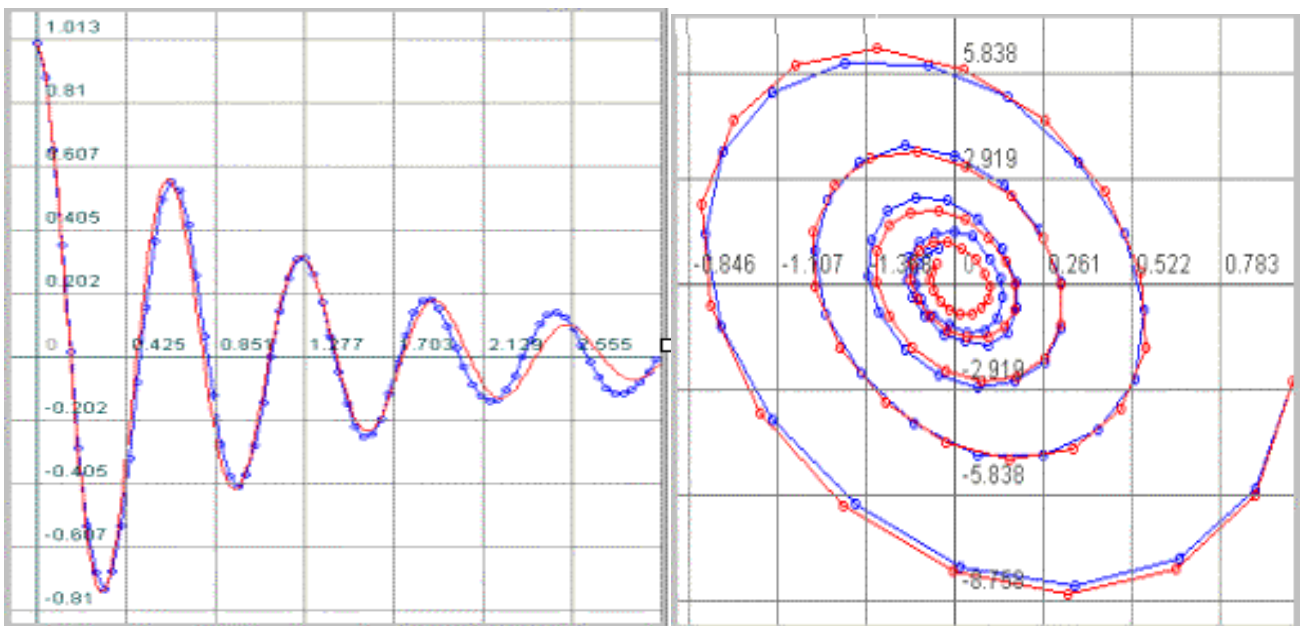


**С.А. Прохоров**

# **АПРОКСИМАТИВНЫЙ АНАЛИЗ СЛУЧАЙНЫХ ПРОЦЕССОВ**



**Самара 2001 г.**

## *Об авторе*



**Прохоров Сергей Антонович** – доктор технических наук, профессор, академик Международной академии информатизации, член-корреспондент Российской академии естественных наук, заслуженный работник высшей школы Российской Федерации, декан факультета информатики, заведующий кафедрой информационных систем и технологий Самарского государственного аэрокосмического университета.

В качестве председателя Головного Совет Минвуза России по автоматизации научных исследований в период 1988-1996 г.г. руководил разработкой и выполнением шести научно-исследовательских программ и подпрограмм АН СССР, Минэлектронпрома СССР, Минвуза России.

Результаты его работы нашли отражение в 180 научных трудах, в том числе, 5 монографиях, 40 авторских свидетельствах, 6 работах опубликованных за рубежом, выступлениях более чем на 50 международных, всесоюзных и республиканских конференциях и симпозиумах.

УДК 681.518.3, 514:681.323/043.3/

ББК 32.965

П 4

Рецензенты:

Заслуженный деятель науки РФ, член-корреспондент РАН, д. т. н., профессор  
Сойфер В.А.;

д. ф.-м. н., профессор Жданов А.И.

**Прохоров С.А.**

**П 4** Аппроксимативный анализ случайных процессов:

Самарский государственный аэрокосмический университет, 2001. 329 с.: ил.

**ISBN 9965-01-958-4**

Рассматриваются методы и алгоритмы, представляющие собой развитие идей метода наименьших квадратов при аппроксимации различных функциональных вероятностных характеристик, обладающих рядом специфических свойств.

Анализируются методы, алгоритмы идентификации случайных процессов, аппроксимации законов распределения, корреляционных функций и спектральных плотностей мощности параметрическими моделями, представляющими собой как функции заданного вида, так и ортогональные функции Лагерра.

Рассматриваются вопросы аппаратурной, аппаратно-программной и программной реализации разработанных алгоритмов аппроксимации, даются рекомендации по расчёту параметров разработанных средств и программного обеспечения.

Приводится описание разработанных автоматизированных информационных систем для аппроксимативного анализа случайных процессов.

Предназначена для научных сотрудников, инженеров, аспирантов и студентов как руководство по основам аппроксимативного анализа случайных процессов.

Издание монографии поддержано грантом 4-Г/2001 76/01/Б по программе «Поддержка важнейших разработок научных коллективов и отдельных учёных на основе системы грантов» по разделу «Важнейшие научные и технические разработки, соответствующие концепции социально-экономического развития Самары».

**ББК 32.965**

Печатается по решению издательского совета Самарского научного центра Российской академии наук.

**П 2402000000  
00 (05) - 01**

**ISBN 9965-01-958-4**

© Самарский научный центр Российской академии наук, 2001

© Самарский государственный аэрокосмический университет

© С.А. Прохоров, 2001

СОДЕРЖАНИЕ	Стр.
СПИСОК СОКРАЩЕНИЙ	6
ПРЕДИСЛОВИЕ	9
ВВЕДЕНИЕ	11
1. АППРОКСИМАЦИЯ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ	25
1.1. Основные понятия и определения	25
1.2. Критерии согласия	28
1.3. Идентификация законов распределения	31
1.4. Метод моментов	35
1.5. Аппроксимация плотностей распределения вероятностей параметрическими моделями	39
1.6. Аппроксимация функций распределения вероятностей параметрическими моделями	47
1.7. АИС для аппроксимативного анализа законов распределения случайных процессов	53
1.7.1. Описание подсистем	53
1.7.2. Описание интерфейса системы	56
1.7.3. Порядок работы с системой	65
2. АППАРАТУРНАЯ АППРОКСИМАЦИЯ КОРРЕЛЯЦИОННЫХ ФУНКЦИЙ ФУНКЦИЯМИ ЗАДАННОГО ВИДА	67
2.1. Основные понятия и определения	67
2.2. Метод аппроксимации корреляционных функций функциями заданного вида	75
2.3. Анализ методических погрешностей аппроксимации корреляционных функций	82
2.4. Анализ статистических погрешностей аппроксимации корреляционных функций	85
2.5. Влияние неточности оценки параметров модели на погрешность аппроксимации корреляционных функций	90
2.6. Принципы построения основных блоков коррелометров с аппроксимацией	91
3. АППАРАТУРНАЯ АППРОКСИМАЦИЯ КОРРЕЛЯЦИОННЫХ ФУНКЦИЙ ОРТОГОНАЛЬНЫМИ ФУНКЦИЯМИ ЛАГЕРРА	98
3.1. Основные понятия и определения	98
3.2. Алгоритм оценки параметров функций Лагерра, обеспечивающий минимум квадратической погрешности аппроксимации	102
3.3. Алгоритм оценки параметра функций Лагерра, обеспечивающий минимум погрешности аппроксимации при ограничениях на модель корреляционной функции	106
3.4. Упрощенный алгоритм оценки параметра функции Лагерра	113
3.5. Применение экспоненциальной аппроксимации корреляционной функции для определения параметра ортогональных функций Лагерра	117
3.6. Минимизация статистической погрешности в коррелометрах с аппроксимацией ортогональными функциями Лагерра	122
3.7. Анализ погрешности коррелометра с аппроксимацией ортогональными функциями Лагерра	126

3.8.	Аппроксимация взаимных корреляционных функций	131
4.	<b>АППРОКСИМАЦИЯ КОРРЕЛЯЦИОННЫХ ФУНКЦИЙ ПАРА- МЕТРИЧЕСКИМИ МОДЕЛЯМИ С ПОМОЩЬЮ ЭВМ</b>	134
4.1.	Основные понятия и определения	134
4.2.	Идентификация корреляционных функций	136
4.3.	Аппроксимация корреляционных функций функциями заданного вида методом Ньютона (с аналитическим взятием производных)	143
4.4.	Аппроксимация корреляционных функций функциями заданного вида методом Ньютона (с конечно-разностными производными)	148
4.5.	Аппроксимация корреляционных функций функциями заданного вида методом деформированного многогранника	152
4.6.	Аппроксимация корреляционных функций функциями заданного вида методом деформированного многогранника	157
4.7.	Аппроксимация корреляционных функций ортогональными функциями Лагерра	162
4.8.	Аппроксимация корреляционных функций ортогональными функциями Лагерра	165
4.9.	Аппроксимация взаимных корреляционных функций параметрическими моделями	168
	Анализ сходимости методов аппроксимации корреляционных функций параметрическими моделями	
5.	<b>АППРОКСИМАЦИЯ КОРРЕЛЯЦИОННЫХ ФУНКЦИЙ НЕЭКВИ- ДИСТАНТНЫХ ВРЕМЕННЫХ РЯДОВ</b>	171
5.1.	Основные понятия и определения	171
5.2.	Метод и алгоритмы корреляционного анализа неэвидистантных временных рядов с использованием интервальной корреляционной функции	176
5.3.	Аппроксимативный метод и алгоритмы измерения корреляционных функций неэвидистантных временных рядов	186
5.4.	Анализ методических погрешностей результатов измерений вероятностных характеристик неэвидистантных временных	196
5.5.	рядов	198
5.6.	Математическое описание неэвидистантных временных рядов	206
5.7.	Процессорные статистические анализаторы	
	Аппроксимация корреляционных функций неэвидистантных	213
5.8.	временных рядов функциями заданного вида	
	Аппроксимация корреляционных функций неэвидистантных	221
	временных рядов ортогональными функциями Лагерра	
6.	<b>АППРОКСИМАТИВНЫЙ АНАЛИЗ КОРРЕЛЯЦИОННО- СПЕКТРАЛЬНЫХ ХАРАКТЕРИСТИК</b>	227
6.1.	Оценка обобщенных корреляционных характеристик	227
6.1.1.	Оценка показателя колебательности	227
6.1.2.	Оценка интервала корреляции	229
6.1.3.	Оценка моментов корреляционных функций	240
6.1.4.	Оценка обобщенных характеристик взаимной корреляционной функции	242
6.2.	Аппроксимация спектральных плотностей мощности	245
6.3.	Оценка обобщенных спектральных характеристик	257

7.	АВТОМАТИЗИРОВАННЫЕ СИСТЕМЫ ДЛЯ АППРОКСИМАТИВНОГО КОРРЕЛЯЦИОННО-СПЕКТРАЛЬНОГО АНАЛИЗА	261
7.1.	Аппроксимативный анализ автокорреляционно-спектральных характеристик	261
7.2.	Аппроксимативный анализ взаимных корреляционно-спектральных характеристик	276
7.3.	Анализ гидрологических параметров в открытой части Балтийского моря	284
	ЗАКЛЮЧЕНИЕ	292
	СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	293
	ПРИЛОЖЕНИЯ	299
П. 1.	Типовые законы распределения	299
П. 2.	Обратные функции законов распределения	304
П. 3.	Характеристики типовых законов распределения ( $P_d=0,999$ )	305
П. 4.	Таблица $\chi^2$ распределения	308
П. 5.	Таблица значений коэффициента $\lambda$ к критерию Колмогорова	308
П. 4.	Производные для однопараметрических плотностей вероятностей	309
П. 7.	Производные для двухпараметрических плотностей распределения вероятностей	310
П. 8.	Производные для однопараметрических функций распределения вероятностей	312
П. 9.	Производные для двухпараметрических функций распределения вероятностей	313
П.10.	Импульсные характеристики формирующих фильтров для генерирования псевдослучайных последовательностей с заданным видом корреляционных функций методом нерекурсивной фильтрации	315
П.11.	Импульсные характеристики формирующих фильтров для генерирования псевдослучайных последовательностей с заданным видом корреляционных функций методом рекурсивной фильтрации	316
П.12.	Параметрические модели корреляционных функций	317
П.13.	Ортогональные функции Лагерра	320
П.14.	Параметры моделей корреляционных функций в ортогональном базисе Лагерра	321
П.15.	Обобщенные спектральные характеристики	325
П.16.	Блок-схема автоматизированной информационной системы аппроксимативного корреляционно-спектрального анализа	327

## СПИСОК СОКРАЩЕНИЙ

- АЦП - аналого-цифровой преобразователь;  
БРП - блок регулировки параметров;  
ВУ - вычитающее устройство;  
ИП - измерительный преобразователь (датчик);  
К - коммутатор;  
КВ - квадратор;  
КФ - корреляционная функция;  
НКФ – нормированная корреляционная функция;  
М - количество дифференциальных коридоров;  
МУ - масштабирующее устройство;  
МУ - множительное устройство;  
НКФ - нормированная корреляционная функция;  
ПО - процессор обработки;  
УУ - усредняющее устройство;  
ФНЧ - фильтр нижних частот;  
 $A_{\Theta} \{ \epsilon(t_k) \}$  - алгоритм оценки вероятностной характеристики  $\Theta$ ;  
 $A_s \{ \epsilon(t_k) \}$  - алгоритм оценки сигнала;  
 $C_x(J)$  - интервальная корреляционная функция;  
 $d$  - параметр усреднения (время  $T$ , совокупность реализаций  $N$  или время и совокупность реализаций  $TN$ );  
 $\hat{D}_x$  - оценка дисперсии;  
 $\text{ent}[\ ]$  - операция взятия целой части числа;  
 $F\{ \}$  - функциональное преобразование полученных оценок вероятностных характеристик;  
 $|F''(x)|_{\max}$  - максимум модуля второй производной функции распределения;  
 $f_a(x, \alpha_1, \dots, \alpha_m)$  - аппроксимирующее выражение плотности распределения вероятностей;  
 $F_a(x, \alpha_1, \dots, \alpha_m)$  - аппроксимирующее выражение функции распределения вероятностей;  
 $F_x(x)$  - функция распределения вероятностей;  
 $f_x(x)$  - плотность распределения вероятностей;  
 $g[\ ]$  - оператор, представляющий собой преобразования, лежащие в основе определения вероятностной характеристики  $\Theta$ ;  
 $h(\tau, \alpha_1, \dots, \alpha_n)$  - импульсная характеристика фильтра с регулируемыми параметрами;  
 $K_a(\tau, \alpha_1, \dots, \alpha_n)$  - аппроксимирующее выражение корреляционной функции;  
 $k_a$  - коэффициент асимметрии;  
 $k_\epsilon$  - коэффициент эксцесса;  
 $k_\phi$  - коэффициент формы;  
 $K_x(\tau)$  - корреляционная функция стационарного случайного процесса;  
 $K_x(t, t')$  - корреляционная функция случайного процесса;  
 $K_{xy}(\tau)$  - взаимная корреляционная функция;

$k_{xy}$  - корреляционный момент;  
 $L_k(\tau, \alpha)$  - ортогональная функция Лагерра k-го порядка;  
 $M[\ ]$  - оператор математического ожидания;  
 $Me$  - медиана;  
 $Mo$  - мода;  
 $\hat{m}_x$  - оценка математического ожидания;  
 $\beta_j$  - частота попадания анализируемой случайной величины в j-ый дифференциальный коридор;  
 $P_x(\tau)$  - полярная (знаковая) корреляционная функция;  
 $R_x(\tau)$  - релейная корреляционная функция;  
 $S_d$  - оператор усреднения;  
 $Signu$  - знаковая функция.  
 $S_x(\omega)$  - спектральная плотность мощности процесса  $\overset{\circ}{x}(t)$ ;  
 $S_{xn} S_{xn}(\omega)$  - нормированная спектральная плотность мощности процесса  $\overset{\circ}{x}(t)$ ;  
 $S_{xy} S_{xy}(\omega)$  - взаимная спектральная плотность мощности;  
 $T_f T_\phi$  - постоянная времени ФНЧ;  
 $W W(j\omega)$  - частотная характеристика фильтра с регулируемыми параметрами;  
 $W_k W_k(j\omega)$  - частотная характеристика фильтра Лагерра k-го порядка;  
 $X_j x_j(t)$  - j-ая реализация случайного процесса;  
 $X_q \bar{x}(\Theta, t)$  - реализация случайного процесса;  
 $\alpha_1, \dots, \alpha_m$  - параметры аппроксимирующего выражения;  
 $\alpha_k$  - начальный момент k-го порядка;  
 $\beta_k$  - коэффициент разложения ортогонального ряда;  
 $\chi$  - коэффициент вариации;  
 $\Delta$  - погрешность аппроксимации;  
 $\Delta u$  - шаг квантования по уровню;  
 $\delta$  - погрешность аппроксимации;  
 $\delta_{ji}$  - индикатор состояния;  
 $\delta(t)$  -  $\delta$ -функция Дирака;  
 $\Delta t_{ji}$  - интервал дискретизации;  
 $\Delta \omega_\phi$  - полоса пропускания фильтра;  
 $\Delta \omega_c$  - эквивалентная ширина спектра мощности сигнала;  
 $\Delta x$  - ширина дифференциального коридора;  
 $\gamma_{см}$  - погрешность от смещенности оценки;  
 $\gamma_{см доп}, \gamma_{м доп}$  - допустимые значения погрешностей оценки;  
 $\gamma_m$  - методическая статистическая погрешность;  
 $\eta$  - пикфактор;  
 $\mu$  - показатель колебательности;



$\mu_k$  - центральный момент k-го порядка;

$\Theta_j[X(t)]$  - j-текущая оценка вероятностной характеристики;

$\Theta[X(t)]$ ,  $\Theta$  - измеряемая вероятностная характеристика;

$\bar{\Theta}$  - вектор информативных параметров случайного процесса;

$\Theta_{cp}[X(t)]$  - средняя оценка вероятностной характеристики;

$\Theta[X(t)]$  - оценка измеряемой вероятностной характеристики;

$\Theta_t[X(t)]$  - t-текущая оценка вероятностной характеристики;

$\rho_a(\tau, \alpha_1, \dots, \alpha_n)$  - аппроксимирующее выражение нормированной корреляционной функции;

$\rho_x(\tau)$  - нормированная корреляционная функция стационарного случайного процесса;

$\rho_x(t, t')$  - нормированная корреляционная функция случайного процесса;

$\rho_{xy}(\tau)$  - взаимная нормированная корреляционная функция;

$\rho_{xy}$  - коэффициент корреляции;

$\tau_k$  - интервал корреляции;

$\tau_{kmax}$  - максимальный интервал корреляции;

$\bar{\Omega}$  - вектор информативных параметров объекта исследований.

## ПРЕДИСЛОВИЕ

Предлагаемая Вашему вниманию монография написана для научных сотрудников, инженеров, аспирантов и студентов как руководство по основам аппроксимативного анализа случайных процессов.

Материалы, представленные в монографии, получены в результате выполнения научно-исследовательских работ на кафедрах «Информационно-измерительная техника», «Информационная техника» Самарского государственного технического университета (СГТУ), «Информационные системы и технологии» Самарского государственного аэрокосмического университета (СГАУ), Самарском филиале Российского НИИ информационных систем, математическом факультете Загребского университета и «Центре исследования моря» института «Руджер Бошкович» (г. Загреб, Хорватия), выполненных под руководством и при непосредственном участии автора.

Отдельные разделы монографии использовались при чтении лекций по ряду дисциплин при подготовке студентов по специальностям «Информационно-измерительная техника», «Автоматизированные системы обработки информации и управления» в СГТУ, СГАУ, а также для научных сотрудников и аспирантов в «Центре исследования моря» института «Руджер Бошкович», на математическом факультете Загребского университета, международном университете подготовки аспирантов (г. Дубровник, Хорватия), Пекинском техническом университете.

