ON PROPERTIES OF THE RIEMANN ZETA DISTRIBUTION

MICHAEL CRANSTON AND ADRIEN PELTZER

1. Abstract

In this paper we examine various properties of positive integers selected according to the Riemann zeta distribution. That is, if $\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s}$, $s > 1$, then we consider the random variable X_s with $P(X_s = n) = \frac{1}{\zeta(s)n^s}, n \geq 1$. We derive various results such as the analog of the Erdös-Kac Central Limit Theorem (CLT) for the number of distinct prime factors, $\omega(X_s)$, of X_s, as $s \searrow 1$, large and moderate deviations for $\omega(X_s)$, a Berry-Esseen result and analogs of Halberstam's CLT for additive functions evaluated at X_s. In addition, we prove analogs of Erdös-Delange type formulas for expectations of additive and multiplicative functions evaluated at X_s. We also examine various applications using Dirichlet series and their associated distributions. Finally, we show how to derive asymptotic distributional results for an integer selected uniformly at random from $[N] = \{1, 2, \cdots, N\}$ or according to the harmonic distribution from their analogous asymptotic results for X_s as $s \searrow 1$.

2. Introduction

In this paper we will examine properties of integers selected according to the Riemann zeta distribution. We will emphasize two aspects of this distribution. The first is its faithful similarity to properties of an integer chosen according to the uniform distribution on a finite interval. The second aspect will be the appearance of Poisson behavior under this distribution. The Riemann zeta function is given for $\operatorname{Re} z > 1$ by

$$\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z}.$$

An alternative description is given by

$$\zeta(z) = \Pi_p \left(1 - \frac{1}{p^z}\right)^{-1},$$

where Π_p and $\sum_p$ will be used to denote the product and sum over primes p, respectively. We shall denote the set of primes by $\mathcal{P}$. Sometimes we will have sums over p, q or $p \neq q$ and it will be understood that p and q are primes. In our discussions we will replace the complex z by a real number $s > 1$. We will denote by X_s a random variable with the distribution

$$P(X_s = n) = \frac{1}{\zeta(s)n^s}, \ n = 1, 2, 3, \cdots. \tag{1}$$

We say X_s has the Riemann zeta distribution or the $\zeta(s)$ distribution for short. Another method for sampling an integer at random is to first take a large integer

N and then select an integer at random in $[N] = \{1, 2, \cdots, N\}$. That is, define a random variable Y_N by stipulating that its distribution is given by

$$P(Y_N = n) = \frac{1}{N},\ n \in [N]. \tag{2}$$

A third method is to use the harmonic distribution. In this case we again take a large integer N and define $a_N = \sum_{n=1}^{N} \frac{1}{n}$ and then stipulating the distribution of Z_N is given by

$$P(Z_N = n) = \frac{1}{a_N n},\ n \in [N]. \tag{3}$$

The random variable Z_N is said to have the harmonic distribution. It will be seen later that by means of Tauberian theorems that asymptotic results valid for X_s as $s \searrow 1$ can be transferred to results for Y_N and Z_N as $N \nearrow \infty$. We make a couple of remarks that will be useful later. First, the asymptotic behavior of $\zeta(s)$ as $s \searrow 1$ is well known and is given by

$$\zeta(s) \sim \frac{1}{s-1},\ s \searrow 1$$

and

$$a_N \sim \log N,\ N \nearrow \infty.$$

A function that comes up often in this paper is the prime zeta function,

$$\mathbb{P}(s) \equiv \sum_p \frac{1}{p^s}.$$

By elementary comparison with the series $\sum_{n=1}^{\infty} \frac{1}{n^s}$ it's clear that $\mathbb{P}(s) < \infty$ for $s > 1$. Also, $\lim_{s \searrow 1} \mathbb{P}(s) = \sum_{p \in \mathcal{P}} \frac{1}{p} = \infty$ is a well-known result due to Euler who made the following observation:

$$\begin{aligned} \log \zeta(s) = & -\sum_p \log(1 - \tfrac{1}{p^s}) \\ = & \sum_p \sum_{k \geq 1} \tfrac{1}{kp^{sk}} \\ = & \sum_{k \geq 1} \sum_p \tfrac{1}{kp^{sk}} \\ = & \sum_{k \geq 1} \tfrac{\mathbb{P}(sk)}{k} \\ = & \mathbb{P}(s) + \sum_{k \geq 2} \tfrac{\mathbb{P}(sk)}{k}. \end{aligned}$$

This implies by a simple comparison argument that

$$\mathbb{P}(s) = \log \zeta(s) + O(1). \tag{4}$$

The last line implies $\sum_{p \in \mathcal{P}} \frac{1}{p} = \infty$ since $\lim_{s \searrow 1} \log \zeta(s) = \infty$. For comparison, a famous theorem of Merten's ([T], Thm 1.10) gives the truncated function $\sum_{p \leq N} \frac{1}{p} = \log \log N + O(1)$ as $N \nearrow \infty$.

An aspect that arises from the present work is the perfect agreement between statistical behavior of X_s as $s \searrow 1$ and Y_N or Z_N as $N \nearrow \infty$. Moreover, by Tauberian theorems given later, one can transfer asymptotic results about the Riemann zeta random variable X_s for $s \searrow 1$ to the uniformly distributed Y_N or the harmonic distributed Z_N as $N \nearrow \infty$. This can be interpreted as a robustness property of integers, in a certain sense, their statistical behavior doesn't depend so much on what distribution is used when they are sampled. We don't know how far this can be extended to other distributions but these three distributions have an important property in common,

$$\lim_{s\searrow 1} P(p \text{ divides } X_s) = \lim_{N\to\infty} P(p \text{ divides } Y_N) = \lim_{N\to\infty} P(p \text{ divides } Z_N) = \frac{1}{p}.$$

The organization of this paper is as follows: Section 3 introduces the basic properties of X_s, Section 4 gives Delange type formulas for computing expectations of multiplicative functions of X_s. In Section 5 we review results of Lloyd,[L] and give some applications, Section 6 contains results about the gcd of independent copies of $\zeta(s)$ random variables and applications, Section 7 has the analogs of the Erdös-Kac and Halberstam CLTs for $\omega(X_s)$, the number of distinct prime factors of X_s, Section 7 gives the Berry-Esseen results for the CLT, Section 9 examines large and moderate deviations for $\omega(X_s)$, Section 10 introduces Dirichlet series and various probabilistic applications, and Section 11 examines how to transfer asymptotic results for X_s to asymptotic results for Y_N and Z_N.

3. Basic Properties of X_s

Many of the basic properties of X_s have been discussed in a variety of places. We recall some of them here with no claim for novelty. A good source for interesting aspects of the $\zeta(s)$ distribution is the paper of Gut [G]. A word about notation, we will denote the event $\{p \text{ divides } X_s\}$ by $\{p|X_s\}$, not to be confused with conditioning. First, for $n \in \mathbf{Z}^+$,

$$\begin{aligned} P(n|X_s) &= \sum_{k=1}^{\infty} P(X_s = nk) \\ &= \frac{1}{\zeta(s)} \sum_{k=1}^{\infty} \frac{1}{(nk)^s} \\ &= \frac{1}{n^s \zeta(s)} \sum_{k=1}^{\infty} \frac{1}{k^s} \\ &= \frac{1}{n^s}. \end{aligned} \tag{5}$$

When one takes a prime p, then $\lim_{s\searrow 1} P(p|X_s) = \frac{1}{p}$ which is appropriate since every p^{th} integer is divisible by p. This is a property shared by the uniform and harmonic distributions as $N \to \infty$. From 5 follows the important property that if $gcd(m,n) = 1$ then

$$\begin{aligned} P(m|X_s, n|X_s) &= P(mn|X_s) \\ &= \frac{1}{m^s} \frac{1}{n^s} \\ &= P(m|X_s) P(n|X_s) \end{aligned} \tag{6}$$

which says the events $\{m|X_s\}$ and $\{n|X_s\}$ are independent when $gcd(m,n) = 1$. In particular this holds when $m = p$ and $n = q$ where p and q are distinct primes. Another consequence of 5 is the independence and exact distribution of the exponents in the prime factorization of X_s. To make this precise, factor X_s,

$$X_s = \Pi_p p^{c_p(s)} \tag{7}$$

where the $c_p(s) : p \in \mathcal{P}$ are nonnegative integer valued random variables. By 6 if $p_1, \cdots, p_k$ are distinct primes then $c_{p_1}(s), \cdots, c_{p_k}(s)$ are independent random variables and by 5

$$P(c_p(s) \geq k) = \frac{1}{p^{sk}},\ k \in \mathbf{N} \tag{8}$$

or what is the same,

$$P(c_p(s) = k) = \frac{1}{p^{sk}}\left(1 - \frac{1}{p^s}\right). \tag{9}$$

An easy computation reveals

$$E[e^{tc_p(s)}] = \frac{1 - \frac{1}{p^s}}{1 - \frac{e^t}{p^s}}. \tag{10}$$

To summarize, the random variables $\{c_p(s) : p \in \mathcal{P}\}$ are independent and $c_p(s)$ has a geometric distribution with parameter $\frac{1}{p^s}$. We now introduce some common functions in number theory.

An observation of Khintchin, [K], is that $\frac{\zeta(s+it)}{\zeta(s)}$ is the characteristic function of a random variable, namely, $-\log X_s$. This follows from the easy computation

$$E[e^{-it\log X_s}] = E[X_s^{-it}] \tag{11}$$

$$= \frac{1}{\zeta(s)}\sum_{n\geq 1}\frac{n^{-it}}{n^s} \tag{12}$$

$$= \frac{\zeta(s+it)}{\zeta(s)}. \tag{13}$$

From this observation we derive the following proposition which gives an idea about the magnitude of X_s as $s \searrow 1$.

Proposition 3.1. Let $\mathcal{E}$ be an exponential random variable with parameter 1. Then as $s \searrow 1$,

$$\frac{1}{\zeta(s)}\log X_s \xrightarrow{d} \mathcal{E}.$$

Proof. Following our computation of the characteristic function of $-\log X_s$ we get as $s \searrow 1$,

$$\begin{aligned} E[e^{t\frac{1}{\zeta(s)}\log X_s}] &= \frac{\zeta(s-\frac{t}{\zeta(s)})}{\zeta(s)} \\ &\sim \frac{s-1}{s-\frac{t}{\zeta(s)}-1} \\ &\to \frac{1}{1-t}, \end{aligned}$$

which is the moment generating function of $\mathcal{E}$.

□

One way to reveal properties of X_s is to check the statistical properties of arithmetic functions of X_s. We now discuss a few of these. The number of distinct prime divisors of X_s is denoted by

$$\omega(X_s) = \sum_p 1_{c_p(s)>0} \tag{14}$$

Notice, $\omega(X_s)$ is the sum of independent Bernoulli random variables. The total number of prime divisors of X_s is denoted by

$$\Omega(X_s) = \sum_p c_p(s). \tag{15}$$

The Euler totient function $\phi(n)$ counts the number of positive integers less than n which are relatively prime to n. That is,

$$\phi(n) = \#\{m < n : gcd(m, n) = 1\}.$$

There is a simple well known formula for this, namely

$$\phi(n) = n\Pi_{p|n}(1 - \frac{1}{p}).$$

Evaluating at $n = X_s$ we get

$$\phi(X_s) = X_s\Pi_p(1 - \frac{1_{c_p(s)>0}}{p}) \tag{16}$$

A couple of simple observations are in order.

Theorem 3.2. Let $\omega(n)$ and $\Omega(n)$ denote the prime counting functions as in 14 and 15. Then

$$E\omega(X_s) = \mathbb{P}(s). \tag{17}$$

$$\text{var}(\omega(X_s)) = \mathbb{P}(s) - \mathbb{P}(2s) \tag{18}$$

$$E[\Omega(X_s)] = \sum_p \frac{1}{p^s - 1} \tag{19}$$

$$E[\phi(X_s)] = \frac{\zeta(s-1)}{\zeta^2(s)} \tag{20}$$

Proof. First, by 9

$$\begin{aligned} E\omega(X_s) &= \sum_p E1_{c_p(s)>0} \\ &= \sum_p \frac{1}{p^s} \\ &\equiv \mathbb{P}(s). \end{aligned} \tag{21}$$

Second, by independence and 9,

$$
\begin{aligned}
E\omega(X_s)^2 &= E\left(\sum_p 1_{c_p(s)>0}\right)^2 \\
&= E\left(\sum_{p,q} 1_{c_p(s)>0}1_{c_q(s)>0}\right) \\
&= E\left(\sum_{p\neq q} 1_{c_p(s)>0}1_{c_q(s)>0}\right) + E\left(\sum_p 1_{c_p(s)>0}\right) \qquad (22)\\
&= \sum_{p\neq q}\frac{1}{p^sq^s} + \sum_p \frac{1}{p^s} \\
&= \left(\sum_p \frac{1}{p^s}\right)^2 - \sum_p \frac{1}{p^{2s}} + \sum_p \frac{1}{p^s} \\
&= \mathbb{P}(s)^2 - \mathbb{P}(2s) + \mathbb{P}(s).
\end{aligned}
$$

From 17 and 22 one derives

$$
\operatorname{var}(\omega(X_s)) = \mathbb{P}(s) - \mathbb{P}(2s). \qquad (23)
$$

As for $\Omega(X_s)$ we have

$$
\begin{aligned}
E[\Omega(X_s)] &= \sum_p E[c_p(s)] \\
&= \sum_p \sum_{k\geq 1} P(c_p(s)\geq k) \qquad (24)\\
&= \sum_p \sum_{k\geq 1} \frac{1}{p^{sk}} \\
&= \sum_p \frac{1}{p^s - 1}.
\end{aligned}
$$

To compute the expected value of $\phi(X_s)$ we rely on a couple of facts from number theory. First, define the convolution of two arithmetic functions f and g by $f\star g = \sum_{d|n} f(\frac{n}{d})g(d)$. There is a nice formula for the product of two Dirichlet series using this convolution, namely, $\sum_{n\geq 1}\frac{f(n)}{n^s}\cdot\sum_{n\geq 1}\frac{g(n)}{n^s} = \sum_{n\geq 1}\frac{f\star g(n)}{n^s}$. Finally a formula of Gauss ([Ga]) gives $\sum_{d|n}\phi(d) = n$. Then, using these facts with $f\equiv 1$ and $g(n)=\phi(n)$, one gets the following formula,

$$
\begin{aligned}
\zeta(s)E[\phi(X_s)] &= \frac{1}{\zeta(s)}\sum_{n\geq 1}\frac{1}{n^s}\sum_{n\geq 1}\frac{\phi(n)}{n^s} \\
&= \frac{1}{\zeta(s)}\sum_{n\geq 1}\frac{1\star\phi(n)}{n^s} \\
&= \frac{1}{\zeta(s)}\sum_{n\geq 1}\frac{\sum_{d|n}\phi(d)}{n^s} \\
&= \frac{1}{\zeta(s)}\sum_{n\geq 1}\frac{n}{n^s} \\
&= \frac{\zeta(s-1)}{\zeta(s)},
\end{aligned}
$$

and dividing both sides by $\zeta(s)$ we get for $s > 2$,

$$E[\phi(X_s)] = \frac{\zeta(s-1)}{\zeta^2(s)}. \tag{25}$$

We can also compute $E[\sigma_0(s)]$, where $\sigma_a(a) = \sum_{d|n} d^a$ is the divisor function. However, we will save this for later when we'll have better techniques for the computation. □

We observe that $\lim_{s\searrow 1} E[\Omega(X_s) - \omega(X_s)] = \sum_p \frac{1}{p(p-1)}$, which indicates that $\Omega(X_s)$ isn't much larger than $\omega(X_s)$. In the next section we'll see that $\Omega(X_s) - \omega(X_s)$ has a non-degenerate limiting distribution.

