Liet. matem. rink., T. 40, spec. nr., 55–60

On the frequency of multisets without some components

Eugenijus MANSTAVIČIUS (VU)
e-mail: eugenijus.manstavicius@maf.vu.lt

We examine unordered samples $\sigma := \{p_1, \ldots, p_w\}$, $w \geqslant 1$, taken with replacement from some set $\mathcal{P}$ of elements which have sizes belonging to the set of natural numbers. Thus, $p_i \in \mathcal{P} = \cup_{j \geqslant 1} \mathcal{P}_j$, $1 \leqslant i \leqslant w$, where $\mathcal{P}_j$ is the set of elements having size j. We assume that $|\mathcal{P}_j| =: \pi(j) < \infty$ and define the size of σ as sum of the sizes of its elements. From now on, following [1], [9], we call σ a *multiset*. Suppose $p(N)$ denotes the number of multisets of size N, $p(0) = 1$. Ascribing the probability $\nu_N(\{\sigma\}) = 1/p(N)$ for each multiset of size N, we obtain a finite probability space. Our interest now lays in the probability of multisets with *a fortiori* given properties. In the present remark, to such results obtained in the papers [1], [5], and [9] we add a lower estimate of the frequency of random multisets which do not contain elements of sizes belonging to some set of natural numbers. The similar problem for random permutations has been dealt with in the author's paper [6] which was further extended [8] for general labeled combinatorial structures called *assemblies*. The multisets do not belong to this class.

Observe that σ may also be understood as the formal symbolic product $p_1 \cdots p_w$. So, taking this point of view, we could reduce our problem to investigating of elements of an additive arithmetical semigroup (see [3] or [5] for the definitions). In this way we could later use the sieve ideas and technical details comming from the paper [2], nevertheless we now have a shorter proof. The desired estimate for multisets will be derived directly from an appropriate result for assemblies.

Let $k_j(\sigma)$ be the multiplicities of elements of size j in σ, having size N, $1 \leqslant j \leqslant N$ and $L(\bar{k}) := 1k_1 + \cdots + Nk_N$ for $\bar{k} = (k_1, \ldots, k_N) \in \mathbb{Z}^{+N}$. Thus $L(\bar{k}(\sigma)) = N$ and

$$\nu_N(k_j(\sigma) = k_j,\ 1 \leqslant j \leqslant N) = \frac{\mathbf{1}(L(\bar{k}) = N)}{p(N)} \prod_{j=1}^{N} \binom{\pi(j) + k_j - 1}{k_j}, \tag{1}$$

where $k_j \in \mathbb{Z}^+$, $1 \leqslant j \leqslant N$. Direct calculation (see [1]) shows that this frequency is also equal to the probability

$$P(\eta_j = k_j,\ 1 \leqslant j \leqslant N \mid L(\bar{\eta}) = N), \tag{2}$$

where η_j, $1 \leqslant j \leqslant N$, are independent negative binomial random variables defined on some probability space by

$$P(\eta_j = k) = \binom{\pi(j) + k - 1}{k} (1 - x^j)^{\pi(j)} x^{jk}, \quad k \geqslant 0, \tag{3}$$

where $0 < x < 1$. As we have shown in [6], such conditioning of independent events substitutes well the typical argumentation in the small sieve approach.

Let $J \subset \{1, \ldots, N\}$ be arbitrary, maybe, depending on N subset, $\nu_N(J) := \nu_N\big(k_j(\sigma) = 0 \ \forall j \in J\big)$. In other words, $\nu_N(J)$ is the frequency of the multisets which does not have any element of size belonging to the set J. One can try to find the asymptotic behaviour of $\nu_N(J)$ as $N \to \infty$ in terms of the sum

$$S(J) := \sum_{j \in J} 1/j.$$

By (1) we have

$$\nu_N(J) = \frac{1}{p(N)} \sum_{\substack{L(\bar{k})=N \\ k_j=0 \, \forall j \in J}} \prod_{j=1}^{N} \binom{\pi(j) + k_j - 1}{k_j}$$

and the following formal identity

$$\sum_{N=0}^{\infty} \nu_N(J) p(N) z^N = \prod_{j \geqslant 1, j \notin J} (1 - z^j)^{-\pi(j)}.$$

