1

The generalized scheme of allocation and the components of random graphs

1.1. The probabilistic approach to enumerative combinatorial problems

The solution to enumerative combinatorial problems consists in finding an exact or approximate expression for the number of combinatorial objects possessing the property under investigation. In this book, the probabilistic approach to enumerative combinatorial problems is adopted.

The fundamental notion of probability theory is the probability space $(\Omega, \mathcal{A}, \mathbf{P})$, where Ω is a set of arbitrary elements, $\mathcal{A}$ is a set of subsets of Ω forming a σ-algebra of events with the operations of union and intersection of sets, and $\mathbf{P}$ is a nonnegative countably additive function defined for each event $A \in \mathcal{A}$ so that $\mathbf{P}(\Omega) = 1$. The set Ω is called the space of elementary events and $\mathbf{P}$ is a probability. A random variable is a real-valued measurable function $\xi = \xi(\omega)$ defined for all $\omega \in \Omega$.

Suppose Ω consists of finitely many elements. Then the probability $\mathbf{P}$ is defined on all subsets of Ω if it is defined for each elementary event $\omega \in \Omega$. In this case, any real-valued function $\xi = \xi(\omega)$ on such a space of elementary events is a random variable.

Instead of a real-valued function, one may consider a function $f(\omega)$ taking values from some set Y of arbitrary elements. Such a function $f(\omega)$ may be considered a generalization of a random variable and is called a random element of the set Y.

In studying combinatorial objects, we consider probability spaces that have a natural combinatorial interpretation: For the space of elementary events Ω, we take the set of combinatorial objects under investigation and assign the same probability to all the elements of the set. In this case, numerical characteristics of combinatorial objects of Ω become random variables. The term "random element of the set Ω" is usually used for the identity function $f(\omega) = \omega$, $\omega \in \Omega$, mapping each element of the set of combinatorial objects into itself. Since the uniform distribution is assumed on Ω, the probability that the identity function f takes any fixed value ω is the same for all $\omega \in \Omega$. Hence the notion of a random combinatorial object of Ω, such as the identity function $f(\omega) = \omega$, agrees with the usual notion of a random element of a set as an element sampled from all elements of the set with equal probabilities.

Note that a random combinatorial object with the same distribution could also be defined on larger probability spaces. For our purposes, however, the natural construction presented here is sufficient for the most part. The exceptions are those few cases that involve several independent random combinatorial objects and in which it would be necessary to resort to a richer probability space, such as the direct product of the natural probability spaces.

Since we use probability spaces with uniform distributions, in spite of the probabilistic terminology, the problems considered are in essence enumeration problems of combinatorial analysis. The probabilistic approach furnishes a convenient form of representation and helps us effectively use the methods of asymptotic analysis that have been well developed in the theory of probability.

Thus, in the probabilistic approach, numerical characteristics of a random combinatorial object are random variables. The main characteristic of a random variable ξ is its distribution function $F(x)$ defined for any real x as the probability of the event $\{\xi \le x\}$, that is,

$$F(x) = \mathbf{P}\{\xi \le x\}.$$

The distribution function $F(x)$ defines a probability distribution on the real line called the distribution of the random variable ξ. With respect to this distribution, given a function $g(x)$, the Lebesgue–Stieltjes integral

$$\int_{-\infty}^{\infty} g(x)\,dF(x)$$

can be defined. The probabilistic approach has advantages in the asymptotic investigations of combinatorial problems. As a rule, we have a sequence of random variables ξ_n, $n = 1, 2, \ldots$, each of which describes a characteristic of the random combinatorial object under consideration, and we are interested in the asymptotic behavior of the distribution functions $F_n(x) = \mathbf{P}\{\xi_n \le x\}$ as $n \to \infty$.

A sequence of distributions with distribution functions $F_n(x)$ converges weakly to a distribution with the distribution function $F(x)$ if, for any bounded continuous function $g(x)$,

$$\int_{-\infty}^{\infty} g(x)\,dF_n(x) \to \int_{-\infty}^{\infty} g(x)\,dF(x)$$

as $n \to \infty$.

The weak convergence of distributions is directly connected with the pointwise convergence of the distribution functions as follows.

Theorem 1.1.1. *A sequence of distribution functions $F_n(x)$ converges to a distribution function $F(x)$ at all continuity points if and only if the corresponding sequence of distributions converges weakly to the distribution with distribution function $F(x)$.*

In a sense, the distribution, or the distribution function $F(x)$, characterizes the random variable ξ. The moments of ξ are simple characteristics. If

$$\int_{-\infty}^{\infty} |x| \, dF(x)$$

exists, then

$$\mathbf{E}\xi = \int_{-\infty}^{\infty} x \, dF(x)$$

is called the mathematical expectation, or mean, of the random variable ξ. Further,

$$m_r = \mathbf{E}\xi^r = \int_{-\infty}^{\infty} x^r \, dF(x)$$

is called the rth moment, or the moment of rth order (if the integral of $|x|^r$ exists).

In probabilistic combinatorics, one usually considers nonnegative integer-valued random variables. For such a random variable, the factorial moments are natural characteristics. We denote the rth factorial moment by

$$m_{(r)} = \mathbf{E}\xi(\xi - 1)\cdots(\xi - r + 1).$$

If a distribution function $F(x)$ can be represented in the form

$$F(x) = \int_{-\infty}^{x} p(u) \, du,$$

where $p(u) \geq 0$, then we say that the distribution has a density $p(u)$. In addition to the distribution function, it is convenient to represent the distribution of an integer-valued random variable ξ by the probabilities of its individual values. For ξ, we will use the notation

$$p_k = \mathbf{P}\{\xi = k\}, \quad k = 0, 1, \ldots,$$

and for integer-valued nonnegative random variables ξ_n,

$$p_k^{(n)} = \mathbf{P}\{\xi_n = k\}, \quad k = 0, 1, \ldots.$$

It is clear that

$$\mathbf{E}\xi = \sum_{n=0}^{\infty} k p_k,$$

if this series converges.

It is not difficult to see that the following assertion is true.

Theorem 1.1.2. *A sequence of distributions* $\{p_k^{(n)}\}$, $n = 1, 2, \ldots$, *converges weakly to a distribution* $\{p_k\}$ *if and only if for every fixed* $k = 1, 2, \ldots$,

$$p_k^{(n)} \to p_k$$

as $n \to \infty$.

If an estimate of the probability $\mathbf{P}\{\xi > 0\}$ is needed for a nonnegative integer-valued random variable ξ, then the simple inequality

$$\mathbf{P}\{\xi > 0\} = \sum_{k=1}^{\infty} \mathbf{P}\{\xi = k\} \leq \sum_{k=1}^{\infty} k p_k = \mathbf{E}\xi \tag{1.1.1}$$

can be useful. In particular, for a sequence ξ_n, $n = 1, 2, \ldots$, of such random variables with $\mathbf{E}\xi_n \to 0$ as $n \to \infty$, it follows that

$$\mathbf{P}\{\xi_n > 0\} \to 0.$$

Since it is generally easier to calculate the moments of a random variable than the whole distribution, one wants a criterion for the convergence of a sequence of distributions based on the corresponding moments. But, first, it should be noted that even if a random variable ξ has moments of all orders, its distribution cannot, in general, be reconstructed on the basis of these moments, since there exist distinct distributions that have the same sequences of moments. For example, it is not difficult to confirm that for any $n = 1, 2, \ldots$,

$$\int_0^{\infty} x^n e^{-1/4} \sin x^{1/4}\, dx = 0.$$

Hence, for $-1 \leq \alpha \leq 1$, the function

$$p_\alpha(x) = \tfrac{1}{24} e^{-1/4}\left(1 + \alpha \sin x^{1/4}\right)$$

is the density of a distribution on $[0, \infty)$ whose moments do not depend on α.

Thus the distribution functions with moments of all orders are divided into two classes: The first class contains the functions that may be uniquely reconstructed from their moments, and the second class contains the functions that cannot be reconstructed from their moments. There are several sufficient conditions for the moment problem to have a unique solution. Let

$$M_n = \int_{-\infty}^{\infty} |x|^n\, dF(x).$$

A distribution function $F(x)$ is uniquely reconstructed by the sequence m_r, $r = 1, 2, \ldots$, of its moments if there exists λ such that

$$\frac{1}{n} M_n^{1/n} \leq \lambda. \tag{1.1.2}$$

The following theorem describing the so-called method of moments is applicable only to the first class of distribution functions.

Theorem 1.1.3. *If distribution functions $F_n(x)$, $n = 1, 2, \dots,$ have the moments of all orders and for any fixed $r = 1, 2, \dots,$*

$$m_r^{(n)} = \int_{-\infty}^{\infty} x^r \, dF_n(x) \to m_r, \quad |m_r| < \infty,$$

as $n \to \infty$, then there exists a distribution function $F(x)$ such that for any fixed $r = 1, 2, \dots,$

$$m_r = \int_{-\infty}^{\infty} x^r \, dF(x),$$

and from the sequence $F_n(x)$, $n = 1, 2, \dots,$ it is possible to select a subsequence $F_{n_k}(x)$, $k = 1, 2, \dots,$ that converges to $F(x)$ as $n \to \infty$ at every continuity point of $F(x)$.

If the sequence m_r, $r = 1, 2, \dots,$ uniquely determines the distribution function $F(x)$, then $F_n(x) \to F(x)$ as $n \to \infty$ at every continuity point of $F(x)$.

Note that the normal (Gaussian) and Poisson distributions are uniquely reconstructible by their moments.

To use the method of moments, it is necessary to calculate moments of random variables. One useful method of calculating moments of integer-valued random variables is to represent them as sums of random variables that take only the values 0 and 1.

Theorem 1.1.4. *If*

$$S_n = \xi_1 + \cdots + \xi_n,$$

and the random variables $\xi_1, \dots, \xi_n$ take only the values 0 and 1, then for any $m = 1, 2, \dots, n$,

$$S_n(S_n - 1) \cdots (S_n - m + 1) = \sum_{\{i_1, \dots, i_m\}} \xi_{i_1} \cdots \xi_{i_m},$$

where the summation is taken over all different ordered sets of different indices $\{i_1, \dots, i_m\}$, the number of which is equal to $\binom{n}{m} m!$.

Generating functions also provide a useful tool for solving many problems related to distributions of nonnegative integer-valued random variables. The complex-valued function

$$\phi(z) = \phi_\xi(z) = \sum_{k=0}^{\infty} p_k z^k = \mathbf{E} z^\xi \qquad (1.1.3)$$

is called the generating function of the distribution of the random variable ξ. It is defined at least for $|z| \leq 1$. For example, for the Poisson distribution with parameter λ, which is defined by the probabilities

$$p_k = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, \dots,$$

the generating function is $e^{\lambda(z-1)}$.

Relation (1.1.3) determines a one-to-one correspondence between the generating functions and the distributions of nonnegative integer-valued random variables, since the distribution can be reconstructed by using the formula

$$p_k = \frac{1}{k!} \phi^{(k)}(0), \quad k = 0, 1, \dots. \tag{1.1.4}$$

Generating functions are especially convenient for the investigation of sums of independent random variables. If $\xi_1, \dots, \xi_n$ are independent nonnegative integer-valued random variables and $S_n = \xi_1 + \cdots + \xi_n$, then

$$\phi_{S_n}(z) = \phi_{\xi_1}(z) \cdots \phi_{\xi_n}(z).$$

The correspondence between the generating functions and the distributions is continuous in the following sense.

Theorem 1.1.5. *Let $\{p_k^{(n)}\}$, $n = 1, 2, \dots$, be a sequence of distributions. If for any $k = 0, 1, \dots$,*

$$p_k^{(n)} \to p_k$$

as $n \to \infty$, then the sequence of corresponding generating functions $\phi_n(z)$, $n = 1, 2, \dots$, converges to the generating function of the sequence $\{p_k\}$ uniformly in any circle $|z| \leq r < 1$.

In particular, if $\{p_k\}$ is a distribution, then the sequence of corresponding generating functions converges to the generating function $\phi(z)$ of the distribution $\{p_k\}$ uniformly in any circle $|z| \leq r < 1$.

Theorem 1.1.6. *If the sequence of generating functions $\phi_n(z)$, $n = 1, 2, \dots$, of the distributions $\{p_k^{(n)}\}$ converges to a generating function $\phi(z)$ of a distribution $\{p_k\}$ on a set M that has a limit point inside of the circle $|z| \leq 1$, then the distributions $\{p_k^{(n)}\}$ converge weakly to the distribution $\{p_k\}$.*

Since a generating function $\phi(z) = \sum_{k=0}^{\infty} p_k z^k$ is analytic, its coefficients can be represented by the Cauchy formula

$$p_n = \frac{1}{n!} \phi^{(n)}(0) = \frac{1}{2\pi i} \int_C \frac{\phi(z)\, dz}{z^{n+1}}, \quad n = 0, 1, \dots,$$

where the integral is over a contour C that lies inside the domain of analyticity of $\phi(z)$ and contains the point $z = 0$.

Thus, if we are interested in the behavior of p_n as $n \to \infty$, then we have to be able to estimate contour integrals of the form

$$G(\lambda) = \frac{1}{2\pi i} \int_C g(z) e^{\lambda f(z)}\, dz,$$

where $g(z)$ and $f(z)$ are analytic in the neighborhood of the curve of integration C and λ is a real parameter tending to infinity.

The saddle-point method is used to estimate such integrals. The contour of integration C may be chosen in different ways. The saddle-point method requires choosing the contour C in such a way that it passes through the point z_0, which is a root of the equation $f'(z) = 0$. Such a point is called the saddle point, since the function $\Re f(z)$ has a graph similar to a saddle or mountain pass. The saddle-point method requires choosing the contour of integration such that it crosses the saddle point z_0 in the direction of the steepest descent. However, finding such a contour and applying it are complicated problems, so for the sake of simplicity one usually does not choose the best contour, hence losing some accuracy in the remainder term when estimating the integral.

A parametric representation of the contour transforms the contour integral to an integral with a real variable of integration. Therefore the following theorem on estimating integrals with increasing parameters, based on Laplace's method, sometimes provides an answer to the initial question on estimating integrals.

Theorem 1.1.7. *If the integral*

$$G(\lambda) = \int_{-\infty}^{\infty} g(t) e^{\lambda f(t)}\, dt$$

converges absolutely for some $\lambda = \lambda_0$, *that is,*

$$\int_{-\infty}^{\infty} |g(t)| e^{\lambda_0 f(t)}\, dt \le M;$$

if the function $f(t)$ *attains its maximum at a point* t_0 *and in a neighborhood of this point*

$$f(t) = f(t_0) + a_2(t - t_0)^2 + a_3(t - t_0)^3 + \cdots$$

with $a_2 < 0$;

if for an arbitrary small $\delta > 0$, *there exists* $h = h(\delta) > 0$ *such that*

$$f(t_0) - f(t) \ge h,$$

for $|t - t_0| > \delta$;

and if, as $t \to t_0$,

$$g(t) = c(t - t_0)^{2m}(1 + O(|t - t_0|)),$$

where c is a nonzero constant and m is a nonnegative integer, then, as $\lambda \to \infty$,

$$G(\lambda) = e^{\lambda f(t_0)}\lambda^{-m-1/2} c c_1^{2m+1}\Gamma(m+1/2)\bigl(1+O(1/\sqrt{\lambda})\bigr),$$

where $\Gamma(x)$ is the Euler gamma function and

$$c_1 = \frac{1}{\sqrt{-a_2}} = \frac{1}{\sqrt{-f''(t_0)/2}}.$$

In particular, if $m = 0$, then $c = g(t_0)$, and as $\lambda \to \infty$,

$$G(\lambda) = e^{\lambda f(t_0)} \frac{g(t_0)}{\sqrt{-f''(t_0)/2}} \sqrt{\pi/\lambda}\bigl(1+O(1/\sqrt{\lambda})\bigr). \qquad (1.1.5)$$

To demonstrate that this rather complicated theorem can really be used, let us estimate the integral

$$\Gamma(\lambda+1) = \int_0^\infty x^\lambda e^{-x}\,dx$$

as $\lambda \to \infty$, and obtain the Stirling formula. The change of variables $x = \lambda t$ leads to the equation

$$\Gamma(\lambda+1) = \lambda^{\lambda+1}e^{-\lambda}\int_0^\infty e^{-\lambda(t-1-\log t)}\,dt.$$

Here $g(t) = 1$, and $f(t) = -(t-1-\log t)$, $f(1) = 0$, $f'(1) = 0$, $f''(1) = -1$. The conditions of the theorem are fulfilled; therefore, by (1.1.5),

$$G(\lambda) = \int_0^\infty e^{\lambda f(t)}\,dt = \sqrt{2\pi/\lambda}\bigl(1+O(1/\sqrt{\lambda})\bigr),$$

and for the Euler gamma function, we obtain the representation

$$\Gamma(\lambda+1) = \lambda^{\lambda+1/2}e^{-\lambda}\sqrt{2\pi}\bigl(1+O(1/\sqrt{\lambda})\bigr)$$

as $\lambda \to \infty$, coinciding with the Stirling formula, except for the remainder term, which can be improved to $O(1/\lambda)$.

Generating functions are only suited for nonnegative integer-valued random variables. A more universal method of proving theorems on the convergence of sequences of random variables is provided by characteristic functions. The characteristic function of a random variable ξ or the characteristic function of its distribution is defined as

$$\varphi(t) = \varphi_\xi(t) = \mathbf{E}e^{it\xi} = \int_{-\infty}^{\infty} e^{itx}\,dF(x), \qquad (1.1.6)$$

where $-\infty < t < \infty$ and $F(x)$ is the distribution function of ξ.

