What is the Binomial Distribution?

Authored by

stats writer

December 27, 2025

RECOMMENDED CITATION

stats writer (2025). *What is the Binomial Distribution?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=109192

The Binomial Distribution is a fundamental concept within statistics and probability distribution theory. It provides a mathematical framework for calculating the likelihood of a specific number of successful outcomes in a fixed sequence of independent tests. Specifically, it models the count of successes resulting from a fixed number (n) of Bernoulli trials, provided that the probability of success (p) remains constant across every trial.

Understanding the Binomial Distribution is critical for fields ranging from quality control and finance to biological research. Whenever a situation involves repeated, identical, and independent trials--each yielding only one of two possible results (success or failure)--the Binomial Distribution is the appropriate tool for prediction and analysis. Before diving into the distribution itself, it is essential to first clearly define the characteristics of the foundational component: the binomial experiment.

```
@import url('https://fonts.googleapis.com/css?family=Droid+Serif|Raleway');

#words {
color: black;
font-family: Raleway;
max-width: 550px;
margin: 25px auto;
line-height: 1.75;
padding-left: 100px;
}

#words label, input {
display: inline-block;
vertical-align: baseline;
width: 350px;
}

#button {
border: 1px solid;
border-radius: 10px;
margin-top: 20px;
padding: 10px 10px;
cursor: pointer;
outline: none;
background-color: white;
color: black;
font-family: 'Work Sans', sans-serif;
border: 1px solid grey;
```

```
/* Green */
}

#button:hover {
background-color: #f6f6f6;
border: 1px solid black;
}

p, li {
color:#000000;
font-size: 19px;
font-family: 'Helvetica';
}

p a {
color: #9b59b6 !important;
}
```

While the binomial distribution describes the overall resulting probabilities, the concept is entirely built upon the conditions of a binomial experiment. Recognizing whether a real-world scenario qualifies as such an experiment is the first crucial step in applying this statistical model correctly.

The Characteristics of a Binomial Experiment

A statistical experiment must satisfy four strict criteria to be classified as a binomial process. These conditions ensure that the resulting outcomes are discrete and predictable using the binomial probability mass function. Failure to meet any one of these criteria means that the binomial distribution cannot be accurately applied.

A **binomial experiment**, sometimes referred to as a set of repeated Bernoulli trials, is characterized by the following required properties:

The experiment consists of a fixed number, denoted as *n*, of repeated trials. This number must be determined in advance and cannot change during the process.
Each individual trial must result in exactly two possible outcomes, conventionally labeled "success" or "failure." These outcomes are mutually exclusive.
The probability of success, symbolized by *p*, must remain identical for every single trial performed. Consequently, the probability of failure is always $1-p$.
All trials must be statistically independent. This means the result of any trial has absolutely no influence on the outcome of any subsequent trial.

A classic and easily understandable example of a binomial experiment is flipping a fair coin multiple times. If we choose to flip a coin 10 times, we can confirm that all four properties are met. There are 10 fixed trials ($n=10$). Each flip results in either heads (success) or tails (failure). Assuming a fair coin, the probability of success ($p=0.5$) is constant. Crucially, the outcome of the fifth flip is not affected by the outcome of the fourth flip, confirming the criterion of independence.

The Binomial Probability Mass Function (PMF)

The primary purpose of the binomial distribution is to quantify the exact likelihood of observing k successes given n total trials. If X is defined as the number of successes, X is considered a random variable that follows a binomial distribution, denoted $X sim B(n, p)$. The probability that X equals exactly k is calculated using the Binomial Probability Mass Function (PMF), which combines combinatorial mathematics with the laws of probability for independent events.

The mathematical expression for calculating the probability of exactly k successes in n trials is:

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

Each term in this formula serves a distinct purpose in determining the overall probability. Understanding these components is key to mastering the application of the binomial distribution:

n: Represents the total, fixed number of trials in the experiment.
k: Represents the exact number of desired successes in those n trials.
p: Is the constant probability of success on any single trial.
(1-p): Represents the probability of failure on any single trial, often denoted as q.
nCk: This term calculates the number of distinct arrangements, or combinations, in which k successes can occur within n trials. It is calculated as $n! / (k! (n-k)!)$.