Thus analytic methods based on Cauchy's theorem, such as Theorem 3 of [7], can be applied to this problem. On this way, certain requirements on reqularity of J or some bounds for the number of its elements are unavoidable. So, if $S(J)$ is small enough, we can get the estimate $\nu_N(J) \asymp \mathrm{e}^{-S(J)}$. However there exist instances when such relations take considerably different form. Consider the arithmetical semigroups as in [4] (the multiset of monic polynomials over a finite field is a particular case) and take $J = (\mathrm{e}^{-K}N, N]$ with $S(J) \sim K$. Corollary 1 in [4] gives the following formula

$$\nu_N(J) = \exp\left\{-K\mathrm{e}^K \left(1 + \mathrm{O}\left(\frac{\log K}{K}\right)\right)\right\} \left(1 + \mathrm{O}\left(\frac{K\mathrm{e}^{2K}}{N}\right)\right)$$

provided that K is large enough and $N \to \infty$. See [8] for more detailed comments. That leads to the question:

What are bounds for $\nu_N(J)$ in terms of $S(J)$?

Set $\mu_N(K) = \min_{S(J) \leqslant K} \nu_N(J)$, $K \geqslant 0$, where minimum is taken over all subsets J satisfying the written condition.

Our purpose is to obtain a lower estimate of $\liminf_{N\to\infty} \mu_N(K) =: \mu(K)$. Having in mind the most popular examples of multisets and our intentions to exploit the results of the paper [8], in what follows we assume

Condition M. *There exist fixed parameters* $q > 1, \theta > 0$ *such that* $\lambda_j := \pi(j)q^{-j} \geqslant c/j$ *with* $c > 0$ *and* $|\lambda_j - \theta j^{-1}| \leqslant \rho(j)$, $j \geqslant 1$, *where* $\rho : \mathbb{R}^+ \to \mathbb{R}^+$ *is a monotonically decreasing function with the properties*

$$\rho(u) \leqslant Cu^{-1}, \qquad \rho(u/2) \leqslant C_1\rho(u), \quad \textit{and} \quad \sum_{j\geqslant 1} \rho(j) \leqslant C_2,$$

where C, C_1, C_2 *are positive constants.*

In the sequel the constants in the symbol $O(\cdot)$ as well as the newly involved c_i, C_j will depend on q, θ, c, C, C_1, and C_2 only.

Theorem. *For the class of multisets satisfying condition M, we have*

$$\mu(K) \geqslant c_0 \exp\{-e^{C_0 K}\}$$

for all $K \geqslant 0$.

Proof. We use independent r.vs η_j, $1 \leqslant j \leqslant N$, defined above by (3) with $x = q^{-1}$. Denote $Y = L(\bar{\eta})$, $\bar{J} = \{1, \ldots, N\} \setminus J$,

$$Y_J = \sum_{j\in J} j\eta_j, \quad \bar{Y}_J = Y - Y_J.$$

From (1) and (2) it follows that

$$\nu_N(J) = P(Y_J = 0\,|Y = N) = \frac{P(Y_J = 0)P(\bar{Y}_J = N)}{P(Y = N)}. \tag{4}$$

The probabilities appearing on the right hand side of (4) will be compared to that of independent Poisson r.vs ξ_j, $1 \leqslant j \leqslant N$, with parameters $\mathbf{E}\xi_j = \lambda_j$. Set as above $X = L(\bar{\xi})$,

$$X_J = \sum_{j\in J} j\xi_j, \quad \bar{X}_J = X - X_J.$$

Via relevant conditioning, interpreting Theorem 2 of [8] probabilistically, we have

Lemma. *If Condition M is satisfied, then*

$$\begin{aligned} P_N(J) :&= P(X_J = 0\,|X = N) \\ &= \frac{P(X_J = 0)P(\bar{X}_J = N)}{P(X = N)} \geqslant c_1 \exp\left\{-e^{C_3 K}\right\}. \end{aligned} \tag{5}$$

Define

$$H(z;A)=\prod_{j\in A}\left(\left(1-(zq^{-1})^j\right)^{-\pi(j)}\mathrm{e}^{-\lambda_j z^j}\right),\quad A\subset\mathbb{N}.$$

