PROBABILISTIC PROOFS OF SOME GENERALIZED MERTENS' FORMULAS VIA GENERALIZED DICKMAN DISTRIBUTIONS

ROSS G. PINSKY

ABSTRACT. The classical Mertens' formula states that $\prod_{p\le N}(1-\frac1p)^{-1}\sim e^\gamma\log N$, where the product is over all primes p less than or equal to N, and γ is the Euler-Mascheroni constant. By the Euler product formula, this is equivalent to the following statement:

$$\lim_{N\to\infty}\frac{\sum_{n:p|n\Rightarrow p\le N}\frac1n}{\sum_{n\le N}\frac1n}=e^\gamma.$$

Via some random integer constructions and a criterion for weak convergence of distributions to so-called generalized Dickman distributions, we obtain some generalized Mertens' formulas that have been proved using number-theoretic tools. We show that if A is a subset of the primes which has natural density $\theta\in(0,1]$ with respect to the set of all primes, then

$$\lim_{N\to\infty}\frac{\sum_{n:p|n\Rightarrow p\le N\text{ and }p\in A}\frac1n}{\sum_{n\le N:p|n\Rightarrow p\in A}\frac1n}=e^{\gamma\theta}\Gamma(\theta+1),$$

and also, for any $k\ge2$,

$$\lim_{N\to\infty}\frac{\sum'^{(k)}_{n:p|n\Rightarrow p\le N\text{ and }p\in A}\frac1n}{\sum'^{(k)}_{n\le N:p|n\Rightarrow p\in A}\frac1n}=e^{\gamma\theta}\Gamma(\theta+1),$$

where $\sum'^{(k)}$ denotes that the summation is restricted to k-free positive integers. We also show that

$$\lim_{N\to\infty}\frac{\sum_{n:p|n\Rightarrow p\le N\text{ and }p\in A}\frac1n}{\sum'^{(k)}_{n\le N:p|n\Rightarrow p\in A}\frac{1}{n_{\{(k-1)-\text{free}\}}\phi(n_{\{(k-1)-\text{power}\}})}}=e^{\gamma\theta}\Gamma(\theta+1),$$

where ϕ is the Euler totient function, and $n_{\{(k-1)-\text{free}\}}$ and $n_{\{(k-1)-\text{power}\}}$ are the $(k-1)$-free part and the $(k-1)$-power part of n.

2010 *Mathematics Subject Classification.* 11N25, 11N37, 60F05, 11K65.

Key words and phrases. Mertens' formula, generalized Mertens' formulas, Dickman distribution, generalized Dickman distributions .

1. Introduction and Statement of Results

The classical Mertens' formula states that

$$\prod_{p\le N}\big(1-\frac{1}{p})^{-1}\sim e^{\gamma}\log N\approx 1.78\log N,\ \text{as } N\to\infty, \tag{1.1}$$

where the product is over all primes p less than or equal to N, and γ is the Euler-Mascheroni constant. (Actually, the classical formula states a little more; namely, $\prod_{p\le N}\big(1-\frac{1}{p})^{-1}=e^{\gamma}\log N+O(1)$.) By the Euler product formula, (1.1) is equivalent to the following statement, which is more in the spirit of the results we present in this paper:

$$\lim_{N\to\infty}\frac{\sum_{n\in\mathbb{N}:p|n\Rightarrow p\le N}\frac{1}{n}}{\sum_{n\le N}\frac{1}{n}}=e^{\gamma}, \tag{1.2}$$

where as usual, $[N]$ denotes the set $\{1,\cdots,N\}$.

One of the results in [5] involved the construction of a sequence of random integers whose distributions were shown to converge weakly to the so-called Dickman distribution. It was noted in that paper that Mertens' formula follows readily as a corollary of this result. In this paper, we make a number of random integer constructions in a similar vein, and use our recent paper [4] to show that their distributions converge weakly to so-called generalized Dickman distributions. From these results, we obtain several generalizations of Mertens' formula, which have been proved via number theoretic methods.

We begin by introducing some notation and constructing the three sequences of random integers that will be used in this paper. Then we state our generalized Mertens' formulas.

Let $\mathbb{P}$ denote the set of prime numbers. Recall that for $k\ge 2$, an integer $n\in\mathbb{N}$ is called *k-free* if $p^k\nmid n$, for all primes p. Let $A\subset\mathbb{P}$ be an infinite set of primes. Denote the primes in A in increasing order by $p_{1;A},p_{2;A},\cdots$. Let $\{T_j\}_{j=1}^{\infty}$ be a sequence of independent random variables with T_j distributed according to the geometric distribution with parameter $\frac{1}{p_{j;A}}$ $\big(T_j\sim\text{Geom}(\frac{1}{p_{j;A}})\big)$, $j=1,\cdots$; that is

$$P(T_j=m)=(1-\frac{1}{p_{j;A}})(\frac{1}{p_{j;A}})^m,\ m=0,1,\cdots. \tag{1.3}$$

For $N \in \mathbb{N}$, we define a random integer by

$$I_{N;A,1} = \prod_{j=1}^{N} p_{j;A}^{T_j}. \tag{1.4}$$

By construction, the support of $I_{N;A,1}$ is $\{n \in \mathbb{N} : p|n \Rightarrow p \le N \,\text{and}\, p \in A\}$. See [5] for a detailed study of the random integer sequence $\{I_{N;A,1}\}_{N=1}^{\infty}$ when $A = \mathbb{P}$.

