

Анализ задачи идентификации закона распределения случайных процессов

С. А. Базуева, А. А. Михайлов

Южно-Российский государственный политехнический университет, Новочеркасск

Аннотация: Предложена модифицированная методика идентификации закона распределения случайных процессов, основанная на известном методе нормированного размаха (*R/S*-анализа). Для модификации метода идентификации осуществлен информационный анализ задачи идентификации, получены выражения для ее эластичности и обусловленности.

Ключевые слова: Диагностика технических систем, метод нормированного размаха (*RS*-анализ), способ Хёрста, информационный анализ, обусловленность задачи, эластичность задачи.

Постановка проблемы

Экспериментальные данные, полученные в результате мониторинга состояния различных технических систем [1–3], обычно имеют случайную составляющую, которая может быть следствием, например, шума измерительной аппаратуры, конечной точности измерений, случайных воздействий на систему в процессе ее функционирования и т.п. Характерной чертой данных процессов является их нестационарность, которая связана с изменяющимися с течением времени внешними условиями, что определяет изменение во времени параметров динамических свойств системы. В связи с этим при разработке интеллектуальных средств измерений, контроля и диагностики технических систем возникают задачи идентификации и классификации исследуемых процессов [1,2], сопровождающих функционирование технических систем, представленных некоторой совокупностью выборочных значений временного ряда наблюдений [3].

Под задачей идентификации закона распределения наблюдаемой случайной величины (структурно-параметрической идентификации) для различных исследуемых процессов понимают задачу выбора такой параметрической модели закона распределения вероятностей, которая

наилучшим образом соответствует результатам экспериментальных наблюдений [4–8]. При этом построенную модель можно использовать для выбора оптимальных алгоритмов обработки экспериментальных данных исследуемого процесса [9,10].

Сравнительный анализ современных методов идентификации распределений вероятности, проведенный на основании результатов, приведенных в [11, 12], показал, что в настоящее время для решения данной задачи широкое распространение получил так называемый метод нормированного размаха (R/S -анализ) [13,14], который характеризуется минимальными предположениями об исследуемом процессе и достаточной устойчивостью при анализе временных рядов, описывающих данный процесс. Этот способ анализа основан на результатах проведенных Хёрстом экспериментальных исследований [13], которые показали, что большинство естественных явлений следуют «смещенному случайному блужданию»– тренду с шумом, пример которого в виде винеровского процесса приведен на рис. 1. Хёрст для оценивания параметров исследуемого процесса ввел показатель Хёрста H_x , который характеризует отношение силы тренда (детерминированный фактор) к уровню шума (случайный фактор).

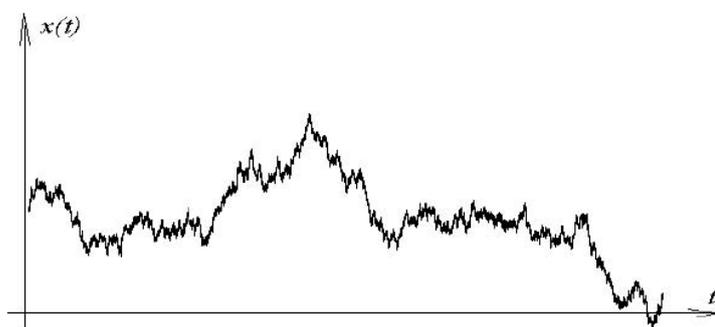


Рис.1. Процесс со «смещенным случайным блужданием»

Целью настоящей работы является исследование свойств показателя Хёрста для теоретического обоснования R/S -способа идентификации закона распределения исследуемых процессов.

Общий анализ R/S -способа идентификации

Цель статистического анализа временного ряда заключается в том, чтобы по имеющейся траектории этого ряда:

- определить, какие из неслучайных функций присутствуют в нем и построить «хорошие» оценки для них;

- подобрать модель, адекватно описывающую поведение случайной составляющей ряда и статистически оценить параметры этой модели.