Expanding the logarithm of this function we observe that, for arbitrary A, $H(z;A)$ is analytic in $|z|<\sqrt{q}$ and the Taylor coefficients of it are positive, the first of them beeing one. Set $h_0=1$ and

$$H(z;\bar{J})=:1+\sum_{k\geqslant 1}h_k z^k.$$

By Cauchy's formula we obtain

$$\begin{aligned}P(\bar{Y}_J=N)&=H(1;\bar{J})^{-1}\frac{1}{2\pi i}\int_{|z|=1}\exp\left\{\sum_{j\in\bar{J}}\lambda_j(z^j-1)\right\}\frac{H(z;\bar{J})\mathrm{d}z}{z^{N+1}}\\&=H(1;\bar{J})^{-1}\sum_{k=0}^{N}h_kP(\bar{X}_J=N-k)\geqslant H(1;\bar{J})^{-1}P(\bar{X}_J=N).\end{aligned}\tag{6}$$

Moreover, we have

$$P(Y_J=0)=H(1;J)^{-1}P(X_J=0).\tag{7}$$

As it follows from Lemma 3 and formula (6) of the paper [8], if Condition M is satisfied, then

$$P(X=N)=\frac{\mathrm{e}^{-\theta\gamma}(1+\mathrm{o}(1))}{\Gamma(\theta)N}\tag{8}$$

as $N\to\infty$. Here γ is the Euler constant and Γ denotes the Euler function.

We now seek for an analogous relation for the probability $P(Y=N)$. It is just the N-th Taylor coefficient of the generating function

$$\prod_{j=1}^{N}(1-q^{-j})^{\pi(j)}\prod_{j=1}^{N}\left(1-(zq^{-1})^j\right)^{-\pi(j)}$$

or that of the function

$$\prod_{j=1}^{N}(1-q^{-j})^{\pi(j)}(1-z)^{-\theta}\exp\left\{\sum_{j\geqslant 1}(\lambda_j-\theta/j)z^j\right\}H(z;\mathbb{N})\tag{9}$$

defined in $|z|<1$. We now use the convolution argument.

It is known that

$$(1-z)^{-\theta} = \sum_{k\geqslant 0} \binom{\theta+k-1}{k} z^k,$$
$$\binom{\theta+k-1}{k} = \frac{k^{\theta-1}}{\Gamma(\theta)}\left(1+\mathrm{O}(k^{-1})\right), \quad k \geqslant 1. \tag{10}$$

Set

$$\exp\Big\{\sum_{j\geqslant 1}(\lambda_j - \theta/j)z^j\Big\} = 1 + \sum_{s\geqslant 1} r_s z^s.$$

If Condition M holds, from Lemma 2 [8] we obtain that the last series is absolutely convergent at the point $z = 1$ and $|r_s| = \mathrm{O}(s^{-1})$ for $s \geqslant 1$. The Taylor coefficients g_l, $l \geqslant 0$ of the analytic in $|z| < \sqrt{q}$ function $H(z;\mathbb{N})$ satisfy the estimate $|g_l| = \mathrm{O}(\alpha^{-l})$, $l \geqslant 0$, with some $1 < \alpha < \sqrt{q}$, and $g_0 = 1$. Hence for the m-th Taylor coefficient f_m of the product of the last two functions in (9), we obtain

$$|f_m| \leqslant \sum_{s=0}^{m} |r_s||g_{m-s}| = \mathrm{O}(m^{-1}), \quad f_0 = 1,$$

and also convergence of the series with the summands $|f_m|, m \geqslant 0$. This suffices to show that

$$Q_N := \sum_{k=0}^{N} \binom{\theta+k-1}{k} f_{N-k} = \frac{N^{\theta-1}}{\Gamma(\theta)}\Big(\sum_{m\geqslant 1} f_m + \mathrm{o}(1)\Big) \tag{11}$$

as $N \to \infty$. We discuss only the more difficult case $0 < \theta < 1$. Let $M \geqslant 1$ and $0 < \varepsilon < 1$ be arbitrary fixed constants, $M \leqslant \varepsilon N$. By virtue of (10) and the other observations above we obtain

$$\begin{aligned} Q_N &= \Big(\sum_{0\leqslant k\leqslant M} + \sum_{M<k\leqslant \varepsilon N} + \sum_{\varepsilon N<k\leqslant N-M} + \sum_{N-M<k\leqslant N}\Big)\binom{\theta+k-1}{k} f_{N-k} \\ &= \mathrm{O}\Big(\frac{M}{N} + \varepsilon^{\theta}N^{\theta-1} + \varepsilon^{1-\theta}N^{\theta-1}\sum_{k\geqslant M}|f_k|\Big) \\ &\quad + \frac{N^{\theta-1}}{\Gamma(\theta)}\Big(1+\mathrm{O}\Big(\frac{M}{N}\Big)\Big)\Big(\sum_{k\geqslant 0} f_k + \mathrm{O}\Big(\sum_{k\geqslant M}|f_k|\Big)\Big). \end{aligned}$$

