Electronic Journal of Probability

Electron. J. Probab. **23** (2018), no. 3, 1–17.
ISSN: 1083-6489 https://doi.org/10.1214/17-EJP126

On the strange domain of attraction to generalized Dickman distributions for sums of independent random variables

Ross G Pinsky*

Abstract

Let $\{B_k\}_{k=1}^\infty, \{X_k\}_{k=1}^\infty$ all be independent random variables. Assume that $\{B_k\}_{k=1}^\infty$ are $\{0,1\}$-valued Bernoulli random variables satisfying $B_k \stackrel{\text{dist}}{=} \text{Ber}(p_k)$, with $\sum_{k=1}^\infty p_k = \infty$, and assume that $\{X_k\}_{k=1}^\infty$ satisfy $X_k > 0$ and $\mu_k \equiv EX_k < \infty$. Let $M_n = \sum_{k=1}^n p_k\mu_k$, assume that $M_n \to \infty$ and define the normalized sum of independent random variables $W_n = \frac{1}{M_n}\sum_{k=1}^n B_kX_k$. We give a general condition under which $W_n \stackrel{\text{dist}}{\to} c$, for some $c \in [0,1]$, and a general condition under which W_n converges weakly to a distribution from a family of distributions that includes the generalized Dickman distributions $\text{GD}(\theta), \theta > 0$. In particular, we obtain the following result, which reveals a strange domain of attraction to generalized Dickman distributions. Assume that $\lim_{k\to\infty} \frac{X_k}{\mu_k} \stackrel{\text{dist}}{=} 1$. Let J_μ, J_p be nonnegative integers, let $c_\mu, c_p > 0$ and let
$\mu_n \sim c_\mu n^{a_0}\prod_{j=1}^{J_\mu}(\log^{(j)} n)^{a_j},\ p_n \sim c_p\big(n^{b_0}\prod_{j=1}^{J_p}(\log^{(j)} n)^{b_j}\big)^{-1},\ b_{J_p} \neq 0$, where $\log^{(j)}$ denotes the jth iterate of the logarithm.
If

$$i.\ J_p \le J_\mu;$$
$$ii.\ b_j = 1,\ 0 \le j \le J_p;$$
$$iii.\ a_j = 0,\ 0 \le j \le J_p - 1,\ \text{and}\ \ a_{J_p} > 0,$$

then $\lim_{n\to\infty} W_n \stackrel{\text{dist}}{=} \frac{1}{\theta}\text{GD}(\theta),$ where $\theta = \frac{c_p}{a_{J_p}}$.

Otherwise, $\lim_{n\to\infty} W_n \stackrel{\text{dist}}{=} \delta_c$, where $c \in \{0,1\}$ depends on the above parameters.
We also give an application to the statistics of the number of inversions in certain random shuffling schemes.

Keywords: Dickman function; generalized Dickman distribution; domain of attraction; normalized sums of independent random variables.
AMS MSC 2010: 60F05.
Submitted to EJP on January 30, 2017, final version accepted on November 14, 2017.

*Technion-Israel Institute of Technology, Israel. E-mail: pinsky@math.technion.ac.il

1 Introduction and statement of results

The Dickman function ρ_1 is the unique function, continuous on $(0,\infty)$, and satisfying the differential-delay equation

$$\begin{aligned}&\rho_1(x)=0,\ x\le 0;\\&\rho_1(x)=1,\ x\in(0,1];\\&x\rho_1'(x)+\rho_1(x-1)=0,\ x>1.\end{aligned}$$

This function has an interesting role in number theory and probability, which we describe briefly in the final section of the paper. With a little work, one can show that the Laplace transform of ρ_1 is given by $\int_0^\infty \rho_1(x)e^{-\lambda x}dx=\exp(\gamma+\int_0^1\frac{e^{-\lambda x}-1}{x}dx)$, where γ is Euler's constant. (See, for example, [6] or [9].) From this it follows that $\int_0^\infty \rho_1(x)dx=e^\gamma$, and consequently, that $e^{-\gamma}\rho_1$ is a probability density on $[0,\infty)$. We will call this probability distribution the *Dickman distribution.* We denote its density by $p_1(x)=e^{-\gamma}\rho_1(x)$, and we denote by D_1 a random variable distributed according to the Dickman distribution. Differentiating the Laplace transform $E\exp(-\lambda D_1)=\exp(\int_0^1\frac{e^{-\lambda x}-1}{x}dx)$ of D_1 at $\lambda=0$ shows that $ED_1=1$. This distribution decays very rapidly; indeed, it is not hard to show that $p_1(x)\le\frac{e^{-\gamma}}{\Gamma(x+1)},\ x\ge 0$ [6].

In fact, for all $\theta>0$, $\exp(\theta\int_0^1\frac{e^{-\lambda x}-1}{x}dx)$ is the Laplace transform of a probability distribution. (We will prove this directly; however, this fact follows from the theory of infinitely divisible distributions, and shows that the distribution in question is infinitely divisible.) This distribution has density $p_\theta=\frac{e^{-\theta\gamma}}{\Gamma(\theta)}\rho_\theta$, where ρ_θ satisfies the differential-delay equation

$$\begin{aligned}&\rho_\theta(x)=0,\ x\le 0;\\&\rho_\theta(x)=x^{\theta-1},\ 0<x\le 1;\\&x\rho_\theta'(x)+(1-\theta)\rho_\theta(x)+\theta\rho_\theta(x-1)=0,\ x>1.\end{aligned}\tag{1.1}$$

We will call such distributions *generalized Dickman distributions* and denote them by GD(θ). We denote by D_θ a random variable with the GD(θ) distribution. Differentiating its Laplace transform at $\lambda=0$ shows that $ED_\theta=\theta$. It is not hard to show that $p_\theta(x)\le\frac{C_\theta}{\Gamma(x+1)}, x\ge 1$, for an appropriate constant C_θ.

In fact, the scope of this paper leads us to consider a more general family of distributions than the generalized Dickman distributions. Let $\mathcal{X}\ge 0$ be a random variable satisfying $E\mathcal{X}\le 1$. Then, as we shall see, for $\theta>0$, there exists a distribution whose Laplace transform is $\exp\left(\theta\int_0^1\frac{E\exp(-\lambda x\mathcal{X})-1}{x}dx\right)$. We will denote this distribution by $GD^{(\mathcal{X})}(\theta)$ and we denote a random variable with this distribution by $D_\theta^{(\mathcal{X})}$. (When $\mathcal{X}\equiv 1$, we revert to the previous notation for generalized Dickman distributions.) Differentiating the Laplace transform at $\lambda=0$ shows that $ED_\theta^{(\mathcal{X})}=\theta E\mathcal{X}$.

It is known that the generalized Dickman distribution $GD(\theta)$ arises as the limiting distribution of $\frac{1}{n}\sum_{k=1}^n kY_k$, where the $\{Y_k\}_{k=1}^\infty$ are independent random variables with Y_k distributed according to the Poisson distribution with parameter $\frac{\theta}{k}$ [1]. It is also known that the Dickman distribution $GD(1)$ arises as the limiting distribution of $\frac{1}{n}\sum_{k=1}^n kY_k$ as $n\to\infty$, where the $\{Y_k\}_{k=1}^\infty$ are independent Bernoulli random variables satisfying $P(Y_k=1)=1-P(Y_k=0)=\frac{1}{k}$. Such behavior is in distinct contrast to the law of large numbers behavior of a "well-behaved" sequence of independent random variables $\{Z_k\}_{k=1}^\infty$ with finite first moments; namely, that $\frac{1}{M_n}\sum_{k=1}^n Z_k$ converges in distribution to 1 as $n\to\infty$, where $M_n=\sum_{k=1}^n EZ_k$.

The purpose of this paper is to understand when the law of large numbers fails and a distribution from the family GD$^{(\mathcal{X})}(\theta)$ arises in its stead. From the above examples, we see that generalized Dickman distributions sometimes arise as limits of normalized sums from a sequence $\{V_k\}_{k=1}^\infty$ of independent random variables which are non-negative and satisfy the following three conditions: (i) $\lim_{k\to\infty} P(V_k=0)=1$, (ii) $\lim_{k\to\infty} \frac{V_k|V_k>0}{E(V_k|V_k>0)} \stackrel{\text{dist}}{=} 1$ and (iii) $\sum_{k=1}^{\infty} EV_k=\infty$. (In the above examples, kY_k plays the role of V_k.) It turns out that these three conditions are very far from sufficient for a generalized Dickman distribution to arise. In fact, as we shall see in Theorem 1.2 below, such distributions arise only in a strange sequence of very narrow windows of opportunity.

In light of the above discussion, we will consider the following setting. Let $\{B_k\}_{k=1}^{\infty}$, $\{X_k\}_{k=1}^{\infty}$ be mutually independent sequences of independent random variables. Assume that $\{B_k\}_{k=1}^{\infty}$ are Bernoulli random variables satisfying:

$$P(B_k=1)=1-P(B_k=0)=p_k\in[0,1), \tag{1.2}$$

and assume that $\{X_k\}_{k=1}^{\infty}$ satisfy:

$$X_k>0, \quad \mu_k\equiv EX_k<\infty. \tag{1.3}$$

Let

$$M_n=\sum_{k=1}^{n} p_k\mu_k, \tag{1.4}$$

and define

$$W_n=\frac{1}{M_n}\sum_{k=1}^{n} B_kX_k. \tag{1.5}$$

We will be interested in the limiting behavior of W_n. In order to avoid trivialities, we will assume that

$$\lim_{n\to\infty} M_n=\infty \ \text{ and } \ \sum_{k=1}^{\infty} p_k=\infty, \tag{1.6}$$

since otherwise $\sum_{n=1}^{\infty} B_kX_k$ is almost surely finite.

