

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ СТАТИСТИЧЕСКОЙ ОБРАБОТКИ ДАННЫХ ВЫБОРОК ОГРАНИЧЕННОГО ОБЪЕМА

**© 2007 г. Н.В. Колыхан, В.С. Тюрчев
Таганрогский технологический институт
Южного федерального университета**

Вопросы разработки и применения методов и средств исследования динамического состояния малоинерционных технологических объектов (ТО) всегда оставались актуальными [1, 2]. Важное значение имеет исследование динамики поведения ТО в нештатных режимах функционирования.

Очевидно, что длительность пребывания технологических объектов в нештатных режимах кратковременна и отражается весьма ограниченным объемом данных. В функции известных управляющих систем и комплексов входит обработка статистической информации с использованием классических методов, требующих больших объемов информации для получения приемлемых по точности результатов.

В подобных системах не предусмотрено их функционирование в случае малого объема данных, отсутствует анализ характера выбросов и формирование модели контролируемого параметра и, соответственно, не принимаются решения об управляющих воздействиях.

Это объясняется отсутствием соответствующих информационных технологий статистической обработки эмпирических данных ограниченного объема.

Анализ классических подходов по обработке стохастических массивов

Прежде чем осуществить анализ, классифицируем традиционные информационные технологии по сбору и обработке стохастических массивов. Предлагаемая ниже классификация достаточно условная, однако отражает основные особенности известных методов. Классификация методов производится в зависимости от того, на решение какой основной задачи математической статистики [2] направлен данный метод. Таким образом, методы обработки статистических данных можно разделить на следующие основные классы:

- 1) Методы, позволяющие определять закон распределения по статистическим данным. При обработке обширных по своему объему статистических данных

часто возникает вопрос об определении тех или иных случайных величин. Теоретически, при достаточном количестве опытов, свойственные этим случайным величинам закономерности будут осуществляться сколь угодно точно. На практике всегда приходится иметь дело с ограниченным количеством экспериментальных данных; в связи с этим результаты наблюдения и их обработки всегда содержат больший или меньший элемент случайности;

- 2) Методы проверки правдоподобности гипотез. Статистический материал может с большим или меньшим правдоподобием подтверждать или не подтверждать справедливость той или иной гипотезы. Поэтому проверяется, согласуются ли результаты эксперимента с гипотезой о том, что данная случайная величина подчинена закону распределения $F(x)$; указывает ли наблюденная в опыте тенденция к зависимости между двумя случайными величинами на наличие действительно объективной зависимости между ними или же она объясняется случайными причинами, связанными с недостаточным объемом статистики. Для решения подобных вопросов выработан ряд специальных приемов, используемых проблемно-ориентированными технологиями;
- 3) Методы определения неизвестных параметров распределения. Часто при обработке статистического материала вовсе не возникает вопрос об определении законов распределений исследуемых случайных величин. Обыкновенно это бывает связано с недостаточным объемом экспериментального материала. Иногда же характер закона распределения качественно известен до опыта, из теоретических соображений. Тогда возникает более узкая задача обработки наблюдений – определить только некоторые параметры (числовые характеристики) случайной величины или системы случайных величин. При небольшом числе опытов задача более или менее точного определения этих параметров не может быть решена; в этих случаях экспериментальный материал содержит в себе неизбежно элемент случайности; поэтому случайными оказываются и все параметры, вычисленные на основе этих данных. В таких условиях может быть поставлена только задача об определении так называемых «оценок» или «подходящих значений» для искомых параметров, т.е. таких приближенных значений, которые при массовом применении приводили бы в среднем к меньшим ошибкам, чем всякие другие. С задачей отыскания «подходящих значений» числовых характеристик тесно связана задача оценки их точности и надежности.

Таким образом, из классификации видно, что существующие и разрабатываемые статистические методы по сбору и обработке стохастических массивов, направленные в основном на решение задач математической статистики, предполагают использование больших массивов выборок. Однако практически разработчик поставлен в более жесткие условия и требуется принятие решения на основе малого объема данных. Очевидно, необходимо отыскание таких подходов и методологий, которые способны функционировать в реальном масштабе времени, используя имеющийся статистический материал, как большого, так и малого объема.

