УДК 519.216

# АНАЛИЗ СВОЙСТВ ВЕРОЯТНОСТНЫХ МОМЕНТОВ ДВОИЧНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ДЛЯ ЗАДАЧ МОДЕЛИРОВАНИЯ СТОХАСТИЧЕСКИХ ОБЪЕКТОВ*

***В.М. Кузнецов, В.А. Песошин, А.И. Гумиров, Д.В. Ширшова***

Казанский национальный исследовательский технический университет им. А.Н. Туполева
Россия, 420111, г. Казань, ул. К. Маркса, 10

E-mail: kuznet_evm@mail.ru, pesoshin-kai@mail.ru, neporebrik@mail.ru, einstein_darya@mail.ru

***Аннотация.*** *Обоснована применимость статистического метода определения значимой однородности двоичных последовательностей в соответствии с вероятностными моментами любого конечного порядка для решения задач математического моделирования стохастических объектов. Описаны подходы к построению критерия однородности в условиях разнообразных способов задания вероятностных моментов. Выявлены существенные отличия разложений центральных вероятностных моментов на начальные для систем случайных событий однородного характера и применительно к последовательностям. Представлены краткие тезисы методик нахождения критической длины выборок, в пределах которых двоичные последовательности обладают значимой однородностью по заданным начальным или центральным вероятностным моментам. Приведен иллюстрационный пример сокращения затрат вычислительных ресурсов при реализации алгоритмов имитации вероятностных свойств внешней среды для моделируемого объекта. Отмечены области применения концепции однородности случайных последовательностей в теоретических постановках задач машинной реализации математических моделей и практических разработках средств адекватного сравнения статистических характеристик стохастических объектов.*

***Ключевые слова:*** *имитационное моделирование, вероятностный момент, начальный момент, центральный момент, значимая однородность, критерий однородности, двоичная последовательность.*

## Введение

Анализ задач машинного моделирования реальных объектов и процессов стохастического типа способен выявить ряд требований к случайным последовательностям, обеспечивающим проявление внешней среды [1, 2]. Кроме основных статистически оцениваемых вероятностных свойств [3], достоверно известными становятся объемы требуемых выборок, временные рамки использования средств

**Результаты исследования получены при поддержке гранта РФФИ и АН РТ № 18-47-160001.*

*Кузнецов Валерий Михайлович (д.т.н., доцент), профессор кафедры «Компьютерные системы».*

*Песошин Валерий Андреевич (д.т.н., профессор), профессор кафедры «Компьютерные системы».*

*Гумиров Артем Ильдарович, старший преподаватель кафедры «Компьютерные системы».*

*Ширшова Дарья Вадимовна, старший преподаватель кафедры «Компьютерные системы».*

имитации, а также данные о видах и порядках вероятностных моментов.

Методы статистических испытаний (методы Монте-Карло) в современных постановках требуют формирования многочисленных выборок случайных последовательностей разной длины [4, 5]. Уникальность необходимых вероятностных и корреляционных свойств вынуждает разработчиков имитационных экспериментов на ЭВМ подбирать адекватные этим алгоритмам программы или разрабатывать новые.

Простейшие варианты выполнения требований постановщиков задачи реализуются разработчиками машинного эксперимента выбором минимального набора последовательностей на основе *схемы независимых испытаний* Бернулли. Технически это достигается реализацией детерминированных алгоритмов получения псевдослучайных чисел [6, 7] или аппаратного формирования истинно случайных последовательностей [7–10], которые воспроизводят известные постулаты Голомба [10, 11], приближающие искусственный характер генерируемых последовательностей к идеальной модели случайного сигнала типа «белого шума» [7, 8].

Строгую математическую форму описания бернуллиевских свойств двоичных последовательностей дали А.Н. Колмогоров и В.А. Успенский [12]. Они определили в терминах теоретико-множественного представления и алгоритмической вычислимости три свойства случайности: типичности, хаотичности и стохастичности. Понимая под генеральной выборкой последовательность бесконечной длины, авторы алгоритмической теории случайности отнесли эти три свойства к *частной выборке* в виде цепочки конечной длины как фрагмента бесконечной последовательности.

Современные тенденции усложнения имитационного моделирования требуют задания многопараметрических атрибутов последовательностей на генеральных выборках, существенно отличающихся от схемы Бернулли вероятностными моментами высоких порядков. Такие программно-алгоритмические инструменты обеспечивают адекватность моделирования внешних возмущений в пределах оценок заданных моментов на конечных длинах частных выборок. Сами же вероятностные моменты, характеризуя исключительность генеральных выборок, делают практически невозможной параллельную реализацию уникальных алгоритмов формирования псевдослучайных отсчетов в условиях фон-неймановской архитектуры ЭВМ.

Разветвление и параллелизация алгоритмов генераторов псевдослучайных последовательностей за счет организации многоядерной, векторной, кластерной, мультипроцессорной обработки, безусловно, сокращает временные издержки имитационного моделирования, но это происходит ценой затрат аппаратных ресурсов весьма дорогих вычислителей. Возникает вопрос: нельзя ли использовать фактор конечности выборок имитирующих последовательностей для упрощения алгоритмов их формирования и за счет этого объединения нескольких разных последовательностей в одну?

Действительно, чем короче выборка, тем шире дисперсионный разброс статистических оценок вероятностных параметров разных случайных последовательностей, среди которых возникает возможность установления факта неразличимости двух и более выборок с длинами, не превышающими некоторой критической величины, с заданной уверенностью.

**Пример экономии ресурсов**

Рассмотрим пример воображаемой ревизии последовательностей в целях снижения ресурсной нагрузки на машинный эксперимент, связанный с реализацией имитационной модели. Обобщенная схема эксперимента представлена на рис. 1.

