Methods, algorithms and programs of computer algebra in problems of registration and analysis of random point structures

A.L. Reznik [1], A.A. Soloviev [1]
[1] Institute of Automation and Electrometry of the Siberian Branch of the Russian Academy of Sciences, 630090, Novosibirsk, Russia, Academician Koptyug ave. 1

Abstract

An original approach to solving difficult time-consuming problems of registration and analysis of random point images is described. The approach is based on the development and application of high-performance specialized computer algebra systems. Three software packages have been created specifically for carrying out equivalent analytical transformations on a computer. The first software system is designed to calculate formulas describing the volumes of convex polyhedra with parametrically specified boundaries in *n*-dimensional space. The second system is based on the calculation of multidimensional integral expressions by the method of cyclic differentiation of the integral with respect to the parameter. The third system is based on the accelerated implementation of complex combinatorial-recursive transformations on a computer. Another distinctive feature of the work is the extension of the classical Catalan numbers to the multidimensional case (they were required to solve a number of intermediate probabilistic-combinatorial problems). The implementation of the above computer algebra software systems on a multi-core cluster of Novosibirsk State University, together with the direct use of the explicit form of generalized Catalan numbers, allowed the authors to obtain several new previously unknown probabilistic formulas and dependencies required for solving problems in the field of analysis of random point images.

Keywords: methods of computer algebra, algorithms of computer algebra, programs of computer algebra, random point structures.

Citation: Reznik AL, Soloviev AA. Methods, algorithms and programs of computer algebra in problems of registration and analysis of random point structures. Computer Optics 2023; 47(6): 1011-1017. DOI: 10.18287/2412-6179-CO-1330.

Acknowledgements: This work was supported by Ministry of Science and Higher Education of the Russian Federation (project no. № AAA-A17-117052410034-6).

Introduction

The first distinguishing feature of the research approaches used in this work is the development of methods (that are quite rare in this field of informatics) based not on numerical calculations, but on analytical transformations on a computer to solve complex fundamental and applied problems that arise in the analysis of random point images. To date, a large number of programming languages and computer algebra software systems are known [1 – 7]. In Russia (more precisely, in the USSR), the authors of the first software systems for analytical manipulation, successfully applied for scientific research, were academicians L.V. Kantorovich, N.N. Yanenko, V.M. Glushkov, D.V. Shirkov (1960-1970), I.V. Pottosin [8 – 12]. The algorithms and programs developed by them turned out to be effective for finding solutions to algebraic-differential equations, calculating polynomial expressions, solving problematic and time-consuming problems of theoretical physics. Although most of the above-mentioned computer algebra software systems become well-known and confirmed their effectiveness, nevertheless, the universality of such systems does not allow them to be an effective research tool in all areas of application. In particular, in one of the intermediate tasks related to the reliability of digital image registration, we needed a cyclic calculation of formulas that describe the volumes of parametrically given polyhedra in *n*-dimensional space. Efficient algorithms for calculating such integrals did not exist in any universal system of computer algebra. As a result, we ourselves created a mathematical basis, computational algorithms and programs implemented on their basis for calculating the analytical formulas we need.

The second important distinguishing feature of the methods used by us is that with their help, in a number of specific cases, it is possible to confirm the idea expressed by J. von Neumann: when a researcher is faced with a problem that he is unable to solve, he can use a computer to carry out time-consuming calculations that can lead to "correct" answer, and after that (with a solution "prompted" by the computer) it becomes possible to find a rigorous and constructive proof. In the work, using computer calculations, we managed to establish and then mathematically rigorously prove a number of analytical formulas that describe the probability of error-free reading of random discrete-point images. To do this without serious software support would be impossible or very difficult.

The third distinguishing feature of the presented work is that for a strict mathematical proof of the probabilistic dependences and relations obtained in it, we proposed an extension of the classical Catalan numbers to the multidimensional case (including their explicit form for a number of specific cases).

All of the above approaches turned out to be a useful tool for solving many scientific problems [13 – 14], which, by their physical nature, are associated with probabilistic processes and, at the same time, require the use of highly efficient tools of computational informatics.

1. Analytical transformations on a computer in the problems of registration and analysis of random point images

The main problem considered here is related to estimating the probability of error-free reading of a random two-dimensional field in the case when the coordinates of point objects forming such a field are recorded using a reading aperture having a limited number of threshold levels.

By a random two-dimensional field is meant a random discrete-point distribution on a plane, created by a certain Poisson source with intensity λ. The classic way to read the coordinates of such point objects is to use a "television" scheme for readout a random image in accordance with the scheme shown in Fig. 1.

