УДК 519.2: 530.1: 600.1

И.И. ГОРБАНЬ

КРИТЕРИИ И ПАРАМЕТРЫ СТАТИСТИЧЕСКОЙ НЕУСТОЙЧИВОСТИ

Анотація. *Розширено поняття статистичної нестійкості процесу (послідовності). Введено визначення статистично стійкого процесу у вузькому та широкому сенсі. Розвинута методика оцінки статистичної нестійкості на обмеженому інтервалі спостереження. Запропоновано додаткові параметри статистичної нестійкості, що характеризують порушення статистичної стійкості по відношенню до вибіркового середньоквадратичного відхилення.*
Ключові слова: *статистична стійкість, параметр статистичної нестійкості, теорія гіпервипадкових явищ.*

Аннотация. *Расширено понятие статистической неустойчивости процесса (последовательности). Введены определения статистически устойчивого процесса в узком и широком смыслах. Развита методика оценки статистической неустойчивости на ограниченном интервале наблюдения. Предложены дополнительные параметры статистической неустойчивости, характеризующие нарушение устойчивости по отношению к выборочному среднеквадратическому отклонению.*
Ключевые слова: *статистическая устойчивость, параметр статистической неустойчивости, теория гиперслучайных явлений.*

Abstract. *Statistical instability conception of the process (of the sequence) has been extended. Definitions of process's statistical stability in the narrow and wide sense have been introduced. The methodology of statistical instability estimation for limited observation interval has been developed. Additional parameters of statistical instability characterizing the disturbance of statistical stability to standard deviation have been proposed.*
Keywords: *statistical stability, parameter of statistical instability, theory of hyper-random phenomena.*

1. Введение

Одним из удивительных физических явлений окружающего мира является феномен статистической устойчивости частоты массовых событий. Долгое время полагали, что при неограниченном увеличении объема выборки частота любого массового события стремится к пределу – определенной вероятности.

Однако экспериментальные исследования на больших интервалах наблюдения процессов различной физической природы показали [1], что феномен статистической устойчивости частоты носит не идеальный характер: вначале с увеличением объема выборки дисперсия флуктуации выборочного среднего, как правило, уменьшается, но затем, достигнув определенного значения, практически перестает меняться, а в некоторых случаях начинает расти.

Таким образом, вероятность, оказывается, не более как математическая абстракция, удобная для описания реальных массовых явлений окружающего мира при относительно небольших интервалах наблюдения, однако не имеющая смысла при больших интервалах.

Для описания реальных физических событий, величин, процессов и полей с учетом ограниченного характера их статистической устойчивости была разработана физико-математическая теория гиперслучайных явлений [1, 2].

В рамках этой теории формализовано понятие статистической устойчивости процесса. К статистически устойчивым отнесены процессы, у которых выборочные дисперсии среднего стремятся к нулю при неограниченном увеличении объема выборки, а к неустойчивым – процессы, не удовлетворяющие этому требованию.

Для оценки статистической устойчивости реальных процессов была разработана специальная методика, основанная на расчете параметров статистической неустойчивости, характеризующих степень нарушения устойчивости на ограниченном интервале наблюдения.

С ее помощью на больших интервалах наблюдения были зафиксированы нарушения статистической устойчивости целого ряда физических процессов: колебаний напряжения городской сети, магнитного поля Земли, курса валют, высоты и периода морских волн, температуры воздуха, температуры воды в океане, скорости ветра и других процессов [1, 3–10].

На основе результатов проведенных экспериментальных исследований была выдвинута гипотеза [1, 2], что нарушение статистической устойчивости присуще всем реальным физическим явлениям. Исключение могут составлять, возможно, лишь фундаментальные физические постоянные, такие как скорость света, гравитационная постоянная и пр., рассматриваемые современной наукой как мировые константы.

Дальнейшие исследования показали, что в природе, хотя и редко, но встречаются физические явления, параметры статистической неустойчивости которых практически неотличимы от параметров идеально статистически устойчивых процессов. Таковыми оказались, например, колебания количества осадков на протяжении столетия [5] и колебания интенсивности излучения некоторых пульсаров на протяжении полутора десятков лет [10].

Естественно, возникает вопрос: можно ли считать эти физические процессы статистически устойчивыми? В соответствии с принятым в [1] критерием устойчивости, они формально устойчивы, однако не исключено, что при другой формализации понятия статистической устойчивости ответ на поставленный вопрос может оказаться иным.