To illustrate this formula, let us revisit the simple experiment of flipping a fair coin three times ($n=3$). We define success as landing on heads, so $p=0.5$. We can meticulously calculate the probability for every possible outcome ($k=0, 1, 2, 3$ heads). For instance, the probability of exactly two heads ($k=2$) involves calculating the number of ways to get two heads in three flips (HHT, HTH, THH, so 3C2=3) multiplied by the probability of that sequence occurring.

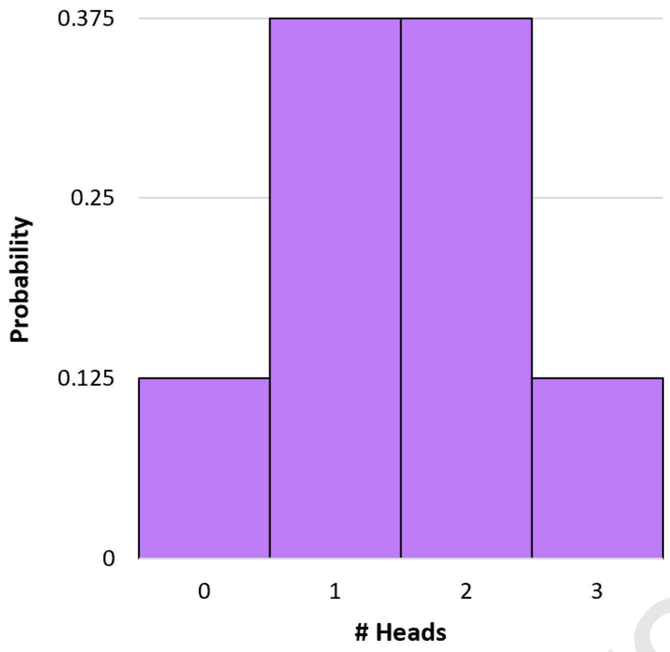
P(X=0) = 3C0 * 0.50 * (1-0.5)3-0 = 1 * 1 * (0.5)3 = **0.125**

P(X=1) = 3C1 * 0.51 * (1-0.5)3-1 = 3 * 0.5 * (0.5)2 = **0.375**

P(X=2) = 3C2 * 0.52 * (1-0.5)3-2 = 3 * 0.25 * (0.5)1 = **0.375**

P(X=3) = 3C3 * 0.53 * (1-0.5)3-3 = 1 * 0.125 * (0.5)0 = **0.125**

The sum of all these probabilities must equal 1 (0.125 + 0.375 + 0.375 + 0.125 = 1.0), which is a fundamental requirement for any probability distribution. This set of discrete probabilities can be visually represented using a histogram, showcasing the shape and symmetry of the distribution:



Understanding the Cumulative Distribution Function (CDF)

While the PMF calculates the probability of exactly k successes, often in real-world applications, we are interested in the probability of a range of outcomes. This is where the Cumulative Distribution Function (CDF) comes into play. The CDF, which calculates the cumulative probability, answers questions like: "What is the probability of observing k or fewer successes?"

Calculating the cumulative probability involves summing the individual probabilities found using the PMF for all values from zero up to the target value, k. For a discrete distribution like the binomial, the cumulative probability $P(X le k)$ is the sum of $P(X=i)$ for $i=0$ through k.

Using our previous coin flip example ($n=3$, $p=0.5$), we can calculate the cumulative probabilities for each outcome. If we want to know the probability of getting 1 head or less, we sum the probabilities for 0 heads and 1 head. This calculation demonstrates how the probability accumulates across the range of possible outcomes:

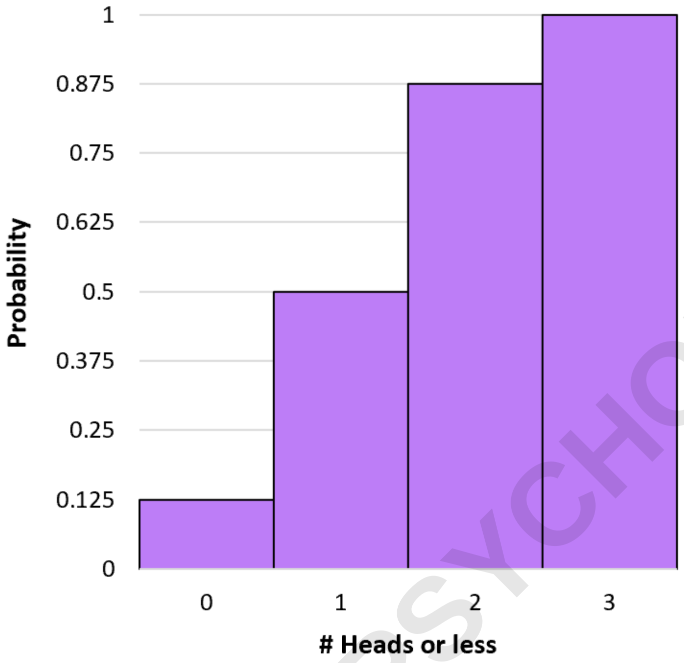
P(X≤0) = P(X=0) = **0.125**. (The probability of getting 0 or fewer heads.)

