	0,3	1,00	1,25	0,00442
R	0,3	1,00	1,50	0,01560
R	0,3	1,00	1,75	0,03136
R	0,3	1,00	2,00	0,05035
R	0,5	1,00	1,25	0,00979
R	0,5	1,00	1,50	0,03348
R	0,5	1,00	1,75	0,06547
R	0,5	1,00	2,00	0,10250
R	0,9	1,00	1,25	0,00719

Продолжение таблицы 7.2

MOI	p_a	ψ_1	ψ_2	ρ_{KL}
<i>R</i>	0,9	1,00	1,50	0,02248
<i>R</i>	0,9	1,00	1,75	0,04084
<i>R</i>	0,9	1,00	2,00	0,06010
<i>A</i>	0,1	1,25	1,56	0,00970
<i>A</i>	0,1	1,50	2,25	0,03409
<i>A</i>	0,1	1,75	3,06	0,06831
<i>A</i>	0,1	2,00	4,00	0,10925
<i>A</i>	0,3	1,25	1,56	0,02169
<i>A</i>	0,3	1,50	2,25	0,07339
<i>A</i>	0,3	1,75	3,06	0,14208
<i>A</i>	0,3	2,00	4,00	0,22033
<i>A</i>	0,5	1,25	1,56	0,02479
<i>A</i>	0,5	1,50	2,25	0,08110
<i>A</i>	0,5	1,75	3,06	0,15266
<i>A</i>	0,5	2,00	4,00	0,23116
<i>A</i>	0,9	1,25	1,56	0,00826
<i>A</i>	0,9	1,50	2,25	0,02552
<i>A</i>	0,9	1,75	3,06	0,04594
<i>A</i>	0,9	2,00	4,00	0,06711
<i>D</i>	0,1	1,25	1,25	0,00814
<i>D</i>	0,1	1,50	1,50	0,02814
<i>D</i>	0,1	1,75	1,75	0,05558
<i>D</i>	0,1	2,00	2,00	0,08780
<i>D</i>	0,3	1,25	1,25	0,01237
<i>D</i>	0,3	1,50	1,50	0,04046
<i>D</i>	0,3	1,75	1,75	0,07618
<i>D</i>	0,3	2,00	2,00	0,11539
<i>D</i>	0,5	1,25	1,25	0,00886
<i>D</i>	0,5	1,50	1,50	0,02793
<i>D</i>	0,5	1,75	1,75	0,05108
<i>D</i>	0,5	2,00	2,00	0,07557
<i>D</i>	0,9	1,25	1,25	0,00045
<i>D</i>	0,9	1,50	1,50	0,00136
<i>D</i>	0,9	1,75	1,75	0,00244
<i>D</i>	0,9	2,00	2,00	0,00354

7.4 Критерии ассоциаций

Рассмотрим наиболее известные и применяемые критерии проверки гипотезы однородности при проведении полногеномного анализа ассоциаций, которые далее будут называться просто *критерии ассоциаций*.

7.4.1 Критерий Хи-квадрат

Классический критерий для проверки однородности двух выборок – это критерий χ^2 . Его статистика [39] имеет вид

$$T_{\chi^2} = n \left(\sum_{i=0}^2 \frac{r_i^2}{n_i r} + \sum_{i=0}^2 \frac{s_i^2}{n_i s} - 1 \right). \quad (7.8)$$

При верной основной гипотезе в пределе статистика T_{χ^2} подчиняется χ^2 распределению с двумя степенями свободы.

7.4.2 Критерий отношения правдоподобия с ограничением (CLRT)

Пусть

$$I_0 = r \ln \frac{r}{n} + s \ln \left(1 - \frac{r}{n} \right), \quad (7.9)$$

$$I_2(f_0, f_1, f_2) = \sum_{i=0}^2 (r_i \ln f_i + s_i \ln(1 - f_i)). \quad (7.10)$$

Тогда статистика критерия CLRT может быть вычислена как [275]

$$T_{CLRT} = \begin{cases} I_2 \left(\frac{r_0}{n_0}, \frac{r_1}{n_1}, \frac{r_2}{n_2} \right) - I_0, & \text{если } \frac{r_0}{n_0} \leq \frac{r_1}{n_1} \leq \frac{r_2}{n_2} \text{ и } \frac{r_0}{n_0} < \frac{r_2}{n_2}; \\ I_2 \left(\frac{r_0}{n_0}, \frac{r_1}{n_1}, \frac{r_2}{n_2} \right) - I_0, & \text{если } \frac{r_0}{n_0} \geq \frac{r_1}{n_1} \geq \frac{r_2}{n_2} \text{ и } \frac{r_0}{n_0} > \frac{r_2}{n_2}; \\ \max \left\{ I_2 \left(\frac{r_0 + r_1}{n_0 + n_1}, \frac{r_0 + r_1}{n_0 + n_1}, \frac{r_2}{n_2} \right), I_2 \left(\frac{r_0}{n_0}, \frac{r_1 + r_2}{n_1 + n_2}, \frac{r_1 + r_2}{n_1 + n_2} \right) \right\} - I_0, & \text{иначе.} \end{cases} \quad (7.11)$$

При верной основной гипотезе асимптотическим распределением статистик (7.11) является смесь распределений χ^2 [275].

7.4.3 Критерий тренда Кокрена-Армитеджа

Статистика критерия тренда Кокрена-Армитеджа (**САТТ**) [160, 161] может быть определена как [255]

$$T_{САТТ} = \frac{\sum_{i=0}^2 \omega_i (sr_i - rs_i)}{\left\{ rs \left(\sum_{i=0}^2 \omega_i^2 n_i - \frac{1}{n} \left(\sum_{i=0}^2 \omega_i n_i \right)^2 \right) \right\}^{1/2}}. \quad (7.12)$$

Для рецессивной модели наследования статистика T_r использует вектор коэффициентов $\bar{\omega}=(0,0,1)$, для аддитивной модели наследования статистика T_a использует вектор коэффициентов $\bar{\omega}=(0,1,2)$ и для доминантной модели наследования статистика T_d имеет вектор коэффициентов $\bar{\omega}=(0,1,1)$.

Эти статистики в пределе подчиняются стандартному нормальному закону [255] при верной основной гипотезе. Основная гипотеза отвергается при больших абсолютных значениях статистики $T_{САТТ}$.

Критерий тренда Кокрена-Армитеджа использует вектор коэффициентов, который является оптимальным с точки зрения максимизации асимптотической относительной эффективности по Питмену (2.29) для логистической регрессии [270]. Статистические свойства оптимального критерия для аддитивной модели были исследованы Слагером и Шейдом [266].

Известно, что критерий тренда Кокрена-Армитеджа имеет существенную потерю мощности, когда используются оптимальные коэффициенты для одной генетической модели, а данные подчиняются другой генетической модели [266].

Распределение статистики критерия тренда T_a имеет явно выраженные ступеньки в середине распределения (в области нуля), как показано на рисунке 7.1. С ростом объемов выборок эти ступеньки уменьшаются, и гипотеза о согласии со стандартным нормальным распределением не отвергается (см. таблицу 7.3). Так как при проверке гипотезы однородности наиболее важны именно хвосты закона распределения статистики, то на практике можно использовать стандартное нормальное распределение для вычисления достигаемого уровня значимости.

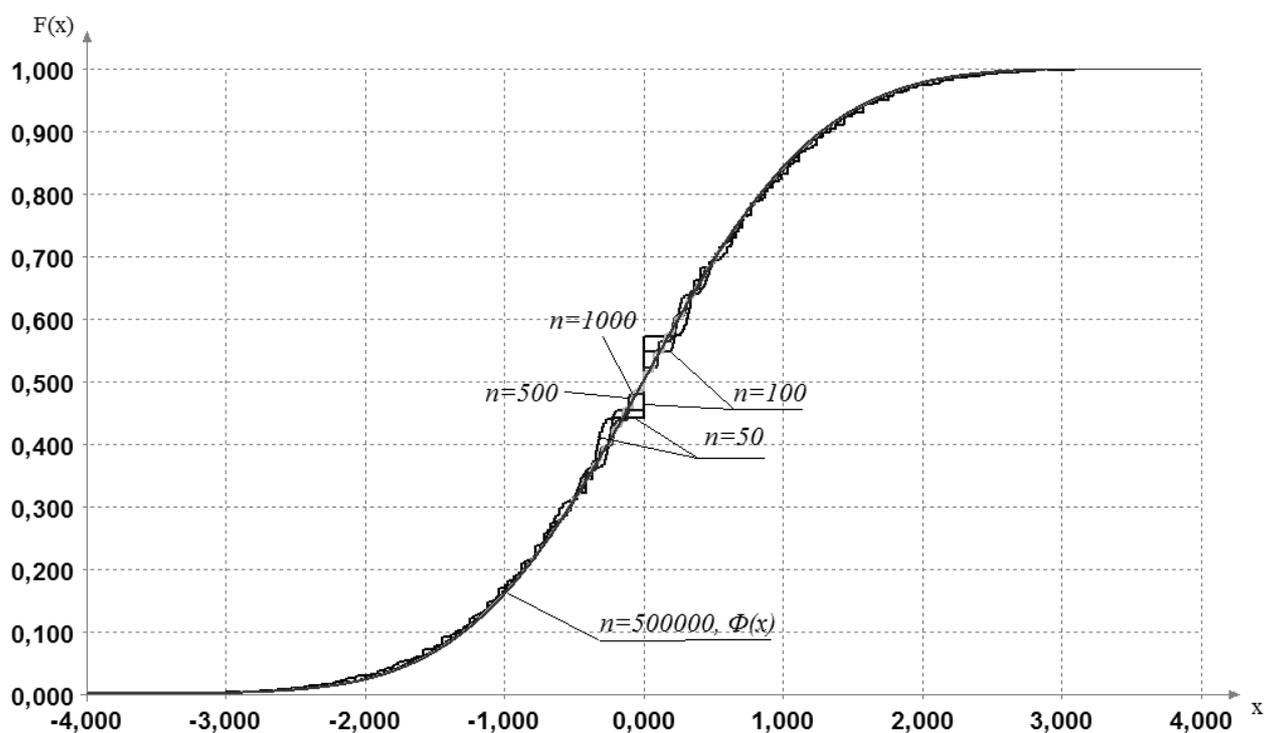


Рисунок 7.1 - Функция распределения статистики T_a для объемов 50, 100, 1000 и 500000 наблюдений ($p_a = 0,3$)

Таблица 7.3 - Достижимый уровень значимости при проверке согласия распределения статистики T_a со стандартным нормальным законом распределения (объем выборки статистик равен $N = 1700000$)

n	Критерий Рао-Робсона-Никулина	Критерий Колмогорова	Критерий Крамера-Мизеса-Смирнова	Критерий Андерсона-Дарлингга
500000	0.5365	0.02200	0.34465	0.412542
1000	1.36e-09	0	0	0

7.4.4 MERT критерий

Критерий является *максиминным эффективным робастным критерием* (MERT, *maximin efficiency robust*) для семейства всевозможных моделей, которым подчиняются наблюдения, если нет никакого другого критерия, который имеет большую минимальную относительную эффективность к асимптотически оптимальному критерию для каждой модели [179].

Статистика критерия MERT, предложенного Гаствиртом (Gastwirth) в 1985 г. имеет вид [179]:

$$T_{MERT} = \frac{T_r + T_d}{\sqrt{2(1 + \rho(T_r, T_d))}}, \quad (7.13)$$

где

$$\rho(T_r, T_d) = \frac{\sqrt{n_0 n_1}}{\sqrt{(n_0 + n_1)(n_1 + n_2)}}. \quad (7.14)$$

Статистика (7.13) подчиняется стандартному нормальному распределению при достаточно большом объеме выборки (рисунок 7.2).

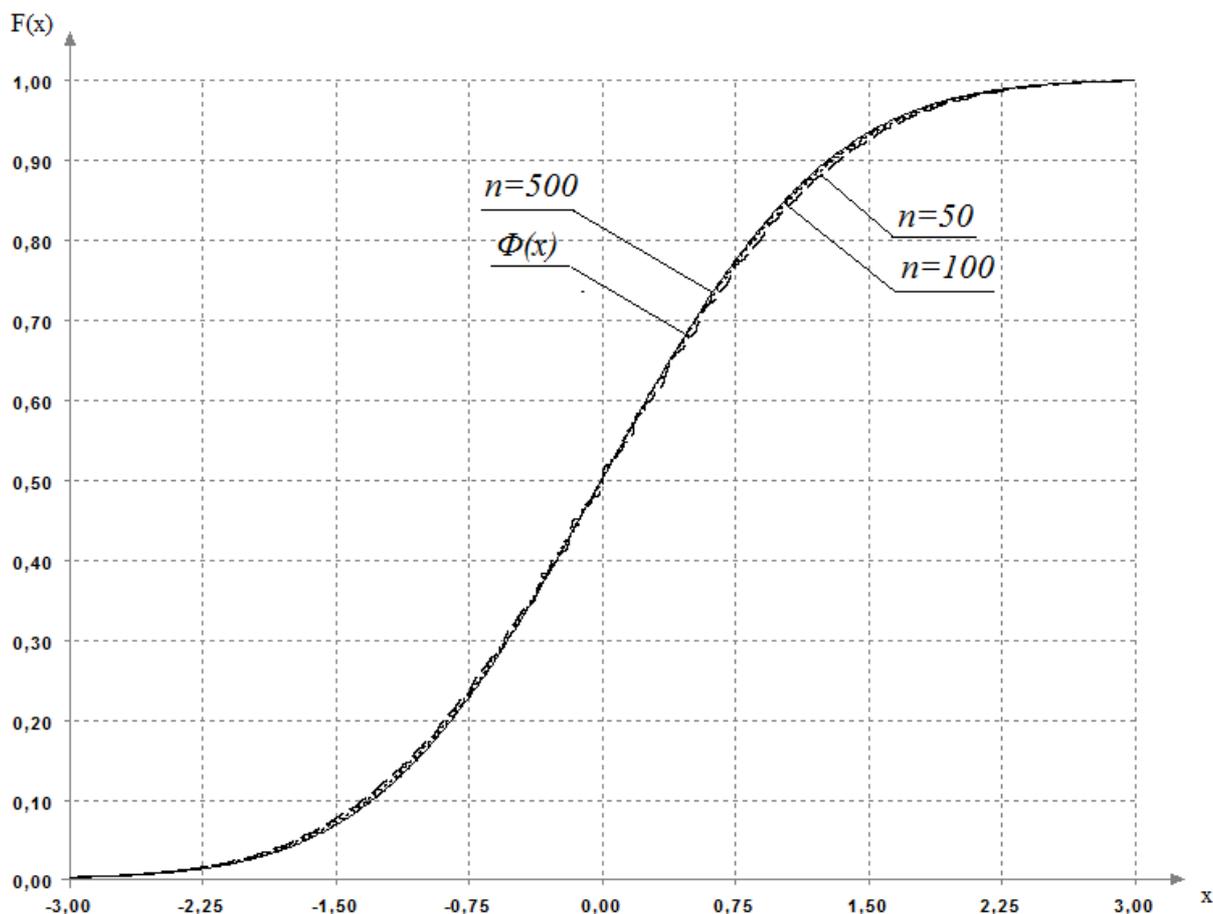


Рисунок 7.2 - Функция распределения статистики T_{MERT} для объемов 50, 100 и 500 наблюдений ($p_a = 0,3$)

7.4.5 Критерий максимума (MAX3)

Как уже было отмечено в п. 7.4.3 критерии тренда Кокрена-Армitedжа существенно теряют мощность при неправильном выборе коэффициентов. Поэтому в работе [179] был предложен мощный и в тоже время робастный (т.е. независимый от предположения о конкурирующей гипотезе) критерий, статистика которого равна максимальному значению статистик критерия тренда Кокрена-Армitedжа с наборами коэффициентов, оптимальными для трех моделей наследования (рецессивной, доминантной и аддитивной):

$$T_{\max} = \max(|T_r|, |T_a|, |T_d|). \quad (7.15)$$

7.4.6 Критерий минимума (MIN2)

В 2007 “World trust case-control consortium” (WTCCC) использовал для проверки ассоциаций другой робастный критерий, основанный на двух критериях – Хи-квадрат и критерии тренда Кокрена-Армитеджа с оптимальным набором коэффициентов для аддитивной модели [278]. Статистика критерия имеет вид

$$T_{MIN2} = \min \left\{ pv(T_a), pv(T_{\chi^2}) \right\}, \quad (7.16)$$

где $pv(T_a)$ – это достигаемый уровень значимости по критерию тренда Кокрена-Армитеджа с оптимальным набором коэффициентов для аддитивной модели, а $pv(T_{\chi^2})$ – это достигаемый уровень значимости критерия Хи-квадрат. Распределение T_{MIN2} было найдено и исследовано Јоо [191].

7.4.7 Критерий выбора генетической модели (GMS)

Другой подход к построению робастного критерия использовали Zheng и Ng [281], предложившие двухэтапную процедуру проверки гипотезы. На первом этапе определяется генетическая модель, а на втором этапе используется оптимальный для этой модели критерий тренда Кокрена-Армитеджа. Генетическая модель определяется на основе статистики T_{HWDIT} [267]

$$T_{HWDIT} = \frac{\sqrt{\frac{rs}{n}} \left[\left(\frac{r_2}{r} - \left(\frac{r_2}{r} + \frac{r_1}{2r} \right)^2 \right) - \left(\frac{s_2}{s} - \left(\frac{s_2}{s} + \frac{s_1}{2s} \right)^2 \right) \right]}{\left(1 - \frac{n_2}{n} - \frac{n_1}{2n} \right) \left(\frac{n_2}{n} + \frac{n_1}{2n} \right)}. \quad (7.17).$$

Статистика критерия GMS может быть записана как [Јоо_2009]

$$T_{GMS} = \begin{cases} |T_r|, & \text{if } (T_a > 0 \text{ and } T_{HWDIT} > c) \text{ or } (T_a \leq 0 \text{ and } T_{HWDIT} < -c); \\ |T_d|, & \text{if } (T_a > 0 \text{ and } T_{HWDIT} < -c) \text{ or } (T_a \leq 0 \text{ and } T_{HWDIT} > c); \\ |T_a|, & \text{otherwise,} \end{cases} \quad (7.18)$$

где c – это константа, обычно равная 1,645. В работе [279] Zang, Fung и Zheng предложили алгоритм для вычисления асимптотических распределений статистик критериев MAX3 и GMS при верной основной гипотезе.

7.4.8 Критерий Кульбака-Лейблера

Эмпирическая симметричная дивергенция Кульбака-Лейблера может быть найдена с использованием оценок $\hat{p}_i = r_i/r$ and $\hat{q}_i = s_i/s$:

$$\hat{\rho}_{KL} = \frac{1}{rs} \sum_{i=0}^2 (sr_i - rs_i) \ln \frac{r_i}{s_i}. \quad (7.19)$$

На базе этой статистики автор предложил построить новый критерий ассоциаций, который логично назвать критерием Кульбака-Лейблера. Эта статистика может рассматриваться как статистика критерия тренда Кокрена-Армитеджа с вектором коэффициентов $\bar{\omega} = (\ln(r_0/s_0), \ln(r_1/s_1), \ln(r_2/s_2))$. Для практических целей лучше использовать статистику

$$T_{KL} = \frac{n\hat{\rho}_{KL}}{8}, \quad (7.20)$$

которая при верной основной гипотезе в пределе подчиняется стандартному экспоненциальному распределению:

$$\lim_{n \rightarrow \infty} P\{T_{KL} < t | H_0\} = 1 - e^{-t}. \quad (7.21)$$

Функция распределения T_{KL} для разных объемов выборки приведена на рисунке 7.3. При малых объемах выборки распределение имеет дискретный вид ($n = 30$). При $n > 100$ ступеньки сглаживаются, распределение смещается вправо. При $n = 1000$ распределение находится правее предельного, а затем

начинает к нему приближаться, и при $n = 10000$ оно визуально неотличимо от стандартного экспоненциального закона распределения.

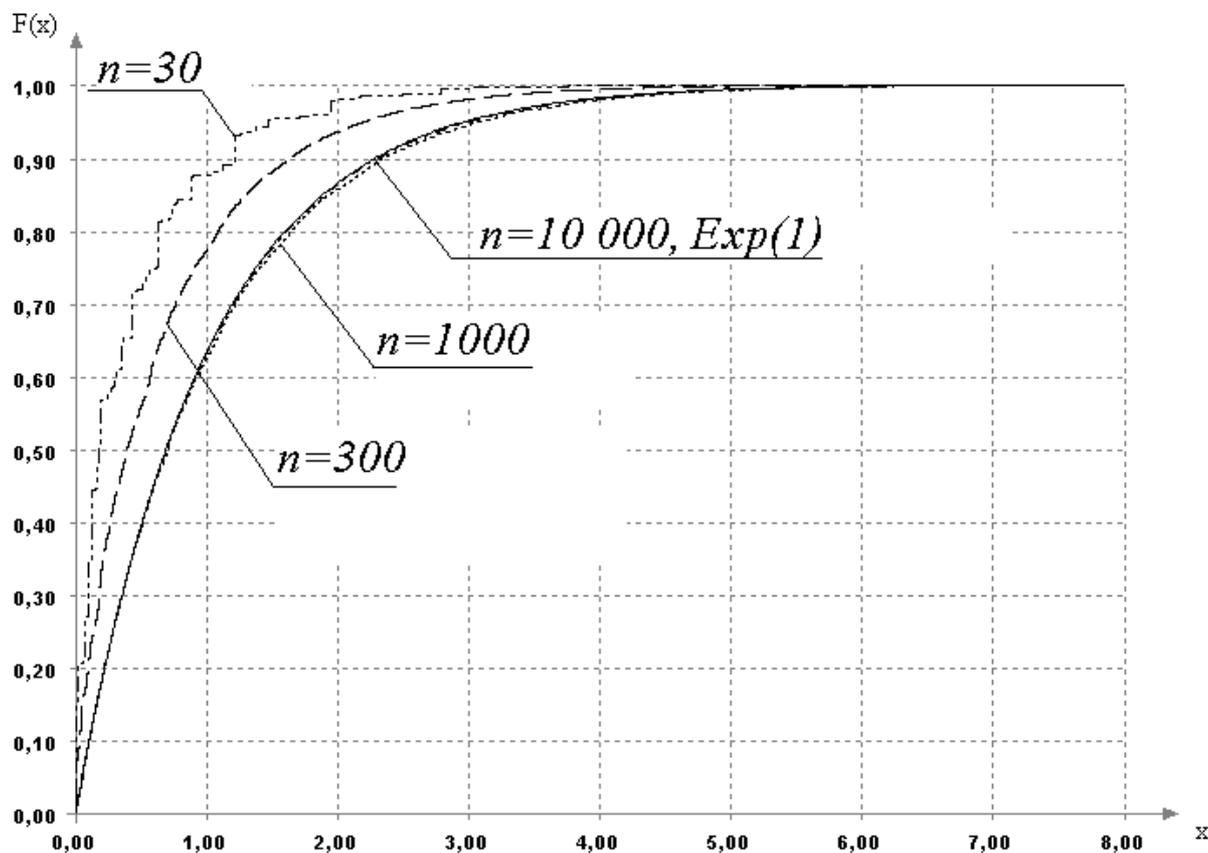


Рисунок 7.3 - Функция распределения T_{KL} для объемов выборки 30, 300, 1000 и 10000

7.5 Сравнительный анализ мощности критериев ассоциаций

Сравнение мощности критериев ассоциаций проведено для трех конкурирующих гипотез с разными видами наследуемости (таблица 7.3). Результаты моделирования методом Монте-Карло с использованием 10^7 повторений показаны в таблице 7.4 и на рисунках 7.4 и 7.5.