Целью настоящей статьи является расширение понятия статистической неустойчивости процесса и развитие методики оценки статистической неустойчивости на ограниченном интервале наблюдения.

2. Известные критерии и параметры статистической неустойчивости

Определение 1а. По определению [1] последовательность $X_1, X_2, \ldots$ случайных величин (случайная выборка) считается статистически устойчивой (статистически стабильной), если при устремлении объема выборки N к бесконечности математическое ожидание выборочной дисперсии

$$\bar{D}_{Y_N} = \frac{1}{N-1}\sum_{n=1}^{N}(Y_n - \bar{m}_{Y_N})^2 \qquad (1)$$

флуктуации выборочного среднего

$$Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad (n = \overline{1, N}) \qquad (2)$$

стремится к нулю, где $\bar{m}_{Y_N} = \frac{1}{N}\sum_{n=1}^{N} Y_n$ – выборочное среднее флуктуации среднего. Последовательности, не удовлетворяющие этому условию, статистически неустойчивы.

Уменьшение выборочного среднего по мере увеличения объема данных может быть вызвано не только стабилизацией среднего, но также уменьшением дисперсии исходного процесса. Для подавления влияния изменения дисперсии процесса в работе [10] предложено следующее определение понятия статистической устойчивости процесса.

Определение 1б. Последовательность случайных величин $X_1, X_2, ...$ названа статистически устойчивой, если при устремлении объема выборки N к бесконечности параметр статистической неустойчивости

$$\gamma_N = \frac{\mathrm{M}\left[\bar{D}_{Y_N}\right]}{ND_{y_N}} = \frac{\mathrm{M}\left[\bar{D}_{Y_N}\right]}{\bar{D}_{x_N}} \tag{3}$$

стремится к нулю, где D_{y_N} – дисперсия выборочного среднего:

$$D_{y_N} = \frac{1}{N^2}\sum_{n=1}^{N} D_{x_n} ,$$

$\mathrm{M}[\cdot]$ – оператор математического ожидания, $D_{x_n} = \mathrm{M}[(X_n - m_{x_n})^2]$ – дисперсия случайной величины X_n, а $m_{x_n} = \mathrm{M}[X_n]$ – ее математическое ожидание, $\bar{D}_{x_N}$ – средняя дисперсия случайных величин X_n $(n = \overline{1, N})$: $\bar{D}_{x_N} = \frac{1}{N}\sum_{n=1}^{N} D_{x_n}$.

При выполнении условий Определения 1а или Определения 1б последовательность естественно называть статистически устойчивой по отношению к среднему, а при невыполнении этих условий – статистически неустойчивой по отношению к среднему.

Процесс обычно описывается последовательностью значений в фиксированные моменты времени, поэтому, говоря о процессах, можно отождествлять их с соответствующими последовательностями.

Заметим, что в частном случае постоянной дисперсии понятия статистически устойчивого процесса, согласно Определениям 1а и 1б, совпадают.

По результатам наблюдения процесса на конечном интервале невозможно точно установить факт нарушения статистической устойчивости. Однако можно количественно оценить степень флуктуации выборочного среднего в фиксированные моменты времени и, анализируя динамику происходящих изменений, выявить некоторые тенденции, ведущие к нарушению статистической устойчивости.

Такие тенденции характеризуют параметр статистической неустойчивости γ_N и параметр μ_N, связанный с параметром γ_N, соотношением

$$\mu_N = \sqrt{\gamma_N / (1 + \gamma_N)} . \tag{4}$$

В отличие от параметра γ_N, ограниченного лишь снизу нулевым значением, параметр μ_N ограничен как снизу, так и сверху: минимально возможное его значение равно нулю, а максимально возможное – единице.

В качестве единицы измерения статистической неустойчивости параметра γ_N при фиксированном N может выступать величина γ_{0N} [8], представляющая собой параметр γ_N, рассчитанный для идеальной статистически устойчивой последовательности N некоррелированных отсчетов с постоянной дисперсией $D_{x_n} = D_x$ и нулевым математическим ожиданием. Поэтому параметром статистической неустойчивости является также параметр

$$h_N = \gamma_N / \gamma_{0N}$$

[8], диапазон изменения которого – $[0, \infty)$.

Для практических расчетов можно использовать оценки γ_N^*, μ_N^*, h_N^* параметров γ_N, μ_N, h_N.