Методы и алгоритмы рассматриваемые в монографии, представляют собой, по сути, развитие идей метода наименьших квадратов при аппроксимации различных функциональных вероятностных характеристик, обладающих рядом специфических свойств и требующих отдельного рассмотрения.

**В первом разделе**, посвященном выравниванию статистических рядов, обсуждаются критерии приближения, методы идентификации по фазовым портретам и способы определения неизвестных параметров аналитических выражений законов распределения, приводится описание разработанной автоматизированной информационной системы.

**Во втором разделе** рассмотрены вопросы аппаратурной аппроксимации корреляционных функций параметрическими моделями, анализа методических и статистических погрешностей.

Аппаратурная аппроксимация корреляционных функций ортогональными функциями Лагерра, методы повышения точности аппроксимации, основанные на оптимизации параметра ортогональных функций по разным критериям, рассматриваются в **третьем разделе**.

Вопросы аппроксимации корреляционных функций параметрическими моделями с помощью ЭВМ, в том числе идентификации корреляционных функций по фазовым портретам, сходимости методов, рассматриваются в **четвертом разделе**.

**В пятом разделе** анализируются особенности аппроксимации корреляционных функций неэквидистантных временных рядов.

**В шестом разделе** рассматриваются аппроксимативные методы анализа обобщенных корреляционных характеристик, спектральных плотностей мощности и обобщенных спектральных характеристик.

**В седьмом разделе** описываются автоматизированные информационные системы для аппроксимативного анализа случайных процессов, в разработке которых активное участие приняли студенты-дипломники.

В разработке и моделировании аппаратно-программных средств на разных этапах работы принимали участие аспиранты и студенты указанных кафедр, выполнявшие под руководством автора кандидатские диссертации, курсовые и дипломные проекты: Батищев В.И., Зеленко Л.С., Дмитриева Е.В., Крупец Н.Г., Мирзоев Р.К., Снигерева-Давыденко Л.Б., Кудрина М.А., Новиченкова И.Ю., Козлов А.Н. и Учеватов С.В. и Шевченко Д., Иващенко А.В. и др.

Автор выражает благодарность всем сотрудникам, аспирантам и студентам кафедры информационных систем и технологий СГАУ за обсуждение материалов монографии, критические замечания, которых по мере возможности были учтены, особенно профессору Коварцеву А.Н. и доценту Иоффе В.Г.

Большую благодарность автор выражает своим соавторам: доцентам Мотову В.В., Батищеву В.И., Сухнину В.П., Белолипецкому В.Н., Иванову С.Г., а также профессорам Куликовскому К.Л., Кловскому Д.Д., Курочкину Е.П., Мартяшину А.И., Иванову В.Н., Рожкову В.А., Трапезникову Ю.А., Тельскнису Л.А. (Литва), М. Бранице (Хорватия), И. Ружичу (Хорватия), Л. Ефтичу (Хорватия), Д. Р. Кушину (Австралия), Ф Вайде (Венгрия)- за обсуждение работы на различных её этапах.

Особую благодарность автор выражает своим учителям: доценту Волкову И.И., профессорам Виттиху В.А., Карпову Е.М., Куликовскому Л.Ф., Самарину Ю.П., Фремке А.В., Цветкову Э.И., Бранко Соучеку (Хорватия), - за постоянное внимание, консультации и особенно за то, что привели в чудесный мир теории вероятностей и случайных процессов.

Автор считает своим долгом выразить глубокую признательность ректору СГАУ член-корреспонденту РАН, профессору Сойферу В.А. за постоянную поддержку, рецензию и неоценимую помощь при издании монографии, а также д.ф.м.н., профессору Жданову А.И. за ценные замечания.

Неоценимую помощь в редактировании монографии оказала инженер кафедры «Информационные системы и технологии» СГАУ Седьмова Е.В., которой автор выражает свою благодарность.

Издание монографии поддержано грантом по программе «Поддержка важнейших разработок научных коллективов и отдельных учёных на основе системы грантов» по разделу «Важнейшие научные и технические разработки, соответствующие концепции социально-экономического развития Самары».

Замечания и пожелания по книге просьба направлять по адресу:

Россия, 443086, г. Самара, Московское шоссе, 34, Самарский государственный аэрокосмический университет имени академика С.П. Королёва, кафедра «Информационных систем и технологий».

E-mail: INTERNET: sp @ smr. ru

## ВВЕДЕНИЕ

На пути создания образцов новой техники, технологических процессов научные исследования являются первым шагом, в процессе которого исследователь открывает новые законы, закономерности, совершает научные открытия.

Научные исследования представляют собой сложный, итерационный процесс, представляющий сочетание теоретических, включая методы моделирования, и экспериментальных методов.

Не умаляя достоинств теоретических методов исследования, значение экспериментальных методов трудно переоценить. Только с помощью эксперимента возможно получение достоверной информации об исследуемом объекте в реальном масштабе времени, после обработки которой возможно построение её модели. Открыв новый эффект, новое явление экспериментальным путем, которые невозможно объяснить на базе существующих теорий, экспериментатор стимулирует развитие фундаментальной науки. В то же время, получив новый теоретический научный результат, исследователь, с целью подтверждения основных положений новой теории, нуждается в его экспериментальной проверке.

При проведении экспериментальных научных исследований исследователь:

1. ставит задачу исследований в терминах предметной области;
2. строит модель исследуемого объекта и определяет вектор информативных параметров  $\vec{\Omega}$ , адекватно описывающий ее в рамках поставленной задачи;
3. с помощью технических средств осуществляет измерение, регистрацию и обработку мгновенных значений наблюдаемых процессов  $\vec{X}(\vec{\Theta}, t)$ , с целью определения вектора информативных параметров  $\vec{\Theta}$ , описывающих модель процесса;
4. по результатам обработки информации устанавливает взаимно однозначное соответствие между векторами  $\vec{\Omega}$  и  $\vec{\Theta}$ :

$$\vec{\Omega} = \Phi(\vec{\Theta}),$$

используемое для построения искомой модели объекта;

5. анализирует полученные результаты;
6. если результаты его устраивают - эксперимент окончен, в противном случае необходимо повторить пункты 3, 4 (точность полученных результатов неудовлетворительна), или пункты 2-4 (вектор параметров  $\vec{\Omega}$  не полно описывает поведение объекта), а иногда и пункты 1-4 (ставится другая задача).

Следует отметить, что задачи 1, 2 и 4, 5, как правило, решает специалист данной предметной области, формулируя и интерпретируя ее в терминах предметной области, а 3 задачу - специалисты в области измерения и обработки измерительной информации.

Такое разделение функций между исследователем и специалистом в области измерения и обработки измерительной информации позволяет последнему абстрагироваться от конкретных физических объектов и вектора физических параметров  $\vec{\Omega}$  и непосредственно перейти:

1. к математическому описанию исследуемых процессов и определению вектора параметров  $\vec{\Theta}$ , достаточных для решения поставленной задачи;
2. сбору информации с помощью первичных преобразователей;
3. оценке вектора параметров  $\vec{\Theta}$  с помощью технических средств;
4. анализу точности полученных результатов;

5. аппроксимации полученных функциональных зависимостей с помощью параметрических моделей.

Каждая из перечисленных задач имеет свои специфические особенности, а эффективность решения четвертой и пятой - зависит от применяемых технических средств, построенных, как правило, на базе современных средств информационно-измерительной и вычислительной техники.

Решая разнообразные задачи научных исследований, исследователь на основании физических представлений и задачи исследований определяет составляющие вектора параметров случайного процесса  $\bar{\Theta}$ , дающие его исчерпывающее описание.

Все вероятностные характеристики, определяемые во временной области, можно условно разделить на характеристики положения и формы кривой распределения вероятностей случайного процесса и характеристики взаимосвязи (см. рис. В.1).

При этом наиболее часто определяются (в порядке возрастания материальных и вычислительных затрат):

- числовые характеристики случайного процесса;
- авто и взаимные корреляционные функции;
- спектральные плотности мощности;
- законы распределения.



Рисунок В.1. Классификация вероятностных характеристик случайных процессов

На основании общей теории статистических измерений [106] измеряемая вероятностная характеристика определяется как предел выборочного среднего функционально преобразованного случайного процесса:

$$\Theta[X(t)] = \lim_{d \rightarrow \infty} S_d g[x_j(t)], \quad (B.1)$$

где  $\Theta$  - измеряемая вероятностная характеристика;

$S_d$  - оператор идеального усреднения;

$d$  - параметр усреднения (время  $T$ , совокупность реализаций  $N$  или время и совокупность реализаций  $TN$ );

$g$  - оператор, представляющий собой преобразования, лежащие в основе определения вероятностной характеристики  $\Theta$ ;

$x_j(t)$  -  $j$ -ая реализация случайного процесса.

В зависимости от вида усреднения получаем следующие вероятностные характеристики:

1. При усреднении по совокупности:

$$\Theta[X(t)] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N g[x_j(t)]. \quad (B.2)$$

2. При усреднении по времени:

$$\Theta[X(t)] = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T g[x_j(t)] dt. \quad (B.3)$$

3. При усреднении по времени и совокупности

$$\Theta[X(t)] = \lim_{\substack{N \rightarrow \infty \\ T \rightarrow \infty}} \frac{1}{NT} \sum_{j=1}^N \int_0^T g[x_j(t)] dt. \quad (B.4)$$

На практике исследователь имеет дело с ограниченной совокупностью выборочных данных (результатов измерения). Результат определения значения вероятностной характеристики по ограниченной совокупности выборочных данных носит название **оценки**:

$$\mathfrak{G}[X(t)] = S_d g[x_j(t)] (j = 1, 2 \dots N). \quad (B.5)$$

К основным свойствам оценок относятся **несмещенность, состоятельность и эффективность**.

Оценка называется **несмещенной**, если ее математическое ожидание равно истинному значению оцениваемой характеристики:

$$M[\mathfrak{G}[X(t)]] = \Theta[X(t)]. \quad (B.6)$$

При невыполнении равенства оценка будет смещенной.

Оценка называется **состоятельной**, если при бесконечном увеличении объема выборочных данных она сходится по вероятности к истинному значению оцениваемой характеристики:

$$\lim_{d \rightarrow \infty} P(|\mathfrak{G}[X(t)] - \Theta[X(t)]| < \varepsilon) = 1 \quad (B.7)$$

при любом  $\varepsilon$ .

При невыполнении этого условия оценка будет несостоятельной.

**Эффективными** называются оценки, дисперсия которых минимальна.

Следует подчеркнуть, что свойства оценок, которые описываются несмещенностью, состоятельностью и эффективностью, тесно связаны с характером ошибок, которые определяются методами математической статистики [45,46].

Таким образом, при ограниченном наборе выборочных данных выражения (B.2)-(B.4) при анализе случайных процессов примут вид:

- при усреднении по совокупности

$$\Theta_t[X(t)] = \frac{1}{N} \sum_{j=1}^N g[x_j(t)]; \quad (B.8)$$

- при усреднении по времени

$$\Theta_j[X(t)] = \frac{1}{T} \int_0^T g[x_j(t)]dt; \quad (B.9)$$

- при усреднении по времени и совокупности

$$\Theta_{cp}[X(t)] = \frac{1}{NT} \sum_{j=1}^N \int_0^T g[x_j(t)]dt. \quad (B.10)$$

Следует отметить, что этими соотношениями определяются разные вероятностные характеристики. При усреднении только по совокупности реализаций (при фиксированном моменте времени) вероятностная характеристика  $\Theta[X(t)]$  будет зависеть от текущего времени и называется  $t$ -текущей характеристикой  $\Theta_t$  [106]. При усреднении только по времени, когда выборочные значения относятся к одной реализации  $j$ , вероятностная характеристика  $\Theta[X(t)]$  будет зависеть от номера реализации и называется  $j$ -текущей характеристикой  $\Theta_j$ . При усреднении и по времени и совокупности значение  $\Theta[X(t)]$  не зависит ни от текущего времени, ни от номера реализации и называется средней характеристикой -  $\Theta_{cp}$ .

Наличие или отсутствие зависимости значений вероятностных характеристик от времени или номера реализации определяет такие фундаментальные свойства процесса, как **стационарность** и **эргодичность**.

**Стационарным** называются процесс, вероятностные характеристики которого не зависят от времени. **Эргодическим** называется процесс, вероятностные характеристики которого не зависят от номера реализации.

В теории случайных процессов различают стационарность в узком и широком смысле. Данное выше определение относится к случайным процессам, стационарным в узком смысле. Для этих процессов равенство  $\Theta_t = \text{const}$  выполняется для любой вероятностной характеристики. Когда от времени не зависят только одно-и двумерные вероятностные характеристики, случайный процесс считается стационарным в широком смысле. Если условие стационарности не выполняется хотя бы для одной вероятностной характеристики, процесс называется **нестационарным** по этой характеристике.

По аналогии, процесс считается эргодическим в узком смысле, если  $\Theta_j = \text{const}$ , где  $j$  - номер реализации, для любой вероятностной характеристики, и в широком смысле, если независимость значений вероятностных характеристик от номера реализации имеет место лишь для характеристик первых двух порядков. Если условие эргодичности не выполняется, процесс называется **неэргодическим**.

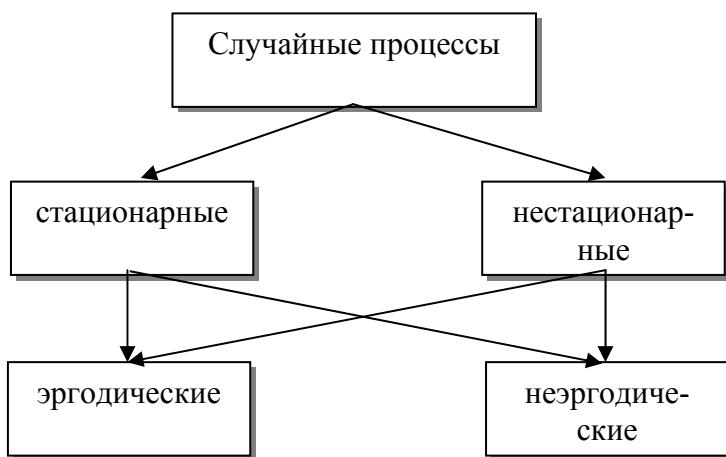


Рисунок В.2. Классификация случайных процессов

Таким образом, случайные процессы на основе свойств стационарности и эргодичности можно представить в виде четырех классов (см. рис.В.2):

- стационарные эргодические;
- стационарные неэргодические;
- нестационарные эргодические;
- нестационарные неэргодические.

Каждый из перечисленных классов имеет своё характерное описание - **математическую модель**, параметры которой подлежат определению как с помощью теоретических, так и экспериментальных методов исследования.

Различные комбинации этих процессов совместно с детерминированными дают возможность построить более сложные модели, используемые как при исследованиях с целью определения их характеристик, так и при генерировании процессов с заданными свойствами, используемых при имитационном моделировании средств измерения и обработки с целью определения их метрологических характеристик.

Ответ на вопрос, какие характеристики определять:  $\Theta_t$ ,  $\Theta_j$ ,  $\Theta_{cp}$ , - во многом определяется свойствами исследуемого процесса и способом формирования выборочных данных.

Таким образом, прежде чем выбрать тип вероятностной характеристики (вид оператора усреднения), необходимо решить вопрос о стационарности и эргодичности случайного процесса. Этот вопрос самостоятельный и выходит за рамки монографии. Отметим, что существенные результаты по классификации случайных процессов получены в работах [96, 99, 101]. В [99] разработано программное обеспечение, позволяющее решить вопрос о стационарности и эргодичности практически.

В [106] показано, что для стационарного эргодического случайного процесса  $\Theta_t = \Theta_j = \Theta_{cp}$ , для стационарного неэргодического процесса  $\Theta_t = \Theta_{cp}$ , для нестационарного эргодического -  $\Theta_j = \Theta_{cp}$ , а для нестационарного неэргодического процесса все виды вероятностных характеристик различны.

При фиксированном типе  $S_d$  вопросы организации эксперимента и принципы организации массивов выборочных данных о мгновенных значениях исследуемого случайного процесса подробно рассмотрены в [106]. Результаты измерений могут формироваться с использованием:

- различных временных интервалов одной и той же совокупности реализаций;
- одних и тех же временных интервалов различных совокупностей реализаций;
- различных временных интервалов различных совокупностей реализаций.

Причем, от эксперимента к эксперименту возможно изменение  $N$ ,  $M$  и  $T$ , т.е. объема выборочных данных.



Выделим три метода статистических измерений: прямые, косвенные и совокупные.

**Прямым методом** статистических измерений будем называть метод получения оценки вероятностной характеристики в соответствии с выражением (В. 1).

**Косвенным методом** статистических измерений будем называть метод получения оценки вероятностной характеристики с использованием функционального преобразования оценок других вероятностных характеристик, полученных с помощью прямых методов статистических измерений:

$$\mathcal{G}[z(t)] = F\left\{S_{d1}g_x[x_j(t_i^{(x)})], S_{d2}g_y[y_1(t_i^{(y)})], \dots\right\} \quad (\text{В. 11})$$

где  $F\{\}$  представляет собой функциональное преобразование полученных оценок  $\mathcal{G}[x(t)]$  и  $\mathcal{G}[y(t)]$  и т.д. с целью получения оценки  $\mathcal{G}[z(t)]$ .

Под **совокупными статистическими измерениями** будем понимать метод получения оценок в результате решения системы уравнений, содержащей оценки других вероятностных характеристик, полученных с помощью прямых, косвенных методов статистических измерений или их комбинацией:

$$\Xi_i \left\{ S_{d1}g^{(1)}[x_j^{(1)}(t_i^{(1)})], \dots, S_{dm}g^{(m)}[x_j^{(m)}(t_i^{(m)})] \right\} = 0; \quad (\text{В. 12})$$

$$\Xi_i \left\{ F^{(1)} \left\{ S_{d1}g^{(1)}[x_j^{(1)}(t_i^{(1)})] \right\}, \dots, F^{(m)} \left\{ S_{dm}g[x_j^{(m)}(t_i^{(m)})] \right\} \right\} = 0. \quad (\text{В. 13})$$

Основной подсистемой любого технического средства, предназначенного для получения и обработки измерительной информации является измерительно-вычислительный канал.

Под **измерительно-вычислительным каналом** понимается совокупность аппаратно-программных средств, предназначенных для измерения мгновенных значений соответствующей физической величины, обработки результатов измерения и представления конечных результатов в форме, удобной для дальнейшего использования.

Рассмотрим структуру отдельного измерительно-вычислительного канала.

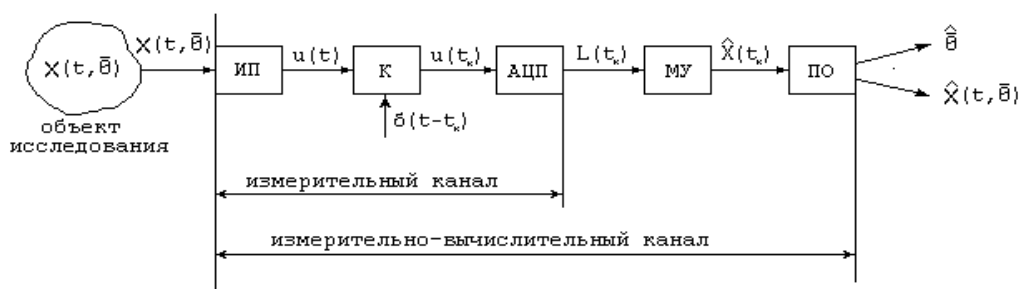


Рисунок В.3. Измерительно-вычислительный канал

На рис. В.3 приняты следующие обозначения.

- **ИП** - измерительный преобразователь (датчик);
- **К** - коммутатор;
- **АЦП** – аналого-цифровой преобразователь;
- **МУ** - масштабирующее устройство;
- **ПО** - процессор обработки.

Рассмотрим преобразования, происходящие с сигналами в измерительно-вычислительном канале.

Независимо от природы измеряемой физической величины на выходе **ИП** получаем электрический сигнал. При этом, каждому значению физической величины ста-

вится в соответствие вполне определенное значение электрической величины:  $x(t, \vec{\Theta}) \rightarrow u(t)$ .

Основным требованием, предъявляемым к **ИП**, является линейность:

$$u(t) = k x(t, \vec{\Theta}), \quad (\text{B.14})$$

где  $k = \text{const}$  - коэффициент преобразования.

Следует отметить, что если связь между  $u(t)$  и  $x(t, \vec{\Theta})$  нелинейная, то производят линеаризацию функции преобразования, воспользовавшись, например, методом наименьших квадратов [44].

После **ИП** исследуемый сигнал поступает на вход коммутатора.

В коммутаторе непрерывный сигнал  $u(t)$  преобразуется в последовательность отсчетов, отстоящих друг от друга на интервале  $\Delta t_i = t_{i+1} - t_i$ . Т.е. выполняется операция **дискретизации**:

$$u(t_i) = \int_{-\infty}^{\infty} u(t) \delta(t-t_i) dt, \quad (\text{B.15})$$

$$\text{где } \delta(t) = \begin{cases} \infty, & \text{если } t = t_i \\ 0, & \text{если } t \neq t_i \end{cases} \text{ - } \delta\text{-функция Дирака.} \quad (\text{B.16})$$

Отметим, что при коммутации возможны два варианта:

- $\Delta t_i = \text{const}$  - регулярная дискретизация;
- $\Delta t_i = \text{random}$  - нерегулярная дискретизация.

После коммутации сигнал поступает на аналого-цифровое преобразование, где последовательно подвергается процедурам **квантования** и **кодирования**.

**Квантование** - процедура отнесения непрерывного значения процесса  $u(t_i)$  к ближайшему разрешенному целому уровню.

$$L(t_i) = \text{ent} \left[ \frac{|u(t_i)|}{\Delta u} + \lambda \right] \text{Sign } u(t_i), \quad (\text{B.17})$$

где  $\text{ent}[\ ]$  - операция взятия целой части числа;

$\Delta u$  - шаг квантования по уровню;

$$\lambda = \begin{cases} 0, & \text{при усечении} \\ 0,5, & \text{при округлении} \end{cases}$$

$$\text{Sign } u = \begin{cases} 1, & \text{если } u > 0; \\ 0, & \text{если } u = 0; \\ -1, & \text{если } u < 0; \end{cases} \text{ - знаковая функция.} \quad (\text{B.18})$$

В результате квантования сигнала получим целое число квантов, которое может кодироваться различными способами. При использовании двоичной системы счисления с весами 8-4-2-1 число двоичных разрядов, необходимых для представления  $L(t_i)$  определяется выражением:

$$n = \log_2 \max |L(t_i)|. \quad (\text{B.19})$$

Следует подчеркнуть, что в случае одноканальной системы операции коммутации и аналого-цифрового преобразования совпадают. В многоканальных же системах, как правило, применяется один коммутатор на несколько каналов.

После аналого-цифрового преобразователя сигнал поступает на вход масштабирующего устройства, выходной сигнал которого равен:

$$\mathfrak{K}(t_k) = \frac{L(t_k)\Delta u}{k}. \quad (\text{B.20})$$

Т.е. при этом происходит обратное преобразование: приведение электрического сигнала к измеряемой физической величине.

Далее сигнал поступает в процессор обработки - устройство, реализующее тот или иной алгоритм получения оценки параметров физического процесса  $x(t, \bar{\Theta})$ . При этом возможны два подхода к решению задачи оценки  $\bar{\Theta}$ :

1. в режиме экспресс-анализа с помощью алгоритма  $A_{\Theta}$  оценивают вектор неизвестных параметров  $\bar{\Theta} = A_{\Theta} \{\mathfrak{K}(t_k)\}$ ;

2. в режиме контроля и регистрации с помощью алгоритма  $A_s$  получают оценку сигнала  $\mathfrak{K}(t, \bar{\Theta}) = A_s \{\mathfrak{K}(t_k)\}$ , записывают его на какой-либо промежуточный носитель, а затем обрабатывают. При этом происходит задержка в обработке информации и, следовательно, в получении результата.

Первый подход будем называть **статистическими измерениями**, второй - **статистической обработкой**.

Под **статистическими измерениями** будем понимать измерение вероятностных характеристик случайных процессов с помощью специальных технических средств, работающих в реальном масштабе времени.

Под **статистической обработкой** будем понимать оценку вероятностных характеристик случайных процессов на ЭВМ, записанных на промежуточный носитель, с задержкой в обработке информации.

Таким образом, измерительно-вычислительный канал даёт возможность обрабатывать **временные последовательности случайных процессов - временные ряды**.

В этом случае выражения (B.2)-(B.4) при представлении случайного процесса  $X(t)$  ансамблем последовательностей примут вид:

$$\Theta[X(t)] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N g[x_j(t_i)]; \quad (\text{B.21})$$

$$\Theta[X(t)] = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{i=1}^M g[x_j(t_i)]; \quad (\text{B.22})$$

$$\Theta[X(t)] = \lim_{\substack{N \rightarrow \infty \\ M \rightarrow \infty}} \frac{1}{NM} \sum_{j=1}^N \sum_{i=1}^M g[x_j(t_i)], \quad (\text{B.23})$$

где  $t_i$  -  $i$ -ый отсчёт  $j$ -ой реализации случайного процесса.

При ограниченном наборе данных при анализе последовательностей выражение (B.5) примет вид:

$$\mathfrak{G}[X(t)] = S_d g[x_j(t_i)] (j = 1, 2, \dots, N; i = 1, 2, \dots, M). \quad (\text{B.24})$$

Выражения (B.8)-(B.10) для оценки вероятностных характеристик при анализе последовательностей (временных рядов) запишем в виде:

- при усреднении по совокупности

$$\mathfrak{G}_i[X(t)] = \frac{1}{N} \sum_{j=1}^N g[x_j(t_i)]; \quad (\text{B.25})$$

- при усреднении по времени

$$\mathfrak{G}_j[X(t)] = \frac{1}{M} \sum_{i=1}^M g[x_j(t_i)]; \quad (\text{B.26})$$

- при усреднении по времени и совокупности

$$\mathfrak{G}_{\text{cp}}[X(t)] = \frac{1}{NM} \sum_{j=1}^N \sum_{i=1}^M g[x_j(t_i)]. \quad (\text{B.27})$$

Выделяют первичную и вторичную статистическую обработку. Под **первичной** статистической обработкой будем понимать оценку вероятностных характеристик по ограниченному набору данных, под **вторичной** - построение аналитических моделей исследуемых процессов и их характеристик.

Если не принимать во внимание фактор реального времени или задержку в обработке информации, то с точки зрения получения оценки по одному и тому же объему данных два подхода с методической точки зрения не отличаются друг от друга. Существенное отличие заключается в том, что работа в реальном масштабе времени накладывает жесткие ограничения на быстродействие технических средств. Это заставляет, в свою очередь, при статистических измерениях применять упрощенные алгоритмы оценивания интересующих параметров, обладающих значительным быстродействием. Кроме того, работа в реальном масштабе времени не дает возможность получить требуемые оценки, воспользовавшись другими алгоритмами, так как данную реализацию повторить нельзя - она **случайна**. Требуемую оценку необходимо получить за время, отводимое на эксперимент. Реализация же, записанная на промежуточный носитель или в память, становится **детерминированной**, и с ней можно экспериментировать сколь угодно долго. Это важное преимущество статистической обработки позволяет:

- с целью повышения точности оценивания осуществить оценку одного параметра с помощью различных алгоритмов обработки информации;
- выбрать оптимальный алгоритм оценивания, соответствующий выбранному критерию;
- с целью построения новой или уточнённой модели осуществить оценку других параметров, описывающих эту модель.

При исследовании сложных объектов проводят, как правило, большое число испытаний. При этом происходит и накопление большого числа массивов числовых и функциональных характеристик, что в значительной степени затрудняет хранение, анализ и интерпретацию полученных результатов.

Один из возможных способов решения этой проблемы заключается в применении аппроксимативных методов (от латинского слова **approximo** - приближаюсь), суть которых заключается в нахождении подходящего аналитического выражения  $\varphi(x(t), \alpha_0, \alpha_1, \dots, \alpha_n)$  с неизвестными параметрами  $\alpha_0, \alpha_1, \dots, \alpha_n$ , удовлетворяющими заданному критерию оптимальности, которое бы описывало найденные экспериментальные результаты. Аппроксимативный подход оказывается эффективным и при обработке результатов имитационного моделирования (вычислительного эксперимента).

Предположим, что в результате опыта мы получили ряд экспериментальных точек и построили график зависимости  $\xi_i = f(\xi_i)$ . Желательно обработать экспериментальные данные таким образом, чтобы по возможности точно отразить общую тенденцию зависимости  $y$  от  $x$  и вместе с тем сгладить случайные отклонения, вызванные погрешностями самого эксперимента.

Наиболее эффективным методом решения этой задачи является метод наименьших квадратов [45].

Пусть задан некоторый класс функций  $\varphi_j(x(t), \alpha_0, \alpha_1, \dots, \alpha_n)$  с одинаковым числом неизвестных параметров. Тогда наилучшей будет та функция, для которой сумма квадратов

$$\Delta_j = \sum_{i=1}^M [\xi_i - \varphi_j(\xi_i, \alpha_0, \alpha_1, \dots, \alpha_n)]^2 = \min. \quad (\text{B.28})$$

Если в сравнении участвуют функции с различным числом неизвестных параметров, это приводит к увеличению числа связей и к уменьшению числа степеней свободы  $m-1$ , что, в свою очередь, может ухудшить дисперсию [45].

$$D_j = \frac{\Delta_j}{m - r}. \quad (\text{B.29})$$

Под числом связей  $r$  понимают число неопределенных выражений, входящих в аналитическое выражение.

Так как  $\Delta_j$  является функцией  $\alpha_0, \alpha_1, \dots, \alpha_n$  то, как известно [62], необходимыми условиями минимума дифференцируемой функции многих переменных является система нормальных уравнений:

$$\frac{\partial \Delta_j(\alpha_0, \alpha_1, \dots, \alpha_n)}{\partial \alpha_k} = 0, \quad k = 0, 1, \dots, n. \quad (\text{B.30})$$

решив которую с помощью того или иного численного метода найдем искомые параметры  $\alpha_0, \alpha_1, \dots, \alpha_n$ .

Классификация методов аппроксимации функциональных характеристик случайных процессов и последовательностей представлена на рис. В.4.

Основными преимуществами аппроксимативного подхода являются:

- наглядность и компактность полученного аналитического выражения, легкость визуализации;
- возможность использования аналитического выражения для дальнейших аналитических исследований и преобразований, с целью получения обобщенных вероятностных характеристик;
- сокращения объёма хранимых данных.

К недостаткам метода следует отнести наличие методической погрешности, возникающей при замене полученных экспериментальных данных или другой функции более простым аналитическим выражением.

В общем случае, для реализации аппроксимативного подхода необходимо выполнить следующие этапы:

1. на основании анализа решаемой задачи определить требования к реализации входного процесса (последовательности): длине реализации (объёму выборки), интервалу дискретизации, числу уровней квантования и т.д. с целью получения оце-

нок вероятностных характеристик для последующей аппроксимации с допустимыми погрешностями;

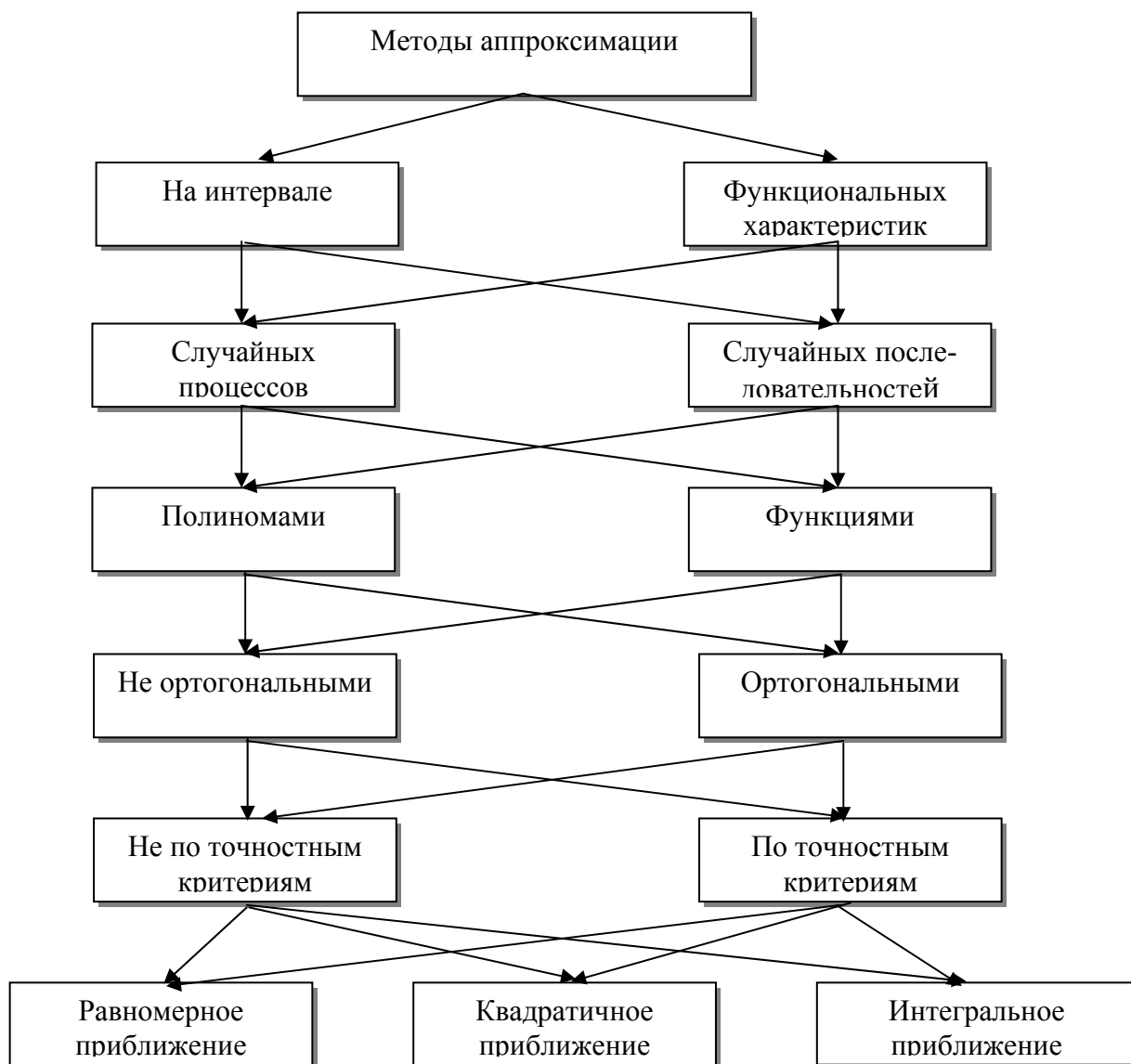


Рисунок В.4. Классификация методов аппроксимации случайных процессов и последовательностей

2. по измеренным значениям входного процесса (последовательности) оценить значения функциональной характеристики в заданных точках;

3. проанализировать полученный результат и выбрать подходящее аналитическое выражение, по возможности, с минимальным количеством неизвестных параметров, подлежащих определению, так как количество параметров в значительной мере определяет сложность аппаратуры или вычислений, его сходимость и устойчивость;

4. выбрать и обосновать критерий приближения;

5. составить и решить систему уравнений относительно неизвестных параметров аналитического выражения и определить погрешность приближения;

6. разработать структурную схему аппаратуры (программы) и рассчитать её параметры;

7. изготовить и отладить аппаратуру (написать и отладить программное обеспечение);
8. провести экспериментальные исследования (обработать полученные экспериментальные данные).
9. аналитически определить все интересующие обобщённые вероятностные характеристики.

Следует отметить, что определять аналитические выражения возможно как при анализе стационарных, так и нестационарных процессов. В первом случае анализируются функциональные характеристики, во втором - и моментные, являющиеся функциями времени.

Определять параметры аналитических выражений возможно как с помощью **статистических измерений**, так и в результате **статистической обработки**.

Под **статистическими измерениями с аппроксимацией** будем понимать измерение (оценку) параметров аппроксимирующего выражения вероятностной функциональной характеристики случайных процессов с помощью специальных технических средств, работающих в реальном масштабе времени. Самыми популярными среди таких технических средств являются **коррелометры** и **спектроанализаторы** с аппроксимацией параметрическими моделями. В литературе их часто называют **статистическими анализаторами**. Статистические анализаторы, как правило, специализированные аппаратно-программные средства, определяющие **параметры реального процесса** (см. рис. В.5).

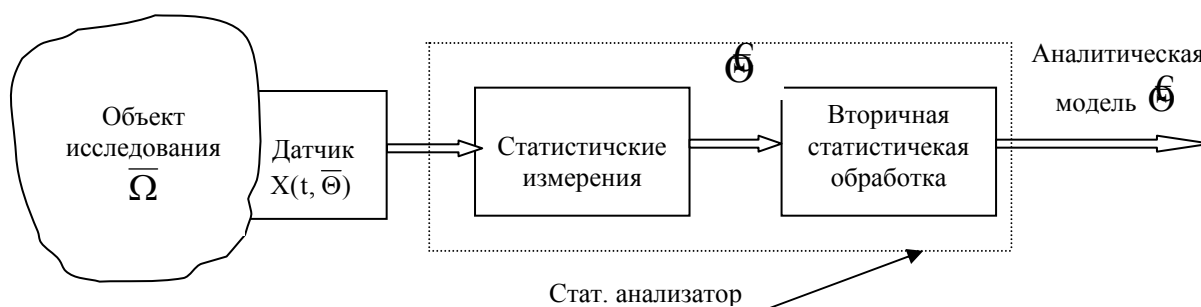


Рисунок В.5. Аппаратно-программные средства статистических измерений и обработки информации

Под **статистической обработкой с аппроксимацией** будем понимать оценку параметров аппроксимирующего выражения вероятностной функциональной характеристики случайных процессов с помощью ЭВМ, записанных на промежуточный носитель или память. При этом происходит временная задержка в обработке информации.

Аппроксимативные методы, основанные на применении ЭВМ, можно отнести к методам **вторичной обработки** информации. Однако, воспользовавшись принципом аппаратно-программного дуализма, возможно алгоритмы, положенные в основу работы статистических анализаторов, реализовать и чисто программным способом. Поэтому в работе рассматриваются решения аппроксимативных статистических задач, как с помощью статистических анализаторов, так и ЭВМ.

Классификация средств вычислительной техники, предназначенных для аппроксимации различных функциональных характеристик, представлена на рис. В.6.

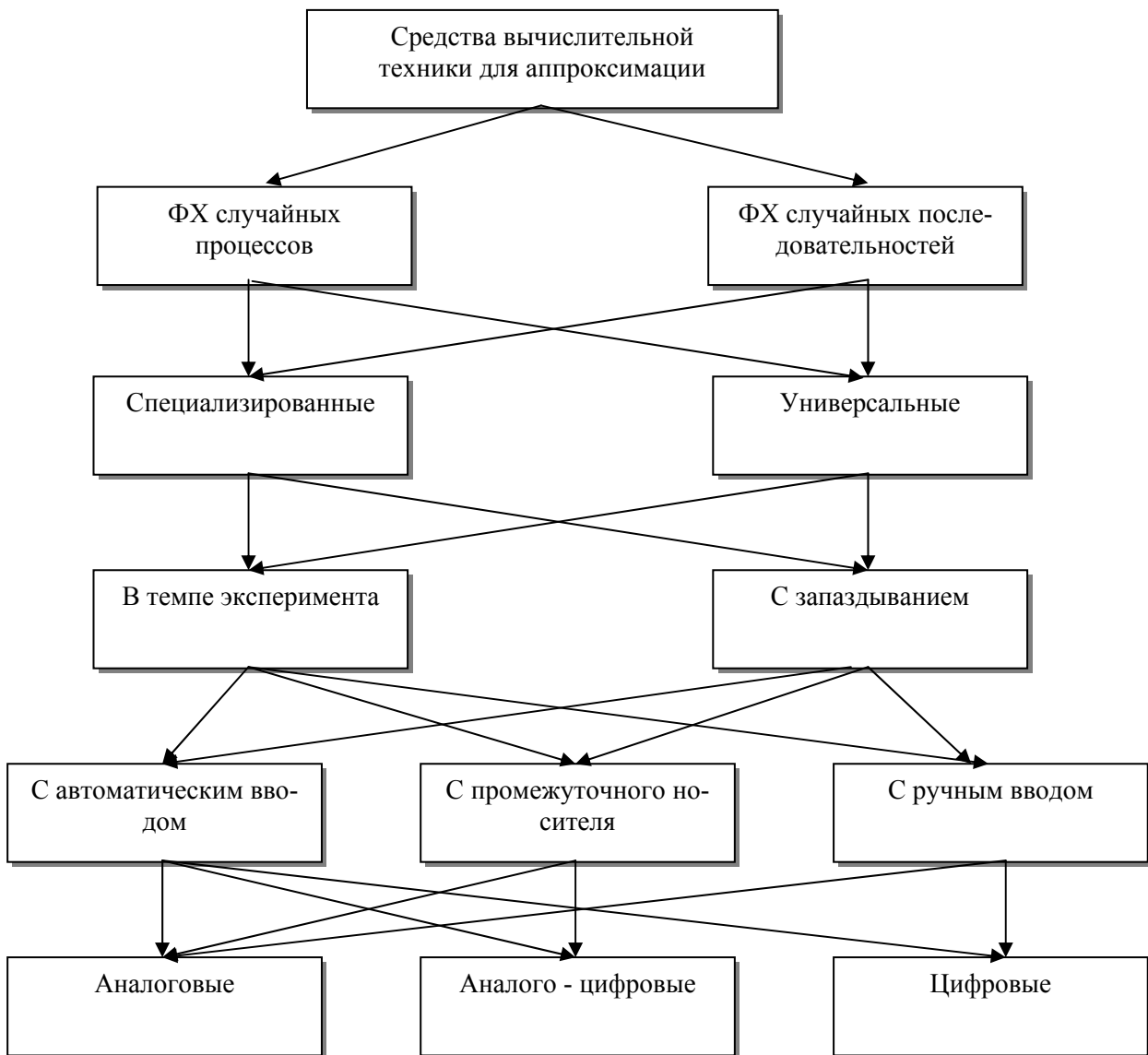


Рисунок В.6. Классификация средств вычислительной техники для аппроксимации функциональных характеристик случайных процессов и последовательностей

При проведении исследований аппроксимативных алгоритмов кроме аналитических методов применялся метод имитационного моделирования, блок-схема которого представлена на рис. В.7.

При этом следует подчеркнуть, что затраты на моделирование, достоверность полученных результатов во многом зависят от принятых решений на этапе планирования эксперимента, особенно при определении необходимого числа испытаний, выборе входных воздействий и т.д.

Согласно методике, изложенной в РТМ 25139-74 [78], в качестве метрологической характеристики может выбираться максимальное значение модуля погрешностей оценки  $\Delta$ :

$$\Delta = \max \left\{ \left| \Delta_j \right| \right\} j = 1, \dots, N, \quad (\text{В.31})$$



где  $N$  - число испытаний, зависящее от доверительной информации  $P_d$ . Так, если  $P_d = 0,95$ , то число испытаний равно 29 независимо от закона распределения погрешностей.



Рисунок В.7.

Одним из важных этапов имитационного моделирования является выбор, обоснование и моделирование сигналов, используемых в модельном эксперименте. Решение этой задачи определяется целевой функцией моделирования, назначением исследуемой системы и т.д. Так как при моделировании алгоритмов основной задачей является определение его метрологических характеристик, то существенным требованием, предъявляемым к образцовому (испытательному или тестовому) сигналу, является возможность оценки с его помощью погрешности результата измерения данным алгоритмом на заданном классе входных воздействий.

Учитывая большое разнообразие решаемых задач и соответствующих им алгоритмов, однозначного ответа о виде образцового сигнала быть не может. Окончательное решение о выборе вида образцового сигнала для конкретных типов средств измерения должно приниматься по результатам экспериментальных исследований.

В самом общем виде выбор образцового сигнала осуществляется:

- выбором наилучшего сигнала из множества возможных входных сигналов, для обеспечения гарантированной погрешности результата измерения;
- формированием набора типовых сигналов, то есть наиболее часто встречающихся входных сигналов или сигналов, наиболее интересных исследователя;
- формированием набора типовых сигналов, включающих в себя наилучший сигнал.

Основными требованиями, предъявляемыми к

образцовым сигналам, являются следующие:

- заданный вид вероятностных характеристик;
- принадлежность к классу входных сигналов, для которых предназначено данное средство;
- стабильность во времени;
- отклонение текущих характеристик от расчетных не должно быть более допустимого.

Результаты моделирования сравнивались, когда это было возможно, с результатами, полученными с помощью классических алгоритмов.

# 1. АППРОКСИМАЦИЯ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ

## 1.1. Основные понятия и определения

При обработке статистического материала часто приходится решать вопрос о том, как подобрать для данного статистического ряда теоретическую кривую распределения, выражающую лишь существенные черты статистического материала. Такая задача, относящаяся к классу аппроксимативных задач, в данном случае называется задачей выравнивания (сглаживания) статистических рядов.

Последовательность этапов решения задачи сглаживания представлена на рис.1.1.

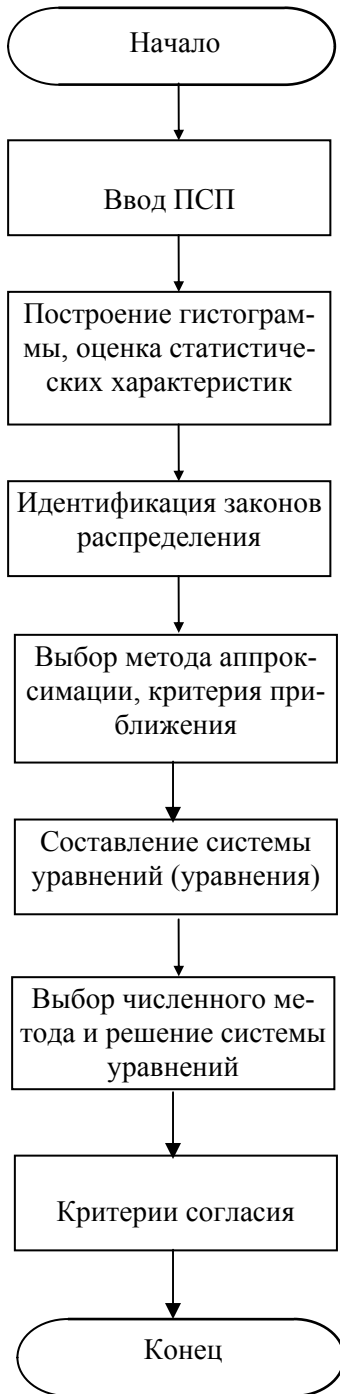


Рисунок 1.1. Блок-схема сглаживания

После ввода исходных данных (ПСП) первым шагом решения этой задачи является построение гистограммы наблюдаемого статистического ряда  $\{x_i\}_{i=1}^N$ . Для этого необходимо выполнить следующие этапы:

1. Определить диапазон изменения статистического ряда  $x_{\min}-x_{\max}$ .

2. При неизвестном законе распределения определить ширину дифференциального коридора:

$$\Delta x = \frac{x_{\max} - x_{\min}}{M}, \quad (1.1)$$

где  $M$  - количество дифференциальных коридоров [45].

Из выражения (1.1) следует, что значение величины  $M$  не определено и, следовательно, исследователь должен задать её самостоятельно, ориентируясь на рекомендации специалистов. Как следует из литературных источников, значение  $M$  рекомендуется выбирать в диапазоне от 10 до 20 [45].

При известном или предполагаемом законе распределения при линейной интерполяции функции распределения величина  $\Delta x$  определяется выражением:

$$\Delta x = \sqrt{\frac{8\delta}{|F''(x)|_{\max}}}, \quad (1.2)$$

где  $\delta$  - погрешность аппроксимации функции распределения;

$|F''(x)|_{\max}$  - максимум модуля второй производной функции распределения.

В приложении П.3 для типовых законов приведены результаты определения  $|F''(x)|_{\max}$  и  $dx$ , воспользовавшись которыми необходимо определить:

$$M = \text{ent} \left[ \frac{x_{\max} - x_{\min}}{\Delta x} \right]. \quad (1.3)$$

3. Определить частоту попадания анализируемой случайной величины в  $j$ -ый дифференциальный коридор:

$$\mathfrak{F}_j = \frac{1}{N} \sum_{i=1}^N \delta_{ij}, \quad (1.4)$$

где

$$\delta_{ij} = \begin{cases} 1, & \text{если } \text{ent} \left[ \frac{x_i - x_{\min}}{\Delta x} \right] + 1 = j \wedge x_i = x_{\max}; \\ \frac{1}{2}, & \text{если } x_i = j \Delta x; \\ 0, & \text{иначе,} \end{cases} \quad (1.5)$$

- индикатор состояния.

Следует отметить, что  $\delta_{i, j+1} = 1/2$ , если  $x_i = j \Delta x \wedge x_i \neq x_{\max}$ , т.е. в этом случае в  $j$  и  $j+1$  коридоры добавляется по  $1/2$ .

4. Если частота попадания в какой-либо  $k$ -ый дифференциальный коридор мала ( $p_j < 0,01 \div 0,02$ ), для уменьшения влияния случайности его объединяют с  $k+1$  коридором. Эта операция может быть применена неоднократно.

Исходным материалом для построения гистограммы является сгруппированный по дифференциальным коридорам статистический ряд, представленный, как правило, в виде таблицы (см. таблицу 1.1), где  $\mathfrak{F}_j = \mathfrak{F}_j / \Delta x_j$ .

Статистический ряд

Таблица 1.1

$\mathfrak{F}_j$	0,099	0,1006	0,1003	0,0989	0,099	0,1067	0,0954	0,1008	0,0997	0,0996
$j \Delta x$	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
$\mathfrak{F}_j$	0,99	1,006	1,003	0,989	0,99	1,067	0,954	1,008	0,997	0,996

На рис.1.2 - 1.3 для разного количества дифференциальных коридоров приведены гистограммы случайной величины, распределённой по равномерному закону в диапазоне  $[0,1]$ , называемой «белым шумом», сгенерированной с помощью стандартной программы на ЭВМ ( $N=10000$ ). Гистограмма, представленная на рис.1.2, соответствует таблице 1.1.

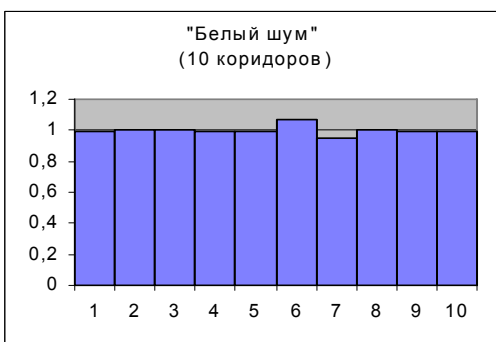


Рисунок 1.2.

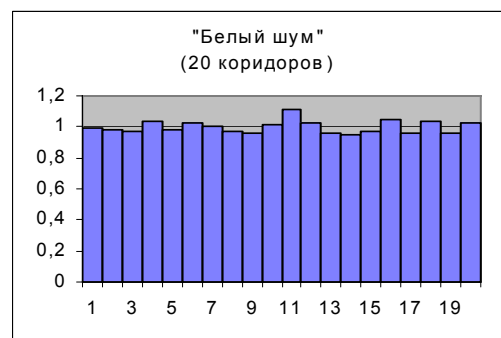


Рисунок 1.3.

Сравнительный анализ рисунков показывает, что при меньшем числе дифференциальных коридоров и большей частоте попадания случайной величины в каждый из них, гистограмма лучше приближается к теоретической плотности распределения вероятностей –  $f_x(x)=1$ .

Отметим, что «белый шум», являясь исходным шумом при решениях самых разных задач, широко применяется при моделировании случайных величин с требуемыми вероятностными характеристиками [112].

Построив гистограмму, можно приступить к задаче идентификации, заключающейся в подборе кривой распределения  $f_a(x, \alpha_1, \dots, \alpha_m)$  с неизвестными параметрами  $\alpha_1, \dots, \alpha_m$ , с той или иной точки зрения наилучшим образом описывающей данное статистическое распределение.

Решение этой задачи возможно различными методами, например: по совпадению моментов [45], с использованием фазовых портретов [72, 103] и т. д. (см. подраздел 1.3).

После выбора аналитического выражения с неизвестными параметрами, приступают к решению задачи аппроксимации в соответствии с выбранным критерием приближения.

Классификация методов аппроксимации законов распределения, предложенная на основании общей классификации методов аппроксимации (см. рис. В.3) и анализа литературных источников, представлена на рис. 1.4.



Рисунок 1.4. Классификация методов аппроксимации законов распределения

Следует при этом иметь в виду, что, независимо от метода, аналитические функции  $f_a(x, \alpha_1, \dots, \alpha_m)$  или  $F_a(x, \alpha_1, \dots, \alpha_m)$ , с помощью которых происходит аппроксимация статистического распределения, должна удовлетворять основным свойствам законов распределения:

$$\left\{ \begin{array}{l} f_a(x, \alpha_1, \dots, \alpha_m) \geq 0; \\ \int_{-\infty}^{\infty} f_a(x, \alpha_1, \dots, \alpha_m) dx = 1. \end{array} \right. \quad \left\{ \begin{array}{l} 0 \leq F_a(x, \alpha_1, \dots, \alpha_m) \leq 1; \\ F_a'(x, \alpha_1, \dots, \alpha_m) > 0. \end{array} \right. \quad (1.6)$$

Вопросы аппроксимации законов распределения различными методами рассмотрены в разделах 1.4-1.6.

Заключительным шагом решения задачи сглаживания статистических рядов является проверка качества аппроксимации с использованием критериев согласия. Если получено удовлетворительное решение задачи, то процесс сглаживания завершен. Если нет, необходимо повторить все этапы, начиная с идентификации.

Для возможности анализа различных методов сглаживания критерии согласия рассматриваются в разделе 1.2.

## 1.2. Критерии согласия

Предположим, что, исходя из вида гистограммы и теоретических предпосылок, выбрана функция  $f_a(x, \alpha_1, \dots, \alpha_m)$  с неизвестными параметрами  $\alpha_1, \dots, \alpha_m$ . Требуется подобрать эти параметры так, чтобы функция  $f_a(x, \alpha_1, \dots, \alpha_m)$  наилучшим образом описывала данный статистический материал в рамках принятого критерия.

Как бы хорошо ни была подобрана функция  $f_a(x, \alpha_1, \dots, \alpha_m)$ , между ней и статистическим распределением неизбежны расхождения. Для ответа на вопрос, связаны ли эти расхождения с ограниченным числом наблюдений, или они связаны с неправильным выбором  $f_a(x, \alpha_1, \dots, \alpha_m)$ , служат «критерии согласия».

Идея применения критериев согласия заключается в следующем. На основании данного статистического материала необходимо проверить гипотезу  $H$ , состоящую в том, что случайная величина  $X$  подчиняется некоторому закону распределения. Введем некоторую случайную величину  $U$ , являющуюся некоторой мерой расхождения теоретического и статистического распределений. Закон распределения этой случайной величины  $f_u(u)$  зависит как от закона распределения случайной величины  $X$ , так и от числа опытов  $N$ . Если гипотеза  $H$  верна, то  $f_u(u)$  определяется законом распределения  $f_a(x)$  и числом опытов  $N$ .

Вычислим вероятность события  $P(u \leq U) = P_d$ . Если эта вероятность мала, то гипотезу следует отвергнуть как маловероятную, если значительна - экспериментальные данные не противоречат гипотезе  $H$ .

Выберем в качестве меры расхождения между теоретическим и статистическим распределениями случайную величину

$$U = \sum_{j=1}^M c_j (\hat{p}_j - p_j)^2, \quad (1.7)$$

где  $c_j$  - веса;

$p_j = F_a(x_{j+1}, \alpha_1, \dots, \alpha_m) - F_a(x_j, \alpha_1, \dots, \alpha_m)$  – теоретические вероятности, соответствующие  $f_j$ .

Коэффициенты  $c_j$  вводятся для учёта веса отклонений, относящихся к разным разрядам. Так отклонения могут быть малозначительными, если вероятность  $p_j$  велика, и наоборот.

К. Пирсон показал, что если положить  $c_j = N/p_j$ , то при большом  $N$   $f_u(u)$  не зависит от  $f_x(x)$  и  $N$ , а зависит только от числа разрядов  $M$ . Этот закон при увеличении  $N$  приближается к закону  $\chi^2$  с  $r$  степенями свободы, плотность распределения вероятностей которого определяется выражением [45]:

$$f_r(u) = \begin{cases} \frac{1}{2^{\frac{r}{2}} \Gamma\left(\frac{r}{2}\right)} u^{\frac{r}{2}-1} e^{-\frac{u}{2}}, & \text{если } u > 0; \\ 0, & \text{если } u < 0, \end{cases} \quad (1.8)$$

где  $\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1} e^{-t} dt$  - гамма-функция [58]. (1.9)

В этом случае мера расхождения обозначается  $\chi^2$  и определяется выражением:

$$\chi^2 = N \sum_{j=1}^M \frac{(f_j - p_j)^2}{p_j}. \quad (1.10)$$

Иногда для удобства вычислений, чтобы избежать вычислений с промежуточными малыми числами, выражение (1.9) представляют в виде

$$\chi^2 = \sum_{j=1}^M \frac{(n_j - N p_j)^2}{N p_j}, \quad (1.11)$$

где  $n_j$  - число попаданий  $x$  в  $j$ -ый дифференциальный коридор.

Распределение  $\chi^2$  зависит от параметра  $r$ , называемого числом «степеней свободы». Число степеней свободы равно  $r = M - k$ , где  $k$  – число независимых наложенных условий (связей), например, условие нормировки, совпадение теоретических и статистических моментов и т.д.

Для распределения  $\chi^2$  составлены специальные таблицы (см. приложение П.4) [45]. Пользуясь ими, можно для каждого значения  $\chi^2$  и числа степеней свободы  $r$  найти вероятность того, что величина, распределенная по закону  $\chi^2$ , превзойдет это значение. Если эта вероятность мала, то результат опыта следует признать противоречащим гипотезе о том, что случайная величина распределена по предполагаемому закону. Если на практике она оказывается меньше, чем 0,1, рекомендуется проверить результаты эксперимента и, если это возможно, повторить его. В противном случае, для описания статистических данных следует искать более подходящее аналитическое выражение.

#### Схема применения критерия сводится к следующему:

1. определяется мера расхождения  $\chi^2$  по формуле (1.10) или (1.11);
2. определяется число степеней свободы  $r = M - k$ ;

3. по  $r$  и  $\chi^2$  по таблице П.4. определяется вероятность того, что величина, имеющая распределение  $\chi^2$  с  $r$  степенями свободы, превзойдет данное значение  $\chi^2$ ;

4. если вероятность  $P_d$  велика, то гипотеза  $H$  принимается.

Следует отметить, что с помощью критериев согласия можно только в некоторых случаях опровергнуть гипотезу  $H$ . Если вероятность  $P_d$  велика, то это указывает лишь на то, что гипотеза  $H$  не противоречит опытным данным [45].

Кроме критерия Пирсона на практике применяется критерий А.Н. Колмогорова [45].

В качестве меры расхождения между статистическим и теоретическим распределениями рассматривается величина, равная

$$\delta = \max \left| \hat{F}_x(x) - F_a(x, \alpha_1, \dots, \alpha_m) \right|. \quad (1.12)$$

Обоснованием выбора этой меры расхождения являются следующие обстоятельства:

- простота определения  $\delta$ ;
- при неограниченном числе испытаний  $N$  вероятность неравенства

$\delta \sqrt{N} \geq \lambda$  стремится к пределу

$$P(\lambda) = 1 - \sum_{k=-\infty}^{\infty} (-1)^k e^{-k^2 \lambda^2}. \quad (1.13)$$

Значения  $P(\lambda)$ , подсчитанные по формуле (1.13), приведены в приложении П.5.

