

МЕТОД ОБРАБОТКИ, АНАЛИЗА И ПРОГНОЗА ХАРАКТЕРИСТИК ПОТОКОВ ИЗЛУЧЕНИЙ ПО ВЫБОРКАМ МАЛОГО ОБЪЕМА

© 2025 г. В. А. Работкин^а, Н. М. Близняков^а, В. М. Вахтель^а, Д. Е. Костомаха^а

^аВоронежский государственный университет, Воронеж, 394018 Россия

*E-mail: vaktel@phys.vsu.ru

Поступила в редакцию 21.05.2024 г.

После доработки 21.05.2024 г.

Принята к публикации 31.05.2024 г.

Предложен метод обработки и анализа близости по типоопределяющим характеристикам больших совокупностей $M > 10^4$ различных типов $Q > 10^2$ эмпирических дискретных векторов случайных частот ВСЧ $\mathbf{v}(\cdot) = v_0, \dots, v_l$, полученных по выборкам малого объема $10 \geq n = \sum_{i=1}^l v_i(k = i)$ случайных отсчетов $k = 0, 1, \dots, l$ со средними по всем выборкам $\bar{k} < 5$. Метод основан на взаимно однозначном соответствии ВСЧ и его типоопределяющей характеристики–идентификатора $I(\mathbf{v}, a) > 0$ – линейной статистики в форме скалярного произведения \mathbf{v} и неслучайного вектора a . Дискретные многомодальные эмпирические распределения $C(I(\mathbf{v}, a))$, представляющие последовательности упорядоченных по $I(\mathbf{v}, a)$ сгруппированных пиков, позволяют анализировать и прогнозировать характеристики пиков и образующих их ВСЧ с малыми частотами их реализаций при данном M .

Ключевые слова: излучение, вектор случайных частот, выборка, распределение, идентификатор, прогноз

DOI: 10.56304/S2079562924060290

ВВЕДЕНИЕ

В радиометрии различных излучений при циклическом многократном повторении измерений длительностью одного цикл $\theta \ll T$ в течение времени T могут регистрироваться небольшие количества $v(k) \leq 10$ отсчетов $k(\Delta t) = 0; 1; 2; \dots, l$ в последовательных интервалах длительностью $\Delta t \sim 0.5/I$ со средним значением $\bar{k}(\Delta t) < 5$ при интенсивности I в течение интервала стационарности $\theta = n \cdot \Delta t$. Каждой из $M = T/\theta$ полученных случайных последовательностей выборок $\{k_i(\Delta t)\}_n$ соответствует эмпирическое распределение частот $v_j(k_j(\Delta t) = j), j = 0, 1, \dots, l$, т.е. вектор случайных частот (ВСЧ) $\mathbf{v}(\cdot) = v(k, n, l, \theta) = (v_0, v_1, \dots, v_l)$, $\sum_0^l v_j = n, l \leq 10$. Число возможных типов реализаций ВСЧ, т.е. случайных комбинаций компонентов v_j векторов $\mathbf{v}(\cdot)$ при данных n, l , не превышает $Q(n, l) = (n + (l - 1))! / (n! (l - 1)!)$ [1]. В частности, при $n = 10, l = 10$ и числе реализаций векторов $M \gg 10$ возможно $Q = 923778$. Обработка и статистический анализ таких совокупностей ВСЧ, например, для оценки степени однородности или близости их распределений, является не простой задачей, так как известные процедуры

обработки и критерии оценки близости ВСЧ мало эффективны при $n \leq 10, Q \geq 10$ [2, 3]. Особый интерес представляют “хвосты” эмпирических распределений, т. к. они обусловлены редкими значениями случайных величин k с вероятностями $P(|k - \bar{k}| \geq 1) \ll 10^{-N}, N = T/\Delta t$. Значения $k_j \gg \bar{k}$ с $v_j(k_j) \sim 1$ могут быть или значимыми большими уклонениями в “хвостах” распределений или же артефактами [4]. Однако именно оценки “хвостов”, использующие различные семейства распределений, дают существенно различные результаты.

МЕТОД ОБРАБОТКИ И АНАЛИЗА ВСЧ

В данной работе предложен метод обработки и анализа последовательностей из M выборок, каждая из n отсчетов k_i , и соответствующим ВСЧ при указанных выше условиях $n \leq 10, \bar{k} < 5$, основанный на взаимно однозначном соответствии между вектором $\mathbf{v}(\cdot)_h$ и дискретной линейной функцией $I(\mathbf{v}(\cdot))_h = I_h$ в форме скалярного произведения $\mathbf{v}(\cdot)_h = (v_0, \dots, v_l)_h$ и заданного проектирующего вектора $a = (a_0, \dots, a_l)$, постоянного для всех M анализируемых ВСЧ:

$$0 < I(v(\cdot)|a)_h = a_0 v_{0h} + \dots + a_i v_{ih} + \dots + a_l v_{lh}.$$

Компоненты a_i проектирующего вектора $a(\cdot)$ – целые неотрицательные числа, не имеющие общих делителей. В [5] доказана сформулированная ниже в приложении лемма, что при этих условиях имеется взаимно однозначное соответствие между совокупностями $I(v(\cdot)|a)_h = I_h$ и $v(\cdot)_h$. Поэтому исходной случайной последовательности из M ВСЧ – $\{v(\cdot)_h\}_{1,M}$ однозначно соответствует случайная последовательность целочисленных значений $\{I_h\}_{1,M}$, состоящая как из многократно повторяющихся, так и единичных значений I_h . Это позволяет обрабатывать и анализировать характеристики ВСЧ, используя I_h как числовой идентификатор типа вектора $v(\cdot)_h$.

Варьируя значения и соотношения между компонентами a_i вектора $a(\cdot)$ можно получать соответствующие различные распределения $I(v(\cdot))_h$, что позволяет детально анализировать их статистические характеристики.

Случайные совокупности из $M > 10^4$ ВСЧ с $n \leq 10$, $l = 10$ и $\bar{k} < 5$, были получены в работе компьютерным сканированием экспериментальных последовательностей единичных отсчетов альфа-частиц, нейtronов, а также стандартным способом генерации случайных чисел [4, 6].