P(X≤1) = P(X=0) + P(X=1) = 0.125 + 0.375 = **0.5**. (The probability of getting 1 head or fewer.)

P(X≤2) = P(X=0) + P(X=1) + P(X=2) = 0.125 + 0.375 + 0.375 = **0.875**. (The probability of getting 2 heads or fewer.)

P(X≤3) = P(X=0) + P(X=1) + P(X=2) + P(X=3) = 0.125 + 0.375 + 0.375 + 0.125 = **1**. (The probability of getting 3 heads or fewer, which must equal 1.)

The resulting cumulative distribution, when visualized, always forms a non-decreasing, step-like function that starts at zero and ends at one, reflecting the certainty that the number of successes will fall within the total range of possible outcomes.



Using Computational Tools for Large Binomial Datasets

While calculating binomial probabilities by hand provides an excellent understanding of the underlying mathematical principles, this method quickly becomes impractical when dealing with large numbers of trials (n). For example, calculating the probability of observing exactly 43 successes in 100 trials requires complex calculations involving large factorials, which are prone to computational errors if done manually. For scenarios involving $n ge 20$, the use of statistical software, programming languages (like R or Python), or specialized online calculators becomes essential for efficiency and accuracy.

A **binomial probability calculator** automates the PMF and CDF calculations, allowing users to rapidly determine probabilities for exact outcomes, upper bounds, lower bounds, or ranges. These tools are indispensable when applying the binomial distribution to practical problems involving large sample sizes, such as industrial quality control or large-scale polling data.

Consider a scenario where an analyst needs to calculate the probability of a coin landing on heads 43 times or less after 100 flips. Here, $n=100$, $p=0.5$, and $k=43$. The calculator below provides all the relevant probabilities for $k=43$. Note that these calculations involve the use of cumulative sums and complements (1 minus the probability) to cover various inequality scenarios (k, $ge k$).

p (probability of success on a given trial)
n (number of trials)
k (number of successes)

P(X=43) = 0.03007

P(X<43) = 0.06661

P(X≤43) = 0.09667

P(X>43) = 0.90333

P(X≥43) = 0.93339

```
function pvalue() {

//get input values
var p = document.getElementById('p').value*1;
var n = document.getElementById('n').value*1;
var k = document.getElementById('k').value*1;

//assign probabilities to variable names
var exactProb = jStat.binomial.pdf(k,n,p);
var lessProb = jStat.binomial.cdf(k-1,n,p);
var lessEProb = jStat.binomial.cdf(k,n,p);
var greaterProb = 1-jStat.binomial.cdf(k,n,p);
var greaterEProb = 1-jStat.binomial.cdf(k-1,n,p);

//output probabilities
document.getElementById('k1').innerHTML = k;
document.getElementById('k2').innerHTML = k;
document.getElementById('k3').innerHTML = k;
document.getElementById('k4').innerHTML = k;
document.getElementById('k5').innerHTML = k;

document.getElementById('exactProb').innerHTML = exactProb.toFixed(5);
document.getElementById('lessProb').innerHTML = lessProb.toFixed(5);
```

```
document.getElementById('lessEProb').innerHTML = lessEProb.toFixed(5);
document.getElementById('greaterProb').innerHTML = greaterProb.toFixed(5);
document.getElementById('greaterEProb').innerHTML = greaterEProb.toFixed(5);
}
```

The interpretation of these calculated values must be precise, distinguishing between "exactly," "less than," and "less than or equal to" the critical value k. Utilizing a calculator provides these five critical interpretations instantly:

The probability that the coin lands on heads **exactly 43 times** ($P(X=43)$) is **0.03007**. This is the PMF value.
The probability that the coin lands on heads **less than 43 times** ($P(X<43)$ or $P(Xle 42)$) is **0.06661**. This is a cumulative probability up to $k-1$.
The probability that the coin lands on heads **43 times or less** ($P(Xle 43)$) is **0.09667**. This is the CDF value.
The probability that the coin lands on heads **more than 43 times** ($P(X>43)$ or $P(Xge 44)$) is **0.90333**. This is calculated as $1 - P(Xle 43)$.
The probability that the coin lands on heads **43 times or more** ($P(Xge 43)$) is **0.93339**. This is calculated as $1 - P(Xle 42)$.