Источник основных трудностей при непараметрическом оценивании законов распределения лежит в различной структуре исходных данных, представленных конечной выборкой, и множества результатов, которые образуют функцию. Получается, что по конечной выборке - точкам p -мерного пространства - нужно восстановить вид поверхности $W^*(x) = W_p^*(x_1, x_2, \dots, x_n)$ в $(p+1)$ -мерном пространстве и сделать это так, чтобы она была наиболее близка к априорно неизвестной поверхности $w(x)$ в каждой точке x области своего определения. Поэтому главные усилия должны быть направлены на поиск минимальных допущений, которые позволили построить, по крайней мере, одну оценку плотности, сходящуюся к истинной плотности по вероятности или в каком-либо вероятностном смысле. В настоящее время при непараметрическом оценивании, в основном, используют гистограммный метод, метод Парзена, метод разложения по базисным функциям, метод полигонов Смирнова, метод локального оценивания по k -ближайшим соседям, а также ряд специальных методов нелинейного оценивания. Объемы выборок при традиционном подходе обработки статистического материала, как отмечалось ранее, составляют десятки и сотни значений. Так применение статистического моделирования, основанного на классическом методе обработки данных, требует проведения большого числа испытаний для достижения заданной точности результата, что приводит зачастую к чрезвычайно большим затратам машинного времени на решение задач анализа и синтеза даже на быстродействующих ЭВМ. В связи с этим возникает проблема усовершенствования процедуры статистического моделирования таким образом чтобы сократить затраты машинного времени.

Статистические погрешности зависят от трех параметров, с которыми оперируют при построении функции плотности и интегральной функции распреде-

ления, один из этих параметров – объем выборки N . Величина N должна выбираться по возможности большей. Так, например, при $N=200$ для средней по оси ординат части распределений относительные статистические погрешности составляют 20–40%. Для значений плотности распределения вблизи моды точность выше, а на «хвостах» распределений относительные погрешности могут составлять сотни процентов [1]. Практически всегда приходится иметь дело с ограниченным объемом выборок случайных величин. Поэтому возникает необходимость в определении того минимального объема статистических данных, который обеспечивает достаточную для практики точность полученных результатов. При определении объема данных необходимо учитывать вид конечного результата – эмпирическую вероятность, выборочное среднее, закон распределения вероятностей.

При ограниченном ресурсе времени или нецелесообразности получения больших объемов выборок стоит вопрос обоснования предельного значения их объемов. Поскольку разбросы выборочных значений случайны, то обусловленная ими некоторая неточность результата эксперимента в значительной мере определяется размером выборки. Иногда задача определения такого размера выборки, который позволяет обеспечить желаемый уровень точности и в то же время минимизировать затраты на получение этой выборки решается следующим образом.

Необходимый объем статистических данных определяют на основе вычисления вероятности приближения оценок, получаемых по выборкам, к соответствующим вероятностным характеристикам. Используя неравенство Чебышева к выборочному среднему, получают:

$$\beta = P\left\{ \left| X - m_x \right| \leq \varepsilon \right\} \geq \frac{D_x}{\varepsilon^2}, \quad (1)$$

где X – эмпирическое среднее;

D_x – дисперсия случайной величины X ;

m_x – математическое ожидание величины X ;

$\varepsilon > 0$ – вероятность расхождения величин X и m_x .

Дальнейшие преобразования (1) приводят к формуле для вычисления доверительной вероятности:

$$\beta \cong 1 - \frac{1}{N\Delta_\beta^2}, \quad (2)$$

где $\varepsilon = |X - m_x|$;

$$\Delta_\beta = \frac{|X - m_x|}{\sigma_x}$$

σ – дисперсия случайной величины X .

