*Combinatorics, Probability and Computing* (2000) 9, 519–527. Printed in the United Kingdom
© 2000 Cambridge University Press

# Expected Number of Distinct Part Sizes in a Random Integer Composition

P. HITCZENKO[1] and G. STENGLE[2]

[1] MCS Department, Drexel University,
Philadelphia, PA 19104, USA
(e-mail: phitczen@mcs.drexel.edu)

[2] Department of Mathematics, Lehigh University,
Bethlehem, PA 18015, USA
(e-mail: gas0@lehigh.edu)

*Received 8 July 1999; revised 7 December 1999*

An asymptotics, as $n \to \infty$, for the expected number of distinct part sizes in a random composition of an integer $n$ is obtained.

## 1. Introduction

In this note we obtain precise asymptotics, as $n \to \infty$, for the expected number of distinct part sizes in a random composition of an integer $n$. Let us recall that a multiset $\lambda = \{\lambda_1, \ldots, \lambda_k\}$ is a partition of an integer $n$ if the $\lambda_j$ are positive integers, called parts, such that $\sum \lambda_j = n$. The values of the $\lambda_j$ s are called part sizes. Compositions are partitions in which the order of parts is significant. Thus, for example, the integer 3 admits three partitions, $\{1,1,1\}$, $\{2,1\}$ and $\{3\}$, and four compositions, namely $(1,1,1)$, $(1,2)$, $(2,1)$ and $(3)$. According to our terminology $(1,2)$ is a composition of 3 in two parts with sizes 1 and 2. By analogy with random partitions, by a random composition of an integer $n$ we mean a composition of $n$ that is chosen uniformly at random out of the set of all $2^{n-1}$ compositions of an integer $n$. More formally, one considers the probability space consisting of the set $C(n)$ of all compositions of $n$ equipped with the uniform probability measure. In this setting, the number of distinct part sizes (or other characteristics) becomes a random variable whose probabilistic behaviour is to be studied.

Investigation of random partitions from this probabilistic perspective originated with a paper by Erdős and Lehner [5], who studied the limiting distribution of the total number of parts in a random partition. Subsequently, Wilf [11] found an asymptotic formula for the expected number of distinct part sizes. Goh and Schmutz [7] obtained more precise information on the distribution of the number of distinct part sizes: namely,

https://doi.org/10.1017/S0963548300004399 Published online by Cambridge University Press

they established the central limit theorem. Recently Corteel, Pittel, Savage and Wilf [3] obtained a refined version of Wilf's result concerning the expectation of the number of distinct part sizes in a random partition. Their result allows one to obtain as many terms for the asymptotic expansion of this expectation as one wishes. For example, on '$o(1)$ level' this expectation is

$$\frac{\sqrt{6n}}{\pi} + \frac{3}{\pi^2} - \frac{1}{2} + o(1).$$

The aim of this note is to obtain an asymptotics for the same quantity in the case of random compositions. In order to state our result we need some notation. For an integer $n$ consider the set $C(n)$ of all compositions $\kappa$ of $n$ equipped with the uniform probability measure $\mathbf{P}_n = \mathbf{P}$ (that is, $\mathbf{P}(\kappa) = 2^{-n+1}$ for every $\kappa \in C(n)$.) For a composition $\kappa = (\gamma_1, \ldots, \gamma_k)$, the number of distinct part sizes, $D_n(\kappa)$ is defined by the formula

$$D_n(\kappa) = 1 + \sum_{i=2}^{k} I_{\{\gamma_i \neq \gamma_j,\ j=1,\ldots,i-1\}},$$

where $I_A$ is the indicator function of the set $A$. We denote the integration with respect to $\mathbf{P}$ on $C(n)$ by $\mathbf{E}$. We have the following result.