R/S -способ идентификации может отличить случайный ряд от неслучайного, даже если случайный ряд не гауссовский. Для формализации представления исследуемых процессов используют временной ряд, составленный по результатам наблюдений. Для сравнения различных типов временных рядов Хёрст ввел следующие соотношения:

для интервала времени τ

$$R(\tau)/S(\tau)=(a_{\tau}\tau)^{H_X},$$

где $R(\tau)=\max_{t \in [1, \tau]} X_{\tau}(t) - \min_{t \in [1, \tau]} X_{\tau}(t)$, $X_{\tau}(t)=\sum_{i=1}^{\tau} (x_i - \langle x \rangle_{\tau})$ – размах от накопленного среднего

на интервале наблюдения τ , $\langle x \rangle_{\tau}$ – среднее на интервале наблюдения $[1, \tau]$, t – мгновенное время временного ряда, $1 \leq t \leq \tau$, a_{τ} – соответствующая константа,

$S(\tau)=[\tau^{-1} \sum_{i=1}^{\tau} (x_i - \langle x \rangle_{\tau})^2]^{1/2}$, H_X – показатель Хёрста,

или на выборке N

$$R(N)/S(N)=(a_N \cdot N)^{H_X},$$

где $R(N)=\max_N X_N - \min_N X_N$, $X_N=\sum_{i=1}^N (x_i - \langle x \rangle_N)$ – размах от накопленного среднего

на выборке N , $S(N)=[N^{-1} \sum_{i=1}^N (x_i - \langle x \rangle_N)^2]^{1/2}$, $\langle x \rangle_N$ – среднее на множестве N , a_N –

соответствующая константа, N – число наблюдений.

Оба соотношения Хёрста эквивалентны друг другу и отличаются лишь интервалом определения этих соотношений. Поэтому проведем анализ соотношений Хёрста на примере выражения

$$H_X = \log_{a_N} \frac{R(N)}{S(N)}. \quad (1)$$

При этом выделяют четыре классификации для показателя Хёрста:

1). *Случайный ряд* ($H_X=0,5$). События случайны и некоррелированы. R/S -анализ может классифицировать произвольный ряд, безотносительно к тому, какой вид распределения ему соответствует.

2). *Антиперсистентный (эргодический) ряд* ($0 \leq H_X < 0,5$). Такой тип модели исследуемого процесса называют «возврат к среднему». Если модель исследуемого процесса демонстрирует «рост» в предыдущий период, то, скорее всего, в следующем периоде начнется спад. И наоборот, если шло снижение, то вероятен близкий подъем. Устойчивость такого антиперсистентного поведения зависит от того, насколько H_X близко к нулю. Такой ряд более изменчив, чем ряд случайный, так как состоит из частых реверсов спад-подъем.

3). *Персистентный (трендоустойчивый) ряд* ($0,5 < H_X < 1,0$) – это обобщенное броуновское движение или смещенные случайные блуждания. Если ряд возрастает (убывает) в предыдущий период, то, вероятно, что он будет сохранять эту тенденцию какое-то время в будущем. Чем ближе H_X к 0,5, тем более зашумлен ряд и тем менее выражен его тренд.

4). *Ряд с фрактальным временем* ($H_X > 1$). В этом случае для исследуемого процесса характерны независимые скачки амплитуды (разрывы производной процесса), распределенные по статистике Леви за время, определенное величиной скачка, и растущее вместе с ним. Дисперсия приращения за данный интервал времени становится конечной, траектория в фазовом пространстве сохраняет свой вид.

Проведенный общий анализ R/S -способа идентификации требует определения параметра $S(N)$, для интерпретации смысла которого осуществим информационный анализ рассматриваемого способа

идентификации закона распределения исследуемого процесса и определим понятие интервала неопределенности исследуемого параметра.

Оценка эквивалентного (по энтропии) интервала неопределенности

Для оценки энтропийного интервала неопределенности воспользуемся показателем качества оценивания исследуемого параметра x_t в виде апостериорного риска $I(p_{x_t}, \hat{x}_t | Y_0^r)$ оценки \hat{x}_t . Данный риск определяется как условное математическое ожидание для квадратичной функции потерь $(x_t - \hat{x}_t)^2$ от принятого решения по оцениванию исследуемого параметра x_t в виде оценки \hat{x}_t , при условии, что известна выборка наблюдений Y_0^r :

$$I(p(x_t, \hat{x}_t | Y_0^r)) = \int_{\Omega_x} (x_t - \hat{x}_t)^2 p(x_t | Y_0^r) dx_t = \int_{\Omega_x} (x_t - \hat{x}_t)^2 p(x_t) \frac{p(x_t | Y_0^r)}{p(x_t)} dx_t$$

где $p(x_t | Y_0^r)$ – условная плотность вероятности параметра x_t при условии, что известна выборка наблюдений Y_0^r ; $p(x_t)$ – плотность вероятности параметра x_t , Ω_x – область определения параметра x_t .