СТРУКТУРА И ХАРАКТЕРИСТИКИ РАСПРЕДЕЛЕНИЙ ИДЕНТИФИКАТОРА I_h

Исходная последовательность из M значений идентификатора после ее ранжирования представляет вариационный ряд, которому однозначно соответствует дискретное многомодальное распределение целочисленных частот $C_h(I(v)_h) = C_h(I_h)$ значений I_h и соответствующих им ВСЧ в виде последовательности сгруппированных пиков, характеризуемых структурными индексами m, g, Ψ, ϕ . Индекс $m = 0, 1, 3, \dots, 9$ – это номер пика в соответствующей группе с индексом $g = 0; 1; 3; \dots; 9$, которая в свою очередь входит по значению I_h в определенный блок с индексом $\Psi = 0; 1; 3; \dots$ и супер блок $\phi_i = 0, 1, 3, \dots, 9$.

Значения индексов m, g, Ψ, ϕ_i каждого пика m определяют компоненты a_i задаваемого проектирующего вектора a и компоненты $v_{h,i,m,g,\Psi,\phi}$ ВСЧ, определяемыми, в свою очередь, исследуемыми эмпирическими распределениями отсчетов $P(k)$ выборок.

В случае вектора a с компонентами $a_0 < a_1 < \dots < a_l$ выполняются следующие структурные соотношения: индекс пика $m = v_{2,h,m,g,\Psi}$ в группе g ; индекс группы $g = v_{3,h,m,g,\Psi}$ в блоке Ψ ; и

индекс блока: $\Psi = v_{h=4,5,6,m,g}$. Количество различных типов ВСЧ, образующих один пик m , равно $q_m = v_{0,1,m,g,\Psi} + v_{1,1,m,g,\Psi} + 1$, где $v_{0,1,m,g,\Psi} = 0, 1, \dots, l$ – значение первого компонента первого ВСЧ в пике, т.е. ВСЧ с минимальным $I(v)_h$ в составе пика m . Количество пиков в одной группе g равно $q_g = n + 1 - g - \Psi = v_{0,m=0,g} + 1$. Количество типов ВСЧ в одной группе g блока Ψ определяется комбинаторикой значений $v_0, v_1, v_2; v_0 + v_1 + v_2 = \text{const}(g)$ и равно: $Q(g, \Psi) = \frac{1}{2}(n + 2 - g - \Psi) \times (n + 1 - g - \Psi) + 1$. Количество групп в одном блоке равно $v_{2,m=1} + 1$.

В случае варианта компонент вектора a : $a_0 > a_1 > \dots > a_l$ количество различных ВСЧ в пике m равно $q_m = v_{l,m} + 1$, $l = l$. Индекс пика m в группе равен $g = v_{l-2,m}$; индекс g группы в блоке Ψ равен $g = v_{l-3,m}$; а индекс блока $\Psi = v_{l-4\dots}$.

Статистически значимый пик m с $\sum_1^{q_m} C_{h,m} > q_m \cdot 5$ является локальным распределением $C_m(I(v)_h) = C_m(.) = (C_{1,m}(I_{1,m}), \dots, C_{h,m}(I_{h,m}), \dots, C_{q,m}(I_{q,m}))$, и при $a_0 < a_1 < \dots < a_l$ он включает $q_m = v_{0,1,m} + v_{1,1,m} + 1$ близких типов ВСЧ, однородных по “правым” хвостам компонент $v_{j,m}$, $j > 2$, удовлетворяющих условию $v_{j,m} = \text{const}(j, m)$. В каждом пике m при $a_0 < a_1 < \dots < a_l$ интервалы между последовательными значениями $I_{h,m}(v)$ составляют $\Delta I = I_{j+1,m} - I_{j,m} = a_1 - a_0$. При $a_0 > a_1 > \dots > a_l$ условия однородности ВСЧ по их хвостам имеют вид: $v_{l,m} + v_{l-1,m} = \text{const}(l, m, g, \Psi)$; $v_{0,m} = \text{const}(0, m)$, а соотношения между компонентами вектора a определяют вид распределений $C(I_h)$ в целом и локальные распределения в виде пиков. Максимум распределения пика $C(I)_{h,m}$ образуется при $I_h = I_{1,m}(v_{0,1,m}(\cdot)) + v_{0,1,m}(\cdot)/2$, если $v_{0,1,m}$ – четное или $I_h = I_{1,m}(v_{0,1,m}(\cdot)) + (v_{0,1,m}(\cdot) + 1)/2$, если нечетное.

Интервал ΔI между соседними пиками с $\Delta m = 1$ в одной группе, т.е. $\Delta g = 0; \Delta \Psi = 0$ составляет $\Delta I = I_{h,m+1} - I_{h,m} = a_2 - a_0$. Интервал $\Delta I(\cdot)$ между близкими пиками последовательных групп $\Delta g = g_j - g_{j+1} = 1$ одного блока $\Delta \Psi = 0$ составляет $\Delta I = a_3 - a_2$. Интервал ΔI между близкими (в смысле статистик [5, 13]) пиками из последовательных блоков, т.е. $\Delta \Psi = 1, \Delta g = 1, \Delta \Psi + \Delta g = 0$ и $\Delta m = 0$ составляет $\Delta I(\cdot) = \dots = a_4 - a_3$.

Таким образом, упорядоченная последовательность идентификаторов $I_{h,m,g,\Psi}$ ВСЧ имеет однородную периодическую структуру характеризуемую параметрами m, h, g, Ψ, a . Рассмотрен-

ные средства описания, анализа структуры и характеристик распределений $C(I_h)$ позволяют выполнять online процедуры: компьютерной обработки, классификации более $Q > 10^3$ типов ВСЧ, анализа однородности и близости больших совокупностей ВСЧ $M > 10^4$, образующих более сотен пиков.