На рис. 1 приведен график зависимости объема выборки N от доверительной вероятности β и относительной погрешности $\Delta\beta$. Следует отметить, что использование неравенства Чебышева при вычислении объема выборки дает завышенную оценку для N , однако порядок объема выборки остается достаточно большим.

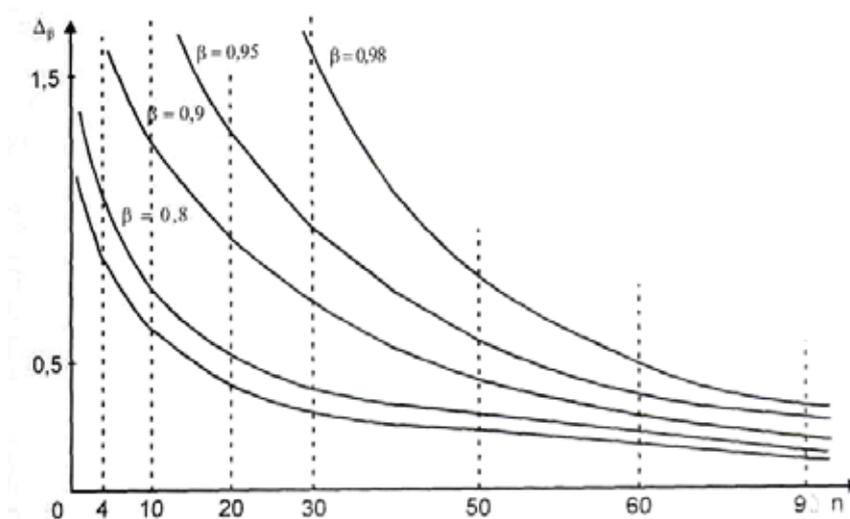


Рис 1. График зависимости объема выборки N от $\Delta\beta$ и β

На основании анализа традиционных технологий статистической обработки эмпирических данных можно сделать следующие выводы:

- В функции известных анализаторов, построенных по классическим статистическим методам, входит обработка статистической информации с использованием классических методов, требующих больших объемов информации для получения приемлемых по точности результатов;
- Большинство из них ограничиваются лишь первичной обработкой, вычислением простейших статистических характеристик параметров без проведения полной статистической обработки экспериментальных данных.
- В подобных системах не предусмотрено их функционирование в случае малого объема данных; отсутствует анализ характера выбросов и форми-

рование модели контролируемого параметра и, соответственно, не принимаются решения об управляющих воздействиях.

- Требуемые объемы выборок при традиционных технологиях обработки стохастических массивов составляют от 40 до 50 значений при уровне доверительной вероятности 80%.

Предпосылки и постановка задачи синтеза

На всех этапах жизненного цикла неравновесные системы антропогенной среды (НСАС) (например, судовые, авиационные двигатели, энергетические установки), представляющие собой сложный технический комплекс, фрагментарно и в целом, постоянно подвергается различного рода испытаниям, объемы и трудоемкость которых постоянно возрастают.

Анализ функциональных, информационных характеристик испытательных систем, позволил выделить ряд их особенностей и сформировать постановку задачи работы. По нашему мнению, испытательные системы, наряду с достоинствами, обладают рядом недостатков: непрерывная регистрация измерительной информации позволяет повысить точность и достоверность контроля параметров, однако это чаще всего приводит к значительной информационной избыточности и к неоправданно большим аппаратным и временным затратам на ее хранение, поиск, обработку; реализация централизованного контроля НСАС на основе циклического опроса датчиков, однако, при таком алгоритме отсутствует функциональная зависимость между заданной программой обслуживания контролируемых параметров и их состояниями, что делает систему инерционной, приводит к излишней информационной избыточности, «старению» информации за цикл и снижению достоверности контроля; обработка накопленной измерительной информации осуществляется с использованием классических статистических методов, требующих больших объемов информации для получения приемлемых по точности результатов; из-за значительной стоимости летных испытаний, их нестабильности в большинстве случаев не удается собрать необходимый объем однородного статистического материала для первичной обработки с использованием классических методов.

