На правах рукописи

Головчинер Ольга Николаевна

МОДИФИЦИРОВАННЫЕ ОЦЕНКИ ЛИНЕЙНЫХ ФУНКЦИОНАЛОВ ОТ РАСПРЕДЕЛЕНИЙ ВЕРОЯТНОСТЕЙ С УЧЕТОМ ДОПОЛНИТЕЛЬНОЙ ИНФОРМАЦИИ

05.13.01 – Системный анализ, управление и обработка информации
(по отраслям информатики, вычислительной техники и автоматизации)

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата физико-математических наук

Томск – 2007

Работа выполнена в ГОУ ВПО "Томский государственный университет", кафедра теоретической кибернетики, и ТНЦ СО РАН

Научный руководитель:
доктор физико-математических наук,
профессор Дмитриев Юрий Глебович

Официальные оппоненты:
доктор физико-математических
наук, профессор
кафедры ВМиММ ТГУ Воробейчиков Сергей Эрикович
доктор технических наук,
профессор
кафедры АСУ ТУСУР Сергеев Виктор Леонидович

Ведущая организация:
Томский политехнический университет

Защита состоится:
20 декабря 2007 г. в 10.30 на заседании диссертационного совета Д 212.267.12 при Томском государственном университете по адресу: 634050, г. Томск, пр. Ленина, 36.

С диссертацией можно ознакомиться
в Научной библиотеке Томского государственного университета.

Автореферат разослан: 12 ноября 2007 г.

Ученый секретарь
диссертационного совета
д.т.н., профессор В.И. Смагин

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность работы. Одной из основных задач статистической обработки данных является определение вероятностных характеристик исследуемого явления или системы. Математическая формулировка таких задач обычно сводится к оцениванию функционалов от неизвестного распределения вероятностей наблюдаемой случайной величины, которое приходится оценивать по результатам проводимых экспериментов, наблюдений и измерений.

Практически всегда исследователь, кроме выборки, обладает какой-либо дополнительной информацией об оцениваемом функционале или распределении. Например, о функционале или других, с ним связанных, может быть известно, что они могут принимать значения из заданного множества, а распределение может быть симметричным, иметь известные моменты заданных уровней и т.п. Стремление повысить качество оценок или уменьшить объем экспериментальных данных, требуемых для достижения заданной точности, приводит к необходимости рационального учета всех имеющихся сведений.

Начиная с середины прошлого века, проблема привлечения дополнительной априорной информации в процедуры статистического оценивания широко обсуждается в научной литературе. В работах H.H. Hansen, H.O. Hartley, Ю.Н. Тюрина, E.F. Schuster, Б.Я. Левита, Ю.А. Кошевника, В.Н. Пугачева, Ф.П. Тарасенко, Ю.Г. Дмитриева, Г.М. Кошкина, Ю.К. Устинова, J. Chen, B. Zhang, A. Arcos, J.N.K. Rao, В.А.Гуревича и многих других исследуются как теоретические аспекты проблемы, так и прикладные вопросы, возникающие в различных приложениях: в радиофизике, статистической радиотехнике, теории надежности, обработке медицинских, социологических, демографических, экономических данных и др.

Но, рассматривая различные виды дополнительной информации, практически все авторы исходят из предположения, что имеющиеся сведения являются достоверными, точными и однозначными. Однако на практике исследователь не всегда может быть абсолютно уверен в полноте и точности априорной информации, особенно когда речь идет об оценках экспертов.

Данная диссертационная работа является логическим продолжением исследований, проводимых на кафедре теоретической кибернетики ТГУ (Ю.Г. Дмитриев, П.Ф. Тарасенко), посвященных проблеме учета при статистической обработке данных многозначной дополнительной информации и информации со смещениями. Многозначной здесь называется информация, заданная в виде конечных множеств возможных значений некоторых функционалов. Смещения появляются, если истинные значения функционалов

не содержатся в заданных множествах. Такую информацию еще называют априорной догадкой.

Цель работы. Построение статистических оценок функционала с учетом многозначности в априорных условиях; исследование свойств этих оценок при конечных объемах наблюдений методом имитационного моделирования.

Методика исследования. При решении поставленных задач применялись методы математического анализа, теории вероятностей, математической статистики и имитационного моделирования на ЭВМ.

Научная новизна работы состоит в

- обобщении постановки задачи условного оценивания на случай разного количества значений дополнительных функционалов, задающих априорную информацию;
- построении оценок, обладающих свойством сходимости к истинному значению оцениваемого параметра в среднеквадратическом;
- построении адаптивной оценки функционала для учета дополнительных условий со смещениями;
- исследовании свойств условных оценок линейного функционала при конечных объемах наблюдений;
- построении оценок функционалов от симметричного распределения с центром симметрии, заданным с точностью до конечного множества значений.

Практическое значение работы состоит в том, что полученные результаты могут быть использованы для построения более точных по среднеквадратической ошибке оценок различных вероятностных характеристик систем или сокращения объема выборки, необходимого для достижения заданной точности оценок, в задачах выборочного контроля качества, обработки технических, социологических и других экспериментальных наблюдений.

Достоверность полученных результатов подтверждается строгими математическими выкладками, проведнными с применением аппарата теории вероятности, математической статистики и теории матриц. Правильность и работоспособность полученных формул подтверждена имитационным моделированием на ЭВМ.

Результаты, выносимые на защиту.

1. Регуляризованные оценки функционала для учета несмещенной априорной информации, доказательство их сходимости в среднеквадратическом.
2. Адаптивная оценка функционала для дополнительных условий со смещениями, сочетающая оценивание параметра с проверкой априорных условий на несмещенность.
3. Результаты исследования свойств оценок при конечных объемах наблюдений, полученные методом имитационного моделирования.

4. Оценки функционалов от симметричного распределения с центром симметрии, заданным с точностью до конечного множества значений, результаты исследования их свойств.

Апробация работы. Работа докладывалась и обсуждалась на научных семинарах кафедры теоретической кибернетики факультета прикладной математики и кибернетики ТГУ, а также на следующих научных конференциях и симпозиумах: VIII Всероссийская научно-практическая конференция "Научное творчество молодежи"(Томск, 2004); V Всероссийский симпозиум по прикладной и промышленной математике (Сочи, 2004, Осенняя сессия); Международная конференция, посвященная 70-летию профессора, доктора физ.-мат. наук Г.А. Медведева (Минск, 2005); III Всероссийская научно-практическая конференция "Информационные технологии и математическое моделирование"(Анжеро-Судженск, 2005); VI Всероссийский симпозиум по прикладной и промышленной математике. Весенняя сессия (С.-Петербург, 2005); I Всероссийская научно-практическая конференция студентов, аспирантов и молодых ученых "Инноватика-2005"(Томск, 2005); VI Всероссийский симпозиум по прикладной и промышленной математике. Осенняя сессия (Сочи, 2005); IV Всероссийская научно-практическая конференция "Информационные технологии и математическое моделирование"(Анжеро-Судженск, 2005);

Публикации. По результатам выполненных исследований опубликовано 12 печатных работ.

