Relations Between The Erdos Kac Theorem and Ergodic Theory

Syed Ismail

July 13, 2026

Introduction

- My idea with this project is to discuss the relations between the Erdos–Kac theorem and Ergodic Theory.
- The scope of this talk will be to discuss basic preliminaries, followed by defining ergodic theory and some of its key notable theorems.
- In addition to this, we will discuss the heuristic derivation of the Erdos–Kac theorem with some examples.
- In the end, I discuss some fields where both topics relate.
- However, due to the time limit of the talk, I will be unable to talk about every relation I wrote in my expository paper.

Key Theorems of Probability

Theorem (σ-algebra)

suppose X is a set and S is a set of subsets of X. Then S is called a σ-algebra on X if the following conditions are satisfied :

- $\emptyset \in S$
- *if $E \in S$, then $X \setminus E \in S$*
- *if $E_1, E_2, \ldots$ is a sequence of elements of S, then $\bigcup_{k=1}^{\infty} E_k \in S$*

Theorem (Probability Space)

We define the probability measure by the ordered pair $(\Omega, \mathcal{F}, P)$ where Ω represents the collection of all possible outcomes, $\mathcal{F}$ is the set of events and P is a function defined $P : \mathcal{F} \to [0, 1]$. $\mathcal{F}$ is a σ-algebra of Ω.

Continued

Definition (Random Variable)

A random variable X is defined as a measurable function that maps from a probability space to the set of all real numbers. So, if for every Borel set B $\in \mathbb{R}$ we have

$$X^{-1}(B) = \{w \mid X(w) \in B\} \in \mathcal{F}$$

Definition (Expected Value)

For a random variable X on the probability space (Ω,$\mathcal{F}$, P). We define its expected value as :

$$E(X) = \int_{\Omega} X \, dP$$

it is the weighted average of all possible outcomes

Continued

Definition (Variance)

The variance of a random variable measures how far its values tend to spread out from their expected value. Var(X) $=E(X^2) - [E(X)]^2$.

Theorem (Central Limit Theorem)

Let $Y_1, Y_2, ..., Y_n$ be independent and identically distributed variables such that$E(Y_i) = \mu$ and $Var(Y_i) = \sigma^2 < \infty$. Then we define

$$U_n = \frac{\overline{Y} - \mu}{\sigma/\sqrt{n}}$$

Then the distributive function U_n converges to the standard normal distribution as $n \to \infty$.

The Erdös-Kac Theorem

The Erdös-kac theorem is one of the most popular results in the field of probabilistic number theory. The theorem was proven by Paul Erdös and Mark kac in 1940 by combining the ideas of the central limit theorem with Brun's sieve. (However for the heuristic derivation, we shall not use Brun's sieve approximation).

Theorem (Erdös–Kac Theorem)

Let $\omega(n)$ denote the number of distinct prime divisors of n. Then, as $n \to \infty$, the normalized arithmetic function

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

converges in distribution to the standard normal distribution.

Heuristic derivation of the Erdös-Kac Theorem

let the arithmetic function $w(n)$ denote the number of prime divisors that n contains.

Definition (Bernoulli Random Variable)

$$B_p = \begin{cases} p & \text{if } p|n \\ 1-p & \text{if } p \nmid n \end{cases}$$

The function $w(n)$ can be viewed as a sum of Bernoulli random variables, where each variable indicates whether a prime divides n. Summing these indicators counts the total number of distinct prime factors of n. Therefore ,

$$w(n) = \sum_{p \leq n} B_p(n)$$

Continued

The expectation of a binomial variable is always $\frac{1}{p}$. Therefore the expectation for $\mathbb{E}(w(n))$ equals to

$$\mathbb{E}(w(n)) = \sum_{p \leq n} \mathbb{E}[B_p(n)] = \sum_{p \leq n} \frac{1}{p}$$

Using Mertens estimate which states $\sum_{p \leq x} \frac{1}{p} = \ln(\ln(x)) + A + o\left(\frac{1}{\ln(x)}\right)$
We approximate $\mathbb{E}(w(n)) \sim \mathrm{loglog} n$.

We shift our focus to the variance of the Bernoulli random variable

Continued

the variance for a Bernoulli random variable is always equal to $\frac{1}{p}\left(1-\frac{1}{p}\right)$. Therefore the variance of $w(n)$ equals to

$$Var(w(n)) = \sum_{p\leq n} Var(B_p(n)) = \sum_{p\leq n} \frac{1}{p}\left(1-\frac{1}{p}\right) = \sum_{p\leq n} \frac{p-1}{p^2}$$

which converges. Hence by using Merten's estimate again we get ,

$$Var(w(n)) \sim \text{loglog} n$$

By applying the central Limit Theorem we get

$$\frac{w(n) - loglogn}{\sqrt{loglogn}} \to N(0,1)$$

Hence proving the theorem.

Ergodic Theory

A dynamical system in mathematics consists of functions defined $f : X \to X$ such that its domain and codomain are the same. Ergodic theory is a subfield, studying the statistical behavior of these dynamical systems.

Definition (A Measure Preserving Dynamical System)

A measure preserving system is the tuple $(X, \mathcal{A}, \mu, T)$ where $(X, \mathcal{A}, \mu)$ and $T : X \to X$ is a measure preserving transformation.

Ergodicity

Definition (Ergodicity)

A measure-preserving transformation T on a measure space $(X, \mathcal{A}, \mu)$ is said to be ergodic if whenever $A \in \mathcal{A}$ is a T-invariant set. Then either, $\mu(A) = 0$ or $\mu(A^c) = 0$. If T is ergodic, then $(X, \mathcal{A}, \mu, T)$ is an ergodic system.

Note: A set A is T-invariant if whenever $x \in A$, then $T(x) \in A$ aswell.

Birkhoff's Ergodic Theorem

Theorem (Birkhoff's Ergodic Theorem)

Let $(X, \mathcal{A}, \mu, T)$ *be an ergodic-measure preserving system . Then for any* $L^1(\mu) wehave$

$$lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1} f(T^n(x)) = \int_X fd\mu$$

as $n \to \infty$

In simpler terms, It means for an ergodic measure-preserving dynamical system, the long-term average of a function along a trajectory is equal to its average over the entire space.

Example of Birkhoff's Ergodic Theorem

- Assume an interval X = [0,1) and let $T(x) = 2x \pmod 1$ with usual measure μ.
- We shall take the observable $f(x) = x$. Then according to Birkhoff's Ergodic Theory :

$$\frac{1}{n}\sum_{i=0}^{n-1} f(T^n(x)) = \int_X f d\mu$$

$$\frac{x + T(x) + T^2(x) + ... + T^{N-1}(x)}{N} = \int_0^1 dx$$

$$\frac{x + T(x) + T^2(x) + ... + T^{N-1}(x)}{N} = 0.5$$

We come to an observation that the time average equates to the space average.

Additive Functions as Dynamical Observables

An arithmetic function is a function

$$f : \mathbb{N} \to \mathbb{R}$$

which assigns a value to each positive integer. It is called additive if

$$f(mn) = f(m) + f(n), \qquad \gcd(m, n) = 1.$$

A key example is the function studied in the Erdös–Kac theorem:

$$\omega(n) = \sum_{p \le n} 1_{\{p|n\}},$$

where $1_{\{p|n\}}$ indicates whether the prime p divides n.

In ergodic theory, an observable is a measurable function that assigns a quantity to each state of a dynamical system.

Therefore, $\omega(n)$ can be viewed as an observable that measures the arithmetic structure of an integer through its prime divisors.

Deterministic Systems with Statistical Behavior

A central idea connecting ergodic theory and number theory is that deterministic objects can exhibit statistical patterns.

In ergodic theory, Birkhoff's theorem studies the long-term average of an observable along a trajectory:

$$\frac{1}{N}\sum_{k=0}^{N-1} f(T^k(x)).$$

This shows how repeated observations of a deterministic system can produce predictable statistical behavior.

Similarly, the Erdös–Kac theorem studies the statistical distribution of

$$\omega(n) = \sum_{p \leq n} 1_{\{p|n\}},$$

as n ranges over the integers.

Thank You