THE DICKMAN FUNCTION IN PROBABILITY

Seth Hardy

January 2023

1 The Dickman function

The Dickman function, ρ, occurs in number theory in the study of 'smooth numbers', that is, numbers whose largest prime factor is below some threshold. Specifically, we find that for $u \in [0, \infty)$,

$$\lim_{x\to\infty} \frac{1}{x} \#\{1 \leq n \leq x : \, P(n) \leq x^{1/u}\} = \rho(u), \tag{1}$$

where $P(n)$ denotes the largest prime factor of n. We may may view the Dickman function ρ as being defined by such a formula, if we assume that this limit exits. We may then try and study the properties of ρ. Number Theoretic heuristics starting from (1) are rather tenuous; the following is from Koukoulopoulos [Kou19],

$$\begin{aligned}\log x \sum_{\substack{n\leq x\\ P(n)\leq x^{1/u}}} 1 \approx \sum_{\substack{n\leq x\\ P(n)\leq x^{1/u}}} \log n \approx \sum_{p\leq x^{1/u}} \log p \sum_{\substack{n\leq x/p\\ P(n)\leq x^{1/u}}} 1 \\ \approx \int_1^{x^{1/u}} \sum_{\substack{n\leq x/t\\ P(n)\leq x^{1/u}}} 1 \, dt = \frac{1}{u} \log x \int_{u-1}^{u} \frac{1}{x^{v/u}} \sum_{\substack{n\leq x^{v/u}\\ P(n)\leq x^{1/u}}} 1 \, dv\,,\end{aligned}$$

leading to the conjecture that $u\rho(u) = \int_{u-1}^{u} \rho(x)\, dx$. Furthermore, by (1), it is clear that we should have $\rho(u) = 1$ for $u \in [0, 1]$. Due to uniqueness of solutions to initial value problems, these properties uniquely define a continuous function.

Definition 1.1 (Dickman function)**.** *The Dickman function (referred to as ρ throughout) is the unique continuous function satisfying the differential equation $u\rho'(u) + \rho(u-1) = 0$ for $u \geq 1$, with initial conditions $\rho(u) = 1$ for $u \in [0, 1]$.*

An undesirable amount of technical manipulations are required for the above heuristic argument (indeed, a 'nice' heuristic doesn't seem to exist). We may view (1) as probabilistic statement; if you pick a random integer uniformly between 1 and x, the probability that such an integer will have all prime factors below $x^{1/u}$ approaches $\rho(u)$ as x goes to infinity. Subsequently, we hope that a more probabilistic view can motivate why the Dickman function in Definition 1.1 should also satisfy (1). We will not prove the relation (1) (such a proof may be found, for example, in [Ram49]), but hope to inspire a deeper understanding of when the Dickman function ρ occurs.

Much injustice is done to the Dickman function if it is viewed purely as a number theoretic function. We present the following 3 problems, each of which share the same answer, $\rho(u)$:

1. Pick a random integer, n, uniformly at random between 1 and N. What is the probability that all prime factors of n are below $N^{1/u}$ (in the limit, as $N \to \infty$)?
2. We begin with a stick of length 1, and countably many people. Person 1 chooses a number U_1 uniformly from $[0, 1]$, determining where they snap the stick. They keep the stick of length U_1, and pass to Person 2 the remaining stick of length $1 - U_1$. Person 2 snaps this stick by choosing U_2 uniformly from $[0, 1]$, independent of U_1, and keeps a U_2 proportion of the stick (so their stick will be of length $U_2(1 - U_1)$). They then pass the remaining stick to Person 3, and this process continues, so that person n has a stick of length $U_n(1 - U_{n-1})...(1 - U_1)$. What is the probability that each person has a stick of length shorter than $1/u$?
3. (Generalisation of the 100 prisoner problem) There are n prisoners each given a distinct number from 1 to n. The director of the prison offers them a challenge for their freedom. In a room of n boxes, the director randomly puts each prisoners number in a different box. One after the other, each prisoner is allowed to open $\lfloor n/u \rfloor$ boxes in the room. The prisoners succeed if (and only if) **all** prisoners find find their number. Assuming the prisoners use the best possible strategy, what is their chance at success as n goes to infinity?