4. Delange/Erdos-Wintner Theorem

In this section we consider $\zeta(s)$ distribution analogs of results of Delange and Erdös-Wintner established for the uniform distribution on $[N]$. A good reference for these results is [T]. We derive the limiting distribution and limiting mean of arithmetic functions of the $\zeta(s)$ random variable as $s \searrow 1$. It is standard to declare a function $f : \mathbf{N} \to \mathbf{R}$ to be multiplicative if $f(mn) = f(m)f(n)$ whenever $gcd(m,n) = 1$ and additive if $f(mn) = f(m) + f(n)$ whenever $gcd(m,n) = 1$. There are many examples of such functions. A common example of a multiplicative function is the Euler totient function $\phi(n)$ defined above.

An example of an additive function is $\omega(n)$. We say f is completely multiplicative if $f(mn) = f(m)f(n)$ for all m, n and completely additive if $f(mn) = f(m) + f(n)$ for all m, n. If the function f satisfies $f(p^m) = f(p)^m$ for all primes p and nonnegative integers m, and for $n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$ one defines $f(m) = \Pi_{i=1}^k f(p_i)^{a_i}$ then f is completely multiplicative. The function Ω is completely additive.

By taking logarithms of multiplicative functions one gets additive functions. Thus, we see that $g(m) = \log \frac{\phi(m)}{m}$ is an additive function. This follows from the formula $\frac{\phi(m)}{m} = \Pi_{p|m}(1 - \frac{1}{p})$. A Theorem of Delange gives a formula for the average of a multiplicative function with respect to the uniform measure on $[N]$.

Theorem 4.1. (Delange, Wirsing, Halasz [E], Thm 6.3, [T], Thm 4.4) If $g : \mathbf{N} \to \{z \in \mathbf{C}, |z| \le 1\}$ is multiplicative and $\sum_p (1 - g(p))/p$ converges, then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{x \le N} g(x) = \Pi_p (1 - \frac{1}{p}) \sum_{m \ge 0} \frac{g(p^m)}{p^m}.$$

Similar to this is a result of Erdös and Wintner, [T], on the asymptotic distribution of additive functions under the uniform distribution.

Theorem 4.2. (Erdös, Wintner [T], Thm 4.1) A necessary and sufficient condition that an additive function have a limiting distribution is that the following three series converge for at least one value of $R > 0$,

- $\sum_{|f(p)| > R} \frac{1}{p}$
- $\sum_{|f(p)| \le R} \frac{f(p)^2}{p}$
- $\sum_{|f(p)| \le R} \frac{f(p)}{p}$

When these three conditions are satisfied then the characteristic function of the limiting distribution is given by

$$\psi(\tau) = \Pi_p(1 - \frac{1}{p}) \sum_{m\geq 0} \frac{e^{i\tau f(p^m)}}{p^m}.$$

The analog's of these results for the Riemann zeta distribution hold exactly for all $s > 1$ and follow easily from the independence of the sequence $\{c_p(s) : p \in \mathcal{P}\}$.

Theorem 4.3. Let g be a multiplicative function with values in $\{z \in \mathbf{C}, |z| \leq 1\}$, and $s > 1$. Then

$$E\left[g(X_s)\right] = \Pi_p\left(\left(1 - \frac{1}{p^s}\right)\sum_{m\geq 0}\frac{g(p^m)}{p^{ms}}\right). \tag{26}$$

Assume f is additive. Then for $\tau \in \mathbb{R}$,

$$E\left[e^{i\tau f(X_s)}\right] = \Pi_p(1 - \frac{1}{p^s})\sum_{m\geq 0}\left(\frac{e^{i\tau f(p^m)}}{p^{ms}}\right). \tag{27}$$

If, in addition, we also assume that $f(p^a) = f(p)$ for all $a \geq 1$, then

$$E\left[e^{i\tau f(X_s)}\right] = \Pi_p\left(1 + (1 - e^{i\tau f(p)})\frac{1}{p^s})\right). \tag{28}$$

Proof. Given a multiplicative g with values in the unit disc, by the independence of the $c_p(s)$'s, we have

$$\begin{aligned} E\left[g(X_s)\right] &= E\left[g(\Pi_p p^{c_p(s)})\right] \\ &= \Pi_p E\left[g(p^{c_p(s)})\right] \\ &= \Pi_p\left(\sum_{m\geq 0} g(p^m)(1 - \frac{1}{p^s})\frac{1}{p^{ms}}\right) \\ &= \Pi_p\left(\left(1 - \frac{1}{p^s}\right)\sum_{m\geq 0}\frac{g(p^m)}{p^{ms}}\right) \end{aligned} \tag{29}$$

We can now apply 29 to compute the characteristic function of an additive function. So, let f denote an additive function and set $g(m) = e^{i\tau f(m)}$. First we point out that $g(1) = 1$. Then, by dint of the fact that f is additive, it follows that g is multiplicative. Thus 29 applies to give the following expression for the characteristic function of $f(X_s)$,

$$E\left[e^{i\tau f(X_s)}\right] = \Pi_p(1 - \frac{1}{p^s})\sum_{m\geq 0}\left(\frac{e^{i\tau f(p^m)}}{p^{ms}}\right). \tag{30}$$

Then, if also $f(p^a) = f(p)$, $a \geq 1$ and keeping in mind that $f(1) = 0$, then 30 becomes,

$$\begin{aligned} E\left[e^{i\tau f(X_s)}\right] &= \Pi_p(1-\tfrac{1}{p^s})\textstyle\sum_{m\geq 0}\left(\frac{e^{i\tau f(p^m)}}{p^{ms}}\right) \\ &= \Pi_p\left((1-\tfrac{1}{p^s})(1+e^{i\tau f(p)}\textstyle\sum_{m\geq 1}\frac{1}{p^{ms}}\right) \\ &= \Pi_p\left((1-\tfrac{1}{p^s})(1+e^{i\tau f(p)}\tfrac{1}{p^s}(1-\tfrac{1}{p^s})^{-1})\right) \\ &= \Pi_p\left(1+(e^{i\tau f(p)}-1)\tfrac{1}{p^s}\right). \end{aligned}$$

□

We now apply these results to the multiplicative function $g(m) = \frac{\phi(m)}{m}$.

Proposition 4.4. For the Euler totient function,

$$E\left[\frac{\phi(X_s)}{X_s}\right] = \Pi_p\left(1-\frac{1}{p^{s+1}}\right)$$

and

$$\lim_{s\searrow 1} E\left[\frac{\phi(X_s)}{X_s}\right] = \frac{6}{\pi^2}.$$

In addition, $\frac{\phi(X_s)}{X_s}$ has a limiting distribution at $s \searrow 1$.

Proof. Since $g(m) = \frac{\phi(m)}{m}$ satisfies $g(p^m) = 1-\frac{1}{p}$, for $m \geq 1$ and $g(1) = 1$, from 29 we get

$$\begin{aligned} E\left[\frac{\phi(X_s)}{X_s}\right] &= \Pi_p\left(1-\tfrac{1}{p^s}\right)\left(\textstyle\sum_{m\geq 0}\frac{g(p^m)}{p^{sm}}\right) \\ &= \Pi_p\left(1-\tfrac{1}{p^s}\right)\left(1+(1-\tfrac{1}{p})\textstyle\sum_{m\geq 1}\frac{1}{p^{sm}}\right) \\ &= \Pi_p\left(1-\tfrac{1}{p^s}\right)\left(1+(1-\tfrac{1}{p})((1-\tfrac{1}{p^s})^{-1}-1)\right) \\ &= \Pi_p\left(1-\tfrac{1}{p^s}+(1-\tfrac{1}{p})\tfrac{1}{p^s}\right) \\ &= \Pi_p\left(1-\tfrac{1}{p^{s+1}}\right). \end{aligned}$$

Thus,

$$\begin{aligned} \lim_{s\searrow 1} E\left[\frac{\phi(X_s)}{X_s}\right] &= \Pi_p\left(1-\tfrac{1}{p^2}\right) \\ &= \tfrac{1}{\zeta(2)} \\ &= \tfrac{6}{\pi^2}. \end{aligned}$$

From 27, we get the following characteristic function for $f(m) = \log g(m)$,

$$
\begin{aligned}
\lim_{s\searrow 1} E\left[e^{i\tau \log \frac{\phi(X_s)}{X_s}}\right] =&\Pi_p(1-\frac{1}{p})\sum_{m\geq 0}\left(\frac{e^{i\tau f(p^m)}}{p^m}\right)\\
=&\Pi_p(1-\frac{1}{p})\sum_{m\geq 0}\left(\frac{g^{i\tau}(p^m)}{p^m}\right)\\
=&\Pi_p(1-\frac{1}{p})\left(1+\sum_{m\geq 1}\left(\frac{(1-\frac{1}{p})^{i\tau}}{p^m}\right)\right)\\
=&\Pi_p(1-\frac{1}{p})\left(1+(1-\frac{1}{p})^{i\tau}\sum_{m\geq 1}\frac{1}{p^m}\right)\\
=&\Pi_p(1-\frac{1}{p})\left(1+(1-\frac{1}{p})^{i\tau}\left((1-\frac{1}{p})^{-1}-1\right)\right)\\
\equiv \psi(\tau).
\end{aligned} \tag{31}
$$

We observe that

$$
\begin{aligned}
\psi(0) = & \quad \Pi_p(1-\tfrac{1}{p})\left(1+\left((1-\tfrac{1}{p})^{-1}-1\right)\right)\\
= & \quad \Pi_p(1-\tfrac{1}{p})\left(1+\left((1-\tfrac{1}{p})^{-1}-1\right)\right)\\
= & \quad 1.
\end{aligned}
$$

And by a simple estimate,

$$\lim_{\tau\to 0}\psi(\tau) = 1.$$

Thus, by Lévy's Continuity Theorem ([D], Thm 3.3.6), $\log g(X_s)$ and therefore $g(X_s)$ has a limiting distribution. The characteristic function of the limiting distribution of $\log g(X_s)$ is given by

$$\psi(\tau) = \Pi_p\left((1-\frac{1}{p}) + (1-\frac{1}{p})^{i\tau}\frac{1}{p})\right).$$

□

There are also revealing expressions for the moment generating functions of $\omega(X_s)$, $\Omega(X_s)$, and $\Omega(X_s) - \omega(X_s)$. They are derived using independence of the $c_p(s)$ and their distribution as geometric random variables. A weighted variant of ω that has appeared elsewhere, say in Halberstam [H], is obtained as follows.

Definition 1. (strongly additive functions) Take a function $f : \mathbb{P} \to \mathbf{R}$ and extend f by declaring that $f(p^a) = f(p)$ for $a \geq 1$ and for arbitrary $m \geq 1$, set

$$f(m) = \sum_{p|m} f(p).$$

Then for $gcd(m,n) = 1$, $f(mn) = f(m)+f(n)$, that is f is an additive function. We call an additive function satisfying $f(p^a) = f(p)$ for $a \geq 1$ and primes p a strongly additive function.

When $f \equiv 1$ we just get ω back and for $f \not\equiv 1$ we get a weighted version of ω. A specific example of this would be to set $f(p) = 1_{p\equiv 1 \mod 4}$ in which case $f(m)$ would be the number of distinct prime divisors of m which are congruent to 1 mod 4.

Theorem 4.5. For $t < \ln 1$ and $s > 1$, if f is a bounded strongly additive function such that $f(p^m) = f(p),\ p \in \mathcal{P},\ m \geq 1$, then

$$Ee^{tf(X_s)} = \exp\{\textstyle\sum_p \frac{(e^{tf(p)}-1)}{p^s}\}\exp\{-\sum_{m=2}^{\infty}\sum_p \frac{(1-e^{tf(p)})^m}{mp^{ms}})\}, \tag{32}$$

and

$$Ee^{t\omega(X_s)} = \exp\{(e^t-1)\mathbb{P}(s)\}\exp\{-\textstyle\sum_{m\geq 2}\frac{(1-e^t)^m}{m}\mathbb{P}(ms)\}. \tag{33}$$

For $t < 0$,

$$Ee^{t\Omega(X_s)} = \exp\left\{\left(e^t-1\right)\mathbb{P}(s)\right\}\exp\left\{\textstyle\sum_{m\geq 2}\left(\frac{e^{kt}-1}{k}\right)\mathbb{P}(ks)\right\}. \tag{34}$$

For $t < \ln 2$

$$Ee^{t(\Omega(X_s)-\omega(X_s))} = \Pi_p\left(1-\tfrac{1}{p^s}\right)\left(1+\tfrac{1}{p^s-e^t}\right). \tag{35}$$

Proof. One has $f(p^m) = f(p)$ for $m \geq 1,\ p \in \mathcal{P}$ and $f(1) = 0$, so on replacing $i\tau$ with t by 27, we have by starting with 28,

$$\begin{aligned} Ee^{tf(X_s)} &= \Pi_p\left((e^{tf(p)}-1)\tfrac{1}{p^s}+1\right) \\ &= \exp\{\textstyle\sum_p \log\left((e^{tf(p)}-1)\frac{1}{p^s}+1\right)\} \\ &= \exp\{\textstyle\sum_p\sum_{m=1}^{\infty}(-1)^{m+1}\frac{(e^{tf(p)}-1)^m}{m}\frac{1}{p^{ms}}\} \\ &= \exp\{\textstyle\sum_{m=1}^{\infty}(-1)^{m+1}\sum_p\frac{(e^{tf(p)}-1)^m}{m}\frac{1}{p^{ms}}\} \\ &= \exp\{\textstyle\sum_p\frac{(e^{tf(p)}-1)}{p^s}\}\exp\{-\sum_{m=2}^{\infty}\sum_p\frac{(1-e^{tf(p)})^m}{mp^{ms}}\}. \end{aligned}$$

In the special case $f(p) \equiv 1$ this gives

$$Ee^{t\omega(X_s)} = \exp\{(e^t-1)\mathbb{P}(s)\}\exp\{-\sum_{m=2}^{\infty}\frac{(1-e^t)^m}{m}\mathbb{P}(ms)\}.$$

In order to compute the Laplace transform of $\Omega(X_s)$ we observe that $\Omega(p^m) = m$ for $m \geq 0$ so by 27 we get, for $t > 0$,

$$\begin{aligned}
Ee^{-t\Omega(X_s)} &= \Pi_p(1-\tfrac{1}{p^s})\textstyle\sum_{m\geq 0}\frac{e^{-t\Omega(p^m)}}{p^{sm}} \\
&= \Pi_p(1-\tfrac{1}{p^s})\textstyle\sum_{m\geq 0}\frac{e^{-tm}}{p^{sm}} \\
&= \Pi_p(1-\tfrac{1}{p^s})(1-\tfrac{e^{-t}}{p^s})^{-1} \\
&= \exp\left\{\textstyle\sum_p\left(\ln(1-\frac{1}{p^s})-\ln(1-\frac{e^{-t}}{p^s})\right)\right\} \\
&= \exp\left\{\textstyle\sum_p\sum_{m=1}^{\infty}\left(\frac{e^{-mt}}{mp^{ms}}-\frac{1}{mp^{ms}}\right)\right\} \\
&= \exp\left\{\textstyle\sum_p\sum_{m=1}^{\infty}\left(\frac{e^{-mt}-1}{mp^{ms}}\right)\right\} \\
&= \exp\left\{\textstyle\sum_{m=1}^{\infty}\left(\frac{e^{-mt}-1}{m}\right)\mathbb{P}(ms)\right\} \\
&= \exp\left\{\left(e^{-t}-1\right)\mathbb{P}(s)\right\}\exp\left\{\textstyle\sum_{m=2}^{\infty}\left(\frac{e^{-mt}-1}{m}\right)\mathbb{P}(ms)\right\}
\end{aligned}$$

In order to compute $Ee^{t(\Omega(X_s)-\omega(X_s))}$, we observe that $\Omega(m)-\omega(m)$ is an additive functional with $\Omega(p^m)-\omega(p^m) = (m-1)^+$. Thus, by 27,

$$\begin{aligned}
Ee^{t(\Omega(X_s)-\omega(X_s))} &= \Pi_p(1-\tfrac{1}{p^s})\textstyle\sum_{m\geq 0}\frac{e^{t(\Omega(p^m)-\omega(p^m))}}{p^{sm}} \\
&= \Pi_p(1-\tfrac{1}{p^s})\textstyle\sum_{m\geq 0}\frac{e^{t(m-1)^+}}{p^{sm}} \\
&= \Pi_p(1-\tfrac{1}{p^s})\left(1+\tfrac{1}{p^s}+\textstyle\sum_{m\geq 2}\frac{e^{t(m-1)}}{p^{sm}}\right) \\
&= \Pi_p(1-\tfrac{1}{p^s})\left(1+\tfrac{1}{p^s}+\tfrac{e^t}{p^{2s}}\textstyle\sum_{m\geq 0}\left(\frac{e^t}{p^s}\right)^m\right) \\
&= \Pi_p(1-\tfrac{1}{p^s})\left(1+\tfrac{1}{p^s}+\frac{e^t}{p^{2s}(1-\frac{e^t}{p^s})}\right) \\
&= \Pi_p\left(1-\tfrac{1}{p^s}\right)\left(1+\tfrac{1}{p^s-e^t}\right).
\end{aligned}$$

□

Remark.
- *It follows from the fact that* $\sum_{p\in\mathcal{P}}\frac{1}{p^{2s}} < \infty$ *that the infinite product in (35) converges for* $t < \ln 2$. *In fact, this remains true even for* $s = 1$.
- *The term* $e^{(e^t-1)\mathbb{P}(s)}$ *is the moment generating function of a Poisson random variable with parameter* $\mathbb{P}(s)$. *This is the "dominant" term in* $Ee^{t\omega(X_s)}$.
- *Note that*

$$\lim_{s\searrow 1}\sum_{n=2}^{\infty}\frac{(e^t-1)^n}{n}\mathbb{P}(ns) = \sum_{n=2}^{\infty}\frac{(e^t-1)^n}{n}\mathbb{P}(n)$$

  is finite for all $t < \ln 2$. *This follows from the estimate* $\mathbb{P}(n) < \sum_{k=2}^{\infty}\frac{1}{k^n} < \frac{1}{n-1}$.
- *An even more striking thing happens with* $\Omega(X_s)$. *Again the pre-factor in the moment generating function for* $\Omega(X_s)$, $\exp\left\{\left(e^{-t}-1\right)\mathbb{P}(s)\right\}$ *is the*

moment generating function of a Poisson random variable with parameter $\mathbb{P}(s)$. *The somewhat amazing fact here is that* $\exp\left\{\sum_{k=2}^{\infty}\left(\frac{e^{-kt}-1}{k}\right)\mathbb{P}(ks)\right\}$ *is also the moment generating function of a random variable parametrized by s and this "sequence" of random variables has a limit in distribution as* $s \searrow 1$. *This is a consequence of an observation due to Lloyd* [L]. *We shall develop this fact below.*

- *Renyi* [MV] *showed that* $\lim_{N\to\infty}\frac{1}{N}\sum_{m=1}^{N}e^{t(\Omega(m)-\omega(m))} = \Pi_p\left(1-\frac{1}{p^s}\right)\left(1+\frac{1}{p^s-e^t}\right)$. *That is* $\Omega(X_s)-\omega(X_s)$ *and* $\Omega(Y_N)-\omega(Y_N)$ *have the same limiting distribution as* $s \searrow 1$ *and* $N \to \infty$, *respectively. This coincidence always seems to happen and is actually not a coincidence as will be proved later.*

5. Lloyd's Decomposition

We now describe the decomposition of Lloyd [L], by noting that embedded in the above computation one finds the moment generating function of a geometric random variable with parameter $\frac{1}{p^s}$;

$$Ee^{tc_p(s)} = \frac{1-\frac{1}{p^s}}{1-\frac{e^t}{p^s}}.$$

If we divide this moment generating function by the moment generating function of a Poisson random variable with parameter $\frac{1}{p^s}$, namely divide by $e^{(e^t-1)\frac{1}{p^s}}$ we get by a straightforward calculation

$$\frac{1-\frac{1}{p^s}}{(1-\frac{e^t}{p^s})e^{(e^t-1)\frac{1}{p^s}}} = \sum_{m=1}^{\infty}\lambda_m(s)e^{mt} \tag{36}$$

where

$$\lambda_m(s) = (1-\frac{1}{p^s})e^{\frac{1}{p^s}}\frac{1}{p^{ms}}\sigma(m),\ m = 0,1,2,\cdots$$

and

$$\sigma(m) = \sum_{j=0}^{m}\frac{(-1)^j}{j!}.$$

Three important observations can be summarized in a Theorem due to Stuart.

Theorem 5.1. (Lloyd, 5.1)

- $\lambda_m(s)$, $m = 0,1,2,\cdots$ is a probability distribution on $\mathbf{N}$,
- $\lambda_1(s) \equiv 0$
- $$\lim_{s\searrow 1}\lambda_m(s) = (1-\frac{1}{p})e^{\frac{1}{p}}\frac{1}{p^m}\sigma(m),\ m = 0,1,2,\cdots$$
  is a well-defined probability distribution on $\mathbf{N}$.

Proof. These are but routine calculations. □

We can thus construct two independent sequences, such that for each $p \in \mathcal{P}$ and $s > 1$ we may write

$$c_p(s) \stackrel{d}{=} e_p(s) + d_p(s),$$

where the sequences $\{e_p(s) : p \in \mathcal{P}\}$ and $\{d_p(s) : p \in \mathcal{P}\}$ are also independent, $d_p(s)$ is Poisson with parameter $\frac{1}{p^s}$ and $P(e_p(s) = m) = \lambda_m(s)$. Writing

$$M_s = \Pi_p p^{d_p(s)}$$

and

$$N_s = \Pi_p p^{e_p(s)}.$$

We then have that

Proposition 5.2. (Lloyd, [L]) For any $s > 1$, there exist independent random variables M_s, N_s taking values in $\mathbf{N}$ such that

$$\begin{aligned}
X_s &\stackrel{d}{=} M_s N_s \\
\Omega(X_s) &= \Omega(M_s) + \Omega(N_s) \\
\Omega(M_s) &= \textstyle\sum_{p\in\mathcal{P}} d_p(s) \text{ is a Poisson random variable with parameter } \mathbb{P}(s). \\
E\left[e^{t\Omega(M_s)}\right] &= \exp\left\{\left(e^t - 1\right)\mathbb{P}(s)\right\} \\
E\left[e^{t\Omega(N_s)}\right] &= \exp\left\{\textstyle\sum_{k=2}^{\infty}\left(\frac{e^{kt}-1}{k}\right)\mathbb{P}(ks)\right\}. \\
\Omega(N_s) &\quad \text{has a limiting nondegenerate distribution as } s \searrow 1.
\end{aligned}$$

Proof. We'll give the computation for the moment generating function of $\Omega(N_s)$. Starting with 71, since $\Omega(N_s) = \sum_p e_p(s)$.

$$\begin{aligned}
E\left[e^{t\Omega(N_s)}\right] &= \Pi_p \frac{1 - \frac{1}{p^s}}{(1 - \frac{e^t}{p^s})e^{(e^t-1)\frac{1}{p^s}}} \\
&= \exp\left\{\sum_p \left(\log(1 - \frac{1}{p^s}) - \log(1 - \frac{e^t}{p^s}) - (e^t - 1)\frac{1}{p^s}\right)\right\} \\
&= \exp\left\{\sum_p \left(-\sum_{m\geq 1} \frac{1}{mp^{ms}} + \sum_{m\geq 1} \frac{e^{mt}}{mp^{ms}} - (e^t - 1)\frac{1}{p^s}\right)\right\} \qquad (37) \\
&= \exp\left\{-\sum_{m\geq 1} \frac{\mathbb{P}(ms)}{m} + \sum_{m\geq 1} \frac{e^{mt}\mathbb{P}(ms)}{m} - (e^t - 1)\mathbb{P}(s)\right\} \\
&= \exp\left\{\sum_{m\geq 2} \frac{(e^{mt} - 1)\mathbb{P}(ms)}{m}\right\}
\end{aligned}$$

Finally,

$$\lim_{s\searrow 1} E\left[e^{t\Omega(N_s)}\right] = \exp\left\{\sum_{m\geq 2} \frac{(e^{mt} - 1)\mathbb{P}(m)}{m}\right\}$$

and the right hand side is well defined so the random variables $\Omega(N_s)$ converge in distribution as $s \searrow 1$.

$\square$

The $e_p(s)$ don't contribute very often in the sense that $P(e_p(s) > 0, i.o.\text{ in } p) = 0$. Indeed, $P(e_p(s) \geq 1) = 1 - \lambda_0(s) = 1 - (1 - \frac{1}{p^s})e^{\frac{1}{p^s}}$. And using a first order Taylor expansion for $e^{\frac{1}{p^s}}$ we get $1 - (1 - \frac{1}{p^s})e^{\frac{1}{p^s}} \sim \frac{1}{p^{2s}}$, $p \to \infty$. Since for $s > 1$,

$$\sum_p \frac{1}{p^{2s}} < \sum_p \frac{1}{p^2} < \infty,$$

by Borel-Cantelli, for all $s \geq 1$,

$$P(e_p(s) \geq 1, \text{for infinitely many } p \in \mathcal{P}) = 0.$$

In fact, $P(N_1 = 1) \approx .729$ ([L], Section 3). We see by 11.3 that $\Omega(X_s) - \Omega(M_s) = \Omega(N_s)$ has a limiting distribution as $s \searrow 1$. Consequently, $\Omega(X_s)$ is nearly a Poisson random variable with parameter $\mathbb{P}(s)$. We formalize this in a variant of an observation due to Lloyd [L]

The Lloyd decomposition allows an alternative proof of the Erdös-Kac CLT. Since $\Omega(N_s)$ is Poisson with parameter $\mathbb{P}(s)$ it follows that

$$\frac{\Omega(M_s) - \mathbb{P}(s)}{\sqrt{\mathbb{P}(s)}} \xrightarrow{d} \mathcal{N}(0,1).$$

Since

$$\frac{\Omega(N_s)}{\sqrt{\mathbb{P}(s)}} \xrightarrow{d} 0,$$

Slutsky's Theorem implies

$$\frac{\Omega(X_s) - \mathbb{P}(s)}{\sqrt{\mathbb{P}(s)}} \xrightarrow{d} \mathcal{N}(0,1).$$

It also holds that

$$\frac{\omega(X_s) - \Omega(X_s)}{\sqrt{\mathbb{P}(s)}} \xrightarrow{d} 0,$$

and since $\omega(X_s) = \omega(X_s) - \Omega(X_s) + \Omega(X_s)$, another application of Slutsky's Theorem gives

$$\frac{\omega(X_s) - \mathbb{P}(s)}{\sqrt{\mathbb{P}(s)}} \xrightarrow{d} \mathcal{N}(0,1).$$

6. Greatest Common Divisors, Visible Points, Parity of $\Omega(X_s)$

We now digress to give an amusing aside. The greatest common divisor of two integers m, n will be denoted by $gcd(m,n)$ as before. If their prime factorizations are $m = \Pi_p p^{c_p(m)}, n = \Pi_p p^{d_p(n)}$ then $gcd(m,n) = \Pi_p p^{c_p(m) \wedge d_p(n)}$. The fact that $c_p(s) \wedge d_p(s)$ is a geometric random variable with parameter $\frac{1}{p^{2s}}$ when $c_p(s)$ and $d_p(s)$ are independent geometric random variables with parameter $\frac{1}{p^s}$ leads to an interesting observation about the gcd of two independent copies of $\zeta(s)$ random variables X_s^1, X_s^2. If we take their prime factorizations to be $X_s^i = \Pi_p p^{c_p^i(s)}$, $i = 1,2$, then $gcd(X_s^1, X_s^2) = \Pi_p p^{c_p^1(s) \wedge c_p^2(s)}$. Since $c_p^1(s)$ and $c_p^2(s)$ are independent, $c_p^1(s) \wedge c_p^2(s)$ is a geometric random variable with parameter $\frac{1}{p^{2s}}$, that is $P(c_p^1(s) \wedge c_p^2(s) \geq k) = \frac{1}{p^{2ks}}$. This means that $X_{2s} \stackrel{d}{=} gcd(X_s^1, X_s^2)$. What's interesting about this is that while neither X_s^1 nor X_s^2 have limiting distributions as $s \searrow 1$, their gcd has a limiting distribution given by that of X_2. This can be generalized easily as follows.

Proposition 6.1. If X_s^j, $j \in [k]$ are *iid* copies of a $\zeta(s)$ random variable X_s, then $X_{ks} \stackrel{d}{=} \gcd(X_s^1, X_s^2, \cdots, X_s^k)$.

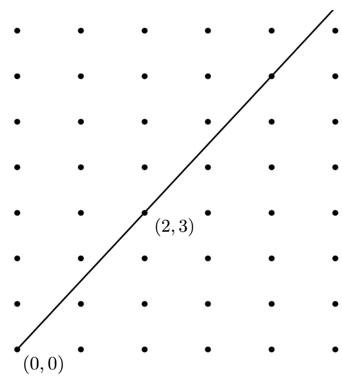

FIGURE 1. Points in $\mathbf{Z^2}$ that are visible from the origin. The point $(2, 3)$ is visible from the origin, but the point $(4, 6)$ is not as it is eclipsed by $(2, 3)$.

We now give a geometric interpretation of this result. In $\mathbf{R^2}$ we say an integer lattice point Q is visible (from the origin) if there is no integer lattice point on the line through the origin and Q and lying between the origin and Q (see Figure 1). If $Q = (m, n)$ then Q is visible from the origin if and only if $\gcd(m, n) = 1$. If $Q_s = (X_s^1, X_s^2)$ where X_s^1 and X_s^2 are independent, then by Proposition 6.1, $\lim_{s \searrow 1} P(Q_s \text{ is visible}) = P(X_2 = 1) = \frac{6}{\pi^2}$ since $P(X_2 = 1) = \frac{1}{\zeta(2)}$ and $\zeta(2) = \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$. It's well known that a similar result holds when selecting a pair of independent random variables according to the uniform distribution on N and letting $N \to \infty$. In higher dimensions, the probability that $Q_s^k = (X_s^1, X_s^2, \cdots, X_s^k)$ is visible, where $X_s^1, \cdots, X_s^k$ are independent $\zeta(s)$ random variables has limit given by

$$\begin{aligned} \lim_{s \searrow 1} P(Q_s^k \text{ visible from origin}) &= \lim_{s \searrow 1} P(gcd(X_s^1, X_s^2, \cdots, X_s^k) = 1) \\ &= \lim_{s \searrow 1} P(X_{sk} = 1) \\ &= \zeta(k)^{-1}. \end{aligned}$$

When k is even, say $k = 2n$, then $\lim_{s \searrow 1} P(Q_s^{2n} \text{ visible from origin}) = \frac{2(2n)!}{(2\pi)^{2n}(-1)^{n+1}B_{2n}}$, where B_{2n} is the $2n^{th}$ Bernoulli number.

7. ERDÖS-KAC, HALBERSTAM CLT

In this section we prove analogs of the famous Erdös-Kac CLT, [EK], and its relatives established by Halberstam [H] for the $\zeta(s)$ distribution. We first review the original central limit theorem of Erdös and Kac. Denote

$$\Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}} dx.$$

Recalling that $\omega(n)$ is the number of distinct primes in the prime factorization of n then the Erdös-Kac central limit theorem states that

$$\lim_{N\to\infty} \frac{1}{N} \left| \{n \le N : a \le \frac{\omega(n) - \log\log N}{\sqrt{\log\log N}} \le b\} \right| = \Phi(b) - \Phi(a). \tag{38}$$

Halberstam derived an extension of this theorem to other additive functions besides the number of distinct prime factors function ω. In this section we shall derive the analogous result for $\omega(X_s)$. In fact, it will be possible to derive a central limit theorem for the number of distinct prime factors of X_s in certain arithmetic sequences. The arithmetic sequences in question are those of the form $A(a,b) = \{a + kb : k \in \mathbf{N}\}$ where $gcd(a,b) = 1$. According to a Theorem of Dirichlet, there are infinitely many primes in any such arithmetic sequence. Fixing b the set $A(a,b)$ is the equivalence class of numbers congruent to $a \mod b$. There are $\phi(b)$ such equivalence classes where ϕ is the Euler totient function, namely, $\phi(b)$ is the number of positive integers less than b and relatively prime to b. Each equivalence class $\mod b$ has a $1/\phi(b)$ proportion of the primes. Using **??** we can separate out the contribution of each equivalence class to $\omega(X_s)$. If $a_1, a_2, \cdots a_{\phi(b)}$ are the distinct representatives of the $\phi(b)$ equivalence classes then we can define

$$\omega_j(X_s) = \sum_{p\in A(a_j,b)} 1_{c_p(s)>0}.$$

Then $\omega(X_s) = \sum_{j=1}^{\phi(b)} \omega_j(X_s)$ and by the independence of the $c_p(s)$'s,

$$Ee^{t\omega(X_s)} = \Pi_{j=1}^{\phi(b)} Ee^{t\omega_j(X_s)}.$$

Alternatively we can achieve this by appropriate use of additive functions. Fixing b, define $f_j : \mathbb{P} \to \mathbb{N}$ by setting $f_j(p) = 1_{A(a_j,b)}(p)$, $f_j(p^m) = f_j(p)$ for $m \ge 1$ and finally setting $f_j(m) = \sum_{p|m} f_j(p)$. Then f_j is an additive function on $\mathbf{Z}^+$ and the value of $f_j(m)$ is exactly the number of primes in the arithmetic progression $A(a_j,b)$ which divide m. In other words, $f_j = \omega_j$. We note that $\sum_p \frac{(e^{tf_j(p)}-1)}{p^s} = (e^t-1)\sum_{p\equiv a_j \mod b} \frac{1}{p^s}$.

Thus, 32 gives that if

$$\mathbb{P}_j(s) = \sum_{p\in A(a_j,b)} \frac{1}{p^s}$$

then

$$Ee^{t\omega_j(X_s)} = \exp\{(e^t-1)\mathbb{P}_j(s)\}\exp\{-\sum_{k=2}^{\infty} \frac{(1-e^t)^k}{k}\mathbb{P}_j(ks)\}.$$

It's also immediate that

$$E\omega_j(X_s) = P_j(s)$$

and

$$var\,\omega_j(X_s) = \mathbb{P}_j(s) - \mathbb{P}_j(2s).$$

With the analogous definition for $\Omega_j(X_s) = \sum_{p\in A(a_j,b)} c_p(s)$ it follows that

$$Ee^{t\Omega_j(X_s)} = \exp\left\{\left(e^t-1\right)\mathbb{P}_j(s)\right\}\exp\left\{\sum_{k=2}^{\infty}\left(\frac{e^{kt}-1}{k}\right)\mathbb{P}_j(ks)\right\}$$

and

$$Ee^{t(\Omega_j(X_s)-\omega_j(X_s))} = \Pi_{p\in A(a_j,b)}\left(1-\frac{1}{p^s}\right)\left(1+\frac{1}{p^s-e^t}\right).$$

We can present the analog of the Erdös-Kac Theorem

Theorem 7.1. The random variables $\omega_j(X_s)$ are independent and

$$\lim_{s\searrow 1} P\left(\frac{\omega_j(X_s)-\mathbb{P}_j(s)}{\sqrt{\mathbb{P}_j(s)}}\le x\right) = \Phi(x),\ j=1,2,\cdots,\phi(b)$$

Also,

$$\lim_{s\searrow 1} P\left(\frac{\omega(X_s)-\mathbb{P}(s)}{\sqrt{\mathbb{P}(s)}}\le x\right) = \Phi(x).$$

Proof. All we need to do is check the limit of the moment generating function for $\frac{\omega(X_s)-\mathbb{P}(s)}{\sqrt{\mathbb{P}(s)}}$ as $s\searrow 1$. But this is straightforward: By (33) we have

$$E[e^{t\frac{\omega(X_s)-\mathbb{P}(s)}{\sqrt{\mathbb{P}(s)}}}] = \exp\{(e^{\frac{t}{\sqrt{\mathbb{P}(s)}}}-1)\mathbb{P}(s)-t\sqrt{\mathbb{P}(s)}\} \times\exp\{-\sum_{k=2}^{\infty}\frac{(1-\exp\{\frac{t}{\sqrt{\mathbb{P}(s)}}\})^k}{k}\mathbb{P}(ks)\} \tag{39}$$

Since $\lim_{s\searrow 1}\mathbb{P}(s)=\infty$ we can use the Taylor expansion $(\exp\{\frac{t}{\sqrt{\mathbb{P}(s)}}\}-1)\mathbb{P}(s) = \frac{t}{\sqrt{\mathbb{P}(s)}}+\frac{t^2}{2}+O(\frac{1}{\sqrt{\mathbb{P}(s)}})$ to get

$$\lim_{s\searrow 1}\exp\{(\exp\{\frac{t}{\sqrt{\mathbb{P}(s)}}\}-1)\mathbb{P}(s)-t\sqrt{\mathbb{P}(s)}\} = \frac{t^2}{2}.$$

Using this Taylor expansion again we obtain

$$\lim_{s\searrow 1}(1-\exp\{\frac{t}{\sqrt{\mathbb{P}(s)}}\})^k = 0,\ k\ge 1.$$

Since $\lim_{s\searrow 1}\sum_{k=2}^{\infty}\frac{\mathbb{P}(ks)}{k}<\infty$, we conclude that

$$\lim_{s\searrow 1}\exp\{-\sum_{k=2}^{\infty}\frac{(1-\exp\{\frac{t}{\sqrt{\mathbb{P}(s)}}\})^k}{k}\mathbb{P}(ks)\} = 1.$$

The same proof works for the $\omega_j(X_s)$. The fact that the $\omega_j(X_s)$ are independent follows from the fact that the arithmetic sequences are disjoint and the independence in p of the random variables $c_p(s)$. □

We can also prove in a similar manner, an analog of a result of Halberstam, [H] in the context of strongly additive functions. Halberstam proved that if f if s a bounded strongly additive function and $A_n=\sum_{p<n}\frac{f(p)}{p}$ and $B_n=\sum_{p<n}\frac{f^2(p)}{p}$ then

$$\frac{1}{n}|\{1\le m\le n:\ \frac{f(m)-A_n}{\sqrt{B_n}}\le b\}|\to\Phi(b),\ n\to\infty.$$

As an example, one may take $f(m)=\omega_j(m)$ as defined above to see that the distinct number of prime divisors congruent to $a_j \mod b$ for a random integer selected uniformly from $[N]$ satisfies the Central Limit Theorem. The analog for the $\zeta(s)$ distribution is the following. The analog of $A(n)$ is

$$\begin{aligned} E[f(X_s)] = &\quad E[f(\Pi_p p^{c_p(s)})] \\ = &\quad E[\textstyle\sum_p f(p) 1_{c_p(s)>0}] \\ = &\quad \textstyle\sum_p f(p) P(c_p(s) > 0) \\ = &\quad \textstyle\sum_p \frac{f(p)}{p^s}. \end{aligned}$$

The analog of $B(n)$ is the variance of $f(X(s))$ which we now compute. First

$$\begin{aligned} E[f^2(X_s)] = &\quad E[f^2(\Pi_p p^{c_p(s)})] \\ = &\quad E[(\textstyle\sum_p f(p) 1_{c_p(s)>0})^2] \\ = &\quad \textstyle\sum_{p\neq q} f(p)f(q)P(c_p(s) > 0)P(c_q(s) > 0) + \sum_p \frac{f^2(p)}{p^s} \\ = &\quad \textstyle\sum_{p\neq q} \frac{f(p)f(q)}{(pq)^s} + \sum_p \frac{f^2(p)}{p^s} \\ = &\quad \textstyle\sum_{p,q} \frac{f(p)f(q)}{(pq)^s} - \sum_p \frac{f^2(p)}{p^{2s}} + \sum_p \frac{f^2(p)}{p^s} \\ = &\quad \left(\textstyle\sum_p \frac{f(p)}{p^s}\right)^2 - \sum_p \frac{f^2(p)}{p^{2s}} + \sum_p \frac{f^2(p)}{p^s}. \end{aligned}$$

So, we then have

$$\begin{aligned} var(f(X_s)) = &\quad \textstyle\sum_p \frac{f^2(p)}{p^s} - \sum_p \frac{f^2(p)}{p^{2s}} \\ \sim &\quad \textstyle\sum_p \frac{f^2(p)}{p^s},\ s \searrow 1, \end{aligned}$$

since the term $\sum_p \frac{f^2(p)}{p^{2s}} = O(1)$. So we set $B(s) = \sum_p \frac{f^2(p)}{p^s}$.

Theorem 7.2. Let f be a bounded strongly additive function for which $\sum_p \frac{f(p)}{p} = \infty$ and set $A(s) = \sum_p \frac{f(p)}{p^s}$ and $B(s) = \sum_p \frac{f^2(p)}{p^s}$ then for all $b \in \mathbf{R}$,

$$\lim_{s\searrow 1} P(\frac{f(X_s) - A(s)}{\sqrt{B(s)}} \leq b) = \Phi(b).$$

Proof. Briefly, by 32,

$$Ee^{itf(X_s)} = \exp\{\sum_p \frac{(e^{itf(p)} - 1)}{p^s}\} \exp\{-\sum_{m\geq 2} \frac{(1 - e^{itf(p)})^m}{mp^{ms}})\},$$

so

$$Ee^{it\frac{f(X_s)-A(s)}{\sqrt{B(s)}}} = \exp\{\sum_p \frac{(e^{\frac{it}{\sqrt{B(s)}}f(p)} - 1)}{p^s} - it\frac{A(s)}{\sqrt{B(s)}}\} \exp\{-\sum_{m\geq 2} \frac{(1 - e^{i\frac{t}{\sqrt{B(s)}}f(p)})^m}{mp^{ms}})\},$$

Using the Taylor expansion of $e^{\frac{it}{\sqrt{B(s)}}f(p)}$ reveals that the exponent in the first factor is $-\frac{t^2}{2} + o(1)$. The exponent in the second factor is $o(1)$. Thus,

$$\lim_{s\searrow 1} Ee^{it\frac{f(X_s)-A(s)}{\sqrt{B(s)}}} = e^{-\frac{t^2}{2}}.$$

This concludes the proof. ∎

8. Total Number of Divisors of X_s

We now turn attention to a Central Limit Theorem for the log of the number of divisors of X_s. The total number of divisors of n denoted by $\sigma_0(n)$ and defined by

$$\sigma_0(n) = \sum_{d|n} 1. \tag{40}$$

In this case the expected value of $\sigma_0(X_s)$ is given by the Dirichlet series $\sum_{n\geq 1} \frac{\sigma_0(n)}{n^s}$. This we can compute and it yields

$$E[\sigma_0(X_s)] = \zeta(s) \tag{41}$$

as we now show. Since σ_0 is multiplicative we can apply 29. As $\sigma_0(p^m) = m + 1$, this gives

$$\begin{aligned}
E[\sigma_0(X_s)] &= \Pi_p \left(1 - \tfrac{1}{p^s}\right) \textstyle\sum_{m\geq 0} \frac{\sigma_0(p^m)}{p^{ms}} \\
&= \Pi_p \left(1 - \tfrac{1}{p^s}\right) \textstyle\sum_{m\geq 0} \frac{m+1}{p^{ms}} \\
&= \Pi_p \left(1 - \tfrac{1}{p^s}\right) \textstyle\sum_{m\geq 0} \sum_{j=0}^{m} \frac{1}{p^{ms}} \\
&= \Pi_p \left(1 - \tfrac{1}{p^s}\right) \textstyle\sum_{j=0}^{\infty} \sum_{m\geq j} \frac{1}{p^{ms}} \\
&= \Pi_p \left(1 - \tfrac{1}{p^s}\right) \textstyle\sum_{j=0}^{\infty} \frac{1}{p^{js}}(1 - \frac{1}{p^s})^{-1} \\
&= \Pi_p \left(1 - \tfrac{1}{p^s}\right)^{-1} \\
&= \zeta(s).
\end{aligned}$$

Using a different approach we can compute the second moment of $\sigma_0(X_s)$. By the multiplicative property of σ_0, we can express

$$\sigma_0(X_s) = \Pi_p(1 + c_p(s))$$

and so using a standard computation for geometric random variables,

$$\begin{aligned}
E[\sigma_0^2(X_s)] =& \Pi_p E[(1 + c_p(s))^2] \\
=& \Pi_p(1 + \frac{1}{p^s})(1 - \frac{1}{p^s})^{-2} \\
=& \zeta(s)^2 \Pi_p(1 + \frac{1}{p^s}) \\
=& \zeta(s)^2 \Pi_p(1 + \frac{1}{p^s})(1 - \frac{1}{p^s})(1 - \frac{1}{p^s})^{-1} \\
=& \zeta(s)^3 \Pi_p(1 - \frac{1}{p^{2s}}) \\
=& \frac{\zeta(s)^3}{\zeta(2s)}.
\end{aligned} \tag{42}$$

It's interesting perhaps to note that the gap between the expected number of divisors and the expected number of prime factors of X_s is on the order of $\zeta(s) - \log \zeta(s)$ since

$$E[\sigma_0(X_s)] = \zeta(s) > \mathbb{P}(s) = E[\omega(X_s)].$$

As an aside, we mention that the function σ_0 arises in the computation of the determinant of a 2×2 matrix with *iid* $\zeta(s)$ random entries. The determinant then takes the form $D = W_sX_s - Y_sZ_s$ with all four entries W_s, X_s, Y_s, Z_s being independent $\zeta(s)$ distributed random variables. Then we can give an explicit expression for the probability that the matrix is singular. Using 42 on line 9 we get

$$\begin{aligned} P(D=0) &= P(W_sX_s - Y_sZ_s = 0) \\ &= P(W_sX_s = Y_sZ_s) \\ &= \textstyle\sum_{n\geq 1} P(W_sX_s = n)^2 \\ &= \textstyle\sum_{n\geq 1}\left(\sum_{d|n} P(W_s = d)P(X_s = n/d)\right)^2 \\ &= \tfrac{1}{\zeta(s)^4}\textstyle\sum_{n\geq 1}\left(\sum_{d|n}\tfrac{1}{d^s}\tfrac{1}{(n/d)^s}\right)^2 \\ &= \tfrac{1}{\zeta(s)^4}\textstyle\sum_{n\geq 1}\left(\sum_{d|n}\tfrac{1}{n^s}\right)^2 \\ &= \tfrac{1}{\zeta(s)^4}\textstyle\sum_{n\geq 1}\tfrac{\sigma_0^2(n)}{n^{2s}} \\ &= \tfrac{\zeta(2s)}{\zeta(s)^4}E[\sigma_0^2(X_{2s})] \\ &= \tfrac{\zeta^4(2s)}{\zeta(s)^4\zeta(4s)}. \end{aligned}$$

Thus, using $\zeta(2) = \frac{\pi^2}{6}$, $\zeta(4) = \frac{\pi^4}{90}$, and $\zeta(s) \sim (s-1)^{-1}$, one gets

$$P(D=0) \sim \frac{5\pi^4}{72}(s-1)^4,\ s \searrow 1.$$

It would be interesting to determine this probability for $n\times n$ matrices. It's tempting to conjecture that the correct asymptotic would involve Bernoulli numbers and a factor of $(s-1)^{n^2}$.

We now turn our attention to computing the mean and variance for $\log\sigma_0(X_s)$. First recall the summation by parts formula

$$\sum_{m\geq 2}(a_m - a_{m-1})b_m = -a_1b_2 + \sum_{m\geq 2} a_m(b_m - b_{m-1}).$$

Then taking $a_m = \frac{1}{p^{sm}}, b_m = \log m$ we can compute

$$\begin{aligned} E\log(c_p(s)+1) =&(1-\frac{1}{p^s})\sum_{m\geq 2} p^{-s(m-1)}\log(m) \\ =&\sum_{m\geq 2}\left(p^{-s(m-1)} - p^{-sm}\right)\log m \\ =&\frac{1}{p^s}\log 2 + \sum_{m\geq 2} p^{-sm}\log(1+\frac{1}{m}). \end{aligned} \tag{43}$$

Therefore,

$$E\log\sigma_0(X_s) = \mathbb{P}(s)\log 2 + \sum_{m\geq 2}\mathbb{P}(sm)\log(1+\frac{1}{m}). \tag{44}$$

Straightforward estimates reveal that $\sum_{m\geq 2}\mathbb{P}(sm)\log(1+\frac{1}{m}) = O(1), s\searrow 1$ so

$$E\log\sigma_0(X_s) = \mathbb{P}(s)\log 2 + O(1),\ s\searrow 1. \tag{45}$$

The second moment poses slightly more challenge. At one point in its calculation we use the following estimate

$$\sum_p \sum_{m\geq 2} \log^2(m)\left(\frac{1}{p^{s(m-1)}} - \frac{1}{p^{sm}}\right)^2 = \sum_p \sum_{m\geq 2} \log^2(m)\left(\frac{1}{p^{2s(m-1)}} - 2\frac{1}{p^{(2m-1)s}} + \frac{1}{p^{2sm}}\right)$$

Using the prime number theorem we see that

$$\begin{aligned} \sum_p \sum_{m\geq 2} \frac{\log^2(m)}{p^{2m}} &\sim \sum_{m\geq 2} \sum_{n\geq 2} \frac{\log^2(m)}{(n\log n)^{2m}} \\ &\leq C\frac{1}{(n\log n)^2} \sum_{n\geq 2} \frac{m}{((n\log n)^2)^{m-1}} \\ &= C\frac{1}{(n\log n)^2}(1-\frac{1}{n\log n})^{-2} \\ &\leq C'\frac{1}{(n\log n)^2} \end{aligned}$$

Which implies

$$\sum_p \sum_{m\geq 2} \frac{\log^2(m)}{p^{2sm}} = O(1), \ , s\to 1.$$

Similar estimates lead to the conclusion that

$$\sum_p \sum_{m\geq 2} \log^2(m)\left(\frac{1}{p^{s(m-1)}} - \frac{1}{p^{sm}}\right)^2 = O(1),\ s\to 1.$$

Thus for the second moment we get

$$
\begin{aligned}
E\log^2\sigma_0(X_s) &= E\left(\sum_p \log(c_p(s)+1)\right)^2\\
&= E\sum_{p,q}\log(c_p(s)+1)\log(c_q(s)+1)\\
&= E\sum_{p\neq q}\log(c_p(s)+1)\log(c_q(s)+1) + E\sum_p \log^2(c_p(s)+1)\\
&= \sum_{p\neq q}\sum_{m,l\geq 2}\log(m)\log(l)\left(\frac{1}{p^{s(m-1)}}-\frac{1}{p^{sm}}\right)\left(\frac{1}{q^{s(l-1)}}-\frac{1}{q^{sl}}\right)\\
&\quad + \sum_p\sum_{m\geq 2}\log^2(m)\left(\frac{1}{p^{s(m-1)}}-\frac{1}{p^{sm}}\right)\\
&= \left(\sum_{p\neq q}\sum_{m,l\geq 2}\log(m)\left(\frac{1}{p^{s(m-1)}}-\frac{1}{p^{sm}}\right)\right)^2\\
&\quad -\sum_p\sum_{m\geq 2}\log^2(m)\left(\frac{1}{p^{s(m-1)}}-\frac{1}{p^{sm}}\right)^2 + \sum_p\sum_{m\geq 2}\log^2(m)\left(\frac{1}{p^{s(m-1)}}-\frac{1}{p^{sm}}\right)\\
&= \left(\mathbb{P}(s)\log 2 + \sum_{m\geq 2}\mathbb{P}(sm)\log(1+\frac{1}{m})\right)^2 - \sum_p\sum_{m\geq 1}\log^2(m)\left(\frac{1}{p^{s(m-1)}}-\frac{1}{p^{sm}}\right)^2\\
&\quad +\mathbb{P}(s)\log^2 2 + \sum_{m\geq 2}\mathbb{P}(sm)\log(1+\frac{1}{m})\log m(m+1)\\
&= \mathbb{P}^2(s)\log^2 2 + 2\mathbb{P}(s)\log 2\sum_{m\geq 2}\mathbb{P}(sm)\log(1+\frac{1}{m}) + \mathbb{P}(s)\log^2 2 + o(1).
\end{aligned}
$$

This leads to

$$
var(X_s) = \mathbb{P}(s)\log 2\left(2\sum_{m\geq 2}\mathbb{P}(sm)\log(1+\frac{1}{m}) + \log 2\right) + o(1). \tag{46}
$$

We make some remarks.

Remark.

- *In the special case $f(p)\equiv 1$ this gives the Erdös-Kac CLT. The Lindeberg-Feller Central Limit Theorem can be applied with ease in this setting.*
- *Since $\lim_{s\searrow 1}\frac{\mathbb{P}(s)}{\mathbb{P}_j(s)} = \phi(b)$ the CLT for $\omega_j(X_s)$ may be rephrased*

$$
\lim_{s\searrow 1} P\left(\frac{\omega_j(X_s)-\frac{\mathbb{P}(s)}{\phi(b)}}{\sqrt{\frac{\mathbb{P}(s)}{\phi(b)}}}\leq x\right) = \Phi(x),\ j=1,2,\cdots,\phi(b)
$$

- *The random variables $\Omega(X_s)$ and $\Omega_j(X_s)$ also satisfy a CLT. This case is quite easy. Due to Proposition 5.2 we have*

$$
\Omega(X_s) = \Omega(M_s) + \Omega(N_s)
$$

where $\Omega(M_s)$ is a Poisson random variable with parameter $\mathbb{P}(s)$, $\Omega(N_s)$ has a limiting distribution and M_s and N_s are independent. Since $\lim_{s\searrow 1}\mathbb{P}(s) = \infty$, it follows that $(\Omega(M_s)-\mathbb{P}(s))/\sqrt{\mathbb{P}(s)}$ has a limiting standard normal distribution and this is inherited by $\Omega(X_s)$ since $\Omega(N_s)/\sqrt{\mathbb{P}(s)}$ tends to zero. A similar argument can be made for the $\Omega_j(X_s)$.

9. Berry-Esseen Estimates

In this section we consider the rate of convergence to the normal distribution for $\frac{\omega(X_s)-\mathbb{P}(s)}{\sqrt{\mathbb{P}(s)}}$. Our result is the analog of the uniform case which was established by Rényi and Turán. See the book by Elliot [E] for an account of the history of this problem for the uniform distribution. The rate of convergence to the normal distribution is essentially the same. One matches up the uniform and the $\zeta(s)$ case by, given N, selecting s so that $E[\omega(X_s)] = E[\omega(Y_N)]$. This leads to $\mathbb{P}(s) = \log\log N$. Due to the fact that $\mathbb{P}(s) + O(1) = \log\zeta(s)$ we could also choose s so that $\zeta(s) = \log N$, or equivalently $s = 1 + \frac{1}{\log N}$. The rate of convergence in the uniform case is $1/\sqrt{\log\log N}$ and for the $\zeta(s)$ case it is $1/\sqrt{\mathbb{P}(s)}$. As in (39), the characteristic function of $\frac{\omega(X_s)-\mathbb{P}(s)}{\sqrt{\mathbb{P}(s)}}$ is

$$\psi_s(\tau) = \exp\left\{(\exp\{\frac{i\tau}{\sqrt{\mathbb{P}(s)}}\}-1)\mathbb{P}(s) - i\tau\sqrt{\mathbb{P}(s)}\right\}\exp\left\{-\sum_{k=2}^{\infty}\frac{(1-\exp\{\frac{i\tau}{\sqrt{\mathbb{P}(s)}}\})^k}{k}\mathbb{P}(ks)\right\}$$

A Berry-Esseen estimate will rely on the difference between the characteristic function of $\frac{\omega(X_s)-\mathbb{P}(s)}{\sqrt{\mathbb{P}(s)}}$ and $e^{-\frac{\tau^2}{2}}$.

Lemma 9.1. For all $|\tau| \le \sqrt{\mathbb{P}(s)}$,

$$|\psi_s(\tau) - e^{-\frac{\tau^2}{2}}| \le e^{-\frac{\tau^2}{2}}\left(\frac{|\tau|^3}{\sqrt{\mathbb{P}^3(s)}} + \frac{\tau^2\mathbb{P}(2s)}{\mathbb{P}(s)}\right)(1+O(1)). \tag{47}$$

Proof. To this end, write

$$M_s(\tau) = \exp\left\{-\sum_{k=2}^{\infty}\frac{(1-\exp\{\frac{i\tau}{\sqrt{\mathbb{P}(s)}}\})^k}{k}\mathbb{P}(ks)\right\}$$

so that

$$\begin{aligned}
\psi_s(\tau) - e^{-\frac{\tau^2}{2}} &= \exp\{(\exp\{\frac{i\tau}{\sqrt{\mathbb{P}(s)}}\}-1)\mathbb{P}(s) - i\tau\sqrt{\mathbb{P}(s)}\}M_s(\tau) - e^{-\frac{\tau^2}{2}} \\
&= \exp\{(\exp\{\frac{i\tau}{\sqrt{\mathbb{P}(s)}}\}-1)\mathbb{P}(s) - i\tau\sqrt{\mathbb{P}(s)}\} - e^{-\frac{\tau^2}{2}} \\
&+ \exp\{(\exp\{\frac{i\tau}{\sqrt{\mathbb{P}(s)}}\}-1)\mathbb{P}(s) - i\tau\sqrt{\mathbb{P}(s)}\}(M_s(\tau)-1) \\
&= I + II.
\end{aligned} \tag{48}$$

As for I, we have

(49)

$$
\begin{aligned}
|I| =& | \exp\{(\exp\{\frac{i\tau}{\sqrt{\mathbb{P}(s)}}\}-1)\mathbb{P}(s) - i\tau\sqrt{\mathbb{P}(s)}\} - e^{-\frac{\tau^2}{2}}| \\
=& | \exp\{(\exp\{\frac{i\tau}{\sqrt{\mathbb{P}(s)}}\}-1)\mathbb{P}(s) - i\tau\sqrt{\mathbb{P}(s)}\} - e^{-\frac{\tau^2}{2}}| \\
=& e^{-\frac{\tau^2}{2}} | \exp\left\{(\cos(\frac{\tau}{\sqrt{\mathbb{P}(s)}})-1)\mathbb{P}(s) + \frac{\tau^2}{2} + i(\mathbb{P}(s)\sin(\frac{\tau}{\sqrt{\mathbb{P}(s)}}) - \tau\sqrt{\mathbb{P}(s)})\right\} - 1| \\
\leq& e^{-\frac{\tau^2}{2}} \left(\frac{|\tau|^3}{\mathbb{P}(s)^{3/2}3!} + O(\frac{|\tau|^4}{\mathbb{P}^2(s)4!})\right) \\
\leq& e^{-\frac{\tau^2}{2}} \frac{|\tau|^3}{\mathbb{P}(s)^{3/2}3!}(1+O(1)).
\end{aligned}
$$

Now for II we have the estimate

$$
\begin{aligned}
|II| =& | \exp\{(\exp\{\frac{i\tau}{\sqrt{\mathbb{P}(s)}}\}-1)\mathbb{P}(s) - i\tau\sqrt{\mathbb{P}(s)}\}(M_s(\tau)-1)| \\
\leq& e^{-\frac{\tau^2}{2} + O(1/\sqrt{\mathbb{P}(s)})|\tau|^3} |M_s(\tau)-1| \\
=& e^{-\frac{\tau^2}{2} + O(1/\sqrt{\mathbb{P}(s)})|\tau|^3} | \exp\{-\sum_{k=2}^{\infty} \frac{(1-\exp\{\frac{i\tau}{\sqrt{\mathbb{P}(s)}}\})^k}{k}\mathbb{P}(ks)\} - 1|
\end{aligned}
\tag{50}
$$

And thinking of $\exp\{\frac{i\tau}{\sqrt{\mathbb{P}(s)}}\}$ as a point on the unit circle, it follows that

$$
\left|1-\exp\{\frac{i\tau}{\sqrt{\mathbb{P}(s)}}\}\right| \leq \frac{|\tau|}{\sqrt{\mathbb{P}(s)}}. \tag{51}
$$

From 51 it follows that $\lim_{\tau\to 0} M_s(\tau) = 1$. Thus on taking the derivative in τ of the right hand side of 50 we see

$$
M_s'(\tau) = \quad M_s(\tau)\frac{i}{\sqrt{\mathbb{P}(s)}}\sum_{k=2}^{\infty}(1-\exp\{\frac{i\tau}{\sqrt{\mathbb{P}(s)}}\})^{k-1}\mathbb{P}(ks) \tag{52}
$$

Using 51 in 52 we see that $\lim_{\tau\to 0} M_s'(\tau) = 0$. Differentiating one more time,

$$
\begin{aligned}
M_s''(\tau) =& M_s'(\tau)\frac{i}{\sqrt{\mathbb{P}(s)}}\sum_{k=2}^{\infty}(1-\exp\{\frac{i\tau}{\sqrt{\mathbb{P}(s)}}\})^{k-1}\mathbb{P}(ks) \\
& - M_s(\tau)\frac{1}{\mathbb{P}(s)}\sum_{k=2}^{\infty}(k-1)(1-\exp\{\frac{i\tau}{\sqrt{\mathbb{P}(s)}}\})^{k-2}\mathbb{P}(ks).
\end{aligned}
\tag{53}
$$

Using 51 in 53 we see that $\lim_{\tau\to 0} M_s''(\tau) = -\frac{\mathbb{P}(2s)}{\mathbb{P}(s)}$.

This implies that as $\tau \to 0$, one has

$$
|M_s(\tau)-1| \leq \frac{\tau^2\mathbb{P}(2s)}{\mathbb{P}(s)}(1+O(1)). \tag{54}
$$

Thus, returning to 50 and using 54 the lemma follows. □

We can now give the Berry-Esseen Theorem in the present setting.

Theorem 9.2. There exists a positive constant C such that if F_s is the distribution function of $\frac{\omega(X_s)-\mathbb{P}(s)}{\sqrt{\mathbb{P}(s)}}$ then for all x,

$$|F_s(x)-\Phi(x)| \le \frac{C}{\sqrt{\mathbb{P}(s)}}.$$

Proof. By a standard computation for the first inequality, see the proof of the Berry-Esseen Theorem in ([D], Theorem 3.4.9), and using 47, we have

$$\begin{aligned}
|F_s(x)-\Phi(x)| \le & \quad \tfrac{1}{\pi}\int_{-\sqrt{\mathbb{P}(s)}}^{\sqrt{\mathbb{P}(s)}} |\psi_s(\tau)-e^{-\frac{\tau^2}{2}}|\tfrac{d\tau}{|\tau|}+\tfrac{24}{\pi\sqrt{\mathbb{P}(s)}} \\
\le & \quad \tfrac{1}{\pi}\int_{-\sqrt{\mathbb{P}(s)}}^{\sqrt{\mathbb{P}(s)}} e^{-\frac{\tau^2}{2}}\left(\tfrac{|\tau|^3}{\sqrt{\mathbb{P}^3(s)}}+\tfrac{\tau^2\mathbb{P}(2s)}{\mathbb{P}(s)}\right)(1+O(1))\tfrac{d\tau}{|\tau|}+\tfrac{24}{\pi\sqrt{\mathbb{P}(s)}} \\
\le & \quad \tfrac{C}{\sqrt{\mathbb{P}(s)}}\int_{-\sqrt{\mathbb{P}(s)}}^{\sqrt{\mathbb{P}(s)}} e^{-\frac{\tau^2}{2}}\left(\tfrac{\tau^2}{\mathbb{P}(s)}+\tfrac{|\tau|\mathbb{P}(2s)}{\sqrt{\mathbb{P}(s)}}\right)d\tau+\tfrac{24}{\pi\sqrt{\mathbb{P}(s)}} \\
\le & \quad \tfrac{C}{\sqrt{\mathbb{P}(s)}}.
\end{aligned}$$

□

Recall that $\log\zeta(s)=\mathbb{P}(s)+O(1)$ so $\sqrt{\mathbb{P}(s)}\sim\sqrt{-\log(s-1)}$ as $s\searrow 1$.

10. Large and Moderate Deviations

In this section we establish large and moderate deviations for $\omega_j(X_s)$, $\omega(X_s)$, $\Omega_j(X_s)$, and $\Omega(X_s)$. Our results are the $\zeta(s)$ analogs of those of Mehrdad and Zhu [MZ] which were established for the case of the uniformly distributed Y_N. Due to Proposition 5.2 it becomes particularly transparent that the large deviation functional for $\omega(X_s)$ is the same as for the Poisson distribution. We can establish large and moderate deviation results using the Gärtner-Ellis Theorem, [DZ]. For this we already have the logarithmic moment generating function of $\omega(X_s)$. Define

$$\Lambda_s(t)=\log E\left[e^{t\frac{\omega(X_s)}{\mathbb{P}(s)}}\right].$$

Then, for t in a suitable interval containing 0,

$$\Lambda_s(t)=\mathbb{P}(s)\left(e^{t/\mathbb{P}(s)}-1-\frac{1}{\mathbb{P}(s)}\sum_{k=2}^{\infty}\mathbb{P}(ks)\frac{(1-e^{t/\mathbb{P}(s)})^k}{k}\right).$$

Thus,

$$\lim_{s\searrow 1}\frac{1}{\mathbb{P}(s)}\Lambda_s(\mathbb{P}(s)t)=e^t-1 \tag{55}$$

Then one easily checks that

$$\lim_{s\searrow 1}\frac{1}{\mathbb{P}(s)}\log E\left[e^{\mathbb{P}(s)t\frac{\omega(X_s)}{\mathbb{P}(s)}}\right]=e^t-1. \tag{56}$$

is the analog of Assumption 2.3.2 in [DZ]. The Fenchel-Legendre transform of e^t-1 is

$$\Lambda^*(x)=x\log x-x+1, \tag{57}$$

which satisfies the conditions of Lemma 2.3.9 again in [DZ]. Thus a direct application of the Gärtner-Ellis Theorem yields

Theorem 10.1. For any closed subset F of $\mathbf{R}$,

$$\limsup_{s\searrow 1} \frac{1}{\mathbb{P}(s)} \log P\left(\frac{\omega(X_s)}{\mathbb{P}(s)} \in F\right) \leq -\inf_{x\in F}(x\log x - x + 1)$$

and for any open subset G of $\mathbf{R}$,

$$\liminf_{s\searrow 1} \frac{1}{\mathbb{P}(s)} \log P\left(\frac{\omega(X_s)}{\mathbb{P}(s)} \in G\right) \geq -\inf_{x\in G}(x\log x - x + 1).$$

Using the logarithmic moment generating function of $\omega(X_s)$ we can also derive a moderate deviation principle. In this case, using 27 one has for $\alpha \in (\frac{1}{2}, 1)$,

$$\lim_{s\searrow 1} \mathbb{P}(s)^{1-2\alpha} \log E\left[e^{\frac{\omega(X_s)-\mathbb{P}(s)}{\mathbb{P}^{\alpha}(s)}}\right] = \frac{t^2}{2}. \tag{58}$$

The Fenchel-Legendre transform of this is

$$\Lambda^*(\lambda) = \frac{\lambda^2}{2}.$$

Again, Assumption 2.3.2 in [DZ] holds and the conditions of Lemma 2.3.9 in [DZ] are satisfied. Thus we have the following moderate deviation result.

Theorem 10.2. Let $\alpha \in (\frac{1}{2}, 1)$. For any closed subset F of $\mathbf{R}$,

$$\limsup_{s\searrow 1} \frac{1}{\mathbb{P}^{2\alpha-1}(s)} \log P\left(\frac{\omega(X_s)-\mathbb{P}(s)}{\mathbb{P}^{\alpha}(s)} \in F\right) \leq -\inf_{x\in F}\frac{x^2}{2}$$

and for any open subset G of $\mathbf{R}$,

$$\liminf_{s\searrow 1} \frac{1}{\mathbb{P}^{2\alpha-1}(s)} \log P\left(\frac{\omega(X_s)-\mathbb{P}(s)}{\mathbb{P}^{\alpha}(s)} \in G\right) \geq -\inf_{x\in G}\frac{x^2}{2}.$$

Similar large and moderate deviation results can be easily proven for $\Omega(X_s), \omega_j(X_s)$ and $\Omega_j(X_s)$.

11. Dirichlet Series and Biased Sampling

In this section we consider sampling methods related to Dirichlet series. Let $\mathbf{a} = \{a_n : n = 1, 2, 3, \cdots\}$ be a completely multiplicative sequence and for simplicity, assume $a_n \in \mathbf{D} \equiv \{z \in \mathbf{C} : |a_n| \leq 1\}$. The most general form of such completely multiplicative functions are given by starting with a general function $\mathbf{a} : \mathbb{P} \to \mathbf{D}$, simply extend to arbitrary $n = \Pi_{i=1}^k p_i^{j_i}$ by setting $a_n = \Pi_{i=1}^k a_{p_i}^{j_i}$ to obtain a completely multiplicative function defined for all values of n. We will sometimes use the fact that if $\mathbf{a}$ is completely multiplicative then so is the sequence $\mathbf{a^2}$ where $\mathbf{a}^2 = (a_n^2)$. One example will be $a_p = a \in [-1, 1]$, $p \in \mathcal{P}$. This leads to the completely multiplicative function $a_m = a^{\Omega(m)}$. Another example would be given by taking $0 \leq a < b \leq 1$ and setting $a_p = a$ if $p \equiv 1 \mod 4$ and $a_p = b$ if $p \equiv 3 \mod 4$. Then for general m we have $a_m = a^{\Omega_1(M)} b^{\Omega_3(m)}$, where $\Omega_i(m)$ is the number of prime factors of m congruent to $i \mod 4$. Such multiplicative sequences determine a Dirichlet series defined by

$$L(\mathbf{a}, s) = \sum_{n\geq 1} \frac{a_n}{n^s} = \Pi_p \left(1 - \frac{a_p}{p^s}\right)^{-1}.$$

When the a_n are non-negative real numbers, consider sampling a positive integer, $X_s^{\mathbf{a}}$, according to the following distribution

$$P(X_s^{\mathbf{a}} = n) = \frac{1}{L(\mathbf{a}, s)} \frac{a_n}{n^s}. \tag{59}$$

We call this distribution the $\zeta(\mathbf{a}, s)$ distribution.

In the case of the example $a_n = a^{\Omega(n)}$, $a > 0$, the integer is sampled with the Riemann zeta distribution with a size bias that discriminates against integers with more prime factors. For general completely multiplicative $\mathbf{a}$ we shall say $X_s^{\mathbf{a}}$ has the $\zeta(\mathbf{a}, s)$ distribution. In the case of the example $a_n = a^{\Omega_1(n)} b^{\Omega_3(n)}$, the integer is sampled with the Riemann zeta distribution with a bias that prefers numbers with prime factors congruent to $3 \mod 4$ over prime factors congruent to $1 \mod 4$. The completely multiplicative property of $\{a_n\}$ allows a similar analysis to the case $a_n \equiv 1$ carried out previously.

We now give four examples of results using Dirichlet series in determining statistical properties of the $\zeta(s)$ distributed random variable X_s.

- Our first example of how Dirichlet series come into play is in considering the parity of $\Omega(X_s)$. Define $a_n = (-1)^{\Omega(n)}$. Then $a_{mn} = a_m a_n$ for all pairs of positive integers m, n. Thus, for this choice of $\mathbf{a}$,

$$\begin{aligned} E[(-1)^{\Omega(X_s)}] = & \quad \tfrac{1}{\zeta(s)} \textstyle\sum_{n \geq 1} \frac{(-1)^{\Omega(n)}}{n^s} \\ = & \quad \Pi_p \left(1 - \tfrac{a_p}{p^s}\right)^{-1} \\ = & \quad \Pi_p \left(1 + \tfrac{1}{p^s}\right)^{-1} \\ = & \quad \tfrac{\zeta(2s)}{\zeta^2(s)}. \end{aligned}$$

But $E[(-1)^{\Omega(X_s)}] = P(\Omega(X_s) \text{ is even}) - P(\Omega(X_s) \text{ is odd})$ so we obtain the result

$$\begin{aligned} P(\Omega(X_s) \text{ is even}) = & \quad P(\Omega(X_s) \text{ is odd}) + \tfrac{\zeta(2s)}{\zeta^2(s)} \\ \sim & \quad P(\Omega(X_s) \text{ is odd}) + \tfrac{\pi^2}{6}(s-1)^2,\ s \searrow 1. \end{aligned}$$

- Our second example arises in the context of "square full" numbers. A number n is called square full if it satisfies $p|n$ implies $p^2|n$. That is every prime factor of n appears at least twice in the prime factorization of n. Using Dirichlet series we can show

$$P(X_s \text{ is square full}) = \frac{\zeta(2s)\zeta(3s)}{\zeta(s)\zeta(6s)}.$$

Indeed, the multiplicative function we need for this is simply $a_n = 1$ when n is square full and $a_n = 0$ otherwise. Then a_n is a multiplicative function and

$$\begin{aligned}
P(X_s \text{ is square full}) &= \tfrac{1}{\zeta(s)} \textstyle\sum_{n\geq 1} \tfrac{a_n}{n^s} \\
&= \tfrac{1}{\zeta(s)} \Pi_p \left(1 + \textstyle\sum_{m\geq 1} \tfrac{a_{p^m}}{p^{ms}}\right) \\
&= \tfrac{1}{\zeta(s)} \Pi_p \left(1 + \tfrac{1}{p^{2s}} + \tfrac{1}{p^{3s}} + \tfrac{1}{p^{4s}} + \cdots\right) \\
&= \tfrac{1}{\zeta(s)} \Pi_p \left(1 + \tfrac{1}{p^{2s}} \textstyle\sum_{m\geq 0} \tfrac{1}{p^{ms}}\right) \\
&= \tfrac{1}{\zeta(s)} \Pi_p \left(1 + \tfrac{1}{p^{2s}} (1 - \tfrac{1}{p^s})^{-1}\right) \\
&= \tfrac{1}{\zeta(s)} \Pi_p \left(1 + \tfrac{1}{p^s(p^s-1))}\right) \\
&= \tfrac{1}{\zeta(s)} \Pi_p \left(\tfrac{p^{2s}-p^s+1}{p^s(p^s-1))}\right) \\
&= \tfrac{1}{\zeta(s)} \Pi_p \left(\tfrac{p^{3s}+1}{(p^s+1)p^s(p^s-1))}\right) \\
&= \tfrac{1}{\zeta(s)} \Pi_p \left(\tfrac{(p^{3s}+1)p^{-3s}}{(p^{2s}-1)p^{-2s}}\right) \\
&= \tfrac{1}{\zeta(s)} \Pi_p \tfrac{1+\frac{1}{p^{3s}}}{1-\frac{1}{p^{2s}}} \\
&= \tfrac{1}{\zeta(s)} \Pi_p \tfrac{1-\frac{1}{p^{6s}}}{(1-\frac{1}{p^{3s}})(1-\frac{1}{p^{2s}})} \\
&= \tfrac{\zeta(2s)\zeta(3s)}{\zeta(s)\zeta(6s)}.
\end{aligned}$$

- Our third example involves the statistics of square free numbers. One can easily express $P(X_s \text{ is square free})$ using multiplicative functions and Dirichlet series. For this we take μ to be the Möbius function. That is $\mu(n) = (-1)^k$ if n is the product of k distinct primes and otherwise, $\mu(n) = 0$. Then $\mu(n)$ is a multiplicative function, as is μ^2 which becomes the indicator function for square free numbers. So,

$$\begin{aligned}
P(X_s \,\text{is}\,\text{square}\,\text{free}) =& E[\mu^2(X_s)] \\
=& \frac{1}{\zeta(s)} \sum_{n\geq 1} \frac{\mu^2(n)}{n^s} \\
=& \frac{1}{\zeta(s)} \Pi_p \left(1 + \frac{\mu^2(p)}{p^s}\right) \\
=& \frac{1}{\zeta(s)} \Pi_p \frac{1 - \frac{1}{p^{2s}}}{1 - \frac{1}{p^s}} \\
=& \frac{1}{\zeta(2s)}.
\end{aligned} \tag{60}$$

We can't refrain from noting that

$$P(X_s \text{ is square free}) \sim \frac{6}{\pi^2},\ s \searrow 1.$$

We now examine some generic properties of the $\zeta(\mathbf{a},\mathbf{s})$ distribution when the sequence $\mathbf{a}$ is completely multiplicative. First,

$$\begin{aligned} P(k|X_s^{\mathbf{a}}) =& \frac{1}{L(\mathbf{a},s)}\sum_{k|n}\frac{a_n}{n^s} \\ =& \frac{1}{L(\mathbf{a},s)}\sum_{n\geq 1}\frac{a_{nk}}{(nk)^s} \\ =& \frac{a_k}{k^s}. \end{aligned} \tag{61}$$

From this we derive the crucial property that for $p, q \in \mathcal{P}$, there is the independence

$$\begin{aligned} P(p|X_s^{\mathbf{a}}, q|X_s^{\mathbf{a}}) = & \quad P(pq|X_s^{\mathbf{a}}) \\ = & \quad \tfrac{a_{pq}}{(pq)^s} \\ = & \quad \tfrac{a_p}{p^s}\tfrac{a_q}{q^s} \\ = & \quad P(p|X_s^{\mathbf{a}})P(q|X_s^{\mathbf{a}}). \end{aligned}$$

We have established the following proposition.

Proposition 11.1. Let $X_s^{\mathbf{a}}$ be defined as in (59). Then $X_s^{\mathbf{a}}$ has a prime factorization

$$X_s^{\mathbf{a}} = \Pi_p p^{r_p(s)}$$

where the random variables $\{r_p(s) : p \in \mathcal{P}\}$ are independent geometrically distributed with parameter $(1 - \frac{a_p}{p^s})$,

$$P(r_p(s) \geq k) = \left(\frac{a_p}{p^s}\right)^k.$$

As a consequences of Proposition 11.1, we have

$$\begin{aligned} E[X_s^{\mathbf{a}}] = & \quad \tfrac{1}{L(\mathbf{a},s)}\textstyle\sum_{n\geq 1}\tfrac{na_n}{n^s} \\ = & \quad \tfrac{L(\mathbf{a},s-1)}{L(\mathbf{a},s)} \end{aligned}$$

from which we deduce that $X_s^{\mathbf{a}}$ has a finite expectation at least for $s > 2$ which is the case for $a_n \equiv 1$. Define

$$\mathbb{P}_{\mathbf{a}}(s) = \sum_p \frac{a_p}{p^s}.$$

Theorem 11.2.

$$E[\omega(X_s^{\mathbf{a}})] = \mathbb{P}_{\mathbf{a}}(s). \tag{62}$$

$$E\left[\omega(X_s^{\mathbf{a}})^2\right] = \mathbb{P}_{\mathbf{a}}(s)^2 - \mathbb{P}_{\mathbf{a}^2}(2s) + \mathbb{P}_{\mathbf{a}}(s). \tag{63}$$

$$var(\omega(X_s^{\mathbf{a}})) = \mathbb{P}_{\mathbf{a}}(s) - \mathbb{P}_{\mathbf{a}^2}(2s). \tag{64}$$

$$E[\Omega(X_s^{\mathbf{a}})] = \sum_p \frac{a_p}{p^s - a_p}. \tag{65}$$

For $t < \ln 1$ and $s > 1$, if f is a bounded strongly additive function such that $f(p^m) = f(p),\ p \in \mathcal{P},\ m \geq 1$, then

$$Ee^{tf(X_s^{\mathbf{a}})} = \Pi_p \left((1 - \frac{a_p}{p^s}) \sum_{m \geq 0} \frac{a_p e^{tf(p)}}{p^{sm}} \right) \tag{66}$$

and

$$Ee^{t\omega(X_s^{\mathbf{a}})} \quad = \exp\{(e^t - 1)\mathbb{P}_a(s)\} \exp\{-\textstyle\sum_{m\geq 2} \frac{(1-e^t)^m}{m} \mathbb{P}_a(ms)\} \tag{67}$$

For $t < 0$,

$$Ee^{t\Omega(X_s^{\mathbf{a}})} \quad = \exp\left\{\left(e^t - 1\right)\mathbb{P}_a(s)\right\} \exp\left\{\textstyle\sum_{m\geq 2} \left(\frac{e^{mt}-1}{k}\right) \mathbb{P}_a(ms)\right\}. \tag{68}$$

For $t < \ln 2$

$$Ee^{t(\Omega(X_s^{\mathbf{a}}) - \omega(X_s^{\mathbf{a}}))} \quad = \Pi_p \left(1 - \tfrac{a_p}{p^s}\right)\left(1 + \tfrac{a_p}{p^s - a_p e^t}\right). \tag{69}$$

Proof. The proof of this Theorem is a minor variation of the proof of Theorem 32 and we place the computations in the Appendix for the interested reader. □

We observe, that as in the case $a_p \equiv 1$ the random variables $\Omega(X_s^{\mathbf{a}}) - \omega(X_s^{\mathbf{a}})$ have a non-degenerate limiting distribution as $s \searrow 1$.

We now turn to the decomposition of Lloyd [L] in the present context. Since the random variables $\{r_p(s) : p \in \mathcal{P}\}$ are independent geometrically distributed with parameter $(1 - \frac{a_p}{p^s})$, this results in

$$Ee^{tr_p(s)} = \frac{1 - \frac{a_p}{p^s}}{1 - \frac{a_p e^t}{p^s}}. \tag{70}$$

Dividing by the moment generating function of a Poisson random variable $u_p(s)$ with parameter $\frac{a_p}{p^s}$ gives

$$\frac{1 - \frac{a_p}{p^s}}{(1 - \frac{a_p e^t}{p^s}) e^{(e^t - 1)\frac{a_p}{p^s}}} = \sum_{m=1}^{\infty} \lambda_m^{\mathbf{a}}(s) e^{mt} \tag{71}$$

where

$$\lambda_m^{\mathbf{a}}(s) = (1 - \frac{a_p}{p^s}) e^{\frac{a_p}{p^s}} (\frac{a_p}{p^s})^m \sigma(m),\ m = 0, 1, 2, \cdots$$

and

$$\sigma(m) = \sum_{j=0}^{m} \frac{(-1)^j}{j!}.$$

Three important observations are

- $\{\lambda_m^{\mathbf{a}}(s) : m = 0, 1 \cdots\}$ is a probability mass function on $\mathbf{Z}^+$ for each $s \geq 1$ which is non-degenerate at $s = 1$.
- $\lambda_1^{\mathbf{a}}(s) \equiv 0$
- $$\lim_{s \searrow 1} \lambda_m^{\mathbf{a}}(s) = (1 - \frac{a_p}{p}) e^{\frac{a_p}{p}} \frac{a_p}{p^m} \sigma(m),\ m = 0, 1, 2, \cdots$$

  is a well-defined probability distribution on $\mathbf{N}$.

We can thus construct two independent sequences, such that for each $p \in \mathcal{P}$ and $s > 1$ we may write

$$r_p(s) \stackrel{d}{=} u_p(s) + v_p(s),$$

where the sequences $\{u_p(s) : p \in \mathcal{P}\}$ and $\{v_p(s) : p \in \mathcal{P}\}$ are also independent, $u_p(s)$ is Poisson with parameter $\frac{a_p}{p^s}$ and $P(v_p(s) = m) = \lambda_m(s)$. Writing

$$M_s^{\mathbf{a}} = \prod_p p^{u_p(s)}$$

and

$$N_s^{\mathbf{a}} = \prod_p p^{v_p(s)}.$$

We then have the following Dirichlet series variant of an observation due to Lloyd [L].

Proposition 11.3. For any $s > 1$, there exist independent random variables $M_s^{\mathbf{a}}$, $N_s^{\mathbf{a}} \in \mathbf{N}$ such that

$$\begin{aligned}
X_s^{\mathbf{a}} &\stackrel{d}{=} M_s^{\mathbf{a}} N_s^{\mathbf{a}}, \\
\Omega(X_s^{\mathbf{a}}) &= \Omega(M_s^{\mathbf{a}}) + \Omega(N_s^{\mathbf{a}}), \\
\Omega(M_s^{\mathbf{a}}) &= \textstyle\sum_{p\in\mathcal{P}} u_p(s) \text{ is a Poisson random variable with parameter } \mathbb{P}_{\mathbf{a}}(s), \\
E\left[e^{t\Omega(M_s^{\mathbf{a}})}\right] &= \exp\left\{\left(e^t - 1\right)\mathbb{P}_{\mathbf{a}}(s)\right\}, \\
E\left[e^{t\Omega(N_s^{\mathbf{a}})}\right] &= \exp\left\{\textstyle\sum_{k=2}^{\infty}\left(\frac{e^{kt}-1}{k}\right)\mathbb{P}_{\mathbf{a}}(ks)\right\}, \\
\Omega(N_s^{\mathbf{a}}) & \text{ has a limiting nondegenerate distribution as } s \searrow 1.
\end{aligned}$$

Proof. We'll give the computation for the moment generating function of $\Omega(N_s^{\mathbf{a}})$. Starting with 71, since $\Omega(N_s^{\mathbf{a}}) = \sum_p v_p(s)$,

$$\begin{aligned}
E\left[e^{t\Omega(N_s^{\mathbf{a}})}\right] &= \Pi_p \frac{1 - \frac{a_p}{p^s}}{(1 - \frac{a_p e^t}{p^s}) e^{(e^t-1)\frac{1}{p^s}}} \\
&= \exp\left\{\sum_p \left(\log(1 - \frac{a_p}{p^s}) - \log(1 - \frac{a_p e^t}{p^s}) - (e^t - 1)\frac{a_p}{p^s}\right)\right\} \\
&= \exp\left\{\sum_p \left(-\sum_{m\ge1} \frac{a_p}{mp^{ms}} + \sum_{m\ge1} \frac{e^{mt}}{mp^{ms}} - (e^t - 1)\frac{a_p}{p^s}\right)\right\} \qquad (72) \\
&= \exp\left\{-\sum_{m\ge1} \frac{\mathbb{P}_{\mathbf{a}}(ms)}{m} + \sum_{m\ge1} \frac{e^{mt}\mathbb{P}_{\mathbf{a}}(ms)}{m} - (e^t - 1)\mathbb{P}_{\mathbf{a}}(s)\right\} \\
&= \exp\left\{\sum_{m\ge2} \frac{(e^{mt} - 1)\mathbb{P}_{\mathbf{a}}(ms)}{m}\right\}
\end{aligned}$$

Finally,

$$\lim_{s\searrow1} E\left[e^{t\Omega(N_s)}\right] = \exp\left\{\sum_{m\ge2} \frac{(e^{mt} - 1)\mathbb{P}_{\mathbf{a}}(m)}{m}\right\}$$

and the right hand side is well defined since $|a_p| \leq 1$ so the random variables $\Omega(N_s)$ converge in distribution as $s \searrow 1$.

□

We see by 11.3 that $\Omega(X_s^{\mathbf{a}}) - \Omega(M_s^{\mathbf{a}}) = \Omega(N_s^{\mathbf{a}})$ has a limiting distribution as $s \searrow 1$. Consequently, $\Omega(X_s^{\mathbf{a}})$ is nearly a Poisson random variable with parameter $\mathbb{P}_{\mathbf{a}}(s)$.

The $gcd(X_s^{\mathbf{a}}, Y_s^{\mathbf{a}})$ where $X_s^{\mathbf{a}}$ and $Y_s^{\mathbf{a}}$ are independent has the $\zeta(\mathbf{a}^2, 2s)$ distribution. To see this, write $X_s^{\mathbf{a}} = \Pi_p p^{r_p(s)}$ and $Y_s^{\mathbf{a}} = \Pi_p p^{u_p(s)}$, where $\{u_p(s) : p \in \mathcal{P}\} \stackrel{d}{=} \{r_p(s) : p \in \mathcal{P}\}$ and the two sequences are independent. Then one can express $gcd(X_s^{\mathbf{a}}, Y_s^{\mathbf{a}}) = \Pi_p p^{r_p(s) \wedge u_p(s)}$. But, $r_p(s) \wedge u_p(s)$ is a geometric random variable with parameter $\frac{a_p^2}{p^{2s}}$. This establishes the claim that $gcd(X_s^{\mathbf{a}}, Y_s^{\mathbf{a}})$ has the $\zeta(\mathbf{a}^2, 2s)$ distribution. This corresponds to our earlier result in the case $a_p \equiv 1$. As a consequence, $\lim_{s \searrow 1} P(gcd(X_s^{\mathbf{a}}, Y_s^{\mathbf{a}}) = 1) = \left(\sum \frac{a_n^2}{n^2}\right)^{-1} = \frac{1}{L(\mathbf{a^2}, 2)}$. This extends easily to $gcd(X_s^{\mathbf{a},1}, \cdots, X_s^{\mathbf{a},k})$ for k independent copies of the $\zeta(\mathbf{a}, s)$ random variables, the gcd has a $\zeta(\mathbf{a}^k, ks)$ distribution.

The Erdös-Kac Central Limit Theorem holds for $\omega(X_s^{\mathbf{a}})$ in the following form

Theorem 11.4.

$$\lim_{s \searrow 1} P\left(\frac{\omega(X_s^{\mathbf{a}}) - \mathbb{P}_{\mathbf{a}}(s)}{\sqrt{\mathbb{P}_{\mathbf{a}}(s)}} \leq b\right) = \frac{1}{2\pi} \int_{-\infty}^{b} e^{-\frac{x^2}{2}} dx. \tag{73}$$

The proof is a minor variant of our earlier version of this theorem for the case $a_n \equiv 1$ using the moment generating function for $\omega(X_s^{\mathbf{a}})$. One can also use Lloyd's decomposition, the CLT holds for $(\Omega(M_s^{\mathbf{a}}) - \mathbb{P}_{\mathbf{a}}(s))/\sqrt{\mathbb{P}_{\mathbf{a}}(s)}$ and $(\omega(X_s^{\mathbf{a}}) - \Omega(M_s^{\mathbf{a}}))/\sqrt{\mathbb{P}_{\mathbf{a}}(s)}$ converges to 0 in distribution as $s \searrow 1$. The Berry-Esseen Theorem as well as Large and Moderate Deviation results follow in essentially as in the case $a_n \equiv 1$.

12. Harmonic and Uniform Distribution, Tauberian theorems

Recall the harmonic distributed random variable Z_N has the distribution $P(Z_N = n) = \frac{1}{\alpha_N n}$ where $\alpha_N = \sum_{n=1}^{N} \frac{1}{n}$, $n \in [N]$. Using a Tauberian Theorem it's possible to transfer asymptotic results about the $\zeta(s)$ random variable as $s \searrow 1$ to the same result for the harmonic distribution as $N \to \infty$. The relevant Tauberian Theorem, found in Diaconis [D], is the following.

Theorem 12.1. Let $\{c_n\}$ be a bounded sequence of complex numbers. Then

$$\lim_{s \searrow 1} \frac{1}{\zeta(s)} \sum_{n \geq 1} \frac{c_n}{n^s} = c$$

if and only if

$$\lim_{N \to \infty} \frac{1}{\alpha_N} \sum_{n=1}^{N} \frac{c_n}{n} = c.$$

(**NB** : the value of c is the same in both cases.)

Translating this to probabilistic language, we have

Corollary 12.2. For X_s and Z_N, the $\zeta(s)$ and harmonic distributed random variables respectively, the existence of

$$\lim_{s\searrow 1} E[c_{X_s}] = c$$

is equivalent to the existence of the limit

$$\lim_{N\to\infty} E[c_{Z_N}] = c$$

with the same value of c

We cite some examples.