МОДЕЛИ И АНАЛИЗ РАСПРЕДЕЛЕНИЙ ИДЕНТИФИКАТОРА ВСЧ

Каждый пик распределения $C(I_h)$ является локальным условным распределением, образованным q_m случайными количествами $C_{h,m}(v(\cdot))$ соответствующих ВСЧ. Вероятность образования вектора $v(\cdot)_{h,m} = (v_{0,h,m}, v_{1,h,m}, \dots, v_{l,h,m})$ с определенной комбинацией его компонент $v_{j,h,m} \geq 0$, удовлетворяющих условию $\sum_{i=0}^l v_{i,h,m} = n$, соответствует невырожденному условному полиномиальному распределению [8, 9]:

$$(v(\cdot)_{h,m}) = \frac{n!}{v_0! \dots v_{l,q}!} P_0^{v_0} P_1^{v_1} \dots P_l^{v_l} P_{l+1}^0, \quad (1)$$

$$l+1 = 1 - \sum_0^l P_{i,m}, \quad v_{l+1} = 0,$$

где $P_{i,h}$ – вероятность наблюдения значений k_i в стационарных условиях. Если величина $I(v(\cdot))$ каждого типа ВСЧ распределены равномерно среди других, то оценками P_i – могут являться величины $N(k_i)/N$, $N = T/\theta$, $\sum N_i(k_i) = N$ или P_i можно оценить как средневзвешенные по выборочным распределениям $C(I(v(\cdot)))_h$, т.е. $\tilde{P}_h(v(\cdot)_h) = \tilde{C}_h(I(v(\cdot)_h))/M$. Ожидаемые средние частоты образования ВСЧ при $M \gg 1$ очевидно равны $C_{h,m}(I(v_h)) = M$ ($v(\cdot)_{h,m}$). Полученные в эксперименте эмпирические величины $\tilde{v}(\cdot)_h = \tilde{C}_h(v(\cdot)_h)/M$ при $M > 1 \cdot 10^6$ изменяются в диапазоне $10^{-2} - 10^{-7}$ и при $\tilde{M}_h(v(\cdot)_h) > 10^2$ согласуются с теоретическими ($v(\cdot)$) при разбросе в основном в несколько процентов (табл. 1). Характеристики многомодального распределения идентификатора $C(I(v(\cdot)))$ обусловлены соотношением компонент вектора a и соотношением вероятностей $v_{i,j,m}$, $v_{i+1,j,m}$ пар последовательных компонент ВСЧ $v_{i,j,m}$, $v_{i+1,j,m}$ и в частности $v_{0,h,m}$, $v_{1,h,m}$ каждого пика.

Многократно повторяющиеся значения I_h при фиксированных значениях Δt , n , \bar{k} распределены равномерно в исходной последовательности из других значений I_h в согласии с дискретным равномерным полиномиальным распределением на уровне значимости $\alpha < 0.05$, например по крите-

рию χ^2 . Интервалы $r_{i,j}$, τ_{hj} между $C(I_h)$ многократно повторяющимися одинаковыми I_h в последовательности всех других значений I_h распределены в согласии с геометрическим распределением на уровне значимости $\alpha > 0.05$. Средние значения интервалов \bar{r}_j согласуются с оценкой $\bar{r}_j = \left(\frac{M}{C_{hj}} - 1 \right) > 1$, полученной для всех типов ВСЧ со среднеквадратическим разбросом $S(r_{hj}) = 1/\bar{r}_j$ и асимметрией $A = 2$.

Эти свойства распределений I_h и r_h каждого типа ВСЧ позволяют считать последовательности величины I_h и соответствующих $v(\cdot)_h$ не противоречащими модели пуассоновского потока [7]. Неточность потоков ВСЧ из-за конечности неперекрывающихся интервалов θ несущественна при $1 \ll \tilde{r}$, как и при известном эффекте мертвого времени [12]. Отметим, что каждому вектору $v(\cdot)_h$ однозначно соответствует величина $L_h = \sum_i^l k_i v_{i,h}$, $0 \leq k_i \leq l$, значения которой равно количеству отсчетов в интервале времени θ , а суммарное количество k_i всех q_m образующих пик m векторов равно $S_m = q_m (L(v_{1,h,m}) + 1)$. Однако L_h и S_m не являются однозначными идентификаторами в отличие от I_h .

АНАЛИЗ ХАРАКТЕРИСТИК РАСПРЕДЕЛЕНИЙ ПИКОВ

Анализ близости характеристик пиков как локальных распределений в спектрах $C(I(v(\cdot)))$ можно выполнять на основе устойчивых соотношений между типохарактеризующими близость и однородность компонентами ВСЧ, образующих пики. Каждый пик m образован упорядоченной по величине последовательностью идентификатора $I_{i+1}(v(\cdot)) > I_i(v(\cdot))$ ВСЧ, компоненты которых удовлетворяют следующим соотношениям:

$$v_{0,j,m} + v_{1,j,m} = \text{const}(m, g, \Psi), \quad 1 < i < q_m, \quad q_m = v_{0,1,m} + 1; \quad (2)$$

$$v_{i,j,m} = \text{const}(i, j, m), \quad 2 < i \leq l$$

при заданном векторе a и значениях n , l . С другой стороны, каждый пик m также, как и ВСЧ – можно рассматривать как случайный вектор $\tilde{C}(m, g, \Psi) = (\tilde{C}_{1,m,g,\Psi}, \dots, \tilde{C}_{h,m,g,\Psi}, \dots, \tilde{C}_{q_m,m,g,\Psi})$ из $q_m = v_{0,1,m} + 1$ компонент, представляющий собой реализацию усеченного распределения дискретных эмпирических частот $\tilde{C}_{h,m,g,\Psi}(I(v(\cdot)))$, которому соответствует вектор вероятностей $(m, g, \Psi) = (v_{1,m,g,\Psi}, \dots, v_{h,m,g,\Psi}, \dots, v_{q_m,m,g,\Psi})$, где $v_{h,m,g,\Psi}$ – ве-

Таблица 1

Тип ВСЧ (v_0, \dots, v_l)	\tilde{C}_j	$\tilde{C}'_j = \tilde{C}_j/M$	$P_j(\cdot)$ полином вер.	$\tilde{P}_{j_{\min}} - \tilde{P}_{j_{\max}}$ – доверительный интервал $\alpha = 0.05$
$\Delta t = 0.05c; \theta = 0.5c; n = 10; \bar{K}(\Delta t) = 1.1738; M = 5904724; l = 10$				
3421	144594	2.449-2	2.443-2	2.436–2.461 $\cdot 10^{-2}$
14401	10114	1.1713-3	1.701-3	1.680–1.747 $\cdot 10^{-3}$
250201	998	1.690-4	1.815-4	1.589–1.798 $\cdot 10^{-4}$
402301	100	1.694-5	1.737-5	1.393–2.060 $\cdot 10^{-5}$
42030001	10	1.694-6	1.645-6	0.919–3.118 $\cdot 10^{-6}$
070102	5	8.468-7	8.204-7	0.362–1.922 $\cdot 10^{-6}$
$\Delta t = 0.1c; \theta = 1c; n = 10; \bar{K}(\Delta t) = 2.3477; M = 2952362$				
13222	10353	3.507-3	3.311-3	3.440–3.575 $\cdot 10^{-3}$
23032	1000	3.387-4	3.177-4	3.184–3.604 $\cdot 10^{-4}$
13201201	100	3.387-5	3.300-5	2.785–1.849 $\cdot 10^{-5}$
1120312	34	1.152-5	1.119-5	0.824–1.609 $\cdot 10^{-5}$
0214003	10	3.387-6	3.316-6	1.848–6.236 $\cdot 10^{-6}$
302220001	5	1.694-6	1.627-6	0.723–3.965 $\cdot 10^{-6}$
$\Delta t = 0.2c; \theta = 2c; n = 10; \bar{K}(\Delta t) = 4.6953; M = 1476181$				
0012322	1015	6.876-4	6.698-4	6.466–7.312 $\cdot 10^{-4}$
0030223	100	6.774-5	6.698-4	5.570–8.238 $\cdot 10^{-5}$
0210122101	34	2.303-5	2.215-5	1.648–3.218 $\cdot 10^{-5}$
1020122101	17	1.152-5	1.107-5	0.719–1.844 $\cdot 10^{-5}$