На рисунке 7.2 хорошо видно, что критерии тренда Кокрена-Армита действительно сильно теряют в мощности при неправильном выборе МОИ. Остальные критерии более или менее робастны по отношению к выбору конкурирующей гипотезы.

Мощность критерия Кульбака-Лейблера больше, чем мощность критерия χ^2 , но меньше, чем мощность критериев MAX3, GMS и CLRT, которые наиболее предпочтительны при использовании на практике.

Таблица 7.3 - Близкие альтернативы для анализа ассоциаций

Конкурирующая гипотеза	МОИ	K	p_a	ψ_1	ψ_2	ρ_{KL}
H_r	Рецессивная	0,05	0,1	1,000	1,490	0,001880
H_a	Аддитивная	0,05	0,5	1,066	1,136	0,002042
H_d	Доминантная	0,05	0,9	1,700	1,700	0,002217

Таблица 7.4 - Мощность критериев ассоциаций при близких конкурирующих гипотезах, $\alpha = 10^{-4}$ и $n=30000$

Критерий	H_a	H_d	H_r	Min
CATT(a)	0,512	0,010	0,015	0,010
CATT(d)	0,232	0,551	0,001	0,001
CATT(r)	0,245	0,001	0,439	0,001
Chi2	0,401	0,453	0,332	0,332
CLRT	0,451	0,479	0,350	0,350
GMS	0,446	0,470	0,363	0,363
MAX	0,439	0,484	0,352	0,352
MERT	0,510	0,223	0,168	0,168
MIN2	0,460	0,393	0,280	0,280
KLDT	0,399	0,456	0,338	0,338

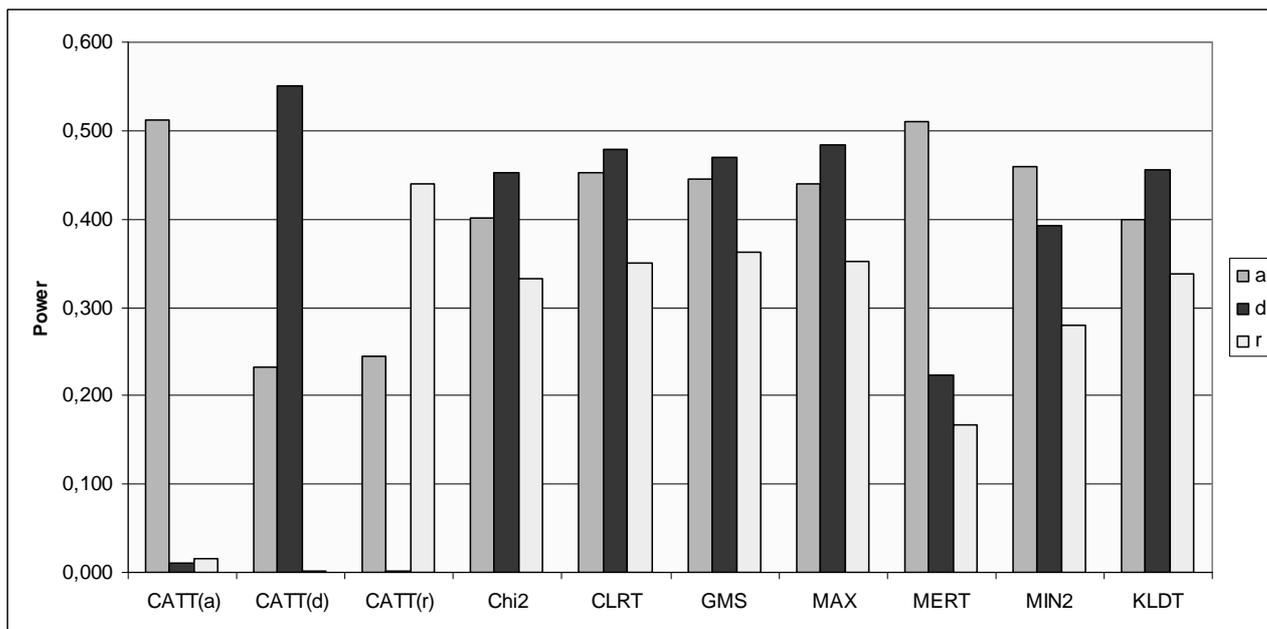


Рисунок 7.4 – Мощность критериев ассоциаций относительно близких конкурирующих гипотез, $\alpha = 10^{-4}$ и $n=30000$

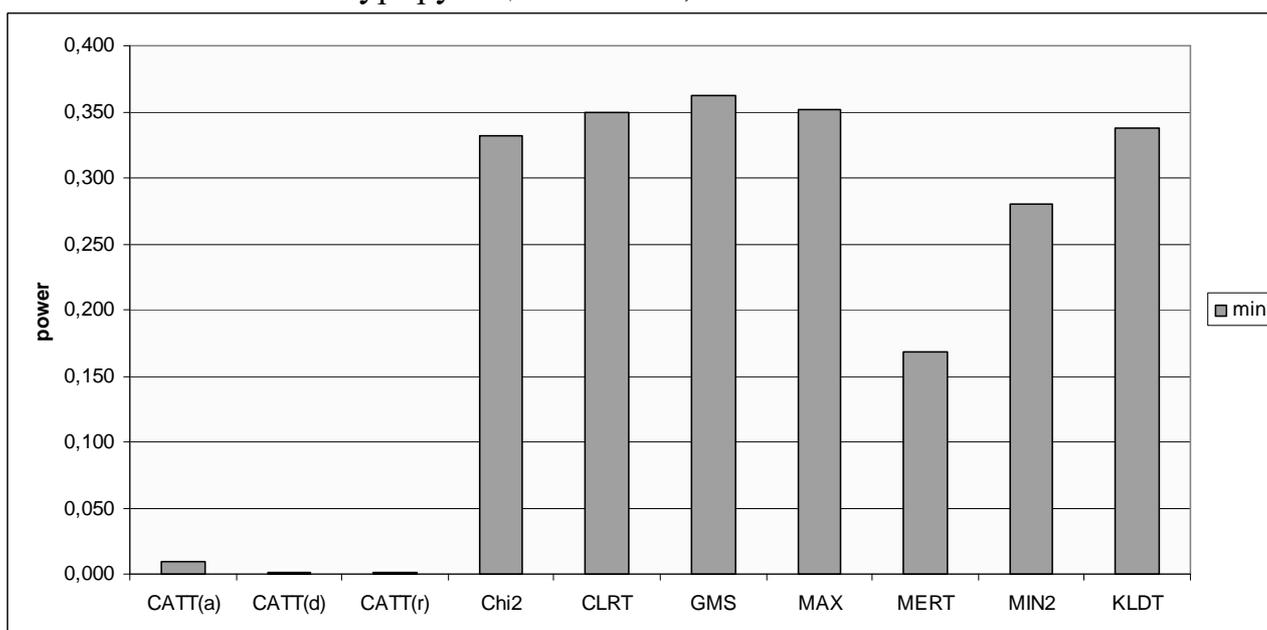


Рисунок 7.5 – Минимальные мощности критериев ассоциаций относительно конкурирующих гипотез H_a , H_r и H_d

7.6 Зависимость необходимого объема выборки от дивергенции Кульбака-Лейблера

В работе [179] рассчитаны необходимые объемы выборок для критерия тренда Кокрена-Армитеджа для различения конкурирующих гипотез с заданными вероятностями ошибок 1-го и 2-го рода. Для данных из таблиц 2 и 3

работы [179] были рассчитаны значения симметричной дивергенции ρ_{KL} , $n_{T_a}\rho_{KL}$ и $n_{T_r}\rho_{KL}$ (таблицы 7.5 и 7.6). Когда модель соответствует оптимальному вектору коэффициентов, то произведение $n\rho_{KL}$ близко к константе. Но если модель определена неверно, тогда величины $n\rho_{KL}$ различаются при разных ρ_{KL} . Это показано на рисунке 7.6. Ромбиками и квадратами показаны точки, когда вектор коэффициентов выбран неправильно, а кружочками и кружками – когда вектор коэффициентов соответствовал модели наследования. В обоих случаях зависимость $n(\rho_{KL})$ аппроксимируется гиперболой. В логарифмической шкале наклон кривой $n(\rho_{KL})$ больше, когда модель определена неверно.

Необходимый объем выборок для критерия МАХЗ и для разных генетических моделей при $\alpha = 0,05 / m$ ($m = 610000$ – число маркеров), и $\beta = 0,1$ рассчитан в работе [236]. Данные из таблиц 1-3 этой работы вместе с ρ_{KL} приведены в таблице 7.7. Зависимость необходимого объема выборки от ρ_{KL} показана на рисунке 7.7.

Таблица 7.5. Объемы выборки, необходимые для достижения мощности критерия 80% при $\alpha = 0.05$ относительно конкурирующей гипотезы H_r .

K	p_a	γ_1	γ_2	ρ_{KL}	n_{T_a} (add. score)	$n_{T_a}\rho_{KL}$	n_{T_r} (rec. score)	$n_{T_r}\rho_{KL}$
0,01	0,01	1	2	0,00007	16000000	1136	460089	33
0,01	0,10	1	2	0,00696	18881	131	4684	33
0,01	0,50	1	2	0,10625	413	44	298	32
0,10	0,01	1	2	0,00009	13000000	1171	367396	33
0,10	0,10	1	2	0,00882	15483	137	3743	33
0,10	0,50	1	2	0,13069	339	44	243	32
0,01	0,01	1	3	0,00023	3960000	895	153113	35
0,01	0,10	1	3	0,02193	5170	113	1568	34
0,01	0,50	1	3	0,28084	151	42	113	32
0,10	0,01	1	3	0,00030	3270000	981	120118	36
0,10	0,10	1	3	0,02898	4212	122	1232	36
0,10	0,50	1	3	0,34799	124	43	92	32

Таблица 7.6. Объемы выборки, необходимые для достижения мощности критерия 80% при $\alpha = 0.05$, относительно конкурирующей гипотезы H_d

К	p_a	γ_1	γ_2	ρ_{KL}	n_{T_a} (add. score)	$n_{T_a} \rho_{KL}$	n_{T_d} (dom. score)	$n_{T_d} \rho_{KL}$
0,01	0,01	2	2	0,01358	2408	33	2396	33
0,01	0,10	2	2	0,09166	367	34	347	32
0,01	0,50	2	2	0,07564	691	52	422	32
0,10	0,01	2	2	0,01717	1926	33	1916	33
0,10	0,10	2	2	0,11343	298	34	281	32
0,10	0,50	2	2	0,08996	576	52	356	32
0,01	0,01	3	3	0,04237	811	34	807	34
0,01	0,10	3	3	0,25083	136	34	128	32
0,01	0,50	3	3	0,16768	337	57	196	33
0,10	0,01	3	3	0,05577	638	36	635	35
0,10	0,10	3	3	0,31437	109	34	103	32
0,10	0,50	3	3	0,19760	282	56	167	33

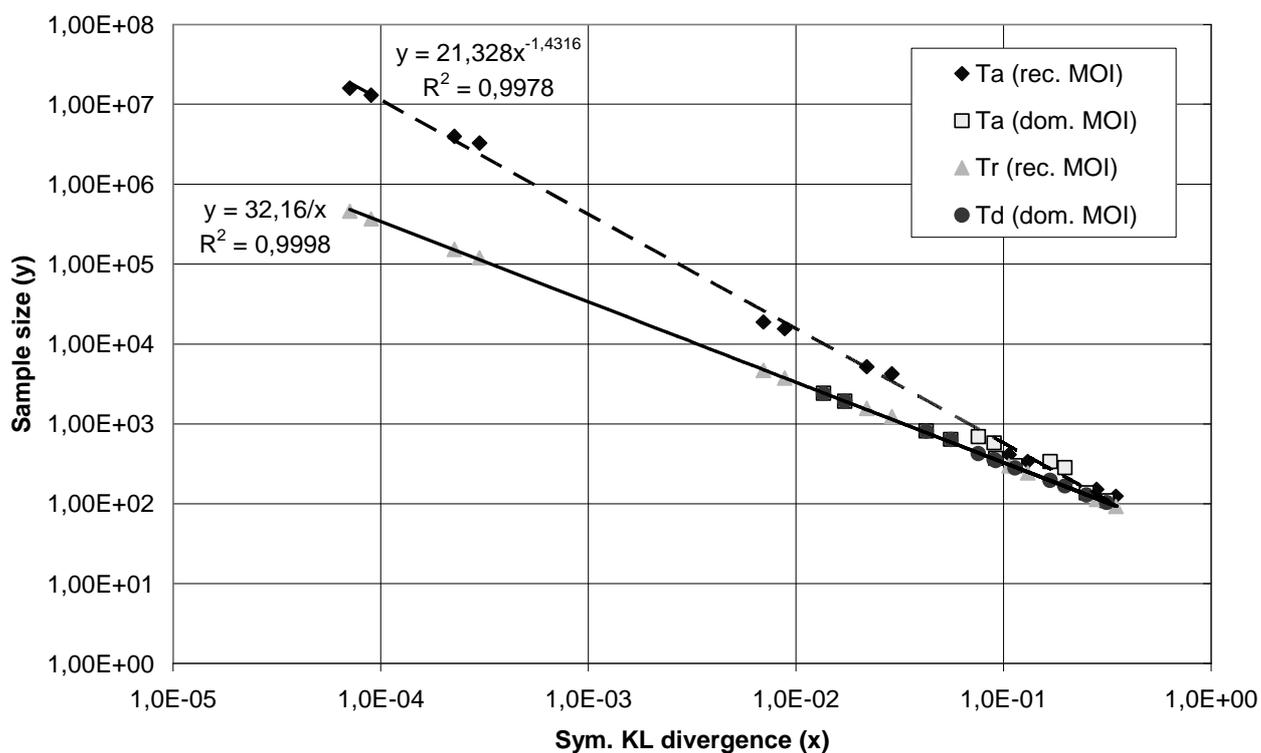


Рисунок 7.6 – Зависимость объема выборки от ρ_{KL} для САТТ ($\alpha = 0.05$ и $\beta = 0.2$)

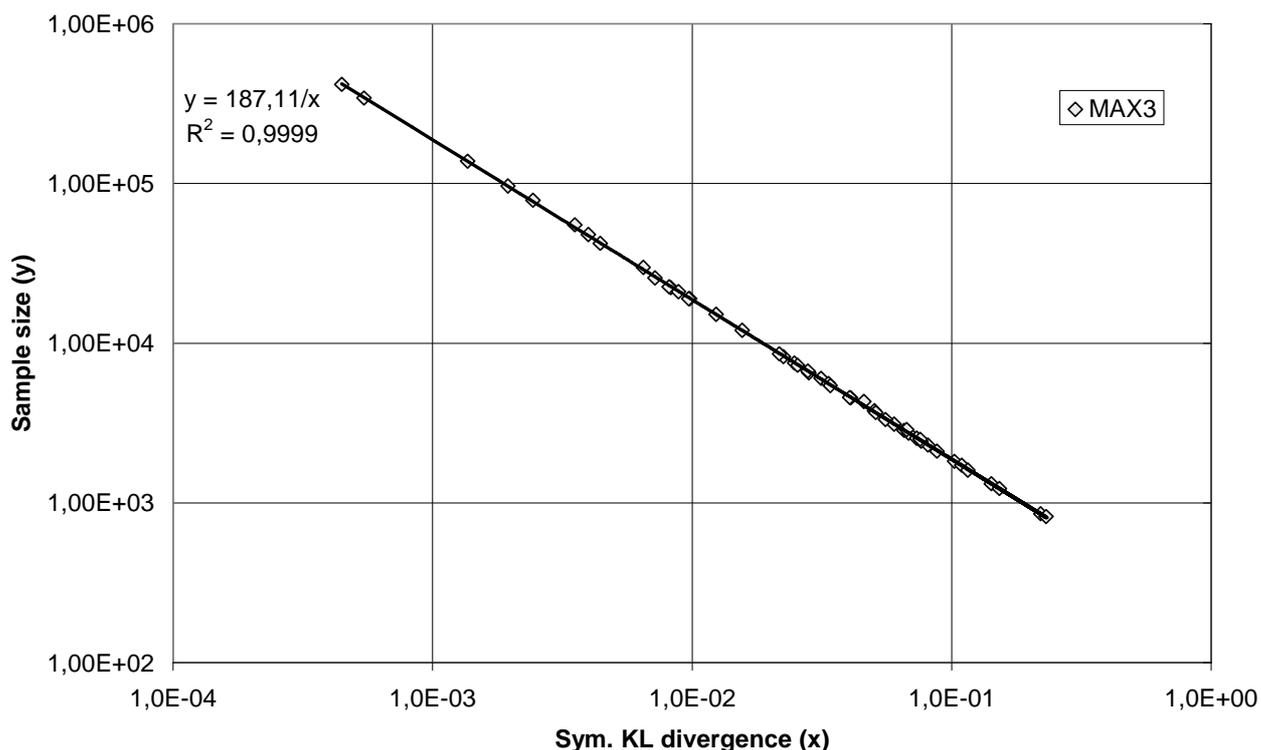


Рисунок 7.7 – Зависимость необходимого объема выборки от ρ_{KL} для критерия МАХЗ

Таблица 7.7 – Объем выборок, необходимый для того, чтобы критерий МАХЗ достигал 90% мощности при $\alpha = 0.05 / 610000$. Распространенность заболевания $K=0.05$.

MOI	p_a	ψ_1	ψ_2	ρ_{KL}	$n_{T_{\max}}$	$n_{T_{\max}} \rho_{KL}$
R	0,1	1,00	1,25	0,00054	342039	186
R	0,1	1,00	1,50	0,00195	96454	188
R	0,1	1,00	1,75	0,00398	47863	190
R	0,1	1,00	2,00	0,00648	29802	193
R	0,3	1,00	1,25	0,00442	42195	187
R	0,3	1,00	1,50	0,01560	12041	188
R	0,3	1,00	1,75	0,03136	6028	189
R	0,3	1,00	2,00	0,05035	3780	190
R	0,5	1,00	1,25	0,00979	18962	186
R	0,5	1,00	1,50	0,03348	5594	187
R	0,5	1,00	1,75	0,06547	2845	186
R	0,5	1,00	2,00	0,10250	1821	187
R	0,9	1,00	1,25	0,00719	25576	184
R	0,9	1,00	1,50	0,02248	8215	185
R	0,9	1,00	1,75	0,04084	4548	186

Продолжение таблицы 7.7

МОИ	p_a	ψ_1	ψ_2	ρ_{KL}	$n_{T_{\max}}$	$n_{T_{\max}} \rho_{KL}$
R	0,9	1,00	2,00	0,06010	3113	187
A	0,1	1,25	1,56	0,00970	18945	184
A	0,1	1,50	2,25	0,03409	5410	184
A	0,1	1,75	3,06	0,06831	2730	186
A	0,1	2,00	4,00	0,10925	1719	188
A	0,3	1,25	1,56	0,02169	8588	186
A	0,3	1,50	2,25	0,07339	2539	186
A	0,3	1,75	3,06	0,14208	1318	187
A	0,3	2,00	4,00	0,22033	854	188
A	0,5	1,25	1,56	0,02479	7503	186
A	0,5	1,50	2,25	0,08110	2302	187
A	0,5	1,75	3,06	0,15266	1229	188
A	0,5	2,00	4,00	0,23116	823	190
A	0,9	1,25	1,56	0,00826	22275	184
A	0,9	1,50	2,25	0,02552	7272	186
A	0,9	1,75	3,06	0,04594	4316	198
A	0,9	2,00	4,00	0,06711	2880	193
D	0,1	1,25	1,25	0,00814	22549	184
D	0,1	1,50	1,50	0,02814	6522	184
D	0,1	1,75	1,75	0,05558	3326	185
D	0,1	2,00	2,00	0,08780	2108	185
D	0,3	1,25	1,25	0,01237	15165	188
D	0,3	1,50	1,50	0,04046	4575	185
D	0,3	1,75	1,75	0,07618	2429	185
D	0,3	2,00	2,00	0,11539	1603	185
D	0,5	1,25	1,25	0,00886	21040	186
D	0,5	1,50	1,50	0,02793	6694	187
D	0,5	1,75	1,75	0,05108	3677	188
D	0,5	2,00	2,00	0,07557	2501	189
D	0,9	1,25	1,25	0,00045	416414	186
D	0,9	1,50	1,50	0,00136	137908	188
D	0,9	1,75	1,75	0,00244	78445	191
D	0,9	2,00	2,00	0,00354	55050	195

В предыдущем случае была рассмотрена ситуация, когда объемы выборки случаев и контрольной выборки совпадали. Но что произойдет, если пропорция не будет равна $\frac{1}{2}$? Женг и Гаствирт [280] нашли необходимые

объем выборок для критерия тренда Кокрена-Армитеджа с различными оценками дисперсии при различных пропорциях между объемами выборки случаев и контрольной выборки r : 0.5 (1:1), 0.33 (1:2), 0.20 (1:4). Данные из таблицы II [280] ρ_{KL} и $n_{T_a}\rho_{KL}$ приведены в таблице 7.8. График зависимости необходимого объема выборки от симметричной дивергенции Кульбака-Лейблера показан на рисунке 7.8. Полученная зависимость также хорошо аппроксимируется гиперболой, при этом из полученных графиков следует вывод, что наименьший объем выборок требуется при одинаковом объеме выборки случаев и контрольной выборок.