Let $k \ge 2$. We define a second random integer sequence by replacing the random variables $\{T_j\}_{j=1}^{\infty}$ by a sequence $\{U_j\}_{j=1}^{\infty}$ of independent random variables, where U_j is distributed as T_j conditioned on being less than k $\big(U_j \overset{\text{dist}}{=} T_j|\{T_j < k\}\big)$; that is

(1.5)

$$P(U_j = m) = P(T_j = m|T_j < k) = \frac{1 - \frac{1}{p_{j;A}}}{1 - (\frac{1}{p_{j;A}})^k}(\frac{1}{p_{j;A}})^m, \ m = 0, 1, \cdots, k-1.$$

For $N \in \mathbb{N}$, define a random integer by

$$I_{N;A,2} = \prod_{j=1}^{N} p_{j;A}^{U_j}. \tag{1.6}$$

By construction, the support of $I_{N;A,2}$ is the set of k-free integers in $\{n \in \mathbb{N} : p|n \Rightarrow p \le N \,\text{and}\, p \in A\}$.

Finally, for $k \ge 2$, we construct a third random integer sequence from a sequence $\{V_j\}_{j=1}^{\infty}$ of independent random variables, where V_j is distributed as T_j truncated at $k-1$ $\big(V_j \overset{\text{dist}}{=} T_j \wedge (k-1)\big)$; that is

$$\begin{aligned} &P(V_j = m) = (1 - \frac{1}{p_{j;A}})(\frac{1}{p_{j;A}})^m, \ m = 0, \cdots, k-2; \\ &P(V_j = k-1) = (\frac{1}{p_{j;A}})^{k-1}. \end{aligned} \tag{1.7}$$

For $N \in \mathbb{N}$, define a random integer by

$$I_{N;A,3} = \prod_{j=1}^{N} p_{j;A}^{V_j}. \tag{1.8}$$

By construction, the support of $I_{N;A,3}$ is the set of k-free integers in $\{n \in \mathbb{N} : p|n \Rightarrow p \le N \,\text{and}\, p \in A\}$. In the notation, we suppress the dependence on k.

We now present our generalized Mertens' formulas. For $A \subset \mathbb{P}$, denote by

$$D_{\text{nat-prime}}(A) := \lim_{N\to\infty} \frac{|A\cap[N]|}{|\mathbb{P}\cap[N]|}$$

the natural density of A in $\mathbb{P}$, if it exists. For $B \subset \mathbb{N}$, let $\sum_B^{\prime(k)}$ denote the summation restricted to the k-free powers in B. Let $\Gamma(x) := \int_0^\infty t^{x-1}e^{-t}dt,\ x>0$, denote the Gamma function.

Theorem 1. *Let $A \subset \mathbb{P}$ be a subset of primes whose natural density in $\mathbb{P}$ is $D_{nat\text{-}prime}(A) = \theta \in (0,1]$. Then*
i.

$$\lim_{N\to\infty} \frac{\sum_{n\in\mathbb{N}:p|n\Rightarrow p\le N\ and p\in A} \frac{1}{n}}{\sum_{n\le N:p|n\Rightarrow p\in A}\frac{1}{n}} = e^{\gamma\theta}\,\Gamma(\theta+1); \tag{1.9}$$

ii. for $k\ge 2$,

$$\lim_{N\to\infty} \frac{\sum^{\prime(k)}_{n\in\mathbb{N}:p|n\Rightarrow p\le N\ and p\in A} \frac{1}{n}}{\sum^{\prime(k)}_{n\le N:p|n\Rightarrow p\in A}\frac{1}{n}} = e^{\gamma\theta}\,\Gamma(\theta+1).$$

Remark 1. Part (i) of Theorem 1 in a slightly different but equivalent form appears in [10], and a refined version appears in [6] . Part (ii) seems to be new.
Remark 2. The function $\theta \to e^{\gamma\theta}\Gamma(\theta+1)$, $\theta\in(0,1]$, is increasing, is equal to 1 at $\theta=0^+$ and is equal to e^γ at $\theta=1$.
Remark 3. When $A=\mathbb{P}$, (i) reduces to the classical Mertens' formula.
Remark 4. For $l\in\mathbb{N}$ and j satisfying $1\le j<l$ and $(j,l)=1$, let $A_{l;j} = \{p\in\mathbb{P} : p = j \bmod l\}$ denote the set of primes that are equal to j modulo l. Dirichlet's arithmetic progression theorem states that $D_{\text{nat-prime}}(A_{l;j}) = \frac{1}{\phi(l)}$, where ϕ is Euler's totient function. In [9], it was proved that

$$\prod_{p\le N, p=j \bmod l}(1-\frac{1}{p})^{-1} = \sum_{n:p|n\Rightarrow p\le N \text{ and } p\in A_{l;j}} \frac{1}{n} \sim \Big(e^\gamma \frac{\phi(l)}{l}C(l,j)\Big)^{\frac{1}{\phi(l)}} (\log N)^{\frac{1}{\phi(l)}}, \tag{1.10}$$

