How can the Multinomial Distribution be used in R?

Authored by

stats writer

July 1, 2024

The Multinomial Distribution is a statistical concept that can be used in the R programming language to model the probability of multiple outcomes occurring simultaneously. It is a generalization of the Binomial Distribution, which only considers the probability of two outcomes. In R, the Multinomial Distribution can be used to analyze data with multiple categories or to simulate discrete events that have multiple possible outcomes. This distribution is particularly useful in fields such as genetics, market research, and quality control, where there are multiple possible outcomes for a given event. By utilizing the Multinomial Distribution in R, researchers and analysts can better understand the likelihood of various outcomes and make more informed decisions based on the data.

Use the Multinomial Distribution in R

The describes the probability of obtaining a specific number of counts for *k* different outcomes, when each outcome has a fixed probability of occurring.

If a *X* follows a multinomial distribution, then the probability that outcome 1 occurs exactly x1 times, outcome 2 occurs exactly x2 times, etc. can be found by the following formula:

Probability = n! * (p1x1 * p2x2 * ... * pkxk) / (x1! * x2! ... * xk!)

where:

n: total number of eventsx1: number of times outcome 1 occursp1: probability that outcome 1 occurs in a given trial

To calculate a multinomial probability in R we can use the dmultinom() function, which uses the following syntax:

dmultinom(x=c(1, 6, 8), prob=c(.4, .5, .1))

where:

x: A vector that represents the frequency of each outcomeprob: A vector that represents the probability of each outcome (the sum must be 1)

The following examples show how to use this function in practice.

Example 1

In a three-way election for mayor, candidate A receives 10% of the votes, candidate B receives 40% of the votes, and candidate C receives 50% of the votes.

If we select a random sample of 10 voters, what is the probability that 2 voted for candidate A, 4 voted for candidate B, and 4 voted for candidate C?

We can use the following code in R to answer this question:

```
#calculate multinomial probability
dmultinom(x=c(2, 4, 4), prob=c(.1, .4, .5))
```

0.0504

The probability that exactly 2 people voted for A, 4 voted for B, and 4 voted for C is 0.0504.

Example 2

If we randomly select 4 balls from the urn, with replacement, what is the probability that all 4 balls are yellow?

We can use the following code in R to answer this question:

```
#calculate multinomial probability
dmultinom(x=c(4, 0, 0), prob=c(.6, .2, .2))
```

0.1296

The probability that all 4 balls are yellow is 0.1296.

Example 3

Suppose two students play chess against each other.

The probability that student A wins a given game is 0.5, the probability that student B wins a given game is 0.3, and the probability that they tie in a given game is 0.2.

If they play 10 games, what is the probability that player A wins 4 times, player B wins 5 times, and they tie 1 time?

We can use the following code in R to answer this question:

```
#calculate multinomial probability
dmultinom(x=c(4, 5, 1), prob=c(.5, .3, .2))
```

```
0.0382725
```

The probability that player A wins 4 times, player B wins 5 times, and they tie 1 time is about 0.038.

Additional Resources

The following tutorials provide additional information about the multinomial distribution: