ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
Государственное образовательное учреждение
высшего профессионального образования
«Южно-Российский государственный университет экономики и сервиса»
(ГОУ ВПО «ЮРГУЭС»)

Марчук В.И., Токарева С.В.

СПОСОБЫ ОБНАРУЖЕНИЯ АНОМАЛЬНЫХ ЗНАЧЕНИЙ ПРИ АНАЛИЗЕ НЕСТАЦИОНАРНЫХ СЛУЧАЙНЫХ ПРОЦЕССОВ

Монография

ШАХТЫ
ЮРГУЭС
2009

УДК 519.251/251.9
ББК 65.050.2
Ц274

Рецензенты:

д.т.н., профессор, зав. кафедры ««Теоретические основы радиотехники» Таганрогского технологического института Южного Федерального университета

Федосов А.П.

д.т.н., профессор, начальник кафедры «Антенные устройства и теоретические основы радиоэлектронных систем» Ростовского военного института ракетных войск

Габриэльян Д.Д.

д.т.н., профессор кафедры «Физика» Южно-российского государственного университета экономики и сервиса

Санников Н.И.

В монографии представлены способы обнаружения аномальных значений при анализе нестационарных случайных процессов, представленных единственной реализацией в условиях ограниченного объема априорной информации и исследована эффективность разработанных способов на различных моделях входного процесса.

Монография рассчитана на инженеров и научных работников, специализирующихся на цифровой обработке сигналов в различных областях науки и техники. Также она может быть полезна для специалистов, занимающихся обработкой результатов измерений временных рядов социально-экономической, медико-биологической, авиационно-космической, геодезической информации и др.

Монографию можно рекомендовать аспирантам и студентам вузов технических и экономических специальностей.

УДК 519.251/251.9
ББК 681.3.06

ISBN 5–8327–0160–7

© ГОУ ВПО «Южно-российский государственный университет экономики и сервиса», 2009
© Марчук В.И., Токарева С.В., 2009

СОДЕРЖАНИЕ

ВВЕДЕНИЕ 6

ГЛАВА 1 Анализ методов обнаружения аномальных значений 9

1.1 Параметрические методы и алгоритмы обнаружения аномальных значений 9

1.2 Непараметрические методы и алгоритмы обнаружения аномальных значений 32

ГЛАВА 2 Способы обнаружения аномальных значений 52

2.1 Математические модели сигналов 52

2.2 Адаптивный способ обнаружения аномальных значений 60

2.3 Структурная схема устройства, реализующего адаптивный способ обнаружения аномальных значений 76

2.4 Способ обнаружения аномальных значений с дискретно-меняющимся уровнем анализа процесса 78

2.5 Структурная схема устройства, реализующего способ обнаружения аномальных значений с дискретно-меняющимся уровнем анализа процесса 95

ГЛАВА 3 Анализ эффективности адаптивного способа обнаружения аномальных значений 98

3.1 Исследование эффективности адаптивного способа обнаружения аномальных значений для нестационарных случайных процессов с аддитивной шумовой составляющей 98

3.2 Исследование эффективности адаптивного способа обнаружения аномальных значений для нестационарных случайных процессов с мультипликативной шумовой составляющей 105

3.3 Сравнительный анализ эффективности адаптивного способа обнаружения аномальных значений для нестационарных случайных процессов с аддитивной и мультипликативной шумовой составляющей 117

3.4 Анализ эффективности адаптивного способа обнаружения аномальных значений для стационарных случайных процессов 120

ГЛАВА 4 Анализ эффективности способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа нестационарного случайного процесса 128

4.1 Исследование эффективности способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа нестационарного случайного процесса с аддитивной шумовой составляющей 128

4.2 Исследование эффективности способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа нестационарного случайного процесса при мультипликативной шумовой составляющей 140

4.3 Сравнительный анализ эффективности способа с обнаружения аномальных значений для нестационарных случайных процессов с аддитивной и мультипликативной шумовой составляющей 152

4.4 Исследование эффективности способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа для стационарных случайных процессов 157

ГЛАВА 5 Исследование эффективности способов обнаружения аномальных значений при обработке реализаций, полученных в результате реальных экспериментов 162

5.1 Сравнительный анализ эффективности адаптивного способа и способа с дискретно-меняющимся уровнем анализа процесса при обнаружении аномальных значений 162

5.2 Анализ эффективности способов обнаружения аномальных значений при выделения полезной составляющей нестационарного случайного процесса 171

5.3 Обработка экспериментальных реализаций 183

ЗАКЛЮЧЕНИЕ 192

Список используемой литературы 193

ВВЕДЕНИЕ

Динамика развития научно-технического прогресса привела к появлению сложных радиотехнических систем. Использование таких систем предъявляет серьезные требования к обработке получаемой измерительной информации и точности принимаемых решений. Эти системы по существу можно отнести к категории систем, основной отличительной особенностью которых является не только их сложность и разновидность элементов, сколько вероятностный, случайный характер их поведения во времени. Этим свойством обладает любая радиотехническая система, так как все её параметры являются реализациями случайных величин.

Большинство реальных сигналов представляют собой реализации нестационарных случайных процессов аддитивно-мультипликативной природы. И их можно рассматривать как выборку нестационарного случайного процесса ограниченного объема, состоящего из полезной и случайной составляющих. Случайная (шумовая) составляющая может включать и аномальные значения. На сегодняшний день в литературе отсутствует строгое определение аномальных значений, так как их статистические характеристики зависят от условий и причин возникновения, которые весьма разнообразны. Аномальными называют значения, резко отличающиеся по величине и статистическим свойствам на фоне основной группы значений реализации процесса. Природа возникновения и источники аномальных значений различны, это может быть импульсная помеха, кратковременные повышения уровня шумов на входах приемников, сбой в работе регистрирующей аппаратуры, отказ оборудования, кратковременное внешнее воздействие на измерительный элемент, «залипание» старшего разряда цифрового счетчика, атмосферные воздействия при передаче радиосигналов, индустриальные помехи и т.д.

Аномальные значения, возникающие в каналах передачи и обработки сигналов, даже при небольшой частоте их появления, вносят большие погрешности при дальнейшей оценке полезной составляющей процесса. Поэтому первым этапом обработки нестационарных случайных процессов является проверка его значений на достоверность, т.е. обнаружение и последующее устранение аномальных значений. В общем случае обнаруженные аномальные значения могут восстанавливаться или не восстанавливаться. Восстановление значений в некоторый момент времени может осуществляться различными известными методами (экстраполяция, интерполяция, аппроксимация).

До недавнего времени на практике для обнаружения аномальных значений широко применялись ручные способы, основанные на визуальном просмотре зарегистрированных реализаций процессов и сравнение их с контрольными процессами известной формы. Помимо субъективизма в критериях обнаружения аномальных значений, основанных главным образом на опыте и интуиции экспериментатора, подобные способы не допускают автоматизации процедур анализа нестационарных случайных процессов. Для преодоления отмеченных недостатков в работах Микешина Н.Г., Орлова А.И., Лемешко Б.Ю., Куллдорфа Г., Савченко В.В., Демина Н.С., Михайлюка В.В., Фомина А.Ф., Паламарчука С.Н., Жадан Л.И., Большева Л. Н., Переверткина С.М., Новоселова О.Н., Марчука В.И. и ряда других предлагается использовать теорию статистических решений, которая позволяет формализовать алгоритмы проверок и выбрать критерий обнаружения аномальных значений. Возможно применение как параметрических, так и непараметрических методов теории статистических решений. В первом случае необходимо располагать априорными сведениями, как о функции полезной составляющей, так и о законе распределения шумовой

составляющей, а так же и о его параметрах (математическом ожидании, дисперсии, корреляционной функции).

Использование непараметрических методов обработки требует значительно меньше априорной информации, но их эффективность определяется параметрами обработки, которые в свою очередь зависят от функции полезной составляющей и закона распределения шумовой составляющей.

ГЛАВА 1 Анализ методов обнаружения аномальных значений

При регистрации, обработке и обмене данными в современных измерительно-вычислительных и информационных системах потоки сигналов искажены действием помех (шумов) природа возникновения которых различна и зачастую носит случайный характер. Для решения задачи выделения полезной составляющей сигнала применяются различные классические процедуры фильтрации, результаты которых зависят от наличия в исследуемом процессе аномальных значений.

1.1 Параметрические методы и алгоритмы обнаружения аномальных значений

По оценкам различных авторов [1 – 8], входные процессы могут содержать до 5-10% аномальных значений. Однако, даже при меньшем их количестве опасность представляет каждое такое значение, поскольку его проникновение в итоговый результат обработки может привести к неправильным оценкам и выводам по отношения к исследуемому процессу. Следовательно, одной из задач первичной обработки процессов является решение вопросов обнаружения и устранения аномальных значений. Это особенно важно, когда исследуемый входной процесс представлен единственной реализацией.

Для решения этой задачи на практике широко используется теория статистических решений с применением параметрических и непараметрических методов и алгоритмов. Если имеются априорные сведения о плотности вероятности распределения значений процесса и его параметрах (математическом ожидании, дисперсии или корреляционной функции), то для решения задачи обнаружения и устранения аномальных значений используются параметрические методы.

Одна из первых работ по исследованию аномальных значений принадлежит астроному Бесселю. Согласно его теории предполагалось использование всех данных, но с различными весами [9]. Первые критерии для обнаружения, аномальных значений, основанные на вероятностных моделях, предложены в монографии Шора и Бендерского [10] и рассматриваются для случая, когда известны параметры плотности распределения вероятности исследуемого процесса.

Если стационарный случайный процесс $Y(t)$ можно представить в виде вариационного ряда $Y_1, Y_2, ...Y_{\mathrm{N}}$, где Y_1 - наименьшее значение; Y_{N} – наибольшее значение; Y_k – гауссовская распределенная случайная величина, с некоррелированными отсчетами, $Y_k \sim \mathrm{N}(m_Y, \sigma_Y)$, $k = \overline{1, N}$, то гипотезу об аномальности крайних значений при известной величине дисперсии ${\sigma_Y}^2$ можно проверить с помощью U - статистик [11]:

$$U = \frac{Y_k - \overline{m}_Y}{\sigma_Y}, \qquad \text{где } \overline{m}_Y = \frac{1}{\mathrm{N}} \sum_{k=1}^{\mathrm{N}} Y_k, \tag{1.1}$$

представляющих собой случайные величины, распределенные по гауссовскому закон плотности распределения вероятностиу $U_k \sim N^*(0;1)$.

Таким образом, согласно материалам [11], значение процесс считается аномальным, если значение статистики (1.1) превышает при выбранном уровне значимости квантиль стандартного гауссовского распределения: $U_k > U_\beta$, где U_β – квантиль стандартного гауссовского распределения; $\alpha = 1 - \beta$ - уровень значимости. В случае, когда известны оба параметра гауссовского распределения случайной величины, т.е. математическое ожидание m_Y и дисперсия σ_Y^2, тогда для проверки принадлежности аномальных значений к исходной выборке процесса применяют статистику χ^2 и модифицированную статистику Фишера –

Снедекора [12]: $\tilde{F} = \frac{(\sigma_Y{}^*)^2}{\sigma_Y{}^2}$, где $\sigma_Y{}^{*2}$ – оценка дисперсии без учета аномального значения.

Однако, априорная информация о параметрах m_Y и σ_Y редко встречается на практике. Если известен один параметр σ_Y, то вместо второго параметра m_Y можно использовать его оценку $\overline{m}_Y$. В этом случае можно считать аномальным крайний элемент выборки, если выполняется неравенство:

$\frac{Y_i - \overline{m}_Y}{\sigma_Y} > U_\beta \sqrt{\frac{\mathrm{N}-1}{\mathrm{N}}}$, где $\overline{m}_Y$ – оценка математического ожидания.

При отсутствии априорной информации о параметрах m_Y и σ_Y, одновременная замена этих параметров их оценками $\overline{m}_Y$ и $\overline{\sigma}_Y$ заставляет отказаться от использования квантилей гауссовского распределения. Для этого случая, в работе Томсона [13] в качестве статистики для обнаружения аномальных значений в выборке стационарного случайного процесса, применяется стандартизированное экстремальное отклонение:

$$\tau_k = \frac{Y_k - \overline{m}_Y}{\overline{\sigma}_Y},\ k = \overline{1,\mathrm{N}},\quad \text{где} \quad \overline{\sigma}_Y = \sqrt{\frac{1}{\mathrm{N}-1}\sum_{k=1}^{\mathrm{N}}(Y_k - \overline{m}_Y)^2}\ . \qquad (1.2)$$

Величина τ_k, которая выражает отклонение выборочного значения Y_k от выборочного среднего значения $\overline{m}_Y$, отнесенное к оценке среднеквадратического отклонения $\overline{\sigma}_Y$, имеет специальное распределение, которое зависит только от объема выборки N:

$$f(\tau) = \frac{1}{\sqrt{(\mathrm{N}-1)^\pi}} \frac{\Gamma(\frac{\mathrm{N}-1}{2})}{\Gamma(\frac{\mathrm{N}-2}{2})} (1 - \frac{\tau^2}{\mathrm{N}-1})^{\frac{\mathrm{N}-4}{2}}\ .$$

При $\mathrm{N} \to \infty$ статистика (1.2) стремится к U – статистики гауссовского распределения.

Томсон в своей работе [13] предлагает исключать из выборки значения, для которых τ_i превышает по модулю некоторую константу K. Априорно задается процент отбрасываемых значений Φ и тогда вероятность исключения k-того значения определяется как Φ / N; отсюда находится критическое значение K. Однако этот критерий является неполным, поскольку заранее предполагается, что исключается только одно аномальное значение. И как представлено в работе Пирсона [14], критерий Томсона обладает маскирующим эффектом, который состоит в том, что при наличии в выборке более одного аномального значения критерий их не обнаруживает, особенно когда выборка содержит меньше 30-40 отсчетов. Таким образом, критерий Томсона эффективен в случаях, когда есть только одно аномальное значение.

В тоже время, если существует одно аномальное значение, то оно является крайним членом вариационного ряда и согласно критерию. Смирного [15] выборку можно проверять при помощи статистики:

$$R = \frac{\max_k (Y_k - \overline{m}_Y)}{\overline{\sigma}_Y{}^2}.$$

Распределение этой статистики изучено в работе Граббса [16]. Для обнаружения одного аномального значения им использована статистика:

$$\frac{\overline{\sigma}_{(Y)}{}^2}{\overline{\sigma}_Y{}^2} = \frac{\sum_{k=1}^{N-1}(Y_{(k)} - \sum_{k=1}^{N-1}\frac{Y_{(k)}}{N-1})^2}{\sum_{k=1}^{N}(Y_k - \sum_{k=1}^{N}\frac{Y_k}{N})^2} = 1 - \frac{\max_k (Y_k - \overline{m}_Y)}{\overline{\sigma}_Y{}^2},$$

где $Y_{(k)}$ – k-тый член вариационного ряда. Если отношение $\overline{\sigma}_{(Y)}{}^2 / \overline{\sigma}_Y{}^2$ меньше табличного, при этом проверяется максимальное значение, и оно признается аномальным. В работе Брофмана М.Я и Римена В.Х. [18] исследуется эффективность критерия Граббса, если для выборок составляет N=20 и два крайних члена являются аномальными.

Вариационный критерий Диксона [18] удобный и достаточно эффективный. При его использовании полученные значения стационарного случайного процесса записываются в вариационный возрастающий ряд $Y_1, Y_2, ..., Y_N$ $(Y_1 < Y_2 < ..... < Y_N)$. Критерий Диксона определяется как $K_Д = (Y_N - Y_{N-1})/(Y_N - Y_1)$. Критическая область для этого критерия $P(K_Д > Z_q) = q$, где значения Z_q определяются в соответствии с заданным уровнем значимости.

В случае двусторонней проблемы, Бендерский А.М. в работе [19] показал, что в случае, когда неизвестно какое из крайних значений вариационного ряда ошибочно, критерии Граббса и Диксона принимают вид:

$$\min\left(\frac{S_1^2}{S^2}; \frac{S_N^2}{S^2}\right) \text{ и } \max\left(\frac{Y_N - Y_{N-1}}{Y_N - Y_1}; \frac{Y_2 - Y_1}{Y_N - Y_1}\right).$$

Критерии Шовене, Граббса, Диксона, Смирного и Томсона основаны на статистиках, которые характеризуют отклонение одного или нескольких элементов от основной группы значений или от соседних членов вариационного ряда. Все эти критерии инвариантны относительно прибавления постоянной ко всем выборочным значениям, умножения на постоянную величину и перестановку индексов.

Среди инвариантных критериев существуют критерии, предложенные Фергюсоном, которые обладают сильным оптимальным свойством [20]. Для односторонней гипотезы критерий основан на коэффициенте асимметрии:

$$\sqrt{b_1} = \frac{\frac{1}{N}\sum_{k=1}^{N}(Y_k - \overline{m}_Y)^3}{\overline{\sigma}_Y{}^2}.$$

Максимальное выборочное значение признается ошибочным, если $b_1 \geq K_1$; это наилучший локальный критерий среди инвариантных

критериев. В случае двусторонней проблемы используется коэффициент эксцесса:

$$b_2 = \frac{1}{N}\sum_{k=1}^{N}\frac{(Y_k - \overline{m}_Y)^4}{\overline{\sigma}_Y{}^4}.$$

Все выборочные значения отвергаются, если $b_2 > K_2$; это наилучший локальный несмещенный критерий.

Оптимальность критериев Граббса и Фергюсона рассмотрена в работе [20]. Первый критерий Фергюсона оптимален, когда есть только одно аномальное значение, и не зависит от величины сдвига, в то время как второй оптимален лишь для малых сдвигов и не зависит от числа аномальных значений.

С помощью параметрических статистик Стьюдента или Румшинского, которые исследованы в работе Дэйвида Г. [21] проверяется гипотеза о достоверности крайних отсчетов некоррелированного стационарного случайного процесса $Y(t)$. Если выборку из N отсчетов расположить в вариационный ряд $Y_1, Y_{2,}.....,Y_{\mathrm{N}}$, где $Y_{\min} = Y_1$ и $Y_{\max} = Y_{\mathrm{N}}$, тогда:

$$t = \frac{Y_{\mathrm{N}} - \overline{m}_Y}{\overline{\sigma}_Y}, \quad t = \frac{\overline{m}_Y - Y_{\mathrm{N}}}{\overline{\sigma}_Y}, \qquad \mathrm{Z} = \frac{Y_{\mathrm{N}} - \overline{m}_Y{}^*}{\overline{\sigma}_Y}, \quad \mathrm{Z} = \frac{\overline{m}_Y{}^* - Y_{\mathrm{N}}}{\overline{\sigma}_Y},$$

где $\overline{m}_Y{}^* = \frac{1}{\mathrm{N}-1}\sum_{k=1}^{\mathrm{N}-j} Y_k$, j -число аномальных отсчетов.

Если проверку на аномальность необходимо применить к нескольким элементам исследуемого процесса, тогда для каждого сомнительного элемента вычисляется максимальное значение относительного отклонения t. Далее эти значения располагаются в порядке возрастания, и выбирается элемент, которому соответствует наименьшее значение t. Остальные элементы временно исключаются из выборки и заново определяются значения $\overline{m}_Y$ и $\overline{\sigma}_Y$ при уменьшенном

объёме выборки и соответствующее значение статистики t для исследуемого значения и так далее. Однако, данная процедура применима для реализаций процессов большого объема, когда исключение из выборки ряда случайных значений не приводит к изменению вероятностных свойств исследуемого стационарного процесса.

Критерий Шарлье используется, если объем реализации стационарного случайного процесса $N > 20$. Тогда по теореме Бернулле [22] число значений, превышающих по абсолютной величине среднее арифметическое значение, на величину $K_{Ш}\overline{\sigma}_Y$, будет $N[1-\Phi(K_ш)]$, где $\Phi(K_ш)$ - значение нормированной функции Лапласа для $Y = K_ш$. Если сомнительным в реализации процесса является одно значение, то $N[1-\Phi(K_ш)]=1$. Пользуясь критерием Шарля, исключается значение, для которого выполняется условие $|Y_k - \overline{m}_Y| > K_ш\overline{\sigma}_Y$.

Если объем реализации стационарного случайного процесса $N < 20$, то используется критерий Романовского. При этом вычисляется отношение $\left|(Y-\overline{m}_Y)/\overline{\sigma}_Y\right| = \beta$ и сравнивается с критерием β_T. Если $\beta \geq \beta_T$, то значение Y_i считается аномальным. Этот критерий применим только для одиночных аномальных значений, что и является его недостатком.

Недостатком методов, основанных на критериях Романовского, Шарля в том, что оценка среднеквадратического отклонения процесса при задании порога принятия решения проводится без учета экстремальных значений (визуальный контроль) и эти методы применимы для ограниченного объема выборки процесса.

Для обнаружения аномальных значений могут быть так же использованы статистики: нормированный размах $\omega = \dfrac{Y_{\mathrm{N}} - Y_1}{\sigma_Y}$;

стъюдентизированный размах $\nu = \frac{Y_{\mathrm{N}} - Y_1}{S_Y}$. Данные статистики мало используют в практике обработки исследуемых процессов, поскольку они имеют достаточно сложный закон плотности распределения вероятности.

Рассмотренные статистики, применимы для обнаружения аномальных значений, только если исследуемый стационарный случайный процесс имеет гауссовский закон плотности распределения вероятности, а его выборочные значения – статистически независимыми. Обнаружение аномальных значений для выборок произвольного закона плотности распределения вероятности можно провести в соответствии с неравенством Чебышева [21], согласно которому любое выборочное значение Y_k отклоняется от математического ожидания m_Y не более чем на величину $\frac{\sigma_Y{}^2}{\sqrt{1-\beta}}$, т. е.

$$\left| \frac{Y_k - m_Y}{\sigma_Y{}^2} \right| \le \frac{1}{\sqrt{1-\beta}},$$

где m_Y, $\sigma_Y{}^2$ - известные значения математического ожидания и дисперсии процесса соответственно, β - доверительная вероятность.

Замена значений математического ожидания и дисперсии на их оценки, что чаще всего встречается на практике, позволяет использовать неравенство Чебышева только для выборок большого объема. Это и является недостатком использования данной статистики.

Математический критерий Шовенэ для обнаружения аномальных значений предложен в работах [23 – 25]. Изучается эмпирическое распределение, формально построенное по N взаимно-независимым случайным величинам с гауссовским законом плотности распределения вероятности. Было установлено, что любое значение функции такого формально построенного эмпирического распределения в окрестностях

концевых точек интервала изменения независимой переменной (т.е. «на хвостах») при $N \to \infty$ и надлежащей нормировке имеет предельный пуассоновский закон плотности распределения вероятности, совпадающий с предельным распределением соответствующего значения нормированной функции эмпирического распределения, построенного в том случае, когда исходный процесс имеет гауссовский закон плотности распределения вероятности. Однако, как отмечено в статье Убайдулловой М. [26], критерий Шовенэ не позволяет выявлять аномальные значения, он лишь «сигнализирует» об их наличии.

В статье [24] Пагуровой В.И. исследуется поведение статистики критерия Шовенэ обнаружения аномальных значений в стационарном случайном процессе с гауссовским законом плотности распределения вероятности и рассматривается три случая комбинации параметров сдвига и масштаба. Этим же автором в работе [27] исследовались асимптотические свойства статистики критерия Шовенэ для равномерного закона плотности распределения вероятности на «хвостах» в предположении, что количество аномальных значений неизвестно, но оно не превосходит заранее заданного фиксированного числа. Позднее Пагуровой В.И. [28], с Чижиковой И.А. в соавторстве, рассматривается задача обнаружения «резко выделяющихся наблюдений» в многомерной выборке в присутствии «мешающих параметров» и строятся асимптотические критерии обнаружения аномальных значений. В качестве оценок неизвестных параметров предполагается использовать робастные оценки, которые сохраняют высокую эффективность и уменьшают «маскирующий эффект» в присутствии аномальных значений.

Работы [23 – 28] носят теоретический характер и не дают практического применения критериев по обнаружению аномальных значений.

Метод Дарлинга можно использовать для обнаружения аномальных значений при любом исходном законе плотности распределении вероятности, преобразовав его к равномерному распределению на интервале $[0,1]$. Это преобразование основано на факте, что интегральная функция распределения любой непрерывной случайной величины равномерно распределена в интервале (0.1), т. е. для любой случайной величины Y с плотностью распределения $f(Y)$ случайная величина $y = F(Y) = \int_{-\infty}^{Y} f(\xi)d\xi$ имеет равномерный закон плотности распределения вероятности в интервале $[0,1]$. В этом случае распределение статистики Дарлинга имеет вид: $Z_{\mathrm{N}} = \sum_{k=1}^{\mathrm{N}} \frac{Y_k}{Y_{\mathrm{N}}}$, где N значений процесса Y_k представляют собой одинаково распределенные независимые положительные случайные величины с известной функцией плотности распределения вероятности. Данный метод применим, если закон плотности распределения вероятности исходного процесса известен [21].

В работе Цветкова Э.И. [29] рассмотрен метод обнаружения аномальных значений с помощью задания порога принятия решения об аномальности. Значение считается аномальным, если выполняется условие: $\left|\overline{m}_Y^{\,*} - Y_k\right| \geq 3 \cdot \overline{\sigma}_Y^{\,*}$, где в качестве $\overline{m}_Y^{\,*}$ и $\overline{\sigma}_Y^{\,*}$ оценки математического ожидания и среднеквадратического отклонения, полученные без учета аномальных значений Y_k, либо априорно задаваемые. Этот метод эффективен при числе отсчетов $\mathrm{N} \geq 20...50$ с доверительной вероятностью $\rho = 0{,}003$. Данный метод, как предложено в работах Новицкого П.В., Зографа И.А., Лабунца В.С. [30, 31], рекомендуется применять с учетом объема исследуемой выборки: при $6 < \mathrm{N} \leq 100$ пороговое значение равно 4σ; при $100 < \mathrm{N} \leq 1000$ - 4.5σ; при $1000 < \mathrm{N} \leq 10000$ -5σ. Такое правило применимо, только когда

стационарный случайный процесс имеет гауссовский закон плотности распределения вероятности.

В общем случае порог принятия решения о аномальности значения можно определить из выражения:

$$\varepsilon = \mathrm{A}\overline{\sigma}_Y \qquad (1.3)$$

Величина порога (1.3), зависит не только от объема N, но и от закона плотности распределения вероятности стационарного случайного процесса. В работе Новицкого П.В., Зографа И.А [30] приводится выражение для приближенного расчета коэффициента А при уровне значимости $\rho < 1/(\mathrm{N}+1)$:

$$\mathrm{A} = 1{,}55 + 0{,}8\sqrt{\mu - 1} * \lg(N/10),$$

где μ - эксцесс распределения. Выражения применимы для следующих законов плотности распределения вероятности стационарного случайного процесса:

- кругловершинных двухмодальных распределений с $\mu = 1{,}5....3$, являющихся композицией дискретного двузначного и гауссовского распределений;
- островершинных двухмодальных распределений с $\mu = 1{,}5....6$, являющихся композицией дискретного двухзначного распределения Лапласа;
- композиций равномерного и экспоненциальных с показателем степени $\alpha = 1/2$ при $\mu = 1{,}8....6$;
- экспоненциальных распределений с $\mu = 1{,}8....6$.

Величина порога, определяемого выражением (1.3) зависит так же от погрешности оценки среднеквадратического отклонения $\overline{\sigma}_Y$ стационарного случайного процесса.

В работе Чернышева С.Л. [32] исследуется поток событий, наблюдаемых с ошибками. Автор предлагает каждую реализацию неотрицательной случайной величины Y, заданной распределением $F(Y)$,определять в результате ее поэтапных сравнений с детерминированными значениями (мерами) $k\tau$, $k = 1,2,...,N$ и принятия одного из четырех решений: истинное совпадение, соответствующее принятию решения $Y \le k\tau$, при $Y \le k\tau$; истинное несовпадение - решение $Y > k\tau$ при $Y > k\tau$; ложное несовпадение - решение $Y > k\tau$ при $Y \le k\tau$; ложное совпадение - решение $Y \le k\tau$, при $Y > k\tau$. Для всех k вероятность ложного совпадения и ложного несовпадения принимаются одинаковыми и равными α и β. Процесс наблюдения, таким образом, представляется в виде дискретного процесса случайного блуждания с четырьмя экранами, соответствующим указанным состояниям (принимаемым решениям). Обозначая как $P(0)$ вероятность нулевого значения величины Y, определяются начальные вероятности нахождения Y_k в каждом из четырех состояний на нулевом шаге сравнения:

$$P1(0;\alpha,\beta) = P(0)(1-\beta)$$
$$P2(0;\alpha,\beta) = [1-P(0)](1-\alpha)$$
$$P3(0;\alpha,\beta) = P(0)\beta$$
$$P4(0;\alpha,\beta) = [1-P(0)](1-\alpha)$$

После построения матрицы вероятностей P и некоторых математических выкладок автор делает вывод, что рассмотренная модель наблюдения с ошибками приводит к преобразованию распределения $P(n;0;0) \cong P(n)$ к распределению:

$$P(k;\alpha,\beta) = (1-\beta)\beta^k \sum_{n=0}^{k} \left(\frac{1-\alpha}{\beta}\right)^n P(n) + \alpha(1-\alpha)^k \left(1 - \sum_{n=0}^{k} P(n)\right)$$

посредством оператора ошибок $M_{\alpha,\beta}$:

$$P(k;\alpha,\beta) = M_{\alpha,\beta}[P(n;0,0)].$$

В статье Чернышева С.Л. [32] исследована модель потока событий, наблюдаемых с ошибками, в случае стационарного пуассоновского потока. Анализ полученных результатов показал, что оценки, и экспериментальные зависимости наблюдаемых величин, от обобщенных параметров, указывают на существенное влияние этих параметров, на смещение оценки максимального правдоподобия параметра потока событий [32].

В работах Савченко В.В. и Ширяева А.Н. [33, 34], момент появления аномальных значений квалифицируется как момент разладки случайного процесса. При этом t_s - момент появления аномального значения, а t_τ - момент обнаружения разладки процесса. Тогда задача состоит в следующем: найти t_τ из условия минимума среднего времени задержки в обнаружении момента появления аномального значения, т.е. из условия

$$M\left\{t_\tau - t_s \middle| t_\tau \geq t_s \right\} \to \min$$

при ограничении на вероятность ложной тревоги: $P\{t_\tau \leq t_s\} \leq \alpha$.

Данный алгоритм сильно чувствителен к значениям априорной вероятности наличия аномалий P^0, а именно с ростом P^0 увеличивается частота ложных тревог и с уменьшением P^0 она уменьшается. Это свидетельствует о том, что алгоритм хорошо реагирует на реализацию маловероятной априори ситуации. Эффективность алгоритма также зависит от интенсивности аномальных значений, которые должны превышать интенсивность аддитивного шума более чем в 3 раза, а анализируемый процесс должен быть стационарен в широком смысле.

К числу методов, позволяющих учесть корреляционную зависимость между отсчетами, относятся методы, использующие вычисление текущих разностей между соседними ординатами и сравнение их с выбранным порогом (Фомин А.Ф., Новоселов О.Н., Плющев А.В. [35]):

$$\Delta Y_k \geq \Delta_{\beta/2}, \qquad \text{где } \Delta Y_k = Y_k - Y_{k-1},\ k = \overline{1,N}. \tag{1.4}$$

При выполнении условия (1.4) проверяемое значение процесса признается аномальным. Случайная величина ΔY_k выбрана из одномерного гауссовского закона плотности распределения вероятности: $\Delta Y_k \sim \mathrm{N}(0,\sigma_\Delta{}^2)$, где $\sigma_\Delta{}^2$ дисперсия разностей процесса ΔY_k. Причем дисперсия разностей процесса σ_Δ^2 связана с дисперсией исходного процесса $\sigma_Y{}^2$ и его нормированной корреляционной функцией $R(\tau)$ следующим соотношением [35]:

$$\sigma_\Delta^2 = \sigma_Y{}^2 \sum_{r=0}^{k} \sum_{j=0}^{k} (-1)^{r-j} \binom{k}{r}\binom{k}{j} \rho_{r-j}, \quad \text{где } i \text{ - порядок разности}$$

$$\Delta Y_k = \sum_{j=0}^{i} (-1)^j \binom{i}{j} Y_{k+i-j}.$$

Для первой и второй разности можно записать:

$$\sigma_{\Delta 1}^2 = 2\sigma_Y^2(1-\rho_1); \quad \sigma_{\Delta 1,2}^2 = 2\sigma_Y^2(3-4\rho_1+\rho_2) \tag{1.5}$$

Анализ выражения (1.5) показывает, что дисперсия возрастает с увеличением порядка разности и уменьшением корреляции между соседними отсчетами. Поэтому рекомендуется применять разности не выше первого порядка [35].

В работе Полякова В.А. и Толпарева Р.Г. [36] предложен способ повышения надежности обнаружения полезной составляющей процесса за счет устранения одиночных аномальных значений из шумовой составляющей. Устранение одиночных аномальных значений производится путем выбора решающего порога с помощью аппроксимации неизвестного распределения шумовой составляющей только в области больших положительных значений аргумента (правого «хвоста» плотности распределения). В работе [36] утверждается о неизвестности закона плотности распределения вероятности шумовой

составляющей процесса, однако, далее уточняется, что распределение относится к экспоненциальному типу. Процедура обнаружения состоит в аппроксимации «хвостов» неизвестных функций распределением $F(x)$ экспоненциального типа:

$$F(x) = 1 - \frac{1}{N} \cdot \exp\left(\frac{-\alpha_1 (x^\tau - u_1^\tau)}{\tau \cdot u_1^{\tau-1}} \right),$$

где α_1 - соответственно экстремальная функция интенсивности;

u_1 - характеристическое наибольшее значение;

τ - дополнительный параметр, определяемый из опытных данных;

N - объем реализации.

Для оценки указанных величин, соответствующих неизвестному распределению шумовой составляющей процесса, согласно методу моментов, выполняются следующие операции: с помощью опорного шумового канала, m_1 раз производится выборка из n_1 независимых членов (общий объем $N = m_1 n_1$). В каждой из m_1 частных выборок выделяются максимальные по величине члены Y_j, где $j = 1,...,m_1$. Вычисляются среднее значение максимальных членов $\overline{m}_1$ и их дисперсия $\overline{\sigma}_1^{\,2}$, на основании которых определяются оценки [36]:

$$\overline{\alpha}_1 = \frac{\pi}{\sqrt{6\overline{\sigma}_1^{\,2}}},\ \overline{u}_1 = \overline{m}_1 - \frac{C}{\overline{\alpha}_1},\ \text{где } C = 0.5772 \text{- постоянная Эйлера.}$$

Полная выборка из N членов разбивается на m_2 выборок по n_2 членов ($N = m_2 n_2$, $m_2 > m_1$). В каждой из m_2 выборок выделяются максимальные по величине члены Y_i, где $k = 1,...,m_2$ и определяются $\overline{m}_2, \overline{\sigma}_2^{\,2}, \overline{\alpha}_2, \overline{u}_2$ соответствующие новому разбиению полной исходной выборки. Согласно предлагаемой в работе процедуре обнаружения определяется параметр:

$$\tau = \frac{\overline{\alpha}_1 \overline{u}_1 - \overline{\alpha}_2 \overline{u}_2}{\ln n_1 - \ln n_2}.$$

По ошибке вероятности первого рода P_0, которая задается априорно для выполнения предлагаемой процедуры и найденным значениям параметров $\overline{\alpha}_1$ и $\overline{u}_1$, определяется требуемое пороговое значение:

$$\Pi = \left(\overline{u}_1^{\tau} - \tau \cdot \overline{u}_1^{\tau-1} \cdot \ln \frac{n_1 P_0}{\overline{\alpha}_1} \right)^{\frac{1}{\tau}}.$$

Приведенная процедура, по утверждению авторов Поляковым В.А. и Толпаревым Р.Г. [36], «обладает высокой сходимостью, то есть, экономна в требуемом объеме выборки» и «обеспечивает достоверные пороговые значения даже при малых P_0, вплоть до $P_0 = 10^{-7}$, чего трудно добиться другим путем».

Авторы Поляков В.А. и Толпарев Р.Г. там же [36] отмечают существенный недостаток данной процедуры обнаружения, а именно, что она «слабо устойчива к аномальным шумовым выбросам, которые в силу импульсных помех различного происхождения встречаются достаточно часто». Присутствие аномальных выбросов делает выборку неоднородной, а это приводит к ошибкам при определении параметров распределения шумовой составляющей процесса. В рассматриваемой процедуре сложность состоит в том, что определение параметров процесса и порога основано на отборе только максимальных членов частных выборок, в которые аномальные попадут автоматически, поскольку в любой частной выборке они будут максимальными.

И, кроме отмеченных в работе недостатков, предлагаемый метод обладает следующими ограничениями: 1) метод обнаруживает только одиночные аномальные значения, в случае, если их величина не превышает в два раза её среднее $\overline{m}_Y$. Это означает, что аномальные

одиночные значения, а тем более групповые, имеющие значительную величину, данной процедурой обнаружены не будут; 2) определение порогового значения Π связано с априорной установкой величины вероятности ошибки первого рода P_0, что может не всегда соответствовать его истинному значению. А так же недостатком предлагаемой методики является зависимость вероятности ошибки первого рода P_0 от числа наблюдений N.

Егорова Н.Ю. и Фабер В.Е. в своей работе [37], ссылаясь и обобщая ранее представленные работы Кузьмина С.З. [38, 39], Фомина А.Ф., Новоселова О.Н., Плющева А.В [35], Репина В.Г., Тартаковского Г.П. [40], предлагают алгоритм фильтрации при негауссовской модели ошибок, т.е. с учетом наличия аномальных значений. Предлагается адаптивный байесовский подход к решению задачи фильтрации, который состоит из двух этапов: «оценке методом максимального правдоподобия значений номеров интервалов неоднозначного отсчета и решении задачи фильтрации в предположении, что интервалы неоднозначности известны». При этом предпочтение отдается квазиоптимальному алгоритму, утверждая что, оптимальное байесовское решение задачи фильтрации при наличии аномальных значений в процессе практически нереализуемо. В то же время отмечается, что трудоемкость вычислений предлагаемого квазиоптимального алгоритма в два раза выше, чем при использовании фильтра Кальмана [37].

В работе [41] Катиковым В.М. решается задача совместной оптимизации параметров обнаружителей сигнала и контроле достоверности значений в радиотехнических измерительных системах. За аномальные значения принимаются «ложные измерения», связанные с ложными захватами сигналов и необнаруженными срывами слежения, а также отсутствие значений, обусловленное пропусками полезного сигнала.

В постановке задачи автор вводит функции ошибки первого рода $F_n, F_к$ и вероятности правильного обнаружения $D_n, D_к$, которые характеризуют качество алгоритмов обнаружения сигналов в устройствах поиска и контроля, которые являются функциями некоторых совокупностей параметров, обозначаемых векторами $\mu_n, \mu_к$ на множестве $M = M_n \oplus M_к$. К наиболее важным элементам множества М автор [41] относит пороговые значения в обнаружителях сигнала при поиске и контроле полезного сигнала. Вероятности аномальных ошибок $a_n, a_л$ являются негативными характеристиками непрерывности и достоверности измерительного процесса и зависят от характеристик $F_n, D_n, F_к, D_к$. Функциональные зависимости $a_n(F_n, D_n, F_к, D_к)$ и $a_л(F_n, D_n, F_к, D_к)$ определяются путем анализа взаимодействия составных частей измерительной системы на основании методики ранее разработанной автором [41]. Отношение наблюдателя измерительного процесса к непрерывности и достоверности характеризуется функцией средних потерь вида:

$$Q = C_n a_n(F_n, D_n, F_к, D_к) + C_л a_л(F_n, D_n, F_к, D_к),$$

где коэффициенты C_n и $C_л$ отражают вес двух видов аномальных ошибок. Задача оптимизации состоит в минимизации целевой функции Q при известных зависимостях $F_n = f_{n0}(\mu_n)$, $D_n = f_{n1}(\mu_n, \alpha)$, $F_к = f_{к0}(\mu_к)$, $D_к = f_{к1}(\mu_к, \alpha)$, определяющих статистические характеристики обнаружителей на множестве параметров $M = M_n \oplus M_к$ (параметр α характеризует отношение сигнал/помеха и не является оптимизируемым). Однако в работе [41] Катиков В.М. приходит к выводу, что «оптимизация взаимодействия обнаружителей сигнала сводится к байесовской оптимизации пороговых значений при условии, что веса ошибок обнаружения определяются, исходя из «стоимости» аномальных значений

C_n и $C_л$ в исходном процессе, и априорные вероятности отсутствия и наличия сигнала полагаются равными» [41]. Следовательно, для оптимизации параметров предлагаемого алгоритма необходима априорная информация о значениях решающих порогов и вероятностях отсутствия и наличия сигнала, а также весе аномальных значений.

В статье [42] Демин Н.С. и Михайлюк В.В. предлагают решение задачи обнаружения аномальных значений в дискретном канале для многомерной стохастической системы. В гибридных непрерывно-дискретных по времени измерительных системах дискретный канал является, как правило, более точным, информация, поступающая с него, более ценной и используется для коррекции информации, поступающая с непрерывного канала, а поэтому обнаружение аномальных значений в дискретном канале и учет их действия является особо важной задачей. Тогда задача обнаружения аномальных значений, с точки зрения теории статистических решений, как показано в работах Кульбак С. и Фукунага К. [43, 44], является задачей различения гипотез $H_o\{\theta = \theta_o\}$ (нормальный режим функционирования) и $H_1\{\theta = \theta_1\}$ (аномальный режим функционирования). Формируется логарифм отношения правдоподобия в момент времени t_m, т.е. $\widetilde{\Lambda}(\theta_1 : \theta_0)$. Обнаружитель аномальных помех на основе $\widetilde{\Lambda}(\theta_1 : \theta_0)$ имеет пороговую структуру вида [44]:

$$\widetilde{\Lambda}(\theta_1 : \theta_0) >^{\theta_1} <^{\theta_0} d(t_m),$$

где значения порога определяется видом критерия. В случае байесовского критерия

$$d(t_m) = \frac{P_{t_m}^0(\theta_0)}{P_{t_m}^0(\theta_1)} \frac{w(0,1) - w(0,0)}{w(1,0) - w(1,1)},$$

где $w(i, j)$ - элементы матрицы потерь $(0 \le i, j \le 1)$, $P_{t_m}^0(\theta_0), P_{t_m}^0(\theta_1)$ - априорные вероятности соответственно отсутствия и наличия аномальных

значений в момент времени t_m, которые определяются исходя из содержания конкретной задачи.

На практике распространены системы с резервированием измерительных комплексов (например, в навигационных системах подвижных объектов (работа Шакина В.В. [45]), когда зарезервированные датчики работают одновременно и их показания используются для построения оценки вектора состояния системы. При исследовании подобных систем возникают задачи, связанные с влиянием глубины резервирования на точность и на возможность обнаружения отказов датчиков. В работе Чемодановой О.А. [46] исследуются вопросы, связанные с отказом датчиков, когда отказ приводит к появлению аномальных значений. Эта задача в работах Кульбака С. [43] и Фукунага К. [44], рассматривается как статистическая задача распознания двух гипотез: $\mathrm{H}_{\mathrm{I}}^{0}$-гипотеза отсутствия отказов, $\mathrm{H}_{\mathrm{I}}^{1}$- гипотеза наличия отказов, обнаруживаемость отказов зависит от «расстояния» между гипотезами. Известны три меры, характеризующие «расстояния» между гипотезами: 1) расстояние Махаланобиса; 2) расстояние Бхаттачартя; 3) дивиргеция [43,44]. В работе [43] в качестве расстояния между гипотезами выбрана дивиргеция. Ограничением в применении данного метода является то, что вектор аномальных значений является гауссовской случайной величиной с нулевым математическим ожиданием и матрица структуры воздействия аномальных значений должна быть известна.

Одно из решений повышения помехозащищенности телеметрических сигналов приведено в работах Голубкова А.С., Коротаева В.П. [47] и Зеленского В.П. [48]. Задача выявления аномального значения рассматривается как задача статистического обнаружения случайного процесса $\varepsilon(n)$ в совокупности:

$$\lambda_k + \eta_k + \varepsilon_k = z_k + \varepsilon_k = Y_k, \qquad k = \overline{1, N} \tag{1.6}$$

где λ_k и η_k - соответственно полезная и шумовая составляющие нестационарного случайного процесса. Тогда согласно теории статистических решений, предложенной в работе Миддлтона Д. [49], процесс обнаружения аномальных значений можно представить в виде двух задач, решаемых последовательно: 1) увеличение соотношение сигнал/шум; 2) принятие решения по некоторому правилу.

Члены модели (1.6) являются «независимыми случайными гауссовскими последовательностями» с известными характеристиками [47, 48]. Для телеметрических систем в большинстве случаев выполняется соотношения: для среднеквадратического отклонения $\sigma_\varepsilon \geq \sigma_\lambda \geq \sigma_\eta$ и коэффициентов корреляции $r_\eta << r_\lambda$. Решения принимается на основе сравнения разности между анализируемым измерением и его оценкой, полученной при помощи полиномиальной экстраполяции по n предшествующим измерениям (работы Аминова И.Н., Фенедбаума А.А., Дудыкина А.В и др. [50, 51]). При экстраполяции на один шаг для исходного процесса Y_k, искомое выражение представляет:

$$\Delta^n Y_k = Y_k - \sum_{m=1}^{n} (-1)^m C_n^m Y_{k-m\tau} ,$$

где n - число предыдущих измерений, используемых для экстраполяции; τ - интервал дискретизации. Эффективность подобного преобразования определяется из соотношения: $\lambda = \dfrac{\alpha_\Delta}{\alpha}$, где $\alpha_\Delta = \dfrac{\sigma_{\varepsilon\Delta}^2}{\sigma_{\lambda\Delta}^2}$ -отношение сигнал/шум для разностного сигнала; $\alpha \dfrac{\sigma_\varepsilon^2}{\sigma_\lambda^2}$ - отношение сигнал/шум для исходного сигнала и зависит от порядка экстраполятора или от коэффициентов корреляции. Это справедливо только при условии $\sigma_\lambda >> \sigma_\varepsilon$, величина σ_q начинает оказывать заметное влияние на эффективность обработки только при $\sigma_\varepsilon > 0{,}2\sigma_\lambda$. Следующий этап

обработки - обнаружение аномальных значений в разностном сигнале, при этом реализуется некоторое решающие правило: $\Delta^n Y \geq K$ - есть аномальное значение, в противном случае оно отсутствует. Значение параметра K предлагается определить через функцию правдоподобия для независимых разностных процессов (Тихонов В.И. [52]): $L = \varphi(\Delta^n z)\varphi(\Delta^n \varepsilon)$. После ряда преобразований:

$$K^2 = -\frac{2\alpha_\Delta \sigma_{z\Delta}^2}{1+\alpha_\Delta} \ln(L^* \sigma_{\varepsilon\Delta}), \text{ где } L^* = 2\pi L.$$

В работе Голубкова А.С. и Коротаева В.П. [47] указано, что использование экстраполяции, для получения разности приводит к тому, что точки, соответствующие хорошо коррелированному полезному сигналу, прогнозируются с более высокой достоверностью, Это способствует созданию большего контраста между ним и слабо коррелированными аномальными значениями для разностных сигналов по сравнению с исходным входным процессом. Исследовалось также влияние величины коэффициента корреляции $r_\lambda(\tau)$ на эффективность обработки при использовании разносности второго порядка и получено выражение для его предельного значения $r_\lambda(\tau) \geq 0{,}7$. Указывалось, что за счет своего быстродействия алгоритм, построенный на предложенном методе, можно использовать в режиме реального времени. В работе Зеленского В.П. [48] предложен ряд решений обнаружения и исключения аномальных значений для некоторых видов законов плотности распределения вероятности приведены формулы для расчета коэффициента правдоподобия. Однако, в работах [47, 48] нет конкретных рекомендаций по выбору L, оговаривается лишь, что «его нужно подбирать с особой осторожностью».

В работе [35], авторами Фоминым А.Ф. и Новоселовым О.Н. рассматриваются вопросы теории и практики методов и алгоритмов

обнаружения одиночных и групповых аномальных значений непрерывных процессов при непрерывном и дискретном наблюдении.

В качестве основного показателя достоверности предлагается использовать вероятность появления аномальных значений за определенное время или за один отсчет: $P_{ан} = P(|\gamma| \geq |\gamma_{ан}|_{\min})$, где $|\gamma_{ан}|_{\min}$ - минимальное значение погрешности, начиная с которого она может считаться аномальной. Определение значения $|\gamma_{ан}|_{\min}$ представляет особый интерес и «является достаточно сложной задачей» [35].

В предлагаемых примерах алгоритмической и аппаратурной реализации методов обнаружения раздельно обрабатываются одиночные и групповые аномальные значения, причем используется априорная информация об их числе в группе как защита от обнаружения значений процесса. Для оптимизации порогового значения используется априорная информация о вероятности появления аномальных измерений за определенное время или на один отсчет сообщения. Оценка дисперсий порогового значения обнаружений является при этом функцией закона плотности распределения вероятности и длины интервала группирования аномальных значений [35]. При оценке контролируемой выборки в случае обнаружения аномальных значений, авторы Фомин А.Ф. и Новоселов О.Н. [35], предлагают использовать интерполяционные алгоритмы и алгоритмы на базе метода наименьших квадратов. Поскольку интерполяционные алгоритмы значительно проще в реализации, а их обнаруживающая способность не ниже алгоритмов метода наименьших квадратов (Новоселов О.Н., Фомин А.Ф. [53]), поэтому интерполяционные алгоритмы получили распространенное применение в виде полиномов Лагранжа. Авторы отдают предпочтение полиномам Лагранжа, так как их представление носит общий характер, и рассматриваемые в работе полиномиальные алгоритмы основывают на этом представлении.

Основным недостатком, предлагаемых методик, является необходимость наличия априорной информации о количестве аномальных значений в выборке исследуемого процесса.

Использование параметрических методов позволяет обнаруживать в основном единичные аномальные значения, а так же они требуют априорных сведений, как о законе плотности распределения вероятности процесса так и о его параметрах (математическом ожидании, дисперсии, корреляционной функции). Использование рассмотренных методов основывается на знании априорной информации об аномальных значениях (вероятность появления, величина аномалий и ряда других), что на практике встречается достаточно редко [54].

1.2 Непараметрические методы и алгоритмы обнаружения аномальных значений

Если неизвестны параметры распределения исследуемого процесса, то используются непараметрические методы, позволяющие определить, принадлежат ли исследуемые значения одной совокупности или нет. Непараметрические методы также применяют, если объемы данных недостаточны для применения параметрических методов (основанных на предположении гауссовского распределения исследуемых процессов). При использовании непараметрических критериев и методов априорная информация сводится только к заданию различий между конкурирующими гипотезами, сами же распределения, охватываемые той или иной гипотезой, не конкретизируются. Так же, условием применения непараметрических методов является, то, что функции распределения величин должны быть непрерывны [37, 55].

Непараметрический критерий для сдвинутых значений рассмотрен в работе Мостеллера [54]. Если даны выборки объема m из n непрерывных

генеральных совокупностей, то необходимо проверить гипотезу, что распределения всех генеральных совокупностей одинаковы, против альтернативы, что одно из них сдвинуто вправо. Непараметрическая процедура Мостеллера состоит в том, чтобы взять выборку с наибольшим значением и посчитать число ее элементов, которые превосходят все наблюдения остальных выборок. Если это число достаточно велико, то нулевая гипотеза опровергается. Данный критерий можно применять и для выборок неодинакового объема (работа Пирсона [14]). Если Y_k обозначает значение в k-й выборке ($k = 1,2,...N$), то сдвиг вправо может быть заменен следующим условием: $P\{Y_I > Y_J\} > \frac{1}{2}$ для некоторого I и $J = 1,2,...,\mathrm{I}-1,\mathrm{I}+1,...n$, где X_J – одинаково распределенные случайные величины.

После упорядочивания по возрастанию всех nm наблюдений и вычислении суммы рангов T_{I} для каждой выборки, выборка считается сдвинутой вправо при условии превышения $\max T$ некоторого критического значения. Такой метод впервые был предложен Дорнбосом и Пирсоном, которые и дали приближения для критических значений [14]. При использовании данного метода требуется условие стационарности входного процесса, некоррелированность отсчетов, а так же априорная информация о закон плотности распределения вероятностие распределения входной реализации процесса.

В работах Дэйвида Г. и Поулсона Е. [21, 56], предложен метод, который применим в случае совокупностей взаимно-независимых случайных величин Y_{kj} (k=1,2,…n; j=1,2,…m) с гауссовским законом плотности распределения вероятности $\mathrm{N}(\mu,\sigma^2)$. Решение D_0, состоящее в равенстве n математических ожиданий. Если все m_k равны, то D_0 должно быть выбрано с вероятностью $1-\alpha$. Поскольку фиксируется вероятность

правильного выбора D_0, то свойство оптимальности состоит в максимизации вероятности принять правильное решение, когда одна из генеральных совокупностей сдвинуто вправо. Это достигается если

$$\frac{m(Y_{\max} - \overline{m}_Y)}{\left[\sum_{k=1}^{n}\sum_{j=1}^{m}(Y_{kj} - \overline{m}_Y)^2\right]^{1/2}} < b_\alpha \quad ,$$

где $Y_{\max}$ – максимальное из n выборочных средних $\overline{m}_Y$, b_α – верхняя α-значимая точка статистики в левой части. В случае $m = 1$, то процедура является вариантом решения задачи об аномальности одного значения справа. В этом случае задача сводится к процедуре:

$$\frac{(Y_{(n)} - \overline{m}_Y)}{\left[\sum_{k=1}^{n}(Y_k - \overline{m}_Y)^2\right]^{1/2}} < b_\alpha .$$

При выполнении условия (с заданным уровнем значимости) принимается решение, что аномальных значений нет, в противном случае за аномальное принимается значение $Y_{(n)}$. Это по существу критерий Пирсона и Чандра Секара.

