On the Multiplicity of Parts in a Random Partition

Sylvie Corteel *
Laboratoire de Recherche en Informatique
Bât 490, Université Paris-Sud
91405 Orsay, FRANCE
Sylvie.Corteel@lri.fr

Boris Pittel †
Department of Mathematics
Ohio State University
Columbus, Ohio 43210
bgp@math.ohio-state.edu

Carla D. Savage ‡
Department of Computer Science
North Carolina State University
Raleigh, NC 27695-8206
cds@cayley.csc.ncsu.edu

Herbert S. Wilf
Department of Mathematics
University of Pennsylvania
Philadelphia, PA 19104-6395
wilf@math.upenn.edu

June 10,1998

Abstract

Let λ be a partition of an integer n chosen uniformly at random among all such partitions. Let $s(\lambda)$ be a part size chosen uniformly at random from the set of all part sizes that occur in λ. We prove that, for every fixed $m \geq 1$, the probability that $s(\lambda)$ has multiplicity m in λ approaches $1/(m(m+1))$ as $n \to \infty$. Thus, for example, the limiting probability that a random part size in a random partition is unrepeated is $1/2$.

In addition, (a) for the average number of different part sizes, we refine an asymptotic estimate given by Wilf, (b) we derive an asymptotic estimate of the average number of parts of given multiplicity m, and (c) we show that the expected multiplicity of a randomly chosen part size of a random partition of n is asymptotic to $(\log n)/2$.

The proofs of the main result and of (c) use a conditioning device of Fristedt.

AMS Subject Classifications: 05A17, 11P81, 11P82

*Supported in part by NSF grant DMS9302505

†Supported by the Focus on Discrete Probability Program at Dimacs Center of Rutgers University, and by NSA grant MDA-904-96-1-0053

‡Supported in part by NSF grant DMS96227722

1 Notation

We will be studying several counting functions related to integer partitions. For clarity, we begin with the definitions and notation.

- If n is a positive integer, then by *a partition*, λ, of n, we mean a representation

$$\lambda : \; n = \sum_i i\mu(i), \tag{1}$$

  in which the μ's are nonnegative integers. We use $\Lambda(n)$ to denote the set of all partitions of n and let $p(n) = |\Lambda(n)|$.

- In (1) the quantity $\mu(i)$ is *the multiplicity* of the part i in the partition. If it is important to designate the partition explicitly then we will use $\mu_\lambda(i)$ for the multiplicity of the part i in the partition λ.

- The *number of parts* in the partition λ in (1) is $\kappa(\lambda) = \sum_i \mu_\lambda(i)$.

- $p(n,k)$ will be the number of partitions of n with $\kappa(\lambda) = k$.

- The number of *distinct part sizes* in the partition λ is $\delta(\lambda) = |\{i : \mu_\lambda(i) > 0\}|$.

- $p_\delta(n,k)$ will be the number of partitions of n with $\delta(\lambda) = k$.

- We will use angle brackets $\langle \ldots \rangle$ to denote averages of these quantities. In particular, $\langle\kappa\rangle_n$ and $\langle\delta\rangle_n$ will denote the *average* values of $\kappa(\lambda)$ and $\delta(\lambda)$ among all partitions of n. That is, $\langle\kappa\rangle_n = \sum_k k \cdot p(n,k)/p(n)$ and $\langle\delta\rangle_n = \sum_k k \cdot p_\delta(n,k)/p(n)$.

- $\delta_m(\lambda) = |\{i : \mu_\lambda(i) = m\}|$ will be the number of part sizes i of a given multiplicity m, and $p(n,j,m)$ will be the number of partitions λ of n with $\delta_m(\lambda) = j$.

- $\langle\delta_m\rangle_n$ will be the average value, over all partitions λ of n, of $\delta_m(\lambda)$; that is $\langle\delta_m\rangle_n = \sum_j j \cdot p(n,j,m)/p(n)$.

- Finally, $[x^n]\{\ldots\}$ will denote the coefficient of x^n in the expression "...".