Оценка γ_N^* может быть вычислена по множеству реализаций или как $\gamma_N^* = \dfrac{\bar{D}_{Y_N}}{\bar{D}_{X_N}}$, где $\bar{D}_{X_N} = \dfrac{1}{N}\sum_{n=1}^{N} D_{X_n}^*$ – среднее оценок дисперсии $D_{X_n}^*$, сформированное по отдельным фрагментам реализации последовательности X_n ($n = \overline{1,N}$). Оценки же μ_N^* и h_N^* могут быть рассчитаны при наличии оценки γ_N^* по формулам $\mu_N^* = \sqrt{\gamma_N^* / (1+\gamma_N^*)}$, $h_N^* = \gamma_N^* / \gamma_{0N}$.

Факт нарушения устойчивости процесса устанавливается по тенденции изменения параметров γ_N, μ_N, h_N (или γ_N^*, μ_N^*, h_N^*). Если при больших значениях N наблюдается рост этих параметров или стабилизация значения на ненулевом уровне, процесс считается неустойчивым, в противном случае – устойчивым.

Известная методика оценки статистической неустойчивости [1] основана на расчете этих параметров.

3. Альтернативные критерии и параметры статистической неустойчивости

Рассмотренные критерии и параметры статистической неустойчивости процесса отслеживают динамику изменения важнейшей величины процесса – выборочного среднего. Однако при этом остаются без внимания функция распределения и другие выборочные моменты.

Стремление к нулю математического ожидания выборочной дисперсии среднего процесса (или какого-либо связанного с ним параметра) не гарантирует сходимость выборочной функции распределения к какой-либо функции распределении. Поэтому Определения 1а и 1б нельзя признать безукоризненными.

Принимая во внимание сказанное, можно предложить несколько альтернативных вариантов определения понятия статистически устойчивой последовательности (процесса).

Рассмотрим последовательность случайных величин $X_1, X_2, \ldots$ (с различными в общем случае законами распределения). Пусть $N(x)$ – количество членов последовательности $\{X_N\} = X_1, X_2, \ldots, X_N$ объема N, меньших x, а $F_N^*(x) = \dfrac{N(x)}{N}$ – эмпирическая (выборочная) функция распределения, представляющая собой неубывающую ступенчатую функцию.

При неограниченном увеличении числа членов N эмпирическая функция распределения $F_N^*(x)$ не обязательно сходится к определенной функции.

Определение 1в. Последовательность случайных величин $X_1, X_2, \ldots$ назовем статистически устойчивой в узком смысле почти наверное (с вероятностью единица), если существует случайная величина X, описываемая функцией распределения $F(x)$, к которой сходится почти наверное (с вероятностью единица) эмпирическая функция распределения $F_N^*(x)$ при неограниченном увеличении N: $\lim\limits_{N\to\infty} F_N^*(x) = F(x)$, т.е.

$$P\left\{\lim_{N\to\infty} \sup_{-\infty<x<\infty} \left|F_N^*(x) - F(x)\right| = 0\right\} = 1. \qquad (5)$$

Если такая случайная величина X не существует, последовательность будем называть статистически неустойчивой в узком смысле.