В работах Большева Л.Н., Смирнова Н.В. и Орлова А.И. [57 – 59] рассматривается ситуация, когда из $Y_1, Y_2, ...Y_N$ резко выделяется одно значение, обозначенное условно $Y_{\max}$. При нулевой гипотезе H_0 значения $Y_1, Y_2, ...Y_N$ рассматриваются как реализация независимых одинаково распределенных случайных величин с функцией распределения $F(Y)$. При альтернативной гипотезе H_1 случайные величины $Y_1, Y_2, ...Y_{N-1}$ имеют распределения $F(Y)$, а $Y_{\overline{1,N-1}}$ - распределение $G(Y)$, которое «существенно сдвинуто вправо», т.е $G(Y) = F(Y - A)$, где A достаточно велико. Если $Y_{\max} \le d$, то принимается гипотеза H_0, в противном случае -

гипотеза H_1. При справедливости нулевой гипотезы $P\left\{\max_{1\le k\le N} Y_k \le d\right\} = [F(d)]^N = 1-\alpha$, и критической значение $d = d(\alpha, N)$ определяется из уравнения $F(d) = \sqrt[N]{1-\alpha}$. При больших N и малых α

$$F(d) = \sqrt[N]{1-\alpha} = 1 - \frac{\alpha}{N} + O\left(\frac{\alpha^2}{N^2}\right),$$

поэтому в качестве хорошего приближения к $d(\alpha, N)$ рассматривается $(1-\frac{\alpha}{N})$ – квантиль распределения $F(Y)$.

Неустойчивость такой процедуры обнаружения, как показано в работе Орлова А.И. [57], связана с возможным неточным определением закон плотности распределения вероятностиа $F(Y)$ и трудность различия близких закон плотности распределения вероятностиов распределения с помощью критериев согласия и, в частности, с помощью критерия Колмогорова. Способность критерия различать близкие альтернативные гипотезы определяется его мощностью. Именно недостаточная мощность критерия не позволяет принять решение.

В работах Шакина В.В., Чемодановой О.А., Русинова Л.А., Паламарчук С.Н., Фомина А.Ф.и Новоселова О.Н., [45, 46, 60 – 62] привлекается внимание к проблеме фильтрации данных, содержащих аномальные значения - быстрые и значительные флуктуации процесса, обусловленные, например, случайными продолжительными и кратковременными неполадками в устройствах аналого-цифрового преобразования данных или в чисто цифровой части системы обработки сигналов. От одиночных, либо k-кратных изолированных аномальных значений можно избавиться, используя алгоритм скользящей медианы (работа Ефимова Л.Н. [63]).

Для случая групповых аномальных значений в исследованиях в работе. [64] приводится алгоритм фильтрации и сжатия данных кусочными многочленами первой степени, который дает возможность выявить начало и конец групповых аномальных значений, а, при необходимости, и избавиться от них. Пусть в течение интервала времени $0 \le t \le T$ регистрируется зависимость $Y(t_k)$, представляющая собой сумму детерминированного полезного сигнала $\lambda(t_k)$ и случайной шумовой составляющей $\eta(t_k)$, причем шумовая составляющая состоит из белого шума с нулевым средним и постоянной дисперсией σ^2 и протяженных аномальных значений типа ступенек. Относительно сигнала $\lambda(t_k)$ предполагается, что он имеет плавно изменяющиеся пологие участки, на которых его можно удовлетворительно описать полиномом первой степени (осуществляя, таким образом, сжатие данных). Выбирается один такой участок $t_r, t_{r+1}, ... t_{r+m}$ без точек разрыва и апроксимируются значения $Y(t_k)$, $k = i,...,i+m$, по методу наименьших квадратов при помощи функции $p_m(t_k) = a_m + \lambda_m t_k$. Алгоритм сжатия состоит в приписывании значения $k_1 = 0$.

Пусть найдено такое значение m_1, для которого впервые отношение $\bar{\sigma}^2_{m_1+1} / \sigma^2$ значимо отличается от ожидаемого согласно распределению χ^2 при заданном уровне значимости. Это соответствует тому, что отношение $\bar{\sigma}^2_{m_1+1} / \sigma^2$ не противоречит выдвинутой гипотезе при $2 \le m \le m_1$ и противоречит ей при $m = m_1 + 1$. Тогда точка t_{m_1+1} принимается в качестве начала следующего участка, соответствующего $k_2 = m_1 + 1$; в результате отыскивается точка со значением $k_3 = k_2 + m_2 + 1$ и т.д. В итоге получается разрывная кусочно-линейная функция. Если значение параметра σ^2

априорно неизвестно или меняется во времени, тогда в качестве статистики используется отношение $\overline{\sigma}_m^2 / \overline{\sigma}_{m_{j-1}}^2$, т.е. предлагается использовать предположение квазистационарности исследуемого процесса. В качестве начального значения σ^2 можно взять оценку дисперсии на участке, где присутствует только гауссовский шум и полезный сигнал.

Однако автор [64] в своей статье не указывает, как выбирается длина первоначально исследуемого участка, а так же не приводятся рекомендации о количестве всех последующих участков. Отмечено также, что алгоритм может реализовываться в режиме реального времени.

Если выборку из N значений, расположить в вариационный ряд $Y_1, Y_2, \ldots, Y_N$, то принадлежность минимального $Y_{\min} = Y_1$, или максимального $Y_{\max} = Y_N$ к рассматриваемой совокупности $\{Y_2, Y_3, \ldots Y_{N-1}\}$ можно проверить с помощью непараметрической статистики Диксона, имеющую следующую плотность распределения (Дэйвид Г., Дуров А.М., Мхиторян В.С., Трошин Л.И. [21, 65]):

$$f(Y) = \frac{N!}{(N-3)!} \int_{-\infty}^{\infty} \int_{0}^{\infty} \left[\int_{y-v}^{y-rv} f(t)dt \right]^{N-3} f(Y-tv) f(Y) v dv dY, \qquad (1.7)$$

где Y - переменная, имеющая гауссовский закон плотности распределения вероятности распределения. Если крайние значения Y_1 или Y_N являются одиночными сбоями, то статистика r представляет собой отнесенную к размаху $(Y_N - Y_1)$ исходящую первую разность, когда проверяется значение:

$$r_{1,N} = \frac{Y_N - Y_{N-1}}{Y_N - Y_1},$$

или восходящую первую разность, когда проверяется значение $Y_{\min} = Y_1$:

$$r_{1,1} = \frac{Y_1 - Y_2}{Y_N - Y_1}.$$

Из выражения (1.7) видно, что распределение переменной Диксона зависит только от объема выборки N и не зависит от среднего значения и дисперсии исходного распределения $f(Y)$. Непараметрические свойства распределения (1.7) позволяют использовать его для проверки достоверности информации при произвольном характере распределения исходных данных. Другое достоинство переменной r - возможность идентификации не только одиночных, но и групповых аномальных значений с помощью статистики вида:

$$r_{i,j} = \frac{Y_N - Y_{N-i}}{Y_N - Y_i},$$

где i - число значений в группе, подозреваемые на аномальность, N - объем анализируемой выборки.

Достоверность проверяемой информации определяется в соответствии с решающим правилом:

$$r_{i,N} > r_\alpha, \quad \text{где } \alpha \text{ - уровень значимости.} \qquad (1.8)$$

Правило (1.8) позволяет принять решение с гарантируемой величиной вероятности ошибки первого рода α. Однако вычислить вероятность ошибки второго рода β с его помощью невозможно, так как альтернатива при этом не проверяется [66].

В программной системе, которая представлена в статье Лемешко Б.Ю. [67], разработанной в Новосибирском государственном техническом университете, все задачи статистического анализа данных рассматриваются с точки зрения наиболее общего представления экспериментальных наблюдений в виде частично группированных выборок (ссылаясь на работы Куллдорфа Г, Денисова В.И., Лемешко Б.Ю., Цой Е.Б. [68, 69]), частными случаями которых являются

негруппированные, группированные и цензурированные выборки. Основным методом оценивания параметров распределения является метод максимального правдоподобия, который предложен в работе Кендалла М. и Стьюарта А. [70]. Оценки параметров распределений находятся в результате максимизации функции правдоподобия по частично группированной выборке, которая имеет вид:

$$L(\theta) = \prod_{(1)} \mathrm{P}_i^n(\theta) \prod_{(2)} \prod_{j=1}^{N} f(Y_{ij}, \theta),$$

где $f(Y,\theta)$ - функция плотности случайной величины, $P(\theta) = \int_{Y_{i-1}}^{Y_i} f(Y,\theta)dY$ - вероятность попадания наблюдений в i-й интервал значений, (1) и (2) означают, что умножение осуществляется по интервалам с группированными и с негруппированными данными соответственно.

Проверка гипотез о согласии осуществляется по ряду критериев: отношения правдоподобия, χ^2 Пирсона, Колмогорова, Смирного, ω^2 и Ω^2 Мизеса. Решение о степени соответствия выборки с закон плотности распределения вероятностиом распределения принимается по их совокупности. Гипотеза о согласии не отвергается, если $P(S > S^*) > \alpha$, где задаваемый уровень значимости α определяется из выражения:

$$\alpha = \int_{S^*}^{\infty} g(S)dS.$$

В работе Лемешко Б.Ю. [71] отмечается два достоинства данного метода. Во-первых, свойства получаемых оценок по группированным выборкам позволяет резко снизить влияние аномальных значений, а иногда даже исключить их влияние. Во-вторых - асимптотически оптимальное группирование минимизирует потери информации, связанные с группированием и, следовательно, гарантирует максимальную мощность различения близких альтернатив для этих критериев. Автор статьи [71]

делает вывод, что при идентификации (при оценивании параметров распределения) необходимо использовать устойчивые к наличию аномальных значений методы оценивания (робастные), а на последующем этапе желательна максимальная мощность критерия для различия близких альтернатив (чувствительность к аномальным значениям). В этой связи на первом этапе оценку необходимо производить по группированным выборкам, причем для большей устойчивости оценок разбиение выборки на интервалы равной вероятности (равночастотные интервалы), а на втором этапе при проверке согласия - разбиение на асимптотически оптимальные интервалы.

В работе [72] рассматривается задача обнаружения аномальных значений в многомерных выборках с использованием оценок неизвестных параметров. На численных примерах демонстрируются робастные свойства предложенных оценок, которые проявляют свою нечувствительность к присутствию аномальных значений. Предложенный метод заключается в обнаружении значений, которые превышают пороговое значение. Для обнаружения нескольких аномальных значений, при гауссовском закон плотности распределения вероятностие распределения сигнала, предлагается правило, при котором не рассматриваются те значения Y_k, для которых среднеквадратическое отклонение $\sigma_k \geq \sqrt{\chi^2_{\rho;0.975}}$ (здесь $\chi^2_{\rho;\alpha}$ означает α - квантиль распределения χ^2 с ρ степенями свободы). Но рекомендаций по выбору такого порогового значения не приводится. Очевидно, что пороговое значение должно зависеть от числа объема выборки исследуемого процесса, потому что в противном случае, при достаточно большой длине выборки, ошибка вероятности первого рода будет возрастать. Это значит, что предложенный «критерий будет показывать наличие аномальных значений, когда на самом деле они отсутствуют» [72].

Методы оценивания, учитывающие наличие «грубых ошибок», и позволяющие достаточно точно определять оценки параметров называются робастными или устойчивыми [7, 58, 72, 73 –76]. Например, методы Хубера, Винзора, Пуанкоры, для симметричных распределений, метод джекккнайф оценок и Хинкли для асимметричных распределений и ряд других методов [65, 73].

Робастные процедуры, лишь незначительно уступая в эффективности классическим оптимальным процедурам при точном выполнении условий их оптимальности, сохраняют в отличие от них высокую эффективность при нарушении этих условий. Робастные статистические процедуры не должны быть чувствительны к ошибкам, которые могут попадать в исходный процесс или искажать используемые вероятностно-статистические модели.

Известна наиболее простая модель изучения устойчивости к аномальным значениям — это модель засорения вида:

$$F(Y) = (1-\alpha)F_0(Y) + \alpha \cdot H(Y), \qquad (1.9)$$

где $F(Y)$ функция распределения.

Модель (1.9) впервые предложена Тьюком и Хубером [73] (модель Тьюки-Хубера). Эта модель (1.9) показывает, что с вероятностью близкой к единице, а именно, с вероятностью $(1-\alpha)$, наблюдения берутся из совокупности с функцией распределения $F_0(Y)$, которая предполагается, обладает «хорошими» свойствами, то есть она имеет известный вид распределения (хотя бы с точностью до параметров), у нее существуют все моменты. С вероятностью α появляются наблюдения из совокупности с «плохим» распределением, например, взятые из распределения Коши, не имеющего математического ожидания, резко выделяющиеся аномальные значения.

Наличие аномальных значений может сильно исказить результаты анализа процессов. Если функция распределения элементов выборки процесса имеет вид (1.10), где первое слагаемое соответствует случайной величине с конечным математическим ожиданием, а второе - такой, для которого математического ожидания не существует (например, если $H(Y)$ - функция распределения Коши), то для итоговой функции распределения (1.10) не будет существовать математического ожидания. Среднее арифметическое результатов наблюдений модели (1.10) не будет иметь никакого предела (это – строгое математическое утверждение, вытекающее из того, что математическое ожидание не существует [73]).

Построена достаточно обширная и развитая теория, посвященная разработке и изучению методов анализа процессов в модели (1.10) [57,73]. Теоретически, обычно, предполагается известной вероятность ошибки первого рода α, что на практике не выполняется. Кроме того, теоретические исследования обычно направлены на защиту от воздействий, предположительно угрожающих из бесконечности, а на самом деле реальные значения процессов финитны. Все это объясняет, почему теория робастности, исходящая из модели (1.9), популярна среди теоретиков, но мало интересна тем, кто анализирует реальные данные.

В работе Дурова А.М. и Мхиторяна В.С. и др.[65] приведен метод выявления аномальных значений, который называется метод Смирнова-Граббса. По результатам наблюдений $Y_1, Y_2, \ldots, Y_N$ строится вариационный ряд $Y_1 \le Y_2 \le \ldots \le Y_N$, если известно среднеквадратическое отклонение σ_Y и математическое ожидание μ_Y исходного процесса, то можно воспользоваться критерием:

$$t_N = \frac{Y_N - \mu_Y}{\sigma_Y}$$

и сравнить его с данными таблицы процентных точек критерия Смирнова-Граббса:

$$t_{\max} = \frac{\max\limits_{k} \left| Y_k - m_Y \right|}{\sigma_Y}.$$

Если же σ_Y и μ_Y неизвестны, то необходимо вычислить статистику:

$$T_N = \frac{Y_N - \overline{m}_Y}{\overline{\sigma}_Y}, \quad \text{где } \overline{m}_Y = \frac{1}{N}\sum_{k=1}^{N} Y_k; \quad \overline{\sigma}_Y{}^2 = \frac{1}{N}\sum_{k=1}^{N}(Y_k - \overline{m}_Y)^2.$$

Если $T_N < C_\alpha$, то верна гипотеза H_0 о том, что Y_N не является аномальным значением. При $T_N > C_\alpha$, Y_N значимо отличается от $\overline{m}_Y$, следовательно, является аномальным значением.

При использовании критерия Граббса для обнаружения одного аномального значения в случае проверки максимального значение на наличие грубой ошибки, используют следующий критерий:

$$G_N = \frac{\sum\limits_{k=1}^{N-1}(Y_k - \overline{m}_{Y_{\overline{1,N-1}}})^2}{\sum\limits_{k=1}^{N}(Y_k - \overline{m}_Y)^2}, \text{ где } \overline{m}_{Y_{\overline{1,N-1}}} = \frac{1}{N-1}\sum_{k=1}^{N-1} Y_k;$$

$$\overline{m}_Y = \frac{1}{N}\sum_{k=1}^{N} Y_k.$$

Если необходимо проверить минимальное значение на наличие грубой ошибки применяют критерий вида:

$$G_1 = \frac{\sum\limits_{k=2}^{N}(Y_k - \overline{m}_{Y_{\overline{2,N}}})^2}{\sum\limits_{k=1}^{N}(Y_k - \overline{m}_Y)^2}, \quad \text{где } \overline{m}_{Y_{\overline{2,N}}} = \frac{1}{N-1}\sum_{k=2}^{N} Y_k.$$

Значения процесса Y_1 или Y_N относят к аномальным, если наблюдаемые статистики G_1 и G_N окажутся больше критического C_α.

Критические значения C_α для критериев G_1 и G_N приведены в работе Дэйвида Г. [21]. Причем статистики G_N и T_N связаны соотношением:

$$G_N = 1 - \frac{1}{N-1}\left(\frac{Y_N - \overline{m}_Y}{s}\right)^2 = 1 - \frac{1}{N-1}T_N^2.$$

На основе критерия Граббса можно записать критерий исключения нескольких аномальных значений (подход Титьена-Мура – обобщение критерия Граббса на несколько экстремальных наблюдений). Если по выборке объемом N построить вариационный ряд, то тогда решающее правило для отношения k наибольших наблюдений к грубым ошибкам основано на статистике [21]:

$L_i = \dfrac{\sum\limits_{k=1}^{N-i}(Y_k - \overline{m}_{Y_{\overline{1,N-i}}})^2}{\sum\limits_{k=1}^{N}(Y_k - \overline{m}_Y)^2}$, где $\overline{m}_{Y_{\overline{1,N-i}}} = \dfrac{\sum\limits_{k=1}^{N-i} Y_k}{N-k}$, $\overline{m}_{Y_{\overline{1,N-i}}}$ – средняя арифметическая $(N-i)$ значений после исключения i наибольших значений; $\overline{m}_Y$ - выборочная средняя арифметическая по всей выборке.

Решающее правило для отнесения i наименьших наблюдений к аномальным значениям основано на статистике:

$$\widetilde{L}_i = \frac{\sum\limits_{k=i+1}^{N}(Y_k - \overline{m}_{Y(i)})^2}{\sum\limits_{k=1}^{N}(Y_k - \overline{m}_Y)^2},$$

где $\overline{m}_{Y_{\overline{i+1,N}}} = \dfrac{\sum\limits_{k=i+1}^{N} Y_k}{N-k}$, $(N-i)$ – число наибольших значений в выборке; $\overline{m}_{Y_{\overline{i+1,N}}}$ – средняя арифметическая $(N-i)$ наблюдений после исключения i наименьших значений; $\overline{m}_Y$ – выборочная средняя арифметическая по всей выборки; L_k - эквивалентна G_N, а $\widetilde{L}_k$ – величине G_1.

На основе полученных критериев обнаружения аномальных значений для максимальных и минимальных значений можно записать общий критерий для одновременного обнаружения наибольших и наименьших экстремальных значений. По выборке вычисляем среднее арифметическое значение $\overline{m}_Y$ и проводим расчет абсолютных отклонений от среднего:

$$r_1 = \left|Y_1 - \overline{m}_Y\right|,\ r_2 = \left|Y_2 - \overline{m}_Y\right|,\ldots,\ r_N = \left|Y_N - \overline{m}_Y\right|.$$

Строится возрастающий ряд с переиндексацией: Z_k - наблюдение r_k – k-е по величине. Самое близкое к $\overline{m}_Y$ значение обозначим Z_1, а через Z_N – самое удаленное от $\overline{m}_Y$ значение.

Для проверки гипотезы о том, что i наибольших по модулю значений являются аномальными, используется следующий критерий:

$$E_i = \frac{\sum_{k=1}^{N-i}(Z_k - \overline{m}_{Z(i)})^2}{\sum_{k=1}^{N}(Z_k - \overline{m}_Z)^2},$$

где $\overline{m}_{Z_{\overline{i+1,N}}} = \frac{\sum_{k=1}^{N-i} Y_k}{N-i}$ ($\overline{m}_{Z_{\overline{i+1,N}}}$ – средняя арифметическая из $N-i$ наблюдений, оставшихся после исключения из выборки i подозрительных элементов); $\overline{m}_Z$ – средняя арифметическая всей выборки.

Основной недостаток критерия Граббса: необходима априорная информация о количестве аномальных значений в выборке; при малом объеме выборке, оценки среднего и среднеквадратического отклонения будут смещены, что является основном недостатком робастных методов, т.е. уровень засорения выборки исследуемого процесса должен быть априорно известен, а так же распределение должно быть симметричным. При асимметричности распределения оценки Хубера, Пуанкаре и Винзора теряют свою эффективность, становясь несостоятельными и смещенными.

Поэтому при несимметричных распределениях наиболее удобными можно считать джеккнайф-оценку Тьюки и Квенсулла. Суть их заключается в том, что значения процесса разбиваются на группы. Эффект каждой группы оценивается по результату, полученному при исключении данной группы из рассмотрения. Такая оценка позволяет уменьшить смещение параметра положения при асимметричных распределениях. Взвешенная джеккнайф-оценка, которую предложил Хинкли заключается в том, что при оценке параметров процесса вклад значений с большими весами уменьшается. Весовые коэффициенты выбираются как расстояние, отражающие недостаток симметрии, и определят вклад каждого значения в дисперсию. Таким образом, априорно должен быть известен коэффициент асимметрии для правильного подбора весовых коэффициентов [21].

Робастные методы позволяют использовать информацию о том, что реальные наблюдения лежат «около» тех или иных параметрических семейств распределения случайной величины. В этом их преимущество по сравнению с непараметрическими методами, которые предназначены для анализа данных из всех возможных распределений. Однако количественных подтверждений этим утверждениям не удается найти [21, 65].

Непараметрические критерии и методы применимы тогда, когда реализация процесса ограниченного объема. Если реализация процесса содержит больше 100 значений, то не имеет смысла использовать непараметрические критерии. Согласно центральной предельной теореме, когда выборки становятся очень большими, то выборочные средние подчиняются гауссовскому закон плотности распределения вероятности, даже если исследуемый процесс не является гауссовским. Таким образом, параметрические методы, являются более чувствительными (имеют большую статистическую мощность), всегда подходят для большого объема выборки процесса, даже при наличии аномальных значений в ней.

В работе Егорова В.А., Морозова Ю.Г., Чекмаревой А.П. [77] представлен метод обнаружения аномальных значений для нестационарных случайных процессов [77].

Если наблюдаемый процесс представлен в виде модели:

$$Y(t) = \lambda(t) + \eta(t) + \varepsilon(t),$$

где $\lambda(t)$ – функция полезного сигнала, $\eta(t)$ – аддитивная шумовая составляющая, $\varepsilon(t)$ – аномальные значения, то «ошибки значений представляют собой центрированный гауссовский случайный процесс» [77]. При этом последовательность действий по определению и устранению аномальных значений заключается в следующем: общее исходное множество результатов измерений случайным образом, независимо друг от друга разбивается, на MP подмножеств, в среднем по M элементов каждое:

$$\left\{k_p^{(j)}\right\}_{p=1}^{M}, \qquad k = 1,2,\ldots,MP,$$

где k – номер значения, которое соответствует некоторому моменту времени.

Для каждого из подмножеств MP, строится линейная функция, причём: $\overline{Y}^{(j)}\left(t_{k_p}^{(j)}\right) = Y_{k_p}^{(j)}$, то есть в точках k_p линейная функция совпадает со значениями исходной реализации процесса. В остальных точках значение функции $\widetilde{F}^{(j)}\left(t_k^{(j)}\right)$ строится с помощью линейной интерполяции. Сравнение полученной оценки $\widetilde{F}^{(j)}\left(t_k^{(j)}\right)$ с исходным процессом и с некоторым априорно заданным значением ошибки ε, позволяет определить аномальное значение на данном подмножестве, то есть если выполняется неравенство:

$$\left|\frac{Y_k - \overline{Y}^{(j)}\left(t_k^{(j)}\right)}{Y_k}\right| > \varepsilon,$$

то данному измерению Y_k приписывается штрафное очко. Данное действие осуществляется на каждом из MP подмножеств. Затем исходная реализация процесса вновь разбивается на MP подмножеств случайным образом, и все вышеуказанные действия повторяются p раз. При этом общее количество штрафов полученных экстремальными значениями накапливаются. После p - кратного анализа исходной реализации, сумма штрафов, полученных каждым значением, сравнивается с некоторым пороговым числом K. При превышении порога K, значение процесса считается аномальным.

Таким образом, определяется наличие в исходной реализации процесса, как единичных аномальных значений, так и групповых. Для данного метода требуется априорные данные о требуемом количестве подмножеств MP, количестве повторений p, значения допустимой ошибки отклонения ε, и пороге K, то есть необходимо задавать четыре параметра исходя из требований максимальной вероятности правильного обнаружения аномального значения и минимальной вероятности ложной тревоги. Конкретное значение каждого параметра зависит от вида функции полезного сигнала и вероятностных характеристик аддитивной шумовой составляющей, закон плотности распределения вероятности распределения которой должен быть только гауссовским и отсчеты шумовой составляющей сигнала независимы. Все это не позволяет использовать данный алгоритм в автоматизированных системах сбора и обработки телеметрической информации. А так же аномальное значение не штрафуется в тех случаях, когда Y_k является одной из вершин аппроксимирующей ломанной.

В работах Марчука В.И. [54, 78, 79] рассмотрен адаптивный метод обнаружения аномальных значений при анализе нестационарных случайных процессов в условиях непараметрической априорной

неопределенности. Предложенный метод основывается на методе размножения оценок исходного нестационарного процесса для выделения полезной составляющей при непараметрической априорной неопределенности [54, 79].

Пусть нестационарный случайный процесс представляет собой единственную реализацию дискретного процесса $\{Y_k\}_{k=1}^{N}$, полученного в равноотстоящие моменты времени $t_1, t_2, ..., t_N$, где $Y(t_k) = Y_k$ и $k = \overline{1, N}$. Математическая модель реализации исходного процесса может быть представлена выражением

$$Y_k = \lambda_k + \eta_k + \varepsilon_k, \tag{1.10}$$

где λ_k, η_k, ε_k – значения низкочастотной составляющей сигнала, аддитивной высокочастотной и аномальной составляющих входного процесса соответственно.

Суть предлагаемого метода [54, 79] заключается в разбиении исходной дискретной реализации значений Y_k на интервалы ∇_j, $j = \overline{1, p}$ случайной длины, в дальнейшем именуемые покрытиями и аппроксимации на каждом из них полиномом второй степени, где p - количество покрытий.

Многократное разбиение на покрытия ∇_j и аппроксимация на них значений процесса Y_k позволяют размножать оценки единственной реализации результатов измерений. На каждом покрытии ∇_j проводят аппроксимацию значений исходного сигнала полиномом не выше второй степени по методу наименьших квадратов и находят разность Δ_k между значениями оценки аппроксимирующей функции S_k и исходной реализацией процесса Y_k. Для разностного процесса Δ_k на каждом

покрытии определяются оценки математического ожидания $\overline{m}_j$ и среднеквадратического отклонения $\overline{\sigma}_j$, далее устанавливается некоторое пороговое значение

$$\varepsilon_j = A \cdot \overline{\sigma}_j, \tag{1.11}$$

где $0 < A \leq 3$. Превышение установленного порогового значения разностным процессом Δ_k штрафуется на каждом из покрытий, т.е. если

$$\Delta_k + \varepsilon_j < Y_k < \Delta_k - \varepsilon_j, \tag{1.12}$$

то Y_k получает одно штрафное значение. В соответствии с методом размножения оценок изложенная выше процедура определения штрафов повторяется для каждой из размноженных оценок исходной реализации нестационарного случайного процесса [54, 79]. При выполнении условия (1.12) происходит накопление значений штрафов для элементов исходной реализации. По окончанию обработки, для всех оштрафованных значений исходной реализации процесса находят среднее значение штрафов. Отсчеты, у которых суммарное число штрафов превышает среднее значение, будут определяться как аномальные. Единичные и групповые аномальные значения устраняются с помощью линейной интерполяции проводимой по двум точкам, соответствующим значениям полученных измерений до аномального значения (группы аномальных значений) и после аномального значения (группы аномальных значений).

В работе [54, 79] аналитически доказаны рекомендуемые значения параметров предложенного метода. Количество размножений $K \geq 10$, количество покрытий $p \geq 8$, закон плотности распределения вероятности распределения длин покрытий предпочтительно равномерный, а минимальная длина покрытия порядка $\nabla \approx 8$. Автор метода обнаружения аномальных значений Марчук В.И. [54, 79] предлагает ввести адаптацию порогового значения (1.11) относительно величины оценки

среднеквадратического отклонения $\overline{\sigma}_j$, говоря о том, что значения некоторого коэффициента A должно быть, фиксировано в интервале $2 < A < 3$. Однако автор не дает точный рекомендаций по выбору этого значения коэффициента A. Правильный выбор коэффициента A, будет влиять на точность задания порогового значения (1.2), а оно в свою очередь на эффективность предлагаемого метода [54, 79].

В монографии [54] приведен подробный анализ существующих методов обнаружения аномальных значений, где показано, что не существует эффективного метода обнаружения аномальных значений при условии нестационарности случайного процесса по математическому ожиданию.

Таким образом, анализ существующих методов обнаружения аномальных значений показал, что не существует универсального подхода к решению задачи обнаружения аномальных значений в выборке как стационарного, так и нестационарного случайных процессов. Каждый из рассмотренных методов имеет ряд ограничений, которые являются зачастую решающими для их использования на практике.

ГЛАВА 2 Способы обнаружения аномальных значений

Одним из наиболее важных требований, предъявляемых к системам цифровой передачи и обработки сигналов, является достоверность передаваемой информации от источника сообщения к получателю. На практике выполнению этого требования неизбежно препятствуют ошибки, которые вызваны внешними воздействиями помех, воздействующих на вход приемного устройства по каналам связи, внутренними шумами, возникающими в самом приемном устройстве, искажениями сигнала, связанными непосредственно с его прохождением по физическому каналу связи.

2.1 Математические модели сигналов

В зависимости от причин возникновения или типа источника, помехи можно разделить на следующие классы: космические, атмосферные, промышленные, обусловленные особенностями распространения в канале (эхо, замирания, флуктуации затухания при распространении и т.д.), искусственные, внутренние тепловые шумы устройств приема сигналов и т.п.

В зависимости от способа взаимодействия шумовой составляющей и полезного сигнала различают аддитивные и мультипликативные помехи. Например, аддитивными шумами являются: излучение посторонних источников; помехи линейных переходов, обусловленные конечной величиной защищенности между парами кабелей; помехи типа попутного потока и нелинейных переходов; тепловой шум; проявляющиеся на макроскопическом уровне; дискретность электромагнитного или светового излучения при слабых уровнях сигнала и т.д. В качестве мультипликативных, можно назвать непредсказуемость изменения

коэффициентов передачи канала, перерывы, замирания сигналов и т.д. [80].

Большое количество работ, посвящено описанию моделей как аналоговых, так и дискретных каналов связи, например работы Миддлтона Д., Финка Л.К., Зюко А.Г., Левина Б.Р., Блоха Э.Л., Попова О.В. и др. [81 - 85]. Модель канала связи, в общем случае, сводится к заданию математической модели принимаемых сигналов.

В общем случае влияние помехи на полезный сигнал можно выразить, согласно предложению Прокиса Дж. [86], оператором:

$$Y_k = L[\lambda_k(\mu_k, \eta_k, \varepsilon_k)] \tag{2.1}$$

В частном случае, когда оператор вырождается в сумму, процесс называют аддитивным и представляют моделью вида [86]:

$$Y_k = \lambda_k + \eta_k + \varepsilon_k, \tag{2.2}$$

λ_k ,η_k ,ε_k – полезная, аддитивная шумовая и аномальная составляющие входного процесса соответственно, где $k = \overline{1, N}$, N – объем выборки исследуемого сигнала.

Если же оператор может быть представлен в виде произведения, процесс называют мультипликативным и представляют моделью вида [86]:

$$Y_k = \lambda_k \cdot \mu_k + \varepsilon_k, \tag{2.3}$$

μ_k – мультипликативная шумовая составляющая процесса.

В системах передачи и обработки сигналов обычно имеют место и аддитивные, и мультипликативные шумовые составляющие, поэтому:

$$Y_k = \lambda_k \cdot \mu_k + \eta_k + \varepsilon_k \tag{2.4}$$

Наряду с шумами, сигналы содержат аномальные значения, в том числе импульсные помехи ε_k [80, 87 – 89]. В зависимости от причин возникновения и характера реализации, исходный процесс $Y(t_k)$, можно представить в виде следующих моделей (работа Фомина А.Ф., Новоселова О.Н. и Плющева А.В. [35]):

$$
\begin{aligned}
&Y(t_k) = [1 - a(t_k)] \cdot [\lambda(t_k) + \eta(t_k)] + a(t_k)\varepsilon(t_k), \\
&Y(t_k) = [1 - a(t_k)] \cdot \lambda(t_k) + \eta(t_k) + a(t_k)\varepsilon(t_k), \\
&Y(t_k) = \lambda(t_k) + [1 - a(t_k)] \cdot \eta(t_k) + a(t_k)\varepsilon(t_k), \\
&Y(t_k) = \lambda(t_k) + \eta(t_k) + \varepsilon(t_k),
\end{aligned}
\tag{2.5}
$$

где $k = \overline{1, N}$, $a(t_k)$ – реализация двоичной переключательной случайной функции (потока событий) или случайной величины, принимающей с вероятностью $P_{ан}$ значение $a(t_k) = 1$ и с вероятностью $(1 - P_{ан})$ значение $a(t_k) = 0$, при этом считаем $a(t_k)$ некоррелированное со значениями полезной составляющей $\lambda(t_k)$, значениями аддитивной шумовой составляющей $\eta(t_k)$ и $\varepsilon(t_k)$ – значениями аддитивной аномальной составляющей.

Первые две модели в (2.5) относятся к случаю, когда аномальные значения полностью подавляют сигнал. Такое подавление наблюдается, например, при срывах синхронизации, при передаче по системам с угловой модуляцией и т.д. Третья и четвертая модели (2.5) характеризуют неполное подавление сигнала аномальными значениями. Такой случай имеет место, например, в цифровых системах при искажениях отдельных символов в кодовом слове [35].

Шумовая составляющая процесса, в общем случае, имеет различные законы плотности распределения вероятности. Это зависит от причины или типа источника, например тепловой шум, возникающий в радиоприемных устройствах можно характеризовать как гауссовский случайный процесс. Такой шум наиболее изучен и представляет интерес, как в теоретическом, так и в практическом отношении. Одномерная плотность вероятности гауссовского распределения имеет вид [90]:

$$
W(\eta, m, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(\eta - m)^2}{2\sigma^2}\right], \tag{2.6}
$$

где область допустимых значений случайной величины $-\infty < \eta < +\infty$, m – математическое ожидание, $\sigma > 0$ среднеквадратическое отклонение.

Использование в качестве модели шумовой составляющей гауссовский случайный процесс является более сложным случаем с точки зрения выделения полезной составляющей процесса в сравнении с другими симметричными законами распределения [54, 91 – 93]. Основная причина этого заключается в том, что случайные значения неравномерно распределены по вероятности появления в интервале определения и значительная их часть сосредоточена на «хвостах» распределения.

Однако, как сказано в работах Орлова А.И., Переверткина С.М. и др. [54, 91 – 93], в случаи ограниченного объема выборки исследуемого процесса, не всегда можно полагаться на то, что шумовая составляющая процесса имеет гауссовский закон плотности распределения вероятности.

Равномерный закон плотности распределения вероятности в основном характеризует шумы квантования в результате аналого-цифрового преобразования сигнала [96]. Одномерная плотность вероятности равномерного распределения и его числовые характеристики имеет вид [104, 105]:

$$W(\eta, m, \sigma) = \frac{1}{\sigma}, \qquad W(\eta, m, \sigma) = \begin{cases} 0 & при \quad \eta < m, \\ \dfrac{\eta - m}{\sigma} & при\ m \le \eta \le m + \sigma, \\ 1 & при \quad \eta \ge m + \sigma, \end{cases} \tag{2.7}$$

где $\eta \subseteq (m, m+\sigma)$ область допустимых значений случайной величины, $|m| < \infty$ – математическое ожидание, $\sigma > 0$ – среднеквадратическое отклонение.

Распределение Рэлея, относится к экспоненциальному типу, но отличается несимметричностью и в отличие от гауссовского закона плотности распределения вероятности его одномерная плотность вероятности имеет вид [90]:

$$W(\eta, m, \sigma) = \frac{\eta}{\sigma^2} \exp\left[\frac{-\eta^2}{2\sigma^2}\right], \qquad (2.8)$$

где $\eta \geq 0$ область допустимых значений случайной величины, $\sigma > 0$ – среднеквадратическое отклонение и $W(\eta, m, \sigma) = 0$, при $\eta < 0$.

В прикладных задачах рэлеевская модель закона плотности распределения вероятности шумовой составляющей процесса используется, например, при приеме телеметрического сигнала с использованием радиотехнического канала связи. При этом, если на входе полезный сигнал был подвержен воздействию гауссовского шума, то на выходе детектора закон плотности распределение вероятности шума является рэлеевским [97].

Аномальные значения, которые могут быть как одиночными, так и групповые, искажающие несколько подряд следующих отсчетов. При дискретном процессе отдельное аномальное значение представляет собой дискретную дельта - функцию, обозначаемую символом Кронекера δ_{kh} [35].

Поэтому аномальную составляющую входного процесса можно представить следующими моделями [35]:

– для одиночного аномального значения считаем ε_k мгновенным выбросом со случайной величиной Amp_k:

$$\varepsilon_k = Amp_k \delta_{kh}; \qquad \delta_{kh} = \begin{cases} 1, k = h; \\ 0, k \neq h, \end{cases} \qquad (2.9)$$

где h – момент возникновения аномального значения;

– для групповых, некоррелированных аномальных значений считаем ε_k отрезком реализации случайной последовательности $\varepsilon_k = Amp_k$ с некоррелированными значениями Amp_k [80]:

$$\varepsilon_k = \begin{cases} Amp_k, k = h, h+1, ..., h+(n_П - 1); \\ 0, k \notin h, ..., h+(n_П - 1), \end{cases} \qquad (2.10)$$

где $n_{П}$ – число аномальных значений в группе.

Вероятностные свойства аномальной составляющей процесса, для практических целей, описываются законами плотности распределения величины аномальных значений и распределением временных интервалов между их появлением. Для последних, часто применяют модель Пуассона [80]:

$$W(\rho) = \frac{1}{\beta\sqrt{2\pi}} \exp\left[\frac{(\rho - a)^2}{2\beta^2}\right]. \qquad (2.11)$$

Распределение величины аномальных значений часто описывается логнормальным законом [80]:

$$W(Amp) = \frac{1}{Amp\beta\sqrt{2\pi}} \exp\left[\frac{(\ln Amp - a)^2}{2\beta^2}\right], \qquad (2.12)$$

где a, β^2 – математическое ожидание и дисперсия процесса, $\rho = \ln Amp$. Для закона (2.12) характерно наличие существенных значений в области больших амплитуд, что соответствует реальной модели импульсной помехи [80].

В качестве моделей функции полезной составляющей λ_k можно выбрать группу моделей, с помощью которых можно представить большинство реальных сигналов. Наиболее часто встречающейся моделью полезного сигнала является гармоническая функция вида [98]:

$$\lambda(t_k) = C\sin(\omega t_k). \qquad (2.13)$$

Выбор гармонической функции в качестве модели полезной составляющей процесса обусловлено тем, что ряд сигналов носят колебательный характер, например гармонические колебания, являются единственной функцией времени, сохраняющей свою форму при прохождении через любую линейную цепь, при этом меняется лишь амплитуда и фаза колебаний.

Экспоненциальная модель функции полезного сигнала вида [98]:

$$\lambda(t_k) = A\exp(-\alpha t_k),\ \alpha > 0, \tag{2.14}$$

широко используется в радиотехнических устройствах при описании переходных процессов. Данной зависимостью описываются фронты видеоимпульса при прохождении его через апериодические цепи, а так же часто используется для моделирования переходных процессов различных радиотехнических устройств [98].

Другой вид модели функции полезной составляющей представляет собой полиномиальную функцию первой и второй степени, т.е. вида [98]:

$$\lambda(t_k) = a + bt, \tag{2.15}$$

$$\lambda(t_k) = a + bt + ct^2 \qquad \text{или} \qquad \lambda(t_k) = ct^2 \quad . \tag{2.16}$$

С помощью моделей полезной составляющей процесса (2.15) и (2.16) можно, в ряде случаев, представить, участки функционирования промышленных объектов, работа, которых контролируется с помощью радиотехнических каналов [99].

В данной работе исследуются модели нестационарных и стационарных случайных процессов, которые представляет собой единственную реализацию дискретного процесса $\{Y_k\}_{k=1}^{N}$, полученного в равноотстоящие моменты времени $t_1, t_2, ..., t_N$, где $Y(t_k) = Y_k$ и $k = \overline{1, N}$, т.е. модели (2.2), (2.3).

Исследование для нестационарных случайных процессов проводятся, когда полезная составляющая процесса $\lambda(t_k)$ представлена простыми моделями функций: гармонической (2.13), экспоненциальной (2.14), полиномиальными (2.15) и (2.16) , а так же составной и сложной моделями вида:

$$\lambda(t_k) = \begin{cases} a + bt + ct^2, & 0 \le t \le t_2; \\ \sin(\omega t), & t_2 \le t \le t_3; \\ C, & t_3 \le t \le t_4; \\ \exp(-\alpha t), & t_4 \le t \le t_5. \end{cases} \tag{2.17}$$

и

$$\lambda(t_k) = C + A\sin(\omega t) \qquad (2.18)$$

Составная модель функции вида (2.17), состоящая из параболы, синусоиды, константы и экспоненты – модель огибающей радиоимпульса на выходе резонансного усилителя при расстройке относительно резонансной частоты. Модель сложной функции (2.18) представляет собой сумму некоторой константы и синусоиды.

Для оценки эффективности предлагаемых способов обнаружения аномальных значений при анализе нестационарных случайных процессов рассматривается аддитивная (2.2) и мультипликативная (2.3) модели процессов. Шумовая составляющая процесса η_k представленна гауссовским (2.6), равномерным (2.7) и рэлееевским (2.7) законами плотности распределения вероятности.

В качестве аномальной составляющей процесса рассматривалась модель вида (2.9), т.е. одиночного аномального значения ε_k с различной величиной Amp_k и местом расположения в выборке исследуемого нестационарного случайного процесса.

Критерием эффективности предлагаемых в данной работе способов обнаружения аномальных значений в реализации нестационарного случайного процесса используются оценки вероятности правильного обнаружения β и вероятности ошибки первого рода α. Вероятность ошибки первого рода α (вероятность ложной тревоги) определяет вероятность принятия значения процесса за аномальное значение. Вероятность правильного обнаружения β, определяет вероятность правильного решения о наличии аномального значения в исходной реализации нестационарного случайного процесса. Использование выше сказанного критерия для оценки эффективности предлагаемых в работе способов осуществляются по усредненным значениям, т.е. в качестве

оценок вероятности правильного обнаружения β и вероятности ошибки первого рода α рассмотрены их средние значения, полученные по множеству реализаций (порядка 1000).

2.2 Адаптивный способ обнаружения аномальных значений

В работах [54, 79], представлен метод выделения полезной составляющей нестационарного случайного процесса, который имеет высокую эффективность в условиях априорной неопределенности. Суть метода состоит в размножении не самой реализации исходного процесса, а оценок, получаемых определенным образом. Автор работ [54, 79], основываясь на основных принципах метода размножения оценок, предлагает и метод обнаружения аномальных значений при анализе нестационарных случайных процессов. В работах [54, 79, 100, 101 и др.] аналитически определены значения основных параметров метода обнаружения аномальных значений и показана его эффективность при анализе как стационарных, так и нестационарных случайных процессов с аддитивной шумовой составляющей.

К одному из достоинств метода обнаружения аномальных значений можно отнести так же следующее: применение двухпорогового критерия принятия решения об аномальности значения процесса позволяет получить результаты, при которых с увеличением величины аномальных значений, оценки вероятности ошибки первого рода α стремятся к минимальным значениям, в то время как оценки вероятности правильного обнаружения β стремятся к максимальным значениям [79, 100, 101].

Наряду с достоинствами предлагаемого метода обнаружения аномальных значений, представленного в работах [54, 79], выявлено, что он обладает весьма существенными недостатками, одним из которых является зависимость порогового значения от некоторого постоянной

коэффициента A. Правильный выбор коэффициента A позволит повысить эффективность обнаружения аномальных значений.

Поэтому в данной работе, на основе проведенных исследований, предлагается модификация уже существующего метода обнаружения аномальных значений, которая заключается в выборе правила определения коэффициента A при задании порогового значения (1.11).

Модификация предлагаемого в работе способа обнаружения аномальных значений предполагает введение адаптации порогового значения относительно коэффициента A при априорно фиксированном значении вероятности ошибки первого рода α.

Предлагаемый в данной работе способ предполагает наличие единственной дискретной реализации исследуемого нестационарного случайного процесса, Y_k, который представлен аддитивной моделью (2.2). Априорная информация об исследуемом процессе заключается в том, что на некоторых интервалах времени полезная составляющая процесса является гладкой функцией [101], т.е. достаточно точно описывается полиномом не выше второй степени:

$$\lambda_k = a + bt_k + ct_k^2 . \qquad (2.19)$$

Реализация исследуемого процесса разбивается на интервалы случайной длины, получаемые следующим образом: с помощью генератора случайных чисел, равномерно распределенных в интервале (0;1), получают $l-1$ чисел $\xi_1^{(1)}, \xi_2^{(1)}, ..., \xi_{l-1}^{(1)} \in (0;1)$. Используя выражение $y = t_1 + (t_N - t_1)x$, осуществляется взаимнооднозначное отображение промежутка (0;1) на интервале значений исследуемого нестационарного случайного процесса $(t_1; t_N)$, получая при этом соответствующее разбиение числами $\alpha_1^{(1)}, \alpha_2^{(1)}, ... \alpha_{l-1}^{(1)}$ промежутка $(t_1; t_N)$ на l непересекающихся интервалов, где

$$\alpha_j^{(1)} = t_1 + (t_N - t_1)\xi_j^{(1)}\ ,\ j = 1,2,...,l-1. \qquad (2.20)$$

Вводятся обозначения для интервалов разбиения:

$$\Delta_1^{(1)} = \left[t_1;\alpha_1^{(1)}\right],\ \Delta_2^{(1)} = \left(\alpha_1^{(1)};\alpha_2^{(1)}\right],\ \ldots,\ \Delta_l^{(1)} = \left(\alpha_{l-1}^{(1)};t_N\right]. \qquad (2.21)$$

Каждый интервал разбиения $\Delta_j^{(1)}$ $(j = 1,2,...,l)$ содержит не менее ∇ отсчетов (минимальная длина интервала разбиения) исходного нестационарного случайного процесса из набора $\{t_1, t_2,...,t_N\}$, в противном случае, случайные числа, формирующие данный интервал разбиения $(\xi_1^{(1)}, \xi_2^{(1)},...,\xi_{l-1}^{(1)})$, отбрасываются и генерируются заново. Наличие этого условия означает, что $l\nabla \le 0{,}6 \cdot N$.

Для получения каждой новой оценки процедура разбиения отрезка $(t_1;t_N)$ на l интервалов случайной длины (с проверкой выше указанного условия) повторяется. В результате получаем p разбиений временного отрезка $(t_1;t_N)$ [80] .

$$\begin{array}{cccc} \Delta_1^{(1)} & \Delta_2^{(1)} & \cdots & \Delta_l^{(1)} \\ \Delta_1^{(2)} & \Delta_1^{(2)} & \cdots & \Delta_l^{(2)} \\ \text{----------} & & & \\ \Delta_1^{(p)} & \Delta_2^{(p)} & \cdots & \Delta_l^{(p)} \end{array}. \qquad (2.22)$$

На каждом интервале разбиения $\Delta_j^{(i)}$, где $i = \overline{1,p}$ и $j = \overline{1,l}$ с помощью метода наименьших квадратов находятся оценки $\bar{a}_j^{(i)}$, $\bar{b}_j^{(i)}$, $\bar{c}_j^{(i)}$ коэффициентов аппроксимирующего полинома $a + bt_k + ct_k^2$ как решение системы линейных уравнений:

$$\begin{cases} aN + b\sum t_k + c\sum t_k^2 = \sum Y_k, \\ a\sum t_k + b\sum t_k^2 + c\sum t_k^3 = \sum t_k Y_k, \\ a\sum t_k^2 + b\sum t_k^3 + c\sum t_k^4 = \sum t_k^2 Y_k, \end{cases} \qquad (2.23)$$

где символ $\sum$ означает суммирование по всем k-ым значения, которые принадлежат интервалу разбиения $\Delta_j^{(i)}$ и $j=\overline{1,l}$, $i=\overline{1,p}$.

Результатом действия способа будет набор определенных на отрезке $(t_1;t_N)$ сглаживающих функций $\overline{\lambda}_{kj}^{(i)}$, где $i=\overline{1,p}$, каждая из которых является «кусочно-квадратичной»:

$$\overline{\lambda}_{kj}^{(i)}=\begin{cases}\overline{a}_1^{(i)}+\overline{b}_1^{(i)}t_k+\overline{c}_1^{(i)}t_k^2,\\ \overline{a}_2^{(i)}+\overline{b}_2^{(i)}t_k+\overline{c}_2^{(i)}t_k^2,\\ \dots\dots\dots\dots\dots\dots\dots\\ \overline{a}_l^{(i)}+\overline{b}_l^{(i)}t_k+\overline{c}_l^{(i)}t_k^2,\end{cases} \qquad (2.24)$$

где $t_k\in\Delta_l^{(i)}$, $i=\overline{1,p}$ и $j=\overline{1,l}$.

Определяются значения разности $S_{kj}^{(i)}$ между исходным нестационарным случайным процессом $Y_{kj}^{(i)}$ и оценкой сглаживающих функций $\overline{\lambda}_{kj}^{(i)}$:

$$S_{kj}^{(i)}=Y_{kj}^{(i)}-\overline{\lambda}_{kj}^{(i)}, \qquad k=\overline{1,N}. \qquad (2.25)$$

При оценке параметров разностного процесса $S_{kj}^{(i)}$, на каждом интервале разбиения $\Delta_j^{(i)}$, используется один из методов стабильногоого оценивания [5], т.е. оценка параметров математического ожидания $\overline{m}$ и среднеквадратического отклонения $\overline{\sigma}$ производится по α^*- усеченной выборке. Для этого на каждом интервале разбиения $\Delta_j^{(i)}$ получаем ряд ранжированных значений и оценку математического ожидания $\overline{m}$ и среднеквадратического отклонения $\overline{\sigma}$, которая проводится без учета первого и последнего значения ранжированного ряда. Тогда выражение

для оценок математического ожидания $\overline{m}_j^{(i)}$ и среднеквадратического отклонения $\overline{\sigma}_j^{(i)}$ принимают следующий вид [102]:

$$\overline{m}_j^{(i)} = \frac{1}{r}\sum_{k=2}^{r-1} S_{kj}^{(i)} \text{ и } \overline{\sigma}_j^{(i)} = \sqrt{\frac{1}{r-1}\sum_{k=2}^{r-1}(\overline{m}_j^{(i)} - S_{kj}^{(i)})^2}, \tag{2.26}$$

где $r = \nabla - 2$.

Далее на каждом интервале разбиения исследуемого нестационарного случайного процесса, устанавливается пороговое значение

$$\varepsilon_j^{(i)} = A \cdot \overline{\sigma}_j^{(i)}, \tag{2.27}$$

где A – некоторый коэффициент, $0 < A \le 3$ и $j = \overline{1,l}$. Превышение значений разностного процесса $S_{kj}^{(i)}$, на каждом интервале разбиения, установленного порогового значения (2.27), штрафуется, т.е. если выполняется условие:

$$S_{kj}^{(i)} + \varepsilon_j^{(i)} < Y_{kj}^{(i)} < S_{kj}^{(i)} - \varepsilon_j^{(i)}, \tag{2.28}$$

то $Y_{kj}^{(i)}$ получает одно штрафное значение. В соответствии с методом размножения оценок [79], вышеизложенная процедура определения штрафов, повторяется p- раз и для каждого повторения проверяется условие (2.28) для каждого значения $Y_{kj}^{(i)}$, где $i = \overline{1,p}$; $k = \overline{1,N}$; $j = \overline{1,l}$; N – объем выборки исследуемого нестационарного случайного процесса; l – количество интервалов разбиения; p- количество повторений процедур (2.22) – (2.28).

Таким образом, происходит накопление ряда штрафных значений для элементов исходной реализации исследуемого процесса, т.е.:

$$g_k^{(i)} = \begin{cases} 1, \left| S_{kj}^{(i)} \pm \varepsilon_j^{(i)} \right| > Y_{kj}^{(i)}, \\ 0, \left| S_{kj}^{(i)} \pm \varepsilon_j^{(i)} \right| < Y_{kj}^{(i)} \end{cases}, \tag{2.29}$$

где $g_k^{(i)}$ – ряд штрафных значений $k = \overline{1, N}$ и $i = \overline{1, p}$, $j = \overline{1, l}$.

