

How to Perform a Binomial Test in R and Interpret the Results

Authored by
stats writer

March 14, 2026

RECOMMENDED CITATION

stats writer (2026). *How to Perform a Binomial Test in R and Interpret the Results*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=135627>

Introduction to the Binomial Test in R

The **Binomial Test** serves as a cornerstone of **non-parametric statistics**, specifically designed to analyze **binary outcomes**. In the realm of data science and research, we frequently encounter variables that can only take two forms, such as success or failure, heads or tails, or the presence or absence of a specific trait. When a researcher wishes to determine if the **observed proportion** of these outcomes deviates significantly from a theoretically expected value, the **Binomial Test** provides the necessary mathematical rigor to reach a conclusion. This test is particularly valuable because it does not rely on the assumption of a **normal distribution**, making it an "exact" test that is reliable even with small sample sizes.

Performing this analysis within the **R programming language** is remarkably efficient due to the built-in functionality of the **binom.test()** function. By leveraging this tool, users can input their **observed data** and receive a comprehensive summary that includes the **p-value**, **confidence intervals**, and **point estimates**. Understanding how to interpret these outputs is crucial for professionals in fields ranging from **biostatistics** to **econometrics**, where making data-driven decisions based on **statistical significance** is a daily requirement.

Before executing the test, it is essential to define the **significance level**, commonly denoted as alpha (α). This threshold represents the probability of rejecting the **null hypothesis** when it is actually true, a scenario known as a **Type I error**. Typically, an alpha of 0.05 is chosen, though this can vary depending on the stringency required by the specific discipline or study. Once the parameters are set, the **Binomial Test** evaluates the likelihood of the observed result occurring by chance, allowing the researcher to either support or challenge their initial **hypothesis** with mathematical confidence.

The Mathematical Framework: Null and Alternative Hypotheses

The foundation of any **statistical hypothesis test** lies in the formulation of the **null hypothesis** (H_0) and the **alternative hypothesis** (H_A). For a **Binomial Test**, these hypotheses are centered around the **population proportion**, often represented by the Greek letter π . The null hypothesis generally posits that the observed **success rate** is equal to a specific hypothesized value, denoted as p . This serves as the default assumption that there is no "real" difference or effect, and any variation seen in the sample is purely due to **sampling error**.

Mathematically, the **null hypothesis** is expressed as **$H_0: \pi = p$** . Conversely, the **alternative hypothesis** represents the research claim or the presence of an effect. In a **two-tailed test**, the alternative hypothesis is **$H_A: \pi \neq p$** , suggesting that the true population proportion is either significantly higher or significantly lower than the hypothesized value. This approach is comprehensive and is used when the researcher has no prior expectation about the direction of the

difference.

In addition to two-tailed tests, researchers may employ **one-tailed tests** when they have a directional prediction. A **left-tailed test** assumes that the true proportion is less than the hypothesized value ($H_A: \pi < p$), while a **right-tailed test** assumes the proportion is greater ($H_A: \pi > p$). Choosing the correct **alternative hypothesis** is a critical step in **experimental design**, as it directly influences how the **p-value** is calculated and interpreted. Rigorous adherence to these definitions ensures that the resulting **statistical inference** is both valid and reproducible.

Understanding the `binom.test()` Function Syntax in R

To execute a **Binomial Test** in **R**, one utilizes the `binom.test()` function, which is part of the standard **stats package**. The function is highly versatile, allowing for simple inputs while generating a sophisticated analysis of **Bernoulli trials**. The basic syntax requires three primary **arguments** that define the scope of the data being analyzed. Mastery of these parameters is essential for any user looking to perform accurate **data analysis**.

`binom.test(x, n, p)`

The first argument, **x**, represents the number of observed successes in the sample. A "success" is defined by the researcher and refers to the outcome of interest within the **binary set**. The second argument, **n**, is the total number of trials or the **sample size**. Finally, **p** denotes the probability of success on a single trial as predicted by the **null hypothesis**. For instance, if testing a fair coin, **p** would be set to 0.5. If these arguments are not explicitly defined, **R** defaults to a probability of 0.5 and a two-sided alternative hypothesis.

x: This identifies the specific count of "successful" outcomes recorded during the observation phase.

n: This represents the exhaustive count of all independent trials conducted during the experiment.

p: This is the theoretical **probability** of success against which the sample is being measured.