Личным вкладом диссертанта в совместные работы является вывод теоретических результатов, разработка вычислительных алгоритмов моделирования и анализ полученных результатов. Постановка изложенных в диссертации задач и формулировка общего подхода к их решению принадлежит научному руководителю соискателя.

Структура и объем диссертации. Работа состоит из введения, четырех глав, заключения, списка использованной литературы и пяти приложений. Объем диссертации без приложений — 156 страниц, иллюстрированных 76 рисунками. Объем приложений — 58 страниц, содержащих 113 таблиц. Список использованной литературы включает 82 наименования.

СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обоснована актуальность темы диссертации, проведен обзор литературы по проблеме учета дополнительной информации в статистическом оценивании, определены цели и методы исследования и сформулированы основные положения, выносимые на защиту.

В первой главе рассматривается задача построения оценки функционала $\theta(P) = \mathrm{M}_P\,\varphi\,(X_1) = \int_{R^n} \varphi(x)P(dx)$ по независимой выборке $X_1, \ldots, X_N$ из неизвестного распределения P на R^n, $n \geqslant 1$, при наличии дополнительной информации о том, что каждый из m других функционалов $b_s(P) = \mathrm{M}_P\,\psi_s\,(X_1) = \int_{R^n} \psi_s(x)P(dx)$, $s = \overline{1,m}$, $m \geqslant 1$, принимает одно из k_s известных значений $\beta_{s1}, \beta_{s2}, \ldots, \beta_{sk_s}$. Здесь полагается, что φ и ψ_s, $s = \overline{1,m}$ — заданные скалярные функции на R^n, числа $k_s \geqslant 1$ — количества возможных значений дополнительных функционалов — могут быть различными для разных s, и для каждого из функционалов $b_s(P)$ известны *все* возможные значения.

В силу последнего условия, если обозначить $\Delta_{st}(P) = b_s(P) - \beta_{st}$ и $\Delta_s(P) = \prod_{t=1}^{k_s} \Delta_{st}(P)$ для всех $t = \overline{1,k_s}$, $s = \overline{1,m}$, то

$$\Delta_s(P) = \int\limits_{R^n} \ldots \int\limits_{R^n} \prod_{t=1}^{k_s} (\psi_s(x_t) - \beta_{st})P(dx_1)\ldots P(dx_{k_s}) = 0. \qquad (1)$$

Таким образом, исходная задача сводится к построению *условных* оценок линейного функционала $\theta(P)$ — с учетом полилинейных условий (1). Подчеркивая равенство нулю в (1), будем называть эти условия *несмещенными*. Вектор $\mathbf{\Delta} = (\Delta_1(P), \ldots, \Delta_m(P))^T$ назовем *вектором смещений*. В рассматриваемом случае $\mathbf{\Delta} \equiv 0$.

Для решения поставленной задачи применяется метод коррелированных процессов В.Н. Пугачева, согласно которому строятся условные оценки вида

$$\theta^{(\boldsymbol{\lambda})} = \theta_N - \sum_{s=1}^{m} \lambda_s \widehat{\Delta}_s = \theta_N - \boldsymbol{\lambda}^T \widehat{\mathbf{\Delta}}, \qquad (2)$$

где $\theta_N = N^{-1} \sum_{i=1}^{N} \varphi(X_i)$ — безусловная эмпирическая оценка $\theta(P)$, $\widehat{\mathbf{\Delta}} = \left(\widehat{\Delta}_1, \ldots, \widehat{\Delta}_m\right)^T$ — оценка вектора смещений $\mathbf{\Delta}$, а вектор коэффициентов $\boldsymbol{\lambda}^T = (\lambda_1, \ldots, \lambda_m)$ выбирается из условия минимума среднеквадратической ошибки оценки $\mathrm{S}_P\,\theta^{(\boldsymbol{\lambda})} = \mathrm{M}_P\left[\theta^{(\boldsymbol{\lambda})} - \theta(P)\right]^2$ (СКО). Вид оптимального коэффициента определяется используемым типом оценок компонент вектора смещений. В диссертационной работе рассматриваются два вида оценок величин $\Delta_s(P)$ — U–статистики U_{Ns} и функционалы Мизеса V_{Ns}.

Оценку $\theta^{(\boldsymbol{\lambda}^*)}$ при коэффициенте $\boldsymbol{\lambda}^*$, доставляющем для фиксированного P минимум СКО на классе оценок U_{Ns} и V_{Ns}, назовем оптимальной.

При использовании U–статистик $\widehat{\mathbf{\Delta}} = \widehat{\mathbf{\Delta}}_U = (U_{N1}, \ldots, U_{Nm})^T$, оценка (2) принимает вид $\theta_U^{(\boldsymbol{\lambda}_U)} = \theta_N - \boldsymbol{\lambda}_U^T \widehat{\mathbf{\Delta}}_U$, а главная часть оптимального коэффициента при условиях

$$\mathrm{M}_P\,\varphi^2\,(X_1) < \infty, \quad \mathrm{M}_P\,\psi_s^2\,(X_1) < \infty, \quad s = \overline{1,m}. \qquad (3)$$

определяется выражением

$$\boldsymbol{\lambda}_U = \boldsymbol{V}_U^{-1}\boldsymbol{C}_U = \boldsymbol{A}^{-1}\boldsymbol{V}^{-1}\boldsymbol{C} - \frac{1}{N-1}\boldsymbol{W}\boldsymbol{A}\boldsymbol{C}, \tag{4}$$

где диагональная $(m \times m)$ матрица $\boldsymbol{A} = \left\|a_s^{(1)}\right\|_{s=\overline{1,m}}$, а также матрицы $\boldsymbol{V} = \left\|\operatorname{cov}_P(\psi_s, \psi_l)\right\|_{s,l=\overline{1,m}}$, $\boldsymbol{V}_U = N^{-1}\left[\boldsymbol{A}\boldsymbol{V}\boldsymbol{A} + (N-1)^{-1}\boldsymbol{W}_U\right]$, $\boldsymbol{W}_U = \left\|2\operatorname{cov}_P^2(\psi_s, \psi_l)\, a_s^{(2)} a_l^{(2)}\right\|_{s,l=\overline{1,m}}$ предполагаются невырожденными; $\boldsymbol{A}^{-1}$, $\boldsymbol{V}^{-1}$, $\boldsymbol{W}_U^{-1}$ и $\boldsymbol{V}_U^{-1}$ соответствующие им обратные матрицы, $\boldsymbol{C}_U = N^{-1}\boldsymbol{A}\boldsymbol{C}$, $\boldsymbol{W} = \left[(\boldsymbol{A}\boldsymbol{V}\boldsymbol{A})\boldsymbol{W}_U^{-1}(\boldsymbol{A}\boldsymbol{V}\boldsymbol{A}) + (N-1)^{-1}\boldsymbol{A}\boldsymbol{V}\boldsymbol{A}\right]^{-1}$, $\boldsymbol{C} = \left\|\operatorname{cov}_P(\varphi, \psi_s)\right\|_{s=\overline{1,m}}$ — вектор-столбец,

$a_s^{(1)} = \sum\limits_{j=1}^{k_s} \prod\limits_{t=1, t\neq j}^{k_s} \Delta_{st}$, $\quad a_s^{(2)} = \sum\limits_{j=1}^{k_s} \sum\limits_{q=1}^{k_s} \prod\limits_{t=1, t\neq j,q}^{k_s} \Delta_{st}$.

