How to make a multinomial distribution?

Authored by

stats writer

December 26, 2025

The multinomial distribution is a crucial concept within probability distribution theory, serving as a powerful tool for modeling complex scenarios. Unlike the simpler binomial distribution, which handles only two possible outcomes (success or failure) in a series of trials, the multinomial distribution allows for any finite number of distinct, discrete outcomes, provided these trials are independent. To construct a valid multinomial distribution model, one must precisely define three fundamental parameters: the total number of independent trials (n), the finite set of possible outcomes (k), and the established probability (pi) associated with each outcome. Once these parameters are established, the distribution enables the calculation of the probability for any specific combination of counts for these various outcomes occurring over the total number of trials.

Understanding this distribution is paramount for statisticians and data scientists who frequently deal with classifications and frequency analysis where results fall into multiple categories. This mathematical framework provides the mechanism for determining the likelihood of observing a particular frequency pattern--for example, predicting the vote counts for multiple candidates in an election, or the frequency of different colored objects drawn from a container with replacement. The calculation fundamentally involves determining how many distinct ways a specified set of results can occur, multiplied by the probability of that particular sequence happening, making it essential for advanced statistical inference and modeling in various fields.

Defining the Multinomial Distribution

The multinomial distribution rigorously describes the probability of obtaining a specified vector of counts (x1, x2, ..., xk) for **k** different categories or outcomes, assuming a fixed total number of independent trials, denoted as **n**. Crucially, each of these **k** outcomes must have a constant and known probability (p1, p2, ..., pk) of occurring in any single trial. It is imperative that the sum of these probabilities equals one (i.e., p1 + p2 + ... + pk = 1) and that the sum of the desired counts equals the total number of trials (i.e., x1 + x2 + ... + xk = n). This structure makes the multinomial distribution the direct generalization of the well-known binomial distribution, extending the model from binary outcomes to scenarios involving three or more possible results.

When we state that a random vector **X** = (X1, X2, ..., Xk) follows a multinomial distribution, we are interested in finding the precise probability that outcome 1 occurs exactly x1 times, outcome 2 occurs exactly x2 times, and so forth, up to outcome **k** occurring exactly xk times. This powerful framework allows for the analysis of complex, multivariate counts. For instance, if a manufacturer produces items categorized as "Good," "Defective Minor," or "Defective Major," the multinomial distribution helps calculate the likelihood of drawing a specific mix of these categories in a sample of fixed size, assuming the overall production probabilities remain constant across trials. This adherence to constant probabilities and independence between trials is what defines its applicability in statistical modeling.

The Multinomial Probability Mass Function (PMF)

The probability mass function (PMF) provides the mathematical mechanism to calculate the specific probability for any defined set of outcomes (x1, x2, ..., xk). This formula combines the combinatorial aspect (the number of ways to arrange the outcomes) with the probability aspect (the likelihood of that specific sequence occurring). The resulting calculation yields the probability of the entire event occurring simultaneously. The formula is expressed as follows:

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

Understanding the components of this formula is critical for accurate computation. The numerator, **n!** (n factorial), represents the total number of ways to arrange **n** trials if all outcomes were distinct. However, since we are interested in specific counts (xi), we must divide by the factorials of the counts (the denominator) to account for the indistinguishability of identical outcomes, giving us the multinomial coefficient. The second part of the numerator, the product term **(p1x1 * p2x2 * ... * pkxk)**, represents the probability of one specific sequence of outcomes occurring, based on the product rule of independent events. The PMF thus elegantly combines these elements to yield the final likelihood.

The variables used in this function represent critical elements of the experimental setup:

n: The **total number of events** or independent trials conducted in the experiment. This must be a fixed integer defined beforehand.
x1: The specific **number of times outcome 1 occurs**. The sum of all xi must equal **n**.
p1: The known, constant **probability that outcome 1 occurs** in a given trial. The sum of all pi must equal 1.

Assumptions Governing Multinomial Experiments

For the multinomial distribution to be an appropriate model for a real-world phenomenon, several strict assumptions must be met. Ignoring these assumptions can lead to invalid probability calculations and faulty statistical inferences. Firstly, the experiment must consist of a fixed number of trials, **n**, and these trials must be identical and independent. The concept of **independence** means that the result of one trial does not influence the result of any subsequent trial--a necessity for the product rule of probabilities to hold true in the formula.

Secondly, every trial must result in one and only one of **k** distinct and mutually exclusive discrete outcomes. If outcomes could overlap or if a trial could yield an outcome outside the defined **k** categories, the model breaks down. This mutual exclusivity ensures that the counting process is clean and comprehensive. Furthermore, the probability of each outcome, pi, must remain **constant** from trial to trial. If, for instance, trials are conducted without replacement, the probabilities change dynamically, and the process would then be better modeled by a multivariate hypergeometric distribution, not the multinomial.