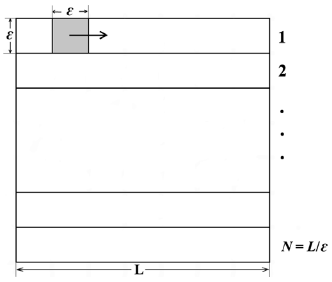

Fig. 1. Television reading scheme of random point image

The reading aperture of the integrator having k threshold levels sequentially scans all horizontal bands of the image. When a point object falls within the reading aperture, the total integrator signal increases by one, and the coordinates of the point object are fixed. Meanwhile, the most important characteristic of the registration process is the probability of its error-free implementation, i.e. the probability that there will never be more than k objects inside the scanning window:

$$P = \left[\exp(-\lambda\varepsilon L)\sum_{n=0}^{\infty}\frac{(\lambda\varepsilon L)^n}{n!}P_{n,k}(\varepsilon, L)\right]^N . \quad (1)$$

Here, $P_{n,k}(\varepsilon,L)$ is the probability of error-free readout of all horizontal bands of the image, provided that exactly n samples fall into separate band. Using the standard normalization, the parameter L can be excluded from consideration, and as a result finding the probability (1) reduces to solving the following one-dimensional problem:

"It is required to find the probability $P_{n,k}(\varepsilon)$ that when randomly throwing n points $x_1, x_2, \dots, x_n$ onto the interval (0,1), no subinterval of length ε containing more than k points is formed inside it."

Unfortunately, the analytical solution of this problem, despite the simplicity of the formulation, is known only for the case $k=1$ [15 – 16] and is given by the formula

$$P_{n,1}(\varepsilon, L) = (L-(n-1)\varepsilon)^n / L^n , \left(0 \le \varepsilon \le L/(n-1) \right). \quad (2)$$

The problem posed is related to a random partition of the interval. Existing methods [17 – 18] are not sufficient to find a general formula for the probability $P_{n,k}(\varepsilon)$, therefore, to find its particular solutions we implemented several ideas based on analytical transformations on a computer [19].

2. Determining the reliability of readout random point fields using analytical transformations on a computer

The relations (3) – (6) given below describe the scheme for calculating the formulas $P_{n,k}(\varepsilon)$ using the analytical calculation on a computer the multidimensional integrals over convex polyhedra in n-dimensional space:

$$P_{n,k}(\varepsilon) = n!\int\limits_{D_{n,k}(\varepsilon)} \dots \int dx_1 \dots dx_n , \quad (3)$$

$$\begin{cases} 0 < x_1 < x_2 < \dots < x_{n-1} < x_n < 1 , \\ x_{k+1} - x_1 > \varepsilon , \\ x_{k+2} - x_2 > \varepsilon , \\ \quad \vdots \\ x_n - x_{n-k} > \varepsilon , \end{cases} \quad (4)$$

$$P_{n,k}(\varepsilon) = n!\int_{-\infty}^{+\infty}\dots\int_{-\infty}^{+\infty} 1[x_1]1[x_2]\dots 1[x_n - x_{n-1}]\,1[1-x_n]\,1[x_{k+1}-x_1-\varepsilon]\times 1[x_{k+2}-x_2-\varepsilon]\dots 1[x_n - x_{n-k}-\varepsilon]dx_1\dots dx_n , \quad (5)$$

$$\left(\prod_{j=1}^{l} 1[x_r-\alpha_j]\right)\left(\prod_{i=1}^{m} 1[\beta_i - x_r]\right) = \sum_{j=1}^{l}\sum_{i=1}^{m} 1[x_r-\alpha_j]1[\beta_i - x_r]\left\{1[\beta_i-\alpha_j]\left(\prod_{\substack{q=1\\ q\ne j}}^{l} 1[\alpha_j-\alpha_q]\right)\left(\prod_{\substack{s=1\\ s\ne i}}^{m} 1[\beta_s-\beta_i]\right)\right\}, \quad (6)$$

$$\begin{aligned} P_{n,k}(\varepsilon) = {} & n!\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty}\left[\int_0^{x_2} dx_1\right]1\left[x_{k+1}-x_2-\varepsilon\right]1\left[x_2\right]\left(\prod_{m=3}^{n}1\left[x_m - x_{m-1}\right]\right)\times 1\left[1-x_n\right]\left(\prod_{m=2}^{n-k}1\left[x_{m+k}-x_m-\varepsilon\right]\right)dx_2\dots dx_n + \\ & + n!\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty}\left[\int_0^{x_{k+1}-\varepsilon} dx_1\right]1\left[x_2+\varepsilon-x_{k+1}\right]1\left[x_{k+1}-\varepsilon\right]\times \\ & \times\left(\prod_{m=3}^{n}1\left[x_m - x_{m-1}\right]\right)1\left[1-x_n\right]\left(\prod_{m=2}^{n-k}1\left[x_{m+k}-x_m-\varepsilon\right]\right)dx_2\dots dx_n , \end{aligned} \quad (7)$$