В ряде случаев результаты испытаний отображаются критически малыми объемами статистических выборок ($n \sim 5-10$ значений параметра), обработка которых требует принципиально новых подходов. Постановка задачи включает-

ся в разработке прикладных алгоритмов и программного обеспечения обработки статистических выборок критически малых объемов, предназначенных автоматизированных рабочих мест (АРМ) испытательных комплексов и стендов для синтеза и верификации статистических моделей исследуемых НСАС [3].

Синтез технологий обработки массивов данных ограниченного объема [3]

Контролируемый объект (ОК) отображается n контролируемыми параметрами $x_i(t) \in X$ с соответствующими допусками в $[x_i^e, x_i^b]$, каждый из которых представляет собой случайную функцию $x_i(t, \xi_i, S_i)$ от неслучайного аргумента времени t , режима функционирования ξ_i , динамического состояния S_i элемента объекта и других флуктуационных факторов (качество конструктивно-технологических решений контролируемого объекта, влияние внешней среды, других компонентов) (рис. 2).

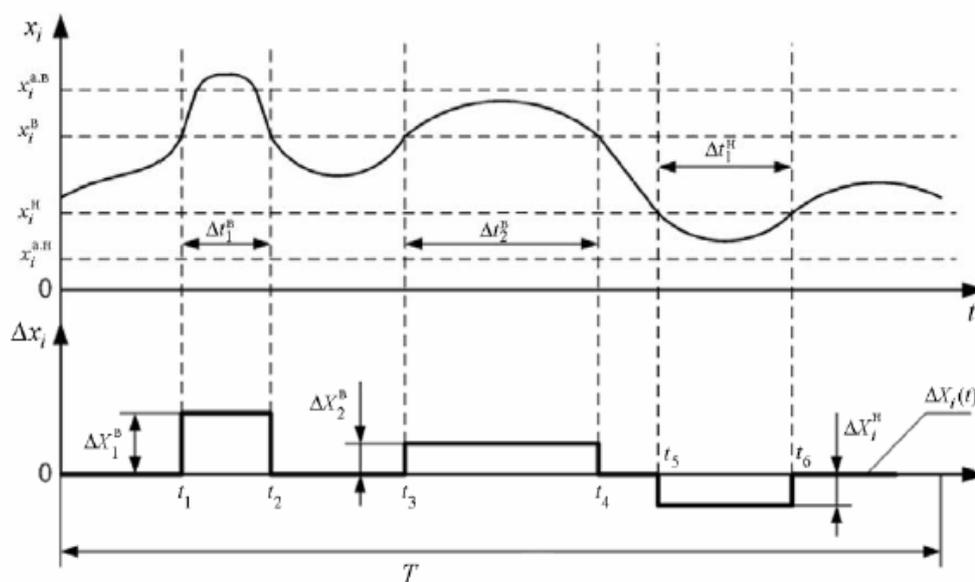


Рис. 2. График реализации $x_i(t)$ процесса относительно заданных $[x_i^e, x_i^b]$ уровней ($x_i^{a.e}, x_i^{a.b}$ - соответственно верхний и нижний аварийные уровни).

Очевидно, что процесс изменения качества контролируемого объекта может быть отображен последовательностью выбросов случайного процесса $X = \{x_1(t_1, \xi_1, S_1), \dots, x_n(t_n, \xi_n, S_n)\}$ параметров над априорно заданными допусками. Диагностика состояния ОК формально может быть сведена к следующим задачам определения: вероятностных характеристик (среднее значение, дисперсия, плотность вероятности и т. д.) числа пересечений $\lambda_i(x_i, t, T)$ заданного уровня $[x_i^e, x_i^b]$

случайным процессом $x_i(t)$ с положительной $x_i'(t)$ (снизу вверх) или с отрицательной (сверху вниз) производной в единицу времени или на интервале T ; вероятностных характеристик числа λ_{\max} (λ_{\min}) максимумов (минимумов) случайного процесса в единицу времени; вероятностных характеристик длительностей выбросов $\Delta t^g(x_i)$, $\Delta t^h(x_i)$ над заданными уровнями, интервалов между выбросами.

