The Erdős-Kac Theorem

Vikram Sarkar

July 13, 2025

Abstract

The Erdős-Kac Theorem states that

$$\text{Prob}\left(\frac{\omega(n) - \log\log n}{\sqrt{\log\log n}} \in [a, b] \mid n \in \mathbb{N}\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}\, \mathrm{d}x.$$

In this paper we provide an accessible proof for this renowned theorem involving random variables, moment generating functions, and complex analysis. We provide an introduction to random variables and moment generating functions as well. Our argument is very similar to that of Steve Fan's. [2]

1 Introduction

Let $\omega(n)$ be the number of distinct prime factors of n. For example, $\omega(12) = 2$ and $\omega(210) = 4$. So, a natural question one might ask: what are the asymptotics of $\omega(n)$? Well, roughly speaking,

$$\omega(n) = \sum_{p|n} 1 = \sum_{p<n} \chi_p,$$

where

$$\chi_p = \begin{cases} 0 & p \nmid n \\ 1 & p \mid n \end{cases}.$$

Loosely speaking, the "expected value" of χ_p where $p \ll n$ is $\frac{1}{p}$, so

$$\sum_{p<n} \chi_p \approx \sum_{p<n} \frac{1}{p}.$$

A famous theorem of Mertens ([4]) says that this quantity is about $\log\log n + O(1)$, and therefore $\omega(n)$ should be close to $\log\log n$. A famous theorem of Túran states the following:

Theorem 1.1 (Túran, 1934)**.**

$$\sum_{n \le x} (\omega(n) - \log\log n)^2 = (1 + o(1))x \log\log x.$$

This suggests that $\omega(n) - \log\log n \approx \sqrt{\log\log x}$. $\sqrt{\log\log x} \approx \sqrt{\log\log n}$ for "most" $n \le x$. So, the error is $\approx \sqrt{\log\log n}$ as well. In fact, we can say more. Much more, in fact.

Theorem 1.2 (Erdős-Kac Theorem). *Let $a < b$ be reals. Then,*

$$\lim_{N\to\infty} \frac{1}{N} \cdot \# \left\{ n \leq N \mid a \leq \frac{\omega(n) - \log\log n}{\sqrt{\log\log n}} \leq b \right\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} \mathrm{d}x.$$

In layman's terms, $\frac{\omega(n)-\log\log n}{\sqrt{\log\log n}}$ behaves like a normal distribution. We will now spend the rest of the paper proving this statement using the machinery we have developed before the introduction.

2 Preliminaries

Definition 2.1. Define the **density** of a set $A \subseteq \mathbb{N}$ as

$$d(A) := \lim_{N\to\infty} \frac{|A \cap \{1, 2, \cdots, N\}|}{N}.$$

Simply put, this is the proportion of naturals that A takes up. For example,

- The density of the evens is $\frac{1}{2}$.
- The density of any finite set is 0.
- The density of the perfect squares is 0.
- The density of the composites is 1.
- The density of the squarefree integers is $\frac{6}{\pi^2}$. (!!)

Now, let f, g be two functions from $\mathbb{C}$ to $\mathbb{C}$.

Definition 2.2. We say $f(z) = O(g(z))$ if there exists some constant c such that $|f(z)| \leq c|g(z)|$ for all z for which $g(z) \neq 0$.

Definition 2.3. We say $f(z) = o(g(z))$ if $\lim_{z\to\infty} \frac{f(z)}{g(z)}$ exists and equals zero.

Now, assume f, g are from $\mathbb{N}$ to $\mathbb{R}$.

Definition 2.4. We say $f(n) \ll g(n)$ if $f(n) = O(g(n))$. For constants a, b, we choose $a \ll b$ to mean "if a is sufficiently smaller than b."

Definition 2.5. We say $f(n) = O(g(n))$ if there exists some constant c for which $|f(n)| \leq c|g(n)|$ for all naturals n.

Definition 2.6. We say $f(n) = o(g(n))$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)}$ exists and equals zero.

Now, later, we will talk about functions $f(N, k)$ of two positive integer variables.

Definition 2.7. We say

$$f(N, k) = O(g(N, k))$$

if there exists a constant c, independent of N or k, such that $f(N, k) \leq cg(N, k)$ for sufficiently large N.

Definition 2.8. We define
$$f(N,k) = o(g(N,k))$$
if for all naturals k,
$$\lim_{N\to\infty} \frac{f(N,k)}{g(N,k)} = 0.$$

3 Probability Theory

Usually, probability theory and definitions of probability and whatnot are very complicated and hyper-general (see: the measure theoretic-definition involving sigma-algebras). In this paper, however, the objects of study do not need this general definition, and are more of a special case. Thus, we will try to keep it very simple and understandable.

3.1 Discrete Random Variables

This is the easy part of probability theory.

Definition 3.1. A **discrete random variable** X has a finite sample space S and fixed **probabilities** of being each element of S. For example, S could be $\{1,2,3\}$, and $\text{Prob}(X = 1) = \frac{1}{3}$, $\text{Prob}(X = 2) = \frac{1}{6}$, $\text{Prob}(X = 3) = \frac{1}{2}$. We must have the following:

- $\sum_{s\in S} \text{Prob}(X = s) = 1$
- $\forall s \in S$, $\text{Prob}(X = s) \in [0,1]$.

We will assume that $S \subseteq \mathbb{R}$, since we will only be working with discrete random variables whose sample space is a subset of $\mathbb{R}$.

Definition 3.2. The **expected value** of a discrete random variable X is defined to be
$$\mathbb{E}[X] := \frac{1}{|S|}\sum_{s\in S} s \cdot \text{Prob}(X = s).$$

Definition 3.3. The k^{th} **moment** of a discrete random variable X is defined to be
$$\mathbb{E}[X^k] := \frac{1}{|S|}\sum_{s\in S} s^k \cdot \text{Prob}(X = s).$$

Definition 3.4. Let X be a discrete random variable with sample space S. Let g be a function from S to $\mathbb{R}$. Then,
$$\mathbb{E}[g(X)] := \frac{1}{|S|}\sum_{s\in S} g(s) \cdot \text{Prob}(X = s)$$
and
$$\text{Prob}(g(X) = a) = \sum_{\substack{s\in S\\ g(s)=a}} \text{Prob}(X = s).$$

Definition 3.5. Let X and Y be discrete random variables with sample spaces S and T, respectively, and let $f : (S,T) \to \mathbb{R}$ be a function. Then, define $Z = f(X,Y)$ to have sample space $f(S,T) = \{f(s,t) \mid s \in S, t \in T\}$, and for all $x \in f(S,T)$,

$$\text{Prob}(Z = x) = \sum_{\substack{s\in S, t\in T \\ f(s,t)=x}} \text{Prob}(X = s) \cdot \text{Prob}(Y = t).$$

Definition 3.6. Two discrete random variables X and Y with sample spaces S and T, respectively, are said to be **independent** if

$$\text{Prob}(X = s \text{ and } Y = t) = \text{Prob}(X = s) \cdot \text{Prob}(Y = t)$$

for all $s \in S, t \in T$.

