УДК 519.2+600.1

И.И. ГОРБАНЬ

ОЦЕНКИ ХАРАКТЕРИСТИК ГИПЕРСЛУЧАЙНЫХ ВЕЛИЧИН

Abstract: *The work continues the cycle of the research in special non-determine phenomena that are hyper random ones. Such phenomena are similar to random ones, however the probable measure absents for them. In the article, the mathematical convergence's conception for hyper-random values has been proposed, methodology for forming estimations of theirs characteristics has been developed, and the convergence of the estimates has been researched.*
Key words: *hyper random values, non-determine phenomena.*

Анотація: *Робота продовжує цикл досліджень особливого класу недетермінованих явищ – гіпервипадкових величин. Ці явища подібні до випадкових, але для них не існує визначеної імовірнісної міри. У статті введено математичні поняття збіжності гіпервипадкових величин, запропонована методологія формування оцінок їх характеристик і досліджена збіжність таких оцінок.*
Ключові слова: *гіпервипадкові величини, недетермінові явища.*

Аннотация: *Работа продолжает цикл исследований особого класса недетерминированных явлений – гиперслучайных величин. Эти явления похожи на случайные, но для них не определена вероятностная мера. В статье введены математические понятия сходимости гиперслучайных величин, предложена методология формирования оценок их характеристик и исследована сходимость таких оценок.*
Ключевые слова: *гиперслучайные величины, недетерминированные явления.*

1. Введение

При исследовании и моделировании сложных информационных систем нередко приходится сталкиваться с неопределенными (недетерминированными) величинами, параметрами, характеристиками и функциями. Под этими словосочетаниями обычно понимают явления, о которых известно очень мало или вообще ничего не известно. Неопределенность, как правило, учитывают, вводя предположение о случайном характере явления. Задаваясь законом распределения, можно описать явление известными методами теории вероятностей и математической статистики. Однако далеко не всегда можно указать закон распределения. Иногда по причине незнания его типа, но чаще по причине отсутствия определенного закона. Недетерминированное явление, для которого не существует конкретный закон распределения, т.е. не определена вероятностная мера [1 – 3], не является случайным. Такое явление относится к совершенно другому классу – классу гиперслучайных явлений [4, 5]. Исследованию различных типов гиперслучайных явлений (событий, величин, функций) и их описанию посвящен цикл работ [4 – 8].

В статьях [4, 5] предложено математическое определение понятия гиперслучайных явлений и разработаны принципы их представления. Применительно к двум классам гиперслучайных явлений (гиперслучайным событиям и гиперслучайным величинам) разработан специальный математический аппарат. Для характеристики гиперслучайной величины X предложено использовать полумеры, представляющие собой границы функций распределения, под которыми подразумеваются верхняя и нижняя границы условной вероятности $P\{X \le x/g\}$ выполнения неравенства $X \le x$ в условиях $g \in G$:

$$F_S(x) = \sup_{g \in G} P\{X \le x/g\}, \quad F_I(x) = \inf_{g \in G} P\{X \le x/g\},$$

где G – множество условий, при которых наблюдается гиперслучайная величина X .

Исследованы свойства этих полумер и предложен ряд дополнительных характеристик и параметров, характеризующих гиперслучайные величины. Основными параметрами являются моменты границ, в частности, математические ожидания границ, дисперсии границ, коэффициенты корреляции и ковариации границ и др.

Полученные результаты обобщены в работе [6] на случай гиперслучайных функций. В статье [7] рассмотрен другой способ описания гиперслучайных явлений – с помощью границ моментов, в частности, границ математического ожидания, границ дисперсии, границ корреляционных и ковариационных моментов и др. В работе [8] формализованы понятия стационарности и эргодичности гиперслучайных процессов.

До настоящего времени остаются малоизученными статистические аспекты гиперслучайных явлений.

Целью настоящей статьи является разработка методологии формирования оценок характеристик гиперслучайных величин и исследование их свойств.