Взяв данный интеграл по частям и учтя условие равенства нулю производной функции правдоподобия

$$\frac{d\left(\frac{p(Y_0^r | x_t)}{p(Y_0^r)}\right)}{dx_t} = 0,$$

получаем

$$I(p_{x_t}, \hat{x}_t | Y_0^r) = \sigma_x^2 \frac{p(Y_0^r | x_t)}{p(Y_0^r)}, \quad (2)$$

где σ_x^2 – дисперсия параметра x_t , $p(Y_0^r | x_t)$ – условная плотность распределения вероятностей по выборке Y_0^r при оценке параметра x_t , $p(Y_0^r)$ – апостериорная плотность распределения вероятностей по выборке Y_0^r .

Таким образом, при идентификации процесса исследуемого параметра x_t апостериорная информация об областях неопределенности оценки \hat{x}_t исследуемого параметра x_t по выборке наблюдений Y_0^r определяется апостериорным риском оценки \hat{x}_t и согласно выражению (2) зависит от дисперсии исследуемого параметра x_t и коэффициента, который определяется максимумом функции правдоподобия соответствующего закона распределения.

Функционал апостериорного риска идентификации процесса исследуемого параметра x_t при неизвестной плотности $p(x_t|Y_0^r)$ может быть оценен по выборке

$$x_1, y_1; \dots; x_N, y_N$$

эмпирическим риском вида [15]

$$I_3(N) = \frac{1}{N} \sum_{i=1}^N (y_i - f(x_i, \alpha))^2, \quad (3)$$

где $y_i \in Y_0^r$, α – параметр, принадлежащий множеству Λ , конкретное значение которого определяет конкретную функцию из класса функций $f(x, \alpha)$; $f(x_i, \alpha) = \langle x \rangle_N$ – выделяемый в виде оценки \hat{x}_t (обычно математическое ожидание) тренд исследуемого параметра x_t .

Данная оценка эквивалентного (по энтропии) интервала неопределенности позволяет перейти к информационному анализу метода нормированного размаха.

Информационный анализ метода нормированного размаха

Количество информации как снятая неопределенность, содержащаяся в результатах измерения наблюдаемого процесса, для любого закона распределения вероятностей погрешности, определяется соотношением [16]

$$H = \log \frac{X_2 - X_1}{d_3} = \log n, \quad (4)$$

где n – число различных градаций измеряемой величины в диапазоне $X_2 - X_1$ при данном законе $p(x)$ распределения погрешности; d_3 – эквивалентный (по энтропии) интервал неопределенности.

Причем, основание логарифма в выражении (4) может быть выбрано любым. При выборе в качестве основания логарифма величины $a_N N$, получаем известное выражение Хёрста (1). Таким образом, введенный Хёрстом показатель имеет информационную природу. В связи с этим исследуем энтропийные свойства показателя Хёрста.

Так, энтропия равномерного распределения с диапазоном изменения случайной величины в интервале (a, b) , которая в соответствии с экстремальным свойством максимальна,

$$H_p(x) = \log \frac{b-a}{d_3},$$

совпадает с энтропией Р. Хартли и с точностью до основания логарифма с соотношением (4).

В [16] введено энтропийное значение погрешности $d_3 = 2\Delta_3$ случайной величины, отношение которого к ее среднему квадратическому значению σ_x определяет энтропийный коэффициент k_3 [9, 16], изменяющийся в зависимости от вида закона распределения:

$$k_3 = d_3 / 2\sigma_x = \Delta_3 / \sigma_x. \quad (5)$$

Проведенный информационный анализ задачи идентификации исследуемого процесса x_t позволил определить интервал неопределенности идентификации через эмпирический риск $I_3(N)$, который совпадает с введенной в выражение для R/S -способа идентификации величиной $S(N)$, и выражение (1) может быть переписано в виде

$$H_X = \log_{a_N N} \frac{R(N)}{\sqrt{I_3(N)}}. \quad (1')$$

Энтропийное значение погрешности d_3 , как и условный риск $I(p_x, \hat{x}_t | Y_0^t)$, определяет неопределенность параметра x_t . Оценив размах $R(N)$ от накопленного среднего на выборке N стандартизованной длиной половины доверительного интервала Δ_m , определяемой через дисперсию

$$\Delta_m = k_m \sigma_x,$$

где k_m – коэффициент доверительного интервала [9], а корень квадратный из эмпирического риска $I_3(N)$ половиной эквивалентного (по энтропии) интервала неопределенности $d_3 = 2\Delta_3$, определяемой согласно (5) как

$$\Delta_3 = k_3 \sigma_x,$$

для типовых законов распределения можно оценить показатель $\frac{R(N)}{\sqrt{I_3(N)}}$ как

$$\frac{R(N)}{\sqrt{I_3(N)}} = \frac{k_m}{k_3}.$$

Результаты расчета $\frac{R(N)}{\sqrt{I_3(N)}}$ для типовых законов распределения [17]

по данному выражению приведены в таблице 1.