We have the following infamous lemma.

Lemma 3.7. *Let X and Y be discrete random variables with sample spaces S and T, respectively, both in $\mathbb{R}$, and let $c \in \mathbb{R}$. Then,*

- $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$.
- $\mathbb{E}[cX] = c\mathbb{E}[X]$.

Furthermore, if X and Y are independent, then

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y].$$

We state these without proof because they are classical.

3.2 Continuous Random Variables

We will only be talking about continuous random variables that have a probability density function and a sample space that is a subset of $\mathbb{R}$. It is true that there exists continuous random variables without probability density functions, but we need not worry about them,

Definition 3.8. A **continuous random variable** X with an infinite sample space $S \subseteq \mathbb{R}$ and a probability density function f has

$$\text{Prob}(a \le X \le b) = \int_a^b f(x)\, \mathrm{d}x$$

for all $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$. We must have the following:

- f must be piecewise continuous.
- $f \ge 0$ on $\mathbb{R}$.
- $\int_{-\infty}^{\infty} f(x)\, \mathrm{d}x = 1$.

Definition 3.9. The **expected value** of a continuous random variable X with probability density function f is defined to be

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x)\, \mathrm{d}x.$$

Definition 3.10. The k^{th} **moment** of a continuous random variable X with probability density function f is defined to be

$$\mathbb{E}[X^k] := \int_{-\infty}^{\infty} x^k f(x)\,\mathrm{d}x.$$

Definition 3.11. Let X be a continuous random variable with sample space S and probability density function f. Let g be a function from S to $\mathbb{R}$. Then,

$$\mathbb{E}[g(X)] := \int_{-\infty}^{\infty} g(x)f(x)\,\mathrm{d}x.$$

3.3 Moment Generating Functions

Definition 3.12. Suppose X is a discrete random variable with sample space $S \subseteq \mathbb{R}$. Define the **moment generating function** $M_X(z) := \mathbb{E}[e^{zX}]$. First of all, note that

$$M_X(t) = \sum_{s\in S} e^{zs} \cdot \mathrm{Prob}(X = s),$$

so M_X is analytic on $\mathbb{C}$.

Lemma 3.13. *For all* $k \geq 1$, $\frac{\mathrm{d}^k}{\mathrm{d}z^k} M_X(z)\Big|_{z=0} = \mathbb{E}[X^k]$.

Proof. Note that from the Taylor Series expansion of e^x,

$$M_X(z) = \mathbb{E}[e^{zX}] = \mathbb{E}\left[\sum_{k\geq 0} \frac{(zX)^k}{k!}\right] = \sum_{k\geq 0} \mathbb{E}\left[\frac{(zX)^k}{k!}\right] = \sum_{k\geq 0} \frac{\mathbb{E}[X^k]}{k!} z^k.$$

Thus $M_X^{(k)}(0) = \mathbb{E}[X^k]$, as desired. □

4 Simplifications

I claim that it suffices to show that

$$\lim_{N\to\infty} \frac{1}{N} \cdot \#\left\{ n \leq N \mid a \leq \frac{\omega(n) - \log\log N}{\sqrt{\log\log N}} \leq b \right\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}\,\mathrm{d}x.$$

Now I will prove this claim. Suppose we knew the above. Let $N_s = N^{1/\log(\sqrt{\log\log N}+3)}$. Note that for $N_s \leq n \leq N$,

$$\begin{aligned}
\log\log N - \log\log n &\leq \log\log N - \log\log N_s \\
&= \log \frac{\log N}{\log N_s} \\
&= \log \frac{\log N}{\log N \cdot \frac{1}{\log(\sqrt{\log\log N}+3)}} \\
&= \log(\sqrt{\log\log N} + 3) \\
&= o(\sqrt{\log\log N}).
\end{aligned}$$

and for N sufficiently large,

$$\begin{aligned}\sqrt{\log\log N}-\sqrt{\log\log n} &= \frac{\log\log N-\log\log n}{\sqrt{\log\log N}+\sqrt{\log\log n}}\\ &< \log\log N-\log\log n\\ &= o(\sqrt{\log\log N}).\end{aligned}$$

Now, fix some $\varepsilon>0$, and some a,b for which the assumed statement is true. Then, note that for sufficiently large N, we have that by the above computations, for all $N_s\le n\le N$,

$$\log\log N+(a-\varepsilon)\sqrt{\log\log N}<\log\log n+a\sqrt{\log\log n}<\log\log N+(a+\varepsilon)\sqrt{\log\log N}$$

and

$$\log\log N+(b-\varepsilon)\sqrt{\log\log N}<\log\log n+b\sqrt{\log\log n}<\log\log N+(b+\varepsilon)\sqrt{\log\log N}.$$

Therefore, for N sufficiently large, since $\frac{N_s}{N}=o(1)$, if we let

$$d_N=\frac{1}{N}\cdot\#\left\{n\le N\mid a\le\frac{\omega(n)-\log\log n}{\sqrt{\log\log n}}\le b\right\},$$

then

$$\begin{aligned}d_N &= \frac{1}{N}\cdot\#\left\{N_s\le n\le N\mid a\le\frac{\omega(n)-\log\log n}{\sqrt{\log\log n}}\le b\right\}+o(1)\\ &= \frac{1}{N}\cdot\#\left\{N_s\le n\le N\mid \log\log n+a\sqrt{\log\log n}\le\omega(n)\le\log\log n+b\sqrt{\log\log n}\right\}+o(1)\\ &\le \frac{1}{N}\cdot\#\left\{N_s\le n\le N\mid \log\log N+(a-\varepsilon)\sqrt{\log\log N}\le\omega(n)\le\log\log N+(b+\varepsilon)\sqrt{\log\log N}\right\}+o(1)\\ &= \frac{1}{N}\cdot\#\left\{n\le N\mid \log\log N+(a-\varepsilon)\sqrt{\log\log N}\le\omega(n)\le\log\log N+(b+\varepsilon)\sqrt{\log\log N}\right\}+o(1).\end{aligned}$$

