CS109 Lecture 13 Worksheet: ANSWER KEY

Multinomial

Summer 2026

Lecture 13 Worksheet: Answer Key

Instructor / TA copy.

Problem 1 (Review): A Two-Node Network

(a) $P(\text{Ov} = 1, \text{Slow} = 0) = P(\text{Ov} = 1)\, P(\text{Slow} = 0 \mid \text{Ov} = 1) = 0.1 \times 0.05 = 0.005$.

(b) $P(\text{Slow} = 1) = 0.95(0.1) + 0.2(0.9) = 0.095 + 0.18 = 0.275$, so

$$P(\text{Ov} = 1 \mid \text{Slow} = 1) = \frac{0.095}{0.275} = \frac{19}{55} \approx 0.345.$$

(c) Throw away every sample in which Slow $\neq 1$; among the kept samples, report the fraction with Overloaded $= 1$.

Problem 2: Counting Orderings of a DNA Strand

(a)

$$\binom{12}{4,\ 3,\ 2,\ 3} = \frac{12!}{4!\,3!\,2!\,3!} = \frac{479{,}001{,}600}{24 \cdot 6 \cdot 2 \cdot 6} = 277{,}200.$$

(b) Each specific sequence has probability $(1/4)^{12}$, and the orderings are mutually exclusive:

$$277{,}200 \times \left(\tfrac{1}{4}\right)^{12} = \frac{277{,}200}{16{,}777{,}216} \approx 0.0165.$$

Problem 3: Load Balancing

(a) $(X_1, X_2, X_3) \sim \text{Multinomial}(n = 10;\ 0.5, 0.3, 0.2)$:

$$P(X_1 = c_1, X_2 = c_2, X_3 = c_3) = \binom{10}{c_1,\ c_2,\ c_3} (0.5)^{c_1} (0.3)^{c_2} (0.2)^{c_3}, \qquad c_1 + c_2 + c_3 = 10.$$

(b)

$$\binom{10}{5,\ 3,\ 2} (0.5)^5 (0.3)^3 (0.2)^2 = 2520 \times 0.03125 \times 0.027 \times 0.04 \approx 0.0851.$$

Problem 4: Loot Boxes

(a)
$$\binom{8}{4,\ 3,\ 1}(0.6)^4(0.3)^3(0.1)^1 = 280 \times 0.1296 \times 0.027 \times 0.1 \approx 0.0980.$$

(b) Each box is "legendary" (prob 0.1) or "not legendary" (prob 0.9), independently, so the legendary count is $\mathrm{Bin}(8, 0.1)$:
$$P(\text{exactly } 1) = \binom{8}{1}(0.1)(0.9)^7 \approx 0.383.$$
Marginally, every count in a multinomial is a binomial.

Problem 5: Midterm Rooms (Winter 2025 Final)

Let X_1, X_2, X_3 be the room counts; $(X_1, X_2, X_3) \sim \text{Multinomial}(300;\ 0.3, 0.2, 0.5)$:
$$P(X_1 = 90, X_2 = 60, X_3 = 150) = \binom{300}{90,\ 60,\ 150}(0.3)^{90}(0.2)^{60}(0.5)^{150} = \frac{300!}{90!\,60!\,150!}(0.3)^{90}(0.2)^{60}(0.5)^{150}.$$
(Numerically this is about 0.00306 — best computed with logs, which foreshadows Problem 6!)

Problem 6: Who Wrote the Pamphlet?

(a) With equal priors, Bayes gives $P(H \mid \text{doc}) \propto P(\text{doc} \mid H)\,P(H) \propto \prod_i P(\text{word}_i \mid H)^{n_i}$. The multinomial coefficient depends only on the word counts, not the author, so it is identical for H and M and cancels in the comparison (as does $P(\text{doc})$ and the equal prior). Words with the same probability under both authors likewise contribute equal factors.

(b) The product of thousands of numbers each $\ll 1$ underflows to 0 in floating point (the computer cannot represent numbers that small), and every author's score becomes 0. Taking logs turns the product into a sum, $\log \prod_i a_i = \sum_i \log a_i$, which stays in a comfortable numeric range (large negative numbers). Since log is monotonically increasing, comparing log-scores gives the same winner as comparing probabilities.

(c) Log-scores (natural log):
$$H:\ 2\ln(0.005) + \ln(0.0001) + 3\ln(0.002) \approx -38.45,$$
$$M:\ 2\ln(0.0005) + \ln(0.001) + 3\ln(0.003) \approx -39.54.$$
H has the larger (less negative) log-score, so H more likely wrote it. The ratio is $e^{-38.45-(-39.54)} = e^{1.09} \approx 3$: author H is roughly 3 times more likely.

Challenge Problem: Multinomials Generalize Binomials

(a) With $r = 2$ we have $X_2 = n - X_1$ and $p_2 = 1 - p_1$:
$$P(X_1 = k, X_2 = n-k) = \frac{n!}{k!\,(n-k)!}\,p_1^k(1-p_1)^{n-k} = \binom{n}{k}p_1^k(1-p_1)^{n-k},$$
which is exactly the $\mathrm{Bin}(n, p_1)$ PMF.

(b) Recolor each trial as "type i" (success, prob p_i) or "anything else" (failure, prob $1 - p_i$). The n trials are independent with the same success probability, so the count of type-i outcomes is $\mathrm{Bin}(n, p_i)$ by definition.