На правах рукописи

Волкова Виктория Михайловна

ИССЛЕДОВАНИЕ РАСПРЕДЕЛЕНИЙ СТАТИСТИК ДИСПЕРСИОННОГО АНАЛИЗА В УСЛОВИЯХ НАРУШЕНИЯ ПРЕДПОЛОЖЕНИЙ НОРМАЛЬНОСТИ

Специальность 05 13 17 – Теоретические основы информатики

Автореферат диссертации на соискание ученой степени
кандидата технических наук

Новосибирск – 2007

Работа выполнена в Государственном образовательном учреждении высшего профессионального образования «Новосибирский государственный технический Университет»

Научный руководитель	доктор технических наук, профессор Лемешко Борис Юрьевич
Официальные оппоненты	доктор технических наук, профессор Загоруйко Николай Григорьевич
	кандидат технических наук, доцент Фаддеенков Андрей Владимирович
Ведущая организация	Томский политехническии университет

Защита состоится 16 мая 2007г в 12^{00} часов на заседании диссертационного совета Д 212 173 06 при Новосибирском государственном техническом университете (630092, Новосибирск-92, пр Карла Маркса, 20)

С диссертацией можно ознакомиться в библиотеке Новосибирского государственного технического университета

Автореферат разослан "12" апреля 2007 г

Ученый секретарь
диссертационного совета Чубич В М

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследований Актуальность и эффективность применения аппарата дисперсионного анализа в различных областях знания (социологии, экономике, биологии, медицине, педагогике, и т п) подчеркивается в работах Г Шеффе, С А Айвазяна, Дж Гласса, Ю Н Тюрина, В П Леонова, Ю П Адлера, В И Денисова, Е В Марковой, А А Попова, И А Полетаевой и других Наряду с научной сферой аппарат дисперсионного анализа используется при анализе процессов, возникающих в ходе промышленного производства (П А Рыжов, В Г Горский) Другими словами, методы дисперсионного анализа могут применяться всюду, где возникает необходимость в анализе влияния различных факторов на исследуемую переменную

Как свидетельствует мировая практика применения методов дисперсионного анализа, отсутствие уверенности в том, что наблюдения распределены по нормальному закону, вынуждают исследователя отказываться от использования параметрических критериев В то же время в отечественной практике широко распространено применение классических параметрических методов без предварительной проверки того, выполняются ли необходимые предположения А одним из основных является предположение нормальности наблюдений Такой подход зачастую приводит к некорректным статистическим выводам

Зачастую исследователи идут по пути построения теоретических оценок, характеризующих робастность имеющегося аппарата статистического анализа Но такие оценки, как, например, предложенные Шеффе для классических методов дисперсионного анализа, в большинстве случаев носят асимптотический характер, при этом далеко не всегда очерчивается граница области корректного применения критерия

Исследование свойств статистических критериев в условиях конкретной ситуации, в которой нарушаются классические предположения, с использованием аналитических методов, как правило, представляет собой очень сложную задачу В то же время, количество и уровень сложности задач, постоянно выдвигаемых практикой, возрастают настолько быстро, что ресурсы человеческого интеллекта, его производительность просто не в состоянии обеспечить решение такого множества задач без создания и использования соответствующих вычислительных технологий

Накопленный опыт в рамках развиваемого направления исследований (Лемешко Б Ю , Постовалов С Н , Чимитова Е В , Помадин С С) показывает, что для исследования статистических закономерностей наиболее эффективно применение компьютерных технологий моделирования С использованием методов статистического моделирования и последующего анализа можно получать результаты, не уступающие по точности аналитическим Применение данного подхода позволяет закрывать многие существующие в прикладной статистике «пробелы», используя при этом относительно простой вычислительный и математический аппарат Кроме того, методика компьютерного моделирования позволяет внедрять получаемые результаты исследований в программное обеспечение по статистическому анализу данных Важность применения современных достижений прикладной статистики в программных продуктах не раз подчеркивалась в работах А А Афифи, С А Айвазяна, Л С Векслера, Ю Н Тюрина, А А Макарова и др

Методика компьютерного моделирования и анализа статистических закономерностей позволяет исследовать поведение распределений различных статистик при нарушении классических предположений о принадлежности ошибок измерений нормальному закону, при нарушении других предположений В результате могут вырабатываться рекомендации по практическому применению исследуемых критериев, строиться аппроксимации распределений статистик критериев при различных законах распределения наблюдений

Цель и задачи исследований. Основная цель диссертационной работы заключалась в исследовании поведения законов распределений статистик дисперсионного анализа в случае принадлежности наблюдаемых случайных величин законам распределения, отличным от нормального

В соответствии с поставленной целью решались следующие задачи

- исследование распределений статистики, используемой в моделях с постоянными уровнями факторов при проверке гипотез о «средних», в условиях нарушения предположений нормальности,
- исследование распределений статистики критерия Т-метода множественных сравнений в зависимости от числа уровней фактора и числа наблюдений на каждом уровне при законах ошибок наблюдения, отличающихся от нормального,
- исследование распределений статистики критерия Хартли в зависимости от закона ошибок наблюдения, исследование мощности критерия Хартли в нормальном случае,
- исследование распределений статистики и мощности критерия Шеффе, применяемого для проверки гипотезы о равенстве дисперсий выборок, при различных законах распределения ошибок наблюдения, выработка рекомендаций по применению критерия,
- исследование распределений статистик и мощности критериев, используемых в однофакторной и двухфакторной моделях компонент дисперсий, при отклонении закона распределения эффектов случайных уровней факторов и ошибок наблюдения от нормального

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

Научная новизна диссертационной работы заключается

- в результатах исследования распределений статистики Т-метода множественных сравнений при различных законах распределения наблюдений в зависимости от числа и объемов сравниваемых выборок,
- в результатах сравнительного анализа мощности исследуемых критериев проверки однородности дисперсий,
- в построенных моделях распределения статистики критерия Хартли, расширяющих возможности критерия на случай ряда отличных от нормального законов распределения ошибок наблюдения,

– в построенных моделях законов распределения статистик, используемых в моделях дисперсионного анализа со случайными уровнями факторов, при различных значениях числа уровней факторов и числа наблюдений на каждом уровне фактора, различных законах распределения случайных эффектов уровней факторов и ошибок наблюдения

Основные положения, выносимые на защиту.