2 The number of different part sizes

The asymptotic relation

$$\langle\delta\rangle_n \sim \frac{\sqrt{6}}{\pi} n^{\frac{1}{2}} \qquad (n \to \infty) \tag{2}$$

for the average number of different part sizes in a partition on n is well known and can be found in [15]. (Erdös and Lehner [3] proved that the number of different part sizes ($\delta(\lambda)$, that is) lies between $(1 \pm \epsilon)\sqrt{6n}/\pi$, for almost all partitions λ.)

Here, following [15], we will first prove a simple identity involving this function, and second we will show that with the aid of that identity one can find as many terms of the asymptotic expansion for $\langle\delta\rangle_n$ as one wishes. We illustrate by displaying one more term of its asymptotic expansion, namely

$$\langle\delta\rangle_n = \frac{\sqrt{6}}{\pi} n^{\frac{1}{2}} + \left(\frac{3}{\pi^2} - \frac{1}{2}\right) + o(1) \qquad (n \to \infty). \tag{3}$$

We claim that the sum that appears in the numerator of the computation of the average number of different part sizes, namely $\sum_k k \cdot p_\delta(n,k)$ is just

$$\sum_{k\geq 1} k \cdot p_\delta(n,k) = p(0) + p(1) + p(2) + \ldots + p(n-1). \tag{4}$$

This identity is well known. It is mentioned in [1], and attributed to Stanley. To prove it combinatorially, we can map

$$\Lambda(0) \cup \Lambda(1) \cup \ldots \cup \Lambda(n-1) \Longrightarrow \Lambda(n)$$

by sending each partition $\lambda \in \Lambda(i)$ to the partition $\lambda \in \Lambda(n)$ with one additional part '$n-i$' adjoined. We will then find that each partition in $\Lambda(n)$ with k different part sizes is the image of exactly k different partitions from the union, namely those that one finds by removing exactly one copy of any of its k different part sizes, which proves (4). □

3 Asymptotic consequences

In view of (4) we have that the average number of different part sizes in a partition of the integer n is exactly

$$\langle\delta\rangle_n = \frac{p(0) + p(1) + \ldots + p(n-1)}{p(n)}. \tag{5}$$

But the complete asymptotic (and convergent) series for the partition function $p(n)$ is known, thanks to the pioneering work of Hardy and Ramanujan [5], and Rademacher [12]. Thus by (5), we can obtain arbitrarily many terms of the asymptotic series for $\langle\delta\rangle_n$ by a simple exercise in summation. We will sketch one step of this process.

For $p(n)$ itself we have

$$p(n) = \frac{1}{4\sqrt{3}} \frac{e^{\pi\sqrt{\frac{2}{3}n}}}{n} - \frac{1}{4\sqrt{2}\pi} \frac{e^{\pi\sqrt{\frac{2}{3}n}}}{n^{\frac{3}{2}}} + O\left(e^{\frac{\pi}{2}\sqrt{\frac{2}{3}n}}\right). \tag{6}$$

To prove (3), we have first, by the Euler-MacLaurin sum formula,

$$\sum_{j=1}^{n} \frac{e^{\alpha\sqrt{j}}}{j^\beta} = e^{\alpha\sqrt{n}} \left\{ \frac{2}{\alpha n^{\beta-\frac{1}{2}}} + \left(\frac{2(2\beta-1)}{\alpha^2} + \frac{1}{2}\right)\frac{1}{n^\beta} + O(n^{-\beta-\frac{1}{2}}) \right\}. \tag{7}$$

Now if we use (7) to sum (6) and thereby estimate (5), the claimed result (3) follows.

4 The number of parts of multiplicity m in a partition

Our goal in this section is to characterize the counting functions $p(n,j,m)$ and $\langle \delta_m \rangle_n$. To this end, for fixed m, we will find the two-variable generating function of $p(n,j,m)$.

Fix n and some set $1 \le i_1 < i_2 < \ldots < i_j$. Following the paradigm of inclusion-exclusion, we want the number of partitions of n that have *at least* $i_1, \ldots, i_j$ as parts of multiplicity m. That is to say, we want the number of partitions λ of n such that $i_1, \ldots, i_j$ do occur as parts of λ, and their multiplicities are all m.