$$
\begin{aligned}
P_{n,k}(\varepsilon) = {} & 1[\varepsilon - a_1]1[b_1 - \varepsilon_1] \times \int\limits_{\alpha_{n,1}(\varepsilon)}^{\beta_{n,1}(\varepsilon)} dx_n \int\limits_{\alpha_{n-1,1}(x_n,\varepsilon)}^{\beta_{n-1,1}(x_n,\varepsilon)} dx_{n-1} \ldots \int\limits_{\alpha_{21}(x_3,\ldots,x_n,\varepsilon)}^{\beta_{21}(x_3,\ldots,x_n,\varepsilon)} dx_2 \int\limits_{\alpha_{11}(x_2,\ldots,x_n,\varepsilon)}^{\beta_{11}(x_2,\ldots,x_n,\varepsilon)} dx_1 + \\
& +1[\varepsilon - a_2]1[b_2 - \varepsilon_1] \times \int\limits_{\alpha_{n,2}(\varepsilon)}^{\beta_{n,2}(\varepsilon)} dx_n \int\limits_{\alpha_{n-1,2}(x_n,\varepsilon)}^{\beta_{n-1,2}(x_n,\varepsilon)} dx_{n-1} \ldots \int\limits_{\alpha_{22}(x_3,\ldots,x_n,\varepsilon)}^{\beta_{22}(x_3,\ldots,x_n,\varepsilon)} dx_2 \int\limits_{\alpha_{12}(x_2,\ldots,x_n,\varepsilon)}^{\beta_{12}(x_2,\ldots,x_n,\varepsilon)} dx_1 + \\
& \ldots \\
& +1[\varepsilon - a_N]1[b_N - \varepsilon_1] \times \int\limits_{\alpha_{n,N}(\varepsilon)}^{\beta_{n,N}(\varepsilon)} dx_n \int\limits_{\alpha_{n-1,N}(x_n,\varepsilon)}^{\beta_{n-1,N}(x_n,\varepsilon)} dx_{n-1} \ldots \int\limits_{\alpha_{2N}(x_3,\ldots,x_n,\varepsilon)}^{\beta_{2N}(x_3,\ldots,x_n,\varepsilon)} dx_2 \int\limits_{\alpha_{1N}(x_2,\ldots,x_n,\varepsilon)}^{\beta_{1N}(x_2,\ldots,x_n,\varepsilon)} dx_1 .
\end{aligned}
\tag{8}
$$

First, the original integral (3) over the domain $D_{n,k}(\varepsilon)$ defined by the system of inequalities (4) is reduced with the help of transformations (5)–(6) to the sum of iterated integrals with the already set integration limits. Relation (7) shows the first step of the algorithm, where the limits of integration with respect to the variable x_1 are determined. Expression (8) describes the resulting set of all n-dimensional iterated integrals into which the original n-fold integral (3) decomposes. At the final stage, each of the iterated integrals is sequentially integrated and the results obtained are combined, taking into account the boundaries of the free parameter ε. We have created a software package for computerized implementation of all the listed analytical transformations on a multi-core cluster of Novosibirsk State University.

It is designed for high-speed calculation of multiple integral expressions with space dimensions up to $n=15$. When calculating manually, it is practically impossible to set the limits of integration, check all intermediate systems of inequalities for compatibility, and directly calculate iterated integrals even for $n=3$.

The software package allows to calculate the formulas $P_{n,k}(\varepsilon)$ for all possible ranges of the parameter ε for any values of $n<16$ and $k<n$. Table 1 is an example of calculation of the formula $P_{n,2}(\varepsilon)$ for $n=15$ using the latest software modification and the capabilities of the Novosibirsk State University multi-core cluster in 2023.