1 Результаты исследования распределений статистики, используемой в моделях с постоянными уровнями факторов при проверке гипотез о «средних», при неоднородности ошибок наблюдений по дисперсиям и по распределению в условиях нарушения предположений нормальности

2 Результаты исследования распределений статистики T-критерия при различных законах распределения ошибок наблюдения в зависимости от числа уровней фактора Построенные модели законов распределений статистик критерия для ряда законов распределений ошибок

3 Результаты исследования распределений статистики критерия Хартли при различных законах распределения ошибок наблюдения в зависимости от числа уровней фактора Построенные модели распределения критерия Хартли для ряда законов распределений ошибок

4 Результаты исследования мощности критериев Хартли и критерия Шеффе для проверки гипотезы о равенстве дисперсий в зависимости от числа сравниваемых выборок и их объемов в нормальном случае

5 Результаты исследований и рекомендации по применению критерия Шеффе в зависимости от закона распределения ошибок наблюдения

6 Результаты исследований распределений статистик и мощности критериев, используемых в моделях со случайными уровнями факторов, построенные модели распределений статистик при различных законах распределения случайных эффектов модели

Обоснованность и достоверность научных положений, выводов и рекомендаций обеспечивается

– корректным применением методов статистического моделирования для исследования распределений статистик критериев,
– совпадением результатов статистического моделирования с известными теоретическими результатами

Личный творческий вклад автора заключается в проведении исследований, обосновывающих основные положения, выносимые на защиту

Практическая ценность и реализация результатов Результаты исследований позволяют корректно применять критерии дисперсионного анализа в более широких границах, по сравнению с условиями, определяемыми классическими предположениями Выделены ситуации, в которых применение классического аппарата приведет к существенным ошибкам в статистических выводах, и возможные последствия таких ошибок Выработаны рекомендации по применению критерия Шеффе Разработано программное обеспечение, позволяющее строить модели распределений статистик в конкретной ситуации

Апробация работы Основные результаты исследований докладывались на Шестой всероссийской НТК «Информационные технологии в науке, проектировании и производстве» (Нижний Новгород, 2002), Международной НТК "Информатика и проблемы телекоммуникаций" (Новосибирск, 2003), Региональной научной конференции (с участием иностранных ученых) «Вероятностные идеи в науке и философии» (Новосибирск, 2003), Российских НТК "Информатика и проблемы телекоммуникаций" (Новосибирск, 2004, 2005, 2006), VI, VII и VIII международных конференциях "Актуальные проблемы электронного приборостроения" (Новосибирск, 2002, 2004, 2006), Седьмой международной конференции «Computer data analysis and modeling robustness and computer intensive methods» (Minsk, 2004), VIII Korea-Russia International Symposium on Science and Technology (Tomsk, 2004), IX Korea-Russia International Symposium on Science and Technology (Novosibirsk, 2005), Международном научно-техническом семинаре "Математическая, статистическая и компьютерная поддержка качества измерений" (Санкт-Петербург, 2006)

Исследования по теме диссертации явились составной частью работ, проводимых в рамках проектов "Математическое и алгоритмическое обеспечение задач статистического анализа данных и исследования статистических закономерностей при нарушении классических предположений", грант Министерства образования Российской Федерации № ТО2-3 3-3356, 2003-2004 гг , "Развитие компьютерных технологий моделирования и исследования фундаментальных закономерностей математической статистики", раздел 3 3 программы "Развитие научного потенциала высшей школы" Министерства образования и науки РФ, код проекта 15378, 2005 г , "Развитие компьютерных технологий исследования статистических закономерностей" (контракт № 2005-РИ-19 0/002/091, 2005 г) и "Применение компьютерных технологий исследования статистических закономерностей в задачах оценивания и различения близких гипотез о виде и свойствах распределений случайных величин" (контракт № 2006-РИ-19 0/001/119, 2006 г), ФЦНТП "Исследования и разработки по приоритетным направлениям развития науки и техники" на 2002-2006 годы по разделу "Проведение научных исследований молодыми учеными", "Расширение прикладных возможностей классических методов математической статистики", грант Российского фонда фундаментальных исследований, № 06-01-00059

Публикации. По теме диссертации опубликовано 13 печатных работ Из них 1 научная статья в рецензируемом журнале, входящем в перечень ВАК РФ, 2 – в сборниках научных трудов, 10 – материалы конференций (5 – международных, 5 – российских) В конце автореферата приведен список основных работ

Структура работы. Диссертация состоит из введения, 6 глав основного содержания и заключения, включает 51 таблицу, 50 рисунков и приложения Общий объем основной части диссертации – 193 страниц, список литературы содержит 138 наименований

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

В первой главе представлены основные определения, виды рассматриваемых моделей, статистических критериев и связанных с ними предположений

Моделью дисперсионного анализа с постоянными уровнями факторов называется модель вида

$$\mathbf{Y} = \mathbf{X}\vartheta + \mathbf{e}, \tag{1}$$

где $\mathbf{Y} = (y_1, \ldots, y_n)^T$ – вектор наблюдений, $\vartheta = (\mu \,|\, \vartheta_1^T \,|\, \ldots \,|\, \vartheta_p^T)^T$ – вектор фиксированных параметров, включающий в себя аддитивную постоянную μ и векторы эффектов уровней p фиксированных факторов, $\mathbf{X}$ – матрица планирования размерности $(n \times m)$, $r = rg(\mathbf{X})$, $\mathbf{e} = (e_1, \ldots, e_n)^T$ – вектор случайных ошибок наблюдения. Компоненты вектора $\mathbf{e}$ предполагаются независимыми случайными величинами, одинаково распределенными с общей функцией распределения, имеющей нулевое математическое ожидание и некоторую дисперсию σ^2. В классическом случае к этим предположениям добавляется требование распределения ошибок $\mathbf{e}$ по нормальному закону.

Двухфакторной сбалансированной моделью дисперсионного анализа со случайными уровнями факторов (двухфакторной моделью компонент дисперсии) без взаимодействий называется модель вида

$$y_{ijl} = \mu + a_i + b_j + e_{ijl}, \quad i = 1, \ldots, I, \; j = 1, \ldots, J, \; l = 1, \ldots, L, \tag{2}$$

где μ – генеральное среднее в двумерной популяции, I – число уровней фактора A, выбранных случайным образом, a_i – эффект i-го уровня, J – число выбранных случайным образом уровней фактора B, b_j – эффект j-го уровня, l - номер опыта, L – число опытов в каждой ячейке. Предполагается, что совокупности случайных величин $\{a_i\}$, $\{b_j\}$, $\{e_{ijl}\}$ независимы и в каждой из групп одинаково распределены с нулевыми средними и дисперсиями σ_A^2, σ_B^2 и σ_e^2 соответственно. В классической постановке предполагается нормальность эффектов уровней факторов и ошибок наблюдений.

Модель наблюдений однофакторной сбалансированной модели со случайными уровнями факторов записывается в виде

$$y_{ij} = \mu + a_i + e_{ij}, \quad i = 1, \ldots, I, \; j = 1, \ldots, J. \tag{3}$$

Линейную гипотезу относительно параметров модели (1) можно представить следующим образом:

$$H_0: \mathbf{K}^T\vartheta = \mathbf{b}, \tag{4}$$

где $\mathbf{K}^T$ – известная матрица размерности $(k \times m)$, $rg(\mathbf{K}) = k \le m$, $\mathbf{b}$ – заданный вектор размерности k. Статистика критерия отношения правдоподобия, используемая для проверки гипотезы вида (4) имеет вид

$$Q = \frac{\left(\mathbf{K}^T\hat{\vartheta} - \mathbf{b}\right)^T \left(\mathbf{K}^T\mathbf{M}^-\mathbf{K}\right)^{-1} \left(\mathbf{K}^T\hat{\vartheta} - \mathbf{b}\right)}{\left(\mathbf{Y} - \mathbf{X}\hat{\vartheta}\right)^T \left(\mathbf{Y} - \mathbf{X}\hat{\vartheta}\right)} \frac{n-r}{k}. \tag{5}$$

Здесь $\hat{\vartheta}$ – оценка вектора параметров модели ϑ, $\mathbf{M}^-$ – матрица, псевдообратная к матрице $\mathbf{M} = \mathbf{X}^T\mathbf{X}$. При выполнении предположений нормальности статистика (5) подчиняется F-распределению Фишера со степенями свободы k и $n-r$ $G(Q|H_0)=F_{k,n-r}$

Чтобы попарно сравнить средние значения вектора $\mathbf{Y}$ в модели (1) на различных уровнях фактора и выявить значимо отличающиеся друг от друга, используются методы множественного сравнения Статистика Т-метода

$$R=\frac{\max\limits_i u_i-\min\limits_i u_i}{s\sqrt{1/n_i}},\quad где\quad u_i=\hat{\mu}_i-\mu_i,\quad s=\sqrt{\frac{1}{I(n_i-1)}\sum_{i=1}^{I}\sum_{j=1}^{n_i}\left(y_{ij}-\hat{\mu}_i\right)^2}, \tag{6}$$

подчиняется распределению стьюдентизированного размаха $q_{I,I(n_i-1)}$ в случае выполнения предположений нормальности В формуле (6) n_i – число наблюдений на i-м уровне фактора, μ_i – математическое ожидание отклика на i-м уровне, $\hat{\mu}_i=\frac{1}{n_i}\sum_{j=1}^{n_i}y_{ij}$, $\sum_{i=1}^{I}n_i=n$, $i=1,\ \ ,I$

В модели (1) также проверяется гипотеза о равенстве дисперсий наблюдений на различных уровнях фактора Такая гипотеза имеет вид

$$H_0\ \ \sigma_1^2=\sigma_2^2=\ \ =\sigma_I^2, \tag{7}$$

а конкурирующая с ней – H_1 $\sigma_{i_1}^2\neq\sigma_{i_2}^2$, $1\le i_1,i_2\le I$, $i_1\neq i_2$ При исследовании мощности рассматривалась альтернатива вида H_1 $\sigma_I^2=C\sigma_1^2, \sigma_1^2=\sigma_2^2=\ \ =\sigma_I^2$

Статистика критерия Хартли для проверки гипотезы (7) имеет вид

$$T_1=\frac{\max\limits_i s_i^2}{\min\limits_i s_i^2},\quad s_i^2=\frac{1}{n_i-1}\sum_{j=1}^{n_i}(y_{ij}-y_{i\bullet})^2,\quad y_{i\bullet}=\frac{1}{n_i}\sum_{j=1}^{n_i}y_{ij} \tag{8}$$

Известны процентные точки распределения статистики (8) в нормальном случае

Критерий Шеффе для проверки гипотезы (7) основан на том, что множество значений каждой i-й выборки наблюдений $\{y_{ij}\}$, $j=1,\ \ ,n_i$, разбивается на J_i групп объемом n_{ij}, так что $n_i=J_in_{ij}$ Если обозначить совокупность значений, полученную путем разбиения $\{y_{ij}\}$ на подвыборки, через $\{x_{ijl}\}$, $i=1,\ \ ,I$, $j=1,\ \ ,J_i$, $l=1,\ \ ,n_{ij}$, то статистика критерия Шеффе может быть записана в виде

$$T_2=\left(\sum_{i=1}^{I}(J_i-1)\right)\sum_{i=1}^{I}J_i\left(z_{i\bullet}-\bar{z}\right)^2\Bigg/\left((I-1)\sum_{i=1}^{I}\sum_{j=1}^{J_i}\left(z_{ij}-z_{i\bullet}\right)^2\right),\quad z_{ij}=\ln s_{ij}^2, \tag{9}$$

$$z_{i\bullet}=\frac{1}{J_i}\sum_{j=1}^{J_i}z_{ij},\ \bar{z}=\frac{1}{\sum\limits_{i=1}^{I}J_i}\sum_{i=1}^{I}J_iz_{i\bullet},\ s_{ij}^2=\frac{1}{n_{ij}-1}\sum_{j=1}^{n_{ij}}(x_{ijl}-x_{ij\bullet})^2,\ x_{ij\bullet}=\frac{1}{n_{ij}}\sum_{j=1}^{n_{ij}}x_{ijl}$$

По предположению Шеффе статистика T_2 будет подчиняться F-распределению Фишера со степенями свободы $I-1$ и $\sum_{i=1}^{I}(J_i-1)$ Причем распределение статистики не должно зависеть от закона ошибок наблюдений $\{e_{ij}\}$

В модели компонент дисперсии относительно случайного фактора A проверяется гипотеза вида

$$H_A \quad \sigma_A^2 \le \theta_0\sigma_e^2, \quad \theta_0 \ge 0, \quad \theta_0 = const \tag{10}$$

В двухфакторной модели (2) для проверки гипотезы (10) при $L>1$ используется статистика (11), которая при выполнении предположений нормальности и справедливости гипотезы $H_A \quad \sigma_A^2 = \theta_0\sigma_e^2$ подчиняется F-распределению Фишера с числом степеней свободы $I-1$ и $IJ(L-1)$

$$S=\frac{1}{(1+JL\theta_0)}\frac{\frac{JL}{I-1}SS_A}{\frac{1}{IJ(L-1)}SS_e}=\frac{1}{(1+JL\theta_0)}\frac{\frac{JL}{I-1}\sum_{i=1}^{I}\left(y_{i\bullet\bullet}-\bar{y}\right)^2}{\frac{1}{IJ(L-1)}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{l=1}^{L}\left(y_{ijl}-y_{ij\bullet}\right)^2}, \tag{11}$$

$$y_{ij\bullet}=\frac{1}{L}\sum_{l=1}^{L}y_{ijl},\ y_{i\bullet\bullet}=\frac{1}{J}\sum_{j=1}^{J}y_{ij\bullet},\ y_{\bullet j\bullet}=\frac{1}{I}\sum_{i=1}^{I}y_{ij\bullet},\ \bar{y}=\frac{1}{I}\sum_{i=1}^{I}y_{i\bullet\bullet}=\frac{1}{J}\sum_{j=1}^{J}y_{\bullet j\bullet}$$

В случае однофакторной модели (3) для проверки гипотезы вида (10) используется статистика вида

$$S=\frac{1}{(1+J\theta_0)}\frac{SS_A/(I-1)}{SS_e/I(J-1)},\quad SS_A=J\sum_{i=1}^{I}(y_{i\bullet}-\bar{y})^2,\quad SS_e=\sum_{i=1}^{I}\sum_{j=1}^{J}(y_{ij}-y_{i\bullet})^2 \tag{12}$$

При выполнении предположений нормальности и справедливости гипотезы $H_A \quad \sigma_A^2=\theta_0\sigma_e^2$ статистика (12) подчиняется распределению $F_{I-1\ I(J-1)}$, где J – это число наблюдений на каждом уровне фактора

Вторая глава содержит исследования распределений статистики (5) критерия отношения правдоподобия при отклонениях наблюдаемого закона от нормального Исследуется, как влияет на распределение статистики (5) дополнительно присутствие неоднородности ошибок наблюдений по дисперсиям и по распределению Проверяется, насколько корректными будут статистические выводы в этих условиях Помимо этого во второй главе приводятся результаты исследования поведения распределения статистики (6) Т-критерия множественных сравнений в зависимости от наблюдаемого закона ошибок наблюдения, оценивается влияние ненормальности наблюдений на границы доверительных интервалов, вычисляемых на основании распределения статистики (6)