as $N \to \infty$, where $C(l,j)$ is a complicated expression involving Dirichlet characters modulo l. Thus, by part (i) we obtain

$$\sum_{n\le N: p|n\Rightarrow p\in A_{l;j}} \frac{1}{n} \sim \frac{1}{\Gamma(1+\frac{1}{\phi(l)})}\Big(\frac{\phi(l)}{l}C(l,j)\Big)^{\frac{1}{\phi(l)}}(\log N)^{\frac{1}{\phi(l)}},\ \text{as } N\to\infty,$$

as was noted in [9]. A much simpler looking form for $C(l,j)$ was obtained in [1]; namely,

$$\frac{\phi(l)}{l}C(l,j) = \prod_p (1-\frac{1}{p})^{-\alpha(p,l,j)},\ \text{where } \alpha(p,l,j) = \begin{cases}\phi(l)-1, & p=j \bmod l;\\ -1, & \text{otherwise.}\end{cases}$$

See also [2] for more on this constant.
From part (ii), we obtain for $k\ge 2$,

$$\lim_{N\to\infty}\frac{\sum'^{(k)}_{n:p|n\Rightarrow p\le N \text{ and } p\in A_{l;j}}\frac{1}{n}}{\sum^{('(k)}_{n\le N:p|n\Rightarrow p\in A_{l;j}}\frac{1}{n}} = e^{\frac{\gamma}{\phi(l)}}\,\Gamma(\frac{1}{\phi(l)}+1).$$

Our second theorem involves the Euler totient function. We will need the following notation. For each $k\ge 2$, every $n\in\mathbb{N}$ can be written uniquely as $n = n_{\{k-\text{free}\}}n_{\{k-\text{power}\}}$, where $n_{\{k-\text{free}\}}$ is k-free and $n_{\{k-\text{power}\}}$ is a kth power.

Theorem 2. *Let $A\subset\mathbb{P}$ be a subset of primes whose density in $\mathbb{P}$ is $D_{\textit{nat-prime}}(A)=\theta\in(0,1]$. Then for all $k\ge 2$,*

$$\lim_{N\to\infty}\frac{\sum_{n:p|n\Rightarrow p\le N \text{ and} p\in A}\frac{1}{n}}{\sum'^{(k)}_{n\le N:p|n\Rightarrow p\in A}\frac{1}{n_{\{(k-1)-free\}}\phi(n_{\{(k-1)-power\}})}} = e^{\gamma\theta}\Gamma(\theta+1). \tag{1.11}$$

Remark. Note that when $k=2$, the result is

$$\lim_{N\to\infty}\frac{\sum_{n:p|n\Rightarrow p\le N \text{ and } p\in A}\frac{1}{n}}{\sum'^{(2)}_{n\le N:p|n\Rightarrow p\in A}\frac{1}{\phi(n)}} = e^{\gamma\theta}\,\Gamma(\theta+1).$$

From Theorem 2, we obtain the following corollary

Corollary 1. *Let $A \subset \mathbb{P}$ be a subset of primes whose density in $\mathbb{P}$ is $D_{nat\text{-}prime}(A) = \theta \in (0,1]$. Then for all $k \geq 2$,*

$$\sum\nolimits'^{(k)}_{n \leq N: p|n \Rightarrow p \in A} \frac{1}{n_{\{(k-1)-free\}}\phi(n_{\{(k-1)-power\}})} \sim \sum\nolimits_{n \leq N: p|n \Rightarrow p \in A} \frac{1}{n}, \ \text{as } N \to \infty.$$

Proof of Corollary. Compare (1.9) to (1.11). □

Remark. When $A = \mathbb{P}$, Corollary 1 reduces to

$$\sum\nolimits'^{(k)}_{n \leq N:} \frac{1}{n_{\{(k-1)-\text{free}\}}\phi(n_{\{(k-1)-\text{power}\}})} \sim \log N, \ \text{as } N \to \infty, \ \text{for all } k \geq 2.$$

When $k = 2$, this reduces to

$$\sum\nolimits'^{(2)}_{n \leq N} \frac{1}{\phi(n)} \sim \log N, \ \text{as } N \to \infty,$$

which is known (see [8] or [3, p. 43, problem 17]. We were unable to find a reference for the generalization to all $k \geq 2$. As an aside, we note that

$$\sum_{n \leq N} \frac{1}{\phi(n)} \sim \frac{\zeta(2)\zeta(3)}{\zeta(6)} \log N \approx 1.94 \log N, \ \text{as } N \to \infty.$$

(see [3, p. 42, problem 13-(d)]).

We prove Theorems 1 and 2 in sections 2 and 3 respectively.