роятности реализации ВСЧ– $v(\cdot)_h$ в полиномиальной модели (1).

Таким образом, векторы $C(m, g, \Psi)$, компоненты которых $\tilde{C}_{h,m,g,\Psi}(I(v(\cdot)))$ согласуются с соотношениями (2) и условием $v_{2,j,m,\Psi} + \dots + v_{l,j,m,\Psi} = \text{const}$, образуют упорядоченные по $I(v(\cdot))$ последовательности пиков близких типов с одинаковыми числами компонент $q_m = v_{0,j,m,\Psi}$, комбинаторика которых и определяет тип пика. Однако эмпирические распределения пиков как случайных векторов $C_{m,g,\Psi}(I(\cdot))$ сложно анализировать и тем более прогнозировать компоненты их “хвостов” образующих их ВСЧ, при $Q > 10^3$, особенно в условиях малой статистики при значительной разнице значений компонент $C_{h,m} > C_{h,r}$ анализируемых пиков m и r .

В работе предлагается исследовать статистические характеристики пиков в совокупности, рассматривая пики, как обсуждалось выше, как случайные векторы, которые представляют локальные

$$C_r = \sum_{h=1}^{q_m} C_{h,r}(\cdot) < q_m \cdot 10; n \leq 10; Q > 10^3,$$

поскольку известные критерии однородности или близости в таких случаях не эффективны [3].

В работе предлагается исследовать статистические характеристики пиков в совокупности, рассматривая пики, как обсуждалось выше, как случайные векторы, которые представляют локальные

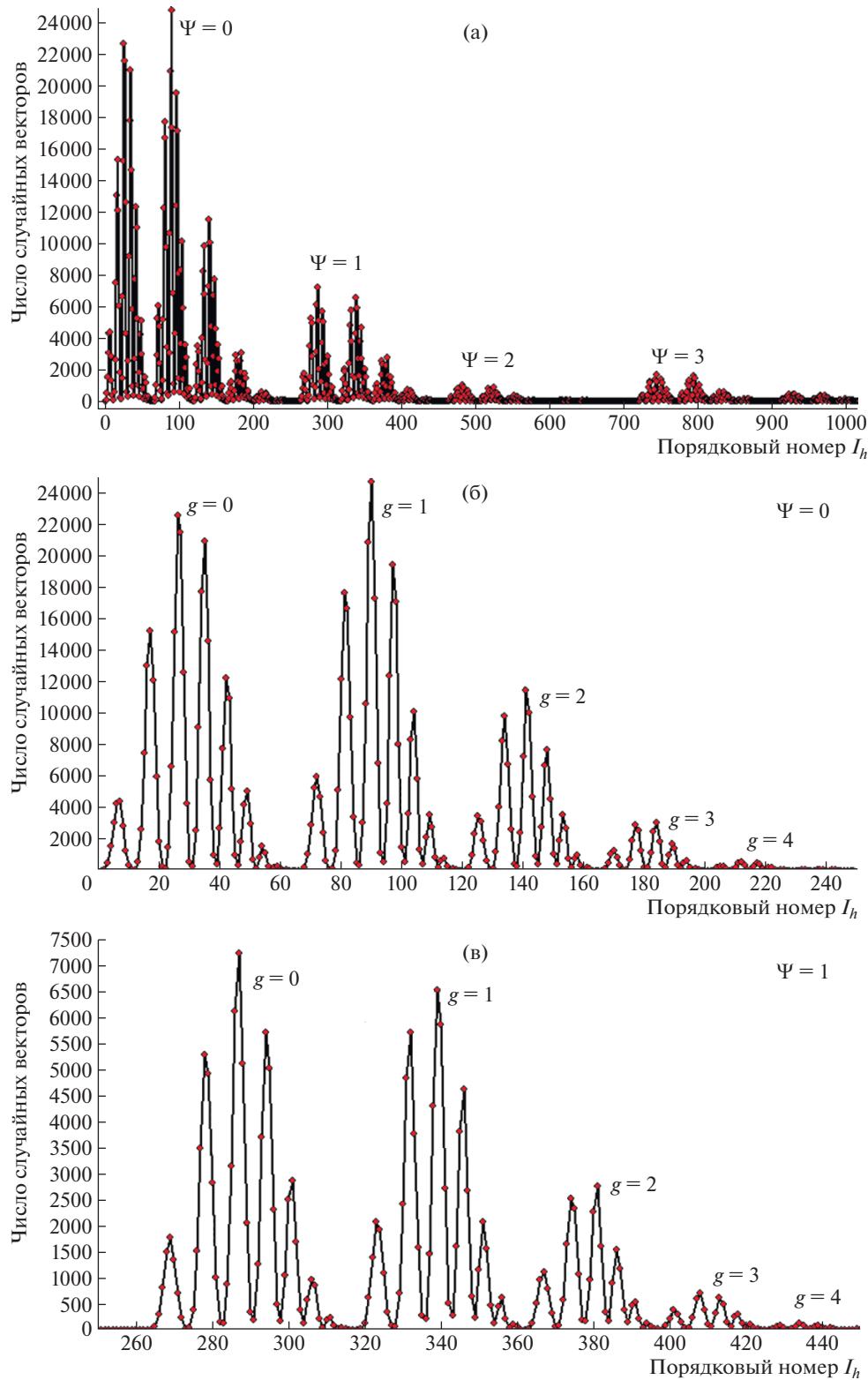


Рис. 1. Распределение эмпирических частот $C(I(v(.)))$ идентификатора $I(v(.))$ для последовательности из $M = 1006458$ ВСЧ, размерности векторов $l = 10$ при $n = 10$, $\bar{k}(\Delta t = 0.05c)$; $\theta = n \cdot \Delta t = 0.5c$. Значению $I(v)$ соответствует порядковый номер №($I(v)$). (а) Общее распределение эмпирических частот. (б) Блок распределения с индексом $\Psi = 0$ из групп с индексами $g = 0; 1; 2; 3; 4; 5; 6$. (в) Блок $\Psi = 1$, группы $g = 0; 1; 2; 3; 4; 5$. (г) Блок $\Psi = 2$, группы $g = 1; 2; 3; 4; 5$. (д) Блок $\Psi = 3$, группы $g = 0; 1; 2; 3; 4$.