Note that for the example brought with the Pois($\frac{\theta}{k}$)-distribution, we have $p_k=1-e^{-\frac{\theta}{k}}$, X_k is distributed according to $kY_k|\{Y_k>0\}$, where Y_k has the Pois($\frac{\theta}{k}$) distribution, $\mu_k=\frac{\theta}{1-e^{-\frac{\theta}{k}}}$ and $M_n=n\theta$. And for the example with the Ber($\frac{1}{k}$)-distribution, we have $p_k=\frac{1}{k}$, $X_k=k$ deterministically, $\mu_k=k$ and $M_n=n$. In the first of these two examples, $\frac{X_k}{\mu_k}\stackrel{\text{dist}}{\to}1$, and in the second one, $\frac{X_k}{\mu_k}\stackrel{\text{dist}}{=}1$ for all k.

Our first theorem gives a general condition for $W_n\stackrel{\text{dist}}{\to}c$ (which is the law of large numbers if $c=1$), and a general condition for convergence to a limiting distribution from the family of distributions GD$^{(\mathcal{X})}(\theta)$. Using this theorem, we can prove our second theorem, which reveals the strange domain of attraction to generalized Dickman distributions. (Of course, we are using the term "domain of attraction" not in its classical sense, since our sequence of random variables, although independent, are not identically distributed.) Let δ_c denote the degenerate distribution at c.

Theorem 1.1. *Let W_n be as in* (1.5)*, where $\{B_k\}_{k=1}^{\infty}$, $\{X_k\}_{k=1}^{\infty}$ and M_n are as in* (1.2)*-*(1.4) *and* (1.6)*.*
i. Assume that $\{\frac{X_k}{\mu_k}\}_{k=1}^{\infty}$ is uniformly integrable (which occurs automatically if $\lim_{k\to\infty}\frac{X_k}{\mu_k}\stackrel{dist}{=}1$).

a. Assume also that

$$\lim_{n\to\infty}\frac{\max_{1\le k\le n}\mu_k}{M_n}=0. \tag{1.7}$$

Then

$$\lim_{n\to\infty} W_n \stackrel{dist}{=} 1.$$

b. Assume also that there exists a sequence $\{K_n\}_{n=1}^{\infty}$ *such that*

$$\lim_{n\to\infty}\sum_{k=K_n+1}^{n} p_k = 0, \tag{1.8}$$

and

$$\lim_{n\to\infty}\frac{\max_{1\le k\le K_n}\ \mu_k}{M_n} = 0. \tag{1.9}$$

If

$$c \equiv \lim_{n\to\infty}\frac{M_{K_n}}{M_n} \text{ exists}, \tag{1.10}$$

then

$$\lim_{n\to\infty} W_n \stackrel{dist}{=} c.$$

If (1.10) *does not hold, then the distributions of* $\{W_n\}_{n=1}^{\infty}$ *form a tight sequence whose set of accumulation points is* $\{\delta_c : c \in A\}$*, where* A *denotes the set of accumulation points of the sequence* $\{\frac{M_{K_n}}{M_n}\}_{n=1}^{\infty}$.

ii. Assume that there exists a random variable $\mathcal{X}$ *such that*

$$\lim_{k\to\infty}\frac{X_k}{\mu_k} \stackrel{dist}{=} \mathcal{X}. \tag{1.11}$$

Assume also that $\{\mu_k\}_{k=1}^{\infty}$ *is increasing, that* $\lim_{k\to\infty} p_k = 0$ *and that there exist* $\theta, L \in (0,\infty)$ *such that*

$$\lim_{k\to\infty}\frac{p_k\mu_k}{\mu_{k+1}-\mu_k} = \theta, \quad \lim_{k\to\infty}\frac{\mu_k}{M_k} = L. \tag{1.12}$$

Then

$$\lim_{n\to\infty} W_n \stackrel{dist}{=} LD^{(\mathcal{X})}(\theta),$$

where $D^{(\mathcal{X})}(\theta)$ *is a random variable with the GD*$^{(\mathcal{X})}(\theta)$ *distribution.*

Remark 1. In (1.12), necessarily $L \le \frac{1}{\theta}$. Indeed, if $\{p_k\}_{k=1}^{\infty}$ and $\{\mu_k\}_{k=1}^{\infty}$ satisfy the conditions of part (ii), and we choose $X_k = \mu_k$, then $W_n \stackrel{\text{dist}}{\to} LD_\theta$. Since $EW_n = 1$ and $ED_\theta = \theta$, it follows from Fatou's lemma that $L \le \frac{1}{\theta}$. In most cases of interest, one has $L = \frac{1}{\theta}$.

Remark 2. By Fatou's lemma, the random variable $\mathcal{X}$ in part (ii) must satisfy $E\mathcal{X} \le 1$.

Remark 3. The uniform integrability of $\{\frac{X_k}{\mu_k}\}_{k=1}^{\infty}$ in part (i) occurs automatically if $\lim_{k\to\infty}\frac{X_k}{\mu_k} \stackrel{\text{dist}}{=} 1$, because if a sequence $\{Y_k\}_{k=1}^{\infty}$ of random variables satisfies $Y_k \stackrel{\text{dist}}{\to} Y$, and $E|Y_k| < \infty$, then $E|Y_k| \to E|Y|$ is equivalent to uniform integrability.

Remark 4. In the case that $X_k = \mu_k$, or more generally, if $EX_k^2 \le C\mu_k^2$, for all k and some $C > 0$, then

$$Var(W_n) \le \frac{C\sum_{k=1}^{N} p_k\mu_k^2}{M_n^2} = C\frac{\sum_{k=1}^{n} p_k\mu_k^2}{(\sum_{k=1}^{n} p_k\mu_k)^2} \le C\frac{\sup_{1\le k\le n}\mu_k}{M_n}.$$

Thus, in this case part (i-a) follows directly from the second moment method.

Using Theorem 1.1, we can prove the following theorem that exhibits the strange domain of attraction to generalized Dickman distributions. Let $\log^{(j)}$ denote the jth iterate of the logarithm, and make the convention $\prod_{j=1}^{0} = 1$.

Theorem 1.2. *Let W_n be as in* (1.5)*, where $\{B_k\}_{k=1}^{\infty}$, $\{X_k\}_{k=1}^{\infty}$ and M_n are as in* (1.2)-(1.4)*. Assume also that $\lim_{k\to\infty}\frac{X_k}{\mu_k}\stackrel{dist}{=}1$. Let J_μ, J_p be nonnegative integers, let $c_\mu, c_p>0$ and define*

$$\mu(x)=c_\mu x^{a_0}\prod_{j=1}^{J_\mu}(\log^{(j)}x)^{a_j},$$
$$p(x)=c_p\big(x^{b_0}\prod_{j=1}^{J_p}(\log^{(j)}x)^{b_j}\big)^{-1},$$

with $b_{J_p}\neq 0$. Assume that

$$\mu_k\sim\mu(k),\ \ p_k\sim p(k);$$
$$\mu_{k+1}-\mu_k\sim\mu'(k).$$

Assume that the exponents $\{a_j\}_{j=0}^{J_\mu}, \{b_j\}_{j=0}^{J_p}$ have been chosen so that (1.6) *holds. If*

$$\begin{aligned}&i.\ J_p\le J_\mu;\\&ii.\ b_j=1,\ 0\le j\le J_p;\\&iii.\ a_j=0,\ 0\le j\le J_p-1,\ \textit{and}\ \ a_{J_p}>0,\end{aligned}\tag{1.13}$$

then

$$\lim_{n\to\infty}W_n\stackrel{dist}{=}\frac{1}{\theta}D_\theta,\ \textit{with}\ \theta=\frac{c_p}{a_{J_p}},$$

where D_θ is a random variable with the GD(θ) distribution.
Otherwise, $\lim_{n\to\infty}W_n\stackrel{dist}{=}c$, where $c\in\{0,1\}$. To determine c, let

$$\kappa_\mu=\min\{0\le j\le J_\mu: a_j\neq 0\}\ \ \textit{and}\ \ \kappa_p=\min\{0\le j\le J_p: b_j\neq 1\}.\tag{1.14}$$

If $\{0\le j\le J_\mu: a_j\neq 0\}$ is not empty, $a_{\kappa_\mu}>0$ and either $\{0\le j\le J_p: b_j\neq 1\}$ is empty and $\kappa_\mu<J_p$, or $\{0\le j\le J_p: b_j\neq 1\}$ is not empty and $\kappa_\mu<\kappa_p$, then $c=0$; otherwise, $c=1$.

Remark 1. Note that if one chooses $\mu_k=\mu(k)$ and $p_k=p(k)$, then the condition $\mu_{k+1}-\mu_k\sim\mu'(k)$ is always satisfied.