2. Выборка гиперслучайной величины

Гиперслучайную величину X в общем случае можно представить множеством случайных величин X/g, наблюдаемых при условии $g \in G$: $X = \{X/g \in G\}$. В частном случае, когда X/g представляет собой детерминированную величину, однозначно связанную с условием $g \in G$, гиперслучайная величина $X = \{X/g \in G\}$ вырождается в множество детерминированных величин.

Генеральной совокупностью гиперслучайной величины $X = \{X/g \in G\}$ будем называть бесконечное множество всех ее реализаций (членов или элементов), наблюдаемых при всех условиях $g \in G$. Это множество может быть как счетным, так и не счетным. Генеральную совокупность можно описать с помощью верхней и нижней границ функции распределения, моментов границ, границ моментов и других характеристик.

Конечное множество членов генеральной совокупности $\vec{x} = (x_1, ..., x_N) = \{x_1, ..., x_N / g \in G\}$ гиперслучайной величины X, полученное при конечном числе N опытов в каждых условиях $g \in G$, будем называть выборкой из генеральной совокупности, а ее элементы $x_1, ..., x_N$ – выборочными значениями или реализациями. Каждая реализация x_n/g $(n = \overline{1, N})$ вектора $\vec{x}/g$ гиперслучайной выборки $\vec{X}$ в условиях $g \in G$ представляет собой детерминированную величину, а каждая реализация x_n вектора $\vec{x}$ без конкретизации условий – множество детерминированных величин, связанных с определенными условиями $g \in G$.

Будем полагать, что выборка $x_1, ..., x_N$ принадлежит гиперслучайной величине X с границами функции распределения $F_S(x)$, $F_I(x)$, если она получена из генеральной совокупности, описываемой границами функции распределения $F_S(x)$, $F_I(x)$.

Бесконечное множество выборок объема N, сформированных из одной генеральной совокупности, представляет собой N-мерный гиперслучайный вектор $\vec{X} = (X_1,...,X_N) = \left\{ \vec{X} / g \in G \right\}$, называемый в дальнейшем гиперслучайной выборкой или выборочной совокупностью. Компоненты X_n / g $(n = \overline{1,N})$ этого вектора в условиях $g \in G$ представляют собой случайные величины, а каждая компонента X_n вектора $\vec{X}$ без конкретизации условий – гиперслучайную величину. При этом все компоненты X_n $(n = \overline{1,N})$ имеют один и тот же закон распределения, совпадающий с законом распределения генеральной совокупности.

Компоненты X_n гиперслучайной выборки $\vec{X}$ будем полагать взаимно независимыми в том смысле, как это понимается [4 – 5] в теории гиперслучайных явлений. Это означает, что границы $F_S(\vec{x})$, $F_I(\vec{x})$ распределения гиперслучайной выборки $\vec{X}$ допускают представление в виде $F_S(\vec{x}) = \prod_{n=1}^{N} F_S(x_n)$, $F_I(\vec{x}) = \prod_{n=1}^{N} F_I(x_n)$.

Статистикой будем называть произвольную функцию $Y = Y(\vec{X})$ выборки $\vec{X}$, вариационным (статистическим) рядом в условиях $g \in G$ – реализации выборки $\vec{x} / g$, упорядоченные по возрастанию или убыванию, а ранжированным рядом в условиях $g \in G$ – реализации выборки $\vec{x} / g$, упорядоченные по убыванию.

По генеральной совокупности гиперслучайной величины можно вычислить различные ее характеристики, например, границы функции распределения $F_S(x)$, $F_I(x)$, математические ожидания границ m_{Sx}, m_{Ix}, границы математического ожидания m_{sx}, m_{ix}, дисперсии границ D_{Sx}, D_{Ix}, границы дисперсии D_{sx}, D_{ix} и пр. По реализациям с использованием определенных статистик можно вычислить оценки этих же характеристик, в частности, оценки границ функции распределения $F_S^*(x)$, $F_I^*(x)$, оценки математических ожиданий границ m_{Sx}^*, m_{Ix}^*, оценки границ математического ожидания m_{sx}^*, m_{ix}^*, оценки дисперсии границ D_{Sx}^*, D_{Ix}^*, оценки границ дисперсии D_{sx}^*, D_{ix}^* и др.