Таблица 7.8 - Объемы выборок, необходимые чтобы критерий тренда САТТ с аддитивным вектором коэффициентов достигал 80% мощности при $\alpha = 0.05$ и пропорцией объемов выборки случаев и контрольной выборки, равной 0,5; 0,33 и 0.2

r	МОИ	К	p_a	γ_1	γ_2	n_{T_a}	ρ_{KL}	$n_{T_a}\rho_{KL}$
0,5	R	0,01	0,1	1	3	1568	0,02193	34
0,5	R	0,01	0,2	1	3	421	0,08027	34
0,5	R	0,01	0,3	1	3	211	0,15646	33
0,5	R	0,01	0,5	1	3	113	0,28084	32
0,5	R	0,01	0,8	1	3	143	0,22608	32
0,5	A	0,01	0,1	2	3	318	0,10194	32
0,5	A	0,01	0,2	2	3	219	0,14842	33
0,5	A	0,01	0,3	2	3	201	0,16258	33
0,5	A	0,01	0,5	2	3	237	0,13999	33
0,5	A	0,01	0,8	2	3	566	0,05804	33
0,5	D	0,01	0,1	2	2	347	0,09166	32
0,5	D	0,01	0,2	2	2	263	0,11989	32
0,5	D	0,01	0,3	2	2	270	0,11699	32
0,5	D	0,01	0,5	2	2	422	0,07564	32
0,5	D	0,01	0,8	2	2	2354	0,01382	33
0,5	R	0,1	0,1	1	3	1232	0,02898	36
0,5	R	0,1	0,2	1	3	332	0,10484	35
0,5	R	0,1	0,3	1	3	168	0,20103	34
0,5	R	0,1	0,5	1	3	92	0,34799	32
0,5	R	0,1	0,8	1	3	121	0,26832	32
0,5	A	0,1	0,1	2	3	257	0,12678	33

Продолжение таблицы 7.8

r	MOI	K	p_a	γ_1	γ_2	n_{T_a}	ρ_{KL}	$n_{T_a} \rho_{KL}$
0,5	A	0,1	0,2	2	3	179	0,18211	33
0,5	A	0,1	0,3	2	3	165	0,19757	33
0,5	A	0,1	0,5	2	3	197	0,16803	33
0,5	A	0,1	0,8	2	3	475	0,06909	33
0,5	D	0,1	0,1	2	2	281	0,11343	32
0,5	D	0,1	0,2	2	2	216	0,14604	32
0,5	D	0,1	0,3	2	2	224	0,14099	32
0,5	D	0,1	0,5	2	2	356	0,08996	32
0,5	D	0,1	0,8	2	2	2007	0,01629	33
0,33	R	0,01	0,1	1	3	1644	0,02193	36
0,33	R	0,01	0,2	1	3	445	0,08027	36
0,33	R	0,01	0,3	1	3	226	0,15646	35
0,33	R	0,01	0,5	1	3	125	0,28084	35
0,33	R	0,01	0,8	1	3	166	0,22608	38
0,33	A	0,01	0,1	2	3	348	0,10194	35
0,33	A	0,01	0,2	2	3	243	0,14842	36
0,33	A	0,01	0,3	2	3	226	0,16258	37
0,33	A	0,01	0,5	2	3	269	0,13999	38
0,33	A	0,01	0,8	2	3	653	0,05804	38
0,33	D	0,01	0,1	2	2	382	0,09166	35
0,33	D	0,01	0,2	2	2	295	0,11989	35
0,33	D	0,01	0,3	2	2	307	0,11699	36
0,33	D	0,01	0,5	2	2	489	0,07564	37
0,33	D	0,01	0,8	2	2	2768	0,01382	38
0,33	R	0,1	0,1	1	3	1274	0,02898	37
0,33	R	0,1	0,2	1	3	347	0,10484	36
0,33	R	0,1	0,3	1	3	178	0,20103	36
0,33	R	0,1	0,5	1	3	111	0,34799	39
0,33	R	0,1	0,8	1	3	140	0,26832	38
0,33	A	0,1	0,1	2	3	280	0,12678	35
0,33	A	0,1	0,2	2	3	197	0,18211	36
0,33	A	0,1	0,3	2	3	185	0,19757	37
0,33	A	0,1	0,5	2	3	223	0,16803	37
0,33	A	0,1	0,8	2	3	549	0,06909	38
0,33	D	0,1	0,1	2	2	309	0,11343	35
0,33	D	0,1	0,2	2	2	242	0,14604	35
0,33	D	0,1	0,3	2	2	254	0,14099	36
0,33	D	0,1	0,5	2	2	413	0,08996	37

Продолжение таблицы 7.8

r	MOI	K	p_a	γ_1	γ_2	n_{T_a}	ρ_{KL}	$n_{T_a} \rho_{KL}$
0,33	<i>D</i>	0,1	0,8	2	2	2365	0,01629	39
0,2	<i>R</i>	0,01	0,1	1	3	2113	0,02193	46
0,2	<i>R</i>	0,01	0,2	1	3	576	0,08027	46
0,2	<i>R</i>	0,01	0,3	1	3	295	0,15646	46
0,2	<i>R</i>	0,01	0,5	1	3	168	0,28084	47
0,2	<i>R</i>	0,01	0,8	1	3	233	0,22608	53
0,2	<i>A</i>	0,01	0,1	2	3	467	0,10194	48
0,2	<i>A</i>	0,01	0,2	2	3	329	0,14842	49
0,2	<i>A</i>	0,01	0,3	2	3	308	0,16258	50
0,2	<i>A</i>	0,01	0,5	2	3	372	0,13999	52
0,2	<i>A</i>	0,01	0,8	2	3	914	0,05804	53
0,2	<i>D</i>	0,01	0,1	2	2	515	0,09166	47
0,2	<i>D</i>	0,01	0,2	2	2	403	0,11989	48
0,2	<i>D</i>	0,01	0,3	2	2	424	0,11699	50
0,2	<i>D</i>	0,01	0,5	2	2	687	0,07564	52
0,2	<i>D</i>	0,01	0,8	2	2	3925	0,01382	54
0,2	<i>R</i>	0,1	0,1	1	3	1611	0,02898	47
0,2	<i>R</i>	0,1	0,2	1	3	442	0,10484	46
0,2	<i>R</i>	0,1	0,3	1	3	229	0,20103	46
0,2	<i>R</i>	0,1	0,5	1	3	134	0,34799	47
0,2	<i>R</i>	0,1	0,8	1	3	196	0,26832	53
0,2	<i>A</i>	0,1	0,1	2	3	373	0,12678	47
0,2	<i>A</i>	0,1	0,2	2	3	266	0,18211	48
0,2	<i>A</i>	0,1	0,3	2	3	252	0,19757	50
0,2	<i>A</i>	0,1	0,5	2	3	308	0,16803	52
0,2	<i>A</i>	0,1	0,8	2	3	768	0,06909	53
0,2	<i>D</i>	0,1	0,1	2	2	414	0,11343	47
0,2	<i>D</i>	0,1	0,2	2	2	330	0,14604	48
0,2	<i>D</i>	0,1	0,3	2	2	351	0,14099	49
0,2	<i>D</i>	0,1	0,5	2	2	580	0,08996	52
0,2	<i>D</i>	0,1	0,8	2	2	3357	0,01629	55

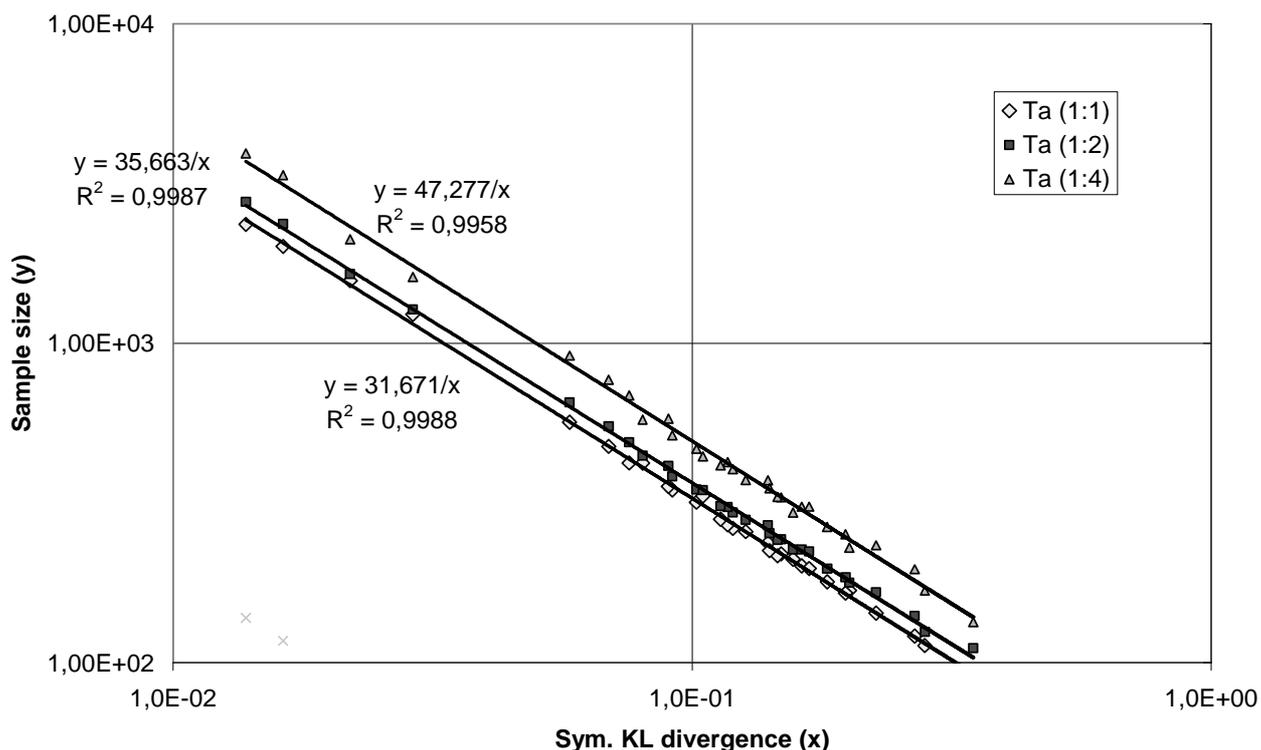


Рисунок 7.8 - Зависимость необходимого объема выборок от ρ_{KL} для САТТ с аддитивным вектором коэффициентов и пропорцией объемов выборки случаев и контрольной выборки, равной 0,5; 0,33; 0,2

7.7 Относительная эффективность критериев ассоциаций

На основании вычисления необходимого объема выборок можно сравнить различные критерии ассоциаций. По формуле из [280] можно предположить, что необходимый объем выборки для критерия тренда Кокрена-Армитеджа пропорционален величине

$$\vartheta(\alpha, \beta) = \left(\Phi^{-1} \left(1 - \frac{\alpha}{2} \right) + \Phi^{-1} (1 - \beta) \right)^2, \quad (7.22)$$

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

$$n_{\text{CATT}} = \frac{4,035 \cdot \vartheta(\alpha, \beta)}{\rho_{KL}}, \quad (7.23)$$

а для критерия МАХЗ соответственно

$$n_{\text{МАХЗ}} = \frac{4,238 \cdot \vartheta(\alpha, \beta)}{\rho_{KL}}. \quad (7.24)$$

В этом случае относительная эффективность критерия МАХЗ по отношению к критерию тренда Кокрена-Армитеджа с оптимальным набором коэффициентов, вычисленная по формуле (2.28), равна 0,95, т.е. критерий МАХЗ требует в среднем на 5% больше наблюдений, чем оптимальный критерий.

7.8 Оптимальное планирование двухэтапного эксперимента

Целью оптимального планирования эксперимента является нахождение таких параметров эксперимента, при которых затраты на его проведение минимальны. Для одноэтапного эксперимента это эквивалентно минимизации объема выборки при заданных вероятностях ошибок первого и второго рода.

При проведении одноэтапного эксперимента, как правило, затраты пропорциональны объему выборки, поэтому оптимальное планирование сводится к нахождению необходимого объема выборки, при котором вероятности ошибок первого и второго рода равны заданным значениям.

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 [254; 262; 276; 201, 202; 233; 265; 236].

Совокупные затраты двухэтапного эксперимента могут быть вычислены как

$$C = n(C_R + \pi m C_{G1} + (1 - \pi)((m - d)\alpha_1 + d(1 - \beta_1))C_{G2}), \quad (7.25)$$

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.

7.8.1 Моделирование вероятностей ошибок первого и второго рода

Для того чтобы определять оптимальные параметры двухэтапного эксперимента, используются алгоритмы моделирования вероятностей ошибок первого и второго рода, представленные на рисунках 7.9 и 7.10. На рисунках $MN(p_1, p_2, p_3)$ обозначает *мультиномиальное распределение* с тремя категориями, каждая из которых имеет вероятность p_i . Случайная величина, подчиненная мультиномиальному закону, легко моделируется при помощи $k - 1$ биномиальной величины, где k – это число категорий. Для моделирования биномиального распределения используется алгоритм BTRD при $np \geq 10$ и алгоритм BINV $np < 10$ [188, 192].

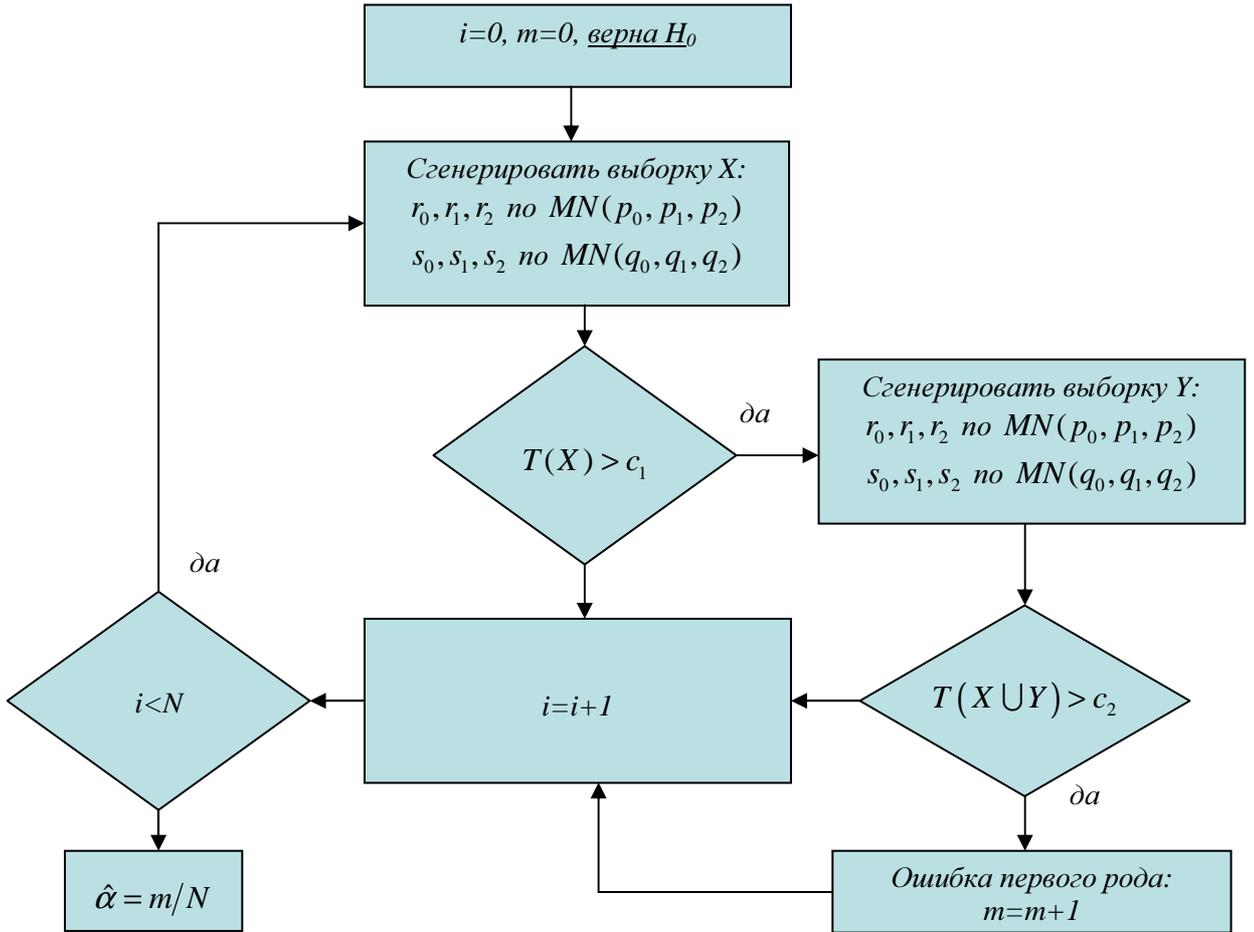


Рисунок 7.9 – Алгоритм моделирования вероятности ошибки первого рода

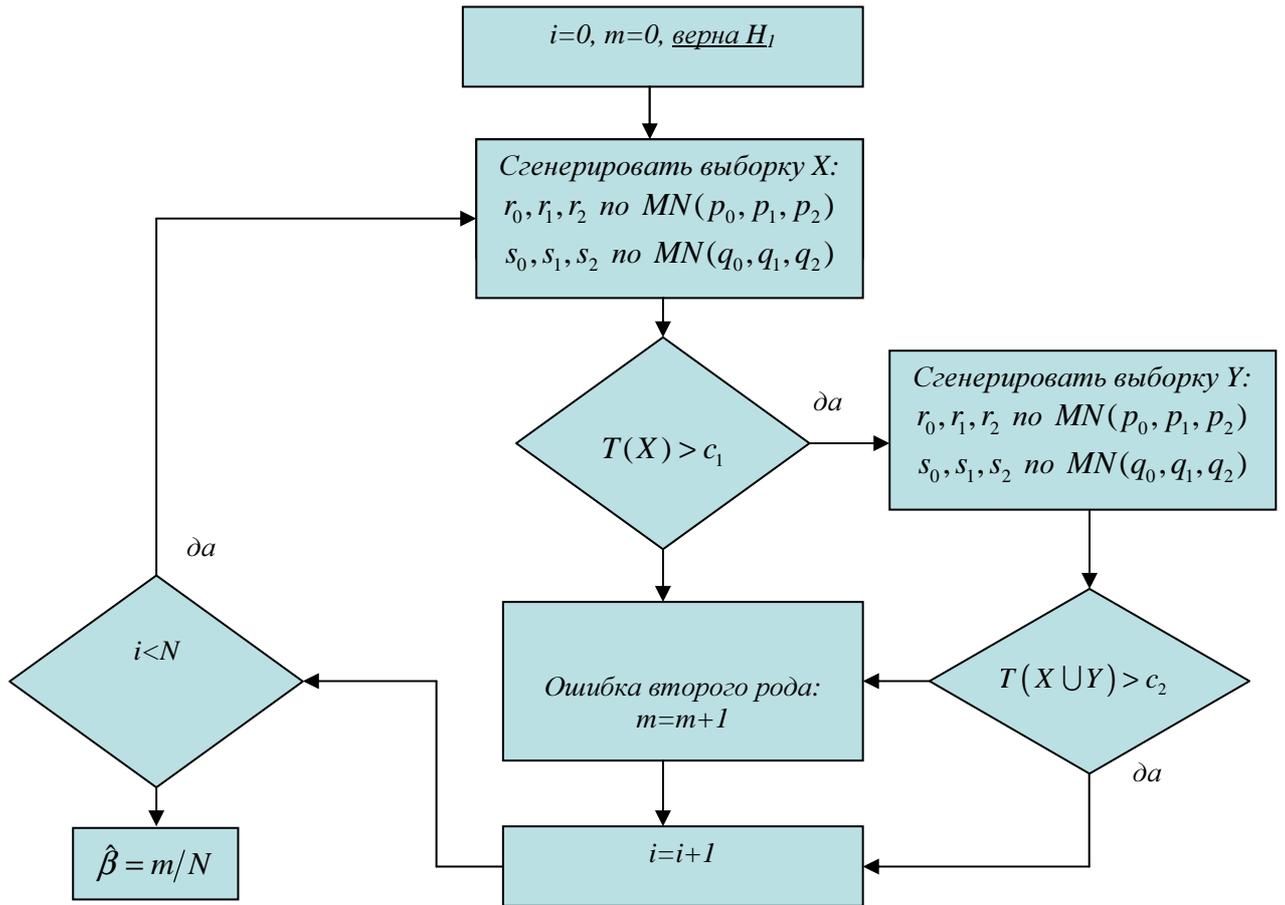


Рисунок 7.10 – Алгоритм моделирования вероятности ошибки второго рода

7.8.2 Оптимальное планирование двухэтапного эксперимента

Оптимальное планирование двухэтапного эксперимента по критерию МАХЗ были выполнено Нгуен, Палом и Шеффером [236] на основании шестимерного нормального распределения компонент статистики T_{\max} на первом и втором этапе. Анализ этого плана с помощью компьютерного моделирования показал, что их результаты не являются точными, как показано в таблице 7.9. Например, для первой строки этой таблицы вероятность ошибки второго рода оказывается равной примерно 0,087, в то время как она должна быть равна 0,1. В результате удалось уточнить оптимальные планы с уменьшением стоимости эксперимента до 3% в некоторых случаях.

Параметры генетической модели, оптимальные n , π , затраты и ρ_{KL} приведены в таблице 7.10. График стоимости проведения эксперимента в зависимости от ρ_{KL} показан на рисунке 7.11. По графику хорошо видно, что и при двухэтапном эксперименте затраты обратно пропорционально зависят от симметричной дивергенции Кульбака-Лейблера между распределениями в выборке случаев и в контрольной выборке.

Это, вообще говоря, позволяет сделать следующий вывод. При планировании эксперимента ПГАА с использованием критерия МАХЗ неважно, какая рассматривается конкурирующая модель: аддитивная, рецессивная или доминантная. Главное при таком планировании – задать расстояние между распределениями в основной и контрольной выборках, на основании которого можно определить необходимый объем выборки и примерную стоимость эксперимента.