2. Proof of Theorem 1

Fix a subset $A \subset \mathbb{P}$ which satisfies $D_{\text{nat-prime}}(A) = \theta \in (0,1]$. Denote the primes in A in increasing order by $p_{1;A}, p_{2;A}, \cdots$, and let

$$A_N = \{p_{1;A}, p_{2;A}, \cdots, p_{N;A}\}.$$

Proof of part (i). Let the random integer $I_{N;A,1}$ be as in (1.4), where $\{T_j\}_{j=1}^{\infty}$ is a sequence of independent random variables with distributions given by (1.3). The support of $I_{N;A,1}$ is $\{n \in \mathbb{N} : p|n \Rightarrow p \in A_N\}$, and for arbitrary $n = \prod_{j=1}^N p_{j;A}^{c_j}$ in the support,

$$\begin{aligned} &P(I_{N;A,1} = n) = P(T_j = c_j,\, j \in [N]) = \prod_{j=1}^N P(T_j = c_j) \\ &= \prod_{j=1}^N (1 - \frac{1}{p_{j;A}})(\frac{1}{p_{j;A}})^{c_j} = \frac{1}{n} \prod_{j=1}^N (1 - \frac{1}{p_{j;A}}). \end{aligned} \tag{2.1}$$

We have

$$\log I_{N;A,1} = \sum_{j=1}^{N} T_j \log p_{j;A}.$$

Noting that the expected value of T_j is given by

$$ET_j = \frac{1}{p_{j;A} - 1}, \tag{2.2}$$

we have

$$E \log I_{N;A,1} = \sum_{j=1}^{N} \frac{\log p_{j;A}}{p_{j;A} - 1}.$$

It follows by the assumption on the density of A and by the prime number theorem that

$$p_{j;A} \sim \frac{j \log j}{\theta}, \text{ as } j \to \infty, \tag{2.3}$$

and thus that

$$E \log I_{N;A,1} \sim \theta \log N, \text{ as } N \to \infty. \tag{2.4}$$

We will demonstrate below that the conditions of a theorem in [4] are satisfied, from which it follows that

$$\lim_{N\to\infty} \frac{\log I_{N;A,1}}{E \log I_{N;A,1}} \stackrel{\text{dist}}{=} \frac{1}{\theta} D_\theta, \tag{2.5}$$

where D_θ is a random variable distributed according to the generalized Dickman distribution GD(θ) with parameter θ. This distribution has density function $p_\theta = \frac{e^{-\gamma\theta}}{\Gamma(\theta)}\rho_\theta$, where ρ_θ satisfies the differential-delay equation

$$\begin{aligned} &\rho_\theta(x) = 0, \ x \le 0; \\ &\rho_\theta(x) = x^{\theta-1}, \ 0 < x \le 1; \\ &x\rho'_\theta(x) + (1-\theta)\rho_\theta(x) + \theta\rho_\theta(x-1) = 0, \ x > 1. \end{aligned} \tag{2.6}$$

(The function ρ_1 is known as the Dickman function; we call ρ_θ a generalized Dickman function.)

On the one hand, by the convergence in distribution in (2.5) and the fact that the limiting distribution is a continuous one, for any sequence $\{\theta_N\}_{N=1}^{\infty}$

satisfying $\lim_{N\to\infty}\theta_N=\theta$, we have

$$\begin{aligned}&\lim_{N\to\infty}P(\frac{\log I_{N;A,1}}{E\log I_{N;A,1}}\le\frac{1}{\theta_N})=P(\frac{1}{\theta}D_\theta\le\frac{1}{\theta})=\int_0^1 p_\theta(x)dx\\&=\frac{e^{-\gamma\theta}}{\Gamma(\theta)}\int_0^1 x^{\theta-1}dx=\frac{e^{-\gamma\theta}}{\theta\Gamma(\theta)}=\frac{e^{-\gamma\theta}}{\Gamma(\theta+1)}.\end{aligned}\tag{2.7}$$

On the other hand, let $\theta_N:=\frac{E\log I_{N;A,1}}{\log p_{N;A}}$ and note from (2.3) and (2.4) that $\lim_{N\to\infty}\theta_N=\theta$. It follows from (2.1) that

$$\begin{aligned}&P(\frac{\log I_{N;A,1}}{E\log I_{N;A,1}}\le\frac{1}{\theta_N})=P\big(I_{N;A,1}\le\exp(\frac{E\log I_{N;A,1}}{\theta_N})\big)\\&=P(I_{N;A,1}\le p_{N;A})=\Big(\prod_{j=1}^N(1-\frac{1}{p_{j;A}})\Big)\sum_{n\le p_{N;A}:p|n\Rightarrow p\in A}\frac{1}{n}\\&=\frac{\sum_{n\le p_{N;A}:p|n\Rightarrow p\in A}\frac{1}{n}}{\sum_{n:p|n\Rightarrow p\in A_N}\frac{1}{n}}.\end{aligned}\tag{2.8}$$

From (2.7) and (2.8), we conclude that

$$\lim_{N\to\infty}\frac{\sum_{n:p|n\Rightarrow p\in A_N}\frac{1}{n}}{\sum_{n\le p_{N;A}:p|n\Rightarrow p\in A}\frac{1}{n}}=e^{\gamma\theta}\Gamma(\theta+1).\tag{2.9}$$

