18. *Blair Jeffrey S.* The Biomedical Engineering handbook. CRC Press Taylor & Francis Group, New York, 2006, pp. 42-4–42-5.
19. *Deng Y., Guo M., Ramos A.F., et al.* Optimal low-latency network topologies for cluster performance enhancement, *J. Supercomput. Published online 02 March 2020*. Available at: https://doi.org/10.1007/s11227-020-03216-y.
20. *Hayes J.P., Mudge T.* Hypercube supercomputers, *Proceedings of the IEEE,* 1989, Vol. 77 (12), pp. 1829-1841.

Статью рекомендовал к опубликованию д.т.н., профессор Ю.М. Вишняков.

Духнич Евгений Иванович – Государственный Морской университет имени адмирала Ф.Ф. Ушакова; e-mail: evgenydukhnich@gmail.com; г. Новороссийск, Россия; тел.: +79184907411; кафедра радиоэлектроники и информационных технологий; д.т.н.; профессор.

Чефранов Александр Гергиевич – Восточный Средиземноморский университет; e-mail: alexander.chefranov@emu.edu.tr; г. Фамагуста, С. Кипр; тел.: +05338673790; кафедра компьютерных технологий; д.т.н.; профессор.

Dukhnich Evgeny Ivanovich – Novorossiysk State Maritime University; e-mail: evgenydukhnich@gmail.com; Novorossiysk, Russia; phone: +79184907411; the department of radio-electronics and information technologies; dr. of eng. sc.; professor.

Chefranov Alexander Georgievich – Eastern Mediterranean University, e-mail: alexander.chefranov@emu.edu.tr; t. Famagusta, N. Cyprus; phone: +05338673790; the department of computer engineering; dr. of eng. sc.; professor.

УДК 519.224.22 DOI 10.18522/2311-3103-2020-7-52-67

А.К. Мельников

АЛГОРИТМИЧЕСКАЯ СЛОЖНОСТЬ РАСЧЕТА ТОЧНЫХ ПРИБЛИЖЕНИЙ РАСПРЕДЕЛЕНИЙ ВЕРОЯТНОСТЕЙ ЗНАЧЕНИЙ СТАТИСТИК МЕТОДОМ РЕШЕНИЯ УРАВНЕНИЯ ПЕРВОЙ КРАТНОСТИ ТИПОВ

Рассматривается алгоритмическая сложность расчета точных распределений вероятностей значений статистик и их точных приближений методом решения уравнения первой кратности. В качестве точных приближений распределений вероятностей значений статистик рассматриваются их Δ–точные распределения, отличающиеся от точных распределений не более чем на заранее заданную, сколь угодно малую величину Δ. Показывается, что основой метода расчета точных распределений вероятностей значений статистик является перечисление элементов области поиска решений линейного уравнения кратности типов, составленной из векторов кратности типов, каждый элемент которого представляет собой число вхождений элементов определенного типа (какого-либо знака алфавита) в рассматриваемую выборку. Одновременно показывается, что для расчета точных приближений распределения вероятностей значений статистик применяется метод ограничения области поиска решений. Приводится выражение определяющее алгоритмическую сложность вычисления точных распределений методом решения уравнения первой кратности. Приведенное выражение является конечным и позволяет для каждого значения мощности алфавита определить максимальный объем выборки, для которой при использовании ограниченного вычислительного ресурса методом решения уравнения первой кратности могут быть рассчитаны точные распределения. Определена область параметров, представляемых объемом выборок и мощностью алфавита, для которых при ограниченном вычислительном ресурсе могут быть рассчитаны точные распределения. Для оценки алгоритмической сложности расчета точных приближений распределений приводится, впервые полученное, выражение для числа решений уравнения первой кратности с ограничением на значения координат векторов решений. Приводится выражение *определяющее алгоритмическую сложность вычисления точных приближений методом решения уравнения первой кратности с ограничением на значения координат векторов решений. В качестве параметра ограничения координат векторов решений используется значение статистики максимальной частоты, вероятность превышения которого меньше заранее заданной, сколь угодно малой величины* Δ*, что позволяет рассчитывать точные приближения распределений, отличающиеся от их точных распределений не более чем на выбранную величину* Δ*. Приведенное выражение является конечным и позволяет для каждого значения алфавита определить максимальный объем выборки, для которой при использовании ограниченного вычислительного ресурса методом решения уравнения первой кратности при ограничениях задаваемых с помощью величины* Δ *могут быть рассчитаны точные приближения. Приводятся результаты вычислений максимальных объемов выборок для которых могут быть рассчитаны точные приближения. Показывается, что алгоритмическая сложность расчета точных распределений на много порядков превосходит сложность расчета их точных приближений. Показано, что применение метода первой кратности для расчета точных приближений позволяет при одинаковых значениях мощности алфавита увеличить, по сравнению с расчетом точных распределений, объём выборок в два и более раз.*

Вероятность; статистика; точное распределение; точное приближение; вектор кратности типов; линейное уравнение; алгоритмическая сложность.

A.K. Melnikov

ALGORITHMIC COMPLEXITY OF CALCULATING EXACT APPROXIMATIONS OF PROBABILITY DISTRIBUTIONS OF STATISTICAL VALUES BY SOLVING THE EQUATION OF THE FIRST MULTIPLICITY OF TYPES

We consider the algorithmic complexity of calculating the exact probability distributions of statistical values and their exact approximations by solving the first multiplicity equation. As exact approximations of probability distributions of statistical values, we consider their $\Delta-$*exact distributions that differ from the exact distributions by no more than a predetermined, arbitrarily small value*Δ*. It is shown that the basis of the method for calculating the exact probability distributions of statistical values is the enumeration of elements of the search area for solutions to a linear equation of multiplicity of types, composed of vectors of multiplicity of types, each element of which is the number of occurrences of elements of a certain type (any sign of the alphabet) in the sample under consideration. At the same time, it is shown that the method of limiting the search area for solutions is used to calculate exact approximations of the probability distribution of statistical values. An expression is given that defines the algorithmic complexity of calculating exact distributions by solving the first multiplicity equation. The given expression is finite and allows for each value of the alphabet power to determine the maximum sample size for which, using a limited computational resource, exact distributions can be calculated by solving the first multiplicity equation. The range of parameters represented by the sample size and alphabet power for which exact distributions can be calculated with a limited computing resource is defined. To estimate the algorithmic complexity of calculating exact approximations of distributions, we present an expression for the first time obtained for the number of solutions to the equation of the first multiplicity with a restriction on the coordinate values of the solution vectors. An expression is given that defines the algorithmic complexity of calculating exact approximations by solving the first multiplicity equation with a restriction on the coordinate values of the solution vectors. As a parameter for limiting the coordinates of solution vectors, the maximum frequency statistic value is used, the probability of exceeding it is less than a pre-set, arbitrarily small value*Δ*, which allows calculating exact approximations of distributions that differ from their exact distributions by no more than the selected value* Δ*. The given expression is finite and allows for each value of the alphabet to determine the maximum sample size for which, when using a limited computational resource, exact approximations can be calculated by solving the equation of the first multiplicity under the restrictions set using the value*Δ*. The results of calculations of the maximum sample volumes for which exact approximations can be calculated are presented. It is shown that the algorithmic complexity of calculating exact distributions exceeds the complexity of calculating their exact ap-*

proximations by many orders of magnitude. It is shown that the use of the first multiplicity method for calculating exact approximations allows for the same values of the alphabet power to increase the sample volume by two or more times compared to the calculation of exact distributions.