Таблица 7.9 – Проверка планов из [236], при $\alpha = 8,2 \cdot 10^{-8}$, $\beta = 0,1$ методом Монте-Карло при количестве повторений $N_\alpha = 8,7 \cdot 10^8$ и $N_\beta = 2,4 \cdot 10^8$

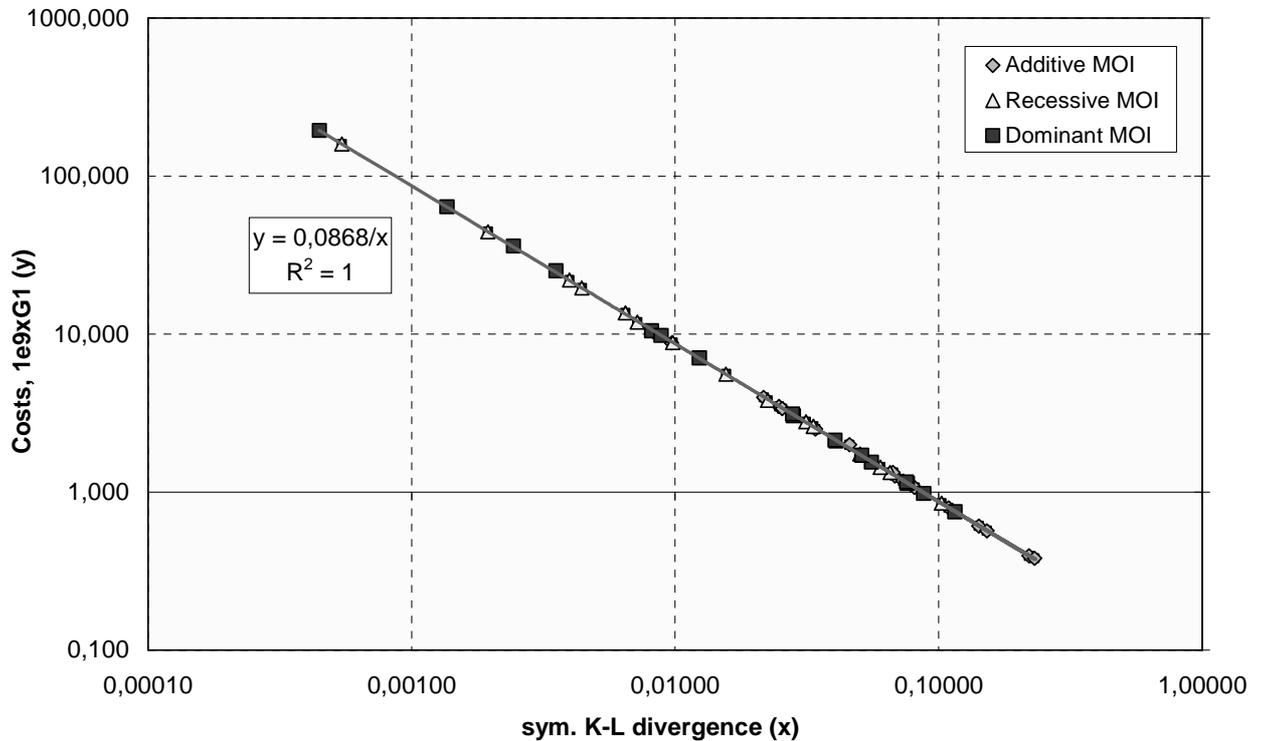
МОI	p_a	ψ_1	ψ_2	n	π	Критические границы	Затраты 10^9	Оценка затрат, $10^9 \times C_{G1}$	$\hat{\alpha}$, 10^{-8}	$\hat{\beta}$
R	0.1	1	2	33555	0.49	3.662260, 5.447788	13.992	13.990	8.16	0.087
R	0.3	1	2	4261	0.49	3.683559, 5.470014	1.782	1.7813	9.19	0.090
A	0.1	2	4	1943	0.48	3.662260, 5.444333	0.791	0.786	5.06	0.093
A	0.3	2	4	969	0.48	3.683559, 5.466117	0.396	0.400	8.96	0.082
D	0.1	2	2	2378	0.49	3.672701, 5.447788	0.989	0.976	7.01	0.092
D	0.3	2	2	1602	0.49	3.683559, 5.466117	0.726	0.671	8.16	0.160

Таблица 7.10 – Робастный двухэтапный план эксперимента при 90% мощности и $\alpha = 0.05/610000$. Распространенность заболевания $K=0.05$. Объемы выборки случаев и контрольной выборки совпадают

MOI	p_a	ψ_1	ψ_2	ρ_{KL}	n	π	Критические границы		Затраты, $10^9 \times C_{G1}$	$n\rho_{KL}$
							c_1	c_2		
R	0,1	1,00	1,25	0,00054	383340	0,489	3,6650	5,4876	159,520	209
R	0,1	1,00	1,50	0,00195	105627	0,490	3,6020	5,5407	44,557	206
R	0,1	1,00	1,75	0,00398	53239	0,486	3,6797	5,3274	22,019	212
R	0,1	1,00	2,00	0,00648	33127	0,482	3,6881	5,2479	13,608	215
R	0,3	1,00	1,25	0,00442	46291	0,489	3,5949	5,3447	19,616	205
R	0,3	1,00	1,50	0,01560	13457	0,489	3,6902	5,4011	5,595	210
R	0,3	1,00	1,75	0,03136	6688	0,490	3,6847	5,4268	2,788	210
R	0,3	1,00	2,00	0,05035	4182	0,488	3,6806	5,4602	1,740	211
R	0,5	1,00	1,25	0,00979	21219	0,489	3,6851	5,4309	8,839	208
R	0,5	1,00	1,50	0,03348	6251	0,489	3,6809	5,4672	2,607	209
R	0,5	1,00	1,75	0,06547	3187	0,489	3,6809	5,4784	1,328	209
R	0,5	1,00	2,00	0,10250	2039	0,489	3,6841	5,4988	0,850	209
R	0,9	1,00	1,25	0,00719	28630	0,488	3,6647	5,4056	11,901	206
R	0,9	1,00	1,50	0,02248	9160	0,489	3,6638	5,4796	3,815	206
R	0,9	1,00	1,75	0,04084	5079	0,486	3,6626	5,4436	2,106	207
R	0,9	1,00	2,00	0,06010	3456	0,487	3,6638	5,4005	1,436	208
A	0,1	1,25	1,56	0,00970	21332	0,479	3,6492	5,4345	8,787	207
A	0,1	1,50	2,25	0,03409	6096	0,480	3,6610	5,4401	2,508	208
A	0,1	1,75	3,06	0,06831	3072	0,480	3,6623	5,4518	1,263	210
A	0,1	2,00	4,00	0,10925	1934	0,479	3,6636	5,4746	0,794	211
A	0,3	1,25	1,56	0,02169	9667	0,480	3,6774	5,4724	3,980	210
A	0,3	1,50	2,25	0,07339	2860	0,480	3,6808	5,4639	1,176	210
A	0,3	1,75	3,06	0,14208	1484	0,480	3,6819	5,4668	0,610	211
A	0,3	2,00	4,00	0,22033	962	0,480	3,6814	5,4735	0,396	212
A	0,5	1,25	1,56	0,02479	8439	0,480	3,6815	5,4651	3,475	209
A	0,5	1,50	2,25	0,08110	2589	0,480	3,6791	5,4695	1,066	210
A	0,5	1,75	3,06	0,15266	1382	0,480	3,6810	5,4700	0,569	211
A	0,5	2,00	4,00	0,23116	924	0,480	3,6800	5,4700	0,381	214
A	0,9	1,25	1,56	0,00826	25084	0,481	3,6541	5,4590	10,343	207
A	0,9	1,50	2,25	0,02552	8198	0,479	3,6581	5,4472	3,370	209
A	0,9	1,75	3,06	0,04594	4854	0,479	3,6600	5,4447	1,996	223
A	0,9	2,00	4,00	0,06711	3227	0,478	3,6570	5,4569	1,326	217
D	0,1	1,25	1,25	0,00814	25324	0,485	3,6608	5,4397	10,488	206
D	0,1	1,50	1,50	0,02814	7305	0,487	3,6644	5,4602	3,033	206
D	0,1	1,75	1,75	0,05558	3715	0,488	3,6632	5,4675	1,546	206

Продолжение таблицы 7.10

MOI	p_a	ψ_1	ψ_2	ρ_{KL}	n	π	Критические границы		Затраты, $10^9 \times C_{G1}$	$n\rho_{KL}$
							c_1	c_2		
<i>D</i>	0,1	2,00	2,00	0,08780	2356	0,489	3,6698	5,4509	0,980	207
<i>D</i>	0,3	1,25	1,25	0,01237	16998	0,488	3,6858	5,4322	7,069	210
<i>D</i>	0,3	1,50	1,50	0,04046	5129	0,489	3,6818	5,4706	2,135	208
<i>D</i>	0,3	1,75	1,75	0,07618	2725	0,489	3,6839	5,4481	1,135	208
<i>D</i>	0,3	2,00	2,00	0,11539	1800	0,490	3,6817	5,4656	0,750	208
<i>D</i>	0,5	1,25	1,25	0,00886	23634	0,485	3,6811	5,4676	9,803	209
<i>D</i>	0,5	1,50	1,50	0,02793	7495	0,487	3,6805	5,4695	3,116	209
<i>D</i>	0,5	1,75	1,75	0,05108	4102	0,488	3,6828	5,5008	1,709	210
<i>D</i>	0,5	2,00	2,00	0,07557	2781	0,489	3,6820	5,4656	1,159	210
<i>D</i>	0,9	1,25	1,25	0,00045	468107	0,487	3,6696	5,5221	194,310	209
<i>D</i>	0,9	1,50	1,50	0,00136	154800	0,485	3,6770	5,5774	63,976	211
<i>D</i>	0,9	1,75	1,75	0,00244	87763	0,483	3,6851	5,2692	36,122	214
<i>D</i>	0,9	2,00	2,00	0,00354	61325	0,481	3,6935	5,7002	25,149	217

Рисунок 7.11 – Зависимость затрат от ρ_{KL} для критерия MAX3

7.8.3 Оптимальное дискретное планирование двухэтапного эксперимента

Современные технологии сканирования ДНК позволяют выполнять параллельную обработку нескольких образцов с помощью микроплашек, которые имеют, как правило, размер 96 или 384 ячеек (рисунок 7.12). Поэтому, вообще говоря, более рационально планировать эксперимент, который загружает максимальное количество микроплашек на первом этапе (на втором этапе сканируется небольшое число участков ДНК, поэтому используется другая технология). Таким образом, к начальной постановке по поиску оптимального плана добавляется условие кратности числа испытуемых на первом этапе числу ячеек в микроплашке.

Решение данной задачи может быть выполнено в два этапа: сначала находится оптимальный непрерывный план, а затем проводится минимизация затрат по узлам дискретной сетки, которая строится вокруг оптимального непрерывного плана.

В таблице 7.11 приведены результаты поиска оптимального плана для аддитивной модели ($p_a=0.3$, $\psi_1=2$, $\psi_2=4$) наследования для дискретного случая с числом ячеек в микроплашке равным 384. Оптимальный непрерывный план в этом случае требует объем выборки в 466 человек на первом этапе и 504 человека на втором этапе. Из двух дискретных вариантов (384 и 768 человек на первом этапе) следует остановиться на первом варианте, так как он дает решение, более близкое к заданным ограничениям, при этом немного уступая оптимальному непрерывному плану по стоимости.

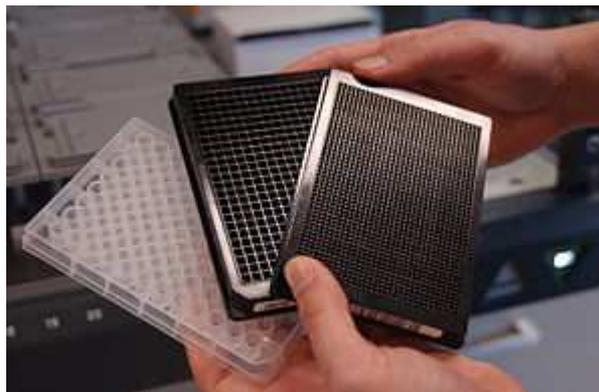


Рисунок 7.12 – Микроплашки для сканирования ДНК

Таблица 7.11 – Оптимальный двухэтапный план эксперимента (первая строка) при 90% мощности и $\alpha = 0,05 / 610000$. Распространенность заболевания $K = 0,05$. Объемы выборки случаев и контрольной выборки совпадают. $C_{G_2} / C_{G_1} = 100$, $C_R / C_{G_1} = 10^5$. Количество ассоциированных SNP $d = 1$, $p_a = 0.3$, $\psi_1 = 2$, $\psi_1 = 4$, число ячеек в микроплашке 384. Количество повторений $N_\alpha = 2,18 \times 10^{10}$, $N_\beta = 9,55 \times 10^6$

Число микроплашек	Объем выборки на первом этапе n_1	Объем выборки на втором этапе n_2	Затраты, $10^8 \cdot C_{G_1}$	Критические границы c_1, c_2	$\hat{\alpha}$, 10^{-8}	$\hat{\beta}$
1	384	578	4,06	3,3489; 5,38015	14,43	0,097
2	768	194	5,65	4,6389; 5,58331	7,47	0,046
Оптимальное непрерывное решение	461	501	3,97	3,68142 5,47351	9.69	0,086

7.9 Выводы

В данном разделе проведено исследование критериев однородности, используемых при проведении полногеномного анализа ассоциаций.

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).

Проведен сравнительный анализ мощности ряда критериев ассоциаций. Показано, что максимальной мощностью обладают критерии МАХЗ, CLRT и GMS. Критерии типа Хи-квдрат и предложенный автором критерий Кульбака-Лейблера имеют немного меньшую мощность [144].

Впервые показано, что при планировании эксперимента ПГАА с использованием критерия МАХЗ неважно, какая рассматривается конкурирующая модель: аддитивная, рецессивная или доминантная. Главное при таком планировании – в задании расстояния между распределениями в выборке случаев и в контрольной выборке, на основании которого можно определить оптимальный объем выборки и примерную стоимость эксперимента [246].

Проведены расчеты, показывающие, что относительная эффективность критерия МАХЗ по отношению к критерию тренда Кокрена-Армитеджа с оптимальным набором коэффициентов, вычисленная по формуле (2.28), равна 0,95, т.е. критерий МАХЗ требует в среднем на 5% больше наблюдений, чем оптимальный критерий [246].

Разработан и применен алгоритм поиска оптимального дискретного плана, опирающийся на использование компьютерного моделирования.

ГЛАВА 8. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ СТАТИСТИЧЕСКОГО АНАЛИЗА

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 [51, 53, 172, 70, 63, 77, 78, 214, 89, 93, 97, 221]. 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 и 5.0 имеются свидетельства о государственной регистрации программы ЭВМ [147, 148].

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 [140]. На систему «НКЦ ИТР: Статистика 1.0» имеются свидетельства о государственной регистрации программы ЭВМ [145, 146], а также сертификат фирмы «1С» о совместимости с системой 1С:Предприятие 8.2 [121, 133, 136, 137, 140, 141].

8.1 Программная система статистического анализа интервальных наблюдений одномерных непрерывных случайных величин "Интервальная статистика" (ISW)

Программная система «ISW» является следующим поколением системы статистического анализа одномерных наблюдений [48].

Программная система обладает рядом достоинств, которые выгодно отличают ее от конкурирующих систем.

1. Широкий выбор моделей теоретических законов распределения, включает более 30 стандартных законов и множество законов распределения, получаемых из стандартных с помощью операций, применяемых к этим моделям: операций сдвига, масштаба, смеси законов, произведения, усечения, логарифмирования [46, 47, 51].
2. Универсальное представление входных данных и возможность обрабатывать точечные, группированные, цензурированные, частично группированные и интервальные выборки [43, 206, 44, 45, 54, 59, 66].
3. Группирование наблюдений в задачах робастного оценивания и проверки статистических гипотез может осуществляться четырьмя различными способами: в соответствии с асимптотически оптимальным (минимизирует потери в информации Фишера), равновероятным, равночастотным и равномерным группированием. Реализована возможность задания интервалов группирования вручную.
4. Для проверки согласия эмпирического распределения с теоретическим используются критерии: отношения правдоподобия, χ^2 Пирсона, χ^2 Рао-Робсона-Никулина, Колмогорова, Смирнова, Крамера-Мизеса-Смирнова (ω^2) и Андерсона-Дарлинга (Ω^2). На базе результатов авторов программной системы корректность применения критериев согласия гарантируется как при проверке простых, так и при проверке сложных гипотез.

5. Для проверки гипотезы о принадлежности выборки нормальному закону используются критерии симметрии, эксцесса, Эппса-Палли, Шапиро-Уилка, Граббса, Д'Агустино, Фросини, Хегази-Грина, Гири, Дэвида-Хартли-Пирсона, Шпигельхальтера.
6. Для проверки гипотезы однородности используются критерии Лемана-Розенблатта, Смирнова и модифицированный критерий Смирнова.
7. Для проверки гипотезы однородности дисперсий используются критерии Бартлетта, Кокрена, Фишера, Хартли, Левене, Ансари-Бредли, Сижела-Тьюки, Зигеля-Тьюки, Муда, Кейпена, Клоца.
8. Оценивание параметров может осуществляться различными методами: максимального правдоподобия, максимального правдоподобия с предварительной группировкой наблюдений, MD-оценивания с минимизацией расстояний, задаваемых статистиками типа Колмогорова, статистиками типа ω^2 и Ω^2 Мизеса, с использованием предложенных авторами оптимальных L-оценок по выборочным квантилям.
9. На базе включенных в систему робастных методов оценивания реализована эффективная параметрическая процедура отбраковки аномальных наблюдений [52].
10. Графическая подсистема позволяет просматривать функции распределения, плотности, гистограммы, ядерные оценки плотности.
11. Разработаны средства для моделирования распределений статистик критериев согласия при различных сложных гипотезах и различных альтернативах. Это позволяет исследовать распределения статистик при различных сложных проверяемых гипотезах, строить приближенные математические модели этих распределений, исследовать мощность критериев относительно различных близких альтернатив.
12. На базе системы возможна организация исследований законов распределения различных одномерных статистик, вычисляемых при статистическом анализе наблюдений.

8.2 Программное обеспечение статистического моделирования в задачах проведения и обработки измерений «НКЦ ИТР: Статистика 1.0»

Сегодня в любой сфере человеческой деятельности накапливается огромное количество разнообразной информации, которая является результатом проведения экспериментов и регистрации наблюдений. Собранные данные требуют анализа, а полученные результаты анализа необходимо уметь интерпретировать. Для этого требуются развитые методы статистического анализа данных и современные статистические инструментальные средства.

8.2.1 Программная платформа 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. В последнем случае пользователь работает с системой из окна браузера.

Другой особенностью платформы 1С:Предприятие 8.2 является наличие встроенного языка программирования и большое количество доступных для использования при разработке прикладных объектов и механизмов. При поставке конфигурации с открытым кодом пользователь может сам вносить изменения в программу, дорабатывая типовую конфигурацию под свои особенности.

Таким образом, выбор платформы 1С:Предприятие 8.2 автоматически снимает целый ряд задач с разработчика, позволяя ему сосредоточиться целиком на предметной области.

В разработанную автором систему встроены методы, которые необходимы для проведения наиболее полного статистического анализа данных. На рисунке. 8.1 приведено главное окно системы статистического анализа.

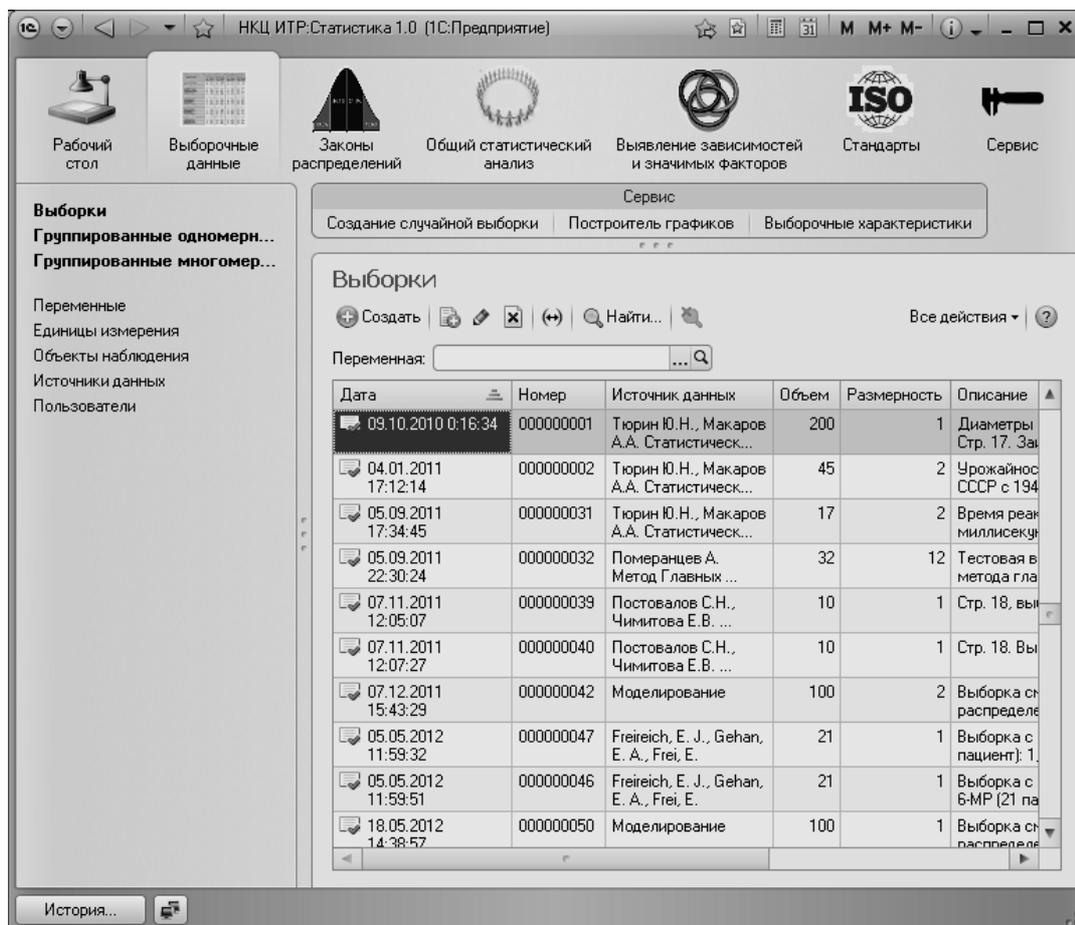


Рисунок 8.1 – Главное меню системы

8.2.2 Объектная модель программной системы

Основной сущностью разработанной и развиваемой системы является понятие «Переменная». Переменная – это любая характеристика наблюдаемого объекта. Например, если мы наблюдаем человека, то такими характеристиками могут быть *рост*, *вес*, *пол*, *возраст* и т.д. Переменные хранятся в справочнике «Переменные» (рисунок 8.2).

Переменные могут иметь разные шкалы измерений. Шкалы измерений принято классифицировать по типам измеряемых данных, которые определяют

допустимые для данной шкалы математические преобразования, а также типы отношений, отображаемых соответствующей шкалой. Современная классификация шкал была предложена в 1946 году Стивенсом [Stevens_1946].

Рисунок 8.2 – Форма элемента справочника «Переменные»

Номинальная шкала. В этом случае результатами наблюдений являются элементы множества допустимых значений, причем эти элементы никак не упорядочены. Примеры номинальных шкал: Пол = {мужской, женский}, Сигнал светофора = {красный, желтый, зеленый}. В программе возможные значения номинальных переменных задаются в подчиненном справочнике «Значения переменных».

Порядковая шкала. В отличие от номинальной шкалы – элементы множества наблюдений можно упорядочить по возрастанию некоторым способом. Например, Успеваемость студентов = {неудовлетворительно, удовлетворительно, хорошо, отлично}. В программе порядковые наблюдения вводятся целыми числами.

Метрическая (интервальная) шкала. Результаты измерений – числа, которые являются результатом сравнения с эталоном. Например, чтобы измерить время, мы сравниваем промежуток времени с заданной единицей измерения (секунда, минута, час и т.п.). Важной характеристикой данной шкалы является отсутствие точки «нуля».

Метрическая (относительная) шкала. Если мы к интервальной шкале добавим точку «нуля» – мы получим относительную шкалу. Точка «нуля» означает отсутствие некоторого свойства, например, нулевая масса, нулевая длина, нулевое количество человек в автобусе и т.д. Для двух измерений допустимо отношение «во столько-то раз больше». Например, объект массой 10 килограмм в два раза больше чем объект массой 5 килограмм.

Первая попытка разработки прототипа конфигурации статистического анализа на платформе 1С:Предприятие 8.2 была предпринята в 2009 [121]. Опыт эксплуатации этой системы показал следующие проблемы.

Во-первых, процедуры трудоемких вычислений, написанные на встроенном языке, обладают очень большим временем выполнения по сравнению с аналогичными процедурами, написанными на языке С [128]. Поэтому естественным решением этой проблемы является вынесение трудоемких вычислительных процедур во внешние компоненты, которые поддерживает платформа 1С:Предприятие 8.2.

Во-вторых, в системе 1С:Предприятие 8.2 нет экспоненциального формата для хранения чисел. При определении реквизита базы данных допускается указывать только длину числа и точность – число знаков после запятой. Для решения этой проблемы был разработан следующий механизм: каждое наблюдение в системе хранится в двух реквизитах: в первом – число длиной 10 символов с тремя знаками после запятой, а во втором – ссылка на элемент справочника «Единицы измерения переменных». За счет этого в

системе можно хранить разномасштабные наблюдения – например, километры и миллиметры (рисунок 8.3).