In the software implementation of any computational algorithms (both numerical and analytical), the most important characteristics are their complexity (i.e., the dependence of the total number of elementary operations on the dimension of the problem being solved) and required resources (i.e., memory costs). On modern multi-core computing clusters, these two indicators are the main parameters that directly affect the total execution time of the computing process. The latest version of implemented parallel algorithm on a multi-core cluster makes it possible to calculate the integral (3) in the case when the integration domain (4) is not necessarily tied to the calculation of the formulas $P_{n,k}(\varepsilon)$, but is specified by an arbitrary system of linear inequalities in the variables x_1, $x_2,\ldots,\ x_n$, ε. The complexity of such an algorithm depends not only on the dimension n of the problem being solved, but also on the parameter $C(n)$, which takes into account the number of inequalities in system (4) and the specifics of the problem being solved, and can be represented as $O(n!\ C(n))$. It should be noted that it is quite difficult to estimate the parameter $C(n)$ in each particular case. For example, in the problem described above related to the calculation of formulas $P_{n,k}(\varepsilon)$, system (4) consists of $(2n-k+1)$ inequalities, i.e. for $k=1$, the number of inequalities is maximum, while the number of iterated integrals into which the original multiple integral decomposes (and hence the execution time of the algorithm) is minimal.

Tab. 1. An example of the program's results for n =15

n	k	Range of ε	$P_{n,k}(\varepsilon)$
15	2	(0,1/9)	$1-1365\varepsilon^2+22750\varepsilon^3+281190\varepsilon^4-12048036\varepsilon^5+122727605\varepsilon^6+112857030\varepsilon^7-14073724665\varepsilon^8+$ $+144064260340\varepsilon^9-680385878172\varepsilon^{10}+1025143195440\varepsilon^{11}+5203298446255\varepsilon^{12}-$ $-30893050642020\varepsilon^{13}+64426034589465\varepsilon^{14}-49991961279274\varepsilon^{15}$
		(1/9,1/8)	$-12+1680\varepsilon-102480\varepsilon^2+3781960\varepsilon^3-96242055\varepsilon^4+1800598800\varepsilon^5-25589273710\varepsilon^6+$ $+280538329500\varepsilon^7-2384010949320\varepsilon^8+15656380639900\varepsilon^9-78629775685461\varepsilon^{10}+$ $+296111707052070\varepsilon^{11}-808664482513160\varepsilon^{12}+1511172218544240\varepsilon^{13}-1727589374382480\varepsilon^{14}+$ $+910833321829088\varepsilon^{15}$
		(1/8,1/7)	$-572+60060\varepsilon-2927925\varepsilon^2+87957870\varepsilon^3-1821799980\varepsilon^4+27567900360\varepsilon^5-14953228590\varepsilon^6+$ $+2767033861020\varepsilon^7-18852497543880\varepsilon^8+99631266835100\varepsilon^9-405144364726101\varepsilon^{10}+$ $+1245107820746790\varepsilon^{11}-2799752684568840\varepsilon^{12}+4349065819578480\varepsilon^{13}-4173546405908880\varepsilon^{14}+$ $+1865389803365088\varepsilon^{15}$

In addition to the programs for calculating multidimensional integrals over convex polyhedra in n-dimensional space, we have developed two computer algebra software packages to solve the problems of registration and analysis of random point images. The first package is designed for specialized recursive analytical calculations. The second package is related to the calculation of probabilistic formulas by the method of differentiating multidimensional integrals with respect to a parameter. Fig. 2 shows one of the one-dimensional schemes for the placement of random point objects, the analysis of which required the use of the two mentioned software systems.

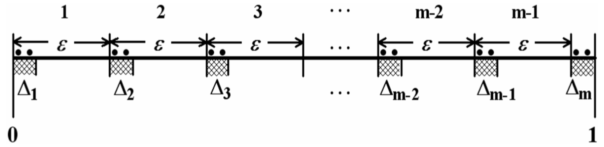

Fig. 2. Allocation scheme for an even (n = 2m) number of random points on the interval (0,1), when no subinterval of length Ω_ε contains more than 2 points. Case $1/m < \varepsilon < 1/(m-1)$

3. Combinatorial-linguistic problems that arise in the study of random point structures and lead to classical and generalized Catalan numbers

In the analysis of random point structures, the word-symbolic formulation of problems turns out to be a very useful tool for ranking interdependent numerical sequences [20]. Below we will give the formulations of several tasks; some of them are direct extensions of the classical Catalan numbers to two and three dimensions.

The first task is formulated as follows: «Different words of length $(N_a+N_b+N_c)$ are composed of N_a symbols "a", N_b symbols "*b*" and N_c symbols "*c*". It is necessary to determine the total number of words $W(N_a, N_b, N_c)$ such that for each of them two conditions are simultaneously fulfilled:

1. When viewing a word from left to right, the number of "*b*" characters encountered never exceeds the number of "*a*" characters encountered.
2. When viewing a word from right to left, the number of "*c*" characters encountered never exceeds the number of "*a*" characters encountered».