**Theorem.** *As $n \to \infty$,*

$$\mathbf{E}D_n = \log_2 n + \frac{\gamma}{\ln 2} - \frac{3}{2} + g(\log_2 n) + o(1),$$

*where $\gamma$ is Euler's constant and $g$ is a mean-zero function of period 1 satisfying $|g| \leqslant 0.0000016$.*

Thus, the expected number of distinct part sizes in a composition of an integer $n$ asymptotically behaves like $\log_2 n$ plus a constant plus a small but periodic oscillation. This oscillatory behaviour, which just a few years ago was considered surprising (to say the least), is by now a well-documented and acknowledged feature of sequences of geometric random variables: see, *e.g.*, [2], [4], [8], [10].

We wish to observe that the asymptotics for the expected number of distinct part sizes is the same as the expected length of the longest run of heads in $n$ tosses of a fair coin: see, *e.g.*, [2] or [8]. Since the size of the largest part is one plus the longest run of heads, it follows that on average one expects to see parts of all but one size between 1 and the largest size (or runs of heads of all but one length between 1 and the longest run).

## 2. Outline of a proof

Quite often results like this are obtained through careful analysis of the the generating function. We will use a different approach. We will view random composition as (essentially) a randomly stopped sequence of i.i.d. geometric random variables and we will express the number of distinct part sizes as a function of this sequence. This will allow

https://doi.org/10.1017/S0963548300004399 Published online by Cambridge University Press

for direct and straightforward estimates. The same approach was used successfully in [9] to handle a problem in which the generating function approach was apparently futile. We believe that this technique will prove useful in many other problems concerning random compositions. Our proof splits naturally into two steps. In the first we will use the aforementioned representation and probabilistic estimates to extract the main contribution to $\mathbf{E}D_n$. Namely, we have the following.

**Proposition 2.1.** *As* $n \to \infty$,

$$\mathbf{E}D_n = \sum_{m=1}^{\infty} \left\{ 1 - \left( 1 - \frac{1}{2^m} \right)^{\frac{n}{2}(1+o(1))} \right\} + o(1).$$

The second, purely analytical step is to analyse the asymptotic behaviour of the infinite sum above. This goal could be accomplished by applying the so-called Rice method (see, *e.g.*, [6] for a very good description and examples). Since this method requires some tools from complex analysis we decided to take a different route. As a result our analysis is completely elementary (thus making this paper fully accessible to advanced undergraduates, for example.) To facilitate our analysis we define

$$f(x) = \sum_{m=1}^{\infty} \left\{ 1 - \left( 1 - \frac{1}{2^m} \right)^{2^x} \right\}.$$

With this definition we will show that $f(x) - x$ tends to a limit as $x$ tends to infinity along sequences of the form $\{x_0 + k\}_{k \in Z}$, but does not possess an unrestricted limit as $x \to \infty$. More specifically, we have the following.

**Proposition 2.2.** *For large positive* $k$,

$$f(x+k) = x + k + \gamma / \ln 2 - 1/2 + g(x) + o\left(2^{-x-k}\right),$$

*where* $\gamma$ *is Euler's constant and*

$$g(x) = -x - \gamma / \ln 2 + 1/2 - \sum_{m=-\infty}^{0} \exp(-2^{-m+x}) + \sum_{m=1}^{\infty} (1 - \exp(-2^{-m+x}))$$

*is a nonconstant, mean-zero function of period* 1 *satisfying* $| g(x) | \leqslant .0000016$.

Clearly, the theorem follows by combining Propositions 2.1 and 2.2.

## 3. Proof of Proposition 2.1

Central to our approach is the following proposition.

https://doi.org/10.1017/S0963548300004399 Published online by Cambridge University Press

**Proposition 3.1.** *Let $\Gamma_1, \Gamma_2 \ldots$ be i.i.d. geometric random variables with parameter $1/2$ (that is, $\mathbf{P}(\Gamma_1 = j) = 2^{-j}$, $j = 1, 2 \ldots$) and define*

$$\tau = \inf\{k \geqslant 1 : \ \Gamma_1 + \Gamma_2 + \ldots + \Gamma_k \geqslant n\}.$$