По окончанию обработки для всех оштрафованных значений исходной реализации определяется суммарное значение штрафов $g_k^* = \sum_{i=1}^{p} g_k^{(i)}$ и максимальное значение ряда $\max(g_k^*) = G$. Далее проверяется условие: если

$$g_k^* > 0{,}7G, \tag{2.30}$$

то k-ое значение из входной реализации нестационарного случайного процесса Y_k будет трактоваться как аномальное. Условие (2.30) получено на основе проведения имитационного моделирования при различных моделях полезной и шумовой составляющих нестационарного случайного процесса.

Основные этапы работы адаптивного способа обнаружения аномальных значений отображаются на рис. 2.1 – 2.3, для случая нестационарного случайного процесса с аддитивной шумовой составляющей $N(0;\sigma_ш)$ и равномерно расположенных аномальных значений по всей исследуемой выборке объемом N.

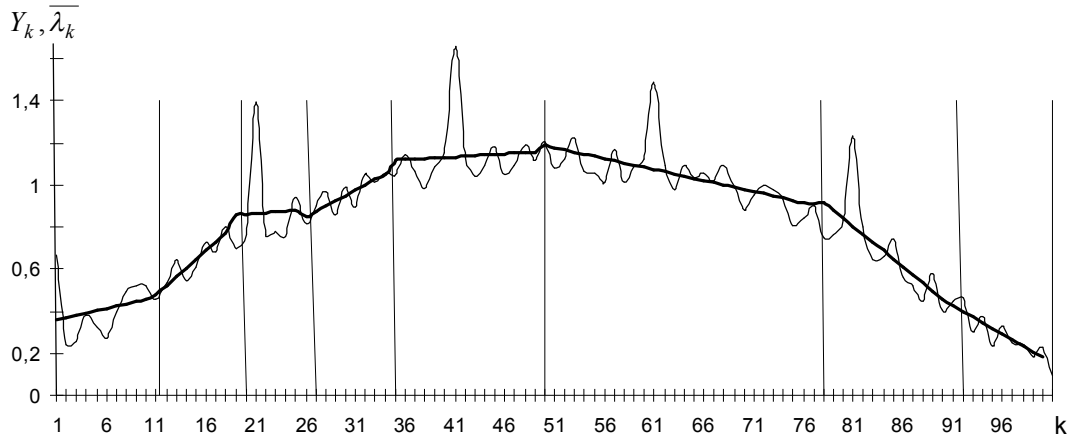
На рисунке 2.1а представлена реализация нестационарного случайного процесса Y_k и сглаживающей функции $\overline{\lambda}_k$, которые получены на каждом интервале разбиения $\Delta_j^{(i)}$.

Границы интервалов разбиений обозначены вертикальными линиями. На рисунке 2.1б, для разностного процесса S_k, на каждом интервале разбиения $\Delta_j^{(i)}$ устанавливаются пороговые значения $\varepsilon_j^{(i)}$, обозначенные горизонтальными линиями.

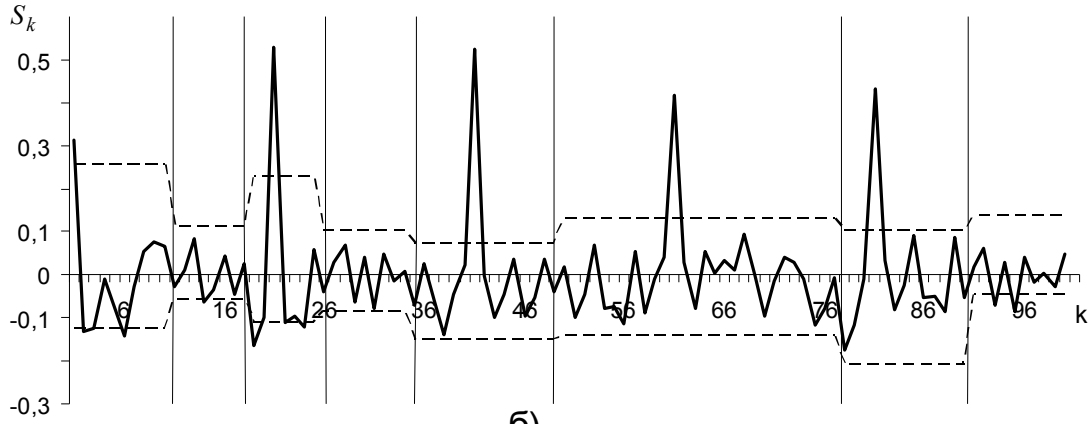
Пороговые значения определяются в соответствии с выражением (2.27) и с использованием уравнения (2.31). На рисунке 2.1в представлен ряд штрафных значений g_k, полученный при превышении каждого значения разностного процесса S_k порогового значения $\varepsilon_j^{(i)}$, т.е. при выполнении условия (2.27) данная процедура повторяется p раз.

На рисунек 2.2а и 2.2б, в качестве примера, представлен результат четвертого разбиения на интервалы ($p = 4$). Рисунке 2.2в показывает как происходит накопление ряда штрафных значений g_k на каждом этапе работы предложенного способа обнаружения аномальных значений. На рисунке 2.3а представлен результирующий ряд штрафов и второе пороговое значение g_k^*, обозначенное вертикальной линией, которое определяется относительно максимального значения ряда штрафов g_k. На рисунке 2.3б – ряд обнаруженных аномальных значений и на рисунке 2.3в представлен исходный нестационарный случайный процесс Y_k и обнаруженные аномальные значения.

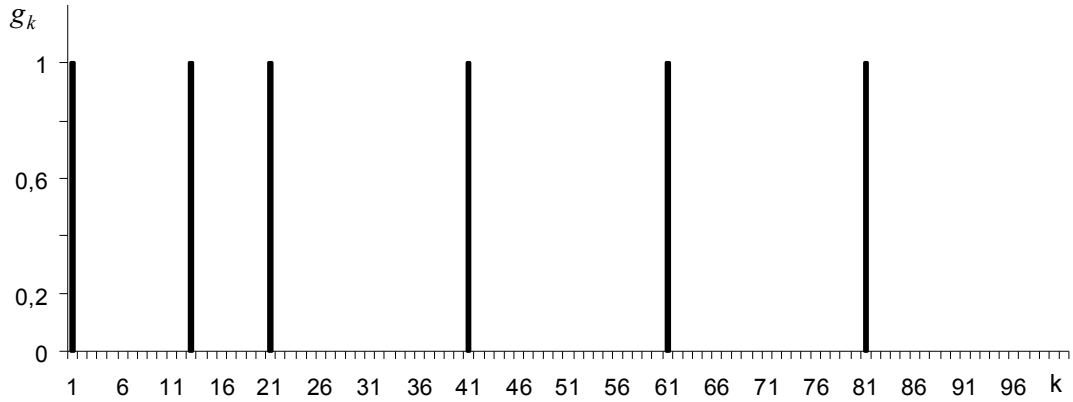
Для случая обнаружения аномальных значений в реализации нестационарного случайного процесса выбирать значение коэффициента A только по оценкам среднеквадратического отклонения шумовой составляющей процесса является нецелесообразным, так как наличие аномальных значений существенно влияет на погрешность оценки полезной составляющей процесса и, как следствие, на оценку среднеквадратического отклонения разностного процесса. Следует так же отметить, что на каждом интервале разбиения значение коэффициента A не может быть фиксированным.



а)

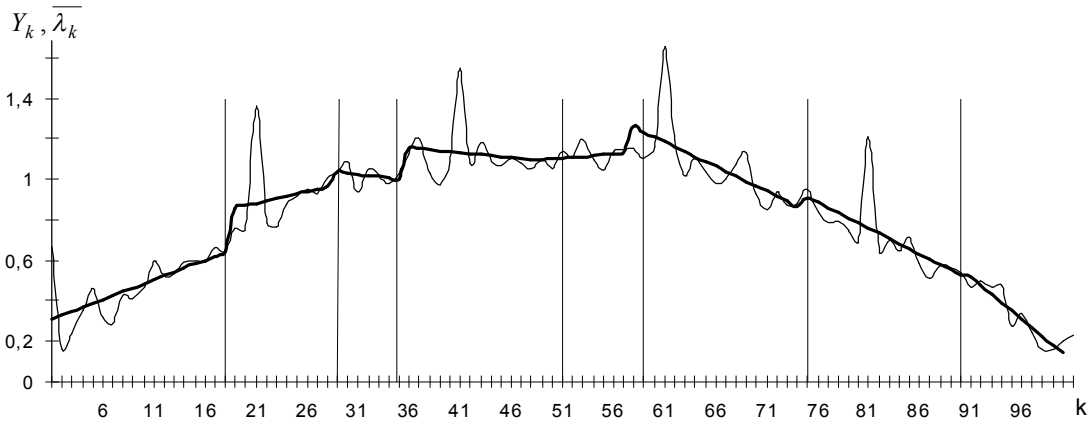


б)

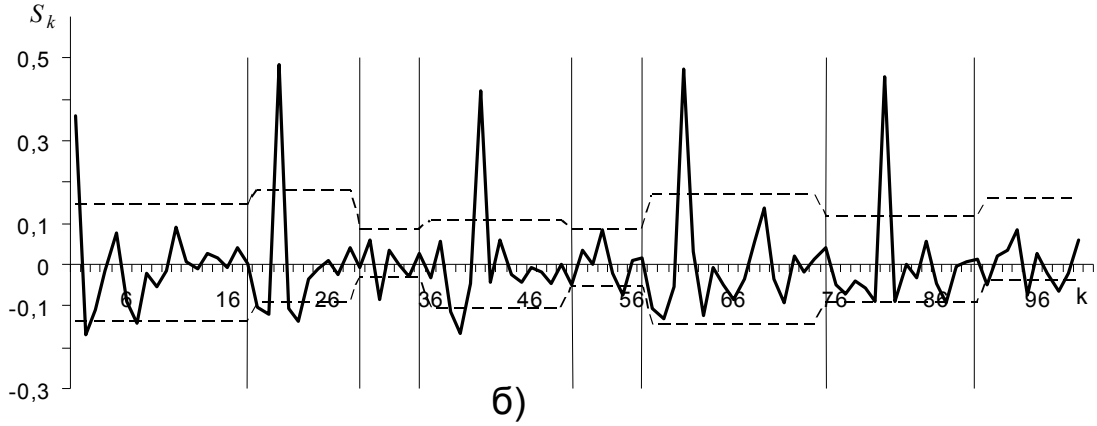


в)

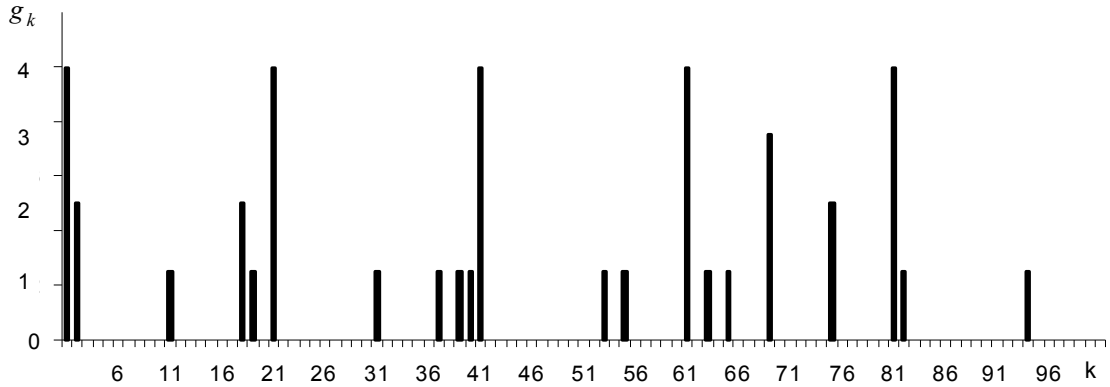
Рис. 2.1. Первый этап адаптивного способа обнаружения аномальных значений:
а) разбиение исходного процесса на интервалы;
б) разностный процесс и пороговые значения на каждом интервале разбиения;
в) значения штрафов при превышении порогового значения



а)



б)



в)

Рис. 2.2. Четвертый этап адаптивного способа обнаружения аномальных значений:
а) разбиение исходного процесса на интервалы;
б) разностный процесс и пороговые значения на каждом интервале разбиения;
в) значения штрафов при превышении порогового значения

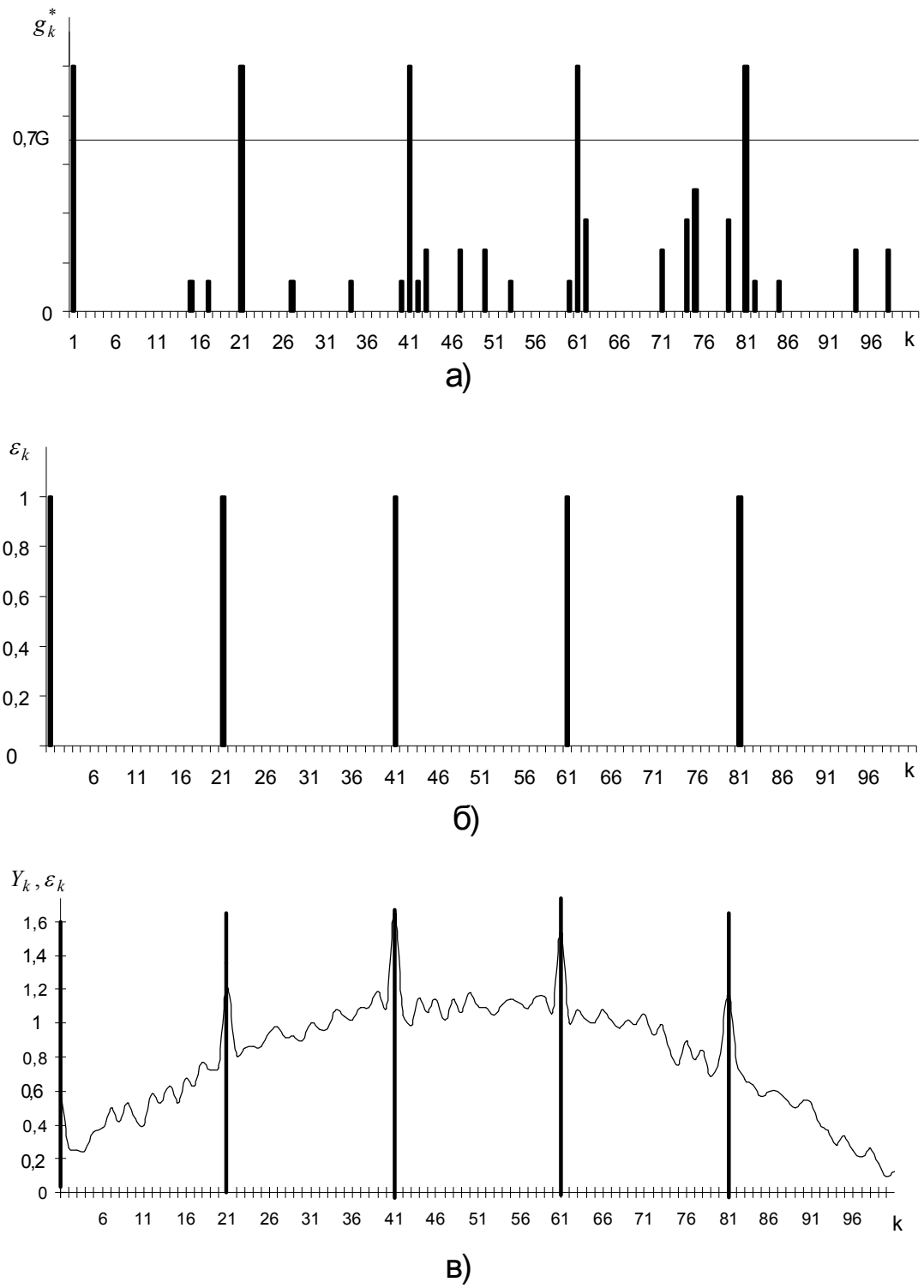


Рис. 2.3. Итоговый этап адаптивного способа обнаружения аномальных значений:
а) результирующие значения штрафов;
б) значения, превысившие второй порог;
в) обнаруженные аномальные значения

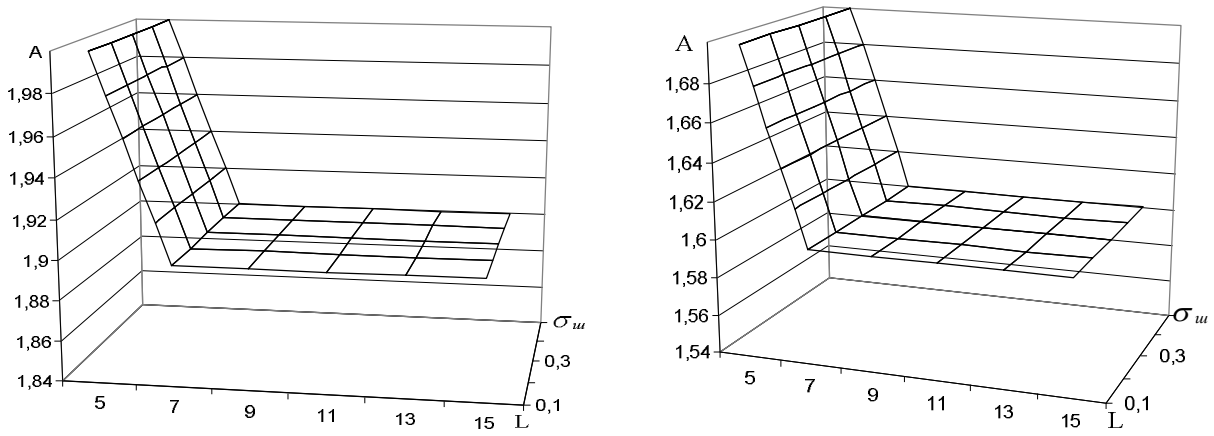
В связи с этим предлагается ввести адаптацию порогового значения о назначении штрафов (2.27) по коэффициенту A относительно априорно фиксированного значения вероятности ошибки первого рода. С этой целью проведены исследования зависимости коэффициента A от объема выборки L, от значения среднеквадратического отклонения случайного процесса $\sigma_ш$ для различных стационарных процессов при априорно фиксированных значениях вероятности ошибки первого рода α.

В результате получены зависимости оценок коэффициента A от объема исследуемой выборки и среднеквадратического отклонения шумовой составляющей процесса, т.е. $A = f(L, \sigma_ш)$. Входная реализация представляет собой стационарный центрированный гауссовский случайный процесс. Исследования проводились на выборках объёмом L= 5, 7, 9, 11, 13 и 15 значений и среднеквадратическом отклонении случайного процесса $\sigma_ш = 0{,}1 - 0{,}5$. В результате проведенных исследований были получены зависимости оценок коэффициента A при различных априорно фиксированных значениях вероятности ошибки первого рода α. Усреднение значений коэффициента A производилось по 1000 выборок [103, 104, 105].

На рисунке 2.4 приведены графики зависимости оценок коэффициента A для объёма выборки L и среднеквадратического отклонения случайного процесса $\sigma_ш$ при априорно фиксированных значениях вероятности ошибки первого рода α (при α=0,05– рис. а) при α=0,1– рис. б)).

Из анализа полученных зависимостей, представленных на рисунке 2.4 следует, что при различных фиксированных значениях вероятности ошибки первого рода α с увеличением объема выборки L оценки коэффициента A стремится к некоторому постоянному значению и практически не зависит от значения среднеквадратического отклонения

случайного процесса. Оценки коэффициента A для выборок $L < 5$ возрастают в среднем на 5%.



а) б)

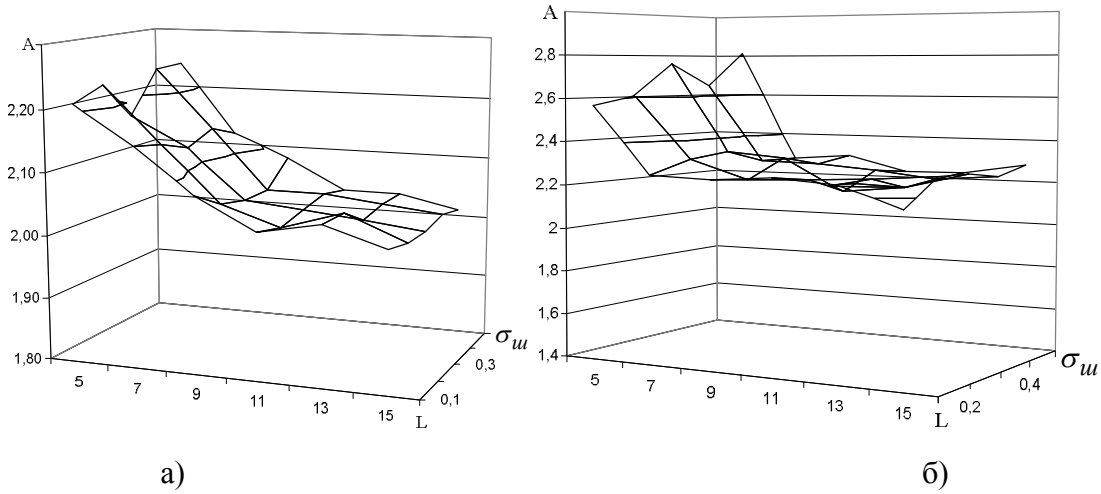
Рис. 2.4. Зависимость $A = f(L, \sigma_ш)$ для гауссовского закона плотности распределения вероятности случайного процесса: а) при $\alpha = 0{,}05$ и б) при $\alpha = 0{,}1$

Так же приведены результаты исследований зависимости коэффициента A от объема выборки L и среднеквадратического отклонения $\sigma_ш$, когда стационарный случайный процесс представлен равномерным и рэлеевским законами плотности распределения вероятности. Результаты полученных зависимостей представлены на рисунке 2.5а – для равномерного и на рисунке 2.5б – для рэлеевского законов плотности распределения вероятности случайного процесса, при априорно фиксированном значении ошибки первого рода α=0,05.

Из анализа графиков, представленных на рисунке 2.5а и 2.5б, видно, что оценки коэффициента A практически не зависит от среднеквадратического отклонения стационарного случайного процесса $\sigma_ш$ и незначительно зависит от объема исследуемой выборки L.

Таким образом, проведенные исследования показывают, что значения оценок коэффициента A для рассмотренных законов распределения случайных процессов практически не зависят от объема исследуемой выборки L и среднеквадратического отклонения

стационарного случайного процесса $\sigma_ш$, а зависят только от априорно задаваемого значения вероятности ошибки первого рода α [103, 104, 105].



а) б)

Рис. 2.5. Зависимость $A = f(L, \sigma_ш)$ при $\alpha = 0{,}05$: а) для равномерного закона, б) для рэлеевского законов плотности распределения вероятности случайных процессов

В связи с этим исследуются зависимости оценок коэффициента A от априорно фиксированного значения вероятности ошибки первого рода α, т.е. $A = f(\alpha)$, для различных законов плотности распределения вероятности стационарных случайных процессов (гауссовского, равномерного, рэлеевского) при значении $\sigma_ш = 0{,}1$. Результаты, которые представлены на рисунках 2.6 и на рис. 2.7 получены при $\alpha = 0{,}05$, $\alpha = 0{,}1$ и $\alpha = 0{,}2$.

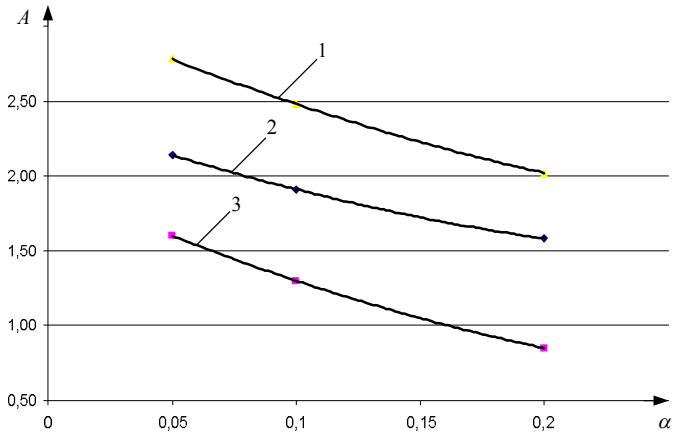


Рис. 2.6. Зависимость $A = f(\alpha)$ для различных законов распределения случайных процессов

На рисунке 2.6 приведены зависимости значения оценок коэффициента A от вероятности ошибки первого рода α: график 1 – рэлеевский; график 2 – равномерный и график 3 –гауссовский законы плотности распределения вероятности стационарного случайного процесса.

Из анализа графиков, представленных на рисунке 2.6, следует, что оценки коэффициента A для всех представленных законов распределения стационарного случайного процесса, существенно зависят от априорно задаваемых значений вероятности ошибки первого рода α.

На рисунке 2.7 представлен усредненный график зависимости $A = f(\alpha)$ для исследуемых стационарных случайных процессов.

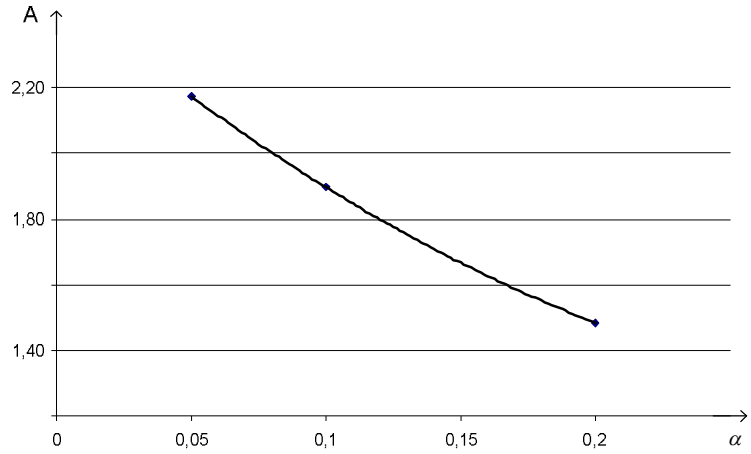


Рис. 2.7. Усредненная зависимость $A = f(\alpha)$ для рассмотренных стационарных случайных процессов

Графическая зависимость, представленная на рисунке 2.7, может быть аппроксимирована полиномом второй степени вида [106, 107]:

$$A(\alpha) = 2{,}16 - 6{,}3\alpha + 7{,}32\alpha^2 . \qquad (2.30)$$

Полученные результаты исследования зависимости коэффициента $A = f(L, \sigma_ш, \alpha)$ позволяют при адаптации порогового значения (2.27), вместо постоянного значения коэффициента A использовать его значение, которое вычисляется в соответствии с (2.30). Использование уравнения (2.30) в оценке порогового значения (2.27) позволяет использовать

предложенный способ обнаружения аномальных значений при фиксированном значении вероятности ошибки первого рода α.

Блок-схема программы [108], реализующая алгоритм адаптивного способа обнаружения аномальных значений при априорно фиксированном значении вероятности ошибки первого рода, представлена на рисунке 2.8. Принцип её работы заключается в следующем:

– запоминается входная реализация значений входного процесса $Y(t_k) = Y_k$, где $t_k = k \times T$ и $k = \overline{1, N}$, T - шаг дискретизации, N - объем входной реализации;

– задаются значения вероятности ошибки первого рода α, минимальный интервал разбиения ∇, количество разбиений l и p;

– определяются оценки $\overline{\lambda}_{kj}^{(i)}$, за счет аппроксимации $Y_{kj}^{(i)}$ на каждом из интервалов разбиения $\Delta_j^{(i)}$ полиномом не выше второй степени, где $k = \overline{1, N}$ и $j = \overline{1, l}$;

– находятся разности $S_{kj}^{(i)} = Y_{kj}^{(i)} - \overline{\lambda}_{kj}^{(i)}$, где $k = \overline{1, N}$ и $j = \overline{1, l}$;

– определяются оценки среднеквадратического отклонения $\overline{\sigma}_j^{(i)}$ для $S_{kj}^{(i)}$ на каждом из интервалов разбиения, $\Delta_j^{(i)}$ и задается пороговое значение $\varepsilon_j^{(i)} = f(\alpha, \overline{\sigma}_j^{(i)})$, где $j = \overline{1, l}$;

– каждое значение входной реализации $Y_{kj}^{(i)}$ сравнивается с пороговым значением $\varepsilon_j^{(i)} = f(\alpha, \overline{\sigma}_j^{(i)})$ на каждом из интервалов разбиения $\Delta_j^{(i)}$ и, если $Y_{kj}^{(i)} \geq \varepsilon_j^{(i)}$, то происходит накопление штрафных значений $g_k^{(i)} = g_k^{(i)} + 1$, где $g_k^{(i)}$ - ряд штрафных значений, $k = \overline{1, N}$, $j = \overline{1, l}$ и $i = \overline{1, p}$;

– данная процедура повторяется p-раз, в результате чего получается суммарный ряд штрафных значений $g_k^* = \sum_{i=1}^{p} g_k^{(i)}$, где $k = \overline{1, N}$ и $i = \overline{1, p}$;

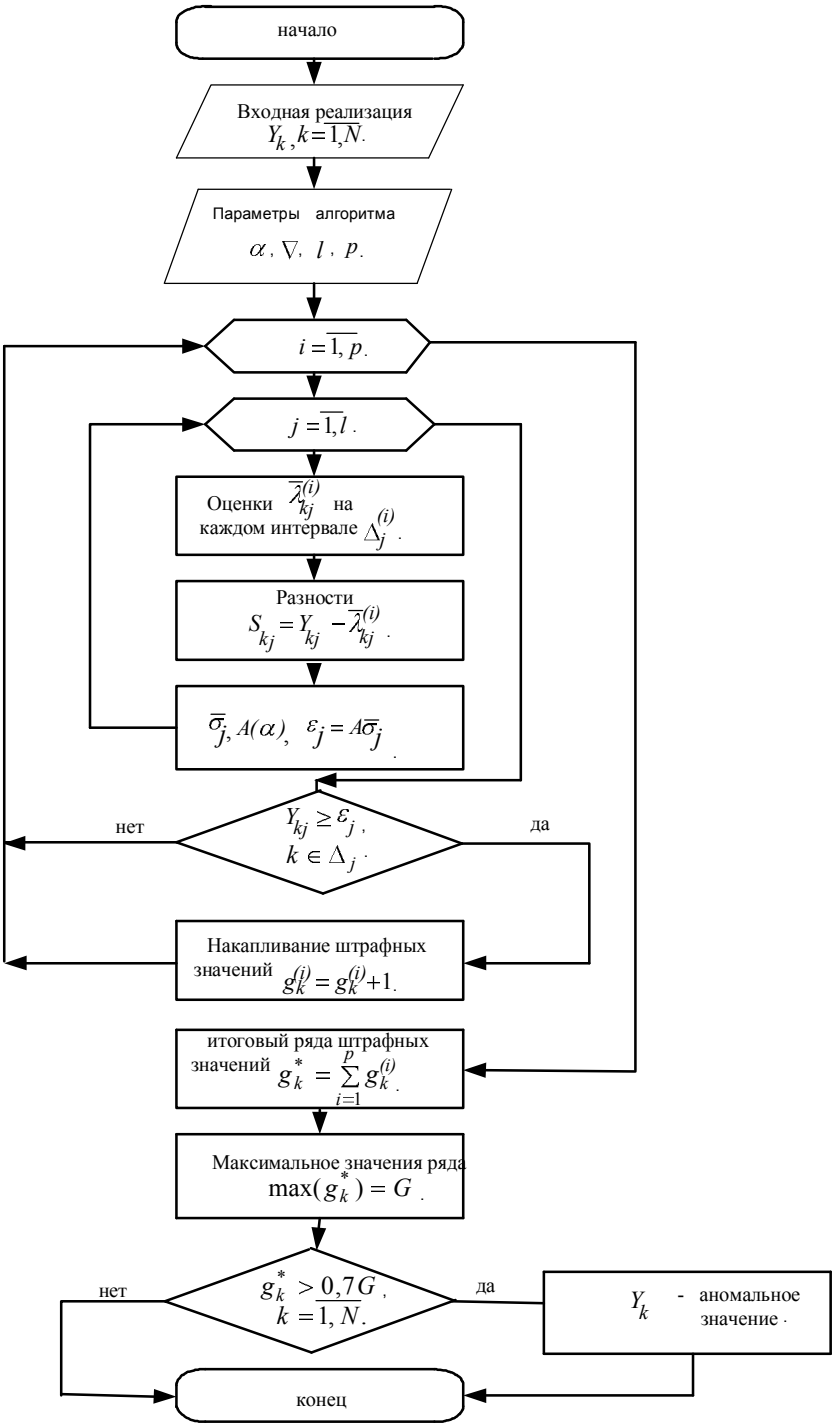


Рис. 2.8. Блок-схема алгоритма реализующего адаптивный способ обнаружения аномальных значений

– определяется максимальное значение ряда $\max(g_k^*)=G$, где $k=\overline{1,N}$;

– проверяется условие: если $g_k^*>0{,}7G$, то k-ое значение из входной реализации Y_k, является аномальным значением.

2.3 Структурная схема устройства, реализующего адаптивный способ обнаружения аномальных значений

В работе предлагается структурная схема устройства, реализующего адаптивный способ обнаружения аномальных значений [109], которая представлена на рисунке 2.9. Устройство являются многоканальным, количество каналов устройства p определяется числом повторений процедур (2.22) – (2.28).

Значения исходного процесса Y_k предварительно записано в буферный каскад. Значения обрабатываемой реализации Y_k с буферного каскада (Б) поступают в блоки разбиения на интервалы случайной длины (БРИ), значения которых поступают из блока формирования интервалов разбиения (БФИ). В блоке аппроксимации (БА) вычисляются значения аппроксимирующей функции $\overline{\lambda}_k$ на полученных интервалах разбиения с помощью метода наименьших квадратов. В блоке нахождения разностного процесса (БР) определяется разность между значениями исходного процесса Y_k и оценками процесса, полученными в результате аппроксимации $\overline{\lambda}_k$ на интервалах разбиения, т.е. S_k. С выхода блока оценки параметров математического ожидания и среднеквадратического отклонения (БОП) разностного процесса S_k поступают в блок расчета порогового значения (П1), где умножаются на значения коэффициента A. Значения коэффициента A определяется в блоке задания коэффициента (БЗК) в соответствии с априорно заданным значением вероятности ошибки первого рода α. В коммутаторе (К1) значениям исходного процесса Y_k, которые превысили пороговое значение, присваиваются штрафы. В блоке суммирования штрафов (БС) суммируются результаты штрафования с каждого канала p и определяется максимальное значение штрафов, на

основании которого в блоке П2 формируется второе пороговое значение. С помощью коммутатора (К2) определяются значения, которые превышают второе пороговое значение и которые считаются аномальными. В блоке (БУ) аномальные значения устраняются за счет интерполяции соседних от них значений исходного процесса Y_k. Синхронность работы устройства обнаружения аномальных значений осуществляется от тактового генератора импульсов (генератора импульсов синхронизации).

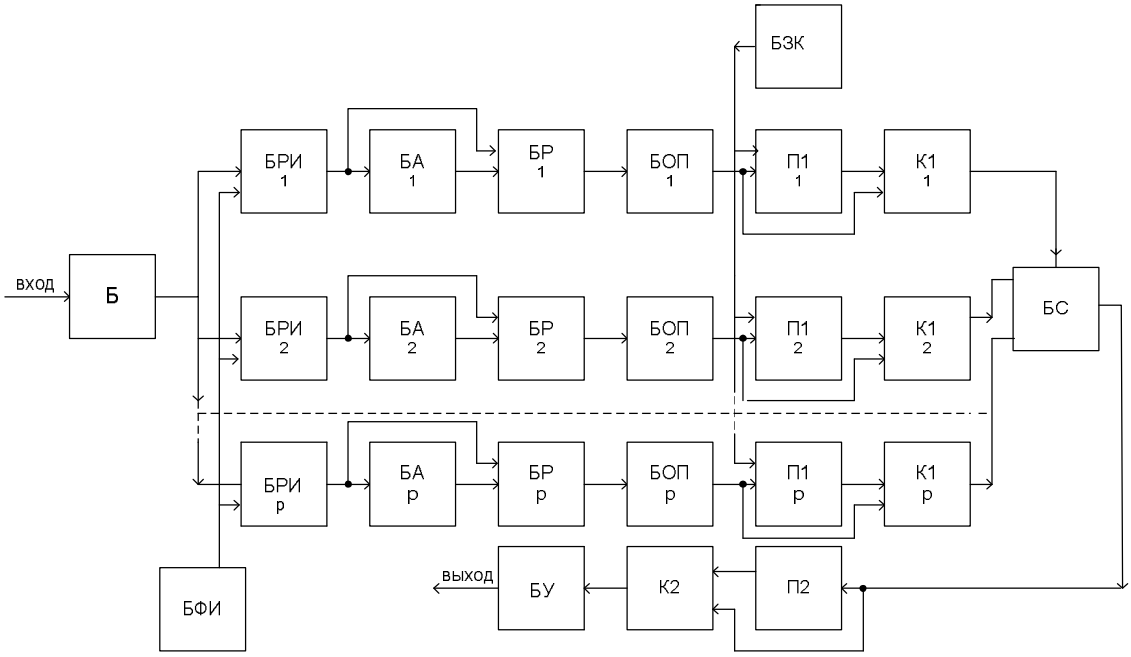


Рис. 2.9. Структурная схема устройства, реализующего адаптивный способ обнаружения аномальных значений

Приведенная структурная схема реализации устройства обнаружения аномальных значений с адаптацией порового значения [109] может быть реализована как в аппаратурном, так и аппаратурно – программном исполнении.

2.4 Способ обнаружения аномальных значений с дискретно-меняющимся уровнем анализа процесса

Анализ литературных источников показал, что для обнаружения аномальных значений в реализации нестационарного случайного процесса используются методы и алгоритмы, основанные на получении оценки функции полезной составляющей процесса с последующим ее вычитанием из реализации исследуемого процесса и установки порогового значения. Обнаружение аномальных значений, основанное на предварительной оценке полезной составляющей процесса, снижает эффективность обнаружения аномальных значений из-за погрешности в оценке полезной составляющей процесса, которая, в свою очередь, зависит от наличия аномальных значений. Таким образом, любому анализу нестационарного случайного процесса, построенному на методе наименьших квадратов, должен предшествовать анализ на содержание аномальных значений.

В связи с этим в данной работе предлагается способ обнаружения аномальных значений без оценки полезной составляющей процесса нестационарного случайного процесса, основанный на обнаружении аномальных значений при использовании дискретно-меняющегося уровня анализа нестационарного случайного процесса. Значения процесса, превышающие уровень анализа штрафуются в том случае, если при этом не выполняется критерий наличия нестационарности на данном уровне анализа.

На практике инженерного проектирования обнаружителей сигналов часто используют эвристические алгоритмы, фиксирующие наличие сигнала по плотности единиц в пределах каждого интервала времени. Наиболее простым и распространенными из этого класса алгоритмов являются так называемые программные обнаружители, фиксирующие начало сигнала по наличию m единиц на d позициях [39, 110 -113].

В работе [111] приведены результаты эффективности цифрового программного обнаружителя, которые часто применяются на практике и работают в соответствии с алгоритмом: начало сигнала фиксируется по наличию m единиц в d позициях (или m единиц в m позициях), конец пачки – по наличию l нулей подряд. Здесь d является длиной цифрового программного обнаружителя.

При этом, начало критерий начала сигнала является одновременно и критерием его обнаружения [111]. Как показал результат численного анализа, который приведен в работах [112, 113], при фиксированном значении ошибки первого рода α выбирались те значения m, при которых вероятность правильного обнаружения β была максимальной, т.е. выбор параметра m определяется на основе критерия Неймана – Пирсона. И в условиях воздействия некоррелированных помех количество подряд следующих единиц, согласно приведенной методике в работе [111], при $l = 2$, определяется из выражения $m = 1{,}5\sqrt{N/2}$, где N – объем выборки нестационарного случайного процесса. Таким образом, если исследовать выборку нестационарного случайного процесса $N = 50$, то $m \approx 7$, а $d \approx 8$ при $l = 2$ [111].

Однако, как указывает автор в работах [110, 111] данная методика применима, только для выборок с $N = 10{,}30{,}50$ значений, «поскольку эффективность обнаружения падает с ростом числа значений в пачке», что является недостатком данного алгоритма.

Кузьмин С.З. в своей работе [39], пишет, что на практике широко используются схемы логических обнаружителей с критериями обнаружения типа $m/d - l$, где m - число подряд следующих единиц, l - число пропусков подряд а d - длина логического обнаружителя, причем $m \le d$ и $d \le 5$. Выбор подходящего критерия срабатыванию логического обнаружителя производится путем сравнения различных вариантов по

вероятности правильного обнаружения β, при фиксированном значении вероятности ошибки первого рода α. Так же приведен сравнительный анализ эффективности таких обнаружителей, который показывает, что наилучшей, с точки зрения вероятности правильного обнаружения β, является логика $3/4-1$ при и фиксированном значении вероятности ошибки первого рода $\alpha = 2 \cdot 10^{-4}$. Позднее в работе [114], Кузьминым С.З. представлены в качестве примера зависимости вероятности правильного обнаружения β программными обнаружителями $3/3-2$, $3/4-2$, $4/5-2$ при фиксированном значении ошибки первого рода $\alpha = 10^{-4}$, которые представлены для случаев обнаружения нефлюктуирующих и быстрофлюктуирующих сигналов. Анализ показывает, что для обоих типов сигналов наиболее эффективные оценки вероятности правильного обнаружения β у программных обнаружителей $3/4-2$ и $3/5-2$. Причем эти логики имеют близкие качественные характеристики по оценкам значений вероятности правильного обнаружения β [114] .

При аппаратурном определении плотности распределения вероятностей стационарных случайных процессов Мирский Г.Я в работе [6] предлагает выбрать число уровней анализа L, исходя из соотношения:

$$L = E(\frac{Y_{\max}}{h}) + 1, \qquad (2.31)$$

где E – целая часть дроби $Y_{\max}/h$; $Y_{\max}$ – пиковое значение входного процесса; h - ширина дифференциального коридора.

Формула (2.31) дает число уровней анализа в положительной области, т.е. когда $0 \le Y \le Y_{\max}$. Если плотность распределения является симметричной, как например, в случае гауссовского закона распределения, количество уровней, вычисленное по формуле (2.31) удваивается.

В своей работе Мирский Г.Я [6] вводит обозначение нормированной ширины дифференциального коридора $h_н = \frac{h}{\sigma}$, где σ – среднеквадратическое отклонение стационарного случайного процесса. Ширину дифференциального коридора $h_н$, предлагается выбирать, учитывая два вида относительных погрешностей измерений: систематическую γ и случайную δ. Первая погрешность γ - это относительная величина смещения оценки, а вторая δ - корень квадратный из относительной дисперсии оценки плотности вероятности. Приведенные расчеты, показывают, что зависимости погрешностей γ и δ от нормированной ширины коридора $h_н$ взаимно противоположны, однако в случае объёма выборки $N \approx 10^6 - 10^7$, γ и δ соизмеримы. При этом возникает задача оптимальной (в смысле минимума суммарной среднеквадратической погрешности) ширины дифференциального коридора. Расчет, оптимального значения для гауссовского закона распределения показал [6]:

$$(h_н)_{opt} = \frac{3{,}5}{\sqrt[5]{N}} \text{ или } \quad (h)_{opt} = \frac{3{,}5\sigma}{\sqrt[5]{N}}.$$

Для рэлеевского закона распределения соответственно [6]:

$$(h_н)_{opt} = \frac{2{,}5}{\sqrt[5]{N}} \text{ или } (h)_{opt} = \frac{2{,}5\sigma}{\sqrt[5]{N}}.$$

Однако данная методика справедлива для реализации стационарных эргодических случайных процессов при объеме выборки исследуемого процесса $N \geq 10^6$.

Рекомендации по выбору количества уровней анализа L для малых объёмов выборки гауссовского случайного процесса $N = 10^1 - 10^4$ в своей работе представляет Жовинский А.Н. [114] , ссылаясь так же на других авторов [115, 116]. В работах [74, 75] имеются ссылки на различные

исследования, устанавливающие оптимальное значение L в зависимости от объема выборки N. Согласно рекомендациям, которые представлены в работе Ершова А.А. [5], зависимость $L = f(N)$ может быть представлена в форме:

$$L = k \cdot \lg N, \tag{2.32}$$

где постоянная k выбирается в диапазоне от 3 до 5.

Автор [115] рекомендует при двухстороннем распределении, при $N = 100$, количество уровней анализа $L = 20$. А шаг изменения уровней анализа h определяется как отношение размаха выборки N к значению уровней анализа, т.е. $h = \frac{Y_{\max} - Y_{\min}}{L}$. Однако в работах [5, 115, 116] рассмотрен только случай когда закон распределения случайного процесса является гауссовским.

В качестве критерия обнаружения сигнала в данной работе используется эвристический метод, который функционирует в соответствии со следующим алгоритмом: начало сигнала фиксируется по наличию m единиц в d позициях (логика m/d, где m – подряд следующие единицы в d позициях) или d единиц в d позициях (логика d/d), конец сигнала по невыполнению этого условия [117]. При этом критерий начала сигнала является одновременно и критерием её обнаружения. Этот критерий широко используется на практике при построении цифровых программных обнаружителей и накопителей сигналов [39, 110–113].

Предлагаемый в данной работе способ обнаружения аномальных значений процесса заключается в следующем: определяется максимальное $Y_{\max} = \max(\{Y_k\})$ и минимальное $Y_{\min} = \min(\{Y_k\})$ значения исходного процесса $\{Y_k\}_{k=1}^{N}$ и рассчитывается шаг изменения уровня анализа по выражению [118, 119, 120]:

$$\Delta Y = \frac{(Y_{\max} - Y_{\min})}{P},$$

где P - априорно заданное количество уровней.

Значения уровней анализа вычисляется в соответствии с выражением

$$L_i = Y_{\max} - \Delta Y \cdot i \text{, где } i = \overline{0, P}$$

Для каждого значения уровня L_i проверяется условие:

$$Y_k \geq L_i \text{, где } k = \overline{1, N}, \quad i = \overline{0, P} \ . \tag{2.33}$$

При выполнении условия (2.33) на каждом шаге k-му отсчету ставится в соответствие штрафное значение, равное $g_k = 1$, $k = \overline{1, N}$. Таким образом, с изменением уровня анализа происходит накопление штрафных значений для элементов исходного процесса, т.е:

$$g_k = g_k + 1, \tag{2.34}$$

где $\{g_k\}_{k=1}^{N}$, ряд штрафных значений с соответствующим номером элемента выборки.

Одновременно, после установки каждого уровня L_i и проверки условия (2.33), исследуемый процесс анализируется на наличие полезной составляющей, т.е. значения входного процесса, превышающие данный уровень L_i анализируются согласно априорно задаваемой логике m/d.

Найденные участки локализуются, то есть, если в ряде $\{g_k\}_{k=1}^{N}$ появится последовательность в d ненулевых элементов, $\{g_k\}_{k=j}^{r} > 0$, $\{g_k\}_{k=j}^{r} \in \{g_k\}_{k=1}^{N}$, $d = r - j + 1$, где d - априорно задаваемая значение, то предполагается, что на отрезке $[j, r]$ ряда $\{Y_k\}_{k=1}^{N}$ присутствует полезная составляющая исходного процесса Y_k, и на этом интервале при дальнейшем изменении уровня анализа L_i штрафы не назначаются.

В результате обработки, ряд значений $\{g_k\}_{k=1}^{N}$, представляет собой ряд штрафов для шумовой η_k и аномальной ε_k составляющих исходного процесса Y_k. При этом, полезная составляющая λ_k исключается из рассмотрения. Штрафные значения, при наличии нестационарности не назначаются.

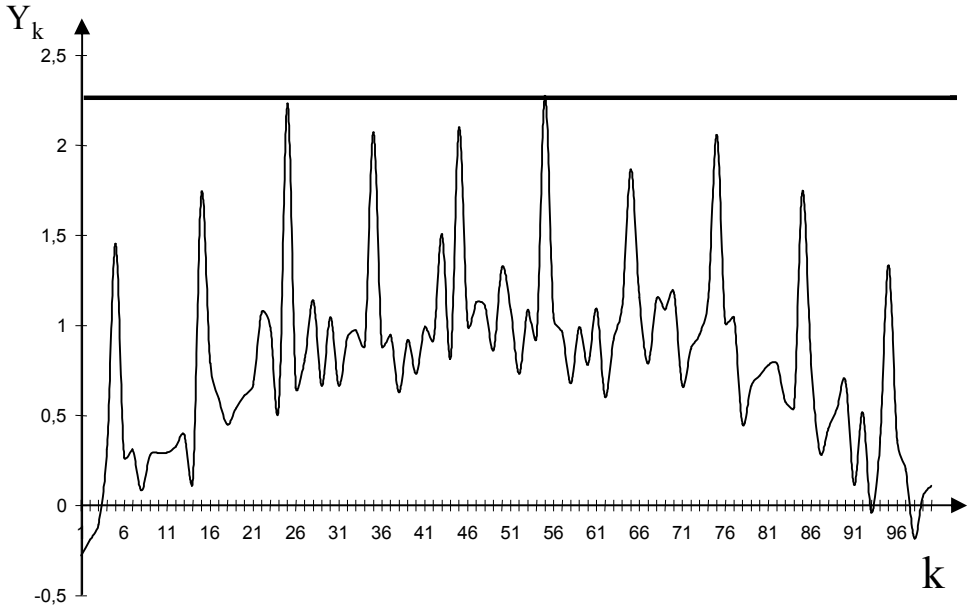
В ряду штрафных значений $\{g_k\}_{k=1}^{N}$ определяется максимальное значение ряда $g_{\max} = \max(\{g_k\})$ и проверяется условие:

$$g_k > 0{,}5 \cdot g_{\max}, \qquad k = \overline{1, N}. \tag{2.35}$$

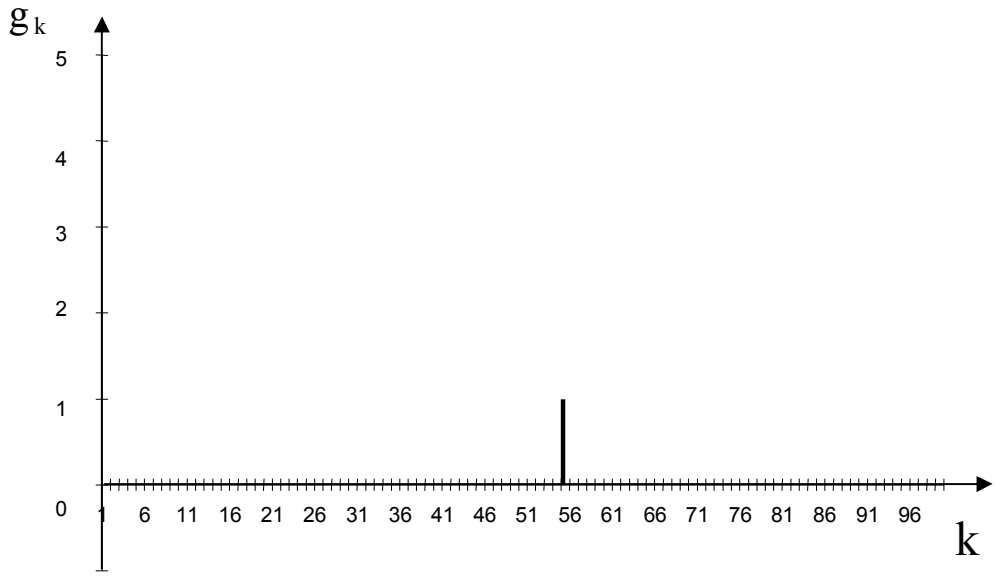
При выполнении условия (2.35) принимается решение, что соответствующее значение Y_k является аномальным. Условие (2.35) получено на основе проведения имитационного моделирования при различных моделях полезной и шумовой составляющих нестационарного случайного процесса [118].

Принцип работы способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа процесса наглядно отображается на рисунках 2.10 – 2.12.

На рисунке 2.10а изображен исходный нестационарный случайный процесс, представленный моделью (2.2), а горизонтальной линией – первый уровень анализа.