#### **Схема применения этого критерия сводится к следующему:**

1. строятся статистическая и теоретическая функции распределения;

2. по графикам определяется  $\delta$ ;

3. вычисляется  $\lambda = \delta \sqrt{N}$ ;

4. по таблице приложения 3 определяется  $P(\lambda)$ ;

5. гипотеза принимается, если для заданного уровня значимости  $P_d$ :

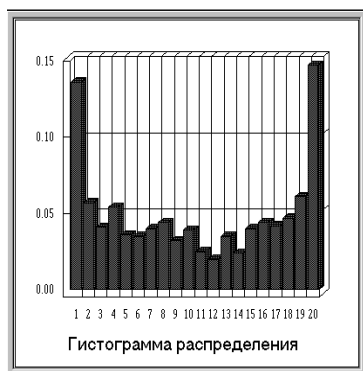
$$\lambda < \lambda_{1-P_d}; \quad (1.14)$$

6. Если вероятность  $P_d$  мала, гипотезу  $H$  следует отвергнуть. При сравнительно больших  $P(\lambda)$  её можно считать совместимой с опытными данными.

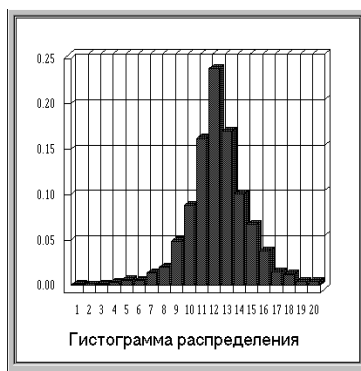
Следует отметить, что критерий Колмогорова можно применять, когда известно теоретическое распределение и численные значения его параметров, например, при проверке качества сгенерированной псевдослучайной последовательности с заданным законом распределения, расчете параметров функций распределения. Если применять этот критерий в случаях, когда параметры теоретического распределения определяются по статистическим данным, критерий даёт завышенное значение  $P(\lambda)$ , и мы рискуем принять неправильную гипотезу.

### 1.3. Идентификация законов распределения

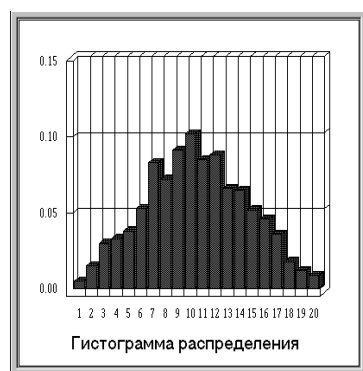
Выбор подходящего аналитического выражения с неизвестными параметрами, подлежащими определению в соответствии с выбранным критерием приближения, является самым сложным этапом процедуры сглаживания. Это объясняется тем, что задача плохо формализована, всегда существует опасность, что рассмотрены не все возможные варианты законов распределения, решение необходимо принимать по ограниченному набору входных данных и. т. д.



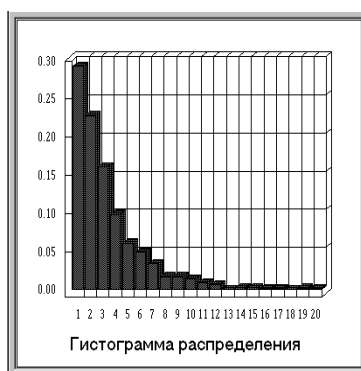
Закон арксинуса



Лапласа



Закон Симпсона



Экспоненциальный

Рисунок 1.5. Гистограммы типовых распределений

На практике задачу идентификации решают на основе анализа:

1. гистограммы и её соответствия кривой теоретического распределения;
2. числовых характеристик и их соответствия значениям теоретических числовых характеристик;
3. фазовых портретов (структурных функций).

При большом числе экспериментальных данных в ряде случаев по виду гистограммы можно высказать предположение о виде закона распределения. На рис. 1.5 в качестве примера приведены гистограммы случайных процессов ( $N=1000$ ), которые с большой долей уверенности можно отнести к тому или иному закону распределения.

При ограниченном наборе данных о законе распределения судят по совпадению значений статистических и теоретических числовых характеристик.

К числовым характеристикам относятся: начальные и центральные моменты  $k$ -го порядка, мода, медиана, коэффициенты вариации и формы, пикфактор.

- Начальным моментом  $k$ -го порядка называется величина, определяемая выражением:

$$\alpha_k = M[x^k] = \int_{-\infty}^{\infty} x^k f_x(x) dx, \quad (1.15)$$

где  $M[ ]$  - операция нахождения математического ожидания.

Из начальных моментов наиболее часто применяются:

- $\alpha_1 = M[x] = m_x$  - математическое ожидание, характеризующие среднее значение оцениваемой случайной величины;
- $\alpha_2 = M[x^2]$  - определяющий мощность случайной величины.



- Центральные моменты  $k$ -го порядка определяются в виде:

$$\mu_k = M[x^k], \quad (1.16)$$

где  $x = x - m_x$  - центрированная случайная величина.

Из них наиболее часто применяются:

- $\mu_2 = M[x^2] = \sigma_x^2$  - дисперсия процесса, характеризующая разброс значений случайной величины относительно математического ожидания;

- $\mu_3 = M[x^3]$  - характеризует симметричность закона распределения. Для устранения зависимости  $\mu_3$  от дисперсии вводят коэффициент асимметрии:

$$K_a = \mu_3 / \sigma_x^3;$$

- $\mu_4 = M[x^4]$  - характеризует острровершинность закона распределения. В качестве базового закона (закона сравнения) был выбран нормальный закон, для которого  $\mu_4 = 3\sigma_x^4$ . Для удобства сравнения ввели понятие коэффициент эксцесса:

$$K_\varepsilon = \mu_4 / \sigma_x^4 - 3.$$

- Мода -  $M_0$  - наиболее вероятное значение случайной величины  $x$ , определяемое в результате решения уравнения:

$$f'_x(x) = 0 \quad (1.17)$$

- Медиана -  $Me$  - определяется из уравнения

$$P\{x < Me\} = 0,5 \quad (1.18)$$

и делит площадь под кривой распределения пополам.

- Коэффициент вариации -  $\chi = \sigma_x / m_x$  - применяется при метрологическом анализе результатов измерения.

- Коэффициент формы -

$K_\phi = M[|x|] / \sigma_x$  - характеризует форму кривой распределения и применяется, например, при оценке релейной корреляционной функции (см. раздел 3).

- Пикфактор -  $\eta = |x_{\max}| / \sigma_x$  - применяется при оценке динамического диапазона случайной величины.

Значения числовых характеристик для типовых законов распределения приведены в приложении 1.

На рис. 1.6 в качестве примера приведены результаты определения числовых характеристик случайного

Выберите закон распределения :	
Арсинуса	
Параметры закона распределения :	
a : 2	
Начальные моменты :	
<u>Статистические :</u>	<u>Теоретические :</u>
$\alpha_1 = 0,07412$	$\alpha_1 = 0,000$
$\alpha_2 = 2,05834$	
$\alpha_3 = 0,16388$	
$\alpha_4 = 6,17581$	
Центральные моменты :	
<u>Статистические :</u>	<u>Теоретические :</u>
$\mu_2 = 2,05284$	$\mu_2 = 2,000$
$\mu_3 = -0,29300$	$\mu_3 = 0,000$
$\mu_4 = 6,19497$	$\mu_4 = 6,000$
Прочие характеристики :	
<u>Статистические :</u>	<u>Теоретические :</u>
$K_a = -0,09962$	$K_a = 0,000$
$K_\varepsilon = -1,52996$	$K_\varepsilon = -1,500$
$K_\phi = 0,90690$	
$\chi = 1,39589$	
$\eta = 19,33030$	
$\sigma = 1,43277$	

Рисунок 1.6. Сравнительный анализ числовых характеристик случайной величины

процесса, сгенерированных на ЭВМ ( $N=1000$ ).

Сравнение теоретических и статистических числовых характеристик позволяет предположить, что случайный процесс распределен по закону арксинуса.

Однако более наглядным способом идентификации является идентификация, основанная на анализе фазовых портретов законов распределения. Под фазовым портретом будем понимать графическую зависимость, построенную в координатах:  $f_x(x)$  и  $f'_x(x)$ . На рисунке 1.7 приведены фазовые портреты типовых законов распределения.

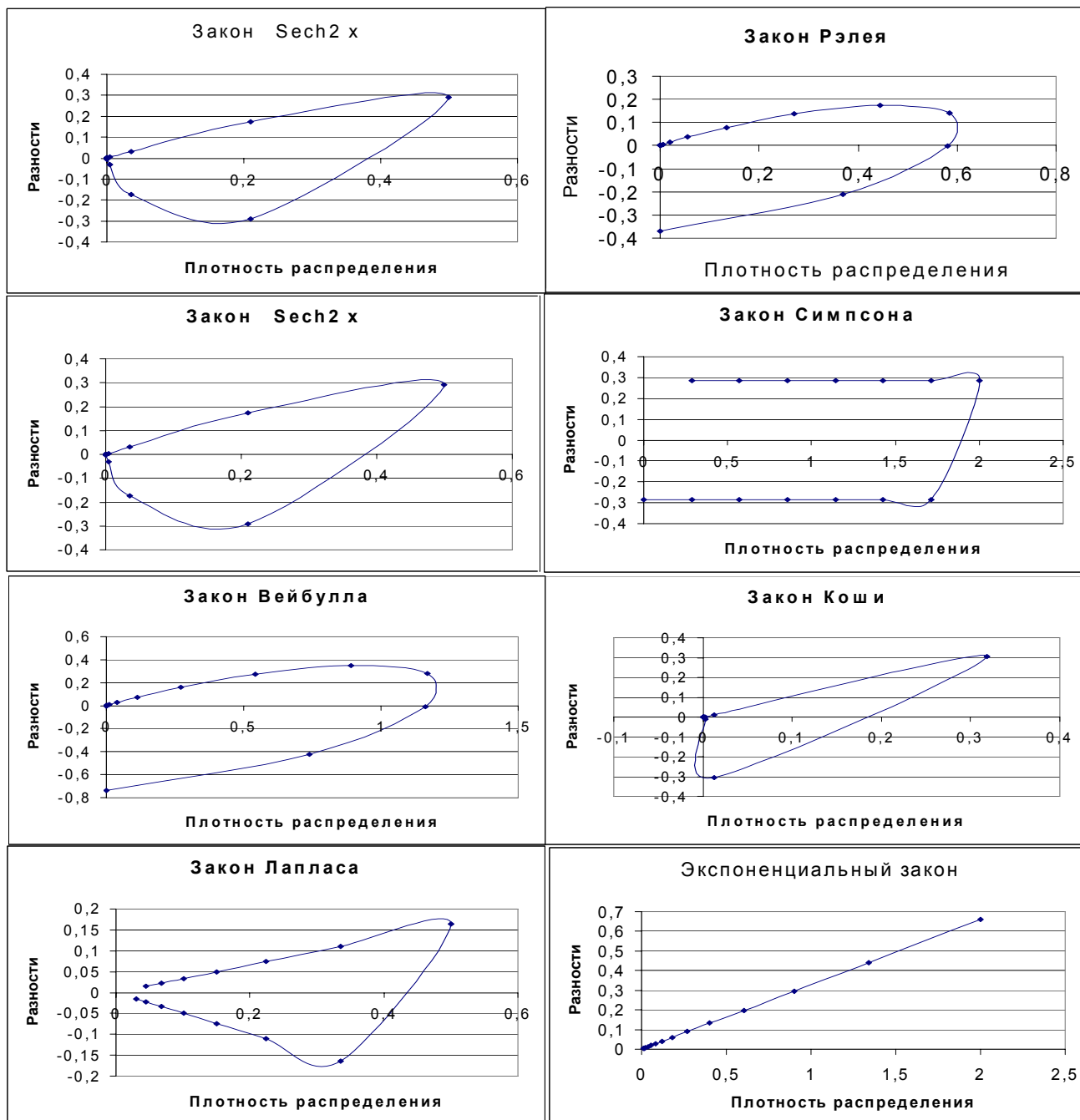
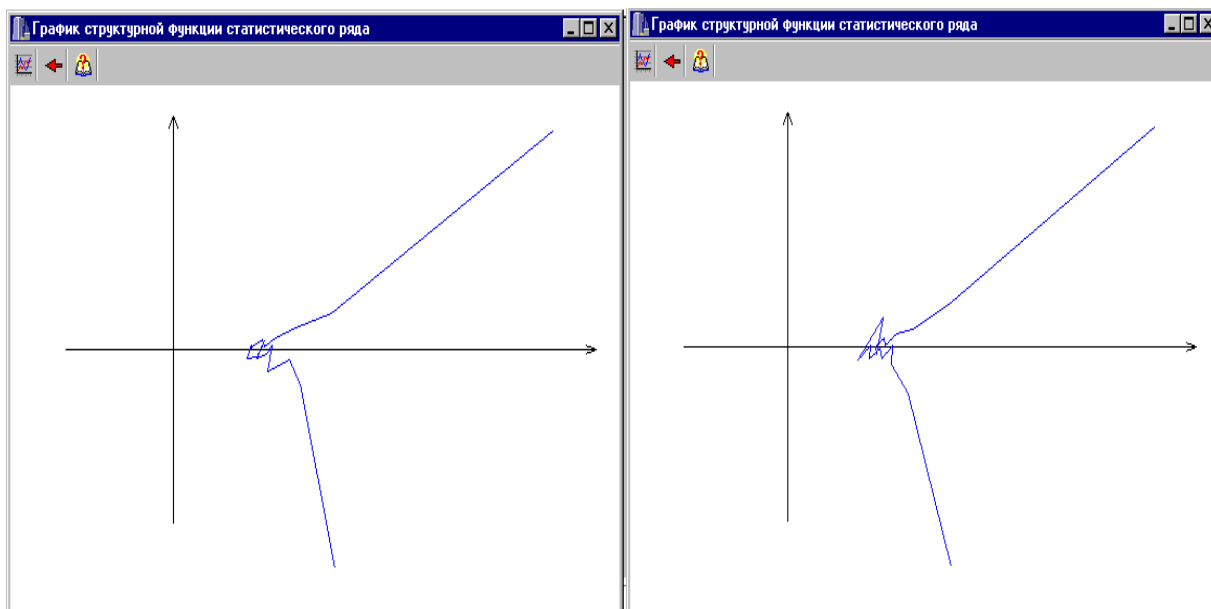


Рисунок 1.7. Фазовые портреты законов распределения

Следует отметить, что каждому закону распределения соответствует свой, уникальный фазовый протрет, не зависящий от значений параметров закона распределения. Заметим, что при расчете координат фазового портрета вместо точного значения

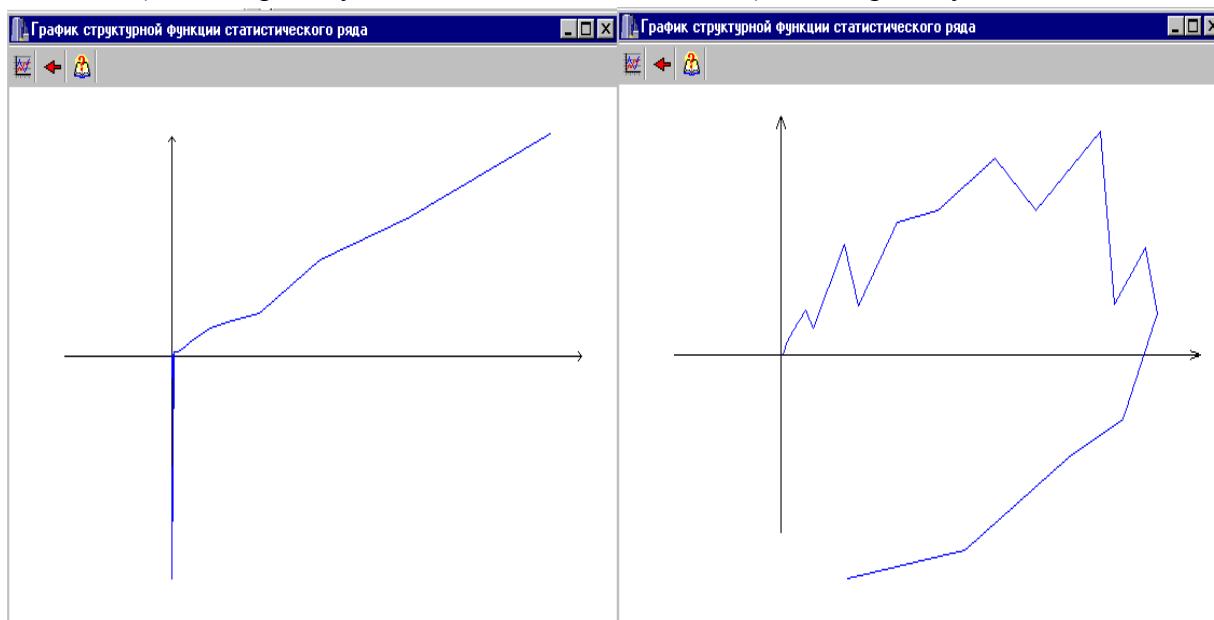
производной плотности распределения вероятностей определялось приращение плотности на заданном интервале.

В силу случайного характера исследуемых случайных величин фазовые портреты, построенные по экспериментальным данным, безусловно, будут отличаться от теоретических. Но эти отличия, тем не менее, позволяют идентифицировать случайные процессы по типу законов распределения. Ниже приведены фазовые портреты для случайного процесса, сгенерированного на ЭВМ по законам распределения арксинуса с разными параметрами, Рэлея и экспоненциальному закону.



а) закон арксинуса –  $a=5$

б) закон арксинуса –  $a=8$



в) экспоненциальное распределение -  $\lambda=1$

г) распределение Рэлея -  $\sigma=1$

Рисунок. 1.8. Фазовые портреты

Построив фазовый портрет по параметрам гистограммы, необходимо сравнить его с теоретическими фазовыми портретами и выбрать подходящий, т.е. выбрать соответствующий закон распределения. После этого необходимо перейти к определению параметров выбранного аналитического выражения.

## 1.4. Метод моментов

Один из методов, применяемый для решения задачи сглаживания статистических рядов, называется методом моментов [45]. Согласно этому методу, параметры  $\alpha_1, \dots, \alpha_m$  выбираются таким образом, чтобы несколько важнейших числовых характеристик (моментов) теоретического распределения были равны статистическим характеристикам. При составлении уравнений для определения неизвестных параметров, как правило, выбирают моменты низших порядков. Общими рекомендациями являются здравый смысл и простота решения полученной системы уравнений. Рассмотрим несколько примеров.

Определим параметры аналитического выражения плотности распределения вероятностей генератора «белого шума» - стандартной программы ПЭВМ. Теоретически закон распределения должен быть равномерным  $f_x(x) = \frac{1}{b-a}$ ,  $a \leq x \leq b$  с параметрами  $a=0$  и  $b=1$ .

Гистограмма приведена на рис. 1.2, а данные для расчётов - в таблице 1.1. Уравнения для определения двух неизвестных параметров распределения могут быть составлены различными способами. Потребуем, например, чтобы у статистического и теоретического распределений совпадали математическое ожидание и дисперсия:

$$\begin{cases} \bar{m}_x = \frac{a+b}{2}; \\ \sigma_x^2 = \frac{(b-a)^2}{12}. \end{cases} \quad (1.19)$$

Отметим, что оценка начальных моментов статистического ряда определяется выражением:

$$\bar{m}_k = \sum_{j=1}^M \bar{x}_j^k \cdot f_j, \quad (1.20)$$

где  $\bar{x}_j$  - среднее значение  $j$  интервала,  
а центральных -

$$\bar{m}_k = \sum_{s=0}^k C_k^s \bar{m}_{k-s} \sigma_x^s. \quad (1.21)$$

Эта система уравнений имеет аналитическое решение:

$$\begin{cases} a = \bar{m}_x - \sqrt{3} \sigma_x; \\ b = \bar{m}_x + \sqrt{3} \sigma_x. \end{cases} \quad (1.22)$$

Для данного статистического распределения

$$\begin{cases} \bar{m}_x = 0,49994; \\ \sigma_x^2 = 0,082208; \\ \sigma_x = 0,286719. \end{cases}$$

Подставив найденные оценки в выражения (1.22), получим:  $a=0,003327$ ,  $b=0,996553$ . Отсюда видно, что рассчитанные параметры закона распределения незначительно, но отличаются от заданного при генерировании. Следовательно, при

проведении статистического моделирования целесообразно проверять качество программных генераторов и оценивать его реальные характеристики.

