AN ELEMENTARY ANALYSIS OF THE ERDŐS-KAC CENTRAL LIMIT THEOREM

BRIAN HONG

Abstract. This expository paper examines the proof of the Erdős-Kac Theorem. We approximate the sum of the reciprocals of primes with the Prime Number Theorem. Next, we examine the Central Limit Theorems to indicate the convergence of the random samples to the normal distribution. We also prove the convergence of the moments of the Erdős-Kac probability distribution to the normal distribution. A probabilistic model of the occurrence of prime numbers is used widely in this proof.

1. Introduction

The Erdős-Kac Central Limit Theorem is a highly celebrated result in classical probabilistic number theory by mathematicians Paul Erdős and Mark Kac in 1940, who showed that a probability distribution involving $\omega(n)$, the number of distinct prime factors of a positive integer that is randomly chosen in the interval $[1, n]$, will converge to a normal distribution as n approaches infinity. The main takeaways from this paper is that not only the probability distribution, but also each individual characteristic of the distribution of the Erdős-Kac Theorem must converge. We prove the convergence of the probability distribution by referencing the Lindeberg-Feller Central Limit Theorem and the Classic Central Limit Theorem. Simply put, these theorems show that the data of a sample of a population of random, independent variables will approach the data of the entire population of those variables as the sample size gets bigger. We evaluate the moments of the Erdős-Kac Theorem by using a moment generating function, which is used very frequently in order to examine the specific features of the distribution. The probability model that we use for counting distinct prime factors is to count by 1 whenever the prime number is a factor of n. We assume in that model that the probability that we count by 1 is $1/p$ (since p is distinct and prime) and the probability we don't count is $1 - 1/p$.

We begin by introducing the theorem that will be proved in the paper, along with notations and definitions that are crucial in understanding the proof of this theorem. In the next section, we view an important theorem, which is Mertens' Second Theorem and prove that, allowing us to approximate sums with reciprocals of primes. Next, we look at the Central Limit Theorem and prove the Lindeberg-Feller CLT, which allows us to show the convergence of the probability distribution as a whole. After that, we examine the moments of the function in efforts to prove its convergence to the normal distribution. Finally, we prove the Erdős-Kac Central Limit Theorem in the final section.

Date: August 25, 2025.

2. Preliminaries

Theorem 2.1. Erdős-Kac Central Limit Theorem
Let x and a both be real numbers and n be a natural number. Then,

$$\lim_{x\to\infty}\frac{1}{x}\cdot\left\{n:3\le n\le x;\frac{\omega(n)-\log\log n}{\sqrt{\log\log n}}\le a\right\}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{a}e^{-t^2/2}dt.$$

In this proof, we will be using **multiplicative functions**, also known as an arithmetic functions. Some important properties of a multiplicative function are that

$$f(p^a)=f(p)^a$$

and

$$f(p_1^{\alpha_1}\cdots p_j^{\alpha_j})=(f(p_1^{\alpha_1})f(p_2^{\alpha_2})\cdots f(p_j^{\alpha_j}).$$

Examples of multiplicative functions that will be used in this paper include

- $\gcd(a,b)$(the Greatest Common Divisor Function),
- $\phi(n)$ (the Euler Totient Function),
- $\tau(n)$ (the Divisor Function),
- $\mu(n)$ (the Mobius Function).

The Divisor Function.
The symbol $\tau(n)$ is the Tau symbol, but more commonly known as the Divisor Function, where $\tau(n)$ counts the number of positive divisors of n. $\tau(n)$ can be mathematically expressed as

$$\tau(n)=\sum_{d|n}1,$$

where d and n are both positive integers. Another way of calculating $\tau(n)$ when $n=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_j^{\alpha_j}$ is

$$\tau(n)=(\alpha_1+1)(\alpha_2+1)\cdots(\alpha_j+1).$$

The Mőbius Function.
The Mőbius Function $\mu(n)$ can be split up into three separate cases:

$$\mu(n)=\begin{cases}1 & \text{if } n \text{ has no squared factors and an even number of prime factors or } n=1,\\ -1 & \text{if } n \text{ has no squared factors and an odd number of prime factors,}\\ 0 & \text{if } n \text{ has squared factors.}\end{cases}$$

The Euler Totient Function
The Euler Totient Function $\phi(n)$ counts the number of positive integers up to n that is relatively prime to n. $\phi(n)$ can be expressed as

$$\phi(n)=n\prod_{p|n}\left(1-\frac{1}{p}\right)$$

and as a floor function

$$\phi(n)=\sum_{a=1}^{n}\left\lfloor\frac{1}{\gcd(a,n)}\right\rfloor.$$

If integers a and n are not relatively prime, the value of $(a, n) > 1$ and so the quantity $\left\lfloor \frac{1}{(a,n)} \right\rfloor = 0$ in these cases. We will now investigate these functions more carefully and prove a relationship between them.

Lemma 2.2. *For any positive integer n, we have*

$$\phi(n) = \sum_{a=1}^{n} \Big(\sum_{k|(a,n)} \mu(k) \Big).$$

Proof. We begin by examining the summation $\sum_{k|n} \mu(k)$. Expanding this summation into individual terms results in

$$\mu(1) + \mu(p_1) + \cdots + \mu(p_j) + \mu(p_1 p_2) + \cdots + \mu(p_{j-1} p_j) + \cdots + \mu(p_1 p_2 \cdots p_j).$$

This is the case because it is not just prime factors that are divisors of an integer, but combinations of them too. Speaking of combinations, we can simplify this expansion by grouping the terms by the product of x number of primes. There are $\binom{j}{x}$ ways to distinctly to group x number of primes from the entire set. Also note that a singular prime number would result in a value of -1 from the Mőbius Function, and so we achieve

$$\sum_{k|n} \mu(k) = 1 + \binom{j}{1}(-1) + \binom{j}{2}(-1)^2 + \binom{j}{3}(-1)^3 + \cdots + \binom{j}{j}(-1)^j.$$

This can be condensed by referencing the binomial theorem, as the result above is an expansion of

$$(1-1)^j,$$

which is equal to 0. This means that for any $n > 1$, the value of $\sum_{k|n} \mu(k)$ will always be 0. Thus, we can generalize that

$$\sum_{k|n} \mu(k) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases}$$

Like we did for the Totient Function, we can represent this quantity as a floor function

$$\sum_{k|n} \mu(k) = \Big\lfloor \frac{1}{n} \Big\rfloor.$$

Recall the statement for the Totient Function

$$\phi(n) = \sum_{a=1}^{n} \Big\lfloor \frac{1}{(a,n)} \Big\rfloor.$$

We can substitute our result for the Mőbius Function in for $\left\lfloor \frac{1}{(a,n)} \right\rfloor$, and it follows that

$$\phi(n) = \sum_{a=1}^{n} \Big(\sum_{k|(a,n)} \mu(k) \Big).$$

■

Other essential notation.