Similarly, we can get that

$$d_N\ge\frac{1}{N}\cdot\#\left\{n\le N\mid \log\log N+(a+\varepsilon)\sqrt{\log\log N}\le\omega(n)\le\log\log N+(b-\varepsilon)\sqrt{\log\log N}\right\}+o(1).$$

Letting the RHS be d_{N^-}, we may note that

$$\lim_{N\to\infty}d_{N^-}=\frac{1}{\sqrt{2\pi}}\int_{a-\varepsilon}^{b+\varepsilon}e^{-x^2/2}\,\mathrm{d}x,$$

but ε was arbitrary, so we can make it tend to 0, and $\lim_{\varepsilon\to0}\int_{a-\varepsilon}^{b+\varepsilon}e^{-x^2/2}\,\mathrm{d}x=\frac{1}{\sqrt{2\pi}}\int_a^b e^{-x^2/2}$. Therefore, if we fix some $\varepsilon_0>0$, we can find some $\varepsilon>0$ and some large N for which

$$d_{N^-}\ge\frac{1}{\sqrt{2\pi}}\int_a^b e^{-x^2/2}\,\mathrm{d}x-\varepsilon_0.$$

Letting $\varepsilon_0 \to 0$. Thus, for N sufficiently large,

$$d_N \geq \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}\, \mathrm{d}x - \varepsilon_0.$$

We can do a very similar thing with the upper bound to get that for N sufficiently large,

$$d_N \leq \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}\, \mathrm{d}x + \varepsilon_0.$$

Since the choice of $\varepsilon_0 > 0$ was arbitrary, it thus follows that

$$\lim_{N\to\infty} d_N = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}\, \mathrm{d}x,$$

as desired. So our claim is proven. Now, let $A(N) = \sum_{p\leq N} \frac{1}{p}$. I claim that it suffices to show that

$$\lim_{N\to\infty} \frac{1}{N} \cdot \# \left\{ n \leq N \mid a \leq \frac{\omega(n) - A(N)}{A(N)^{1/2}} \leq b \right\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}\, \mathrm{d}x.$$

By Mertens' theorem ([4]), $A(N)$ is $\log\log N + O(1)$, and so $A(N)^{1/2}$ is $\sqrt{\log\log N} + O(1)$ as well. Since $O(1)$ is $o(\sqrt{\log\log N})$, we can just use the same logic as above to show this. Therefore, we will show this statement instead.

5 Converting To Moments

Let P_N denote the discrete random variable with sample space $\{1, 2, \cdots, N\}$, with each number being chosen with probability $\frac{1}{N}$. Let $X_N = A(P_N)$ be another discrete random variable. Now, let $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. We have the following two lemmas:

Lemma 5.1. *The standard normal distribution is completely determined by all its moments. That is, if there exists some random variable X with*

$$\mathbb{E}[X^k] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^k e^{-x^2/2}\, \mathrm{d}x,$$

then the probability density function of X must in fact be $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

Lemma 5.2. *Suppose $X_1, X_2, \cdots, X_n, \cdots$ is a sequence of random variables. Let X be a random variable that is completely determined by all its moments. Then, if $\mathbb{E}[X_n^k] \to \mathbb{E}[X^k]$ for all $k \geq 1$, then*

$$\mathit{Prob}(a \leq X_n \leq b) \to \mathit{Prob}(a \leq X \leq b)$$

for all $a < b \in \mathbb{R} \cup \{-\infty, \infty\}$.

The first one is well known, for the second one, see [1]. Now, how does this apply here? Well, let X be a random variable with probability density function φ. By Lemma 5.1, X is completely determined by all its moments. Therefore, if we show that

$$\lim_{N\to\infty} \mathbb{E}[X_N^k] = \mathbb{E}[X^k] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^k e^{-x^2/2}\,\mathrm{d}x,$$

then by Lemma 5.2, we have that for all $a < b \in \mathbb{R} \cup \{-\infty, \infty\}$,

$$\mathrm{Prob}(a \le X_N \le b) \to \mathrm{Prob}(a \le X \le b) = \int_a^b \varphi(x)\mathrm{d}x.$$

However, note that

$$\begin{aligned}
\lim_{N\to\infty} \mathrm{Prob}(a \le X_N \le b) &= \lim_{N\to\infty} \left[\frac{1}{N} \cdot \# \left\{ \frac{\omega(n) - A(N)}{\sqrt{A(N)}} \in [a,b] \mid n \le N \right\} + \frac{O(1)}{N} \right] \\
&= \lim_{N\to\infty} \left[\frac{1}{N} \cdot \# \left\{ \frac{\omega(n) - A(N)}{\sqrt{A(N)}} \in [a,b] \mid n \le N \right\} \right] \\
&= d\left(\left\{ a \le \frac{\omega(n) - A(N)}{\sqrt{A(N)}} \le b \mid n \in \mathbb{N} \right\} \right),
\end{aligned}$$

so it would imply

$$d\left(\left\{ a \le \frac{\omega(n) - A(N)}{\sqrt{A(N)}} \le b \mid n \in \mathbb{N} \right\} \right) = \int_a^b \varphi(x)\mathrm{d}x,$$

which is what we want. Note that the $O(1)$ was because of $n = 1, 2$. Thus, it suffices to show that $\mathbb{E}[X_N^k] \to \mathbb{E}[X^k]$.

6 Further Simplifications

So, we wish to show that

$$\mathbb{E}[X_N^k] \to \mathbb{E}[X^k].$$

From Lemma 8.2, we can get that $\mathbb{E}[X^k] = \mu_k = \begin{cases} (k-1)!! & k \text{ even}, k > 0 \\ 0 & k \text{ odd} \\ 1 & k = 0 \end{cases}$. Therefore, we wish to show that $\mathbb{E}[X_N^k] \to \mu_k$, i.e. $\mathbb{E}[X_N^k] = \mu_k + o(1)$. (**Here the** $o(1)$ **is with respect to** N, **not** k, **meaning think of** 1 **as a function of** N **and** k **and apply the definition of little** o **in that case**.) Note that

$$\begin{aligned}
\mathbb{E}[X_N^k] &= \frac{1}{N} \sum_{n \le N} \left(\frac{\omega(n) - A(N)}{A(N)^{1/2}} \right)^k \\
&= \frac{1}{NA(N)^{k/2}} \sum_{n \le N} (\omega(n) - A(N))^k.
\end{aligned}$$

Thus, it suffices to show that

$$\sum_{n \le N} (\omega(n) - A(N))^k = N(A(N))^{k/2}(\mu_k + o(1)).$$

7 Preliminary Lemmas

We start with a basic lemma from elementary number theory.