*Then, the distribution of a randomly chosen composition in $C(n)$ is given by*

$$\left(\Gamma_1, \Gamma_2, \ldots, \Gamma_{\tau-1}, n - \sum_{j=1}^{\tau-1} \Gamma_j\right).$$

This proposition is nothing more than a reiteration of a known (see, *e.g.*, [1]) connection between compositions of integers and $\{0,1\}$-valued sequences. Namely, a composition $\kappa = (\gamma_1, \ldots, \gamma_k)$ of an integer $n$ into parts $\gamma_1, \ldots, \gamma_k$ is associated with a string of 0 s and 1 s of length $n$ as follows: there is a 1 on the $n$th place and the numbers $\gamma_1, \ldots, \gamma_k$ are 'waiting times' for the first, second, ..., and $k$th appearance of 1. (For example, the composition $(1, 2, 3, 1, 1)$ of 8 corresponds to the string 10100111, while $(4, 2, 2)$ corresponds to 00010101.) Choosing a composition at random amounts to having the 0 s and 1 s on the first $n-1$ places occur according to a binomial $\mathrm{Bin}(n-1, 1/2)$ law. We refer to [9] for more details. Let $\tilde{\Gamma}_i(\kappa)$ denote parts of a randomly chosen composition $\kappa$, that is,

$$\tilde{\Gamma}_i(\kappa) = \Gamma_i(\kappa), \quad \text{for} \quad i < \tau(\kappa) \quad \text{and} \quad \tilde{\Gamma}_{\tau(\kappa)}(\kappa) = n - \sum_{i=1}^{\tau(\kappa)-1} \Gamma_i(\kappa).$$

Note that $\tilde{\Gamma}_\tau \leqslant \Gamma_\tau$. The expected value of $D_n$ is computed as follows:

$$\begin{aligned} \mathbf{E}D_n &= 1 + \mathbf{E}\sum_{i=2}^{\tau} I_{\tilde{\Gamma}_i \neq \tilde{\Gamma}_j;\ j=1,\ldots,i-1} \\ &= 1 + \mathbf{E}\sum_{i=2}^{\tau-1} I_{\Gamma_i \neq \Gamma_j;\ j=1,\ldots,i-1} + \mathbf{P}(\tilde{\Gamma}_\tau \neq \Gamma_1, \ldots, \Gamma_{\tau-1}). \end{aligned}$$

We will first show that the last probability is negligible. This is because $\tau$ being a $1 + \mathrm{Bin}(n-1, 1/2)$ random variable satisfies the bound

$$\mathbf{P}(|\tau - \mathbf{E}\tau| \geqslant t) \leqslant 2\exp(-2t^2/(n-1)),$$

so that

$$\mathbf{P}(|\tau - \mathbf{E}\tau| \geqslant \sqrt{(n-1)\log n}) \leqslant 2\exp\left\{-2\frac{(n-1)\log n}{n-1}\right\} = O(1/n^2).$$

Therefore, letting $t_n = \sqrt{(n-1)\log n}$ and then

$$n_0 \sim \mathbf{E}\tau - t_n = (n+1)/2 - \sqrt{(n-1)\log n},$$

and

$$n_1 \sim \mathbf{E}\tau + t_n = (n+1)/2 + \sqrt{(n-1)\log n},$$

https://doi.org/10.1017/S0963548300004399 Published online by Cambridge University Press