- σ^2 = Variance, also the average of all the squared differences between the mean and each value.

- σ = Standard Deviation, or the measure of variance within a collection of data.
- $\mathbb{E}(X)$ = Expected value of random variable X.
- μ = Average value of the data set.
- $\pi(n)$ = The number of prime numbers up to or equal to positive integer n.

The normal distribution.
The normal distribution is the shape of the graph that the probability distribution of the Erdős-Kac Theorem will converge to. The normal distribution is often referred to as the bell curve, as the shape of a bell represents the graph of the normal distribution. Some important properties of the normal distribution function $f(z)$ include the following:

- $f(z)$ is symmetric about $z = \mu$,
- $f(z)$ increases and then decreases about its apex at $z = \mu$,
- $f(z)$ has inflection points at $z = \mu \pm \sigma$,
- $\lim_{z\to\pm\infty} f(z) = 0$.

The normal distribution density function is expressed by the equation

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp\bigg(-\frac{1}{2}\bigg(\frac{z-\mu}{\sigma}\bigg)^2\bigg).$$

Moments of the distribution.
In the field of probabilistic analysis, the moments of the probability distribution are frequently explored. They are the individual measures that help indicate the features of a probability distribution. They can help describe the asymptotic behavior of a distribution's graph. In mathematical terms, the nth central moment of continuous random variable Z is

$$\mathbb{E}[(Z-\mu)^n] = \int_{-\infty}^{\infty} (z-\mu)^n f(z)dz.$$

In addition, each n represents

- $n = 1$: mean
- $n = 2$: variance
- $n = 3$: skewness
- $n = 4$: kurtosis

of the distribution. For the sake of this paper and the proof of the Erdős-Kac Theorem, we will compute the moments of the probability distribution of the theorem to match the moments of the normal distribution. Another important piece in the proof of this theorem involves the estimation of the higher moments. Halberstam [3] showed that for positive integer k, the moments of the standard normal distribution m_k can be estimated as

$$\frac{1}{x}\sum_{n\le x}(\omega(n) - \log\log(x))^k = m_k(\log\log(x))^{k/2} + o((\log\log(x))^{k/2}). \tag{2.1}$$

Characteristic Functions.
Often times in the study of probabilistic theory, it is useful to examine the characteristic function of random variable X. The characteristic function is useful in the sense that it determines the asymptotic behavior of the probability distribution of X. The characteristic function is often known as the Fourier transform of the probability density function.

Let t be a real number with imaginary unit i and random variable X such that

$$\varphi_X(t) = \mathbb{E}(e^{itX}) = \mathbb{E}\cos tX + i\,\mathbb{E}\sin tX. \tag{2.2}$$

Another important equation concerning these characteristic functions is when it is applied to the normal distribution. The characteristic function of a normal distribution is

$$\exp(-t^2\sigma^2/2). \tag{2.3}$$

This will be explored in Lemma 4.2. more closely.

3. Mertens' Second Theorem

Theorem 3.1. Mertens' Second Theorem. *Let x be a real number greater than 2 and p be a positive integer less than x. Then,*

$$\sum_{p\leq x}\frac{1}{p}\sim \log\log(x))+O(1).$$

Proof. Start off by taking the Prime Number Theorem, which states that for any positive integer n

$$\pi(n)\sim\frac{n}{\log n}.$$

We can rewrite the Prime Number Theorem so that we get an equation for p_n, or the *nth* prime number, such that

$$p_n\sim n\log n.$$

Now, rewrite the original expression so that

$$\sum_{p\leq x}\frac{1}{p}\sim\sum_{p\leq x}\frac{1}{n\log n}$$

$$\sim\int_2^x\frac{1}{n\log n}.$$

The lower boundary for the integral is 2 because the value of p must be at least 2. Next, use u-substitution in which $u=\log n$ in order to solve the integral:

$$\int_2^x\frac{1}{n\log n}=\log\log x-\log\log 2.$$

The quantity $\log\log 2$ is negligible, so we will put it as an error term of $O(1)$.

■

4. The Central Limit Theorems

Theorem 4.1. The Central Limit Theorem. *In the theory of analytic probability, random, independent variables $X_1, X_2, \cdots, X_n$ are drawn from a population with mean μ and variance σ^2. A random sample of size n taken from a population allows for calculations of sample means:*

$$\mu_n=\frac{X_1+X_2+...+X_n}{n}$$

and sample standard deviation:

$$\sigma_n=\frac{\sigma}{\sqrt{n}}.$$

The Central Limit Theorem ultimately states that

$$\lim_{n\to\infty}\left(\frac{\mu_n-\mu}{\sigma_n}\right)$$

converges to a standard normal distribution.

This equation essentially means that a sample of independent, random variables will converge to the expected values of the entire population that the sample was taken out of. Empirically, the Central Limit Theorem will usually hold for samples with a size of at least 30. Graphically, the probability distribution will be shaped like a bell curve, also known as a normal distribution and a Gaussian curve, which we discussed previously.

Theorem 4.2. Lindeberg-Feller Central Limit Theorem. *Let independent random variable $X_{n,m}$ in which $1 \leq m \leq n$ have $\mathbb{E}(X_{n,m}) = 0$. If conditions*

- $\sum_{m=1}^{n} \mathbb{E}(X_{n,m}^2) \to \sigma^2 > 0$ *and*
- $\lim_{n\to\infty} \sum_{m=1}^{n} \mathbb{E}(|X_{n,m}|^2) = 0$

are met, then $X_{n,1} + \cdots X_{n,n} \to \sigma_\chi$ as $n \to \infty$, where χ has a standard normal distribution.