3. Оценки характеристик гиперслучайной величины

Границы функции распределения $F_S(x)$, $F_I(x)$, моменты m_{Sx}, m_{Ix}, m_{sx}, m_{ix}, D_{Sx}, D_{Ix}, D_{sx}, D_{ix} и пр. являются детерминированными характеристиками. Соответствующие же оценки $F_S^*(x)$, $F_I^*(x)$, m_{Sx}^*, m_{Ix}^*, m_{sx}^*, m_{ix}^*, D_{Sx}^*, D_{Ix}^*, D_{sx}^*, D_{ix}^* и пр. являются детерминированными, если получены по конкретной реализации гиперслучайной выборки $\vec{X}$, и случайными, если получены по генеральной совокупности гиперслучайной величины X.

Процедура формирования оценки может строиться по следующей схеме. Для всего множества G условий g формируется выборка $\vec{x} = \{x_1, ..., x_N / g \in G\}$. По этой выборке для каждого условия g в отдельности рассчитывается оценка условных характеристик: оценка условной функции распределения $F^*(x/g)$, оценка условного математического ожидания $m^*_{x/g}$, оценка условной дисперсии $D^*_{x/g}$ или др. По оценкам условной функции распределения $F^*(x/g)$ вычисляются оценки границ функции распределения $F^*_S(x) = \sup_{g \in G} F^*(x/g)$, $F^*_I(x) = \inf_{g \in G} F^*(x/g)$ и оценки характеристик, характеризующие эти границы: математические ожидания границ m^*_{Sx}, m^*_{Ix}, оценки дисперсий границ D^*_{Sx}, D^*_{Ix} и пр. По оценкам условных величин определяются оценки границ соответствующих величин. Например, по оценкам математических ожиданий $m^*_{x/g}$ – оценки границ математического ожидания $m^*_{sx} = \sup_{g \in G} m^*_{x/g}$, $m^*_{ix} = \inf_{g \in G} m^*_{x/g}$, по оценкам условных дисперсий $D^*_{x/g}$ – оценки границ дисперсии $D^*_{sx} = \sup_{g \in G} D^*_{x/g}$, $D^*_{ix} = \inf_{g \in G} D^*_{x/g}$ и т.д.

Определенные трудности можно ожидать при формировании выборки $\vec{x} = \{x_1, ..., x_N / g \in G\}$ из-за сложности обеспечения, контроля и поддержания условий $g \in G$. Однако вопрос облегчается тем, что для расчета искомых характеристик не требуются знания того, в каких именно условиях получены условные характеристики. Главное, чтобы были представлены на уровне условных характеристик все возможные условия g множества G и в массив данных, используемый для расчета этих характеристик, не попадали данные, соответствующие другим условиям. Обычно последнее требование можно обеспечить ограничением объема данных N, поскольку условия, хотя и меняются зачастую непрерывно, но меняются достаточно медленно, и поэтому на основе некоторой априорной информации возможно указать максимальное число последовательных элементов $N_{\max}$, для которых условия можно считать практически неизменными.

Это позволяет собирать данные на достаточно большом интервале наблюдения, не заботясь о том, каковы в конкретный момент времени условия и в какой последовательности они чередуются. Полученные данные далее можно разделять на фрагменты по $N_{\max}$ последовательных элементов в каждом и использовать для расчета промежуточных условных и затем конечных искомых оценок.

4. Сходимость гиперслучайных величин

Для дальнейшего изложения необходимо ввести понятие сходимости последовательности гиперслучайных величин.