But that is the number of partitions of the integer $n - mi_1 - \ldots - mi_j$ into parts that do not include any of $i_1, \ldots, i_j$ but are otherwise unrestricted. This number is evidently

$$\begin{aligned}
[x^{n-mi_1-\ldots-mi_j}] \prod_{i \neq i_1, \ldots, i \neq i_j} \frac{1}{1-x^i} &= [x^{n-mi_1-\ldots-mi_j}] \frac{\prod_{\nu=1}^{j}(1-x^{i_\nu})}{\prod_{i \ge 1}(1-x^i)} \\
&= [x^n] \frac{\prod_{\nu=1}^{j} x^{mi_\nu}(1-x^{i_\nu})}{\prod_{i \ge 1}(1-x^i)} \\
&= [x^n]\mathcal{P}(x) \left\{ \prod_{\nu=1}^{j} x^{mi_\nu}(1-x^{i_\nu}) \right\}
\end{aligned} \tag{8}$$

where $\mathcal{P}(x) = \prod_k (1-x^k)^{-1}$.

Note that the factor in brackets contributes 1 to the coefficient of t^j in the product

$$\prod_{i \ge 1} (1 + tx^{mi}(1-x^i)).$$

If we now sum over all sets $1 \le i_1 < i_2 < \ldots < i_j$ we obtain

$$p(n,j,m) = [x^n]\mathcal{P}(x) \cdot [t^j] \prod_{i \ge 1} (1 + tx^{mi}(1-x^i)).$$

The principle of inclusion-exclusion can, however, be stated in the following form (here we follow [14]): if $h(t)$ is the generating function for the numbers N_r, defined to be the sum over all sets of r properties of the number of objects that have *at least* that set of properties, then $h(t-1)$ is the generating function for the number of objects that have *exactly* each number of properties.

Hence we have the following result.

Theorem 1 *If $p(n,j,m)$ is the number of partitions of n that have exactly j parts of multiplicity m, then*

$$\sum_{n,j \ge 0} x^n t^j \cdot p(n,j,m) = \mathcal{P}(x) \prod_{i \ge 1} \{1 + (t-1)x^{mi}(1-x^i)\}. \tag{9}$$

We now use this generating function to find an interesting identity. We claim that

$$\sum_{j\geq 1} j\cdot p(n,j,m) = \sum_{k\geq 0}(p(n-mk)-p(n-(m+1)k)) \tag{10}$$

in which $p(\cdot)$ is the unrestricted partition function. Indeed the identity follows at once by the usual method of logarithmic differentiation of (9) w.r.t. t, setting $t=1$, clearing of fractions, and matching the coefficients of x^n on both sides. We omit the details.

To prove (10) combinatorially, let $\Lambda_i(n)$ be the set of partitions of n with no part equal to i and let $q_i(n) = |\Lambda_i(n)|$. It is easy to see that $q_i(n) = p(n) - p(n-i)$, for example by partitioning $\Lambda(n)$ into those partitions which do have a part i, and those which do not. Thus, (10) becomes

$$\sum_{j\geq 1} j\cdot p(n,j,m) = \sum_{i\geq 1} q_i(n-mi). \tag{11}$$

For the mapping, take

$$\Lambda_1(n-m)\cup\Lambda_2(n-2m)\cup\Lambda_3(n-3m)\cup\ldots \quad\Longrightarrow\quad \Lambda(n)$$

as follows. If λ is a partition of $n-im$, in which no part 'i' occurs, add m copies of i to λ to obtain a partition of n in which part i has multiplicity m.

Conversely, if λ is a partition of n in which part i occurs with multiplicity m, delete the m copies of i to obtain a partition of $n-mi$ in which no part has size i. Thus, each partition $\lambda\in\Lambda(n)$ with exactly j parts of multiplicity m is the image under this mapping of exactly j different partitions, namely those obtained by deleting all copies of one of the parts of multiplicity m. This proves (11). □

A consequence is the following.