The main reason for our need to use this theorem lies in the fact that the distribution of $X(p) - 1/p$, although independent and random, are not necessarily identically distributed as the classical Central Limit Theorem requires in its conditions. For this reason, it is necessary to prove this theorem, which shows that independent and random variables with a finite variance will converge to a normal distribution. It is also helpful to visualize random variables in the Lindeberg-Feller CLT as a triangular array, such that

$$\begin{array}{llll} X_{1,1} & & & \\ X_{2,1} & X_{2,2} & & \\ X_{3,1} & X_{3,2} & X_{3,3} & \\ \cdots & \cdots & \cdots & \cdots \\ X_{n,1} & X_{n,2} & X_{n,3} & \cdots\ X_{n,n}. \end{array}$$

Precursory Observation. Before we go into the proof of this theorem, it is important to see how the original Central Limit Theorem is implied by the Lindeberg-Feller. Let independent and identically distributed variables $Y_1, Y_2, \cdots$ be the variables for the Central Limit Theorem. Assume that $\mathbb{E}(Y_n) = 0$ and $\mathbb{E}(Y_n^2) = \sigma^2$. Now, set up each sample by letting $X_{n,m} = Y_m/n^{1/2}$. In Theorem 4.2, we stated that

$$\lim_{n\to\infty} \sum_{m=1}^{n} \mathbb{E}(|X_{n,m}|^2) = 0.$$

Then,

$$\sum_{m=1}^{n} \mathbb{E}(X_{n,m}^2) = n\mathbb{E}((Y_1/n^{1/2})^2)$$

$$= \mathbb{E}(Y_1^2)$$

will converge to 0 as $n \to \infty$, showing that the Central Limit Theorem does indeed apply to the Lindeberg-Feller Theorem.

Lemma 4.3. *The characteristic function of a normal distribution with mean μ and variance σ^2 is $\varphi_X(t) = \exp(i\mu t - \sigma^2 t^2/2)$, where t is any real number.*

Proof. Start off by recalling the continuous and cumulative distribution function for a normal distribution:

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} \int e^{itx - x^2/2}$$
$$= \frac{1}{\sqrt{2\pi}} \int e^{itx} e^{-x^2/2}.$$

Substitute in Equation 2.2, and we can omit $i \sin tx$ because it is an odd function, meaning that its value when integrated is 0. This will give us

$$\frac{1}{\sqrt{2\pi}} \int (\cos tx) e^{-x^2/2}.$$

Now, differentiate this equation with respect to t such that

$$\varphi'(t) = \frac{1}{\sqrt{2\pi}} \int -(x \sin tx) e^{-x^2/2}.$$

Integration by parts of $\varphi'(t)$ lead us to

$$-\frac{1}{\sqrt{2\pi}} \int t(\cos tx) e^{-x^2/2} = -t\varphi(t).$$

Thus, we can assume that $\frac{d}{dt}\big[\varphi(t) exp(t^2/2)\big] = 0$ and consequently that means $\varphi(t) exp(t^2/2) = \varphi(o) = 1$.

■

Lemma 4.4. *Let t be a real number, with imaginary unit i and random variable X. The absolute value of the differences between the 2 expected values of characteristic functions can be maximized as shown:*

$$\left| \mathbb{E}(e^{itX}) - \sum_{m=0}^{n} \mathbb{E}\left(\frac{(itX)^m}{m!}\right) \right| \leq (|tX|^{n+1}, 2|tX|^n)$$

.

Proof.

$$\left| \mathbb{E}(e^{itX}) - \sum_{m=0}^{n} \mathbb{E}\left(\frac{(itX)^m}{m!}\right) \right| \leq \mathbb{E}\left| (e^{itX}) - \sum_{m=0}^{n} \left(\frac{(itX)^m}{m!}\right) \right|$$

This is the consequence of Jensen's Inequality, where it states that

$$\mathbb{E}(\varphi(X)) \leq \varphi(\mathbb{E}(X))$$

. This inequality is applicable to this scenario because we are dealing with expected values of characteristic functions. A deeper analysis of this inequality and its proof can be found in Theorem 1.1 of [2].

Simplify as shown:

$$\mathbb{E}\left| (e^{itX}) - \sum_{m=0}^{n} \left(\frac{(itX)^m}{m!}\right) \right| \leq \mathbb{E}\left| (e^{itX}) - \sum_{m=0}^{n} \left((itX)^m\right) \right|$$

Finally, notice that the second term is the dominant term. Since the sum runs from 0 to n, we can assume that it will be smaller than $|tX|^{n+1}$, thus concluding this proof. ■

Lemma 4.5. *Let real number c_j converge to 0, as well as n_j which will converge to ∞ as integer j increases. Also, $c_j n_j$ will converge to real number γ. If $max_{1\leq j\leq n}|c_{j,n}| \to 0$ and $\sum_{j=1}^{n} c_{j,n} \to \gamma$, then*

$$\prod_{j=1}^{n}(1+c_{j,n}) \to e^{\gamma}.$$

Proof. Since c_j will converge to 0, $\log(1+c_j)$ will also converge to 0 as j increases. Since $c_j n_j$ will converge to γ,

$$n_j \log(1+c_j) \to \gamma.$$

exponentiate both sides to obtain

$$(1+c_j)^{n_j} \to e^{\gamma}$$

$$\prod_{j=1}^{n}(1+c_{j,n}) \to e^{\gamma}.$$

■

Now, we have all the tools to prove the Lindeberg-Feller CLT!

Proof of the Lindeberg-Feller CLT. Let characteristic function $\varphi_{n,m}(t) = \mathbb{E}\, exp(itX_{n,m})$. It would suffice to prove that

$$\prod_{m=1}^{n} \varphi_{n,m}(t) \to exp(-t^2\sigma^2/2).$$

This is because the characteristic function of a normal distribution is $\exp(-t^2\sigma^2/2)$. As long as the product of all the characteristic functions converge to that quantity, we would be able to prove the Lindeberg-Feller CLT.