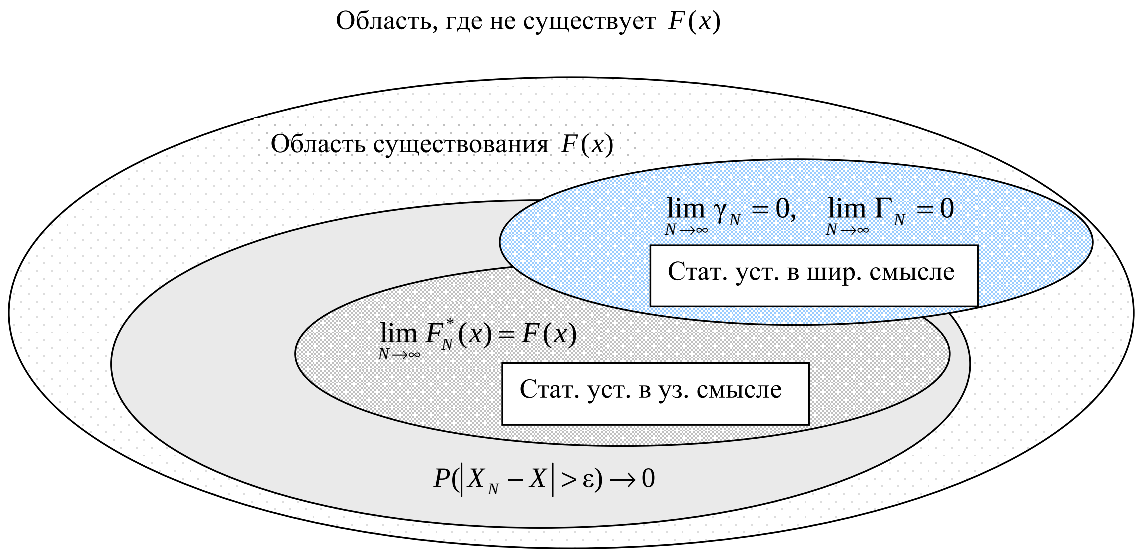

Рис. 1. Статистически устойчивые и неустойчивые случайные последовательности (процессы) в узком и широком смыслах

Поскольку сходимость с вероятностью единица более сильная, чем сходимость по вероятности, то в случае, когда $F_N^*(x)$ сходится к $F(x)$ почти наверное, имеет место сходимость $F_N^*(x)$ к $F(x)$ и по вероятности (рис. 1).

Определение 1в может быть полезным при проведении теоретических исследований, однако оно мало пригодно для оценки нарушений статистической устойчивости реальных процессов, поскольку оценка сходимости эмпирической функции распределения представляет собой практически непреодолимую проблему.

Задачу можно существенно упростить, если ограничиться оценкой математических ожиданий выборочных дисперсий двух первых выборочных моментов: среднего и дисперсии (или среднеквадратического отклонения (СКО)).

Определение 1г. Последовательность $X_1, X_2, ...$ случайных величин назовем статистически устойчивой в широком смысле, если при устремлении объема выборки N к бесконечности

1) математическое ожидание выборочной дисперсии (1) флуктуации выборочного среднего (2) и

2) математическое ожидание выборочной дисперсии

$$\bar{D}_{Z_N} = \frac{1}{N-2}\sum_{n=2}^{N}(Z_n - \bar{m}_{Z_N})^2 \qquad (6)$$

выборочного СКО

$$Z_n = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - Y_n)^2} \quad (n = \overline{2,N}) \qquad (7)$$

стремятся к нулю, где $\bar{m}_{Z_N} = \frac{1}{N-1}\sum_{n=2}^{N} Z_n$ – среднее флуктуации выборочного СКО.

Последовательности, не удовлетворяющие этому условию, будем называть статистически неустойчивыми в широком смысле.

Таким образом, статистически устойчивый процесс в широком смысле – процесс, статистически устойчивый по отношению к среднему и к СКО.

Словосочетания «в узком смысле» и «широком смысле» используются здесь и далее по аналогии с тем, как они применяются применительно к понятию стационарного процесса [11].

Обратим внимание, что области статистической устойчивости (как в широком, так и в узком смысле) располагаются внутри области существования функции распределения $F(x)$, а области неустойчивости охватывают не только части области, в которой существует функция распределения $F(x)$, но и область, в которой она не существует (рис. 1).

Не всякий процесс, устойчивый в узком смысле, является устойчивым в широком смысле, и, наоборот, не всякий процесс, устойчивый в широком смысле, является устойчивым в узком смысле.

Для подавления влияния изменения дисперсии исследуемого процесса можно предложить еще один, более конструктивный, вариант определения понятия статистической устойчивости в широком смысле.

Определение 1д. Последовательность случайных величин $X_1, X_2, ...$ назовем статистически устойчивой в широком смысле, если при устремлении объема выборки N к бесконечности стремятся к нулю

1) параметр статистической неустойчивости по отношению к среднему (3) и
2) параметр статистической неустойчивости по отношению к СКО:

$$\Gamma_N = \frac{\mathrm{M}\left[\bar{D}_{Z_N}\right]}{ND_{y_N}} = \frac{\mathrm{M}\left[\bar{D}_{Z_N}\right]}{\bar{D}_{x_N}}. \tag{8}$$

Последовательности, не удовлетворяющие этому условию, будем называть статистически неустойчивыми в широком смысле (рис. 1).

Наряду с параметром Γ_N можно использовать параметр $\mathrm{M}_N = \sqrt{\Gamma_N / (1+\Gamma_N)}$, аналогичный параметру μ_N, и параметр $H_N = \Gamma_N / \gamma_{0N}$, аналогичный параметру h_N.