This essay hopes to shed light on the connection between these problems. Those who have seen variants of problem 3 before will know that the optimal solution is connected to the longest cycle of a random permutation [War+06]: in fact, the question is just asking 'what is the probability that a uniformly chosen random permutation from S_n has all cycles of length $\leq n/u$'. Therefore, comparing problems 1 and 3 above, one can ask if there is an analogy between integers and permutations; prime numbers and cycles. Indeed, both integers and permutations have unique factorisation up to reordering, and the parallels one can draw from this are quite fascinating. So much so, in fact, they were even subject of a graphic novel called 'Prime Suspects' by A. Granville and J. Granville [GG19].

Regarding the Dickman function, there are two interesting distributions that may be derived from it. The first of which, sometimes called the max-Dickman distribution (for example, in [MP20]), is derived by noticing that ρ is a decreasing continuous function on $[0, \infty)$ with $\rho(0) = 1$ and $\rho(u) \to 0$ as $u \to \infty$.

Definition 1.2 (max-Dickman Distribution)**.** *The max-Dickman Distribution is defined by the following cumulative distribution function*

$$F(u) = \begin{cases} 0 & \text{if } u \leq 0 \\ \rho(1/u) & \text{if } u > 0 \end{cases},$$

Noting that $F(u) = 1$ for $u \geq 1$.

Studying this distribution will unveil the link between the above 3 problems, but it will not be the only focus of this essay. There is another, equally interesting distribution relating to the Dickman function. This distribution appears in insurance mathematics (perpetuities) [Dev01], algorithms [HT02], and many other fields (see [Dev01] for more references). First, we note that

$$\int_0^\infty \rho(u)\, du = e^\gamma,$$

where $\gamma = 0.5772...$ is the Euler-Mascheroni[1] constant. Therefore, we can normalise appropriately to obtain the following cumulative distribution function:

Definition 1.3 (Dickman Distribution)**.** *A random variable X is said to have Dickman distribution or Dickman density if X has cumulative distribution function*

$$D(x) = \begin{cases} 0 & \text{if } x \leq 0\,, \\ \int_0^x e^{-\gamma}\rho(u)\, du & \text{if } x > 0\,. \end{cases} \tag{2}$$

We will return to this distribution in Section 3.

2 The max-Dickman Distribution and Partitions of Unity

Suppose that we know that the largest cycle length of a permutation σ, denoted $L(\sigma)$, indeed satisfies $\frac{1}{n!}\#\{\sigma \in S_n : L(\sigma) \leq n/u\} \to \rho(u)$ (that is, if you uniformly pick a random permutation from S_n, the probability that all cycles in its cyclic decomposition are of length $\leq n/u$ approaches $\rho(u)$). Furthermore, suppose that we know (1) is true. In this section, we search for a deeper understanding as to why the Dickman function arises in these situations. The following is motivated by a blog-post by Tao [Tao], and a fantastic survey paper on the Dickman Distribution by Molchanov and Panov [MP20].

First, one should notice that if p is a prime factor of m, the condition that $p \leq m^{1/u}$, [2] is equivalent to $\log p / \log m \leq 1/u$. Along a similar line, if $\sigma = C_1...C_r$ is the cyclic decomposition of σ, then the condition $|C_i| \leq n/u$ (where $|C_i|$ denotes the length of cycle i) can instead be written as $|C_i|/n \leq 1/u$. The probabilities of these equations being satisfied for all prime factors of a random integer / all cycles of a random permutation, converge to the Dickman function, and so we want to dig deeper into the structures of $\log p/\log m$ and $|C_i|/n$ for randomly chosen integers and permutations.

To begin, we fix $m \in \mathbb{N}$, and write its prime factorisation $m = p_1 p_2 ... p_r$, where these primes are possibly not distinct. We then have

$$\frac{\log p_1}{\log m} + \frac{\log p_2}{\log m} + ... + \frac{\log p_r}{\log m} = 1.$$

Similarly, for fixed $\sigma \in S_n$ with cyclic decomposition $\sigma = C_1...C_r$, we have

$$\frac{|C_1|}{n} + \frac{|C_2|}{n} + ... + \frac{|C_r|}{n} = 1.$$

[1]This constant is another quantity commonly found in number theory, $\gamma = \lim_{n\to\infty}(\sum_{k\leq n} \frac{1}{k} - \log n)$.

[2]Note that, in (1), if we instead consider $m \leq x$ such that $P(n) \leq m^{1/u}$, then the limit is unchanged.

Therefore, the multisets (sets that may have elements repeated)

$$\Big\{\frac{\log p_1}{\log m}, \frac{\log p_2}{\log m}, ..., \frac{\log p_r}{\log m}\Big\} \text{ and } \Big\{\frac{|C_1|}{n}, \frac{|C_2|}{n}, ..., \frac{|C_r|}{n}\Big\}, \tag{3}$$

both form **partitions of unity**, that is, they are countable multisets of non-negative elements that sum to 1. This interpretation allows us to directly compare the distributions of primes and cycle-lengths via (3), as we know that for randomly chosen $m \in \{1, 2, ..., n\}$,

$$\mathbb{P}\Big(\Big\{\frac{\log p_1}{\log m} \le \frac{1}{u}, \frac{\log p_2}{\log m} \le \frac{1}{u}, ..., \frac{\log p_r}{\log m} \le \frac{1}{u}\Big\}\Big) \to \rho(u) \text{ as } n \to \infty,$$

and for randomly chosen $\sigma \in S_n$,

$$\mathbb{P}\Big(\Big\{\frac{|C_1|}{n} \le \frac{1}{u}, \frac{|C_2|}{n} \le \frac{1}{u}, ..., \frac{|C_r|}{n} \le \frac{1}{u}\Big\}\Big) \to \rho(u) \text{ as } n \to \infty.$$

Therefore the largest elements of both objects in (3) have the same distribution. Here, it is nice to mention the fact (originally proved by Ramanujan [AB18], Sec 8.2) that

$$\rho(u) = 1 + \sum_{k \ge 1} \frac{(-1)^k}{k!} \int \cdots \int_{I_k(u)} \frac{1}{y_1 \ldots y_k} dy_1 \ldots dy_k,$$
$$I_k(u) = \{y_1 > 1/u, \ldots, y_k > 1/u,\, y_1 + \cdots + y_k < 1\},$$

suggesting that the Dickman function is inherently linked to the size of elements of partitions of unity. With this in mind, we want to create the most natural random partition of unity. It seems sensible to try and use uniformly distributed random variables for such a construction.

Let $\mathcal{U}_1, \mathcal{U}_2, ...$ be independent random variables with uniform distribution $U([0, 1])$. We take $\mathcal{U}_1$ to be the first element of our partition of unity. Now, the remaining elements should sum to $1 - \mathcal{U}_1$, and so the next element should be no bigger than this. Subsequently, we take the second element to be $\mathcal{U}_2(1 - \mathcal{U}_1)$. Now, the remaining elements should sum to $1 - \mathcal{U}_1 - \mathcal{U}_2(1 - \mathcal{U}_1) = (1 - \mathcal{U}_1)(1 - \mathcal{U}_2)$. Therefore, we take the third element to be $\mathcal{U}_3(1 - \mathcal{U}_1)(1 - \mathcal{U}_2)$, and so on, with the n-th element being $\mathcal{U}_n(1 - \mathcal{U}_{n-1})...(1 - \mathcal{U}_1)$. Note that this is identical to the list of stick lengths in the stick breaking problem mentioned in problem 2 of Section 1. This process gives us a natural partition of unity

$$\{\mathcal{U}_1, \mathcal{U}_2(1 - \mathcal{U}_1), \ldots, \mathcal{U}_k(1 - \mathcal{U}_1) \ldots (1 - \mathcal{U}_{k-1}), \ldots\}. \tag{4}$$

We would then hope, according to our discussion, that the probability that the largest element of this partition is $\le 1/u$ also converges to the Dickman function $\rho(u)$. This would suggest that the largest prime factor of a randomly selected integer and the largest cycle length of a random permutation behave in a very natural way, analogously to the largest element of (4). Let

$$\mathcal{M} = \sup\{\mathcal{U}_1, \mathcal{U}_2(1 - \mathcal{U}_1), \ldots, \mathcal{U}_k(1 - \mathcal{U}_1) \ldots (1 - \mathcal{U}_{k-1}), \ldots\}.$$

Note that, by self-similarity, if $\mathcal{U} \sim U([0, 1])$ is independent of $\mathcal{M}$, then we can equivalently define $\mathcal{M}$ as being the random variable satisfying

$$\mathcal{M} \stackrel{\text{Law}}{=} \max\{\mathcal{U}, (1 - \mathcal{U})\mathcal{M}\}. \tag{5}$$

