What are the similarities and differences between binomial and poisson distribution?

Authored by
stats writer

December 7, 2025

RECOMMENDED CITATION

stats writer (2025). *What are the similarities and differences between binomial and poisson distribution?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=106604

The fields of statistics and data science frequently rely on specific mathematical frameworks to model real-world phenomena. Among the most fundamental tools for analyzing countable outcomes are the Binomial distribution and the Poisson distribution. Both fall under the category of discrete probability distributions, meaning they quantify the likelihood of specific, countable outcomes rather than continuous measurements.

While sharing this fundamental characteristic of discreteness, their applications and underlying assumptions diverge significantly. The **Binomial distribution** is tailored for scenarios involving a fixed number of independent trials, each resulting in one of two outcomes (success or failure). Conversely, the **Poisson distribution** excels at modeling the frequency of rare events occurring within a specified unit of time or space, where the number of potential occurrences is theoretically infinite or unknown.

This comprehensive guide delves into the definitions, core similarities, crucial differences, and practical decision-making processes required when choosing between these two powerful statistical models. Understanding the nuances between fixed trials and events in an interval is essential for accurate statistical modeling and inference.

Two distributions that are similar in statistics are the **Binomial distribution** and the **Poisson distribution**. This tutorial provides a detailed explanation of each distribution along with the similarities and differences between the two.

The Binomial Distribution: Definition and Parameters

The Binomial distribution is the cornerstone model for sequences of independent experiments, often referred to as Bernoulli trials. It calculates the exact probability of observing a specific number of "successes," denoted as k, within a predetermined, fixed quantity of attempts, denoted as n. Crucially, the outcome of each attempt must be binary--either success or failure--and the probability of success, p, must remain constant across all trials. This fixed structure makes the Binomial model indispensable for quality control, survey analysis, and genetics where the population size is fixed or the number of measurements is predetermined.

For a distribution to be classified as Binomial, four strict conditions must be met. First, the experiment must consist of a fixed number of trials, n. Second, each trial must be independent, meaning the outcome of one trial does not influence the outcome of any subsequent trial. Third, every trial must result in one of two possible outcomes (success or failure). Fourth, the probability of success, p, must be identical for every trial. When these conditions hold, the corresponding random variable X is said to follow a binomial distribution, commonly written as $X \sim B(n, p)$.

The parameters necessary to define a Binomial distribution are simple yet powerful: **n**, the total number of trials, and **p**, the probability of success on any single trial. These parameters govern the shape and center of the distribution. For instance, if *p* is close to 0.5, the distribution will be symmetric. If *n* is large, the distribution begins to approximate a normal distribution, a fundamental concept known as the De Moivre-Laplace theorem.

The Binomial Distribution Formula Explained

To determine the probability that a Binomial random variable *X* equals exactly *k* successes, we utilize a precise formula incorporating combinatorial mathematics. This formula not only accounts for the likelihood of obtaining *k* successes and *n-k* failures but also calculates the number of distinct ways those successes and failures can be ordered within the *n* trials.

If a random variable *X* follows a binomial distribution, then the probability that *X* = *k* successes can be found by the following formula:

P(X=k) = nCk * pk * (1-p)n-k

The components of this formula are defined as follows:

n: The total, fixed number of trials.
k: The specific number of successes we are interested in calculating the probability for.
p: The fixed probability of success on a given trial.
nCk: The binomial coefficient, representing the number of ways to obtain *k* successes in *n* trials.

For example, suppose we flip a fair coin 3 times. We can use the formula above to determine the probability of obtaining 0 heads (successes) during these 3 flips, demonstrating that *n* = 3 and *k* = 0:

P(X=0) = 3C0 * .50 * (1-.5)3-0 = 1 * 1 * (.5)3 = **0.125**

The Poisson Distribution: Definition and Context

In contrast to the structured nature of the Binomial model, the Poisson distribution is designed to model the count of events occurring randomly and independently over a continuous interval of time, space, or other defined unit. It is used when we know the average rate of occurrence (λ, or lambda), but we do not know the maximum possible number of events that could occur during that interval. This makes it ideal for modeling phenomena such as website traffic spikes per minute, defects per square meter of fabric, or the number of emergency calls received per hour.

The Poisson distribution is sometimes referred to as the "Law of Rare Events" because it accurately models scenarios where the probability of any single event occurring is very low, but the total number of opportunities for the event to occur is very large (theoretically infinite). The primary parameter governing the Poisson distribution is λ **(lambda)**, which represents the mean rate of occurrence for the specified interval. For a random variable X following this pattern, it is denoted as $X \sim P(\lambda)$.

The key assumptions underlying the Poisson process include the requirement that the events must occur independently, and the rate (λ) must be constant over the interval of interest. Furthermore, events cannot occur simultaneously; they are instantaneous occurrences. If the underlying rate changes or if events are clustered, the Poisson model may not provide an accurate fit, requiring more complex statistical techniques.

The Poisson Distribution Formula Explained

The calculation of probability using the Poisson distribution relies on the mean rate of occurrence (λ) and the mathematical constant *e* (Euler's number). The formula determines the likelihood of observing exactly *k* events when the long-run average rate is known.

If a random variable *X* follows a Poisson distribution, then the probability that *X* = *k* events can be found by the following formula:

$P(X=k) = \lambda k * e- \lambda / k!$

The variables within this powerful formula are crucial for determining event frequencies:

λ: The mean number of successes (events) that occur during a specific interval.
k: The desired number of successes (events) for which the probability is being calculated.
e: Euler's number, a constant equal to approximately 2.71828.