Finally, we are calculating the probability of specific counts (x1, x2, ..., xk). Since the counts must sum up exactly to the total number of trials **n** ($sum_{i=1}^{k} x_i = n$), the constraints on the counts are directly linked to the fixed nature of the trials. These assumptions--fixed **n**, independent trials, mutually exclusive outcomes, and constant probabilities--are what mathematically define a true multinomial experiment, ensuring the validity of the probability mass function.

Practical Application: The Urn Example Calculation

To illustrate how the multinomial distribution is applied, consider a classic probability problem involving an urn. Suppose we have an urn containing 10 marbles: 5 red marbles, 3 green marbles, and 2 blue marbles. Since the drawing is performed **with replacement**, the overall probabilities for each color remain constant throughout the trials. If we decide to randomly select 5 marbles from the urn, with replacement, we want to determine the specific probability of obtaining exactly 2 red marbles, 2 green marbles, and 1 blue marble.

First, we must establish the fixed parameters for our experiment. The total number of trials, **n**, is 5. We have three possible outcomes (k=3): Red, Green, and Blue. The initial probabilities (pi) are derived from the total population of 10 marbles:

p1 (prob. red) = 5/10 = 0.5
p2 (prob. green) = 3/10 = 0.3
p3 (prob. blue) = 2/10 = 0.2

Note that p1 + p2 + p3 = 0.5 + 0.3 + 0.2 = 1.0. Next, we define the specific counts (xi) we are interested in calculating:

x1 (number of red marbles) = 2
x2 (number of green marbles) = 2
x3 (number of blue marbles) = 1

Confirming the constraint, x1 + x2 + x3 = 2 + 2 + 1 = 5 = **n**. Now, we plug these established numbers into the multinomial formula:

Probability = 5! * (0.52 * 0.32 * 0.21) / (2! * 2! * 1!)

The factorial term calculation (the multinomial coefficient) is: 5! / (2! * 2! * 1!) = 120 / 4 = 30. The probability component is: 0.52 * 0.32 * 0.21 = 0.25 * 0.09 * 0.2 = 0.0045. Multiplying these two results yields the final probability: 30 * 0.0045 = **0.135**. This means there is a 13.5% chance of observing exactly this combination of marbles in 5 trials with replacement.

Comparison to the Binomial Distribution

The relationship between the multinomial distribution and the binomial distribution is hierarchical; the binomial distribution is a specific, simpler case of the multinomial distribution. When an experiment has only two possible outcomes--often termed "success" and "failure"--the multinomial framework simplifies perfectly into the binomial framework. If we define k=2, where outcome 1 is "success" (with probability p1) and outcome 2 is "failure" (with probability p2 = 1 - p1), and we are looking for x1 successes and x2 failures (where x1 + x2 = n), the multinomial formula naturally reduces to the standard binomial PMF.

Specifically, in the binomial case, the combinatorial term $n! / (x_1! x_2!)$ is simply the binomial coefficient $binom{n}{x_1}$, which counts the number of ways to achieve x_1 successes in n trials. The probability term $p_1^{x_1} p_2^{x_2}$ becomes $p^{x} (1-p)^{n-x}$. Therefore, whenever an experimental scenario involves categorical results beyond a simple dichotomy, such as classification into three or more groups (e.g., small, medium, large; or A, B, C, D), the multinomial model becomes the necessary and appropriate statistical tool, offering a robust way to analyze the joint probabilities of multiple counts occurring simultaneously.

This distinction highlights the necessity of correctly identifying the underlying structure of the data. Using a binomial model inappropriately for data that has three or more natural categories will lead to loss of information and potentially incorrect probabilistic conclusions. The multinomial distribution thus offers increased flexibility and realism when modeling phenomena across political science, quality control, genetics, and ecology, where outcomes rarely fit into a binary framework.

Multinomial Distribution Practice Problems

To solidify the understanding of the multinomial distribution and its application, the following practice problems are presented. These examples demonstrate various real-world scenarios where the PMF (or a specialized calculator derived from it) is used to find specific probabilities related to multiple outcomes. Note that while manual calculation is essential for grasping the theory, statistical software or a specialized calculator is typically used for rapid, accurate computation of these complex factorials and products in professional practice.

Note: *We will use the Multinomial Distribution Calculator to swiftly compute the answers to these application questions.*

Problem 1: Political Election Analysis

Question: In a three-way election for mayor, historical data indicates that candidate A receives 10% of the votes (pA = 0.10), candidate B receives 40% of the votes (pB = 0.40), and candidate C receives 50% of the votes (pC = 0.50). If we select a random sample of 10 voters (n=10) from the population, what is the probability that exactly 2 voted for candidate A (xA=2), 4 voted for candidate B (xB=4), and 4 voted for candidate C (xC=4)?