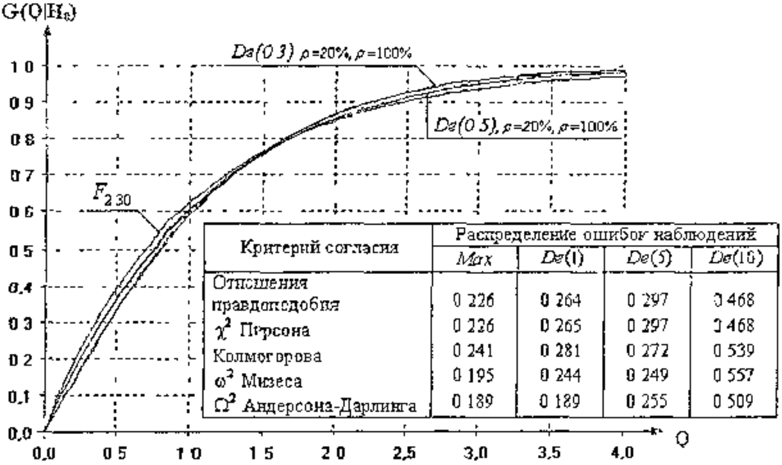
Исследования (преимущественно) проводились при следующих законах распределения случайных величин распределении максимальных значений с плотностью $f(x,\theta_2,\theta_1)=\exp(-w-\exp(-w))/\theta_2$, $w=(x-\theta_1)/\theta_2$, нормальном законе, и семействе распределений с плотностью

$$De(\lambda) = f(x,\lambda,\theta_2,\theta_1) = \frac{\lambda}{2\sqrt{2}\theta_2\Gamma(1/\lambda)}\exp\left(-\left(\frac{|x-\theta_1|}{\sqrt{2}\theta_2}\right)^{\lambda}\right) \qquad (13)$$

при различных значениях параметра формы λ (в работе приводятся результаты при значениях параметра формы λ, равном 0 3, 0 5, 1, 3, 5 и 10)

Численные исследования распределения статистики (5) в случае нарушения предположений нормальности подтвердили предположения Шеффе распределение данной статистики устойчиво к существенным отклонениям распределения ошибок от нормального закона Значимые отклонения эмпирического распределения статистики (5) от соответствующего распределения Фишера наблюдаются только при законах ошибок с тяжелыми хвостами (De(0 3) и De(0 5))

Рис 1 иллюстрирует поведение распределения статистики в случае принадлежности ошибок наблюдения законам распределения с тяжелыми хвостами и различным уровнем шума На рис 1 приведена таблица с достигнутыми уровнями значимости при проверке гипотезы о согласии эмпирических законов распределения статистики (5) (при различных законах распределения ошибок) с распределением статистики (5) в нормальном случае (с распределением Фишера $F_{2\,30}$) при уровне шума 20%



Критерий согласия	Распределение ошибок наблюдений			
	Max	De(1)	De(5)	De(10)
Отношения правдоподобия	0 226	0 264	0 297	0 468
χ^2 Пирсона	0 226	0 265	0 297	0 468
Колмогорова	0 241	0 281	0 272	0 539
ω^2 Мизеса	0 195	0 244	0 249	0 557
Ω^2 Андерсона-Дарлинга	0 189	0 189	0 255	0 509

Рис 1 Функции распределения статистики (5) при различных законах ошибок наблюдений и различных уровнях шума ρ и соответствующее нормальному случаю теоретическое распределение Фишера $F_{2\,30}$ ($k = 2, n = 36, r = 6$)

Исследования влияния неоднородности ошибок показали, что сильная неоднородность в дисперсиях может приводить к отклонению распределения статистики (5) от соответствующего распределения Фишера Рассматриваемый вид неоднородности в дисперсиях может быть охарактеризован параметром $C=\sigma_1^2/\sigma_i^2$, $\sigma_i^2=2,\quad ,I$, где σ_i^2 - это дисперсия ошибок наблюдений на i-м уровне фактора В таблице 1 представлены результаты исследований, связанных с влиянием неоднородности по дисперсиям для модели (1) при $k=2, n=36, r=6$ В ней приведены два типа погрешностей, возникающих при вычислении вероятности ошибки 1-го рода по сравнению с заданным α Погрешности первого типа связаны с нарушением предположений нормальности и однородности Они обозначены как ε_P и определяются соотношением $\varepsilon_P=P_{0\,9}^*-0\,9$, где $P_{0\,9}^*=P\{S\le S_\alpha\}$ – вероятность, вычисленная по модели распределения статистики, а S_α соответствующая квантиль $F_{2\,30}$-распределения при $\alpha=0\,1$ ε_P показывает характер влияния на распределение статистики При $\varepsilon_P<0$ применение вместо действительного распределения статистики (5) "классического" F-распределения Фишера будет приводить к увеличению вероятности ошибки 1-го рода по сравнению с заданным α И наоборот Погрешности второго типа связаны с точностью моделирования получаемой вероятности $P_{0\,9}^*$ Такая погрешность в таблице обозначается как ε_M Она представляет собой точность оценки $P_{0\,9}^*$ с заданной достоверностью $\gamma=0\,99$ при используемом объеме выборок статистик $N=100000$ Прочерк в таблице означает, что в этом случае наблюдалось хорошее согласие эмпирического распределения статистики (5) с $F_{2\,30}$-распределением

Таблица 1

Распределение ошибок наблюдения	C							
	0 2		1		2		5	
	ε_P	ε_M	ε_P	ε_M	ε_P	ε_M	ε_P	ε_M
De(0 3)	0 0186	*0 0022*	0 0238	*0 0022*	0 0237	*0 0022*	0 0205	*0 0022*
De(0 5)	0 0021	*0 0024*	0 0112	*0 0023*	0 0053	*0 0024*	-0 0007	*0 0025*
De(1)	-0 0059	*0 0025*	–	–	0 0015	*0 0024*	-0 0154	*0 0026*
Max	-0 0089	*0 0025*	–	–	-0 0089	*0 0025*	-0 0177	*0 0026*
Norm	-0 0087	*0 0025*	–	–	-0 0051	*0 0025*	-0 0087	*0 0025*
De(5)	-0 0092	*0 0025*	–	–	-0 0025	*0 0025*	-0 0132	*0 0026*
De(10)	-0 0071	*0 0025*	–	–	-0 0026	*0 0025*	-0 0108	*0 0026*

Из таблицы 1 видно, что при очень большой степени неоднородности, когда дисперсии различаются в 5 раз, это различие сказывается на степени согласия распределения статистики (5) с теоретическим распределением, что выражается в изменении значения погрешности ε_P При этом характер влияния зависит от закона распределения ошибок наблюдения при распределении ошибок по закону $De(0\,3)$ неоднородность в дисперсиях приводит к снижению степени расхождения распределения статистики (5) от F-распределения ($|\varepsilon_P|$ уменьшается), при значении $\lambda \geq 1$ неоднородность в дисперсиях приводит к увеличению степени отклонения от F-распределения ($|\varepsilon_P|$ увеличивается)

В процессе исследований выявлено, что неоднородность по распределениям (как таковая) не оказывает существенного влияния на распределение статистики