Пусть имеется последовательность гиперслучайных величин $\mathrm{X}=\{X_1,...,X_N\}$ и гиперслучайная величина X. Для всех $X_1,\ldots,X_N$ и X определены множество условий G и условные функции распределения $F_1(x/g),\ldots,F_N(x/g)$, $F(x/g)$ $(g\in G)$. Тогда последовательность X:

1) сходится к X по функции распределения, если при всех условиях $g\in G$ в каждой точке x, где $F(x/g)$ непрерывны, $F_N(x/g)\to F(x/g)$ при $N\to\infty$;

2) сходится к X в среднеквадратическом, если при всех условиях $g\in G$ условные математические ожидания $\mathrm{M}[|X_N-X|^2/g]$ стремятся к нулю при $N\to\infty$, т. е при фиксированных условиях g случайная последовательность $\mathrm{X/g}$ сходится в среднеквадратическом к случайной величине X/g. При такой сходимости можно писать $\lim\limits_{N\to\infty} X_N = X$;

3) сходится к X почти наверное, если при всех условиях $g\in G$ условная вероятность $P(X_N\to X/g)$ равна единице при $N\to\infty$, т. е. при фиксированных условиях g случайная последовательность $\mathrm{X/g}$ сходится с вероятностью единица к случайной величине X/g. При такой сходимости можно писать $\lim\limits_{N\to\infty} X_N = X$;

4) сходится к X по вероятности, если $P(|X_N-X|>\varepsilon/\mathrm{g})$ стремится к нулю при всех условиях $g\in G$, $\varepsilon>0$ и $N\to\infty$.

Отметим, что понятия сходимости последовательности гиперслучайных величин допускают обобщения на последовательность гиперслучайных функций.

Пусть имеется последовательность гиперслучайных функций $\mathrm{X}(t)=\{X_1(t),...,X_N(t)\}$ и гиперслучайная функция $X(t)$ $(t\in T)$, для которых определены условные функции распределения $F_1(x;t/g),...,F_N(x;t/g)$, $F(x;t/g)$. Тогда последовательность $\mathrm{X}(t)$:

1) сходится к $X(t)$ в среднеквадратическом, если для всех $t\in T$ и всех $g\in G$ $\mathrm{M}[|X_N(t)-X(t)|^2/g]\to 0$ при $N\to\infty$, т.е. $\lim\limits_{N\to\infty} X_N = X\,(t)$;

2) сходится к $X(t)$ почти наверное, если для всех $t\in T$ и всех $g\in G$ $P(X_N(t)\to X(t)/g)=1$ при $N\to\infty$, т. е. $\lim\limits_{N\to\infty} X_N(t)=X(t)$.

Сходимость последовательности гиперслучайных функций по функции распределения и по вероятности можно определить аналогично сходимости последовательности гиперслучайных величин.

5. Сходимость оценок

Важным свойством ряда гиперслучайных оценок является то, что при увеличении объема выборки эти оценки сходятся по вероятности к соответствующим характеристикам. Нетрудно убедиться, что справедливо следующее положение.

Положение. Пусть $X_1,...,X_N$ – выборка гиперслучайной величины X объемом N, $\theta^*_{xN/g}$ – сформированная по выборке в условиях g случайная оценка, обладающая свойством сходимости по вероятности к параметру $\theta_{x/g}$. Тогда гиперслучайная оценка $\theta^*_{xN} = \{\theta^*_{xN/g}, \quad g \in G\}$ сходится по вероятности к множеству детерминированных величин $\theta_x = \{\theta_{x/g}, \quad g \in G\}$, а границы оценки $\theta^*_{sxN} = \sup\limits_{g \in G} \theta^*_{xN/g}$, $\theta^*_{ixN} = \inf\limits_{g \in G} \theta^*_{xN/g}$ – к соответствующим границам $\theta_{sx} = \sup\limits_{g \in G} \theta_{x/g}$, $\theta_{ix} = \inf\limits_{g \in G} \theta_{x/g}$.