Now, for $u \geq 1$,

$$\begin{aligned}\phi(u) := \mathbb{P}(\mathcal{M} \leq 1/u) &= \mathbb{P}(\max\{\mathcal{U}, (1-\mathcal{U})\mathcal{M}\} \leq 1/u) \\ &= \int_0^1 \mathbb{P}(\max\{t, (1-t)\mathcal{M}\} \leq 1/u)\, dt \\ &= \int_0^{1/u} \mathbb{P}(\mathcal{M} \leq 1/u(1-t))\, dt \\ &= \int_0^{1/u} \phi(u(1-t))\, dt = \frac{1}{u}\int_{u-1}^{u} \phi(y)\, dy,\end{aligned}$$

where we have used independence on the second line. Differentiating the first and last terms, we find that $u\phi'(u) + \phi(u-1) = 0$. Furthermore, if $u \in (0, 1]$, the fact that $\mathcal{M} \leq 1$ implies that $\phi(u) = 1$. This is precisely the definition of the Dickman function given in Definition 1.1, and so $\phi = \rho$, and $\mathcal{M}$ has max-Dickman distribution according to Definition 1.2, as suspected.

We have found that the largest element of the partition of unity in (4) has max-Dickman distribution. This distribution is in fact a specific case of a Poisson-Dirichlet process/distribution, which takes the shape of a maximum over infinitely many beta-distributed random variables. Donnelly and Grimmett [DG93] showed directly that the distribution of prime factors of an integer (note that this is not just the largest prime factor) should converge in distribution to a Poisson Dirichlet distribution.

We finish the section by giving some (not entirely rigorous) motivation as to why the distribution of the largest cycle of a random permutation should have the same distribution as $\mathcal{M}$. A full explaination can be found in [MP20]. Let us first consider the value distribution of $|C_1|$ for a uniformly chosen random permutation $\sigma = C_1 ... C_r$ from S_n, where $C_1, ..., C_r$ are ordered so that C_1 includes the number 1, C_2 includes the smallest number not in C_1 and so on. If $|C_1| = 1$, then 1 is fixed by S_n. We must have $(n-1)!$ of these permutations, since we are just counting the number of permutations of $\{2, 3, ..., n\}$. Therefore, $\mathbb{P}(|C_1| = 1) = 1/n$. Now if $|C_1| = 2$, then we have $n-1$ possible elements that 1 is mapped to. Furthermore, we count all permutations of the remaining $n-2$ numbers. This gives $\mathbb{P}(|C_1| = 2) = (n-1)(n-2)!/n! = 1/n$. Continuing, we find that in general $\mathbb{P}(|C_1| = r) = 1/n$ for $r \in \{1, 2, ..., n\}$.

We now consider the other cycles. Suppose that we know $|C_1| = m < n$. Then for $r \leq n - m$, we can use identical calculations to above to find that $\mathbb{P}(|C_2| = r \,\big|\, |C_1| = m) = \frac{1}{n-m}$, as C_2 can be viewed as being a cycle from a random permutation of $n - m$ objects. This extends more generally to

$$\mathbb{P}\Big(|C_k| = r \,\Big|\, |C_1| = m_1, |C_2| = m_2, ..., |C_{k-1}| = m_{k-1}\Big) = \frac{1}{n - m_1 - m_2 - ... - m_{k-1}},$$

for any $m_1 + ... + m_{k-1} < n$. Now, for $a \in \{1, 2, ..., n\}$,

$$\mathbb{P}\Big(\frac{|C_1|}{n} = \frac{a}{n}\Big) = \frac{1}{n} \quad \Rightarrow \quad \frac{|C_1|}{n} \xrightarrow{d} \mathcal{U}_1,$$

where $\mathcal{U}_1 \sim U([0,1])$. Then for $b \in \{1, 2, ..., n-a\}$, we have

$$\mathbb{P}\Big(\frac{|C_2|}{n} = \frac{b}{n} \,\Big|\, \frac{|C_1|}{n} = \frac{a}{n}\Big) = \frac{1}{n-a} \quad \Rightarrow \quad \frac{|C_2|}{n} \xrightarrow{d} \mathcal{U}_2(1 - \mathcal{U}_1),$$

where $\mathcal{U}_2 \sim U([0,1])$ is independent of $\mathcal{U}_1$. The intuition here is that, once we know $|C_1|/n$, the distribution of $|C_2|/n$ is uniform, and $|C_2|/n$ cannot be any larger than $1 - |C_1|/n$. This continues and we find that

$$\left\{\frac{|C_1|}{n}, \frac{|C_2|}{n}, ..., \frac{|C_n|}{n}\right\} \xrightarrow{d} \{\mathcal{U}_1, \mathcal{U}_2(1-\mathcal{U}_1), \ldots, \mathcal{U}_k(1-\mathcal{U}_1)\ldots(1-\mathcal{U}_{k-1}), \ldots\}$$

Therefore, the largest cycle will have distribution according the the max-Dickman distribution, as we hoped!