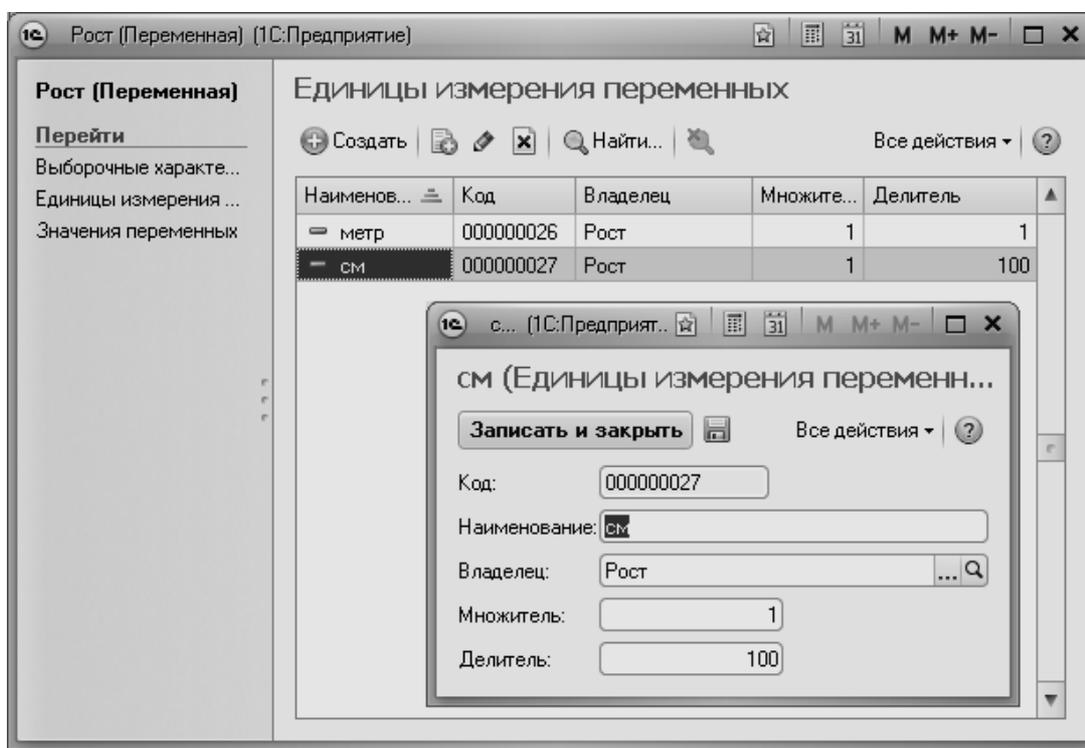


Рисунок 8.3 – Единицы измерения переменных

В-третьих, проблемой была организация хранения наблюдений переменных в выборках. Поскольку система ориентирована на многомерный статистический анализ, то в выборке могут присутствовать наблюдения произвольного числа переменных. Однако в структуре данных системы 1С:Предприятия 8.2 нельзя динамически менять число колонок таблиц базы данных. Поэтому было решено, что наблюдения будут храниться в табличной части документа «Выборка» с четырьмя колонками: в первой колонке номер наблюдения, во второй – переменная, в третьей колонке – значение наблюдения, а в четвертой – единица измерения переменной.

Естественно, что такой формат удобен для хранения наблюдений, но неудобен для ввода данных пользователем, поэтому для ввода данных разработана специальная форма в классическом табличном виде (рисунок 8.4).

Выборка 000000031 от 05.09.2011 17:34:45

Провести и закрыть | Провести | Все действия ▾ ?

Номер: 000000031 | Дата: 05.09.2011 17:34:45

Пользователь: Постовалов Сергей

Источник данных: Тюрин Ю.Н., Макаров А.А. Статистический анали: ...

Объем: 17 | Размерность: 2

Время реакции на свет и звук, в миллисекундах Стр. 103.

Данные | Переменные

Номер	Объект	Время реакции на звук, миллисекунда	Время реакции на свет, миллисекунда
1		223	181
2		104	194
3		209	173
4		183	153
5		180	168
6		168	176
7		215	163
8		172	152
9		200	155
10		191	156
11		197	178
12		183	160
13		174	164
14		176	169
15		155	155
16		115	122
17		163	144

Рисунок 8.4 – Форма документа «Выборка»

Если наблюдения являются группированными, т.е. потеряна информация об индивидуальных наблюдениях, то для их хранения предназначен документ «Группированная многомерная выборка» (рисунок 8.5). На вкладке «Переменные» задается список переменных и их групп (интервалов группирования), а на вкладке «Количество наблюдений» для каждой комбинации групп вводится количество наблюдений.

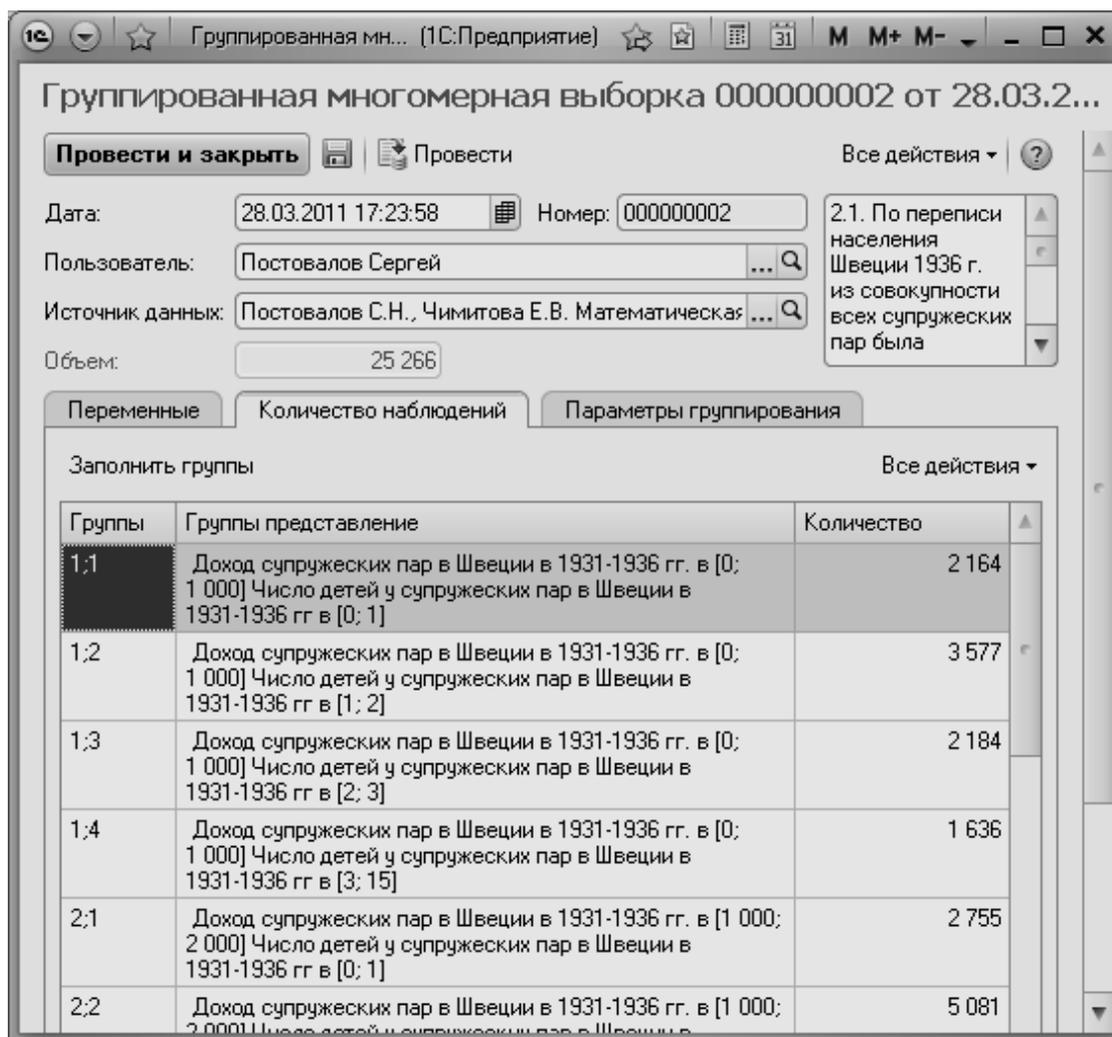


Рисунок 8.5 – Форма документа «Группированная многомерная выборка»

8.2.3 Идентификация закона распределения

Целью первичной обработки экспериментальных наблюдений обычно является установление закона распределения, наиболее хорошо описывающего случайную величину, выборку которой наблюдают. Насколько хорошо наблюдаемая выборка описывается теоретическим законом, проверяют с использованием различных критериев согласия.

Процедура идентификации закона распределения, как правило, состоит из двух этапов. На первом этапе оцениваются параметры закона распределения, а на втором проводится проверка гипотезы о согласии.

Оценивание параметров закона в системе реализовано по методу максимального правдоподобия, методам минимального расстояния и по порядковым статистикам. Проверка гипотезы о согласии в программной системе реализована по критериям Колмогорова, Смирнова, Крамера-Мизеса-Смирнова и Андерсона-Дарлинга в соответствии с ГОСТ Р 50.1.037-2002 [69, 25, 135] и критериям хи-квадрат Пирсона, отношения правдоподобия в соответствии ГОСТ Р 50.1.033-2001 [32, ГОСТ_Р_2001, 83, 135].

На рисунке 8.6 показан пример работы программы. В реквизите «Переменная» указывается анализируемая переменная. В списке выборок необходимо отметить, какие выборки необходимо объединить для анализа. В реквизите «Закон распределения» указывается соответствующее распределение. Параметры закона распределения, которые требуется оценивать, отмечаются флажками. Список критериев согласия отображается в поле внизу формы, флажками можно отметить критерии, которые требуется использовать для проверки гипотезы.

Идентификация закона распределения

Выполнить | Записать | График | Все действия

Переменная: Возраст | Единица измерения: Год

Закон распределения: Нормальное

Параметр	Оценивать	Значение
Масштаба	<input checked="" type="checkbox"/>	9,36729848
Сдвига	<input checked="" type="checkbox"/>	34,43749523

Метод оценивания: Метод максимального правдоподобия

Критерий выбран	Критерий	Значение статистики	Достигаемый уров...	Результат
<input checked="" type="checkbox"/>	Андерсона-Дарлинга	0,22	0,826	НЕ ОТВЕРГАЕТСЯ
<input checked="" type="checkbox"/>	Колмогорова	0,59	0,058	НЕ ОТВЕРГАЕТСЯ
<input checked="" type="checkbox"/>	Крамера-Мизеса-Смирнова	0,03	0,088	НЕ ОТВЕРГАЕТСЯ
<input checked="" type="checkbox"/>	Смирнова (левосторонний)	1,39	0,967	НЕ ОТВЕРГАЕТСЯ
<input checked="" type="checkbox"/>	Смирнова (правосторонн...	0,68	1	НЕ ОТВЕРГАЕТСЯ

Рисунок 8.6 – Обработка «Идентификация закона распределения»

Для выполнения процедуры оценивания параметров и проверки гипотезы следует нажать на кнопку «Выполнить». Результаты проверки гипотезы отображаются в колонках «Значение статистики», «Достигаемый уровень значимости» и «Результат».

8.2.4 Проверка статистических гипотез

Большинство статистических методов обработки данных имеет целью либо оценивание параметров статистических моделей, либо проверку каких-либо предположений о виде или свойствах наблюдаемых случайных величин. Такие предположения в математической статистике называются гипотезами. Как правило, выделяют следующие виды статистических гипотез:

- гипотезы о виде распределения,
- гипотезы однородности,
- гипотезы независимости.

Проверка данных гипотез может быть также необходима для некоторых специализированных методов анализа данных, например, одни методы требуют независимости рассматриваемых величин, другие предполагают, что данные подчинены нормальному закону распределения. Игнорирование подобных требований может привести к ошибкам интерпретации результата анализа этими методами.

Существует несколько видов критериев проверки многомерных выборок на независимость. Для точечных наблюдений в системе реализованы критерии, основанные на коэффициентах корреляции Кендалла, Спирмена, Пирсона. Для группированных выборок реализован критерий независимости Хи-квадрат Пирсона.

Для проверки гипотезы однородности средних в данной программе реализованы три известных критерия для проверки гипотезы однородности

средних: критерий Стюдента, критерий Беренса-Фишера (тест Уэлча для проблемы Беренса-Фишера) и критерий Манна-Уитни (U-критерий).

В данной программе реализованы 3 параметрических и 3 непараметрических критерия для проверки гипотезы о *равенстве дисперсий*. Критерии Бартлетта, Левене и Фишера являются параметрическими критериями, которые предполагают, что сравниваемые выборки подчиняются нормальному закону распределения. Критерии Бартлетта и Левене позволяют проверять более чем две выборки на равенство дисперсий, при этом гипотеза отклоняется, если хотя бы одна выборка имеет дисперсию, значительно отличающуюся от дисперсий остальных выборок.

Критерии Ансари-Бредли, Сиджела-Тьюки и Муда относятся к непараметрическим и не зависят от вида распределений, которым подчиняются сравниваемые выборки. Эти тесты позволяют сравнивать только две выборки одновременно. В программе используются приближения распределений статистик непараметрических критериев с помощью нормального закона, что требует достаточно большого количества наблюдений (не менее 10 в каждой выборке).

8.2.5 Выявление зависимостей и значимых факторов

Многомерный статистический анализ, как правило, заключается в выявлении статистических закономерностей и определении значимых факторов. В программе реализованы наиболее распространенные методов многомерного статистического анализа:

- корреляционный анализ;
- регрессионный анализ;
- дисперсионный анализ;
- факторный анализ.

Модуль факторного анализа включает в себя метод главных компонент. Модуль дисперсионного анализа включает однофакторный дисперсионный анализ и обобщенную линейную модель (GLM). Модуль регрессионного анализа включает в себя различные методы линейной регрессии, оценки которой находятся по методу наименьших квадратов. Модуль корреляционного анализа включает вычисление парной корреляции, частной корреляции, множественной корреляции, проверку гипотез о значимости коэффициентов корреляции.

8.2.6 Поддержка ГОСТов по статистике

8.2.6.1 Проверка нормальности

Нормальное распределение, также называемое распределением Гаусса, – распределение вероятностей, которое играет важнейшую роль во многих областях знаний: физике, медицине, экономике. Например, физическая величина, подверженная влиянию значительного числа случайных помех, часто подчиняется нормальному распределению, поэтому из всех распределений чаще всего встречается нормальное. Но бывают случаи, когда рассматриваемая величина распределена по какому-либо другому – не нормальному закону, или имеет засоренное (смешанное) распределение. В этом случае результат анализа будет некорректным и может привести к серьезным ошибкам. Поэтому проверка данных на нормальность является важной практической задачей.

В работе рассматриваются методы и критерии для проверки отклонения распределения вероятностей от нормального распределения при независимых наблюдениях. Известно много различных критериев нормальности, однако на данном этапе реализованы критерии, включенные в ГОСТ Р 5479-2002 [21, 137], такие как: графический метод, направленные критерии проверки на асимметрию и кривизну, многосторонние критерии (критерий Шапиро-Уилка и Эппса-Палли. Эти критерии в основном рассчитаны на полные,

негруппированные данные и удобны для использования специалистами в промышленности для проверки на нормальность различных данных в ходе проведения измерений, контроля и испытаний.

На рисунке 8.7 представлена форма обработки проверки нормальности. В поле «Переменная» нужно выбрать анализируемую случайную величину, после чего ниже отобразится список выборок, в которых присутствует заданная переменная. Флажками следует отметить выборки, относительно которых требуется проверить нормальность, и критерии нормальности, а затем нажать на кнопку «Проверить». Результат проверки отобразится в колонках «Значение статистика» и «Результат». Более подробный результат можно увидеть, нажав на кнопку «Печать» в левом верхнем углу формы. Полученный отчет содержит информацию об исходной выборке, и результат проведенного анализа с детальной информацией.

Критерий	Источник	Значение стати...	Крит...	ДУЗ	Результат
<input checked="" type="checkbox"/> Направленный критерий проверки на асимметрию	ГОСТ Р ИСО 5479-...	-0,897	0,3	-	НЕ ОТВЕРГАЕТ
<input checked="" type="checkbox"/> Направленный критерий проверки на кривизну	ГОСТ Р ИСО 5479-...	3,522	4,098	-	НЕ ОТВЕРГАЕТ
<input checked="" type="checkbox"/> Шапиро-Уилка	ГОСТ Р ИСО 5479-...	0	0,892	-	ОТВЕРГАЕТСЯ
<input checked="" type="checkbox"/> Эппса-Палли	ГОСТ Р ИСО 5479-...	26,81	0,3668	-	ОТВЕРГАЕТСЯ

Рисунок 8.7 – Форма обработки проверки нормальности

8.2.6.2 Планирование статистического приемочного контроля

Одной из разновидностей статистического приёмочного контроля является контроль по альтернативному признаку. Все процедуры данного типа контроля стандартизированы и описаны в ГОСТ Р ИСО 2859-1-2007 [20, 133].

Пусть рассматривается большая совокупность предметов 2-х типов. Речь идет о партии N изделий, в которой n изделий оказываются бракованными (т.е. имеют хотя бы один дефект), а остальные $(N - n)$ являются годными. При неизвестном n можно оценить это число по сравнительно маленькой выборке, выбрав наугад (случайно) m изделий и определив из них все бракованные (пусть их число равно k).

По наблюдаемому числу k можно дать оценку доли брака (входной уровень дефектности) во всей партии. Партии считаются хорошими при $q < AQL$ и плохими при $q > LQ$, где AQL - приёмочный уровень дефектности, и LQ - браковочный уровень дефектности. При $AQL < q < LQ$ качество партии считается ещё допустимым.

На рисунке 8.8 приведена форма документа «Планирование статистического приемочного контроля». Для получения результата необходимо указать объем партии, риски поставщика и потребителя, уровень контроля, тип плана и AQL , а затем нажать на кнопку «Рассчитать». В таблице будут выведены приемочные и браковочные числа для разных объемов выборки.

8.2.6.3 Определение правильности и прецизионности стандартного метода измерений

Для описания точности метода измерений используют два термина: «правильность» и «прецизионность». Термин «правильность» характеризует степень близости среднего арифметического значения большого числа результатов измерений к истинному или принятому опорному значению,

термин «прецизионность» – степень близости результатов измерений друг к другу.

Планирование статистического приемочного контроля 000...

Провести и закрыть | Провести | Печать | Все действия ▾

Задайте параметры:

Объем партии (N): Риск поставщика, %: Риск потребителя, %:

Уровень контроля: ...

Тип плана: ...

AQL: ...

Результат

Добавить | Все действия ▾

N	Объем выборки	Приемочное число	Браковочное число	AQL	LQ
1	5	1	5	1	3
2	10	3	8	2	6
3	15	6	10	4	9
4	20	9	12	7	13
5	25	12	13	9	16

Рассчитать | Печать

Рисунок 8.8 – Форма документа «Планирование статистического приемочного контроля»

Необходимость рассмотрения «прецизионности» возникает из-за того, что измерения, выполняемые на предположительно одинаковых материалах при предположительно одинаковых обстоятельствах, не дают, как правило, одинаковых результатов. Это объясняется неизбежными случайными погрешностями, присущими каждой измерительной процедуре, а факторы, оказывающие влияние на результат измерения, не поддаются полному контролю.

На изменчивость результатов измерений, выполненных по одному методу, помимо различий между предположительно идентичными образцами, могут влиять многие различные факторы, в том числе: оператор; используемое оборудование; калибровка оборудования; параметры окружающей среды (температура, влажность, загрязнение воздуха и т.д.); интервал времени между измерениями.

Прецизионность является общим термином для выражения изменчивости повторяющихся измерений. Два условия прецизионности, называемые условиями повторяемости и воспроизводимости, были признаны необходимыми и, во многих практических случаях, полезными для представления изменчивости метода измерений.

В условиях повторяемости (сходимости) факторы, перечисленные выше, считают постоянными, и они не влияют на изменчивость, в то время как в условиях воспроизводимости все эти факторы переменны и влияют на изменчивость результатов испытаний. Таким образом, повторяемость и воспроизводимость представляют собой два крайних случая прецизионности, где первый характеризует минимальную, а второй - максимальную изменчивость результатов.

Прочие промежуточные условия между этими двумя экстремальными условиями прецизионности допустимы, когда один или несколько факторов могут изменяться, и использоваться при определенных обстоятельствах. Прецизионность, как правило, выражают в терминах стандартных отклонений. Термин «стандартное отклонение» означает корень квадратный из дисперсии, определяемый в отечественных нормативных документах как «среднеквадратическое отклонение».

В данной работе было разработано программное обеспечение для определения повторяемости и воспроизводимости на основании стандарта ГОСТ Р ИСО 5725-2-2002 [22, 23, 136].

8.3 Выводы

Разработанные системы базируются на фундаментальных результатах, представленных в монографии [97]. Программное обеспечение может быть использовано в промышленности в отделах по контролю качества и метрологических лабораториях; в медицинских, научных и образовательных учреждениях [89, 93, 141].

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 $O(n^{-0,93})$. 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 МАХЗ 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 модели наследования. Показано, что критерий МАХЗ 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.

СПИСОК ЛИТЕРАТУРЫ

1. Айвазян С. А. Сравнение оптимальных свойств критериев Неймана-Пирсона и Вальда. — «Теория вероятностей и ее применения», т.IV (1959), №1.
2. Айвазян С. А. Различение близких гипотез о виде плотности распределения в схеме обобщенного последовательного критерия. — Теория вероятностей и ее применения, 1965, т. X, № 4.
3. Айвазян С. А., Енюков И. С., Мешалкин Л. Д. Прикладная статистика: Основы моделирования и первичная обработка данных. Справочное изд. / — М.: Финансы и статистика, 1983.
4. Акушерство. Справочник Калифорнийского университета / Под ред. К. Нисвандер, А. Эванс / перевод с англ. Н.А.Тимонина. Москва, Изд-во Практика, 1999. – 712 с.
5. Андрейченко А. В. Выбор интервалов группирования в критерии Рао-Робсона-Никулина при максимизации параметра нецентральности / А. В. Андрейченко, С. Н. Постовалов // Материалы IX международной конференции "Актуальные проблемы электронного приборостроения" АПЭП-2008, Новосибирск, 24-26 сентября 2008г. - Новосибирск: Изд-во НГТУ, 2008. - Т. 6. - С. 13-17.
6. Биндер К., Хеерман Д.В. Моделирование методом Монте-Карло в статистической физике. Пер. с англ. В. Н. Задкова. - М.: Наука. Физматлит, 1995. - 144 с. - (Компьютеры в физике).
7. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. – М.: Наука, 1983. – 416 с.
8. Вентцель Е.С. Исследование операций. М.: Советское радио, 1972.— 552 с.
9. Волкова Виктория Михайловна. Исследование распределений статистик дисперсионного анализа в условиях нарушения предположений

- нормальности: диссертация... кандидата технических наук: 05.13.17 Новосибирск, 2007. - 227 с.
10. Володин И. Н. О числе наблюдений, необходимых для различения двух близких гипотез // ТВП, 12:3 (1967), – С. 575–582.
 11. Володин И. Н. О числе наблюдений, необходимых для различения двух гипотез о параметре биномиального распределения // ТВП, 14:2 (1969), – С. 327–332
 12. Володин И. Н. Планирование эксперимента при сравнении параметров двух нормальных совокупностей // ТВП, 18:1 (1973), – С. 206–210
 13. Володин И. Н. Оптимальный объем выборки в процедурах статистического вывода // Изв. вузов. Матем., 1978, № 12, – С. 33–45
 14. Володин И. Н. Нижние границы для среднего объема выборки в критериях согласия и однородности // ТВП, 24:3 (1979), – С. 637–645
 15. Володин И. Н. Нижние границы для среднего объема выборки и эффективность процедур статистического вывода // ТВП, 24:1 (1979), 119–129
 16. Володин И. Н. Нижние границы для среднего объема выборки в критериях инвариантности // ТВП. – 1980, 25:2. С. 359–364.
 17. Володин И. Н. Асимптотика необходимого объема выборки при d -гарантийном различении двух близких гипотез / Володин И. Н., Новиков А. А. // Изв. вузов. Матем. – 1983, № 11. – С. 59–66.
 18. Володин И. Н. Гарантийные процедуры статистического вывода (определение объема выборки) // Исслед. по прикл. матем. – 1984. - №10. – С. 13–53.
 19. Гнеденко Б.В. Курс теории вероятностей. Учебник. 8-е изд., испр. и доп.— М.: Едиториал УРСС, 2005.— 448 с.
 20. ГОСТ Р ИСО 2859-1-2007 Статистические методы. Процедуры выборочного контроля по альтернативному признаку. Часть 1. Планы

- выборочного контроля последовательных партий на основе приемлемого уровня качества. – 2007. – 105 с.
- 21.ГОСТ Р 5479-2002. Проверка отклонения распределения вероятностей от нормального распределения – М. : Издательство стандартов, 2002. – 30 с.
- 22.ГОСТ Р ИСО 5725-1-2002. Точность (правильность и прецизионность) методов и результатов измерений. Часть 1. Основные положения и определения. – Введ. 2002–04–23. – М.: Издательство стандартов, 2002 . – 57 с.
- 23.ГОСТ Р ИСО 5725-2-2002. Точность (правильность и прецизионность) методов и результатов измерений. Часть 2. Основной метод определения повторяемости и воспроизводимости стандартного метода измерений. – Введ. 2002–04–23. – М. : Издательство стандартов, 2002 . – 108 с.
- 24.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.
- 25.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.
- 26.Гродзенская И. С. Метод "постепенного усечения" последовательной процедуры в случае распределения Рэля / И. С. Гродзенская, С. Я. Гродзенский. // Измерительная техника. - 2006. - N 11. - С. 15-16
- 27.Гродзенская, И. С. Применение последовательных критериев для проверки статистических гипотез при существенно неравноценных ошибках первого и второго рода / И. С. Гродзенская, С. Я. Гродзенский // Метрология. - 2006. - N 11. - С. . 3-10.