Отсюда следует, что при увеличении объема выборки ($N \to \infty$) оценки границ выборочной функции распределения $F^*_{SN}(x)$, $F^*_{IN}(x)$ сходятся к границам функции распределения $F_S(x)$, $F_I(x)$, оценки моментов границ – к моментам границ, а оценки границ моментов – к границам моментов. В частности, оценки математического ожидания границ m^*_{Sx}, m^*_{Ix} сходятся к математическому ожиданию границ m_{Sx}, m_{Ix}, оценки дисперсии границ D^*_{Sx}, D^*_{Ix} – к дисперсии границ D_{Sx}, D_{Ix}, оценки границ математического ожидания m^*_{sx}, m^*_{ix} – к границам математического ожидания m_{sx}, m_{ix}, а оценки границ дисперсии D^*_{sx}, D^*_{ix} – к границам дисперсии D_{sx}, D_{ix}.

Назовем функцию распределения случайной величины фрагментарно-гауссовской, если она описывается фрагментами гауссовских функций распределений. Фрагментарно-гауссовское распределение аналитически можно представить в виде

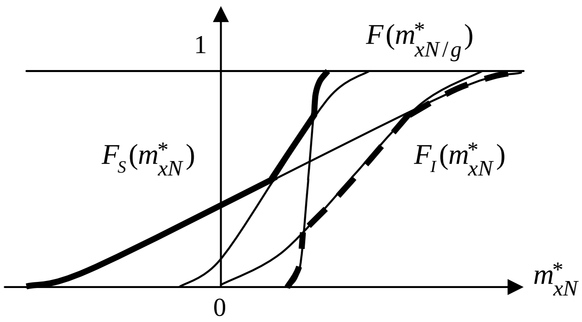
$$G(x) = \sum_{i=0}^{I} F(x/m_i, D_i)\mathrm{rect}\left[\frac{x - x_i}{x_{i+1} - x_i}\right],$$

где $F(x/m_i, D_i) = \Phi\left(\frac{x - m_i}{\sqrt{D_i}}\right)$ – гауссовская функция распределения с математическим ожиданием m_i и дисперсией D_i, $\mathrm{rect}[x]$ – П-образная функция:

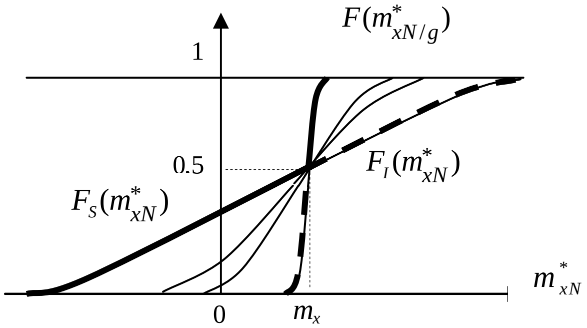
$$\mathrm{rect}[x] = \begin{cases} 0 & \text{при} \quad x < 0, \quad x > 1, \\ 1 & \text{при} \quad 0 \le x \le 1, \end{cases}$$

$x_0, x_1, \ldots, x_I$ – последовательный ряд точек на оси x, в которых происходит изменение вида распределения, $x_0 \to -\infty$, $x_I \to \infty$.

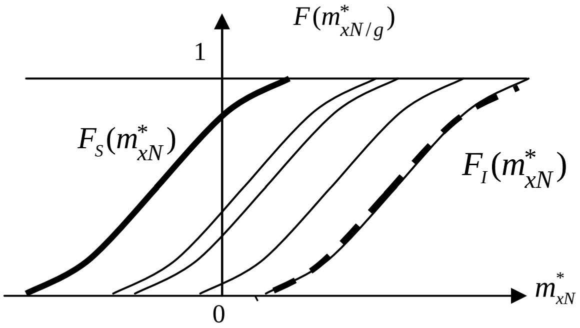
Для гиперслучайных величин справедлива теорема, аналогичная центральной предельной теореме теории вероятностей.