Таким образом, проведенные во второй главе исследования поведения распределения статистики (5) расширяют сферу корректного применения классических результатов При проверке гипотез о «средних» в моделях с постоянными уровнями факторов нет настоятельной необходимости в проверке нормальности Достаточно убедиться в отсутствии «тяжелых хвостов» закона распределения ошибок и отсутствии существенной неоднородности в дисперсиях групп

Проведенные исследования статистики (6) показали, что предположение Шеффе об отсутствии устойчивости к нарушению предположений нормальности у Т-метода является оправданным Выявлено, что устойчивость Т-метода зависит от числа уровней факторов При малом числе уровней факторов значительное влияние на распределение статистики оказывает принадлежность ошибок законам с "тяжелыми" хвостами В этом случае построенные с использованием традиционного аппарата доверительные интервалы будут шире действительных

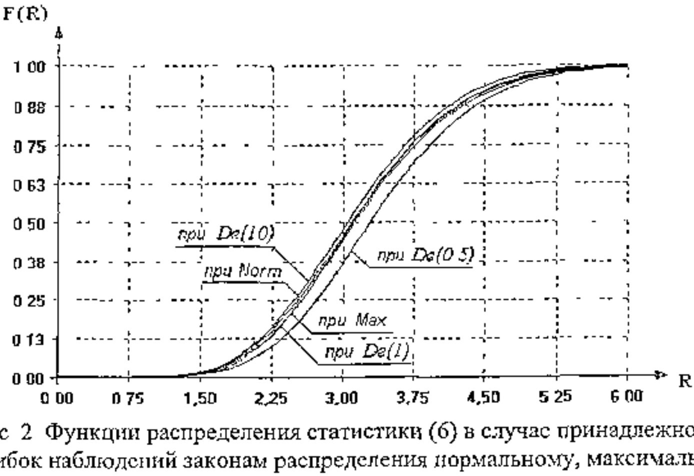


Рис 2 Функции распределения статистики (6) в случае принадлежности ошибок наблюдений законам распределения нормальному, максимальных значений и $De(\lambda)$ с различными значениями параметра λ при $I=10$ и $n_i=40$

При большом числе уровней факторов наблюдается высокая степень неустойчивости к нарушению предположений нормальности На рис 2 приведены эмпирические распределения статистики (6) при различных распределениях ошибок наблюдения для случая $I=2$, $n_i=40$

Для ряда законов распределений ошибок наблюдений и тех чисел уровней фактора I, чисел наблюдений n_i, при которых наблюдалось отклонение распределения статистики (6) от соответствующего распределения стьюдентизированного размаха, построены модели распределения статистики

В третьей главе методами компьютерного моделирования исследуются распределения статистик критериев Хартли и Шеффе, применяемых для проверки гипотезы о равенстве дисперсий нескольких выборок, в случае принадлежности наблюдений случайных величин различным законам распределения Проводится сравнительный анализ мощности критериев однородности

Проведенные исследования показали, что критерий Хартли крайне чувствителен к виду закона, которому принадлежат ошибки наблюдений Эта зависимость проиллюстрирована на рис 3

Для диапазона значений $I=\{2,3,4,5\}$, $n_i=\{4,5,6,8,10,20\}$ построены модели законов распределений статистики (8) в случае принадлежности ошибок наблюдений законам распределения $De(1)$, $De(5)$, $De(10)$, экстремальных значений

Критерий Шеффе (в отличие от критерия Хартли) действительно весьма устойчив к нарушению предположений нормальности Однако вид закона распределения ошибок наблюдений влияет на объем выборки, начиная с которого достигается достаточно хорошее согласие распределения статистики с распределением Фишера Чем "легче" хвосты распределений, тем при больших объемах выборок достигается достаточно хорошее согласие

В нормальном случае минимальное значение n_i, при котором допустимо использовать F-распределение в качестве распределения статистики (9), составляет 10-12 наблюдений в группе Необходимый для приемлемого согласия объем n_i в наибольшей степени определяется выбором числа J_i Например, при $J_i=2$ минимально допустимый объем n_i составляет 10-12, при $J_i=4$ – 16-20, а при $J_i=6$ – 20-25 Проведенные исследования показали, что при нормальном законе ошибок и числе наблюдений n_i в группах около 30, можно без риска совершения больших ошибок использовать F -распределение в качестве распределения статистики (9) при условии, что будет выбрано разбиение, в котором $n_{ij} \geq 4$ Но, если объем выборок n_i меньше этого числа, то для того, чтобы распределение статистики (9) хорошо согласовалось с соответствующим F-распределением, следует выбирать такое разбиение на подгруппы, чтобы n_{ij} было наибольшим, а J_i - наименьшим из возможных

Исследования мощности критерия Шеффе в зависимости от закона распределения ошибок наблюдения показали, что на мощность критерия Шеффе влияет также закон распределения наблюдений $\{x_{ijl}\}$

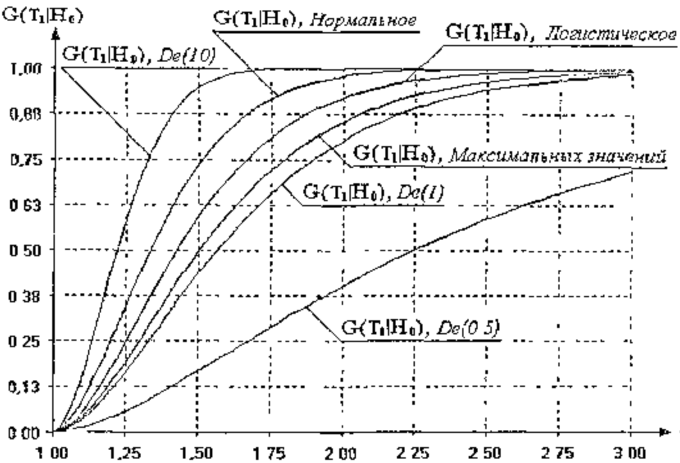


Рис 3 Эмпирические функции распределения статистики (8) при справедливости гипотезы (7) в случае I=3, n_i=61 при различных законах ошибок наблюдений

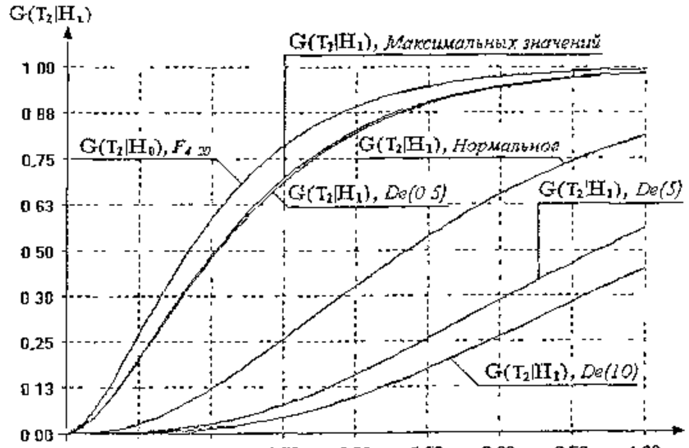


Рис 4 Функции распределения статистики (9) при справедливости проверяемой гипотезы H_0 вида (7) и при справедливости конкурирующей гипотезы H_1 вида $\sigma_I^2 = 1{,}2\sigma_1^2$ при различных законах наблюдений при $I=5$, $J_i=10$, $n_{ij}=50$