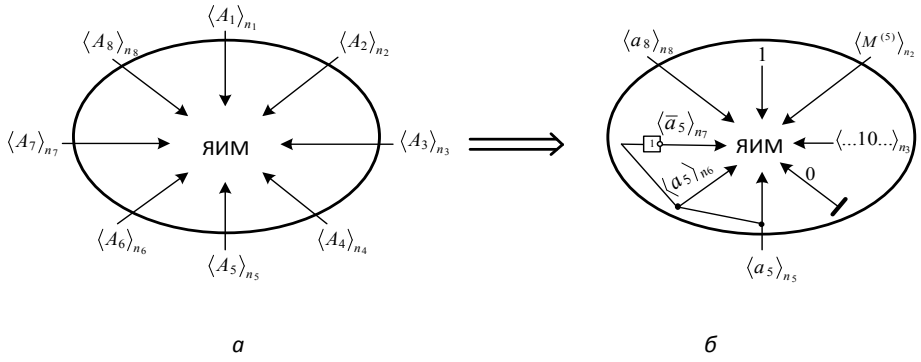



Рис. 1. Обобщенная структура имитационной модели статистического типа:
*а* – исходная форма; *б* – результат ревизии входных последовательностей

Выделим в структуре общей постановки машинного эксперимента логико-алгоритмические блоки и связи, ядро имитационной модели (ЯИМ), находящееся под воздействием некоторой области внешних возмущений и внутренних вариаций свойств ядра как динамического объекта [2]. В случае статистического характера модели эти возмущения должны проявлять случайное поведение, адекватное моделируемым реальным процессам [1].

Пусть такая структура требует совокупности из 8 фрагментов двоичных последовательностей $\langle a_1\rangle_{n_1}$, $\langle a_2\rangle_{n_2}$, …, $\langle a_8\rangle_{n_8}$, являющейся реализацией набора случайных последовательностей $\langle A_1\rangle_{n_1}$, $\langle A_2\rangle_{n_2}$, …, $\langle A_8\rangle_{n_8}$ размерностью $n_1$, $n_2$, …, $n_8$ соответственно, с необходимыми оригинальными теоретико-вероятностными свойствами. Для их формирования надо создать алгоритмы и программы, реализация которых связана с необходимыми затратами машинной памяти и времени.

Предположим, что имеется инструмент попарного сравнения имитирующих последовательностей на предмет неразличимости их по вероятностным свойствам в пределах указанных длин выборок между собой. При проведении такого сравнения, с добавлением также в качестве альтернативных имеющихся в арсенале исследователя эффективных низкозатратных программных средств, включая примитивные источники типа констант и коротких циклов, было предположительно выяснено следующее:

- $\langle a_1\rangle$ и $\langle a_4\rangle$ на длинах $n_1$ и $n_4$ статистически неразличимы с константами 1 и 0 соответственно;
- $\langle a_2\rangle$ на длине $n_2$ статистически не отличается от М-последовательности 5-го порядка $\langle M^{(5)}\rangle$;
- $\langle a_3\rangle$ на длине $n_3$ статистически не отличается от элементарной последовательности типа …0101…;

– $\langle a_5 \rangle$ на длине $n_6$ и $n_7$ статистически неразличима с $\langle a_6 \rangle$ и $\langle \overline{a}_7 \rangle$ соответственно;
– $\langle a_8 \rangle$ на длине $n_8$ статистически оригинальна (отлична) от констант 1 и 0, последовательности типа …0101…, $\langle M^{(5)} \rangle$, $\langle a_5 \rangle$ и $\langle \overline{a}_5 \rangle$.

Таким образом, для машинной реализации математической модели достаточно сформировать две оригинальные последовательности: $\langle a_5 \rangle$, совпадающую с $\langle a_6 \rangle$, $\langle \overline{a}_7 \rangle$, на длине $\max\{n_5, n_6, n_7\}$ и $\langle a_8 \rangle$ длиной $n_8$, а также три низкозатратных в ресурсном отношении типа $\langle a_3 \rangle$ = …0101…, на протяжении $n_3$ тактов, вырожденных в константы $\langle a_1 \rangle$ = 1 и $\langle a_4 \rangle$ = 0 в течение $n_1$ и $n_4$ тактов модельного времени (см. рис. 1 *б*).

Приведенный пример демонстрирует актуальность создания инструмента тестирования на статистическую неразличимость случайных последовательностей.

**Статистически значимая однородность последовательностей**

Ставится задача определения однородности двух последовательностей вероятностно-статистической природы $\langle a \rangle$ и $\langle b \rangle$ по выбранным квалифицирующим параметрам. Такими параметрами могут быть, например, математические ожидания базовой и альтернативной последовательностей (БП и АП). Для двоичных последовательностей они совпадают с вероятностями появления единицы соответственно, $P_a$ и $P_b$. Покажем универсальность вероятности элементарных, а также сложных событий как квалифицирующего параметра при тестировании на однородность по моментам не только первого, но и более высокого конечного порядка.

Основная независимая переменная $n$ – длина частной выборки. Имитационная модель воспринимает от БП вероятность $P_a$ в форме статистических оценок на момент реализации $n$-го модельного такта, т. е. $P_a \to P_a^*(n)$. То же самое для АП: $P_b \to P_b^*(n)$. Предусмотрим ограничения вида $\tau_{\max}$ при исследовании на однородность автокорреляционных свойств, через моментные функции второго порядка, и предел $n_{\max}$ для $n$. Из условий парадигмы, принятой в машинном моделировании конкретного класса стохастических объектов, задается типичный уровень *значимости критерия однородности* $\alpha$ как вероятность ошибки первого рода.

Используемый в непараметрических критериях подход [13–15] основывается на подсчетах эмпирической статистики, прямо пропорциональной средним значениям расхождений квалифицирующих параметров и обратно пропорциональной величине дисперсионного разброса оценок расхождений. Завершающей процедурой критерия является проверка гипотезы об однородности путем сравнения величины полученной статистики с критическим уровнем, учитывающим заданную значимость.