Remark 2.1. *Interestingly, the max-Dickman distribution also makes an appearance when looking at the maximal edge length of random graphs in* $(0,1]^2$ *(Section 3.1 of [MP20]). Hopefully, this is no longer very surprising!*

3 The max-Dickman Distribution and the Dickman Distribution

In this section, we motivate why one may also be interested in the Dickman distribution (Definition 1.3). We will expose a link between the max-Dickman distribution and the Dickman distribution that arises when studying the largest cycle of permutations using an alternate method to those in the previous section. This will be taken from Arratia, Barbour and Tavaré [ABT03], mostly without proof. In this method, we instead look at cycle counts. For a uniformly randomly chosen permutation $\sigma \in S_n$, we let $C_i^{(n)}$ denote the random variable equal to the number of cycles of length i of σ. We must then have $C_1^{(n)} + 2C_2^{(n)} + \cdots + nC_n^{(n)} = n$. It is shown in Theorem 1.3 of [ABT03] that

$$(C_1^{(n)}, C_2^{(n)}, ..., C_n^{(n)}) \xrightarrow{d} (Z_1, ..., Z_n), \tag{6}$$

where $(Z_j)_{j\geq 1}$ are independent Poisson random variables with $Z_j \sim \text{Pois}(1/j)$. This convergence is in the sense that each $C_j^{(n)} \xrightarrow{d} Z_j$ as $n \to \infty$, and each component behaves independently in the limit. The proof is quite long due to some technical combinatorics.

Once (6) is established, we may analyse the distribution of cycle lengths using Poisson random variables. One thing we have to do here is condition on the fact that $C_1^{(n)} + 2C_2^{(n)} + \cdots + nC_2^{(n)} = n$. Therefore, with $T_j = Z_1 + 2Z_2 + ... + jZ_j$, it can be shown (Lemma 5.3 of [ABT03]) that

$$\begin{aligned}
\mathbb{P}(L(\sigma) \leq n/u) &= \mathbb{P}(C_{\lceil n/u\rceil+1}^{(n)} = 0, ..., C_n^{(n)} = 0) \\
&= \mathbb{P}(Z_{\lceil n/u\rceil} = 0, ..., Z_n = 0 |\, T_n = n) \\
&= \Big(\prod_{i=\lceil n/u\rceil}^{n} \mathbb{P}(Z_i = 0)\Big)\frac{\mathbb{P}(T_{\lceil n/u\rceil} = n)}{\mathbb{P}(T_n = n)} \\
&= \exp\Big(\sum_{i=\lceil n/u\rceil}^{n} \frac{1}{i}\Big)\frac{\mathbb{P}(T_{\lceil n/u\rceil} = n)}{\mathbb{P}(T_n = n)}.
\end{aligned} \tag{7}$$

We have seen in the previous section that the left hand side converges to $\rho(1/u)$. But it happens that T_n/n actually converges to the Dickman Distribution according to Definition 1.3! This is fairly straightforward to show. The Laplace Transform of T_n/n is

$$\begin{aligned}\mathbb{E}[e^{-sT_n/n}] &= \prod_{i=1}^{n}\mathbb{E}[e^{-sZ_i/n}] = \prod_{i=1}^{n}\sum_{m=0}^{\infty}\frac{(1/i)^m e^{-1/i}}{m!}e^{-sm/n} \\ &= \prod_{i=1}^{n}e^{-1/i}\sum_{m=0}^{\infty}\Big(\frac{e^{-s/n}}{i}\Big)^m\frac{1}{m!} = \exp\Big(\sum_{i=1}^{n}\frac{e^{-si/n}-1}{i}\Big).\end{aligned}$$

As noted in Theorem 4.8 of [ABT03], this sum inside the exponential can be written as

$$\int_0^1 \frac{e^{-sx}-1}{x}d\mu_n(x),$$

where μ_n is the probability measure that puts mass $1/n$ at i/n for $i \in \{1,2,3,...,n\}$. This measure converges weakly to Leb on $[0,1]$, hence we find that as $n \to \infty$, we have

$$\mathbb{E}[e^{-sT_n/n}] \to \exp\Big(\int_0^1 \frac{e^{-sx}-1}{x}\Big).$$

We later see (in (10), when $\theta = 1$ there) that this is the Laplace transform of the Dickman Distribution. When it exists, the Laplace transform (which is just the moment generating function) uniquely defines a distribution in the same way that the characteristic function does, and so we have found that T_n/n does indeed converge in distribution to the Dickman distribution.