Let variable $z_{m,n} = \varphi_{n,m}(t)$ and $y_{m,n} = (1 - t^2\sigma^2/2)$. Now, apply Lemma 4.4. so that

$$|z_{m,n} - y_{m,n}| \leq \mathbb{E}(|tX_{n,m}|^3, 2|tX_{n,m}|^2).$$

$$\leq t^3\, \mathbb{E}(|tX_{n,m}|^3 + 2t^2\, \mathbb{E}(|X_{n,m}|^2)$$

Using the 2 conditions of the Lindeberg-Feller in Theorem 4.2 leads to

$$\lim_{n\to\infty} \sum_{m=1}^{n} |z_{m,n} - y_{m,n}| \leq t^3\sigma^2$$

$$\left| \prod_{m=1}^{n} \varphi_{n,m}(t) - \prod_{m=1}^{n}(1 - t^2\sigma_{n,m}^2/2) \right| \to 0.$$

The process behind that last step can be further explored and proved in Lemma 3.4.3 of [1]. This converges to 0 because we defined in Theorem 4.2. that σ^2 would converge to 0.

To finish, apply Lemma 4.5. to the second term $\prod_{m=1}^{n}(1 - t^2\sigma_{n,m}^2/2)$, which means that

$$\prod_{m=1}^{n}(1 - t^2\sigma_{n,m}^2/2) \to exp(-t^2\sigma_{n,m}^2/2).$$

Plug this result back into

$$\left| \prod_{m=1}^{n} \varphi_{n,m}(t) - \prod_{m=1}^{n} (1 - t^2\sigma_{n,m}^2/2) \right| \to 0$$

such that

$$\left| \prod_{m=1}^{n} \varphi_{n,m}(t) - exp(-t^2\sigma_{n,m}^2/2) \right| \to 0,$$

and thus we have the convergence

$$\prod_{m=1}^{n} \varphi_{n,m}(t) \to exp(-t^2\sigma_{n,m}^2/2).$$

5. Observing the moments

The proof of this theorem involves the assumption that since the family of distributions of the Erdős-Kac Theorem converges to the normal distributions, the moments should as well. Unfortunately, this cannot be completely certain due to the case of periodical outliers. Consider a X_n to be uniformly distributed along $(0, 1/n)$. $f_n(x)$ is equivalent to n on interval $(0, 1/n)$ and 0 in every other interval. This would mean that $\int f_n(x) = 1$ for any n because we have a rectangle of length $1/n$ and width n. However, the limit of the function converges to 0. This is problematic because outliers in a sequence of distributions can create a discrepancy between the limit and the integral, which can impact the calculation of the expected value, typically derived from an integral. Ultimately this can impact the moments of the sequence of distributions, although the sequence itself may still converge. We will take a look at the moments of the Erdős-Kac Theorem and show their convergence to the normal distribution by using moment generating function ζ, which is used to directly compute the moments of random variables. It would suffice to show that

$$\mathbb{E}\left(\exp\left(\zeta \frac{\sum_{p\le q}(X(p) - \frac{1}{p})}{\sqrt{\log\log(q)}} \right) \right) \to \exp(\zeta^2/2)$$

as $q \to \infty$ and p and q are positive integers.

Lemma 5.1. *The moment generating function can be written as*

$$\mathbb{E}\left(\exp\left(\zeta \frac{\sum_{p\le q}(X(p) - \frac{1}{p})}{\sqrt{\log\log(q)}} \right) \right) = \prod_{p\le q} \mathbb{E}\left(1 + \zeta \frac{(X(p) - \frac{1}{p})}{\sqrt{\log\log(q)}} + \zeta^2 \frac{(X(p) - \frac{1}{p})^2}{2\log\log(q)} + O\left(\frac{\zeta^3}{\log\log(q)^{3/2}} \right) \right).$$

Proof. Since the summation is in the exponent of the function, we can bring it out and replace it with the product.

$$\mathbb{E}\left(\exp\left(\zeta \frac{\sum_{p\le q}(X(p) - \frac{1}{p})}{\sqrt{\log\log(q)}} \right) \right) = \prod_{p\le q} \mathbb{E}\left(\exp\left(\zeta \frac{(X(p) - \frac{1}{p})}{\sqrt{\log\log(q)}} \right) \right)$$

The function in this scenario can be expanded by using the Taylor series for exponential function around $c = 0$. We will use 3 terms and an additional error term in this proof.

$$\exp(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6}$$

Replacing x for

$$\prod_{p\leq q}\mathbb{E}\left(\exp\left(\zeta\frac{(X(p)-\frac{1}{p})}{\sqrt{\log\log(q)}}\right)\right),$$

we obtain

$$\prod_{p\leq q}\mathbb{E}\left(1+\zeta\frac{(X(p)-\frac{1}{p})}{\sqrt{\log\log(q)}}+\zeta^2\frac{(X(p)-\frac{1}{p})^2}{2\log\log(q)}+\frac{\zeta^3(X(p)-\frac{1}{p})^3}{6(\log\log(q))^{3/2}}\right).$$

The quantities $(X(p)-\frac{1}{p})^3$ and $1/6$ are negligible in the final term, which will be an error term. Thus we can eliminate them and move on. ■

Lemma 5.2.

$$\mathbb{E}\left(\left(X(p)-\frac{1}{p}\right)^2\right)=\log\log(p)+O(1)$$

Proof. We will use a probabilistic model for the occurrence of prime numbers by defining $X(p)$ as independent variables in such a manner that:

$$X(p)=\begin{cases}1 & \text{if } p|n\\ 0 & \text{otherwise}\end{cases}$$

We can expect that the probability that a random integer is divisible by a prime number is $1/p$, since the distinct factor p is only occurs once. We can expect that the probability of an integer to be divisible by p with probability $\frac{1}{p}$ and not divisible with probability $1-\frac{1}{p}$. The expected value of $X(p)$ can be calculated such that

$$\mathbb{E}(X(p))=1\Big(\frac{1}{p}\Big)+0\Big(1-\frac{1}{p}\Big)=\frac{1}{p}.$$

Now, we calculate:

$$\mathbb{E}\left(\left(X(p)-\frac{1}{p}\right)^2\right)=\mathbb{E}(X(p)^2)-\frac{2}{p}\mathbb{E}(X(p))+\frac{1}{p^2}.$$

Notice that squaring $X(p)$ makes no difference in its expected value since $1^2=1$ and $0^2=0$, while the probabilities say the same:

$$\mathbb{E}(X(p)^2)-\frac{2}{p}\mathbb{E}(X(p))+\frac{1}{p^2}=\frac{1}{p}-\frac{2}{p^2}+\frac{1}{p^2}=\frac{1}{p}\Big(1-\frac{1}{p}\Big).$$

Simply distribute so that

$$\sum_{p\leq q}\left(\frac{1}{p}\right)\left(1-\frac{1}{p}\right)=\sum_{p\leq q}\frac{1}{p}-\sum_{p\leq q}\frac{1}{p^2}.$$

Since we established the value of the first term as approximately $\log\log(q)$, substitute that for the first term. $1/p^2$ will converge to a negligible value so we can replace it with an error term.

$$\sum_{p\leq q}\frac{1}{p}-\sum_{p\leq q}\frac{1}{p^2}=\log\log(q)+O(1)$$

■

Lemma 5.3. *The expanded moment generating function converges to a normal distribution as $q \to \infty$. Let both q and p be positive integers.*

$$\prod_{p\leq q}\mathbb{E}\left(1+\zeta\frac{(X(p)-\frac{1}{p})}{\sqrt{\log\log(q)}}+\zeta^2\frac{(X(p)-\frac{1}{p})^2}{2\log\log(q)}+O\left(\frac{\zeta^3}{(\log\log(q))^{3/2}}\right)\right)=\exp(\zeta^2/2)\left(1+O\left(\frac{\zeta^3}{(\log\log(q))^{3/2}}\right)\right)$$

Proof. Distribute the expected value to the terms. Notice that we can eliminate the second term in the parenthesis:

$$\mathbb{E}(X(p)-\frac{1}{p})=\mathbb{E}X(p)-\frac{1}{p}=0$$

Leaving us with

$$\prod_{p\leq q}\left(1+\zeta^2\frac{(X(p)-\frac{1}{p})^2}{2\log\log(q)}+O\left(\frac{\zeta^3}{(\log\log(q))^{3/2}}\right)\right).$$

Now, revert back to the sigma function and use the results from Lemma 5.2.

$$\prod_{p\leq q}\left(1+\zeta^2\frac{(X(p)-\frac{1}{p})^2}{2\log\log(q)}+O\left(\frac{\zeta^3}{(\log\log(q))^{3/2}}\right)\right)$$

$$=\exp\left(\zeta^2\frac{\sum_{p\leq q}(X(p)-\frac{1}{p})^2}{2\log\log(q)}+O\left(\frac{\zeta^3}{(\log\log(q))^{3/2}}\right)\right.$$

$$=\exp\left(\frac{\zeta^2}{2}+O\left(\frac{\zeta^3}{(\log\log(q))^{3/2}}\right)\right)$$

$$=exp(\zeta^2/2)\left(1+O\left(\frac{\zeta^3}{(\log\log(q))^{3/2}}\right)\right).$$

As q approaches infinity, the moments of the distribution will converge closer and closer to $\exp(\zeta^2/2)$, thus concluding this proof. ■

6. Erdős-Kac Central Limit Theorem

Now, we can finally begin the proof of this theorem. There are various ways to prove this theorem, but we will be looking at the computation of moments in this paper, a strategy first suggested by Halberstam in 1955 and followed up by Granville and Soundarajan.

We will reuse the previous probabilistic model for primes:

$$X(p)=\begin{cases}1 & \text{if } p|n\\ 0 & \text{otherwise.}\end{cases}$$

Notice that a summation of this model would equate to $\omega(n)$. To account for the quantity $-\log\log(x)$, we can apply Mertens' Second Theorem that we proved earlier. We will set function:

$$f_p(n)=\begin{cases}1-\frac{1}{p} & \text{if } p|n\\ -\frac{1}{p} & \text{otherwise}\end{cases}$$

Lemma 6.1. *Let $M = \prod_{r=1}^{s} p_r$ and $m = \prod_{r=1}^{s} p_r^{\alpha_r}$ where p_r represents the distinct prime factors and α_r is any positive integer. Then*

$$\sum_{n\leq x} f_m(n) = \prod_{j=1}^{s} \left(\left(\frac{1}{p}\right)\left(1-\frac{1}{p}\right)^{\alpha_j} + \left(\frac{-1}{p}\right)^{\alpha_j}\left(1-\frac{1}{p}\right)\right) + O\Big(3^s\Big).$$

Proof. Start off by letting $\gcd(n, M) = d$. Notice that we can split the summation into two parts:

$$\sum_{n\leq x} f_m(n) = \sum_{d|M} \sum_{\substack{n\leq x \\ (M,n)=d}} f_m(n).$$

Note that since d is a factor of n, we can also assume that $f_m(n) = f_m(d)$, leading us to

$$\sum_{n\leq x} f_m(n) = \sum_{d|M} \sum_{\substack{n\leq x \\ (M,n)=d}} f_m(d)$$

$$= \sum_{d|M} f_m(d) \sum_{\substack{n\leq x \\ (M,n)=d}} 1.$$

Since d is a factor of n, there must be some positive integer N so that $Nd = n$. The previous expression becomes

$$\sum_{d|M} f_m(d) \sum_{\substack{Nd\leq x \\ (M,Nd)=d}} 1$$

$$= \sum_{d|M} f_m(d) \sum_{\substack{N\leq x/d \\ (M/d,N)=1}} 1.$$

Since the second sum counts on the condition that M/d and N are distinctly prime, we can recall the Euler Totient Equation in Lemma 2.2, leading to

$$\sum_{d|M} f_m(d)\Bigg(\sum_{N\leq x/d} \sum_{k|(M/d,N)} \mu(k)\Bigg).$$

Recall that $M = \prod_r^s p_r$, which means that M cannot have factors that are squares. This expression can now be rewritten as

$$\sum_{d|M} f_m(d)\Bigg(\sum_{\substack{N\leq x/d \\ k|N}} \sum_{k|(M/d)} \mu(k)\Bigg).$$