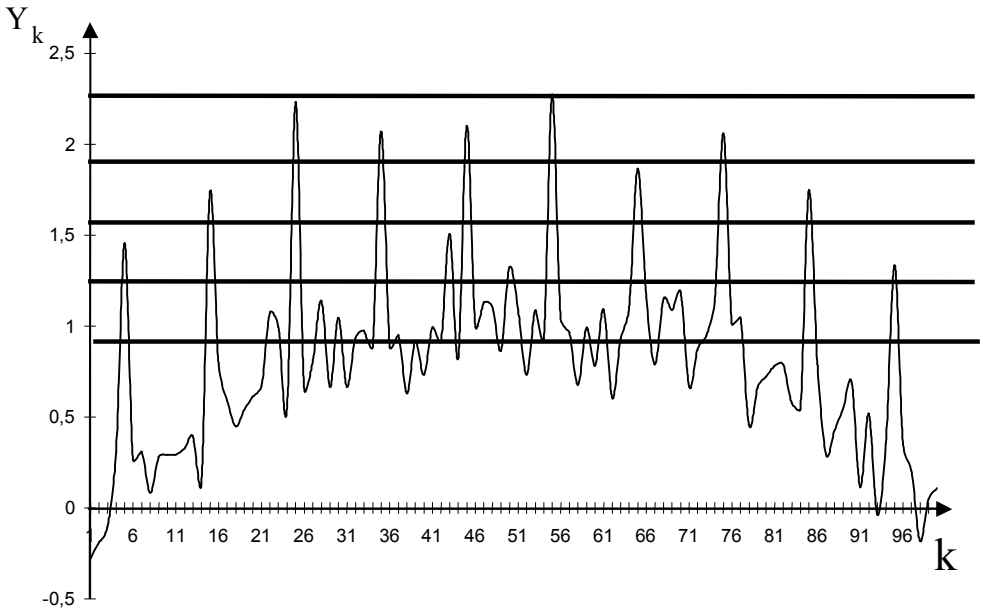


а)

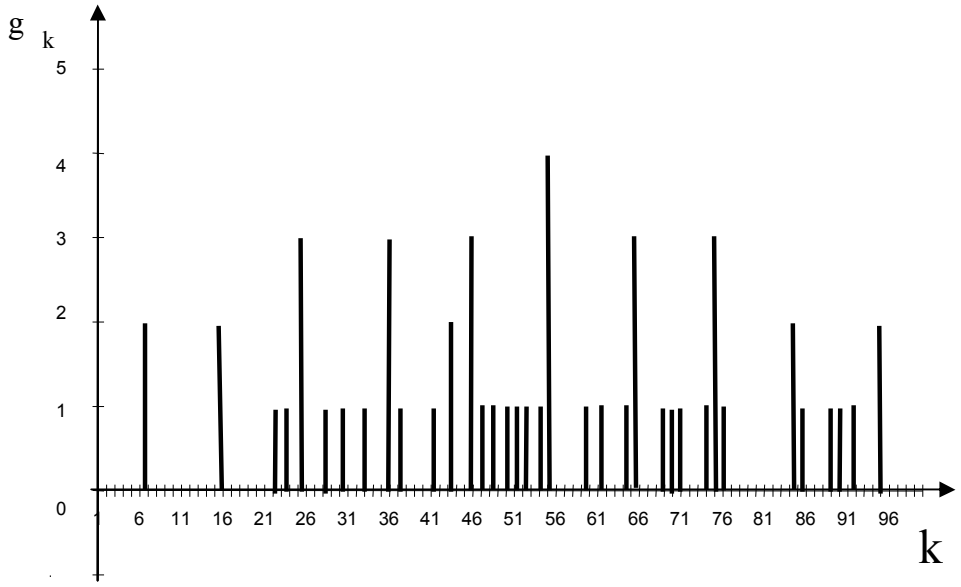


б)

Рис. 2.10. Первый этап работы способа обнаружения
аномальных значений с дискретно-меняющимся уровнем анализа процесса:
а) первый уровень анализа;
б) ряд штрафов

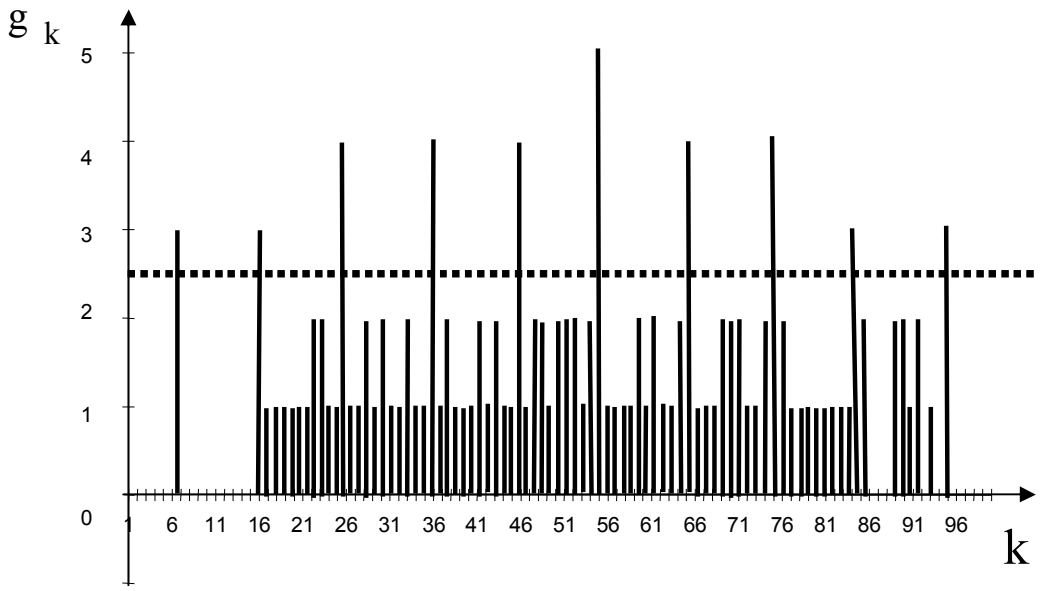


а)

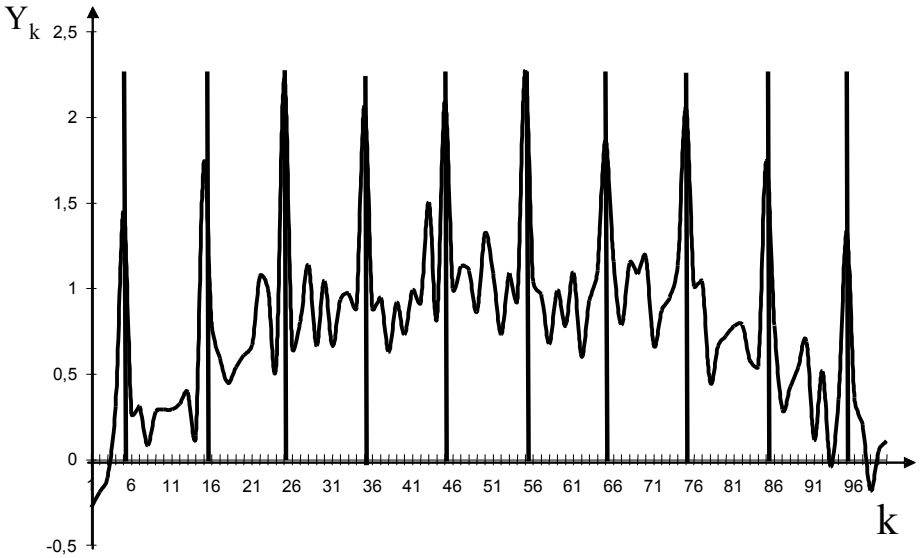


б)

Рис. 2.11. Пятый этап работы предлагаемого способа обнаружения с аномальных значений дискретно-меняющимся уровнем анализа процесса:
а) пятый уровень анализа;
б) ряд штрафов.



а)



б)

Рис. 2.12. Заключительный этап работы способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа процесса:
а) результирующий ряд штрафов;
б) обнаруженные аномальные значения

На рисунке 2.10б представлен ряд штрафов, полученных при превышении каждым из отсчётов $\{Y_k\}_{k=1}^{N}$ уровня анализа L_1. Далее изменяется значение уровня анализа $L_i = Y_{\max} - \Delta Y \cdot i$ и проверяется условие (2.33).

На рисунке 2.11а представлен исходный процесс и пятый уровень анализа, а на рисунке 2.11б соответственно ряд штрафов. После последовательной установки всех уровней анализа $\{L_i\}_{i=0}^{P}$ и проверки условия (2.33) формируется ряд результирующих штрафов изображенных на рисунке 2.12а, где штриховой линией показано пороговое значение, устанавливаемое на уровне 0,5 относительно максимального значения ряда штрафов. На рисунке 2.12б представлены исходный процесс и аномальные значения, обнаруженные по предложенному способу.

Эффективность обнаружения аномальных значений определяется вероятностью ошибки первого рода α и вероятностью правильного обнаружения β. Оценки этих параметров существенно зависят от количества априорно задаваемых уровней анализа нестационарного случайного процесса P и длины интервала нестационарности d, т.е. интервалов на которых нарушается условие стационарности исследуемого входного процесса [112, 113].

Для определения рекомендуемых значений количества уровней анализа процесса P и длины интервалов нестационарности d, с помощью имитационного моделирования получены зависимости оценок вероятности ошибки первого рода α и вероятности правильного обнаружения β от параметров P и d.

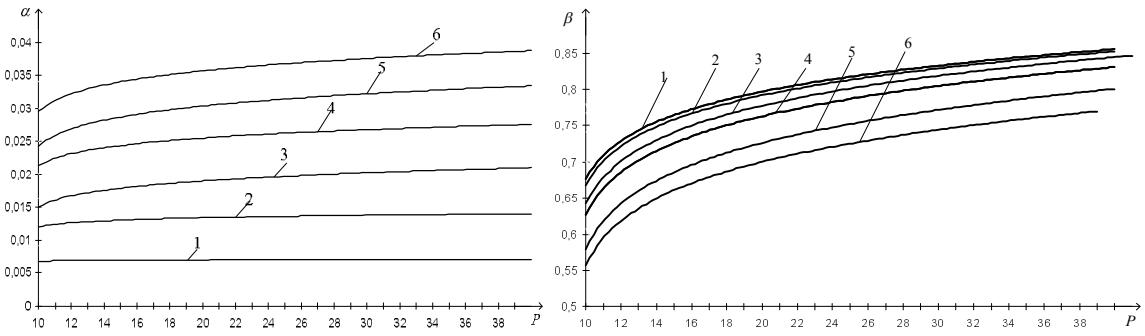
Исследования проводятся для нестационарных случайных процессов, которые представлены моделью (2.2). Аддитивная шумовая составляющая η_k, имеет гауссовский и равномерный законы плотности распределения вероятности. В качестве модели полезной составляющей λ_k

рассматриваются нормированные функции: гармоническая, экспоненциальная и параболическая. Величина аномальных значений выбирается фиксированной и равной $Amp = 5\sigma_{ш}$, где среднеквадратическое отклонение шумовой составляющей процесса $\sigma_{ш} = 0{,}1$. Аномальные значения составляют 5% от исходной выборки N и располагаются равномерно по всей реализации нестационарного случайного процесса Y_k. Значения уровней анализа P изменяется от 10 до 40 с шагом 1, а длина интервалов нестационарности d, принимает дискретные значения от 5 до 10.

Результаты исследований представлены для различных функций полезной составляющей λ_k и гауссовского закона плотности распределения вероятности шумовой составляющей η_k: на рисунках. 2.13а, 2.14а, 2.15а (оценки вероятности ошибки первого рода α), на рисункх 2.13б, 2.14б, 2.15б (оценки вероятности правильного обнаружения β).

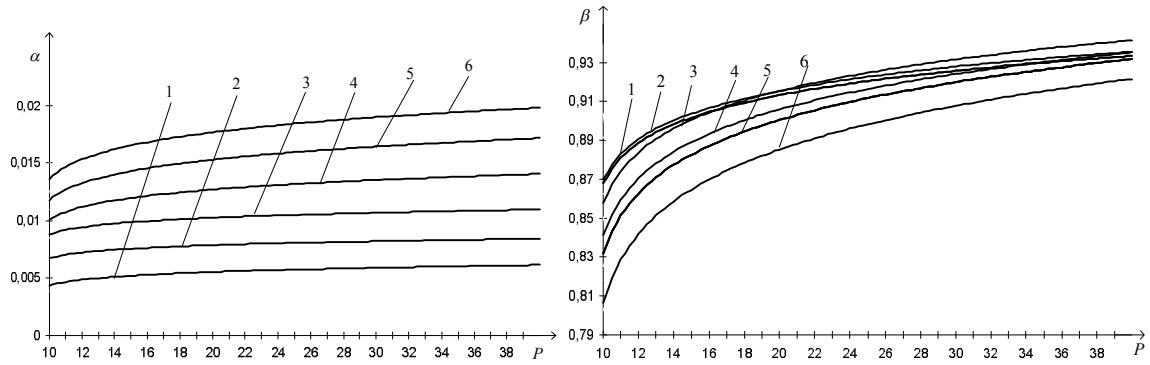
Для случая, когда шумовая составляющая η_k имеет равномерный закон плотности распределения вероятности, результаты исследований представлены: на рисунках 2.16а, 2.17а, 2.18а (оценки вероятности ошибки первого рода α), на рисунках 2.16б, 2.17б, 2.18б (оценки вероятности правильного обнаружения β).

Оценки вероятности ошибки первого рода α и вероятности правильного обнаружения β представлены: на рисунках 2.13а и 2.13б и 2.16а и 2.16б модель полезной составляющей λ_k имеет гармоническую функцию; на рисунках 2.14а и 2.14б и рисунках 2.17а и 2.17б – экспоненциальную функцию; на рисунках 2.16а и 2.16б и на рисунках 2.18а и 2.18б – параболическую функцию.



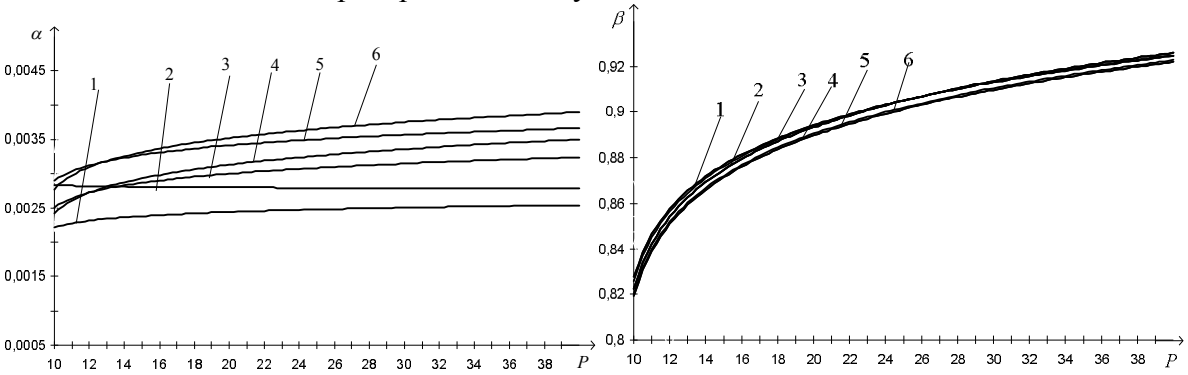
а) б)

Рис. 2.13. Оценки вероятности ошибки первого рода α и вероятности правильного обнаружения β: а) $\alpha = f(P,d)$ и б) $\beta = f(P,d)$ для гармонической функции полезной составляющей и гауссовским законом распределения шумовой составляющей



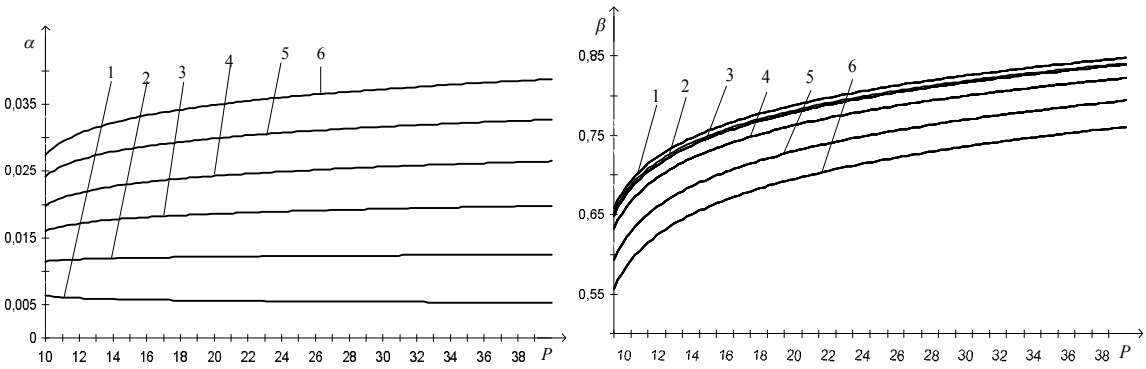
а) б)

Рис. 2.14. Оценки вероятности ошибки первого рода α и вероятности правильного обнаружения β: а) $\alpha = f(P,d)$ и б) $\beta = f(P,d)$ для экспоненциальной функции полезной составляющей процесса и гауссовским законом распределения шумовой составляющей



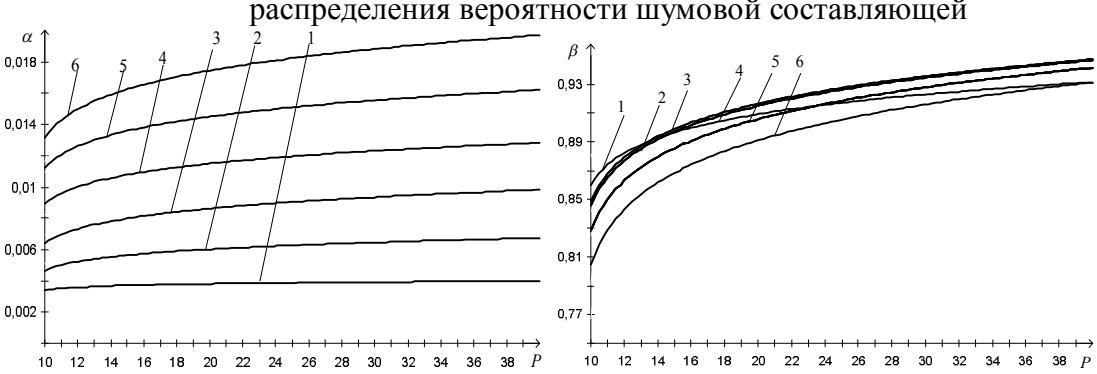
а) б)

Рис. 2.15. Оценки вероятности ошибки первого рода α и вероятности правильного обнаружения β: а) $\alpha = f(P,d)$ и б) $\beta = f(P,d)$ для параболической функции полезной составляющей процесса и гауссовским законом плотности распределения вероятности шумовой составляющей



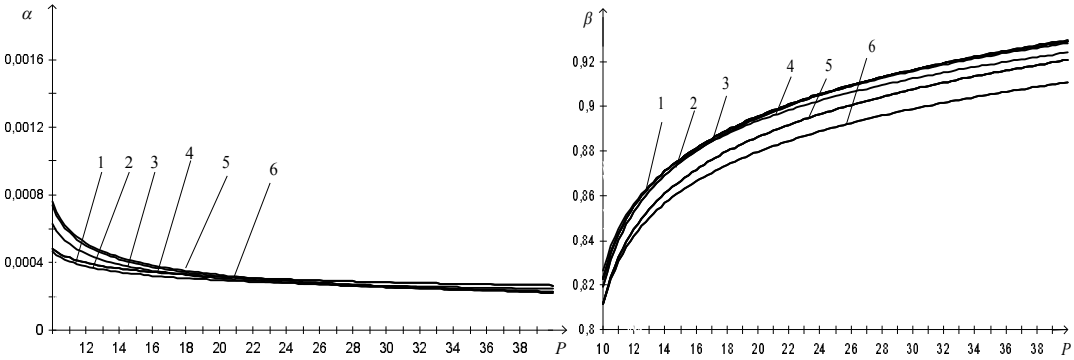
а) б)

Рис. 2.16. Оценки вероятности ошибки первого рода α и вероятности правильного обнаружения β: а) $\alpha = f(P,d)$ и б) $\beta = f(P,d)$ для гармонической функции полезной составляющей процесса и равномерным законом плотности распределения вероятности шумовой составляющей



а) б)

Рис. 2.17. Оценки вероятности ошибки первого рода α и вероятности правильного обнаружения β: а) $\alpha = f(P,d)$ и б) $\beta = f(P,d)$ для экспоненциальной функции полезной составляющей процесса и равномерным законом плотности распределения вероятности шумовой составляющей



а) б)

Рис. 2.18. Оценки вероятности ошибки первого рода α и вероятности правильного обнаружения β: а) $\alpha = f(P,d)$ и б) $\beta = f(P,d)$ для параболической функции полезной составляющей процесса и равномерным законом плотности распределения вероятности шумовой составляющей

На всех представленных рисунках 2.13 – 2.18: график 1 – для $d = 5$; график 2 – для $d = 6$; график 3 – для $d = 7$; график 4 – для $d = 8$; график 5 – для $d = 9$; график 6 – для $d = 10$.

Анализ результатов исследований, представленных на рисунках 2.13 – 2.18 показывает, что при использовании предлагаемого способа обнаружения аномальных значений, оценки вероятности ошибки первого рода α уменьшаются с уменьшением d, а оценки вероятности правильного обнаружения β увеличиваются с уменьшением d. При этом, для всех рассматриваемых функций полезной составляющей λ_k, наилучшие значения лежат в пределах $d = 5 - 7$ (графики 1, 2, 3 на рис. 2.13 – 2.18), т.е. когда оценки вероятности ошибки первого рода α минимальны, а оценки вероятности правильного обнаружения β максимальны.

Проведенные исследования показывают, что для параметров длины интервала нестационарности $d = 5 - 7$ значения количества уровней анализа при $P > 16$ практически не влияет на оценки вероятности ошибки первого рода α (рис. 2.13,а, 2.14,а, 2.15,а, 2.16,а, 2.17,а). Оценки вероятности правильного обнаружения β при равномерном законе плотности распределения распределении вероятности шумовой составляющей процесса η_k так же практически не меняются при значении количества уровней анализа $P > 18$ для всех моделей функций полезной составляющей λ_k (рис. 2.16,б, 2.17,б, 2.18,б).

Для более точного оценивания значения количества уровней анализа P рассмотрены зависимости оценки вероятности правильного обнаружения β от длины интервала нестационарности d. Исследование проводятся для случая, когда модель полезной составляющей процесса λ_k является гармонической, а закон распределения шумовой составляющей η_k представлена гауссовским законом распределения. Значения d

выбирается в пределах от 5 до 10. На рисунке 2.19 представлены результаты такого исследования.

Результаты, представленные на рисунке 2.19, соответствуют следующим значениям уровней анализа процесса: график 1 – $P = 13$, график 2 – $P = 12$, график 3 – $P = 11$, график 4 – $P = 15$, график 5 – $P = 10$, график 6 – $P = 20$, график 7 – $P = 25$, график 8 – $P = 30$, график 9 – $P = 35$, график 10 – $P = 40$.

Из анализа результатов, представленных на рисунке 2.19 видно, что максимальные оценки вероятности правильного обнаружения β получаются, при значениях уровней анализа $P > 25$ при $d = 5 - 7$ [118].

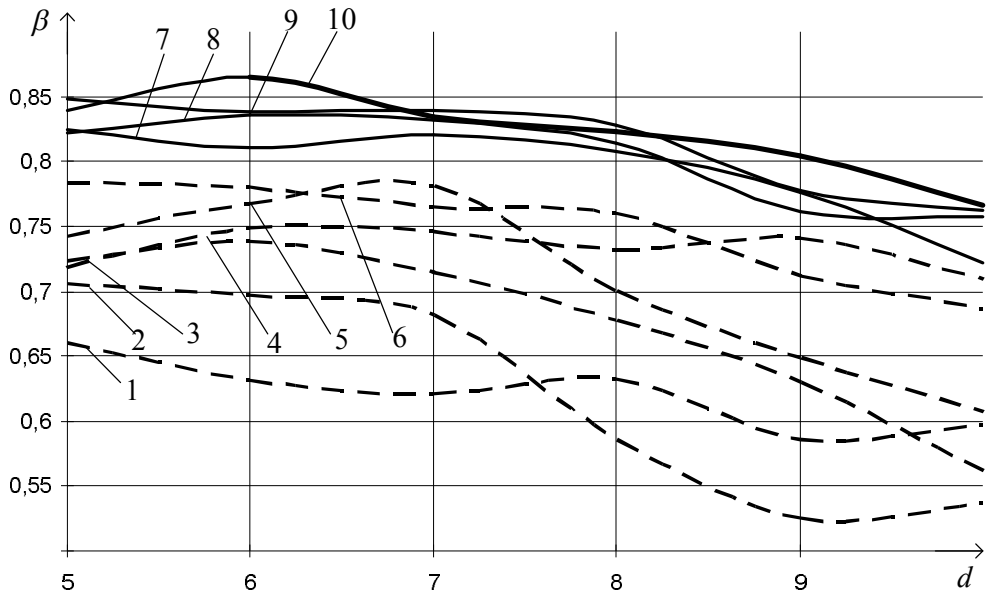


Рис. 2. 19. Зависимость $\beta = f(d, P)$ для гармонической функции полезной составляющей процесса

Программная реализация алгоритма способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа процесса, представленного на рисунке 2.20, осуществляется следующим образом:

– запоминается входная реализация $Y_1, Y_2, .., Y_N$;

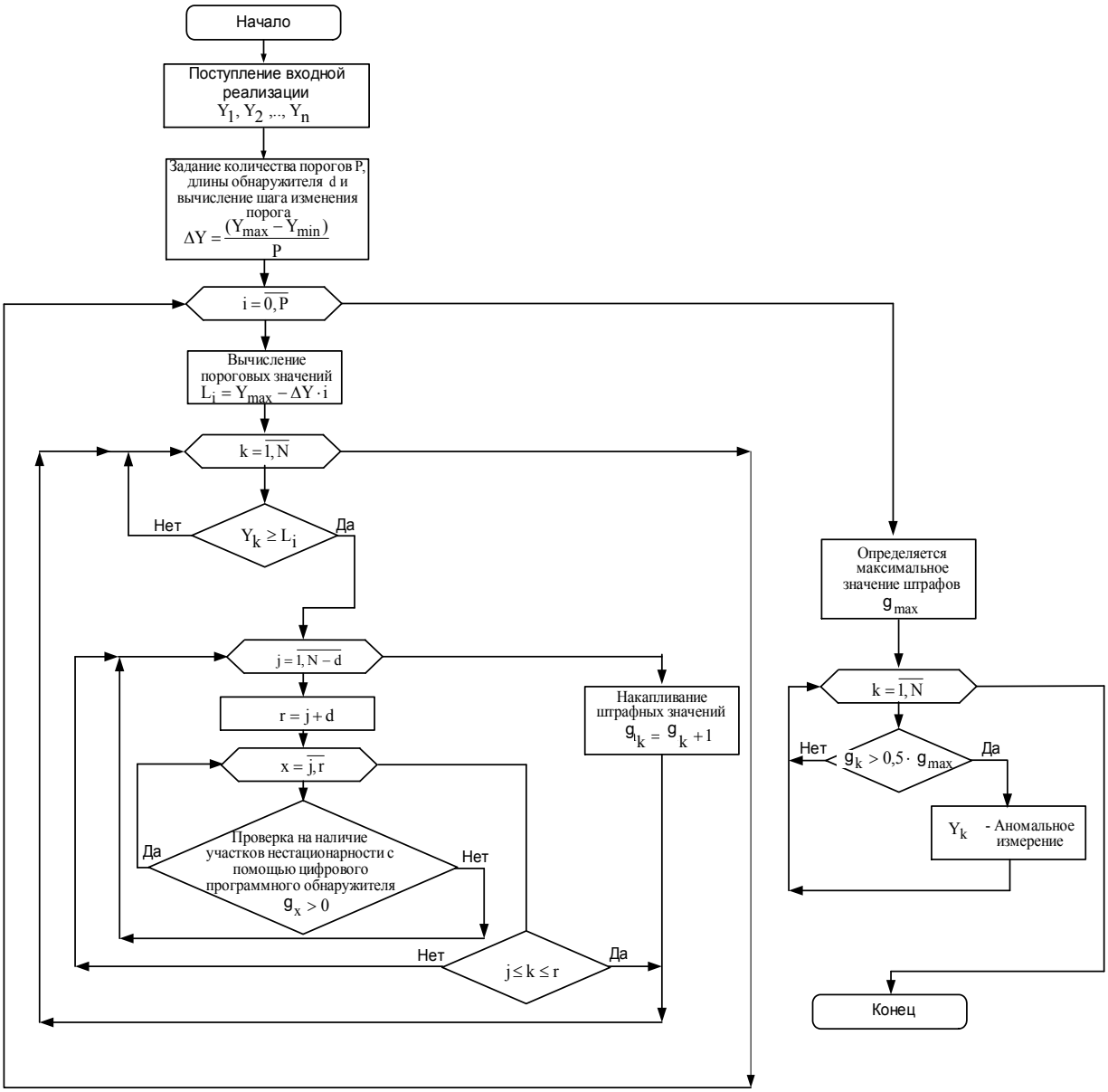


Рис. 2.20. Блок-схема алгоритма способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа

– задается количество уровней анализа P и длина интервал нестационарности d, согласно которой определяются участки наличия полезной составляющей, а так же вычисляется изменения шага уровня ΔY;

– вычисляются значения уровней анализа $\{L_i\}_{i=0}^{P}$;

– каждое значение входной реализации $Y_1, Y_2, .., Y_N$ сравнивается со значением уровня анализа, L_i и определяются штрафные значения, которые накапливаются в виде ряда $\{g_k\}_{k=1}^{N}$;

– на каждом уровне L_i проводится анализ штрафных значений на наличие участков нестационарности, то есть обнаружение в реализации штрафов последовательности в d ненулевых элементов и их локализация (штрафы не назначаются на этом интервале при дальнейшем уменьшении значения уровня анализа);

– определяется $g_{\max}$ ряда $\{g_k\}_{k=1}^{N}$ и проверяется условие, $g_k > 0{,}5 \cdot g_{\max}$, $k = \overline{1, N}$.

При выполнении данного условия принимается решение, что соответствующее значение нестационарного случайного процесса Y_k является аномальным значением [119, 120].

2.5 Структурная схема устройства, реализующего способ обнаружения аномальных значений с дискретно-меняющимся уровнем анализа процесса

На рисунке 2.21 представлена структурная схема устройства, реализующего способ обнаружения аномальных значений с дискретно-меняющимся уровнем анализа процесса [119]. Одноканальное устройство для обнаружения аномальных значений с дискретно-меняющимся уровнем анализа процесса может быть реализовано следующим образом: значения обрабатываемого нестационарного случайного процесса Y_k с буферного каскада (Б) поступают в блок определения количества уровней анализа (БУА), которые определяются в зависимости от максимального и минимального значений процесса Y_k, и одновременно в коммутатор (К1).

С помощью коммутатора (К1) определяются значения входного процесса, которые превысили задаваемый уровень анализа. Значения, которые превышают задаваемый уровень анализа, проверяются в блоке (ПН) на наличие участков нестационарности. Во втором коммутаторе (К2) осуществляется назначение штрафов при превышении уровня анализа и отсутствие процесса с выхода блока (ПН). При наличии процесса с выхода блока (ПН) штрафы не назначаются даже в случае их превышения уровня анализа. В суммирующем устройстве (СУ) определяется общее количество штрафов, которые получены при изменении уровней анализа от максимального до минимального. С выхода (СУ) суммарное число штрафных значений поступает в блок вычисления второго порового значения (БУП). Вычисленное второе пороговое значение с выхода блока (БУП) поступает на вход третьего коммутатора (К3), где определяются аномальные значения из числа оштрафованных исследуемого нестационарного случайного процесса Y_k. С выхода коммутатора (К3) реализация анализируемого процесса поступает в блок устранения аномальных значений (БУ), за счет интерполяции по соседним значениям. Синхронность работы устройства обнаружения аномальных значений без оценки функции полезной составляющей процесса осуществляется от тактового генератора импульсов (генератора импульсов синхронизации).

В том случае, когда априорно известно (априорно фиксировано) количество уровней анализа исследуемых нестационарных случайных процессов возможно построение многоканального устройства, реализующего предложенный способ обнаружения аномальных значений.

Тогда при построении устройства количество каналов будет определяться количеством уровней анализа исследуемых процессов и определение участков нестационарности будет происходить по каждому каналу одновременно, а суммарное количество штрафов определяться как сумма штрафов, полученных после анализа значений в каждом канале.

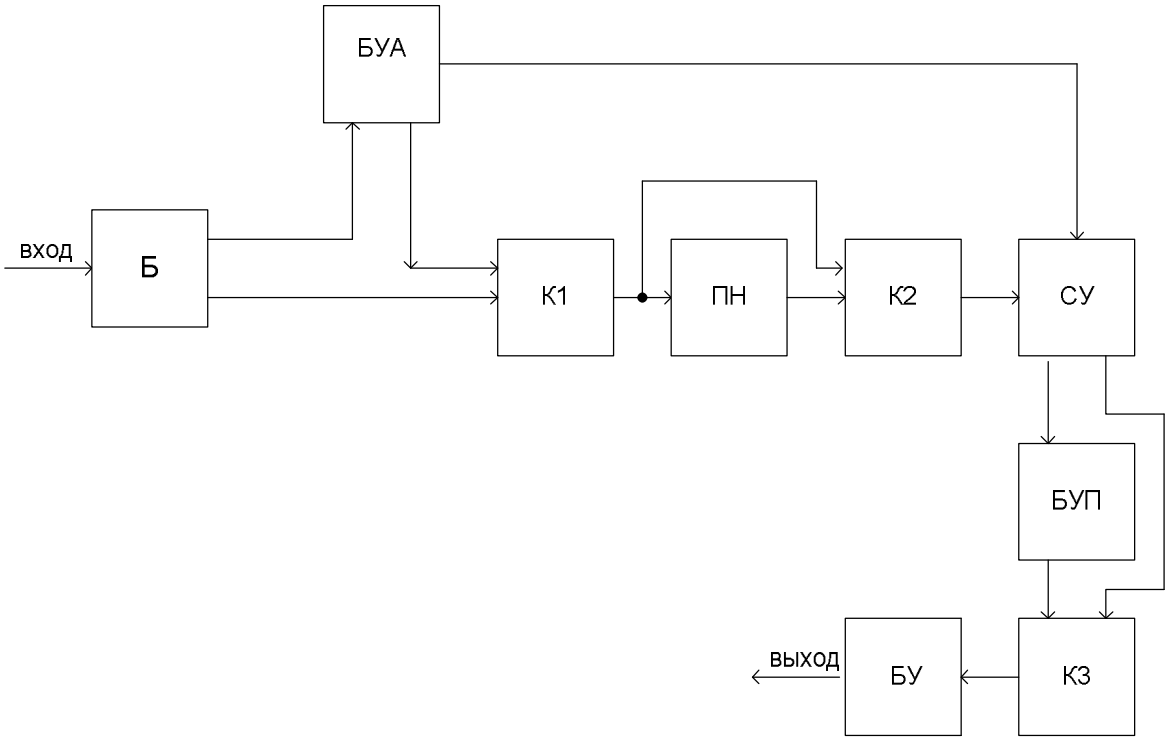


Рис. 2.21. Структурная схема устройства, реализующего способ обнаружения аномальных значений с дискретно-меняющимся уровнем анализа процесса

Реализация, предложенного способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа процесса [119], возможна как в аппаратном исполнении, так и в программном.

ГЛАВА 3 Анализ эффективности адаптивного способа обнаружения аномальных значений

Для определения эффективности адаптивного способа обнаружения аномальных значений проводится сравнительный анализ предлагаемого способа и способа обнаружения аномальных значений без адаптации порогового значения (2.27), который представлена в работах [54, 79].

3.1 Исследование эффективности адаптивного способа обнаружения аномальных значений для нестационарных случайных процессов с аддитивной шумовой составляющей

На основе имитационного моделирования, в работах [54, 79], при анализе нестационарных случайных процессов, получены зависимости оценок вероятности ошибки первого рода α и вероятности правильного обнаружения β для способа обнаружения аномальных значений без адаптации, т.е. когда значение коэффициента A в пороговом значении (2.27) задается фиксированным $A = 2{,}5$, и с адаптацией порогового значения (2.27), т.е. когда значение коэффициента A определяется выражением $A = f(\alpha)$ (2.30) [121].

Исследования эффективности предлагаемого способа проводятся для случая, когда модель нестационарного случайного процесса является аддитивной (2.2) [122]. В качестве моделей полезной составляющей λ_k исходного процесса Y_k, где $k = \overline{1, N}$, используются следующие нормированные функции: гармоническая, экспоненциальная, параболическая, составная и сложная. Аддитивная шумовая составляющая процесса ε_k имеет гауссовский закон плотности распределения вероятности (2.6). Одиночные аномальные значения распределены

равномерно по всей реализации нестационарного случайного процесса Y_k и составляют 5% от выборки N. Исследования проводятся для различных значений величины аномальных значений Amp, т.е: $3\sigma_ш, 4\sigma_ш, 5\sigma_ш, 6\sigma_ш, 7\sigma_ш$, $\sigma_ш = 0{,}1$ - среднеквадратическое отклонение аддитивной шумовой составляющей. Значения вероятности ошибки первого рода для способа с адаптацией порогового значения априорно фиксируется $\alpha = 0{,}05$.

В результате проведенных исследований, для нестационарных случайных процессов получены зависимости оценок вероятности правильного обнаружения $\beta = f(Amp)$. Для случая, когда не используется адаптация порогового значения – графики $1''$, $2''$, $3''$, $4''$, $5''$ и с применением адаптация порогового значения – графики 1, 2, 3, 4, 5 (рис. 3.1).

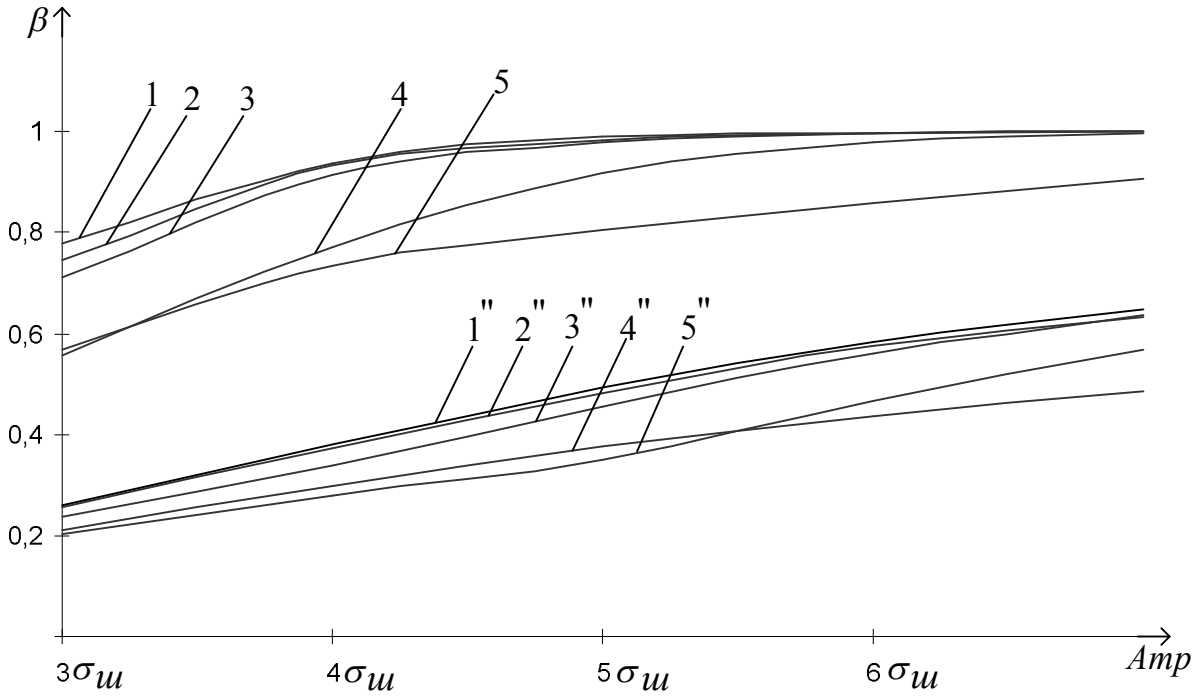


Рис. 3.1. Зависимость оценок вероятности правильного обнаружения $\beta = f(Amp)$ для способа без адаптации и способа с адаптацией порогового значения при $\alpha = 0{,}05$

Зависимости $\beta = f(Amp)$ на рисунке 3.1 представлены для различных моделей функций полезной составляющей λ_k: графики 1, 1" – экспоненциальной; графики 2, 2" – параболической; графики 3, 3" – гармонической; графики 4, 4" – составной и графики 5, 5" – сложной функции.

Анализ результатов, представленных на рисунке 3.1 показывает, что при введении адаптации порогового значения, оценки вероятности правильного обнаружения β возрастают для всех рассмотренных функций полезной составляющей λ_k. Причем, для параболической, гармонической и экспоненциальной модели функций, при величине аномальных значений порядка $3\sigma_ш$, оценки вероятности правильного обнаружения β возрастают примерно на 66%. С увеличением величины аномальных значений ($Amp > 5\sigma_ш$), оценки вероятности правильного обнаружения β увеличивается примерно на 54%. Из анализа зависимостей $\beta = f(Amp)$ так же следует, что при использовании адаптации порогового значения, с увеличением величины аномальных значений Amp, вероятность правильного обнаружения β асимптотически стремится к единице независимо от модели функции полезной составляющей λ_k [121, 122].

Применяя адаптацию порогового значения, так же получены зависимости оценок вероятности ошибки первого рода $\alpha = f(Amp)$, которые представлены на рисунке 3.2.

Зависимости $\alpha = f(Amp)$ на рисунке 3.2 представлены для следующих моделей функций полезной составляющей сигнала λ_k: график 1 – параболической; график 2 – составной; график 3 – экспоненциальной; график 4 – гармонической; график 5 – сложной.

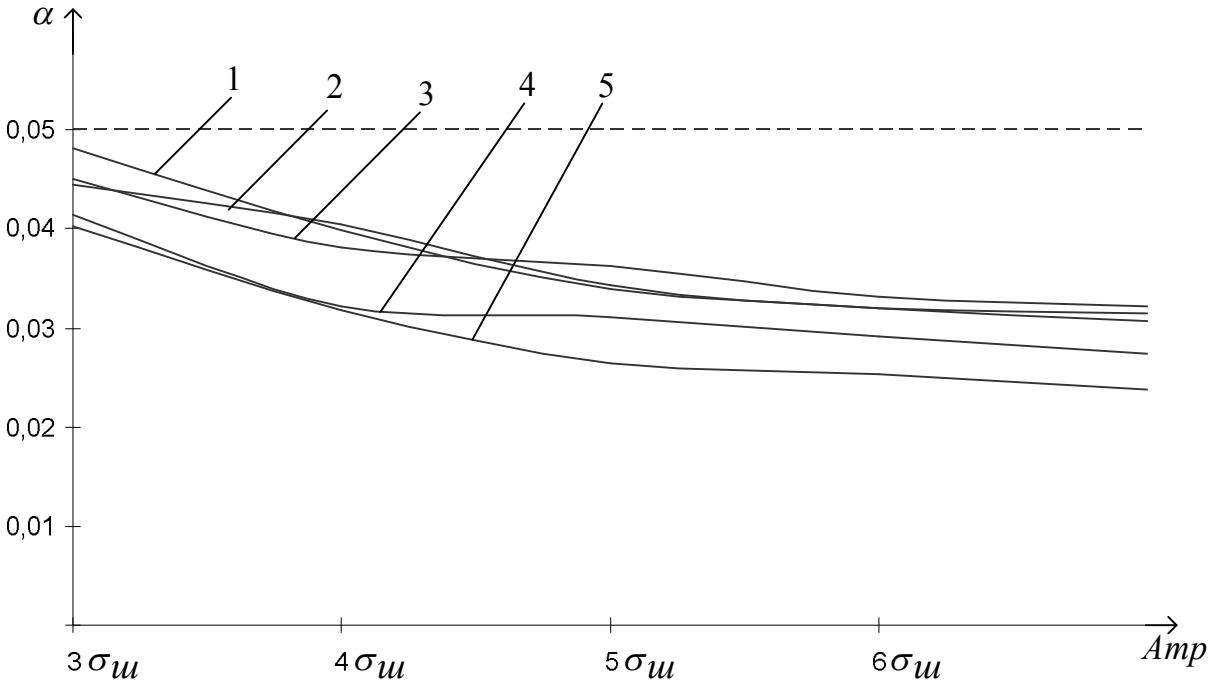


Рис. 3.2. Зависимость оценок ошибки первого рода $\alpha = f(Amp)$ для способа с адаптацией порогового значения при $\alpha = 0{,}05$

Из анализа полученных зависимостей $\alpha = f(Amp)$ следует, что при использовании адаптации порогового значения оценки вероятности ошибки первого рода α практически не превосходят априорно задаваемого значения, т.е. $\alpha = 0{,}05$, для всех исследуемых нестационарных случайных процессов (рис. 3.2).

Так же исследуется влияние использования одного из способов стабильного оценивания параметров выборки нестационарного случайного процесса, т.е. оценок математического ожидания $\overline{m}_j^{(i)}$ и среднеквадратического отклонения $\overline{\sigma}_j^{(i)}$ (2.28), на оценки вероятности правильного обнаружения β [123]. Результаты исследования представлены на рисунке 3.3.

Для случая с адаптацией порогового значения, но без стабильного оценивания параметров выборки: график $1''$ для экспоненциальной; график $2''$ для параболической; график $3''$ для гармонической функций полезной

составляющей λ_k; с применением стабильного оценивания параметров выборки график 1 для экспоненциальной; график 2 для параболической; график 3 для гармонической функций полезной составляющей λ_k.

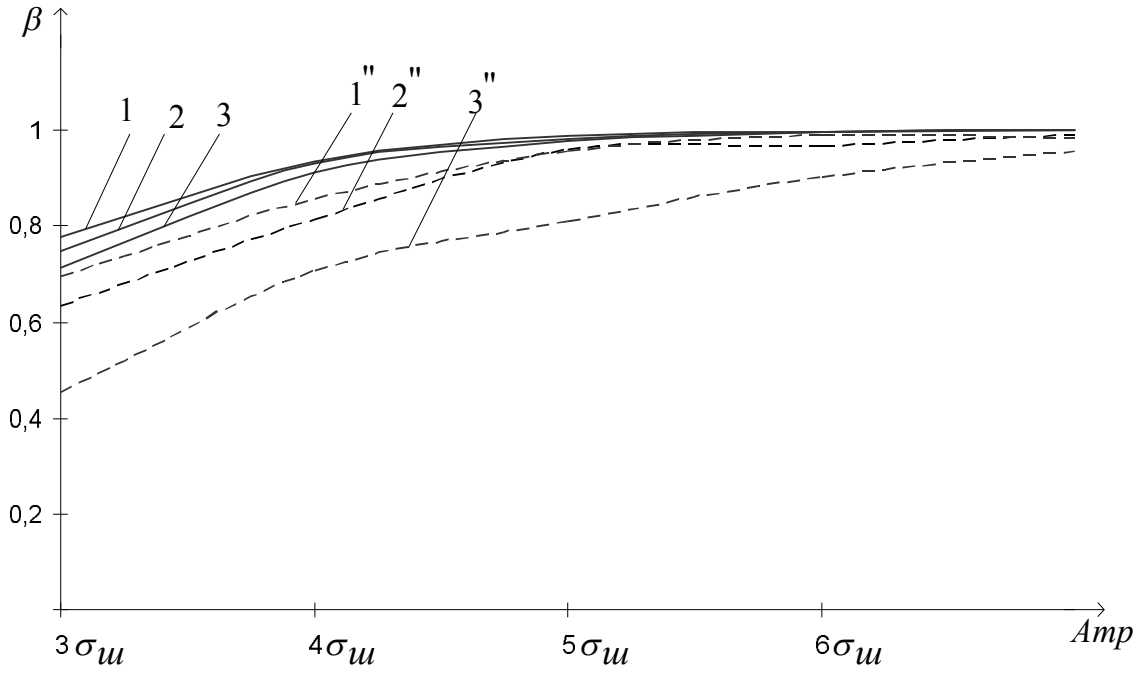


Рис. 3.3. Зависимость оценок вероятности правильного обнаружения $\beta = f(Amp)$ при $\alpha = 0{,}05$

Анализ зависимостей, которые представлены на рисунке 3.3, показывает, что при использовании стабильного оценивания параметров выборки нестационарного случайного процесса, оценки вероятности правильного обнаружения β возрастают для всех исследуемых функций полезной составляющей.

При величине аномальных значений $3\sigma_ш$ и $4\sigma_ш$, для гармонической функции полезной составляющей λ_k оценки вероятности правильного обнаружения β увеличиваются в среднем на 36-23 % соответственно.

При тех же условиях: для параболической функции полезной составляющей λ_k увеличение оценок β составляет примерно 16-13%, а

для экспоненциальной функции полезной составляющей λ_k оценки β увеличиваются примерно на 9 – 11 %.

При дальнейшем увеличении величины аномальных значений Amp оценки β практически не меняются, т.е. не зависят от способа оценивания параметров выборки исследуемого процесса. Однако, для случая, когда полезная составляющая сигнала λ_k является гармонической моделью, даже при величине аномальных значений порядка $5\sigma_{ш}$, использование стабильного оценивания параметров выборки исследуемого нестационарного случайного процесса позволяет получить значение оценок β на 17% больше, чем если бы такая поцедура оценивания не была бы применима [123].

В работе исследуется эффективность адаптивного способа обнаружения аномальных значений в зависимости от места расположения аномальных значений в выборке нестационарного случайного процесса.

Рассматривается модель с аддитивной шумовой составляющей ε_k, закон плотности распределения вероятности которой является центрированным гауссовским случайным процессом со среднеквадратическим отклонением $\sigma_{ш} = 0{,}1$. В качестве модели функции полезной составляющей λ_k используются следующие нормированные функции: экспоненциальная, гармоническая, составная функции.

Аномальные значения с фиксированной величиной $Amp = 5\sigma_{ш}$, составляют 5% от объема исследуемой выборке N. Рассматриваются случаи, когда аномальные значения располагаются в начале выборки, в середине выборки, в конце выборки и равномерно по всей выборке нестационарного случайного процесса Y_k , где $k = \overline{1, N}$.

В результате проведенных исследований, получены оценки вероятности правильного обнаружения β для случая без адаптации и с адаптацией порогового значения, которые представлены в таблице 3.1.

Таблица 3.1

Значение оценок вероятности правильного обнаружения

Расположение аномальных значений	Гармоническая функция		Экспоненциальная функция		Составная функция	
	с адаптацией порогового значения	без адаптации пороговог о значения	с адаптацией порогового значения	без адаптации порогового значения	с адаптацией порогового значения	без адаптации пороговог о значения
В начале выборки	0,796	0,457	0,861	0,204	0,694	0,199
В середине выборки	0,930	0,201	0,928	0,269	0,842	0,251
В конце выборки	0,898	0,252	0,925	0,254	0,828	0,241
Равномерно расположены по всей выборке	0,979	0,204	0,925	0,254	0,920	0,369

Анализ результатов, представленных в таблице 3.1, показывает, что оценки вероятности правильного обнаружения β при использовании способа с адаптацией порогового значения более высокие, чем при использовании способа с адаптацией порогового значением. Например, для экспоненциальной модели функции полезной составляющей λ_k, при расположении аномальных значений в начале выборки, оценки вероятности правильного обнаружения β возрастают примерно на 76% и на 70% при расположении аномальных значений в конце выборки. Причем при адаптации порогового значения, оценки вероятности правильного обнаружения β практически не зависят от места расположения аномальных значений в выборке нестационарного случайного процесса Y_k.

При тех же условиях для способа с адаптацией порогового значения определяются оценки вероятности ошибки первого рода α. В таблице 3.2

приведены оценки вероятности ошибки первого рода, которые получены для случая априорного значения $\alpha = 0{,}05$.

Таблица 3.2

Значение оценок вероятности ошибки первого рода

Расположение аномальных значений	Гармоническая функция	Экспоненциальная функция	Составная функция
В начале выборки	0,037	0,044	0,037
В середине выборки	0,044	0,045	0,035
В конце выборки	0,042	0,042	0,036
Равномерно расположены по всей выборке	0,040	0,044	0,034

Из анализа таблицы 3.2, следует, что оценки вероятности ошибки первого рода α не превосходят своего априорно задаваемого значения независимо от места расположения аномальных значений в выборке нестационарного случайного процесса Y_k.

3.2 Исследование эффективности адаптивного способа обнаружения аномальных значений для нестационарных случайных процессов с мультипликативной шумовой составляющей

В данном разделе работы рассматривается влияние применения адаптации порогового значения, для случая когда нестационарный случайный процесс Y_k $(k = \overline{1, N})$ является мультипликативным, т.е. представлен моделью (2.3). Для этого анализируется эффективность использования способа с адаптацией порогового значения для нестационарных случайных процессов с мультипликативной шумовой составляющей μ_k при различных значениях величины аномальных значений Amp [124].

В качестве модели функции полезной составляющей λ_k используются следующие нормированные функции: экспоненциальная, гармоническая и составная функции. В качестве модели мультипликативной шумовой составляющей сигнала μ_k – гауссовский случайный процесс. Величина аномальных значений Amp: $3\sigma_ш, 4\sigma_ш, 5\sigma_ш, 6\sigma_ш, 7\sigma_ш$ ($\sigma_ш = 0{,}1$ – среднеквадратическое отклонение мультипликативной шумовой составляющей μ_k). Значение вероятности ошибки первого рода для способа с адаптацией порогового значения априорно выбирается $\alpha = 0{,}05$. Одиночные аномальные значения распределены равномерно по всей выборке нестационарного случайного процесса Y_k, где $k = \overline{1, N}$ и составляют 5% от N.

В результате исследований получены зависимости оценок вероятности правильного обнаружения $\beta = f(Amp)$, которые представлены на рисунке 3.4. При применении способа обнаружения аномальных значений с адаптацией порогового значения: графики 1, 2 и 3 и способа обнаружения аномальных значений без адаптации порогового значения: графики $1''$, $2''$ и $3''$.

Функциональные зависимости соответствуют следующим моделям функции полезной составляющей сигнала λ_k, в случае с адаптацией порогового значения: график 1 – экспоненциальной; график 2 – гармонической; график 3 – составной функциям и в случае без адаптации порогового значения: график $1''$ – экспоненциальной; график $2''$– гармонической; график $3''$ – составной функциям.