Remark 2. Theorem 1.2 shows that to obtain a generalized Dickman distribution, $\{p_k\}_{k=1}^{\infty}$ in particular must be set in a very restricted fashion. For some intuition regarding this phenomenon, take the situation where $X_k=\mu_k$, and consider the sequence $\{\sigma^2(W_n)\}_{n=1}^{\infty}$ of variances. This sequence converges to 0 in the cases where W_n converges to 1, converges to ∞ in the cases where W_n converges to 0, and converges to a positive number in the cases where W_n converges to a generalized Dickman distribution.

We now state explicitly what Theorem 1.2 yields in the cases $J_p=0,1$.

$\mathbf{J_p=0}$. We have

$$p_n\sim\frac{c_p}{n^{b_0}},\ b_0>0,\ \ \mu_n\sim c_\mu n^{a_0}\prod_{j=1}^{J_\mu}(\log^{(j)}n)^{a_j}.$$

In order that (1.6) hold, we require $b_0\le 1$. We also require either: $a_0-b_0>-1$; or $a_0-b_0=-1$ and $a_1>-1$; or $a_0-b_0=a_1=-1$ and $a_2>-1$; etc.
If $b_0=1$ and $a_0>0$, then

$$\lim_{n\to\infty}W_n\stackrel{dist}{=}\frac{1}{\theta}D_\theta,\ \textit{where}\ \theta=\frac{c_p}{a_0}.$$

Otherwise, $\lim_{n\to\infty}W_n\stackrel{dist}{=}1$.

$\mathbf{J_p = 1}$. We have

$$p_n \sim \frac{c_p}{n^{b_0}(\log n)^{b_1}},\ b_1 \neq 0, \quad \mu_n \sim c_\mu n^{a_0} \prod_{j=1}^{J_\mu} (\log^{(j)} n)^{a_j}.$$

In order that (1.6) hold, we require either $b_0 = 0$ and $b_1 > 0$, or $0 < b_0 < 1$, or $b_0 = 1$ and $b_1 \leq 1$. We also require either: $a_0 - b_0 > -1$; or $a_0 - b_0 = -1$ and $a_1 - b_1 > -1$; or $a_0 - b_0 = a_1 - b_1 = -1$ and $a_2 > -1$; etc.
If $J_\mu \geq 1$, $b_0 = b_1 = 1$, $a_0 = 0$ *and* $a_1 > 0$, *then*

$$\lim_{n\to\infty} W_n \stackrel{dist}{=} \frac{1}{\theta} D_\theta, \text{ where } \theta = \frac{c_p}{a_1}.$$

If $b_0 = 1$ *and* $a_0 > 0$, *then* $\lim_{n\to\infty} W_n \stackrel{dist}{=} 0$.
Otherwise, $\lim_{n\to\infty} W_n \stackrel{dist}{=} 1$.

Remark. In [3] and [8], where the GD(1) distribution arises, one has $J_p = 1$ with $J_\mu = 1$, $b_0 = b_1 = 1, a_0 = 0, a_1 = 1, c_p = c_\mu = 1$.

The organization of the rest of the paper is as follows. In section 2 we use Theorems 1.1 and 1.2 to investigate a question raised in [5] concerning the statistics of the number of inversions in certain random shuffling schemes. In sections 3 and 4 respectively we prove Theorems 1.1 and 1.2. In section 5 we prove a couple basic facts about generalized Dickman distributions. In particular, we provide a rather probabilistic proof that the distribution whose Laplace transform is given by $\exp(\theta \int_0^1 \frac{e^{-\lambda x}-1}{x} dx)$ possesses a density p_θ of the form $p_\theta = c_\theta \rho_\theta$, where ρ_θ satisfies (1.1). We also give a reference for the formula $c_\theta = \frac{e^{-\theta\gamma}}{\Gamma(\theta)}$. Finally, in section 6, we offer a little historical background concerning the Dickman function ρ_1 and its connection to number theory and probability.

2 An application to random permutations

We consider a setup that appeared in [5], and which in the terminology of this paper can be described as follows. For each $k \in \mathbb{N}$, let $E_k \subset \{1, \ldots, k-1\}$. Let X_k be uniformly distributed on E_k, and let $B_k \stackrel{\text{dist}}{=} \text{Ber}(\frac{|E_k|}{k})$. So

$$\mu_k = \frac{1}{|E_k|} \sum_{l \in E_k} l, \quad p_k = \frac{|E_k|}{k}.$$

Define

$$I_n = \sum_{k=1}^{n} B_k X_k.$$

We allow $E_k = \emptyset$, in which case $B_k = 0$ and X_k is not defined. In such a case, we define $B_k X_k = 0$ and $\mu_k = 0$. We always have $E_1 = \emptyset$.

Consider first the case that $E_k = \{1, \ldots, k-1\}$. Then $B_1 X_1 = 0$ and for $2 \leq k \leq n$, $B_k X_k$ is uniformly distributed over $\{0, 1 \ldots, k-1\}$. In this case, I_n has the distribution of the number of inversions in a uniformly random permutation from S_n. (The authors in [5] have a typo and wrote $E_k = \{1, \ldots, k\}$ instead.) To see this, consider the following shuffling procedure for n cards, numbered from 1 to n. The cards are to be inserted in a row, one by one, in order of their numbers. At step one, card number 1 is set down. The number of inversions created by this step is zero, which is given by $B_1 X_1$. At step k, for $k \in \{2, \ldots, n\}$, card number k is randomly inserted in the current row of cards, numbered 1 to $k-1$. Thus, for any $j \in \{0, 1, \ldots, k-1\}$, card number k has probability $\frac{1}{k}$

of being placed in the position with j cards to its right (and $k-1-j$ cards to its left), in which case this step will have created j new inversions, and this is represented by $B_k X_k$. It is clear from the construction that the random variables $\{B_k X_k\}_{k=1}^n$ are independent. Thus, I_n indeed gives the number of inversions in a uniformly random permutation from S_n. It is well-known that the law of large numbers and the central limit theorem hold for I_n in this case. (In particular, using the above representation, a direction calculation shows that $EI_n = \frac{n(n-1)}{4}$ and that $\text{Var}(I_n) = O(n^3)$; thus the law of large numbers follows from the second moment method.)

Consider now the general case that $E_k \subset \{1, \ldots, k-1\}$. Then I_n gives the number of inversions in a random permutation created by a shuffling procedure in the same spirit as the above one. At step k, with probability $1 - \frac{|E_k|}{k}$, card number k is inserted at the right end of the row, thereby creating no new inversions, and for each $j \in E_k$, with probability $\frac{1}{k}$ it is inserted in the position with j cards to its right, thereby creating j new inversions.

In particular, as a warmup consider the cases $E_k = \{1\}$ and $E_k = \{k-1\}$, $2 \le k \le n$. In each of these two cases, at step k, $2 \le k \le n$, card number k is inserted at the right end of the row with probability $1 - \frac{1}{k}$. In the first case, with probability $\frac{1}{k}$ card number k is inserted immediately to the left of the right most card, thereby creating one new inversion, while in the second case, with probability $\frac{1}{k}$ card number k is inserted at the left end of the row, thereby creating $k-1$ new inversions. In both cases $\frac{X_n}{\mu_n} \overset{\text{dist}}{=} 1$ for all n, and in both cases, $p_k = \frac{1}{k}$. In the first case, $\mu_k = 1$ while in the second case, $\mu_k = k-1$. Thus, in the first case, $M_n = \sum_{k=1}^n p_k \mu_k \sim \log n$, and in the second case, $M_n \sim n$. Therefore, it follows from Theorem 1.1 or 1.2 that in the first case $\frac{I_n}{\log n}$ converge in distribution to 1, while in the second case, $\frac{I_n}{n}$ converges in distribution to $\text{GD}(1)$.

The authors of [5] ask which choices of $\{E_k\}_{k=1}^\infty$ lead to the Dickman distribution and which choices lead to the central limit theorem. Of course, the law of large numbers is a prerequisite for the central limit theorem. The following theorem gives sufficient conditions for the law of large numbers to hold and sufficient conditions for convergence to a distribution from the family $\text{GD}^{(\mathcal{X})}(\theta)$. In order to avoid trivialities, we need to assume that (1.6) holds. Recalling that $\mu_k = 0$ when $|E_k| = 0$, and that $\mu_k \ge 1$ otherwise, note that

$$M_n = EI_n = \sum_{k=1}^\infty \frac{|E_k|}{k}\mu_k \ge \sum_{k=1}^\infty \frac{|E_k|}{k} = \sum_{k=1}^\infty p_k.$$

Thus, in the present context the requirement (1.6) is

$$\sum_{k=1}^\infty \frac{|E_k|}{k} = \infty, \tag{2.1}$$

which holds in particular if $E_k \neq \emptyset$ for all sufficiently large k.

Theorem 2.1. *Assume that* (2.1) *holds.*

i. Assume that at least one of the following conditions holds:

a. $\lim_{k\to\infty} |E_k| = \infty$ *and* $\{\frac{\max_{1\le k\le n}\mu_k}{\sum_{k=1}^n \frac{\mu_k}{k}}\}_{n=1}^\infty$ *is bounded;*

b. $\lim_{n\to\infty} \frac{\max_{1\le k\le n}\mu_k}{\sum_{k=1}^n \frac{\mu_k}{k}} = 0$.