Using this, it can be shown (using a technique called size biasing, Theorem 4.13 of [ABT03]) that the right hand side of (7) is related to $u\mathbb{P}(1/u-1 \le T_n/n \le 1/u) \to ue^{-\gamma}\int_{1/u-1}^{1/u}\rho(x)\,dx = e^{-\gamma}\rho(1/u)$. It should not be a surprise that the max-Dickman distribution and the Dickman distribution are linked in such a way. The first of these has cumulative distribution relating to the Dickman function, the latter has density relating the the Dickman function. However, the Dickman function is, by definition, inherently tied to its derivative. Therefore one would expect that random variables with max-Dickman distribution and Dickman density should certainly be related, since density is just the derivative of the cumulative distribution.

4 Infinite Divisibility of Dickman Distribution

In this section we show that the Dickman distribution (Definition 1.3) is infinitely divisible, whilst also proving some nice relations to other interesting distributions. As in Section 2, we follow the work of Molchanov and Panov [MP20], utilising the Laplace transform to compare distributions. First we will show that if $\mathcal{D}$ is the random variable defined by

$$\mathcal{D} = \mathcal{U}_1 + \mathcal{U}_1\mathcal{U}_2 + \mathcal{U}_1\mathcal{U}_2\mathcal{U}_3 + \cdot. \tag{8}$$

For $(\mathcal{U}_i)_{i\ge1}$ independent with $\mathcal{U}_i \sim U([0,1])$, then $\mathcal{D}$ has Dickman distribution. Similarly to (5), we certainly have

$$\mathcal{D} \overset{\text{Law}}{=} \mathcal{U}(1+\mathcal{D}),$$

where $\mathcal{U} \sim U([0,1])$ is independent of $\mathcal{D}$. A variation on (8) will then give a very neat proof that the Dickman distribution is infinitely divisible.

Proposition 4.1. *$\mathcal{D}$ as defined above has Dickman distribution according to Definition 1.3, assuming $\mathcal{D}$ has density $f_{\mathcal{D}}$ that is differentiable.*

Proof. We know that $\mathcal{D} \overset{\text{Law}}{=} \mathcal{U}(1+\mathcal{D})$, so that for $x > 0$ we have cumulative distribution function

$$\begin{aligned} F_{\mathcal{D}}(x) = \mathbb{P}(\mathcal{D} \le x) = \mathcal{P}(\mathcal{U}(1+\mathcal{D}) \le x) &= \mathbb{P}(\mathcal{D} \le x/\mathcal{U} - 1) \\ &= \int_0^1 \mathbb{P}(\mathcal{D} \le x/t - 1)\, dt = x \int_0^{1/x} F_{\mathcal{D}}(1/y - 1)\, dy. \end{aligned}$$

Differentiation both sides with respect to x, we find that

$$f_{\mathcal{D}}(x) = \int_0^{1/x} F_{\mathcal{D}}(1/y-1)\, dy \; - \frac{1}{x} F_{\mathcal{D}}(x-1),$$

and so, using the first result to say that this integral is equal to $F_{\mathcal{D}}(x)/x$, we have

$$x f_{\mathcal{D}}(x) = F_{\mathcal{D}}(x) - F_{\mathcal{D}}(x-1).$$

Finally, differentiating again with respect to x, we find that

$$x f'_{\mathcal{D}}(x) + f_{\mathcal{D}}(x-1) = 0.$$

Therefore, by Definition 1.1, $f_{\mathcal{D}}$ must be some scalar multiple of the Dickman function ρ. To be a valid distribution function, we must have the appropriate normalisation $f_{\mathcal{D}}(u) = e^{-\gamma}\rho(u)$ for $u > 0$. Furthermore, by (8), $F_{\mathcal{D}}(x) = 0$ for $x \le 0$. Subsequently, $\mathcal{D}$ has Dickman distribution according to Definition 1.3. □