For example, suppose a particular hospital experiences an average of λ = 2 births per hour. We can use the formula above to determine the probability of experiencing 3 births in a given hour (k=3), illustrating the model's predictive capability in interval-based counting:

$P(X=3) = 23 * e- 2 / 3!$

Core Similarities Between Binomial and Poisson Models

Despite their distinct applications, the Binomial and Poisson models share several fundamental properties rooted in the definition of discrete probability distributions. Recognizing these commonalities helps solidify the understanding of when these models may overlap or when one can approximate the other.

The Binomial and Poisson distribution share the following **similarities**:

Both distributions can be used to model the number of occurrences of some event, ensuring the

output variable is always an integer (a count).
In both distributions, events are assumed to be **independent**. In the Binomial model, each trial must be independent; in the Poisson model, the occurrence of one event during an interval must not affect the likelihood of another event occurring during that same interval.
Both models are defined by a small set of parameters which completely characterize the distribution and allow for calculation of the mean, variance, and specific probabilities.

Furthermore, an important conceptual similarity is that the Poisson distribution can often serve as an excellent approximation of the Binomial distribution when the number of trials n is very large and the probability of success p is very small, such that the expected number of events, $\lambda = np$, is small and constant. This relationship is highly valuable in computation, as calculating probabilities for very large binomial coefficients can be computationally complex, making the simpler Poisson calculation a practical alternative.

Key Differences: Fixed Trials vs. Intervals

The primary distinguishing factor between the two distributions lies in the nature of the observation space and the parameters involved. This difference dictates whether we are analyzing a predetermined set of discrete tests or monitoring events over a continuous medium.

The distributions share the following key **differences**:

Number of Trials (N): In a **Binomial distribution**, there is a fixed and known number of trials (e.g., flip a coin 3 times, test 50 components). This parameter n is always finite and defined prior to the experiment.
Observation Space: In a **Poisson distribution**, there is no fixed upper limit on the number of events that could occur during a certain time interval. The count is defined over an interval where the number of possible events is theoretically unlimited (e.g., how many customers will arrive at a store in a given hour?).
Parameters Required: The Binomial distribution requires two parameters (n and p). The Poisson distribution requires only one parameter (λ, the average rate).
Success Definition: The Binomial model requires knowledge of both successes and failures within a closed set of trials. The Poisson model only tracks the occurrence of the event itself; there is no implicit concept of a "failure" or non-event count.

Choosing the correct model hinges entirely on identifying whether the experiment has a predetermined number of opportunities for success (Binomial) or if you are simply measuring the frequency of occurrences where the total number of opportunities is effectively unlimited (Poisson). Misidentifying this core difference leads to fundamentally incorrect statistical inferences.

Practice Problems: Applying the Concepts

To solidify the distinction between these two models, we examine practical scenarios and determine which distribution best fits the underlying process. The decision rests solely on whether the process involves a fixed number of trials or an unbounded observation interval.

Problem 1: Network Failures

A tech company wants to model the probability that a certain number of network failures occur in a given week. Suppose it's known that an average of 4 network failures occur each week. Let *X* be the number of network failures in a given week. What type of distribution does the random variable *X* follow?

Analysis: In this scenario, we are counting events (network failures) within a fixed time interval (one week). While the average rate ($\lambda=4$) is known, there is no fixed upper boundary for the number of failures that could occur. Since we are modeling the frequency of occurrences in an interval based on an average rate, we use the Poisson model.

Answer: *X* follows a **Poisson distribution** because we're interested in modeling the number of network failures in a given week, defined by an average rate, and there is no upper limit on the number of failures that could occur. This is not a Binomial distribution because there is not a fixed number of trials.

Problem 2: Shooting Free-Throws

Tyler makes 70% of all free-throws he attempts. Suppose he shoots 10 free-throws. Let *X* be the number of times Tyler makes a basket during the 10 attempts. What type of distribution does the random variable *X* follow?

Analysis: This scenario defines the exact number of attempts, $n = 10$. Each attempt is an independent trial with a binary outcome (success/failure), and the probability of success, $p = 0.70$, is constant. Since all criteria for a Bernoulli process are met, and the number of trials is fixed, the Binomial distribution is the appropriate choice.

Answer: *X* follows a **Binomial distribution** because there is a fixed number of trials (10 attempts), the probability of "success" on each trial is the same, and each trial is independent.

Conclusion: Selecting the Appropriate Model

Selecting between the Binomial and Poisson distributions requires careful evaluation of the experimental design. If the total number of opportunities for the event to occur is **fixed and known-**

-such as the number of coin flips, survey respondents, or manufactured items checked--the **Binomial distribution** (B(*n*, *p*)) is the correct choice. This model emphasizes structure and containment.

Conversely, if the total number of opportunities is **unknown or potentially infinite**, and the data concerns the frequency of events within a defined unit (time, area, volume) based on an average rate--such as accidents per month or calls per minute--the **Poisson distribution** (P(λ)) is the appropriate tool. This model emphasizes rate and randomness.

Mastery of these two fundamental discrete probability distributions is essential for anyone engaged in statistical inference, modeling, and predictive analytics, ensuring that conclusions drawn from data accurately reflect the underlying stochastic process.