Probability; statistics; exact distribution; an accurate approximation of the vector of multiplicity of types; the linear equation algorithmic complexity.

Введение. Применение в критериях согласия с равновероятным распределением в качестве эталонного распределения точных распределений вероятностей значений статистик при заданном уровне значимости позволяет строить критерии с наибольшей относительной эффективностью [1]. Однако расчет точных распределений для многих значений параметров мощности алфавита N и объема выборки n (длины последовательности) не всегда возможен [2, 3], что, для решения задачи по статистическому анализу последовательностей (текстов) [4], заставляет пользоваться точными или предельными приближениями данных распределений [5, 6]. Применение точных приближений (Δ-точных распределений [7]) предпочтительнее перед применением предельных приближений, так как дает возможность строить критерии с большей относительной эффективностью [8]. При заданном уровне значимости и невозможности рассчитать точное распределение большую эффективность дает обработка текстов, построенная на критерии использующем точное приближение эталонного распределения, чем применение предельного распределения . Рассмотрение относительной алгоритмической сложности методов расчета точных распределений и их точных приближений показывает, что расчет точных приближений во много порядков проще расчета точных распределений [9]. Но для расчета значений параметров мощности алфавита и объема выборки, для которых могут быть рассчитаны точные приближения, учитывая используемый вычислительный ресурс, необходимо уметь рассчитывать алгоритмическую сложность методов расчета точных распределений и их точных приближений. Одним из таких методов [10] является метод решения уравнения первой кратности типов [11]. Данная работа посвящена оценке алгоритмической сложности расчета точных распределений вероятностей значений статистик и их точных приближений методом решения уравнения первой кратности типов.

Постановка задачи. Рассмотрим статистику $S(N,n) = f(h_1, ..., h_N)$, где h_i – частота встречаемости знака (исхода) a_i , n – длина текста (объем выборки), N – число исходов полиномиальной схемы (мощность алфавита $A_N=\{a_1, a_2, \dots , a_N\}$) и p_i – вероятность a_i -го исхода.

Одним из методов расчета точных значений распределения вероятности значений статистики $S(N,n)$ – $P\{ S(N,n) \geq c \}$ является расчет её значений для всех решений $(h_1^{(\nu)}, h_2^{(\nu)}, \dots, h_N^{(\nu)})$ в неотрицательных целых числах линейного уравнения

$$h_1 + \dots + h_N = n, \tag{1}$$

называемого ещё уравнением кратности типов или уравнением первой кратности типов. Поэтому данный метод справедливо можно назвать методом решения уравнения первой кратности типов или сокращенно *методом первой кратности* (МПК).

Решения уравнения (1) находятся путем перебора в лексикографическом порядке области поиска решений $R_{n,n}^N = \{h_i \mid i = \overline{1, N},\ h_i \in \mathrm{N},\ 0 \leq h_i \leq n\}$. При этом для расчета точных распределений $P_T\{ S(N,n) \geq c \}$ перебор координат вектора $(h_1^{(\nu)}, h_2^{(\nu)}, \dots, h_N^{(\nu)})$ возможных решений производится в условии

$$\{h_i \mid i=\overline{1,N},\ h_i \in \mathrm{N},\ 0 \le h_i \le n\}\,. \tag{1.1}$$

Основой методики расчета точных приближений $P_\Delta\ \{\ S(N,n) \ge c\ \}$ [12] является ограничение области поиска решений до

$$R_{n,r}^{N} = \{h_i \mid i=\overline{1,N},\ h_i \in \mathrm{N},\ 0 \le h_i \le r\},$$

где $r < n$ является параметром ограничения. Соответственно для расчета точных приближений перебор координат вектора $(h_1^{(\nu)}, h_2^{(\nu)}, \ldots, h_N^{(\nu)})$ возможных решений уравнения (1) производится в условии

$$\{h_i \mid i=\overline{1,N},\ h_i \in \mathrm{N},\ 0 \le h_i \le r\}\,. \tag{1.2}$$

Для оценка и сравнение между собой сложностей вычисления, числа обобщенных операций (алгоритмической сложности), точного распределения $P_{ext}\ \{\ S(N,n) \ge c\ \}$ – $C(P_{ext}\ \{\ S(N,n) \ge c\ \})$ и его точного приближения $P_\Delta\ \{\ S(N,n) \ge c\ \}$ – $C(P_\Delta\ \{\ S(N,n) \ge c\ \})$ необходимо получить их аналитические выражения.

Алгоритмическая сложность расчета точных распределений вероятностей значений статистик методом решения уравнения первой кратности типов. Количество векторов $(h_1^{(\nu)}, h_2^{(\nu)}, \ldots, h_N^{(\nu)})$ – $K_h(N,n,n)$, равное равно числу целочисленных неотрицательных решений уравнения (1) в условиях (1.1), равное числу сочетаний с повторениями $\overline{C}_N^{\,n}$, связанному с числом сочетаний следующим соотношением [11] (стр. 127)

$$K_h(N,n,n) = \overline{C}_N^{\,n} = C_{N+n-1}^{n} = \frac{(N+n-1)!}{n!\cdot(N-1)!}\,. \tag{2}$$

Тогда алгоритмическая сложность расчета точных распределений методом первой кратности может быть выражена следующим образом

$$C_{МПК}(P_T\{S_{N,n} \ge c\}) = N^n \cdot C(N,n,n,\overrightarrow{h(N,n)}^{(\nu)}) + \tag{3}$$

$$+\,C_{N+n-1}^{n} \cdot (C(S_{N,n}(\overrightarrow{h(N,n)})) + C(P(S_{N,n} = S_{N,n}(\overrightarrow{h(N,n)}^{(i)})) +$$

$$+\,C(P\{S_{N,n} \ge c\}, \{S_{N,n}^{(i)}, P_{N,n}^{(i)} \mid i=1,\ldots,C_{N+n-1}^{n}\}),$$

где $C(N,n,n,\overrightarrow{h(N,n)}^{(\nu)})$ есть алгоритмическая сложность генерации очередного, отличающегося от уже сгенерированных до этого, вектора частот $\overrightarrow{h(N,n)}$ без ограничения на координаты (N,n,n) и его проверки на удовлетворение уравнению (1), а $C(S_{N,n}(h(N,\ n))$ есть алгоритмическая сложность вычисления значения статистики $S_{N,n}$ от вектора первой кратности типов $\overrightarrow{h(N,n)}^{(i)}$ от первой (частотной) маркировки, $C(P(S_{N,n} = S_{N,n}(\overrightarrow{h(N,n)}^{(i)})$ сложность вычисление вероятности, с которой статистика $S_{N,n}$ примет значений $S_{N,n} = S_{N,n}(\overrightarrow{h(N,n)}^{(i)})$

$$P_{N,n}^{(i)} = P(S_{N,n} = S_{N,n}(\overrightarrow{h(N,n)}^{(i)}) = \frac{n!}{h_1^{(i)}!\cdot h_2^{(i)}!\cdot\ldots\cdot h_N^{(i)}}$$

и $C(P\{S_{N,n} \geq c\},\{S_{h(N,n)}^{(i)}, P_{h(N,n)}^{S(i)} \mid i=1,\ldots,C_{N+n-1}^{n}\})$ есть алгоритмическая сложность расчета распределения вероятностей $P_T\{S_{N,n} \geq c\}$ (таблицы распределения) используя рассчитанные значения статистики и их вероятности $\{S_{h(N,n)}^{(i)}, P_{h(N,n)}^{S(i)} \mid i=1,\ldots,C_{N+n-1}^{n}\})$.