- 28.Гродзенская И. С. Поиск наиболее экономичного способа проверки статистических гипотез о параметрах двух нормальных распределений // Метрология. - 2006. - N 4. – С. 9-14
- 29.Гродзенский Я.С. Применение подхода Кифера-Вейсса для контроля технологических процессов // Метрология. 2009. № 12. С. 3-6.
- 30.Гродзенский С.Я., Гродзенский Я.С. Задача Кифера-Вейсса в науке и технике // Методы менеджмента качества. 2012. № 5. - С. 48-52.
- 31.Денисов В.И., Лемешко Б.Ю., Цой Е.Б. Оптимальное группирование, оценка параметров и планирование регрессионных экспериментов. В 2-х ч. / Новосиб. гос. техн. ун-т. - Новосибирск, 1993. - 347 с.
- 32.Денисов В.И., Лемешко Б.Ю., Постовалов С.Н. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Методические рекомендации. Часть I. Критерии типа Хи-квадрат. - Новосибирск: Изд-во НГТУ, 1998. - 126 с.
- 33.Джекел П. Применение метода Монте-Карло в финансах. М., 2004. – 256 с.
- 34.Джунгурова О. А., Володин И. Н. Асимптотика необходимого объема выборки при проверке гипотез о параметре формы распределения, близкого к нормальному // Изв. вузов. Матем., 2007, № 5. – С. 46–52
- 35.Ермаков С.М. Метод Монте-Карло и смежные вопросы. М.: ФИЗМАТЛИТ, 1975. - 471 с. - 2-е изд.
- 36.Ермаков С.М., Михайлов Г.А. Статистическое моделирование. М.: ФИЗМАТЛИТ, 1982. - 296 с. 2-е изд., дополн.
- 37.Жиенбаева А. А. Оценивание параметров и проверка гипотезы о согласии для закона распределения Ципфа / А. А. Жиенбаева, С. Н. Постовалов // Материалы Российской научно-технической конференции "Обработка информационных сигналов и математическое моделирование", Новосибирск, 23–24 мая 2013 г. – Новосибирск: СибГУТИ, 2013. – С. 50–52.

- 38.Зайцева Е.А. 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.
- 39.Ивченко Г.И., Медведев Ю.И. Математическая статистика. Учеб. пособие для вузов. - М.: Высш. шк.,1984. - 248 с.
- 40.Кнут Д.Э. Искусство программирования 3-е изд. — М.: Вильямс, 2000. — Т. 2. Получисленные алгоритмы. — 832 с.
- 41.Куллдорф Г. Введение в теорию оценивания по группированным и частично группированным выборкам. - М.: Наука, 1966. - 176 с.
- 42.Кульбак С. Теория информации и статистика. М.: Наука, 1967. - 408 с
- 43.Лемешко Б.Ю., Постовалов С.Н. Статистический анализ одномерных наблюдений по частично группированным данным // Изв. вузов. Физика. - Томск, 1995. - № 9. - С. 39-45.
- 44.Лемешко Б.Ю., Постовалов С.Н. К использованию непараметрических критериев по частично группированным данным // Сб. научных трудов НГТУ. - Новосибирск: Изд-во НГТУ, 1995. - № 2. - С. 21-30.
- 45.Лемешко Б.Ю., Постовалов С.Н. Проверка непараметрических гипотез по группированным данным // Материалы международной НТК "Информатика и проблемы телекоммуникаций". - Новосибирск, 1995. - Т.1. - С. 63-65.
- 46.Лемешко Б.Ю., Постовалов С.Н. Статистический анализ смесей распределений по частично группированным данным // Сб. научных трудов НГТУ. - Новосибирск: Изд-во НГТУ, 1995. - № 1. - С. 25-31.
- 47.Лемешко Б.Ю., Постовалов С.Н. Статистический анализ смесей распределений по группированным данным // Материалы международной НТК "Информатика и проблемы телекоммуникаций". - Новосибирск, 1995. - Т.1. - С. 83-85.

48. Лемешко Б.Ю. Статистический анализ одномерных наблюдений случайных величин: Программная система. – Новосибирск: Изд-во НГТУ, 1995. – 125 с.
49. Лемешко Б.Ю., Постовалов С.Н. Вопросы обработки выборок одномерных случайных величин // Научный вестник НГТУ. - Новосибирск, 1996. - № 2. - С. 3-24.
50. Лемешко Б.Ю., Постовалов С.Н. К вопросу о робастности оценок по группированным данным // Сб. научных трудов НГТУ. - Новосибирск: Изд-во НГТУ. - 1996. - № 2(4). - С. 9-18.
51. Лемешко Б.Ю., Постовалов С.Н. Программное обеспечение статистического анализа смесей случайных величин, представленных частично группированными и интервальными выборками // Тр. третьей МНТК "Актуальные проблемы электронного приборостроения АПЭП-96". - Новосибирск, 1996. - Т. 6. - Ч.1. - С.50-53.
52. Лемешко Б.Ю., Постовалов С.Н. Робастные алгоритмы оценивания и параметрические методы отбраковки аномальных наблюдений // Тр. третьей МНТК "Актуальные проблемы электронного приборостроения АПЭП-96". - Новосибирск, 1996. - Т. 6. - Ч.1. - С.45-49.
53. Лемешко Б.Ю., Постовалов С.Н. Система статистического анализа одномерных непрерывных распределений случайных величин (версия 3.0) // Мат. III международной НТК "Микропроцессорные системы автоматики", Новосибирск, 1996. - С. С-16 - С-17.
54. Лемешко Б.Ю., Постовалов С.Н. Статистический анализ наблюдений, имеющих интервальное представление // Сб. научных трудов НГТУ. - Новосибирск: Изд-во НГТУ, 1996. - № 1. - С. 3-12.
55. Лемешко Б.Ю., Постовалов С.Н. К вопросу о распределениях статистик непараметрических критериев согласия // Сб. научных трудов НГТУ. - Новосибирск: Изд-во НГТУ, 1997. - № 1(6). - С. 23-32.

56. Лемешко Б.Ю., Постовалов С.Н. Моделирование распределений статистик непараметрических критериев согласия при потере свойства "свободы от распределения" // Тр. международной НТК "Научные основы новых технологий". - Новосибирск, НГТУ, 1997. - Т.2. - С. 159-163.
57. Лемешко Б.Ю., Постовалов С.Н. О влиянии способа группирования данных на распределения статистик Хи квадрат Пирсона и отношения правдоподобия // Материалы международной НТК "Информатика и проблемы телекоммуникаций". - Новосибирск, 1997. - С. 120-123.
58. Лемешко Б.Ю., Постовалов С.Н. О распределениях статистик непараметрических критериев при потере свойства "свободы от распределения" // Мат. международной НТК "Информатика и проблемы телекоммуникаций". - Новосибирск, 1997. - С. 117-120.
59. Лемешко Б.Ю., Постовалов С.Н. О решении задач статистического анализа интервальных наблюдений // Вычислительные технологии. - 1997. - Т.2. - № 1. - С. 28-36.
60. Лемешко Б.Ю., Постовалов С.Н. Прикладные аспекты использования критериев согласия в случае проверки сложных гипотез // Надежность и контроль качества. - 1997. - № 11. - С. 3-17.
61. Лемешко Б.Ю., Постовалов С.Н. Программное обеспечение задач статистического анализа одномерных непрерывных случайных величин // Мат. международной научно-методической конференции "Новые информационные технологии в университетском образовании". - Новосибирск, НИИМИОО, 1997. - С. 44-45.
62. Лемешко Б. Ю. Асимптотически оптимальное группирование наблюдений в критериях согласия // Заводская лаборатория. 1998. Т. 64, № 1. С. 56–64.
63. Лемешко Б.Ю., Постовалов С.Н. Исследование допредельных распределений статистик критериев согласия при проверке сложных

- гипотез // Тр.международной конференции "Актуальные проблемы электронного приборостроения". Новосибирск, 1998. Т.3. - С. 12-16.
- 64.Лемешко Б.Ю., Постовалов С.Н. О зависимости предельных распределений статистик Хи-квадрат Пирсона и отношения правдоподобия от способа группирования данных // Заводская лаборатория. 1998. Т. 64. - № 5. - С. 56-63.
- 65.Лемешко Б.Ю., Постовалов С.Н. О распределениях статистик непараметрических критериев согласия при оценивании по выборкам параметров наблюдаемых законов // Заводская лаборатория. 1998. Т. 64. - № 3. - С. 61-72.
- 66.Лемешко Б.Ю., Постовалов С.Н. Об оценивании параметров распределений по интервальным наблюдениям // Вычислительные технологии. 1998. Т.3. - № 2. - С. 31-38.
- 67.Лемешко Б.Ю., Постовалов С.Н. Компьютерные методы моделирования и анализа статистических закономерностей // Материалы международной НТК "Информатика и проблемы телекоммуникаций". - Новосибирск, 1999. - С. 130.
- 68.Лемешко Б.Ю., Постовалов С.Н. О правилах проверки согласия опытного распределения с теоретическим // Методы менеджмента качества. Надежность и контроль качества. - 1999. № 11. - С. 34-43.
- 69.Лемешко Б.Ю., Постовалов С.Н. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Методические рекомендации. Часть II. Непараметрические критерии. - Новосибирск: Изд-во НГТУ, 1999. - 85 с.
- 70.Лемешко Б.Ю., Постовалов С.Н. Система статистического анализа наблюдений и исследования статистических закономерностей // Сб. "Моделирование, автоматизация и оптимизация наукоемких технологий". - Новосибирск: изд-во НГТУ, 2000. - С. 44-46.

71. Лемешко Б.Ю., Постовалов С.Н., Французов А.В. О нахождении параметра размытости непараметрических оценок функции плотности // Тр. V международной конференции "Актуальные проблемы электронного приборостроения" АПЭП-2000. Новосибирск, 2000. - Т. 6. - С. 17-20.
72. Лемешко Б.Ю., Постовалов С.Н. О зависимости распределений статистик непараметрических критериев и их мощности от метода оценивания параметров // Заводская лаборатория. Диагностика материалов. 2001. Т. 67. - № 7. - С. 62-71.
73. Лемешко Б.Ю., Чимитова Е.В. Построение оптимальных L-оценок параметров сдвига и масштаба распределений по выборочным квантилям // Сибирский журнал индустриальной математики. 2001. - Т.4. - № 2. - С. 166-183.
74. Лемешко Б.Ю., Гильдебрант С.Я., Постовалов С.Н. К оцениванию параметров надежности по цензурированным выборкам // Заводская лаборатория. Диагностика материалов. 2001. Т. 67. - № 1. - С. 52-64.
75. Лемешко Б.Ю., Постовалов С.Н. Оптимальные оценки параметров сдвига и масштаба по выборочным квантилям // Мат. межд. научн.-практ. конференции "САКС-2001". - Красноярск: САА. - Ч.2. 2001. - С.302-304.
76. Лемешко Б.Ю., Постовалов С.Н. Применение непараметрических критериев согласия при проверке сложных гипотез // Автометрия. 2001. - № 2. - С. 88-102.
77. Лемешко Б.Ю., Постовалов С.Н. Программное обеспечение статистического анализа одномерных наблюдений случайных величин // Мат. межд. научн.-практ. конференции "САКС-2001". - Красноярск: САА. - Ч.2. 2001. - С.293-295.
78. Лемешко Б.Ю., Постовалов С.Н. Система статистического анализа наблюдений и исследования статистических закономерностей // Материалы международной НТК "Информатика и проблемы телекоммуникаций". - Новосибирск, 2001. - С. 80-81.

79. Лемешко Б.Ю., Постовалов С.Н., Французов А.В. Исследование распределений статистики типа Колмогорова при использовании ядерных оценок // Материалы международной НТК "Информатика и проблемы телекоммуникаций". - Новосибирск, 2001. - С. 82.
80. Лемешко Б.Ю., Постовалов С.Н., Чимитова Е.В. О распределениях статистики и мощности критерия типа Хи-квадрат Никулина // Заводская лаборатория. Диагностика материалов. 2001. Т. 67. - № 3. - С. 52-58.
81. Лемешко Б.Ю., Постовалов С.Н. К использованию критериев согласия в случае проверки сложных гипотез // Мат. межд. научн.-практ. конференции "САКС-2001". - Красноярск: САА. - Ч.2. 2001. - С.305-307.
82. Лемешко Б.Ю., Постовалов С.Н. Непараметрические критерии при проверке сложных гипотез о согласии с распределениями Джонсона // Доклады СО АН ВШ. 2002. - № 1(5). - С.65-74.
83. Лемешко Б.Ю., Постовалов С.Н. Рекомендации ГОСТ Р 50.1-033-2001 и ГОСТ Р 50.1-037-2002 как итог исследований вопросов применения критериев согласия // Труды 10-го Юбилейного Международного Симпозиума по непараметрическим и робастным методам в кибернетике. Томск. 2002. - С.60-74.
84. Лемешко Б.Ю., Постовалов С.Н., Французов А.В. Вопрос проверки адекватности непараметрических моделей // Материалы международной НТК "Информатика и проблемы телекоммуникаций". - Новосибирск, 2002. - С. 123-124.
85. Лемешко Б.Ю., Постовалов С.Н., Французов А.В. К применению непараметрических критериев согласия для проверки адекватности непараметрических моделей // Автометрия. 2002. - № 2. - С.3-14.
86. Лемешко Б.Ю.. Один подход к моделированию псевдослучайных векторов с "заданными" числовыми характеристиками по законам, отличным от нормального / Лемешко Б.Ю., Помадин С.С // Материалы

- международной НТК "Информатика и проблемы телекоммуникаций". - Новосибирск, 2002. - С. 121-122.
87. Лемешко Б.Ю., Постовалов С.Н. Компьютерное моделирование как способ познания статистических закономерностей в технике, экономике, естествознании // Материалы региональной конференции (с участием иностранных ученых) "Вероятностные идеи в науке и философии". – Новосибирск: Ин-т философии и права СО РАН / НГУ. 2003. – С. 110-113.
88. Лемешко Б.Ю., Чимитова Е.В. О выборе числа интервалов в критериях согласия типа χ^2 // Заводская лаборатория. Диагностика материалов. 2003. – Т. 69. – № 1. – С. 61-67.
89. Лемешко Б.Ю. Компьютерные технологии анализа данных и исследования статистических закономерностей. / Постовалов С.Н., Лемешко Б.Ю // Учебное пособие. – Новосибирск: Изд-во НГТУ, 2004. 119 с.
90. Лемешко Б.Ю., Чимитова Е.В. Оптимальные L-оценки параметров сдвига и масштаба распределений по выборочным квантилям // Заводская лаборатория. Диагностика материалов. 2004. – Т.70. – №1. – С. 54-66.
91. 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.
92. Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н. Мощность критериев согласия при близких альтернативах // Измерительная техника. 2007. № 2. – С.22-27.
93. Лемешко Б.Ю., Постовалов С.Н., Лемешко С.Б. Компьютерные технологии анализа данных и исследования статистических закономерностей: Методические указания к выполнению лабораторных работ. – Новосибирск: Изд-во НГТУ, 2007. – 71 с.

94. Лемешко Станислав Борисович. Расширение прикладных возможностей некоторых классических методов математической статистики: диссертация ... кандидата технических наук: 05.13.17.- Новосибирск, 2007.- 305 с
95. 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.
96. Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н. Сравнительный анализ мощности критериев согласия при близких альтернативах. II. Проверка сложных гипотез // Сибирский журнал индустриальной математики. 2008. - Т.11. - № 4(36). - С.78-93.
97. 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").
98. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход / Б. Ю. Лемешко, С. Б. Лемешко, С. Н. Постовалов, Е. В. Чимитова // Научная сессия НИЯУ МИФИ-2011 (3-5 февраля 2011). Аннотации докладов. В 3 томах. Т 1. Инновационные ядерные технологии. М.: НИЯУ МИФИ, 2010. – С.223.
99. 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.

100. Михайлов Г. А., Войтишек А.В. Численное статистическое моделирование. Методы Монте-Карло. М.: Издательский центр «Академия», 2006. – 366 с.
101. Окурин Б. И. Методы отбраковки аномальных наблюдений / Б. И. Окурин, С. Н. Постовалов // Материалы Российской научно-технической конференции "Обработка информационных сигналов и математическое моделирование", Новосибирск, 23–24 мая 2013 г. – Новосибирск : СибГУТИ, 2013. – С. 81–83.
102. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука, 1979. – 296 с.
103. Орлов А.И. О влиянии погрешностей наблюдений на свойства статистических процедур (на примере гамма-распределения) // Статистические методы оценивания и проверки гипотез: Межвуз. сб. науч. тр. / Перм. ун-т. — Пермь, 1988. — С.45–55.
104. Орлов А.И. Некоторые алгоритмы реалистической статистики // Статистические методы оценивания и проверки гипотез: Межвуз. сб. науч. тр. / Перм. ун-т. — Пермь, 1991. — С.77–86.
105. Орлов А.И. Прикладная статистика. Учебник. - М.: Издательство "Экзамен", 2004. - 656 с.
106. Помадин Сергей Сергеевич. Исследование распределений статистик многомерного анализа данных при нарушении предположений о нормальности: Дис. ... канд. техн. наук : 05.13.17 : Новосибирск, 2004. - 139 с.
107. Постовалов С.Н., Танасейчук А.В. Об использовании критерия инверсий и критерия корреляции при проверке гипотезы случайности// Материалы VII Междунар. конф. "Актуальные проблемы электронного приборостроения" АПЭП–2004. – Новосибирск: Изд-во НГТУ, 2004. – Т. 6. – С. 315-318.

108. Постовалов С.Н., Танасейчук А.В. Построение критерия для проверки гипотезы случайности // Рос. науч.-техн. конф. "Информатика и проблемы телекоммуникаций", Новосибирск, 22 – 23 апр., 2004.: Материалы конф. – Новосибирск: Изд-во СибГУТИ, 2004. – Т. 1. – С. 134-136.
109. Постовалов С.Н., Курлаев С.А., Моделирование статистических закономерностей в задачах массового обслуживания // Материалы Российской НТК "Информатика и проблемы телекоммуникаций", Новосибирск. – 2005. – Т.1. – С. 132-133.
110. Постовалов С.Н., Танасейчук А.В. Ранговые критерии вычисления корреляционных зависимостей // Материалы Российской НТК "Информатика и проблемы телекоммуникаций", Новосибирск. – 2005. – Т.1. – С. 144-146.
111. Постовалов С.Н., Угрюмова М.В. Моделирование статистических закономерностей на параллельном компьютере // Материалы Российской НТК "Информатика и проблемы телекоммуникаций", Новосибирск. – 2005. – Т.1. – С. 124-125.
112. Постовалов С.Н. Сравнение мощности критериев согласия и критерия отношения правдоподобия // Материалы Российской НТК "Информатика и проблемы телекоммуникаций", Новосибирск. 2006. – Т.1. – С.190-193.
113. Постовалов С.Н., Иванова Е.В. О построении точных критических границ последовательного критерия отношения правдоподобия // Материалы Российской НТК "Информатика и проблемы телекоммуникаций", Новосибирск. 2006. – Т.1. – С.194-197.
114. Постовалов С.Н., Танасейчук А.В. Вычисление мощности критериев проверки гипотезы случайности против нескольких разнородных альтернатив // Материалы VIII международной

- конференции «Актуальные проблемы электронного приборостроения», Новосибирск. 2006. – Т.6. – С. 120-123.
115. Постовалов С.Н. Исследование распределений статистик критериев однородности по многократно цензурированным выборкам фиксированного объема / М. А. Ишалина, С. Н. Постовалов // Материалы X международной конференции "Актуальные проблемы электронного приборостроения" АПЭП-2010, Новосибирск, 22-24 сентября 2010г. - Новосибирск: Изд-во НГТУ, 2010. - Т. 6. - С. 62-67.
116. Постовалов С.Н. Исследование распределений статистик критериев однородности многократно цензурированных выборок / М. А. Ишалина, С. Н. Постовалов // Рос. науч.-техн. конф. "Информатика и проблемы телекоммуникаций", Новосибирск, 27 – 28 апр., 2010.: Материалы конф. – Новосибирск: Изд-во СибГУТИ, 2010. – Т. 1. – С.58-60
117. Постовалов С.Н. Исследования распределений статистик обобщенных критериев согласия типа хи-квадрат при многократном цензурировании / Е. С. Зыкина, Е. В. Севостьянова, С. Н. Постовалов // Материалы X международной конференции "Актуальные проблемы электронного приборостроения" АПЭП-2010, Новосибирск, 22-24 сентября 2010г. - Новосибирск: Изд-во НГТУ, 2010. - Т. 6. - С. 59-61.
118. Постовалов С.Н. О возможности применения критериев типа Хи-квадрат по многократно-цензурированным данным / Е. В. Севостьянова, С. Н. Постовалов // Рос. науч.-техн. конф. "Информатика и проблемы телекоммуникаций", Новосибирск, 27 – 28 апр., 2010.: Материалы конф. – Новосибирск: Изд-во СибГУТИ, 2010. – Т. 1. – С.81-84
119. Постовалов С.Н. Определение критических границ критерия Вальда при проверке сложной гипотезы о параметре сдвига нормального распределения / М. Ю. Ходырева, С. Н. Постовалов // Рос. науч.-техн. конф. "Информатика и проблемы телекоммуникаций", Новосибирск, 27 –