В данной задаче предусматриваются многократные испытания гипотез на основе статистики в форме отношения двух функций от *n*-ожидаемого различия оценок вероятностей $\Delta P^* = P_a^*(n) - P_b^*(n)$ и оценки стандартного отклонения этой разности, т. е.

$$t_{\text{эмп}}(n) = M^*_{\Delta P^*}(n) \Big/ \sqrt{D^*_{\Delta P^*}(n)}. \qquad (1)$$

Эмпирический материал для этой статистики обеспечивается формированием достаточного множества выборок элементов разностной последовательности вида

$$\langle d\rangle_n = \langle a\rangle_n - \langle b\rangle_n \tag{2}$$

длиной $n$ при ее возрастания от 1 до $n_{\max}$. Полученные числовые значения (1) сравниваются с некоторой критической величиной $t_{\text{кр}}(\alpha)$, дающей основание принять нуль-гипотезу $\text{H}_0$ об однородности тестируемых объектов или, отвергнув ее, выбрать конкурирующую гипотезу $\text{H}_1$ согласно условиям:

$$\begin{cases} |t_{\text{эмп}}| < t_{\text{кр}}, \text{H}_0, \ \max\lfloor n(|t_{\text{эмп}}| < t_{\text{кр}})\rfloor = n_{\text{кр}}, \\ |t_{\text{эмп}}| \geq t_{\text{кр}}, \text{H}_1, \ n = \overline{1, n_{\max}}. \end{cases} \tag{3}$$

Содержательным результатом тестирования является: «Обе последовательности на длине выборки $n_{\text{кр}}$ (или не менее $n_{\max}$) статистически однородны (относятся к одной генеральной совокупности) со степенью значимости $\alpha$». При таком сравнительном исследовании не требуется определение самих вероятностных моментов и скрытых в них многосвязных условий расположения элементов тестируемых последовательностей на временной оси. Достигается лишь неразличимость, эквивалентность, взаимозаменяемость, подобность последовательностей в определенном смысле относительно заданных вероятностных моментов как по форме, так и по величине порядка.

Рассмотрим применимость классических вероятностных моментов конечных порядков, уделив особое внимание двоично структурированным формам последовательностей.

**Начальные вероятностные моменты**

Принято [14–16] начальный вероятностный момент порядка $r$ дискретной случайной величины $A$ представлять в виде

$$\nu_r(A) = \sum_{i=1}^{u} a_i^r p_i, \tag{4}$$

где $u$ – количество уровней дискретности; $p_i$ – вероятность принятия величиной $A$ уровня $a_i$; порядок $r$ определен на множестве натуральных чисел.

Минимальное значение $u = 2$ соответствует бинарному характеру величины $A$. Уменьшив на единицу оба предела суммирования в (4) и определив бинарность $A$ алфавитом $a_i \in \{0, 1\}$ для $i = \overline{0, 1}$, представим следующее распределение вероятностей: $\begin{array}{c|c|c} A & a_0 = 0 & a_1 = 1 \\ \hline p_i & p_0 & p_1 \end{array}$, где $p_1 = \mathbf{P}\{A = 1\}$.

Если полагать $0^r = 0$ и $1^r = 1$, то справедливо следующее *утверждение*: начальный момент двоичной случайной величины $A$ любого сколь угодно высокого конечного порядка $r$ в виде натурального числа равен начальному моменту этой величины первого порядка, допускающий выражение в форме математического ожидания и вероятности появления единицы вида:

$$\nu_r(A) = \nu_1(A) = \mathbf{M}[A] \text{ и } \mathbf{M}[A] = \mathbf{P}\{A = 1\} = p_1.$$

Система $k$ случайных дискретных величин $A_1, A_2, ..., A_k$ многоуровневого типа ($u > 2$) характеризуется смешанным начальным моментом

$$\nu_r(A_1, A_2, ..., A_k) = \sum_{i_1=0}^{u-1} \sum_{i_2=0}^{u-1} \cdots \sum_{i_k=0}^{u-1} a_{1i_1}^{r_1} a_{2i_2}^{r_2} \cdots a_{ki_k}^{r_k} p_{i_1 i_2 \ldots i_k}, \quad (5)$$

где $r = \sum_{j=1}^{k} r_j$ – порядок момента как сумма порядков $r_j$ элементов $A_j$ системы.

Для системы случайных двоичных величин (при $u = 2$) очевидна фиксация элементарных составляющих $r_j = 1$ для всех $j = \overline{1, k}$, что обеспечивает равенство размерности системы ее порядку, т. е. $r = k$. Тогда смешанный начальный момент двоичной системы выразится на основании формы (5) в следующем виде:

$$\nu_r(A_1, A_2, ..., A_r) = \sum_{i_1=0}^{1} \sum_{i_2=0}^{1} \cdots \sum_{i_r=0}^{1} a_{1i_1} a_{2i_2} \cdots a_{ri_r} p_{i_1 i_2 \ldots i_r}$$

или через математическое ожидание и вероятность появления совокупности из $r$ единиц:

$$\nu_r(A_1, A_2, ..., A_r) = \mathbf{M}\left[A_1 A_2 \ldots A_r\right];$$

$$\mathbf{M}\left[A_1 A_2 \ldots A_r\right] = \mathbf{P}\{A_1 = 1, A_2 = 1, ..., A_r = 1\} = p_{11\ldots1}. \quad (6)$$

Описанный момент ориентирован на единичные значения двоичных переменных. Однако наряду с моментом вида (6) имеют право на самостоятельное использование начальные моменты, допускающие инверсные значения двоичных переменных:

$$\left.\begin{aligned}
&\nu_r\left(\overline{A}_1, A_2, ..., A_r\right) = \mathbf{P}\{A_1 = 0, A_2 = 1, ..., A_r = 1\} = p_{01\ldots1},\\
&\nu_r\left(A_1, \overline{A}_2, ..., A_r\right) = \mathbf{P}\{A_1 = 1, A_2 = 0, ..., A_r = 1\} = p_{10\ldots1},\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,\\
&\nu_r\left(\overline{A}_1, \overline{A}_2, ..., \overline{A}_r\right) = \mathbf{P}\{A_1 = 0, A_2 = 0, ..., A_r = 0\} = p_{00\ldots0}.
\end{aligned}\right\} \quad (7)$$

Такие системы двоичных комбинаций и соответствующие вероятностные моменты $r$-го порядка в количестве $2^r$ образуют полную группу событий.