Таблица 1

Параметры типовых законов распределения

Закон распределения	Энтропийный коэффициент, k_3	Показатель $R(N)/\sqrt{I_3(N)}$
<p>Равномерный (прямоугольный)</p> $f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & x < a, x > b \end{cases}$	$\sqrt{3} = 1,73$ [9, 16]	1
Трапецидальный, отношение оснований		
1/3	1,94	2,19/1,94=1,129
1/2	2,00	2,32/2=1,16

<p>Треугольный (Симпсона)</p> $f(x) = \begin{cases} 0, & x < a \\ \frac{4(x-a)}{(b-a)^2}, & a < x < \frac{b+a}{2} \\ \frac{4(b-x)}{(b-a)^2}, & (b+a)/2 < x < b, x > b \end{cases}$	$\sqrt{3e}/2 = 2,02$ [18]	$2,45/2,02 = 1,21$
<p>Арксинусоидальный</p> $f(x) = \begin{cases} 0, & x < \mu - \lambda \\ \frac{1}{\pi\sqrt{\lambda^2 - (x-\mu)^2}}, & (\mu - \lambda < x < \mu + \lambda) \\ 0, & x > \mu + \lambda \end{cases}$	1,11 [18]	$1,46/1,11 = 1,315$
<p>Вейбулла $f(x) = \begin{cases} \alpha\lambda x^{\alpha-1} \exp(-\lambda x^\alpha), & x \geq 0, \\ 0, & x < 0 \end{cases}$ с параметрами $\lambda > 0, \alpha > 0$</p>		
Вейбулла ($\alpha=1/3$)	0,83 [9, 16]	$1,23/0,83 = 1,48$
Вейбулла ($\alpha=1/2$)	1,35 [9, 16]	$1,786/1,35 = 1,32$
Вейбулла ($\alpha=2/3$)	1,64 [9, 16]	$2/1,64 = 1,22$
Вейбулла ($\alpha=3/4$)	1,9 [9, 16]	$2/1,9 = 1,05$
<p>Симметричный экспоненциальный</p> $f(x) = \frac{\beta}{2\sigma} \exp\left[-\beta\left \frac{x}{\sigma}\right \right] \text{ (Лапласса) } (\alpha=1)$	1,92 [9, 16]	$2/1,92 = 1,04$
Вейбулла ($\alpha=3/2$)	2,05	$1,97/2,05 = 0,96$
<p>Гауссов (нормальный) ($\alpha=2$)</p> $f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right],$ <p>$(-\infty < x < \infty)$</p>	$\sqrt{\pi e}/2 \approx 2,07$ [9, 16]	$3/2,07 = 1,445$
<p>Стьюдента (t-распределение) $f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)\left(1+\frac{x^2}{\nu}\right)^{\frac{\nu+1}{2}}}$</p>		

t – распределение $\nu=4$	1,90 [9, 16]	2,78/1,9=1,46
t – распределение $\nu=5$	1,97 [9, 16]	2,57/1,97=1,3045
t – распределение $\nu=6$	2,0 [9, 16]	2,45/2,0=1,225
t – распределение $\nu=7$	2,023 [9, 16]	2,31/2,023=1,142
t – распределение $\nu=10$	2,047 [9, 16]	2,23/2,047=1,09

Взаимосвязь между видом функции (временного ряда) и показателем Хёрста говорит о возможности его использования для классификации этих функций (рядов).

Идентификация тренда по временным рядам

Временные ряды отражают тенденцию изменения параметров системы во времени, поэтому входным параметром x является момент времени и в случае отсутствия ярко выраженных изменений во времени, общая тенденция сохраняется. Ряд можно описать уравнением вида

$$Y_T = F(t) + E_T,$$

где $F(t)$ – детерминированная функция времени, E_T – случайная величина.

Выделяют моментальные ряды, отражающие абсолютные величины по состоянию на определенный момент времени, и интервальные – относительные величины (показатель за год, месяц, и т.д.). При исследовании данных при помощи рядов для детерминированной функции рассчитываются базисные показатели, которые соотносятся к начальному уровню ряда, и цепные показатели, относящиеся к предыдущему уровню.

Пусть для описания некоторого объекта исследования в последовательные моменты времени используется n -мерный временной ряд $\{x_1(t_1), x_2(t_1+\Delta t), \dots, x_j(t_1+\mu\Delta t), x_n(t_1+(n-1)\Delta t)\}$, $j=1, \dots, n$, $\mu=1, 2, \dots, n$ – размерность вложения (лаг).