Анализ результатов, приведенных на рисунке 3.4, показывает, что с применением способа обнаружения аномальных значений с адаптацией порогового значения, оценки вероятности правильного обнаружения β увеличиваются в зависимости от величины аномальных значений

($Amp = 3\sigma_ш \div 7\sigma_ш$): для гармонической модели функции полезной составляющей λ_k примерно на 51–26%, для экспоненциальной модели функции полезной составляющей λ_k примерно на 28–25%, для составной модели функции полезной составляющей λ_k примерно на 43–24%.

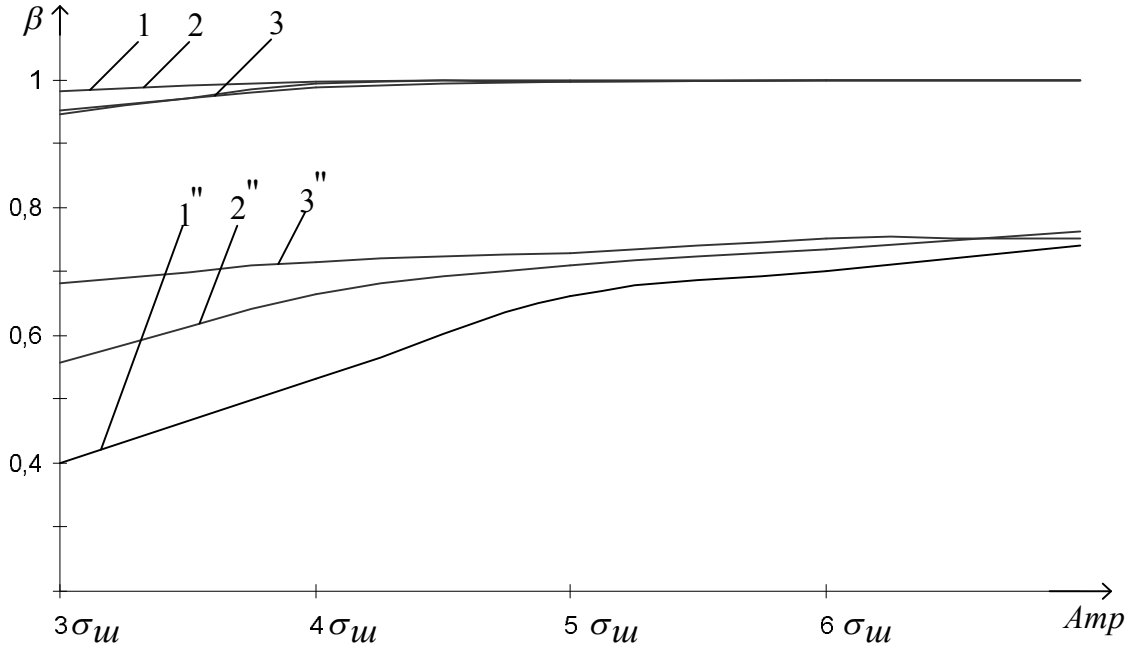


Рис. 3.4. Зависимость оценок вероятности правильного обнаружения $\beta = f(Amp)$ для способа без адаптации и с адаптацией порогового значения при $\alpha = 0{,}05$

Из анализа зависимостей, представленных на рисунке 3.4, так же следует, что при использовании способа с адаптацией порогового значения оценки вероятности правильного обнаружения β практически не зависят от величины аномальных значений Amp и стремятся к единице независимо от модели функции полезной составляющей λ_k.

При тех же условиях определяются оценки вероятности ошибки первого рода α, которые представлены на рисунке 3.5. Соответственно, график 1 – экспоненциальная, график 2 – гармоническая и график 3 – составная функции полезной составляющей λ_k. Анализ результатов, представленных на рис. 3.5, показывает, что адаптация порогового

значения позволяет получить оценки вероятности ошибки первого рода α непревосходящие априорно задаваемого значения $\alpha = 0{,}05$, независимо от модели функции полезной составляющей λ_k, при всех рассмотренных величинах аномальных значений Amp.

При исследовании эффективности способа обнаружения аномальных значений с адаптацией порогового значения для нестационарных случайных процессов с мультипликативной шумовой составляющей, рассматривались зависимости вероятности правильного обнаружения $\beta = f(Amp)$ при гауссовском (2.6) и равномерном (2.7) законах плотности распределения вероятности мультипликативной шумовой составляющей μ_k, которые представлены на рисунках 3.6 и 3.7.

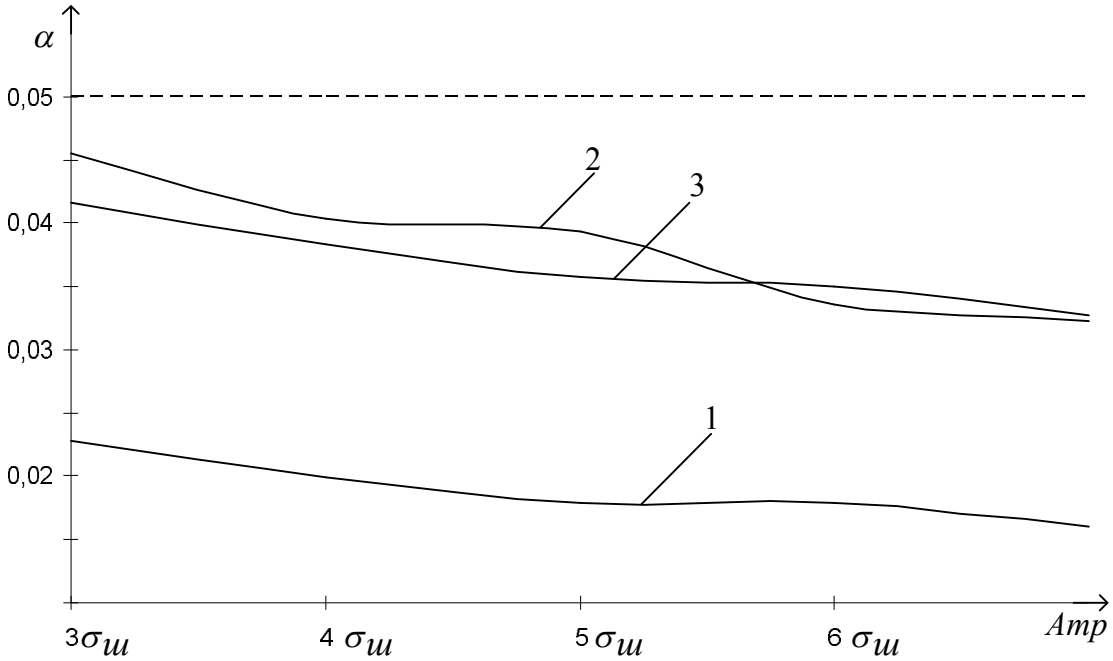


Рис. 3.5. Зависимость оценок ошибки первого рода $\alpha = f(Amp)$ для способа с адаптацией порогового значения при $\alpha = 0{,}05$

Зависимости $\beta = f(Amp)$ рассматриваются при различных моделях функции полезной составляющей λ_k: график 1 – составная, график 2 –

сложная, график 3 – экспоненциальная, график 4 – гармоническая, график 5 – параболическая функции.

Анализ результатов, представленных на рисунке 3.6, показывает, что при гауссовском законе плотности распределения вероятности мультипликативной шумовой составляющей μ_k, оценки вероятности правильного обнаружения β, при величине аномальных значений $Amp = 3\sigma_ш$, достигают $\beta = 0{,}932 - 0{,}980$ в зависимости от функции полезной составляющей λ_k.

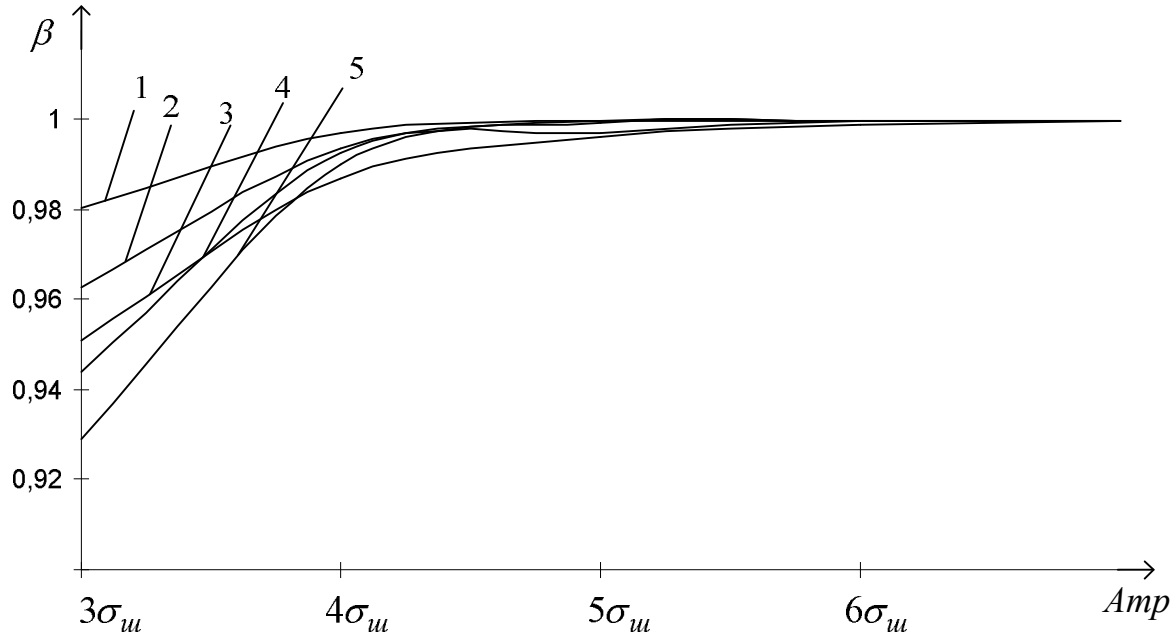


Рис. 3.6. Зависимость оценок вероятности правильного обнаружения $\beta = f(Amp)$ при гауссовском законе плотности распределения вероятности мультипликативной шумовой составляющей

А при $Amp > 5\sigma_ш$ оценки вероятности правильного обнаружения β стремятся к единице, независимо от модели полезной составляющей λ_k. При равномерном законе плотности распределения вероятности мультипликативной шумовой составляющей μ_k оценки вероятности правильного обнаружения β при $Amp = 3\sigma_ш$ достигают $\beta = 0{,}970 - 0{,}985$ в зависимости от функции полезной составляющей λ_k, а при $Amp > 4\sigma_ш$

стремятся к единице независимо от функции полезной составляющей λ_k (рис. 3.7).

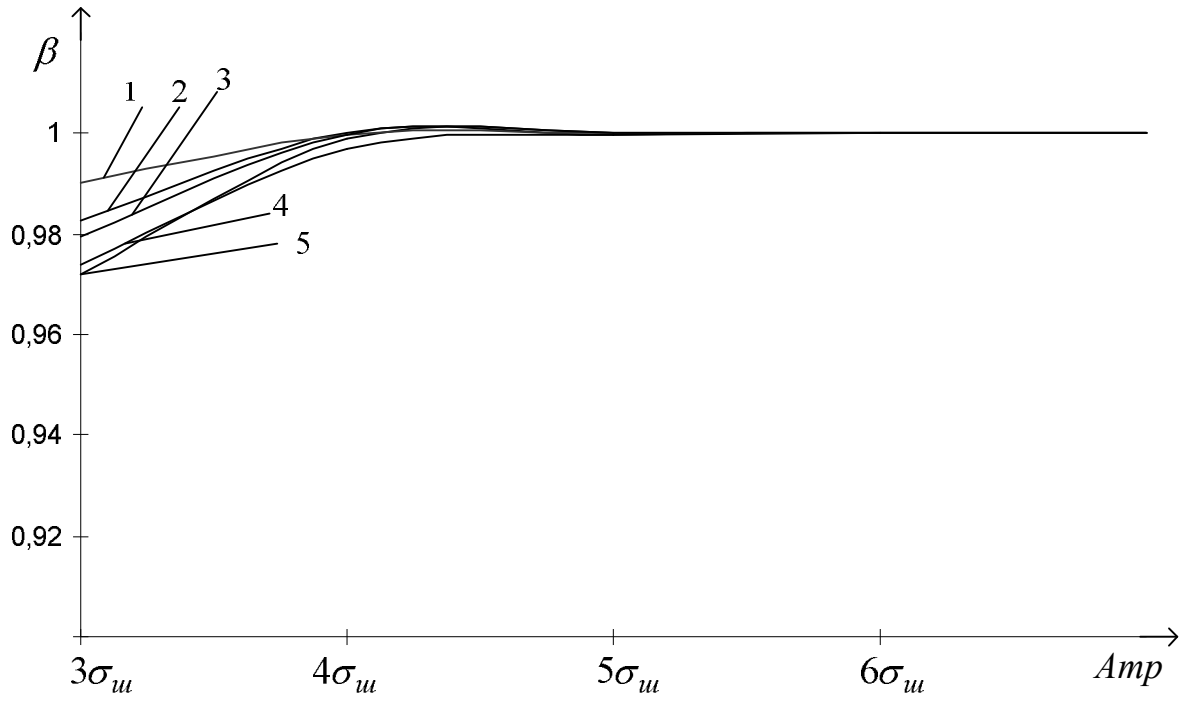


Рис. 3.7. Зависимость оценок вероятности правильного обнаружения $\beta = f(Amp)$ при равномерном законе плотности распределения вероятности мультипликативной шумовой составляющей

На рисунке 3.8 представлены графики зависимости $\beta = f(Amp)$ для случая, когда полезная составляющая λ_k является параболической, а мультипликативная шумовая составляющая сигнала μ_k имеет равномерный – график 1 и гауссовский – график 2 законы плотности распределения вероятности.

Анализ результатов, представленных на рисунке 3.8, показывает, что при величине аномальных значений $Amp = 3\sigma_ш$ оценки вероятности правильного обнаружения β для гауссовского закона плотности распределения вероятности мультипликативной шумовой составляющей μ_k на 4 %, меньше, чем при равномерном законе плотности распределения вероятности мультипликативной шумовой составляющей μ_k, что может

считать незначительным. При увеличении величине аномальных значений Amp оценки вероятности правильного обнаружения β практически не зависят от закона распределения мультипликативной шумовой составляющей μ_k и стремятся к единице [124].

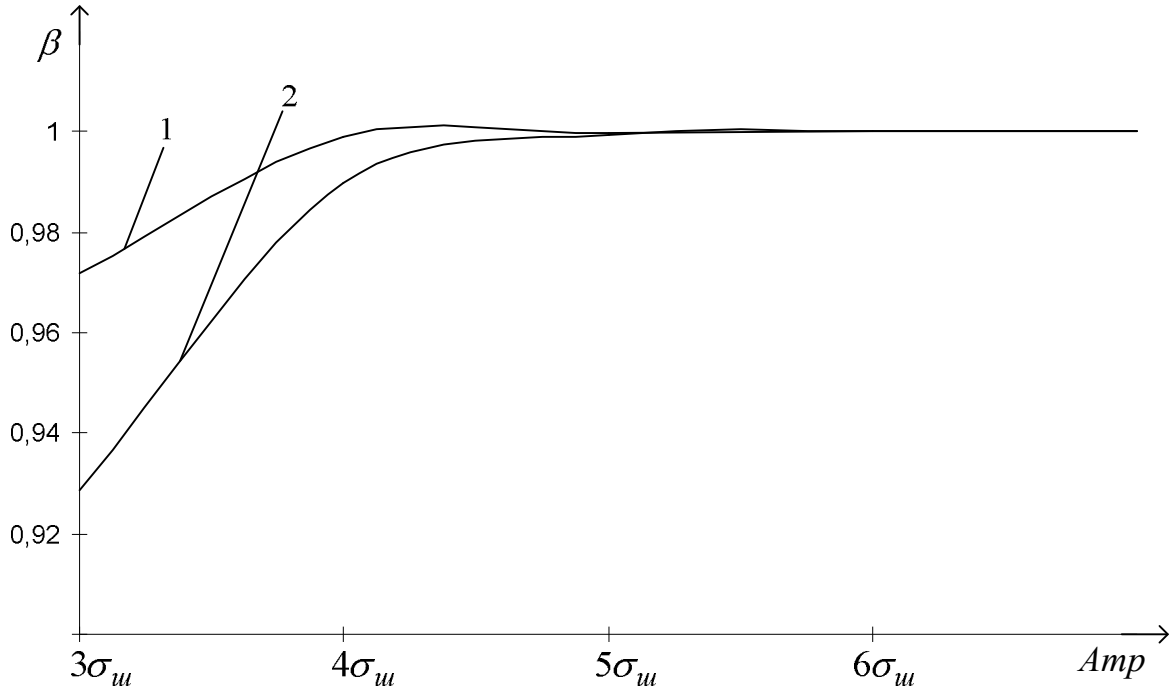


Рис. 3.8. Зависимость оценок вероятности правильного обнаружения $\beta = f(Amp)$

Так же исследуются зависимости оценок вероятности правильного обнаружения β и вероятности ошибки первого рода α от места расположения аномальных значений в выборке нестационарного случайного процесса Y_k с мультипликативной шумовой составляющей μ_k для различных моделей функций полезной составляющей λ_k, которые приведены соответственно в таб. 3.3 и в таб. 3.4.

В таб. 3.3 и таб. 3.4, оценки представлены для случая, когда мультипликативная шумовая составляющая μ_k имеет гауссовский закон плотности распределения вероятности. Аномальные значения составляют 5% от объема выборки N и их величина $Amp = 5\sigma_ш$.

Среднеквадратическое отклонение шумовой составляющей сигнала $\sigma_ш = 0{,}1$, априорно задаваемое значение вероятности ошибки первого рода $\alpha = 0{,}05$.

Анализ оценок вероятности правильного обнаружения β, которые представлены в таб. 3.3, показывает, что использование способа обнаружения аномальных значений с адаптацией порогового значения позволяет получать, оценки вероятности правильного обнаружения β значительно выше, чем при применении способа без адаптации порогового значения. Например, если модель функции полезной составляющей λ_k является гармонической, то для случая, когда аномальные значения расположены в начале выборки, оценки вероятности правильного обнаружения β возрастает примерно на 80%.

Таблица 3.3

Значение оценок вероятности правильного обнаружения

Расположение аномальных значений	Гармоническая функция		Экспоненциальная функция		Составная функция	
	с адаптацией порогового значения	без адаптации порогового значения	с адаптацией порогового значения	без адаптации порогового значения	с адаптацией порогового значения	без адаптации порогового значения
В начале выборки	0,954	0,180	0,903	0,228	0,974	0,211
В середине выборки	0,995	0,298	1	0,339	0,998	0,307
В конце выборки	0,954	0,296	1	0,313	0,975	0,277
Равномерно расположены по всей выборке	1	0,645	0,997	0,736	1	0,712

Если же модель функции полезной составляющей λ_k является экспоненциальной, при расположении аномальных значений в начале выборке Y_k, оценки вероятности правильного обнаружения β увеличиваются на 75%, и примерно на 68% и при расположении

аномальных значений в середине и в конце исследуемого нестационарного процесса Y_k.

Анализ показывает, что оценки вероятности правильного обнаружения β, при применении адаптации порогового, стремятся к единице независимо от места расположения аномальных значений в выборке исследуемого процесса Y_k и модели функций полезной составляющей λ_k.

Результаты, которые представлены в таб. 3.4, показывают, что при использовании способа обнаружения аномальных значений с адаптацией порогового значения, оценки вероятности ошибки первого рода α не превосходят априорно задаваемого значения $\alpha = 0{,}05$ при любом расположении аномальных значений для всех представленных моделей функций полезной составляющей λ_k нестационарного случайного процесса Y_k.

Таблица 3.4

Значение вероятности ошибки первого рода

Расположение аномальных значений	Гармоническая функция	Экспоненциальная функция	Составная функция
В начале выборки	0,048	0,044	0,045
В середине выборки	0,046	0,043	0,045
В конце выборки	0,046	0,042	0,045
Равномерно расположены по всей выборке	0,040	0,018	0,036

Поскольку оценки правильного обнаружения β практически не зависят от расположения аномальных значений в выборке нестационарного случайного процесса Y_k, а оценки вероятности ошибки первого рода α не превосходят своего априорно задаваемого значения, то исследуется эффективность применения способа обнаружения

аномальных значений с адаптацией порогового значений в зависимости от закона распределения мультипликативной шумовой составляющей μ_k.

В качестве модели полезной составляющей λ_k рассматриваются следующие нормированные функции: экспоненциальная, гармоническая и составная функция. Мультипликативная шумовая составляющая μ_k имеет следующие законы плотности распределения вероятности: гауссовский, равномерный и рэлеевский. Аномальная составляющая ε_k представляет собой одиночные значения, расположенные: равномерно, в начале, в середине и в конце реализации нестационарного случайного процесса Y_k. Количество аномальных значений составляет 5% от объема выборки N. Величина аномальных значений равна $Amp = 5\sigma_ш$, где $\sigma_ш$ – среднеквадратическое отклонение мультипликативной шумовой составляющей μ_k, которая равна $\sigma_ш = 0{,}1$. Значение вероятности ошибки первого рода априорно выбирается равной $\alpha = 0{,}05$.

В результате проведенных исследований с адаптацией порогового значения были получены оценки вероятности правильного обнаружения β в зависимости от расположения аномальных значений при гауссовском, равномерном и рэлеевском законах плотности распределения вероятности мультипликативной шумовой составляющей μ_k. Результаты проведенных исследований, обобщены и приведены в таб. 3.5, таб. 3.6 и в таб. 3.7.

Оценки вероятности правильного обнаружения β, в зависимости от места расположения аномальных значений, для экспоненциальной модели функции составляющей λ_k представлены в таб. 3.5, для модели функции сложного λ_k в таб. 3.6 и для модели функции гармонического λ_k в таб. 3.7.

Таблица 3.5

Значение оценок вероятности правильного обнаружения для экспоненциальной модели полезной составляющей

Расположение аномальных	Равномерный закон плотности	Гауссовский закон плотности	Рэлеевский закон плотности

значений	распределения вероятности	распределения вероятности	распределения вероятности
В начале выборки	0,928	0,909	0,968
В середине выборки	1	0,990	1
В конце выборки	0,991	1	0,991
Равномерно по всей выборке	1	0,99	0,990

Таблица 3.6

Значение оценок вероятности правильного обнаружения для сложной модели полезной составляющей

Расположение аномальных значений	Равномерный закон плотности распределения вероятности	Гауссовский закон плотности распределения вероятности	Рэлеевский закон плотности распределения вероятности
В начале выборки	0,976	0,979	0,967
В середине выборки	1	0,990	0,992
В конце выборки	0,990	0,977	0,976
Равномерно по всей выборке	1	0,990	0,990

Таблица 3.7

Значение оценок вероятности правильного обнаружения для гармонической модели полезной составляющей

Расположение номальных значений	Равномерный закон плотности распределения вероятности	Гауссовский закон плотности распределения вероятности	Рэлеевский закон плотности распределения вероятности
В начале выборки	0,967	0,959	0,965
В середине выборки	0,993	0,99	0,990
В конце выборки	0,981	0,967	0,991
Равномерно по всей выборке	1	0,990	1

Анализ результатов, представленных в таблицах 3.5, 3.6 и 3.7 показывает, что для рассмотренных нестационарных случайных процессов, оценки вероятности правильного обнаружения β практически не зависят

от места расположения аномальных значений и от закона распределения мультипликативной шумовой составляющей μ_k.

Аналогичные результаты получаются и при исследовании эффективности, предлагаемого способа обнаружения аномальных значений, с другими моделями функций полезной составляющей λ_k, которые представлены в таб. 3.8 для экспоненциальной, в таб. 3.9 для составной, в таб. 3.10 для гармонической моделей функций.

Таблица 3.8

Значение оценок вероятности правильного обнаружения для экспоненциальной модели полезной составляющей

Расположение аномальных значений	Равномерный закон плотности распределения вероятности	Гауссовский закон плотности распределения вероятности	Рэлеевский закон плотности распределения вероятности
В начале выборки	0,022	0,044	0,039
В середине выборки	0,023	0,042	0,030
В конце выборки	0,023	0,042	0,038
Равномерно по всей выборке	0,012	0,017	0,014

Таблица 3.9

Значение оценок вероятности правильного обнаружения для сложной модели полезной составляющей

Расположение аномальных значений	Равномерный закон плотности распределения вероятности	Гауссовский закон плотности распределения вероятности	Рэлеевский закон плотности распределения вероятности
В начале выборки	0,031	0,045	0,037
В середине выборки	0,031	0,044	0,044
В конце выборки	0,031	0,045	0,037
Равномерно по всей выборке	0,022	0,036	0,029

Таблица 3.10

Значение оценок вероятности правильного обнаружения для гармонической модели полезной составляющей

Расположение аномальных значений	Равномерный закон плотности распределения вероятности	Гауссовский закон плотности распределения вероятности	Рэлеевский закон плотности распределения вероятности
В начале выборки	0,028	0,044	0,045
В средине выборки	0,027	0,043	0,041
В конце выборки	0,024	0,046	0,039
Равномерно по всей выборке	0,019	0,039	0,030

Анализ результатов, представленных в таб. 3.8, в таб. 3.9 и в таб. 3.10 показывает, что оценки вероятности ошибки первого рода α не превосходит априорно задаваемого значения $\alpha = 0{,}05$ независимо от места расположении аномальных значений в выборке нестационарного случайного процесса. Оценки вероятности ошибки первого рода α при этом также практически не зависит от закона распределения мультипликативной шумовой составляющей μ_k.

3.3 Сравнительный анализ эффективности адаптивного способа обнаружения аномальных значений для нестационарных случайных процессов с аддитивной и мультипликативной шумовой составляющей

В данном подразделе исследуется влияние модели исследуемого процесса на эффективность предложенного в работе способа с применением адаптации порогового значения. Для этого рассматриваются нестационарные случайные процессы с аддитивной η_k и мультипликативной шумовой составляющей μ_k [125].

Исследуемые нестационарные случайные процессы представлены моделями вида (2.2) и (2.3). В качестве модели полезной составляющей λ_k

исходного процесса Y_k используются следующие функции: параболическая, экспоненциальная, гармоническая и составная.

В качестве модели аддитивной шумовой составляющей η_k и мультипликативной шумовой составляющей μ_k используется гауссовский закон плотности распределения вероятности.

Аномальная составляющая ε_k представляет собой одиночные значения, равномерно распределенные по всей выборке исходного нестационарного случайного процесса Y_k. Количество аномальных значений составляет 5% от объема выборки N. Величина аномальных значений Amp: $3\sigma_ш, 4\sigma_ш, 5\sigma_ш, 6\sigma_ш, 7\sigma_ш$ ($\sigma_ш = 0{,}1$ - среднеквадратическое отклонение мультипликативной шумовой составляющей η_k и μ_k). Значение вероятности ошибки первого рода априорно выбирается равной $\alpha = 0{,}05$.

В результате проведенных исследований, с адаптацией порогового значения получены зависимости $\beta = f(Amp)$ с аддитивной η_k и мультипликативной шумовой составляющей μ_k. На рисунке 3.9 представлены графики зависимости $\beta = f(Amp)$ для экспоненциальной модели функции полезной составляющей λ_k: график 1 – мультипликативная шумовая составляющая; график 2 – аддитивная шумовая составляющая.

Анализ результатов, представленных на рисунке 3.9, показывает, что при величине $Amp = 3\sigma_ш$, оценки вероятности правильного обнаружения β для мультипликативной модели нестационарного случайного процесса больше, чем при аддитивной модели примерно на 10%. При введении адаптации порогового значения, согласно выражению (2.33), для данной модели полезной составляющей, оценки вероятности правильного обнаружения β при величине аномальных значений $Amp = 4\sigma_ш$ достигают

$0{,}940 - 0{,}981$. Однако при $Amp > 5\sigma_ш$ оценки вероятности правильного обнаружения β стремятся к единице независимо от модели шумовой составляющей исследуемого нестационарного случайного процесса Y_k.

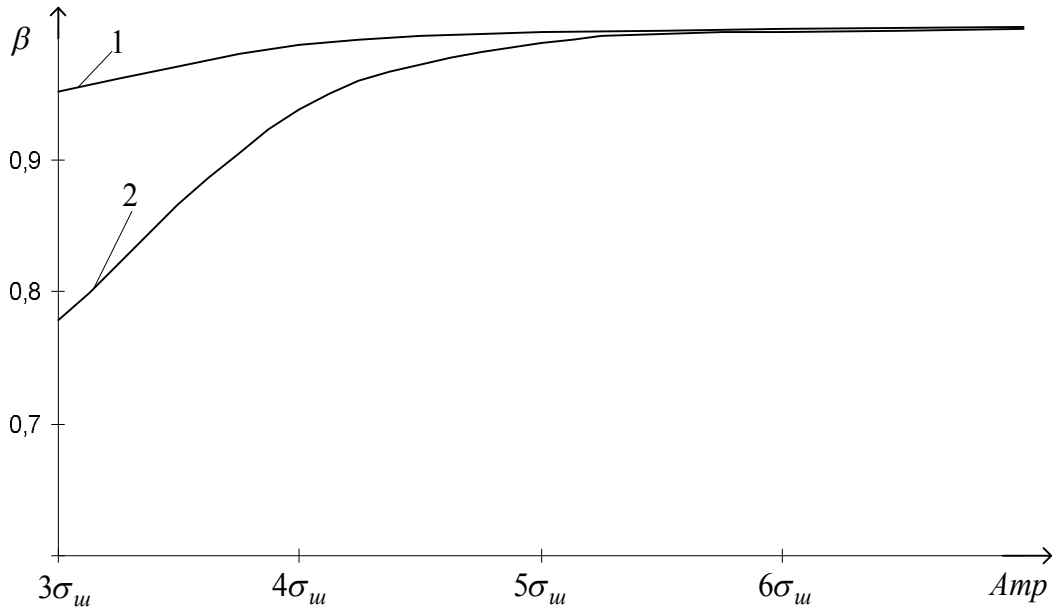


Рис. 3.9. Зависимость оценок вероятности правильного обнаружения $\beta = f(Amp)$

На рисунке 3.10 представлены графики зависимостей оценок вероятности ошибки первого рода α от величины аномальных значений для: аддитивной модели шумовой составляющей – график 1; мультипликативной модели шумовой составляющей – график 2.

Анализ результатов, представленных на рисунке 3.10, показывает, что при введении адаптации порогового значения оценки вероятность ошибки первого рода α не превосходит априорно задаваемого $\alpha = 0{,}05$. Причем, для мультипликативной модели шумовой составляющей оценки α на 57% меньше, чем для аддитивной модели шумовой составляющей.

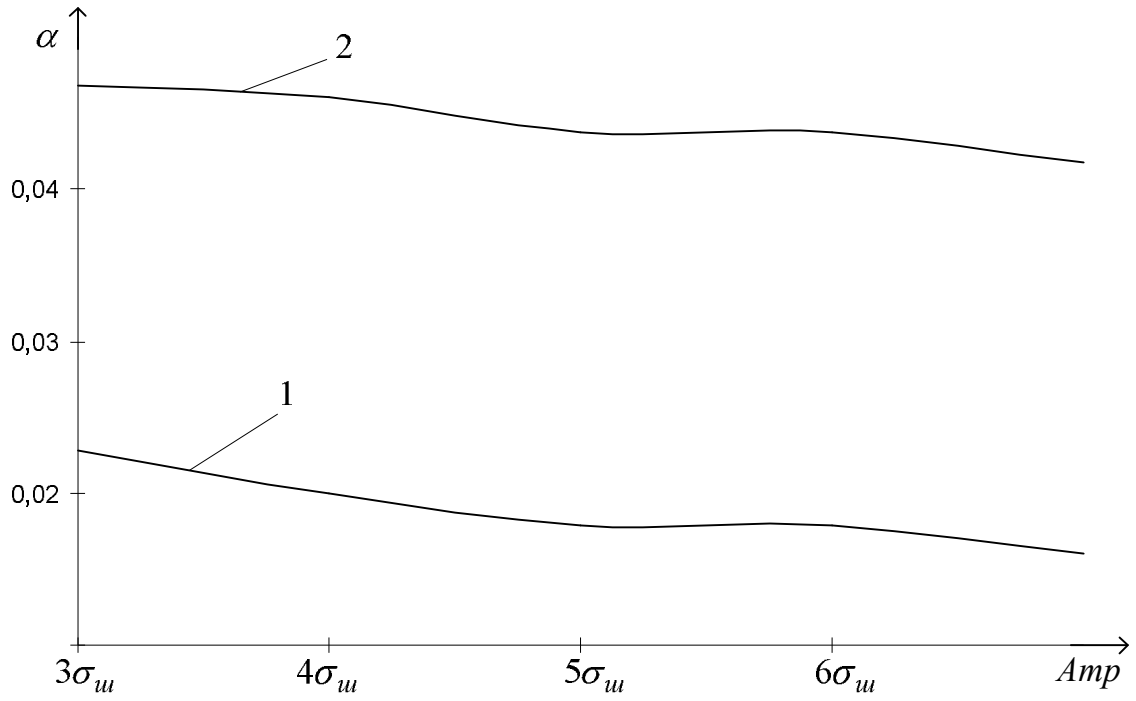


Рис. 3.10. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$

Аналогичная зависимость наблюдается и для других моделях функции полезной составляющей нестационарного случайного процесса [125].

3.4 Анализ эффективности адаптивного способа обнаружения аномальных значений для стационарных случайных процессов

Рассмотрим возможность применения способа обнаружения аномальных значений с адаптацией порогового значения для случая, когда анализируемый сигнал является стационарным случайным процессом [126, 127].

Сигнал Y_k является стационарным случайным процесса, который представлен моделями: гауссовским (2.6), равномерным (2.7) и рэлеевским (2.8) законами плотности распределения вероятности. При проведении исследований рассматривается случай, когда аномальные значения

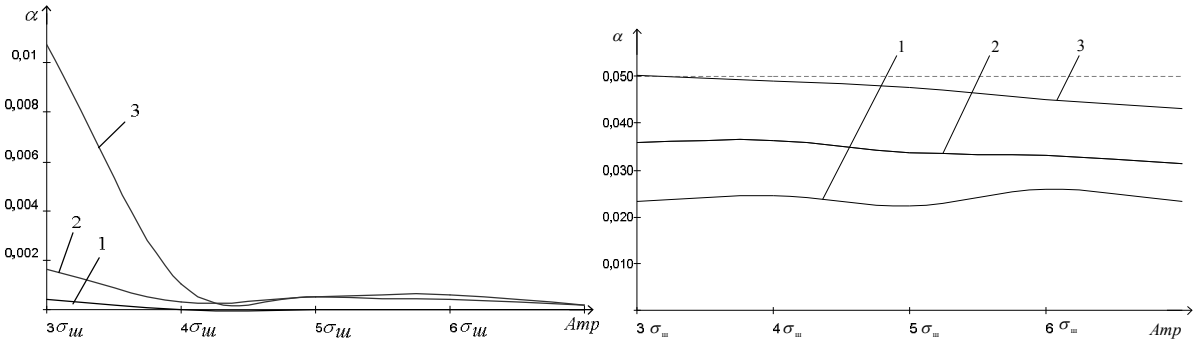
равномерно расположены по всей выборке стационарного случайного процесса. Величина аномальных значений Amp: $3\sigma_ш, 4\sigma_ш, 5\sigma_ш, 6\sigma_ш, 7\sigma_ш$ ($\sigma_ш = 0{,}1$ – среднеквадратическое отклонение шумовой составляющей η_k). Количество аномальных значений составляет 5% от объема исходного процесса Y_k. Значение вероятности ошибки первого рода априорно фиксировано $\alpha = 0{,}05$.

В результате проведения имитационного моделирования получены зависимости оценок вероятности ошибки первого рода α и вероятность правильного обнаружения β, как функции от амплитуды аномальных значений, для способа обнаружения аномальных значений без адаптации и с адаптацией порогового значения, которые представлены на рисунке 3.11 (3.11а – без адаптации порогового значения, 3.11б – с адаптацией порогового значения).

Полученные зависимости, соответствуют: график 1 – равномерному; график 2 – рэлеевскому; график 3 – гауссовскому законам плотности распределения вероятности шумовой составляющей η_k.

Анализ функциональных зависимостей, представленных на рисунке 3.11б показывает, что при использовании способа с адаптацией порогового значения, оценки вероятности ошибки первого рода α не превышают априорно фиксированного значения $\alpha = 0.05$ независимо от закона распределения шумовой составляющей η_k. Как видно из рисунке 3.11б, для равномерного и рэлеевского законов плотности распределения вероятности шумовой составляющей η_k, оценки вероятности ошибки первого рода α практически не зависят от величины аномальных значений Amp (рис. 3.11б, график 2 и график 1). Для гауссовского закона плотности распределения вероятности шумовой составляющей оценки α уменьшаются с увеличением величины аномальных значений. Например,

при величине аномальных значений $Amp = 7\sigma_ш$ оценки вероятности ошибки первого рода α уменьшаются на 16% (рис. 3.11б, график 3).



а) б)

Рис. 3.11. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$ а) способ обнаружения аномальных значений без адаптации порового значения; б) способ обнаружения аномальных значений с адаптации порового значения при $\alpha = 0{,}05$

На рисунке 3.12, представлены зависимости оценок вероятности правильного обнаружения $\beta = f(Amp)$ для способа с адаптацией и без адаптации порогового значения. Графики 1, $1''$– равномерный; графики 2, $2''$ – рэлеевский; графики 3, $3''$ – гауссовский законы плотности распределения вероятности шумовой составляющей η_k соответственно с адаптацией и без адаптации порогового значения.

Анализ функциональных зависимостей представленных на рисунке 3.12 показывает, что в случае использования способа обнаружения аномальных значений с адаптацией порогового значения при увеличении величины аномальных значений Amp, оценки вероятности правильного обнаружения β возрастают для каждого закона плотности распределения вероятности шумовой составляющей η_k. Причем, при величине аномальных значений порядка $Amp \geq 5\sigma_ш$, разность в оценках вероятности правильного обнаружения β для гауссовского законаплотности распределения вероятности шумовой оставляющей η_k (график 3 и график

$3''$) составляет порядка 67%, для равномерного закона плотности распределения вероятности шумовой составляющей η_k (график 1 и график $1''$) – 63%, для рэлеевского закона плотности распределения вероятности шумовой составляющей η_k (график 2 и график $2''$) – 54%.

Так же из анализа функциональных зависимостей $\beta = f(Amp)$, которые представлены на рисунке 3.12, следует, что при использовании адаптации порогового значения, отличие в оценках вероятности правильного обнаружения β при $Amp \geq 3\sigma_ш$ в зависимости от закона плотности распределения шумовой вероятности составляющей η_k составляют порядка 6%,. При дальнейшем увеличении $Amp \geq 5\sigma_ш$, оценки вероятности правильного обнаружения β, для данного случая, практически не зависят от величины аномальных значений Amp и стремятся к единице.

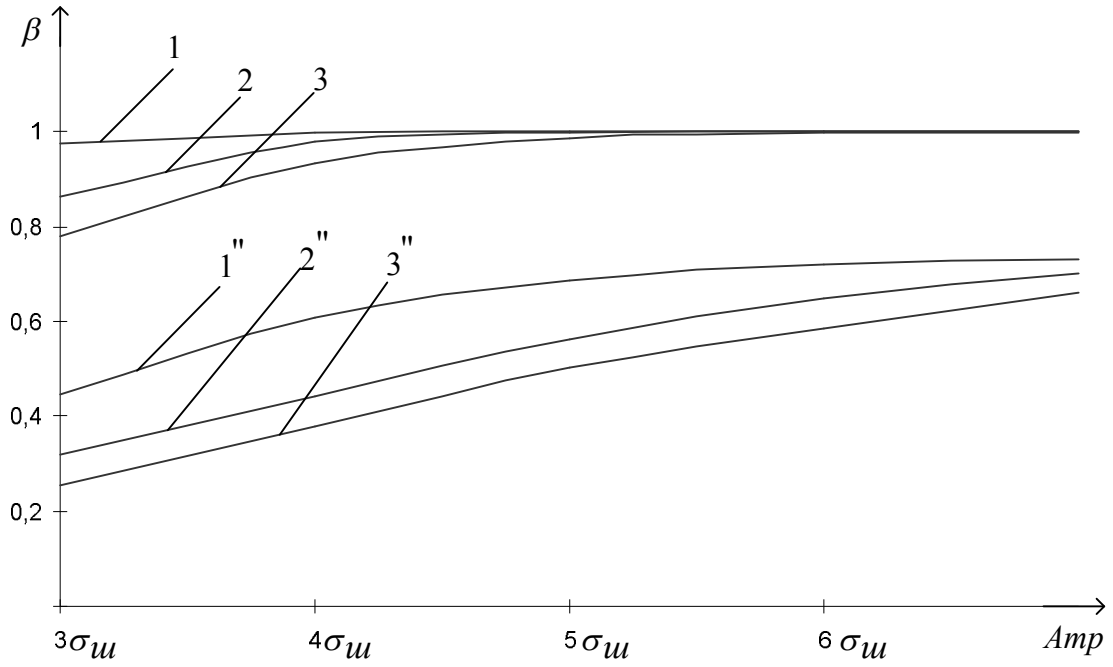


Рис. 3.12. Зависимость оценок вероятности правильного обнаружения $\beta = f(Amp)$

Аналогичные результаты оценок вероятности правильного обнаружения β, при использовании способа с адаптацией порогового значения, получены и при других априорно фиксированных значениях вероятности ошибки первого рода $\alpha = 0{,}1$ и $\alpha = 0{,}2$ [126, 127].

Для определения эффективности способа обнаружения аномальных значений с адаптации порогового значения получены зависимости вероятности ошибки первого рода α и вероятности правильного обнаружения β при различном расположении аномальных значений в выборке исследуемого стационарного случайного процесса Y_k, где $k = \overline{1, N}$. Стационарный случайный процесс представлен равномерным, рэлеевским и гауссовским законами плотности распределения вероятности. Среднеквадратическое отклонение шумовой составляющей $\sigma_ш = 0{,}1$.

Величина аномальных значений выбирается фиксированной и равной $Amp = 5\sigma_ш$, поскольку ранее показано, что при значении величины аномальных значений порядка $Amp = 5\sigma_ш$, оценки вероятности правильного обнаружения β стремятся к единице. Аномальные значения составляют 5% от объема выборки N и могут располагаться равномерно, в начале, в середине или в конце исходного процесса Y_k .

В таб. 3.11 представлены оценки вероятности правильного обнаружения β в зависимости от места расположения аномальных значений для различных законов плотности распределения вероятности шумовой составляющей η_k, полученные при использовании способа без адаптации порогового значения. В таб. 3.12 представлены оценки вероятности правильного обнаружения β в зависимости от места расположения аномальных значений для различных законов плотности распределения вероятности шумовой составляющей η_k, полученные при

использовании способа с адаптации порогового значения. При этом априорно задаваемое значение ошибки первого рода $\alpha = 0{,}05$.

Таблица 3.11

Значение оценок вероятности правильного обнаружения

Расположение аномальных значений	Гауссовский закон плотности распределения вероятности	Равномерный закон плотности распределения вероятности	Рэлеевский закон плотности распределения вероятности
В начале выборки	0,205	0,243	0,215
В середине выборки	0,269	0,293	0,302
В конце выборки	0,270	0,288	0,310
Равномерно расположены по всей выборке	0,502	0,561	0,686

Таблица 3.12

Значение оценок вероятности ошибки первого рода

Расположение аномальных значений	Гауссовский закон плотности распределения вероятности	Равномерный закон плотности распределения вероятности	Рэлеевский закон плотности распределения вероятности
В начале выборки	0,821	0,933	0,866
В середине выборки	0,928	0,994	0,973
В конце выборки	0,934	0,994	0,968
Равномерно расположены по всей выборке	0,986	0,999	0,997

Анализ результатов, представленных в таб. 3.11 и в таб. 3.12 показывает, что использование адаптации порогового значения, в предлагаемом способе обнаружения аномальных значений, позволяет получить оценки вероятности правильного обнаружения β намного выше, чем при использовании способа без адаптации порогового значения. Например, оценки вероятности правильного обнаружения β для гауссовского закона плотности распределения вероятности шумовой составляющей η_k, при расположении аномальных значений в начале стационарного случайного процесса Y_k, возрастают примерно на 75%, а

для равномерного и рэлеевского законов плотности распределения вероятности шумовой составляющей η_k при равномерном расположении аномальных значений оценки вероятности правильного обнаружения β возрастают в среднем на 70%.

Анализ оценок правильного обнаружения β, которые представлены в таб. 3.12 показывает, что они не зависят от закона плотности распределения вероятности шумовой составляющей стационарного случайного процесса Y_k и практически не зависят от места расположения аномальных значений в исследуемой реализации. Например, в случае, когда аномальные значения расположены в начале и равномерно по всей реализации исследуемого процесса, максимальная относительная разница в выборных значениях вероятности правильного обнаружения составляет: при гауссовском – 16%; при рэлеевском – 13%; при равномерном – 6,5% законах плотности распределения вероятности шумовой составляющей.

В результате проведенных исследований эффективности способа обнаружения аномальных значений с адаптацией порогового значения, при тех же условиях, получены значения оценок вероятности ошибки первого рода α, которые представлены в таб. 3.13.

Таблица 3.13

Значение оценок вероятности ошибки первого рода

Расположение аномальных значений	Гауссовский закон плотности распределения вероятности	Равномерный закон плотности распределения вероятности	Рэлеевский закон плотности распределения вероятности
В начале выборки	0,044	0,024	0,048
В середине выборки	0,048	0,026	0,043
В конце выборки	0,049	0,026	0,044
Равномерно расположены по всей выборке	0,049	0,026	0,033

Из анализа данных таблицы 3.13 следует, что при использовании адаптации порогового значения, оценки вероятности ошибки первого рода

α не превосходят априорно задаваемого значения $\alpha = 0{,}05$ независимо от места расположения аномальных значений и закона плотности распределения вероятности шумовой составляющей исследуемого процесса Y_k.

Таким образом, предлагаемый в работе адаптивный способ обнаружения аномальных значений при анализе нестационарных случайных процессов можно применить и для случая, когда исследуемый процесс является стационарным.

ГЛАВА 4 Анализ эффективности способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа нестационарного случайного процесса

В работе, на основе имитационного моделирования определены основные параметры способа обнаружения аномальных значений (количество уровней анализа P и длина d), при которых достигается минимум оценок вероятности ошибки первого рода α и максимум оценок вероятности правильного обнаружения β [118].

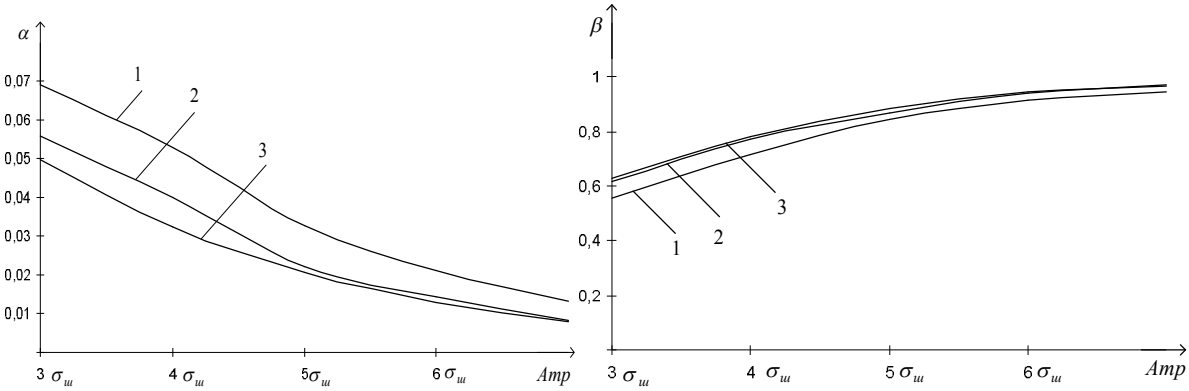
4.1 Исследование эффективности способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа нестационарного случайного процесса с аддитивной шумовой составляющей

При количестве уровней анализа $P = 25$ и $d = 7$ получены зависимости оценок вероятности правильного обнаружения β и оценок вероятности ошибки первого рода α от величины аномальных значений Amp. Т.е для модели (2.2) построены зависимости $\beta = f(Amp)$ и $\alpha = f(Amp)$ при равномерном расположении аномальных значений в выборке нестационарного случайного процесса с аддитивной шумовой составляющей [118, 120, 128].

В качестве модели функции полезной составляющей λ_k выбираются нормированные функции: гармоническая, экспоненциальная и составная. Аддитивная шумовая составляющая η_k имеет гауссовский, равномерный и рэлеевский законы плотности распределения вероятности. Аномальные значения составляют 5% от выборки N исходного нестационарного случайного процесса Y_k, где $k = \overline{1, N}$. Величина аномальных значений

Amp принимает значения $3\sigma_ш, 4\sigma_ш, 5\sigma_ш, 6\sigma_ш, 7\sigma_ш$, где $\sigma_ш = 0{,}1$ - среднеквадратическое отклонение аддитивной шумовой составляющей.

На рисунке 4.1 представлены графики зависимости оценок вероятности ошибки первого рода α (а) и оценок вероятности правильного обнаружения β (б) для случая, когда функция полезной составляющей λ_k является гармонической. Графики на рисунке 4.1 соответствуют случаям, когда аддитивная шумовая составляющая η_k имеет следующие законы плотности распределения вероятности: график 1 – рэлеевский; график 2 – гауссовский; график 3 – равномерный.

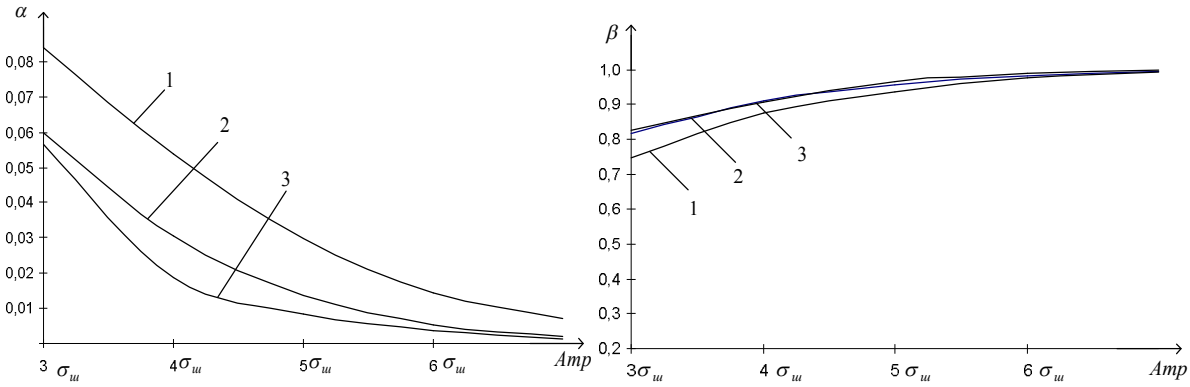


а) б)

Рис.4.1. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$ и вероятности правильного обнаружения $\beta = f(Amp)$ для гармонической функции полезной составляющей

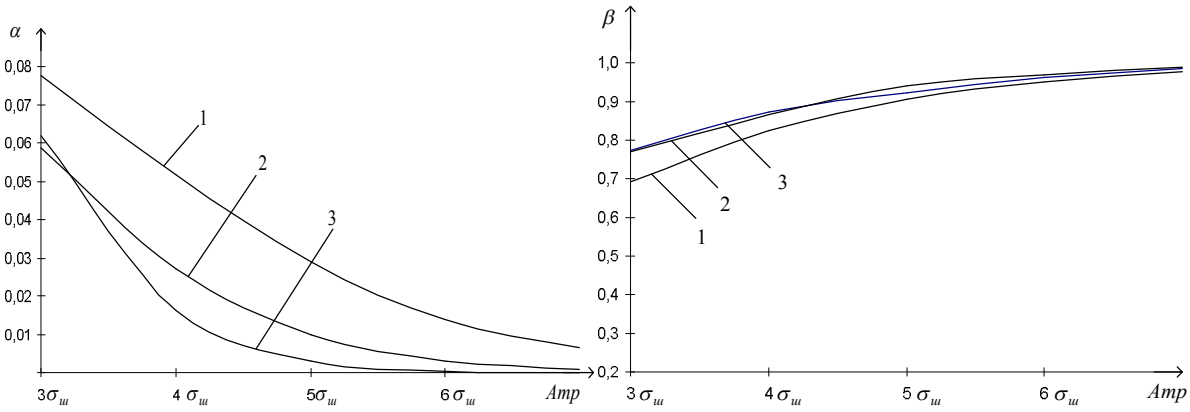
Аналогично на рисунках 4.2 и 4.3 представлены зависимости оценок вероятности ошибки первого рода $\alpha = f(Amp)$ (а) и вероятности правильного обнаружения $\beta = f(Amp)$ (б) для экспоненциальной и составной функций полезной составляющей λ_k нестационарного случайного процесса Y_k. Графики 1, 2 и 3 на рис. 4.2 и 4.3 так же соответствуют рэлеевскому, гауссовскому и равномерному законам плотности распределения вероятности аддитивной шумовой составляющей η_k нестационарного случайного процесса Y_k.

Анализ результатов представленных на рисунке 4.1 показывает, что при величине аномальных значений $Amp = 3\sigma_ш$, для гармонической функции полезной составляющей λ_k, оценки вероятности ошибки первого рода α (рис. 4.1,а) достигают своего максимального значения $\alpha = 0{,}055 - 0{,}069$, а оценки вероятности правильного обнаружения β (рис. 4.1б) минимального значения $\beta = 0{,}570 - 0{,}611$, в зависимости от закона плотности распределения вероятности аддитивной шумовой составляющей η_k нестационарного случайного процесса Y_k.



а) б)

Рис.4.2. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$ и оценок вероятности правильного обнаружения $\beta = f(Amp)$ для экспоненциальной функции полезной составляющей



а) б)

Рис.4.3. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$ и оценок вероятности правильного обнаружения $\beta = f(Amp)$ для составной функции полезной составляющей

Для экспоненциальной и составной функций полезной составляющей процесса λ_k (рис. 4.2 и 4.3), при том же величине аномальных значений $Amp = 3\sigma_ш$ оценки вероятности ошибки первого рода α (рис. 4.2а и 4.3а) и вероятности правильного обнаружения β (рис. 4.2б и 4.3б) составляют соответственно: $\alpha = 0{,}057 - 0{,}084$, $\beta = 0{,}749 - 0{,}815$ для экспоненциальной функции полезной составляющей λ_k и $\alpha = 0{,}059 - 0{,}078$, $\beta = 0{,}694 - 0{,}773$ для составной функции полезной составляющей λ_k, в зависимости от закона плотности распределения вероятности аддитивной шумовой составляющей исследуемого нестационарного случайного процесса η_k.

Анализ графиков, представленных на рисунках 4.1, 4.2 и 4.3 показывает, что при величине аномальных значений $Amp > 5\sigma_ш$ для гармонической, экспоненциальной и составной функции полезной составляющей λ_k оценки вероятности ошибки первого рода принимают значения, $0{,}032 < \alpha < 0{,}022$, а оценки вероятности правильного обнаружения $0{,}847 < \beta < 0{,}869$ в зависимости от закона плотности распределения вероятности аддитивной шумовой составляющей η_k нестационарного случайного процесса.

Можно считать, что выше перечисленные оценки практически не зависят от закона плотности распределения вероятности аддитивной шумовой составляющей η_k. При гармонической функции полезной составляющей λ_k (рис. 4.1б, графики 2, 3), экспоненциальной функции полезной составляющей λ_k (рис. 4.2,б, графики 2, 3), составной функции λ_k (рис. 4.3б, графики 2, 3), оценки вероятности правильного обнаружения β практически совпадают, когда аддитивная шумовая составляющая η_k имеет гауссовский и равномерный законы плотности распределения

вероятности при любых величинах аномальных значений Amp. Например, для гармонической функции полезной составляющей λ_k, (рис. 4.1б, графики 2, 3), при $Amp = 3\sigma_ш$ разница в оценках 11% , а при $Amp = 6\sigma_ш$ разница уже величину 7%. Для составной функции полезной составляющей нестационарного случайного процесса λ_k (рис. 4.3б, графики 2, 3), при $Amp = 3\sigma_ш$ разница порядка 5%, а при $Amp = 6\sigma_ш$ разница в оценках уже порядка 1%.

Анализ представленных зависимостей (рис. 4.1, 4.2 и 4.3) показывает, что с увеличением Amp оценки вероятности ошибки первого рода α уменьшаются, а оценки вероятности правильного обнаружения β стремятся к единице в зависимости от закона плотности распределения вероятности аддитивной шумовой составляющей η_k.

Из анализа результатов исследований оценок вероятности ошибки первого рода $\alpha = f(Amp)$ (рис. 4.1а, 4.2а и 4.3а) и вероятности правильного обнаружения $\beta = f(Amp)$ (рис. 4.1б, 4.2б, 4.3б) видно, что при увеличении величины аномальных значений Amp, вероятность их обнаружения увеличивается, в зависимости от закона плотности распределения вероятности аддитивной шумовой составляющей η_k нестационарного случайного процесса Y_k.

При гауссовском законе плотности распределения вероятности аддитивной шумовой составляющей η_k для всех рассматриваемых функций полезной составляющей λ_k получены минимальные оценки вероятности ошибки первого рода α и максимальные оценки вероятности правильного обнаружения β.

При увеличении величины аномальных значений Amp оценки вероятности ошибки первого рода α и оценки вероятности правильного

обнаружения β практически не зависят от закона плотности распределения вероятности аддитивной шумовой составляющей η_k.

Рассмотрены зависимости оценок вероятности ошибки первого рода α и вероятности правильного обнаружения β относительно функции полезной составляющей λ_k нестационарного случайного процесса Y_k [128].

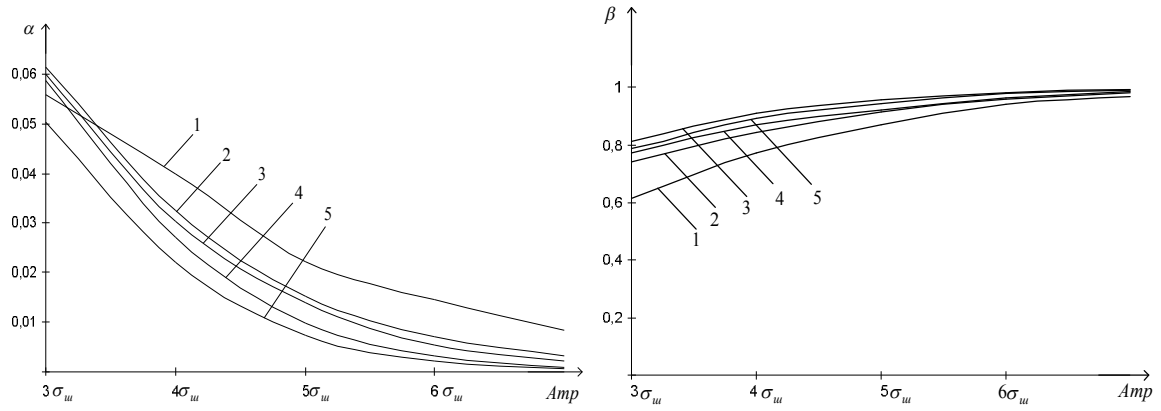
На рисунке 4.4 представлены зависимости оценок вероятности ошибки первого рода α и вероятности правильного обнаружения β при гауссовском законе плотности распределения вероятности аддитивной шумовой составляющей η_k для следующих нормированных функций полезной составляющей λ_k: график 1 – гармонической, график 2 – сложной, график 3 – экспоненциальной, график 4 – составной, график 5 – параболической.

Для рассмотренных моделей функций полезной составляющей λ_k, на рисунках 4.5 и 4.6 представлены зависимости оценок вероятности ошибки первого рода α и вероятности правильного обнаружения β при равномерном и рэлеевском законах плотности распределения вероятности аддитивной шумовой составляющей η_k.

Анализ зависимостей, представленных на рис. 4.4а, при гауссовском законе плотности распределения вероятности аддитивной шумовой составляющей η_k показывает, что оценки вероятности ошибки первого рода α при $Amp = 3\sigma_ш$ в среднем для рассмотренных функций полезной составляющей λ_k, составляют величину порядка $\alpha = 0{,}057$.

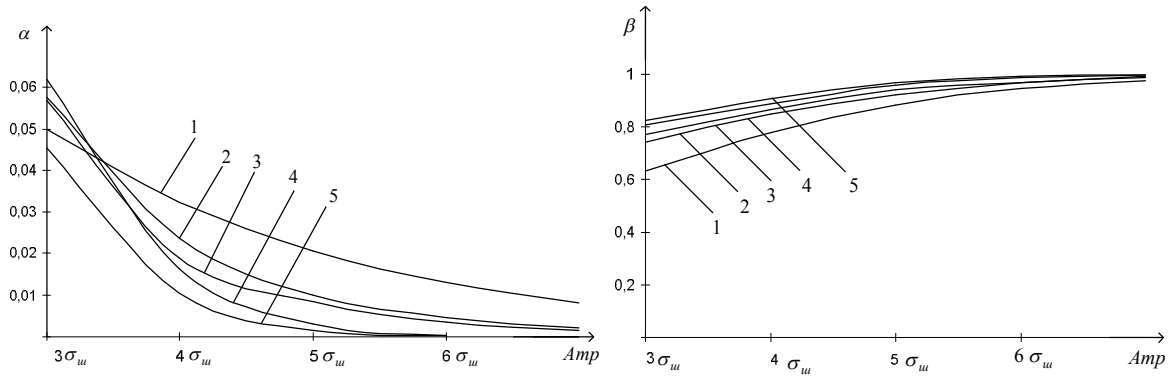
При равномерном законе плотности распределения вероятности аддитивной шумовой составляющей η_k оценки вероятности ошибки первого рода α при величине аномальных значений $Amp = 3\sigma_ш$ составляют $\alpha = 0{,}053$ (рис. 4.5а). При рэлеевском законе плотности

распределения вероятности - $\alpha = 0{,}076$ (рис. 4.6а). С увеличением величины аномальных значений, оценки вероятности ошибки первого рода α стремятся к своим минимальным значениям независимо о функции полезной составляющей λ_k (рис. 4.4а, 4.5а, 4.6а).



а) б)

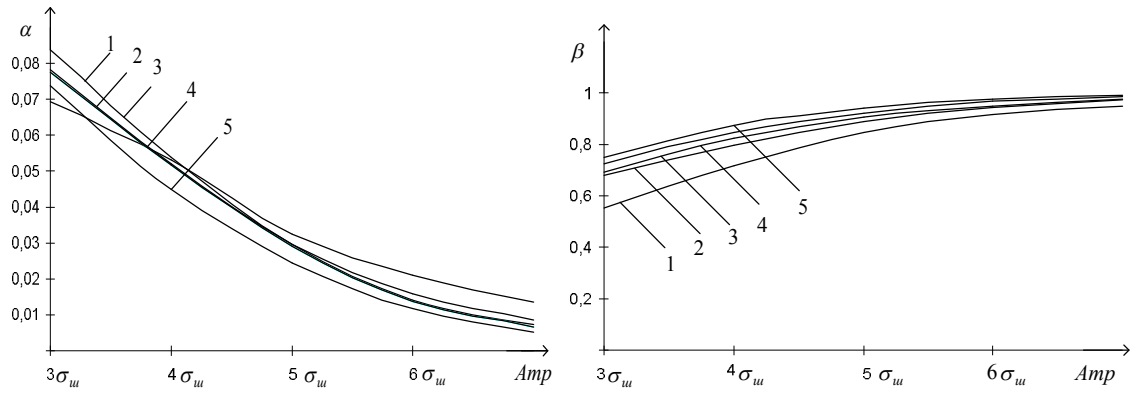
Рис. 4.4. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$ и вероятности правильного обнаружения $\beta = f(Amp)$ при гауссовском законе плотности распределения вероятности аддитивной шумовой составляющей



а) б)

Рис. 4.5. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$ и вероятности правильного обнаружения $\beta = f(Amp)$ при равномерном законе плотности распределения вероятности аддитивной шумовой составляющей

Из анализа рисунка 4.4б следует, что оценки вероятности правильного обнаружения β при величине аномальных значений $Amp = 3\sigma_ш$ составляют в среднем $\beta = 0{,}775$, а при величине аномальных значений $Amp > 5\sigma_ш$ в среднем – $\beta = 0{,}972$.



а) б)

Рис. 4.6. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$ и вероятности правильного обнаружения $\beta = f(Amp)$ при рэлеевком законе плотности распределения вероятности аддитивной шумовой составляющей

Для случаев, когда аддитивная шумовая составляющая η_k нестационарного случайного процесса представлена равномерным и рэлеевским законами плотности распределения вероятности, при $Amp = 3\sigma_ш$, оценки вероятности правильного обнаружения β составляют в среднем $\beta = 0{,}756$ и $\beta = 0{,}680$; при величине аномальных значений больше $Amp > 5\sigma_ш$ в среднем – $\beta = 0{,}972$ и $\beta = 0{,}900$ соответственно (рис. 4.5б и рис. 4.6б).

Таким образом, с увеличением величины аномальных значений оценки стремятся к единице практически независимо от функции полезной составляющей процесса.

Следовательно, для нестационарных случайных процессов с аддитивной шумовой составляющей, оценки вероятности ошибки первого рода α и вероятности правильного обнаружения β практически не зависят от функции полезной составляющей λ_k [118, 128].

Для нестационарных случайных процессов с аддитивной шумовой составляющей η_k получены зависимости оценок вероятности ошибки первого рода α и вероятности правильного обнаружения β от места расположения аномальных значений в исследуемом процессе, при тех же

значениях параметра способа обнаружения аномальных значений, т.е. количестве порогов $P = 25$ и длине участков нестационарности порядка $d = 8$ [129].

В качестве моделей функции полезной составляющей λ_k используются нормированные функции: гармоническая, экспоненциальная и составная. Аддитивная шумовая составляющая сигнала η_k представлена гауссовским, равномерным и рэлеевским законами плотности распределения вероятности. Величина аномальных значений выбирается $Amp = 5\sigma_ш$, где $\sigma_ш = 0{,}1$ – среднеквадратическое отклонение аддитивной шумовой составляющей нестационарного случайного процесса Y_k.

Результаты исследований для оценок вероятности ошибки первого рода α представлены в таб. 4.1 и для оценок вероятности правильного обнаружения в таб. 4.2.

Анализ результатов представленных в таб. 4.1 показывает, что в случае расположения аномальных значений в начале, в середине и равномерно в анализируемом процессе, оценки вероятности ошибки первого рода α при гармонической функции полезной составляющей нестационарного случайного процесса и гауссовом законе плотности распределения вероятности аддитивной шумовой составляющей составляют величину порядка $\alpha = 0{,}023$, т.е в этом случае практически не зависят от места расположения аномальных значений.

При расположении аномальных значений в конце выборки, оценки вероятности ошибки первого рода значительно меньше, т.е. порядка $\alpha \approx 0{,}009$. Причем такие результаты получены и при равномерном, и при рэлеевском законе плотности распределения вероятности аддитивной шумовой составляющей η_k. Например, в случае расположения аномальных значений в начале, в середине и равномерно в исследуемом процессе, оценки вероятности ошибки первого рода α при равномерном

законе плотности распределения вероятности аддитивной шумовой составляющей η_k: $\alpha = 0{,}020$, а при расположении аномальных значений в конце выборки, оценки вероятности ошибки первого рода $\alpha \approx 0{,}009$.

Таблица 4.1

Значение оценок вероятности ошибки первого рода

Расположение аномальных значений	Гармоническая функция			Экспоненциальная функция			Составная функция		
	Гауссовский закон	Равномерный закон	Рэлеевский закон	Гауссовский закон	Равномерный закон	Рэлеевский закон	Гауссовский закон	Равномерный закон	Рэлеевский закон
В начале выборки	0,022	0,019	0,032	0,007	0,004	0,024	0,011	0,003	0,028
В середине выборки	0,024	0,021	0,036	0,014	0,009	0,030	0,007	0,002	0,023
В конце выборки	0,009	0,009	0,016	0,014	0,008	0,028	0,008	0,002	0,022
Равномерно расположены по всей выборке	0,023	0,020	0,033	0,014	0,008	0,028	0,010	0,003	0,028

Для той же гармонической функции полезной составляющей λ_k (таб. 4.1), в случае расположения аномальных значений в начале, в середине и равномерно, в реализации исследуемого процесса, оценки вероятности ошибки первого рода α при рэлеевском законе плотности распределения вероятности аддитивной шумовой составляющей: $\alpha = 0{,}033$, а при расположении аномальных значений в конце выборки, оценки вероятности ошибки первого рода порядка $-\alpha \approx 0{,}016$.

При экспоненциальной функции полезной составляющей нестационарного случайного процесса (таб. 4.1) разница в оценках при рассмотренных законах плотности распределения вероятности наблюдается, когда аномальные значения расположены в начале выборки процесса, причем эти оценки являются минимальными относительно других, полученных при других расположениях аномальных значений.

Например, для гауссовского закона плотности распределения вероятности аддитивной шумовой составляющей, при расположении аномальных значений в начале выборки, оценки вероятности ошибки первого рода – $\alpha \approx 0{,}007$, а при других расположениях $-\alpha = 0{,}013$. Для равномерного закона плотности распределения вероятности аддитивной шумовой составляющей η_k, при расположении аномальных значений в начале выборки, оценки вероятности ошибки первого рода порядка $-\alpha \approx 0{,}004$, а при других расположениях аномальных значений– $\alpha = 0{,}009$.

Таблица 4.2

Значение оценок вероятности правильного обнаружения

Расположение аномальных значений	Гармоническая функция			Экспоненциальная функция			Составная функция		
	Гауссовский закон	Равнмерный закон	Рэлеевский закон	Гауссовский закон	Равнмерный закон	Рэлеевский закон	Гауссовский закон	Равнмерный закон	Рэлеевский закон
В начале выборки	0,908	0,898	0,879	0,888	0,889	0,881	0,925	0,937	0,899
В середине выборки	0,868	0,877	0,844	0,952	0,976	0,945	0,955	0,968	0,942
В конце выборки	0,830	0,825	0,804	0,963	0,976	0,955	0,917	0,930	0,894
Равномерно расположены по всей выборке	0,865	0,892	0,847	0,948	0,965	0,934	0,929	0,933	0,902

Как показывает анализ результатов таб. 4.1, если полезная составляющая нестационарного случайного процесса является составной функцией, то оценки вероятности ошибки первого рода α практически не зависят от места расположения аномальных значений в исследуемом процессе. Например, если аддитивная шумовая составляющая η_k имеет гауссовский закон плотности распределения вероятности, оценки вероятности ошибки первого рода порядка $-\alpha = 0{,}009$; равномерный закон

плотности распределения вероятности $-\alpha = 0{,}003$; рэлеевский закон плотности распределения вероятности – $\alpha = 0{,}025$.

Анализируя, оценки вероятности правильного обнаружения β, которые представлены в таб. 4.2, видно что они практически не зависят от места расположения аномальных значений в исследуемом нестационарном случайном процессе. При гармонической функции полезной составляющей, если закон плотности распределения вероятности аддитивной шумовой составляющей гауссовский, то оценки вероятности правильного обнаружения – $\beta = 0{,}868$, если равномерный $-\beta = 0{,}873$ и если рэлеевский – $\beta = 0{,}844$.

Причем максимальные оценки вероятности правильного обнаружения β (таб. 4.2), получены в случае, когда аномальные значения расположены в начале выборки исследуемого процесса для всех рассмотренных законов плотности распределения вероятности аддитивной шумовой составляющей.

При экспоненциальной функции полезной составляющей (таб. 4.2), оценки вероятности правильного обнаружения $-\beta = 0{,}940$, если закон плотности распределения вероятности аддитивной шумовой составляющей гауссовский$-\beta = 0{,}951$, если равномерный и $\beta = 0{,}929$, если рэлеевский. Максимальные оценки вероятности правильного обнаружения β, в этом случае, получены при расположении аномальных значений в конце анализируемого процесса. Например, при гауссовском законе плотности распределения вероятности аддитивной шумовой составляющей, оценки вероятности правильного обнаружения $-\beta \approx 0{,}963$, при равномерном законе $-\beta \approx 0{,}976$ и при рэлеевском законе $-\beta \approx 0{,}955$ (таб. 4.2).

Если полезная составляющая нестационарного случайного процесса λ_k является составной функцией, то как показывает анализ таб. 4.2, оценки вероятности правильного обнаружения β, так же практически не зависят

от места расположения аномальных значений в выборке. Оценки вероятности правильного обнаружения β принимают следующие значения соответственно при гауссовском $-\beta = 0{,}932$, равномерном – $\beta = 0{,}942$, рэлеевском $-\beta = 0{,}909$ законах плотности распределения вероятности аддитивной шумовой составляющей η_k.

Таким образом, можно считать, что эффективность способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа практически не зависит от места расположения аномальных значений в выборке исследуемого процесса с аддитивной шумовой составляющей [129].

4.2 Исследование эффективности способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа нестационарного случайного процесса при мультипликативной шумовой составляющей

В работе рассматривается возможность применения предлагаемого способа обнаружения аномальных значения для случая, когда нестационарный случайный процесс Y_k является мультипликативным, т.е. представлен моделью (2.3). С этой целью, при использовании способа с дискретно-меняющимся уровнем анализа процесса, получены зависимости оценок вероятности ошибки первого рода α и вероятности правильного обнаружения β для нестационарных случайных процессов с мультипликативной шумовой составляющей μ_k при различных величинах аномальных значений Amp [118, 130].

В качестве модели полезной составляющей λ_k исходного процесса Y_k используются следующие нормированные функции: гармоническая, экспоненциальная и составная функции. В качестве моделей

мультипликативной шумовой составляющей μ_k - гауссовский, равномерный и рэлеевский законы плотности распределения вероятности.

Величина аномальных значений $Amp = 3\sigma_ш, 4\sigma_ш, 5\sigma_ш, 6\sigma_ш, 7\sigma_ш$ ($\sigma_ш = 0{,}1$- среднеквадратическое отклонение мультипликативной шумовой составляющей). Одиночные аномальные значения распределены равномерно по всей реализации и составляют 5% от всего объема N. Исследования проводились при следующих параметрах предлагаемого способа обнаружения аномальных значений: количество порогов $P = 25$ и $d = 8$.

На рисунках 4.7а и 4.7б представлены графики зависимости оценок вероятности ошибки первого рода α и оценок вероятности правильного обнаружения β для случая, когда функция полезной составляющей λ_k является гармонической. Графики на рисунке 4.7 соответствуют случаям, когда мультипликативная шумовая составляющая μ_k имеет следующие законы плотности распределения вероятности: график 1 – рэлеевский; график 2 – гауссовский; график 3 – равномерный.

Аналогично на рисунках 4.8а, 4.8б и 4.9а, 4.9б представлены зависимости оценок вероятности ошибки первого рода $\alpha = f(Amp)$ и вероятности правильного обнаружения $\beta = f(Amp)$ для экспоненциальной и составной функции полезной составляющей нестационарного случайного процесса Y_k. На рисунках 4.8 и 4.9 соответственно график 1 – рэлеевскому, график 2 – гауссовскому и график 3 – равномерному законам плотности распределения вероятности мультипликативной шумовой составляющей μ_k нестационарного случайного процесса.

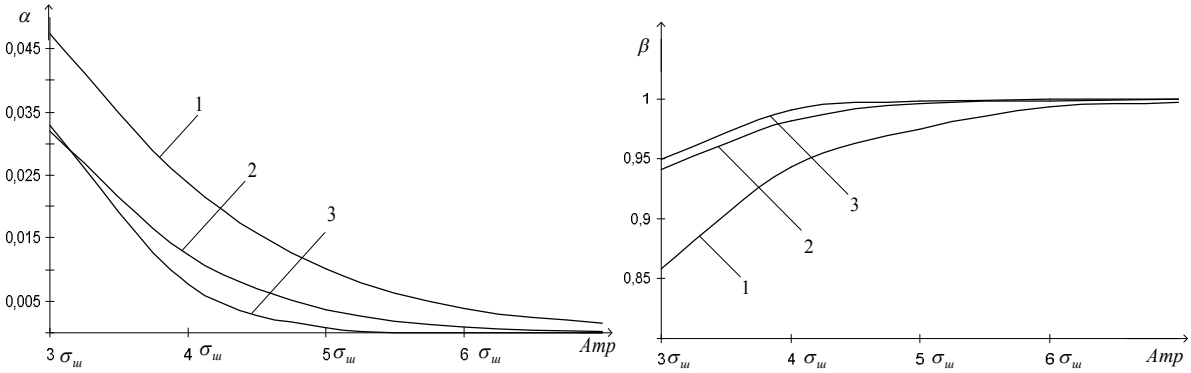
Анализ результатов представленных на рисунке 4.7 показывает, что при величине аномальных значений порядка $Amp = 3\sigma_ш$, для гармонической функции полезной составляющей λ_k, оценки вероятности

ошибки первого рода α (рис. 4.7а) при гауссовском и равномерном законах плотности распределения вероятности мультипликативной шумовой составляющей μ_k практически совпадают, в среднем $\alpha \approx 0{,}032$, а при рэлеевском $-\alpha \approx 0{,}047$. Оценки вероятности правильного обнаружения β ($Amp = 3\sigma_ш$, рис. 4.7б) при гауссовском и равномерном законах плотности распределения вероятности мультипликативной шумовой составляющей μ_k порядка $\beta \approx 0{,}945$, а при рэлеевском $\beta \approx 0{,}859$. С ростом величины аномальных значений ($Amp > 5\sigma_ш$, рис. 4.7б), разница между оценками вероятности ошибки первого рода α в зависимости от закона плотности распределения вероятности мультипликативной шумовой составляющей μ_k уменьшается и составляет $\alpha = 0{,}002$ (рис. 4.7а). Аналогично, для гауссовского, равномерного и рэлеевского законов плотности распределения вероятности мультипликативной шумовой составляющей μ_k, при величине аномальных значений $Amp > 5\sigma_ш$, оценки вероятности правильного обнаружения $\beta = 0{,}996$ (рис. 4.7б).

Для экспоненциальной функции полезной составляющей λ_k (рис. 4.8а), при величине аномальных значений порядка $Amp = 3\sigma_ш$, оценки вероятности ошибки первого рода для гауссовского и равномерного законов плотности распределения вероятности в среднем $\alpha = 0{,}007$, для рэлеевского $-\alpha \approx 0{,}015$. Оценки вероятности правильного обнаружения ($Amp = 3\sigma_ш$) для гауссовского и равномерного законов плотности распределения вероятности в среднем $\beta = 0{,}993$, для рэлеевского – $\beta \approx 0{,}959$ (рис. 4.8б).

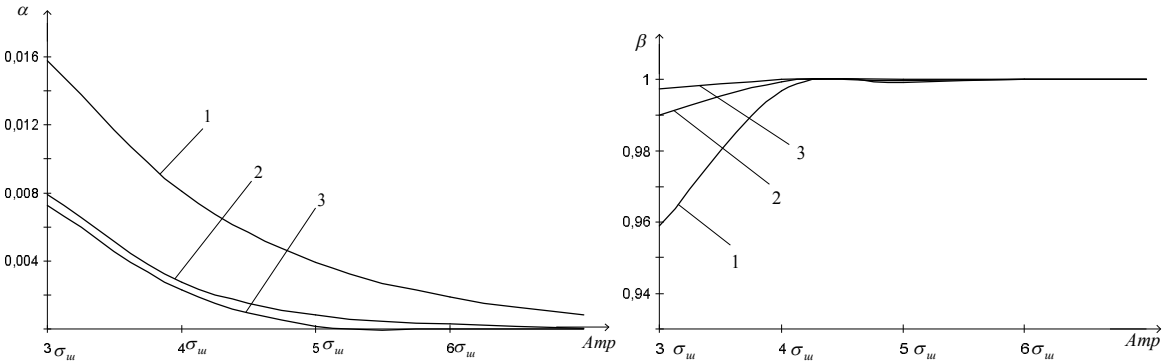
При величине аномальных значений порядка $Amp = 5\sigma_ш$, оценки вероятности ошибки первого рода α при гауссовском и равномерном законе плотности распределения вероятности мультипликативной

шумовой составляющей μ_k практически равны нулю (рис. 4.8а). Оценки вероятности правильного обнаружения уже при величине аномальных значений порядка $Amp = 4\sigma_ш$ составляют в среднем $\beta \approx 0{,}998$ при всех, рассмотренных законах плотности распределения вероятности мультипликативной шумовой составляющей входного процесса (рис. 4.8б).



а) б)

Рис. 4.7. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$ и вероятности правильного обнаружения $\beta = f(Amp)$ для гармонической функции полезной составляющей

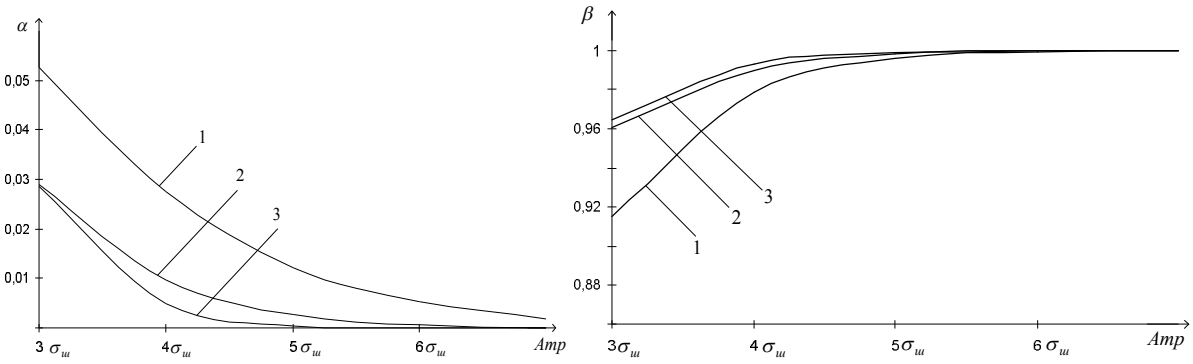


а) б)

Рис. 4.8. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$ и вероятности правильного обнаружения $\beta = f(Amp)$ для экспоненциальной функции полезной составляющей

Для случая, когда полезная составляющая λ_k представлена составной функцией (рис. 4.9), при том же значении величины аномальных значений, т.е. порядка $Amp = 3\sigma_ш$ оценки вероятности ошибки первого рода α и вероятности правильного обнаружения β составляют

соответственно $\alpha = 0{,}029$, $\beta = 0{,}963$ для гауссовского и равномерного законов плотности распределения вероятности мультипликативной шумовой составляющей μ_k. Если же шумовая составляющая μ_k имеет рэлеевский закон плотности распределения вероятности, то при той же величине аномальных значений $Amp = 3\sigma_ш$, оценки вероятности ошибки первого рода и вероятности правильного обнаружения: $\alpha \approx 0{,}053$ и $\beta \approx 0{,}915$ (рис. 4.9а и рис. 4.10б).



а) б)

Рис. 4.9. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$ и вероятности правильного обнаружения $\beta = f(Amp)$ для составной функции полезной составляющей

С ростом величины аномальных значений, как в случае, когда функция полезной составляющей λ_k нестационарного случайного процесса является гармонической и экспоненциальной, для составной функции оценки вероятности правильного обнаружения практически не зависят от закона плотности распределения вероятности мультипликативной шумовой составляющей μ_k.

Как показывает анализ зависимостей, представленных на рисунках 4.7, рис. 4.8 и рис. 4.9 оценки вероятности ошибки первого рода α и вероятности правильного обнаружения β практически совпадают, в случае, когда мультипликативная шумовая составляющая μ_k имеет гауссовский и равномерный законы плот плотности распределения

вероятности. При увеличении величины аномальных значений ($Amp > 5\sigma_ш$) эффективность предлагаемого способа обнаружения аномальных значений практически не зависит от закона плотности распределения вероятности мультипликативной шумовой составляющей.

Следует так же отметить, минимум оценок вероятности ошибки первого рода α и максимум вероятности правильного обнаружения β, для всех рассматриваемых функций полезной составляющей λ_k, получены для случая, когда мультипликативная шумовая составляющая μ_k имеет равномерный закон плотности распределения вероятности.

Для определения эффективности предложенного способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа нестационарных случайных процессов с мультипликативной шумовой составляющей μ_k рассмотрены зависимости оценок вероятности ошибки первого рода α и вероятности правильного обнаружения β относительно функции полезной составляющей λ_k нестационарного случайного процесса Y_k, которые представлены на рисунке 4.10 (график 1 – параболической, график 2 – гармонической, график 3 – составной, график 4 – сложной, график 5 – экспоненциальной функций).

Аналогично, на рисунках 4.11 и рис. 4.12, для исследуемых моделей функций полезной составляющей λ_k, представлены зависимости оценок вероятности ошибки первого рода α и вероятности правильного обнаружения β при равномерном и рэлеевском законах плотности распределения вероятности мультипликативной шумовой составляющей сигнала μ_k.

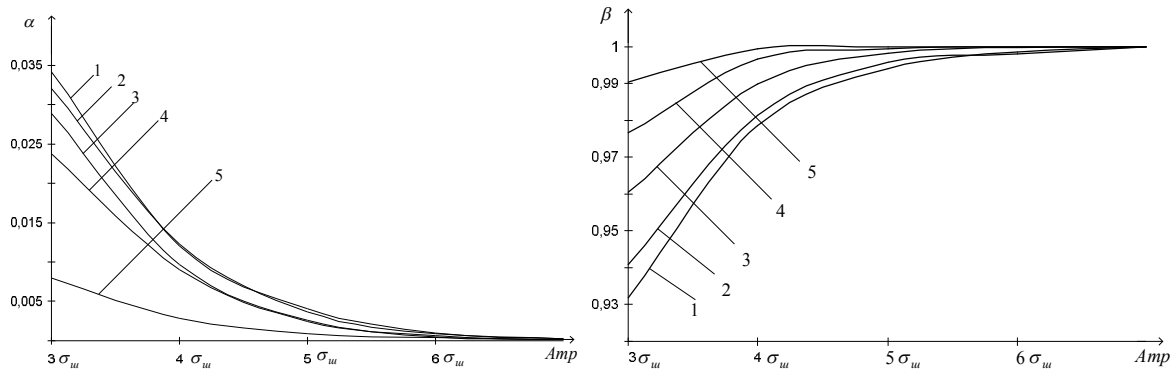
Из анализ полученных зависимостей, представленных на рисунках 4.10а (графики 1, 2, 3, 4), следует, что при гауссовском законе плотности распределения вероятности мультипликативной шумовой составляющей

сигнала μ_k оценки вероятности ошибки первого рода α, при величине аномальных значений $Amp = 3\sigma_ш$, для гармонической, параболической, составной и сложной функций, составляют примерно $\alpha = 0{,}030$. Причем, если полезная составляющая является экспоненциальной функцией, то в этом случае оценки вероятности ошибки первого рода минимальны $\alpha \approx 0{,}007$ (рис. 4.10а, график 5). Оценки вероятности правильного обнаружения при величине аномальных значений $Amp = 3\sigma_ш$ порядка $\beta = 0{,}990 - 0{,}932$, т.е. разница в оценках около 5% (рис. 4.10б).

При равномерном законе плотности распределения вероятности мультипликативной шумовой составляющей оценки вероятности ошибки первого рода α при величине аномальных значений $Amp = 3\sigma_ш$ в среднем, для гармонической, параболической, составной и сложной функций, составляют $\alpha = 0{,}031$ (рис. 4.11а, графики 1, 2, 3, 4). Если полезная составляющая λ_k является экспоненциальной функцией, то в этом случае оценки вероятности ошибки первого рода минимальны $\alpha \approx 0{,}007$ (рис. 4.11а, график 5). Оценки вероятности правильного обнаружения при величине аномальных значений $Amp = 3\sigma_ш$ – $\beta = 0{,}964 - 0{,}942$, т.е. разница в оценках около 2% (рис. 4.11б).

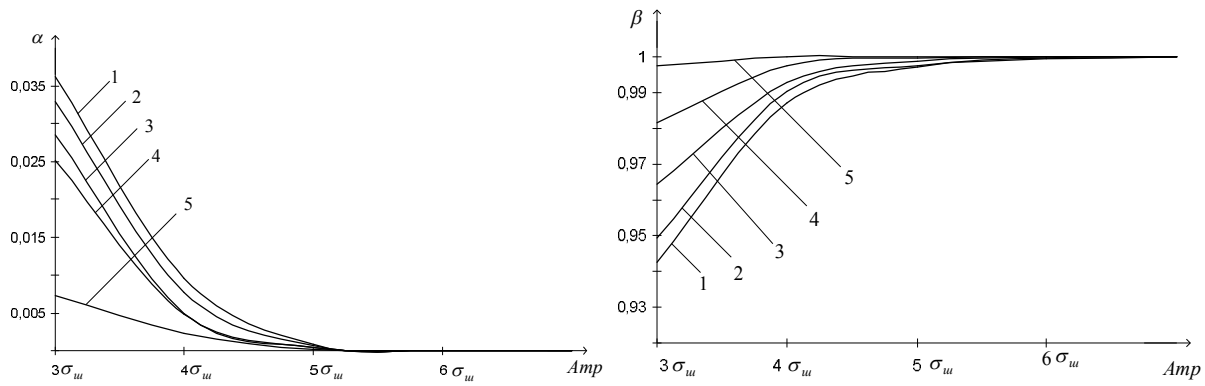
При рэлеевском законе плотности распределения вероятности мультипликативной шумовой составляющей μ_k, для гармонической, параболической, составной и сложной функций полезной составляющей λ_k, составляют ($Amp = 3\sigma_ш$) $\alpha = 0{,}049$ (рис. 4.12а). В том случае, когда функция полезной составляющей λ_k является экспоненциальной оценки вероятности ошибки первого рода минимальны $\alpha \approx 0{,}007$ (рис. 4.12а, график 5). При величине аномальных значений $Amp = 3\sigma_ш$ оценки вероятности правильного обнаружения порядка $\beta = 0{,}859 - 0{,}959$, т.е. разница в оценках около 10% (рис. 4.12,б).

Анализ зависимостей $\alpha = f(Amp)$ и $\beta = f(Amp)$ для различных законов плотности распределения вероятности мультипликативной шумовой составляющей μ_k, которые представлены на рис. 4.10а, 4.11а, 4.12а, показывает, что с увеличением величины аномальных значений, оценки вероятности ошибки первого рода α стремятся к своим минимальным значениям независимо от функции полезной составляющей λ_k.



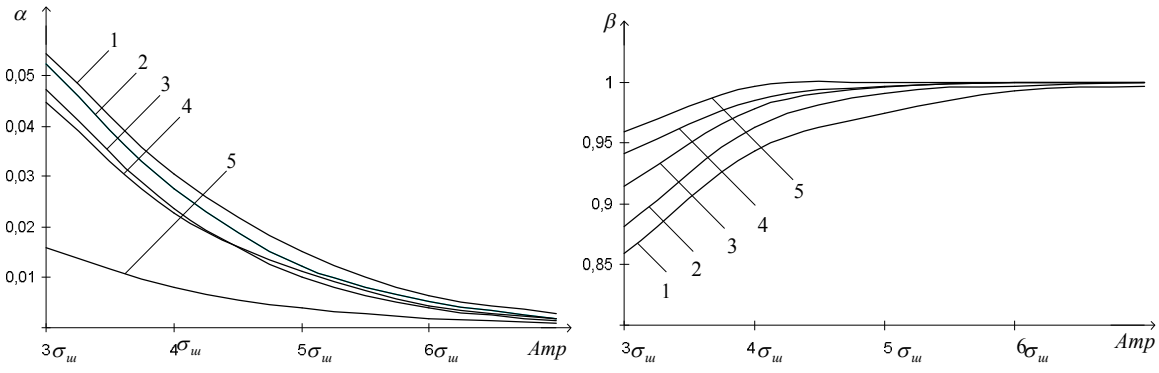
а) б)

Рис. 4.10. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$ и вероятности правильного обнаружения $\beta = f(Amp)$ при гауссовском законе плотности распределения вероятности мультипликативной шумовой составляющей



а) б)

Рис. 4.11. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$ и вероятности правильного обнаружения $\beta = f(Amp)$ при равномерном законе плотности распределения вероятности мультипликативной шумовой составляющей



а) б)

Рис. 4.12. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$ и вероятности правильного обнаружения $\beta = f(Amp)$ при рэлеевком законе плотности распределения вероятности мультипликативной шумовой составляющей

Например, с ростом величины аномальных значений ($Amp > 5\sigma_ш$) разность в оценках вероятности правильного обнаружения β при гауссовском законе плотности распределения вероятности мультипликативной шумовой составляющей μ_k менее 0,7% относительно функций полезной составляющей λ_k (рис. 4.10б).

При равномерном законе плотности распределения вероятности мультипликативной шумовой составляющей μ_k (рис. 4.11а) с увеличением величины аномальных значений оценки вероятности ошибки первого рода α для всех исследуемых функциях полезной составляющей практически равна нулю. Следовательно для нестационарных случайных процессов с мультипликативной шумовой составляющей, с увеличением величины аномальных значений, оценки вероятности ошибки первого рода α и вероятности правильного обнаружения β практически не зависят от функции полезной составляющей λ_k [130].

Для нестационарных случайных процессов с мультипликативной шумовой составляющей μ_k получены зависимости оценок вероятности ошибки первого рода α и вероятности правильного обнаружения β от места расположения аномальных значений в исследуемой процессе. При

тех же значениях параметра способа обнаружения аномальных значений, т.е. количестве уровней анализа $P = 25$ и $d = 8$.

В качестве моделей функции полезной составляющей λ_k используются нормированные функции: гармоническая, экспоненциальная и составная. Мультипликативная шумовая составляющая μ_k представлена гауссовским, равномерным и рэлеевский законами плотности распределения верятности. Величина аномальных значений $Amp = 5\sigma_ш$, где $\sigma_ш = 0{,}1$ – среднеквадратическое отклонение мультипликативной шумовой составляющей μ_k нестационарного случайного процесса.

Результаты исследований для оценок вероятности ошибки первого рода α представлены в таб. 4.3 и для оценок вероятности правильного обнаружения β в таб. 4.4.

Таблица 4.3

Значение оценок вероятности ошибки первого рода

Расположение аномальных значений	Гармоническая функция			Экспоненциальная функция			Составная функция		
	Гауссовский закон	Равномерный закон	Рэлеевский закон	Гауссовский закон	Равномерный закон	Рэлеевский закон	Гауссовский закон	Равномерный закон	Рэлеевский закон
В начале выборки	0,003	0	0,010	0	0	0,002	0,004	0,002	0,022
В середине выборки	0,005	0,001	0,014	0,003	0,001	0,008	0,003	0	0,010
В конце выборки	0,003	0	0,007	0,003	0,002	0,010	0,001	0	0,005
Равномерно расположены по всей выборке	0,004	0	0,009	0,001	0	0,004	0,002	0	0,010

Анализ результатов представленных в таб. 4.3 показал, что оценки вероятности ошибки первого рода α при гармонической функции полезной составляющей нестационарного случайного процесса при

гауссовом и равномерном законах плотности распределения вероятности мультипликативной шумовой составляющей μ_k составляют соответственно $\alpha = 0{,}004$ и $\alpha = 0{,}001$, т.е. в этих случаях практически не зависит от места расположения аномальных значений. При рэлеевском законе плотности распределения вероятности мультипликативной шумовой составляющей максимально возможные оценки вероятности ошибки первого рода, в этом случае, получена при расположении аномальных значений в середине выборки исследуемого процесса – $\alpha = 0{,}014$ (таб.4.3).

Если функция полезной составляющей нестационарного случайного процесса λ_k является экспоненциальной, наблюдается аналогичная закономерность в оценках вероятности ошибки первого рода α относительно расположения аномальных значений при гауссовском и равномерном законах плотности распределения вероятности мультипликативной шумовой составляющей μ_k (таб. 4.3).

Если закон плотности распределения вероятности мультипликативной шумовой составляющей μ_k гауссовский, тогда оценки ошибки первого рода порядка $\alpha = 0{,}002$, если равномерный – $\alpha = 0{,}001$, т.е. слабая зависимость от места расположения аномальных значений. В случае рэлеевского закона распределения, оценки вероятности ошибки первого рода порядка $\alpha = 0{,}006$ и максимальная оценки $\alpha \approx 0{,}010$ при расположении аномальных значений в конце реализации исследуемого процесса (таб. 4.3).

Как показывает анализ результатов таб. 4.3, в случае когда полезная составляющая λ_k нестационарного случайного процесса является составной, то оценки вероятности ошибки первого рода α практически не зависят от места расположения аномальных значений в реализации исследуемого процесса при гауссовском и равномерном законах плотности

распределения вероятности мультипликативной шумовой составляющей процесса μ_k. Если мультипликативная составляющая μ_k имеет гауссовский закон плотности распределения вероятности, оценки вероятности ошибки первого рода порядка $\alpha = 0{,}002$; равномерный – $\alpha = 0{,}001$. При рэлеевском $-\alpha = 0{,}012$, максимальное значение $-\alpha = 0{,}022$ при расположении аномальных значений в начале реализации процесса (таб. 4.3).

Таблица 4.4

Значение оценок вероятности правильного обнаружения

Расположение аномальных значений	Гармоническая функция			Экспоненциальная функция			Составная функция		
	Гауссовский закон	Равномерный закон	Рэлеевский закон	Гауссовский закон	Равномерный закон	Рэлеевский закон	Гауссовский закон	Равномерный закон	Рэлеевский закон
В начале выборки	0,994	0,998	0,984	0,998	1	0,994	1	1	1
В середине выборки	0,998	0,999	0,993	1	1	1	1	1	1
В конце выборки	0,991	0,998	0,966	1	1	1	0,992	0,997	0,988
Равномерно расположены по всей выборке	0,96	0,998	0,973	1	1	0,999	0,999	1	0,998

Анализируя, оценки вероятности правильного обнаружения β, которые представлены в таб. 4.4, видно, что они практически не зависят от места расположения аномальных значений в исследуемом нестационарном случайном процессе. Например, при гармонической функции полезной составляющей оценки вероятности правильного обнаружения в среднем: если закон плотности распределения вероятности мультипликативной шумовой составляющей μ_k гауссовский $-\beta = 0{,}995$, если равномерный – $\beta = 0{,}998$, если рэлеевский $-\beta = 0{,}979$. Максимальные оценки вероятности

правильного обнаружения получены при равномерном законе плотности распределения вероятности мультипликативной шумовой составляющей μ_k.

При экспоненциальной функции полезной составляющей нестационарного случайного процесса (таб. 4.4), оценки вероятности правильного обнаружения: если закон плотности распределения вероятности мультипликативной шумовой составляющей μ_k гауссовский – $\beta \approx 1$, если равномерный – $\beta \approx 1$ и если рэлеевский – $\beta = 0{,}998$.

Если полезная составляющая нестационарного случайного процесса является составной функцией, то как показывает анализ таб. 4.4, оценки вероятности правильного обнаружения β, слабо зависят от места расположения аномальных значений в исследуемой реализации. Оценки вероятности правильного обнаружения β принимают следующие значения соответственно при гауссовском $-\beta = 0{,}998$, равномерном – $\beta = 0{,}999$, рэлеевском $-\beta = 0{,}997$ законах плотности распределения вероятности.

Из проведенного анализа видно, что оценки вероятности ошибки первого рода α и вероятности правильного обнаружения β практически не зависят от места расположения аномальных значений в выборке нестационарного случайного процесса с мультипликативной шумовой составляющей μ_k.

4.3 Сравнительный анализ эффективности способа с обнаружения аномальных значений для нестационарных случайных процессов с аддитивной и мультипликативной шумовой составляющей

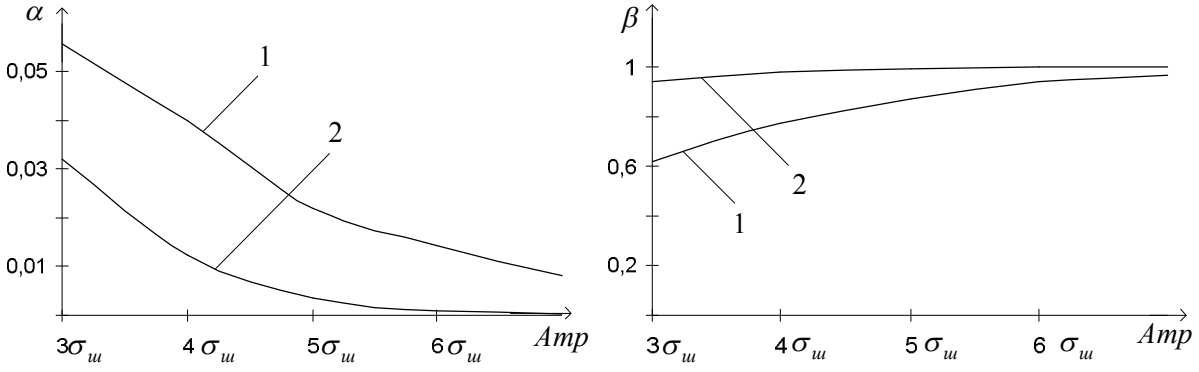
Для определения эффективности способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа исследуется влияние

модели процесса на оценки вероятности ошибки первого рода α и вероятности правильного обнаружения β. Рассматриваются нестационарные случайные процессы с аддитивной η_k и мультипликативной шумовой составляющей μ_k, анализируемые процессы представлены моделями вида (2.2) и (2.3). В качестве модели полезной составляющей исходного процесса Y_k используются гармоническая, экспоненциальная и составная функции [118, 120].

В качестве модели аддитивной шумовой составляющей η_k и мультипликативной шумовой составляющей μ_k рассматривается случай, когда закон плотности распределения вероятности шумовой составляющей является гауссовский центрированным случайным процессом.

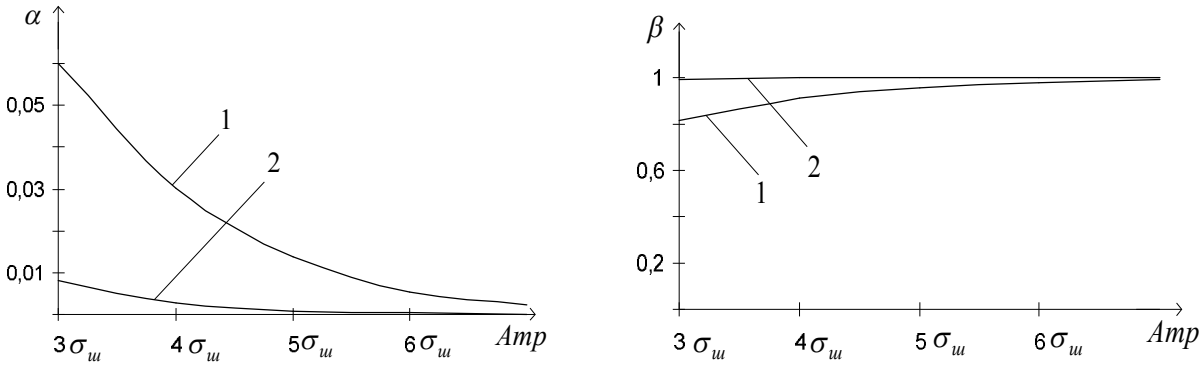
Аномальная составляющая ε_k представляет собой одиночные значения, равномерно распределенные по всей выборке исходного нестационарного случайного процесса Y_k где $k = \overline{1, N}$. Количество аномальных значений составляет 5% от объема выборки N. Величина аномальных измерений Amp, принимает значения $Amp = 3\sigma_ш, 4\sigma_ш, 5\sigma_ш, 6\sigma_ш, 7\sigma_ш$, где $\sigma_ш = 0{,}1$ – среднеквадратическое отклонение шумовой составляющей.

В результате проведенных исследований, получены зависимости оценок вероятности ошибки первого рода α и вероятности правильного обнаружения β от величины аномальных значений Amp с аддитивной η_k и мультипликативной шумовой составляющей μ_k, которые соответственно представлены на рисунках 4.13, 4.14 и 4.15. На рис. 4.13 представлены графики зависимости $\alpha = f(Amp)$ и $\beta = f(Amp)$ для гармонической функции полезной составляющей нестационарного случайного процесса, на рисунках 4.14 и 4.15 аналогичные зависимости для экспоненциальной и составной функции полезного λ_k.



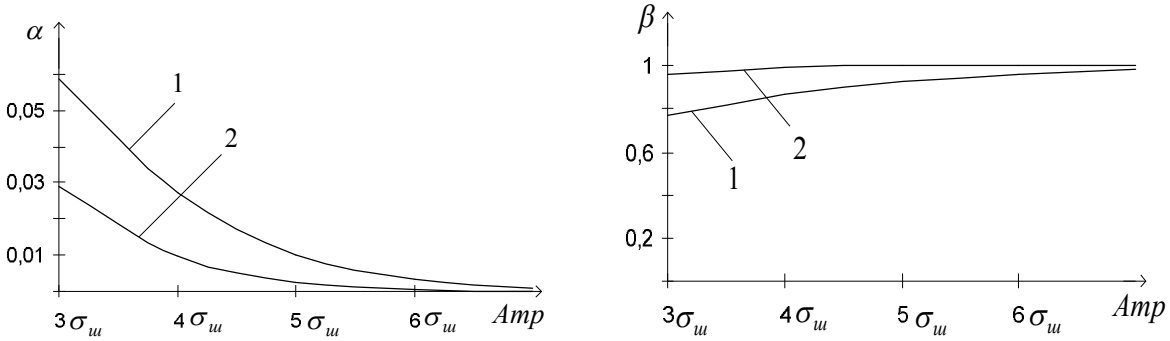
а) б)

Рис. 4.13. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$ и вероятности правильного обнаружения $\beta = f(Amp)$ при гармонической функции полезной составляющей



а) б)

Рис. 4.14. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$ и вероятности правильного обнаружения $\beta = f(Amp)$ при экспоненциальной функции полезной составляющей



а) б)

Рис. 4.15. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$ и вероятности правильного обнаружения $\beta = f(Amp)$ при составной функции полезной составляющей

График 1, на всех рассматриваемых зависимостях, соответствует случаю, когда модель нестационарного случайного процесса является аддитивной (2.2) и график 2, когда является мультипликативной (2.3).