Из одинаково сформированных групп случайных элементарных переменных БП и АП, согласно выбранному начальному вероятностному моменту, достаточно образовать разностную последовательность (2) вида $\langle a\rangle_n - \langle b\rangle_n = \langle d\rangle_n$, чтобы через средние значения и среднеквадратические отклонения множества частных сумм элементов разности $\langle d\rangle_n$ образовать ряд статистик (1) критерия однородности для всех частных выборок размерности $n$ от 1 до критического значения. Алгоритмически очевидно применение описанных вероятностей элементарных и сложных событий $\nu_1 = p_1$ и $\nu_r = \mathbf{P}\{A_1, A_2, ..., A_r\}$ как начальных вероятностных моментов конечного порядка в роли квалифицирующих параметров $P_a$ или $P_b$ статистического критерия, определенного выражениями (1)–(3).

**Связь центральных вероятностных моментов с начальными**

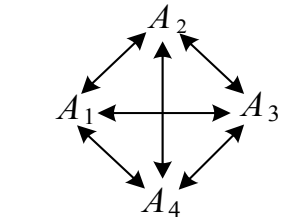
Имитационная модель как объект использования случайных последовательностей может требовать однородности не по начальному, а по центральному вероятностному моменту *r*-го порядка общего вида $\mu_r(A)=\sum_{i=1}^{u}\left(a_i-\nu_1\right)^r p_i$ [13–16]. В этом случае двоичные формы элементарных переменных в событии центрального момента $\mu_r(A)=\sum_{i=0}^{1}\left(a_i-\nu_1\right)^r p_i$ не позволяют свести процедуру тестирования на однородность к таким же простым и технологичным действиям с вероятностями, как это достигается при задании квалифицирующих параметров начальными моментами. Однако, используя выражение центрального момента через оператор математического ожидания $\mu_r(A)=\mathbf{M}\left[(A-\nu_1)^r\right]$, можно получить его разложение на ряд начальных моментов *r*-го и меньших порядков. Так, в литературе по теории вероятностей [14–17] приведены примеры разложения нескольких центральных моментов малого порядка на начальные:

$$\begin{aligned}
&\mu_1=0,\\
&\mu_2=\nu_2-\nu_1^2,\\
&\mu_3=\nu_3-3\nu_1\nu_2+2\nu_1^3,\\
&\mu_4=\nu_4-4\nu_1\nu_3+6\nu_1^2\nu_2-3\nu_1^4.
\end{aligned} \tag{8}$$

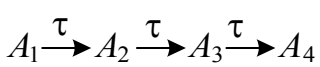
Известно также общее выражение центрального момента через начальные:

$$\mu_r=\sum_{s=0}^{r}\left(-1\right)^s C_r^s \nu_{r-s}\nu_1^s. \tag{9}$$

При этом случайные события в абстрактной форме теоретико-вероятностного описания предполагают полную равноправность по отношению друг к другу, т. е. каждый элемент системы испытывает взаимозависимость от всех остальных, как это показано на рис. 2 а для $r=4$.




а




б

Рис. 2. Образование связей в сложных событиях:
*а* – для равноправных элементарных переменных;
*б* – в последовательностях

Применительно к процессам интересующие нас события $A_j$, где $j=\overline{1,r}$, разделяются однонаправленным дискретным временем $\tau$, дисциплинирующим фор-

мирование системы случайных двоичных величин. Кроме этого, для полноты описания смешанными вероятностными моментами свойства последовательности необходимо ввести управляющие факторы временного расстояния в виде сдвигов элементов относительно друг друга на кратные $\tau$ тактов. Один из вариантов задания неанализируемых позиций двоичных комбинаций реализуется допустимостью условия $r_j = 0$. Однако в этом случае двоичный алфавит элементов последовательности для начальных моментов спровоцирует противоречие «ноль в нулевой степени». В целях его устранения изменим принадлежность $r_j$ с множества натуральных чисел (в частности, с фиксированного значения 1) к множеству неотрицательных целых чисел 0 и 1, что позволит считать $0^0 = 1$ и $0^1 = 0$. На рис. 2 *б* приведен фрагмент расположения событий последовательного характера, участвующий в образовании вероятностного момента, в данном примере 4-го порядка.

Взаимные связи между элементарными событиями последовательности определяются разными временными расстояниями ($\tau, 2\tau, 3\tau$), что нарушает их равноправные отношения. Тем не менее общее представление последовательности, как и в случае образования наборов абстрактных событий, по формуле (9) разложения центрального момента *r*-го порядка, способствует однократному ($C_r^0 = 1$) вхождению начального момента $\nu_r = p_{\underbrace{11\ldots1}_{r}}$ того же *r*-го порядка в качестве слагаемого. Так, например, при $r = 2$ это $\nu_2 = p_{11}$. Формальной инвариантностью разложения (9) также обладают два последних слагаемых при $s = r-1, r$ вида

$$(-1)^{r-1} C_r^{r-1} \nu_1 \nu_1^{r-1} + (-1)^r C_r^r \nu_0 \nu_1^r = (-1)^{r-1}(r-1)\nu_1^r = \begin{cases} (r-1)p_1^r, \text{ когда } r - \text{нечетно}, \\ -(r-1)p_1^r, \text{ когда } r - \text{четно}, \end{cases}$$

дающих неизменное $(r-1)$-кратное вхождение *r*-ой степени вероятности появления 1 со знаком минус или плюс в зависимости от четности/нечетности порядка *r*. Например, при $r = 3$ это $2p_1^3$, а $r = 4$ дает $-3p_1^4$.