Theorem 2 *The average number of parts of multiplicity m in a partition of n is*

$$\langle\delta_m\rangle_n = \sum_{j\geq 1}\frac{p(n-jm)-p(n-j(m+1))}{p(n)}.$$

It is easy, with the aid of (7), to find the asymptotic behavior. The result is that for each fixed m the average number of parts of multiplicity m of a partition of n is

$$\langle\delta_m\rangle_n \sim \frac{\sqrt{6}}{m\pi}n^{\frac{1}{2}} - \frac{\sqrt{6}}{(m+1)\pi}n^{\frac{1}{2}} = \frac{\sqrt{6}}{m(m+1)\pi}n^{\frac{1}{2}} \qquad (n\to\infty). \tag{12}$$

5 The total number of parts

Let $\langle\kappa\rangle_n$ denote the average value of $\kappa(\lambda)$ over all partitions λ of n, where $\kappa(\lambda)$ is the number of parts of λ. Then if $p(n,k)$ is the number of partitions of n with exactly k parts,

$$\langle\kappa\rangle_n = \frac{1}{p(n)} \sum_{k\geq 1} k \cdot p(n,k).$$

The identity

$$\sum_{k\geq 1} k \cdot p(n,k) = \sum_{m\geq 1} \sum_{j\geq 1} p(n-mj) \tag{13}$$

can be established by observing that the following gives a bijection from the set counted by the right-hand-side to the set counted by the left-hand-side: Given $m, j \geq 1$ and $\lambda \in \Lambda(n-mj)$, add m copies of part j to λ to obtain a partition in $\Lambda(n)$. Note that for every $k \geq 1$, each partition of n with exactly k parts will be the image of a partition in $\Lambda(n-mj)$ for exactly k pairs (m,j). A proof of (13) by generating functions can be found in [2], p. 296.

The asymptotic behavior of $\langle\kappa\rangle_n$ was studied in references [7, 8, 9] and is known to be

$$\langle\kappa\rangle_n \sim \frac{\sqrt{6}}{2\pi} n^{\frac{1}{2}} \log n. \tag{14}$$

The reader might note the factor of $\log n/2$ by which this formula for the average number of parts differs from (2), for the average number of part sizes, as a measure of the influence of multiplicities on the averages.

The result (14) appears in [9], although after the calculations in the paper were called into question by the reviewer [10], a rigorous proof was provided by Kessler and Livingston [8]. But in fact, an earlier expansion of $\langle\kappa\rangle_n$ up to an $o(1)$ remainder term can be found in the fascinating paper of Husimi [7]. (According to Husimi, $p(n,k)$ represents the number of complexions of a Bose gas of k particles and of energy n distributed over the energy levels $(\epsilon = 1, 2, 3, \ldots)$ and $\langle\kappa\rangle_n$ can be interpreted as the *"mean number of excited particles"*. Husimi was motivated to investigate the asymptotic behavior of $\langle\kappa\rangle_n$ in order to confirm or refute the conjecture that $\langle\kappa\rangle_n \sim n^{\frac{2}{3}}$ which was supported to within a few percent by experimental evidence for $n \leq 100$.)

Since all we need is the leading term in the formula for $\langle\kappa\rangle_n$, we give a short derivation of that simplified formula below.

In the sum for $\langle\kappa\rangle_n$, the indices m and j are subject to a restriction $mj \leq n-1$. Using the Hardy-Ramanujan formula

$$p(\mu) = \frac{e^{2c\mu^{\frac{1}{2}}}}{4\sqrt{3}\mu}(1 + O(\mu^{-\frac{1}{2}})), \quad (c = \pi 6^{-\frac{1}{2}}),$$

(uniformly for $\mu \geq 1$), after simple algebra we get:

$$\begin{aligned} \frac{p(n-mj)}{p(n)} &= \left[1 + O\left(\frac{mj}{n^{\frac{3}{2}}}\right) + O(n-mj)^{-\frac{1}{2}}\right] \exp\left(-\frac{cmj}{n^{\frac{1}{2}}}\right) \\ &= \begin{cases} (1 + O(n^{-\frac{1}{2}})) \exp(-cmjn^{-\frac{1}{2}}), & \text{if } mj \leq n/2; \\ O(\exp(-cn^{\frac{1}{2}}/2)), & \text{if } mj > n/2. \end{cases} \end{aligned}$$