Now (2.9) is equivalent to part (i) of Theorem 1. Indeed, for any $M\in\mathbb{N}$, let $N^+(M)=\max\{n:p_{n;A}\le M\}$. Then

$$\sum_{n:p|n\Rightarrow p\le M\text{ and }p\in A}\frac{1}{n}=\sum_{n:p|n\Rightarrow p\le p_{N^+(M);A}\text{ and }p\in A}\frac{1}{n},\tag{2.10}$$

and

$$\sum_{n\le M:p|n\Rightarrow p\in A}\frac{1}{n}=\sum_{n\le p_{N^+(M);A}:p|n\Rightarrow p\in A}\frac{1}{n}+H_M,\tag{2.11}$$

where

(2.12)

$$H_M:=\sum_{n\in[p_{N^+(M);A}+1,M]:p|n\Rightarrow p\in A}\frac{1}{n}\le\sum_{n=p_{N^+(M);A}+1}^{p_{N^+(M)+1;A}}\frac{1}{n}\le\log\frac{p_{N^+(M)+1;A}}{p_{N^+(M);A}}=o(1).$$

Thus, (2.9)-(2.12) give

$$\lim_{M\to\infty}\frac{\sum_{n:p|n\Rightarrow p\le M\text{ and }p\in A}\frac{1}{n}}{\sum_{n\le M:p|n\Rightarrow p\in A}\frac{1}{n}}=e^{\gamma\theta}\Gamma(\theta+1),$$

which is part (i) of the theorem.

We now show that (2.5) holds. Let $\{B_j\}_{j=1}^\infty$ and $\{Xj\}_{j=1}^\infty$ be mutually independent random variables with distributions as follows:
$B_j \sim \text{Ber}(\frac{1}{p_{j;A}})$; that is,

$$q_j := P(B_j = 1) = 1 - P(B_j = 0) = \frac{1}{p_{j;A}}. \tag{2.13}$$

$X_j \overset{\text{dist}}{=} \log p_{j;A} \cdot T_j|\{T_j \geq 1\}$; that is

$$P(X_j = m \log p_{j;A}) = P(T_j = m|T_j \geq 1) = (1-\frac{1}{p_{j;A}})(\frac{1}{p_{j;A}})^{m-1},\ m = 1, 2, \cdots .$$

Then

$$\mu_j := EX_j = \frac{p_{j;A}}{p_{j;A} - 1} \log p_{j;A}, \tag{2.14}$$

and it follows that

$$\lim_{j \to \infty} \frac{X_j}{\mu_j} \overset{\text{dist}}{=} 1.$$

By the construction of $\{B_j\}_{j=1}^\infty$ and $\{X_j\}_{j=1}^\infty$, we have

$$\log I_{N;A,1} = \sum_{j=1}^N T_j \log p_{j;A} \overset{\text{dist}}{=} \sum_{j=1}^N B_j X_j. \tag{2.15}$$

From (2.3), (2.13) and (2.14),

$$\mu_j \sim \log j; \quad q_j \sim \frac{\theta}{j \log j}, \text{ as } j \to \infty. \tag{2.16}$$

Let $W_N = \frac{\sum_{j=1}^N B_j X_j}{E \sum_{j=1}^N B_j X_j}$. Now Theorem 1.2 in [4] applies to W_N. Our notation here coincides with the notation in that theorem except that the summation there is over k while here it is over j, and p_k there corresponds to q_j here. In light of (2.16), we have $J_\mu = J_p = 1, a_0 = 0, a_1 = b_0 = b_1 = c_\mu = 1, c_p = \theta$ in the notation of that theorem. For these values, the theorem indicates that W_N converges in distribution to $\frac{1}{\theta} D_\theta$. By (2.15), $\frac{\log I_{N;A,1}}{E \log I_{N;A,1}} \overset{\text{dist}}{=} W_N$; thus (2.5) holds. □

Proof of part (ii). Fix $k \geq 2$. The proof follows the proof of part (i), except that we replace the random integer $I_{N;A,1}$ by the random integer $I_{N;A,2}$ from (1.6), where $\{U_j\}_{j=1}^\infty$ is a sequence of random variables with distributions given by (1.5). The support of $I_{N;A,2}$ is the set of k-free integers in $\{n\in\mathbb{N}: p|n \Rightarrow p\in A_N\}$, and for arbitrary $n=\prod_{j=1}^N p_{j;A}^{c_j}$ in the support,

$$\begin{aligned}P(I_{N;A,2}=n)&=\prod_{j=1}^N P(U_j=c_j)=\prod_{j=1}^N\frac{1-\frac{1}{p_{j;A}}}{1-(\frac{1}{p_{j;A}})^k}(\frac{1}{p_{j;A}})^{c_j}\\ &=\frac{1}{n}\prod_{j=1}^N\big(1+\frac{1}{p_{j;A}}+(\frac{1}{p_{j;A}})^2+\cdots+(\frac{1}{p_{j;A}})^{k-1}\big)^{-1}.\end{aligned}\tag{2.17}$$

We have

$$\log I_{N;A,2}=\sum_{j=1}^N U_j\log p_{j;A}.$$

Since

$$\begin{aligned}EU_j&=\frac{1-\frac{1}{p_{j;A}}}{1-(\frac{1}{p_{j;A}})^k}\sum_{m=0}^{k-1}m(\frac{1}{p_{j;A}})^m\\ &=\frac{1-\frac{1}{p_{j;A}}}{1-(\frac{1}{p_{j;A}})^k}\frac{1+(k-1)(\frac{1}{p_{j;A}})^k-k(\frac{1}{p_{j;A}})^{k-1}}{(1-\frac{1}{p_{j;A}})^2}\frac{1}{p_{j;A}},\end{aligned}$$

we have

$$EU_j\sim\frac{1}{p_{j;A}},\ \text{as } j\to\infty.\tag{2.18}$$

From (2.2) note that EU_j and ET_j have the same asymptotic behavior. From (2.3) and (2.18), we have

$$E\log I_{N;A,2}\sim\sum_{j=1}^N\frac{\log p_{j;A}}{p_{j;A}}\sim\theta\log N,\ \text{as } N\to\infty.\tag{2.19}$$

Note from (2.4) that $E\log I_{N;A,2}$ and $E\log I_{N;A,1}$ have the same asymptotic behavior.

We now give the appropriate redefinition of the mutually independent random variables $\{B_j\}_{j=1}^\infty$ and $\{Xj\}_{j=1}^\infty$ that were defined in the proof of part (i):
$B_j\sim\text{Ber}\big(\frac{\frac{1}{p_{j;A}}-(\frac{1}{p_{j;A}})^k}{1-(\frac{1}{p_{j;A}})^k}\big)$; that is,

$$q_j:=P(B_j=1)=1-P(B_j=0)=\frac{\frac{1}{p_{j;A}}-(\frac{1}{p_{j;A}})^k}{1-(\frac{1}{p_{j;A}})^k}.\tag{2.20}$$

$X_j \stackrel{\text{dist}}{=} \log p_{j;A} \cdot U_j|\{U_j \geq 1\}$; that is

$$P(X_j = m \log p_{j;A}) = P(U_j = m|U_j \geq 1) = \frac{1 - \frac{1}{p_{j;A}}}{1 - (\frac{1}{p})^{k-1}} (\frac{1}{p_{j;A}})^{m-1},\ m = 1, \cdots, k-1.$$

As in part (i), we have $\mu_j := EX_j \sim \log p_{j;A}$, as $j \to \infty$, and $\lim_{j\to\infty} \frac{X_j}{\mu_j} \stackrel{\text{dist}}{=} 1$. By the construction of $\{B_j\}_{j=1}^{\infty}$ and $\{X_j\}_{j=1}^{\infty}$, we have

$$\log I_{N;A,2} = \sum_{j=1}^{N} U_j \log p_{j;A} \stackrel{\text{dist}}{=} \sum_{j=1}^{N} B_j X_j.$$

By the same considerations as in part (i), it follows that $W_N := \frac{\sum_{j=1}^N B_j X_j}{E \sum_{j=1}^N B_j X_j}$ converges in distribution to $\frac{1}{\theta} D_\theta$; thus,

$$\lim_{N\to\infty} \frac{\log I_{N;A,2}}{E \log I_{N;A,2}} \stackrel{\text{dist}}{=} \frac{1}{\theta} D_\theta. \tag{2.21}$$

On the one hand, just as in (2.7), by the convergence in distribution in (2.21) and the fact that the limiting distribution is a continuous one, for any for any sequence $\{\theta_N\}_{N=1}^{\infty}$ satisfying $\lim_{N\to\infty} \theta_N = \theta$, we have

$$\lim_{N\to\infty} P(\frac{\log I_{N;A,2}}{E \log I_{N;A,2}} \leq \frac{1}{\theta_N}) = P(\frac{1}{\theta} D_\theta \leq \frac{1}{\theta}) = \frac{e^{-\gamma\theta}}{\Gamma(\theta+1)}. \tag{2.22}$$

On the other hand, let $\theta_N := \frac{E \log I_{N;A,2}}{\log p_{N;A}}$ and note from (2.3) and (2.19) that $\lim_{N\to\infty} \theta_N = \theta$. It follows from (2.17) that

$P(\frac{\log I_{N;A,2}}{E \log I_{N;A,2}} \leq \frac{1}{\theta_N}) = P\big(I_{N;A,2} \leq \exp(\frac{E \log I_{N;A,2}}{\theta_N})\big) = P(I_{N;A,2} \leq p_{N;A})$

$= \prod_{j=1}^N \big(1 + \frac{1}{p_{j;A}} + (\frac{1}{p_{j;A}})^2 + \cdots + (\frac{1}{p_{j;A}})^{k-1}\big)^{-1} \sum'^{(k)}_{n \leq p_{N;A}: p|n \Rightarrow p \in A} \frac{1}{n}$

$= \frac{\sum'^{(k)}_{n \leq p_{N;A}: p|n \Rightarrow p \in A} \frac{1}{n}}{\sum'^{(k)}_{n: p|n \Rightarrow p \in A_N} \frac{1}{n}}$. From this and (2.22), we conclude that

$$\lim_{N\to\infty} \frac{\sum'^{(k)}_{n: p|n \Rightarrow p \in A_N} \frac{1}{n}}{\sum'^{(k)}_{n \leq p_{N;A}: p|n \Rightarrow p \in A} \frac{1}{n}} = e^{\gamma\theta} \Gamma(\theta+1), \tag{2.23}$$

which is equivalent to part (ii) of Theorem 1, just as (2.9) was equivalent to part (i) of the theorem. □

3. Proof of Theorem 2

Fix a subset $A \subset \mathbb{P}$ which satisfies $D_{\text{nat-prime}}(A) = \theta \in (0,1]$. Denote the primes in A in increasing order by $p_{1;A}, p_{2;A}, \cdots$, and let

$$A_N = \{p_{1;A}, p_{2;A}, \cdots, p_{N;A}\}.$$

Let $I_{N;A,3}$ be as in (1.8), where $\{V_j\}_{j=1}^\infty$ are independent random variables with distributions given by (1.7). The support of $I_{N;A,3}$ is the set of k-free integers in $\{n \in \mathbb{N} : p|n \Rightarrow p \in A_N\}$.