Параметры Γ_N, M_N и H_N – безразмерные величины. Параметры Γ_N и H_N ограничены снизу нулевым значением, а параметр M_N – нулевым значением снизу и единичным значением сверху.

Для практических расчетов на конечном интервале наблюдения вместо параметров Γ_N, M_N и H_N целесообразно использовать их оценки: $\Gamma_N^* = \dfrac{\bar{D}_{Z_N}}{\bar{D}_{X_N}}$, $\mathrm{M}_N^* = \sqrt{\Gamma_N^* / (1+\Gamma_N^*)}$ и $H_N^* = \Gamma_N^* / \gamma_{0N}$.

Таким образом, методика оценки статистической неустойчивости физических процессов, основанная на расчете параметров статистической неустойчивости по отношению к выборочному среднему γ_N^*, μ_N^*, h_N^*, дополняется расчетами параметров статистической неустойчивости по отношению к выборочному СКО Γ_N^*, M_N^*, H_N^*.

Обратим внимание на одно важное обстоятельство. При расчете оценок параметров статистической неустойчивости γ_N^*, μ_N^*, h_N^* и Γ_N^*, M_N^*, H_N^* на основе выборочных дисперсий не используется тот факт, что исследуемые величины $X_1, X_2, ...$ являются случайными (т.е. характеризуются вероятностной мерой). Поэтому понятие статистической устойчивости в широком смысле можно обобщить на любые последовательности, в том числе и неслучайные.

Определение 1е. Последовательность величин $X_1, X_2, \ldots$ назовем статистически устойчивой в широком смысле, если оценки параметров статистической неустойчивости по отношению к среднему $\gamma_N^* = \frac{\bar{D}_{Y_N}}{\bar{D}_{X_N}}$ и по отношению к СКО $\Gamma_N^* = \frac{\bar{D}_{Z_N}}{\bar{D}_{X_N}}$ при неограниченном увеличении объема выборки N стремятся к нулю: $\lim_{N\to\infty} \gamma_N^* = 0$, $\lim_{N\to\infty} \Gamma_N^* = 0$.

Последовательности, не обладающие этим свойством, будем считать статистически неустойчивыми в широком смысле.

4. Исследование статистической устойчивости реальных процессов

Для иллюстрации информативности дополнительных параметров оценки статистической устойчивости Γ_N, M_N и H_N на рис. 2 и 3 приведены результаты расчетов оценок этих параметров для двух астрофизических аккрецирующих источников рентгеновского излучения: GRS 1915+105 (рис. 2) и PSRJ 1012+5307 (рис. 3).