Таким образом, сущность статистической диагностики состояния ОК такова: использование нового диагностического параметра — выбросов случайного процесса, т. е. кратковременных превышений контролируемых параметров пределов априорно заданных допусков $[x_i^g, x_i^b]$, $x_i(t) \in [(x_i^{a.g} - x_i^b), (x_i^h - x_i^{a.h})]$; учет свойств нового параметра не как переменной во времени величины, а как генеральной совокупности случайного процесса; установление связи стохастических свойств выбросов случайного процесса с изменениями параметров диагностируемого ОК (характер отказа, адрес параметра).

За время T сформировано для каждого i -го ($i = \overline{1, n}$) параметра два стохастических массива значений амплитуд $|\Delta x_j^g|, |\Delta x_j^h|, j = \overline{1, n}, k = \overline{1, M}$ и $|\Delta t_j^g|, |\Delta t_j^h|$ длительности превышений траектории случайного процесса, именуемых малыми выборками, которые не могут быть обработаны традиционным методом ввиду их малости ($N, M \sim 5-10$). Отообразим процесс формирования выборок по значению и времени графиком реализации случайного процесса X на рис. 2. В основе информационных технологий статистической обработки данных эмпирических выборок ограниченного объема положен метод аддитивной аппроксимации значений выборки линейной суммой определенных (базисных) функций (вкладов) симметричных вероятностных распределений:

$$f^*(x) = \frac{1}{n+1} \left[f_0(x) + \sum_{i=1}^n \Psi_{x_i}(x) \right], \quad (3)$$

где $f_0(x)$ и $\sum_{i=1}^n \Psi_{x_i}(x)$ соответственно априорная и эмпирическая компоненты:

$$f_0(x^{(i)}) = \begin{cases} \frac{1}{x_j^{a.g} - x_j^g}, x_j^g \leq x_j \leq x_j^{a.g}, \\ \frac{1}{x_j^h - x_j^{a.h}}, x_j^{a.h} \leq x_j \leq x_j^h, \\ 0, x_j \in [x_j^g, x_j^h]; \end{cases} \quad \Psi_{x_i}(x^{(j)}) = \begin{cases} \frac{1}{d}, x^{(j)} \in \left[x^{(j)} - \frac{1}{d}, x^{(j)} + \frac{1}{d} \right] \\ 0, x^{(j)} \notin \left[x^{(j)} - \frac{1}{d}, x^{(j)} + \frac{1}{d} \right] \end{cases} \quad (4)$$

где d – ширина вклада, определяемая как $d = k(x^e - x^{a.e})$; $d = k(x^{a.n} - x^n)$;

k – коэффициент, $k \in [0,1]$, значение которого принимается эмпирическим путем.

Отметим, что функция вклада задается симметрично относительно значения x_i , т. е. математическое ожидание базисного вклада $m_{x_i} = x_i$ в конечном интервале длиной d , являющемся дисперсионным размахом вклада. Однако, формулы (3) и (4) еще не дают приемлемой эмпирической плотности, построенной по $N(M)$ значениям случайной величины X . Дело в том, что для некоторых значений x_i соответствующие функции вклада $\Psi_{x_i}(x)$ будут выходить за пределы $[x_j^{a.e}, x_j^b] \cup [x_j^{a.n}, x_j^n]$ и, значит, функция $f^*(x)$ будет отлична от нуля и может находиться вне этого интервала, что противоречит (11). Поэтому выходящую за пределы часть площади данного вклада $\Psi_{x_i}(x)$ перераспределяют над вкладами, лежащими внутри основания $[a, b]$. Исправленные таким образом функции вкладов вносятся в выражение (8).