Now, we define

$$\mathcal{D}_\theta = \mathcal{U}_1^{1/\theta} + \mathcal{U}_1^{1/\theta}\mathcal{U}_2^{1/\theta} + \mathcal{U}_1^{1/\theta}\mathcal{U}_2^{1/\theta}\mathcal{U}_3^{1/\theta} + \dots . \tag{9}$$

By analysing the Laplace transform of $\mathcal{D}_\theta$, we will prove that $\mathcal{D}$ as defined in (8) is infinitely divisible. We again have the relation $\mathcal{D}_\theta = \mathcal{U}^{1/\theta}(1+\mathcal{D}_\theta)$. Therefore the Laplace transform of $\mathcal{D}_\theta$ is

$$\begin{aligned} L_\theta(s) := \mathbb{E}[e^{-s\mathcal{D}_\theta}] &= \mathbb{E}[e^{-s\mathcal{U}^{1/\theta}(1+\mathcal{D}_\theta)}] \\ &= \theta \int_0^1 \int_{\mathbb{R}} e^{-st(1+x)}\, d\mathcal{D}_\theta(x)\, t^{\theta-1}\, dt \\ &= \theta \int_0^1 e^{-st} t^{\theta-1} \Big(\int_{\mathbb{R}} e^{-stx} d\mathcal{D}_\theta(x) \Big)\, dt \\ &= \theta \int_0^1 e^{-st} t^{\theta-1} L_\theta(st)\, dt \\ &= \frac{\theta}{s^\theta} \int_0^s e^{-y} y^{\theta-1} L_\theta(y)\, dy. \end{aligned}$$

Multiplying both sides by s^θ and differentiating, we find that L_θ satisfies the following differential equation:

$$\begin{aligned} s^\theta L'_\theta(s) + \theta s^{\theta-1} L_\theta(s) &= \theta e^{-s} s^{\theta-1} L_\theta(s) \\ \Rightarrow s L'_\theta(s) + \theta L_\theta(s) &= \theta e^{-s} L_\theta(s). \end{aligned}$$

The unique solution of this equation with $L_\theta(0) = 1$ (which is clearly implied by the definition of L_θ) is

$$L_\theta(s) = \exp\Big(\theta \int_0^1 \frac{e^{-sx} - 1}{x}\Big). \tag{10}$$

For $n \in \mathbb{N}$, if we take $\mathcal{D}^{(1)}_{1/n}, \mathcal{D}^{(2)}_{1/n}, ..., \mathcal{D}^{(n)}_{1/n}$ independent copies of $\mathcal{D}_{1/n}$, then by (10) we have that the Laplace transform of their sum is

$$\mathbb{E}[e^{-s(\mathcal{D}^{(1)}_{1/n}+\mathcal{D}^{(2)}_{1/n}+...+\mathcal{D}^{(n)}_{1/n})}] = L_{1/n}(s)^n = \Big(\exp\Big(\frac{1}{n}\int_0^1 \frac{e^{-sx}-1}{x}\Big)\Big)^n = L_1(s).$$

As discussed, the Laplace transform uniquely defines a distribution, so noting that $\mathcal{D}_1 = \mathcal{D}$, we find that for any $n \in \mathbb{N}$,

$$\mathcal{D} = \mathcal{D}^{(1)}_{1/n} + \mathcal{D}^{(2)}_{1/n} + ... + \mathcal{D}^{(n)}_{1/n},$$

for i.i.d $(\mathcal{D}^{(i)}_{1/n})_{i=1}^n$ defined by (9). This proves that the Dickman distribution is infinitely divisible.

5 Stein's method for the Dickman Distribution

It is stated in Arras et al. [Arr+16] that a random variable Z has Dickman density according to Definition 1.3 if and only if for all bounded functions f we have

$$\mathbb{E}[Zf(Z)] = \mathbb{E}[f(Z+U)],$$

where $U \sim U([0,1])$ is independent of Z. This is, ofcourse, another outcome of the differential equation that characterises the Dickman function. Indeed, suppose that Z has Dickman distribution according to Definition 1.3, then for any bounded function f, we have