Lemma 7.1. *Let N, m be positive integers. Then,*

$$\#\{n \le N \mid \gcd(n,m)=1\} = \frac{\phi(m)}{m}N + O(\tau(m)).$$

Proof. Let $N = ma + b$, where a, b are nonnegative integers with $b < m$. Let $S = \{p_1, p_2, \cdots, p_r\}$ be the set of prime divisors of m. Then,

$$\begin{aligned}
\#\{n \le N \mid \gcd(n,m)=1\} &= \#\{n \le ma \mid \gcd(n,m)=1\} + \#\{ma < n \le ma+b \mid \gcd(n,m)=1\} \\
&= a \cdot \#\{n \le m \mid \gcd(n,m)=1\} + \#\{n \le b \mid \gcd(n,m)=1\} \\
&= a\phi(m) + \sum_{T \subseteq S} (-1)^{|T|} \left\lfloor \frac{b}{\prod_{t \in T} t} \right\rfloor \\
&= a\phi(m) + \sum_{T \subseteq S} \left[(-1)^{|T|} \frac{b}{\prod_{t \in T} t} + O(1) \right] \\
&= a\phi(m) + b\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_r}\right) + O(\tau(m)) \\
&= a\phi(m) + b\frac{\phi(m)}{m} + O(\tau(m)) \\
&= \frac{\phi(m)}{m}N + O(\tau(m)),
\end{aligned}$$

as desired. □

Lemma 7.2. $\mu_k = \begin{cases} (k-1)!! & k \text{ even}, k > 0 \\ 0 & k \text{ odd} \\ 1 & k = 0 \end{cases}.$

Proof. Let X be a continuous random variable with probability density function $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$. Then,

$$\mu_k := \mathbb{E}[X^k] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^k e^{-x^2/2} \mathrm{d}x.$$

For k odd, $x^k e^{-x^2/2}$ is an odd function, so the integral is zero. For k even, we will use induction to show that $\mu_k = \begin{cases} (k-1)!! & k > 0 \\ 1 & k = 0 \end{cases}$. Our base case is $k = 0$, which works because $\mathbb{E}[X^0] = 1$. Now, assume the result is true for some even $k \ge 0$. We will prove it true for $k+2$. Note that by integration by parts,

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^k e^{-x^2/2} \mathrm{d}x &= \left[\frac{1}{\sqrt{2\pi}} \frac{x^{k+1}}{k+1} e^{-x^2/2} \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{x^{k+1}}{k+1} (-xe^{-x^2/2}) \mathrm{d}x \\
&= 0 + \frac{1}{k+1} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{k+2} e^{-x^2/2} \mathrm{d}x \right),
\end{aligned}$$

so

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^{k+2}e^{-x^2/2}\mathrm{d}x = (k+1)\cdot\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^{k}e^{-x^2/2}\mathrm{d}x.$$

If $k = 0$, then this implies $\mu_2 = 1 = (2-1)!!$. Else, this implies $\mu_{k+2} = (k+1)\mu_k = (k+1)\cdot(k-1)!! = (k+1)!!$, as desired. □

Definition 7.3. For each prime p, define

$$L_p(n) := \begin{cases} -1/p+1 & p \mid n \\ -1/p & p \nmid n \end{cases}.$$

For any natural $m = p_1^{e_1}p_2^{e_2}\cdots p_k^{e_k}$, define

$$L_m(n) := \prod_{i=1}^{k}(L_{p_i}(n))^{e_i}.$$

The key here is that $\sum_{p\le N} L_p(n) = \omega(n) - A(N)$.

Lemma 7.4. *We have*

$$\sum_{N_s<p\le N} L_p(n) = O(\log(\sqrt{\log\log N}+3)),$$

for all $n \le N$.

Proof. Note

$$\sum_{N_s<p\le N} L_p(n) = \sum_{\substack{N_s<p\le N\\ p\mid n}} 1 - \sum_{N_s<p\le N}\frac{1}{p}$$

from the same logic as above. By Mertens ([4]),

$$\begin{aligned} O\left(\sum_{N_s<p\le N}\frac{1}{p}\right) &= O(\log\log N - \log\log N_s) \\ &= O\left(\log\frac{\log N}{\log N_s}\right) \\ &= O(\log\log(\sqrt{\log\log N}+3)). \end{aligned}$$

Now, suppose there were k primes in $(N_s, N]$ dividing n. Then, $N > N_S^k$, and since $n \le N$, $N > N_S^k$. So $k < \log_{N_S} N = \log(\sqrt{\log\log N}+3)$. Thus,

$$\sum_{\substack{N_S<p\le N\\ p\mid n}} 1 = O(\sqrt{\log\log N}+3),$$

so the result now follows.

□

Definition 7.5. Define the sequence of independent discrete random variables $\{L(p)\}_{p \text{ prime}}$ such that $L(p) = -1/p$ with probability $1 - \frac{1}{p}$ and $L(p) = -1/p + 1$ with probability $\frac{1}{p}$. Note that $L(p)$ "mimics" L_p, since $L_p(n) = -1/p + 1$ when $p \mid n$, which has "probability" $\frac{1}{p}$, and $L_p(n) = -1/p$ when $p \nmid n$, which has "probability" $1 - \frac{1}{p}$. So the following result should not be so surprising.

Lemma 7.6. *Suppose $m = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$ is fixed. Then,*

$$\sum_{n \le N} L_m(n) = N\mathbb{E}\left[\prod_{i=1}^{k} (L(p_i))^{e_i}\right] + O(3^k)$$

for all $N \ge 1$.