The `binom.test()` function also includes optional parameters such as **alternative**, which can be set to "two.sided", "less", or "greater" to match the research **hypothesis**. Additionally, the **conf.level** parameter allows users to adjust the width of the **confidence interval**, providing flexibility for different **significance levels**. This level of customization makes **R** an ideal environment for complex **statistical modeling**.

Example 1: Performing a Two-Tailed Binomial Test

Consider a scenario where you are investigating the **randomness** of a six-sided die. Specifically, you want to determine if the die is fair regarding the number "3". In a perfectly **uniform**

distribution, any single side of a six-sided die should land face up with a **probability** of 1/6. To test this, you conduct an experiment where you roll the die 24 times. During these trials, the number "3" appears 9 times. This **frequency** seems higher than expected, so a **Binomial Test** is employed to see if this deviation is statistically significant.

#perform two-tailed Binomial test

```
binom.test(9, 24, 1/6)
```

```
#output
```

```
Exact binomial test
```

```
data: 9 and 24
```

```
number of successes = 9, number of trials = 24, p-value = 0.01176
```

```
alternative hypothesis: true probability of success is not equal to 0.1666667
```

```
95 percent confidence interval:
```

```
0.1879929 0.5940636
```

```
sample estimates:
```

```
probability of success
```

```
0.375
```

Upon executing the code in **R**, the output provides a **p-value** of **0.01176**. Because we are using a standard **significance level** of 0.05, we compare our result to this threshold. Since 0.01176 is significantly lower than 0.05, we have sufficient evidence to **reject the null hypothesis**. This suggests that the die is likely biased, as the rate of landing on "3" (37.5%) is statistically different from the expected rate of 16.67%.

The **confidence interval** provided in the output (approximately 0.188 to 0.594) further supports this conclusion. Since the hypothesized value of 0.1667 (1/6) falls outside of this interval, it reinforces the decision to reject the **null hypothesis**. This example illustrates how a two-tailed test captures deviations in either direction, providing a robust check against the assumption of a fair process.

Example 2: Analyzing Data with a Left-Tailed Binomial Test

A **left-tailed test** is utilized when the research objective is to determine if a proportion is significantly *lower* than a hypothesized value. Suppose you suspect that a specific coin is weighted to favor tails over heads. To investigate this potential bias, you perform 30 coin flips and record that the coin lands on heads only 11 times. Given that a **fair coin** has a 50% chance of landing on heads, you apply the **Binomial Test** with a directional alternative to see if the success rate is indeed less than 0.5.

```
#perform left-tailed Binomial test  
binom.test(11, 30, 0.5, alternative="less")
```

```
#output
```

```
Exact binomial test
```

```
data: 11 and 30
```

```
number of successes = 11, number of trials = 30, p-value = 0.1002
```

```
alternative hypothesis: true probability of success is less than 0.5
```

```
95 percent confidence interval:
```

```
0.0000000 0.5330863
```

```
sample estimates:
```

```
probability of success
```

```
0.3666667
```

The resulting **p-value** for this analysis is **0.1002**. In the context of a 5% **significance level**, this value exceeds the 0.05 cutoff. Consequently, we **fail to reject the null hypothesis**. Although the observed proportion of 36.67% is lower than the expected 50%, the **sample size** and the margin of difference are not substantial enough to rule out random chance as the cause of this discrepancy.

In this instance, the **confidence interval** extends from 0.00 to 0.533. Because the hypothesized value of 0.5 is contained within this interval, the **statistical inference** remains that the coin could still be fair. This highlights a critical aspect of **hypothesis testing**: failing to reject the null does not prove the null is true, but rather indicates that the evidence is insufficient to conclude otherwise.

Example 3: Verifying Improvements with a Right-Tailed Binomial Test

In industrial **quality control**, **one-tailed tests** are frequently used to validate system improvements. Imagine a manufacturing shop that produces widgets with a historical effectiveness rate of 80%. The management implements a new production system designed to increase this rate. To evaluate the success of the new system, 50 widgets are randomly sampled, and 46 are found to be effective. The goal is to determine if this 92% success rate represents a significant improvement over the baseline 80%.

```
#perform right-tailed Binomial test  
binom.test(46, 50, 0.8, alternative="greater")
```

```
#output
```

```
Exact binomial test
```

```
data: 46 and 50
```

```
number of successes = 46, number of trials = 50, p-value = 0.0185
alternative hypothesis: true probability of success is greater than 0.8
95 percent confidence interval:
0.8262088 1.0000000
sample estimates:
probability of success
0.92
```

The **p-value** generated by **R** is **0.0185**. Since this is less than the standard 0.05 **significance level**, we **reject the null hypothesis**. We now have sufficient statistical evidence to conclude that the new system has successfully improved the production effectiveness rate beyond the original 80% threshold. This type of analysis is vital for **operations management** when deciding whether to scale a new process.