To find a solution, the given word-symbolic problem was translated into a geometric form (see Fig. 3). It was required on a three-dimensional discrete lattice in the coordinate system (X, Y, Z) to find the number of monotone paths from the point (0,0,0) to the point (N_a, N_b, N_c) that do not intersect any of the $\mathcal{P}_1$ plane given by the equation $X-Y=0$, nor the plane $\mathcal{P}_2$ given by the equation $X-Z+N_c-N_a=0$ (in the problems we considered, the planes $\mathcal{P}_1$ and $\mathcal{P}_2$ could not have intersections inside the parallelepiped shown in Fig. 3).

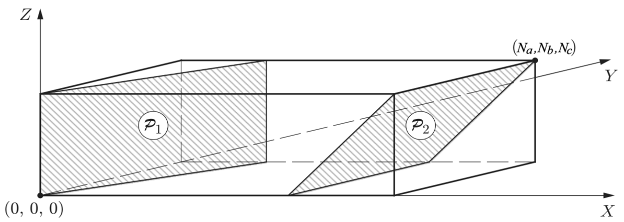

Fig. 3. Box with bounding planes $\mathcal{P}_1$ and $\mathcal{P}_2$

The number of such paths corresponds exactly to the number of three-character words $W(N_a, N_b, N_c)$ that satisfy the conditions of the problem. For the case $N_a > N_b + N_c - 2$ its solution is the relation

$$W(N_a,N_b,N_c)= \\ =\frac{(N_a+N_b+N_c)!}{N_a!N_b!N_c!}-\frac{(N_a+N_b+N_c)!}{(N_a+1)!(N_b-1)!N_c!}- \\ -\frac{(N_a+N_b+N_c)!}{(N_a+1)!N_b!(N_c-1)!}+\frac{(N_a+N_b+N_c)!}{(N_a+2)!(N_b-1)!(N_c-1)!}. \quad (9)$$

This formula was obtained with the help of a double application of the mirror reflection method, first proposed by the French mathematician D. Andre back in 1887 to solve the Bertrand ballot problem [21]. We named these numbers as «generalized three-dimensional Catalan numbers», since the resulting formula generalizes the one-dimensional Catalan sequence known from many applications (it is obtained from (9) with $N_a=N_b$ and $N_c=0$). The use of multidimensional generalized Catalan numbers allowed us to prove a number of analytical probabilistic relations that characterize the reliability of registration of random point images. In particular, we have proved the exact relation (10):

$$P_{2m,2}(\varepsilon)=(C_{2m}^{m}-C_{2m}^{m-1})(1-(m-1)\varepsilon)^{2m} \quad (10)$$

for $1/m<\varepsilon<1/(m-1)$;

To do this, we first managed to track the structure of relation (10) using the existing set of computer-calculated formulas $P_{2m,2}(\varepsilon)$, then convert the original continuous problem into a discrete one, and then, using the program-combinatorial approach, find its exact solution. By the way, the first factor in formula (10) is the classical Catalan number.

The basis for the proof of formula (11) below

$$P_{2m+1,2}(\varepsilon)=C_{2m+1}^{m+1}(1-m\varepsilon)^{m+1}(1-(m-1)\varepsilon)^{m}- \\ -2C_{2m+1}^{m+2}(1-m\varepsilon)^{m+2}(1-(m-1)\varepsilon)^{m-1}+ \\ +C_{2m+1}^{m+3}(1-m\varepsilon)^{m+3}(1-(m-1)\varepsilon)^{m-2} \quad (11)$$

for $1/(m+1)<\varepsilon<1/m$;
was the formula (9) of the above word-symbolic formulation of the problem (as already noted, this is a three-dimensional extension of the classical Catalan sequence).

To prove formula (12)

$$P_{2m,2}(\varepsilon)=C_{2m}^{m}(1-(m-1)\varepsilon)^{2m}-C_{2m}^{m-1}(1-(m-1)\varepsilon)^{2m}- \\ -C_{2m}^{m-2}(1-m\varepsilon)^{m+2}(1-(m-2)\varepsilon)^{m-2}+ \\ +2C_{2m}^{m-3}(1-m\varepsilon)^{m+3}(1-(m-2)\varepsilon)^{m-3}- \\ -C_{2m}^{m-4}(1-m\varepsilon)^{m+4}(1-(m-2)\varepsilon)^{m-4} \quad (12)$$

for $1/(m+1)<\varepsilon<1/m$,
we needed to solve two additional word-symbolic problems. One of them is quite simple and formulated as follows: «Different words of length (N_a+N_b) are composed of N_a symbols "*a*" and N_b symbols "*b*". It is required to find the total number of words such that when they are viewed from left to right, the number of "*a*" characters

encountered never exceeds the number of "*b*" characters encountered by more than k (it is assumed that $N_b \le N_a$ and $0 \le k \le N_a - N_b$».

$$W(N_a, N_b) = \frac{(N_a + N_b)!}{N_a!\,N_b!} - \frac{(N_a + N_b)!}{(N_a - k - 1)!(N_b + k + 1)}. \quad (13)$$

The decision was found using the already mentioned mirror reflection method. It is easy to see that formula (13) is a two-dimensional extension of the classical

Catalan numbers, which are obtained from it when $N_a = N_b$ and $k = N_a - N_b$ (or $k = 0$).