**Анализ разложений центральных моментов на начальные для двоичных последовательностей**

Описанные особенности первого и двух последних слагаемых в сумме (9) в виде вероятностей характерны как для двоичной последовательности, так и в классическом теоретико-вероятностном случае в виде начальных моментов равноправных абстрактных событий. То же самое касается части $(s+1)$-кратных вероятностей плотных наборов из $r-s$ двоичных символов последовательности и соответствующих начальных моментов абстрактных событий при $1 \le s \le r-2$, т. е. $p_{11\ldots1}$, $p_1$ и $\nu_r$, $\nu_1$.

Недостаточность описания центральных моментов последовательности в (8) и (9) проявляется при $r \ge 3$ и $1 \le s \le r-2$. В этих случаях требуется включение не более ($C_r^s - s - 1$)-кратных вероятностей комбинаций двоичных переменных, собранных в систему элементов с организацией пропусков тактового времени, нарушающих плотность набора анализируемых переменных.

Введенное переопределение элементарных порядков с $r_j = 1$ на $r_j \in \{0, 1\}$ допускает задание произведения $r_1 \cdot r_2 \cdots r_j \cdots r_r = 1$ для плотных *r*-ичных наборов

и $r_1 \cdot r_2 \cdots r_j \cdots r_r = 0$ для разреженных. Общий вид формул определения смешанных моментов как начальных, так и центральных применительно к двоичным значениям $r_j$ и $A_j$ представим в виде:

$$\nu_{r'}(A_1, A_2, ..., A_r) = \sum_{\langle i_1 i_2 ... i_r \rangle = 0}^{2^r - 1} \prod_{j=1}^{r} a_{j\,i_j}^{r_j} p_{i_1 i_2 ... i_r} = \mathbf{M}\left[A_1^{r_1} A_2^{r_2} ... A_r^{r_r}\right]; \qquad (10)$$

$$\mu_{r'}(A_1, A_2, ..., A_r) = \sum_{\langle i_1 i_2 ... i_r \rangle = 0}^{2^r - 1} \prod_{j=1}^{r} (a_{j\,i_j} - \nu_1)^{r_j} p_{i_1 i_2 .. i_r} = \mathbf{M}\left[A_1^{\circ r_1} A_2^{\circ r_2} ... A_r^{\circ r_r}\right], \qquad (11)$$

где $\langle i_1\, i_2 ... i_r \rangle$ – $r$-разрядное двоичное целое число;

$r' = \sum_{j=1}^{r} r_j$ – частный (фактический) порядок момента как сумма единичных значений $r_j$;

$A_1^{\circ}, A_2^{\circ}, ..., A_r^{\circ}$ – центрированные величины $A_1, A_2, ..., A_r$ относительно общего математического ожидания $\nu_1 = p_1$.

Каждое нулевое значение $r_j$ убирает из сложной комбинации элемент $A_j$ (или $A_j^{\circ}$), увеличивает на один такт временное расстояние между оставшимися элементами и уменьшает на единицу фактический порядок всего вероятностного момента в пределах $1 \le r' \le r$.

Например, при $r = 3$ второе слагаемое в разложении (9) для $\mu_3$ обуславливает 3-кратное вхождение $\nu_2$. Для абстрактной системы событий $A_{j-1}, A_j, A_{j+1}$ это обеспечивается комбинациями $A_{j-1} A_j$, $A_j A_{j+1}$ и $A_{j-1} A_{j+1}$. В последовательности (при ее стационарности) первые две комбинации дают компонент $2p_{11}$, третья комбинация соответствует начальному моменту второго порядка, в общем случае отличному от моментов первых двух комбинаций того же второго порядка. Обозначим этот начальный момент $p_{1 \cdot 1}$ как вероятность совпадения по 1 двух переменных, разделенных двумя тактами реализации последовательности в дискретном времени за счет задания $r_j = 0$.

Нетрудно показать, что реальная случайная переменная $A$ эквивалентна абстрактной системе $A_1, A_2, ..., A_k$ равноправных событий (см. рис. 2 *а*) в случае применимости к ним двоичного алфавита. Согласно замечанию о начальном моменте двоичной случайной величины допустимо считать $\nu_r = p_1$, где $r$ – натуральное число. Тогда разложение (9) центральных моментов на начальные запишется в следующей вероятностной форме:

$$\mu_r = p_1\left[1 + \sum_{s=1}^{r-1} (-1)^s C_r^s p_1^s\right] + p_1^r. \qquad (12)$$

Конкретные виды связей центральных вероятностных моментов первых четырех порядков с начальными моментами для двоичной переменной и $r$-ичной системы двоичной последовательности представлены в табл. 1. Из таблицы видно, что эти разложения существенно отличаются при $r \ge 3$ слагаемыми, записанными в суммах на позициях между $p_1$ или $p_{11...1}$ и $(r-1)p_1^r$.