При использовании статистик Мизеса получаем оптимальную оценку $\theta_V^{(\boldsymbol{\lambda}_V)} = \theta_N - \boldsymbol{\lambda}_V^T \widehat{\boldsymbol{\Delta}}_V$ с вектором $\widehat{\boldsymbol{\Delta}} = \widehat{\boldsymbol{\Delta}}_V = (V_{N1}, \ldots, V_{Nm})^T$ и коэффициентом, главная часть которого совпадает с главной частью (4) (точное выражение не приводится вследствие его громоздкости).

Эта оценка определена, если

$$\mathrm{M}_P\, \varphi^2(X_1) < \infty, \quad \mathrm{M}_P\, \psi_s^4(X_1) < \infty, \quad s = \overline{1,m}. \tag{5}$$

Вычисление оптимальных коэффициентов с точностью до слагаемых порядка N^{-1} потребовалось для исследования свойств оценок при конечных объемах наблюдений.

В действительности оптимальные коэффициенты, как правило, неизвестны, что затрудняет практическое применение оценок $\theta_U^{(\boldsymbol{\lambda}_U)}$ и $\theta_V^{(\boldsymbol{\lambda}_V)}$. Этот факт приводит к построению соответствующих адаптивных оценок $\widehat{\theta}_U = \theta_N - \widehat{\boldsymbol{\lambda}}_U^T \widehat{\boldsymbol{\Delta}}_U$ и $\widehat{\theta}_V = \theta_N - \widehat{\boldsymbol{\lambda}}_V^T \widehat{\boldsymbol{\Delta}}_V$ с коэффициентами $\widehat{\boldsymbol{\lambda}}_U$ и $\widehat{\boldsymbol{\lambda}}_V$, вычисляемыми путем замены неизвестных матриц $\boldsymbol{A}$, $\boldsymbol{C}$, $\boldsymbol{V}$, $\boldsymbol{W}$, $\boldsymbol{Q}$ и $\boldsymbol{H}$ их эмпирическими оценками $\widehat{\boldsymbol{A}}$, $\widehat{\boldsymbol{C}}$, $\widehat{\boldsymbol{V}}$, $\widehat{\boldsymbol{W}}$, $\widehat{\boldsymbol{Q}}$ и $\widehat{\boldsymbol{H}}$, построенными по исходной выборке.

Теорема 1 *Пусть для распределения P выполняются условия* (3), $\det \boldsymbol{V} \neq 0$ *и* $\sigma_0^2 = \mathrm{D}_P\, \varphi - \boldsymbol{C}^T \boldsymbol{V}^{-1} \boldsymbol{C} > 0$.
Тогда, при $N \to \infty$ $\quad \mathcal{L}\left(\sqrt{N}\left(\widehat{\theta}_U - \theta\right)\right) \to \mathcal{N}\left(0, \sigma_0^2\right)$, *где* $\mathcal{N}\left(0, \sigma_0^2\right)$ — *нормальное распределение с нулевым математическим ожиданием и дисперсией* σ_0^2.

Если выполняются условия (5), аналогичное утверждение справедливо для оценки $\widehat{\theta}_V$: $\quad \mathcal{L}\left(\sqrt{N}\left(\widehat{\theta}_V - \theta\right)\right) \to \mathcal{N}\left(0, \sigma_0^2\right)$.

При конечном объеме наблюдений $\det \widehat{\boldsymbol{V}}$ может принимать нулевое значение с положительной вероятностью, и тогда у адаптивных оценок не существуют моменты. Применение метода кусочно-гладкой аппроксимации оценок позволяет получить регуляризованные оценки, обладающие конечными моментами до второго порядка включительно. Рассмотрены два вида регуляризации:

1) Кусочно-гладкая аппроксимация всей оценки $\widehat{\theta}_U$, приводящая к
$$\tilde{\theta}_U = \frac{N d_U^3 \left(d_U\, \theta_N - \widehat{\boldsymbol{C}}_U^T \widehat{\boldsymbol{V}}_U^* \widehat{\boldsymbol{\Delta}}_U \right)}{N d_U^4 + \left(d_U\, \theta_N - \widehat{\boldsymbol{C}}_U^T \widehat{\boldsymbol{V}}_U^* \widehat{\boldsymbol{\Delta}}_U \right)^4}, \tag{6}$$
где $d_U = \det \widehat{\boldsymbol{V}}_U$ и $\widehat{\boldsymbol{V}}_U^*$ — матрица, присоединенная к $\widehat{\boldsymbol{V}}_U$.

2) Кусочно-гладкая аппроксимация коэффициента $\widehat{\boldsymbol{\lambda}}_U$, приводящая к оценке $\theta_U^{(\tilde{\boldsymbol{\lambda}}_U)} = \theta_N - \tilde{\boldsymbol{\lambda}}_U^T \widehat{\boldsymbol{\Delta}}_U$ с регуляризованным коэффициентом $\tilde{\boldsymbol{\lambda}}_U = \left(\tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_m\right)^T$, где
$$\tilde{\lambda}_s = \frac{N\, d_U^3\, \widehat{\boldsymbol{V}}_s^* \widehat{\boldsymbol{C}}_U}{N\, d_U^4 + \left(\widehat{\boldsymbol{V}}_s^* \widehat{\boldsymbol{C}}_U \right)^4}, s = \overline{1, m}. \tag{7}$$

Теорема 2 *Пусть* $\det \boldsymbol{V} \neq 0$ *и для распределения* P *существуют моменты функций* $\mathrm{M}_P\, \varphi^4(X_1) < \infty$, $\mathrm{M}_P\, \varphi^4(X_1) \psi_s^4(X_1) < \infty$, $\mathrm{M}_P\, \psi_s^4(X_1) \psi_l^4(X_1) < \infty$, *и для всех* $k_s > 3$ $\mathrm{M}_P\, \psi_s^{4(k_s-1)}(X_1) < \infty$; $s, l = \overline{1, m}$. *Тогда при* $N \to \infty$ *регуляризованная оценка* $\tilde{\theta}_U$ *сходится к истинному значению оцениваемого параметра в среднеквадратическом и ее среднеквадратическая ошибка определяется выражением*
$$\mathrm{S}_P\, \tilde{\theta}_U = \mathrm{M}_P \left[\tilde{\theta}_U - \theta \right]^2 = N^{-1} \left[\mathrm{D}_P\, \varphi - \boldsymbol{C}^T \boldsymbol{V}^{-1} \boldsymbol{C} \right] + O\left(N^{-3/2}\right).$$

Аналогичная теорема доказывается для оценки $\theta_U^{(\tilde{\boldsymbol{\lambda}}_U)}$.