- Let $c_n = \frac{\phi(n)}{n}$ where we've established $\lim_{s\searrow 1} E[\frac{\phi(X_s)}{X_s}] = \frac{6}{\pi^2}$ so it follows then that $\lim_{N\to\infty} E[\frac{\phi(Z_N)}{Z_N}] = \frac{6}{\pi^2}$ as well.
- Another example is gotten by taking $c_n = e^{it(\Omega(n)-\omega(n))}$. This sequence is bounded and 69 gives

  $$\lim_{s\searrow 1} Ee^{it(\Omega(X_s)-\omega(X_s))} = \Pi_p\left(1-\frac{1}{p}\right)\left(1+\frac{1}{p-e^{it}}\right).$$

  Consequently,

  $$\lim_{s\searrow 1} Ee^{it(\Omega(Z_N)-\omega(Z_N))} = \Pi_p\left(1-\frac{1}{p}\right)\left(1+\frac{1}{p-e^{it}}\right).$$

- Recalling the result of Renyi [MV], we see that the limiting distributions of $\Omega(X_s)-\omega(X_s)$ and $\Omega(Z_N)-\omega(Z_N)$ are identical.
- Another result along these lines involves $P(Z_N \text{ is square free})$. In the previous section at 60 we showed $\lim_{s\searrow 1} E[\mu^2(X_s)] = \lim_{s\searrow 1} P(X_s \text{ is square free}) = \frac{6}{\pi^2}$. This immediately gives that $\lim_{N\to\infty} P(Z_N \text{ is square free}) = \frac{6}{\pi^2}$.

We can also extend asymptotic results for the $\zeta(s)$ distribution to the uniform by means of another Tauberian Theorem. For this, let $L(\mathbf{a}, s) = \sum_{n\geq 1}\frac{a_n}{n^s}$ as before where $a_n \geq 0$ and define $A(N) = \sum_{n=1}^N a_n$. Then we have the Wiener-Ikehara Tauberian Theorem [MV],

Theorem 12.3. Assume that

- $L(\mathbf{a}, s)$ extends to an analytic function in an open set containing

  $$\{Re\, s \geq 1\}\setminus\{1\}$$

  for which $\lim_{s\searrow 1}\frac{L(\mathbf{a},s)}{\zeta(s)} = \alpha$.
- There is a constant C such that for all $N \geq 1$,

  $$A(N) \leq CN.$$

- $L(\mathbf{a}, s)$ converges for $Re\, s > 1$.

Then

$$\lim_{N\to\infty}\frac{1}{N}A(N) = \alpha.$$

We can restate Theorem 12.3 in probabilistic language as follows.

Corollary 12.4. Suppose $\mathbf{a} : \mathbf{Z}^+ \to \mathbf{R}$ is a bounded multiplicative function and assume that $L(\mathbf{a}, s)$ extends to an analytic function in an open set containing

$$\{Re\, s \geq 1\}\setminus\{1\}$$

If for some $c \in \mathbf{R}$, $\lim_{s\searrow 1} E[a_{X_s}] = c$ then $\lim_{N\to\infty} E[a_{Y_N}] = c$.

Using Corollary 12.4 we can derive asymptotic results for uniformly sampled integers from the companion result for integers sampled according to the $\zeta(s)$ distribution. The advantage of this is that computations for the $\zeta(s)$ distribution are easier than they are for the uniform distribution. So, one can make the computation for the $\zeta(s)$ random variable and transfer the result to the uniform random variable by means of the Tauberian Theorem. We consider some examples.

- Take $a_n = \frac{\phi(n)}{n}$. We proved that

$$\lim_{s\searrow 1} E[a_{X_s}] = \lim_{s\searrow 1} \frac{1}{\zeta(s)} \sum_{n\geq 1} \frac{a_n}{n^s} = \frac{6}{\pi^2}.$$

  Thus, by Theorem 12.3 we have

$$\lim_{N\to\infty} E[a_{Y_N}] = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} a_n = \frac{6}{\pi^2}.$$

  The latter is a result of Kac [K]. The proof for the $\zeta(s)$ distribution was much easier than the proof in the uniform case. We can also consider $a_n = e^{it\frac{\phi(n)}{n}}$ for which we proved that

$$\lim_{s\searrow 1} E\left[e^{i\tau \log \frac{\phi(X_s)}{X_s}}\right] = \Pi_p(1-\frac{1}{p})\left(1+(1-\frac{1}{p})^{i\tau}\left((1-\frac{1}{p})^{-1}-1\right)\right).$$

  Thus,

$$\lim_{N\searrow\infty} E\left[e^{i\tau \log \frac{\phi(Y_N)}{Y_N}}\right] = \Pi_p(1-\frac{1}{p})\left(1+(1-\frac{1}{p})^{i\tau}\left((1-\frac{1}{p})^{-1}-1\right)\right)$$

  and so $\frac{\phi(X_s)}{X_s}$ and $\frac{\phi(Y_N)}{Y_N}$ have the same limiting distribution.
- We can also apply this to the sequence $a_n = e^{-t(\Omega(n)-\omega(n))}, t > 0$. The Tauberian Theorem 12.3 then implies

$$\lim_{N\to\infty} E[a_{Y_N}] = \Pi_p\left(1-\frac{1}{p}\right)\left(1+\frac{1}{p-e^{-t}}\right).$$

  As mentioned earlier, this asymptotic result for the sequence Y_N is originally due to Renyi [MV]. It's proof in the X_s was a straightforward calculation.
- Finally, we mention the application to square free numbers, and for this we take $a_n = \mu^2(n)$. Then, application of the Tauberian Theorem 12.3 gives

$$\lim_{N\to\infty} E[a_{Y_N}] = \lim_{N\to\infty} P(Y_N \text{ is square free}) = \frac{6}{\pi^2}.$$

13. Appendix

Simple variants of earlier computations reveal the following:

$$\begin{aligned} E[\omega(X_s^{\mathbf{a}})] &= \sum_p E[1_{r_p(s)>0}] \\ &= \sum_p \frac{a_p}{p^s} \\ &\equiv \mathbb{P}_{\mathbf{a}}(s). \end{aligned}$$

$$
\begin{aligned}
E\omega(X_s^{\mathbf{a}})^2 &= E\left(\sum_{p\in\mathcal{P}} 1_{r_p(s)>0}\right)^2 \\
&= E\left(\sum_{p,q\in\mathcal{P}} 1_{r_p(s)>0}1_{r_q(s)>0}\right) \\
&= E\left(\sum_{p\neq q\in\mathcal{P}} 1_{r_p(s)>0}1_{r_q(s)>0}\right) + E\left(\sum_{p\in\mathcal{P}} 1_{r_p(s)>0}\right) \\
&= \sum_{p\neq q\in\mathcal{P}} \frac{a_p a_q}{p^s q^s} + \sum_{p\in\mathcal{P}} \frac{a_p}{p^s} \\
&= \left(\sum_{p\in\mathcal{P}} \frac{a_p}{p^s}\right)^2 - \sum_{p\in\mathcal{P}} \frac{a_p^2}{p^{2s}} + \sum_{p\in\mathcal{P}} \frac{a_p}{p^s} \\
&= \mathbb{P}_{\mathbf{a}}(s)^2 - \mathbb{P}_{\mathbf{a}^2}(2s) + \mathbb{P}_{\mathbf{a}}(s).
\end{aligned} \tag{74}
$$

Consequently

$$
var(\omega(X_s^{\mathbf{a}})) = \mathbb{P}_{\mathbf{a}}(s) - \mathbb{P}_{\mathbf{a}^2}(2s). \tag{75}
$$

$$
\begin{aligned}
E[\Omega(X_s^{\mathbf{a}})] &= \sum_p E[r_p(s)] \\
&= \sum_p \sum_{k\geq 1} P(r_p(s) \geq k) \\
&= \sum_p \sum_{k\geq 1} \left(\frac{a_p}{p^s}\right)^k \\
&= \sum_p \frac{a_p}{p^s - a_p}.
\end{aligned} \tag{76}
$$

We compute the moment generating functions for the random variables just as in the case $a_n \equiv 1$.

$$
\omega(X_s^{\mathbf{a}}) = \sum_p 1_{r_p(s)>0}
$$

and

$$
\Omega(X_s^{\mathbf{a}}) = \sum_p r_p(s)
$$

Recall that

$$
\mathbb{P}_a(s) = \sum_p \frac{a_p}{p^s},
$$

and then

$$
\begin{aligned}
Ee^{t\omega(X_s^{\mathbf{a}})} &= \Pi_p(1-\tfrac{a_p}{p^s})\sum_{m\geq 0}\frac{a_p^m e^{t\omega(p^m)}}{p^{sm}}\\
&= \Pi_p(1-\tfrac{a_p}{p^s})\left(1+\sum_{m\geq 1}\frac{a_p^m e^t}{p^{sm}}\right)\\
&= \Pi_{p\in\mathcal{P}}\left(e^t\tfrac{a_p}{p^s}+(1-\tfrac{a_p}{p^s})\right)\\
&= \Pi_{p\in\mathcal{P}}\left((e^t-1)\tfrac{a_p}{p^s}+1\right)\\
&= \exp\{\sum_{p\in\mathcal{P}}\log\left((e^t-1)\tfrac{a_p}{p^s}+1\right)\}\\
&= \exp\{\sum_{p\in\mathcal{P}}\sum_{m=1}^{\infty}(-1)^{m+1}\tfrac{(e^t-1)^m}{m}\tfrac{a_p^m}{p^{ms}}\}\\
&= \exp\{\sum_{m=1}^{\infty}(-1)^{m+1}\tfrac{(e^t-1)^m}{m}\sum_{p\in\mathcal{P}}\tfrac{a_p^m}{p^{ms}}\}\\
&= \exp\{\sum_{m=1}^{\infty}(-1)^{m+1}\tfrac{(e^t-1)^m}{m}\mathbb{P}_a(ms)\}\\
&= \exp\{(e^t-1)\mathbb{P}_a(s)\}\exp\{-\sum_{m=2}^{\infty}\tfrac{(1-e^t)^m}{m}\mathbb{P}_a(ms)\}.
\end{aligned}
$$

In a similar fashion, for $t<0$,

$$
\begin{aligned}
Ee^{t\Omega(X_s^{\mathbf{a}})} &= \Pi_p(1-\tfrac{a_p}{p^s})\sum_{m\geq 0}\frac{a_p^m e^{t\Omega(p^m)}}{p^{sm}}\\
&= \Pi_p(1-\tfrac{a_p}{p^s})\sum_{m\geq 0}\frac{(a_p e^t)^m}{p^{sm}}\\
&= \Pi_{p\in\mathcal{P}}(1-\tfrac{a_p}{p^s})(1-\tfrac{a_e^t}{p^s})^{-1}\\
&= \exp\left\{\sum_{p\in\mathcal{P}}\left(\ln(1-\tfrac{a_p}{p^s})-\ln(1-\tfrac{a_p e^t}{p^s})\right)\right\}\\
&= \exp\left\{\sum_{p\in\mathcal{P}}\sum_{m=1}^{\infty}\left(\frac{(a_p e^t)^m}{mp^{ms}}-\frac{a_p^m}{mp^{ms}}\right)\right\}\\
&= \exp\left\{\sum_{p\in\mathcal{P}}\sum_{m=1}^{\infty}\frac{a_p^m(e^{mt}-1)}{mp^{ms}}\right\}\\
&= \exp\left\{\sum_{m=1}^{\infty}\left(\frac{e^{mt}-1}{m}\right)\mathbb{P}_a(ms)\right\}\\
&= \exp\left\{\left(e^t-1\right)\mathbb{P}_a(s)\right\}\exp\left\{\sum_{m=2}^{\infty}\left(\frac{e^{mt}-1}{m}\right)\mathbb{P}_a(ms)\right\}.
\end{aligned}
$$

Again, similar to the computation for $a_n\equiv 1$,

$$
\begin{aligned}
Ee^{t(\Omega(X_s^{\mathbf{a}})-\omega(X_s^{\mathbf{a}}))} &= \Pi_p(1-\tfrac{a_p}{p^s})\textstyle\sum_{m\geq 0}\frac{a_p^m e^{t(\Omega(p^m)-\omega(p^m))}}{p^{sm}} \\
&= \Pi_p(1-\tfrac{a_p}{p^s})\textstyle\sum_{m\geq 0}\frac{a_p^m e^{t(m-1)^+}}{p^{sm}} \\
&= \Pi_p(1-\tfrac{a_p}{p^s})\left(1+\tfrac{a_p}{p^s}+\textstyle\sum_{m\geq 2}\frac{a_p^m e^{t(m-1)}}{p^{sm}}\right) \\
&= \Pi_p(1-\tfrac{a_p}{p^s})\left(1+\tfrac{a_p}{p^s}+\tfrac{a_p^2 e^t}{p^{2s}}\textstyle\sum_{m\geq 0}\left(\frac{a_p e^t}{p^s}\right)^m\right) \\
&= \Pi_p(1-\tfrac{a_p}{p^s})\left(1+\tfrac{a_p}{p^s}+\frac{a_p^2 e^t}{p^{2s}(1-\frac{a_p e^t}{p^s})}\right) \\
&= \Pi_p\left(1-\tfrac{a_p}{p^s}\right)\left(1+\tfrac{a_p}{p^s-a_p e^t}\right).
\end{aligned}
$$

References

[DZ] Dembo, A., Zeitouni, O. *Large Deviations Techniques and Applications.* Jones and Bartlett, 1992

[D] Durrett, R. *Probability: Theory and Examples.* Duxbury Press, 2nd edition, 1995

[E] Elliott, P.D.T.A. *Probabilistic number theory I. Springer Verlag, N.Y. 1980*

[EK] Erdös, P., Kac, M. *The Gaussian law of errors in the theory of additive number theoretic functions. Amer. J. Math., 62, 1940, 738-742.*

[MZ] Mehrdad, B., Zhu, L. *Moderate and large deviations for the Erdös-Kac theorem. The Quarterly Journal of Mathematics, Volume 67, Issue 1, March 2016, Pages 147?160.*

[Ga] Gauss, C. *Disquisitiones Arithmeticae.* Yale University Press, 1966.

[G] Gut, A. *Some remarks on the Riemann zeta distribution. Rev. Roumaine Math. Pures et Appl. 51 (2006), 205-217.*

[H] Halberstam, H. *On the Distribution of Additive Number-Theoretic Functions (II).* The Journal of the London Mathematical Society, Vol. 31, Part 1, January, 1956, No. 121

[K] Kac, M. *Statistical Independence in Probability, Analysis and Number Theory Carus Mathematical Monographs, Mathematical Association of America, 1959.*

[L] Lloyd, S. *Ordered Prime Divisors of a Random Integer.* Annals of Probability, 1984, Vol.12, No. 4, 1205-1212.

[MV] Montgomery, H., Vaughn, R. *Multiplicative Number Theory I. Classical Theory Cambridge University Press, February 2010.*

[RT] Rényi, A., Turán, P. 1. *On a theorem of Erdös-Kac. Acta Arithmetica 4(1958), 71?84.*

[T] Tenenbaum, G. *Introduction to Analytic and Probabilistic Number Theory.* AMS, Graduate Studies in Mathematics, Vol. 163..

Department of Mathematics, University of California, Irvine, mcransto@uci.edu

Department of Mathematics, University of California, Irvine, apeltzer@uci.edu