Анализ зависимостей, представленных на рисунке 4.13а показывает, что при гармонической функции полезной составляющей, оценки вероятности ошибки первого рода α для процесса с мультипликативной шумовой составляющей μ_k (график 2) значительно меньше, чем для процесса с аддитивной шумовой составляющей η_k, примерно на 44% (график 1, при $Amp = 3\sigma_ш$). Однако с ростом величины аномальных значений, разница в оценках вероятности ошибки первого рода α уменьшается. Например, при величине аномальных значений порядка $Amp = 3\sigma_ш$ они отличаются на 44%, при величине аномальных значений порядка $Amp = 5\sigma_ш$ – 36%, а при величине аномальных значений порядка $Amp = 7\sigma_ш$ она уже составляет величину 23%. При гармонической функции полезной составляющей, оценки вероятности правильного обнаружения максимальны в случае, когда шумовая составляющая μ_k является мультипликативной (рис. 4.13б, график 2). Если величина аномальных значений $Amp > 5\sigma_ш$, то разница в оценках вероятности правильного обнаружения β отличается на 6% (рис. 4.13б, график 1 и график 2).

Из результата анализа графиков на рисунке 4.13а видно, что оценки вероятности ошибки первого рода α при экспоненциальной функции полезной составляющей и мультипликативной шумовой составляющей μ_k значительно меньше (рис. 4.14а, график 2), чем при аддитивной шумовой составляющей η_k (рис. 4.14а, график 1, при $Amp = 3\sigma_ш$), примерно на 84%. Оценки вероятности ошибки первого рода α при величине аномальных

значений $Amp = 3\sigma_ш$ отличаются на 84%, а при величине аномальных значений $Amp = 6\sigma_ш$ – 15% (рис. 4.14а).

Оценки вероятности правильного обнаружения β при экспоненциальной функции полезной составляющей нестационарного случайного процесса максимальны, когда шумовая составляющая μ_k является мультипликативной (рис. 4.14б, график 2). С ростом величины аномальных значений разница в оценках вероятности правильного обнаружения для исследуемых моделей нестационарного случайного процесса уменьшается (рис. 4.14б график 1, график 2). Например, если величина аномальных значений $Amp = 3\sigma_ш$, оценки вероятности правильного обнаружения β при мультипликативной и аддитивной шумовой составляющей отличаются на 26%, а при величине аномальных значений $Amp = 5\sigma_ш$ на 9% (рис. 4.14б).

Для составной функции полезной составляющей разница в оценках вероятности ошибки первого рода α при мультипликативной и аддитивной шумовой составляющей практически незначительна, так, например, при величине аномальных значений $Amp = 6\sigma_ш$ она составляет примерно 5% (рис. 4.15а график 1, график 2).

Оценки вероятности правильного обнаружения β при мультипликативной шумовой составляющей μ_k значительно отличаются от тех же оценок, полученных при аддитивной шумовой составляющей. Так, например, при величине аномальных значений $Amp = 3\sigma_ш$ отличаются примерно на 24%. С ростом величины аномальных значений эта разница значительно уменьшается и составляет, например, при $Amp = 5\sigma_ш$ – 8% (рис. 4.15б).

Таким образом, при величине аномальных значений $Amp = 3\sigma_ш - 5\sigma_ш$, эффективность их обнаружения выше, когда шумовая

составляющая является мультипликативной. С ростом величины аномальных значений эффективность предлагаемого способа обнаружения аномальных значений практически не зависит от формы взаимодействия полезной и шумовой составляющей нестационарного случайного процесса.

4.4 Исследование эффективности способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа для стационарных случайных процессов

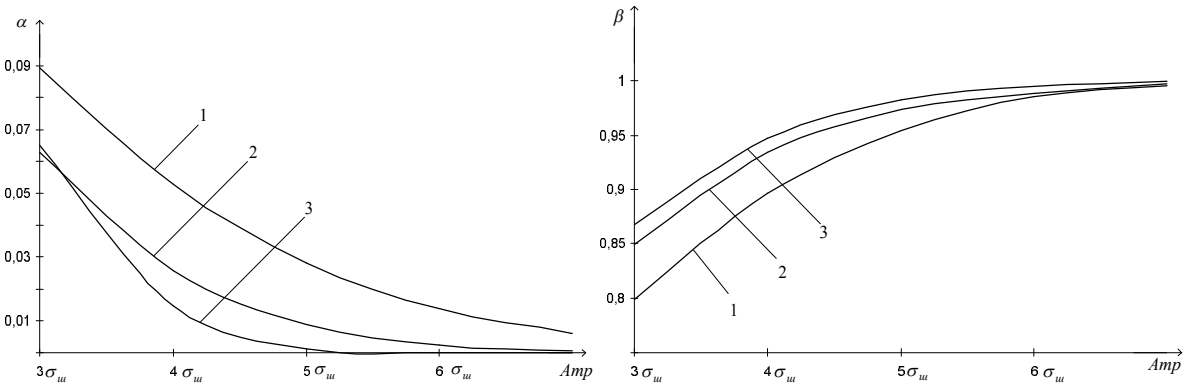
Проверка эффективности способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа также проводится для стационарного случайного процесса. Параметры способа обнаружения аномальных значений те же, что и для нестационарного случайного процесса, т.е. $P = 25$ и $d = 8$.

Рассматривается случай, когда шумовая составляющая стационарного процесса, представлена гауссовским (2.6), рэлеевским (2.8) и равномерным (2.7) законами плотности распределения вероятности . Аномальная составляющая ε_k представляет собой одиночные значения, равномерно распределенные по всей выборке исходного стационарного случайного процесса Y_k где $k = \overline{1, N}$. Количество составляет аномальных значений 5% от объема выборки N. Величина аномальных значений $Amp =$ $3\sigma_ш, 4\sigma_ш, 5\sigma_ш, 6\sigma_ш, 7\sigma_ш$, где $\sigma_ш = 0{,}1$ - среднеквадратическое отклонение шумовой составляющей.

В результате проведения имитационного моделирования получены зависимости оценок вероятности ошибки первого рода α и вероятность правильного обнаружения β от величины аномальных значений, которые представлены на рисунке 4.16 (график 1 – рэлеевский; график 2 –

гауссовский и график 3 – равномерный законы плотности распределения вероятности шумовой составляющей η_k).

Анализ результатов исследований, представленных на рис. 4.16а, показывает, что оценки вероятности ошибки первого рода α максимальны для случая, когда закон плотности распределения вероятности шумовой составляющей η_k является рэлеевским (график 1). Причем максимальность оценок для данного закона распределения относительно двух других (график 2 и график 3) сохраняется при всех величинах аномальных значений. Оценки значений вероятности ошибки первого рода α при гауссовском законе плотности распределения вероятности стационарного случайного процесса (рис. 4.16а, график 2) уменьшаются с увеличением величины аномальных значений и при $Amp \geq 6\sigma_ш$ близко к нулю.



а) б)

Рис. 4.16. Зависимость оценок вероятности ошибки первого рода $\alpha = f(Amp)$ и вероятности правильного обнаружения $\beta = f(Amp)$

Если закон плотности распределения вероятности шумовой составляющей η_k является равномерным, то, как видно из анализа рис. 4.16а (график 3), оценки вероятности ошибки первого рода α являются минимальными относительно других зависимостей (график 1 и график 2) и практически равны нулю при величине аномальных значений порядка $Amp = 5\sigma_ш$.

Необходимо так же отметить, что при величине аномальных значений $Amp = 3\sigma_ш$, оценки вероятности ошибки первого рода α при гауссовском и равномерном законах плотности распределения вероятности шумовой составляющей η_k практически совпадают, т.е. $\alpha = 0{,}064$ (рис. 4.16а график 2 и график 3).

С ростом величины аномальных значений разница в оценках вероятности ошибки первого рода α для равномерного и гауссовского законов плотности распределения вероятности шумовой составляющей η_k увеличиваются и составляет, например, при $Amp = 4\sigma_ш$ – $\alpha = 0{,}020$ (рис. 4.16а график 2 и график 3). При дальнейшем увеличении величины аномальных значений оценки вероятности ошибки первого рода α стремятся к нулю, как для гауссовского, так и для равномерного законов плотности распределения вероятности стационарного случайного процесса.

Оценки вероятности правильного обнаружения β (рис. 4.16б), максимальны, при всех рассмотренных величинах аномальных значений, для равномерного закона плотности распределения вероятности шумовой составляющей η_k – график 3. При $Amp = 3\sigma_ш$, оценки вероятности правильного обнаружения для гауссовского и равномерного законов плотности распределения вероятности шумовой составляющей η_k составляют в среднем $\beta = 0{,}858$, а для рэлеевского, при той же величине аномальных значений – $\beta = 0{,}799$ (рис. 6.16б, график 3).

С ростом величины аномальных значений, оценки вероятности правильного обнаружения стремятся к единице (график 1, график 2, график 3) и при $Amp \geq 5\sigma_ш$ практически совпадают для всех рассмотренных законов плотности распределения вероятности шумовой составляющей стационарного случайного процесса – $\beta = 0{,}970$ (рис. 4.16б).

Следовательно, при обнаружении аномальных в стационарном случайном процессе, оценки вероятности правильного обнаружения β практически не зависят от закона плотности распределения вероятности шумовой составляющей стационарного случайного процесса и с ростом величины аномальных значений стремятся к единице.

Так же для определения эффективности обнаружения аномальных значений, при использовании способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа, получены зависимости оценок вероятности ошибки первого рода α и вероятности правильного обнаружения β при различном расположении аномальных значений в выборке исследуемого стационарного случайного процесса.

Стационарный случайный процесс представлен равномерным, рэлеевским и гауссовскими законами плотности распределения вероятности ($\sigma_ш = 0{,}1$). Величина аномальных значений выбирается равной $Amp = 5\sigma_ш$, поскольку ранее показано, что при значении величины аномальных значений порядка $5\sigma_ш$, оценки вероятности правильного обнаружения β практически не зависят от законов плотности распределения вероятности шумовой составляющей. Аномальные значения составляют 5% от исходной выборки и могут располагаться равномерно, в начале, в середине или в конце исходной реализации.

В таб. 4.5 представлены оценки вероятности ошибки первого рода α для различных законов плотности распределения вероятности шумовой η_k, а в таб. 4.6 оценки вероятности правильного обнаружения β в зависимости от места расположения аномальных значений.

Анализ значений представленных в таб. 4.5 показывает, что оценки вероятности ошибки первого рода α практически не зависят от места расположения аномальных значений при всех рассмотренных законах

плотности распределения вероятности шумовой составляющей стационарного процесса.

Таблица 4.5

Значение оценок вероятности ошибки первого рода

Расположение аномальных значений	Гауссовский закон плотности распределения вероятности	Равномерный закон плотности распределения вероятности	Рэлеевский закон плотности распределения вероятности
В начале выборки	0,009	0,002	0,029
В середине выборки	0,009	0,002	0,030
В конце выборки	0,009	0,001	0,030
Равномерно расположены по всей выборке	0,009	0,001	0,028

Таблица 4.6

Значение оценок вероятности правильного обнаружения

Расположение аномальных результатов измерений	Гауссовский закон плотности распределения вероятности	Равномерный закон плотности распределения вероятности	Рэлеевский закон плотности распределения вероятности
В начале выборки	0,972	0,983	0,956
В середине выборки	0,970	0,985	0,957
В конце выборки	0,970	0,985	0,962
Равномерно расположены по всей выборке	0,972	0,983	0,954

Оценки вероятности правильного обнаружения β, как видно из таб. 4.6, при гауссовском и равномерном законах плотности распределения вероятности шумовой составляющей η_k практически не зависят от места расположения аномальных значений в исследуемой реализации процесса.

ГЛАВА 5 Исследование эффективности способов обнаружения аномальных значений при обработке реализаций, полученных в результате реальных экспериментов

При регистрации, обработке и обмене информации в современных радиотехнических системах потоки сигналов, кроме полезной составляющей и флуктуационных шумов, содержат, как правило, аномальные значения (импульсные помехи) различной интенсивности с регулярной или хаотичной структурой. Причины появления аномальных значений могут быть как внешние импульсные электромагнитные помехи, так и наводки, сбои и помехи в работе самих систем. Таким образом, шумовая составляющая нестационарного случайного процесса представляет собой совокупность статистически распределенного шума и квазидетерминированных аномальных значений, т.е. носит комбинированный характер. Поэтому, для повышения точности обработки сигналов в современных системах и комплексах необходимо учитывать именно комбинированный характер шумовой составляющей нестационарного случайного процесса [54, 118].

5.1 Сравнительный анализ эффективности адаптивного способа и способа с дискретно-меняющимся уровнем анализа процесса при обнаружении аномальных значений

Разработанные способы обнаружения аномальных значений обладают относительно высокой эффективностью при анализе нестационарных случайных процессов в условиях ограниченного объема априорной информации об исследуемом процессе. Для определения более конкретных рекомендаций при практическом использовании

разработанных способов необходимо провести сравнительный анализ их эффективности по обнаружению аномальных значений.

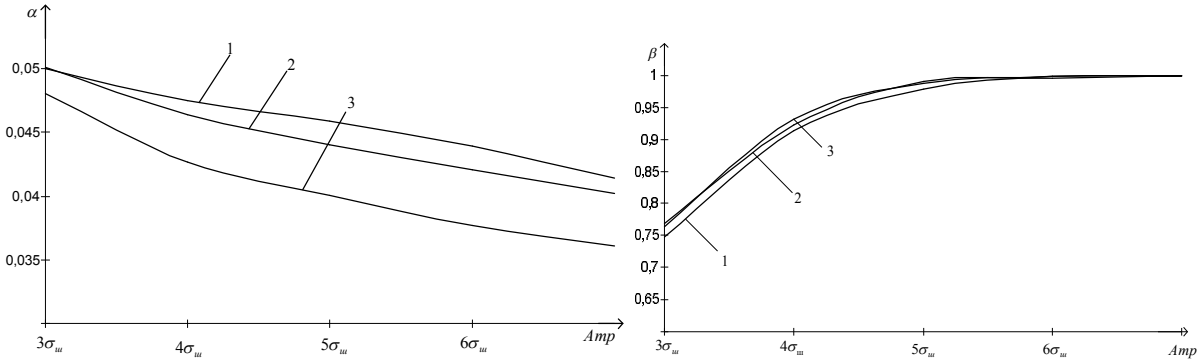
В частности, при использовании разработанных способов необходимо рассмотреть влияние величины среднеквадратического отклонения аддитивной шумовой составляющей $\sigma_ш$ на эффективность обнаружения аномальных значений. С этой целью получены оценки значения вероятности ошибки первого рода α и вероятности правильного обнаружения β для нестационарных случайных процессов при различных значениях среднеквадратического отклонения аддитивной шумовой составляющей $\sigma_ш$.

В качестве исследуемого процесса Y_k, где $k = \overline{1, N}$, рассмотрена аддитивная модель (2.2). Шумовая составляющая η_k имеет гауссовский закон плотности распределения вероятности, аномальная составляющая ε_k представляет собой одиночные аномальные значения, равномерно расположенные по всей исследуемой реализации. Количество аномальных значений фиксировано и составляет 5% от объёма выборки Y_k. Величина аномальных значений принимает целые, значения: $Amp = 3\sigma_ш, 4\sigma_ш, 5\sigma_ш, 6\sigma_ш, 7\sigma_ш$, где $\sigma_ш$ - среднеквадратическое отклонение аддитивной шумовой составляющей исследуемого процесса.

На рисунках 5.1 и 5.2 представлены зависимости оценок вероятности ошибки первого рода α и вероятности правильного обнаружения β, которые получены при использовании способа с адаптацией порогового значения. Результаты исследований, представленные на рисунке 5.1, соответствуют случаю, когда полезная составляющая процесса λ_k является гармонической функцией и на рис. 5.2 результаты, когда полезная составляющая процесса λ_k является параболической. Зависимости на рисунках 5.1 и 5.2 получены при

величине среднеквадратического отклонения аддитивной шумовой составляющей процесса: график 1 – $\sigma_ш = 0{,}1$, график 2 – $\sigma_ш = 0{,}2$ и график 3 – $\sigma_ш = 0{,}3$.

Анализ графиков, представленных на рисунке 5.1а показывает, что с ростом среднеквадратического отклонения аддитивной шумовой составляющей нестационарного случайного процесса апостериорная вероятность ошибки первого рода α уменьшается. Например, при величине аномальных значений порядка $Amp = 4\sigma_ш$, разница в оценках значений апостериорной вероятности ошибки первого рода α, полученной при $\sigma_ш = 0{,}1$ и при $\sigma_ш = 0{,}3$, составляет примерно 10%. А при величине аномальных значений порядка $Amp = 7\sigma_ш$, эта разница составляет уже примерно 14%. В тоже время оценки значения вероятности правильного обнаружения β (рис. 5.1б) практически не зависят от величины среднеквадратического отклонения аддитивной шумовой составляющей. Максимальная разница между оценками вероятности правильного обнаружения β получена при величине аномальных значений порядка $Amp = 3\sigma_ш$ и составляет менее 4%.

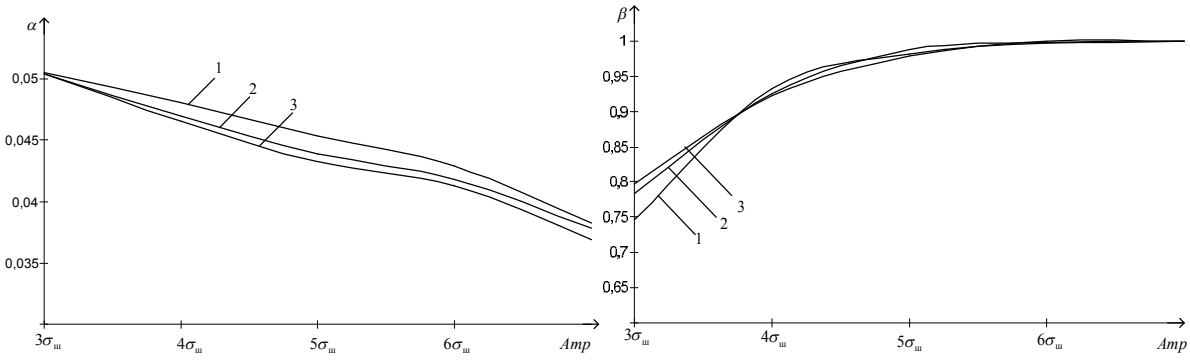


а) б)

Рис. 5.1. Оценки значения $\alpha = f(Amp)$ и $\beta = f(Amp)$ для гармонической функции от среднеквадратического отклонения аддитивной шумовой составляющей

Анализ результатов эффективности обнаружения аномальных значений при параболической функции полезной составляющей λ_k

нестационарного случайного процесса (рис. 5.2) показывает, что апостериорные оценки значения ошибки первого рода практически не зависят от величины среднеквадратического отклонения аддитивной шумовой составляющей $\sigma_ш$ и не превосходят априорно задаваемой величины $\alpha = 0{,}05$ (рис. 5.2а).



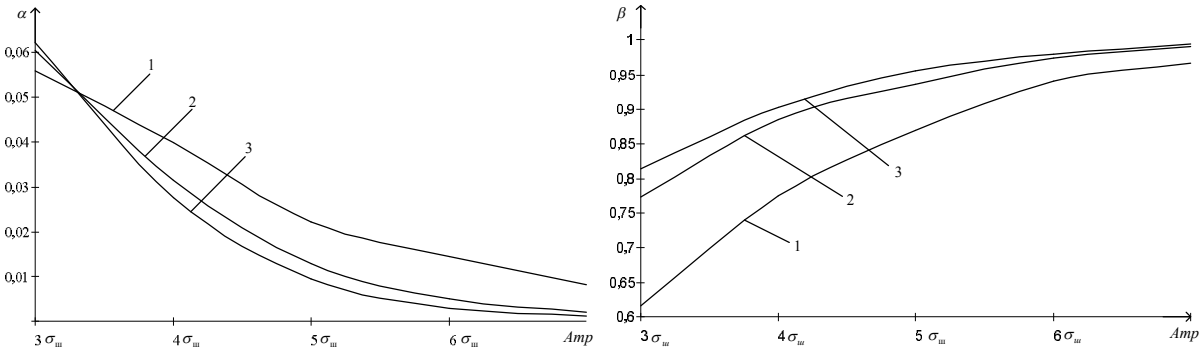
а) б)

Рис. 5.2 . Оценки значения $\alpha = f(Amp)$ и $\beta = f(Amp)$ для параболической функции от среднеквадратического отклонения аддитивной шумовой составляющей

Таким образом, для рассмотренного способа, с ростом величины среднеквадратического отклонения аддитивной шумовой составляющей исследуемого процесса, эффективность обнаружения аномальных значений практически не меняется. Это связано с тем, что в данном способе, среднеквадратическое отклонение аддитивной шумовой составляющей определяет пороговое значение (2.27): $\varepsilon_j^{(i)} = A \cdot \overline{\sigma}_j^{(i)}$, где $i = \overline{1, p}$ и $j = \overline{1, l}$, т.е. для каждого интервала разбиения, пороговое значение, помимо коэффициента A, зависит так же от значения среднеквадратического отклонения аддитивной шумовой составляющей.

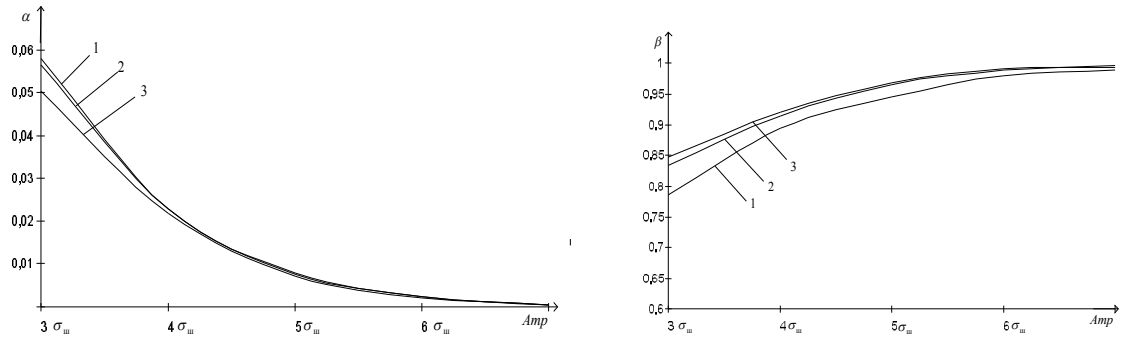
На рисунках 5.3 и 5.4 представлены зависимости вероятности ошибки первого рода α и вероятности правильного обнаружения β, полученные при применении способа обнаружения аномальных значений с дискретно-меняющимся уровнем анализа нестационарного случайного процесса.

На рисунке 5.3 представлены результаты при гармонической функции полезной составляющей процесса λ_k, на рисунке 5.4 – при параболической функции полезной составляющей процесса λ_k. Полученные зависимости представлены для следующих значений аддитивной шумовой составляющей: график 1 – $\sigma_ш = 0{,}3$, график 2 – $\sigma_ш = 0{,}2$ и график 3 – $\sigma_ш = 0{,}1$.



а) б)

Рис. 5.3. Оценки значения $\alpha = f(Amp)$ и $\beta = f(Amp)$ для гармонической функции от среднеквадратического отклонения аддитивной шумовой составляющей



а) б)

Рис. 5.4. Оценки значения $\alpha = f(Amp)$ и $\beta = f(Amp)$ для параболической функции от среднеквадратического отклонения аддитивной шумовой составляющей

Анализ результатов, представленных на рисунке 5.3 показывает, что с увеличением величины среднеквадратического отклонения аддитивной шумовой составляющей $\sigma_ш$ эффективность предлагаемого способа обнаружения аномальных значений возрастает. При $Amp = 4\sigma_ш$ и $\sigma_ш = 0{,}3$, $\sigma_ш = 0{,}1$ оценки значения вероятности ошибки первого рода α

уменьшаются на 30%. С ростом аномальных величины значений ($Amp \geq 6\sigma_ш$) относительная разница в оценках α уменьшается и составляет в среднем 12 % (рис. 5.3а, графики 1 и 3). Оценки значения вероятности правильного обнаружения β, с увеличением величины среднеквадратического отклонения аддитивной шумовой составляющей $\sigma_ш$, возрастают. Например, при $Amp = 3\sigma_ш$ и $\sigma_ш = 0{,}1$ вероятность правильного обнаружения аномальных значений $\beta = 0{,}614$, а при $\sigma_ш = 0{,}3$ величина $\beta = 0{,}818$, что составляет относительную разницу порядка 25%. При величине аномальных значений $Amp = 7\sigma_ш$ относительная разница в оценках вероятности правильного обнаружения уменьшается до 2% (рис. 5.3б, графики 1 и 3).

Для параболической модели функции полезной составляющей нестационарного случайного процесса наблюдается такая же закономерность в поведении оценок вероятности ошибки первого рода α и вероятности правильного обнаружения β, т.е. с увеличением величины среднеквадратического отклонения аддитивной шумовой составляющей $\sigma_ш$, эффективность предлагаемого способа возрастает. При $Amp = 3\sigma_ш$ и $\sigma_ш = 0{,}3$, $\sigma_ш = 0{,}1$ оценки вероятности ошибки первого рода α уменьшаются на 17%. С увеличением величины аномальных значений относительная разница значительно уменьшается, например, уже при $Amp = 5\sigma_ш$ она составляет 1,5% (рис. 5.4а, графики 1 и 3). Анализируя результаты, приведенные на рис. 5.4б, видно, при $Amp = 3\sigma_ш$ и $\sigma_ш = 0{,}1$ вероятность правильного обнаружения аномальных значений $\beta = 0{,}786$, а при $\sigma_ш = 0{,}3$ величина $\beta = 0{,}850$, т.е. относительная разница β порядка 8 %. С ростом величины аномальных значений эта разница уменьшается и уже при $Amp = 5\sigma_ш$ она составляет 2% (графики 1 и 3).

Аналогичные результаты получены при исследовании нестационарных случайных процессов с другими моделями функции полезной составляющей.

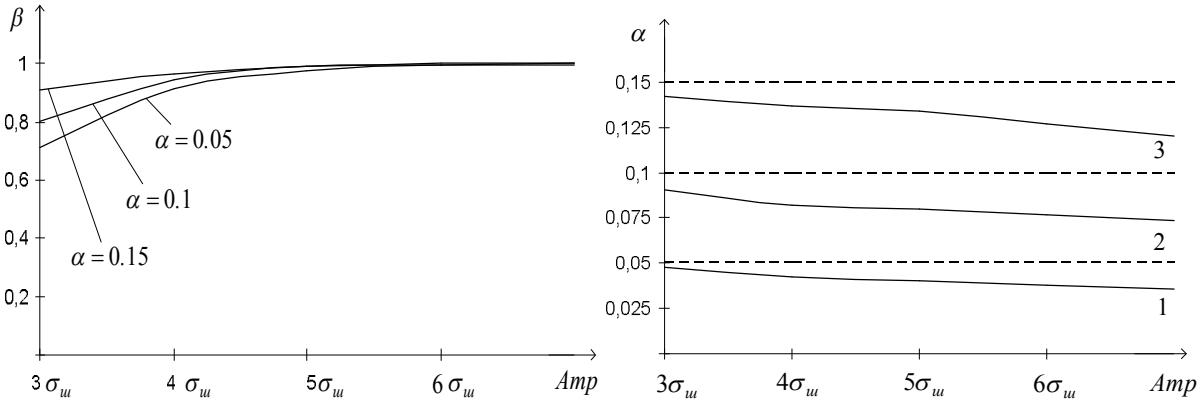
Таким образом, эффективность способа с дискретно-меняющимся уровнем анализа нестационарного случайного процесса возрастает с увеличении величины среднеквадратического отклонения $\sigma_ш$ аддитивной шумовой составляющей. Причем, при $Amp = 3\sigma_ш - 4\sigma_ш$ наблюдается более выраженная зависимость эффективности предлагаемого способа от величины $\sigma_ш$.

Для адаптивного способа обнаружения аномальных значений проводятся исследование зависимостей оценок вероятности правильного обнаружения β и ошибки первого рода α от априорно задаваемого α^* при анализе нестационарных случайных процессов с аддитивной шумовой составляющей (2.2).

Например, на рисунках 5.5 и 5.6, представлены зависимости оценок вероятности правильного обнаружения β и апостериорной вероятности ошибки первого рода α от величины аномальных значений Amp, для случая, когда полезная составляющая процесса является гармонической (рис. 5.5) и параболической (рис. 5.6).

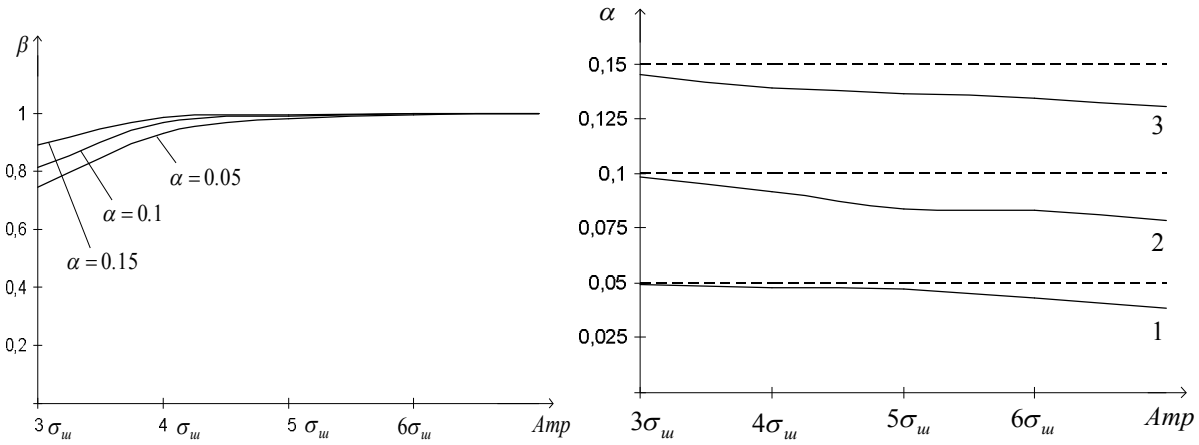
Оценки значения получены при следующих значениях априорно задаваемой ошибки первого рода $\alpha = 0{,}05$, $\alpha^* = 0{,}1$ и $\alpha^* = 0{,}15$. Аддитивная шумовая составляющая η_k имеет гауссовский закон плотности распределения вероятности (2.6). Аномальные значения ε_k составляют 5% от выборки N исходного нестационарного случайного процесса Y_k, $k = \overline{1, N}$ (2.2). Величина аномальных значений принимает дискретные значения $Amp = 3\sigma_ш, 4\sigma_ш, 5\sigma_ш, 6\sigma_ш, 7\sigma_ш$, где $\sigma_ш = 0{,}1$ - среднеквадратическое отклонение аддитивной шумовой составляющей

нестационарного случайного процесса. Зависимости апостериорной вероятности ошибки первого рода $\alpha = f(Amp)$, на рисунках 5.4б и 5.6б, соответственно получены: график 1 – при $\alpha^* = 0{,}05$; график 2 – при $\alpha^* = 0{,}1$; график 3 – при $\alpha^* = 0.15$.



а) б)

Рис. 5.5. Оценки значения $\beta = f(Amp)$ и $\alpha = f(Amp)$ для гармонической функции от априорно задаваемой вероятности ошибки первого рода



а) б)

Рис.5.6. Оценки значения $\beta = f(Amp)$ и $\alpha = f(Amp)$ для параболической функции от априорно задаваемой вероятности ошибки первого рода

Анализ полученных результатов показывает, что минимальные оценки значения вероятности правильного обнаружения β при $\alpha = 0{,}05$ – график 1 (рис. 5.5а и рис. 5.6а) соответствуют, случаю, когда величина аномальных значений $Amp = 3\sigma_ш$. С увеличением априорно фиксированного значения α^*, при величине аномальных значений

$Amp = 3\sigma_ш$ оценки значения вероятности правильного обнаружения β возрастают. При гармонической функции полезной составляющей нестационарного случайного процесса ($Amp = 3\sigma_ш$), возрастание значений оценок вероятности правильного обнаружения, с изменением фиксированного α^*, составляет примерно 11,5% (рис. 5.5а графики 1, 2, 3). Для параболической функции полезной составляющей нестационарного случайного процесса, при величине аномальных значений $Amp = 3\sigma_ш$, возрастание значений оценок вероятности правильного обнаружения составляет примерно $8{,}5\%$. С ростом величины аномальных значений разница в оценках вероятности правильного обнаружения β уменьшается, и при $Amp \geq 5\sigma_ш$ они практически совпадают и составляют, для рассмотренной модели нестационарного случайного процесса величину равную $\beta = 0{,}988$.

Оценки значения апостериорной вероятности ошибки первого рода α не превосходят своих априорных значений (рис. 5.5б и рис. 5.6б). Причем, с ростом величины аномальных значений разница между значениями оценки вероятности ошибки первого рода α и априорно задаваемыми значениями α увеличивается. Так, например (рис. 5.5б, график 1), при величине аномальных значений $Amp = 3\sigma_ш$ уменьшается на 4%, а при величине аномальных значений $Amp = 6\sigma_ш$ уменьшается уже на 15%. Стоит заметить, что оценки значения вероятности ошибки первого рода α уменьшаются относительно задаваемого значения с увеличением фиксированного α^*.

График 3 на рисунке 5.5,б соответствует оценкам вероятности ошибки первого рода α, полученным при $\alpha^* = 0{,}15$. В этом случае, если $Amp = 3\sigma_ш$ относительная разница в оценках вероятности ошибки первого

рода 5%, а при величине аномальных значений $Amp = 6\sigma_ш$ относительная разница 20%.

Аналогичные закономерности в поведении оценок вероятности правильного обнаружения β и вероятности ошибки первого рода α получены для других функций и законов полезной и аддитивной шумовой составляющих нестационарного случайного процесса.

Анализ полученных результатов показывает, что с ростом величины аномальных значений, $Amp \geq 5\sigma_ш$, оценки значения вероятности правильного обнаружения β слабо зависят от априорно заданного фиксированного значения вероятности ошибки первого α, а оценки значения апостериорной вероятности ошибки первого рода, с увеличением априорно задаваемого значения вероятности ошибки первого рода α^*, практически линейно уменьшаются.

Таким образом, оценки значения вероятности ошибки первого рода α^* не превосходят априорно задаваемых α, а с увеличением априорно задаваемых α значительно уменьшаются. При этом вероятность правильного обнаружения β стремиться к единице.

5.2 Анализ эффективности способов обнаружения аномальных значений при выделения полезной составляющей нестационарного случайного процесса

Проведенное исследование, предложенных в работе способов обнаружения аномальных значений показало их работоспособность как для случая стационарных, так и нестационарных случайных процессов. Рассмотрим одновременное обнаружение аномальных значений и выделение полезной составляющей нестационарного случайного процесса при аддитивной шумовой составляющей.

На практике часто возникают ситуации, когда функция полезной составляющей нестационарного случайного процесса априорно неизвестна. В этом случае, при решении задачи выделения полезной составляющей процесса применяют различные методы сглаживания. Задачей сглаживания принято считать задачу ослабления действия шумовой составляющей, т.е. уменьшение ее дисперсии. К наиболее используемым на практике методам уменьшения дисперсии шумовой составляющей можно отнести: метод скользящего среднего, метод взвешенного скользящего среднего, медианный метод, метод экспоненциального сглаживания и т.д.

При использовании метода скользящего среднего, для уменьшения дисперсии шумовой составляющей, достаточно одной реализации Y_k (2.2) исходного нестационарного случайного процесса, где $k = \overline{1, N}$. Определяется интервал сглаживания (ширина «скользящего окна») m, т.е. натуральное число $m < N$, для которого производится вычисление среднего арифметического, $\bar{\lambda}_i = \frac{1}{m}\sum_{i=1}^{m} Y_i$ значений Y_k, замену центрального из значений $Y_1, Y_2, ..., Y_m$ найденным средним $\bar{\lambda}_i$, сдвиг «скользящего окна» на одно значение вправо (т.е. выбор вместо отрезка $Y_k, Y_{k+1}, \ldots, Y_{k+m-1}$ другого отрезка $Y_{k+1}, Y_{k+2}, \ldots, Y_{k+m}$), вычисление среднего арифметического выбранных значений реализации, и так до тех пор, пока не будет достигнут правый конец исходной дискретной реализации. Ширину «скользящего окна» выбирают нечетной. Выражение для вычисления сглаженных значений исходной дискретной реализации результатов измерений записывается в виде: $\bar{\lambda}_j = \frac{1}{m}\sum_{k=j-p}^{j+p} Y_k$, где $j \geq p$ и $p = (m-1)/2$. В литературе рекомендации по выбору ширины «скользящего окна» m даются приблизительно. Некоторые авторы в своих работах [131-133]

указывают, что наиболее часто используется на практике значение ширины «скользящего окна», которая составляет 3 – 7 отсчетов.

Метод экспоненциального сглаживания [97, 132] заключается в том, что в процедуре нахождения оценки полезной составляющей нестационарного случайного процесса Y_k ($k = \overline{1, N}$), используются только предшествующие значения исследуемого процесса Y_k, взятые с определенным «весом», причем значение «весов» убывает к началу процесса. Рассматриваемый способ, предполагает выбор параметра сглаживания α, ($0{,}1 < \alpha < 1$), значения $\overline{\lambda}_0$, вычисление оценки полезной составляющей сигнала по рекуррентной формуле:

$$\overline{\lambda}_k = \alpha Y_k + (1-\alpha)\overline{\lambda}_{k-1} = \alpha \sum_{j=0}^{k-j} (1-\alpha)^j Y_{k-j} + (1-\alpha)^k \overline{\lambda}_0, \quad k = 1,2,..., N.$$

Для использования метода экспоненциального сглаживания задаются: начальным значением оценки полезной составляющей $\overline{\lambda}_0$ и параметр сглаживания α. Неверный выбор начальных условий может оказать существенное влияние на результат выделения полезной составляющей нестационарного случайного процесса. В практических рекомендациях по применению экспоненциального сглаживания [97] предлагается выбирать в качестве начального значения $\overline{\lambda}_0$ либо первое значение результатов измерений, либо среднее арифметическое нескольких первых членов результатов измерений, например, $\overline{\lambda}_0 = (Y_1 + Y_2 + Y_3)/3$. Влияние начальных условий является существенным при анализе сигналов малого объема.

Значения параметров метода скользящего среднего и метода экспоненциального сглаживания, как показывают результаты их практической реализации, зависят не только от вида функции полезного сигнала, но и от статистических характеристик шумовой составляющей.

Практические рекомендации для каждой модели процесса по выбору оптимальных параметров этих методов обоснованы в работе [78] .

В работах [54, 78, 134] предложен метод размножении оценок (РАЗОЦ), который обладает большой эффективностью при выделении полезного сигнала в условиях априорной неопределенности. Метод основан на получении множества оценок полезного сигнала при выделении его из реализации на фоне аддитивного шума. При этом объем исходной реализации ограничен, как и априорная информация о полезной и шумовой составляющих. Исходная реализация нестационарного случайного процесса Y_k ($k=\overline{1,N}$) (2.2) разбивается на подинтервалы случайными числами, имеющими равномерный закон распределения, на каждом подинтервале с помощью метода наименьших квадратов определяется значение оценок коэффициентов аппроксимирующего полинома $a+bt+ct^2$. Процедура разбиения полезного сигнала Y_k повторяется P- раз и для каждого разбиения, оценка полезного сигнала $\overline{\lambda}_k^{(j)}$, где $j=1,2,\ldots,l$, (l-количество подитервалов разбиения). В итоге получается кусочно-ломанная кривая, которая описывается на каждом подинтервале полиномом не выше второй степени. Результирующая оценка Y_k находится как среднее арифметическое функций $\overline{\lambda}_k^{(j)}$ по всей реализации N [54, 78], т.е. $\overline{\lambda_k}=\frac{1}{N}\sum_{j=1}^{N}\overline{\lambda}_k^{(j)}$. Автором работ [54, 78] экспериментально получены и математически доказаны, рекомендуемые значения параметров метода размножения оценок (минимальная длина выборки, количество разбиений на подинтервалы и число размножений оценок). Эти параметры практически инварианты в широком классе моделей шумовой составляющей и моделей полезной составляющей обрабатываемого процесса.

Метод медианного сглаживания предполагает для исходного дискретного процесса Y_k ($k=\overline{1,N}$) (2.2), определение длины m^* интервала ряда $Y_1, Y_2, ..., Y_N$ (или ширины «скользящего окна»), для которого будет производиться вычисление медианы, то есть ранжирование выбранного интервала входной реализации нестационарного случайного процесса, определение медианы $\underset{1\le j\le m^*}{med}\{Y_j\}$, где $j=\overline{1,m^*}$, замена центрального значения интервала $Y_1, Y_2, ..., Y_{m^*}$ медианой $\underset{1\le j\le m^*}{med}\{Y_j\}$, сдвиг «скользящего окна» на одно значение вправо (т.е. выбор вместо интервала $Y_k, Y_{k+1}, \ldots, Y_{k+m^*-1}$ реализации другого интервала $Y_{k+1}, Y_{k+2}, \ldots, Y_{k+m^*}$), вычисление медианы на новом интервале исследуемого процесса, и так до тех пор, пока не будет достигнуто правого конца исходной реализации. Медианная фильтрация, отличается явной избирательностью к элементам массива с немонотонной составляющей последовательности значений исходного процесса в пределах ширины «скользящего окна» m^* и хорошо подавляет импульсные помехи, а так же случайные шумовые выбросы [135 – 138]. В этой связи большое интерес представляет сравнение медианного метода фильтрации с предлагаемыми способами обнаружения аномальных значений, как предварительного этапа сглаживания исследуемого процесса.

Для количественной оценки качества выделения полезной составляющей нестационарного случайного процесса могут быть использованы различные критерии. В данной работе используется наиболее часто встречающийся критерий [54] в виде среднеквадратического отклонения полученных усреднением оценок значений процесса от исходной функции полезной составляющей:

$$\overline{\sigma}_0 = \sqrt{\frac{1}{N-1}\sum_{k=1}^{N}(\lambda_k - \overline{\lambda}_k)^2}\ , \qquad (5.1)$$

где λ_k – исходная модель функции полезной составляющей анализируемого процесса, $\overline{\lambda}_k$ – усредненные оценки значения анализируемого процесса.

Рассмотрим эффективность использования способа обнаружения аномальных значений с адаптацией порогового значения (С1) и способа с дискретно-меняющимся уровнем анализа нестационарного случайного процесса (С2), как предварительного этапа обработки процесса в целях анализа результата сглаживания сигнала λ_k (2.1). В качестве исследуемых функций полезной составляющей λ_k нестационарного случайного процесса исследуются: гармоническая (2.13), параболическая (2.16), экспоненциальная (2.14), составная (2.17) и сложная (2.18) функции. Шумовая составляющая представлена гауссовским (2.6), равномерным (2.7) и рэлеевским законами (2.8) плотности распределения вероятности, отсчеты шумовой составляющей – некоррелированные. Количество аномальных значений в исследуемой реализации 5% ($N = 100$). При проведении исследований, величина аномальных значений изменяется дискретно от $3\sigma_ш$ до $9\sigma_ш$ с шагом $2\sigma_ш$ ($\sigma_ш = 0{,}1$ - среднеквадратическое отклонение шумовой составляющей). Аномальные значения расположены равномерно вдоль исходного процесса. Усредненные оценки значения среднеквадратического отклонения результата сглаживания процесса, $\overline{\sigma}_0$ получены по 1000 реализациям. Результаты экспоненциального, медианного и скользящего среднего сглаживания получены с учетом априорной информации о модели полезной, законе распределения плотности вероятности шумовой составляющих нестационарного случайного процесса и без учета «концевых эффектов» [54, 78]. Устранение, обнаруженных аномальных значений, осуществляется за счет интерполяции по соседним отсчетам.

Результаты сравнительного анализа при использовании рассмотренных способов сглаживания исследуемого процесса, представлены в таб. 5.1 для гармонической, в таб. 5.2 для параболической и в таб. 5.3 для экспоненциальной функций полезной составляющей нестационарного случайного процесса с аддитивной шумовой составляющей.

Таблица 5.1

Оценка эффективности способов при гармонической модели полезной составляющей

Гауссовский закон плотности распределения вероятности аддитивной шумовой составляющей					
Способы сглаживания		Размножения оценок	Скользящее среднее	Медианный	Экспоненциаль-ный
Без аномальных значений		**0,032**	**0,036**	**0,046**	**0,071**
С аномальными значениями и величиной	$3\sigma_ш$	0,034	0,042	0,049	0,086
	$5\sigma_ш$	0,037	0,046	0,050	0,104
	$7\sigma_ш$	0,041	0,055	0,051	0,129
	$9\sigma_ш$	0,048	0,062	0,052	0,156
С устранением аномальных значений С 1		**0,035**	**0,040**	**0,049**	**0,071**
С устранением аномальных значений С 2		**0,034**	**0,041**	**0,050**	**0,074**
Равномерный закон плотности распределения вероятности аддитивной шумовой составляющей					
Способы сглаживания		Размножения оценок	Скользящее среднее	Медианный	Экспоненциаль-ный
Без аномальных значений		**0,028**	**0,034**	**0,043**	**0,060**
С аномальными значениями и величиной	$3\sigma_ш$	0,032	0,038	0,046	0,076
	$5\sigma_ш$	0,035	0,046	0,047	0,098
	$7\sigma_ш$	0,039	0,053	0,047	0,123
	$9\sigma_ш$	0,044	0,059	0,047	0,149
С устранением аномальных значений С 1		**0,029**	**0,034**	**0,044**	**0,062**
С устранением аномальных значений С 2		**0,028**	**0,034**	**0,044**	**0,061**

Таблица 5.2

Оценка эффективности способов при параболической модели полезной составляющей

Гауссовский закон плотности распределения вероятности аддитивной шумовой составляющей					
Способы сглаживания		Размножения оценок	Скользящее среднее	Медианный	Экспоненциальный
Без аномальных значений		**0,024**	**0,027**	**0,036**	**0,067**
С аномальными значениями и величиной	$3\sigma_ш$	0,027	0,032	0,041	0,081
	$5\sigma_ш$	0,031	0,037	0,045	0,0101
	$7\sigma_ш$	0,036	0,046	0,048	0,124
	$9\sigma_ш$	0,042	0,054	0,051	0,151
С устранением аномальных значений С 1		**0,026**	**0,030**	**0,038**	**0,066**
С устранением аномальных значений С 2		**0,025**	**0,029**	**0,037**	**0,067**
Равномерный закон плотности распределения вероятности аддитивной шумовой составляющей					
Способы сглаживания		Размножения оценок	Скользящее среднее	Медианный	Экспоненциальный
Без аномальных значений		**0,021**	**0,023**	**0,034**	**0,055**
С аномальными значениями и величиной	$3\sigma_ш$	0,023	0,028	0,035	0,071
	$5\sigma_ш$	0,028	0,036	0,037	0,093
	$7\sigma_ш$	0,034	0,044	0,039	0,119
	$9\sigma_ш$	0,040	0,053	0,043	0,146
С устранением аномальных значений С 1		**0,024**	**0,024**	**0,036**	**0,057**
С устранением аномальных значений С 2		**0,022**	**0,025**	**0,035**	**0,056**

Анализ полученных результатов показывает, что при всех рассмотренных функциях полезной составляющей нестационарного случайного процесса (таб. 5.1, таб. 5.2, таб. 5.3) усредненные оценки значения среднеквадратического отклонения $\overline{\sigma}_0$ (5.1) при наличии

аномальных значений превышают значения $\overline{\sigma}_0$, если аномальные значения отсутствуют в выборке. Причем такая зависимость наблюдается при всех рассмотренных способах сглаживания исследуемого процесса.

Следует отметить, что при величине аномальных значений порядка $Amp = 3\sigma_ш$ усредненные оценки значения среднеквадратического отклонения $\overline{\sigma}_0$ увеличиваются. Например, при использовании способа «размножения оценок», для гармонической функции полезной составляющей процесса и гауссовском законе плотности распределения вероятности аддитивной шумовой составляющей значения $\overline{\sigma}_0$ увеличивается примерно на 7% и на 12%, если закон распределения аддитивной шумовой составляющей равномерный (таб. 5.1). Аналогично, при той же величине аномальных значений, для параболической функции полезной составляющей процесса усредненные оценки значения среднеквадратического отклонения $\overline{\sigma}_0$ увеличивается соответственно на 11% при гауссовском и на 9% при равномерном законах плотности распределения вероятности аддитивной шумовой составляющей нестационарного случайного процесса (таб. 5.2).

Методы экспоненциального сглаживания и скользящего среднего еще более «чувствительны» к наличию аномальных значений. При наличии аномальных значений с величиной порядка $Amp = 3\sigma_ш$ их эффективность уменьшается в среднем для параболической функции полезной составляющей на 15-17% и для экспоненциальной на 10-12% (таб. 5.2 и таб. 5.3).

Как показывают проведенные исследования, при малых величинах аномальных значений порядка $Amp = 3\sigma_ш$, эффективность метода медианного сглаживания, относительно других рассмотренных методов, практически остается постоянной. Увеличение среднеквадратического

отклонения $\overline{\sigma}_0$ составляет в среднем не более 6% для всех рассматриваемых моделей нестационарного случайного процесса.

Таблица 5.3

Оценка эффективности способов при экспоненциальной модели полезной составляющей

Гауссовский закон плотности распределения вероятности аддитивной шумовой составляющей					
Способы сглаживания		Размножения оценок	Скользящее среднее	Медианный	Экспоненциаль-ный
Без аномальных значений		**0,025**	**0,027**	**0,034**	**0.067**
С аномальными значениями и величиной	$3\sigma_ш$	0,028	0,030	0,036	0,076
	$5\sigma_ш$	0,030	0,035	0,038	0,094
	$7\sigma_ш$	0,036	0,43	0,037	0,118
	$9\sigma_ш$	0,042	0,050	0,039	0,142
С устранением аномальных значений С 1		**0,026**	**0,030**	**0,035**	**0,070**
С устранением аномальных значений С 2		**0,025**	**0,029**	**0,034**	**0,069**
Равномерный закон плотности распределения вероятности аддитивной шумовой составляющей					
Способы сглаживания		Размножения оценок	Скользящее среднее	Медианный	Экспоненциаль-ный
Без аномальных значений		**0,021**	**0,023**	**0,036**	**0,056**
С аномальными значениями и величиной	$3\sigma_ш$	0,024	0,026	0,036	0,065
	$5\sigma_ш$	0,028	0,032	0,038	0,085
	$7\sigma_ш$	0,034	0,040	0,040	0,111
	$9\sigma_ш$	0,040	0,049	0,044	0,137
С устранением аномальных значений С 1		**0,023**	**0,028**	**0,037**	**0,061**
С устранением аномальных значений С 2		**0,022**	**0,025**	**0,038**	**0,059**

Однако, с ростом величины аномальных значений, при медианном сглаживании, усредненные оценки значения среднеквадратического

отклонения $\overline{\sigma}_0$ значительно возрастают. Например, при величине аномальных значений порядка $Amp = 9\sigma_ш$, увеличение $\overline{\sigma}_0$ составляет в среднем 12% (таб. 5.1, таб. 5.2 и таб. 5.3). При использовании других рассматриваемых методов сглаживания, их эффективность так же уменьшается (относительно усредненных оценок среднеквадратического отклонения) с ростом величины аномальных значений. Метод скользящего среднего при величине аномальных значений порядка $Amp = 9\sigma_ш$, для гармонической функции полезной составляющей сигнала, дает оценки значения среднеквадратического отклонения $\overline{\sigma}_0$ на 40-42%, в зависимости от закона плотности распределения вероятности шумовой составляющей, чем при сглаживании процесса без аномальных значений (таб. 5.1). Этот же способ, при тех же условиях, для параболической и экспоненциальной функции уменьшает свою эффективность в среднем на 55-57% соответственно (таб. 5.2, таб. 5.3).

Таким образом, наличие аномальных значений, даже при незначительном их количестве и относительно небольшой величине, сильно влияет на погрешность результатов сглаживания процессов.

Анализ усредненных оценок среднеквадратического отклонения $\overline{\sigma}_0$, которые получены после применения предлагаемых способов обнаружения аномальных значений и их дальнейшем устранении из выборки, представленных в таб. 5.1, таб. 5.2 и таб. 5.3, показывает, что эффективность способов сглаживания значительно возрастает в отличие от той ситуации, когда методы сглаживания применялись без учета наличия аномальных значений в исследуемых процессов.

Например, в результате обработки способом обнаружения аномальных значений с априорно фиксированном значением вероятности ошибки первого рода С1 и способом С2 усредненные оценки значения среднеквадратического отклонения $\overline{\sigma}_0$ практически совпадают, как при

применении метода «размножения оценок» так и при применении методов экспоненциального и медианного сглаживания. Максимальная разница $\overline{\sigma}_0$ после удаления аномальных значений наблюдается при скользящем среднем и составляет величину порядка 5%. Причем, такая закономерность для экспоненциальной функции полезной составляющей наблюдается независимо от закона плотности распределения вероятности аддитивной шумовой составляющей (таб. 5.3).

Для гармонической функции полезной составляющей разница в усредненных оценок значениях среднеквадратического отклонения $\overline{\sigma}_0$ после удаления аномальных значений порядка 3-6% в зависимости закона плотности распределения вероятности аддитивной шумовой составляющей. Максимальное отличие значений $\overline{\sigma}_0$, порядка 12%, в случае, когда для сглаживания процесса применяется метод скользящего среднего (таб. 5.1).