Применив критерий Пирсона, вычислим по формуле (1.10) значение  $\chi^2 = 7,77$ , что соответствует вероятности  $P_d > 0,3$  (см. приложение 4). Таким образом, можно принять гипотезу о том, что данный статистический ряд соответствует равномерному распределению с найденными параметрами.

Преимуществом метода моментов является простота определения параметров распределения, недостатком - неоднозначность в выборе уравнений, которых может быть большое количество. Так, например, оценив  $\mu_1, \mu_2, \mu_3, \mu_4$ , при определении параметра распределения Релея  $\sigma$  возможно использование следующих уравнений (см. приложение 1):

$$\sigma = \frac{\mu_1}{1,25}; \quad \sigma = \sqrt{\frac{\mu_2}{0,43}}; \quad \sigma = \sqrt[3]{\frac{\mu_3}{0,1775}}; \quad \sigma = \sqrt[4]{\frac{\mu_4}{0,5978}}. \quad (1.23)$$

Результаты расчетов при определении параметра закона распределения Релея при различном числе дифференциальных коридоров  $M$  и различном объеме выборки приведены в таблице 1.2. При этом, с помощью ЭВМ генерировалась псевдослучайная последовательность, распределенная по закону Релея с параметром распределения  $\sigma = 1$ , а параметр распределения определялся в результате решения одного из уравнений (1.23) через соответствующие моменты.

Результаты экспериментальных исследований при определении параметра закона распределения Рэлея

Таблица 1.2

M	N	$\mu_1$	$\mu_2$	$\mu_3$	$\mu_4$
10	300	0,9944	1,0585	1,0626	1,0502
	1000	1,0051	1,0320	1,0565	1,040
	5000	0,9947	0,9848	1,0040	0,9931
	10000	1,0086	1,0039	1,0000	1,0074
20	300	1,0209	1,0162	1,0692	1,0378
	1000	0,9989	1,0068	1,0130	0,9975
	5000	1,0086	1,0079	1,0201	1,0141
	10000	1,0072	1,0065	1,0220	1,0165

Анализ результатов показывает, что при составлении уравнений для определения параметров теоретического распределения по методу моментов необходимо выбирать моменты более низких параметров.

Приведем аналитические решения систем уравнений для типовых законов распределений (см. таблицу 1.3)

Системы уравнений для типовых законов распределений

Таблица 1.3

№	Название	Функции распределения	Системы уравнений
1	Арксинуса	$\begin{cases} 0, (-\infty < x < -a) \\ \frac{1}{2} + \frac{1}{\pi} \arcsin \frac{x}{a}, (-a < x < a) \\ 1, (a < x < \infty) \end{cases}$	$a = \sqrt{2} \sigma_x$
2	Лапласа	$\begin{cases} \frac{1}{2} e^{\lambda(x-\mu)}, (-\infty < x < \mu) \\ 1 - \frac{1}{2} e^{-\lambda(x-\mu)}, (\mu < x < \infty) \end{cases}$	$\begin{cases} \lambda = \frac{\sqrt{2}}{\sigma_x} \\ \mu = m_x \end{cases}$
3	Симпсона	$\begin{cases} 0, (-\infty < x < a) \\ \frac{2(x-a)^2}{(b-a)^2}, \left(a < x < \frac{a+b}{2}\right) \\ 1 - \frac{2(b-x)^2}{(b-a)^2}, \left(\frac{a+b}{2} < x < b\right) \\ 1, (b < x < \infty) \end{cases}$	$\begin{cases} a = m_x - \sqrt{6} \sigma_x \\ b = m_x + \sqrt{6} \sigma_x \end{cases}$
4	Sech <sup>2</sup>	$\frac{1}{2} + \frac{1}{2} \operatorname{th} ax$	$a = \frac{\pi}{2\sqrt{3} \sigma_x}$
5	Рэля	$\begin{cases} 0, (-\infty < x < 0) \\ 1 - e^{-\frac{x^2}{2\sigma^2}}, (0 < x < \infty) \end{cases}$	$\sigma = \frac{m_x}{1,25}$
6	Экспоненциальный	$\begin{cases} 0 < (-\infty < x < 0) \\ 1 - e^{-\alpha x}, (0 < x < \infty) \end{cases}$	$\alpha = \frac{1}{m_x}$

На рис. 1.8-1.9 представлены результаты эксперимента - гистограмма, полигон частот и плотность распределения вероятностей для закона Рэля при N=300, M=10,  $\sigma=1,00084$ .

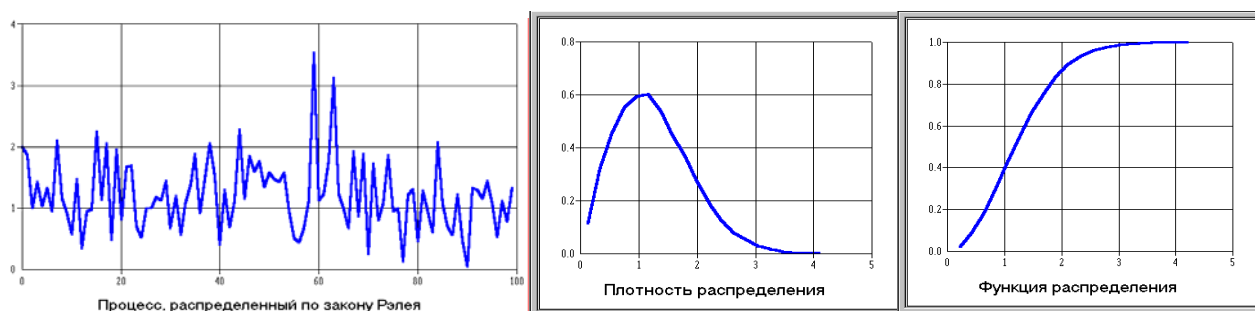


Рисунок 1.8. Результаты экспериментальных исследований

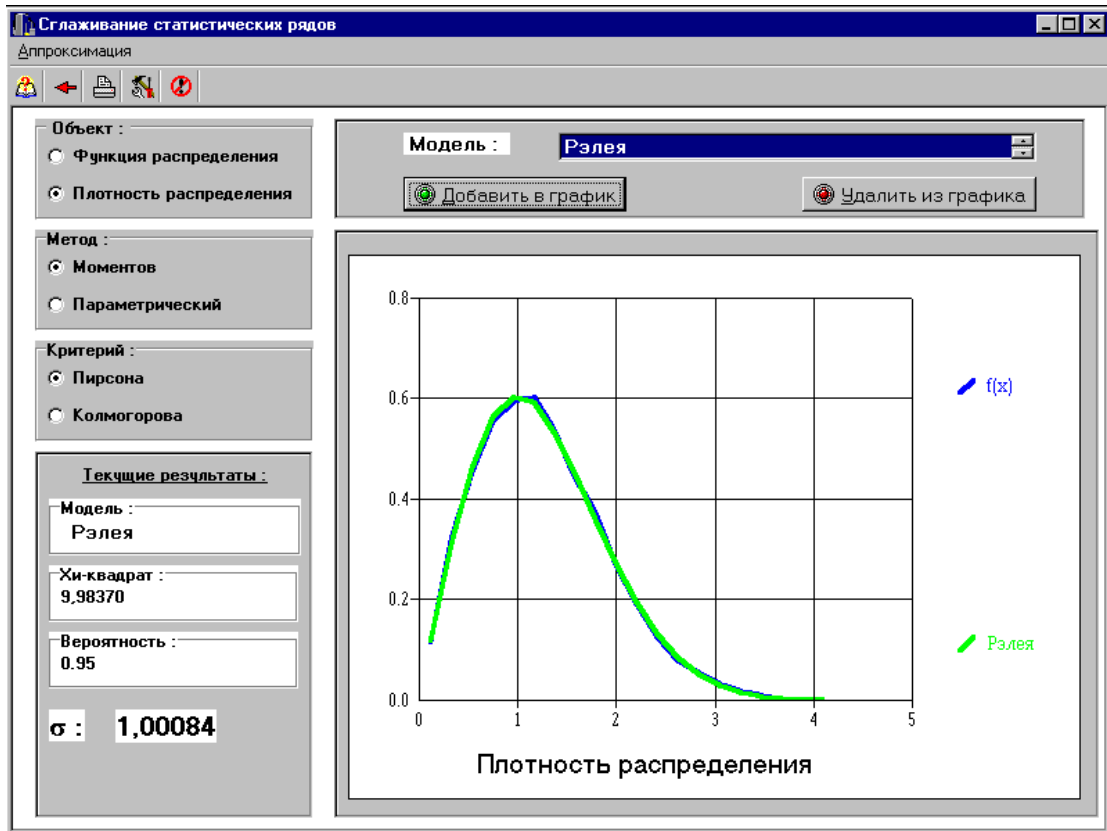
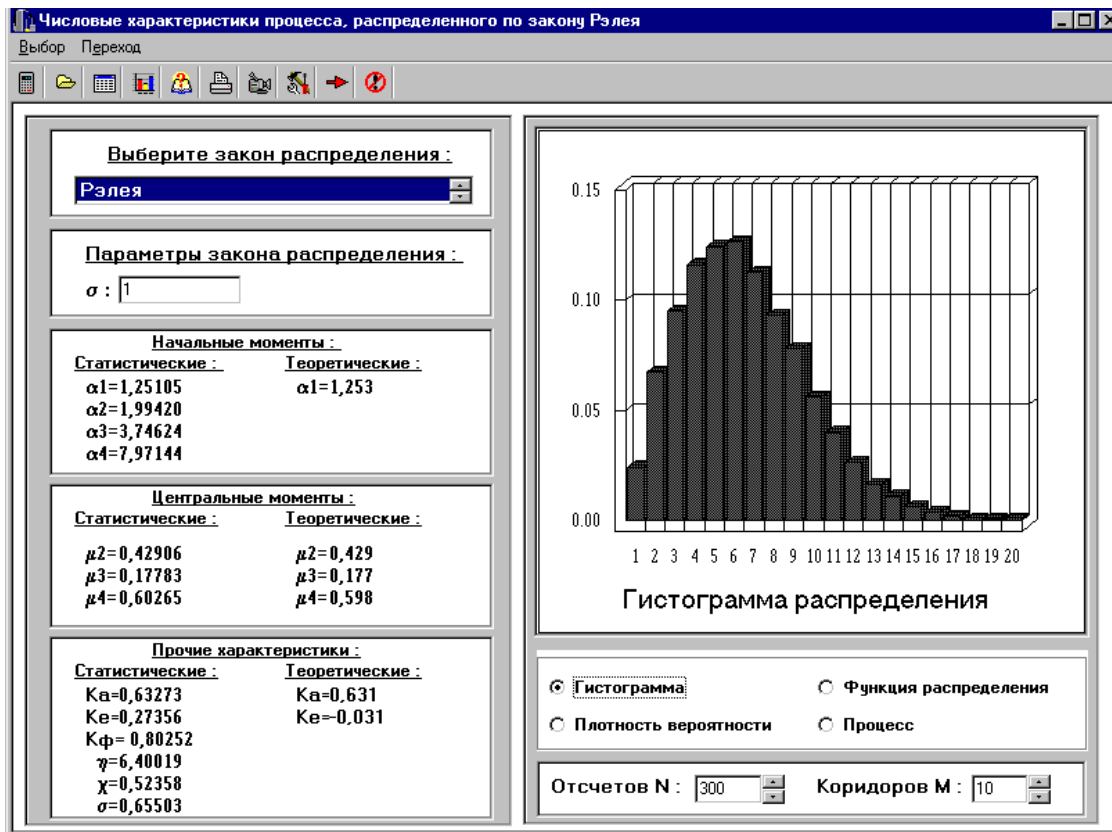


Рисунок 1.9. Определения параметров Рэлея методом моментов

## 1.5. Аппроксимация плотностей распределения вероятностей параметрическими моделями

Другим способом решения задачи сглаживания статистических рядов является определение параметров аналитического выражения, удовлетворяющих минимуму квадратической погрешности аппроксимации:

$$\Delta = \sum_{j=1}^M [\hat{f}_x(\bar{x}_j) - f_a(\bar{x}_j, \beta_1, \beta_2, \dots)]^2 = \min, \quad (1.24)$$

где  $M$  - число дифференциальных коридоров;

$\hat{f}_x(\bar{x}_j) = \hat{f}_j / \Delta_j$  - значение плотности распределения вероятностей в середине  $j$ -го дифференциального коридора  $\bar{x}_j$ ;

$f_a(\bar{x}_j, \beta_1, \beta_2, \dots)$  - аналитическое выражение с неизвестными параметрами.

Условиями минимума погрешности  $\Delta$  является следующая система уравнений:

$$\left\{ \begin{array}{l} \frac{\partial \Delta}{\partial \beta_1} = \sum_{j=1}^M [\hat{f}_x(\bar{x}_j) - f_a(\bar{x}_j, \beta_1, \beta_2, \dots)] \frac{\partial f_a(\bar{x}_j, \beta_1, \beta_2, \dots)}{\partial \beta_1} = 0 \\ \frac{\partial \Delta}{\partial \beta_2} = \sum_{j=1}^M [\hat{f}_x(\bar{x}_j) - f_a(\bar{x}_j, \beta_1, \beta_2, \dots)] \frac{\partial f_a(\bar{x}_j, \beta_1, \beta_2, \dots)}{\partial \beta_2} = 0. \\ \dots \dots \dots \end{array} \right. \quad (1.25)$$

Сложность этой системы зависит от вида аналитического выражения и числа неизвестных параметров, подлежащих определению. Как правило, решение этой системы возможно лишь приближенными методами.

Так, например, при однопараметрической аппроксимации с использованием метода Ньютона [62], неизвестный параметр определяется в результате решения следующего уравнения:

$$\beta_{n+1} = \beta_n - \frac{\sum_{j=1}^M [\hat{f}_x(\bar{x}_j) - f_a(\bar{x}_j, \beta_n)] \frac{\partial f_a(\bar{x}_j, \beta)}{\partial \beta}}{\sum_{j=1}^M \left\{ [\hat{f}_x(\bar{x}_j) - f_a(\bar{x}_j, \beta_n)] \frac{\partial^2 f_a(\bar{x}_j, \beta)}{\partial \beta^2} - \left[ \frac{\partial f_a(\bar{x}_j, \beta)}{\partial \beta} \right]^2 \right\}_{\beta=\beta_n}}. \quad (1.26)$$

В качестве начального приближения можно выбрать значение параметра, определенное по методу моментов.

Алгоритм завершает свою работу, когда выполняется следующее условие:

$$|\beta_{n+1} - \beta_n| \leq \varepsilon, \quad (1.27)$$

где  $\varepsilon$  - погрешность вычисления параметра, задаваемая исследователем.

Приведем пример расчета по методу Ньютона параметра аппроксимации плотности распределения вероятностей экспоненциального закона распределения.

Плотностью вероятности экспоненциального закона распределения является функция следующего вида:

$$f_a(\bar{x}_j, \alpha) = \alpha e^{-\alpha \bar{x}_j}, \quad \bar{x}_j > 0,$$

где  $\alpha$  - параметр, значение которого необходимо найти, решив уравнение:



$$\frac{\partial \Delta}{\partial \alpha} = \sum_{j=1}^M \left[ \mathbb{E}_x(\bar{x}_j) - f_a(\bar{x}_j, \alpha) \right] \frac{\partial f_a(\bar{x}_j, \alpha)}{\partial \alpha} = 0.$$

Для решения этого уравнения применим метод Ньютона (1.26). В этом уравнении присутствуют первая и вторая производные аппроксимирующей функции по неизвестному параметру. Вид этих производных приведен ниже.

$$\frac{\partial f_a(\bar{x}_j, \alpha)}{\partial \alpha} = e^{-\alpha \bar{x}_j} - \alpha \bar{x}_j e^{-\alpha \bar{x}_j} = e^{-\alpha \bar{x}_j} [1 - \alpha \bar{x}_j],$$

$$\frac{\partial^2 f_a(\bar{x}_j, \alpha)}{\partial \alpha^2} = -\bar{x}_j e^{-\alpha \bar{x}_j} [1 - \alpha \bar{x}_j] - \bar{x}_j e^{-\alpha \bar{x}_j} = -\bar{x}_j e^{-\alpha \bar{x}_j} [2 - \alpha \bar{x}_j].$$

Подставив выражения производных аппроксимирующей модели в уравнение (1.26), получим:

$$\alpha_{n+1} = \alpha_n - \frac{\sum_{j=1}^M \left[ \mathbb{E}_x(\bar{x}_j) - \alpha_n e^{-\alpha_n \bar{x}_j} \right] e^{-\alpha_n \bar{x}_j} (1 - \alpha_n \bar{x}_j)}{\sum_{j=1}^M \left\{ \left[ \mathbb{E}_x(\bar{x}_j) - \alpha_n e^{-\alpha_n \bar{x}_j} \right] [-\bar{x}_j e^{-\alpha_n \bar{x}_j} (2 - \alpha_n \bar{x}_j)] - e^{-2\alpha_n \bar{x}_j} (1 - \alpha_n \bar{x}_j) \right\}}.$$

Аналогично рассчитываются параметры аналитических моделей и для других типовых законов распределения, выражения для производных которых приведены в таблице 1.4 и приложениях 6-7.

#### Производные законов распределения

Таблица 1.4

N	Название закона	$\frac{\partial f_a(\bar{x}_j, \beta)}{\partial \beta}$	$\frac{\partial^2 f_a(\bar{x}_j, \beta)}{\partial \beta^2}$
1	Арксинуса	$-\frac{\alpha}{\pi} (\alpha^2 - \bar{x}_j^2)^{-3/2}$	$3 \frac{\alpha^2}{\pi} (\alpha^2 - \bar{x}_j^2)^{-5/2} - \frac{1}{\pi} (\alpha^2 - \bar{x}_j^2)^{-3/2}$
2	Релея	$e^{-\frac{\bar{x}_j^2}{2\sigma^2}} \left( \frac{\bar{x}_j^3}{\sigma^5} - \frac{2\bar{x}_j}{\sigma^3} \right)$	$e^{-\frac{\bar{x}_j^2}{2\sigma^2}} \left( \frac{\bar{x}_j^5}{\sigma^8} - \frac{7\bar{x}_j^3}{\sigma^6} + \frac{6\bar{x}_j}{\sigma^4} \right)$
3	Sech <sup>2</sup>	$\frac{1}{2} \text{ch}^{-2}(a\bar{x}_j) - a\bar{x}_j \times$ $\times \text{ch}^{-3}(a\bar{x}_j) \text{sh}(a\bar{x}_j)$	$\text{ch}^{-3}(a\bar{x}_j) \text{sh}(a\bar{x}_j) (1 + \bar{x}_j) +$ $a\bar{x}_j^2 \text{ch}^{-2}(a\bar{x}_j) (3\text{ch}^{-2}(a\bar{x}_j) \text{sh}^2(a\bar{x}_j) - 1)$
4	Экспоненциальный	$e^{-\alpha \bar{x}_j} (1 - \alpha \bar{x}_j)$	$-\bar{x}_j e^{-\alpha \bar{x}_j} (2 - \alpha \bar{x}_j)$

На рис. 1.10 последовательно приведены результаты определения параметра экспоненциального распределения рассматриваемым способом.

