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: $\{A, A, A, B, B, B, B, B, C, C\}$.

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.

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 $X_1, \ldots, X_6$?

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).

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**?

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.

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

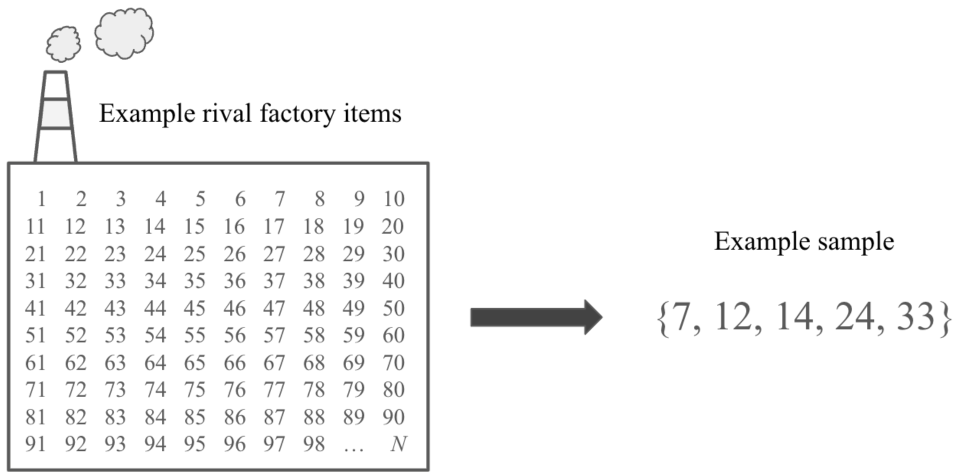
d) In practice, we often work with the "Log Likelihood Ratio" to avoid underflow. Write an expression for $\log\left(\frac{P(H|D)}{P(M|D)}\right)$ using the simplified ratio from part (c).

Midterm Practice (Problem from real midterm last winter)

4. Rival Production [19 Points]

A rival is producing items. We would like to estimate the number of items, N, that they have produced. We notice that each item has a unique serial number and we assume that when we acquire (sample) items each serial number on the item is a positive integer equally likely to be any number from the set $\{1, 2, \ldots, N\}$.

For example, if you randomly acquired (sampled) 5 items produced at the factory, you might see the serial numbers $\{7, 12, 14, 24, 33\}$ which should give you a clue as to what N could be!



a. (7 points) For part (a) only, assume $N = 100$. We sample 5 items. What is the probability that the largest serial number in our sample is 33?

b. (10 points) Your prior belief is that every value of N between 33 and 100 (inclusive) is equally likely. What is your updated probability mass function for N, given that you sampled 5 items and the largest serial number was 33?

c. (3 points) Given that you sampled 5 items and the largest serial number was 33, what is the probability that $N < 50$?