Примем следующие ограничения. Будем считать, что:

♦ для генерации в лексикографическом порядке очередного, отличающегося от уже сгенерированных до этого векторов, вектора частот $\overrightarrow{h(N,n)}$ длины *N*, у которого каждая координата изменяется от 0 до *n*, без ограничения на координаты (N,n,n), и его проверки на удовлетворение уравнению (1) потребуется не более 3*N* операций

$$C(N,n,n,\overrightarrow{h(N,n)}^{(v)}) \leq 3N;$$

♦ не нарушая общности можно рассмотреть в качестве статистики $S_{N,n}$ статистику хи-квадрат $f=\chi_{N,n}$, имеющую вид [13]

$$\chi_{N,n} = \chi_n(h_1,h_2,\ldots,h_N) = \sum_{i=1}^{N} \frac{\left(h_i - np_i\right)^2}{np_i}.$$

Тогда сложность расчета значения статистики $S_{N,n} = S_{N,n}(\overrightarrow{h(N,n)}^{(i)})$ от вектора первой кратности типов $\overrightarrow{h(N,n)}^{(i)}$ и записи её значения в результирующий массив $R_{h(N,n)}[1:K_h(N,n,n)]$ потребуется не более 6*N* операций

$$C(S_{N,n}(\overrightarrow{h(N,n)})) \leq 6N;$$

♦ для вычисления вероятности $P_{h(N,n)}^{S(i)}$, с которым статистика $S_{h(N,n)}$ примет значение $S_{N,n}(\overrightarrow{h(N,n)}^{(i)})$ и записи её значения в результирующий массив потребуется не более 3*N* операций с учетом того, что значения *n*! и N^n будут вычислены заранее, а значения частот $h_j^{(i)}$ распределены равномерно и можно считать, что $h_j^{(i)} \approx n/N$.

Будем считать, что для вычисления $h_j^{(i)}!$ в условиях $h_j^{(i)} \approx n/N$ требуется не более *n*/*N* операций. Тогда для вычисления всех *N* значений $\{h_j^{(i)}! \mid j=\overline{1,N}\}$ с учетом того, что все $h_j^{(i)}$ удовлетворяют условию (1) потребуется не менее (*N*+*n*) операций, а для вычисления их произведения $h_1^{(i)}! \cdot h_2^{(i)}! \cdot \ldots \cdot h_N^{(i)}!$ потребуется еще не более *N* операций. Следовательно для вычисления $P_{h(N,n)}^{S(i)}$ в соответствии с формулой (3) при заранее рассчитанных значениях $n!$ и $1/N^n$ потребуется не более (2*N*+*n*+3) операций

$$C(P(S_{h(N,n)} = S_{N,n}(\overrightarrow{h(N,n)}^{(i)})) \le 2 \cdot N + n + 3 \cdot$$

♦ для вычисления распределения $P\{S_{N,n} \ge c\}$ по массиву $R_{h(N,n)}[1:K_h(N,n,n)]$ необходимо вначале его отсортировать, а потом однопроходным суммированием по формуле

$$P_T(S_{N,n} \ge c_i) = 1 - \sum_{j=1}^{K_h(N,n,n)} (P^{S(j)}_{h(N,n)}, \text{ }если\text{ } S^{(j)}_{h(N,n)} < c_i)$$

получить все составляющие $P_T(S_{N,n} \ge c_i)$. Для сортировки методом Шелла [14] по возрастанию массива $R_{h(n,N)}$, состоящего из $K_h(N,n,n)$ двойных элементов $(S^{(i)}_{h(N,n)}, P^{S(i)}_{h(N,n)})$ по ключу $S^{(i)}_{h(N,n)}$ необходимо порядка $2 \times K_h(N,n,n) \cdot \log_2 K_h(N,n,n)$ операций. Для однопроходного просмотра отсортированного массива $R_{h(n,N)}$ на сравнение $S^{(i)}_{h(N,n)}$ с c_j и при необходимости суммирования и получения всех $P_T(S_{N,n} \ge c_j)$ для $c_j = \overline{1{,}100}$ необходимо $2 \times K_h(N,n,n)$ операций.

Теперь в соответствии с принятыми ограничениями, выясненным фактом $K_h(N,n,n) = C^n_{N+n-1}$ и (3) поучаем значение для ограничения алгоритмической сложности метода первой кратности

$$C_{МПК}(P_T\{S_{N,n} \ge c\}) \le (n+1)^N \cdot 3N + C^n_{N+n-1} \cdot (6 \cdot N + 2 \cdot N + n + 3) +$$
$$+ 2 \times C^n_{N+n-1} \cdot \log_2 C^n_{N+n-1} + 2 \times C^n_{N+n-1} =$$
$$= (n+1)^N \cdot 3N + C^n_{N+n-1} \cdot (8 \cdot N + n + 2 \cdot \log_2 C^n_{N+n-1} + 5) \cdot \qquad (4)$$

Выражение (4) позволяет определить алгоритмическую сложность вычисления точных распределений вероятностей значений статистик методом решения уравнения первой кратности для любых значений длин последовательностей n и мощностей алфавита N.

Так же выражение (4) может быть использовано для определения значений параметров распределений при фиксированном вычислительном ресурсе, выделенном для его расчета.

Оценка максимальных значений параметров выборок для расчета точных распределений на современной вычислительной технике. Пусть $P_{МВС}$ доступная производительность в операциях в секунду многопроцессорной вычислительной системы (МВС) и t предоставляемое время в секундах её использования для поведения расчетов. Тогда предоставляемый вычислительный ресурс $R_{МВС}$ равен

$$R_{МВС} = P_{МВС} \cdot t$$

и значения параметров N и n вычисляются из соотношений

$$C_{МПК}(P_T\{S_{N,n} \ge c\}) \le R_{МВС} \cdot \qquad (5)$$

Анализ 55-й редакции от 22 июня 2020 года списка Top500 [15] 500 наиболее мощных компьютеров мира показывает, что пиковая производительность находящегося на первом месте списка японского суперкомпьютера Fugaku производства фирмы Fujitsu на базе процессоров ARM A64FX, установленного в RIKEN Center for Computational Science (R-CCS), равна 513.9 PFlop/s, (0,514 x 10^{18} оп./сек.) а производительность на тесте Linpack - 415.5 PFlop/s (0,41614 x 10^{18} оп./сек.).