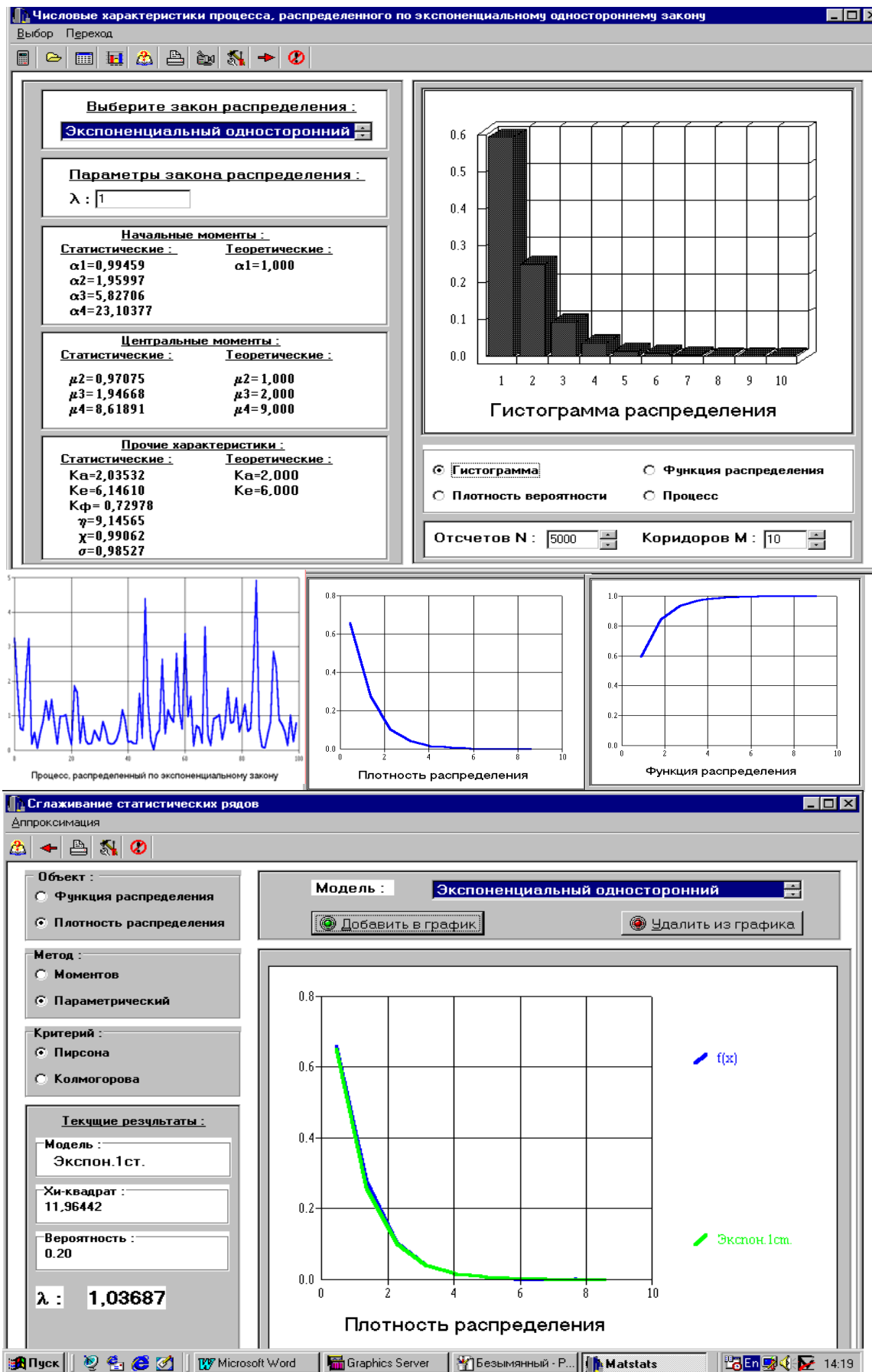


Рисунок 1.10. Определение параметра модели экспоненциального распределения

Несмотря на положительные качества метода Ньютона, позволяющие решать уравнения с достаточной скоростью и точностью, применение его не всегда оправдано, поскольку этот метод очень критичен к начальным условиям. Подобрать же удовлетворяющие сходимости метода Ньютона значения удается не для всех законов распределения, поэтому в некоторых случаях задачу нахождения неизвестных параметров приходится решать более простым методом хорд.

Известно, что если функция  $f_x(x)$  имеет разные знаки на концах интервала  $[a, b]$ , и производная этой функции сохраняет на данном интервале неизменный знак, то внутри промежутка находится один-единственный корень уравнения  $f_x(x)=0$  (если производная не сохраняет знака, то корень тоже есть, но он может быть не единственным).

Суть метода хорд заключается в последовательном сужении интервала, на котором находится корень [62]. Для этого следует определить интервал, на концах которого исследуемая функция  $f_x(x)$  имеет противоположные знаки.

В качестве первого приближения корня уравнения  $f_x(x)=0$  принимаем точку  $x_1$  пересечения оси ОХ и отрезка, соединяющего значение функции на концах интервала (этот отрезок является хордой). Значение  $x_1$  вычисляется по формулам:

$$x_1 = a - \frac{(b-a)f_x(a)}{f_x(b) - f_x(a)}, \quad (1.28)$$

или

$$x_1 = b - \frac{(b-a)f_x(b)}{f_x(b) - f_x(a)}. \quad (1.29)$$

Затем вычисляется значение функции  $f_x(x)$  в точке  $x_1$ , и выбирается тот интервал  $(a, x_1)$  или  $(x_1, b)$ , на концах которого  $f_x(x)$  имеет противоположные знаки. Искомый корень лежит в этом интервале и т.д.

Приведем, применительно к рассматриваемому случаю нахождения параметра аппроксимирующей модели, формулы, необходимые для работы метода хорд.

В качестве начального приближения значений  $a$  и  $b$  можно взять значения параметра, рассчитанные по методу моментов для данного закона распределения. В противном случае необходимо подбирать значения начальных границ рабочего интервала, исходя из здравого смысла и вида закона распределения.

Определим

$$f(a) = \sum_{j=1}^M \left( f_x(\bar{x}_j) - f_a(\bar{x}_j, a) \right) \frac{\partial f_a(\bar{x}_j, a)}{\partial a}, \quad (1.30)$$

$$f(b) = \sum_{j=1}^M \left( f_x(\bar{x}_j) - f_a(\bar{x}_j, b) \right) \frac{\partial f_a(\bar{x}_j, b)}{\partial b}. \quad (1.31)$$

Полученные значения  $f(a)$  и  $f(b)$  подставляем в формулу (1.28) или (1.29) для вычисления первой точки  $x_1$  пересечения хорды и оси ОХ.

Найденное значение  $x_1$  используем в формуле (1.32) для установления значения функции  $f(x_1)$ .

$$f(x_1) = \sum_{j=1}^M \left( f_x(\bar{x}_j) - f_a(\bar{x}_j, x_1) \right) \frac{\partial f_a(\bar{x}_j, x_1)}{\partial x_1}. \quad (1.32)$$

Приведем пример расчетов параметра аппроксимирующей модели методом хорд для закона распределения арксинуса (см. рис. 1.11)

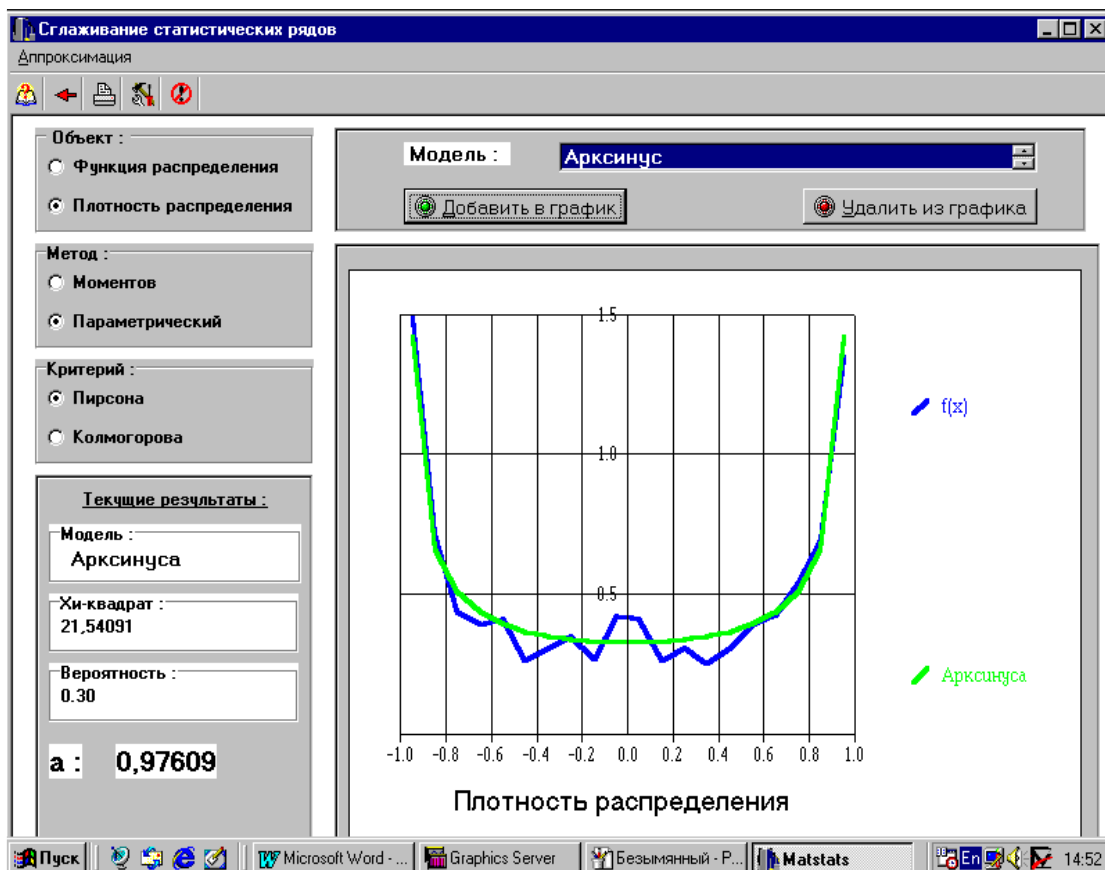
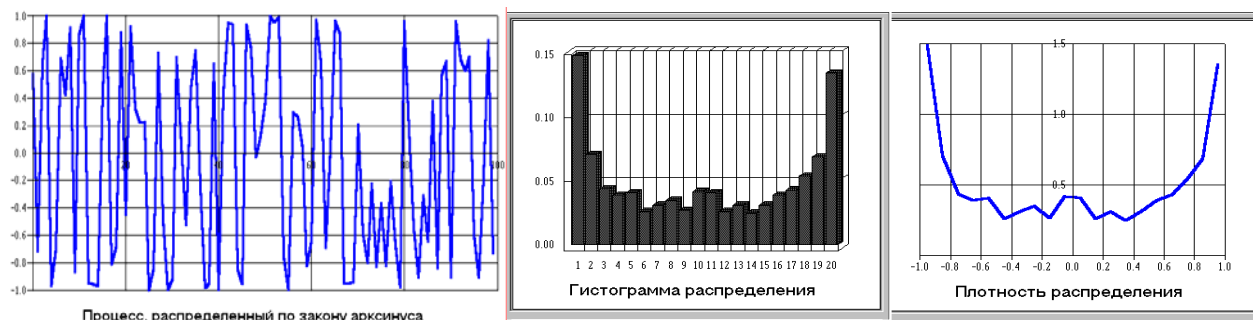


Рисунок 1.11. Результаты определения параметров закона распределения арксинуса

В качестве начальных значений  $a$  и  $b$  возьмем значения, рассчитанные по методу моментов:

$$a = \sqrt{2\widehat{D}_X} - 0.5,$$

$$b = \sqrt{2\widehat{D}_X} + 0.5.$$

Определим значения  $f(a)$  и  $f(b)$  в точках  $a$  и  $b$ :

$$f(a) = \sum_{j=1}^M \left( \epsilon_x(\bar{x}_j) - \frac{1}{\pi\sqrt{a^2 - \bar{x}_j^2}} \right) \frac{-a}{\pi\sqrt{(a^2 - \bar{x}_j^2)^3}},$$

$$f(b) = \sum_{j=1}^M \left( \mathbb{E}_x(\bar{x}_j) - \frac{1}{\pi\sqrt{b^2 - \bar{x}_j^2}} \right) \frac{-b}{\pi\sqrt{(b^2 - \bar{x}_j^2)^3}}.$$

Находим значение  $x_1$  и вычисляем функцию в этой точке:

$$x_1 = a - \frac{(b-a)f(a)}{f(b) - f(a)},$$

$$f(x_1) = \sum_{j=1}^M \left( \mathbb{E}_x(\bar{x}_j) - \frac{1}{\pi\sqrt{x_1^2 - \bar{x}_j^2}} \right) \frac{-x_1}{\pi\sqrt{(x_1^2 - \bar{x}_j^2)^3}}.$$

Проверяем знаки у функций и итерационно добиваемся требуемой точности.

Для нахождения параметров двухпараметрического закона распределения необходимо решить систему уравнений (1.25) для двумерного случая:

$$\begin{cases} f_1 = \sum_{j=1}^M \left[ \mathbb{E}_x(\bar{x}_j) - f_a(\bar{x}_j, \beta_1, \beta_2) \right] \frac{\partial f_a(\bar{x}_j, \beta_1, \beta_2)}{\partial \beta_1} = 0 \\ f_2 = \sum_{j=1}^M \left[ \mathbb{E}_x(\bar{x}_j) - f_a(\bar{x}_j, \beta_1, \beta_2) \right] \frac{\partial f_a(\bar{x}_j, \beta_1, \beta_2)}{\partial \beta_2} = 0 \end{cases}. \quad (1.33)$$

Решить эту систему можно только приближенными методами, например, методом Ньютона. Воспользовавшись формулой для решения системы двух уравнений с двумя неизвестными по методу Ньютона, получим:

$$\beta_1^{n+1} = \beta_1^n - \frac{1}{\Delta'} \left( \frac{\partial f_2}{\partial \beta_2} f_1(\beta_1^n, \beta_2^n) - \frac{\partial f_1}{\partial \beta_2} f_2(\beta_1^n, \beta_2^n) \right), \quad (1.34)$$

$$\beta_2^{n+1} = \beta_2^n - \frac{1}{\Delta'} \left( \frac{\partial f_1}{\partial \beta_1} f_2(\beta_1^n, \beta_2^n) - \frac{\partial f_2}{\partial \beta_1} f_1(\beta_1^n, \beta_2^n) \right), \quad (1.35)$$

где,

$$\Delta' = \frac{\partial f_1}{\partial \beta_1} \frac{\partial f_2}{\partial \beta_2} - \frac{\partial f_1}{\partial \beta_2} \frac{\partial f_2}{\partial \beta_1}.$$

Для вычислений необходимо знать значения частных производных по неизвестным параметрам функций  $f_1$  и  $f_2$ . Их выражения приведены в формулах (1.36) - (1.39).

$$\frac{\partial f_1}{\partial \beta_1} = \sum_{j=1}^M \left[ \mathbb{E}_x(\bar{x}_j) - f_a(\bar{x}_j, \beta_1, \beta_2) \right] \frac{\partial^2 f_a(\bar{x}_j, \beta_1, \beta_2)}{\partial \beta_1^2} - \left[ \frac{\partial f_a(\bar{x}_j, \beta_1, \beta_2)}{\partial \beta_1} \right]^2, \quad (1.36)$$

$$\frac{\partial f_2}{\partial \beta_2} = \sum_{j=1}^M \left[ \mathbb{E}_x(\bar{x}_j) - f_a(\bar{x}_j, \beta_1, \beta_2) \right] \frac{\partial^2 f_a(\bar{x}_j, \beta_1, \beta_2)}{\partial \beta_2^2} - \left[ \frac{\partial f_a(\bar{x}_j, \beta_1, \beta_2)}{\partial \beta_2} \right]^2, \quad (1.37)$$

$$\frac{\partial f_1}{\partial \beta_2} = \frac{\sum_{j=1}^M [\bar{f}_x(\bar{x}_j) - f_a(\bar{x}_j, \beta_1, \beta_2)] \frac{\partial^2 f_a(\bar{x}_j, \beta_1, \beta_2)}{\partial \beta_1 \partial \beta_2}}{\frac{\partial f_a(\bar{x}_j, \beta_1, \beta_2)}{\partial \beta_1} \frac{\partial f_a(\bar{x}_j, \beta_1, \beta_2)}{\partial \beta_2}}, \quad (1.38)$$

$$\frac{\partial f_2}{\partial \beta_1} = \frac{\sum_{j=1}^M [\bar{f}_x(\bar{x}_j) - f_a(\bar{x}_j, \beta_1, \beta_2)] \frac{\partial^2 f_a(\bar{x}_j, \beta_1, \beta_2)}{\partial \beta_2 \partial \beta_1}}{\frac{\partial f_a(\bar{x}_j, \beta_1, \beta_2)}{\partial \beta_2} \frac{\partial f_a(\bar{x}_j, \beta_1, \beta_2)}{\partial \beta_1}}. \quad (1.39)$$

Приведем пример расчета параметров аппроксимативной модели для закона распределения Симпсона.

Плотность распределения закона Симпсона:

$$f_a(\bar{x}_j, a, b) = \begin{cases} 0, & (-\infty < \bar{x}_j < a) \\ \frac{4(\bar{x}_j - a)}{(b-a)^2}, & \left( a < \bar{x}_j < \frac{a+b}{2} \right) \\ \frac{4(b - \bar{x}_j)}{(b-a)^2}, & \left( \frac{a+b}{2} < \bar{x}_j < b \right) \\ 0, & (b < \bar{x}_j < \infty) \end{cases}.$$

Приведем необходимые для расчетов частные первые, вторые и смешанные производные функций  $f(x)$ .

а)  $(a < x < (a+b)/2)$

$$\begin{aligned} \frac{\partial f}{\partial a} &= \frac{-4}{(b-a)^2} + 8 \frac{(\bar{x}_j - a)}{(b-a)^3}, & \frac{\partial f}{\partial b} &= 8 \frac{(\bar{x}_j - a)}{(b-a)^3}, \\ \frac{\partial^2 f}{\partial a^2} &= \frac{-16(\bar{x}_j - a)}{(b-a)^3} + \frac{24(\bar{x}_j - a)}{(b-a)^4}, & \frac{\partial^2 f}{\partial b^2} &= \frac{24(\bar{x}_j - a)}{(b-a)^4}, \\ \frac{\partial^2 f}{\partial a \partial b} &= \frac{8}{(b-a)^3} - \frac{24(\bar{x}_j - a)}{(b-a)^4}, & \frac{\partial^2 f}{\partial b \partial a} &= \frac{8}{(b-a)^3} - \frac{24(\bar{x}_j - a)}{(b-a)^4}. \end{aligned}$$

б)  $((a+b)/2 < x < b)$

$$\frac{\partial f}{\partial a} = \frac{8(b - \bar{x}_j)}{(b-a)^3}, \quad \frac{\partial f}{\partial b} = \frac{4}{(b-a)^2} - \frac{8(b - \bar{x}_j)}{(b-a)^3}, \quad \frac{\partial^2 f}{\partial a^2} = \frac{24(b - \bar{x}_j)}{(b-a)^4},$$

$$\frac{\partial^2 f}{\partial b^2} = \frac{-16}{(b-a)^3} + \frac{24(b-\bar{x}_j)}{(b-a)^4}, \quad \frac{\partial^2 f}{\partial a \partial b} = \frac{8}{(b-a)^3} - \frac{24(b-\bar{x}_j)}{(b-a)^4},$$

$$\frac{\partial^2 f}{\partial b \partial a} = \frac{8}{(b-a)^3} - \frac{24(b-\bar{x}_j)}{(b-a)^4}.$$

Полученные выражения необходимо подставить в выражения (1.36)-(1.39) и рассчитать частные производные функций  $f_1$  и  $f_2$ , которые используются в вычислениях неизвестных параметров путем последовательного приближения итерационных формул (1.34) и (1.35).

В качестве примера на рис. 1.12 приведены результаты определения параметров распределения Симпсона.