Thus

$$\begin{aligned} \langle \kappa \rangle_n &= \sum_{j=1}^{n} \sum_{m \leq n/(2j)} \exp(-cmjn^{-\frac{1}{2}}) + O(\exp(-c'n^{\frac{1}{2}})) \qquad (\forall\, c' < c/2) \\ &= \sum_{j=1}^{n} \frac{\exp(-cjn^{-\frac{1}{2}})}{1 - \exp(-cjn^{-\frac{1}{2}})} + O(\exp(-c'n^{\frac{1}{2}})) \\ &= \int_1^n \frac{\exp(-cxn^{-\frac{1}{2}})}{1 - \exp(-cxn^{-\frac{1}{2}})}\, dx + O(n^{\frac{1}{2}}) \\ &= n^{\frac{1}{2}} c^{-1} \log \frac{1}{1 - e^{-cn^{-\frac{1}{2}}}} + O(n^{\frac{1}{2}}) \\ &= n^{\frac{1}{2}} (2c)^{-1} \log n + O(n^{\frac{1}{2}}). \end{aligned}$$

(We have used here the Maclaurin formula with the remainder term.)

6 The probability that a part size has multiplicity m

The fact that the average number of parts of multiplicity m and the average number of distinct part sizes are both proportional to $\sqrt{n}$ make plausible the following conjecture. Consider a two-step sampling procedure, in which we first sample uniformly at random (uar) a partition λ of n and second sample uar one of the different part sizes in λ. *Then the unconditional probability that the chosen part size has multiplicity m approaches a universal constant, β_m, as n tends to infinity*. We prove this with $\beta_m = 1/(m(m+1))$ in Theorem 3 below.

Let X_j be the multiplicity of the part j in a random partition λ of n, (that is, $X_j(\lambda) = \mu_\lambda(j)$); let I_j be the indicator of the event $\{X_j \geq 1\}$, and let $I_{j,m}$ be the indicator of the event $\{X_j = m\}$; $j \geq 1$. Then $D_n = \sum_{j \geq 1} I_j$ is the total number of different part sizes in the random partition. (Of course, $\mathbf{E}\ D_n = \langle \delta \rangle_n \sim \sqrt{6n}/\pi$ see Sections 2 and 3. Goh and Schmutz [6] proved the asymptotic normality of D_n, from which it follows that D_n is asymptotic to $\sqrt{6n}/\pi$ in probability as well.) Likewise, $D_{n,m} = \sum_{j \geq 1} I_{j,m}$ is the total number of part sizes of multiplicity m, and

$$\mathbf{E} D_{n,m} = \langle \delta_m \rangle_n$$

is sharply estimated in Section 4. We see that, given the random variables $\mathbf{X} = (X_1, X_2, \ldots)$, the *conditional* probability $\rho(\mathbf{X})$ that the randomly selected size has multiplicity m is given by

$$\rho(\mathbf{X}) = \frac{D_{n,m}}{D_n}.$$

Now ρ_n, the probability in question, is

$$\rho_n = \mathbf{E}\rho(\mathbf{X}) = \mathbf{E}\left(\frac{D_{n,m}}{D_n}\right), \tag{15}$$

that is, it equals the expected value of $\rho(\mathbf{X})$.

Theorem 3 *The probability that, for a fixed m, a randomly chosen part size of a random partition of n occurs with multiplicity m approaches $1/(m(m+1))$ as $n \to \infty$. In short, $\lim_{n\to\infty} \rho_n = 1/(m(m+1))$.*

Proof. We know from (2) and (12) that

$$\mathbf{E}D_n \sim \frac{n^{\frac{1}{2}}}{c}; \quad \mathbf{E}D_{n,m} \sim \frac{n^{\frac{1}{2}}}{cm(m+1)}; \quad c = \frac{\pi}{\sqrt{6}}, \tag{16}$$

and that D_n is typically close to $\mathbf{E}\ D_n$. So, intuitively, one is justified in replacing D_n in (15) by $n^{\frac{1}{2}}c^{-1}$. To do this rigorously though, we need to know how unlikely is the event

$$A_n := \left\{\left|\frac{D_n}{n^{\frac{1}{2}}c^{-1}} - 1\right| \geq \varepsilon\right\};$$