We will prove the theorem using the following equation, whose proof will be given at the end.

$$P(I_{N;A,3} = n) = \Big(\prod_{j=1}^N (1 - \frac{1}{p_{j;A}})\Big) \frac{1}{n_{\{(k-1)-\text{free}\}}\phi(n_{\{(k-1)-\text{power}\}})}. \tag{3.1}$$

We have

$$\log I_{N;A,3} = \sum_{j=1}^N V_j \log p_{j;A}.$$

It is easy to check that as with ET_j and EU_j, we have

$$EV_j \sim \frac{1}{p_{j;A}}, \text{ as } j \to \infty. \tag{3.2}$$

From (2.3) and (3.2), we have

$$E \log I_{N;A,3} \sim \sum_{j=1}^N \frac{\log p_{j;A}}{p_{j;A}} \sim \theta \log N, \text{ as } N \to \infty. \tag{3.3}$$

As in the proof of Theorem 1, we define mutually independent random variables $\{B_j\}_{j=1}^\infty$ and $\{Xj\}_{j=1}^\infty$:
$B_j \sim \text{Ber}(\frac{1}{p_{j;A}})$; that is,

$$q_j := P(B_j = 1) = 1 - P(B_j = 0) = \frac{1}{p_{j;A}}. \tag{3.4}$$

$X_j \stackrel{\text{dist}}{=} \log p_{j;A} \cdot V_j | \{V_j \geq 1\}$; that is

$$P(X_j = m \log p_{j;A}) = P(V_j = m | V_j \geq 1) = \begin{cases} (1 - \frac{1}{p_{j;A}})(\frac{1}{p_{j;A}})^{m-1}, \ m = 1, 2, \cdots, k-2. \\ (\frac{1}{p_{j:A}})^{k-2}, \ m = k-1. \end{cases}$$

As in the proof of Theorem 1, we have $\mu_j := EX_j \sim \log p_{j;A}$ and $\lim_{j\to\infty} \frac{X_j}{\mu_j} \stackrel{\text{dist}}{=} 1$. By the construction of $\{B_j\}_{j=1}^{\infty}$ and $\{X_j\}_{j=1}^{\infty}$, we have

$$\log I_{N;A,3} = \sum_{j=1}^{N} V_j \log p_{j;A} \stackrel{\text{dist}}{=} \sum_{j=1}^{N} B_j X_j.$$

By the same considerations as in the proof of parts (i) and (ii) of Theorem 1, it follows that $W_N := \frac{\sum_{j=1}^N B_j X_j}{E\sum_{j=1}^N B_j X_j}$ converges in distribution to $\frac{1}{\theta}D_\theta$; thus,

$$\lim_{N\to\infty} \frac{\log I_{N;A,3}}{E\log I_{N;A,3}} \stackrel{\text{dist}}{=} \frac{1}{\theta}D_\theta. \tag{3.5}$$

On the one hand, just as in (2.7) and (2.22), by the convergence in distribution in (3.5) and the fact that the limiting distribution is a continuous one, for any sequence $\{\theta_N\}_{N=1}^{\infty}$ satisfying $\lim_{N\to\infty}\theta_N = \theta$, we have

$$\lim_{N\to\infty} P(\frac{\log I_{N;A,3}}{E\log I_{N;A,3}} \le \frac{1}{\theta_N}) = P(\frac{1}{\theta}D_\theta \le \frac{1}{\theta}) = \frac{e^{-\gamma\theta}}{\Gamma(\theta+1)}. \tag{3.6}$$

On the other hand, let $\theta_N := \frac{E\log I_{N;A,3}}{\log p_{N;A}}$ and note from (2.3) and (3.3) that $\lim_{N\to\infty}\theta_N = \theta$. It follows from (3.1) that

$$P(\tfrac{\log I_{N;A,3}}{E\log I_{N;A,3}} \le \tfrac{1}{\theta_N}) = P\big(I_{N;A,3} \le \exp(\tfrac{E\log I_{N;A,3}}{\theta_N})\big) = P(I_{N;A,3} \le p_{N;A})$$

$$= \textstyle\prod_{j=1}^N (1 - \frac{1}{p_{j;A}}) \sum_{n\le p_{N;A}: p|n\Rightarrow p\in A}^{\prime(k)} \frac{1}{n_{\{(k-1)-\text{free}\}}\phi(n_{\{(k-1)-\text{power}\}})}$$

$$= \frac{\sum_{n\le p_{N;A}: p|n\Rightarrow p\in A}^{\prime(k)} \frac{1}{n_{\{(k-1)-\text{free}\}}\phi(n_{\{(k-1)-\text{power}\}})}}{\sum_{n:p|n\Rightarrow p\in A_N} \frac{1}{n}}.$$

From this and (3.6) it follows that

$$\lim_{N\to\infty} \frac{\sum_{n:p|n\Rightarrow p\in A_N} \frac{1}{n}}{\sum_{n\le p_{N;A}: p|n\Rightarrow p\in A}^{\prime(k)} \frac{1}{n_{\{(k-1)-\text{free}\}}\phi(n_{\{(k-1)-\text{power}\}})}} = e^{\gamma\theta}\Gamma(\theta+1),$$

which is equivalent to (1.11) just as (2.9) and (2.23) were equivalent to parts (i) and (ii) respectively of Theorem 1 □

We now return to the proof of (3.1).