The first sum in the parenthesis can be represented by $\frac{x}{dk}$ because integers up to x/d is counted every k integers as $k|N$. One scenario that we cannot forget here is in the case that x/d is not a factor of k, meaning that the quantity $\frac{x}{dk}$ may be a slightly inaccurate estimate of the actual sum. To take into account for this discrepency, we use an error term $O(1)$:

$$\sum_{d|M} f_m(d) \sum_{k|(M/d)} \mu(k)\left(\frac{x}{dk} + O(1)\right)$$

$$= \sum_{d|M} f_m(d)\Big(\sum_{k|(M/d)} \frac{x\mu(k)}{dk} + \sum_{k|(M/d)} \mu(k)O(1)\Big)$$

$$= \sum_{d|M} f_m(d)\Big(\frac{x}{d} \sum_{k|(M/d)} \frac{\mu(k)}{k} + \sum_{k|(M/d)} \mu(k)O(1)\Big)$$

Backtrack to Lemma 2.2. We notice that the second sum counts all k that divide a and n. Since $a \leq n$), we can use a positive integer r and set up $a = rk$. This would mean that r is bounded by n/k :

$$\phi(n) = \sum_{a=1}^{n} \Big(\sum_{k|(a,n)} \mu(k)\Big)$$

$$= \sum_{k|n} \Big(\sum_{r=1}^{n/k} \mu(k)\Big)$$

$$= \sum_{k|n} \mu(k)\frac{n}{k}.$$

Now, we can apply this to our previous statement in the first term in the parenthesis to obtain

$$\sum_{n\leq x} f_m(n) = \sum_{d|M} f_m(d)\Big(\frac{x}{M}\phi\Big(\frac{M}{d}\Big) + O\Big(\sum_{k|(M/d)} 1\Big)\Big)$$

$$= \sum_{d|M} f_m(d)\Big(\frac{x}{M}\phi\Big(\frac{M}{d}\Big) + O\Big(\tau\Big(\frac{M}{d}\Big)\Big)\Big).$$

Note that we can combine $\mu(k)O(1)$ into 1 because both terms are bounded by 1. Next, simply distribute:

$$\sum_{n\leq x} f_m(n) = \frac{x}{M} \sum_{d|M} f_m(d)\phi\Big(\frac{M}{d}\Big) + O\Big(\sum_{d|M} f_m(d)\tau\Big(\frac{M}{d}\Big)\Big).$$

The important idea to understand here is that the first term has multiplicative functions. We can use the function $f_m(d)$ to split up into cases of being prime and having divisors.

$$\sum_{n\leq x} f_m(n) = x \sum_{d|M} \Big(\prod_{p_j|d} \Big(1 - \frac{1}{p}\Big)^{\alpha_j} \prod_{p_j|R/d} \Big(\frac{-1}{p}\Big)^{\alpha_j} \phi\Big(\frac{M}{d}\Big)\Big) + O\Big(\sum_{d|M} f_m(d)\tau\Big(\frac{M}{d}\Big)\Big)$$

$$= x \sum_{d|M} \Big(\prod_{p_j|d} \Big(\frac{1}{p}\Big)\Big(1 - \frac{1}{p}\Big)^{\alpha_j} \prod_{p_j|R/d} \Big(\frac{-1}{p}\Big)^{\alpha_j} \Big(1 - \frac{1}{p}\Big)\Big) + O\Big(\sum_{d|M} f_m(d)\tau\Big(\frac{M}{d}\Big)\Big).$$

Notice that M has distinct factors, d would also have some distinct factors as well, which means we can assume that $d = \prod_{j\subset(1,2,\cdots,s)}$:

$$\prod_{j=1}^{s} \Big(\Big(\frac{1}{p}\Big)\Big(1 - \frac{1}{p}\Big)^{\alpha_j} + \Big(\frac{-1}{p}\Big)^{\alpha_j} \Big(1 - \frac{1}{p}\Big)\Big) + O\Big(\sum_{d|M} f_m(d)\tau\Big(\frac{M}{d}\Big)\Big).$$

Let's take a closer look at the error term. We can eliminate $f_m(d)$ because $|f_m(d)|$ is bounded by 1:

$$O\bigg(\sum_{d|M}\tau\bigg(\frac{M}{d}\bigg)\bigg).$$

Recall that $M=\prod_r^s p_r$, and the fact that p_r has a power of 1. This means that $\tau(M)=(1+1)_1(1+1)_2(1+1)_3\cdots(1+1)_s=2^s$. Something to notice about the sum is that it only counts when $d|M$. We can consider all the combinations of the factors of M that can be in d by using a binomial expansion as follows:

$$\sum_{d|M}\tau\bigg(M\bigg)=\binom{s}{0}(2^s)+\binom{s}{1}(2)^s+\binom{s}{2}(2)^s+\cdots+\binom{s}{s}(2)^s.$$

Now, we can take into account the quantity (M/d) in the tau function, which brings us to

$$\begin{aligned}\sum_{d|M}\tau\bigg(\frac{M}{d}\bigg)&=\binom{s}{0}(2^s)+\binom{s}{1}(2)^{s-1}+\binom{s}{2}(2)^{s-2}+\cdots+\binom{s}{s}(2)^{s-s}\\ &=\binom{s}{0}(2^s)(1^0)+\binom{s}{1}(2)^{s-1}(1^1)+\binom{s}{2}(2)^{s-2}(1^2)+\cdots+\binom{s}{s}(2)^{s-s}(1^s)\\ &=\sum_{q=0}^{s}\binom{s}{q}2^q1^{s-q}\\ &=\big(2+1\big)^s\\ &=3^s,\end{aligned}$$

finishing the proof. ■

Lemma 6.2. *Before we move on, note that we are able to simplify the statement of Lemma 6.1. so that*

$$\prod_{j=1}^{s}\mathbb{E}\bigg(\bigg(X(p_j)-\frac{1}{p}\bigg)^{\alpha_j}\bigg)=\prod_{j=1}^{s}\bigg(\bigg(\frac{1}{p_j}\bigg)\bigg(1-\frac{1}{p}\bigg)^{\alpha_j}+\bigg(\frac{-1}{p}\bigg)^{\alpha_j}\bigg(1-\frac{1}{p}\bigg)\bigg).$$

Proof. Recall that

$$X(p)=\begin{cases}1 & \text{if } p|n\\ 0 & \text{otherwise.}\end{cases}$$

This means that we can consider two cases, which is when $X(p_j)=1$ and $X(p_j)=0$. This would mean that

$$\bigg(X(p_j)-\frac{1}{p_j}\bigg)^{\alpha_j}$$

can equal

$$\bigg(1-\frac{1}{p}\bigg)^{\alpha_j}\text{ with probability }(1/p_j)$$

and

$$\bigg(\frac{-1}{p}\bigg)^{\alpha_j}\text{ with probability }(1-1/p_j).$$

We can take these individual values with their respective probabilities in order to find the expected value, thus concluding the proof of this lemma. ■

Lemma 6.3. *Given positive integer k, we have*

$$\sum_{n\leq x}\Big(\sum_{p\leq z} f_p(n)\Big)^k = \mathbb{E}\Big(\sum_{p\leq z}\Big(X(p)-\frac{1}{p}\Big)\Big)^k + O\Big(3^s\frac{z^k}{(\log z)^k}\Big).$$

Proof. Start off by manipulating the sums so that

$$\sum_{n\leq x}\Big(\sum_{p\leq z} f_p(n)\Big)^k = \sum_{p_1\leq z}\sum_{p_2\leq z}\cdots\sum_{p_k\leq z}\Big(\sum_{n\leq z} f_{p_1\cdots p_k}(n)\Big).$$

Apply Lemma 6.2 and 6.1 to obtain

$$\sum_{p_1\leq z}\sum_{p_2\leq z}\cdots\sum_{p_k\leq z}\Big(\prod_{j=1}^{s}\Big(\Big(\frac{1}{p}\Big)\Big(1-\frac{1}{p}\Big)^{\alpha_j}+\Big(\frac{-1}{p}\Big)^{\alpha_j}\Big(1-\frac{1}{p}\Big)\Big)+O\Big(3^s\Big)\Big)$$

$$=\sum_{p_1\leq z}\sum_{p_2\leq z}\cdots\sum_{p_k\leq z}\Big(\mathbb{E}\Big(\prod_{j=1}^{s}\Big(X(p_j)-\frac{1}{p_j}\Big)^{\alpha_j}\Big)+O\Big(3^s\Big)\Big)$$

Next, revert the summations back to the singular sum $\Big(\sum_{p\leq z} f_p(n)\Big)^k$ in order to attain

$$\sum_{n\leq x}\Big(\sum_{p\leq z} f_p(n)\Big)^k = \mathbb{E}\Big(\sum_{p\leq z}\Big(X(p)-\frac{1}{p}\Big)\Big)^k + O\Big(3^s\pi(z)^k\Big).$$

The error term gains $\pi(z)^k$ because it is included in the parenthesis. The number of primes up to z is represented by $\pi(z)$ and that sum happens k times. Finally, we can apply the Prime Number Theorem to complete the proof of this lemma. ■

Lemma 6.4. *Given integer $z = x^{1/\log\log\log(x)}$, and assuming $n \leq x$ we have*

$$\omega(n)-\log\log(x) = \sum_{p\leq z} f_p(n) + O(\log\log\log(x)).$$

Proof. Rewrite as

$$\omega(n)-\log\log(x) = \sum_{p|n} 1-\log\log(x)$$

$$=\sum_{\substack{p|n\\ p\leq z}} 1+\sum_{\substack{p|n\\ p>(z)}} 1-\log\log(x)$$

$$=\sum_{\substack{p|n\\ p\leq z}} 1-\log\log(x)+\sum_{\substack{p|n\\ p>(z)}} 1.$$

By Mertens' Second Theorem, $\log\log(x)$ is approximately 1/p, meaning we can substitute in $f_p(n)$, which is equivalent to $1-1/p$, as we do below:

$$\sum_{p\leq z} f_p(n)+\sum_{\substack{p|n\\ p>(z)}} 1.$$

Since we know that p is a factor of n and $p > (z)$ in the second sum, the number of factors p that divide n must be $\mathrm{O}(\log\log\log(x))$ because that would limit n to not being greater than x.

■

Lemma 6.5.

$$\frac{1}{x}\sum_{n\leq x}(\omega(n)-\log\log(x))^k$$
$$= m_k(\log\log(x))^{k/2} + O\left(3^s\frac{z^k}{(\log z)^k}\right) + O_k(\log\log\log(x))^k(\log\log(x))^{\frac{k-1}{2}})$$

Proof. Start off by taking quantity

$$\sum_{n\leq x}(\omega(n)-\log\log(x))^k$$

and replacing the inner difference such that

$$\sum_{n\leq x}\left(\sum_{p\leq z}f_p(n)+R(x)\right)^k$$

$R(x)$ is used to indicate a remainder term in which its value increases as x increases. This remainder term is on the order of a particular value that we will explore later. For the next step, apply the Binomial Theorem as follows:

$$\sum_{n\leq x}\left(\sum_{p\leq z}f_p(n)+R(x)\right)^k = \sum_{n\leq x}\left(\sum_{l=0}^{k}\binom{k}{l}\left(\sum_{p\leq z}f_p(n)\right)^l R(x)^{k-l}\right).$$

We will break up the sum into the case when $l = k$ and the other case in which l goes up to $k-1$ as shown:

$$\sum_{n\leq x}\left(\sum_{p\leq z}f_p(n)\right)^k + \sum_{n\leq x}\left(\sum_{l=0}^{k-1}\binom{k}{l}\left(\sum_{p\leq z}f_p(n)\right)^l R(x)^{k-l}\right).$$