The second task is much more difficult: «Different words of length $(N_a + N_b + N_c + N_d)$ are composed of N_a symbols "*a*", N_b symbols "*b*", N_c symbols "*c*" and N_d symbols "*d*" (they must meet the conditions $0 \le N_b < N_a \le m$, $0 \le N_d < N_c \le m$, $(N_a + N_d) \le m$, $(N_b + N_c) \le m$, $(N_a + N_c) \ge m + 1$). It is required to find the total number of words $W_k(N_a, N_b, N_c, N_d)$ that, when viewed from left to right, simultaneously satisfy three conditions:

1. The number of "*b*" characters encountered never exceeds the number of "*a*" characters encountered.
2. The number of "*d*" characters encountered never exceeds the number of "*c*" characters encountered.
3. The number of "*a*" characters encountered never exceeds the number of "*c*" characters encountered by more than $k = (m+1) - N_c$».

Formula (14)

$$W_{(m+1)-N_c}(N_a, N_b, N_c, N_d) = (N_a + N_b + N_c + N_d)! \times \left\{ \left[\frac{1}{N_a!\,N_b!\,N_c!\,N_d!} - \frac{1}{N_a!\,N_b!(N_c+1)!(N_d-1)!} \right] + \right.$$
$$+ \left[\frac{1}{(N_a+1)!(N_b-1)!(N_c+1)!(N_d-1)!} - \frac{1}{(N_a+1)!(N_b-1)!N_c!\,N_d!} \right] + \left[\begin{array}{l} \dfrac{1}{(m+3)!N_b!(N_d-1)!(N_a+N_c-(m+2))!} - \\ -\dfrac{1}{(m+2)!N_b!N_d!(N_a+N_c-(m+2))!} \end{array} \right] + \quad (14)$$
$$\left. + \left[\begin{array}{l} \dfrac{1}{(m+3)!(N_b-1)!N_d!(N_a+N_c-(m+2))!} - \\ -\dfrac{1}{(m+4)!(N_b-1)!(N_d-1)!(N_a+N_c-(m+2))!} \end{array} \right] \right\},$$

is the solution to this problem. It was obtained using the technique of searching for random paths in Weyl chambers [22].

Using program-analytical methods of direct integration (for two free parameters), we have obtained and proved closed analytical relations (15–18):

$$P_{n,n-1}(\varepsilon) = 1 - \varepsilon^n - n\varepsilon^{n-1}(1-\varepsilon), \quad (15)$$

$$P_{n,n-2}(\varepsilon) = 1 - 2C_n^2\varepsilon^{n-2}(1-\varepsilon)^2 - 2\varepsilon^n, \quad 0 \le \varepsilon \le (1/2) \quad (16)$$

$$P_{n,n-2}(\varepsilon) = 1 - 2\varepsilon^n + (2\varepsilon-1)^n - 2C_n^2\varepsilon^{n-2}(1-\varepsilon)^2, \quad (1/2) \le \varepsilon \le 1 \quad (17)$$

$$\underset{(n>6)}{P_{n,n-3}(\varepsilon)} = \begin{cases} 1 - 2\varepsilon^n + C_n^1(6\varepsilon^n - 4\varepsilon^{n-1}) + C_n^2(-3\varepsilon^n + \varepsilon^{n-2}) + C_n^3(9\varepsilon^n - 18\varepsilon^{n-1} + 12\varepsilon^{n-2} - 3\varepsilon^{n-3}), \\ 0 \le \varepsilon \le (1/2); \\ 1 - 2\varepsilon^n + (2\varepsilon-1)^n + C_n^1(1-\varepsilon)(-2\varepsilon^{n-1} + 2(2\varepsilon-1)^{n-1}) + C_n^2(1-\varepsilon)^2(\varepsilon^{n-2} + (2\varepsilon-1)^{n-2}) - 3C_n^3\varepsilon^{n-3}(1-\varepsilon)^3, \\ (1/2) \le \varepsilon \le 1; \end{cases} \quad (18)$$