Задача определения модели тренда $f(x, \alpha)$ по эмпирическим данным сводится к минимизации неопределенности решения задачи восстановления

зависимостей на классе интегрируемых с квадратом функций тренда $f(x, \alpha)$, которая оценивается величиной среднего риска [15]

$$I(p_x, \hat{x}_t | Y_0^\tau) = \int_{\Omega_x} (x_t - \hat{x}_t)^2 p(x_t | Y_0^\tau) dx_t.$$

Причем, данный функционал при неизвестной плотности $p(x_t | Y_0^\tau)$, как уже отмечалось выше, может быть оценен эмпирическим риском вида

$$I_3(\alpha) = \frac{1}{N} \sum_{i=1}^N (y_i - f(x_i, \alpha))^2$$

по выборке

$$x_1, y_1; \dots; x_N, y_N.$$

Введенная оценка интервала неопределенности исследуемого случайного процесса позволяет осуществить оценку качества решающего правила, минимизирующего на заданной последовательности $x_1, y_1; \dots; x_N, y_N$, эмпирический риск [15] выбора модели тренда $f(x_i, \alpha)$ исследуемого случайного процесса, в виде математического ожидания

$$I_3(\alpha) = \frac{1}{N} \sum_{i=1}^N (y_i - f(x_i, \alpha))^2$$

и определение условий определения тренда

$$\frac{dI_3}{d\alpha} = 2(y_i - f(x_i, \alpha)) \frac{df(x_i, \alpha)}{d\alpha} = 0.$$

Полученное условие минимума эмпирического риска определяет возможность воспользоваться методом “скользящего контроля” [15], сформировав величину

$$T(x_1, y_2; \dots; x_N, y_N) = \frac{1}{N} \sum_{i=1}^N (y_i - f(x_i; \alpha(x_1, y_1; \dots, \overline{x_i, y_i; \dots; x_N, y_N})))^2,$$

где знак $\overline{x_i, y_i}$ указывает на то, что из заданной последовательности исключается пара x_i, y_i .

Методика идентификации нестационарных случайных процессов

Для формирования алгоритма идентификации нестационарных случайных процессов исследуем параметры качества задачи расчета показателя Хёрста.

Исследование эластичности задачи расчета показателя Хёрста.

Выражение (1), определяющее показатель Хёрста H_X , содержит второй неизвестный параметр (константу a_N), входящий в основание логарифма, который в системе координат $\log[R(N)/\sqrt{I_3(N)}]$ (ось ординат) и $\log N$ (ось абсцисс) определяет сдвиг функции $\log[R(N)/\sqrt{I_3(N)}]$ по оси абсцисс. Для определения показателя Хёрста H_X воспользуемся понятием эластичности функции $R(N)/\sqrt{I_3(N)}$ [19]. Для этого учтем, что эластичность степенной функции $(a_N N)^{H_X}$ по N постоянна и совпадает с показателем степени, поскольку

$$E_N [(a_N N)^{H_X}] = a_N^{H_X} H_X N^{H_X-1} \frac{N}{a_N^{H_X} N^{H_X}} = H_X,$$

а с другой стороны эластичность переменной R/I по переменной N равна производной логарифма R/I по логарифму N , поскольку

$$\frac{d \ln[R(N)/\sqrt{I_3(N)}]}{d \ln N} = \frac{N}{R(N)/\sqrt{I_3(N)}} \frac{d[R(N)/\sqrt{I_3(N)}]}{dN} = E_N [R(N)/\sqrt{I_3(N)}].$$

На основании последних двух выражений, перейдя от дифференциала к приращениям, получаем выражение для показателя Хёрста в виде

$$H_X = \frac{N}{R(N)/\sqrt{I_3(N)}} \frac{\Delta[R(N)/\sqrt{I_3(N)}]}{\Delta N}. \quad (6)$$

Исследование обусловленности задачи расчета показателя Хёрста.

В реальных вычислениях показателя Хёрста переменные выражения (1), являясь параметрами измерения, известны с определенной погрешностью, и соответственно вместо точного значения показатель Хёрста определяется приближенным значением. Чтобы результаты вычисления имели

практический смысл, необходимо выполнение требования сходимости к точному значению. Таким образом, корректность задачи расчета показателя Хёрста сводится к выяснению вопроса существования и единственности решения, исследованию устойчивости решения к погрешностям входных данных, т.е. обусловленности вычислительных алгоритмов [20]. Здесь под обусловленностью задачи понимается «чувствительность» погрешности решения к погрешности входных данных. Задача хорошо обусловлена, если число обусловленности порядка 1 или меньше. Если оно $\gg 1$, то задача плохо обусловлена. Бесконечно плохо обусловленная задача – неустойчива в вычислительном смысле. Из выражения (6) может быть получено выражение для абсолютного числа обусловленности в виде

$$\Delta_V = \frac{\Delta[R/I]}{\frac{R/I}{\Delta N}} = \frac{H_x}{N}. \quad (7)$$

Данное выражение позволяет определить основание логарифма в выражении (1), обеспечивающее достаточную точность вычисления показателя Хёрста.

Проведенные исследования качества алгоритма определения показателя Хёрста позволяет сформулировать **методику идентификации нестационарных случайных процессов**, которая включает следующие этапы:

Идентификация тренда исследуемого процесса. При идентификации тренда исследуемого процесса может быть использован метод наименьших квадратов или выделение математического ожидания исследуемого процесса методом “скользящего контроля”.

При этом выполняются следующие шаги.

1. Выбирают из множества характеристических функций $f(x, \alpha)$ конкретную i -ю функцию $f(x, \alpha_i)$, $i = \overline{0, m}$, где m – мощность множества характеристических функций $f(x, \alpha)$.

2. В рамках метода “скользящего контроля” [15] составляют функционал $T_i(x_1, y_2; \dots; x_l, y_l)$ и определяют его значения для конкретной характеристической функции $f(x, \alpha_i)$.

3. Если $T_{i+1}(x_1, y_2; \dots; x_l, y_l) < T_i(x_1, y_2; \dots; x_l, y_l)$, то в качестве эталона для анализа на следующем этапе выбирают $T_{i+1}(x_1, y_2; \dots; x_l, y_l)$.

Если $T_{i+1}(x_1, y_2; \dots; x_l, y_l) \geq T_i(x_1, y_2; \dots; x_l, y_l)$, то в качестве эталона для анализа на следующем этапе сохраняют $T_i(x_1, y_2; \dots; x_l, y_l)$.

4. Если $i < m$, то $i = i + 1$ и переходят к п. 1.1.

5. Если $i = m$ функционал $T_i(x_1, y_2; \dots; x_l, y_l)$ объявляется минимальным и для него идентифицируют $f(x, \alpha_i)$, для которой риск идентификации тренда исследуемого процесса также минимален.

II. *Идентификация закона распределения исследуемого процесса по результатам вычисления показателя Хёрста производится по схеме, приведенной в [13]:*

1. Вычисляется отклонение от среднего значения $\langle x \rangle_{N_{CO}}$, которое в отличие от известной схемы [13] определяется не на всем временном ряде, а на части N_{CO} общей совокупности результатов измерения N , образующих анализируемый временной ряд, и приравнивается соответствующему значению выделенного тренда $f(x, \alpha_i)$:

$$X_{N_{CO}, N} = \sum_{i=1}^{N_{CO}} (x_i - \langle x \rangle_{N_{CO}}),$$

где N —длина периода, меняющаяся от 2 до <длины временного ряда>; N_{CO} —переменная, меняющая своё значение от 1 до $N-1$; $\langle x \rangle_{N_{CO}}$ —среднее по N_{CO} элементов; x_i —конкретный элемент временного ряда.

2. На каждой итерации получаем $N-1$ значений $X_{N_{CO}, n}$, которые используем в следующей формуле:

$$R = \max (X_{N_{CO}, N}) - \min (X_{N_{CO}, N}),$$

где R —размах отклонения X .

3. Определяется эмпирический риск идентификации тренда временного ряда I_3 (3).

4. Нормируется размах R делением на эмпирический риск идентификации тренда временного ряда I_3 , который вычисляется по N_{CO} значениям.

5. Логарифмируя $R/\sqrt{I_3}$ и N_{CO} строят на основании полученных данных, например, график зависимости $\log(R/\sqrt{I_3})$ от $\log(N_{CO})$.

6. По графику функции $\log(R/\sqrt{I_3})$ от $\log(N_{CO})$ находят наклон путём линейной аппроксимации (эластичность задачи расчета показателя Хёрста), тангенс угла этого наклона определяют показатель Хёрста H_X и совместно с величиной эмпирического риска I_3 по табл. 1 идентифицируют закон распределения случайной величины.

На базе выражения 6 для эластичности пункт 5 и 6 методики могут быть модифицированы к аналитической форме.

7. Сдвигают период N_{CO} по всему временному ряду, при этом N_{CO} меняется уже не от единицы, а от начала периода N_{CO} , который не совпадает с началом временного ряда, т. е. формируется “окно скользящего контроля” вдоль всего временного ряда. В результате получают несколько значений показателя $R/\sqrt{I_3}$.

Проведенные исследования способа идентификации случайного процесса позволяют повысить качество решения задачи путем модификации п. II.1 второго этапа идентификации, сводящейся к тому, что по выражению (7) для абсолютного числа обусловленности решаемой задачи определяют число наблюдений N_{CO} “окна скользящего контроля” временного ряда, обеспечивающее допустимое значение погрешности определения показателя Хёрста от погрешности входных данных.

В результате применения предложенной методики осуществляется идентификация как тренда временного ряда, так и закона распределения случайной величины, определяющей модель исследуемого процесса.

Выводы

1. Минимальные априорные предположения об исследуемом процессе и достаточная устойчивость при анализе временных рядов определяет широкое использование при идентификации случайного процесса эмпирического метода Хёрста ((R/S)–метода), который ориентирован на выявление и оценку нестационарных компонент случайного процесса.

2. При идентификации процесса исследуемого параметра x_t апостериорная информация об областях неопределенности оценки \hat{x}_t исследуемого параметра x_t по выборке наблюдений Y_0^t определяется апостериорным риском оценки \hat{x}_t , который зависит от дисперсии исследуемого параметра x_t и максимума функции правдоподобия соответствующего закона распределения.

3. Проведенный информационный анализ метода нормированного размаха показал, что показатель Хёрста H_X соответствует максимальной энтропии исследуемого процесса, т. е. энтропии Хартли, оцененную в единицах a_N/N .

4. Исследование свойств показателя Хёрста H_X показало, что он равен эластичности функции $R/\sqrt{I_s}$ по переменной N , а величина показателя Хёрста H_X определяет точность его расчета.

5. Выявленная в результате информационного подхода информационная природа показателя Хёрста H_X позволило предложить методику общего анализа исследуемых процессов для идентификации, как нестационарных случайных процессов, так и при фрактальном представлении исследуемых процессов.

Литература

1. Meyer R., Christensen N. Bayesian reconstruction of chaotic dynamical systems // Phys. Rev. E. 2000. V. 62. pp. 3535–3542.

2. Колыхан Н. В., Тюрязев В.С. Информационные технологии статистической обработки данных выборок ограниченного объема //



Инженерный вестник Дона, 2007, № 2. URL: ivdon.ru/magazine/archive/n2y2007/21/

3. Farmer J. D., Sidorowich J. J. Predicting chaotic time series//Phys. Rev. Lett. 1987. V. 59. pp. 845–848.

4. Аль-Хулайди А. А. Разработка нового стохастического метода управления очередями заданий с использованием Марковских процессов для параллельных вычислений на кластере//Инженерный вестник Дона, 2011, №1 URL: ivdon.ru/magazine/archive/n1y2011/332/.

5. Д. В. Степанов, А. В. Кузовников Исследование фрактальных свойств радиосигналов // Вестник Сибирского государственного аэрокосмического университета имени академика М. Ф. Решетнева. С. 35–39

6. А. Г. Масловская, Т. Р. Осокина, Т. К. Барабаш Применение фрактальных методов для анализа динамических данных // Вестник Амурского государственного университета. 2010. Вып. 51: Сер. Естеств. и экон. науки. С. 13–20

7. Кроновер Р. М. Фракталы и хаос в динамических системах. Основы теории. М.: Постмаркет. 2000. 353 с.

8. Н. Ф. Бахарева, И. В. Карташевский, В. Н. Тарасов Анализ и расчет непуассоновских моделей трафика в сетях ЭВМ // Инфокоммуникационные технологии. 2009. Том 7, № 4. С. 61–66

9. Новицкий П. В., Зограф И. А. Оценка погрешностей результатов измерений. Л.: Энергоатомиздат. 1991. 304 с.

10. Бахрушин В.Е. Проблемы идентификации моделей распределения случайных величин с применением современного программного обеспечения //Успехи современного естествознания. 2011. № 11. С. 50–54

11 Лемешко Б. Ю. О задаче идентификации закона распределения случайной составляющей погрешности измерений // Метрология. 2004. № 7. С. 8–17.

12. Губарев В. В. Вероятностные модели: Справочник в 2-х частях. Новосибирск.: НЭТИ. 1992. 422 с.

13. В. Бутаков, А. Граковский Оценка уровня стохастичности временных рядов произвольного происхождения при помощи показателя Хёрста// Computer Modelling and New Technologies, 2005, V.9, Riga, Latvia, №2, pp. 27–32.

14. Е. С. Филатова, Д. М. Филатов, А. Д. Стоцкая Анализ временного ряда электропотребления методом нормированного размаха // Современные проблемы науки и образования. 2014. № 5; URL: science-education.ru/119-15063

15. Вапник В. Н. Восстановление зависимостей по эмпирическим данным. М.: Наука. 1979. 448 с.