Then $\frac{I_n}{EI_n} \overset{\text{dist}}{\to} 1$.

ii. Assume that $|E_k| = N \ge 1$*, for all large* k*, and that* $\frac{X_k}{\mu_k} \overset{\text{dist}}{\to} \mathcal{X}$*. Also assume that* $\mu_k \sim \mu(k)$ *and* $\mu_{k+1} - \mu_k \sim \mu'(k)$*, where* $\mu(x) = c_\mu x^{a_0} \prod_{j=1}^{J_\mu} (\log^{(j)} x)^{a_j}$*, with* $a_0 > 0$*. Then* $\frac{I_n}{EI_n} \overset{\text{dist}}{\to} \frac{1}{\theta} D_\theta^{(\mathcal{X})}$*, with* $\theta = \frac{N}{a_0}$*, where* $D_\theta^{(\mathcal{X})}$ *is a random variable with the* $\text{GD}^{(\mathcal{X})}(\theta)$ *distribution.*

Remark 1. The condition on $\{\mu_k\}$ in part (i-a) is just a very weak regularity requirement on its growth rate (recall that $1 \le \mu_k < k-1$). The condition in part (i-b) is fulfilled if $\mu_k \sim \mu(k)$ and $\mu_{k+1} - \mu_k \sim \mu'(k)$, where $\mu(x) = c_\mu \prod_{j=1}^{J_\mu} (\log^{(j)} x)^{a_j}$ with $J_\mu \ge 0$.

Remark 2. Note that the random variable $\mathcal{X}$ in part (ii) takes on no more than N distinct values.

Proof. Assume first that the condition in part (i-a) holds. We claim that since $\{\frac{\max_{1\le k\le n}\mu_k}{\sum_{k=1}^n \frac{\mu_k}{k}}\}_{n=1}^\infty$ is bounded, there exists a sequence of positive integers $\{\gamma_n\}_{n=1}^\infty$ satisfying $\lim_{n\to\infty}\gamma_n = \infty$ and such that $\{\frac{\max_{1\le k\le n}\mu_k}{\sum_{k=\gamma_n+1}^n \frac{\mu_k}{k}}\}_{n=1}^\infty$ is also bounded. Indeed, assume to the contrary that the above sum is unbounded for all such choices of $\{\gamma_n\}_{n=1}^\infty$. Then necessarily, $\{\mu_n\}_{n=1}^\infty$ is unbounded. (Indeed, since by assumption $|E_k| \ge 1$ for sufficiently large k, the same is true for μ_k, and thus, choosing, for example, $\gamma_n = [n^{\frac12}]$, it follows that for sufficiently large n, $\sum_{k=\gamma_n+1}^n \frac{\mu_k}{k} \ge \sum_{k=\gamma_n+1}^n \frac1k \sim \frac12 \log n$. Thus $\{\mu_n\}_{n=1}^\infty$ must be unbounded if $\{\frac{\max_{1\le k\le n}\mu_k}{\sum_{k=\gamma_n+1}^n \frac{\mu_k}{k}}\}_{n=1}^\infty$ is unbounded.) Also, since $\mu_k < k$, we have $\sum_{k=1}^n \frac{\mu_k}{k} < \gamma_n + \sum_{k=\gamma_n+1}^n \frac{\mu_k}{k}$, and it follows from the boundedness of $\{\frac{\max_{1\le k\le n}\mu_k}{\sum_{k=1}^n \frac{\mu_k}{k}}\}_{n=1}^\infty$ and the unboundedness of $\{\frac{\max_{1\le k\le n}\mu_k}{\sum_{k=\gamma_n}^n \frac{\mu_k}{k}}\}_{n=1}^\infty$ that the unbounded sequence $\{\max_{1\le k\le n}\mu_k\}_{n=1}^\infty$ has the property that $\{\frac{\max_{1\le k\le n}\mu_k}{\gamma_n}\}_{n=1}^\infty$ is bounded for all sequences $\{\gamma_n\}_{n=1}^\infty$ satisfying $\lim_{n\to\infty}\gamma_n = \infty$, which is impossible.

Now let $\{\gamma_n\}_{n=1}^\infty$ be a sequence such that $\lim_{n\to\infty}\gamma_n = \infty$ and $\{\frac{\max_{1\le k\le n}\mu_k}{\sum_{k=\gamma_n+1}^n \frac{\mu_k}{k}}\}_{n=1}^\infty$ is bounded. Since

$$M_n = \sum_{k=1}^n \frac{|E_k|}{k}\mu_k \ge (\min_{k>\gamma_n} |E_k|) \sum_{k=\gamma_n}^n \frac{\mu_k}{k}$$

and $\{\frac{\max_{1\le k\le n}\mu_k}{\sum_{k=\gamma_n+1}^n \frac{\mu_k}{k}}\}_{n=1}^\infty$ is bounded, it follows from the first condition in (i-a) that (1.7) holds.

Now assume that the condition in part (i-b) holds. Since $M_n \ge \sum_{k=1}^n \frac{\mu_k}{k}$, it follows again that (1.7) holds.

Thus, assuming either (i-a) or (i-b), it follows from part (i-a) of Theorem 1.1 that $\frac{I_n}{EI_n} \stackrel{\text{dist}}{\to} 1$.

Now assume that the condition in part (ii) holds. Then $p_k = \frac{N}{k}$, for large k, and $\mu_k \sim c_\mu k^{a_0} \prod_{j=1}^{J_\mu} (\log^{(j)} k)^{a_j}$, with $a_0 > 0$. Thus,

$$M_n = \sum_{k=1}^n \frac{|E_k|}{k}\mu_k \sim \frac{Nc_\mu}{a_0} n^{a_0} \prod_{j=1}^{J_\mu} (\log^{(j)} n)^{a_j},$$

and $\lim_{k\to\infty} \frac{\mu_k}{M_k} = \frac{a_0}{N}$. Also, if the condition in part (ii) holds, then $\mu_{k+1} - \mu_k \sim a_0 c_\mu k^{a_0-1} \prod_{j=1}^{J_\mu} (\log^{(j)} k)^{a_j}$. Thus, $\lim_{k\to\infty} \frac{p_k\mu_k}{\mu_{k+1}-\mu_k} = \frac{N}{a_0}$. We conclude from part (ii) of Theorem 1.1 that $\frac{I_n}{EI_n} \stackrel{\text{dist}}{\to} \frac1\theta \mathrm{GD}^{(\mathcal{X})}(\theta)$, where $\theta = \frac{N}{a_0}$. □

3 Proof of Theorem 1.1

Since $EW_n = 1$, for all n, the distributions of $\{W_n\}_{n=1}^\infty$ are tight. Thus, since the random variables are nonnegative, it suffices to show that their Laplace transforms $E\exp(-\lambda W_n)$ converge under the conditions of part (i) to $\exp(-\lambda c)$, for the specified value of c, and under the conditions of part (ii) to $\exp(\theta \int_0^1 \frac{Ee^{-L\lambda x\mathcal{X}}-1}{x} dx)$, which is the Laplace transform of $LD^{(\mathcal{X})}(\theta)$.

Proof of part (i). Note that part (i-a) is the particular case of part (i-b) in which one can choose $K_n = n$, and then (1.10) holds with $c = 1$. Thus, it suffices to consider part (i-b).

We have for $\lambda > 0$,

$$\begin{aligned} &E\exp(-\lambda W_n) = \\ &= \prod_{k=1}^{n} E\exp(-\frac{\lambda}{M_n}B_kX_k) = \prod_{k=1}^{n}\Big(1-p_k\big(1-E\exp(-\frac{\lambda}{M_n}X_k)\big)\Big) = \\ &\prod_{k=1}^{K_n}\Big(1-p_k\big(1-E\exp(-\frac{\lambda}{M_n}X_k)\big)\Big)\prod_{k=K_n+1}^{n}\Big(1-p_k\big(1-E\exp(-\frac{\lambda}{M_n}X_k)\big)\Big). \end{aligned} \tag{3.1}$$

Since

$$\prod_{k=K_n+1}^{n}(1-p_k) \le \prod_{k=K_n+1}^{n}\Big(1-p_k\big(1-E\exp(-\frac{\lambda}{M_n}X_k)\big)\Big) \le 1,$$

it follows from assumption (1.8) that

$$\lim_{n\to\infty}\prod_{k=K_n+1}^{n}\Big(1-p_k\big(1-E\exp(-\frac{\lambda}{M_n}X_k)\big)\Big) = 1. \tag{3.2}$$

Applying the mean value theorem to $E\exp(-\frac{\lambda}{M_n}X_k)$ as a function of λ, and recalling that $\mu_k = EX_k$, we have

$$\frac{\lambda}{M_n}EX_k\exp(-\frac{\lambda}{M_n}X_k) \le 1-E\exp(-\frac{\lambda}{M_n}X_k) \le \lambda\frac{\mu_k}{M_n}. \tag{3.3}$$