Рис 4 иллюстрирует влияние закона распределения наблюдений на распределение статистики при справедливости конкурирующей гипотезы и большом объеме выборки n_i =500

Картина, представленная на рис 4, показывает, что чем “легче” хвосты распределений наблюдений $\left\{x_{ij}\right\}$, тем выше мощность Влияние несимметричности закона распределения наблюдений на распределение статистики (9) при верной альтернативе схоже с влиянием распределения с более “тяжелыми” чем у нормального закона хвостами

Таблица 2

Мощность критериев Кокрена, Бартлетта, Шеффе и Хартли относительно альтернативы вида H_1 $\sigma_I^2 = 1\ 44\sigma_1^2$ при I=5 и различных значениях n_i в случае принадлежности наблюдений нормальному закону

α	n_i	Разбиение на подгруппы в случае критерия Шеффе	Критерий Шеффе	Критерий Хартли	Критерий Бартлетта	Критерий Кокрена
0 05	200	J_i=2, n_{ij}=100	0 362	0 760	0 757	0 837
		J_i=4, n_{ij}=50	0 578			
		J_i=5, n_{ij}=40	0 609			
		J_i=10, n_{ij}=20	0 647			
		J_i=20, n_{ij}=10	0 619			
	500	J_i=5, n_{ij}=100	0 967	0 993	0 993	0 997
		J_i=10, n_{ij}=50	0 981			
		J_i=20, n_{ij}=25	0 982			
		J_i=25, n_{ij}=20	0 8677			
		J_i=50, n_{ij}=10	0 172			

В целом по результатам проведенных исследований поведения распределения статистики критерия Шеффе можно заключить следующее При использовании критерия сложен выбор “оптимальных” значений n_{ij} и J_i при заданном n_i, при которых распределение статистики близко к соответствующему F-распределению Этот выбор зависит от вида закона ошибок наблюдения Если ошибки подчиняются закону с «легкими» хвостами (например, $De(10)$ или $De(5)$), следует выбирать разбиение с минимальным значением J_i Это позволит с большей уверенностью использовать распределение Фишера в качестве распределения статистики (9) В случае законов ошибок с “тяжелыми” хвостами, следует выбирать разбиение с достаточно большим значением J. Это позволит увеличить мощность критерия, которая мала в этих ситуациях

Таким образом, свойства критерия Шеффе зависят от вида закона ошибок наблюдения и от того, насколько удачно выбрано разбиение на подвыборки. По результатам проведенных исследований были выработаны рекомендации по практическому применению критерия Шеффе в тех или иных ситуациях

Полученные результаты сравнительного анализа мощности критериев однородности дисперсий в нормальном случае иллюстрирует таблица 2

Мощность критерия Хартли при нормальном законе ошибок наблюдения сравнима с мощностью критерия Бартлетта, ниже мощности критерия Кокрена и выше мощности критерия Шеффе. Мощность критерия Шеффе ниже мощности других критериев и зависит от того, каким образом наблюдения разбиваются на подгруппы

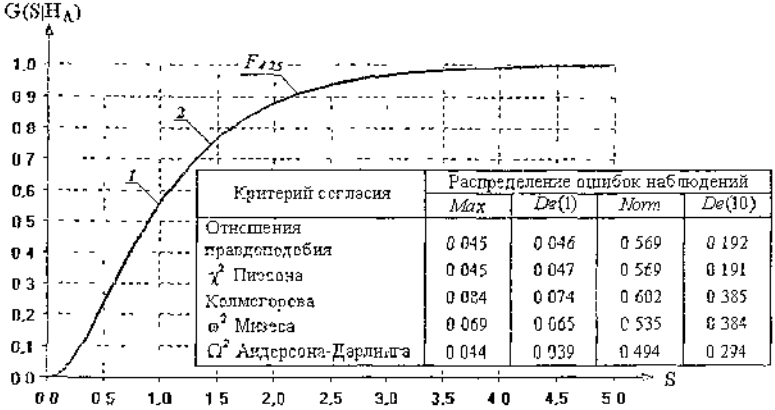
Четвертая глава содержит результаты исследования распределений статистик, применяемых в однофакторной модели (3) со случайными уровнями фактора при различных законах распределения входящих в модель случайных величин

На первом этапе методами статистического моделирования исследовались распределения статистик SS_A и SS_e, входящих в соотношение (12) и подчиняющихся в нормальном случае распределению χ^2 со степенями свободы $I-1$ и $I(J-1)$. Исследование распределений статистик SS_A и SS_e при нарушении предположений нормальности во многом объясняет поведение в этих условиях распределения статистики (12)

Далее проводились исследования распределения статистики (12) при выполнении условий гипотезы вида H_A: $\sigma_A^2 = 0$. Показано, что распределение статистики (12) в этом случае устойчиво к нарушению предположений нормальности. Данную ситуацию иллюстрирует рис. 5. В таблице, приведенной на рисунке, приведены достигнутые уровни значимости при проверке согласия эмпирических распределений статистики (12) при различных законах распределения ошибок с распределением статистики (12) в нормальном случае (с распределением $F_{4,25}$). Из рисунка видно, что даже в тех случаях, когда согласие эмпирического распределения с F-распределением ниже, чем в нормальном случае (законы распределения ошибок $De(1)$ и максимальных значений), применение соответствующего распределения Фишера в качестве закона распределения статистики (12) при проверке гипотезы вида H_A: $\sigma_A^2 = 0$ не приведет к существенным ошибкам в статистических выводах. Этот факт существенно расширяет область корректного применения данного критерия при проверке гипотез такого вида

Напротив, распределение статистики (12) при справедливости гипотезы H_A: $\sigma_A^2 = \theta_0 \sigma_e^2$, $\theta_0 \neq 0$ неустойчиво к нарушениям предположений нормальности: на поведение распределения статистики (12) оказывает влияние как распределение случайных эффектов уровней фактора, так и распределение ошибок наблюдений Причем распределение случайных эффектов уровней фактора оказывает большее влияние Для диапазона значений $I = \{2, 3, 4, 5, 6\}$, $J = \{3, 4, 5, 6, 15\}$ (кроме случая $I = 2$, $J = 2$), $\theta_0 = 0.5, 1, 2$ при всех возможных сочетаниях законов (из числа $De(1)$, $Norm$, $De(5)$, $De(10)$), которым подчиняются ошибки наблюдений и эффекты уровней случайных факторов, были построены модели законов распределений статистики (12)

Было показано, что мощность критерия зависит от законов распределения ошибок и случайных факторов уменьшение величины эксцесса распределений $\{e_{ij}\}$ или $\{a_i\}$ приводит к увеличению мощности



Критерий согласия	Распределение ошибок наблюдений			
	Max	*De*(1)	*Norm*	*De*(10)
Отношения правдоподобия	0 045	0 046	0 569	0 192
χ^2 Пирсона	0 045	0 047	0 569	0 191
Колмогорова	0 084	0 074	0 602	0 385
ω^2 Мизеса	0 069	0 065	0 535	0 384
Ω^2 Андерсона-Дарлинга	0 044	0 039	0 494	0 294

