

Understanding the Multinomial Test: A Guide to Comparing Observed and Expected Frequencies

Authored by
Mohammed loot

November 5, 2025

RECOMMENDED CITATION

Mohammed loot (2025). *Understanding the Multinomial Test: A Guide to Comparing Observed and Expected Frequencies*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=10546>

The Fundamentals of the Multinomial Test

The [multinomial test](#) stands as a cornerstone in inferential statistics, providing a robust methodology for determining whether observed frequency counts from a finite experiment align with a predefined theoretical framework. Specifically, this powerful statistical tool assesses if the frequencies of a [categorical variable](#)--one that can take on two or more distinct, mutually exclusive outcomes--are consistent with a specific set of expected probabilities. This collection of expectations is formally referred to as the [hypothesized distribution](#).

A key differentiator of the multinomial test is its applicability to scenarios involving multiple outcomes, setting it apart from simpler statistical procedures like the binomial test, which is strictly limited to dichotomous (two-outcome) results. When an experiment can yield k different outcomes, where k is greater than two, the multinomial test provides the necessary mathematical rigor to compare actual experimental results against theoretical expectations. This capability makes it indispensable across various scientific and commercial disciplines, from validating genetic models to analyzing consumer preferences.

To grasp the utility of this test, consider a classic example: examining the fairness of a standard six-sided die. If the die is perfectly fair, the probability of rolling any specific number (1, 2, 3, 4, 5, or 6) should be exactly $1/6$ for each outcome. By conducting numerous trials, the multinomial test allows researchers to statistically compare the recorded observations against this expected uniform distribution, providing a definitive answer regarding the die's balance.

Formulating the Hypotheses: Null versus Alternative

As with all procedures within [frequentist hypothesis testing](#), the multinomial test requires the establishment of two opposing statements: the null hypothesis and the alternative hypothesis. These statements represent the two possible statistical conclusions regarding the underlying distribution of the data, forming the basis upon which all subsequent calculations and decisions rest.

The hypotheses are formally structured to articulate whether the observed data conforms to the expected probabilities:

H₀ (Null Hypothesis): The [categorical variable](#) follows the specified hypothesized distribution. Stated differently, the observed frequencies obtained from the experiment are statistically consistent with the expected probabilities set forth in the theoretical model.

H_A (Alternative Hypothesis): The categorical variable *does not* follow the hypothesized distribution. This means the observed frequencies deviate significantly from the expected probabilities, suggesting that the underlying process is different from the one assumed by H₀.

The critical step in the testing process is the evaluation of the calculated **p-value**. The p-value quantifies the probability of observing the experimental data (or data more extreme) if the null hypothesis were true. The decision rule is straightforward: if the calculated p-value is less than the predetermined **significance level**, typically denoted as α (often set to 0.05), we reject H_0 . Rejecting H_0 provides compelling statistical evidence to conclude that the actual distribution of outcomes is significantly different from the expected distribution, supporting the alternative hypothesis.

Case Study 1: Analyzing the Fairness of a Six-Sided Die

To demonstrate the practical application of the multinomial test, let us revisit the scenario of assessing die fairness. Our objective is to determine if a specific six-sided die is balanced, meaning the probability of landing on any side is equally likely, set at $1/6$. We commence the experiment by rolling the die 30 times and meticulously recording the frequency of each result.

Under the null hypothesis (H_0), the expected probabilities are uniform: $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$. The observed data collected from the 30 rolls are crucial for the analysis. The summary of these observed frequencies is presented in the image below, showing the actual counts obtained during the trials:

Outcome	Probability	Frequency
1	1/6	4
2	1/6	5
3	1/6	2
4	1/6	9
5	1/6	5
6	1/6	5

We execute the exact multinomial test using the statistical programming language **R**, leveraging the specialized `EMT` package designed for this purpose. The following code snippet defines both our expected probabilities (`prob`) and the actual observed frequencies (`actual`) before initiating the calculation:

```
library(EMT)
```

```
#specify probability of each outcome  
prob <- c(1/6, 1/6, 1/6, 1/6, 1/6, 1/6)
```

```
#specify frequency of each outcome from experiment
```

```
actual <- c(4, 5, 2, 9, 5, 5)
```

```
#perform multinomial test
```

```
multinomial.test(actual, prob)
```

Exact Multinomial Test, distance measure: p

Events pObs p.value

```
324632 0 0.4306
```

The output of the statistical analysis yields a calculated **p-value** of **0.4306**. Given that this value is substantially larger than the conventionally accepted significance level ($\alpha = 0.05$), we must conclude that we lack sufficient statistical evidence to reject the null hypothesis (H_0). In practical terms, based on these 30 rolls, there is no statistically significant reason to believe that the die is unfair; the observed deviations are likely due to random chance.

Case Study 2: Testing Uniformity in Commercial Product Sales

The multinomial test holds significant value in commercial analysis, particularly when assessing market uniformity or product preference. Imagine a retail shop owner who wishes to verify if customer purchasing preferences are equally distributed among four unique products: A, B, C, and D. The null hypothesis posits that sales are uniform, meaning the expected purchase probability for each product is $1/4$ (or 25%).

The owner diligently tracks sales over a one-week period, resulting in a set of observed frequencies that show some variation. The total sales and their distribution across the four categories are provided in the figure below:

Product	Probability	Sales
A	$1/4$	40
B	$1/4$	20
C	$1/4$	30
D	$1/4$	50

While a visual inspection of the observed frequencies (40, 20, 30, 50) clearly suggests a deviation from the expected uniform distribution, the multinomial test is necessary to confirm whether this deviation is statistically meaningful or merely random noise. We input the expected uniform

probabilities and the actual observed counts back into the [R](#) environment for computation:

library(EMT)

```
#specify probability of each outcome
prob <- c(1/4, 1/4, 1/4, 1/4)

#specify frequency of each outcome from experiment
actual <- c(40, 20, 30, 50)

#perform multinomial test
multinomial.test(actual, prob)
```

Exact Multinomial Test, distance measure: p

```
Events pObs p.value
477191 0 0.00226
```

Upon execution, the resulting [p-value](#) is determined to be **0.00226**. Since this value is significantly lower than the standard [significance level](#) ($\alpha = 0.05$), we are compelled to reject the null hypothesis (H_0). This statistical outcome confirms that the sales distribution is definitively *not* uniform across the four products; the observed preference for certain items is statistically significant, requiring the shop owner to adjust inventory or marketing strategies accordingly.

Case Study 3: Verifying a Claimed Population Proportion

Beyond testing uniformity, the multinomial test is invaluable for verifying theoretical claims regarding population proportions or expected distributions derived from prior knowledge. Consider a scenario where an individual, Tom, claims that a bag contains marbles of three colors--red, green, and purple--with specific drawing probabilities (with replacement) of 0.2, 0.5, and 0.3, respectively. To test this claim, Mike conducts an experiment, drawing a marble 100 times.

Tom's claim establishes the null hypothesis. The total counts of observed outcomes gathered by Mike are summarized in the following results table, showing how frequently each color was drawn:

Color	Probability	Frequency
Red	0.2	25
Green	0.5	45
Purple	0.3	30

We define the expected probabilities (0.2, 0.5, 0.3) as the basis for H_0 and proceed with the analysis using the statistical capabilities of [R](#). The test compares the actual frequencies observed by Mike against the hypothesized distribution to calculate the appropriate test statistic and the resulting p-value.

library(EMT)

```
#specify probability of each outcome
prob <- c(.2, .5, .3)

#specify frequency of each outcome from experiment
actual <- c(40, 20, 30, 50)

#perform multinomial test
multinomial.test(actual, prob)
```

Exact Multinomial Test, distance measure: p

```
Events pObs p.value
5151 0.0037 0.3999
```

The resulting [p-value](#) calculated for this specific set of observations is **0.3999**. Since 0.3999 far exceeds the conventional critical threshold of $\alpha = 0.05$, we must fail to reject the null hypothesis. This outcome indicates that there is insufficient statistical evidence to conclude that the actual distribution of marbles in the bag differs from the probabilities claimed by Tom (0.2, 0.5, 0.3). The observed fluctuations are statistically insignificant.

Conclusion and Practical Applications

The [multinomial test](#) is an indispensable and versatile tool for the rigorous analysis of discrete, [categorical data](#) involving multiple possible outcomes. Its primary function is to provide a statistical confirmation of whether observed experimental results align with, or significantly deviate from, a theoretical expectation or a previously established hypothesized distribution.

Successful application of this test hinges on meeting a few key requirements: the necessity of having fixed, expected probabilities defined by the null hypothesis, and ensuring that the total number of experimental trials is large enough to adequately observe the frequencies across all potential outcome categories. When these conditions are met, the multinomial test offers a definitive mechanism for drawing strong statistical conclusions.

Mastering the application of the multinomial test--whether verifying the fairness of equipment, analyzing consumer sales patterns, or validating population proportion claims--is fundamental for

achieving accurate data interpretation across numerous quantitative fields, including market research, quality control, biology, and social sciences.

Further Statistical Resources

For researchers and students seeking to deepen their understanding of related concepts and alternative statistical testing methods, the following resources are highly recommended:

[The Chi-Squared Test for Goodness of Fit](#) (This is a closely related asymptotic test often employed for similar purposes, especially when dealing with very large sample sizes).

[EMT Package Documentation](#) (The official technical documentation for the R package utilized throughout the illustrative examples in this article).