CS109: Probability for Computer Scientists
Lecture 13 — Multinomial Distribution
Feb 4, 2026

---

1 Counting Warmup

How many ways are there to order 3 A's, 5 B's, and 2 C's?

a) Calculate the specific number of ways to arrange the 10 letters: $\{A, A, A, B, B, B, B, B, C, C\}$.

Solution

This is a permutation with indistinguishable objects:

$$\frac{10!}{3!5!2!} = \frac{3,628,800}{6 \times 120 \times 2} = \frac{3,628,800}{1,440} = 2,520$$

b) State the general formula for permutations of a multiset of n elements where there are $c_1, c_2, \ldots, c_k$ elements of each type.

Solution

The multinomial coefficient:

$$\binom{n}{c_1, c_2, \ldots, c_k} = \frac{n!}{c_1!c_2! \ldots c_k!}$$

where $\sum_{i=1}^{k} c_i = n$.

2 Joint Probability Tables

Suppose you roll 100 fair dice. Let $X_1, X_2, \ldots, X_6$ be the number of 1s, 2s, …, 6s rolled.

- $X_1 + X_2 + X_3 + X_4 + X_5 + X_6 = 100$

a) How many entries/outcomes are in the joint distribution table of $X_1, \ldots, X_6$?

Solution

The size of this table depends on how the variables are related. This problem illustrates how joint tables can quickly become very large.
1. Without the sum constraint: If the number of 1s did not impact the number of 2s, it would be as if we rolled 100 dice and counted the 1s, then rolled 100 dice again and counted the 2s, and so on. In this case, each of the 6 random variables has 101 options $(0, 1, \ldots, 100)$. The number of entries would be:

$$101^6$$

2. With the sum constraint $(X_1 + \cdots + X_6 = 100)$: Since we are rolling 100 dice once and counting the amount of each number, we must satisfy the constraint that the number of 1s 2s 3s... must sum to 100. We need to split the number 100 into 6 parts (the number of 1s, 2s, 3s, ...). A helpful visualization is to imagine the 100 dice as 100 identical objects, and we want to divide them into 6 buckets.
To do this, we place 5 dividers between the objects to separate the buckets. For example, a particular placement of dividers might look like:

$$* * * \mid * * * * * * * \mid * \mid * * * * * * * * * \mid * * * * \mid * * * * * * * * * * * * * * * *$$

Every valid placement of the 5 dividers gives a unique way to split the 100 dice into 6 buckets. To find the total entries, we consider that there are now 105 total positions (100 dice + 5 dividers). We want to choose all the ways we could order the 5 dividers:

$$\frac{105!}{5!100!} = \binom{105}{5}$$

3 Building Intuition

Consider an experiment where you roll 6 fair, independent six-sided dice.

a) Which outcome is more probable? **A:** Rolling six "6s" OR **B:** Rolling exactly one of each number. Explain **why** (no math needed, just intuition).

Solution

B is much more probable. While any single *sequence* (e.g., $1, 2, 3, 4, 5, 6$ vs $6, 6, 6, 6, 6, 6$) is equally likely, there is only 1 way to get all sixes, whereas there are many ways, specifically 6! ways, to arrange exactly one of each number.

b) **(Funny Shaped Dice)** Suppose a die has: $P(1) = 0.2, P(2) = 0.3, P(3) = 0.1, P(4) = 0.1, P(5) = 0.1, P(6) = 0.2$. You roll it 6 times. What is the probability of getting exactly: **two 2s, two 4s, and two 6s**?

Solution

Using the Multinomial PMF:

$$P(X_2 = 2, X_4 = 2, X_6 = 2) = \frac{6!}{2!2!2!}(0.3)^2(0.1)^2(0.2)^2$$

$$= 90 \times 0.09 \times 0.01 \times 0.04 = 90 \times 0.000036 = 0.00324$$

Problem 5: Multinomial Inference (The Federalist Papers)

Suppose we are trying to determine if Alexander Hamilton (H) or James Madison (M) wrote a specific Federalist Paper. The number of words in the federalist paper is n and there are k unique words.

a) Write an expression for the posterior probability $P(H \mid D)$ using Bayes' Rule.

Solution

By Bayes' Rule:

$$P(H \mid D) = \frac{P(D \mid H)P(H)}{P(D)}$$

b) Write an expression for the likelihood $P(D \mid H)$, where h_i is the probability that Hamilton uses word i.

Solution

Using the multinomial PMF:

$$P(D \mid H) = \binom{n}{c_1, c_2, \ldots, c_k} \prod_{i=1}^{k} h_i^{c_i}$$

c) Write an expression for the ratio of the posteriors $\frac{P(H|D)}{P(M|D)}$. Assume $P(H) = P(M) = 0.5$ and simplify your answer as much as possible.

Solution

When we take the ratio, the priors $P(H)$ and $P(M)$, the evidence $P(D)$, and the multinomial coefficient all cancel out:

$$\frac{P(H \mid D)}{P(M \mid D)} = \frac{\prod_{i=1}^{k} h_i^{c_i}}{\prod_{i=1}^{k} m_i^{c_i}}$$

d) In practice, we often work with the "Log Likelihood Ratio" to avoid underflow. Write an expression for $\log\left(\frac{P(H|D)}{P(M|D)}\right)$ using the simplified ratio from part (c).

Solution

Using the property $\log(A/B) = \log A - \log B$ and expanding the products into sums:

$$\log\left(\frac{\prod_{i=1}^{k} h_i^{c_i}}{\prod_{i=1}^{k} m_i^{c_i}}\right) = \log\left(\prod_{i=1}^{k} h_i^{c_i}\right) - \log\left(\prod_{i=1}^{k} m_i^{c_i}\right)$$

$$= \sum_{i=1}^{k} \log(h_i^{c_i}) - \sum_{i=1}^{k} \log(m_i^{c_i})$$

$$= \sum_{i=1}^{k} c_i \log(h_i) - \sum_{i=1}^{k} c_i \log(m_i)$$

4 Midterm Practice

For the midterm practice problem solution - see Midterm Winter 2025 [Soln] on the midterm page.