Proof. Note that

$$L_m(n) = \prod_{i=1}^{k} (L_{p_i}(n))^{e_i}.$$

Let $F_p(n) = \begin{cases} 1 - p & p \mid n \\ 1 & p \nmid n, \end{cases}$ then $L_{p_i}(n) = -\frac{1}{p_i} F_{p_i}(n)$. Then,

$$L_m(n) = \frac{(-1)^{\sum e_i}}{m} \prod_{i=1}^{k} (F_{p_i}(n))^{e_i}.$$

Therefore,

$$\sum_{n \le N} L_m(n) = \frac{(-1)^{\sum e_i}}{m} \sum_{n \le N} \prod_{i=1}^{k} (F_{p_i}(n))^{e_i}.$$

Let $P = \{p_1, p_2, \cdots, p_k\}$. Then, suppose $S \subseteq P$. The amount of $n \le N$ such that the set of prime divisors of n in P is exactly S is the number of integers less than or equal to $\left\lfloor \frac{N}{\prod_{s \in S} s} \right\rfloor$ relatively prime to all primes in $T := P \setminus S$. By Lemma 8.1, this is

$$\begin{aligned} \frac{\phi\left(\prod_{t \in T} t\right)}{\prod_{t \in T} t} \left\lfloor \frac{N}{\prod_{s \in S} s} \right\rfloor + O\left(\tau\left(\prod_{t \in T} t\right)\right) &= \frac{\prod_{t \in T}(t-1)}{\prod_{t \in T} t} \cdot \frac{N}{\prod_{s \in S} s} + O(2^{|T|}) \\ &= \frac{N \prod_{t \in T}(t-1)}{\prod_{p \in P} p} + O(2^{k-|S|}). \end{aligned}$$

Now, if $s = p_i$, let $e_s = e_i$. Then,

$$\begin{aligned}
\sum_{n\le N} L_m(n) &= \frac{(-1)^{\sum e_i}}{m} \sum_{n\le N} \prod_{i=1}^{k} (F_{p_i}(n))^{e_i} \\
&= \frac{(-1)^{\sum e_i}}{m} \sum_{S\subseteq P} \left[\frac{N \prod_{t\in T}(t-1)}{\operatorname{rad}(m)} + O(2^{k-|S|}) \right] \prod_{s\in S} (1-s)^{e_s} \\
&= \frac{N(-1)^{\sum e_i}}{m \operatorname{rad}(m)} \sum_{S\subseteq P} \left(\frac{\prod_{p\in P}(p-1)}{\prod_{s\in S}(s-1)} \prod_{s\in S}(1-s)^{e_s} \right) + \sum_{S\subseteq P} O(2^{k-|S|}) \frac{\prod_{s\in S}(1-s)^{e_s}}{m(-1)^{\sum e_i}} \\
&= \frac{N(-1)^{\sum e_i}\phi(m)}{m^2} \sum_{S\subseteq P} (-1)^{|S|} \prod_{s\in S} (s-1)^{e_s-1} + \sum_{S\subseteq P} O(2^{k-|S|}) \\
&= \frac{N(-1)^{\sum e_i}\phi(m)}{m^2} \prod_{i=1}^{k} (1-(p_i-1)^{e_i-1}) + \sum_{i=0}^{k} O\left(2^{k-i}\binom{k}{i}\right) \\
&= \frac{N(-1)^{\sum e_i}\phi(m)}{m^2} \prod_{i=1}^{k} (1-(p_i-1)^{e_i-1}) + O(3^k).
\end{aligned}$$

On the other hand, note that

$$\begin{aligned}
N\mathbb{E}\left[\prod_{i=1}^{k} (L(p_i))^{e_i} \right] &= N \prod_{i=1}^{k} \mathbb{E}[(L(p_i)^{e_i}] \\
&= N \prod_{i=1}^{k} \left[\left(1-\frac{1}{p_i}\right)\left(-\frac{1}{p_i}\right)^{e_i} + \frac{1}{p_i}\left(1-\frac{1}{p_i}\right)^{e_i} \right] \\
&= N \prod_{i=1}^{k} \frac{(p_i-1)(-1)^{e_i} + (p_i-1)^{e_i}}{p_i^{e_i+1}} \\
&= \frac{N}{m \operatorname{rad}(m)} \prod_{i=1}^{k} \left[((-1)^{e_i}(p_i-1)) \left(1-(p_i-1)^{e_i-1}\right) \right] \\
&= \frac{N(-1)^{\sum e_i}}{m} \cdot \frac{\prod_{i=1}^{k}(p_i-1)}{\operatorname{rad}(m)} \prod_{i=1}^{k} (1-(p_i-1)^{e_i-1}) \\
&= \frac{N(-1)^{\sum e_i}\phi(m)}{m} \prod_{i=1}^{k} (1-(p_i-1)^{e_i-1}),
\end{aligned}$$

so the result follows. □

Lemma 7.7. *Let $\pi(n)$ be the number of primes $\le n$. Then, for any positive integers k, we have*

$$\sum_{n\le N} \left(\sum_{p\le N_S} L_p(n) \right)^k = N\mathbb{E}\left[\left(\sum_{p\le N_S} L(p) \right)^k \right] + O(3^k \pi(N_S)^k)$$

Proof. Note that

$$\left(\sum_{p\le N_S} L_p(n)\right)^k = \sum_{p_1,p_2,\cdots,p_k\le N_S} L_{p_1p_2\cdots p_k}(n),$$

since $L_m(n)$ is completely multiplicative in m. Note that the p_i are not necessarily distinct or ordered or anything like that. Therefore,

$$\sum_{n\le N}\left(\sum_{p\le N_S} L_p(n)\right)^k = \sum_{n\le N}\sum_{p_1,p_2,\cdots,p_k\le N_S} L_{p_1p_2\cdots p_k}(n) = \sum_{p_1,p_2,\cdots,p_k\le N_S}\sum_{n\le N} L_{p_1p_2\cdots p_k}(n).$$

Now, by Lemma 8.6,

$$\sum_{n\le N} L_{p_1p_2\cdots p_k}(n) = N\mathbb{E}\left[\prod_{i=1}^k L(p_i)\right] + O(3^k).$$

So the LHS is equivalent to

$$\begin{aligned}\sum_{p_1,p_2,\cdots,p_k\le N_S}\left[N\mathbb{E}\left[\prod_{i=1}^k L(p_i)\right] + O(3^k)\right] &= N\mathbb{E}\left[\sum_{p_1,p_2,\cdots,p_k\le N_S}\prod_{i=1}^k L(p_i)\right] + O(3^k\pi(N_S)^k)\\ &= N\mathbb{E}\left[\left(\sum_{p\le N_S} L(p)\right)^k\right] + O(3^k\pi(N_S)^k),\end{aligned}$$

as desired. □

The next lemma will involve complex analysis (Sorry, it's too vast a topic for me to make a section about the fundamentals. Try Rudin or something.)