we get

$$\begin{aligned}
\mathbf{P}(\tilde{\Gamma}_\tau \neq \Gamma_1,\dots,\Gamma_{\tau-1}) &\leqslant \sum_{j=1}^{n} \mathbf{P}(\tilde{\Gamma}_\tau = j,\ \Gamma_1,\dots,\Gamma_{\tau-1} \neq j)\\
&\leqslant \sum_{j=1}^{n} \mathbf{P}(\tilde{\Gamma}_\tau \geqslant j,\ \Gamma_1,\dots,\Gamma_{\tau-1} \neq j)\\
&\leqslant \sum_{j=1}^{n} \mathbf{P}(\Gamma_\tau \geqslant j,\ \Gamma_1,\dots,\Gamma_{\tau-1} \neq j)\\
&\leqslant \sum_{j=1}^{n} \mathbf{P}(\Gamma_\tau \geqslant j,\ \Gamma_1,\dots,\Gamma_{\tau-1} \neq j,\ |\tau-\mathbf{E}\tau| \leqslant t_n) + n\mathbf{P}(|\tau-\mathbf{E}\tau| \geqslant t_n)\\
&\leqslant \sum_{j=1}^{n}\sum_{k=n_0}^{n_1} \mathbf{P}(\Gamma_k \geqslant j, \Gamma_1,\dots,\Gamma_{k-1} \neq j,\ \tau = k) + O(1/n)\\
&\leqslant \sum_{k=n_0}^{n_1}\sum_{j=1}^{\infty} \mathbf{P}(\Gamma_k \geqslant j, \Gamma_1,\dots,\Gamma_{k-1} \neq j) + O(1/n)\\
&= \sum_{k=n_0}^{n_1}\sum_{j=1}^{\infty} \frac{1}{2^{j-1}}\left(1-\frac{1}{2^j}\right)^{k-1} + O(1/n)\\
&\leqslant \sum_{k=n_0}^{n_1} C\int_0^{\infty} \frac{1}{2^x}\left(1-\frac{1}{2^{x+1}}\right)^{k-1} dx + O(1/n)\\
&\leqslant \sum_{k=n_0}^{n_1} \frac{C}{k} + O(1/n) \leqslant C\frac{\sqrt{n\log n}}{n} \to 0,
\end{aligned}$$

as $n \to \infty$. As for the other term, we have

$$\begin{aligned}
\mathbf{E}\sum_{i=2}^{\tau-1} I_{\Gamma_i \neq \Gamma_j,\ j<i} &\leqslant \mathbf{E}\left(\sum_{i=2}^{\tau-1} I_{\Gamma_i \neq \Gamma_j,\ j<i}\right) I_{|\tau-\mathbf{E}\tau| \leqslant t_n}\\
&\quad + \mathbf{E}\left(\sum_{i=2}^{\tau-1} I_{\Gamma_i \neq \Gamma_j,\ j<i}\right) I_{|\tau-\mathbf{E}\tau| > t_n}\\
&\leqslant \mathbf{E}\left(\sum_{i=2}^{\tau-1} I_{\Gamma_i \neq \Gamma_j,\ j<i}\right) I_{|\tau-\mathbf{E}\tau| \leqslant t_n} + (n-2)\mathbf{P}(|\tau-\mathbf{E}\tau| \geqslant t_n).
\end{aligned}$$

The second term is bounded above by $C/n$ and, of course, tends to 0 as $n \to \infty$. For the first one we have

$$\begin{aligned}
\mathbf{E}\left(\sum_{i=2}^{\tau-1} I_{\Gamma_i \neq \Gamma_j,\ j<i}\right) I_{|\tau-\mathbf{E}\tau| \leqslant t_n} &\leqslant \mathbf{E}\left(\sum_{i=2}^{n_1} I_{\Gamma_i \neq \Gamma_j,\ j<i}\right) I_{|\tau-\mathbf{E}\tau| \leqslant t_n}\\
&\leqslant \mathbf{E}\sum_{i=2}^{n_1} I_{\Gamma_i \neq \Gamma_j,\ j<i} = \sum_{i=2}^{n_1} \mathbf{P}(\Gamma_i \neq \Gamma_j,\ j<i).
\end{aligned}$$

https://doi.org/10.1017/S0963548300004399 Published online by Cambridge University Press