Синтез модели посредством компьютерной геометрии

Задаваясь определенным значением «ширины» d функции вклада, нетрудно синтезировать алгоритм вычисления функции плотности $f^*(x)$ сразу по всем значениям $x_i \in X, i = \overline{1, N}$, затем графическим или численным интегрированием можно получить функцию распределения $F^*(x)$.

Таким образом, (3), (4) есть эмпирическая функция плотности вероятности верхних выбросов значений параметра $x^{(j)}(t)$ над допустимой зоной x_j^e , которую необходимо затем отнести к какой-либо стандартной функции распределения, т. е. верифицировать. Для этого интервалы $[x_j^{a.e}, x_j^b], [x_j^{a.n}, x_j^n]$ разбиваются на k равных подынтервалов и вычисляются частичные $S_i (i = \overline{1, k})$ площади, т. е. секторов, ограниченные i -м подынтервалом, а затем определяются частоты $P_i^*(x) = \frac{S_i}{S}$, где S – площадь, ограниченная эмпирической функцией $f^*(x)$ плотности распределения, соответствует условию

плотности распределения, соответствует условию $\sum_{i=1}^k P_i^*(x) = 1$.

Приведенный процесс синтеза гистограммы плотности эмпирического распределения отображен рис. 3.

Следующий шаг алгоритма – идентификация эмпирической функции $f^*(x^{(j)})$ стандартными распределениями. Для оценки их близости воспользуемся одним из критериев адекватности, например критерием Пирсона.

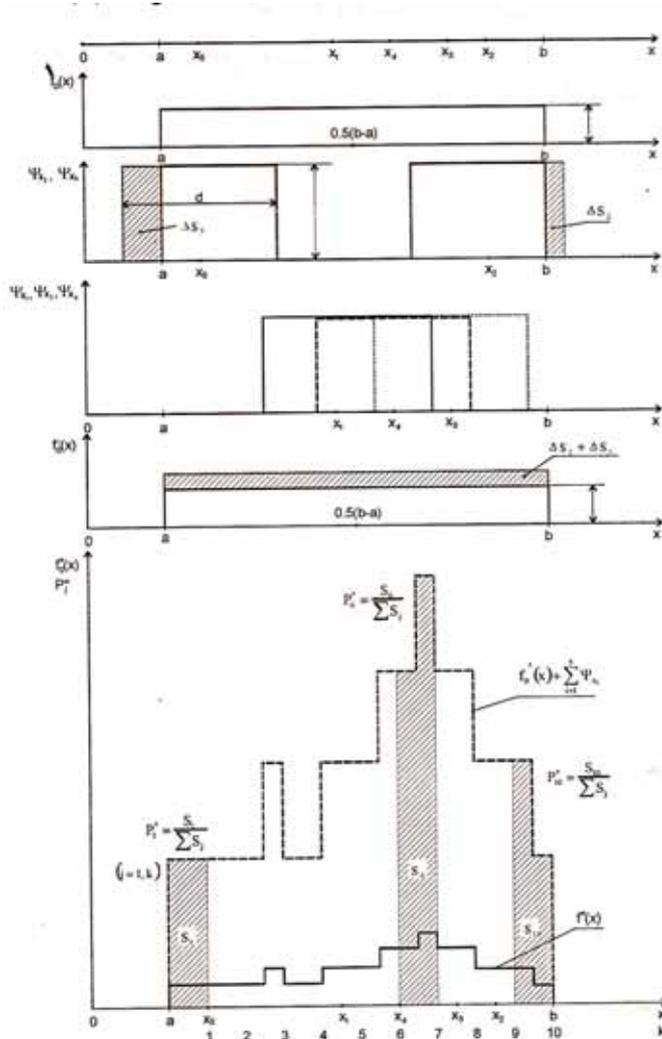


Рис. 3. Диаграмма построения гистограммы

Приведем обобщенный алгоритм аддитивной аппроксимации эмпирического распределения прямоугольными вкладами.