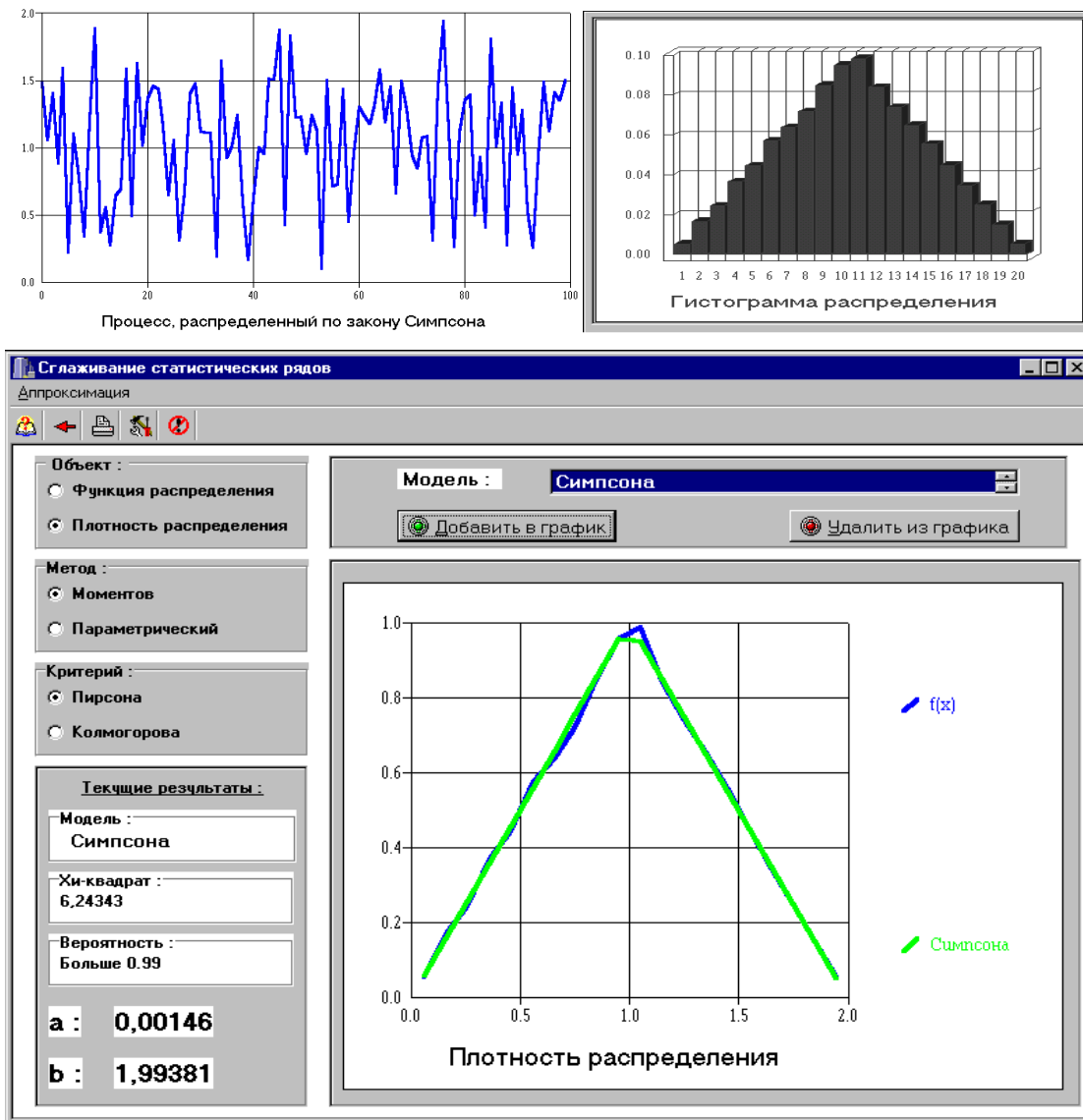


Рисунок 1.12. Результаты определения параметров закона распределения Симпсона (N=500)

Из результатов раздела видно, что метод аппроксимации плотностей распределения вероятностей параметрическими моделями значительно сложнее метода моментов. Применение его оправдано при небольших объемах выборки.

## 1.6. Аппроксимация функций распределения вероятностей параметрическими моделями

При аппроксимации плотностей распределения вероятностей в качестве аргумента используется середина дифференциального коридора, что, в свою очередь, вносит дополнительные погрешности при анализе асимметричных законов распределения. От этого недостатка свободна аппроксимация функций распределения вероятностей.

Задача аппроксимации статистического ряда функциями распределения вероятностей ставится аналогично задаче аппроксимации плотностей распределения вероятностей:

$$\Delta = \sum_{j=1}^M \left[ \bar{F}_x(x_j) - F_a(x_j, \beta_1, \beta_2, \dots) \right]^2 = \min, \quad (1.40)$$

где  $M$  - число дифференциальных коридоров;

$$\bar{F}_x(x_j) = \sum_{s=1}^j \bar{F}_s - \text{значение функции распределения вероятностей в конце } j\text{-го}$$

дифференциального коридора  $x_j$ ;

$F_a(x_j, \beta_1, \beta_2, \dots)$  - аналитическое выражение с неизвестными параметрами  $\beta_1, \beta_2, \dots$

Условиями минимума погрешности  $\Delta$  является следующая система уравнений:

$$\left\{ \begin{array}{l} \frac{\partial \Delta}{\partial \beta_1} = \sum_{j=1}^M \left[ \bar{F}_x(x_j) - F_a(x_j, \beta_1, \beta_2, \dots) \right] \frac{\partial F_a(x_j, \beta_1, \beta_2, \dots)}{\partial \beta_1} = 0 \\ \frac{\partial \Delta}{\partial \beta_2} = \sum_{j=1}^M \left[ \bar{F}_x(x_j) - F_a(x_j, \beta_1, \beta_2, \dots) \right] \frac{\partial F_a(x_j, \beta_1, \beta_2, \dots)}{\partial \beta_2} = 0. \\ \dots \dots \dots \end{array} \right. \quad (1.41)$$

При однопараметрической аппроксимации с использованием метода Ньютона, неизвестный параметр определяется в результате решения следующего уравнения:

$$\beta_{n+1} = \beta_n - \frac{\sum_{j=1}^M \left[ \bar{F}_x(x_j) - F_a(x_j, \beta_n) \right] \frac{\partial F_a(x_j, \beta)}{\partial \beta}}{\sum_{j=1}^M \left\{ \left[ \bar{F}_x(x_j) - F_a(x_j, \beta_n) \right] \frac{\partial^2 F_a(x_j, \beta)}{\partial \beta^2} - \left[ \frac{\partial F_a(x_j, \beta)}{\partial \beta} \right]^2 \right\}_{\beta=\beta_n}}, \quad (1.42)$$

и дальше все расчеты производятся аналогично случаю с плотностями вероятностей (см. раздел 2.4).

Рассмотрим метод хорд применительно к задаче аппроксимации функций распределения вероятностей.

Приведем все формулы, необходимые для работы метода.



$$F(a) = \sum_{j=1}^M \left( \hat{F}_x(x_j) - F_a(x_j, a) \right) \frac{\partial F_a(x_j, a)}{\partial a}, \quad (1.43)$$

$$F(b) = \sum_{j=1}^M \left( \hat{F}_x(x_j) - F_a(x_j, b) \right) \frac{\partial F_a(x_j, b)}{\partial b}. \quad (1.44)$$

Далее вычисляется первая точка  $x_1$  пересечения хорды и оси OX.

$$x_1 = a - \frac{(b-a)F(a)}{F(b) - F(a)}, \quad (1.45)$$

или

$$x_1 = b - \frac{(b-a)F(b)}{F(b) - F(a)}. \quad (1.46)$$

Найденное значение  $x_1$  используется для нахождения значения функции  $F(x_1)$ .

$$F(x_1) = \sum_{j=1}^M \left( \hat{F}_x(x_j) - F_a(x_j, x_1) \right) \frac{\partial F_a(x_j, x_1)}{\partial x_1}. \quad (1.47)$$

Проверяем знаки получившихся функций. Если  $F(a) \cdot F(x_1) \geq 0$ , то левая граница интервала  $a$  переходит в  $x_1$ . В противном случае, в  $x_1$  переходит правая граница  $b$ .

Новые значения  $a$  и  $b$  вновь подставляются в формулы (1.43) и (1.44), и итерационный процесс повторяется до достижения необходимой точности.

Приведем пример расчетов параметра аппроксимирующей модели для закона распределения арксинуса.

В качестве начальных значений  $a$  и  $b$  выберем значения, рассчитанные по методу моментов.

$$a = \sqrt{2\hat{D}_x} - 0.5,$$

$$b = \sqrt{2\hat{D}_x} + 0.5.$$

Рассчитываем значения  $F(a)$  и  $F(b)$  в точках  $a$  и  $b$ :

$$F(a) = \sum_{j=1}^M \left( F_a(x_j) - \left( \frac{1}{2} + \frac{1}{\pi} \arcsin \frac{x_j}{a} \right) \right) \frac{-x_j}{\pi \sqrt{(a^4 - x_j^2 a^2)}},$$

$$F(b) = \sum_{j=1}^M \left( F_a(x_j) - \left( \frac{1}{2} + \frac{1}{\pi} \arcsin \frac{x_j}{b} \right) \right) \frac{-x_j}{\pi \sqrt{(b^4 - x_j^2 b^2)}}.$$

Находим значение  $x_1$  и вычисляем функцию в этой точке:

$$x_1 = a - \frac{(b-a)F(a)}{F(b) - F(a)},$$

$$F(x_1) = \sum_{j=1}^M \left( F_a(x_j) - \left( \frac{1}{2} + \frac{1}{\pi} \arcsin \frac{x_j}{x_1} \right) \right) \frac{-x_j}{\pi \sqrt{(x_1^4 - x_j^2 x_1^2)}}.$$

Результаты определения параметра случайного процесса, распределенного по закону арксинуса, представлены на рис. 1.13.

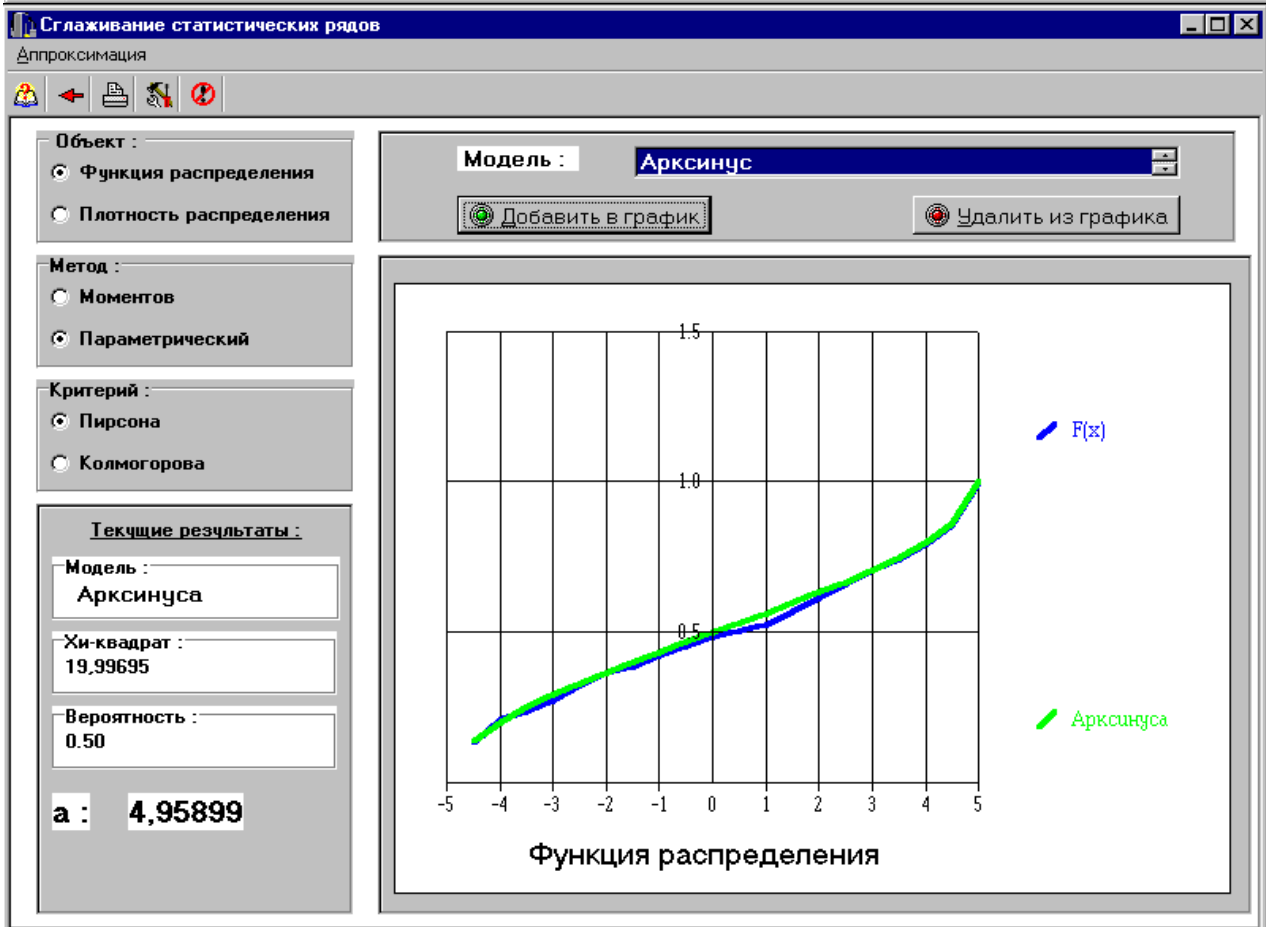
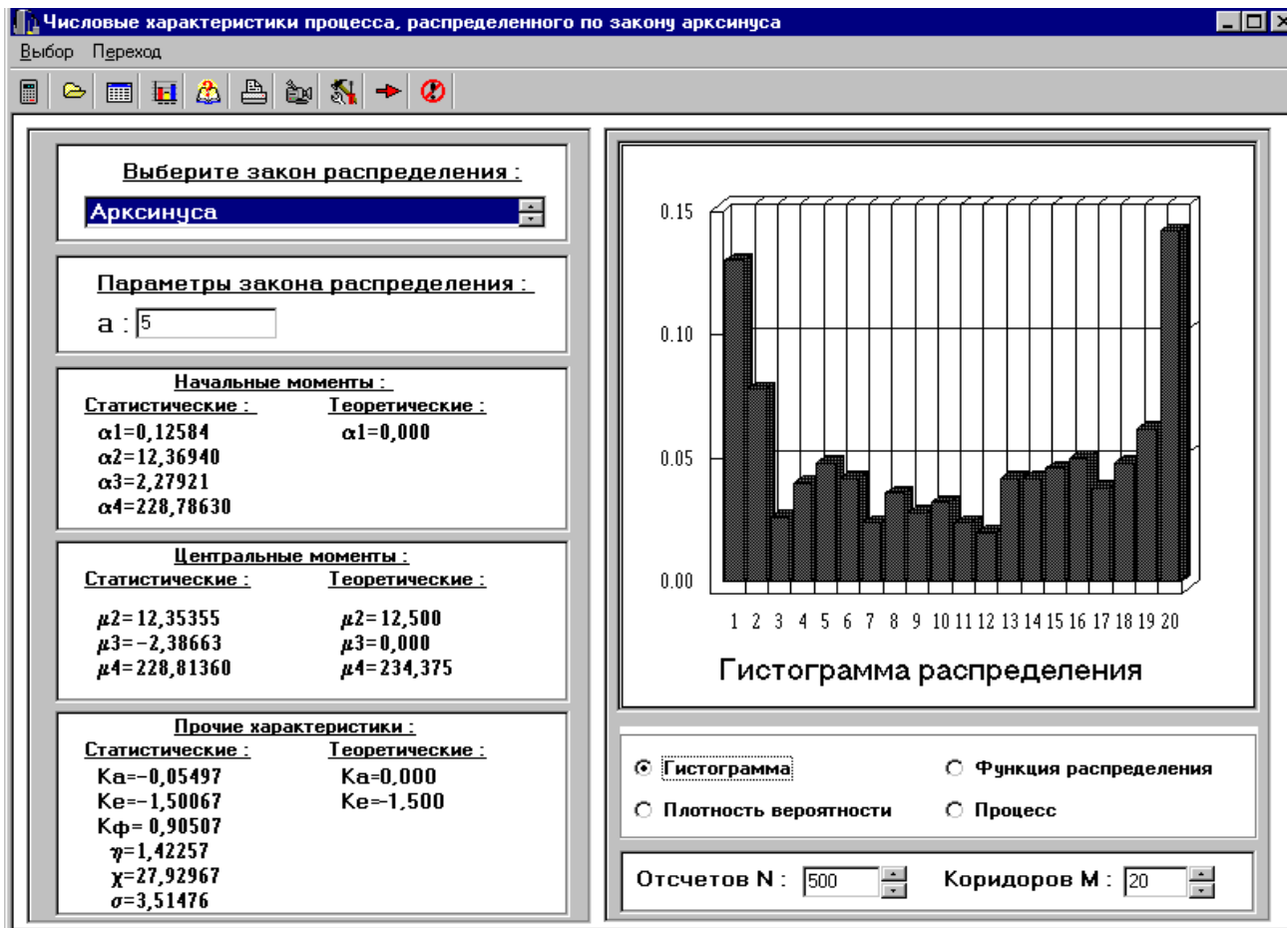


Рисунок 1.13. Результаты определения параметра закона арксинуса

Для нахождения параметров двухпараметрического закона распределения необходимо решить уравнение (1.40) для двумерного случая.

Составим систему из двух уравнений для нахождения неизвестных параметров аппроксимации. Эту систему можно получить, продифференцировав выражение (1.40) по неизвестным параметрам.

$$\begin{cases} F_1 = \sum_{j=1}^M [\hat{F}_x(x_j) - F_a(x_j, \beta_1, \beta_2)] \frac{\partial F_a(x_j, \beta_1, \beta_2)}{\partial \beta_1} = 0 \\ F_2 = \sum_{j=1}^M [\hat{F}_x(x_j) - F_a(x_j, \beta_1, \beta_2)] \frac{\partial F_a(x_j, \beta_1, \beta_2)}{\partial \beta_2} = 0 \end{cases} \quad (1.48)$$

Для решения системы (1.48) воспользуемся приближенным методом Ньютона. Способ нахождения неизвестных параметров аналогичен случаю с плотностями распределения вероятностей (смотрите раздел 1.5) – по формулам (1.34) и (1.35).

Для вычислений необходимо определить частные производные по неизвестным параметрам  $\beta_1, \beta_2$  функций  $F_1$  и  $F_2$ :

$$\frac{\partial F_1}{\partial \beta_1} = \sum_{j=1}^M [\hat{F}_x(x_j) - F_a(x_j, \beta_1, \beta_2)] \frac{\partial^2 F_a(x_j, \beta_1, \beta_2)}{\partial \beta_1^2} - \left[ \frac{\partial F_a(x_j, \beta_1, \beta_2)}{\partial \beta_1} \right]^2, \quad (1.49)$$

$$\frac{\partial F_2}{\partial \beta_2} = \sum_{j=1}^M [\hat{F}_x(x_j) - F_a(x_j, \beta_1, \beta_2)] \frac{\partial^2 F_a(x_j, \beta_1, \beta_2)}{\partial \beta_2^2} - \left[ \frac{\partial F_a(x_j, \beta_1, \beta_2)}{\partial \beta_2} \right]^2, \quad (1.50)$$

$$\frac{\partial F_1}{\partial \beta_2} = \sum_{j=1}^M [\hat{F}_x(x_j) - F_a(x_j, \beta_1, \beta_2)] \frac{\partial^2 F_a(x_j, \beta_1, \beta_2)}{\partial \beta_1 \partial \beta_2} - \frac{\frac{\partial F_a(x_j, \beta_1, \beta_2)}{\partial \beta_1} \frac{\partial F_a(x_j, \beta_1, \beta_2)}{\partial \beta_2}}{\partial \beta_1 \partial \beta_2}, \quad (1.51)$$

$$\frac{\partial F_2}{\partial \beta_1} = \sum_{j=1}^M [\hat{F}_x(x_j) - F_a(x_j, \beta_1, \beta_2)] \frac{\partial^2 F_a(x_j, \beta_1, \beta_2)}{\partial \beta_2 \partial \beta_1} - \frac{\frac{\partial F_a(x_j, \beta_1, \beta_2)}{\partial \beta_2} \frac{\partial F_a(x_j, \beta_1, \beta_2)}{\partial \beta_1}}{\partial \beta_2 \partial \beta_1}. \quad (1.52)$$

Значения неизвестных параметров вычисляются итерационно до достижения заданной точности.

Приведем пример расчета параметров аппроксимационной модели для закона распределения Симпсона: