Juliette Woodrow
CS109

Section #2
Jan 21st, 2026

Section 2: Core Probability and Random Variables

---

1 Prickly Dictators and Plausible Deniability

You are trying to estimate the fraction of people in a country who support an authoritarian regime. Directly asking "do you support the dictator?" is unsafe and respondents will not respond truthfully. Luckily, you've taken CS109, so, you use the following survey instead:

1. Is your birthday in Jan–Jun?
2. Is the last digit of your ID number even?
3. Do you enjoy surprises?
4. Would you rather have a salad over a soup?
5. Does your first name have an even number of letters?
6. Do you support the dictator?

Instead of collecting answers to each question, you only collect **the total number of "yes" answers** from each person. Because question 6 is sensitive, reporting only the total number of "yes" answers gives respondents plausible deniability. Even then, using our probability muscles, we can still estimate how often people answered "yes" to question 6. Assume the following:

- For each question 1–5, a person answers "yes" with probability 0.5.
- The answers to questions 1–5 are independent of each other. The answer to question 6 is independent of the answers to questions 1–5.
- Let p be the probability that a person answers "yes" to question 6 (supporting the dictator).

a. What is the probability that a randomly chosen person answers "yes" to exactly 4 questions?

b. Suppose you survey 10,000 people, and let 2,500 be the number of respondents who report a total of 4 "yes" answers. Give an estimate for p, the probability that a person answers "yes" to supporting the dictator.

c. (Optional Challenge) Instead of focusing only on the event that a randomly chosen person answers "yes" to exactly 4 of the 6 questions, we can estimate p using the **average** number of "yes" answers.

   Let Y be the total number of "yes" answers out of 6. Find $\mathbb{E}[Y]$ in terms of p, and give an estimate for p.

d. (Optional Challenge) Suppose the dictator learned that a particular respondent reported a total of 4 "yes" answers out of 6. Can the dictator find out whether this person answered "yes" to question 6?

2 Betcoin

Preamble: When a random variable fits neatly into a family we've seen before (e.g. Binomial), we get its expectation for free. When it does not, we have to use the definition of expectation.

Problem: Your friend tells you about the hottest new investment craze: Betcoin! The cryptocurrency is so popular that its value doubles every day. That means that if you invested \$100 today, it would be worth \$200 tomorrow, \$400 the day after tomorrow, etc.

However, market research says that Betcoin's value is likely to crash any day now. Each day, it has a 40% chance of crashing – and if it crashes, you'll lose all of the money you invest!

a. Imagine you decide to invest \$100 today. What is the expected value of your investment tomorrow? What does it mean for this number to be greater or less than \$100?

b. (Challenge) You devise a scheme: each day that Betcoin doesn't crash and your \$100 investment doubles in value, you sell the \$100 surplus, leaving \$100 still invested (to potentially be doubled again the next day). You repeat this process daily, selling \$100 and leaving \$100 still invested, until Betcoin crashes and you lose the invested \$100. Generally, if Betcoin crashes on day i, you'd earn \$100 for $i - 1$ days, then lose \$100 the last day, giving you a net profit of $100(i - 2)$. What is your expected profit with this scheme?

3 Independent Infants

Each child in a daycare has a 0.2 probability of having disease A, and has an independent 0.4 probability of having disease B. A child is sick if they have either disease A or disease B.

a. What is the probability that a child is sick?

b. If there are 10 children in a daycare, what is the probability that 3 or more are sick?

4 Review (Optional): Conditional Probabilities - Missing Not at Random

Preamble: We have three big tools for manipulating conditional probabilities:

- Chain Rule: $P(EF) = P(E|F)P(F)$
- Law of Total Probability: $P(E) = P(EF) + P(EF^C) = P(E|F)P(F) + P(E|F^C)P(F^C)$
- Bayes Rule: $P(E|F) = \frac{P(F|E)P(E)}{P(F)} = \frac{P(F|E)P(E)}{P(F|E)P(E)+P(F|E^C)P(E^C)}$

We're going to practice inferring which formula is best for solving a problem.

Problem: You recently tried out a new collaborative note-taking tool in class and want to know if students like it. You email all 100 people in class, asking them to reply saying if they liked it or not.

User Response	**Count**
Responded that they liked your tool	40
Responded that they didn't like your tool	45
Did not respond	15

Let L be the event that a person liked your tool. Let R be the event that a person responded. We are interested in estimating $P(L)$; however, that is hard, given that 15 people did not respond.

a. What is the probability that a user liked your tool and that they responded to the email $P(L \text{ and } R)$?

b. Which formula would you use to calculate $P(L)$? Consider that people who like your tool are in one of two (mutually exclusive) groups: those that replied, and those that did not.

c. You estimate that the probability that someone did not respond, given that they liked the tool, is $P(R^C|L) = \frac{1}{5}$. Calculate $P(L)$.