Анализ полученных усредненных значений среднеквадратического отклонения $\overline{\sigma}_0$ показывает, что они мало отличаются от значений $\overline{\sigma}_0$ полученных без наличия аномальных значений в выборке нестационарного случайного процесса и значительно отличаются от тех значений среднеквадратического отклонения, которые были получены при наличии аномальных значений. Обнаружение аномальных значений может повысить эффективность сглаживания в среднем на 40-60% в зависимости от функции полезной составляющей нестационарного случайного процесса.

Аналогичные исследования проведены и для случаев сложной и составной моделей полезной составляющей исследуемого процесса, а так же для случая, когда закон плотности распределения вероятности аддитивной шумовой составляющей является рэлеевским случайным процессом.

Использование способов обнаружения аномальных значений при анализе нестационарных случайных процессов позволяет повысить точность выделения полезной составляющей сигнала нестационарного случайного процесса в условиях комбинированной шумовой составляющей, т.е. при учете наличия аномальных значений (импульсной помехи).

5.3 Обработка экспериментальных реализаций

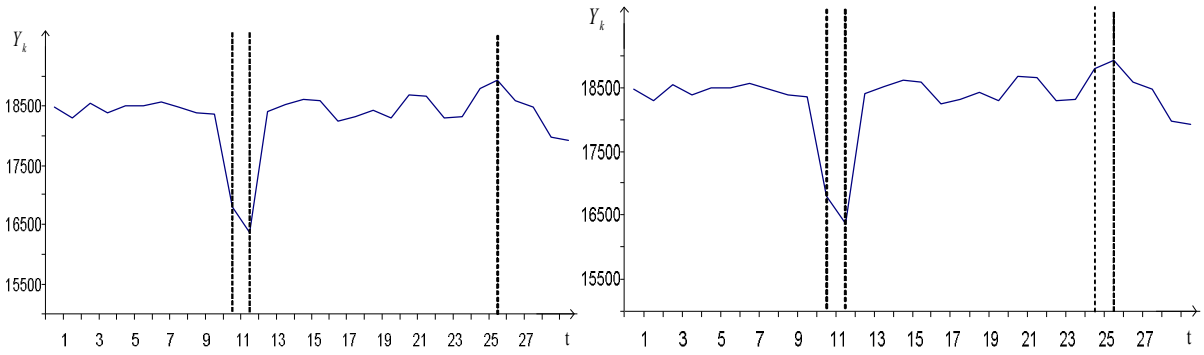
При решении многих практических задач, актуальной является задача обнаружения аномальных (резко выделяющихся) значений, их выделение или устранение. Это необходимо для улучшения качества последующей обработки процесса, либо для анализа работы самой системы, поскольку наличие резко выделяющихся значений может характеризовать изменения, происходящие в самой системе. Такая задача встречается, например, при анализе временных рядов в различных областях науки (экономике, астрофизике, электрохимии и др.) и им посвящено большое количество работ [139 – 145]. В связи с этим, рассмотрим возможность использования предлагаемых способов обнаружения аномальных значений при решении некоторых практических задач.

К настоящему времени в астрофизике существуют две основные гипотезы-сценария, объясняющие происхождение "убегающих" звезд [143]. В первом сценарии происходит взрыв сверхновой в тесной двойной системе. При этом система не распадается, а лишь приобретает высокую пространственную пекулярную скорость. Согласно второму сценарию, высокие скорости приобретаются благодаря динамическому взаимодействию двойных и кратных звезд, происходящему в центрах молодых звездных скоплений. Причем звезды могут вылетать из

родительского скопления как одиночные, так и двойные и даже кратные. Обработка, анализ и интерпретация фотометрических и спектральных наблюдений большого количества звезд указанного класса позволят уточнить теорию их происхождения и эволюционные процессы, протекающие в них.

В качестве данных для обработки использовались результаты фотометрических наблюдений «убегающей» звезды HD 218915 (спектральный класс O 9.5 Iab, видимая звездная величина V = $7^m.2$), выполненные на 60-см рефлекторе Крымской лаборатории ГАИШ МГУ с использованием электрофотометра «Одесса» на счете фотонов [143]. Обрабатываемый процесс представляет цикл наблюдений, охватывающий август 2002 года и содержит 30 индивидуальных измерения блеска в V-фильтре (видимый диапазон) в течение 5 ночей.

Результаты обработки, предлагаемыми способами, исследуемого стационарного случайного процесса, представлены на рисунке 5.7. Рисунок 5.7а – результат применения адаптивного способа обнаружения аномальных значений, априорная вероятность ошибки первого рода фиксирована и составляла $\alpha = 0{,}05$. Рисунок 5.7б – результат применения способа с дискретно-меняющимся уровнем анализа исследуемого процесса.



а) б)

Рис. 5.7. Результаты обнаружения аномальных значений при анализе фотометрических данных: а) адаптивный способ; б) способ с дискретно-меняющимся уровнем анализа

При использовании адаптивного способа обнаружения аномальных значений выявлено, что результаты измерений, соответствующие 10, 12 и 26 отсчетам является аномальными (рис. 5.7а). При использовании способа с дискретно-меняющимся уровнем анализа, к числу аномальных значений добавился еще и 24 отсчет (рис. 5.7б).

Для оценки значений рассматриваемого процесса на практике применяется визуальный контроль или так называемая «ручная» оценка, т.е. если текущее значение превышает предыдущее на 2%, то его считают аномальным. Значения визуально определенные как аномальные, соответствуют тем результатам, которые найдены предлагаемыми способами обнаружения аномальных значений.

Таким образом, используя представленные в работе способы обнаружения аномальных значений можно частично автоматизировать процесс обработки результатов некоторых фотометрических исследований применяемых в астрофизике.

Анализ переменности блеска «убегающей» HD 218915 по результатам фотометрических наблюдений и визуальной оценке результатов приведен в работе [142]. В результате обработки данных исследователи обнаружили переменность звезды, соответствующую периоду 9,1 суток, с уровнем значимости 93 %. Эффективность предлагаемых способов обнаружения аномальных значений исследовалась и при анализе нестационарных случайных сигналов, представленных единственной реализацией.

Многие современные промышленные объекты в качестве источника энергии используют кислотные аккумуляторные батареи. Актуальной задачей при этом является определение остаточной емкости батареи или времени, в течение которого объект может работать при данной степени заряженности батареи и при данных условиях разряда. Для этого необходимо определить момент времени, когда кривая спада напряжения

для отстающего аккумулятора достигнет заранее определенного критического значения [144, 145].

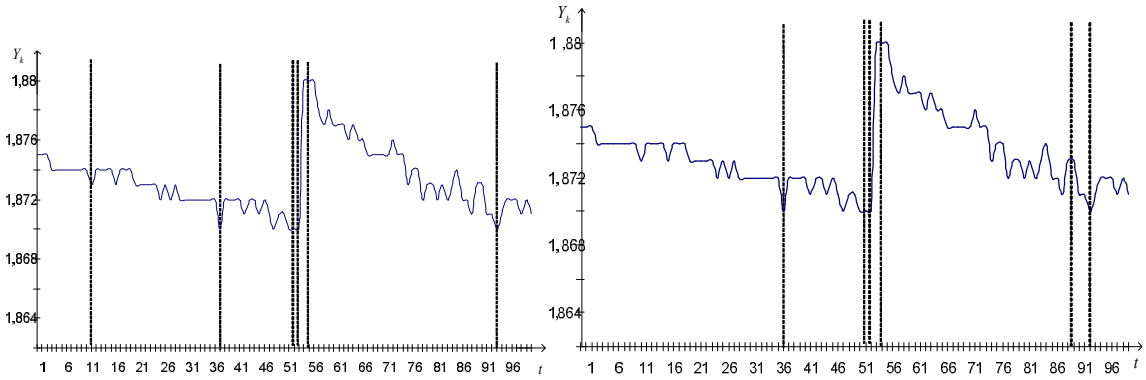
Разрядный ток, как правило, является нестационарным случайным процессом и содержит шумы, всплески и выбросы, что в свою очередь отражается и на поведении измеряемых характеристик батареи. Кроме этого отсутствует возможность повторного снятия той же характеристики, как, например, это делается во многих других методах. В данном случае исследователь, к сожалению, располагает лишь единственной реализацией измеряемого процесса.

Следовательно, необходимо использовать способ обнаружения экстремальных значений, который практически не зависит от априорной информации о поведении исследуемого объекта. Задачу обнаружения экстремальных значений необходимо проводить при наличии единственной кривой разряда, снятой с обрабатываемого аккумулятора в течение некоторого интервала времени.

На рисунке 5.8 представлена кривая зависимости, характеризующая работу аккумуляторной батарее при изменении внешней нагрузки. Результаты исследуемого процесса получены в равноотстоящие промежутки времени.

Результат обработки адаптивным способом представлен на рисунке 5.8а, априорная вероятность ошибки первого рода фиксирована и составляет $\alpha = 0{,}05$. На рис. 5.8б приведены результаты обработки способом с дискретно-меняющимся уровнем анализа.

Анализ рисунка 5.8 показывает, что при применении предлагаемых способов обнаружения аномальных значений в обоих случаях определён момент времени изменения кривой напряжения. Причем апостериорная вероятность ошибки первого рода не превышает 0,03% как для адаптивного способа, так и для способа, в котором применяется дискретно-меняющийся уровень анализа исследуемого процесса.



а) б)

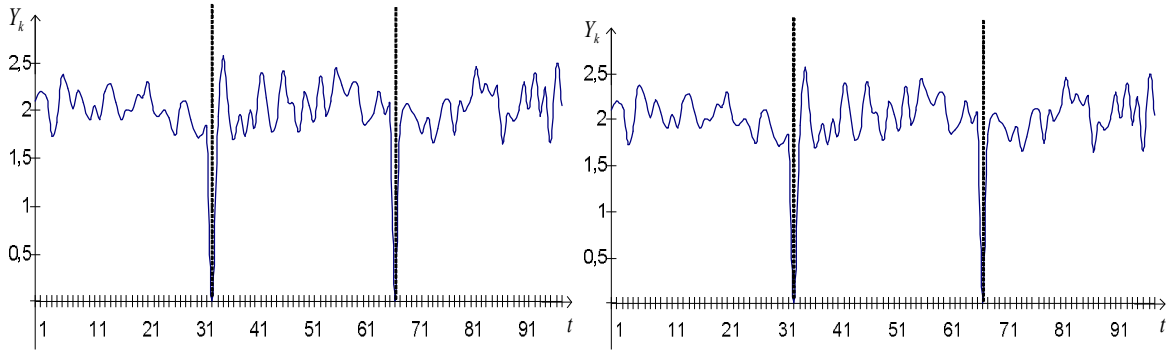
Рис. 5.8. Результаты обнаружения аномальных значений при анализе работы аккумуляторной батарее: а) адаптивный способ;
б) способ с дискретно-меняющимся уровнем анализа

На рисунке 5.9 представлена выборка, которая характеризует работу аккумуляторной батареи в режиме непланируемого короткого замыкания. Режим работы устройства при коротком замыкании характеризуется наличием резких выбросов значения сигнала. При автоматизированном режиме диагностики работы устройства, важность определения этих выбросов, позволит определить резкое изменение параметров либо самого устройства, либо параметров и режимов работы, подключенных к нему внешних систем.

Из анализа рисунка 5.9 видно, что при использовании как адаптивного способа обнаружения аномальных значений (рис. 5.9а), так и способа с дискретно-меняющимся уровнем анализа процесса (рис. 5.9б), вероятность обнаружения выбросов практически равна единице. Отсчеты 34 и 68 определены как аномальные, что видно при визуальном анализе кривых (рис. 5.9).

Таким образом, при работе аккумуляторной батареи применение предлагаемых способов обнаружения аномальных значений позволяет в автоматическом режиме обнаруживать резкие изменения в поведении параметров процесса. Это в перспективе, возможно, поможет точнее выбирать параметры самой аккумуляторной батареи или позволит

объективно оценить работу внешних, подключенных к ней устройств и систем.



а) б)

Рис. 5.9. Результаты обнаружения аномальных значений при анализе значений при анализе работы аккумуляторной батарее в режиме короткого замыкания: а) адаптивный способ; б) способ с дискретно-меняющимся уровнем анализа

В качестве натурной реализации для оценки эффективности предлагаемых способов рассмотрен сигнал, поступающий с металлодетектора. Это устройство позволяет обнаруживать металлические предметы, находящиеся в земле. Существует несколько видов металлодетекторов, способных не только обнаружить металлические предметы, но классифицировать их по размеру и химическому составу. Детекторы металлов могут быть оборудованы различными устройствами считывания информации: цифровой дисплей, индикация на стрелочном приборе, и другие устройства, помогающие идентифицировать объект.

Металлодетектор, в данном случае, имеет два инверсных канала. Входной процесс содержит в себе полезный сигнал и шумы, т.е. является нестационарным. Полезный сигнал имеет строго определённую форму во времени, но различную величину. Его наличие указывает на то, что сквозь датчик прошёл металл. Именно этот момент наличия полезного сигнала и необходимо выявить. Графический пример такого входного процесса с одного канала металлодетектора представлен на рисунке 5.10.

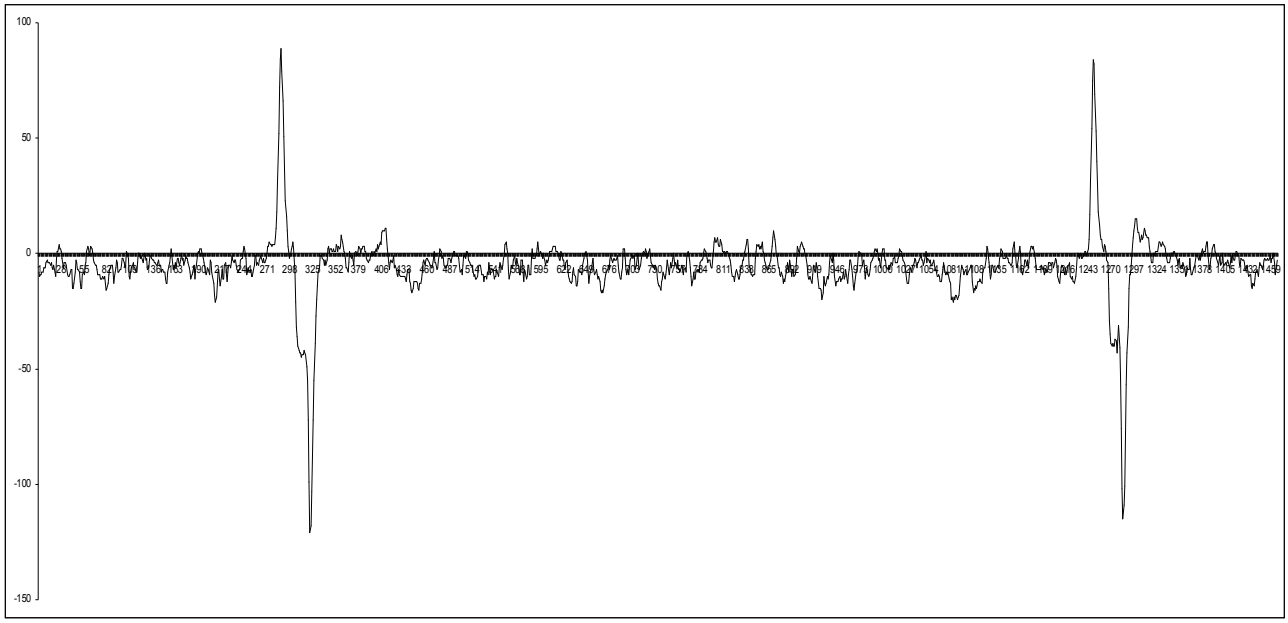


Рис. 5.10. Входной сигнала с металлодетектора

Кривая, на рисунке 5.10, отражает реальную форму выходного сигнала металлодетектора, в количестве 1460 отсчетов, от двух металлических частиц, проходящих сквозь датчик с интервалом около 10с. Вначале, при прохождении металла мимо первой катушки, вырабатывается максимальный положительный сигнал. Затем при прохождении металла мимо второй катушки вырабатывается максимальный отрицательный сигнал.

Шумы, как составляющая входного сигнала, представляют собой высокочастотные колебания. Форма этой составляющей процесса во времени отлична от процесса, образуемого обнаруживаемыми металлами. Причем, отсчеты шумовой составляющей входного процесса являются коррелированными. Для определения наличия корреляции между отсчетами шумовой составляющей построена корреляционная функция, график которой представлен на рисунке 5.11. Из анализа кривой, представленной, на рисунке 5.11 следует, что исходные отсчеты можно представить как гауссовский случайный процесс с экспоненциальной корреляционной функцией.

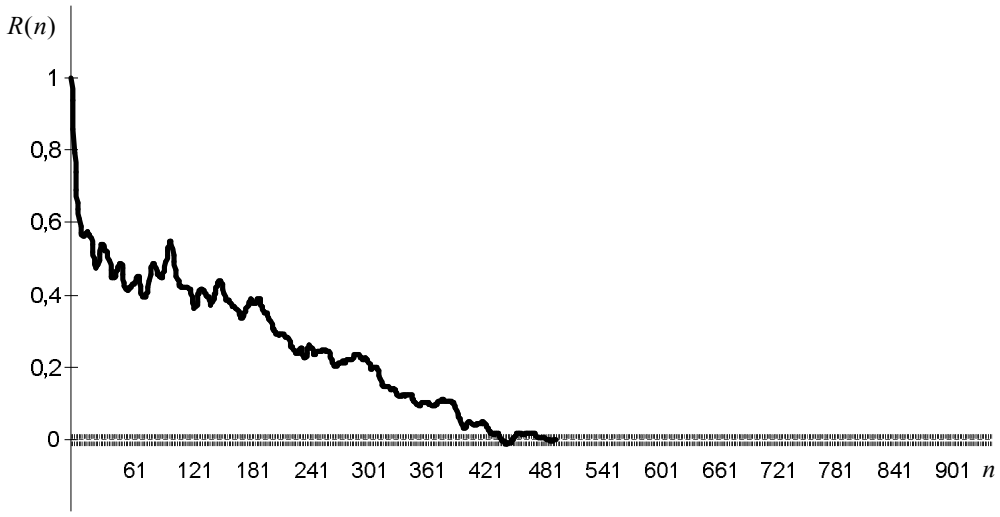
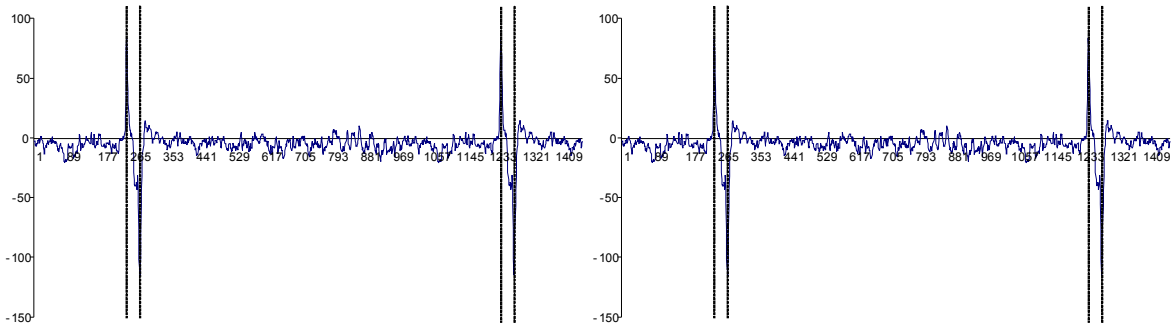


Рис. 5.11. График корреляционной функции

Для обнаружения наличия полезного сигнала в определенный момент времени, применяется адаптивный способ и способ с дискретно-меняющимся уровнем анализа случайного процесса. На рисунке 5.12а представлен результат применения адаптивного способа, а на рисунке 5.12б – результат применения способа с дискретно-меняющимся уровнем анализа.



а) б)

Рис. 5.12. Результаты обнаружения аномальных значений при анализе сигнала с металлодетектора: а) адаптивный способ; б) способ с дискретно-меняющимся уровнем анализа

В контексте решение данной задачи обнаружение положения аномальных значений трактуется в дальнейшем, как обнаружение полезной составляющая входного процесса.

Анализ результатов обнаружения аномальных значений, представленных на рисунках 5.12а и 5.12б, показывает, что оба способа

определяют положение экстремальных значений процесса с вероятность практически равной единице. Экстремальные значения соответствуют 240, 284, 1250 и 1284 номерам отсчетам. По положениям экстремальных значений исследуемого нестационарного процесса определяется и полезный сигнал. Таким образом, предлагаемые в работе способы обнаружения аномальных значений, позволяют определить экстремальные значения процесса на входе металлодетектора.

ЗАКЛЮЧЕНИЕ

Проведенный анализ существующих методов и алгоритмов обнаружения аномальных значений показал, что их эффективность зависит от объема априорной информации об исследуемом процессе. Большинство рассмотренных методов используют для случаев, когда процесс стационарный и имеет гауссовский закон плотности распределения вероятности. При анализе нестационарных случайных процессов, существующие методы и алгоритмы опираются на то, что шумовая составляющая является аддитивной и априорно известны характеристики аномальных значений (величина, частота появления, место расположения).

В монографии предложен новый адаптивный способ, позволяющий при анализе нестационарных случайных процессов аддитивно-мультипликативной природы обнаруживать аномальные значения в условиях ограниченного объема априорной информации. Применение предлагаемого адаптивного способа для анализа нестационарных случайных процессов, позволяет обнаруживать аномальные значения при априорно фиксированном значении вероятности ошибки первого рода. Предложен так же новый способ обнаружения аномальных значений с дискретно-меняющимся уровнем анализа процесса.

В работе экспериментально показано, что при наличии аномальных значений в исследуемом нестационарном случайном процессе, использование предлагаемых способов в задаче выделения полезной составляющей, позволяет повысить их эффективность в среднем на 40-60%.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ:

1. Tukey, J.W. The future of data analysis [Text] / J.W. Tukey. – Annals of mathematical statistics, 1962. – 1-67.p.
2. Микешина, Н.Г. Выявление и исключение аномальных значений [Текст] / Н.Г. Микешина// Заводская лаборатория. – 1966,- Т.38 № 3. – С.310-318.
3. Клигенэ, Н. Методы обнаружения моментов изменений свойств случайных процессов [Текст] / Н. Клигенэ, Л. Телькснис// Автоматика и телемеханика. - 1983. - №10. – С.25.
4. Переверткин, С.М. Бортовая телеметрическая аппаратура космических летательных аппаратов [Текст] / С.М. Переверткин [и др.]. – М.: Машиностроение, 1977. – 208 с.
5. Венецкий, И.Г. Теория вероятности и математическая статистика [Текст] / И.Г. Венецкий, Г.С. Кильдишев. - М.: Статистика, 1975. – 388 с.
6. Мирский, Г.Я. Аппаратное определение характеристик случайных процессов. Изд. 2-е переработ. и доп. [Текст] / Г.Я. Мирский. - М.: Энергия, 1972. – 456 с.
7. Мановцев, А.П. Введение в цифровую радиометрию [Text] / А.П. Мановцев.- М.: Энергия, 1973. – 317 с.
8. Мановцев, А.П. Основы теории радиотелеметрии [Text] / А.П. Мановцев.- М.: Энергия, 1967. – 343 с.
9. Киричук, В.С. Исключение недостоверных данных [Текст] / В.С. Киричук, Б.Н. Луценко // Автометрия. - 1970. - №6. – С.47-54.
10. Шор, Я.Б. Методы оценки анормальности результатов измерений [Текст] / Я.Б. Шор, А.М. Бендерский. – М: Сов. радио, 1953. – 60 с.

11. Bliss, C.J. Biometrika [Текст] / C.J. Bliss, W.G. Cochen, J.W. Tukey. - Boston, 1956. – 340 p.
12. Беляев, Ю.К. Вероятностные методы выборочного контроля [Text] / Ю.К.Беляев.- М.: Наука, 1975. – 408 с.
13. Thompson, W.R. The problem of negative estimates of variance components [Text] / W.R. Thompson. – Annals of mathematical statistics, 1962. – 214-265.p.
14. Pearson, E.S. Biometrika [Text] / E.S. Pearson, S.C. Chandra . – Annals of mathematical statistics, 1936. – 30-49 p.
15. Смирнов, Н.В. Доклады АН СССР [Текст] / Н.В. Смирнов. – ГОРОД: ИЗД-ВО, 1941. – 130 с.
16. Grubbs, F.G. Sample criteria for testing outlying observations [Text] / F.G. Grubbs. – Annals of mathematical statistics, 1950. –27-58 p.
17. Брофман, М.Я. Анализ критериев обнаружения экстремальных значений [Текст] / М.Я. Брофман, В.Х. Римен // Заводская лаборатория. – 1964. - Т.30, № 7. – С.861-865.
18. Bliss, C.J. Biometrika [Text] / C.J. Bliss, W.G. Cochen, J.W. Tukey. – Annals of mathematical statistics, 1956. – 418-503.p.
19. Бендерский, А.М. Доклады АН СССР, 85,1,5 (1952).
20. Ferguson, T.S. Proceeding of the 4 Berkeley symposium on mathematical statistics fnd probadiliti [Text] / T.S. Ferguson. – Annals of mathematical statistics, 1963. – 253-260p.
21. Дэйвид, Г. Порядковые статистики [Текст]: пер. с анг. под редакцией Петрова В.В. / Г. Дэйвид. – М.: Наука, 1979. – 336 с.
22. Корн, Г. Справочник по математике для научных работников и инженеров [Текст]: пер. с анг. / Г. Корн, Т. Корн. – М.: Наука, 1973. – 832 с.

23. Большев, Л.Н. Критерий Шовенэ в классической теории ошибок [Текст] / Л.Н. Большев, М. Убайдуллова // Теория вероятностей и ее применение. – 1974. Т.XIX, вып. 4. – С.714-723.
24. Пагурова, В.И. О критерии Шовенэ для обнаружения нескольких выбросов [Текст] / В.И. Пагурова // Теория вероятностей и ее применение. – 1985. Т.XXX, вып. 3. – С.558-561.
25. Большев, Л.Н. Обнаружение грубых ошибок в результате наблюдений [Текст] / Л.Н. Большев // Международная летняя школа по теории вероятности и математической статистике (Варна, 1974): Сб.науч.трудов. - София: ИЗД.БАН, 1974. – С.120.
26. Убайдуллова, М. Об отбраковке резко выделяющихся наблюдений [Текст] / М. Убайдуллова // Теория вероятностей и ее применение. – 1974. Т.XIX, вып.4. – С.864-868.
27. Пагурова, В.И. Об асимптотических свойствах критерия обнаружения выбросов [Текст] / В.И. Пагурова, К.Д. Родионов // Теория вероятностей и ее применение. – 1986. Т.XXX, вып. 3. – С.764-767.
28. Пагурова, В.И. Критерии обнаружения выбросов, использующие робастные оценки мешающих параметров [Текст] / В.И. Пагурова, И.А. Чижикова // Теория вероятностей и ее применение. – 1995. – Т. XX, вып. 2. – С.535-538.
29. Цветков, Э.И. Основы теории статистических измерений [Текст] / Э.И. Цветков. – Л.: Энергия. Ленингр. отд-ние, 1979. – 288 с.
30. Новицкий, П.В. Оценка погрешностей результатов измерений [Текст] / П.В. Новицкий, И.А. Зограф. - М.: Энергоатомиздат, 1985. – 248 с.
31. Новицкий, П.В., Динамика погрешностей средств измерений [Текст] / П.В. Новицкий, И.А. Зограф, В.С. Лабунец. - Л.: Энергоатомиздат, 1990. – 190 с.

32. Чернышев, С.Л. Исследование потоков событий наблюдаемых с ошибками [Текст] / С.Л. Чернышев // Радиотехника и электроника. - 1993. - Т.38, № 1. – С.310-314.
33. Савченко, В.В. Обнаружение и прогнозирование разладки случайного процесса на основе спектрального оценивания [Текст] / В.В. Савченко // Автометрия. – 1996. - №2. – С.11-16.
34. Ширяев, А.Н. Статистический последовательный анализ [Текст] / А.Н. Ширяев. – М.: Наука, 1976. – 232 с.
35. Фомин, А.Ф. Отбраковка аномальных результатов измерений [Текст] / А.Ф. Фомин, О.Н. Новоселов, А.В. Плющев. – М.: Энергоатомиздат, 1985. – 200 с.
36. Поляков, К.В. Повышение надежности для одной из процедур обнаружения [Текст] / К.В. Поляков, Р.Г. Толпарев // Радиотехника. - 1990. - №6. – С.57-59.
37. Егорова, Н.Ю. Решение задачи нелинейной фильтрации при наличии негауссовских ошибок измерений [Текст] / Н.Ю. Егорова, В.Е. Фабер // Радиотехника и электроника. - 1995. - Т.40, №4. - С.604-610.
38. Кузьмин, С.З. Основы теории цифровой обработки радиолокационной информации [Текст] / С.З. Кузьмин. – М.: Советское радио, 1974. – 432 с.
39. Кузьмин, С.З. Цифровая обработка радиолокационной информации [Текст] / С.З. Кузьмин. – М.: Советское радио, 1967. – 345 с.
40. Репин, В.Г. Статистический синтез при априорной неопределенности и адаптации информационных систем [Текст] / В.Г. Репин, Г.П. Тартаковский. - М.: Советское радио, 1977. – 432 с.
41. Катиков В.М. Мининимация аномальных ошибок в измерительных системах с устройством поиска и контроля [Текст] / В.М. Катиков // Радиотехника и электроника. - 1984. - Т.29, № 5. – С.920-927.

42. Демин, Н.С. Обнаружение аномальных помех в случае непрерывно-дискрентных каналов измерения [Текст] / Н.С. Демин, В.В. Михайлюк // Автометрия. - 1994. - №1. – С.21-26.
43. Кульбак, С. Теория информации и статистика [Текст] / С. Кульбак. – М.: Наука, 1967. – 408 с.
44. Фукунага, К. Введение в статистическую теорию распознания образов [Текст] / К. Фукунага. – М.: Наука, 1979. – 368 с.
45. Шакин, В.В. Вычислительная электрокардиография [Текст] / В.В. Шакин. – М.: Наука, 1981. – 68 с.
46. Чемоданова, О.А. Сравнение различных методов обнаружения аномальных погрешностей [Текст] / О.А. Чемоданова // Науч. труды / Моск. лесотехн. Ин-т. – М., 1981. - вып. 136. - С.168-170.
47. Голубков, А.С. Отбраковка недостоверных результатов телеизмерений [Текст] / А.С. Голубков, В.П. Коротаев // Метрология. - 1973. - №2. – С.10-18.
48. Зеленский, В.П. Применение методов теории статистических решений при исключении анормальных измерений [Текст] / В.П. Зеленский // Техническая кибернетика. - 1969. - №2. – С.139-142.
49. Миддлтон, Д. Введение в статистическую теорию связи [Текст] / Д. Миддлтон. – М.: Советское радио, 1961. -Т1. – 830 с.
50. Аминов, И.Н. Применение теории решения к задачам обнаружения сигналов и выделения сигналов из шумов [Текст] / И.Н. Аминов.- М.: Изд.ВВИА им. Жуковского, 1956. – 186 с.
51. Фенедбаум, А.А., Дудыкин А.В., Мановцев А.П., Миролюбов Н.Н. Теоретические основы связи и управлений [Текст] / А.А. Фенедбаум [и др.]. – М.: Физматгиз, 1963. – 280 с.
52. Тихонов, В.И. Выбросы случайных процессов [Текст] / В.И. Тихонов. - М.: Наука, 1970. – 392 с.

53. Новоселов, О.Н. Основы теории и расчета информационно измерительных систем [Текст] / О.Н. Новоселов, А.Ф. Фомин. – М.: Машиностроение, 1980. – 280 с.
54. Марчук, В.И. Первичная обработка результатов измерений при ограниченном объеме априорной информации [Текст]: Монография / Под ред. К.Е. Румянцева. – Таганрог: Изд-во ТРТУ, 2003. – 160с.
55. Тарасенко, Ф.П. Непараметрическая статистика [Текст] / Ф.П. Тарасенко. - Томск: Изд.Томского ТГУ, 1976. – 294 с.
56. Paulson, E. An optimum solution to the k-sample slippage problem for the normal distribution / E. Paulson. - Ann. Math. Statist, 1952. - 610-616.p.
57. Орлов, А.И. Неустойчивость параметрических методов отбраковки резко выделяющихся наблюдений [Текст] / А.И. Орлов // Заводская лаборатория. – 1992. - Т.58, №7. – С.40-42.
58. Орлов, А. И. Эконометрика [Текст] / А. И. Орлов. – М.: Экзамен, 2002. – 575 с.
59. Большев, Л.Н. Таблицы математической статистики [Текст] / Л.Н. Большев, Н.В. Смирнов. – М.: Наука, 1983. – 416 с.
60. Фомин, А.Ф. Методы и средства повышения достоверности измерений непрерывных процессов [Текст] / А.Ф. Фомин, О.Н. Новоселов, А.В. Плющев // Измерения, контроль, автоматизация. - 1981. - №4. – С.3-10.
61. Русинов, Л.А. Алгоритмическое обеспечение информационно-измерительных систем с аналитическими приборами [Текст] / Л.А. Русинов, Гуркевич А.Л. // Измерения, контроль, автоматизация. - 1982. - №3. – С.9-16.
62. Паламарчук, С.Н. Алгоритм оценивания параметров режима и анализа грубых ошибок на основе линеализованных выражений измеряемых величин [Текст] / С.Н. Паламарчук // Алгоритмы обработки данных в электроэнергетике / С.Н. Паламарчук. – Иркутск: СЭИ, 1982. - С.178-184.

63. Ефимов, Л.Н. Методы порядковых статистик и рангов в задачах обработки надлюдений [Текст] / Л.Н. Ефимов // Измерения, контроль, автоматизация. - 1981. - №5. – С.19-27.
64. Жадан, Л.И. К процедуре исключения аномальных измерений [Текст] / Л.И. Жадан // Автометрия. - 1985. - №2. – С.25-29.
65. Дуров, А.М. Многомерные статистические методы [Текст] / А.М. Дуров, В.С. Мхиторян, Л.И. Трошин. – М.: Финансы и статистика, 2000. – 352 с.
66. Химмельблау, Д. Анализ процессов статистическими методами [Текст] / Д. Химмельблау. - М.: Мир, 1973. – 957 с.
67. Лемешко, Б.Ю. Статистический анализ одномерных наблюдений случайных величин: Программная система [Текст] / Б.Ю. Лемешко. – Новосибирск: Новосиб.гос.техн.ун-т, 1995. – 125 с.
68. Куллдорф, Г. Введение в теорию оценивания по группированным и частично группированным выборкам [Текст] / Г. Куллдорф. – М.: Наука, 1966. – 176 с.
69. Денисов, В.И. Оптимальное группирование, оценка параметров и планирование регрессионных экспериментов: В 2ч. [Текст] / В.И Денисов, Б.Ю. Лемешко, Е.Б. Цой. – Новосибирск: Новосиб.гос.техн.ун-т, 1993. – 346 с.
70. Кендалл, М., Статистические выводы и связи [Текст] / М. Кендалл, А. Стьюарт. – М.: Наука, 1973. – 900 с.
71. Лемешко, Б.Ю. Робастные методы оценивания и отбраковки аномальных измерений [Текст] / Б.Ю. Лемешко // Заводская лаборатория. – 1997. - Т.63, №5. – С.43-49.
72. Rousseeuw, P. Unmasking multivariate outliers and leverage points [Текст]/ P. Rousseeuw, B. van Zomeren // J.Amer.Statist. Assoc., 1990. - v.85. - №411. - P.633-639.

73. Хьюбер, П. Робастность в статистике [Текст]: пер. с анг. / П. Хьюбер. – М.: Мир, 1984. - 304с.
74. Гуда, А.Н. Модели, методы и средства анализа данных в затрудненных условиях: автореф. дис….д-ра техн. наук: [Текст] /А.Н. Гуда; Таганрогский государственный радиотехнический университет. - Таганрог, 1997. - 38 с.
75. Мудров, В.И. Методы обработки измерений [Текст] / В.И. Мудров, В.Л. Кушко. – М.: Советское радио, 1976. – 192 с.
76. Мудров, В.И. Методы обработки измерений [Текст] / В.И. Мудров, В.Л. Кушко. – М.: Советское радио, 1976. – 192 с.
77. Егоров, В.А. Алгоритм удаления случайных сбоев в экспериментальном материале при автоматизированной регистрации данных [Текст] / В.А. Егоров, Ю.Г. Морозов, А.П. Чекмарева // Алгоритмы и математическое обеспечение для физических задач. - 1977. - №12. - С.21-35.
78. Марчук, В.И. Адаптивный алгоритм устранения аномальных значений случайного нестационарного процесса [Текст] / В.И. Марчук, А.М. Муратов // Элементы приёмно-усилительных устройств: Межвузовский тематический сборник. – Таганрог: ТРТИ, 1982. - Вып.1. - С. 87-90.
79. Марчук, В.И. Методы выделения полезной составляющей и обнаружение аномальных значений при анализе нестационарных случайных сигналов в условиях непараметрической априорной неопределенности и ограниченном объеме измерений [Текст] / Владимир Иванович Марчук. – Таганрог., 2006. - 372 с.
80. Фомичев, С.М., Абилов А.В. Обзор математических моделей каналов связи и их применение в телекоммуникационных системах [Текст] / С.М. Фомичев, А.В. Абилов.- Ижевск: Изд. Ижевского ГУ, 2001. С.60.
81. Миддлтон, Д. Введение в статистическую теорию связи [Текст]: пер. с англ./ Под ред. Б.Р. Левина / Д. Миддлтон. - М.: Советское радио, 1962. Т.2. - 260 с.

82. Финк, Л.К. Теория передачи дискретных сообщений [Текст] / Л.К. Финк. - М.: Советское радио, 1970. – 728 с.
83. Зюко, А.Г. Помехоустойчивость и эффективность систем связи [Текст] / А.Г. Зюко. – М.: Связьиздат, 1963. – 360 с.
84. Блох, Э.Л. Модели источника ошибок в каналах передачи цифровой информации [Текст] / Э.Л. Блох, О.В. Попов, В.Я. Турин. – М.: Связь, 1971. – 312 с.
85. Статистическая теория связи и ее практическое приложение [Текст] / Под ред. Левина Б.Р. – М.: Связь, 1979. – 287 с.
86. Прокис, Дж. Цифровая связь [Текст]: пер. с англ. / Под ред. Д.Д. Кловского / Дж. Прокис. - М.: Радио и связь, 2000. – 800 с.
87. Кловский, Д.Д. Передача дискретных сообщений по радиоканалам [Текст] / Д.Д. Кловский. - М.: Радио и связь, 1982. – 304 с.
88. Чесноков, М.Н. Оптимальный прием дискретных сообщений с переменными параметрами на фоне импульсных и флуктуационных помех [Текст] / М.Н. Чесноков // Известия вузов. Радиоэлектроника. - 1983. - №7. - С.3-11.
89. Кириллов, Н.Е. Помехоустойчивая передача сообщение по линейным каналам со случайно изменяющимися параметрами [Текст] / Н.Е. Кириллов. - М.: Связь, 1971. – 256 с.
90. Севостьянов, Б.А. Курс теории вероятности и математической статистики [Текст] / Б.А. Севостьянов. - М.: Мир, 1988. - 360 с.
91. Солонина, А.И. Основы цифровой обработки сигналов. Курс лекций [Текст] / А.И. Солонина [и др.]. 2-е изд., испр. и перераб. – СПб.: БХВ, 2005. – 768с.
92. Айвазян, С.А. Прикладная статистика. Основы эконометрики: Учебник для вузов: В 2 т. 2 -е изд., испр. Т.1: Теория вероятностей и прикладная статистика [Текст] / Айвазян С.А., Мхитарян В.С. – М.: ЮНИТИ-ДАНА, 2001. – 656 с.

93. Орлов, А.И. О современных проблемах внедрения прикладной статистики и других статистических методов (Обобщающая статья) [Текст] / А.И. Орлов // Заводская лаборатория. - 1992. - № 1. - С.67-74.
94. Налимов, В.В. Применение математической статистики при анализе веществ [Текст] / В.В. Налимов. - М.: Физматгиз, 1960. - 430 с.
95. Венцель, Е.С. Теория вероятностей [Текст] / Е.С. Венцель. - М.: Высш. шк., 1998. - 576 с.
96. Сиберт, У.М. Цепи, сигналы, системы: В 2-х ч. [Текст]: пер. с анг. / У.М. Сиберт. - М.: Мир, 1988. - Ч.2. - 360 с.
97. Левин, Б.Р. Теоретические основы статистической радиотехники [Текст] / Б.Р. Левин. – М.: Советское радио, 1969. - Т.1. – 752 с.
98. Бендат, Дж. Прикладной анализ случайных данных [Текст]: пер. с анг. / Дж. Бендат, А. Пирсон. - М.: Мир, 1989. – 540 с.
99. Баррер, М. Ракетные двигатели [Текст]: пер. с франц. / М. Баррер [и др.]. - М.: Оборонгиз, 1962. - 799 с.
100. Марчук, В.И. Новый способ повышения достоверности результатов измерений [Текст] / В.И. Марчук, К.Е. Румянцев // Авиакосмическое приборостроение. - 2004. - №2. - С.51-55.
101. Марчук, В.И. Анализ методов адаптации порогового значения при обнаружении аномальных измерений [Текст] / В.И. Марчук, К.Е. Румянцев / Известия вузов России. Радиоэлектроника. - 2006. - Вып.1. - С.20-25.
102. Справочник по радиолокации [Текст] / Под ред. М.Скольника. Нью-Йорк, 1970 / Пер. с анг. (в 4-х томах) под общей ред.К.Н. Трофимова. Т.1. Основы радиолокации.- М.: Сов. радио, 1976.-456 с.
103. Токарева, С.В. Сравнительный анализ модификации нового метода обнаружения аномальных измерений [Текст] / С.В. Токарева, А.И. Шерстобитов // Системный подход в науках о природе, человеке и

технике: Материалы международной научной конференции. - Таганрог: ТРТУ, 2003. - Ч.5. - С.99-103.

104. Токарева, С.В. Способ уменьшения погрешности при использовании метода размножения оценок [Текст] / С.В. Токарева, А.И. Шерстобитов, // Моделирование. Теория, методы и средства: Материалы 4-ой международной научно-практической конференции. - Новочеркасск: ЮРГТУ, 2004. - С.4-5.

105. Токарева, С.В. Повышение вероятности правильного обнаружения аномальных измерений при параметрической неопределённости [Текст] / С.В. Токарева, А.И. Шерстобитов // Известия ВУЗов Северо-Кавказкий регион «Техника, Технология и экономика сервиса». Технические науки. - Новочеркасск: ЮРГТУ, - 2004. - приложение № 6. – С.197-199.

106. Марчук, В.И. Анализ вероятностных характеристик адаптивного порога при обнаружении аномальных измерений [Текст] / В.И. Марчук, С.В. Токарева // Цифровая обработка сигналов и её применение: Труды Российского научно-технического общества радиотехники, электроники и связи им. А.С.Попова. -, Москва, 2005. - Вып.VII-2. - С.387-389.

107. Токарева, С.В. Возможность адаптации порогового значения при симметричных и несимметричных законах распределения шумовой составляющей [Текст] / С.В. Токарева // Цифровые методы и технологии: Материалы международной научной конференции. - Таганрог: ТРТУ, 2005. - Ч.3. – С.61-63.

108. Свидетельство об официальной регистрации программ для ЭВМ № 2006611185. Обнаружение аномальных измерений при возможности фиксирования вероятности ложной тревоги / Марчук В. И., В.В. Воронин, Токарева С.В., Шерстобитов А.И.; заявитель и патентообладатель Южно-Рос. Гос. ун-т экономики и сервиса. – заявл. 13.02.06 г., г. Москва. – опуб. 3.04.06 г.

109. Пат. 2005131632 Российская Федерация, С1, МПК G06F 17/18. Устройство для обнаружения и устранения аномальных измерений / Марчук В. И., В.В. Воронин, Токарева С.В., Шерстобитов А.И.; заявитель и патентообладатель Южно-Рос. Гос. ун-т экономики и сервиса. – № 2301445; заявл12.10.2005; опубл. 20.06.2007, Бюл. № 17. – 16 с. 9 ил.
110. Лихарев, В.А. Проектирование радиолокационных систем с цифровой обработкой информации [Текст] / В.А. Лихарев. – М.: Советское Радио, 1968. – 161 с.
111. Лихарев, В.А. Цифровые методы и устройства в радиолокации [Текст] / В.А. Лихарев. – М.: Советское Радио, 1973. – 451 с.
112. Лихарев, В.А. Выбор оптимальных параметров цифровых обнаружителей пачек некогерентных эхо-сигналов [Текст] / В.А. Лихарев // Проблемы оптимальной фильтрации: труд второй научно-технической конференции; под ред.канд.тех.наук, доцента А.Д. Дымовой. – М: Сов.радио, 1968. – С.456.
113. Бусленко, Н.П., Голенко Д.И., Соболь И.М., и др. Метод статистических испытаний (метод Монте-Карло) [Текст] / Н.П. Бусленко [и др.]. – М: ГИФМЛ, 1962. – 236 с.
114. Жовинский, А.Н. Инженерный экспресс–анализ случайных процессов [Текст] / А.Н. Жовинский, В.Н. Жовинский. – М.: Энергия, 1979. – 113 с.
115. Бендат, Дж. Измерение и анализ случайных процессов [Текст]: пер. с анг. . Г.В. Матушевского и В.Е. Привальского. / Дж. Бендат, А. Пирсон. - М.: Мир, 1974. – 536 с.
116. Ван дер Варден, Б.Л. Математическая статистика [Текст]: пер с нем. Л.Н. Большева / Под ред. Н.В. Смирнова. / Б.Л. Ван дер Варден. - М.: Иностранная литература, 1960. – 384с.

117. Стратонович, В.Л. Принципы адаптивного приема [Текст] / В.Л. Стратонович. - М.: Советское радио, 1973. – 340 с.
118. Практические аспекты цифровой обработки сигналов [Текст]: Монография / под ред.В.И.Марчука. - Шахты: Изд-во ЮРГУЭС, 2007. – 207 с.
119. Пат. 2005132148 Российская Федерация, С1, МПК G06F 15/00. Способ обнаружения аномальных измерений без оценки функции тренда и устройство его реализующее / Марчук В. И., В.В. Воронин, Токарева С.В., Шерстобитов А.И.; заявитель и патентообладатель Южно-Рос. Гос. ун-т экономики и сервиса. – №2302655; заявл. 17.10.2005; опубл. 10.07.2007, Бюл. № 19. – 14 с. 9 ил.
120. Марчук, В.И. Способ обнаружения аномальных значений при анализе нестационарных случайных сигналов [Текст] / В.И. Марчук, С.В. Токарева // Известия ЮФУ. Технические науки. Тематический выпуск “Безопасность телекоммуникационных систем”.- Таганрог: Издательство ТТИ ЮФУ, 2008, №3(80), С. 66-72.
121. Токарева, С.В. Использование адаптивного метода обнаружения аномальных измерений при заданной вероятности ошибки первого рода [Текст] / С.В. Токарева // Современные проблемы радиоэлектроники: Материалы первой межрегиональной научной конференции. - Ростов-на-Дону, 2006. - С.132-134.
122. Токарева, С.В. Использование адаптивного метода обнаружения аномальных измерений для нестационарных случайных процессов [Текст] / С.В. Токарева // Статистические методы в естественных, гуманитарных и технических науках: Материалы международной научной конференции. - Таганрог: ТРТУ, 2006. - Ч.3. - С.83-86.
123. Токарева, С.В. Оценка эффективности обнаружения аномальных значений при наличии корреляции между отсчетами шумовой составляющей [Текст] / С.В. Токарева // Информация, сигналы,

системы: вопросы методологии, анализа и синтеза: Материалы международной научной конференции. - Таганрог: ТИЮФУ, 2008. - Ч.4. - С.72-75.

124. Токарева, С.В. Исследование эффективности адаптивного метода обнаружения аномальных значений при мультипликативной шумовой составляющей [Текст] / С.В. Токарева // Цифровая обработка сигналов и её применение: Труды Российского научно-технического общества радиотехники, электроники и связи им. А.С.Попова. - Москва, 2007. - Вып.IX-2. - С.386-389.

125. Токарева, С.В. Сравнительный анализ обнаружения аномальных значений при аддитивной и мультипликативной шумовой составляющей [Текст] / С.В. Токарева, В.В. Воронин, Е.А. Семенищев // Перспектива-2007: Материалы международного конгресса студентов, аспирантов и молодых ученых. - Нальчик: Каб.-Балк. Ун-т, 2007. - Том II. - С.79-81.

126. Токарева, С.В. Исследование модификации метода обнаружения аномальных значений [Текст] / С.В. Токарева, В.В. Воронин // Наука и образование без границ: Материалы 3 международной научно практической конференции. – София: «Бял ГРАД-БГ» ООД, 2007. – Т.16.Технологии. – С.74-77.

127. Марчук, В.И. Использования адаптивного метода при анализе стационарных случайных процессов [Текст] / В.И. Марчук, С.В. Токарева, А.И. Шерстобитов // Информационные технологии в современном мире: Материалы международной научной конференции. - Таганрог: ТРТУ, 2006. - Ч.2. - С.76-78.

128. Токарева, С.В. Анализ эффективности метода обнаружения аномальных значений без выделения функции полезной составляющей процесса [Текст] / С.В. Токарева, В.В. Воронин, А.И. Шерстобитов // Методы и средства управления технологическими процессами:

Материалы IV Международной конференции. – Саранск: Мордовский государственный университет, 2007. - С.225-227.

129. Токарева, С.В Обнаружение аномальных измерений при многопороговом уровне анализа нестационарного случайного процесса [Текст] / С.В. Токарева // Цифровая обработка сигналов и её применение: Труды Российского научно-технического общества радиотехники, электроники и связи им. А.С.Попова. – Москва, 2008. - Вып.X-2. - С.537-540.

130. Токарева, С.В Обнаружение аномальных измерений в реализации нестационарного случайного процесса с мультипликативной шумовой составляющей [Текст] / С.В. Токарева, В.В. Воронин // Радиолокация и навигация: Научно-техническая конференция. – Воронеж, 2007. – Том 1. – С.29–31.

131. Брандт, З. Статистические методы анализа наблюдений [Текст] / З. Брандт. – М.: Мир, 1975. – 312 с.

132. Дуброва, Т.А. Статистические методы прогнозирования: Учеб. Пособие для вузов [Текст] / Т.А. Дуброва. – М.: ЮНИТИ – ДАНА, 2003. – 206 с.

133. Кокс, Д. Прикладная статистика. Принципы и примеры [Текст]: пер. с анг. / Д. Кокс, Э. Снелл. – М.: Мир, 1984. – 200 с.

134. Токарева, С.В Сравнительный анализ метода экспоненциального сглаживания и размножения оценок при выделении полезного сигнала [Текст] / С.В. Токарева, А.И. Шерстобитов // Измерения и контроль. Информация: Материалы международной технической конференции. – Барнаул, 2003. – С.75–80.

135. Большаков, И.А. Прикладная теория случайных потоков [Текст] / И.А. Большаков, Ракошиц В.С. – М.: Советское радио, 1978. – 248 с.

136. Яровой, Н.И. Адаптивная медианная фильтрация [Электронный ресурс] / Н.И. Яровой // Информационный портал Центрального

научно-исследовательского института ВВВ. – 2007. – Режим доступа – http://www. controstyle.ru

137. Черненко, С.А. Медианный фильтр [Электронный ресурс] / С.А. Черненко // Информационный портал Центрального научно-исследовательского института ВВВ.–2008.–Режим доступа – http://www.logis-pro.kiev.ua/math_medianfilter_ru.html

138. Радченко, Ю.С. Эффективность приема сигналов на фоне комбинированных помехи с дополнительной обработкой в медианном фильтре [Текст] / Ю.С. Радченко // Радиоэлектроника. - 2001. - № 7. – С.30-37.

139. Хеннан, Э. Анализ временных рядов [Текст]: пер.с анг. / Э. Хеннан. – М.: Мир, 1974. – 546 с.

140. Марчук, В.И. Выделение тренда цен и других экономических показателей методом размножения оценок при единственной исходной реализации экономического процесса [Текст] / В.И. Марчук, Б.Ю. Сербиновский // Организация и управление производительностью производственных систем: Материалы международной научной конференции. - Новочеркасск: Юж.-Рос.гос.техн. ун-т (НПИ), 2003. - Ч.2. - С.85-97.

141. Марчук, В.И. Анализ погрешностей прогнозирования экономических временных рядов и способы их уменьшения [Текст] / В.И. Марчук // Научная мысль Кавказа. Спец. Выпуск. - 2004. - №2. - С.85-90.

142. Троцкий, И.Н. Оптимальная обработка информации: (становление и развитие принципов) [Текст] / И.Н. Троцкий – М.: Знание, 1990. – 64 с.

143. Баранников, А.А. Изучение периодической переменности оптического блеска и лучевых скоростей "убегающей" звезды HD 218915 [Текст] / А.А. Баранников, А.В. Меркулова // Известия вузов. Северо-Кавказский регион, Естественные Науки. – Ростов на дону, 2007. - № 2. – С.54-56.

144. Марчук, В.И. Способ обработки измерений для алгоритмов диагностики аккумуляторов [Текст] / В.И. Марчук, Н.В. Рудов // Математические методы в технике и технологиях: Материалы XVI международной конференции. - Ростов на Дону, 2003. - Т.5. - С.85-90.
145. Маслов, М.Д. Способ определение остаточной емкости свинцового аккумулятора. Ф.С. №2120158, – Опуб.10.10.1998.