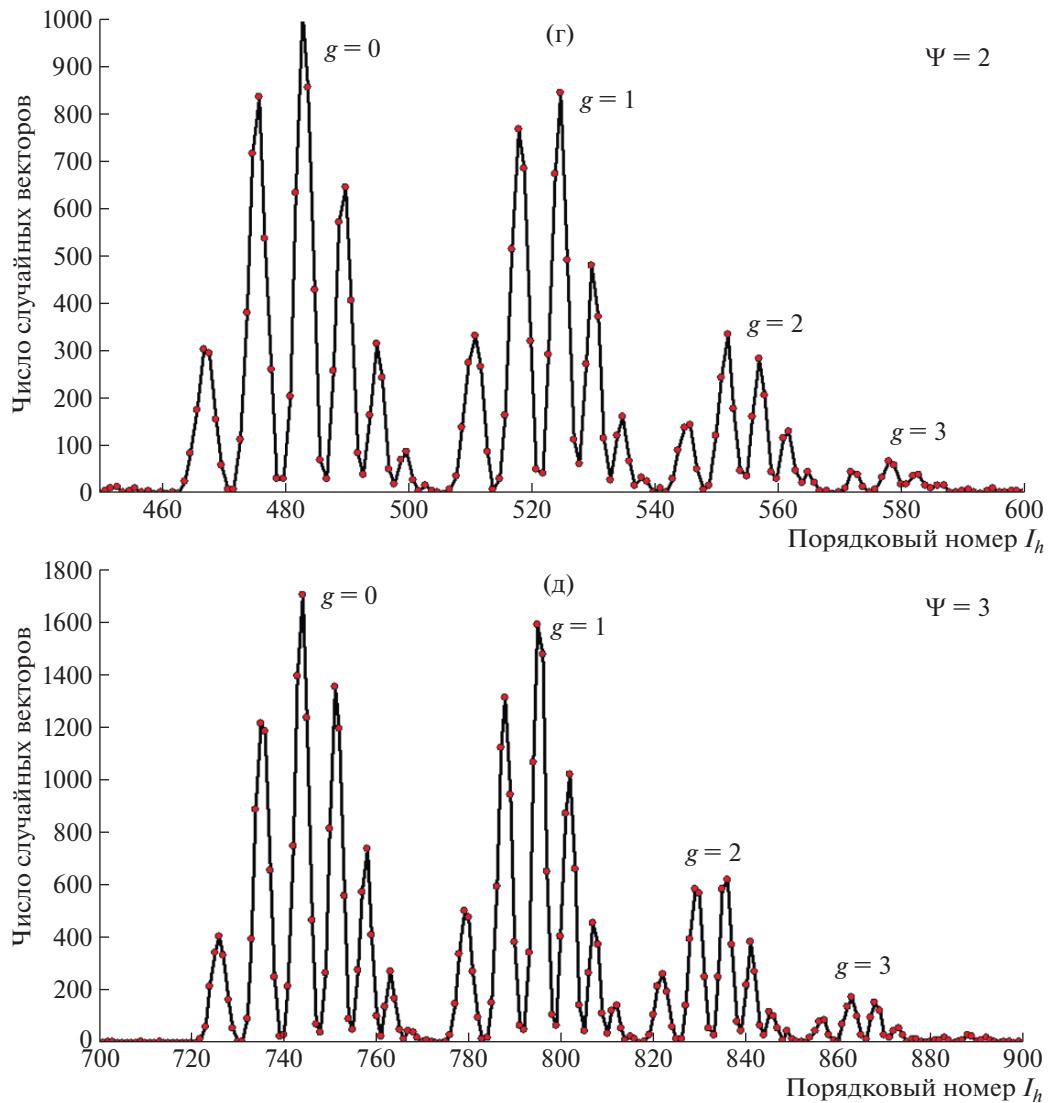


Рис. 1. Окончание

усеченные условные распределения относительных эмпирических частот $\tilde{C}'_m = (\tilde{C}'_{1,m}, \dots, \tilde{C}'_{q,m})$:

$$\tilde{C}'_{h,j,m} \tilde{C}_{h,j,m} \left/ \sum_{j=1}^{q_m} \tilde{C}_{h,j,m} \right. ; \quad \sum_{j=1}^{q_m} \tilde{C}'_{h,j,m} = 1; \quad (3)$$

$$q_m = v_{0,h,j,m} + v_{1,h,j,m} + 1$$

и соответствующих им векторов вероятностей:

$$P'_m = (P'_{1,m}, \dots, P'_{h,m}, \dots, P'_{q,m}).$$

Как известно [11], условное распределение некоторой части компонент полиномиального вектора при определенных значениях остальных его компонент также является полиномиальным рас-

пределением, которое зависит от фиксируемых компонент в виде их суммы, а полиномиальные вероятности могут быть представлены в виде произведения биномиальных вероятностей.