Measures of Central Tendency and Dispersion

Like any probability distribution, the binomial distribution is fully characterized by specific parameters that define its center, spread, and shape. These fundamental descriptive statistics--the mean, variance, and standard deviation--allow us to understand the expected outcomes and the variability surrounding those expectations without having to calculate every individual probability.

For a binomial distribution defined by n trials and probability p, these measures are calculated using remarkably straightforward formulas:

The Mean (mu): Also known as the expected value $E(X)$, the mean represents the average number of successes anticipated over the long run.

$$ mu = np $$

The Variance ($sigma^2$): This measures the spread or dispersion of the distribution around the mean. A larger variance indicates a wider distribution of outcomes.

$$ sigma^2 = np(1-p) $$

The Standard Deviation ($sigma$): This is the square root of the variance and is often preferred as it is expressed in the same units as the random variable X.

$$ sigma = sqrt{np(1-p)} $$

Applying these formulas provides immediate insight. Returning to the example of flipping a coin 3 times ($n=3$, $p=0.5$), we calculate:

The expected number (mean) of heads: $mu = 3 times 0.5 = mathbf{1.5}$. This result makes intuitive sense, as we expect half of the 3 flips to be heads.
The variance of the number of heads: $sigma^2 = 3 times 0.5 times (1-0.5) = 3 times 0.5 times 0.5 = mathbf{0.75}$.

Practical Application Problems

To solidify your understanding of both the Binomial PMF, CDF, and the distribution's properties, review the following practical examples which demonstrate how to correctly identify n, k, and p, and interpret the results using the relevant formulas or computational tools.

Problem 1: Calculating Exact Probability

Question: Bob makes 60% of his free-throw attempts. If he shoots 12 free throws, what is the probability that he makes exactly 10?

Answer: Using the Binomial Distribution Calculator above with $p = 0.6$, $n = 12$, and $k = 10$, we find that $P(X=10) = mathbf{0.06385}$.

Problem 2: Calculating Cumulative Probability (Upper Bound)

Question: Jessica flips a fair coin 5 times. What is the probability that the coin lands on heads 2 times or fewer?

Answer: Using the Binomial Distribution Calculator above with $p = 0.5$, $n = 5$, and $k = 2$, we find that $P(Xle 2) = mathbf{0.5}$.

Problem 3: Calculating Complementary Probability (Lower Bound)

Question: The probability that a given student gets accepted to a certain college is 0.2. If 10 students apply, what is the probability that more than 4 get accepted?

Answer: Using the Binomial Distribution Calculator above with $p = 0.2$, $n = 10$, and $k = 4$, we find that $P(X>4) = mathbf{0.03279}$.

Problem 4: Calculating Expected Value

Question: You flip a coin 12 times. What is the mean expected number of heads that will show up?

Answer: Recall that the mean of a binomial distribution is calculated as $mu = np$. Thus, $mu = 12 times 0.5 = mathbf{6}$ heads.

Problem 5: Calculating Variance

Question: Mark hits a home run during 10% of his attempts. If he has 6 attempts in a given game, what is the variance of the number of home runs he'll hit?

Answer: Recall that the variance of a binomial distribution is calculated as $sigma^2 = np(1-p)$. Thus, $sigma^2 = 6 times 0.1 times (1-0.1) = mathbf{0.54}$.

Further Resources and Statistical Software Implementation

Implementing the binomial distribution calculations in statistical software packages is a necessary skill for advanced analysis. These functions often utilize built-in commands for the PMF (probability density function) and CDF (cumulative distribution function). The following resources provide detailed guides on applying these principles using popular tools:

How to calculate binomial probabilities in Excel
How to calculate binomial probabilities on a TI-84 calculator
How to calculate binomial probabilities in R
How to plot a binomial distribution in R