Применение метода кусочно-гладкой аппроксимации для регуляризации оценки $\widehat{\theta}_V$, основанной на функционалах Мизеса, приводит к регуляризованной оценке $\tilde{\theta}_V$, подобной (6). Если выполнены условия теоремы 2, $\mathrm{M}_P\, \varphi^4(X_1) \psi_s^8(X_1) < \infty$, $\mathrm{M}_P\, \psi_s^4(X_1) \psi_l^8(X_1) < \infty$, и для всех $k_s > 4$ $\mathrm{M}_P\, \psi_s^{4(k_s-1)}(X_1) < \infty$, $s, l = \overline{1, m}$, то при $N \to \infty$ распределение $\tilde{\theta}_V$ совпадает с распределением оценок с U–статистиками.

Таким образом, разные методы оценивания вектора смещений приводят к построению оценок с одинаковыми асимптотическими свойствами. Оценки, основанные на U–статистиках, гораздо удобнее для аналитических исследований, чем оценки с функционалами Мизеса, вследствие большей простоты и компактности соответствующих выражений, но практическое

применение этих оценок оказывается затруднительным из-за исключительной вычислительной трудоемкости U–статистик. Сравнение свойств оценок при конечных объемах наблюдений проводилось методом имитационного моделирования, описанного в главе 3.

Во **второй главе** рассматривается проблема учета дополнительной информации при наличии смещений в условиях (1).

Такая ситуация возникает, когда исследователю известны *не все* возможные значения дополнительных функционалов, и равенств нулю в (1) может не быть. В этом случае $\Delta_s(P)$, $s = \overline{1,m}$ принимают неизвестные, отличные от нуля значения, *вектор смещений* в априорных условиях $\boldsymbol{\Delta} \not\equiv 0$, причем исследователю неизвестно, какие из компонент вектора отличны от нуля. Те дополнительные условия, для которых равенство нулю в (1) не выполняется, назовем *смещенными.*

Применение метода коррелированных процессов с U–статистиками в качестве оценок компонент вектора смещений приводит к оптимальной оценке $\theta_N^{(\boldsymbol{\lambda}^*)} = \theta_N - (\boldsymbol{\lambda}^*)^T \widehat{\boldsymbol{\Delta}}$ с коэффициентом

$$\boldsymbol{\lambda}^* = \boldsymbol{A}^{-1}\boldsymbol{V}^{-1}\boldsymbol{C} - \frac{\boldsymbol{W}\boldsymbol{A}\boldsymbol{C}}{N-1} - \frac{\boldsymbol{V}_U^{-1}\boldsymbol{\Delta}\boldsymbol{\Delta}^T\boldsymbol{V}_U^{-1}\boldsymbol{A}\boldsymbol{C}}{N^{-1} + \boldsymbol{\Delta}^T\boldsymbol{V}_U^{-1}\boldsymbol{\Delta}} + O\big(N^{-2}\big), \qquad (8)$$

и среднеквадратическим отклонением

$$N\mathrm{S}_P\,\theta_N^{(\boldsymbol{\lambda}^*)} = \mathrm{D}_P\varphi - \boldsymbol{C}^T\boldsymbol{V}^{-1}\boldsymbol{C} + \frac{\boldsymbol{C}^T\boldsymbol{A}\boldsymbol{W}\boldsymbol{A}\boldsymbol{C}}{N-1} + \frac{N\big(\boldsymbol{C}^T\boldsymbol{A}\boldsymbol{V}_U^{-1}\boldsymbol{\Delta}\big)^2}{1+N\boldsymbol{\Delta}^T\boldsymbol{V}_U^{-1}\boldsymbol{\Delta}} + O\big(\frac{1}{N^2}\big) \quad (9)$$

Свойства этой оценки зависят от наличия ненулевых компонент в векторе смещений:

Теорема 3 *Пусть для распределения P выполняются условия* (3), $\det \boldsymbol{V} \neq 0$ *и* $\acute{\sigma}_0^2 > 0$, *где*

$$\acute{\sigma}_0^2 = \begin{cases} \mathrm{D}_P\,\varphi - \boldsymbol{C}^T\boldsymbol{V}^{-1}\boldsymbol{C} & \textit{при } \boldsymbol{\Delta} \equiv 0, \\ \mathrm{D}_P\,\varphi - \boldsymbol{C}^T\boldsymbol{V}^{-1}\boldsymbol{C} + \dfrac{\big(\boldsymbol{C}^T\boldsymbol{A}\boldsymbol{V}_U^{-1}\boldsymbol{\Delta}\big)^2}{\boldsymbol{\Delta}^T\boldsymbol{V}_U^{-1}\boldsymbol{\Delta}} & \textit{при } \boldsymbol{\Delta} \not\equiv 0. \end{cases}$$

Тогда, при $N \to \infty$: $\quad \mathcal{L}\left(\sqrt{N}\big(\theta_N^{(\boldsymbol{\lambda}^*)} - \theta\big)\right) \to \mathcal{N}\left(0, \acute{\sigma}_0^2\right).$

Адаптивная оценка, построенная методом замены неизвестных матриц их выборочными оценками, не сходится по распределению к оптимальной, поэтому строится оценка

$$\widetilde{\theta}_\delta = \theta_N - \widetilde{\boldsymbol{\lambda}}_\delta^T\widehat{\boldsymbol{\Delta}} = \theta_N - \widehat{\boldsymbol{\Delta}}^T\widetilde{\boldsymbol{\lambda}}_\delta, \qquad (10)$$

с коэффициентом

$$\widetilde{\boldsymbol{\lambda}}_\delta = \widehat{\boldsymbol{V}}_U^{-1}\widehat{\boldsymbol{C}}_U - \frac{\widehat{\boldsymbol{V}}_U^{-1}\widetilde{\boldsymbol{\Delta}}\widetilde{\boldsymbol{\Delta}}^T\widehat{\boldsymbol{V}}_U^{-1}\widehat{\boldsymbol{C}}_U}{N^{-1} + \widetilde{\boldsymbol{\Delta}}^T\widehat{\boldsymbol{V}}_U^{-1}\widetilde{\boldsymbol{\Delta}}}. \qquad (11)$$

Компоненты вектора $\widetilde{\boldsymbol{\Delta}}$ вычисляются по формуле

$$\widetilde{\Delta}_s = \widehat{\Delta}_s \left(1 - \frac{\widehat{v}_{ss}}{\widehat{v}_{ss} + N^{\delta}\widehat{\Delta}_s^2}\right), \tag{12}$$

где $\widehat{v}_{ss}$ — эмпирическая оценка s-го диагонального элемента матрицы $\boldsymbol{V}_U$, а $1/2 < \delta < 1$. Особенностью этих оценок является наличие статистики $N^{\delta}\widehat{\Delta}_s^2$, осуществляющей "проверку"равенств $\Delta_s = 0$. Если равенство выполняется, $\widetilde{\Delta}_s$ сходится к нулю с большей скоростью, чем U–статистика $\widehat{\Delta}_s$: $\widetilde{\Delta}_s = O\left(N^{\delta-3/2}\right)$ при $N \to \infty$ и $\Delta_s = 0$.