Однако решение поставленной задачи наталкивается на достаточно серьезные затруднения [4]. Эмпирическая функция распределения (3), (4) – аддитивная конструкция из $(N+1)$ симметричных распределений — на основании центральной предельной теоремы в различных формах А. М. Ляпунова должна быть описана нормальным законом. Формально это выглядит так:

$$\limsup_{n \rightarrow \infty} \left(\left| F \left(\sum_{i=1}^n X_i \right) - \Phi(x) \right| \right) = 0, \quad (5)$$

При таком условии аддитивная аппроксимация малой выборки по (3), (4) приводила бы всегда к нормальному закону и идентификация эмпирического распределения была бы проблематичной. Однако анализ работ в области статистики за последние 20 лет показал, что современная теория суммирования случайных величин предполагает отказ от использования результатов классических предельных теорем при определенных условиях: при разных математических ожиданиях, т. е. по теореме Берри — Эссена:

$$\sup_{n \rightarrow \infty} \left(\left| F \left(\sum_{i=1}^n X_i \right) - \Phi(x) \right| \right) \neq 0 \quad (6)$$

где показана возможность аномальных распределений $\sum_{i=1}^n X_i$ случайных величин [4].

Было проведено большое количество экспериментов на имитационных моделях, показавших возможность получения любых аномальных распределений в базисе аддитивных аппроксимаций значений выбросов параметров массива малой выборки стандартными симметричными однородными распределениями и подтвердивших выводы (10) [4].

При аппроксимации значений выборки симметричными вкладами возникает вопрос об оптимальном значении d ширины вклада. По результатам проведенных исследований, d_{opt} для различных распределений различно ($k_{opt} \equiv 0.2 \div 0.6$). Поскольку априори неизвестно распределение X , то в алгоритмах предусмотрен оператор определения $d_{opt} = k_{opt} (x^g - x^{a.g})$ по наилучшему значению критерия адекватности эмпирической плотности $f^*(x)$ и плотности $f(x)$ традиционного распределения путем перебора значений $k_{opt} \in (0,1)$ задаваясь при этом величиной d шага. Малый объем выборки, процедура ее аппроксимации симметричными вкладами являются основными источниками «размытости» эмпирической гистограммы, и, в отличие от интервального метода, по виду гистограммы идентифицировать ее одним распределением в процессе одного цикла алгоритма трудно ввиду «неудовлетворительного» значения крите-

рия согласия. Разрешение этого конфликта заключается в формировании упорядоченного ряда значений соответственно одноименных (Пирсона, энтропийного и наименьших квадратов) критериев согласия по результатам идентификации эмпирической гистограммы посредством перебора ряда классических распределений, выбранных по эмпирическим значениям критериев вариации, асимметрии и эксцесса. Обоснованием выбора гипотезы принятия того или иного закона распределения для описания ОК является наилучшее (наибольшее или наименьшее) ему соответствующее значение критерия согласия упорядоченного (вариационного) ряда, число элементов которого соответствует числу классов распределений. Формально задача принятия гипотезы о выборе класса закона распределения из $l = \overline{1, l}$ может быть отображена в виде

$$H_1 : F^*(x) \rightarrow F_l(x) \text{ при } \begin{cases} P(\chi^2, r) \rightarrow \max; \\ |H - H^*| \rightarrow \min; \\ F_k(x) - F^*(x) \rightarrow \min, \end{cases}$$

где l – классы законов распределений, например: 1 – экспоненциальный, 2 – нормальный, ..., l – Релея.

Если в результате анализа выборки малого объема различные критерии согласия обеспечивают оптимальные значения для нескольких классов законов распределения, то необходим переход в интерактивный режим работы программного процессора и дальнейшие исследования выборки проводятся путем варьирования ширины вклада d , изменения диапазона возможных значений.

С целью повышения достоверности метода (7), (8) произведем обработку тех же значений малой выборки $X = \{x_1, \dots, x_n\}$ по другому алгоритму аддитивной аппроксимации, основанному на принципе ядерного имитационного моделирования. Совпадение гипотез по выбору одноименного закона распределения при идентификации по результатам реализации исследуемых алгоритмов будет подтверждением адекватности выбранной модели динамике изменения реального ОК.