Taking $N \to \infty$, $M \to \infty$, and $\varepsilon \to 0$ in the relevant order, we complete the proof of (11).

Now, from (9) and (11) we derive the desired asymptotic formula

$$P(Y = N) = \prod_{j=1}^{N}(1-q^{-j})^{\pi(j)} Q_N = \prod_{j=1}^{N}(1-q^{-j})^{\pi(j)} \frac{N^{\theta-1}}{\Gamma(\theta)}\Big(\sum_{k\geqslant 0} f_k + \mathrm{o}(1)\Big)$$

$$= \prod_{j=1}^{N}(1-q^{-j})^{\pi(j)}\frac{N^{\theta-1}}{\Gamma(\theta)}\exp\Big\{\sum_{j\geqslant 1}\Big(\lambda_j-\frac{\theta}{j}\Big)\Big\}H(1;\mathbb{N})(1+\mathrm{o}(1))$$
$$= \frac{\mathrm{e}^{-\theta\gamma}(1+\mathrm{o}(1))}{\Gamma(\theta)N}. \qquad (12)$$

Inserting the estimates (6), (7), and (12) into the expression (4), using (8) and relation (5) in Lemma, we obtain

$$\nu_N(J) \geqslant H(1;J\cup\bar{J})^{-1}P(X_J=0)P(\bar{X}_J=N)\Big(\frac{\mathrm{e}^{-\gamma\theta}}{\Gamma(\theta)N}\Big)^{-1}(1+\mathrm{o}(1))$$
$$= H(1;\mathbb{N})^{-1}P_N(J)(1+\mathrm{o}(1)).$$

This by Lemma completes the proof of Theorem.

An analogue of the inequality in Theorem is also true for selections, e.g., the samples σ where no repetition of its elements is allowed.

References

[1] R. Arratia and S. Tavaré, Independent process approximations for random combinatorial structures, *Adv. Math.*, **104**(1), 90–154 (1994).

[2] P. Erdös and I.Z. Ruzsa, On the small sieve. I. Sifting by primes, *J. Number Th.*, **12**, 385–394 (1980).

[3] J. Knopfmacher, *Analytic Arithmetic of Algebraic Function Fields*, Lecture Notes in Pure Appl. Math., **50**, Marcel Dekker, New York (1979).

[4] E. Manstavičius, Remarks on the semigroup elements free of large prime factors, *Lith. Math. J.*, **32**(4), 400–409 (1992).

[5] E. Manstavičius, On analytic problems of combinatorial structures, in *Trans. Lith. Math. Soc.*, **III**, 75–80 (1999).

[6] E. Manstavičius, On random permutations without cycles of some lengths, *Periodica Math. Hung.* (to appear, 2000).

[7] E. Manstavičius, Decomposable mappings on combinatorial structures. Analytic approach, Preprint 98-15, Vilnius University, Department of Mathematics (1998).

[8] E. Manstavičius, On probability of combinatorial structures without some components, Preprint 2000-5, Vilnius University, Department of Mathematics and Informatics (2000).

[9] D. Stark, Total variation asymptotics for independent process approximations of logarithmic multisets and selections, *Random Structures and Algorithms*, **11**(1), 51–80 (1997).

Kartotinių aibių be tam tikrų komponenčių dažnis

E. Manstavičius

Nagrinėjamos atsitiktinės didėjančios kartotinės aibės, neturinčios elementų, kurių svoriai priklauso tam tikram poaibiui. Iš apačios įvertintas tokių kombinatorinių struktūrų dažnis. Palyginami sąlyginiai neigiamų binominių atsitiktinių dydžių bei Poissono dydžių skirstiniai.