Lemma 7.8. *Let $f_n : \mathbb{C}\to\mathbb{C}$ for $n = 1, 2, \cdots$ be a sequence of functions which are analytic on an open disk D. Furthermore, suppose that $f_n \to f$ uniformly. Then, f is analytic on D as well and $f_n' \to f'$ uniformly on any open subdisk $E \subseteq D$ that has the same center as D and smaller radius than D.*

Proof. Let Γ be a closed contour in D. Note that since each f_n is analytic in D,

$$\oint_\Gamma f_n = 0$$

for all $n \ge 1$. Note that since all f_n are continuous, so is f due to a well known theorem: [3]. So f is integrable. Now fix an $\varepsilon > 0$. Then, there exists some N such that for all $n \ge N$, $|f - f_n| < \varepsilon$ on D. Now, note that

$$\begin{aligned}\left|\oint_\Gamma f\right| &= \left|\oint_\Gamma f - \oint_\Gamma f_n\right|\\ &= \left|\oint_\Gamma (f_n - f)\right|\\ &< \ell(\Gamma)\cdot\varepsilon,\end{aligned}$$

where ℓ denotes arc length. Thus, it follows that $\int_\Gamma f = 0$, for all closed contours Γ in D, so by Morera's theorem ([5]), f is analytic as well. Now, fix some $z_0 \in E$, and let r be the difference in radii of E and D. Then, $C_r := \{z \mid |z - z_0| < r\}$ lies completely within D. Then by Cauchy's Generalized Integral Formula,

$$\begin{aligned} f'(z_0) &= \frac{1}{2\pi i} \oint_{C_r} \frac{f(z)}{(z-z_0)^2}\,\mathrm{d}z \\ &= \frac{1}{2\pi i} \oint_{C_r} \frac{f_n(z)}{(z-z_0)^2}\,\mathrm{d}z + \frac{1}{2\pi i} \oint_{C_r} \frac{f(z) - f_n(z)}{(z-z_0)^2}\,\mathrm{d}z \\ &= f_n'(z_0) + \frac{1}{2\pi i} \oint_{C_r} \frac{f(z) - f_n(z)}{(z-z_0)^2}\,\mathrm{d}z. \end{aligned}$$

Thus,

$$f'(z_0) - f_n'(z_0) = \frac{1}{2\pi i} \oint_{C_r} \frac{f(z) - f_n(z)}{(z-z_0)^2}.$$

Now, fix some $\varepsilon > 0$, then for sufficiently large n, $|f(z) - f_n(z)| < \varepsilon$. Then, for sufficiently large n,

$$\begin{aligned} |f'(z_0) - f_n'(z_0)| &= \left| \frac{1}{2\pi i} \oint_{C_r} \frac{f(z) - f_n(z)}{(z-z_0)^2} \right| \\ &< \frac{1}{2\pi} \frac{\varepsilon}{r^2} \cdot 2\pi r \\ &= \frac{\varepsilon}{r}. \end{aligned}$$

Since r is fixed over all z_0, it follows that $f_n' \to f'$ uniformly on E, as desired. □

Corollary 7.9. *Let $f_n : \mathbb{C} \to \mathbb{C}$ for $n = 1, 2, \cdots$ be a sequence of functions which are analytic on an open disk D, and let $k \geq 1$. Furthermore, suppose that $f_n \to f$ uniformly. Then, f is analytic on D as well and $f_n^{(k)} \to f^{(k)}$ uniformly on any open subdisk $E \subseteq D$ that has the same center as D and smaller radius than D.*

8 Moment Generating Functions

Our aim in this section will be to compute the moments of $E_N = \sum_{p \leq N} L(p)$. We will do this by computing the moment generating function $M_{E_N}(z) = \mathbb{E}[e^{zE_n}]$, and then computing the k^{th} derivatives of this and then using Lemma 3.14. Note that

$$\begin{aligned} M_{E_N}(z) &= \mathbb{E}[e^{zE_n}] \\ &= \mathbb{E}\left[e^{z \sum_{p \leq N} L(p)}\right] \\ &= \prod_{p \leq N} \mathbb{E}\left[e^{zL(p)}\right]. \end{aligned}$$

Note that from Taylor Series, we have for $|z| \leq \frac{1}{2}$,

$$e^{zL(p)} = 1 + zL(p) + \frac{z^2}{2} L(p)^2 + O\left(|z|^3 |L(p)|^3\right),$$

since

$$
\begin{aligned}
\left|\sum_{k\geq 3}\frac{1}{k!}z^k L(p)^k\right| &\leq \sum_{k\geq 3}\frac{1}{k!}|z|^k|L(p)|^k \\
&\leq \sum_{k\geq 3}|z|^k|L(p)|^k \\
&= \frac{|z|^3|L(p)|^3}{1-|z|^3|L(p)|^3} \\
&\leq \frac{|z|^3|L(p)|^3}{1-\frac{1}{8}} \\
&= O(|z|^3|L(p)|^3).
\end{aligned}
$$

Note that

$$
\begin{aligned}
\mathbb{E}\left[L(p)\right] &= \left(1-\frac{1}{p}\right)\left(-\frac{1}{p}\right)+\frac{1}{p}\left(1-\frac{1}{p}\right) = 0 \\
\mathbb{E}\left[L(p)^2\right] &= \left(1-\frac{1}{p}\right)\left(-\frac{1}{p}\right)^2+\frac{1}{p}\left(1-\frac{1}{p}\right)^2 = \frac{p-1}{p^2}, \\
\text{and } \mathbb{E}\left[|L(p)|^3\right] &= \left(1-\frac{1}{p}\right)\left(\frac{1}{p}\right)^3+\frac{1}{p}\left(1-\frac{1}{p}\right)^3 = O\left(\frac{1}{p}\right).
\end{aligned}
$$

Therefore,

$$
\mathbb{E}[e^{zL(p)}] = 1+\frac{p-1}{2p^2}z^2+O\left(\frac{|z|^3}{p}\right).
$$