*Таблица 1*

**Выражения центральных моментов через начальные для двоичной случайной величины и системы двоичных величин последовательности ($\overline{r=1,4}$)**

| $\mu_r$ | Для двоичной $A$ | Для двоичной системы $A_1, A_2, \ldots, A_r$ |
|---|---|---|
| $\mu_1$ | $0,$ | $0,$ |
| $\mu_2$ | $p_1 - p_1^2,$ | $p_{11} - p_1^2,$ |
| $\mu_3$ | $p_1 - 3p_1^2 + 2p_1^3,$ | $p_{111} - (2p_{11} + p_{1\cdot 1})\, p_1 + 2p_1^3,$ |
| $\mu_4$ | $p_1 - 4p_1^2 + 6p_1^3 - 3p_1^4$ | $p_{1111} - (2p_{111} + p_{1\cdot 11} + p_{11\cdot 1})\, p_1 + (3p_{11} + 2p_{1\cdot 1} + p_{1\cdot\cdot 1})\, p_1^2 - 3p_1^4$ |

Для численного определения центрального момента $r$-го порядка двоичной последовательности требуется найти все $2^{r-1}$ начальных моментов частных порядков от 1 до $r$. Например, при $r=4$ необходимо определить $2^{4-1}=8$ вероятностей и в кратном количестве включить их в $\mu_4$ как слагаемые и вычитаемые (см. верхние скобки) в следующем виде:

$$\underbrace{p_{\,1111}}_{\nu_4},\ \overbrace{\underbrace{p_{111}, p_{1\cdot 11}, p_{11\cdot 1}}_{\nu_3}}^{-2},\ \overbrace{\underbrace{p_{11}, p_{1\cdot 1}, p_{1\cdot\cdot 1}}_{\nu_2}}^{3\ \ 2}\ \text{и}\ \overbrace{\underbrace{p_1}_{\nu_1}}^{-4\to -3},$$

что свидетельствует о неприменимости известного представления (8) и в общей форме (9) к рассматриваемым последовательным событиям.

Наборы из символов 1 и неопределенных символов (обозначенных точками), образующие необходимые сложные события последовательности для определения начальных моментов, входящих в центральный момент до 6-го порядка, приведены в табл. 2. Центральный момент $\mu_r$ наряду с $\nu_r = p_{\underbrace{11\ldots1}_{r}}$ включает в себя наборы из всех меньших порядков. Количество дополнительных наборов для данного порядка $r$ относительно $r-1$ обозначено $b_r$, а общая сумма $\sum_{i=2}^{r} b_i$ приведена как $c_r$.

*Таблица 2*

**Комбинации двоичных переменных сложных событий при определении связи начальных моментов двоичной последовательности с центральными для $r=\overline{2,6}$**

| $r$ | Частный порядок $\tau'=\overline{1,r}$ | | | | | | $b_r$ | $c_r$ |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | | |
| 2 | 1 | 11 | – | – | – | – | 2 | 2 |
| 3 | – | 1.1 | 111 | – | – | – | 2 | 4 |
| 4 | – | 1..1 | 1.11, 11.1 | 1111 | – | – | 4 | 8 |
| 5 | – | 1…1 | 1..11, 11..1, 1.1.1 | 1.111, 111.1, 11.11 | 11111 | – | 8 | 16 |
| 6 | – | 1….1 | 1…11, 1..1.1, 11…1, 1.1..1 | 1..111, 1.1.11, 1.11.1, 111..1, 11.1.1, 11..11 | 1.1111, 1.111, 1.1111, 111.1, | 111111 | 16 | 32 |

Нетрудно заметить в табл. 2, что при увеличении порядка центрального момента на единицу к двоичной комбинации соответствующего начального момента из плотного набора $r$ единиц добавляются только $r-1$ неплотных наборов,

с кратно увеличенными задержками между внутренними символами. Это замечание делает очевидным заполнение табл. 2 для любого конечного *r*.

**Формирование последовательностей сложных событий совпадения двоичных элементов для критерия однородности**

Следует подчеркнуть, что для работы критерия однородности по вероятностным моментам не требуется вычислять числовые значения моментов, например в виде оценок соответствующих вероятностей. Достаточно воспроизвести сами последовательности событий по обеим тестируемым последовательностям и использовать их в качестве входных данных для работы критерия: образовать разность $\langle d \rangle_n$ по (2), оценить среднее, дисперсию $\Delta p^*$ для статистики (1), получить цепочку реализации испытаний гипотез (3) и выбрать результат в виде $n_{\text{кр}}$ (или «не менее $n_{\max}$»).

Последовательно формируемые случайные отсчеты $A_{j-1}, A_j, A_{j+1} \ldots$, принимающие двоичные значения, объединяются конъюнкцией в группы, образующие сложные события (6) и/или (7), с учетом задержек между ними и инверсий. Варианты аппаратного формирования событий 4-го порядка представлены на рис. 3. Образование произведения элементарных переменных в прямой форме показано на рис. 3 *а*. Участие элементарных переменных в сочетании прямой и инверсной форм изображено схемой на рис. 3 *б*. Пример формирования неплотного набора элементарных переменных в прямой форме путем добавления задержки между первой и третьей переменными представлен на рис. 3 *в* для начального момента 3-го (частного) порядка.

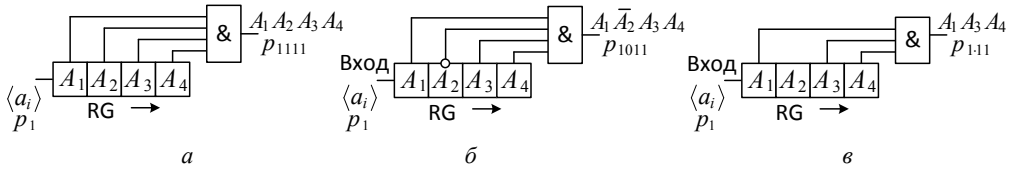



Рис. 3. Аппаратное формирование сложных событий двоичных последовательностей для определения начальных моментов, включаемых в состав заданного центрального момента

*Краткая методика определения критической длины значимо однородных последовательностей по начальному вероятностному моменту.* Заданным порядком *r* и конкретным видом начального вероятностного момента определяем необходимую функцию преобразования вида (6) обеих последовательностей на основе операций сдвигов, конъюнкции и, в необходимых случаях для вида (7), инверсии. Вероятность полученных комбинаций двоичных переменных представляет заданный начальный момент. Из последовательностей полученных событий от обеих тестируемых последовательностей образуем разностную последовательность, элементы которой используем как входные данные для статистического критерия. Проводим серию итераций по вычислению статистики, величины которой сравниваются с критическим уровнем, образуя испытания гипотез. Результатом этих итераций является цепочка (3) принятия нуль-гипотезы, классифицирующая свойство однородности тестируемых последовательностей с заданной значимостью по за-

данному начальному вероятностному моменту. Событие нарушения условия непротиворечивости нуль-гипотезы фиксируется как наступление критической длины частной выборки, превышение которой интерпретируется как прекращение однородности.