- 28 апр., 2010.: Материалы конф. – Новосибирск: Изд-во СибГУТИ, 2010. – Т. 1. – С.84-87
120. Постовалов С.Н. Применение последовательного критерия отношения правдоподобия по многократно цензурированным данным III типа / М. В. Масюта, С. Н. Постовалов // Материалы X международной конференции "Актуальные проблемы электронного приборостроения" АПЭП-2010, Новосибирск, 22-24 сентября 2010г. - Новосибирск: Изд-во НГТУ, 2010. - Т. 6. - С. 68-72.
121. Постовалов С.Н. Разработка подсистемы корреляционного анализа на платформе "1С:Предприятие 8.2" / С. В. Чебукина, С. Н. Постовалов // Материалы X международной конференции "Актуальные проблемы электронного приборостроения" АПЭП-2010, Новосибирск, 22-24 сентября 2010г. - Новосибирск: Изд-во НГТУ, 2010. - Т. 6. - С. 80-84.
122. Постовалов С. Н. Проверка простых и сложных гипотез с использованием последовательного критерия Вальда// ДОКЛАДЫ АН ВШ РФ. - 2011. - № 2(17). - С.140-150.
123. Постовалов С. Н. Проверка простых и сложных гипотез с использованием последовательных критериев Лордена и Айвазяна / С. Н. Постовалов, М. Р. Шахмаметова // Научный вестник НГТУ. - Новосибирск, 2011. - № 3 (44). - С. 17-28.
124. Постовалов С.Н. О возможности применения последовательного анализа по многократно цензурированным данным / М. В. Масюта, 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.102-106
125. 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
126. Постовалов С.Н. Исследование мощности критериев однородности по многократно цензурированным выборкам / М. А. Ишалина, 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.67-69
127. Постовалов С.Н. Исследование сходимости критерия экспоненциальности, основанного на эмпирическом преобразовании Лапласа, к предельному закону / А. И. Тимофеева, 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.95-96
128. Постовалов С.Н. Об использовании системы "1С:Предприятие 8.2" при проведении математических и статистических расчетов // 11-я международная научно-практическая конференция "Новые информационные технологии в образовании", Москва, 1-2 февраля, 2011: Сборник научных трудов. - М.: Изд-во 1С-Публишинг, 2011. - Ч.1. - С.523-525.
129. Постовалов С.Н. Определение критических границ последовательных критериев Лордена и Айвазяна / М. Р. Шахмаметова, 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.97-99.
130. Постовалов С. Н. Использование последовательных критериев проверки гипотез о виде распределения по полным и случайно цензурированным данным / С. Н. Постовалов, М. Р. Шахмаметова // Рос. науч.-техн. конф. "Обработка информационных сигналов и

- математическое моделирование", Новосибирск, 26 – 27 апр., 2012.:
Материалы конф. – Новосибирск: Изд-во СибГУТИ, 2012. – С.54-57
131. Постовалов С. Н. Поиск критических границ последовательного t -критерия Стьюдента / С. Н. Постовалов, Н. А. Гализин // Рос. науч.-техн. конф. "Обработка информационных сигналов и математическое моделирование", Новосибирск, 26 – 27 апр., 2012.: Материалы конф. – Новосибирск: Изд-во СибГУТИ, 2012. – С.21-23
132. 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
133. Постовалов С. Н. Планирование статистического приемочного контроля на основе приемлемого уровня качества AQL и LQ на платформе 1С:Предприятие 8.2 / С. Н. Постовалов, А. В. Поздеева // Рос. науч.-техн. конф. "Обработка информационных сигналов и математическое моделирование", Новосибирск, 26 – 27 апр., 2012.: Материалы конф. – Новосибирск: Изд-во СибГУТИ, 2012. – С.50-52
134. Постовалов С. Н. Исследование мощности обобщенного критерия типа Хи-квадрат при случайном цензурировании в случае простой гипотезы / С. Н. Постовалов, Е. С. Черемных // Рос. науч.-техн. конф. "Обработка информационных сигналов и математическое моделирование", Новосибирск, 26 – 27 апр., 2012.: Материалы конф. – Новосибирск: Изд-во СибГУТИ, 2012. – С.57-60
135. Постовалов С. Н. Программная поддержка ГОСТ Р 50.1.037, ГОСТ Р 50.1.033-2002 на платформе 1С:Предприятие 8.2 / С. Н. Постовалов, Д. Г. Демин // Рос. науч.-техн. конф. "Обработка информационных сигналов и математическое моделирование", Новосибирск, 26 – 27 апр., 2012.: Материалы конф. – Новосибирск: Изд-во СибГУТИ, 2012. – С.26-28

136. Постовалов С. Н. Программная поддержка ГОСТ Р ИСО 5725-1-2-2002 на платформе 1С:Предприятие 8.2 / С. Н. Постовалов, Б. И. Окурин // Рос. науч.-техн. конф. "Обработка информационных сигналов и математическое моделирование", Новосибирск, 26 – 27 апр., 2012.: Материалы конф. – Новосибирск: Изд-во СибГУТИ, 2012. – С.48-50
137. Постовалов С. Н. Разработка подсистемы проверки нормальности на платформе 1С:Предприятие 8.2 / С. Н. Постовалов, Т. В. Тимошенко // Рос. науч.-техн. конф. "Обработка информационных сигналов и математическое моделирование", Новосибирск, 26 – 27 апр., 2012.: Материалы конф. – Новосибирск: Изд-во СибГУТИ, 2012. – С.52-54.
138. Постовалов С. Н. Сравнение мощности критериев согласия и критерия отношения правдоподобия Неймана-Пирсона / С. Н. Постовалов, Е. А. Наумова // Рос. науч.-техн. конф. "Обработка информационных сигналов и математическое моделирование", Новосибирск, 26 – 27 апр., 2012.: Материалы конф. – Новосибирск: Изд-во СибГУТИ, 2012. – С.40-41
139. Постовалов С. Н. Сравнение мощности различных критериев при проверке гипотезы экспоненциальности / С. Н. Постовалов, А. И. Тимофеева // Рос. науч.-техн. конф. "Обработка информационных сигналов и математическое моделирование", Новосибирск, 26 – 27 апр., 2012.: Материалы конф. – Новосибирск: Изд-во СибГУТИ, 2012. – С.73-75
140. 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

141. Постовалов С. Н. Об использовании конфигурации "НКЦ ИТР:Статистика 1.0" в учебном процессе / С. Н. Постовалов // Сборник научных трудов 13-й международной научно-практической конференции "Новые информационные технологии в образовании" (Технологии 1С для эффективного обучения и подготовки кадров в целях повышения производительности труда) 29 - 30 января 2013 года. - Москва: 1С-Публишинг, 2013. том 2. - С. 702-706
142. 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
143. 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.
144. 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.
145. 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
146. 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
147. 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
148. 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
149. 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
150. Рогожников, Андрей Павлович. Исследование свойств некоторых критериев проверки статистических гипотез и обеспечение корректности их применения методами компьютерного моделирования: автореферат дис. ... кандидата технических наук : 05.13.17 / Рогожников Андрей Павлович; Новосибирск, 2012.- 24 с.:
151. Сархан А.Е., Гринберг Б.Г. Введение в теорию порядковых статистик. – М.: Статистика, 1970. – 414 с.

152. Смирнов Н.В. Оценка расхождения между эмпирическими кривыми распределения в двух независимых выборках // Бюллетень МГУ, серия А. – 1939. – Т.2. №2. – С.3-14.
153. Соболев И.М. Численные методы Монте-Карло. М: Наука, 1973, 312 с.
154. Тимошенко Т. В. Исследование критериев многомерной нормальности, основанных на эмпирической характеристической функции. / Т. В. Тимошенко, С. Н. Постовалов // Материалы Российской научно-технической конференции "Обработка информационных сигналов и математическое моделирование", Новосибирск, 23–24 мая 2013 г. – Новосибирск : СибГУТИ, 2013. – С. 102–103.
155. Федоров В.В. Теория оптимального эксперимента. М.: Наука, 1971. - 312 с.
156. Чимитова Екатерина Владимировна. Исследование методами компьютерного моделирования свойств оценок и статистик критериев согласия по группированным и цензурированным выборкам: диссертация ... кандидата технических наук : 05.13.17.- Новосибирск, 2003.- 162 с.: ил.
157. Шахов В. В. Обзор и сравнительный анализ библиотек генераторов псевдослучайных чисел // Пробл. информатики. 2010. № 2. С. 66-74.
158. Allen, D. L. Hypothesis testing using an L1-distance bootstrap // The American Statistician. – 1997. – Vol. 51. – P. 145–150.
159. Anderson, T.W. A modification of the sequential probability ratio test to reduce the sample size // Ann. Math. Statist. – 1960. – Vol. 31. – P.165–197.
160. Armitage, P. Test for linear trends in proportions and frequencies // Biometrics. – 1955. – Vol. 11. – P. 375-386.
161. Armitage, P. Sequential Medical Trials. 2nd edition. – Oxford: Blackwell, 1975.

162. Bagdonavičius, V. B., Levulienė, R. J., Nikulin, M. S. and Zdorova-Cheminade, O. Tests for equality of survival distributions against non-location alternatives // *Lifetime Data Analysis*. – 2004. – Vol. 10, №4. – P. 445-460.
163. Bahadur R. Rates of convergence of estimates and test statistics // *Ann. Math. Stat.* – 1967. Vol. 38, №2. – P. 303–324.
164. Bangdiwala Sh. I. Sequential and time-sequential procedures for generalized models: dis. PhD in Department of Biostatistics University of North Carolina at Chapel Hill / Bangdiwala Sh. I.; adv. Sen P.K. – North Carolina: Institute of Statistics Mimeo Series № 1322, 1980. – 165 p.
165. Baringhaus, L. Limit distributions for Mardia's measure of multivariate skewness. / Baringhaus, L. and Henze, N. // *Ann. Statist.* – 1992. – Vol. 20. – P. 1889-1902.
166. Berk R. Asymptotic efficiencies of sequential tests // *Ann. Statist.* – 1976. Vol. 4, №5. – P. 891–911.
167. Berk R. Sequential Bahadur efficiency / R. Berk, L. Brown // *Ann. Statist.* – 1978. – Vol. 6, №3. – P. 567–581.
168. Breslow N.E. Design and analysis of two-phase studies with binary outcome applied to Wilms tumor prognosis / Breslow N.E., Chatterjee N. // *JRSS C*. – 1999. Vol. 48. – P. 457–468.
169. Canner, P.L. Monitoring treatment differences in long-term clinical trials // *Biometrics*. – 1977. Vol. 33. – P. 603–615.
170. Conover, W.J., Iman, R.L. The rank transformation as a method of discrimination with some examples // *Communications in Statistics - Theory and Methods*. – 1980. – Vol. 9. P. 465-487.
171. Cox D.R. Note on grouping // *J. of the Amer. Statist. Ass.* 1957. – Vol. 52. – P. 543-547.
172. Denisov V.I. Software for statistical analysis of grouped data / Denisov V.I., Lemeshko B.Yu., Tsoi Y.B., Tishkovskaya S.V., Postovalov S.N. //

- Proceedings The First Korea-Russia International Symposium on Science and Technology (KORUS'97). - Republic of Korea, Ulsan, 1997. - P. 239-243.
173. Dunn, Ch.L. Critical values and powers for tests of uniformity of directions under multivariate normality // *Comm. Statist. - Th. Meth.* – 1995. Vol. 24. – P. 2541-2560.
174. Eguchi S. Interpreting Kullback-Leibler divergence with the Neyman-Pearson lemma / S. Eguchi and J. Copas // *J. Multivariate Analysis.* – 2006. Vol. 97. P. 2034-2040.
175. Elston R.C. Two-stage global search designs for linkage analysis: use of the mean statistic for affected sib pair / Elston R.C., Guo X., Williams L.W. // *Genet Epidemiol.* – 1996. – Vol. 18. – P. 97-110.
176. Enomoto R. On the distribution of test statistic using Srivastava's skewness and kurtosis. / Enomoto R, Okamoto N. and Seo T. [Электронный ресурс] // Hiroshima Statistical Research Group, Technical Report, 2011, TR–10-07. – Режим доступа: <http://www.math.sci.hiroshima-u.ac.jp/stat/TR/>
177. Farrell, P.J. On tests for multivariate normality and associated simulation studies / Farrell, P.J., Salibian-Barrera, M. and Naczki, K. // *Journal of Statistical Computation and Simulation* – 2007. – Vol. 77. P. 1065-1080.
178. Fisz, M. On a result by M. Rosenblatt concerning the Mises-Smirnov test // *Annals of Mathematical Statistics.* – 1960. – Vol. 31. – P. 427–429.
179. Freidlin B, Zheng G, Li Z, Gastwirth JL. Trend tests for case-control studies of genetic markers: power, sample size and robustness // *Human Heredity.* – 2002. – Vol. 53. – P. 146-152.
180. Gastwirth, J. L. The use of maximin efficiency robust tests in combining contingency tables and survival analysis // *Journal of the American Statistical Association.* – 1985. – Vol. 80, - P. 380–384.
181. Hammersley, J.M., Handscomb, D.C. Monte Carlo Methods. Methuen, London (Halsted, NY), 1964. - ISBN 0-416-52340-4.

182. Hara A. Normal Approximation to Multivariate Sample Measures of Kurtosis./ Hara A. and Seo T. [Электронный ресурс] // Hiroshima Statistical Research Group, Technical Report, 2011, TR-11-02. – Режим доступа: <http://www.math.sci.hiroshima-u.ac.jp/stat/TR/>
183. Harase S. Maximally equidistributed pseudorandom number generators via linear output transformations // Math. Comput. Simulation. – 2009. Vol 79, no. 5. P. 1512-1519.
184. Henze, N.. On Mardia's kurtosis test for multivariate normality. / Henze, N. // Comm. Statist. Theory Methods. 1994. – Vol. 23. - P. 1031 - 1045.
185. Henze N. Invariant tests for multivariate normality: a critical review. / Henze N. // Statistical Papers. – 2002. Vol. 43. - P. 468 – 506.
186. Hodges J. The efficiency of some nonparametric competitors of the t-test / J. Hodges, E. Lehmann // Ann. Math. Stat. – 1956. – Vol. 27, №2. – P. 324–335.
187. Hopkins, J.W., Clay, P.P.F. Some empirical distributions of bivariate T2 and homoscedasticity criterion M under unequal variance and leptokurtosis // Journal of the American Statistical Association. – 1963. – Vol. 58. – P. 1048-1053.
188. Hörmann, W. The generation of binomial random variates. // J. Stat. Comput. Simulation. – 1993. – Vol. 46, No.1-2. – P. 101-110.
189. Jacoboni C., Lugli P. The Monte Carlo Method for Semiconductor Device Simulation, Springer Verlag, New York, 1989. – 104 p.
190. Johnson A.D., O'Donnell C.J. An Open Access Database of Genome-wide Association Results // BMC Med. Genet. – 2009. Vol. 10, N6.
191. Joo J, Kwak M, Ahn K, Zheng G. A robust genome-wide scan statistic of the Wellcome Trust Case-Control Consortium // Biometrics. – 2009.
192. Kachitvichyanukul, V. and Schmeiser, B. W. Binomial random variate generation // Communications of the ACM. – 1988. – Vol. 31. – P. 216-222.

193. Kahn, H., Marshall, A.W. Methods of reducing sample size in Monte Carlo computations // *J.Oper.Res.Soc.Amer.* – 1953. Vol. 1. – P. 263–271.
194. Kiefer, J., Weiss, L. Some Properties of Generalized Sequential Probability Ratio Tests. // *The Annals of Mathematical Statistics.* – 1957. Vol 28. – P. 57-75.
195. Klar, B. Classical and new goodness-of-fit tests (in German). / Klar, B. // Doctoral dissertation, University of Karlsruhe, Germany, 1998.
196. Klein R.J. Complement Factor H Polymorphism in Age-Related Macular Degeneration / Klein R.J., Zeiss C., Chew E.Y., Tsai J.Y., Sackler R.S., Haynes C., Henning A.K., SanGiovanni J.P., Mane S.M., Mayne S.T., Bracken M.B., Ferris F.L., Ott J., Barnstable C., Hoh J. // *Science.* – 2005. - Vol. 308, no. 5720. - P. 385-389.
197. Koziol, J.A. A class of invariant procedures for assessing multivariate normality // *Biometrika.* – 1982. – Vol. 69. – P. 423-427.
198. Koziol, J.A. On assessing multivariate normality // *J. Roy. Statist. Soc. Set. B.* – 1983. – Vol. – 45. – P. 358-361.
199. Koziol, J.A. An alternative formulation of Neyman's smooth goodness of fit tests under composite alternatives. / J.A. Koziol // *Metrika.* - 1987. – Vol. 34. - P. 17-24.
200. Koziol, J.A. A note on measures of multivariate kurtosis. / J.A. Koziol // *Biometrical Journal.* – 1989. – Vol.31. – P. 619-624.
201. Kraft P. Efficient two-stage genome-wide association designs based on false positive report probabilities // *Pac Symp Biocomputing.* – 2006. – Vol. 11. – P. 523–534.
202. Kraft P., Cox D.G. Study designs for genome-wide association studies // *Adv Genet.* – 2008. – Vol. 60. – P. 465–504.
203. Kullback, S and Leibler, R. A. On information and sufficiency // *Ann. Math.Statist.* – 1951. – Vol. 55. – P. 79-86.

204. Kullback S. Information theory and statistics. John Wiley and Sons, NY, 1959.
205. Lehmann E.L. Consistency and unbiasedness of certain nonparametric tests // Ann. Math. Statist. – 1951. V.22. №1. – P.165-179.
206. Lemeshko B.Yu., Postovalov S.N. Statistical analysis of one-dimensional observations from partially grouped data // Russian Physics Journal (Historical Archive). Vol. 38, Number 9, September 1995. – P. 901 – 906 (Translated from Izvestiya Vysshikh Uchebnykh Zavedenii, Fizika, No. 9, pp. 39–45, September, 1995)
207. Lemeshko B.Yu., Postovalov S.N. Limit distributions of the Pearson chi² and likelihood ratio statistics and their dependence on the mode of data grouping // Industrial laboratory. 1998, vol. 64, no5, pp. 344-351. (Consultants Bureau, New York, NY, ETATS-UNIS)
208. Lemeshko B.Yu., Postovalov S.N. Statistical distributions of nonparametric goodness-of-fit tests as estimated by the sample parameters of experimentally observed laws // Industrial laboratory (Ind. lab.). 1998, vol. 64, no3, pp. 197-208 (Consultants Bureau, New York)
209. Lemeshko B.Yu., Postovalov S.N. Nonparametric goodness-of-fit tests when checking composite hypotheses // Proceedings The Third Russian-Korean International Symposium on Science and Technology (KORUS-99). Novosibirsk. Russia. June 22-25, 1999. Vol.2. - P.501-504.
210. Lemeshko B.Yu., Postovalov S.N. Application of the nonparametric goodness-of-fit Tests in testing composite hypotheses // Optoelectronics, Instrumentation and Data Processing. 2001. - № 2. - P. 76-88
211. Lemeshko B.Yu., Postovalov S.N., Frantsuzov A.V. Application of the nonparametric goodness-of-fit tests to testing nonparametric model adequacy // Optoelectronics, Instrumentation and Data Processing. 2002. - № 2. - P. 3-12.
212. Lemeshko B.Yu., Postovalov S.N., Chimitova E.V. Rules of application of goodness-of-fit tests in simple and composite hypothesis testing // The 7th

- Korea-Russia International Symposium on Science and Technology (KORUS 2003). University of Ulsan. June 28-July 6, 2003. – Vol.3. – P.126-132.
213. Lemeshko B.Yu., Postovalov S.N., Chimitova E.V. Numerical Research Of The Theoretical Recommendations Of Mathematical Statistics In Non-Standard Conditions // Proceedings of the Seventh International Conference “Computer Data Analysis and Modeling: Robustness and Computer Intensive Methods”, September 6-10, 2004, Minsk. Vol. 1. – P. 94-97.
214. Lemeshko B.Yu., Postovalov S.N., Chimitova E.V., Pomadin S.S., Ponomarenko V.M., Frantsuzov A.V., Mirkin E.P., Lemeshko S.B. Computer simulation technique on the investigation of statistical regularities // Proceedings of 8th Korea-Russia International Symposium on Science and Technology (KORUS 2004). Tomsk. Russia. June 26-July 3, 2004. – Vol.2. – P.149-152.
215. Lemeshko B.Yu., Lemeshko S.B., Postovalov S.N. The power of goodness of fit tests for close alternatives // Measurement Techniques, 2007. V.50, № 2. – P. 132-141.
216. Lemeshko B.Yu. Computer methods for investigating statistical regularities in problems of statistical data analysis and reliability / B.Yu. Lemeshko, S.B. Lemeshko, S.N. Postovalov, E.V. Chimitova // MMR 2009 - Mathematical Methods in Reliability. Theory. Methods. Applications. VI International Conference. Extended Abstracts. Moskow, 22-29 June, 2009. – P.418-422.
217. Lemeshko B.Yu., Lemeshko S.B., and Postovalov S.N. Comparative Analysis of the Power of Goodness-of-Fit Tests for Near Competing Hypotheses. I. The Verification of Simple Hypotheses // Journal of Applied and Industrial Mathematics, 2009, Vol. 3, No. 4, P. 462–475.
218. Lemeshko B.Yu. Distributed computing system for simulation of classical test statistic distributions under nonstandard conditions / E.V. Chimitova, B.Yu. Lemeshko, S.B. Lemeshko, S.N. Postovalov, A.P.

- Rogozhnikov // Proceedings Third International Conference on Accelerated Life Testing, Reliability-based Analysis and Design. 19-21 May 2010, Clermont-Ferrand, France. – P.107-109
219. Lemeshko B.Yu. Lemeshko S.B. and Postovalov S.N. Comparative analysis of the power of goodness-of-fit tests for near competing hypotheses. II. Verification of complex hypotheses // Journal of Applied and Industrial Mathematics, 2010, Vol. 4, No. 1, – P. 79–93.
220. Lemeshko B.Yu., Lemeshko S.B. and Postovalov S.N. Statistic Distribution Models for Some Nonparametric Goodness-of-Fit Tests in Testing Composite Hypotheses // Communications in Statistics - Theory and Methods, 2010. Vol. 39, No. 3. – P. 460-471.
221. 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.
222. Lorden G. 2-SPRT's and the modified Kiefer-Weiss problem on minimizing an expected sample size // Annals of Statistics. – 1976. – Vol. 4. – P. 281-291
223. Mardia, K.V.. Measures of multivariate skewness and kurtosis with applications. / K.V.Mardia // Biometrika. – 1970. – Vol. 57. – P. 519-530.
224. Mardia, K.V.. Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. / Mardia, K.V. // Sankhya: The Indian Journal of Statistics. – 1974. – Vol. 36. – P. 115-128.
225. Mardia, K.V. Assessment of multinormality and the robustness of Hotelling's T2 test // Applied Statistics. – 1975. – Vol. 24. – P. 163-171.