Рис 5 Распределения статистики (12) при справедливости гипотезы вида H_A $\sigma_A^2 = 0$ 1 – в случае принадлежности $\{e_{ij}\}$ распределению максимального значения, 2 – в случае принадлежности $\{e_{ij}\}$ закону $De(1)$, $I = 5$, $J = 6$

В пятой главе методами компьютерного моделирования исследуются распределения статистики (11), используемой в двухфакторной модели (2) со случайными уровнями факторов при проверки гипотезы (10), в условиях нарушения предположений нормальности

Показано, что, так же как и в случае однофакторной модели, распределение статистики устойчиво к отклонению от предположений нормальности при проверке гипотез о равенстве нулю дисперсии случайного фактора Повторяется ситуация, наблюдавшаяся в поведении распределения статистики однофакторной модели на распределение статистики (11) оказывает влияние распределение случайных эффектов уровней того фактора, относительно которого справедливы условия гипотезы вида H_A $\sigma_A^2 = \theta_0\sigma_\varepsilon^2$, $\theta_0 \neq 0$, а также распределение ошибок наблюдения модели Исследования показали, что, как и ожидалось, распределение случайных эффектов уровней того фактора, относительно которого не проверяется гипотеза (10), не оказывает значительного влияния на распределение статистики (11) При этом в поведении статистики (11) сохраняются все тенденции, наблюдавшиеся в однофакторной модели компонент дисперсии

Более того, модели распределений статистики, полученные в случае однофакторной модели, могут использоваться в качестве моделей распределений статистики в соответствующей двухфакторной модели Это справедливо для случаев, в которых при выполнении предположений нормальности теоретические распределения соответствующих статистик в однофакторной и двухфакторной моделях совпадают Эту ситуацию иллюстрирует рис 6, на котором для некоторых отличных от нормального законах распределения случайных эффектов уровней фактора и ошибок наблюдения представлены получаемые эмпирические распределения статистики (12) и (11) моделей (3) и (2) соответственно На рисунке показано также распределение Фишера $F_{4,20}$, которому в нормальном случае подчиняется статистика (12) при $I=5$, $J=5$ и статистика (13) при $I=5$, $J=2$, $L=3$ Используемые в подрисуночной надписи обозначения расшифровываются, например, следующим образом $G(S(De(5),De(10)))$ соответствует распределению статистики S, построенному при справедливости проверяемой гипотезы для случая, когда ошибки ($\{e_{ij}\}$ в однофакторном случае, $\{e_{ijl}\}$ в двухфакторном случае) подчиняются закону распределения $De(5)$, $\{a_i\}$ – закону распределения $De(10)$

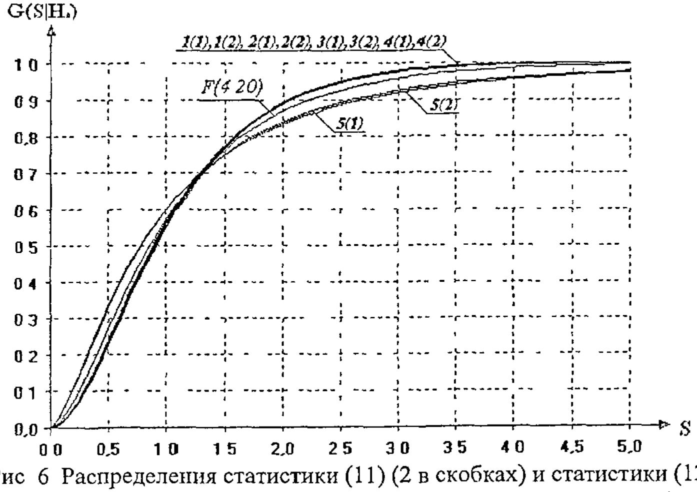


Рис 6 Распределения статистики (11) (2 в скобках) и статистики (12) (1 в скобках) при справедливости гипотезы вида H_A $\sigma_A^2=\sigma_e^2$
1 – *G(S(De(10),De(10)))*, 2 – *G(S(De(10), De(5)))*, 3 – *G(S(De(5),De(10)))*, 4 – *G(S(De(5), De(5)))*, 5 – *G(S(De(1),De(1)))*

Для ряда случаев, в которых наблюдалось недостаточно хорошее согласие полученных моделей распределений статистики (12) и статистики (11) с помощью ряда критериев была проведена проверка на однородность выборок статистики (12) и статистики (11) Таблица 3 содержит результаты проверки, полученные для тех сочетаний законов распределений случайных величин, входящих в модели (3) и модели (2), которые представлены на рис 6 Результаты, представленные в таблице 3 говорят о том, что нет оснований для отклонения гипотезы об однородности

В целом полученные в ходе исследования результаты позволяют надеяться, что построенные в однофакторном случае модели распределений статистик можно без опасения совершения больших ошибок использовать для проверки гипотез в двухфакторном случае, если статистики (12) и (11) в нормальном случае подчиняются одному и тому же закону распределения Это существенно расширяет область применения критерия со статистикой (11) за счет разработанного для однофакторной модели компонент дисперсии математического аппарата

Также в пятой главе исследовалось поведение статистики, используемой в модели (2) при числе опытов в каждой ячейке $L=1$ Показано, что все получаемые в этом случае результаты поддерживают тенденцию, наблюдавшуюся в случае $L>1$

Таблица 3

Результаты проверки однородности эмпирических распределений статистики (11), полученных в двухфакторной модели (2), с соответствующими эмпирическими распределениями статистики (12), полученными в случае однофакторной модели (3), при различных законах распределения эффектов $\{a_i\}$ и ошибок наблюдения

Распределение $\{a_i\}$ и $\{e_{ijl}\}$ ($\{e_{ij}\}$ в случае однофакторной модели)	Значения достигнутых уровней значимости	
$\{a_i\} \sim \mathrm{De}(1)$, $\{e_{ijl}\} \sim \mathrm{De}(1)$	Лемана-Розеблатта Смирнова Модифицированный Смирнова	0 177 0 179 0 185
$\{a_i\} \sim \mathrm{De}(1)$, $\{e_{ijl}\} \sim \mathrm{De}(10)$	Лемана-Розеблатта Смирнова Модифицированный Смирнова	0 428 0 432 0 374
$\{a_i\} \sim \mathrm{De}(5)$, $\{e_{ijl}\} \sim \mathrm{De}(1)$	Лемана-Розеблатта Смирнова Модифицированный Смирнова	0 413 0 419 0 421
$\{a_i\} \sim \mathrm{De}(10)$, $\{e_{ijl}\} \sim \mathrm{De}(5)$	Лемана-Розеблатта Смирнова Модифицированный Смирнова	0 152 0 153 0 124

Шестая глава содержит краткое описание разработанного программного обеспечения, которое использовалось для исследования распределений статистик дисперсионного анализа, и примеры его применения для ряда практических задач

Вначале приводится описание возможностей разработанного комплекса программ Во всех программных модулях используются датчики программной системы ISW, что позволяет генерировать псевдослучайные величины, подчиняющиеся одному из 30 заложенных в ISW законов распределений Комплекс состоит из трех частей

1 Программный модуль для моделирования распределений статистик критерия отношения правдоподобия при различных законах распределения ошибок наблюдения, который позволяет задавать различные законы распределения (как по виду закона, так и по значениям параметров) на различных уровнях и сочетаниях уровней факторов при задаваемой пользователем матрице планирования произвольного вида