*Краткое дополнение к методике определения критической длины значимо однородных последовательностей по заданному центральному вероятностному моменту.* Используя оператор математического ожидания (11), получаем разложение заданного смешанного центрального момента *r*-го порядка на множество из $2^{r-1}$ начальных моментов, вид которых определен в табл. 2 как содержимое клеток с единичными наборами частных порядков. Определяем все необходимые функции преобразования видов начальных моментов с размерностью частных порядков от 1 до *r*, входящих в заданный центральный момент. Проводим необходимое количество серий итераций вычислений статистик для всех сформированных начальных моментов по вышеописанной методике. В результате проведенных в полном объеме серий тестовых итераций формируется множество значений критических длин, минимальная величина которых является искомой критической. За ее пределами тестируемые последовательности интерпретируются как неоднородные по заданному центральному вероятностному моменту.

**Заключение**

Получены обоснования применимости статистического метода определения значимой однородности к двоичным последовательностям по вероятностным моментам любого конечного порядка.

Выбор начального момента в качестве основного квалифицирующего ограничения однородности допускает однократное проведение цепочки испытаний гипотез. В случае выбора центрального момента работа критерия существенно усложняется необходимостью проверки однородности для всех порядков от 1 до *r* включительно. Полная аналитическая форма связи центрального момента *r*-го порядка с начальными для двоичной последовательности еще не установлена. Однако логика заполнения полученной в работе табл. 2 на алгоритмическом уровне вполне заменяет математическое выражение в задаче практической реализации критерия.

Анализируя содержание табл. 2, нетрудно сделать вывод, что центральный вероятностный момент порядка $r \geq 2$ включает в себя автокорреляционные зависимости в области определения аргумента $\tau = \overline{1, r-1}$ при условии однородности по вероятности $p_1$. Это позволяет наряду с вероятностными моментами в качестве исходных квалифицирующих данных критерия применять значения автокорреляционной функции.

Процедура определения статистической однородности случайных последовательностей применима в задачах имитационного моделирования, вычислительных методах Монте-Карло, системах защиты информации. Обеспечение области существования значимой однородности в гарантированных пределах заданных вероятностных моментов повышает объективность сравнительного анализа свойств стохастических объектов и способствует повышению достоверности результатов машинной реализации математических моделей.

Приведенный материал может быть использован для синтеза аппаратных или программных анализаторов моментов и моментных функций как на ПЭВМ, так и с использованием многопроцессорных средств. Выявленные особенности анали-

тических связей начальных и центральных вероятностных моментов двоичных последовательностей (6), (7), (10)–(12) и табл. 2 достаточно полно задают арифметико-логические алгоритмы обработки моментных функций, ориентированные, например, на ресурсы программируемых интегральных схем в форме «системы-на-кристалле».

Рассматривается перспектива развития методов сравнительного исследования генераторов физически случайных последовательностей на предмет построения индикаторов поддержания статистически гарантированных штатных режимов работы в реальном времени. Разрабатывается методика оперативной поверки качества источников сложных двоичных сигналов на основе рассмотренной концепции однородности.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК


1. *Советов Б.Я., Яковлев С.А.* Моделирование систем: 3-е изд., перераб. и доп. М.: Высш. шк., 2001. 343 с.
2. *Forrester J.W.* Industrial Dinamics, Portland, OR: Productivity Press, 1961. 464 p.
3. *Knuth D.E.* The Art of Computer Programming: Seminumerical Algorithms. 3nd Edition. Addison-Wesley, 1997. 782 p.
4. *Bangsow S.* Manufacturing Simulation with Plant Simulation and Sim Talk Usage and Programming with Examples and Solutions. Springer, 2010. 300 p.
5. *Robert C.P., Casella G.* Monte Carlo Statistical Methods. 2nd Edition, Springer, 2004. 683 p.
6. *Schneier B.* Applied Cryptography. Protocols, Algorithms, and Source Code in C. New York: John Wiley & Sons, 1996.
7. *Johnston D*. Random Number Generators – Principles and Practices: A Guide for Engineers and Programmers. DEG Press, 2018. 439 p.
8. *Fischer V., Drutarovsky M.* True random number generator embedded in reconfigurable Hardware // Cryptographic Hardware and Embedded Systems – CHES 2002, Redwood Shores, CA, USA, 2002, Revised Papers, ser. LNCS, vol. 2523. Springer, 2002. Pp. 415–430.
9. *Dichtl M., Golic J.* High-speed true number generation with logic gates only. Cryptographic Hardware and Embedded Systems – CHES 2007, Vienna, Austria, 2007, Proceedings, ser. LNCS, vol. 4727. Springer, 2007. Pp. 45–61.
10. Recommendation for the entropy sources used for random bit generation // M.S. Turan, E. Barker, J. Kelsey, K.A. McKay, M.L. Baish, M. Boyle. NIST Special Publication. Computer Science Published. 2018. 84 p.
11. *Golomb S.W.* Shift Register Sequences, San Francisco: Holden-Day, 1967. 224 p.
12. *Колмогоров А.Н., Успенский В.А.* Алгоритмы и случайность // Теория вероятностей и ее приложения. М.: Наука, 1987. Т. XXXII, вып. 3. С. 425–455.
13. *Hossein Pishro-Nik.* Introduction Probability, Statistics and Random Processes. Kappa Research, LLC, 2014. 747 p.
14. *Dodge Y.* The Concise Encyclopedia of Statistics Authors: Springer Science + Business Media, LLC, 2008. 616 p.
15. *Кремер Н.Ш.* Теория вероятностей и математическая статистика. М.: Юнити-Дана, 2002. 543 с.
16. *Петров А.В.* Исчисление смешанных моментов высших порядков при полиномиальной зависимости случайных величин // Вестник ИрГТУ, 2015. № 11. С. 16–22.
17. *Петров А.В.* К вопросу нормирования вероятностных характеристик // Вестник ИрГТУ, 2016. № 1. С. 56–64.