Пусть r из m априорных условий являются несмещенными, а остальные m-r условий имеют смещения. Без потери общности будем считать несмещенными первые r условий. Тогда

$$\boldsymbol{\Delta}^T = \left(\boldsymbol{\Delta}_{[1]}^T, \boldsymbol{\Delta}_{[2]}^T\right),$$

где $\boldsymbol{\Delta}_{[1]} = (\Delta_1, \ldots, \Delta_r)^T \equiv 0$, а $\boldsymbol{\Delta}_{[2]} = (\Delta_{r+1}, \ldots, \Delta_m)^T$, состоит из компонент, не равных нулю, $r = \overline{0, m}$.

В соответствии с этим векторы и матрицы, используемые при построении $\widetilde{\theta}_{\delta}$, могут быть разбиты на блоки: $\boldsymbol{C}_U^T = \left(\boldsymbol{C}_{[1]}^T, \boldsymbol{C}_{[2]}^T\right)$,

$$\boldsymbol{V}_U = \left\| \begin{array}{cc} \boldsymbol{V}_{11} & \boldsymbol{V}_{12} \\ \boldsymbol{V}_{21} & \boldsymbol{V}_{22} \end{array} \right\|, \quad \boldsymbol{V}_U^{-1} = \left\| \begin{array}{cc} \boldsymbol{Y}_{11} & \boldsymbol{Y}_{12} \\ \boldsymbol{Y}_{21} & \boldsymbol{Y}_{22} \end{array} \right\| = \left\| \begin{array}{c} \boldsymbol{Y}_{[1]} \\ \boldsymbol{Y}_{[2]} \end{array} \right\|.$$

Теорема 4 *Пусть для распределения P выполняются условия* (3), $\det \boldsymbol{V} \neq 0$ *и* $1/2 < \delta < 1$. *Тогда, при $N \to \infty$ адаптивная оценка $\widetilde{\theta}_{\delta}$ асимптотически нормальна:* $\mathcal{L}\left(\sqrt{N}\left(\widetilde{\theta}_{\delta} - \theta\right)\right) \to \mathcal{N}\left(0, \sigma_{[1]}^2\right)$, *где*

$$\begin{gathered} \sigma_{[1]}^2 = \mathrm{D}_P\,\varphi - \boldsymbol{C}^T\boldsymbol{V}^{-1}\boldsymbol{C} + \frac{\left(\boldsymbol{C}_U^T\boldsymbol{Y}_{[2]}^T\boldsymbol{\Delta}_{[2]}\right)^2}{\boldsymbol{\Delta}_{[2]}^T\boldsymbol{Y}_{22}\boldsymbol{\Delta}_{[2]}} + \\ + \boldsymbol{C}_U^T\boldsymbol{Y}_{[2]}^T\left[\boldsymbol{V}_{22} - \frac{\boldsymbol{H}\boldsymbol{V}_{22} + \boldsymbol{V}_{22}\boldsymbol{H}^T}{\boldsymbol{\Delta}_{[2]}^T\boldsymbol{Y}_{22}\boldsymbol{\Delta}_{[2]}} + \frac{\boldsymbol{H}\boldsymbol{V}_{22}\boldsymbol{H}^T}{(\boldsymbol{\Delta}_{[2]}^T\boldsymbol{Y}_{22}\boldsymbol{\Delta}_{[2]})^2}\right]\boldsymbol{Y}_{[2]}\boldsymbol{C}_U, \\ \boldsymbol{H} = \boldsymbol{\Delta}_{[2]}\boldsymbol{\Delta}_{[2]}^T\boldsymbol{Y}_{22}. \end{gathered} \tag{13}$$

Следствие. В случае полностью несмещенных априорных условий, т.е. при $\boldsymbol{\Delta}_{[1]} = \boldsymbol{\Delta} \equiv 0$, $\sigma_{[1]}^2 = \acute{\sigma}_0^2$ (для $\boldsymbol{\Delta} \equiv 0$). Если же все компоненты вектора смещений – не нулевые, т.е. $\boldsymbol{\Delta}_{[2]} = \boldsymbol{\Delta}$, $\sigma_{[1]}^2 = \mathrm{D}_P\,\varphi$.

Кроме того, если только одно из априорных условий является смещенным ($\boldsymbol{\Delta}_{[2]} = \Delta_m$) при любом m, то асимптотическая дисперсия адаптивной оценки $\sigma_{[1]}^2$ совпадает с дисперсией оптимальной оценки $\acute{\sigma}_0^2$ для ненулевого вектора смещений.

Таким образом, построенная адаптивная оценка $\widetilde{\theta}_\delta$ сочетает оценивание неизвестного параметра $\theta(P)$ с проверкой априорных условий на несмещенность, при этом она имеет асимптотически нормальное распределение с дисперсией, соответствующей несмещенным условиям.

В **третьей главе** проводится имитационное моделирование построенных в предыдущих главах оценок с целью исследования их свойств при конечных объемах наблюдений.

Все численные эксперименты проводились для функционалов, определенных через c-функции на основе псевдослучайных выборок из стандартного нормального распределения, сгенерированных средствами пакета Statistica 6.0. Результаты представлены в виде таблиц и графиков, отображающих значения среднеквадратических ошибок безусловных, адаптивных, регуляризованных и оптимальных оценок функционала $\theta(F) = \theta(\Phi) = \Phi(z_0)$, $z_0 = \overline{-1,1}$ при различных дополнительных условиях и объемах исходных выборок. Дополнительные функционалы задавались в виде $b_s(F) = b_s(\Phi) = \Phi(z_s)$.

Анализ полученных результатов показал:

1. Рассмотренные процедуры условного оценивания позволяют получить оценки, обладающие меньшими среднеквадратическими ошибками, чем безусловные, при конечных объемах наблюдений (до 100 включительно), несмотря на неопределенность в задании дополнительной информации (2-3 возможных значения). В некоторых (наилучших) случаях точность построенных оценок оказалась в 2-3 раза выше, чем у безусловной, при $N = 25$, и до 10 раз выше при $N = 100$.
2. Согласно полученным аналитически выражениям, среднеквадратическая ошибка оценок, построенных с учетом дополнительной информации, зависит от вида оценки, объема выборки и от "ценности"информации, которая определяется величиной ковариации между подынтегральными функциями φ и ψ_s. При конечных объемах наблюдений значительное влияние на точность условных оценок оказывают и другие особенности привлекаемой дополнительной информации:

а) Число возможных значений каждого из функционалов $b_s(F)$ – числа k_s. Большее количество значений функционала может трактоваться как более высокая неопределенность в имеющейся дополнительной информации, которая существенно уменьшает выигрыш в точности оценивания.