The assumption that $\{\frac{X_k}{\mu_k}\}_{k=1}^{\infty}$ is uniformly integrable means that $\lim_{N\to\infty}\sup_{1\le k<\infty}E(\frac{X_k}{\mu_k}1_{\frac{X_k}{\mu_k}>N}) = 0$. Thus, in light of (1.9) and the uniform integrability assumption, it follows that for all $\epsilon > 0$, there exists an n_ϵ such that

$$\begin{aligned} &\frac{\lambda}{M_n}EX_k\exp(-\frac{\lambda}{M_n}X_k) = \lambda\frac{\mu_k}{M_n}E\frac{X_k}{\mu_k}\exp(-\lambda\frac{\mu_k}{M_n}\frac{X_k}{\mu_k}) \ge (1-\epsilon)\lambda\frac{\mu_k}{M_n}, \\ &1\le k\le K_n,\ n\ge n_\epsilon. \end{aligned} \tag{3.4}$$

Thus, (3.3) and (3.4) yield

$$(1-\epsilon)\lambda\frac{\mu_k}{M_n} \le 1-E\exp(-\frac{\lambda}{M_n}X_k) \le \lambda\frac{\mu_k}{M_n},\ 1\le k\le K_n,\ n\ge n_\epsilon. \tag{3.5}$$

Since for any $\epsilon > 0$, there exists an $x_\epsilon > 0$ such that $-(1+\epsilon)x \le \log(1-x) \le -x$, for $0 < x < x_\epsilon$, it follows from (3.5) and (1.9) that there exists an n'_ϵ such that

$$\begin{aligned} &-(1+\epsilon)\lambda p_k\frac{\mu_k}{M_n} \le \log\Big(1-p_k\big(1-E\exp(-\frac{\lambda}{M_n}X_k)\big)\Big) \le -(1-\epsilon)\lambda p_k\frac{\mu_k}{M_n}, \\ &1\le k\le K_n,\ n\ge n'_\epsilon. \end{aligned} \tag{3.6}$$

From (3.6) we have

$$\begin{aligned} &-(1+\epsilon)\lambda\frac{\sum_{k=k_\epsilon}^{K_n}p_k\mu_k}{M_n} \le \log\prod_{k=1}^{K_n}\Big(1-p_k\big(1-E\exp(-\frac{\lambda}{M_n}X_k)\big)\Big) \le \\ &-(1-\epsilon)\lambda\frac{\sum_{k=k_\epsilon}^{K_n}p_k\mu_k}{M_n},\ n\ge n'_\epsilon. \end{aligned} \tag{3.7}$$

If

$$c \equiv \lim_{n\to\infty}\frac{M_{K_n}}{M_n} = \lim_{n\to\infty}\frac{\sum_{k=1}^{K_n}p_k\mu_k}{M_n} \tag{3.8}$$

exists, then from (3.1), (3.2), (3.7) and (3.8), along with the fact that $\epsilon > 0$ is arbitrary, we conclude that

$$\lim_{n\to\infty} E\exp(-\lambda W_n) = \exp(-\lambda c),$$

which proves that $\lim_{n\to\infty} W_n \stackrel{\text{dist}}{=} c$. The rest of the results in part (i-b), concerning accumulation points, follow in the same manner.

Proof of part (ii). From (3.1), we have

$$\log E\exp(-\lambda W_n) = \sum_{k=1}^{n} \log\Big(1 - p_k\big(1 - E\exp(-\frac{\lambda}{M_n}X_k)\big)\Big). \tag{3.9}$$

Since by assumption $\lim_{k\to\infty} p_k = 0$, for any $\epsilon > 0$ there exists a k_ϵ such that

$$\begin{aligned}&-(1+\epsilon)p_k\big(1 - E\exp(-\frac{\lambda}{M_n}X_k)\big) \le \log\Big(1 - p_k\big(1 - E\exp(-\frac{\lambda}{M_n}X_k)\big)\Big) \le \\ &-p_k\big(1 - E\exp(-\frac{\lambda}{M_n}X_k)\big),\ k \ge k_\epsilon.\end{aligned} \tag{3.10}$$

We now show that for any $\epsilon > 0$ there exists a k'_ϵ such that

$$(1-\epsilon)E\exp(-\lambda\frac{\mu_k}{M_n}\mathcal{X}) \le E\exp(-\frac{\lambda}{M_n}X_k) \le (1+\epsilon)E\exp(-\lambda\frac{\mu_k}{M_n}\mathcal{X}),\ k \ge k'_\epsilon. \tag{3.11}$$

By assumption (1.12) and the assumption that $\{\mu_n\}_{n=1}^\infty$ is increasing, there exists a C such that $\frac{\mu_k}{M_n} \le C$, for $1 \le k \le n$ and $n \ge 1$. By assumption, $\frac{X_k}{\mu_k} \stackrel{\text{dist}}{\to} \mathcal{X}$. Without loss of generality, we assume that all of these random variables are defined on the same space and that $\frac{X_k}{\mu_k} \to \mathcal{X}$ a.s. For $\delta > 0$, let

$$A_{k;\delta} = \{\sup_{l\ge k} |\frac{X_l}{\mu_l} - \mathcal{X}| \le \delta\}.$$

Then $A_{k;\delta}$ is increasing in k and $\lim_{k\to\infty} P(A_{k;\delta}) = 1$. We have

$$\int_{A^c_{k;\delta}} \exp(-\frac{\lambda}{M_n}X_k)dP \le P(A^c_{k;\delta}), \tag{3.12}$$

and

$$\begin{aligned}&\exp(-\lambda C\delta)\int_{A_{k;\delta}} \exp(-\lambda\frac{\mu_k}{M_n}\mathcal{X})dP \le \int_{A_{k;\delta}} \exp(-\lambda\frac{\mu_k}{M_n}\frac{X_k}{\mu_k})dP \le \\ &\exp(\lambda C\delta)\int_{A_{k;\delta}} \exp(-\lambda\frac{\mu_k}{M_n}\mathcal{X})dP.\end{aligned} \tag{3.13}$$

Now (3.11) follows from (3.12) and (3.13).

Letting $k''_\epsilon = \max(k_\epsilon, k'_\epsilon)$, it follows from (3.10) and (3.11) that

$$\begin{aligned}&-(1+\epsilon)p_k\Big(1 - (1-\epsilon)E\exp(-\lambda\frac{\mu_k}{M_n}\mathcal{X})\Big) \le \log\Big(1 - p_k\big(1 - E\exp(-\frac{\lambda}{M_n}X_k)\big)\Big) \le \\ &-p_k\Big(1 - (1+\epsilon)E\exp(-\lambda\frac{\mu_k}{M_n}\mathcal{X})\Big),\ k \ge k''_\epsilon.\end{aligned} \tag{3.14}$$

From (3.9) and (3.14) we have

$$\begin{aligned}&-\sum_{k=k''_\epsilon}^{n} p_k(1+\epsilon)\Big(1 - (1-\epsilon)E\exp(-\lambda\frac{\mu_k}{M_n}\mathcal{X})\Big) + o(1) \le \log E\exp(-\lambda W_n) \le \\ &-\sum_{k=k''_\epsilon}^{n} p_k\Big(1 - (1+\epsilon)E\exp(-\lambda\frac{\mu_k}{M_n}\mathcal{X})\Big),\ \text{as } n\to\infty.\end{aligned} \tag{3.15}$$

Define $x_k^{(n)}=\frac{\mu_k}{M_n}$, $k_\epsilon^{''}\le k\le n$, and $\Delta_k^{(n)}=x_{k+1}^{(n)}-x_k^{(n)}=\frac{\mu_{k+1}-\mu_k}{M_n}$, $k_\epsilon^{''}\le k\le n-1$. Then we have

$$\begin{aligned}&\sum_{k=k_\epsilon^{''}}^{n}p_k\Big(1-(1\pm\epsilon)E\exp(-\lambda\frac{\mu_k}{M_n}\mathcal{X})\Big)=\\&\sum_{k=k_\epsilon^{''}}^{n}\frac{1-(1\pm\epsilon)E\exp(-\lambda x_k^{(n)}\mathcal{X})}{x_k^{(n)}}\Delta_k^{(n)}\big(p_k\frac{\mu_k}{\mu_{k+1}-\mu_k}\big).\end{aligned}\tag{3.16}$$

By assumption, $\{\mu_k\}_{k=1}^{\infty}$ is increasing; thus $\{x_k^{(n)}\}_{k=k_\epsilon^{''}}^{n}$ is a partition of $[\frac{\mu_{k_\epsilon^{''}}}{M_n},\frac{\mu_n}{M_n}]$. By assumption, $\lim_{n\to\infty}\frac{\mu_{k_\epsilon^{''}}}{M_n}=0$ and $\lim_{n\to\infty}\frac{\mu_n}{M_n}=L$. We now show that the mesh, $\max_{k_\epsilon^{''}\le k\le n-1}\Delta_k^{(n)}$, of the partition converges to 0 as $n\to\infty$. Let $\Delta_{j_n}^{(n)}=\max_{k_\epsilon^{''}\le k\le n-1}\Delta_k^{(n)}$, where $k_\epsilon^{''}\le j_n\le n$. Without loss of generality, assume either that $\{j_n\}$ is bounded or that $\lim_{n\to\infty}j_n=\infty$. In the former case it is clear that $\max_{k_\epsilon^{''}\le k\le n-1}\Delta_k^{(n)}=\Delta_{j_n}^{(n)}=\frac{\mu_{j_n+1}-\mu_{j_n}}{M_n}\stackrel{n\to\infty}{\to}0$. Now consider the latter case. From assumption (1.12) and the assumption that $\lim_{k\to\infty}p_k=0$, it follows that $\lim_{n\to\infty}\frac{\mu_{n+1}-\mu_n}{M_n}=0$. Then we have

$$\max_{k_\epsilon^{''}\le k\le n-1}\Delta_k^{(n)}=\Delta_{j_n}^{(n)}=\frac{\mu_{j_n+1}-\mu_{j_n}}{M_n}=\frac{\mu_{j_n+1}-\mu_{j_n}}{M_{j_n}}\frac{M_{j_n}}{M_n}\le\frac{\mu_{j_n+1}-\mu_{j_n}}{M_{j_n}}\stackrel{n\to\infty}{\to}0.$$

Finally, we note that from (1.12) we have $\lim_{k\to\infty}p_k\frac{\mu_k}{\mu_{k+1}-\mu_k}=\theta$. In light of these facts, along with (3.15), (3.16) and the fact that $\epsilon>0$ is arbitrary, it follows that

$$\lim_{n\to\infty}\log E\exp(-\lambda W_n)=\theta\int_0^L\frac{E\exp(-\lambda x\mathcal{X})-1}{x}dx=\theta\int_0^1\frac{E\exp(-\lambda Lx\mathcal{X})-1}{x}dx,\tag{3.17}$$

□

4 Proof of Theorem 1.2

We will assume that $J_p,J_\mu\ge1$ so that we can use a uniform notation, leaving it to the reader to verify that the proof also goes through if J_p or J_μ is equal to zero.