*Статья поступила в редакцию 10 января 2020 года*

# ANALYSIS OF PROPERTIES OF PROBABILISTIC MOMENTS OF BINARY SEQUENCES FOR STOCHASTIC OBJECTS MODELING

***V.M. Kuznetsov, V.A. Pesoshin, A.I. Gumirov, D.V. Shirshova***

Kazan National Research Technical University named after A.N. Tupolev
10, K. Marx St., Kazan, Tatarstan, 420111, Russian Federation

***Abstract.*** *The applicability of the statistical method for determining the significant homogeneity of binary sequences under probabilistic moments of any finite order for solving problems of mathematical modeling of stochastic objects is proved. Approaches to the construction of the homogeneity criterion in the framework of various ways of setting probabilistic moments are described. Significant differences in the decomposition of central probabilistic moments into raw ones for systems of random events of a homogeneous nature and concerning sequences are revealed. Brief theses of methods for finding the critical length of samples within which binary sequences have significant homogeneity over the specified raw or central probabilistic moments are presented. Areas of application of the concept of homogeneity of random sequences in theoretical statements of problems of machine implementation of mathematical models and practical development of means of adequate comparison of statistical characteristics of stochastic objects are determined.*

***Keywords:*** *simulation, probabilistic moment, raw moment, central moment, significant homogeneity, homogeneity criterion, binary sequence.*

REFERENCES

1. *Sovetov B.Ya., Yakovlev S.A.* System modeling: 3-e izd., pererab. i dop. M.: Vyssh. shk., 2001. 343 p.
2. *Forrester J.W.* Industrial Dinamics, Portland, OR: Productivity Press, 1961. 464 p.
3. *Knuth D.E.* The Art of Computer Programming: Seminumerical Algorithms. 3rd Edition. Addison-Wesley, 1997. 782 p.
4. *Bangsow S.* Manufacturing Simulation with Plant Simulation and Sim Talk Usage and Programming with Examples and Solutions. Springer, 2010. 300 p.
5. *Robert C.P., Casella G.* Monte Carlo Statistical Methods. 2nd Edition, Springer, 2004. 683 p.
6. *Muller M.E.* Some continuons Monte-Karlo methods for the Dirichlet problem. Annals Math. Statistics, 1956, v. 27, № 3. P. 569–589.
7. *Amann H.* Monte-Karlo methoden und lineare randwerprobleme. ZAMM, 1968, № 48. S. 109–116.
8. *Schneier B.* Applied Cryptography. Protocols, Algorithms, and Source Code in C. New York: John Wiley & Sons, 1996. 758 p.
9. *Neuman F.* Autocorrelation peaks in congruential pseudorandom number gene-rators. IEEE Transactions on Computers, 1976, Vol. 25, № 5. P. 457–460.
10. *Johnston D.* Random Number Generators – Principles and Practices: A Guide for Engineers and Programmers. DEG Press, 2018. 439 p.
11. *Fischer V., Drutarovsky M.* True random number generator embedded in reconfigurable Hardware. Cryptographic Hardware and Embedded Systems – CHES 2002, Redwood Shores, CA, USA, 2002, Revised Papers, ser. LNCS, vol. 2523. Springer, 2002. P. 415–430.
12. *Dichtl M., Golic J.* High-speed true number generation with logic gates only. Cryptographic Hardware and Embedded Systems – CHES 2007, Vienna, Austria, 2007, Proceedings, ser. LNCS, vol. 4727. Springer, 2007. P. 45–61.
13. Recommendation for the entropy sources used for random bit generation. M.S. Turan, E. Barker, J. Kelsey, K.A. McKay, M.L. Baish, M. Boyle. NIST Special Publication. Computer Science Published. 2018. 84 p.

*Valery M. Kuznetsov (Dr. Sci. (Techn.)), Professor.*
*Valery A. Pesoshin (Dr. Sci. (Techn.)), Professor.*
*Artyom I. Gumirov, Senior Lecture.*
*Darya V. Shirshova, Senior Lecture.*

14. *Golomb S.W*. Shift Register Sequences, San Francisco: Holden Day, 1967. 224 p.
15. *Kolmogorov A.N., Uspenskij V.A.* Algorithms and randomness. Probability Theory and its Applications. M.: Nauka. 1987, Vol. XXXII, No. 3. P. 425–455.
16. *Hossein Pishro-Nik*. Introduction Probability, Statistics and Random Processes. Kappa Research, LLC, 2014. 747 p. https://www.probabilitycourse.com.
17. *Dodge Y.* The Concise Encyclopedia of Statistics Authors: Springer Science + Business Media, LLC, 2008. 616 p.
18. *Kremer N.Sh.* Theory of Probability and Mathematical Statistics. M.: Yuniti-Dana, 2002. 543 p.
19. *Petrov A.V*. Calculation of mixed moments of higher orders in the polynomial dependence of random variables. Vestnik IrGTU, 2015, № 11. P. 1622.
20. *Petrov A.V.* To the issue of standardization of probabilistic characteristics. Vestnik IrTU, 2016, № 1. P. 56–64.