Следовательно предположение, что нам может быть доступна МВС производительностью 10^{18} операций в секунду обоснован. Это могут быть МВС как кластерного типа, основанные на взаимодействии универсальных процессоров с классической фон-неймановской архитектурой [16, 17] (предложенной ещё в 1945 году Джо фон Нейманом), так и реконфигурируемые [18], основанные на программируемых логических схемах (ПЛИС). Пусть этот доступ может осуществляться в течении 1 (одного) месяца (2 592 000 секунды), тогда предоставленный вычислительный ресурс МВС R_{MBC} имеет значение

$$R_{MBC} = 2{,}592 \cdot 10^{24}$$

и параметры N и n, для которых могут быть вычислены точные распределения вычисляются из соотношений

$$C_{МПК}(P_T\{S_{N,n} \geq c\}) \leq 2{,}592 \cdot 10^{24}.$$

Значения параметров N и n частично приведены в таблице 1, область их значений показана на рис. 1.

Таблица 1

Параметры N и n, для которых при использовании МВС производительностью 10^{18} оп/сек. за один месяц методом первой кратности могут быть рассчитаны точные распределения (BP=2,592 x 10^{24})

№ п/п	Параметры		$C_{МПК}(P_T\{S_{N,n}\geq c\})$	$C_{МПК}(P_T\{S_{N,n+1}\geq c\})$
	N	n		
1.	12	79	2,4739E+24	2,87159E+24
2.	14	41	2,23223E+24	3,10319E+24
3.	16	25	2,09322E+24	3,82879E+24
4.	18	19	1,41558E+25	3,40676E+25
5.	20	12	1,1403E+24	5,0201E+24
6.	22	9	6,60000E+23	5,37258E+24
7.	24	7	3,4001E+23	5,74318E+24
8.	26	6	7,32223E+23	2,35741E+25
9.	28	5	2,35741E+25	3,86389E+25
10.	30	4	8,3819E+22	1,98967E+25
11.	32	4	2,23517E+24	7,64031E+26
12.	34	3	3,01051E+22	5,93718E+25
13.	36	3	5,10016E+23	1,57161E+27
14.	38	2	1,53997E+20	8,6136E+24
15.	40	2	1,45892E+21	1,45071E+26
16.	46	2	1,22309E+24	6,83343E+29
17.	48	1	4,05324E+16	1,14864E+25
18.	50	1	1,68885E+17	1,07685E+26
19.	58	1	5,01521E+19	8,19562E+29
20.	64	1	3,54177E+21	6,59267E+32
21.	68	1	6,02102E+22	5,67382E+34
22.	72	1	1,02003E+24	4,86613E+36
23.	74	0	4,19346E+24	4,50117E+37

Рис. 1. Область параметров выборок, для которых методом первой кратности могут быть посчитаны точные распределения

Анализ значений параметров N и n, приведенных в табл. 1 и области их изменений, показанных на рис. 1, говорит о том, для значений параметров N>46, метод первой кратности не позволяет рассчитывать точные распределения вероятностей значений статистик. Поэтому перейдем к рассмотрению алгоритмической сложности расчета точных приближений распределения вероятностей значений статистик методом первой кратности, в качестве которых рассматриваются Δ-точные распределения [7].

Алгоритмическая сложность расчета точных приближений распределения методом решения уравнения первой кратности. При ограничении области поиска решений уравнения (1) до

$$R_{n,r}^{N} = \{h_i \mid i = \overline{1, N},\ h_i \in \mathrm{N},\ 0 \le h_i \le r\},$$

где $r < n$ возникает вопрос об оценке числа этих решений $K_h(N,n,r)$.

Опираясь на результаты Сачкова В.Н., определяющие число композиций [19] (выражение 3.23 стр. 215) и используя известный метод установления взаимно однозначного соответствия, получаем выражение определяющее число целочисленных неотрицательных решений уравнения (1) в условиях (1.2), равное

$$K_h(N,n,r) = \sum_{i=0}^{\min(N,\frac{n}{r+1})} (-1)^i \binom{N}{i} \binom{N+n-1-(r+1)i}{N-1}. \tag{6}$$

Обращаем внимание, что выражение (6), в прямой постановке, получено автором впервые. Опираясь на результат (6) получаем выражение для алгоритмической сложности расчета точных приближений распределений методом МПК с ограничениями r.

$$\begin{gathered} C_{МПК}(P_\Delta\{S_{N,n} \ge c\}) = (r+1)^N \cdot C(N,n,r,\overrightarrow{h(N,n)}^{(\nu)}) + \\ + \cdot K_h(N,n,r) \cdot (C(S_{N,n}(\overrightarrow{h(N,n)}^{(i)}) + C(P(S_{N,n} = S_{N,n}(\overrightarrow{h(N,n)}^{(i)}))) + \\ + C(P\{S_{N,n} \ge c\}, \{S_{N,n}^{(i)}, P_{N,n}^{(i)} \mid i = 1,\dots,K_h(N,n,r)\}) \end{gathered} \tag{7}$$

где $\cdot C(N,n,r,\overrightarrow{h(N,n)}^{(\nu)})$ есть алгоритмическая сложность генерации очередного, отличающегося от уже сгенерированных до этого, вектора частот $\overrightarrow{h(N,n)}$ с ограничениями на координаты (N,n,r) и его проверки на удовлетворение уравнению (1), а остальные обозначения аналогичны обозначениям используемым в (3).

Как было показано в [7], для того чтобы точное приближение отличалось от точного значения распределения не более чем на сколь угодно малую величину Δ

$$| P_T\{S_{N,n} \ge c\} - P_T\{S_{N,n} \ge c\} | \le \Delta \tag{8}$$

необходимо, чтобы $r = m(N,n,\Delta)$, удовлетворяло условию

$$P\{M_n > m(N,n,\Delta)\} \le 1 - \Delta,$$

где M_n статистика максимальной частоты, а рекуррентная формула для вычисления значения её вероятности приведена в [8]. Тогда используя (7) и принятые ограничения получаем аналитическое выражение для $C_{МПК}(P_\Delta\{S_{N,n} \ge c\})$ алгоритмической сложности вычисления точного приближения распределений $P_\Delta\{S_{N,n} \ge c\}$ методом МПК

$$C_{МПК}(P_\Delta\{S_{N,n} \ge c\}) = (r+1)^N \cdot 3N + \tag{9}$$

$$+ K_h(N,n,r)\cdot(8\cdot N + n + 2\cdot \log_2 K_h(N,n,r) + 5)\cdot$$

Учитывая то, что для выполнения условия (8) r обязательно должно выбираться из условия $r = m(N,n,\Delta)$ выражение (9) должно быть преобразовано в выражение

$$C_{МПК}(P_\Delta\{S_{N,n} \ge c\}) = (m(N,n,\Delta)+1)^N \cdot 3N + \tag{10}$$

$$+ K_h(N,n,m(N,n,\Delta)\cdot(8\cdot N + n + 2\cdot \log_2 K_h(N,n,m(N,n,\Delta)) + 5)\cdot$$

Полученное выражение (10), определяет алгоритмическую сложность расчета точных приближений распределений вероятностей значений статистик методом первой кратности, конечно и может быть использовано для вычисления их величины при любых конкретных значениях параметров N и n и задаваемой пользователем метода точности Δ.