This problem requires us to calculate P(XA=2, XB=4, XC=4) given n=10 and the defined probabilities. Since the sum of the counts (2+4+4=10) equals the total trials, and the sum of probabilities (0.10+0.40+0.50=1.00) is one, the multinomial model is appropriate. Plugging these values into the PMF would involve calculating the multinomial coefficient 10! / (2! 4! 4!) and multiplying it by the probability component $(0.10)^2 cdot (0.40)^4 cdot (0.50)^4$.

Answer: Using the Multinomial Distribution Calculator with the specified inputs (n=10, x=(2, 4, 4), p=(0.1, 0.4, 0.5)), we find that the probability of observing exactly this voter distribution in the sample is **0.0504**.

Outcome	Probability	Frequency
Outcome 1	0.10	2
Outcome 2	0.40	4
Outcome 3	0.50	4
Outcome 4		
Outcome 5		
Outcome 6		
Outcome 7		
Outcome 8		
Outcome 9		
Outcome 10		

CALCULATE

Multinomial Probability: **0.050400**

Problem 2: Sampling with a Single Outcome

Question: Suppose an urn contains 10 balls in total: 6 yellow marbles, 2 red marbles, and 2 pink marbles. The resulting probabilities are pY=0.6, pR=0.2, and pP=0.2. If we randomly select 4 balls from the urn (n=4), with replacement, what is the probability that all 4 balls selected are yellow (xY=4, xR=0, xP=0)?

This scenario is an important test case because one of the counts is equal to the total number of trials, and the other counts are zero. The condition "with replacement" is crucial, as it maintains the independence of trials and the constancy of the probabilities (0.6, 0.2, 0.2). We are calculating P(XY=4, XR=0, XP=0). Applying the formula, the multinomial coefficient simplifies significantly: 4! / (4! 0! 0!) = 1 (since 0! = 1). The probability calculation then becomes $(0.6)^4 cdot (0.2)^0 cdot (0.2)^0 = (0.6)^4$.

Answer: Using the Multinomial Distribution Calculator with the inputs (n=4, x=(4, 0, 0), p=(0.6, 0.2, 0.2)), we confirm the calculated probability is **0.1296**. This demonstrates how the multinomial distribution can handle edge cases where all outcomes fall into a single category.

Outcome	Probability	Frequency
Outcome 1	0.6	4
Outcome 2	0.2	0
Outcome 3	0.2	0
Outcome 4		
Outcome 5		
Outcome 6		
Outcome 7		
Outcome 8		
Outcome 9		
Outcome 10		

CALCULATE

Multinomial Probability: **0.129600**

Problem 3: Competitive Games and Ties

Question: Consider a scenario where two students play chess against each other. The possible outcomes are Player A wins, Player B wins, or they tie. The established probabilities based on their skill levels are: probability that student A wins is 0.5 (pA=0.5), probability that student B wins is 0.3 (pB=0.3), and probability of a tie is 0.2 (pT=0.2). If they play a series of 10 games (n=10), what is the probability that player A wins 4 times (xA=4), player B wins 5 times (xB=5), and they tie 1 time (xT=1)?

This example perfectly showcases the multinomial model's utility in competitive scenarios where outcomes are non-binary, requiring the analysis of three distinct discrete outcomes. We are seeking P(XA=4, XB=5, XT=1). We confirm the counts sum to 10 (4+5+1=10) and the probabilities sum to 1.0 (0.5+0.3+0.2=1.0). The required calculation involves the multinomial coefficient 10! / (4!

5! 1!) multiplied by the probability term $(0.5)^4 cdot (0.3)^5 cdot (0.2)^1$. This combination represents the likelihood of observing this specific distribution of wins and ties over the 10 games.

Answer: Using the Multinomial Distribution Calculator with the inputs (n=10, x=(4, 5, 1), p=(0.5, 0.3, 0.2)), we determine that the probability of this specific outcome set is **0.038272**. This probability is relatively low, reflecting the specificity of requiring exactly four wins for A, five wins for B, and one tie, given the underlying probabilities.

Outcome	Probability	Frequency
Outcome 1	0.5	4
Outcome 2	0.3	5
Outcome 3	0.2	1
Outcome 4		
Outcome 5		
Outcome 6		
Outcome 7		
Outcome 8		
Outcome 9		
Outcome 10		

CALCULATE

Multinomial Probability: **0.038272**

Further Exploration in Statistical Distributions

Mastering the multinomial distribution provides a strong foundation for understanding multivariate analysis and complex counting processes. While the binomial and multinomial distributions cover discrete, independent trials, other distributions are required for different data structures (e.g., continuous data, dependency between trials, or rare events).

The following tutorials provide an introduction to other common distributions in statistics, such as the Poisson distribution for event rates, the Normal distribution for continuous variables, and the Hypergeometric distribution for sampling without replacement scenarios, all of which are essential tools in a statistician's toolkit for comprehensive data analysis.