Such an estimate does not follow from the results in [3], [6], and [15]. Instead, we get a good bound by using the conditioning device, suggested for the integer partitions by Fristedt [4] (see also [11]) and patterned after the analogous treatment of random permutations by Shepp and Lloyd [13]. Namely, introduce the sequence of independent random variables $\mathbf{Y} = (Y_1, Y_2, \ldots)$, where Y_j is geometrically distributed with a parameter q^j,

$$\mathbf{Pr}\ \{Y_j = k\} = (1 - q^j)q^{jk}.$$

Then, for every fixed q, the sequence $\mathbf{X}$ has the same distribution as the sequence $\mathbf{Y}$, *conditioned* on the event

$$B_n := \left\{\sum_{j\geq 1} jY_j = n\right\}$$

(see [4] for a proof). It is natural to pick q for which $\mathbf{Pr}\ (B_n)$ is as large as possible, and Fristedt's almost optimal choice was to set $q = e^{-cn^{-\frac{1}{2}}}$. For this q,

$$\mathbf{Pr}\ (B_n) \sim \text{const} \cdot n^{-\frac{3}{4}}.$$

Let

$$\begin{aligned}
\mathcal{D}_n &:= |\{j \geq 1 : Y_j \geq 1\}| = \sum_{j \geq 1} \mathcal{I}_j, \\
\mathcal{I}_j &:= \begin{cases} 1, & \text{if } Y_j \geq 1, \\ 0, & \text{if } Y_j = 0, \end{cases} \\
\mathcal{A}_n &= \left\{ \left| \frac{\mathcal{D}_n}{n^{\frac{1}{2}} c^{-1}} - 1 \right| \geq \varepsilon \right\}.
\end{aligned}$$

Then

$$\begin{aligned}
\mathbf{Pr}(A_n) &= \mathbf{Pr}(\mathcal{A}_n | B_n) \\
&= \frac{\mathbf{Pr}(\mathcal{A}_n \cap B_n)}{\mathbf{Pr}(B_n)} \\
&\leq \frac{\mathbf{Pr}(\mathcal{A}_n)}{\mathbf{Pr}(B_n)} \\
&= O\left(n^{\frac{3}{4}}\ \mathbf{Pr}(\mathcal{A}_n)\right). \qquad (17)
\end{aligned}$$

So we need to bound $\mathbf{Pr}(\mathcal{A}_{\mathrm{n}})$. This is easy since $Y_1, Y_2, \ldots$ are independent. Here is a standard argument. Let $u \in \mathbf{R}$ be given. Then

$$\begin{aligned}
\mathbf{E}(e^{u\mathcal{D}_n}) &= \prod_{j \geq 1} \mathbf{E}(e^{u\mathcal{I}_j}) \\
&= \prod_{j \geq 1} \left(1 - q^j + e^u q^j\right) \\
&\leq \exp\left((e^u - 1) \sum_{j \geq 1} q^j\right) \\
&= \exp\left((e^u - 1) \frac{q}{1-q}\right).
\end{aligned}$$

So, for every $u > 0$, by Chebyshev's inequality,

$$\mathbf{Pr}\left\{\mathcal{D}_n \geq \frac{n^{\frac{1}{2}}}{c}(1+\varepsilon)\right\} \leq \frac{\exp\left((e^u - 1)\frac{q}{1-q}\right)}{\exp(u\frac{n^{\frac{1}{2}}}{c}(1+\varepsilon))},$$

and an almost optimal u (that minimizes the bound) is $\log(1+\varepsilon)$, which gives an upper bound

$$\exp\left(-n^{\frac{1}{2}} a_1\right), \quad a_1 = a_1(\varepsilon) \sim c^{-1}\left[(1+\varepsilon)\log(1+\varepsilon) - \varepsilon\right] > 0.$$

Analogously, using $u = \log(1-\varepsilon) < 0$, we obtain

$$\begin{aligned}
\mathbf{Pr}\left\{\mathcal{D}_n \leq \frac{n^{\frac{1}{2}}}{c}(1-\varepsilon)\right\} &\leq \exp\left(-n^{\frac{1}{2}} a_2\right), \\
a_2 &= a_2(\varepsilon) \sim c^{-1}\left[\varepsilon + (1-\varepsilon)\log(1-\varepsilon)\right].
\end{aligned}$$