Similarly,

$$\begin{aligned}\mathbf{E}\left(\sum_{i=2}^{\tau-1} I_{\Gamma_i\neq\Gamma_j,\ j<i}\right) I_{|\tau-\mathbf{E}\tau|\leqslant t_n} &\geqslant \mathbf{E}\left(\sum_{i=2}^{n_0} I_{\Gamma_i\neq\Gamma_j,\ j<i}\right) I_{|\tau-\mathbf{E}\tau|\leqslant t_n}\\ &= \mathbf{E}\sum_{i=2}^{n_0} I_{\Gamma_i\neq\Gamma_j,\ j<i} - \mathbf{E}\sum_{i=2}^{n_0}\mathbf{E}\left(\Gamma_i\neq\Gamma_j,\ j<i\right) I_{|\tau-\mathbf{E}\tau|>t_n}\\ &\geqslant \sum_{i=2}^{n_0}\mathbf{P}(\Gamma_i\neq\Gamma_j,\ j<i) - (n_0-1)\mathbf{P}(|\tau-\mathbf{E}\tau|\geqslant t_n)\\ &= \sum_{i=2}^{n_0}\mathbf{P}(\Gamma_i\neq\Gamma_j,\ j<i) - O(n^{-1}).\end{aligned}$$

We will now fix $k$ and approximate $\sum_{i=2}^{k}\mathbf{P}(\Gamma_i\neq\Gamma_j,\ j<i)$ as follows:

$$\mathbf{P}(\Gamma_i\neq\Gamma_j,\ j<i) = \sum_{m=1}^{\infty}\mathbf{P}(\Gamma_i=m,\ \Gamma_j\neq m;\ j<i) = \sum_{m=1}^{\infty}\frac{1}{2^m}\left(1-\frac{1}{2^m}\right)^{i-1}.$$

Hence, by summing up over $i$ we get

$$\begin{aligned}\sum_{i=2}^{k}\sum_{m=1}^{\infty}\frac{1}{2^m}\left(1-\frac{1}{2^m}\right)^{i-1} &= \sum_{m=1}^{\infty}\frac{1}{2^m}\sum_{i=1}^{k-1}\left(1-\frac{1}{2^m}\right)^{i}\\ &= \sum_{m=1}^{\infty}\frac{1}{2^m}\left(1-\frac{1}{2^m}\right)\frac{1-(1-2^{-m})^{k-1}}{1-(1-2^{-m})}\\ &= \sum_{m=1}^{\infty}\left(1-\frac{1}{2^m}\right)\left(1-\left(1-\frac{1}{2^m}\right)^{k-1}\right)\\ &= \sum_{m=1}^{\infty}\left\{\left(1-\frac{1}{2^m}\right)-\left(1-\frac{1}{2^m}\right)^{k}\right\}\\ &= \sum_{m=1}^{\infty}\left\{1-\left(1-\frac{1}{2^m}\right)^{k}\right\}-\sum_{m=1}^{\infty}\frac{1}{2^m}\\ &= \sum_{m=1}^{\infty}\left\{1-\left(1-\frac{1}{2^m}\right)^{k}\right\}-1.\end{aligned}$$

It follows that after ignoring terms of order $o(1)$ we have

$$\sum_{m=1}^{\infty}\left\{1-\left(1-\frac{1}{2^m}\right)^{n_0}\right\}\leqslant \mathbf{E}D_n\leqslant\sum_{m=1}^{\infty}\left\{1-\left(1-\frac{1}{2^m}\right)^{n_1}\right\}.$$

Since both $n_0$ and $n_1$ are of the form $\frac{n}{2}(1+o(1))$, Proposition 2.1 follows. □

https://doi.org/10.1017/S0963548300004399 Published online by Cambridge University Press

## 4. Proof of Proposition 2.2

Recall that

$$f(y) = \sum_{m=1}^{\infty} \left\{ 1 - \left( 1 - \frac{1}{2^m} \right)^{2^y} \right\}.$$

We first give a simple argument that gives the limiting behaviour of $f$ without, however, yielding an estimate for the rate of convergence. Let $y = x + k$, where $0 \leqslant x < 1$. We re-index the sum by $m + k$ to obtain

$$f(k+x) - k - x = -x - \sum_{m=-k+1}^{0} (1 - 2^{-m-k})^{2^{k+x}} + \sum_{m=1}^{\infty} \left(1 - (1 - 2^{-m-k})^{2^{k+x}}\right).$$