б) Количество учитываемых дополнительных функционалов – величина m. Каждое из дополнительных условий содержит информацию об оцениваемом распределении, привлечение которой позволяет снизить неопределенность, обусловленную множественностью заданных значений функ-

ционалов, и повысить точность оценок.

в) Расхождение между возможными значениями дополнительных функционалов – абсолютные значения величин Δ_{st}, $s = \overline{1,m}$, $t = \overline{1,k_s}$. Чем дальше друг от друга расположены заданные значения, тем меньше СКО полученных оценок. Это свойство объясняется особенностями применяемого алгоритма: далеко расположенные значения четко различимы, поэтому истинное значение распознается лучше, характеристики адаптивных и регуляризованных оценок с ростом объема исходной выборки сходятся к оптимальным значениям значительно быстрее. В случае близких значений повышение точности условной оценки может быть незначительным даже при известном оптимальном коэффициенте.

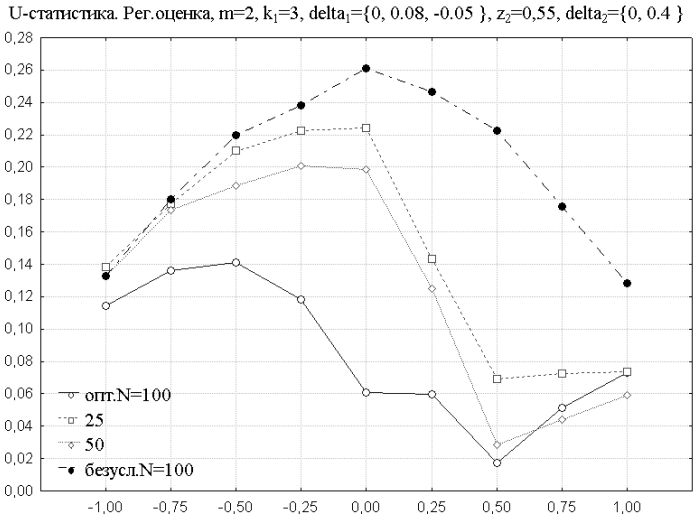

Рис. 1: СКО оценки $\widetilde{\theta}_U$ при $m = 2$, $k_1 = 3$, $k_2 = 2$

Графики, приведенные на рисунке 1, позволяют сравнить среднеквадратические ошибки регуляризованной оценки $\tilde{\theta}_U$ с двумя дополнительными условиями при 25 и 50 наблюдениях с СКО безусловной и оптимальной оценок при $N = 100$ (по оси абсцисс указаны значения z_0).

3. Замена оптимальных коэффициентов эмпирическими оценками их главных частей (адаптация) заметно ухудшает точность оценок, особенно при небольших объемах наблюдений ($N = 25$). Негативное влияние адаптации убывает с ростом N, но скорость убывания зависит от расстояния между заданными значениями дополнительных функционалов $\beta_{s1}, \beta_{s2}, \ldots, \beta_{sk_s}$.

4. Регуляризация оценок не снижает их точность по сравнению с адаптивными, а в некоторых случаях приводит к заметному уменьшению СКО оценок.
5. Применение формул, учитывающих возможные ненулевые смещения, даже к полностью несмещенным условиям заметно (на 5-20%) снижает точность оценок при конечных объемах наблюдений (до 100) по сравнению с ранее рассмотренными условными оценками. Учет смещенных априорных условий, в некоторых случаях, приводит к оценке даже менее точной, чем безусловная (при указанных объемах наблюдений). Графики СКО оценок, учитывающих два дополнительных условия, одно из которых – смещенное, приведены на рисунке 2.

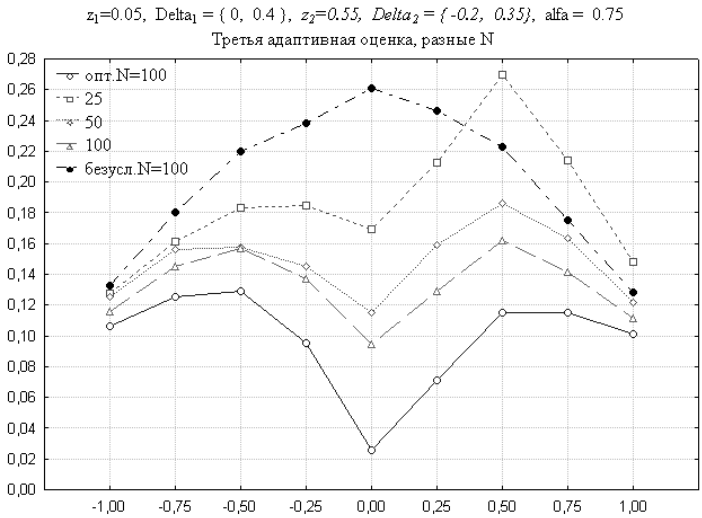

Рис. 2: СКО оценки $\widetilde{\theta}_\delta$ при $m = 2$, $k_1 = k_2 = 2$ и разных N

В **четвертой главе** метод коррелированных процессов применяется для привлечения дополнительной информации о том, что оцениваемое распределение является симметричным относительно одной из нескольких заданных точек.

Рассматривается задача оценивания линейного функционала $\theta(F) = \mathrm{M}_F\,\varphi(X) = \int_{R^1} \varphi(x)dF(x)$ от заданной скалярной функции $\varphi(x)$ на R^1 по результатам N независимых наблюдений $X_1, \ldots, X_N$ над случайной величиной X с функцией распределения $F(x)$, симметричной относительно точки α, принимающей одно из m заданных значений: $\alpha \in \{\alpha_i\}_{i=\overline{1,m}}$ и

$F(x) = 1 - F(2\alpha - x)$, $\forall x \in R^1$.

При известном центре симметрии, то есть при $m = 1$ и $\alpha = \alpha_1$, для оценивания $\theta(F)$ с учетом дополнительной информации применяется несмещенная оценка

$$\theta_{N\alpha} = \int\limits_{R^1} \frac{\varphi(x) + \varphi(2\alpha - x)}{2} dF_N(x) = \sum_{i=1}^{N} \frac{\left[\varphi(X_i) + \varphi(2\alpha - X_i)\right]}{2N} \qquad (14)$$

с дисперсией

$$\mathrm{D}_F\, \theta_{N\alpha} = \frac{\mathrm{D}_F\, \varphi(X) - \mathrm{cov}_F\big(\varphi(X), \varphi(2\alpha - X)\big)}{2\,N} = \frac{\sigma_0^2}{N},$$

не превышающей дисперсию безусловной эмпирической оценки (здесь $F_N(x)$ — эмпирическая функция распределения).