The following asymptotic dependence (19) has been established and confirmed by all software calculations (including the results of 2023 obtained on the multi-core cluster of Novosibirsk State University):

$$\begin{aligned} P_{n,2}(\varepsilon) &= C_n^0 + C_n^2(-n+2)\varepsilon^2 + C_n^3(4n-10)\varepsilon^3 + \\ &+ C_n^4(2n^2 - 37n + 86)\varepsilon^4 + C_n^5(-40n^2 + 394n - 922)\varepsilon^5 + \\ &+ C_n^6(-15n^3 + 625n^2 - 5171n - 12086)\varepsilon^6 + \\ &+ C_n^7(420n^3 - 10724n^2 + 79996n - 187002)\varepsilon^7 + \\ &+ C_n^8(105n^4 - 10570n^3 + 205499n^2 - 1426841n + \\ &+ 3336406)\varepsilon^8 + C_n^9(5040n^4 - 155708n^3 + 2267664n^2 - \\ &- 17317506n + 52315558)\varepsilon^9 + C_n^{10}(-945n^5 + 189000n^4 - \\ &- 15794525n^3 + 389687181n^2 - 3798029823n + \\ &+ 12998966646)\varepsilon^{10} + o(\varepsilon^{10}). \end{aligned} \quad (19)$$

Conclusion

It is important to note that the exact probabilistic formulas $P_{n,k}(\varepsilon)$ obtained as a result of the software-analytical calculations are required for solving many practical and applied scientific problems related to digital image processing. In particular, their knowledge is necessary for selecting the optimal number of threshold levels of a scanning detector receiver for television reading of random point images [23]. In biomedical research, these dependencies are required when classifying the analyzed images into standard and abnormal ones in order to promptly identify tumors and other disease-causing formations [24]. Methods and programs developed for intellectual analysis of random point images turn out to be a useful tool in solving problems related to the construction of optimal algorithms for localizing point sources with a random discipline of pulse generation [25]. The created software-algorithmic complex and the scientific results obtained with its application became the basis for many research projects and scientific works carried out by the Institute of Automation and Electrometry of the Siberian Branch of the Russian Academy of Sciences (Novosibirsk) together with the Joint Institute for Informatics Problems (Minsk). Author's works aimed at creating new methods of software-analytical solution of the problems of registration, intellectual processing and analysis of random point images were supported by the RFBR already at the initial stage of research. In particular, RFBR grants financed the publication of a monograph [26] and a popular science article [27].

Thus, an example has been demonstrated when a computer with specialized software support is effectively used to solve complex probabilistic problems of analysis of random point images. In this case, the computer did not played the role of just a powerful calculator, but an intelligent assistant equipped with tools for conducting branched and time-consuming analytical calculations.

References

[1] Recio T, Losada-Liste R, Tabera LF, Ueno C. Visualizing a cubic linkage through the use of CAS and DGS. Mathematics 2022; 10(15): 2550. DOI: 10.3390/math10152550.

[2] Koepf W. Computer algebra: An algorithm-oriented introduction. Cham: Springer Nature Switzerland AG; 2022. ISBN: 978-3-030-78016-6.

[3] Bright C, Kotsireas I, Ganesh V. Applying computer algebra systems with SAT solvers to the Williamson conjecture. J Symb Comput 2020; 100: 187-209. DOI: 10.1016/j.jsc.2019.07.024.

[4] Meurer A, Smith C, Paprocki M, et al. SymPy: symbolic computing in Python. PeerJ Comput Sci 2017; 3: e103. DOI: 10.7717/peerj-cs.103.

[5] Moses J. Macsyma: A personal history. J Symb Comput 2012; 47(2): 123-130. DOI: 10.1016/j.jsc.2010.08.018.

[6] Peeters K. Introducing Cadabra: a symbolic computer algebra system for field theory problems. arXiv Preview. 2018. Source: <https://arxiv.org/abs/hep-th/0701238>.

[7] McCallum M. Computer algebra in gravity research. Living Reviews in Relativity, 2018; 21(1): 6. DOI:10.1007/s41114-018-0015-6.

[8] Kantorovich LV. On carrying out numerical and analytical calculations on machines with program control [In Russian]. Izvestiya AN ArmSSR, Seriya Fiz-Mat Nauk 1957; 10(2): 3-16.

[9] Shurygin VA, Yanenko NN. On the realization of algebraic-differential algorithms on the computer [In Russian]. Problems of Cybernetics 1961; 1: 33-43.

[10] Bezhanova MM, Katkov VL, Pottosin IV. Work on analytical transformations in the Computing Center of the Siberian Branch of the USSR Academy of Sciences [In Russian]. T 3. Kharjkov: "FTINT AN USSR" Publisher; 1972: 18-20.