Permuting summation and limits as $k \to \infty$ yields

$$f(x+k) - k - x = -x - \sum_{m=-\infty}^{0} \exp(-2^{-m+x}) + \sum_{m=1}^{\infty} (1 - \exp(-2^{-m+x})) + o(1).$$

But this step is justified by dominated convergence using the following majorizing convergent series of positive terms independent of $k$:

$$\sum_{m=-k+1}^{0} (1 - 2^{-m-k})^{2^{k+x}} \ll \sum_{m=-\infty}^{0} \exp(-2^{-m+x})$$

and

$$\sum_{m=1}^{\infty} \left(1 - (1 - 2^{-m-k})^{2^{k+x}}\right) \ll \sum_{m=1}^{\infty} (1 - \exp(-2^{-m+x+1})).$$

These follow from the estimates $\exp(-2ab) \leqslant (1 - b/\lambda)^{a\lambda} \leqslant \exp(-ab)$ if $\lambda a > 0$ and $b/\lambda \leqslant 1/2$ with $\lambda = 2^k, a = 2^x$ and $b = 2^{-m}$.

The series thus established as the limit of $f(k+x) - (k+x)$ defines a function of period 1. Denoting its mean by $c$ and its mean-zero part by $g(x)$, we have

$$f(x+k) = x + k + c + g(x) + o(1) = x + k + c + g(x+k) + o(1).$$

To obtain the finer estimate stated in the proposition we use the higher-order estimate $\exp(-ab - ab^2/\lambda) \leqslant (1 - b/\lambda)^{a\lambda}$ if $\lambda a > 0$ and $b/\lambda \leqslant 1/2$. We must bound

$$\begin{aligned} f(x+k) - (x + c + g(x)) \quad &= \quad \sum_{m=-\infty}^{-k} \exp(-2^{-m+x}) \\ &\quad + \sum_{m=-k+1}^{\infty} \left\{ \exp(-2^{-m+x}) - (1 - 2^{-m-k})^{2^{k+x}} \right\}. \end{aligned}$$

The first sum can be rewritten as

$$\sum_{m=0}^{\infty} \exp(-2^{m+k+x}) = \exp(-2^{k+x}) \sum_{m=0}^{\infty} \exp(-(2^m - 1)2^{k+x}),$$

https://doi.org/10.1017/S0963548300004399 Published online by Cambridge University Press

which is bounded by

$$\exp(-2^{k+x})\sum_{m=0}^{\infty}\exp(-(2^m-1))$$

and thus makes an exponentially small contribution to an error term of $O(2^{-k-x})$.

The second sum consists of positive terms and is bounded above by

$$\sum_{m=-\infty}^{\infty}\{\exp(-2^{-m+x})-\exp(-2^{-m+x}-2^{-2m+x-k})\}.$$

Then the inequality $\exp(-a)-\exp(-a-b)\leqslant b\exp(-a)$ for positive $a$ and $b$ gives the bound

$$2^{-x-k}\sum_{m=-\infty}^{\infty}2^{-2m+2x}\exp(-2^{-j+x}).$$

This bound has the form $2^{-x-k}h(x)$ where $h$ is a periodic function of $x$ and is therefore bounded by a constant. This establishes the asserted rate of convergence.

It remains to calculate the mean. The mean of $-x$ is $-1/2$ and the mean of the residual series is

$$c_0=-\sum_{m=-\infty}^{0}\int_0^1\exp(-2^{-m+x})dx+\sum_{m=1}^{\infty}\int_0^1(1-\exp(-2^{-m+x}))dx.$$

On the $m$th summand the change of variable $u=2^{-m+x}$ gives

$$c_0\ln 2=-\sum_{m=-\infty}^{0}\int_{2^{-m}}^{2^{-m+1}}\exp(-u)\frac{du}{u}+\sum_{m=1}^{\infty}\int_{2^{-m}}^{2^{-m+1}}(1-\exp(-u))\frac{du}{u}$$

or

$$c_0\ln 2=-\int_1^{\infty}\exp(-u)\frac{du}{u}+\int_0^1(1-\exp(-u))\frac{du}{u}.$$

Integrating each integral by parts yields a single integral

$$-\int_0^{\infty}\exp(-u)\ln u\,du,$$

which is a well-known integral representing Euler's constant.