Для вычисления значений алгоритмической сложности $C_{МПК}(P_\Delta\{S_{N,n} \ge c\})$ также необходимо знать значения $m(N,n,\Delta)$, которые были рассчитаны с помощью программы для ПЭВМ [20], результаты докладывались в [21] и частично приведены в таблице 2 и на рис. 2.

Таблица 2

Значения *m* статистики максимальной частоты M_n, принимаемые с вероятностью $P\{M_n > m\} < \Delta$ для $\Delta = 10^{-5}$

		$n = 5N$ объём выборки (длина текста)													
		10	**15**	**50**	**100**	**150**	**180**	**200**	**250**	**300**	**350**	**400**	**450**	**500**	**1280**
N мощность алфавита	**2**	10													
	3	10	14												
	8	9	11												
	10	8	10	18											
	26	7	8	12	17										
	32	6	7	12	16	19									
	36	6	7	11	15	18	20								
	64	6	6	9	12	14	16	18	20						
	128	5	6	8	10	11	13	14	15	16	17	18	19	19	
	256	5	5	7	8	9	10	11	12	12	13	14	14	14	22

На рис. 2 визуально показано отличие значений n в выражении (4) и значений $m(N,n,\Delta)$ в выражении (10), определяющих основное различие в значениях алгоритмической сложности расчета точных распределений и их точных приближений методом первой кратности.

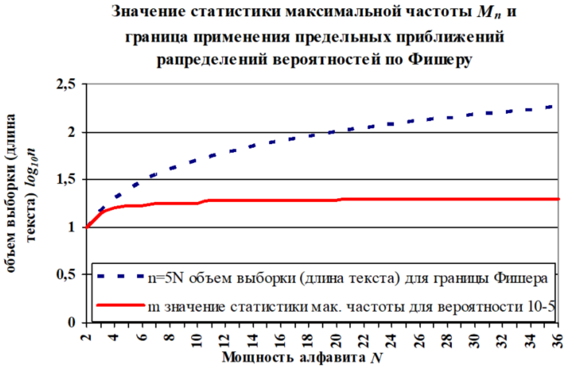

Рис. 2. График значений статистики максимальной частоты для параметров мощности алфавита и объема выборки на границе Фишера

В соответствии со значениями $m(N,n,\Delta)$ и (10) из соотношения

$$C_{МПК}(P_{\Delta}\{S_{N,n} \geq c\}) \leq 2{,}592 \cdot 10^{24}$$

были рассчитаны значения параметров N и n, для которых могут быть вычислены точные приближения распределений. Эти значения частично приведены в таблице 3, область их значений показана на рис. 3.

Таблица 3

Параметры N и n, для которых при использовании МВС производительностью 10^{18} оп/сек. за один месяц методом первой кратности могут быть рассчитаны точные приближения распределения (ВР=2,592 x 10^{24})

№ п/п	Параметры		$C_{МПК}(P_{\Delta}\{S_{N,n}\geq c\})$	$C_{МПК}(P_{\Delta}\{S_{N,n+1}\geq c\})$
	N	n		
1.	18	80	2,12470604E+24	5,62287692E+24
2.	20	41	1,14029783E+24	5,02009533E+24
3.	22	24	6,60000000E+23	5,37258146E+24
4.	24	14	3,40010387E+23	5,74318390E+24
5.	26	10	7,32223466E+23	2,35740535E+25
6.	28	7	5,15839146E+23	3,86388691E+25
7.	30	6	1,98966528E+25	2,02854063E+27
8.	32	5	2,23517418E+24	7,64031467E+26
9.	34	3	3,01050863E+22	5,93718141E+25
10.	36	3	5,10015580E+23	1,57160684E+27
11.	38	2	1,53997096E+20	8,61359646E+24
12.	40	2	1,45891986E+21	1,45071098E+26
13.	42	2	1,37867926E+22	2,43719445E+27
14.	46	2	1,22308546E+24	6,83342902E+29
15.	48	1	4,05323966E+16	1,14863678E+25
16.	50	1	1,68884986E+17	1,07684698E+26
17.	66	1	1,46098213E+22	6,11882457E+33
18.	72	1	1,02003116E+24	4,86613430E+36
19.	74	0	4,19346144E+24	4,50117423E+37

Рис. 3. Область параметров выборок, для которых методом первой кратности могут быть посчитаны точные приближения распределений

Анализ значений параметров N и n, приведенных в табл. 3 и области их изменений, показанных на рис. 3, говорит о том, для значений параметров N>46, метод первой кратности не позволяет рассчитывать точные приближения распределений вероятностей значений статистик.

Перейдем к сравнению алгоритмических сложностей расчета точных распределений и их точных приближений производимых методом первой кратности.

Оценка применения метода первой кратности для расчета точных распределений и их точных приближений. Для сравнения результатов применения метода первой кратности к расчету точных распределений и их точных приближений сравним полученные выражения для алгоритмических сложностей (4) и (10), для этого оценим их разность

$$C_{МПК}(P_T\{S_{N,n}\geq c\})-C_{МПК}(P_\Delta\{S_{N,n}\geq c\})=$$
$$=(n+1)^N\cdot 3N+C^n_{N+n-1}\cdot(8\cdot N+n+2\cdot\log_2 C^n_{N+n-1}+5)-$$
$$-(m(N,n,\Delta)+1)^N\cdot 3N-$$
$$-K_h(N,n,m(N,n,\Delta)\cdot(8\cdot N+n+2\cdot\log_2 K_h(N,n,m(N,n,\Delta))+5)\cdot$$

Учитывая факт того, что $K_h(N,n,m(N,n,\Delta))<\cdot K_h(N,n,n)$, рассматриваемая разность будет больше или равна при замене $K_h(N,n,m(N,n,\Delta))$ на $K_h(N,n,n)=C^n_{N+n-1}$. Тогда

$$C_{МПК}(P_T\{S_{N,n}\geq c\})-C_{МПК}(P_\Delta\{S_{N,n}\geq c\})\geq$$
$$\geq 3N\cdot(\,(n+1)^N-(m(N,n,\Delta)+1)^N\,)$$

и учитывая результаты значений $m(N,n,\Delta)$, полученных с помощью программы для ПЭВМ [10] и частично приведенных в табл. 3 видно что,

$$m(N,n,\Delta)<<n$$

и

$$C_{МПК}(P_T\{S_{N,n}\geq c\})-C_{МПК}(P_\Delta\{S_{N,n}\geq c\})>>0\,.$$

Оценка разности алгоритмических сложностей вычислений методом первой кратности точных распределений и их точных приближений показывает, что для большинства значений параметров распределений N и n сложность вычисления точных приближений намного меньше сложности вычисления самих точных распределений. Относительная алгоритмическая сложность вычислений точных распределений и их точных приближений рассматривалась в работах [4, 11], где показано, что вычисление точных приближений во много порядков проще расчета самих точных распределений.