First assume that (1.13) holds. Then by the assumptions in the theorem,

$$\begin{aligned}&1\le J_p\le J_\mu;\\&\mu_k\sim c_\mu\prod_{j=J_p}^{J_\mu}(\log^{(j)}k)^{a_j},\ a_{J_p}>0;\\&p_k\sim c_p\big(j\prod_{j=1}^{J_p}\log^{(j)}k\big)^{-1};\\&\mu_{k+1}-\mu_k\sim c_\mu a_{J_P}\frac{(\log^{(J_P)}k)^{a_{J_p}-1}}{j\prod_{j=1}^{J_p-1}\log^{(j)}k}\prod_{j=J_p+1}^{J_\mu}(\log^{(j)}k)^{a_j}.\end{aligned}$$

Thus,

$$M_n=\sum_{k=1}^{n}p_k\mu_k\sim c_\mu c_p\frac{(\log^{(J_p)}n)^{a_{J_p}}}{a_{J_p}}\prod_{j=J_p+1}^{J_\mu}(\log^{(j)}n)^{a_j}.$$

Consequently,

$$\lim_{k\to\infty}\frac{\mu_k}{M_k}=\frac{a_{J_p}}{c_p}\ \text{ and }\ \lim_{k\to\infty}\frac{p_k\mu_k}{\mu_{k+1}-\mu_k}=\frac{c_p}{a_{J_p}}.\tag{4.1}$$

Thus, from part (ii) of Theorem 1.1 it follows that $\lim_{n\to\infty} W_n \stackrel{\text{dist}}{=} \frac{1}{\theta}D_\theta$, where $\theta = \frac{c_p}{a_{J_p}}$.

Now assume that (1.13) does not hold. We need to show that $\{K_n\}_{n=1}^\infty$ can be defined so that (1.8) and (1.9) hold, and so that (1.10) holds with $c\in\{0,1\}$. We also have to show when $c=0$ and when $c=1$. Recall the definitions in (1.14). If $\{0\le j\le J_\mu : a_j\ne 0\}$ is empty, or if it is not empty and $a_{\kappa_\mu}<0$, then $\{\mu_k\}_{k=1}^\infty$ is bounded. Therefore, (1.8) and (1.9) hold with $K_n=n$ and it follows from part (i-a) of Theorem 1.1 that $\lim_{n\to\infty} W_n \stackrel{\text{dist}}{=} 1$. Thus, from now on we assume that $\{0\le j\le J_\mu : a_j\ne 0\}$ is not empty and that $a_{\kappa_\mu}>0$. In order to use uniform notation, we will assume that $\kappa_\mu>0$, leaving the reader to verify that the proof goes through if $\kappa_\mu=0$. Thus, we have

$$\mu_k \sim \prod_{j=\kappa_\mu}^{J_\mu} (\log^{(j)} k)^{a_j}, \quad \kappa_\mu\ge 1,\ a_{\kappa_\mu}>0. \tag{4.2}$$

In order to simplify notation, for the rest of this proof, we will let $\mathcal{L}_l(k)$ denote a positive constant multiplied by a product of powers (possibly of varying sign) of iterated logarithms $\log^{(j)} k$, where the smallest j is strictly larger than l. The exact form of this expression may vary from line to line. Sometimes we will need to distinguish between two such expressions in the same formula, in which case we will use the notation $\mathcal{L}_l^{(1)}(k), \mathcal{L}_l^{(2)}(k)$. Thus, we rewrite (4.2) as

$$\mu_k \sim (\log^{(\kappa_\mu)} k)^{a_{\kappa_\mu}} \mathcal{L}_{\kappa_\mu}(k), \quad \kappa_\mu\ge 1,\ a_{\kappa_\mu}>0. \tag{4.3}$$

If $\{0\le j\le J_p : b_j\ne 1\}$ is empty, then the second condition in (1.13) is fulfilled and we have

$$p_k \sim c_p\big(j\prod_{j=1}^{J_p}\log^{(j)} k\big)^{-1}. \tag{4.4}$$

Since we are assuming that (1.13) does not hold, at least one of the other two conditions in (1.13) must fail. This forces $\kappa_\mu\ne J_p$. (Recall that we are assuming that $\{0\le j\le J_\mu : a_j\ne 0\}$ is not empty and that $a_{\kappa_\mu}>0$.)

Consider first the case that $\kappa_\mu>J_p$. Then from (4.3) and (4.4) we have

$$M_n=\sum_{k=1}^n p_k\mu_k \sim \quad (\log^{(J_p+1)} n)(\log^{(\kappa_\mu)} n)^{a_{\kappa_\mu}}\mathcal{L}_{\kappa_\mu}(n),\ \text{where } \kappa_\mu\ge J_p+1. \tag{4.5}$$

From (4.3) and (4.5) it follows that (1.8) and (1.9) hold by choosing $K_n=n$. Thus, from part (i-a) of Theorem 1.1, $\lim_{n\to\infty} W_n \stackrel{\text{dist}}{=} 1$.

Now consider the case $\kappa_\mu<J_p$. Then from (4.3) and (4.4) we have

$$M_n=\sum_{k=1}^n p_k\mu_k \sim \quad (\log^{(\kappa_\mu)} n)^{a_{\kappa_\mu}}\mathcal{L}_{\kappa_\mu}(n),\ \text{where } \kappa_\mu\le J_p-1, \tag{4.6}$$

and for any K_n satisfying $K_n\to\infty$ and $K_n\le n$, we have

$$\sum_{k=K_n}^n p_k \sim c_p\big(\log^{(J_p+1)} n-\log^{(J_p+1)} K_n\big)=c_p\log\frac{\log^{(J_p)} n}{\log^{(J_p)} K_n}. \tag{4.7}$$

From (4.3) and (4.6) we have

$$\begin{aligned}
\frac{\mu_{K_n}}{M_n} &\sim \Big(\frac{\log^{(\kappa_\mu)} K_n}{\log^{(\kappa_\mu)} n}\Big)^{a_{\kappa_\mu}}\frac{\mathcal{L}^{(1)}_{\kappa_\mu}(K_n)}{\mathcal{L}^{(2)}_{\kappa_\mu}(n)},\ \kappa_\mu\le J_p-1,\ a_{\kappa_\mu}>0;\\
\frac{M_{K_n}}{M_n} &\sim \Big(\frac{\log^{(\kappa_\mu)} K_n}{\log^{(\kappa_\mu)} n}\Big)^{a_{\kappa_\mu}}\frac{\mathcal{L}^{(1)}_{\kappa_\mu}(K_n)}{\mathcal{L}^{(2)}_{\kappa_\mu}(n)},\ \kappa_\mu\le J_p-1,\ a_{\kappa_\mu}>0;
\end{aligned} \tag{4.8}$$

As we explain in some detail below, since $\kappa_\mu < J_p$, we can choose $\{K_n\}_{n=1}^\infty$ so that

$$\lim_{n\to\infty}\frac{\log^{(J_p)}K_n}{\log^{(J_p)}n}=1 \ \text{ and } \ \lim_{n\to\infty}\Big(\frac{\log^{(\kappa_\mu)}K_n}{\log^{(\kappa_\mu)}n}\Big)^{a_{\kappa_\mu}}\frac{\mathcal{L}^{(1)}_{\kappa_\mu}(K_n)}{\mathcal{L}^{(2)}_{\kappa_\mu}(n)}=0. \tag{4.9}$$

From (4.3) and (4.7)-(4.9), we conclude that $\{K_n\}$ can be defined so that (1.8) and (1.9) hold, and so that (1.10) holds with $c=0$. This proves that $\lim_{n\to\infty}W_n \stackrel{\text{dist}}{=} 0$.