Поскольку каждый пик m как случайный вектор $C_m = (C_{1,m}, \dots, C_{q_m,m})$ является усеченным распределением частот $C_j = (I(\cdot)_h)_m$, образованным $q_m = v_{0,l,m} + 1$ случайными компонентами — векторами $v(\cdot)_{j,m}$ имеет полиномиальное распределение, то с учетом (1) и (2) и при условии $v_{2,l,m} + v_{l,h_m} = \text{const}_{2,l,m}$, справедливо условное распределение пика m как вектора $\tilde{v}'_m(\cdot) = \left(\tilde{v}'_{1,m}(\cdot), \dots, \tilde{v}'_{q_m,m}(\cdot) \right)$ в виде

$$\tilde{v}'_{h,m}(v(\cdot)_m) = \left(v_{h,m}(v_h(\cdot)_m) \right) \left/ \sum_{h=1}^{q_m} v_{h,m}(v(\cdot)) \right.,$$

где

$$\sum_{h=1}^{q_m} {}_{h,m}(\cdot) = \frac{h!}{v_{0,1,m}! \dots v_{l,1,m}} \times \times P_{2,m}^{v_{2,m}} \dots P_{l,m}^{v_{l,m}} (P_0 + P_1)^{v_{0,1,m} + v_{1,1,m}}. \quad (4)$$

$$\text{Поэтому } {}_{h,m}(v_m(\cdot)) = \binom{v_{0,1,m}}{h} P^{v_{0,1,m}-h} (1-P)^h,$$

$$h = 0, 1, \dots, v_{0,1,m}; \text{ где } P = P'_0 = P_0 / (P_0 + P_1); \quad P'_1 = P_1 / (P_0 + P_1), \sum_{h=1}^{q_m} {}_{h,m} = \left(P'_0 + P'_1 \right)^{v_{0,1,m} + v_{1,1,m}} = 1.$$

Таким образом, векторы ${}_{m}(\mathbf{v}(\cdot)_m) = (v_{1,m}, \dots, v_{q_m,m})$, представленные в форме (4), распределения

${}_{m}(\mathbf{v}(\cdot)_m) = (v'_{1,m}, \dots, v'_{q_m,m})$, являются биномиальными, которым соответствуют эмпирические векторы $\tilde{C}'_{j,m} = (\tilde{C}'_{1,1}, \dots, \tilde{C}'_{h,m}, \dots, \tilde{C}'_{q_m,m})$.

Векторы пиков m и r $C'_m = (C'_{1,m}, \dots, C'_{j,m}, \dots, C'_{l,m})$ и $C'_r = (C'_{1,r}, \dots, C'_{j,r}, \dots, C'_{l,r})$ с одинаковыми фиксированными значениями пар “соседних” компонент $v_{i,j,m} = v_{i,j,r}$, включая компоненты хвостов $v_{0,j}$, $v_{l,j}$, очевидно являются однородными соответствующими модели биномиального распределения $\text{Bi}(N, P)$ с параметрами $N = v_{0,1,m} + v_{1,1,m}$, $P = P'_0 = P_0 / (P_0 + P_1)$.

На основе соотношений (2, 3, 4), значений структурных параметров пиков h, m, g, Ψ , однозначно определяемых вектором a и величинами n, l, M , можно анализировать и прогнозировать характеристики случайных пиков r однородных с m , как эмпирических и биномиальных распределений C'_m и C'_r , включающих статистические незначимые компоненты ВСЧ $0 \leq C_r(I(v(\cdot)_{j,h})) \leq 5$, если в распределении $C(I(\cdot))$ имеются статистические значимые “опорные” пики $C_m(\cdot) = (C_{1,m}, \dots, C_{l,m})$ с $\sum_{h=1}^{q_m} C_{h,m,g,\Psi} \gg 1$. Отметим, если векторы $C'_m(\cdot)$ и $C'_r(\cdot)$ являются однородными распределениями, то их идентификаторы удовлетворяют соотношению

$$I(\cdot)_r = I_m(\cdot) + \Delta_\Psi (a_4 - a_3).$$

Отсутствующие компоненты $C_{h,r}$ предполагаемого прогнозируемого пика r можно оценить, если в пике r имеются $1 \leq R$ отличных от нуля компонент $0 < C_{k,r}$, а в соответствующем “близком” однородном опорном пике m известны все компоненты $C'_{h,m}$ и справедливо $C'_{h,m} = C'_{h,r}$. Прогнози-

руемое возможное среднее значение $\bar{C}_r = \sum_{i=1}^{q_m} C_{i,r}$

оценим $\bar{C}_r = \frac{1}{R} \sum_{i=1}^R \frac{\tilde{C}_{h,i,r}}{\tilde{C}'_{h,m}}$, а прогнозируемые значения $C_{h,r}$ компонент вектора r получим $C_{h,r} = \bar{C}_r C'_{j,m}$.

Если в пике r имеется только одна компонента $C_{h,r} \geq 0$, то

$$C_{h,r} = \frac{C_{i,r}}{\tilde{C}'_{i,m}} C'_{h,m}, \quad h = 1, \dots, q_m, \quad (5)$$

и абсолютная $\sigma(\cdot)$ и относительная погрешности $\delta(\cdot)$ в этом случае могут быть значительными

$$\sigma(C_{h,r}) = \frac{C'_{h,r}}{C'_{i,m}} \left(1 - C'_{i,m} \right)^{1/2} \sqrt{\tilde{C}_{i,r}}, \quad (6)$$

$$\delta(C_{h,r}) = \left(1 - C'_{i,m} \right)^{1/2} \sqrt{\tilde{C}_{i,r}}.$$

Отметим, что для прогноза характеристик статистически незначимых пиков и образующих их ВСЧ достаточно применять только способ опорных пиков, не анализируя соответствие распределений пиков $C'_m(\cdot)$ и $C'_r(\cdot)$ в частности биномиальной модели.

В табл. 2 приведены оценки значений $C'_{j,r}(\cdot)$, полученные по формулам (5), (6) и непосредственно из распределений $\tilde{C}(I(v(\cdot)))$. Таким образом, все пики, согласующиеся с полиномиальным распределением $(v(\cdot)) = (v_{1,m}, \dots, v_{h,m}, \dots, v_{q_m,m})$, будучи приведенными к биномиальной форме распределения $\text{Bi}(N, P)$, образуют совокупности из Q_i типов векторов, каждый из которых удовлетворяет условиям (3), (4) и $Q(\cdot) = \sum_{j=1}^R Q_j(\cdot)$, где величина $v_{i,j} + v_{i+1,j} = \text{const}_j$ определяет значения $Q_i(\cdot)$. Очевидно, прогнозируемые пики r относятся к одному типу $Q_i(\cdot)$.