226. Martinez R.-L. M. C., Naranjo J. D. A pretest for choosing between logrank and wilcoxon tests in the two-sample problem // International Journal of Statistics. – 2010. - Vol. LXVIII, n. 2. – P. 111 – 125.
227. Matsumoto, M.; Nishimura, T. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator // ACM Transactions on Modeling and Computer Simulation. – 1998. – Vol. 8 (1). – P. 3–30.
228. Metropolis N. The Monte Carlo Method / N. Metropolis, S. Ulam // J. Amer. statistical assoc. – 1949. – Vol. 44. - № 247. – P. 335—341.
229. Metropolis, N. The Beginning of the Monte Carlo Method // Los Alamos Science Special Issue. – 1987, No. 15. - P. 125
230. Moore, D., and Stubblebine, J. Chi-square tests for multivariate normality with applications to common stock prices // Comm. Statist. - Th. Meth. – 1981. – Vol. 10. – P. 713-738.
231. Mori, T.F. On multivariate skewness and kurtosis. / T.F. Mori, V.K. Rohatgi, and G.J. Szekely // Theory Probab. Appl – 1993. – Vol. 38. - P. 547-551.
232. Mudholkar, G.S., McDermott, M., and Srivastava, D.K. A test of p-variate normality // Biometrika. – 1992. – Vol. 79. - P. 850-854.
233. Muller, H.-H., Pahl, R. & Schafer, H. Including sampling and phenotyping costs into the optimization of two stage designs for genome wide association studies // Genet Epidemiol. – 2007. – Vol. 31. – P. 844–852.
234. Neumann J. Various techniques used in connection with random digits in A.S. Householder, G.E. Forsythe, and H.H. Germond, eds., Monte Carlo Method, National Bureau of Standards Applied Mathematics Series, vol. 12. Washington, D.C.: U.S. Government Printing Office. – 1951. P. 36-38.
235. Neyman J. On the Problem of the Most Efficient Tests of Statistical Hypotheses" /Jerzy Neyman, Egon Pearson // Philosophical Transactions of the

- Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character. – 1933. – Vol. 231. – P. 289–337.
236. Nguyen T.T., Pahl R., Schäfer H. Optimal robust two-stage designs for genome-wide association studies // *Ann Hum Genet.* – 2009. – Vol. 73 (Pt 6). – P. 638-51.
237. Ogawa J. Contributions to the theory of systematic statistics. // *I. Osaka Math. J.* – 1951. – Vol. 3. – P. 175-213.
238. Okamoto, N. On the distribution of multivariate sample skewness for assessing multivariate normality. / Okamoto, N. and Seo, T. [Электронный ресурс] // Hiroshima Statistical Research Group, Technical Report, 2008, TR-08-01. – Режим доступа: <http://www.math.sci.hiroshima-u.ac.jp/stat/TR/>
239. Panneton, F. Improved Long-Period Generators Based on Linear Recurrences Modulo 2 / Panneton, F, L'Ecuyer, P, Matsumoto // *ACM TRANSACTIONS ON MATHEMATICAL SOFTWARE.* – 2006. – Vol. 32, Issue: 1. P. 1-16.
240. Paulson, A.S., Roohan, P., and Sullo, P. Some empirical distribution function tests for multivariate normality // *J. Statist. Comput. Simul.* – 1987. – Vol. 28. – P. 15-30.
241. Pettitt, A.N. A two-sample Anderson-Darling rank statistic // *Biometrika.* – 1976. – Vol. 63. – P. 161-168.
242. Philonenko P. A power comparison of homogeneity tests for randomly censored data / P. Philonenko, S. Postovalov // *Applied Methods of Statistical Analysis. Simulations and Statistical Inference (AMSA 2013) International Conference. Proceedings.* – 2013. Novosibirsk, Russia – P. 227-237
243. Postovalov S.N. Convergence of two-sample test statistic distributions to the limiting law / M.A. Ishalina, S.N. Postovalov // *Proceedings Third International Conference on Accelerated Life Testing, Reliability-based Analysis and Design.* 19-21 May 2010, Clermont-Ferrand, France. – P.237-242.

244. Postovalov S. Reduction of the Average Sample Number in Sequential Scheme of Testing Hypotheses / S. Postovalov, M. Shakhmametova // Applied Methods of Statistical Analysis. Simulations and Statistical Inference (AMSA 2011) International Conference. Proceedings. 20-22 September, 2011. Novosibirsk, Russia – P. 158-166
245. Postovalov S., Shakhmametova M. Sequential probability ratio tests (2-SPRT and Ayvazyan's test) for simple and composite hypotheses // XIVth Applied Stochastic Models and Data Analysis (ASMDA 2011) International Conference. Proceedings. June 6 – June 11, 2011. Rome, Italy. – P. 1118-1125.
246. 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
247. Quiroz, A.J., and Trabucco, J.C. Kolmogorov-Smirnov type statistics for testing multivariate normality // Publ. Mat. Urug. – 1998. – Vol. 7. - P. 67-82.
248. Rao C.R. Linear statistical inference and its applications, Wiley, 1965.
249. Rayleigh, L. On the problems of random vibrations and of random flights in one, two or three dimensions // Phil. Mag. -1919. – Vol. 37. – P. 321-347.
250. Romeu, J.L., and Ozturk, A. A comparative study of goodness-of-fit tests for multivariate normality // J. Multiv. Anal. – 1993. – Vol. 46. – P. 309-334.
251. Rosenblatt M. Limit theorems associated with variants of the von Mises statistic // Ann. Math. Statist. – 1952. V.23. – P.617-623.
252. Ross S.M., Lin K. Applying Variance Reduction Ideas in Queuing Simulation // Probability Eng. Informational Sci. – 2001. - Vol. 15. – 481-494.
253. Ross S.M. Simulation, fourth edition, Academic Press, 2006.

254. Saito A, Kamatani N. Strategies for genome-wide association studies: optimization of study designs by the stepwise focusing method. // *J Hum Genet.* – 2002. – Vol. 47. – P. 360–365.
255. Sasieni P.D. From genotypes to genes: doubling the sample size // *Biometrics.* – 1997. Vol. 53. – P. 1253-1261.
256. Satagopan J.M., Verbel D.A., Venkatraman E.S., Offit K.E., Begg C.B. Two-stage designs for gene-disease association studies // *Biometrics.* – 2002. – Vol. 58. – P. 163-170.
257. Satagopan J.M., Elston R.C. Optimal two-stage genotyping in population-based association studies // *Genet Epidemiol.* – 2003. – Vol. 25. - P. 149-157.
258. Satagopan J.M., Venkatraman E.S., Begg C.B. Two-stage designs for gene-disease association studies with sample size constraints // *Biometrics.* 2004 - . Vol. 60. – P. 589-597.
259. Scholz F. W. K-Sample Anderson-Darling Tests / F. W. Scholz, M. A. Stephens // *Journal of the American Statistical Association*ю – 1987. - Vol. 82, No. 399. - P. 918-924.
260. Seo, T. On the distribution of kurtosis test for multivariate normality. / Seo, T. and Ariga, M. [Электронный ресурс] // Hiroshima Statistical Research Group, Technical Report, 2006, TR-06-03. – Режим доступа. <http://www.math.sci.hiroshima-u.ac.jp/stat/TR/>
261. Seo, T. On the distribution of kurtosis test for multivariate normality. / Seo, T. and Ariga, M. [Электронный ресурс] // Hiroshima Statistical Research Group, Technical Report, 2009. – Режим доступа <http://www.math.sci.hiroshima-u.ac.jp/stat/TR/>
262. Service S.K., Sandkuijl L.A., Freimer N.B. Cost-effective designs for linkage disequilibrium mapping of complex traits // *Am J Hum Genet.* - 2003. – Vol. 72. P. 1213–1220.

263. Siegmund D. Importance Sampling in the Monte Carlo Study of Sequential Tests // *Ann. Statistics*. – 1976. – Vol. 4. – P. 673-684.
264. Simons G. An Improved Statement of Optimality for Sequential Probability Ratio Tests // *The Annals of Statistics*. - 1976. - Vol. 4, No. 6. – P. 1240-1243.
265. Skol, A. D., Scott, L. J., Abecasis, G. R. & Boehnke, M. Optimal designs for two-stage genome-wide association studies // *Genet Epidemiol*. – 2007. – Vol. 31. – P. 776–788.
266. Slager S.L., Schaid D.J. Case-control studies of genetic markers: Power and sample size approximations for Armitage’s test for trend // *Hum Hered*. – 2001. – Vol. 52. – P. 149–153.
267. Song K., Elston R.C. A powerful method of combining measures of association and Hardy-Weinberg disequilibrium for fine mapping in case-control studies // *Stat Med*. – 2006. – Vol. 25. – P. 105–126.
268. Srivastava, M. S. A measure of skewness and kurtosis and a graphical method for assessing multivariate normality / Srivastava, M. S. // *Statistics & Probability Letters*. – 1984. – Vol. 2. – P. 263–267.
269. Stevens, S. S. On the Theory of Scales of Measurement // *Science*. – 1946. – Vol. 103 (2684). – P. 677–680.
270. Tarone R.E., Gart J.J. On the robustness of combined tests for trends in proportions // *J Am Stat Assoc*. – 1980. – Vol. 75. – P. 110–116.
271. Wald, A. Statistical decision functions which minimize the maximum risk // *The Annals of Mathematics*. – 1945. – Vol. 46(2). – P. 265-280.
272. Abraham Wald, *Sequential Tests of Statistical Hypotheses* // *Annals of Mathematical Statistics*. – 1945. Vol. 16, no 2. - P. 117—186.
273. Abraham Wald, *Sequential Analysis*, New York, John Wiley & Sons, 1947.

274. A. Wald and J. Wolfowitz Optimum Character of the Sequential Probability Ratio Test // *The Annals of Mathematical Statistics*. – 1948. – Vol. 19, No. 3. – P. 326-339.
275. Wang K, Sheffield V.C. A constrained-likelihood approach to marker-trait association studies // *American Journal of Human Genetics*. – 2005. – Vol. 77. – P.768-780.
276. Wang, H., Thomas, D. C., Pe'er, I. & Stram, D. O. Optimal two-stage genotyping designs for genome-wide association scans // *Genet Epidemiol*. – 2006. – 30. – P. 356–368.
277. White J.E. A two stage design for the study of the relationship between a rare exposure and a rare disease // *Am J Epidemiol*. – 1982. – Vol. 115. – P. 119–128.
278. The Welcome Trust Case Control Consortium. Genome-wide association study of 14 000 cases of seven common diseases and 3,000 shared controls // *Nature*. – 2007. Vol. 447. – P. 661-683.
279. Zang Y., Fung W.K. and Zheng G. Simple algorithms to calculate asymptotic null distribution for robust tests in case-control genetic association studies in R // *Journal of Statistical Software*. - . Vol. 33, issue 8. – P.1-24.
280. Zheng, G. & Gastwirth, J.L. On estimation of the variance in Cochran-Armitage trend tests for genetic association using case-control studies // *Statist Med*. – 2006. – Vol. 25. – P. 3150-3159.
281. Zheng G., Ng HKT. Genetic model selection in two-stage analysis for case-control association studies // *Biostatistics*. – 2008. – Vol. 9. – P. 391-399

ПРИЛОЖЕНИЕ 1. АКТ О ВНЕДРЕНИИ В УЧЕБНЫЙ ПРОЦЕСС

“УТВЕРЖДАЮ”

Проректор по учебной работе НГТУ

д.т.н., профессор

“ ” 2012 г.**АКТ****об использовании результатов исследований Постовалова С.Н. в учебном процессе факультета прикладной математики и информатики**

Результаты диссертационных исследований докторанта кафедры прикладной математики Постовалова Сергея Николаевича, полученные при выполнении работ в рамках государственного контракта № П950 от 20 августа 2009 года, используются в учебном процессе на факультете прикладной математики и информатики при выполнении курсовых проектов и лабораторных работ по курсу “Компьютерные технологии анализа данных и исследования статистических закономерностей”, читаемому для магистрантов направления 010400.68 – Прикладная математика и информатика.

Декан ФПМИ
д.т.н., профессор

Лемешко Б.Ю.

**ПРИЛОЖЕНИЕ 2. СВИДЕТЕЛЬСТВА О РЕГИСТРАЦИИ ПРОГРАММЫ
ДЛЯ ЭВМ «ИНТЕРВАЛЬНАЯ СТАТИСТИКА» ISW**

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

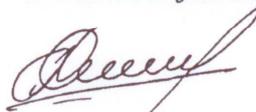
о государственной регистрации программы для ЭВМ
№ 2012613663

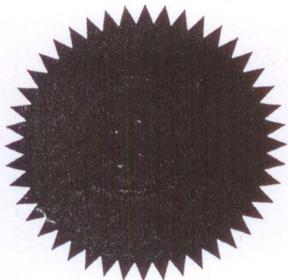
**«Статистический анализ интервальных
наблюдений одномерных непрерывных случайных
величин "Интервальная статистика 4.0"»**

Правообладатель(ли): **Федеральное государственное бюджетное
образовательное учреждение высшего профессионального
образования «Новосибирский государственный
технический университет» (RU)**

Автор(ы): **Лемешко Борис Юрьевич,
Постовалов Сергей Николаевич (RU)**

Заявка № **2011619739**
Дата поступления **19 декабря 2011 г.**
Зарегистрировано в Реестре программ для ЭВМ
19 апреля 2012 г.

Руководитель Федеральной службы
по интеллектуальной собственности

Б.П. Симонов



РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2012613664

«Статистический анализ интервальных наблюдений одномерных непрерывных случайных величин "Интервальная статистика 5.0"»

Правообладатель(ли): *Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Новосибирский государственный технический университет» (RU)*

Автор(ы): *Лемешко Борис Юрьевич, Постовалов Сергей Николаевич, Лемешко Станислав Борисович, Чимитова Екатерина Владимировна, Рогожников Андрей Павлович, Щеглов Алексей Евгеньевич, Горбунова Алиса Александровна (RU)*

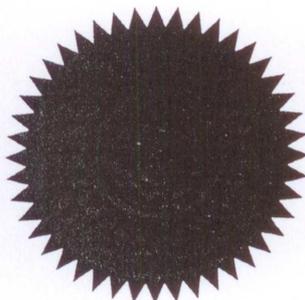
Заявка № 2011619740

Дата поступления 19 декабря 2011 г.

Зарегистрировано в Реестре программ для ЭВМ
19 апреля 2012 г.

Руководитель Федеральной службы
по интеллектуальной собственности

Б.П. Симонов



РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2013615968

Статистический анализ интервальных наблюдений
одномерных непрерывных случайных величин
"Интервальная статистика 5.1"

Правообладатель: *Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Новосибирский государственный технический университет» (RU)*

Авторы: *см. на обороте*

Заявка № 2013612140

Дата поступления 21 марта 2013 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 25 июня 2013 г.

Руководитель Федеральной службы
по интеллектуальной собственности

Б.П. Симонов



**ПРИЛОЖЕНИЕ 3. СВИДЕТЕЛЬСТВА О РЕГИСТРАЦИИ ПРОГРАММЫ
ДЛЯ ЭВМ «НКЦ ИТР: СТАТИСТИКА 1.0»**

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО
о государственной регистрации программы для ЭВМ
№ 2009614173

**Набор классов для статистического анализа интервальных
наблюдений одномерных непрерывных случайных величин**

Правообладатель(ли): **Общество с ограниченной ответственностью
«Научно-консультационный центр инженерно-технических
решений» (RU)**

Автор(ы): **Постовалов Сергей Николаевич (RU)**

Заявка № **2009613046**
Дата поступления **16 июня 2009 г.**
Зарегистрировано в Реестре программ для ЭВМ
10 августа 2009 г.

Руководитель Федеральной службы по интеллектуальной
собственности, патентам и товарным знакам


 **Б.П. Симонов**

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2012610267

Программное обеспечение статистического моделирования
в задачах проведения и обработки измерений
«НКЦ ИТР: Статистика 1.0»

Правообладатель(ли): *Общество с ограниченной ответственностью
«Научно-консультационный центр инженерно-технических
решений» (RU)*

Автор(ы): *Постовалов Сергей Николаевич,
Макеева Полина Марковна, Постовалова Анастасия Юрьевна,
Шиловский Максим Викторович, Чебукина Светлана Вячеславовна,
Каньшин Михаил Викторович, Макеев Александр Геннадьевич,
Косарева Лариса Александровна (RU)*

Заявка № 2011618443

Дата поступления 8 ноября 2011 г.

Зарегистрировано в Реестре программ для ЭВМ
10 января 2012 г.



Руководитель Федеральной службы по интеллектуальной
собственности, патентам и товарным знакам

Б.П. Симонов



СЕРТИФИКАТ

ПРОГРАММНЫЙ ПРОДУКТ

«НКЦ ИТР:Статистика 1.0»,

разработанный фирмой ООО "Научно-консультационный центр инженерно-технических решений",
признан совместимым с системой программ «1С:Предприятие 8.2»
как конфигурация в режиме Управляемого приложения.

Фирма – разработчик получает право использовать логотип



как элемент оформления упаковки и рекламных материалов
программного продукта.

Директор фирмы «1С»



Нуралиев Б.Г./

13 июля 2012 г.

срок действия сертификата 2 года

ПРИЛОЖЕНИЕ 4. ГОСТ Р 50.1.033-2001 (ФРАГМЕНТ)

Р 50.1.033—2001

РЕКОМЕНДАЦИИ ПО СТАНДАРТИЗАЦИИ

Прикладная статистика

**ПРАВИЛА ПРОВЕРКИ СОГЛАСИЯ
ОПЫТНОГО РАСПРЕДЕЛЕНИЯ
С ТЕОРЕТИЧЕСКИМ**

Часть 1

Критерии типа хи-квадрат

Издание официальное

БЗ 2—2001/4

191.00
РУБ

ГОССТАНДАРТ РОССИИ
Москва

Р 50.1.033—2001

Предисловие

1 РАЗРАБОТАНЫ Новосибирским государственным техническим университетом, доработаны с участием Технического комитета по стандартизации ТК 125 «Стандартизация статистических методов управления качеством»

ВНЕСЕНЫ Техническим комитетом по стандартизации ТК 125 «Стандартизация статистических методов управления качеством»

2 ПРИНЯТЫ И ВВЕДЕНЫ В ДЕЙСТВИЕ Постановлением Госстандарта России от 14 декабря 2001 г. № 525-ст

3 ВВЕДЕНЫ ВПЕРВЫЕ

© ИПК Издательство стандартов, 2002

Настоящие рекомендации не могут быть полностью или частично воспроизведены, тиражированы и распространены в качестве официального издания без разрешения Госстандарта России

II

Содержание

1 Область применения	1
2 Теоретические основы рекомендаций	1
2.1 Общие положения	1
2.2 Критерии типа χ^2 при простых гипотезах	3
2.3 Критерии типа χ^2 при сложных гипотезах	4
2.4 Статистика типа χ^2 Никулина	5
2.5 Связь мощности критериев со способом группирования наблюдений	6
2.6 Асимптотически оптимальное группирование	7
2.7 Характер влияния способов группирования и метода оценивания на распределения статистик типа χ^2	8
2.8 Выбор числа интервалов	14
2.9 Рекомендации по использованию асимптотически оптимального группирования в критериях согласия	19
2.10 Примеры использования таблиц асимптотически оптимального группирования	39
3 Порядок применения критериев типа χ^2	45
3.1 Порядок проверки простой гипотезы	45
3.2 Порядок проверки сложной гипотезы	46
Приложение А Таблицы асимптотически оптимального группирования наблюдений в критериях согласия типа χ^2	47
Приложение Б Процентные точки χ^2 -распределений	85
Приложение В Библиография	86

ПРИЛОЖЕНИЕ 5. ГОСТ Р 50.1.037-2002 (ФРАГМЕНТ)

Р 50.1.037—2002

РЕКОМЕНДАЦИИ ПО СТАНДАРТИЗАЦИИ

Прикладная статистика

**ПРАВИЛА ПРОВЕРКИ СОГЛАСИЯ
ОПЫТНОГО РАСПРЕДЕЛЕНИЯ
С ТЕОРЕТИЧЕСКИМ**

Часть II

Непараметрические критерии

Издание официальное

ГОССТАНДАРТ РОССИИ
Москва

193-

40

Р 50.1.037—2002

Предисловие

1 РАЗРАБОТАНЫ Новосибирским государственным техническим университетом, доработаны с участием Технического комитета по стандартизации ТК 125 «Стандартизация статистических методов управления качеством»

ВНЕСЕНЫ Техническим комитетом по стандартизации ТК 125 «Стандартизация статистических методов управления качеством»

2 ПРИНЯТЫ И ВВЕДЕНЫ В ДЕЙСТВИЕ Постановлением Госстандарта России от 22 января 2002 г. № 24-ст

3 ВВЕДЕНЫ ВПЕРВЫЕ

© ИПК Издательство стандартов, 2002

Настоящие рекомендации не могут быть полностью или частично воспроизведены, тиражированы и распространены в качестве официального издания без разрешения Госстандарта России

II

Содержание

1 Область применения	1
2 Общие положения	1
2.1 Простые и сложные гипотезы при проверке согласия опытного распределения с теоретическим	1
2.2 Распределения статистик непараметрических критериев согласия при простых гипотезах	2
2.2.1 Критерий Колмогорова	2
2.2.2 Критерий Смирнова	3
2.2.3 Критерии ω^2	3
2.3 Непараметрические критерии согласия при сложных гипотезах	4
2.3.1 Потеря критериями свойства «свободы от распределения»	4
2.3.2 Методика компьютерного анализа статистических закономерностей	5
2.3.3 Факторы, влияющие на распределения статистик критериев при проверке сложных гипотез	6
2.3.4 Влияние объема выборки на распределения статистик непараметрических критериев при простых и сложных гипотезах	6
2.3.5 Влияние объема выборки на мощность непараметрических критериев при простых и сложных гипотезах	8
2.3.6 Влияние метода оценивания на распределения статистик непараметрических критериев при сложных гипотезах	11
2.3.7 Метод оценивания и мощность непараметрических критериев согласия	17
2.3.8 Зависимость распределений статистик непараметрических критериев от конкретных значений параметра	20
2.3.9 Выводы	21
3 Порядок проверки гипотез о согласии	22
3.1 Порядок проверки простой гипотезы о согласии	22
3.1.1 Критерий Колмогорова при простой гипотезе	23
3.1.2 Критерий Смирнова при простой гипотезе	23
3.1.3 Критерии ω^2 Крамера — Мизеса — Смирнова при простой гипотезе	23
3.1.4 Критерий Ω^2 Андерсона — Дарлинга при простой гипотезе	23
3.2 Порядок проверки сложной гипотезы	23
3.2.1 Проверка сложной гипотезы о согласии по критерию типа Колмогорова	24
3.2.2 Проверка сложной гипотезы о согласии по критерию типа Смирнова	24
3.2.3 Проверка сложной гипотезы о согласии по критерию типа ω^2 Мизеса	25
3.2.4 Проверка сложной гипотезы о согласии по критерию типа Ω^2 Мизеса	25
3.2.5 Проверка сложных гипотез о согласии с гамма-распределением	25
3.2.5.1 Проверка сложной гипотезы о согласии с гамма-распределением по критерию типа Колмогорова	25
3.2.5.2 Проверка сложной гипотезы о согласии с гамма-распределением по критерию типа Смирнова	26
3.2.5.3 Проверка сложной гипотезы о согласии с гамма-распределением по критерию типа ω^2 Мизеса	26
3.2.5.4 Проверка сложной гипотезы о согласии с гамма-распределением по критерию типа Ω^2 Мизеса	26
3.2.6 Проверка сложных гипотез о согласии с распределениями Джонсона	27
3.2.7 Перечень распределений, для которых регламентирована проверка сложных гипотез с использованием настоящих рекомендаций	27
3.2.8 Законы распределения, используемые для аппроксимации предельных распределений статистик непараметрических критериев при проверке сложных гипотез	28
3.2.9 Примеры применения критериев согласия при простых и сложных гипотезах	29
Приложение А Таблицы распределений статистик непараметрических критериев согласия при простых и сложных гипотезах	37
Приложение Б Библиография	59