Для случая $m > 1$, то есть если центр симметрии задан с точностью до конечного множества значений, рассматриваются два подхода, основанные на методе коррелированных процессов, приводящие к двум разным типам оценок параметра $\theta(F)$.

Согласно первому подходу, для каждого из m заданных значений центра симметрии определяется вспомогательная функция

$$\psi_i(x) = \frac{\varphi(x) - \varphi(2\alpha_i - x)}{2}$$

и величина $\Delta_i = \int_{R^1} \psi_i(x) d\,F(x)$, принимающая нулевое значение, если $\alpha_i = \alpha$. Таким образом, имеющаяся дополнительная информация о функции распределения сводится к равенству

$$\prod_{i=1}^{m} \Delta_i = \int\limits_{R^1} \ldots \int\limits_{R^1} \psi_1(x_1) \cdots \psi_m(x_m)\, d\,F(x_1) \cdots d\,F(x_m) = 0. \qquad (15)$$

Если $\mathrm{M}_F\, \varphi^2(X) < \infty$, то оценка $\theta_\alpha^{(\lambda^*)} = \theta_N - \lambda^* \cdot U_N$ с коэффициентом

$$\lambda^* = \frac{\sum_{i=1}^{m} \Delta^{(i)}\ \mathrm{cov}_F(\varphi, \psi_i)}{\sum_{i=1}^{m} \left(\Delta^{(i)}\right)^2 \mathrm{D}_F\, \psi_i + N^{-1}\, Q},$$

где

$$Q = \Big(\sum_{i=1}^{m} \sum_{\substack{j=1 \\ j>i}}^{m} \Delta^{(ij)}\, \mathrm{cov}_F(\psi_i, \psi_j)\Big)^2 + \sum_{i=1}^{m} \sum_{\substack{j=1 \\ j>i}}^{m} \left(\Delta^{(ij)}\right)^2 \mathrm{D}_F\, \psi_i\ \mathrm{D}_F\, \psi_j +$$

$$+ 2 \sum_{i=1}^{m} \sum_{\substack{j=1 \\ j\neq i}}^{m} \sum_{\substack{p\neq i \\ p>j}}^{m} \Delta^{(ij)}\ \Delta^{(ip)}\ \mathrm{D}_F\, \psi_i\ \mathrm{cov}_F(\psi_j, \psi_p) + O(N^{-1});$$

$$\Delta^{(i)} = \prod\nolimits_{k\neq i}^{m} \Delta_k; \qquad \Delta^{(ij)} = \prod\nolimits_{k\neq i,j}^{m} \Delta_k,$$

является оптимальной с точки зрения минимума дисперсии.

Свойства адаптивной оценки $\widehat{\theta}_\alpha = \theta_N - \widehat{\lambda} \cdot U_N$, полученной заменой оптимального коэффициента λ^* на эмпирическую оценку его главной части $\widehat{\lambda}$, определяются теоремой

Теорема 5 *Пусть* $\mathrm{M}_F\, \varphi^2(X) < \infty$ *и*

$$\sigma_0^2 = \frac{1}{2}\left[\mathrm{D}_F\, \varphi(X) - \mathrm{cov}_F\left(\varphi(X), \varphi(2\alpha - X)\right)\right] > 0.$$

Тогда при $N \to \infty$

1. $\widehat{\theta}_\alpha \xrightarrow{p} \theta(F)$

2. $\mathcal{L}\left(\sqrt{N}(\widehat{\theta}_\alpha - \theta)\right) \to \mathcal{N}\left(0, \sigma_0^2\right)$*, где* $\mathcal{N}\left(0, \sigma_0^2\right)$ *- нормальное распределение с нулевым математическим ожиданием и дисперсией* σ_0^2.

Оценка $\widehat{\theta}_{\alpha V} = \theta_N - \widehat{\lambda}' \cdot V_N$, в которой для оценивания условия (15) вместо U–статистики используется функционал Мизеса, обладает такими же асимптотическими свойствами.

Второй подход к построению условных оценок функционала от симметричного распределения заключается в том, что метод коррелированных процессов применяется для оценивания центра симметрии α с учетом дополнительной информации о его возможных значениях: $\alpha \in \{\alpha_i\}_{i=\overline{1,m}}$.

Оценивая функционал $\alpha(F) = \mathrm{M}_F\, X_1 = \int_{R^1} x\, dF(x)$ при условии

$$\prod_{i=1}^{m}(\alpha - \alpha_i) = \int_{R^1} \ldots \int_{R^1} (x_1 - \alpha_1) \ldots (x_m - \alpha_m)\, dF(x_1) \ldots dF(x_m) = 0,$$

получаем оценку центра симметрии вида (2) с оптимальным коэффициентом

$$\gamma^* = \left[a_1 + \frac{2\, a_2^2\ \mathrm{D}_F X}{(N-1)\, a_1}\right]^{-1} + O\left(N^{-2}\right),$$

где $a_1 = \sum_{j=1}^{m} \prod_{i \neq j}^{m} (\alpha - \alpha_i)$, $\quad a_2 = \sum_{j=1}^{m} \sum_{k=1}^{m} \prod_{i \neq j,k}^{m} (\alpha - \alpha_i)$.

Соответствующая адаптивная оценка

$$\alpha_N = \widehat{\alpha} - \widehat{\gamma} \cdot U_\alpha, \tag{16}$$

при $\mathrm{D}_F X = \mathrm{M}_F (X - \alpha)^2 < \infty$ и $N \to \infty$ сходится к оцениваемому значению быстрее, чем обычная эмпирическая оценка: $\sqrt{N}\left(\alpha_N - \alpha\right) \xrightarrow{p} 0$.

Подставляя (16) вместо неизвестного центра симметрии α в (14), получим оценку

$$\widehat{\theta}_{\alpha_N} = \frac{1}{2N} \sum_{i=1}^{N} \left[\varphi(X_i) + \varphi(2\alpha_N - X_i)\right], \tag{17}$$

асимптотические свойства которой совпадают со свойствами оценок первого типа, если функция $\varphi(x)$ непрерывна в окрестности точки α и имеет непрерывные производные до второго порядка включительно, $\mathrm{M}_F\, \varphi^2(X) < \infty$, $\mathrm{M}_F\, \varphi'(X) < \infty$, $\mathrm{M}_F\, \varphi''(X) < \infty$.

Исследование свойств оценок при конечных объемах наблюдений проводилось с помощью численного моделирования в системе Statistica 6.0, при котором адаптивные оценки первого и второго типов сравнивались с безусловной и оценкой с известным центром симметрии, вычисленными для функционала $\theta(F) = \int_{R^1} e^x \, dF(x)$ и стандартного нормального распределения с учетом дополнительной информации о двух или трех возможных значениях центра симметрии.