So

$$\mathbf{Pr}(\mathcal{A}_n) \leq e^{-an^{\frac{1}{2}}}, \quad a > 0,$$

and, combining this bound with (17), we have

$$\mathbf{Pr}(A_n) \leq e^{-bn^{\frac{1}{2}}}, \quad 0 < b < a. \tag{18}$$

The rest is short and easy. We write

$$\begin{aligned} \rho_n &= \mathbf{E}\left(\rho(\mathbf{X})\right) \\ &= \mathbf{E}\left(\rho(\mathbf{X}) I_{A_n^c}\right) + \mathbf{E}\left(\rho(\mathbf{X})) I_{A_n}\right) \\ &= \mathbf{E}_1 + \mathbf{E}_2. \end{aligned}$$

By the definition of $\rho(\mathbf{X})$ and the event A_n^c,

$$\begin{aligned} \mathbf{E}_1 &= (1 + O(\varepsilon)) \frac{c}{n^{\frac{1}{2}}} \mathbf{E}(D_{n,m} I_{A_n^c}) \\ &= (1 + O(\varepsilon)) \frac{c}{n^{\frac{1}{2}}} \left[\mathbf{E}(D_{n,m}) - \mathbf{E}\left(D_{n,m} I_{A_n}\right)\right]. \end{aligned} \tag{19}$$

Furthermore, by (18),

$$\begin{aligned} \mathbf{E}(D_{n,m} I_{A_n}) &= O(n \mathbf{Pr}(A_n)) \\ &= O(n e^{-bn^{\frac{1}{2}}}) \\ &= o(1). \end{aligned}$$

Combining this estimate with (19) and (16), we have

$$\mathbf{E}_1 = \frac{1}{m(m+1)} + O(\varepsilon) + o(1), \quad \text{as } n \to \infty.$$

It remains to notice that

$$\mathbf{E}_2 \leq \mathbf{Pr}(A_n) = o(1), \quad \text{as } n \to \infty,$$

and we conclude, letting $n \to \infty$ and then $\varepsilon \to 0$, that

$$\lim_{n \to \infty} \rho_n = \frac{1}{m(m+1)}, \quad m \geq 1.$$

□

Notes. 1. Analogously to D_n, we could have proved that the distribution of $D_{n,m}$ is concentrated around $\mathbf{E} D_{n,m}$. Consequently, the *empirical* probability $\rho(\mathbf{X}) = D_{n,m}/D_n$ that the random part size has multiplicity m, converges, in probability, to $1/(m(m+1))$. This also implies that $\rho_n \to 1/(m(m+1))$.

2. How will the multiplicity distribution change if we sample a part from the set of *all* parts with the (conditional) probability *proportional* to a part's size? An answer is not that obvious, since in a typical partition the small parts i have multiplicities of order $n^{1/2}$, while the middle range parts are of order $n^{1/2}$, but of low multiplicity. In this case the conditional probability that the selected part's size has multiplicity m is given by

$$\rho(\mathbf{X}) = \frac{m}{n} \sum_{j:X_j=m} j,$$

and the unconditional probability is therefore given by

$$\rho_n = \mathbf{E}\rho(\mathbf{X}) = \frac{m}{n} \sum_{j\geq 1} j\mathbf{Pr}(X_j = m).$$

We leave it to the reader to show that

$$\lim_{n\to\infty} \rho_n = \frac{6}{\pi^2} \frac{2m+1}{m(m+1)^2}, \quad m \geq 1.$$

The sum of the limits is 1 again! And so, as before, the multiplicity is bounded in probability.

3. What if a part is chosen uar among all parts without any size bias? Intuitively, one should expect the multiplicity to be higher in probability. For this sampling, the conditional probability of multiplicity m is

$$\rho(\mathbf{X}) = \frac{mD_{n,m}}{P_n},$$

where $P_n = P(\lambda)$ is the total number of parts in the random partition λ. Erdös and Lehner [3] proved that, in probability, P_n is asymptotic to $\mathbf{E}P_n$ $(= \langle \kappa \rangle_n)$. Extending our argument in the proof of Theorem 3, we can also show, for $m = o(n^{1/2})$, that $mD_{n,m}$ is asymptotic in probability to

$$\begin{aligned} m\mathbf{E}\left(\sum_{j\geq 1} I_{j,m}\right) &= m\sum_{j\geq 1}(1-q^j)q^{jm} \\ &\sim \frac{n^{1/2}}{c(m+1)}. \end{aligned}$$

Therefore, in probability,

$$\rho(\mathbf{X}) \sim \frac{2}{(m+1)\log n},$$

whence

$$\mathbf{Pr}(u_n = m) \sim \frac{2}{(m+1)\log n}, \quad m = o(n^{1/2}).$$

Here u_n denotes the random multiplicity of the chosen part. Consequently

$$\mathbf{Pr}\left(\frac{\log u_n}{\log n} \leq x\right) \to 2x,$$

for every $x < 1/2$. Thus $\log u_n$ is typically of order $\log n$.