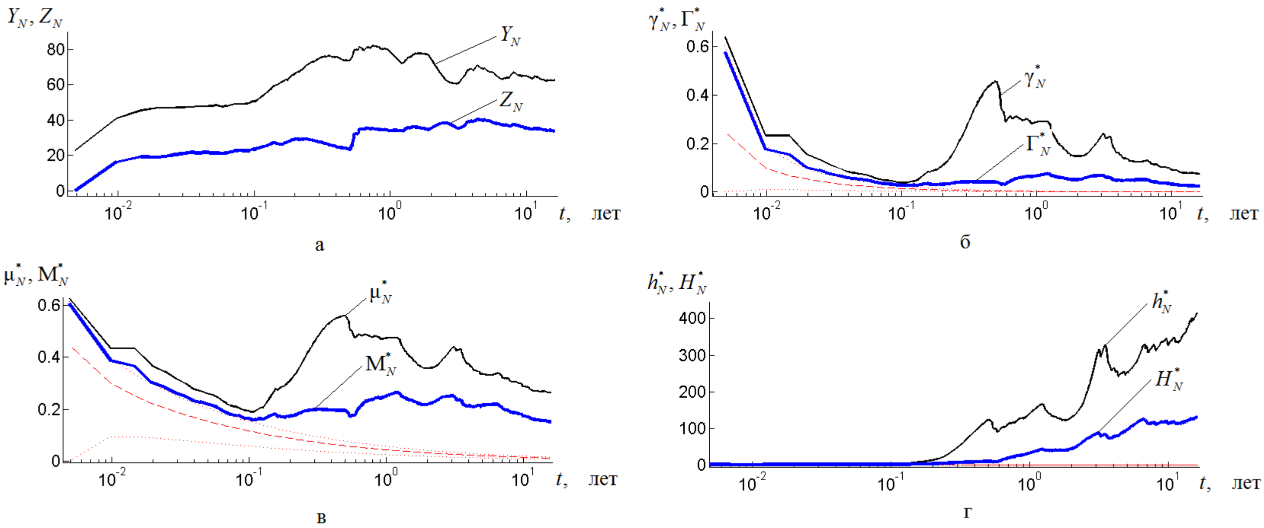

Рис. 2. Выборочное среднее Y_n и выборочное среднеквадратическое отклонение Z_n колебания интенсивности излучения источника GRS 1915+105 (а) и соответствующие параметры статистической неустойчивости γ_N^*, Γ_N^* (б), μ_N^*, M_N^* (в) и h_N^*, H_N^* (г)

Как и в статье [10], исходные данные взяты с сайта [12]. Измерения интенсивности излучения осуществлялись в период с 1 января 1996 г. по 31 декабря 2011 г. Средняя периодичность измерений составляла 2,7 ч для GRS 1915+105 и 2,8 ч для PSRJ 1012+5307.

На рис. 2а и 3а сплошными линиями изображены колебания во времени выборочных средних Y_n и выборочных СКО Z_N, на рис. 2б, 3б – зависимости от времени параметров γ_N^*, Γ_N^*, на рис. 2в, 3в – зависимости от времени параметров μ_N^*, M_N^*, на рис. 2г, 3г – зависимости от времени параметров h_N^*, H_N^*.

Пунктирными линиями на рис. 2б, 3б, а также 2в, 3в и 2г, 3г представлены соответственно параметры γ_N, μ_N и h_N для идеального устойчивого процесса, а точечными линиями – СКО от них.

Для источника GRS 1915+105 (рис. 2) сильные нарушения статистической устойчивости по отношению к среднему сопровождаются значительными нарушениями устойчивости по отношению к СКО. Степень нарушения устойчивости по отношению к среднему выше, чем по отношению к СКО.

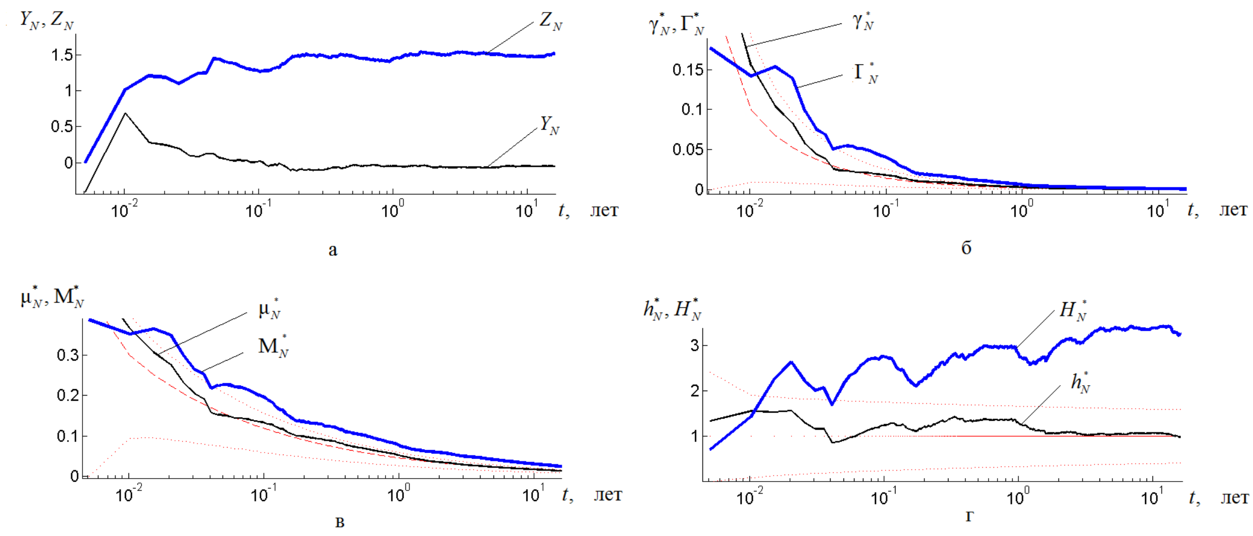

Рис. 3. Выборочное среднее Y_n и выборочное среднеквадратическое отклонение Z_n колебания интенсивности излучения пульсара PSR J1012+5307 (а) и соответствующие параметры статистической неустойчивости γ^*_N, Γ^*_N (б), μ^*_N, M^*_N (в) и h^*_N, H^*_N (г)

Для источника PSRJ 1012+5307 ситуация иная (рис. 3): нарушения статистической устойчивости по отношению к среднему не наблюдаются, однако имеют место нарушения устойчивости по отношению к СКО.

Отсюда следует, что колебания обоих рассматриваемых источников излучения статистически неустойчивы в широком смысле.

Аналогичные исследования флуктуации количества осадков в нескольких регионах выявили нарушения статистической устойчивости по отношению к СКО.

Таким образом, претендентами на роль статистически устойчивых процессов в широком смысле остаются пока лишь фундаментальные физические постоянные (мировые константы).

5. Выводы

1. Расширено понятие статистической неустойчивости процесса (последовательности). Введены определения статистически устойчивого процесса в узком и широком смыслах.
2. Развита методика оценки статистической неустойчивости на ограниченном интервале наблюдения. Предложены дополнительные параметры статистической неустойчивости, характеризующие нарушение устойчивости по отношению к выборочному среднеквадратическому отклонению.
3. На примере астрофизических источников излучения продемонстрирована информативность предложенных дополнительных параметров статистической неустойчивости.
4. Установлено, что претендентами на роль статистически устойчивых процессов в широком смысле остаются пока лишь фундаментальные физические постоянные (мировые константы).

СПИСОК ЛИТЕРАТУРЫ

1. Горбань И.И. Теория гиперслучайных явлений: физические и математические основы [Электронный ресурс] / Горбань И.И. – К.: Наукова думка, 2011. – 318 с. – Режим доступа: http://www.immsp.kiev.ua/perspages/ gorban_i_i/index.html.

2. Горбань И.И. Теория гиперслучайных явлений [Электронный ресурс] / Горбань И.И. – К.: ИПММС, 2007. – 181 с. – Режим доступа: http://www.immsp.kiev.ua/perspages/ gorban_i_i/index.html.

3. Горбань И.И. Нарушение статистической устойчивости физических процессов / И.И. Горбань // Математичні машини і системи. – 2010. – № 1. – С. 171 – 184.

4. Gorban I.I. Disturbance of statistical stability / I.I. Gorban // Information Models of Knowledge. – Sofia: ITHEA, 2010. – P. 398 – 410.

5. Горбань И.И. Статистическая устойчивость колебаний температуры воздуха и осадков в районе Москвы / И.И. Горбань // Математичні машини і системи. – 2011. – № 3. – С. 97 – 104.

6. Исследование статистической устойчивости колебаний температуры шельфовой зоны окраинных морей / И.И. Горбань, Н.И. Горбань, В.В. Новотрясов [и др.] // Седьмой Всероссийский симпозиум «Физика геосфер». – Владивосток, 2011. – С. 542 – 547.

7. Gorban I.I. Disturbance of statistical stability (part II) / I.I. Gorban // International Journal “Information Theories and Applications”. – 2010. – Vol. 18, N 4. – P. 321 – 334.

8. Горбань И.И. Статистическая неустойчивость физических процессов / И.И. Горбань // Известия вузов. Радиоэлектроника. – 2011. – Т. 54, № 9. – С. 40 – 52.

9. Горбань И.И. Исследование нарушений статистической устойчивости колебаний скорости ветра в Чернобыле / И.И. Горбань, А.Д. Скорбун // Системи підтримки прийняття рішень. Теорія і практика: зб. доп. наук.-прак. конф. з міжнар. участю. «СППР’2012». – К., 2012. – С. 39 – 42.

10. Горбань И.И. Статистическая устойчивость излучения астрофизических объектов / И.И. Горбань // Математичні машини і системи. – 2012. – № 2. – С. 155 – 160.

11. Горбань І.І. Теорія ймовірностей і математична статистика для наукових працівників та інженерів [Електронний ресурс] / І.І. Горбань. – К.: ІПММС НАН України, 2003. – 245 с. – Режим доступу: http://www.immsp.kiev.ua/perspages/gorban_i_i/index.html.

12. All-Sky Monitor (ASM) team at the Kavli Institute for Astrophysics and Space Research at the Massachusetts Institute of Technology [Електронний ресурс]. – Режим доступу: http://xte.mit.edu/ASM_lc.html.

Стаття надійшла до редакції 27.07.2012