Результаты моделирования показали, что различия в точности условных оценок, обусловленные их видом и влиянием адаптации, быстро уменьшаются с ростом объема наблюдений и становятся несущественными при$N > 50$. На рисунке 3 приведены графики среднеквадратических ошибок рассмот-

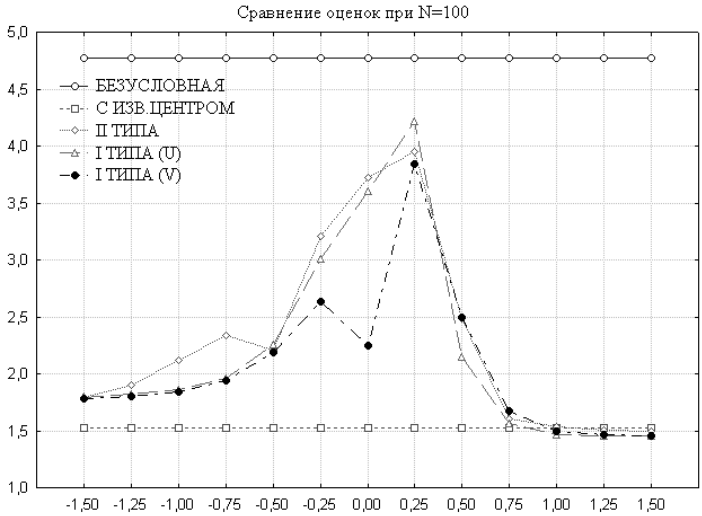

Рис. 3: NСКО разных оценок для $N = 100$

ренных оценок для $N = 100$ и двух возможных значений центра симметрии: $\alpha \in \{\alpha_1,\ \alpha_2\}$, где $\alpha_1 = \alpha = 0$, а величина α_2, отображенная по оси абсцисс, принимает значения в диапазоне $[-1.5, 1.5]$.

В **приложениях** приведены таблицы с данными, полученными в результате имитационного моделирования, описанного в третьей и четвертой главах работы.

ОСНОВНЫЕ ВЫВОДЫ И РЕЗУЛЬТАТЫ РАБОТЫ

В настоящей работе рассмотрена задача условного оценивания функционала при наличии многозначности и смещений в априорных условиях. Получены следующие результаты:

1. Постановка задачи обобщена на случай разного количества возможных значений дополнительных функционалов.
2. Для несмещенных априорных условий найдены оптимальные в смысле минимума среднеквадратической ошибки оценки двух типов — на основе U–статистик и функционалов Мизеса и соответствующие адаптивные оценки — асимптотически нормальные и сходящиеся по распределению к оптимальным.
3. Методом кусочно-гладкой аппроксимации оценок построены регуляризованные оценки, сходящиеся к оцениваемому значению функционала в среднеквадратическом.
4. Для априорных условий со смещениями построена адаптивная оценка, асимптотическая дисперсия которой соответствует несмещенным условиям, если они присутствуют. Показано, что при наличии смещения не более чем в одном условии эта оценка эквивалентна оптимальной (в смысле слабой сходимости), а если все априорные условия имеют смещения, ее предельное распределение совпадает с распределением безусловной оценки.
5. Имитационное моделирование в системе Statistica 6.0 показало, что исследуемые алгоритмы условного оценивания при конечных объемах наблюдений (до 100) позволяют существенно (до 10 раз) повысить точность оценок при несмещенных априорных условиях, несмотря на многозначность в привлекаемой информации.

   Сравнение доасимптотических свойств оценок при разных распределениях и подынтегральных функциях, а также детальное исследование свойств оценок со смещениями в априорных условиях не проводилось вследствие технической сложности. Эти вопросы требуют дополнительных исследований с применением другого программного и аппаратного обеспечения.
6. Для решения задачи условного оценивания функционала от симметричной функции распределения с центром симметрии, заданным с точностью до конечного множества значений, построены адаптивные оценки, асимптотические распределения которых совпадают с предельным распределением оценки с известным центром симметрии. Проведенное имитационное моделирование показало, что для выбранных условий раз-

личия в точности построенных оценок становятся несущественными при $N > 50$.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Результаты работы были опубликованы в следующих статьях и материалах научных конференций:

1. Головчинер О.Н., Дмитриев Ю.Г. Об условной оценке функционала // Вестник Томского государственного университета. Приложение. - 2004.- № 9 (II). - с.145-150.
2. Головчинер О.Н., Дмитриев Ю.Г. Оценивание функционалов от распределений с учетом априорных догадок // Обозрение прикладной и промышленной математики. - 2004. - Т.11. - вып.4. - с.785-786.
3. Головчинер О.Н., Дмитриев Ю.Г. О сходимости в среднеквадратическом оценки функционала // Материалы VIII Всероссийской научно-практической конференции "Научное творчество молодежи". Ч.1. - Томск: Изд-во ТГУ, 2004, с.24-25.
4. Головчинер О.Н., Дмитриев Ю.Г. Условное оценивание функционала на основе U-статистик. // Теория вероятностей, случайные процессы, математическая статистика и приложения: сборник научных статей международной конференции, посвященной 70-летию проф. Медведева. - Минск: Изд-во БГУ, 2005. - с.52-59.
5. Головчинер О.Н., Дмитриев Ю.Г. Оценивание линейного функционала при смещениях в априорных условиях // Обозрение прикл. и промышленной математики. - 2005. - Т.12. - вып.1. - с.138-139.
6. Головчинер О.Н., Дмитриев Ю.Г. Статистики Мизеса в условном оценивании линейного функционала // Обозрение прикладной и промышленной математики. - 2005. - Т.12. - вып. 4. - с.935-936
7. Головчинер О.Н., Дмитриев Ю.Г. Об оценке функционала при наличии смещений в априорных условиях // Вестник Томского государственного университета. Приложение. - 2005. - № 14. - с.280-285
8. Головчинер О.Н., Дмитриев Ю.Г. Условное оценивание функционала на основе статистики Мизеса // Материалы III Всероссийской научно-практической конференции "Информационные технологии и математическое моделирование"Анжеро-Судженск. 2005. - Томск: Изд-во ТГУ, 2005. - Ч.2. - с.11-12.
9. Головчинер О.Н. Об условной оценке доли объектов // Инноватика-2005: сб. материалов I Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых. - Томск: Изд-во ТГУ, 2005. - с.22-24.

10. Головчинер О.Н. Статистическое моделирование условных оценок функционалов // Материалы IV Всероссийской научно-практической конференции "Информационные технологии и математическое моделирование", ч.2 - Томск: Изд-во ТГУ, 2005. - с.6-8.
11. Головчинер О.Н., Дмитриев Ю.Г. Об оценке функционала от симметричного распределения // Вестник Томского государственного университета. Приложение. - 2006. - № 17. - с.280-285.
12. Головчинер О.Н., Дмитриев Ю.Г. Статистическое оценивание функционала с учетом симметрии распределения // Вестник Томского государственного университета. Серия "Информатика. Кибернетика. Математика". - 2006. - № 293. - с.84-88.