[11] Gerdt VP, Tarasov OV, Shirkov DV. Analytical calculations on computers in the application to physics and mathematics [In Russian]. Uspekhi Fizicheskih Nauk 1980; 130 (1): 113-147.

[12] Glushkov VM, Grinchenko TA, Dorodnicyna AA, et al. ANALITIK-74 [In Russian]. Cybernetics 1978; 5: 114-147.

[13] Reznik AL, Efimov VM, Soloviev AA, Torgov AV. Reliability of readout of random point fields with a limited number of threshold levels of the scanning aperture. Optoelectronics, Instrumentation and Data Processing 2014; 50(6): 582-588. DOI: 10.3103/S8756699014060065.

[14] Reznik AL, Soloviev AA. Software and combinatorial-probabilistic tools for the analysis of random point structures. Pattern Recognit Image Anal 2022; 32(3): 636-638. DOI: 10.1134/S1054661822030348.

[15] Wilks S. Mathematical statistics. Princeton: Princeton University Press; 1944.

[16] Parzen E. Modern probability theory and its applications. New York: John Wiley and Sons; 1960.

[17] Darling DA. On class problems related to the random division of an interval. Ann Math Stat 1953; 24: 239-253.

[18] David G. Ordinal statistics [In Russian]. Moscow: "Nauka" Publisher; 1979.

[19] Reznik AL, Tuzikov AV, Soloviev AA, Torgov AV. Analysis of random point images with the use of symbolic computation codes and generalized Catalan numbers. Optoelectronics, Instrumentation and Data Processing 2016, 52(6): 529-536. DOI: 10.3103/S8756699016060017.

[20] Saracevic M, Adamovic S, Macek N, Selimi A, Pepic S. Source and channel models for secret-key agreement based on Catalan numbers and the lattice path combinatorial approach. J Inf Sci Eng 2021, 37(2): 469-482. DOI: 10.6688/JISE.202103_37(2).0012.

[21] André D. Solution directe du problème résola par M. Bertrand. Comptes Rendus de l'Académie des Sciences 1887; 105: 436-437.

[22] Gessel IM, Zeilberger D. Random walk in a Weyl chamber. Proc Am Math Soc 1992; 115(1): 27-31. DOI: 10.1090/S0002-9939-1992-1092920-8.

[23] Efimov VM, Iskoldsky AM, Krendel YM, Livshits ZA. On the characteristics of various methods for discrete structure images readout [In Russian]. Avtometriya 1973; 1: 3-7.

[24] Anishchenko VV, Vankevich PE, Kovalev VA, Kutsan NV, Lapitsky VA, Linev VN. The use of digital scanning devices and advanced telemedicine technologies in the diagnosis of lung diseases. Minsk: "OIPI NAS of Belarus" Publisher; 2010.

[25] Reznik AL, Tuzikov AV, Soloviev AA, Torgov AV, Kovalev VA. Time-optimal algorithms focused on the search for random pulsed-point sources. Computer Optics 2019; 43(4): 605-610. DOI: 10.18287/2412-6179-2019-43-4-605-610.

[26] Reznik AL, Efimov VM. Computer analytics and generalized Catalan numbers in problems of registration of random discrete objects [In Russian]. Novosibirsk: Publishing House Siberian Department of Russian Academy of Sciences; 2013. ISBN: 978-5-7692-1303-8.

[27] Reznik AL, Efimov VM, Soloviev AA, Torgov AV. About a barber from Ohio and analytical computing on a computer [In Russian]. Collection of popular science articles-winners of the RFBR competition in 2012 2013: 8-19.

Authors' information

Aleksandr L'vovich Reznik (b. 1948) graduated from Novosibirsk State University in 1969. Doctor of Technical Sciences. Head of laboratory at the Institute of Automation and Electrometry, Siberian Branch, Russian Academy of Sciences (IA&E SB RAS). Scientific interests: analytical and numerical methods for solving complex probabilistic problems using computer technology. E-mail: reznik@iae.nsk.su .

Aleksandr Anatolievich Soloviev (b. 1980) graduated from Novosibirsk State University in 2004. Candidate of Technical Sciences. Now he works as a Senior Researcher at the Institute of Automation and Electrometry, Siberian Branch, Russian Academy of Sciences (IA&E SB RAS). Scientific interests: image processing, mathematical statistics. E-mail: solowey@rambler.ru .

Code of State Categories Scientific and Technical Information (in Russian – GRNTI)): 28.23.15
Received April 28, 2023. The final version – July 9, 2023.