To explain (4.9), note that $\frac{\mathcal{L}^{(1)}_{\kappa_\mu}(K_n)}{\mathcal{L}^{(2)}_{\kappa_\mu}(n)} \le (\log^{(\kappa_\mu+1)}n)^A$, for some $A>0$ and all large n. (Recall that the powers of the iterated logarithms in $\mathcal{L}^{(2)}_{k_\mu}$ can be negative.) Thus, in place of the second limit in (4.9), it suffices to show that $\delta_n \equiv \Big(\frac{\log^{(\kappa_\mu)}K_n}{\log^{(\kappa_\mu)}n}\Big)^{a_{\kappa_\mu}}(\log^{(\kappa_\mu+1)}n)^A \stackrel{n\to\infty}{\to} 0$. We have

$$\log^{(\kappa_\mu)}K_n=(\delta_n)^{\frac{1}{a_{\kappa_\mu}}}(\log^{(\kappa_\mu+1)}n)^{-\frac{A}{a_{\kappa_\mu}}}\log^{(\kappa_\mu)}n;$$

thus,

$$\frac{\log^{(\kappa_\mu+1)}K_n}{\log^{(\kappa_\mu+1)}n}=\frac{\log\delta_n}{a_{\kappa_\mu}\log^{(\kappa_\mu+1)}n}-\frac{A\log^{(\kappa_\mu+2)}n}{a_{\kappa_\mu}\log^{(\kappa_\mu+1)}n}+1. \tag{4.10}$$

Defining K_n by choosing $\delta_n=(\log^{(\kappa_\mu+1)}n)^{-1}$, it follows from (4.10) and the fact that $J_p\ge\kappa_\mu+1$ that the two equalities in (4.9) hold.

We now consider the case that $\{0\le j\le J_p : b_j\ne 1\}$ is not empty. Then in order to fulfill the second condition in (1.6), we have $b_{\kappa_p}<1$. We write

$$p_k\sim c_p\big(j\prod_{j=1}^{\kappa_p-1}\log^{(j)}k\big)^{-1}\big(\log^{(\kappa_p)}k\big)^{-b_{\kappa_p}}\big(\prod_{j=\kappa_p+1}^{J_p}\log^{(j)}k\big)^{-b_j}. \tag{4.11}$$

From (4.3) and (4.11) it follows that $M_n=\sum_{k=1}^n p_k\mu_k$ satisfies

$$M_n\sim\begin{cases}(\log^{(\kappa_\mu)}n)^{a_{\kappa_\mu}}\mathcal{L}_{\kappa\mu}(n), \ \kappa_\mu<\kappa_p;\\ (\log^{(\kappa_p)}n)^{a_{\kappa_p}-b_{\kappa_p}+1}\mathcal{L}_{\kappa_p}(n), \ \kappa_\mu=\kappa_p;\\ (\log^{(\kappa_p)}n)^{1-b_{\kappa_p}}\mathcal{L}_{\kappa_p}(n), \ \kappa_\mu>\kappa_p,\end{cases} \tag{4.12}$$

and from (4.11) it follows that for any K_n satisfying $K_n\to\infty$ and $K_n\le n$,

$$\sum_{k=K_n}^n p_k\sim$$
$$\frac{c_p}{1-b_{\kappa_p}}\Big[\big(\log^{(\kappa_p)}n\big)^{1-b_{\kappa_p}}\big(\prod_{j=\kappa_p+1}^{J_p}\log^{(j)}n\big)^{-b_j}-\big(\log^{(\kappa_p)}K_n\big)^{1-b_{\kappa_p}}\big(\prod_{j=\kappa_p+1}^{J_p}\log^{(j)}K_n\big)^{-b_j}\Big]. \tag{4.13}$$

From (4.3) and (4.12) we have

$$\frac{\mu_{K_n}}{M_n}\sim\begin{cases}\big(\frac{\log^{(\kappa_\mu)}K_n}{\log^{(\kappa_\mu)}n}\big)^{a_{\kappa_\mu}}\frac{\mathcal{L}^{(1)}_{\kappa_\mu}(K_n)}{\mathcal{L}^{(2)}_{\kappa_\mu}(n)}, \ \kappa_\mu<\kappa_p;\\ \frac{\big(\log^{(\kappa_p)}K_n\big)^{a_{\kappa_p}}}{(\log^{(\kappa_p)}n\big)^{a_{\kappa_p}-b_{\kappa_p}+1}}\frac{\mathcal{L}^{(1)}_{\kappa_p}(K_n)}{\mathcal{L}^{(2)}_{\kappa_p}(n)}, \ \kappa_\mu=\kappa_p;\\ \frac{\big(\log^{(\kappa_\mu)}K_n\big)^{a_{\kappa_\mu}}}{\big(\log^{(\kappa_p)}n\big)^{1-b_{\kappa_p}}}\frac{\mathcal{L}^{(1)}_{\kappa_\mu}(K_n)}{\mathcal{L}^{(2)}_{\kappa_p}(n)}, \ \kappa_\mu>\kappa_p.\end{cases} \tag{4.14}$$

It is immediate (4.3) and (4.14) that if $\kappa_\mu\ge\kappa_p$, then (1.8) and (1.9) hold by choosing $K_n=n$. (For the case $\kappa_\mu=\kappa_p$, recall that $b_{\kappa_p}\in(0,1)$.) Thus, from part (i-a) of Theorem 1.1, $\lim_{n\to\infty}W_n\stackrel{\text{dist}}{=}1$.

Now consider the case $\kappa_\mu < \kappa_p$. For simplicity, we will assume that the higher order iterated logarithmic terms do not appear; that is, we will assume from (4.12)-(4.14) that

$$\begin{aligned}
&\sum_{k=K_n}^{n} p_k \sim \frac{c_p}{1-b_{\kappa_p}}\Big[\big(\log^{(\kappa_p)} n\big)^{1-b_{\kappa_p}} - \big(\log^{(\kappa_p)} K_n\big)^{1-b_{\kappa_p}}\Big];\\
&\frac{\mu_{K_n}}{M_n} \sim \big(\frac{\log^{(\kappa_\mu)} K_n}{\log^{(\kappa_\mu)} n}\big)^{a_{\kappa_\mu}};\\
&\frac{M_{K_n}}{M_n} \sim \big(\frac{\log^{(\kappa_\mu)} K_n}{\log^{(\kappa_\mu)} n}\big)^{a_{\kappa_\mu}}.
\end{aligned} \tag{4.15}$$

The additional logarithmic terms can be dealt with similarly to the way they were dealt with for (4.9), as explained in the paragraph following (4.9). Applying the mean value theorem to the function $x^{1-b_{\kappa_p}}$, we obtain

$$\big(\log^{(\kappa_p)} n\big)^{1-b_{\kappa_p}} - \big(\log^{(\kappa_p)} K_n\big)^{1-b_{\kappa_p}} = \frac{(1-b_{\kappa_p})\log^{(\kappa_p)} \frac{n}{K_n}}{(\log^{(\kappa_p)} n^*)^{b_{\kappa_p}}}, \tag{4.16}$$

where $n^* \in (K_n, n)$. Since $\kappa_\mu < \kappa_p$, we can choose $K_n \to \infty$ such that $\lim_{n\to\infty} \frac{\log^{(\kappa_\mu)} K_n}{\log^{(\kappa_\mu)} n} = 0$, but $\lim_{n\to\infty} \log^{(\kappa_p)} \frac{K_n}{n} = 1$. For such a choice of $\{K_n\}$, it follows from (4.3), (4.15) and (4.16) that (1.8) and (1.9) hold, and that (1.10) holds with $c = 0$; thus, $\lim_{n\to\infty} W_n \stackrel{\text{dist}}{=} 0$. $\square$

5 Basic facts concerning generalized Dickman distributions

The proof of Theorem 1.1 showed in particular that $\exp(\theta\int_0^1 \frac{e^{-\lambda x}-1}{x}dx)$ is the Laplace transform of a probability distribution, which we have denoted by GD(θ).

Proposition 5.1. *Let D_θ ∼GD(θ). Then*

$$D_\theta \stackrel{\text{dist}}{=} U^{\frac{1}{\theta}}(D_\theta + 1), \tag{5.1}$$

where U is distributed according to the uniform distribution on $[0,1]$, and U and D_θ on the right hand side above are independent.

Remark 1. From (5.1) it is immediate that

$$D_\theta \stackrel{\text{dist}}{=} U_1^{\frac{1}{\theta}} + (U_1U_2)^{\frac{1}{\theta}} + (U_1U_2U_3)^{\frac{1}{\theta}} + \cdots,$$

where $\{U_n\}_{n=1}^\infty$ are IID random variables distributed according to the uniform distribution on $[0,1]$.

Remark 2. Our proof of the proposition is rather probabilistic; a more analytic proof can be found in [7].