If the rth moment m_r exists, then the characteristic function $\varphi(t)$ is r times differentiable, and

$$\varphi^{(r)}(0) = i^r m_r.$$

Characteristic functions are convenient for investigating sums of independent random variables, since if $S_n = \xi_1 + \cdots + \xi_n$, where $\xi_1, \ldots, \xi_n$ are independent random variables, then

$$\varphi_{S_n}(t) = \varphi_{\xi_1}(t) \cdots \varphi_{\xi_n}(t).$$

The characteristic function of the normal distribution with parameters (m, σ^2) and density

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-m)^2/(2\sigma^2)}$$

is $e^{imt - \sigma^2 t^2/2}$.

Relation (1.1.6) defines a one-to-one correspondence between characteristic functions and distributions. There are different inversion formulas that provide a formal possibility of reconstructing a distribution from its characteristic function, but they have limited practical applications. We state the simplest version of the inversion formulas.

Theorem 1.1.8. *If a characteristic function $\varphi(t)$ is absolutely integrable, then the corresponding distribution has the bounded density*

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t)\, dt.$$

The correspondence defined by (1.1.6) is continuous in the following sense.

Theorem 1.1.9. *A sequence of distributions converges weakly to a limit distribution if and only if the corresponding sequence of characteristic functions $\varphi_n(t)$ converges to a continuous function $\varphi(t)$ as $n \to \infty$ at every fixed t, $-\infty < t < \infty$. In this case, $\varphi(t)$ is the characteristic function of the limit distribution, and the convergence $\varphi_n(t) \to \varphi(t)$ is uniform in any finite interval.*

For a sequence ξ_n of characteristics of random combinatorial objects, applying Theorem 1.1.9 gives the limit distribution function. But for integer-valued characteristics, one would rather have an indication of the local behavior, that is, the behavior of the probabilities of individual values. To this end the so-called local limit theorems of probability theory are used.

Let ξ be an integer-valued random variable and $p_n = \mathbf{P}\{\xi = n\}$. It is clear that $\mathbf{P}\{\xi \in \Gamma_1\} = 1$, where Γ_1 is the lattice of all integers. If there exists a lattice Γ_d with a span d such that $\mathbf{P}\{\xi \in \Gamma_d\} = 1$ and there is no lattice Γ with span greater than d such that $\mathbf{P}\{\xi \in \Gamma\}$, then d is called the maximal span of the distribution of ξ. The characteristic function $\varphi(t)$ of the random variable ξ is periodic with period $2\pi/d$ and $|\varphi(t)| < 1$ for $0 < t < 2\pi/d$.

For integer-valued random variables, the inversion formula has the following form:

$$p_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-itn}\varphi(t)\,dt.$$

Consider the sum $S_N = \xi_1 + \cdots + \xi_N$ of independent identically distributed integer-valued random variables $\xi_1, \ldots, \xi_N$. When the distributions of the summands are identical and do not depend on N, the problem of estimating the probabilities $\mathbf{P}\{S_N = n\}$, as $N \to \infty$, has been completely solved. If there exist sequences of centering and normalizing numbers A_N and B_N such that the distributions of the random variables $(S_N - A_N)/B_N$ converge weakly to some distribution, then the limit distribution has a density. Moreover, a local limit theorem holds on the lattice with a span equal to the maximal span of the distribution of the random variable ξ_1. If the maximal span of the distribution of ξ_1 is 1, then the local theorem holds on the lattice of integers.

Theorem 1.1.10. *Let $\xi_1, \xi_2, \ldots$ be a sequence of independent identically distributed integer-valued random variables and let there exist A_N and B_N such that, as $N \to \infty$ for any fixed x,*

$$\mathbf{P}\left\{\frac{S_N - A_N}{B_N} \le x\right\} \to \int_{-\infty}^{x} p(u)\,du.$$

Then, if the maximal span of the distribution of ξ_1 is 1,

$$B_N\mathbf{P}\{S_N = n\} - p\big((n - A_N)/B_N\big) \to 0$$

uniformly in n.

Local limit theorems are of primary importance in what follows. Therefore, let us prove a local theorem on convergence to the normal distribution as a model for proofs of local limit theorems in more complex cases, which will be discussed later in the book.

Theorem 1.1.11. *Let the independent identically distributed integer-valued random variables $\xi_1, \xi_2, \ldots$ have a mathematical expectation a and a positive variance σ^2. Then, if the maximal span of the distribution of ξ_1 is 1,*

$$\sigma\sqrt{N}\mathbf{P}\{\xi_1 + \cdots + \xi_N = n\} - \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{(n - aN)^2}{2\sigma^2 N}\right\} \to 0$$

uniformly in n as $N \to \infty$.

Proof. Let

$$z = \frac{n - aN}{\sigma\sqrt{N}} \quad \text{and} \quad P_N(n) = \mathbf{P}\{\xi_1 + \cdots + \xi_N = n\}.$$