We will now work towards maximizing the second term. The first term can simply be replaced by using Lemma 6.3, so we focus on the second term. Because the remainder term grows as x grows, we will make that as large as possible:

$$\sum_{n\leq x}(\omega(n)-\log\log(x))^k \ll \sum_{n\leq x}\left(\sum_{p\leq z}f_p(n)\right)^k + \sum_{n\leq x}\left(\sum_{l=0}^{k-1}\binom{k}{l}\left(\sum_{p\leq z}f_p(n)\right)^l R(x)^{k}\right).$$

Simply rearrange the values and apply the Binomial Theorem so that

$$\sum_{n\leq x}\left(\sum_{p\leq z}f_p(n)\right)^k + \sum_{l=0}^{k-1}\binom{k}{l}\sum_{n\leq x}\left(\left(\sum_{p\leq z}f_p(n)\right)^l R(x)^k\right)$$
$$= \sum_{n\leq x}\left(\sum_{p\leq z}f_p(n)\right)^k + \sum_{l=0}^{k-1}\binom{k}{l}1^{l-k}1^l\sum_{n\leq x}\left(\left(\sum_{p\leq z}f_p(n)\right)^l R(x)^k\right)$$
$$= \sum_{n\leq x}\left(\sum_{p\leq z}f_p(n)\right)^k + (1+1)^k\sum_{n\leq x}\left(\left(\sum_{p\leq z}f_p(n)\right)^l R(x)^k\right)$$

$$= \sum_{n\leq x}\left(\sum_{p\leq z} f_p(n)\right)^k + 2^k R(x)^k \sum_{n\leq x}\left(\sum_{p\leq z} f_p(n)\right)^l.$$

Now, we focus on maximizing quantity $\sum_{n\leq x}\left(\sum_{p\leq z} f_p(n)\right)^l$. To do this, we will apply the Cauchy-Schwarz Inequality so that

$$\frac{1}{x}\sum_{n\leq x}\left(\sum_{p\leq z} f_p(n)\right)^l \ll \frac{1}{x}\left(\sum_{n\leq x}\left(\sum_{p\leq z} f_p(n)\right)^{2l}\right)^{1/2}$$

Substitute in Equation 2.1 and now we get

$$(m_{2l}(\log\log(x)^{2l/2})^{1/2}$$
$$= (\log\log(x))^{\frac{k-1}{2}},$$

leaving us with

$$\sum_{n\leq x}\left(\sum_{p\leq z} f_p(n)\right)^k + 2^k R(x)^k (\log\log(x))^{\frac{k-1}{2}}.$$

We will make the last term into an error term. Since 2^k is a fixed number, we will just exclude from the error term and rewrite as

$$\sum_{n\leq x}\left(\sum_{p\leq z} f_p(n)\right)^k + O_k\left(R(x)^k (\log\log(x))^{\frac{k-1}{2}}\right).$$

To finish, apply Lemma 6.3 and then equation 2.1. Take Lemma 6.4 and the first step of the proof of this Lemma in order to see the value of $R(x) = O(\log\log\log x)$. ■

Proof of the Erdős-Kac CLT. Take the result from Lemma 6.5, in which

$$\frac{1}{x}\sum_{n\leq x}(\omega(n) - \log\log(x))^k$$
$$= m_k(\log\log(x))^{k/2} + O\left(3^s \frac{z^k}{(\log z)^k}\right) + O_k(\log\log\log(x))^k(\log\log(x))^{\frac{k-1}{2}}).$$

It would suffice to prove that this result converges to

$$\frac{1}{x}\sum_{n\leq x}(\omega(n) - \log\log(x))^k = m_k(\log\log(x))^{k/2} + o((\log\log(x))^{k/2}).$$

Notice that $O_k((\log\log(x))^{\frac{k-1}{2}})$ is on the order of $(\log\log(x))^{k/2}$. $(\log\log\log(x))^k$ is smaller than $(\log\log(x))^{k/2}$, thus making both factors within the order of $(\log\log(x))^{k/2}$. This means that as $x \to \infty$, the moments of

$$\frac{\omega(n) - \log\log n}{\sqrt{\log\log n}}$$

approach the asymptotic behavior of the random variable in a normal distribution, thus concluding the proof.

Acknowledgements.
I would like to thank my mentors Simon Rubinstein-Salzedo and Andrew Lin for assisting me in this independent research.

Bibliography.

(1) Rick Durrett. *Probability: theory and examples.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010.
(2) Mikhail Lavrov. *Chapter 2, Lecture 4: Jensen's Inequality.* University of Illinois at Urbana-Champaign, February 2019. https://faculty.math.illinois.edu/ mlavrov/docs/484-spring-2019/ch2lec4.pdf
(3) H. Halberstam. *On the distribution of additive number-theoretic functions.* J. London Math. Soc. 30 (1955), 43–53. MR 0066406 (16,569g)
(4) Steve Lester. *Normal Order Of $\omega(n)$ And The Erdős-Kac Theorem.* http://www.math.tau.ac.il/ rudnick/courses/sieves2015/ErdosKac.pdf
(5) Andrew Granville and K. Soundararajan. *Sieving and the Erdős-Kac theorem, Equidistribution in number theory, an introduction.* NATO Sci. Ser. II Math. Phys. Chem., vol. 237, Springer, Dordrecht, 2007, pp. 15–27. MR 2290492 (2008b:11103)
(6) Scheithauer, Thomas John. *A Thorough and Accessible Proof of the Erdős-Kac Theorem Following Granville and Soundararajan.* (2015), Online Theses and Dissertations. 422. https://encompass.eku.edu/etd/422
(7) Tom M. Apostol. *Introduction to analytic number theory.* Springer-Verlag, New York-Heidelberg, 1976, Undergraduate Texts in Mathematics. MR 0434929 (55 # 7892)
(8) P. Erdős and M. Kac. *The Gaussian law of errors in the theory of additive number theoretic functions.* Amer. J. Math. 62 (1940), 738–742. MR 0002374 (2,42c)

Email address: brianhong73@gmail.com