Для проведения анализа результатов применения метода МПК к расчету точных распределений и их точных приближений данные о максимальных значениях параметров N и n, из табл. 1 и 2 были сведены в сравнительную табл. 4.

Таблица 4

Сравнительная таблица максимальных значений параметров *N* и *n*, для которых при использовании МВС производительностью 10^{18} оп/сек. за один месяц (ВР=2,592 х 10^{24}) методом первой кратности могут быть рассчитаны точные распределения и их точные приближения

№ п/п	Мощность алфавита N	Максимальный объем выборки n для которой при ограниченном ВР методом МПК рассчитывается	
		точное приближение $P_{\Delta}\{S_{N,n}\geq c\}$- $n_{\Delta}(N)$	точное распределение $P_{T}\{S_{N,n}\geq c\}\}$ - $n_{T}(N)$
1.	18	80	19
2.	20	41	12
3.	22	24	9
4.	24	14	7
5.	26	10	6
6.	28	7	5
7.	30	6	4
8.	32	5	4
9.	34	3	3
10.	36	3	3
11.	38	2	2
12.	40	2	2
13.	42	2	2
14.	44	2	2
15.	46	2	2
16.	48	1	1
17.	72	1	1

Области расположения параметров выборок, для которых при ограниченном вычислительном ресурсе ВР методом первой кратности могут быть рассчитаны точные распределения и их точные приближения показаны на рис. 4.

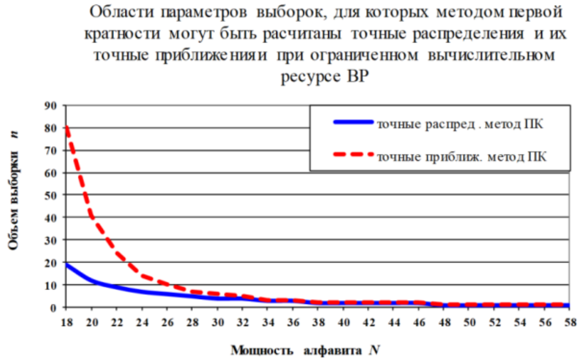

Рис. 4. Область параметров выборок, для которых методом первой кратности могут быть рассчитаны точные распределения и их точные приближения

Анализ данных таблицы 4, представленных на рис. 4 показывает, что при одинаковом выделенном вычислительном ресурсе объемом в $2{,}592 \cdot 10^{24}$ операций:

♦ методом первой кратности могут быть рассчитаны точные распределения и их точные приближения для мощности алфавита $N \leq 46$, при этом объём выборок $n \geq 2$;

♦ применение метода первой кратности не позволяет рассчитывать точные распределения и их точных приближения для значений мощности алфавита $N > 46$, при этом объем выборок $n \leq 1$;

♦ применение метода первой кратности для расчета точных приближений распределений позволяет при одинаковых значениях мощности алфавита N увеличить, по сравнению с расчетом точных распределений, объём выборок n, для которых могут быть рассчитаны точные приближения распределений;

♦ для мощности алфавита $N \leq 46$ объём выборок $n_{\Delta}(N)$, для которых могут быть рассчитаны точные приближения распределений, для каждого значения N превышает объём выборок $n_T(N)$, для которых могут быть рассчитаны точные распределения $n_{\Delta}(N) > n_T(N)$;

♦ для мощности алфавита $N \leq 24$ объём выборок $n_{\Delta}(N)$, для которых могут быть рассчитаны точные приближения распределений, в 2 и более раз превышает объём выборок $n_T(N)$, для которых могут быть рассчитаны точные распределения $(n_{\Delta}(N)/n_T(N)) \geq 2$.

Заключение и выводы. В работе приведены аналитические выражения позволяющие для любых значений параметров распределений мощности алфавита N и объёма выборки n рассчитывать алгоритмическую сложность вычисления точных распределений вероятностей значений статистик и их точных приближений методом решения уравнения первой кратности, учитывающие не только область перебора возможных решений но и число этих решений. В качестве точных приближений распределений рассматриваются Δ–точные распределения, отличающиеся от точных распределений не более чем на заданную заранее величину Δ.

Анализ аналитических выражений, полученных для алгоритмических сложностей расчета распределений, показал, что сложность вычисления точных приближений распределений на много меньше сложности вычисления самих точных распределений и позволил провести вычисление максимальных значений параметров распределений N и n, для которых при фиксированном вычислительном ресурсе, ограниченным современным уровнем развития вычислительных средств, методом решения уравнения первой кратности могут быть рассчитаны точные распределения и их точные приближения.

Анализ вычисленных максимальных значений параметров распределений показал, что методом решения уравнения первой кратности нельзя вычислить точные распределения и их точные приближения для мощностей алфавита N превышающих 46.

Сравнительный анализ вычисленных максимальных значений параметров распределений показал, что применение метода первой кратности для расчета точных приближений распределений позволяет при одинаковых значениях мощности алфавита N увеличить, по сравнению с расчетом точных распределений, объём выборок n, для которых могут быть рассчитаны точные приближения распределений, так для мощностей алфавита меньше 24 это увеличение составляет два и более раз.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Мельников А.К.* Сравнение эффективности обработки текстов при применении в статистических критериях точных и предельных приближений базовых распределений вероятностей значений тестовых статистик // Обозрение прикладной и промышленной математики. – 2018. – Т. 25. – Вып. 4. – С. 375-378. – ISSN 0869-8325. – https://tvp.ru/conferen/vsppmXIX/repso114.pdf (дата обращения: 14.01.2019).