Note that

$$
1+\frac{p-1}{2p^2}z^2+O\left(\frac{|z|^3}{p}\right) = e^{\frac{p-1}{2p^2}z^2+O\left(\frac{|z|^3}{p}\right)},
$$

since by Taylor Series

$$
\begin{aligned}
e^{\frac{p-1}{2p^2}z^2+O\left(\frac{|z|^3}{p}\right)} &= 1+\frac{p-1}{2p^2}z^2+O\left(\frac{|z|^3}{p}\right)+\sum_{n\geq 2}\left[\frac{p-1}{2p^2}z^2+O\left(\frac{|z|^3}{p}\right)\right]^n \\
&= 1+\frac{p-1}{2p^2}z^2+O\left(\frac{|z|^3}{p}\right)+O\left(\sum_{n\geq 2}\frac{|z|^{2n}}{p^n}\right) \\
&= 1+\frac{p-1}{2p^2}z^2+O\left(\frac{|z|^3}{p}\right)+O\left(\frac{|z|^4/p^2}{1-|z|^2/p}\right) \\
&= 1+\frac{p-1}{2p^2}z^2+O\left(\frac{|z|^3}{p}\right)+O\left(\frac{|z|^4}{p^2}\right) \\
&= 1+\frac{p-1}{2p^2}z^2+O\left(\frac{|z|^3}{p}\right).
\end{aligned}
$$

We now write e^x as $\exp(x)$ for easier readability. Note that the sum of $\frac{1}{p^2}$ for p prime converges since $\sum_{n=1}^{\infty} \frac{1}{n^2}$ does.

$$\begin{aligned}
M_{E_N}(z) &= \prod_{p \le N} \mathbb{E}[e^{zL(p)}] \\
&= \prod_{p \le N} \exp\left(\frac{p-1}{2p^2} z^2 + O\left(\frac{|z|^3}{p}\right)\right) \\
&= \exp\left(z^2 \sum_{p \le N} \left(\frac{p-1}{2p^2}\right) + O\left(|z|^3 \sum_{p \le N} \frac{1}{p}\right)\right) \\
&= \exp\left(\frac{1}{2} z^2 \sum_{p \le N} \frac{1}{p} - \frac{1}{2} z^2 \sum_{p \le N} \frac{1}{p^2} + O\left(|z|^3 \sum_{p \le N} \frac{1}{p}\right)\right) \\
&= \exp\left(\frac{1}{2} z^2 \log\log N + O(|z|^2 + |z|^3 \log\log N)\right) \\
&= \exp\left(\frac{1}{2} z^2 \log\log N\right) \cdot \exp\left(O(|z|^2 + |z|^3 \log\log N)\right).
\end{aligned}$$

Let $M_{1E_N}(z) = M_{E_N}\left(\frac{z}{\sqrt{\log\log N}}\right)$. Then, for $|z| \le \frac{1}{2}\sqrt{\log\log N}$,

$$M_{1E_N}(z) = \exp\left(\frac{1}{2} z^2\right) \cdot \exp\left(O\left(\frac{|z|^2}{\log\log N} + \frac{|z|^3}{\sqrt{\log\log N}}\right)\right).$$

Now, suppose $|z| < 1$. Then, since $|e^z| \le e^{|z|}$, we have that

$$\begin{aligned}
M_{1E_N}(z) &= \exp\left(\frac{1}{2} z^2\right) \cdot \exp\left(O\left(\frac{1}{\log\log N} + \frac{1}{\sqrt{\log\log N}}\right)\right) \\
&= \exp\left(\frac{1}{2} z^2\right) \cdot \exp\left(O\left(\frac{1}{\sqrt{\log\log N}}\right)\right) \\
&= \exp\left(\frac{1}{2} z^2\right) + \exp\left(\frac{1}{2} z^2\right) \cdot \left(\exp\left(O\left(\frac{1}{\sqrt{\log\log N}}\right)\right) - 1\right) \\
&= \exp\left(\frac{1}{2} z^2\right) + O(1) \cdot \left(\exp\left(O\left(\frac{1}{\sqrt{\log\log N}}\right)\right) - 1\right).
\end{aligned}$$

Fix some $\varepsilon > 0$. Then, for N sufficiently large,

$$\left| O(1) \cdot \left(\exp\left(O\left(\frac{1}{\sqrt{\log\log N}}\right)\right) - 1\right) \right| < \varepsilon.$$

Thus, for all $|z| < 1$ and all N sufficiently large,

$$\left| M_{1E_N}(z) - e^{\frac{1}{2} z^2} \right| < \varepsilon.$$

Thus, $M_{1E_N}(z)$ converges uniformly to $e^{\frac{1}{2}z^2}$ on $|z| < 1$. Thus, by Corollary 8.9, it follows that $M_{1E_N}^{(k)}(z)$ converges uniformly to $\frac{\mathrm{d}^k}{\mathrm{d}z^k} e^{\frac{1}{2}z^2}$ on $|z| < \frac{1}{2}$. Therefore, $M_{1E_N}^{(k)}(0) = \left[\frac{\mathrm{d}^k}{\mathrm{d}z^k} e^{\frac{1}{2}z^2}\right]_{z=0}$. Note that from the Taylor Series of e^z, we have

$$e^{\frac{1}{2}z^2} = \sum_{n\geq 0} \frac{(z^2/2)^n}{n!} = \sum_{n\geq 0} \frac{z^{2n}}{2^n \cdot n!}.$$

Therefore, for k odd, the k^{th} derivative of $e^{\frac{1}{2}z^2}$ is zero, for $k = 0$, the k^{th} derivative of it is 1, and if k is even, then the k^{th} derivative of it is

$$\frac{k!}{(k/2)!2^{k/2}} = \frac{k!}{k!!} = (k-1)!!.$$

So in fact, the k^{th} derivative of $e^{\frac{1}{2}z^2}$ is μ_k. Thus, we have that

$$\begin{aligned}
\mathbb{E}[E_N^k] &= M_{E_N}^{(k)}(0) \\
&= \left(\sqrt{\log\log N}\right)^k M_{1E_N}^{(k)}(0) \\
&= (\sqrt{\log\log N})^k(\mu_k + o(1)).
\end{aligned}$$

9 The Finish

Note here: when we use small o notation here, it will always be with respect to N, not k. Plugging the above into Lemma 8.7, we have

$$\begin{aligned}
\sum_{n\leq N}\left(\sum_{p\leq N_S} L_p(n)\right)^k &= N\mathbb{E}\left[\left(\sum_{p\leq N_S} L(p)\right)^k\right] + O(3^k\pi(N_S)^k) \\
&= N\left(\sqrt{\log\log N_S}\right)^k(\mu_k + o(1)) + O(3^k\pi(N_S)^k).
\end{aligned}$$

