

# Sums of independent random variables and some combinatorial problems

V. Kolchin

Steklov Mathematical Institute

September 17, 1992

[summary by Philippe Robert]

## 1. Introduction

Consider  $n$  independent random variables uniformly distributed on the set  $\{1, 2, \dots, N\}$  and denote by  $\eta_i$  the number of occurrences of  $i$ ,  $1 \leq i \leq N$ . For any  $N$ -tuple of integers  $n_1, n_2, \dots, n_N$  such that  $\sum_1^N n_i = n$ , then

$$P(\eta_1 = n_1, \dots, \eta_N = n_N) = \frac{n!}{n_1! \dots n_N! N^n}.$$

In the language of allocation of particles into cells (or balls into urns):  $n$  particles are put at random into  $N$  cells,  $\eta_i$  is the number of particles in the  $i$ -th cell.

If  $\xi_1, \dots, \xi_N$  are independent Poisson random variables with parameter  $\lambda$ ,

$$p_k = P(\xi_1 = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, \dots$$

it is easy to check that

$$(1) \quad P(\eta_1 = n_1, \dots, \eta_N = n_N) = P(\xi_1 = n_1, \dots, \xi_N = n_N / \xi_1 + \dots + \xi_N = n).$$

Let  $\mu_r(n, N)$  be the number of cells with exactly  $r$  particles and  $\eta_{(N)} = \max\{\eta_1, \dots, \eta_N\}$ . If  $\xi_1^{(r)}, \dots, \xi_N^{(r)}$  are independent identically distributed (i.i.d.) random variables such that

$$P(\xi_1^{(r)} = k) = P(\xi_1 = k / \xi_1 \neq r),$$

and  $\bar{\xi}_1^{(r)}, \dots, \bar{\xi}_N^{(r)}$  i.i.d variables with

$$P(\bar{\xi}_1^{(r)} = k) = P(\xi_1 = k / \xi_1 \leq r),$$

then using the relation (1), one gets

$$\begin{aligned} P(\eta_{(N)} \leq r) &= (P(\xi_1 \leq r))^N \frac{P(\bar{\xi}_1^{(r)} + \dots + \bar{\xi}_N^{(r)} = n)}{P(\xi_1 + \dots + \xi_N = n)}, \\ P(\mu_r(n, N) = k) &= \binom{N}{k} p_r^k (1 - p_r)^{N-k} \frac{P(\xi_1^{(r)} + \dots + \xi_{N-k}^{(r)} = n - kr)}{P(\xi_1 + \dots + \xi_N = n)}. \end{aligned}$$

**DEFINITION 1.** An  $N$ -tuple  $\eta_1, \dots, \eta_N$  of random variables is a generalized scheme of allocating particles if there exist random variables  $\xi_1, \dots, \xi_N$  such that (1) is satisfied.

EXAMPLE. We consider the partitions of integers  $n$  into  $N$  non-negative integer summands,  $n = n_1 + \dots + n_N$ . There are  $\binom{n - N + 1}{N - 1}$  such partitions; if they are equally likely, then  $n = \eta_1 + \dots + \eta_N$ . If we take independent geometrically distributed random variables  $\xi_1, \dots, \xi_N$ ,

$$P(\xi_1 = k) = p^k(1 - p), \quad k \in \mathbb{N}, \quad 0 < p < 1,$$

then relation (1) is satisfied.

## 2. A general model of application of the generalized scheme of allocation

Let

- $\Gamma_n(R)$  be the set of all graphs with  $n$  vertices which satisfy some property  $R$ ,
- $\Gamma_{n,N}(R)$  the set of the elements of  $\Gamma(R)$  with  $N$  connected components,
- $\bar{\Gamma}_{n,N}(R)$  the set of objects which consists of ordered collections on  $N$  components.

Take the uniform distribution on  $\bar{\Gamma}_{n,N}(R)$  and denote by  $\eta_1, \dots, \eta_N$  the sizes of the components of a random element from  $\bar{\Gamma}_{n,N}(R)$ . Denote by  $a_n, a_{n,N}, \bar{a}_{n,N}$  the respective cardinalities of  $\Gamma_n(R), \Gamma_{n,N}(R), \bar{\Gamma}_{n,N}(R)$ ,  $b_n$  the number of connected graphs in  $\Gamma_n(R)$  and let  $\xi_1, \dots, \xi_N$  be i.i.d. random variables such that

$$P(\xi_1 = k) = \frac{b_k x^k}{k! B(x)}, \quad k \in \mathbb{N},$$

where  $B(x) = \sum_1^{+\infty} \frac{b_k x^k}{k!}$  and  $x$  is in the domain of convergence of this series. Then relation (1) is valid:

$$a_{n,N} = \frac{\bar{a}_{n,N}}{N!} = \frac{1}{N!} \sum_{n_1 + \dots + n_N = n} \frac{n!}{n_1! \dots n_N!} b_1 \dots b_{n_N},$$

$$\begin{aligned} P(\xi_1 + \dots + \xi_N = n) &= \sum_{n_1 + \dots + n_N = n} \prod_{i=1}^N \frac{b_{n_i} x^{n_i}}{n_i! B(x)} \\ &= \frac{x^n}{B(x)^N} \sum_{n_1 + \dots + n_N = n} \prod_{i=1}^N \frac{b_{n_i}}{n_i!}, \end{aligned}$$

hence

$$(2) \quad a_{n,N} = \frac{n!(B(x))^N}{N! x^n} P(\xi_1 + \dots + \xi_N = n).$$

EXAMPLE. 1) Random permutations from  $S_n$ .

$$a_n = n!, \quad b_n = (n - 1)!,$$

$$P(\xi_1 = k) = \frac{-x^k}{k \log(1 - x)}, \quad k \in \mathbb{N}, \quad 0 < x < 1.$$

2) Random mappings from  $\Sigma_n$ .

$$a_n = n^n, \quad b_n = (n - 1)! \sum_{k=0}^{n-1} \frac{n^k}{k!},$$

$$P(\xi_1 = k) = \frac{b_k x^k}{k! B(x)}, \quad 0 < x < 1.$$

3) Random partitions from the set unordered partitions of the set  $\{1, \dots, n\}$ .

$$b_n = 1,$$

$$a_n = \sum_{N=1}^n a_{n,N} = \sum_{N=1}^n \frac{n!}{N!} \sum_{n_1+\dots+n_N=n} \frac{1}{n_1! \cdots n_N!},$$

$$P(\xi_1 = k) = \frac{x^k}{k!(e^x - 1)}, \quad k \in \mathbb{N}, \quad 0 < x < +\infty.$$

4) Random forest from the set of all forests of  $N$  non-rooted trees with  $n$  total number of vertices.

$$b_n = n^{n-2}, \quad B(x) = \sum_1^{+\infty} \frac{n^{n-2} x^n}{n!}, \quad 0 < x < e^{-1},$$

$$P(\xi_1 = k) = \frac{k^{k-2} x^k}{k! B(x)}, \quad k \in \mathbb{N}.$$

The complete investigation of  $a_{n,N}$  was carried out by Britikov in 1990,

$$E(\xi_1) = \frac{1}{B(x)} \sum_1^{+\infty} \frac{k^{k-1} x^k}{k!} = \frac{\theta(x)}{B(x)},$$

$$\sigma^2 E(\xi_1^2) = \frac{1}{B(x)} \sum_1^{+\infty} \frac{k^k x^k}{k!} = \frac{A(x)}{B(x)},$$

where

$$A(x) = \frac{\theta(x)}{1 - \theta(x)}, \quad B(x) = \frac{1}{2}(1 - (1 - \theta(x))^2)$$

and  $\theta(x)$  is the root of  $x = \theta e^{-\theta}$  in the interval  $[0, 1]$ .

As a consequence of the local central limit theorem (see [1] p. 233),

$$P(\xi_1 + \dots + \xi_N = k) = \frac{1}{\sigma \sqrt{2\pi N}} e^{-u^2/2} (1 + o(1)),$$

uniformly on the integers  $k$  such that  $u = \frac{(k - NE(\xi_1))}{\sigma \sqrt{N}}$  is in a finite interval.

The number of edges in the forest is  $T = n - N$ , if  $\theta = \frac{2T}{n}$ , then  $E(\xi_1) = \frac{n}{N}$ ,

$$P(\xi_1 + \dots + \xi_N = n) = \frac{1}{\sigma \sqrt{2\pi N}} (1 + o(1)).$$

Finally if  $n, N \rightarrow +\infty$  such that  $\theta = \frac{2T}{n}$  is constant, then using (2) one gets

$$a_{n,T} = \frac{n^{2T} \sqrt{1 - \theta}}{2^T T!} (1 + o(1))$$

### Bibliography

- [1] Gnedenko (B. V.), Kolmogorov (A. N.), and Chung (K. L.). – *Limit Distributions for Sums of Independent Random Variables*. – Addison-Wesley, 1968.