2. *Зелюкин Н.Б., Мельников А.К.* Сложность расчета точных распределений вероятности значений статистик и область применения предельных распределений // Электронные средства и системы управления: Матер. докладов XIII Междунар. науч.-практ. конф. (29 ноября – 1 декабря 2017 г.): в 2 ч. Ч. 2. – Томск: В-Спектр, 2017. – С. 84-90. – https://storage.tusur.ru/files/115115/2017-2.pdf (дата обращения: 13.07.2018).
3. *Мельников А.К.* Сложность расчета точных распределений вероятности симметричных аддитивно разделяемых статистик и область применения предельных распределений // Доклады ТУСУР. – Томск, 2017. – Т. 20, № 4. – С. 126-130. – ISSN 1818-0442.
4. *Чеповский А.М.* Информационные модели в задачах обработки текстов на естественных языках. – М.: Национальный открытый университет «ИНТУИТ», 2015. – 228 с. – ISBN 978-5-9556-0176-2.
5. *Мельников А.К.* Анализ точных и предельных приближений распределений вероятностей значений статистик // Суперкомпьютерные технологии (СКТ-2018): Матер. 5-й Всероссийской научно-технической конференции: в 2 т. Т. 1. – Ростов-на-Дону; Таганрог: Изд-во ЮФУ, 2018. – С. 100-104. – ISBN 978-9275-2834-9.
6. *Мельников А.К.* Применение точных распределений в процедуре двухэтапной обработки текстов // Обозрение прикладной и промышленной математики. – 2018. – Т. 25. – Вып. 2. – С. 89-95. – ISSN 0869-8325. – https://tvp.ru/conferen/vsppmXIX/repso051.pdf (дата обращения: 24.01.2019).
7. *Мельников А.К., Ронжин А.Ф.* Обобщенный статистический метод анализа текстов, основанный на расчете распределений вероятности значений статистик // Информатика и её применения. – 2016. – Т. 10. – Вып. 4. – С. 91-97. – ISSN 1992-2264.
8. *Мельников А.К.* Применение точных и предельных приближений распределений вероятностей значений статистик при решении задачи обработке текстов // Известия ЮФУ. Технические науки. – 2018. – № 8 (202). – С. 114-135.
9. *Мельников А.К.* Относительная алгоритмическая сложность расчета точных приближений распределений вероятностей значений статистик // Известия ЮФУ. Технические науки. – 2019. – № 7 (209). – С. 35-45.
10. *Мельников А.К.* Направление модернизации частотного метода расчета точных распределений вероятностей значений статистик // Обозрение прикладной и промышленной математики. – 201. – Т. 24. – Вып. 5. – С. 324-325. – https://tvp.ru/conferen/vsppmXVIII/kisso073.pdf (дата обращения: 12.07.2018).
11. *Зуев Ю.А.* Современная дискретная математика: От перечислительной комбинаторики до криптографии XXI века. – М.: ЛЕНАРД, 2019. – 720 с. (Основы защиты информации № 17). – ISBN 978-5-9710-5660-7.
12. *Мельников А.К.* Методика расчета распределения вероятностей значений симметричных аддитивно разделяемых статистик, приближенных к их точному распределению // Научный вестник НГТУ. – 2018. – № 1 (70). – С. 153-166. – ISBN 1814-1196. – Doi: 10.17212/1814-1196-2018-1-153-166.
13. *Крамер Г.* Математические методы статистики. – М.: Мир, 1975. – 648 с.
14. *Кнут Д.Э.* Искусство программирования для ЭВМ. Т. 3. Сортировка и поиск. – М.: Мир, 1978. – 844 с.
15. Список TOP 500 55-й редакция от 22 июня 2020 года.
16. *Воеводин В.В., Воеводин Вл.В.* Параллельные вычисления. – СПб.: БХВ-Петербург, 2002. – 608 с. – ISBN 5-94157-160-7.
17. *Муравьёв С.А.* История отечественной электронной вычислительной техники / под общей редакцией директора Департамента радиоэлектронной промышленности Минпромторга России С.В. Хохлова. – М.: ООО "Издательский дом "Столичная энциклопедия", 2017. – 680 с. – ISBN 978-5-903989-34-8.
18. *Гузик В.Ф., Каляев И.А., Левин И.И.* Реконфигурируемые вычислительные системы / под общ. ред. И.А. Каляева. – Таганрог: Изд-во ЮФУ, 2016. – 472 с. – ISBN 978-5-9275-1918-7.
19. *Сачков В.Н.* Введение в комбинаторные методы дискретной математики. – М.: Наука. Гл. ред. физ.-мат. лит., 1982. – 384 с.

20. Свидетельство о государственной регистрации программы для ЭВМ № 2018664980. Расчет вероятностей значений статистики максимальной частоты. Правообладатель Мельников А.К. Автор: Мельников А.К. и Зелюкин Н.Б. Заявка № 2018662123. Дата поступления 01 ноября 2018 г. Дата государственной регистрации в Реестре программ для ЭВМ 27 ноября 2018 г.
21. *Мельников А.К.* Относительная алгоритмическая сложность расчета точных приближений распределений вероятностей значений статистик // Модели, методы и технологии интеллектуального управления (ИУ-2019): Матер. 12-й Мультиконференции по проблемам управления (МКПУ-2019): в 4 т. Т. 1. – Ростов-на-Дону; Таганрог: Изд-во ЮФУ, 2019.. – С. 108-112. – ISBN 978-5-9275-3189-9.

REFERENCES

1. *Mel'nikov A.K.* Sravnenie effektivnosti obrabotki tekstov pri primenenii v statisticheskikh kriteriyakh tochnykh i predel'nykh priblizheniy bazovykh raspredeleniy veroyatnostey znacheniy testovykh statistik [Comparison of the efficiency of text processing when applying exact and marginal approximations of the basic probability distributions of test statistics in statistical criteria], *Obozrenie prikladnoy i promyshlennoy matematiki* [Review of Applied and Industrial Mathematics], 2018, Vol. 25, Issue 4, pp. 375-378. ISSN 0869-8325. Available at: https://tvp.ru/conferen/vsppmXIX/repso114.pdf (accessed 14 January 2019).
2. *Zelyukin N.B., Mel'nikov A.K.* Slozhnost' rascheta tochnykh raspredeleniy veroyatnosti znacheniy statistik i oblast' primeneniya predel'nykh raspredeleniy [The complexity of calculating accurate probability distributions of statistical values and the scope of limit distributions], *Elektronnye sredstva i sistemy upravleniya: Mater. dokladov XIII Mezhdunar. nauch.-prakt. konf. (29 noyabrya – 1 dekabrya 2017 g.)* [Electronic means and control systems: Materials of reports of the XIII International Scientific and Practical Conference (November 29 – December 1, 2017)]: in 2 parts. Part 2. Tomsk: V-Spektr, 2017, pp. 84-90. Available at: https://storage.tusur.ru/files/115115/2017-2.pdf (accessed 13 July 2018).
3. *Mel'nikov A.K.* Slozhnost' rascheta tochnykh raspredeleniy veroyatnosti simmetrichnykh additivno razdelyaemykh statistik i oblast' primeneniya predel'nykh raspredeleniy [The complexity of calculating the exact probability distributions of symmetric additive-separated statistics and the application of limit distributions], *Doklady TUSUR* [Proceedings of Tomsk State University]. Tomsk, 2017, Vol. 20, No. 4, pp. 126-130. ISSN 1818-0442.
4. *Chepovskiy A.M.* Informatsionnye modeli v zadachakh obrabotki tekstov na estestvennykh yazykakh [Information models in tasks of processing of natural language texts]. Moscow: Natsional'nyy otkrytyy universitet «INTUIT», 2015, 228 p. ISBN 978-5-9556-0176-2.
5. *Mel'nikov A.K.* Analiz tochnykh i predel'nykh priblizheniy raspredeleniy veroyatnostey znacheniy statistik [Analysis of exact and marginal approximations of probability distributions of statistical values], *Superkomp'yuternye tekhnologii (SKT-2018): Mater. 5-y Vserossiyskoy nauchno-tekhnicheskoy konferentsii* [Supercomputer Technologies (SCT-2018): Proceedings of the 5th All-Russian Scientific and Technical Conference]: in 2 vol. Vol. 1. Rostov-on-Don; Taganrog: Izd-vo YuFU, 2018, pp. 100-104. ISBN 978-9275-2834-9.
6. *Mel'nikov A.K.* Primenenie tochnykh raspredeleniy v protsedure dvukhetapnoy obrabotki tekstov [Application of exact distributions in the procedure of two-step text processing], Obozrenie prikladnoy i promyshlennoy matematiki [Review of applied and industrial mathematics], 2018, Vol. 25, Issue 2, pp. 89-95. ISSN 0869-8325. Available at: https://tvp.ru/conferen/vsppmXIX/repso051.pdf (accessed 24 January 2019).
7. *Mel'nikov A.K., Ronzhin A.F.* Obobshchennyy statisticheskiy metod analiza tekstov, osnovannyy na raschete raspredeleniy veroyatnosti znacheniy statistik [Generalized statistical method of text analysis based on calculation of probability distribution of statistical values] *Informatika i ee primeneniya* [Informatics and its applications], 2016, Vol. 10, Issue 4, pp. 91-97. ISSN 1992-2264.
8. *Mel'nikov A.K.* Primenenie tochnykh i predel'nykh priblizheniy raspredeleniy veroyatnostey znacheniy statistik pri reshenii zadachi obrabotke tekstov [Application of exact and limit approximations of statistics probability distributions for the problem of text processing], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2018, No. 8 (202), pp. 114-135.
9. *Mel'nikov A.K.* Otnositel'naya algoritmicheskaya slozhnost' rascheta tochnykh priblizheniy raspredeleniy veroyatnostey znacheniy statistik [Relative algorithmic complexity of exact approximations for probability distributions of statistics values], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2019, No. 7 (209), pp. 35-45.