Let f, g be functions satisfying

$$f(j) = 1,\ j = 0,\cdots,k-2;\quad f(k-1) = 0;$$

$$g(j) = j+1,\ j = 0,\cdots,k-2;\quad g(k-1) = k-1.$$

For definiteness, we take

$$f(x)=1-\binom{x}{k-1},\qquad g(x)=(x+1)-\binom{x}{k-1}.$$

Then for arbitrary $n=\prod_{j=1}^N p_{j;A}^{c_j}$ in the support,

$$P(I_{N;A,3}=n)=\prod_{j=1}^N P(V_j=c_j)=\prod_{j=1}^N\frac{(p_{j:A}-1)^{f(c_j)}}{p_{j;A}^{g(c_j)}}. \tag{3.7}$$

Noting that $g(x)-f(x)\equiv x$, we rewrite the right hand side of (3.7) as

$$\begin{aligned}
&\prod_{j=1}^N\frac{(p_{j:A}-1)^{f(c_j)}}{p_{j;A}^{g(c_j)}}=\frac{\prod_{j=1}^N\big(\frac{p_{j:A}-1}{p_{j;A}}\big)^{f(c_j)}}{\prod_{j=1}^N p_{j;A}^{g(c_j)-f(c_j)}}\\
&=\frac{\prod_{j=1}^N\big(1-\frac{1}{p_{j;A}}\big)^{f(c_j)}}{\prod_{j=1}^N p_{j;A}^{c_j}}=\frac{\prod_{j=1}^N(1-\frac{1}{p_{j;A}})}{\prod_{j:c_j=k-1}(1-\frac{1}{p_{j;A}})}\times\frac{1}{n}\\
&=\Big(\prod_{j=1}^N(1-\frac{1}{p_{j;A}})\Big)\times\frac{1}{n_{\{(k-1)-\text{free}\}}}\times\frac{1}{n_{\{(k-1)-\text{power}\}}\prod_{j:c_j=k-1}(1-\frac{1}{p_{j;A}})}\\
&=\Big(\prod_{j=1}^N(1-\frac{1}{p_{j;A}})\Big)\times\frac{1}{n_{\{(k-1)-\text{free}\}}}\times\frac{1}{n_{\{(k-1)-\text{power}\}}\prod_{j:p_{j;A}|n_{\{(k-1)-\text{power}\}}}(1-\frac{1}{p_{j;A}})}\\
&=\Big(\prod_{j=1}^N(1-\frac{1}{p_{j;A}})\Big)\times\frac{1}{n_{\{(k-1)-\text{free}\}}}\times\frac{1}{\phi(n_{\{(k-1)-\text{power}\}})}.
\end{aligned}\tag{3.8}$$

Now (3.1) follows from (3.7) and (3.8).

References

[1] Languasco, A. and Zaccagnini, A., *A note on Mertens' formula for arithmetic progressions*, J. Number Theory **127** (2007), 37-46.

[2] Languasco, A. and Zaccagnini, A., *On the constant in the Mertens product for arithmetic progressions. I*, Identities. Funct. Approx. Comment. Math. **42** (2010), 17-27.

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

[4] Pinsky, R. G., *On the strange domain of attraction to generalized Dickman distribution for sums of independent random variables*, Electron. J. Probab. **23**, Paper No. 3, 17 pp. (2018).

[5] Pinsky, R. G., *A natural probabilistic model on the integers and its relation to Dickman-type distributions and Buchstab's function*, Probability and Analysis in Interacting Physical Systems, Springer Proc. Math. Stat., **283**, Springer, Cham, (2019) 267-294.

[6] Rieger, G., *Zahlentheoretische Anwendung eines Taubersatzes mit Restglied.* (German) Math. Ann. **182** (1969), 243-248.

[7] Shapiro, H., *Introduction to the Theory of Numbers*, John Wiley & Sons, Inc., New York, (1983).

[8] Ward, D., *Some series involving Euler's function*, J. London Math. Soc. **2** (1927), 210-214.

[9] Williams, K., *Mertens' theorem for arithmetic progressions*, J. Number Theory **6** (1974), 353-359.

[10] Wirsing, E., *ber die Zahlen, deren Primteiler einer gegebenen Menge angehren*, Arch. Math. **7** (1956), 263-272.

Department of Mathematics, Technion—Israel Institute of Technology, Haifa, 32000, Israel

E-mail address: pinsky@math.technion.ac.il

URL: http://www.math.technion.ac.il/~pinsky/