

A New Perspective on Statistics: Combinations and Convolution as Fundamental Principles

Abstract

Statistics, at its core, is the study of uncertainty and data patterns. Traditional approaches often separate combinatorial probability from continuous distributions, treating them as distinct entities. We propose a unifying framework where **combinations (patterns)** and **convolution (joint probabilities)** form the foundation of statistical thinking. By recognizing these two operations as the core principles, we gain a deeper understanding of probability distributions, inferential statistics, and data science applications.

1. Introduction

Statistics is traditionally introduced through descriptive measures and probability theory. However, we argue that the entire discipline can be **reconstructed** using two fundamental mathematical operations:

1. **Combinations (patterns):** The enumeration and arrangement of discrete possibilities.
2. **Convolution (joint probabilities):** The accumulation of independent random variables into new distributions.

These two pillars naturally extend into discrete and continuous probability distributions, forming the basis of inferential statistics, machine learning, and Bayesian inference.

2. Combinations: The Mathematics of Patterns

Combinatorial methods determine the number of ways an event can occur, providing the structure for discrete probability models. The binomial, multinomial, and hypergeometric distributions all arise from counting arrangements of outcomes.

2.1 Example: The Binomial Distribution

The probability of exactly k successes in n independent Bernoulli trials follows:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

Here, the binomial coefficient $\binom{n}{k}$ counts the number of ways to arrange k successes among n trials.

2.2 Connection to the Poisson Process

Taking the limit of the binomial process as $n \rightarrow \infty$, $p \rightarrow 0$, while keeping $\lambda = np$ constant, we obtain the Poisson distribution:

$$P(N = k) = \frac{\overset{\text{stats_approach}}{(\lambda t)^k e^{-\lambda t}}}{k!}.$$

This demonstrates that discrete event counting stems directly from combinatorial structures. However, this explanation does not fully capture the deeper relationship between Poisson and continuous distributions, which we explore in the next section.

3. Convolution: The Mathematics of Joint Probabilities

Convolution governs how independent random variables combine. If X and Y are independent, their sum $Z = X + Y$ has a probability density function given by:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx.$$

This operation underlies many fundamental distributions and statistical principles.

3.1 Example: The Normal Distribution from Convolution

By the Central Limit Theorem, summing independent identically distributed (i.i.d.) random variables leads to a normal distribution, regardless of the original distribution:

$$X_1 + X_2 + \cdots + X_n \approx \mathcal{N}(\mu, \sigma^2).$$

This is a direct consequence of repeated convolution operations.

3.2 Connection to the Gamma and Beta Distributions

- **Gamma Distribution:** The sum of k independent exponential waiting times follows a gamma distribution, forming the basis of Poisson arrival models.
- **Beta Distribution:** Normalizing the sum of two gamma-distributed variables results in the beta distribution, which is central to Bayesian statistics.

3.3 Poisson as the Discrete Inverse of Gamma, Not Exponential

A common misconception is that the Poisson distribution is simply the integral of the exponential distribution. However, this overlooks the fact that Poisson deals with **multiple occurrences**, not just a single waiting time. Instead, Poisson is better understood as the **inverse of the gamma distribution**, because:

- The **gamma distribution** models the waiting time for the k -th event:

$$f_{T_k}(t) = \frac{\lambda^k t^{k-1} e^{-\lambda t}}{(k-1)!}.$$

- The **Poisson probability** of seeing exactly k events by time t is obtained by integrating this gamma density:

$$P(N = k) = \int_0^t f_{T_k}(s) e^{-\lambda(t-s)} ds.$$

Since the gamma distribution arises from the **repeated convolution of exponential distributions**, the factorial term in Poisson naturally emerges from **repeated integration operations** rather than from combinatorial counting.

4. Unified View of Statistics

Recognizing **combinations and convolution** as the primary tools of statistics allows us to:

- **Derive all classical probability distributions systematically.**
- **Explain Bayesian inference as a recursive convolution process.**
- **Model real-world data using a dual discrete-continuous framework.**

This perspective connects discrete probability (counting patterns) with continuous probability (integrating densities), offering a more intuitive and mathematically rigorous approach to learning statistics.

5. Conclusion

We propose that statistics should be taught and understood as the interplay of:

- **Combinations, which define probability spaces and patterns.**
- **Convolution, which determines how distributions evolve through interactions.**

This shift in perspective aligns statistics more closely with raw mathematics, making it not only more intuitive but also more powerful in applications such as Bayesian inference, deep learning, and stochastic modeling.