Такой подход позволяет анализировать и прогнозировать образование и реализацию пиков как случайных однородных или близких векторов $C'_m(\cdot), C'_r(\cdot)$ в условиях малой статистики $\sum_{i=1}^q C_{i,r} < 50$ и при больших значениях $I(v(\cdot))$, когда растет вклад в хвосты распределений $C(I(v(\cdot)))$ ВСЧ- $(v_0(k=0), \dots, v_l(k=l))$ компонент с $k_i \geq \bar{k}$. Отметим, что значительные уклонения k могут быть или артефактами, или значимыми редкими событиями. Если такие $k \geq \bar{k}$ входят в компоненты ВСЧ в составе статистически значимых пиков $C(I(\cdot))$, то это позволяет счи-

Таблица 2. Прогноз частот С (ВСЧ) компонент пиков. “П” – прогнозируемый пик, “О” – опорный пик

Для выборки $\bar{k} = 1.1762; M = 1006458$						
Тип ВСЧ	3061	2161	1261	0361	О	
$\widetilde{cnt} = \tilde{C}$	187	665	810	297		
\tilde{C}'	0.095	0.339	0.413	0.152		
теор. '	0.097	0.342	0.403	0.158		
Тип ВСЧ	3007	2107	1207	0307		
$\widetilde{cnt} = \tilde{C}$		1	1		П	
прогноз С	0.24	0.85	1	0.39		
Тип ВСЧ	30061	21061	12061	03061	П	
$\widetilde{cnt} = \tilde{C}$		2	1			
прогноз С	0.24	0.85	1	0.39		
Тип ВСЧ	2062	1162	0262	Нет		
$\widetilde{cnt} = \tilde{C}$	70	195	109		О	
\tilde{C}'	0.187	0.521	0.291			
теор. '	0.211	0.497	0.292			
Тип ВСЧ	20152	11152	02152	Нет	П	
$\widetilde{cnt} = \tilde{C}$		1	1			
прогноз С	0.42	1.0	0.58			
Тип ВСЧ	2026	1126	0226	Нет		
$\widetilde{cnt} = \tilde{C}$	2	2	3		П	
прогноз С	1.7	3.9	2.3			
Тип ВСЧ	4051	3151	2251	1351	0451	О
$\widetilde{cnt} = \tilde{C}$	409	2117	3555	2787	834	
\tilde{C}'	0.042	0.218	0.366	0.287	0.086	
теор. '	0.044	0.208	0.368	0.287	0.092	
Тип ВСЧ	4006	3106	2206	1306	0406	П
$\widetilde{cnt} = \tilde{C}$		4	8	4		
прогноз С	0.57	3.49	5.86	4.59	1.38	
Тип ВСЧ	40042	31042	22042	13042	04042	П
$\widetilde{cnt} = \tilde{C}$		7	8	4		
прогноз С	0.8	4.1	7.0	5.4	1.6	
Для выборки $\bar{k} = 0.5881; M = 2012918$						
Тип ВСЧ	4060	3160	2260	1360	0460	О
$\widetilde{cnt} = \tilde{C}$	38	71	65	32	5	
\tilde{C}'	0.18	0.336	0.308	0.152	0.024	
теор. '	0.157	0.369	0.326	0.128	0.0188	
Тип ВСЧ	Нет	3124	2224	1324	0424	П
$\widetilde{cnt} = \tilde{C}$		1	1	2		
прогноз С	1.15	2.71	2.4	0.94	0.14	

Таблица 2. Продолжение

Для выборки $\bar{k} = 1.1738; M = 5904724$

Тип ВСЧ	6040	5140	4240	3340	2440	1540	0640	О
$\widetilde{cnt} = \tilde{C}$	2194	15409	46250	72090	63280	29988	5868	
\tilde{C}'	0.0093	0.0655	0.197	0.307	0.269	0.128	0.025	
теор. '	0.0095	0.0674	0.196	0.306	0.270	0.127	0.0248	
Тип ВСЧ	60201000	51201000	42201000	33201000	24201000	15201000	06201000	П
$\widetilde{cnt} = \tilde{C}$	1	1	1			2		
прогноз C	0.048	0.34	0.98	1.53	1.35	0.64	0.12	
Тип ВСЧ	60300000001	513000001	413000001	3330000001	243000001	153000001	06300001	
$\widetilde{cnt} = \tilde{C}$				1	2			П
прогноз C	0.014	0.20	0.59	0.92	0.81	0.38	0.074	

Для выборки $\bar{k} = 1.1738; M = 5904724$

Тип ВСЧ	5032	4132	3232	2332	1432	0532	О
$\widetilde{cnt} = \tilde{C}$	2739	16874	38955	45527	26511	6281	
\tilde{C}'	0.020	0.123	0.284	0.332	0.194	0.0459	
теор. '	0.0206	0.1209	0.2838	0.332	0.195	0.0459	
Тип ВСЧ	503000101	413000101	323000101	233000101	143000101	053000101	П
$\widetilde{cnt} = \tilde{C}$			2		1		
прогноз C	0.12	0.74	1.74	2.02	1.19	0.28	
Тип ВСЧ	502020001	412020001	322020001	232020001	142020001	052020001	
$\widetilde{cnt} = \tilde{C}$		1	1	1	2	2	П
прогноз C	0.14	0.85	2.0	2.32	1.36	0.32	

Для выборки $\bar{k} = 1.1808; M = 5904724$

Тип ВСЧ	7030	6130	5230	4330	343	2530	1630	0730	О
$\widetilde{cnt} = \tilde{C}$	1713	14876	52245	102956	121437	86163	33610	5790	
\tilde{C}'	4.090-3	3.552-2	1.248-1	2.458-1	2.900-1	2.057-1	8.025-2	1.38-2	
теор. '	4.264-3	3.525-2	1.249-1	2.457-1	2.902-1	2.056-1	8.053-2	1.365-2	
Тип ВСЧ	700020001	610020001	520020001	430020001	340020001	250020001	160020001	070020001	П
$\widetilde{cnt} = \tilde{C}$			1	1	2	1			
прогноз C	0.025	0.21	0.75	1.45	1.73	1.20	0.48	0.08	
Тип ВСЧ	70200000001	61200000001	52200000001	43200000001	34200000001	25200000001	16200000001	07200000001	
$\widetilde{cnt} = \tilde{C}$					1		1		П
прогноз C	0.015	0.13	0.47	0.92	2.72	0.797	0.57	0.31	

тать такие значения k большими уклонениями и не исключать их из анализа.

Отметим, что оценки близости пар пиков как векторов, один из которых опорный (обучающий вектор) проводился на основе тестовой статистики в виде функции центральных дробных моментов [5] на уровне значимости критерия 0.005–0.01.