16. Левшина Е. С., Новицкий П. В. Электрические измерения физических величин. Измерительные преобразователи. Л.: Энергоатомиздат. 1983. 320 с.

17. Вадзинский Р.Н. Справочник по вероятностным распределениям. СПб.: Наука. 2001. 295 с.

18. Третьяк Л. Н. Обработка результатов наблюдений: Учебное пособие. Оренбург: ГОУ ОГУ. 2004. 171 с.

19. Малугин В. А. Математика для экономистов: Математический анализ. Курс лекций. М.: Эксмо. 2005. 272 с.

20. Горбаченко В. И. Вычислительная линейная алгебра с примерами на MATLAB. СПб.: БХВ–Петербург. 2011. 320 с.

References

1. Meyer R., Christensen N. Bayesian reconstruction of chaotic dynamical systems//Phys. Rev. E. 2000. V. 62. pp. 3535–3542.

2. Kolyhan N. V., Tjurjaev V.S. Inženernyj vestnik Dona (Rus), 2007, № 2. URL: ivdon.ru/magazine/archive/n2y2007/21/

3. Farmer J. D., Sidorowich J. J. Predicting chaotic time series//Phys. Rev. Lett. 1987. V. 59. pp. 845–848.
4. Al'-Hulajdi A.A. Inženernyj vestnik Dona (Rus), 2011, №1 URL:ivdon.ru/magazine/archive/n1y2011/332/.
5. Stepanov D. V., Kuzovnikov A. V. Vestnik Sibirskogo gosudarstvennogo ajerokosmicheskogo universiteta imeni akademika M. F. Reshetneva. pp. 35–39.
6. Maslovskaja A. G., Osokina T. R., Barabash T. K. Vestnik Amurskogo gosudarstvennogo universiteta. 2010. Vyp. 51: Ser. Estestv. i jekon. nauki. pp. 13–20
7. Kronover R. M. Fraktaly i haos v dinamicheskikh sistemah. Osnovy teorii. [Fractals and chaos in dynamical systems. Fundamentals of the theory]. M.: Post-market. 2000. 353 p.
8. Bahareva N. F., Kartashevskij I. V., Tarasov V. N. Infokommunikacionnye tehnologii. 2009. Tom 7, № 4. pp. 61–66
9. Novickij P. V., Zograf I. A. Ocenka pogreshnostej rezul'tatov izmerenij. [Assessment of uncertainties of measurement results]. L.: Jenergoatomizdat. 1991. 304 p.
10. Bahrushin V.E. Uspehi sovremennogo estestvoznaniya. 2011. № 11. pp. 50–54
11. Lemeshko B. Ju. Metrologija. 2004. № 7. pp. 8–17.
12. Gubarev V. V. Veroyatnostnye modeli: Spravochnik v 2-h chastjah. [Probabilistic models: The Handbook in 2 parts]. Novosibirsk: NJeTI. 1992. 422 p.
13. Butakov V., Grakovskij A. Computer Modelling and New Technologies, 2005, V.9, Riga, Latvia, №2, pp. 27–32.
14. Filatova E. S., Filatov D. M., Stockaja A. D. Sovremennye problemy nauki i obrazovanija. 2014. № 5; URL: science-education.ru/119-15063
15. Vapnik V. N. Vosstanovlenie zavisimostej po jempiricheskim dannym. [The restoration of dependences based on empirical data]. M.: Nauka. 1979. 448 p.



16. Levshina E. S., Novickij P. V. Jelektricheskie izmerenija fizicheskikh velichin. Izmeritel'nye preobrazovateli. [Electrical measurements of physical quantities. Measuring transducers]. L.: Jenergoatomizdat. 1983. 320 p.

17. Vadzinskij R.N. Spravochnik po verojatnostnym raspredelenijam. [Handbook of probability distributions]. SPb.: Nauka. 2001. 295 p.

18. Tret'jak L. N. Obrabotka rezul'tatov nabljudenij: Uchebnoe posobie. [Processing of results of observations: a tutorial]. Orenburg: GOU OGU. 2004. 171 p.

19. Malugin V. A. Matematika dlja jekonomistov: Matematicheskij analiz. Kurs lekcij. [Mathematics for economists: A Mathematical analysis. A course of lectures]. M.: Jeksmo. 2005. 272 p.

20. Gorbachenko V. I. Vychislitel'naja linejnaja algebra s primerami na MATLAB. [Computational linear algebra with examples in MATLAB]. SPb.: BHV–Peterburg. 2011. 320 p.