Обработки данных малой выборки посредством ядерного имитационного моделирования.

Принцип ядерного имитационного моделирования заключается в следующем. Априорную $f_0^*(x)$, составляющую выражения (7), представим значениями N_0 случайных чисел, распределенных по равновероятностному (Симпсона,

нормальному) закону в окрестности x_i точки (ядра) $0.5(b-a)$ – числовой оси X с математическим ожиданием $M_{f_0}(x) = 0.5(b+a)$ и дисперсией $D_{f_0}(x) = (b-a)$.

В окрестности d каждого значения x_i выборки X размещаются $N_i (i = \overline{1, n})$ равновероятно (Симпсона, нормальный закон) распределенных случайных чисел с математическим ожиданием $M_{\Psi_{x_i}} = x_i$ и дисперсией $D_{\Psi_{x_i}} = d$. Сложение чисел не производится, они размещаются на интервале $[a, b]$ согласно своим значениям. Таким образом, осуществляется переход от малой выборки $N = 5 \div 10$ к выборке стандартного объема данных, представляющей собой аддитивную суперпозицию $(N_0 + nN)$ псевдореализаций случайных значений. Величины N_0 и N априорно задаются, исходя из требуемой точности и возможности применения интервального метода оценивания. Дальнейшая обработка производится по стандартному алгоритму интервального метода оценивания. Идентификация стандартным распределением, т. е. верификация модели, осуществляется по ранее описанной схеме.

По рассмотренным технологиям обрабатываются массивы малых выборок по значениям превышений, по множеству состояний X и по множеству времени пребывания траектории случайного процесса $x(t)$ над уровнями в $[x_i^a, x_i^b]$ и модель отображается в виде:

$$f(x) = \int_A f^*(x) dx \int_t f^*\left(\frac{1}{A_i}\right) dt.$$

На основании вышеизложенного предлагается методология синтеза архитектуры анализатора процесса обработки и идентификации массивов укороченной выборки, гипотетический вариант которого приведен в приложении 1.

Статистический программный процессор создан по модульной структуре и реализует метод компьютерной графики, заключающийся в построении эмпирической гистограммы с последующей верификацией ее стандартным распределением по методике вариационного ряда критериев адекватности, и метод ядерного имитационного моделирования, заключающегося в аддитивной композиции $(n+1)$ генераторов случайных чисел.

Выводы:

1. Осуществлен анализ традиционных технологий по обработке стохастических массивов с последующим обоснованием постановки задачи.

2. Предложен алгоритм оценки объема эмпирической выборки данных.
3. Введен новый диагностический параметр – выбросы траектории контролируемого параметра над границами допусковых зон.
4. Реализован принцип аддитивной аппроксимации посредством компьютерной геометрии с последующей верификацией эмпирической диаграммы стандартными вероятностными распределениями.
5. Разработан принцип реализации аддитивной аппроксимации значений малой выборки посредством ядерного имитационного моделирования.

Литература

1. Динамическое моделирование и испытания технических систем / Под. ред. И.Д. Кочубиевского. М.: Энергия, 1978
2. Автоматизированные испытания в авиастроении / Р.И. Адгамов, М.М. Берхеев, И.А. Заляев. М.: Машиностроение, 1989.
3. *Колыхан Н.В., Тюряев В.С., Самойленко А.П.* Статистический анализатор и алгоритмы построения моделей контролируемого объекта в базисе эмпирических данных ограниченного объема // Материалы Всероссийского смотр-конкурса научно-технического творчества студентов высших учебных заведений «Эврика-2006», Новочеркасск, 2006, С. 97-98.
4. *Самойленко А.П., Колыхан Н.В., Тюряев В.С.* Анализ отклонений суммы случайных независимых величин от нормального распределения // Материалы восьмого международного научно-практического семинара «Практика и партнерство в науке и учебном процессе в высшей школе». Донецк: Изд-во ДонНТУ, 2007, С. 286-301.