а)



б)



в)

Рис. 1. Веер функций распределения $F(m^*_{xN/g}) = \Phi\left(\dfrac{(m^*_{xN/g} - m_{x/g})\sqrt{N}}{\sqrt{D_{x/g}}}\right)$ случайных величин $m^*_{xN/g}$, $g = \mathrm{varia}$ (тонкие кривые), верхняя $F_S(m^*_{xN})$ (жирная сплошная кривая) и нижняя $F_I(m^*_{xN})$ (жирная штриховая кривая) границы функции распределения гиперслучайной величины m^*_{xN}, $N = \mathrm{const}$: *а)* – общий случай, *б)* – при постоянном математическом ожидании ($m_{x/g} = m_x$), *в)* – при постоянной дисперсии ($D_{x/g} = D_x$)

Теорема 1. Пусть гиперслучайная величина $X = \{X / g \in G\}$ имеет конечные условные математические ожидания $m_{x/g}$ и конечные условные дисперсии $D_{x/g}$. Проводятся независимые испытания по N испытаний в каждых условиях. Тогда при $N \to \infty$ оценки верхней и нижней границ функции распределения среднего $m_{xN}^{*}=\left\{\frac{1}{N}\sum_{n=1}^{N}X_{n}/g\in G\right\}$ стремятся по вероятности к фрагментарно-гауссовским распределениям.

Доказательство теоремы основано на центральной предельной теореме для случайных величин (теореме Леви – Лиденберга). В соответствии с этой теоремой функция распределения оценки случайной величины $m_{xN/g}^{*}$ при $N\to\infty$ стремится по вероятности к гауссовскому распределению с математическим ожиданием $m_{x/g}$ и дисперсией $\frac{D_{x/g}}{N}$: $\lim\limits_{N\to\infty}P\left\{\frac{(m_{xN/g}^{*}-m_{x/g})\sqrt{N}}{\sqrt{D_{x/g}}}\le x\right\}=F(x/0,1)\quad\forall g\in G.$ Границы функции распределения гиперслучайной величины формируются из фрагментов функций распределения образующих случайных величин $m_{xN/g}^{*}$ (рис. 1 *а*). Отсюда следует справедливость утверждения теоремы 1.

Отметим, что границы функции распределения $F_{S}(m_{x}^{*})$, $F_{I}(m_{x}^{*})$ среднего m_{xN}^{*} при $N\to\infty$ и конечном количестве условий G нетрудно рассчитать по известным функциям распределения $F(m_{x/g}^{*})$ условных средних $m_{xN/g}^{*}$ ($N\to\infty$), принимая во внимание, что любые две кривые, описывающие гауссовские функции распределения, пересекаются не более чем в одной точке.

Для практики могут представлять особый интерес два частных случая, рассматриваемые ниже в виде двух теорем.

Теорема 2. Пусть выполняются условия теоремы 1, и при всех условиях $g\in G$ условные математические ожидания одинаковые ($m_{x/g}=m_{x}$), а условные дисперсии $D_{x/g}$ – разные. Тогда при $N\to\infty$ оценка верхней границы функции распределения среднего m_{xN}^{*} стремится по вероятности к фрагментарно-гауссовскому распределению, описываемому функцией

$$F_{S}(x)=\lim_{N\to\infty}\begin{cases}F\left(x/m_{x},\dfrac{D_{sx}}{N}\right), & \text{если}\quad x\le m_{x},\\ F\left(x/m_{x},\dfrac{D_{ix}}{N}\right), & \text{если}\quad x>m_{x},\end{cases}$$

а оценка нижней границы – к фрагментарно-гауссовскому распределению, описываемому функцией

$$F_{I}(x)=\lim_{N\to\infty}\begin{cases}F\left(x/m_{x},\dfrac{D_{ix}}{N}\right), & \text{если}\quad x\le m_{x},\\ F\left(x/m_{x},\dfrac{D_{sx}}{N}\right), & \text{если}\quad x>m_{x},\end{cases}$$

где D_{sx}, D_{ix} – соответственно верхняя и нижняя границы дисперсии гиперслучайной величины $X=\{X/g\in G\}$.