Furthermore, the 95% **confidence interval** starts at 0.826, which is strictly greater than our **null hypothesis** value of 0.80. This gives the stakeholders confidence that the true performance of the new system is likely above the old standard. By using the **Binomial Test**, the shop can move forward with the new system, backed by a rigorous **quantitative analysis** rather than mere intuition.

Interpreting the Components of the R Output

The output of the **binom.test()** function in **R** is rich with information that requires careful **interpretation**. The first line confirms the type of test performed--the "Exact binomial test"--reminding the user that this is a **non-parametric** procedure. The "data" section recaps the number of successes and trials, ensuring that the inputs were correctly processed by the **compiler**. However, the most critical element for decision-making is the **p-value**.

The **p-value** represents the probability of observing a result at least as extreme as the one obtained, assuming the **null hypothesis** is true. A low p-value suggests that the observed data is unlikely to have occurred by chance alone. In addition to the p-value, **R** provides a **confidence interval**, which offers a range of plausible values for the true **population proportion**. If the hypothesized proportion lies within this interval, the results are typically considered non-significant at the chosen alpha level.

Finally, the "sample estimates" section provides the **point estimate**, which is simply the observed **sample proportion** (x divided by n). While this estimate is the best guess for the true proportion based on the sample, the **Binomial Test** adds depth by quantifying the **uncertainty** around that guess. By synthesizing the p-value, the confidence interval, and the point estimate, researchers can form a holistic view of their data's **statistical significance**.

Practical Applications Across Diverse Fields

The **Binomial Test** is an incredibly versatile tool with applications across many scientific and professional disciplines. In **psychology**, it might be used to determine if a participant in a forced-choice task is performing better than chance, which would suggest some level of underlying **cognition** or perception. For example, if a subject correctly identifies a stimulus in 75 out of 100 trials where the chance of guessing is 50%, the **Binomial Test** would confirm if this performance is statistically significant.

In the field of **genetics**, researchers use binomial testing to analyze **Mendelian inheritance** patterns. If a specific cross of organisms is expected to produce a certain phenotype in 25% of the offspring, the **Binomial Test** can evaluate whether the actual counts observed in the laboratory deviate from these theoretical **genetic ratios**. This helps scientists identify potential deviations such as **genetic linkage** or lethal alleles that might be affecting the survival rates of certain genotypes.

Furthermore, in **marketing research**, companies use binomial tests to evaluate **A/B testing** results. If a marketing team wants to know if a new advertisement leads to a **conversion rate** higher than the industry standard of 2%, they can use a right-tailed binomial test. By analyzing the number of clicks or purchases relative to the total number of impressions, the team can determine if the new strategy is effectively driving consumer behavior or if the results are merely a product of **random variation**.

Assumptions and Limitations of the Binomial Test

While the **Binomial Test** is powerful, its validity relies on several core assumptions regarding the **data collection** process. First and foremost is the assumption of **independence**. Each trial in the experiment must be independent of the others, meaning the outcome of one trial cannot influence the outcome of another. In the case of coin flips or dice rolls, this is easily satisfied, but in social science research, ensuring independence requires rigorous **random sampling** techniques.

Another critical requirement is that the **probability** of success must remain constant across all trials. If the conditions of the experiment change--such as a participant becoming fatigued or a machine wearing down--the underlying probability **p** may shift, violating the assumptions of the **binomial distribution**. Researchers must carefully control the **experimental environment** to ensure that the success rate is stable throughout the duration of the study.

Finally, the **Binomial Test** is strictly for **categorical data** with exactly two levels. If the data involves multiple categories (e.g., "low," "medium," and "high"), a different **goodness-of-fit test**, such as the **Chi-squared test**, would be more appropriate. Despite these limitations, the binomial test remains a preferred choice for **exact testing** because it provides accurate p-values regardless

of how small the sample size might be, unlike many other tests that rely on large-sample **asymptotic approximations**.

ARABPSYCHOLOGY.COM