2 Программный модуль для моделирования распределений статистик критерия Хартли, критерия Шеффе, Т-критерия

3 Программный модуль для моделирования распределений статистик критериев моделей компонент дисперсии, позволяющий задавать различные по виду и значениям параметров законы распределения для различных случайных величин, входящих в модель компонент дисперсии

Все программное обеспечение реализовано на языке C++ в среде Borland C++ 6 с использованием объектно-ориентированного подхода

Далее приводится описание применения разработанного программного обеспечения к данным, приведенным в работе П А Рыжова «Математическая статистика в горном деле»

Дается описание применения разработанного программного обеспечения к медицинским данным, полученным в результате наблюдения за персоналом специализированной туберкулезной больницы № 2 г Новосибирска с целью выявления факторов устойчивости персонала к заболеванию туберкулезом

Приводимые примеры позволяют продемонстрировать принципы работы разработанного программного обеспечения, которое может применяться в любой сфере человеческой деятельности, в которой возникает необходимость решения задач дисперсионного анализа К таким сферам деятельности относятся и различные технические приложения, в которых традиционно широко используется классический аппарат дисперсионного анализа

Заключение

В соответствии с целями исследований получены следующие результаты

1 Методами статистического моделирования исследовано распределение статистики, используемой в моделях с постоянными уровнями факторов при проверке гипотез о «средних» Показано, что распределение статистики устойчиво к существенным отклонениям распределения ошибок наблюдения от нормального закона

2 Показано, что сильная неоднородность наблюдений по дисперсиям может приводить к отклонению распределения статистики критерия отношения правдоподобия от соответствующего распределения Фишера Неоднородность по распределению (при однородности дисперсий) не оказывает существенного влияния

3 Впервые показано, что устойчивость Т-метода к нарушению предположений нормальности зависит от числа уровней факторов Для ряда значений числа уровней фактора и числа наблюдений на каждом уровне построены модели распределения статистики при различных законах ошибок наблюдения

4 Показано, что распределение статистики критерия Хартли неустойчиво к отклонениям закона распределения наблюдений от нормального Для ряда значений числа уровней фактора и числа наблюдений на каждом уровне построены модели распределения статистики при различных законах распределения ошибок наблюдения

5 Экспериментально исследовано поведение распределения статистики Шеффе при нарушении предположения нормальности Впервые показано, что, несмотря на устойчивость распределения статистики критерия к нарушению предположений нормальности (при справедливости проверяемой гипотезы), вид закона распределения наблюдений влияет на объем выборки, начиная с которого можно применять соответствующее распределение Фишера в качестве распределения статистики критерия

6 Показано, что мощность критерия Шеффе существенно зависит от вида разбиения наблюдений на подвыборки Закон распределения ошибок наблюдений также влияет на мощность критерия Даны рекомендации по выбору подразбиения, обеспечивающего максимальную мощность при соблюдении необходимого уровня согласия эмпирического распределения статистики с соответствующим распределением Фишера

7 Проведен сравнительный анализ мощности критериев однородности Хартли и Шеффе в нормальном случае Проведено сравнение мощности данных критериев с мощностью других критериев однородности дисперсий

8 Проведено экспериментальное исследование критериев, используемых для проверки гипотез о дисперсиях в моделях случайных компонент Показано, что распределения статистик данных критериев неустойчивы к нарушению предположений нормальности Исключение составляют случаи проверки гипотез о равенстве нулю дисперсии случайного фактора Для различных сочетаний законов ошибок измерений и уровней случайного фактора по ряду значений I, J и θ_0 построены модели распределений статистики, используемой в однофакторной модели

9 Впервые показано, что модели распределений статистики, полученные в случае однофакторной модели, могут использоваться в качестве моделей распределений статистик в соответствующей двухфакторной модели при условии, что распределение соответствующих статистик в однофакторной и двухфакторной моделях при выполнении предположений нормальности одно и то же

Полученные результаты расширяют сферу корректного применения методов классического дисперсионного анализа в приложениях

Полученные результаты и разработанное программное обеспечение используется при проведении научных исследований, может использоваться при решении различных прикладных задач Разработанное программное обеспечение было использовано при анализе медицинских данных, рассматриваемых в рамках подготовки двух кандидатских и одной докторской диссертации, что подтверждается актами о внедрении

Список основных публикаций

1 **Пономаренко, В М.** Влияние метода оценивания параметров модели на свойства оценок ФДО в условиях нарушения предположений о нормальности / В М Пономаренко // Сборник научных трудов НГТУ – 2004 – № 1(35) – С 35-40

2 **Ponomarenko, V M** Statistical hypotheses testing in variance analysis in case of classical assumptions failure / V M Ponomarenko, B Yu Lemeshko // proceedings of the Seventh International Conference «Computer data analysis and modeling robustness and computer intensive methods» – V 1 – Minsk 2004 – C 110-113 [Проверка статистических гипотез дисперсионного анализа в условиях нарушения основных предположений]

3 **Пономаренко, В М** Проверка гипотез о дисперсиях в рамках дисперсионного анализах при нарушении предположений нормальности / В М Пономаренко // Материалы VII международной конференции «Актуальные проблемы электронного приборостроения» – Т 6 – Новосибирск 2004 – С 312-314

4 **Lemeshko, B Yu** Investigation of dependence variance analysis statistical distributions on error and random factor distribution laws / B Yu Lemeshko, V M Ponomarenko // proceedings of the 9th Russian-Korean International Symposium on Science and Technology KORUS-2005 – Novosibirsk 2005 P 79-82 [Исследование зависимости распределений статистик дисперсионного анализа от распределения ошибок наблюдения и случайных уровней факторов]

5 **Лемешко, Б Ю** Проверка гипотез в моделях дисперсионного анализа со случайными факторами при нарушении предположений о нормальности / Б Ю Лемешко, В М Пономаренко // Доклады Академии наук высшей школы России – 2005 – № 2(5) – С 26-39

6 **Лемешко, Б Ю.** Исследование распределений статистик, используемых для проверки гипотез о равенстве дисперсий при законах ошибок наблюдений, отличных от нормального / Б Ю Лемешко, В М Пономаренко // Научный вестник НГТУ – 2006 – № 2(23) – С 21-33

7 **Лемешко, Б Ю** Исследование распределений статистики критерия Шеффе при законах ошибок наблюдений, отличающихся от нормального / Б Ю Лемешко, В М Пономаренко // Материалы VII международной конференции «Актуальные проблемы электронного приборостроения» – Т 6 – Новосибирск 2006 – С 87-91

8 **Пономаренко, В М** Исследование поведения статистики, используемой в Т-методе множественных сравнений при нарушении предположений о нормальности / В М Пономаренко, Б Ю Лемешко // Материалы Российской НТК «Информатика и проблемы телекоммуникаций» – Т 1 – Новосибирск 2006 – С 190-194

Отпечатано в типографии Новосибирского
государственного технического университета
630092, г Новосибирск, пр К Маркса, 20,
тел /факс (383) 346-08-57
формат 60x84/16, объем 1 5 п л , тираж 100 экз ,
заказ № 635 , подписано в печать 10 04 07 г