CS109 Lecture 2 Worksheet: Conditional Probability and Bayes

Conditioning, the Chain Rule, Total Probability, and Bayes' Theorem

Summer 2026

Lecture 2 Worksheet: Conditional Probability and Bayes

Topics: conditional probability $P(E \mid F) = \frac{P(E \cap F)}{P(F)}$, the chain rule, the law of total probability, and Bayes' theorem (with applications to spam filtering and medical testing).

How to use this sheet. Work in small groups. A TA and the instructor will circulate to help. We will go over selected solutions on the board. You do not need to finish every part; aim for understanding over speed. No calculators or code required, though a couple of problems suggest a quick check you could run later.

Problem 1 (Review of Lecture 1): Sample Spaces, Mutually Exclusive Events, Complement

A single fair six-sided die is rolled once. Let A be the event "the result is even" and B the event "the result is greater than 3."

(a) Write the sample space S, and find $P(A)$ and $P(B)$ using equally likely outcomes.

(b) List $A \cap B$. Are A and B mutually exclusive?

(c) Find $P(A \cup B)$ by directly counting the outcomes in the union.

(d) Use the complement rule to find $P(\text{the result is not a 6})$.

Problem 2: Conditioning Changes the Sample Space

You roll **two fair six-sided dice** (distinguishable).

(a) Find $P(\text{sum} = 8)$.

(b) Find $P(\text{sum} = 8 \mid \text{the first die shows 5})$. Explain how conditioning on the first die shrinks the sample space.

(c) Find $P(\text{sum} = 8 \mid \text{at least one die shows 5})$. Use the definition $P(E \mid F) = \frac{P(E \cap F)}{P(F)}$ carefully, and explain why your answer differs from part (b).

Problem 3: The Chain Rule (Drawing Without Replacement)

You draw two cards, one after another **without replacement**, from a standard 52-card deck.

(a) Using the chain rule $P(E_1 \cap E_2) = P(E_1)\, P(E_2 \mid E_1)$, find the probability that **both** cards are aces.

(b) Find the probability that the first card is a king and the second is a queen.

(c) Extend the chain rule: find the probability that the first three cards drawn are all hearts.

Problem 4: Law of Total Probability (Spam Filter)

In an inbox, 30% of incoming email is spam and 70% is legitimate ("ham"). The word *"free"* appears in 60% of spam messages but in only 10% of ham messages.

(a) Define the events clearly. Use the law of total probability to find P(a message contains "free").

(b) Sketch the probability tree for this scenario (spam/ham, then "free"/no "free") and label each branch.

Problem 5: Bayes' Theorem (Spam Filter, continued)

Using the same spam filter from Problem 4: a new message arrives and it **contains the word "free."**

(a) Use Bayes' theorem to find P(spam | "free").

(b) Identify the *prior*, the *likelihood*, and the *posterior* in your calculation.

(c) In one sentence, explain how seeing the word "free" updated our belief that the message is spam.

Problem 6: Bayes' Theorem and the Base-Rate Effect (Medical Testing)

A disease is present in 1% of a population. A test has the following characteristics: if a person has the disease, the test is positive 98% of the time (sensitivity); if a person does not have the disease, the test is positive 5% of the time (false-positive rate).

(a) Find P(test positive) using the law of total probability.

(b) A randomly selected person tests positive. Find P(has disease | tests positive).

(c) The answer in (b) is surprisingly small. Re-derive it using **natural frequencies**: imagine 1000 representative people, and count how many have the disease and test positive versus how many are healthy and test positive. Explain the intuition for why a positive result is still more likely to be a false alarm.