Приложение

Взаимно-однозначное соответствие случайного вектора $v(\cdot) = (v_0, \dots, v_l)$ и его идентификатора $I(v(\cdot))$ при проектирующем векторе $a = (a, \dots, a_l)$.

Лемма. Пусть n и l – некоторые фиксированные натуральные числа, $n \leq l$ и (v_1, \dots, v_l) – некоторое решение уравнения $x_1 + x_2 + \dots + x_l = n$. Пусть m_1, m_2, \dots, m_l – натуральные числа, удовлетворяющие условиям $nm_1 < m_2; nm_1 + nm_2 < m_3; \dots; nm_1 + \dots + nm_{k>1} < m_k$. Обозначим

$$m_1 v_1 + m_2 v_2 + \dots + m_l v_l = I,$$

тогда система уравнений

$$x_1 + x_2 + \dots + x_l = n,$$

$$m_1 x_1 + m_2 x_2 + \dots + m_l x_l = I,$$

имеет единственное целочисленное неотрицательное решение (v_1, \dots, v_l) [5].

ЗАКЛЮЧЕНИЕ

Предложенный метод многомодального распределения идентификатора $C(I(\cdot))$ случайного вектора в виде линейной дискретной статистики $I(v(\cdot)) = a_0 v_0 + \dots + a_l v_l$ позволяет обрабатывать, анализировать и прогнозировать статистические характеристики больших упорядоченных по $I(v(\cdot))$ последовательностей $M(v_j(\cdot)) > 10^3$ различных типов $Q > 10^3$ случайных векторов $v(\cdot) = (v_0, \dots, v_l)$ для выборок малого объема $n \leq 10$ отсчетов $k = 0, 1, \dots, l$ при $\bar{k} < 5$.

Метод основан на анализе распределений $C(v(\cdot))$ типохарактеризующих устойчивых структурных характеристик пиков как случайных векторов $C_m(\cdot) = (C_{1,m}(\cdot), \dots, C_{l,m}(\cdot))$ с близкими распределениями. Однозначность локализации пиков по I_m и по расстояниям $\Delta I_{m,r} = I_r(\cdot) - I_m(\cdot)$ между пиками близкими по компонентам в различных группах и блоках позволяет анализировать последовательности пиков с однородными компонентами и прогнозировать статистически недостаточные составляющие $C(I(v(\cdot)_h))$.

Метод позволяет из случайной последовательности выборок $\{k_i\}_h, h = 1, M$ после их взаимно однозначного преобразования в последовательность значений их идентификаторов I_h “автоматически” при заданных параметрах n, l, M и векторе a получать распределения в виде пиков из близких типов ВСЧ, которые “автоматически” объединяются в группы и блоки.

СПИСОК ЛИТЕРАТУРЫ/REFERENCES

1. Риордан Дж. Введение в комбинаторный анализ. 1963. Москва: Изд. ин. лит-ры.
2. Айвазян С.А. и др. Прикладная статистика. Исследование зависимостей: Справ. изд. 1985. Москва: Финансы и статистика.
3. Петров А.А. // Теория вероятностей и ее применения. 1956. Т. 1 (2). С. 223–245.
4. Бабенко А.Г., Вахтель В.М., Муратов И.В., Работкин В.А. // Ядерн. физ. инжинир. 2022. Т. 13 (2). С. 200.
<https://doi.org/10.56304/S2079562922010055>
[Babenko A.G., VakhTEL V.M., Muratov I.V., Rabotkin V.A. // Phys. At. Nucl. 2021. V. 84 (12). P. 2041–2047. [https://doi.org/10.1134/S1063778821090076\].](https://doi.org/10.1134/S1063778821090076)
5. Blizniakov N.M., Vahtel V.M., Kostomakha D.E., Rabotkin V.A. // Voronezh Winter Math. School S.G. Crane. Mater. Int. Conf. 2022. P. 30–36.
6. Akindinova E.V., Babenko A.G., VakhTEL V.M., Evseev N.A., Rabotkin V.A., Kharitonova D.D. Proc. Int. Symp. Exotic Nuclei (EXON-2014). 2014. P. 651–658.
7. Кляцкин В.И. Стохастические уравнения и волны в случайно-неоднородных средах. 1980. Москва: Наука.
8. Большев Л.Н. // Теория вероятностей и ее применения. 1965. Т. 10 (3). С. 446–456.
9. Akindinova E.V., Babenko A.G., VakhTEL, Rabotkin V.A., Kharitonova D.D., Muratov I.V. // Voronezh Winter Math. School S.G. Crane. Mater. Int. Conf. 2018. P. 115–117.
10. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. 1965. Москва: Наука.
11. Ивченко Г.И., Медведев Ю.И. Дискретные распределения. Вероятностно-статистический справочник. Многомерные распределения. 2016. Москва: УРСС.
12. Потапов А.В., Чернявский А.Ф. Статистические методы измерений в экспериментальной ядерной физике. 1980. Москва: Атомиздат.
13. Rabotkin V.A., Bliznyakov N.M., VakhTEL V.M., Kostomakha D.E. // NUCLEUS-2022: Fundamental Problems and Applications. Book of Abstracts. 2022. Moscow. P. 302–304.

Method of Processing, Analyzing, and Predicting Characteristics of Radiation Fluxes from Small-Volume Samples

V. A. Rabotkin¹, N. M. Blizniakov¹, V. M. Vahtel¹, *, and D. E. Kostomakha¹

¹ Voronezh State University, Voronezh, 394018 Russia

*e-mail: vakhtel@phys.vsu.ru

Received May 21, 2024; revised May 21, 2024; accepted May 31, 2024

Abstract—A method has been proposed to process and analyze the proximity by type-defining characteristics of large aggregations $M > 10^4$ of various types $Q > 10^2$ of empirical discrete random frequency vectors $v(\cdot) = v_0, \dots, v_l$, obtained from samples of a small volume $10 \geq n = \sum_{i=1}^l v_i(k = i)$ of random counts $k = 0, 1, \dots, l$ with averages over all samples $\bar{k} < 5$. The method is based on a bijection between the random frequency vectors and its type-defining identifier $J(v, a) > 0$, which is a linear statistics in the form of the scalar product of v and the non-random frequency vector a . Discrete multimodal empirical distributions $C(J(v, a))$ representing sequences of arranged and grouped peaks enable us to analyze and forecast the characteristics of peaks and random frequency vectors forming them with low frequencies of their occurrences at the given M value.

Keywords: radiation, random frequency vector, sample, distribution, identifier, forecast