10. *Mel'nikov A.K.* Napravlenie modernizatsii chastotnogo metoda rascheta tochnykh raspredeleniy veroyatnostey znacheniy statistik [Direction of modernization of the frequency method of exact probability distributions for statistics values], *Obozrenie prikladnoy i promyshlennoy matematiki* [Review of applied and industrial mathematics], 2017, Vol. 24, Issue 5, pp. 324-325. Available at: https://tvp.ru/conferen/vsppmXVIII/kisso073.pdf (accessed 12 July 2018).
11. *Zuev Yu.A.* Sovremennaya diskretnaya matematika: Ot perechislitel'noy kombinatoriki do kriptografii XXI veka [Modern discrete mathematics: From enumerative combinatorics to cryptography of the XXI century]. Moscow: LENARD, 2019, 720 p. (Fundamentals of Information security No. 17). ISBN 978-5-9710-5660-7.
12. *Mel'nikov A.K.* Metodika rascheta raspredeleniya veroyatnostey znacheniy simmetrichnykh additivno razdelyaemykh statistik, priblizhennykh k ikh tochnomu raspredeleniyu [A method for calculating the probability distribution of values of symmetric additively separated statistics that are close to their exact distribution], *Nauchnyy vestnik NGTU* [Science bulle-tin of the Novosibirsk state technical university], 2018, No. 1 (70), pp. 153-166. ISBN 1814-1196. Doi: 10.17212/1814-1196-2018-1-153-166.
13. *Kramer G.* Matematicheskie metody statistiki [Mathematical methods of statistics]. Moscow: Mir, 1975, 648 p.
14. *Knut D.E.* Iskusstvo programmirovaniya dlya EVM. T. 3. Sortirovka i poisk [The art of computer programming. Vol. 3. Sorting and Searching]. Moscow: Mir, 1978, 844 p.
15. Spisok TOR 500 55-y redaktsiya ot 22 iyunya 2020 goda [TOP 500 list 55 edition of june 22, 2020].
16. *Voevodin V.V., Voevodin Vl.V.* Parallel'nye vychisleniya [Parallel computing]. St. Petersburg: BKhV-Peterburg, 2002, 608 p. ISBN 5-94157-160-7.
17. *Murav'ev S.A.* Istoriya otechestvennoy elektronnoy vychislitel'noy tekhniki [History of Russian electronic computer technology], under the General editorship of the Director of the Department of radio-electronic industry of the Ministry of industry and trade of Russia S.V. Khokhlov. Moscow: OOO "Izdatel'skiy dom "Stolichnaya entsiklopediya", 2017, 680 p. ISBN 978-5-903989-34-8.
18. *Guzik V.F., Kalyaev I.A., Levin I.I.* Rekonfiguriruemye vychislitel'nye sistemy [Reconfigurable computing systems], under the General ed. of I.A. Kalyayev. Taganrog: Izd-vo YuFU, 2016, 472 p. ISBN 978-5-9275-1918-7.
19. *Sachkov V.N.* Vvedenie v kombinatornye metody diskretnoy matematiki [Introduction to combinatorial methods of discrete mathematics]. Moscow: Nauka. Gl. red. fiz.-mat. lit., 1982, 384 p.
20. Svidetel'stvo o gosudarstvennoy registratsii programmy dlya EVM № 2018664980. Raschet veroyatnostey znacheniy statistiki maksimal'noy chastoty. Pravoobladatel' Mel'nikov A.K. Avtor: Mel'nikov A.K. i Zelyukin N.B. Zayavka № 2018662123. Data postupleniya 01 noyabrya 2018 g. Data gosudarstvennoy registratsii v Reestre programm dlya EVM 27 noyabrya 2018 g. [Certificate of state registration of computer software № 2018664980. Calculation of probabilities of maximum frequency statistics. Proprietor – A.K. Melnikov. Authors: A.K. Melnikov and N.B. Zeliukin. Application № 2018662123. Date of filing – November 01, 2018. Date of state registration in the Register of computer software – November 27, 2018].
21. *Mel'nikov A.K.* Otnositel'naya algoritmicheskaya slozhnost' rascheta tochnykh priblizheniy raspredeleniy veroyatnostey znacheniy statistik [Relative algorithmic complexity of exact approximations for probability distributions of statistics values], *Modeli, metody i tekhnologii intellektual'nogo upravleniya (IU-2019): Mater. 12-y Mul'tikonferentsii po problemam upravleniya (MKPU-2019)* [Models, methods and technologies of intelligent management (IU-2019): Materials of the 12th Multi-Conference on Management Problems (ICPU-2019)]: in 4 vol. Vol. 1. Rostov-on-Don; Taganrog: Izd-vo YuFU, 2019, pp. 108-112. ISBN 978-5-9275-3189-9.

Статью рекомендовал к опубликованию д.т.н., профессор И.И. Левин.

Мельников Андрей Кимович – АО «Вычислительные решения»; e-mail: ak@comp-sol.ru; 117587, г. Москва, Варшавское шоссе, 125, тел.: +74952870035; к.т.н.; доцент ВАК; г.н.с.

Melnikov Andrey Kimovich – SC «Computing solutions»; e-mail: ak@comp-sol.ru; 125, Varshavskoye roag, Moscow, 117587, Russia; phone: +784952870035; cand. of eng. sc.; associate professor of SAC; chief research officer.