Доказательство теоремы основано на теореме 1. При выполнении условий теоремы 2 функции распределения случайных величин m_{xN}^* / g образуют веер гауссовских функций распределений, пересекающихся в точке $x = m_x$ (рис. 1 *б*). Отсюда следует справедливость теоремы 2.

Теорема 3. Пусть выполняются условия теоремы 1, и при всех условиях $g \in G$ условные математические ожидания $m_{x/g}$ – разные, а условные дисперсии – одинаковые ($D_{x/g} = D_x$), m_{sx}, m_{ix} – соответственно верхняя и нижняя границы математического ожидания гиперслучайной величины $X = \{X / g \in G\}$. Тогда при $N \to \infty$ оценка верхней границы функции распределения среднего m_{xN}^* стремится по вероятности к гауссовскому закону распределения с параметрами $\left(m_{ix}, \frac{D_x}{N}\right)$, а оценка нижней границы функции распределения – к гауссовскому закону распределения с параметрами $\left(m_{sx}, \frac{D_x}{N}\right)$, т.е.

$$\lim_{N\to\infty} \sup_{g\in G} P\left\{\frac{(m_{xN/g}^* - m_{ix})\sqrt{N}}{\sqrt{D_x}} \le x\right\} = F(x/0,1), \quad \lim_{N\to\infty} \inf_{g\in G} P\left\{\frac{(m_{xN/g}^* - m_{sx})\sqrt{N}}{\sqrt{D_x}} \le x\right\} = F(x/0,1).$$

Доказательство теоремы основано на теореме 1. При выполнении условий теоремы 3 функции распределения случайных величин $m_{xN/g}^*$ образуют веер непересекающихся гауссовских функций распределений (рис. 1, *в*). Отсюда следует справедливость теоремы 3.

6. Заключение

В результате рассмотрения оценок характеристик гиперслучайных величин формализовано понятие гиперслучайной выборки и определены ее свойства, введены математические понятия сходимости гиперслучайных величин, предложена методология формирования оценок характеристик гиперслучайной величины и исследована сходимость гиперслучайных оценок к соответствующим точным характеристикам. Доказаны для гиперслучайных оценок теоремы устанавливающие факты их сходимости по вероятности к точным значениям.

СПИСОК ЛИТЕРАТУРЫ

1. Колмогоров А.М. Теория вероятностей и математическая статистика. – М.: Наука, 1986. – 535 с.
2. Королюк В.С. и др. Справочник по теории вероятностей и математическая статистика. – М.: Наука, 1985. – 637 с.
3. Горбань І.І. Теорія ймовірностей і математична статистика для наукових працівників і інженерів. – К.: ІПММС НАН України, 2003. – 244 с.
4. Горбань И.И. Случайность, гиперслучайность, хаос и неопределенность // Стандартизація, сертифікація, якість. – 2005. – № 3. – С. 41 – 48.
5. Горбань И.И. Гиперслучайные явления и их описание // Акустичний вісник. – 2005. – Т. 8, № 1 – 2. – С. 16 – 27.
6. Горбань И.И. Гиперслучайные функции и их описание // Радиоэлектроника. – 2006. – № 1. – С. 3 – 15.
7. Горбань И.И. Методы описания гиперслучайных величин и функций // Акустичний вісник. – 2005 (в печати).
8. Горбань И.И. Стационарные и эргодические гиперслучайные функции // Радиоэлектроника. – 2005 (в печати).