7 The expected multiplicity of a part

We now consider, for a random partition λ, the average multiplicity of a part size selected at random from the set of all part sizes occurring in λ. Equations (2) and (14) suggest the following.

Theorem 4 *If M_n is the expected multiplicity of a randomly chosen part size in a random partition of n then*

$$M_n \sim \frac{\log n}{2}.$$

Proof. With X_j, I_j, and D_n as defined in the previous section, clearly

$$M_n = \mathbf{E}\left(\frac{\sum_{j\geq 1} X_j}{D_n}\right).$$

Here $\sum_{j\geq 1} X_j \leq n$; so it can be shown, analogously to the previous argument, that M_n is asymptotic to

$$\tilde{M}_n = \frac{\mathbf{E}\left(\sum_{j\geq 1} X_j\right)}{\mathbf{E} D_n}.$$

Note that from (14),

$$\mathbf{E}\sum_{j\geq 1} X_j = \langle\kappa\rangle_n \sim \frac{\sqrt{6}}{2\pi} n^{\frac{1}{2}} \log n$$

and from (2),

$$\mathbf{E}\ D_n \sim \frac{\sqrt{6}}{\pi} n^{\frac{1}{2}},$$

so

$$M_n \sim \tilde{M}_n \sim \frac{\log n}{2}.$$

□

In contrast, for size-unbiased sampling from the set of all parts discussed in Note 3 of the previous section, it is the logarithm of the multiplicity which is of order $\log n$.

Acknowledgement. We gratefully acknowledge our debt to a referee who carefully read the paper and provided us with very useful comments and thoughtful suggestions.

References

[1] Cohen, Daniel I. A., *Basic Techniques of Combinatorial Theory*, John Wiley & Sons, New York, 1978.

[2] Comtet, L., *Advanced Combinatorics*, Reidel, 1974.

[3] Erdős, P., and Lehner, J., The distribution of the number of summands in the partitions of a positive integer, *Duke Math. J.* **8** (1941), 335-345.

[4] Fristedt, B., The structure of random partitions of large integers, *Trans. Amer. Math. Soc.* **337** (1993), 703-735.

[5] Hardy, G. H., and Ramanujan, S., Asymptotic formulæ in combinatory analysis, *Proc. London Math. Soc.* **17** (1918), 75-115.

[6] Goh, W. M. Y. and Schmutz, E., The number of distinct part sizes in a random integer partition, *J. Combin. Theory Ser. A* **69** (1995), 149-158.

[7] Husimi, K., Partitio numerorum as occurring in a problem of nuclear physics, *Proceedings of the Physico-Mathematical Society of Japan* **20** (1938), 912-925.

[8] Kessler, I. and Livingston, M., The expected number of parts in a partition of n, *Monatshefte für Mathematik* **81** (1976), 203-212.

[9] Luthra, S. M., On the average number of summands in a partition of n, *Proc. Nat. Inst. Sci. India*, Part A **23** (1957), 483–498.

[10] Newman, M., Review #1162, *Math. Rev.* **19** (1958).

[11] Pittel, B., On a likely shape of the random Ferrers diagram, *Adv. Appl. Math.* **18** (1997), 432-488.

[12] Rademacher, H., On the partition function $p(n)$, *Proc. London Math. Soc.* **43** (1937), 241-254.

[13] Shepp, L. A. and Lloyd, S. P., Ordered cycle lengths in a random permutation, *Trans. Amer. Math. Soc.* **121** (1966), 340-357.

[14] Wilf, Herbert S., *generatingfunctionology*, Academic Press, 1994.

[15] Wilf, Herbert S., Three problems in combinatorial asymptotics, *J. Combin. Theory Ser. A* **35** (1983), 199-207.