Note that

$$3^k\pi(N_S)^k = o(N(\sqrt{\log\log N_S})^k).$$

So in fact,

$$\sum_{n\leq N}\left(\sum_{p\leq N_S} L_p(n)\right)^k = N(\sqrt{\log\log N_S})^k(\mu_k + o(1)).$$

Now, note that

$$\begin{aligned}
\sqrt{\log\log N} - \sqrt{\log\log N_S} &\leq \sqrt{\log\log N - \log\log N_S} \\
&= \sqrt{\log\frac{\log N}{\log N_S}} \\
&= \sqrt{\log\log(\sqrt{\log\log N} + 3)},
\end{aligned}$$

so

$$(\sqrt{\log\log N})^k - (\sqrt{\log\log N_S})^k \leq (\sqrt{\log\log N})^k - \left(\sqrt{\log\log N} - \sqrt{\log\log(\sqrt{\log\log N}+3)}\right)^k$$
$$= O\left(2^k(\sqrt{\log\log N})^{k-1}\sqrt{\log\log(\sqrt{\log\log N}+3)}\right),$$

since $\log\log(\sqrt{\log\log N}+3) = o(\sqrt{\log\log N})$. Therefore,

$$\sum_{n\leq N}\left(\sum_{p\leq N_S} L_p(n)\right)^k = N\left[(\sqrt{\log\log N})^k + O\left(2^k(\sqrt{\log\log N})^{k-1}\sqrt{\log\log(\sqrt{\log\log N}+3}\right)\right](\mu_k + o(1))$$
$$= N(\sqrt{\log\log N})^k(\mu_k + o(1)) + NO\left(2^k(\sqrt{\log\log N})^{k-1}\sqrt{\log\log(\sqrt{\log\log N}+3)}\right).$$

Note that

$$2^k(\sqrt{\log\log N})^{k-1}\sqrt{\log\log(\sqrt{\log\log N}+3)} = o((\sqrt{\log\log N})^k).$$

Therefore, we obtain that

$$\sum_{n\leq N}\left(\sum_{p\leq N_S} L_p(n)\right)^k = N(\sqrt{\log\log N})^k(\mu_k + o(1)).$$

Now, note that by Lemma 8.4 and the $k=1$ case from above,

$$\sum_{n\leq N}\left(\sum_{p\leq N} L_p(n)\right)^k = \sum_{n\leq N}\left(\sum_{p\leq N_S} L_p(n) + \sum_{N_S<p\leq N} L_p(n)\right)^k$$
$$= \sum_{n\leq N}\left(\sum_{p\leq N_S} L_p(n) + \log(\sqrt{\log\log N}+3)\right)^k$$
$$= \sum_{n\leq N}\left(\sum_{p\leq N_S} L_p(n)\right)^k + \sum_{n\leq N}\sum_{j=0}^{k-1}\binom{k}{j}\left(\sum_{p\leq N_S} L_p(n)\right)^j(\log(\sqrt{\log\log N}+3))^{k-j}$$
$$= \sum_{n\leq N}\left(\sum_{p\leq N_S} L_p(n)\right)^k + \sum_{j=0}^{k-1}\binom{k}{j}(\log(\sqrt{\log\log N}+3))^{k-j}\sum_{n\leq N}\left(\sum_{p\leq N_S} L_p(n)\right)^j$$
$$= \sum_{n\leq N}\left(\sum_{p\leq N_S} L_p(n)\right)^k + O(k2^k(\log(\sqrt{\log\log N}+3))^k)\max_{0\leq j\leq k-1}\sum_{n\leq N}\left|\sum_{p\leq N_S} L_p(n)\right|^j.$$

Note that by QM-AM,

$$\sqrt{\frac{\sum_{n\leq N}\left(\sum_{p\leq N_S} L_p(n)\right)^{2j}}{N}} \geq \frac{\sum_{n\leq N}\left|\sum_{p\leq N_S} L_p(n)\right|^j}{N}.$$

By our above work,

$$\sum_{n\le N}\left(\sum_{p\le N_S} L_p(n)\right)^{2j} = N(\sqrt{\log\log N})^{2j}(\mu_{2j}+o(1)).$$

Thus,

$$\sum_{n\le N}\left|\sum_{p\le N_S} L_p(n)\right|^{j} \le N(\sqrt{\log\log N})^{j}\sqrt{\mu_{2j}+o(1)} = N(\sqrt{\log\log N})^{k-1}\sqrt{\mu_{2k}}.$$

Thus,

$$\begin{aligned}\sum_{n\le N}\left(\sum_{p\le N} L_p(n)\right)^{k} &= \sum_{n\le N}\left(\sum_{p\le N_S} L_p(n)\right)^{k} + O\left(k2^k(\log(\sqrt{\log\log N}+3)^k N(\sqrt{\log\log N})^{k-1}\right)\\ &= N(\sqrt{\log\log N})^{k}(\mu_k+o(1)) + O\left(k2^k(\log(\sqrt{\log\log N}+3)^k N(\sqrt{\log\log N})^{k-1}\sqrt{\mu_{2k}}\right).\end{aligned}$$

Now, note that

$$k2^k(\log(\sqrt{\log\log N}+3)^k N(\sqrt{\log\log N})^{k-1}\sqrt{\mu_{2k}} = o(N(\sqrt{\log\log N})^k),$$

so

$$\sum_{n\le N}\left(\sum_{p\le N} L_p(n)\right)^{k} = N(\sqrt{\log\log N})^{k}(\mu_k+o(1)) = N(A(N))^{k/2}(\mu_k+o(1)),$$

as $A(N)^{1/2} = \sqrt{\log\log N} + O(1)$ (so the difference, which is $O(N(\sqrt{\log\log N})^{k-1}2^k)$, gets sucked into the $o(1)$), so we have proved the Erdos-Kac theorem. ■

References

[1] P. Billingsley, *Probability and Measure*. 1995, Proves the lemma about a convergent sequence of moments. The relevant lemma is Theorem 30.2.

[2] S. Fan, *The Erdős-Kac Theorem*, Jul. 2023, Expository paper on the Erdős-Kac Theorem.

[3] J. Howie, *Complex Analysis*. May 2003, Complex Analysis Textbook. The relevant citation is at Theorem 5.30.

[4] F. Mertens, *Journal fur die reine und angewandte Mathematik*. 1874, This is ancient, but contains the original proof of Mertens' theorem by Mertens himself.

[5] J. Shurman, *Morera's Theorem*. Feb. 2025, Proof of Morera's Theorem.