$$\begin{aligned}
\mathbb{E}[f(Z+U)] &= \int_0^1 \int_0^\infty f(x+u)e^{-\gamma}\,dx\,du = \int_0^1 \int_0^\infty f(t)e^{-\gamma}\rho(t-u)\,dt\,du \\
&= \int_0^\infty f(t) \int_0^{\min\{1,t\}} e^{-\gamma}\rho(t-u)\,du\,dt \\
&= \int_0^1 f(t) \int_0^t e^{-\gamma}\rho(t-u)\,du\,dt + \int_1^\infty f(t)\int_0^1 e^{-\gamma}\rho(t-u)\,du\,dt \\
&= \int_0^1 f(t) \int_0^t e^{-\gamma}\rho(u)\,du\,dt + \int_1^\infty f(t)\int_{t-1}^t e^{-\gamma}\rho(u)\,du\,dt \\
&= \int_0^\infty tf(t)e^{-\gamma}\rho(t)\,dt = \mathbb{E}[Zf(Z)],
\end{aligned}$$

as required. To obtain the last line we have used the fact that $\rho(t) = 1$ for $t \in [0,1]$ in the first integral, and the fact that $\int_{t-1}^t \rho(u)\,du = t\rho(t)$ in the second integral. We note that f being bounded allows the use of Fubini.

Taking $f(\xi) = e^{it\xi}$ allows us to calculate the characteristic function of the Dickman distribution. Denoting $D(x)$ as the cumulative distribution function of a random variable with Dickman distribution, we have

$$\begin{aligned}\varphi_Z'(t) &= \frac{d}{dt}\mathbb{E}[e^{itZ}] = \mathbb{E}[iZe^{itZ}] \\ &= i\mathbb{E}[e^{it(Z+U)}] = i\int_0^\infty e^{itx}\int_0^1 e^{itu}\,du\,dD(x) \\ &= i\int_0^\infty e^{itx}\Big(\frac{1}{it}(e^{it}-1)\Big)\,dD(x) = \frac{e^{it}-1}{t}\varphi_Z(t).\end{aligned}$$

Therefore, the logarithmic derivative of $\varphi_Z(t)$ is $(e^{it}-1)/t$, and so an explicit form for the characteristic function is

$$\varphi_Z(t) = \exp\Big(\int_0^t \frac{e^{ix}-1}{x}dx\Big) = \exp\Big(\int_0^1 \frac{e^{itx}-1}{x}\,dx\Big). \tag{11}$$

Analysing this characteristic function may lead one to another proof of the infinite divisibility of the Dickman distribution! The above technique also gives an alternate method of computing the Laplace transform of the Dickman distribution, as oppose to our method used to arrive at (10).

6 Closing Remark

It seems that the defining properties of the Dickman function (recall Definition 1.1) give rise to distributions that have a kind of 'self-similarity' property, in that the differential equation defining the Dickman function gives relations between the properties of these distributions. This kind of self-similarity may, at least heuristically, be made explicit for each distribution in (5) and (8). It is this property of the Dickman function that makes it unavoidable in probability.

References

[AB18] George E. Andrews and Bruce C. Berndt. "Ramanujan's lost notebook". In: (2018).

[ABT03] Richard Arratia, A. D. Barbour, and Simon Tavaré. "Logarithmic combinatorial structures: A probabilistic approach". In: *EMS Monographs in Mathematics* (2003).

[Arr+16] Benjamin Arras et al. "Distances between probability distributions via characteristic functions and biasing". In: (2016).

[Dev01] Luc Devroye. In: *Methodology And Computing In Applied Probability* (2001).

[DG93] Peter Donnelly and Geoffrey Grimmett. "On the Asymptotic Distribution of Large Prime Factors". In: *Journal of the London Mathematical Society* (1993).

[GG19] A. Granville and J. Granville. *Prime Suspects: The Anatomy of Integers and Permutations.* Princeton University Press, 2019.

[HT02] HSIEN-KUEI HWANG and TSUNG-HSI TSAI. "Quickselect and the Dickman Function". In: *Combinatorics, Probability and Computing* (2002).

[Kou19] D. Koukoulopoulos. *The Distribution of Prime Numbers.* American Mathematical Society, 2019.

[MP20] S. A. Molchanov and V. A. Panov. "The Dickman–Goncharov distribution". In: *Russian Mathematical Surveys* (2020).

[Ram49] V. Ramaswami. "On the number of positive integers less than and free of prime divisors greater than ". In: *Bulletin of the American Mathematical Society* (1949).

[Tao] Terence Tao. *Dickman's function*. URL: `https://terrytao.wordpress.com/tag/dickmans-function/`.

[War+06] Max Warshauer et al. "The locker puzzle". In: *The Mathematical Intelligencer* (2006).