Proof. The proof of Theorem 1.1 showed in particular that if we let $X_k = \mu_k = k$ and $p_k = \frac{\theta}{k}$, in which case $M_n = \sum_{k=1}^n p_k\mu_k = \theta n$, then

$$\hat{W}_n \equiv \theta W_n = \frac{1}{n}\sum_{k=1}^n kB_k \stackrel{\text{dist}}{\to} D_\theta, \tag{5.2}$$

where $D_\theta \stackrel{\text{dist}}{\sim} GD(\theta)$. Let

$$J_n^+ = \max\{k \le n : B_k \ne 0\},$$

with $\max\emptyset \equiv 0$. We write

$$\hat{W}_n \equiv \frac{1}{n}\sum_{n=1}^{n} kB_k = \frac{J_n^+ - 1}{n}\Big(\frac{1}{J_n^+ - 1}\sum_{k=1}^{J_n^+-1} kB_k\Big) + \frac{J_n^+}{n}, \tag{5.3}$$

where the first of the two summands on the right hand side above is interpreted as equal to 0 if $J_n^+ \le 1$. We have

$$P(\frac{J_n^+}{n} \le x) = \prod_{k=[xn+1]}^{n}(1-\frac{\theta}{k}) \sim x^\theta,\ x \in (0,1). \tag{5.4}$$

Also, by the independence of $\{B_k\}_{k=1}^\infty$, we have

$$\frac{1}{J_n^+ - 1}\sum_{k=1}^{J_n^+-1} kB_k \mid \{J_n^+ = k_0\} \stackrel{\text{dist}}{=} \frac{1}{k_0 - 1}\sum_{k=1}^{k_0-1} kB_k = \hat{W}_{k_0-1},\ k_0 \ge 2. \tag{5.5}$$

Letting $n \to \infty$ in (5.3) and using (5.2), (5.4) and (5.5), we conclude that (5.1) holds, where U is distributed according to the uniform distribution on $[0,1]$, $D_\theta \stackrel{\text{dist}}{\sim}$ GD(θ) and U and D_θ on the right hand side are independent. □

Proposition 5.2. *The GD(θ) distribution has a density function p_θ satisfying $p_\theta = c_\theta\rho_\theta$, for some $c_\theta > 0$, where ρ_θ satisfies* (1.1).

Remark. For a derivation of the formula $c_\theta = \frac{e^{-\theta\gamma}}{\Gamma(\theta)}$, see [1].

Proof. Let $F_\theta(x) = P(D_\theta \le x)$ denote the distribution function for the GD(θ) distribution. Then from (5.1) we have

$$\begin{aligned} &F_\theta(x) = P(D_\theta \le x) = P(U^{\frac{1}{\theta}}(D_\theta+1) \le x) = \int_0^1 P(D_\theta + 1 \le xy^{-\frac{1}{\theta}})dy = \\ &\int_0^1 F_\theta(xy^{-\frac{1}{\theta}} - 1)dy. \end{aligned} \tag{5.6}$$

For $x > 0$, making the change of variables, $v = xy^{-\frac{1}{\theta}} - 1$, we can rewrite (5.6) as

$$F_\theta(x) = \theta x^\theta \int_{x-1}^\infty F_{D_\theta}(v)(1+v)^{-1-\theta}dv,\ x > 0. \tag{5.7}$$

From (5.7) and the fact that $F_\theta(x) = 0$, for $x \le 0$, it follows that F_θ is continuous on $\mathbb{R}$. Also, since $F_\theta(x) = 0$, for $x \le 0$, we have

$$\int_{x-1}^\infty F_{D_\theta}(v)(1+v)^{-1-\theta}dv = \int_0^\infty F_{D_\theta}(v)(1+v)^{-1-\theta}dv,\ x \le 1.$$

Consequently, it follows from (5.7) that $F_\theta(x) = C_\theta x^\theta$, for $x \in [0,1]$, where $C_\theta = \theta\int_0^\infty F_{D_\theta}(v)(1+v)^{-1-\theta}dv$. From this and (5.7) it follows that F is differentiable on $(0,1)$ and on $(1,\infty)$, and that, letting $p_\theta = F'_\theta$,

$$p_\theta = c_\theta x^{\theta-1},\ 0 < x < 1,\ c_\theta = \theta^2\int_0^\infty F_{D_\theta}(v)(1+v)^{-1-\theta}dv, \tag{5.8}$$

and

$$\begin{aligned} &p_\theta(x) = \theta^2 x^{\theta-1}\int_{x-1}^\infty F_{D_\theta}(v)(1+v)^{-1-\theta}dv - \theta x^{-1}F_\theta(x-1) = \\ &\frac{\theta}{x}(F_\theta(x) - F_\theta(x-1)),\ x > 1. \end{aligned} \tag{5.9}$$

From (5.9), it follows that p_θ is differentiable on $x>1$, and that $(xp_\theta(x))' = \theta\big(p_\theta(x) - p_\theta(x-1)\big)$, for $x>1$, or equivalently,

$$xp'_\theta(x) + (1-\theta)p_\theta(x) + \theta p_\theta(x-1) = 0,\ x>1. \tag{5.10}$$

From (5.8) and (5.10) we conclude that $p_\theta(x) = c_\theta \rho_\theta$, where ρ_θ satisfies (1.1). Integrating by parts in the formula for c_θ in (5.8) shows that

$$c_\theta = \theta \int_0^\infty (1+v)^{-\theta} p_\theta(v)dv = \theta E(1+D_\theta)^{-\theta}. \qquad \square$$

6 The Dickman function in number theory and probability

The Dickman function $\rho \equiv \rho_1$ arises in probabilistic number theory in the context of so-called *smooth* numbers; that is, numbers all of whose prime divisors are "small." Let $\Psi(x,y)$ denote the number of positive integers less than or equal to x with no prime divisors greater than y. Numbers with no prime divisors greater than y are called y-*smooth* numbers. Then for $s \geq 1$, $\Psi(N, N^{\frac{1}{s}}) \sim N\rho(s)$, as $N \to \infty$. This result was first proved by Dickman in 1930 [4], whence the name of the function, with later refinements by de Bruijn [2]. See also [6] or [9]. Let $[n] = \{1,\ldots,n\}$ and let $p^+(n)$ denote the largest prime divisor of n. Then Dickman's result states that the random variable $\frac{\log p^+(j)}{\log n}, j \in [n]$, on the probability space $[n]$ with the uniform distribution converges in distribution as $n \to \infty$ to the distribution whose distribution function is $\rho(\frac{1}{x})$, $x \in [0,1]$, and whose density is $-\frac{\rho'(\frac{1}{x})}{x^2} = \frac{\rho(\frac{1}{x}-1)}{x}$, $x \in [0,1]$. It is easy to see that an equivalent statement of Dickman's result is that the random variable $\frac{\log p^+(j)}{\log j}$, $j \in [n]$, on the probability space $[n]$ with the uniform distribution converges in distribution as $n \to \infty$ to the distribution whose distribution function is $\rho(\frac{1}{x})$, $x \in [0,1]$, We note that the length of the longest cycle of a uniformly random permutation of $[n]$, normalized by dividing by n, also converges to a limiting distribution whose distribution function is $\rho(\frac{1}{x})$. If instead of using the uniform measure on S_n, the set of permutations of $[n]$, one uses the Ewens sampling distribution on S_n, obtained by giving each permutation $\sigma \in S_n$ the probability proportional to $\theta^{C(\sigma)}$, where $C(\sigma)$ denotes the number of cycles in σ, then the length of the longest cycle of such a random permutation of $[n]$, normalized by dividing by n, converges to a limiting distribution whose distribution function is $\rho_\theta(\frac{1}{x})$, $x \in [0,1]$. This distribution is also the distribution of the first coordinate of the Poisson-Dirichlet distribution $\mathrm{PD}(\theta)$ (see [1]).

The examples in the above paragraph lead to limiting distributions where the Dickman function arises as a distribution function, not as a density as is the case with the $\mathrm{GD}(\theta)$ distributions discussed in this paper. The $\mathrm{GD}(\theta)$ distribution arises as a normalized limit in the context of certain natural probability measures that one can place on $\mathbb{N}$; see [3], [8].

References

[1] Arratia, R., Barbour, A. and Tavaré, S., *Logarithmic Combinatorial Structures: A Probabilistic Approach*, EMS Monographs in Mathematics, European Mathematical Society (EMS), Zurich, (2003). MR-2032426

[2] de Bruijn, N., *On the number of positive integers $\leq x$ and free of prime factors $> y$*, Nederl. Acad. Wetensch. Proc. Ser. A. 54, (1951), 50–60. MR-0205945

[3] Cellarosi, F. and Sinai, Y., *Non-standard limit theorems in number theory*, Prokhorov and Contemporary Probability Theory, 197–213, Springer Proc. Math. Stat., 33, Springer, Heidelberg, (2013). MR-3070473

[4] Dickman, K., *On the frequency of numbers containing prime factors of a certain relative magnitude*, Ark. Math. Astr. Fys. 22, 1–14 (1930).

[5] Hwang, H.-K. and Tsai, T.-H. *Quickselect and the Dickman function*, Combin. Probab. Comput. 11 (2002), 353–371. MR-1918722

[6] Montgomery, H. and Vaughan, R., *Multiplicative Number Theory. I. Classical Theory*, Cambridge Studies in Advanced Mathematics, 97, Cambridge University Press, Cambridge, (2007). MR-2378655

[7] Penrose, M. and Wade, A., *Random minimal directed spanning trees and Dickman-type distributions*, Adv. in Appl. Probab. 36 (2004), 691–714. MR-2079909

[8] Pinsky R., *A Natural Probabilistic Model on the Integers and its Relation to Dickman-Type Distributions and Buchstab's Function*, preprint, (2016). arXiv:1606.02965

[9] Tenenbaum, G., *Introduction to Analytic and Probabilistic Number Theory*, Cambridge Studies in Advanced Mathematics, 46, Cambridge University Press, Cambridge, (1995). MR1342300 MR-1342300