We remark that similar reasoning shows that the periodic function

$$h(x)=\sum_{m=-\infty}^{\infty}2^{-2m+2x}\exp(-2^{-m+x})$$

appearing in the rate estimate overestimates the error asymptotically by a factor of two and, in fact,

$$f(x+k)=x+k+\gamma/\ln 2-1/2+g(x)+h(x)2^{-x-k-1}+O(2^{-2x-2k}).$$

The bound on $g$ and its nonconstant character are easily checked numerically, although calculations dealing with $f$ rather than analytically derived asymptotic forms are rather sensitive. Alternately its complex Fourier coefficients are easily obtained by a simple

https://doi.org/10.1017/S0963548300004399 Published online by Cambridge University Press

variant of the calculation of $c_0$ and have the form $c_k = \Gamma(2\pi ki/\ln 2)/\ln 2$. These are non-zero, small and decrease geometrically in magnitude. For example, $2 \mid c_1 \mid = .00000157316$ accounts for the maximum contribution of the first harmonic, while for the second harmonic $2 \mid c_2 \mid < 10^{-12}$. □

**Acknowledgement**

Part of the research of the first author was carried out while he was visiting the Department of Mathematics at Lehigh University in the spring semester of 1999. He would like to thank the department for its hospitality.

**Note added in proof**

As we recently found out, the same result has been obtained independently and by an entirely different method in Hwang, H.-K. and Yeh, Y. (1997) Measures of distinctness for random partitions and compositions of an integer, *Adv. Appl. Math.* **19** 378–414. Related results may be found in Knopfmacher, A. and Mays, M. E. (1999) Compositions with $m$ distinct parts, *Ars Combin.* **53** 111–128.

**References**

[1] Andrews, G. E. (1976) *The Theory of Partitions*, Addison-Wesley, Reading, MA.
[2] Boyd, D. W. (1972) Losing runs in Bernoulli trials. Unpublished manuscript.
[3] Corteel, S., Pittel, B., Savage, C. D. and Wilf, H. S. (1999) On the multiplicity of parts in a random partition. *Random Structures and Algorithms* **14** 185–197.
[4] Eisenberg, B., Stengle, G. and Strang, G. (1993) The asymptotic probability of a tie for first place. *Ann. Appl. Probab.* **13** 731–745.
[5] Erdős, P. and Lehner, J. (1941) The distribution of the number of summands in the partitions of positive integer. *Duke Math. J.* **8** 335–345.
[6] Flajolet, P. and Sedgewick, R. (1995) Mellin transform and asymptotics: finite differences and Rice's integrals. *Theoret. Comput. Sci.* **144** (special volume on mathematical analysis of algorithms) 101–124.
[7] Goh, W. M. Y. and Schmutz, E. (1995) The number of different part sizes in a random integer partition. *J. Combin. Theory Ser. A* **69** 149–158.
[8] Guibas, L. J. and Odlyzko, A. M. (1980) Long repetitive patterns in random sequences. *Z. Wahrscheinlichkeitstheorie und Verv. Gebiete* **53** 241–262.
[9] Hitczenko, P. and Savage, C. D. (1999) On the multiplicity of parts in a random composition of a large integer. Preprint.
[10] Kirschenhofer, P. and Prodinger, H. (1996) The number of winners in a discrete geometrically distributed sample. *Ann. Appl. Probab.* **6** 687–694.
[11] Wilf, H. S. (1983) Three problems in combinatorial asymptotics. *J. Combin. Theory Ser. A* **35** 199–207.

https://doi.org/10.1017/S0963548300004399 Published online by Cambridge University Press