University of Washington
Department of Mathematics

Homework 3 Solutions

MATH 394: Probability I

Instructor: Arman Jahangiri

Submission: Single PDF on Gradescope
Deadline: 10:00 PM Pacific Time on the listed due date

This homework is worth **50 points**.
Across the quarter, there are **eight homework assignments**, worth **400 points total**. Homework assignments together account for **40%** of the final course grade. Further course policies can be seen in the **MATH 394 syllabus.**

Homework Policy

Submission and Deadline

Please:

- Submit homework as a single PDF on Gradescope by **10:00 PM (Pacific Time)** on the listed due date.
- Begin each problem on a new page, and ensure that pages are correctly tagged in Gradescope.
- Write solutions clearly and legibly.
- Typesetting solutions in LaTeX is encouraged, but not required.

Late Days

Each student is allotted **six late days** for the quarter. A late day extends a homework deadline by up to 24 hours without penalty. For example, submitting an assignment anytime between 10:01 PM on the due date and 10:00 PM the following day counts as one late day.

The following rules apply:

- No explanation is required to use late days.
- At most **two late days** may be used on any single homework assignment.
- Late days may not be used on the final homework assignment or on any assignment for which solutions have already been released.
- It is the student's responsibility to keep track of how many late days have been used during the quarter.

Once all late days have been exhausted, additional late submissions will incur a penalty of **10% per day**, up to a maximum deduction of **50%**. Assignments submitted more than five days late, or after solutions have been released, will not be accepted.

Submission Issues and Technical Difficulties

If a serious technical issue prevents a timely Gradescope submission, students may temporarily submit their assignment by email to the instructor at armanjg@uw.edu. In such cases:

- The submission must consist of a **single PDF file**. The email subject line must follow the format:

  `HW# Submission - FULL NAME - STUDENT ID.`

- Once the Gradescope issue has been resolved, the student must upload the *exact same file* to Gradescope as soon as possible.
- The student should clearly indicate in the Gradescope submission that the assignment was previously submitted by email before the deadline.

Email submissions are intended only for genuine technical emergencies and should not be used as a substitute for timely Gradescope submission.

Problem 1.

Source: Blitzstein and Hwang, Introduction to Probability, Chapter 2, Problem 1.

A spam filter is designed by looking at commonly occurring phrases in spam. Suppose that 80% of email is spam. In 10% of the spam emails, the phrase "free money" is used, whereas this phrase is only used in 1% of non-spam emails. A new email has just arrived, which does mention "free money." What is the probability that it is spam?

Solution

Let S be the event that the email is spam, and let F be the event that the email contains the phrase "free money." We are given

$$P(S) = 0.80, \qquad P(F \mid S) = 0.10, \qquad P(F \mid S^c) = 0.01.$$

Bayes' rule gives

$$P(S \mid F) = \frac{P(F \mid S)P(S)}{P(F \mid S)P(S) + P(F \mid S^c)P(S^c)}.$$

Substituting the given values,

$$P(S \mid F) = \frac{0.10 \cdot 0.80}{0.10 \cdot 0.80 + 0.01 \cdot 0.20} = \frac{0.08}{0.082} = \frac{40}{41}.$$

So even though the phrase is not especially common among spam emails, it is still strong evidence for spam because spam emails are much more common in the population.

Problem 2.

Source: Blitzstein and Hwang, Introduction to Probability, Chapter 2, Problem 6.

A hat contains 100 coins, where 99 are fair but one is double-headed, always landing Heads. A coin is chosen uniformly at random. The chosen coin is flipped 7 times, and it lands Heads all 7 times.

Given this information, what is the probability that the chosen coin is double-headed?

Solution

Let D be the event that the selected coin is double-headed, and let H be the event that the seven flips all land Heads. The prior probability of selecting the double-headed coin is

$$P(D) = \frac{1}{100}.$$

If the coin is double-headed, then seven Heads are certain, so

$$P(H \mid D) = 1.$$

If the selected coin is fair, then

$$P(H \mid D^c) = \left(\frac{1}{2}\right)^7 = \frac{1}{128}.$$

Bayes' rule gives

$$P(D \mid H) = \frac{P(H \mid D)P(D)}{P(H \mid D)P(D) + P(H \mid D^c)P(D^c)}.$$

Therefore

$$P(D \mid H) = \frac{1 \cdot \frac{1}{100}}{1 \cdot \frac{1}{100} + \frac{1}{128} \cdot \frac{99}{100}} = \frac{1}{1 + \frac{99}{128}} = \frac{128}{227}.$$

The seven Heads make the double-headed coin much more plausible, but the fair coins were so much more numerous at the start that the posterior probability is still not close to 1.

Problem 3.

Source: Blitzstein and Hwang, Introduction to Probability, Chapter 2, Problem 10.

Fred is working on a major project. In planning the project, two milestones are set up, with dates by which they should be accomplished. This serves as a way to track Fred's progress. Let A_1 be the event that Fred completes the first milestone on time, A_2 be the event that he completes the second milestone on time, and A_3 be the event that he completes the project on time.

Suppose that

$$P(A_{j+1} \mid A_j) = 0.8 \qquad \text{but} \qquad P(A_{j+1} \mid A_j^c) = 0.3$$

for $j = 1, 2$, since if Fred falls behind on his schedule it will be hard for him to get caught up. Also, assume that the second milestone supersedes the first, in the sense that once we know whether he is on time in completing the second milestone, it no longer matters what happened with the first milestone. We can express this by saying that A_1 and A_3 are conditionally independent given A_2, and they are also conditionally independent given A_2^c.

(a) Find the probability that Fred will finish the project on time, given that he completes the first milestone on time. Also find the probability that Fred will finish the project on time, given that he is late for the first milestone.

(b) Suppose that $P(A_1) = 0.75$. Find the probability that Fred will finish the project on time.

Solution

(a) Since the second milestone supersedes the first, once we know whether A_2 occurred, the first milestone no longer gives additional information about A_3. Thus we condition on whether Fred completes the second milestone on time.
If Fred completes the first milestone on time, then

$$P(A_2 \mid A_1) = 0.8, \qquad P(A_2^c \mid A_1) = 0.2.$$

Also,

$$P(A_3 \mid A_2) = 0.8, \qquad P(A_3 \mid A_2^c) = 0.3.$$

Therefore

$$P(A_3 \mid A_1) = 0.8 \cdot 0.8 + 0.3 \cdot 0.2 = 0.64 + 0.06 = 0.70.$$

If Fred is late for the first milestone, then

$$P(A_2 \mid A_1^c) = 0.3, \qquad P(A_2^c \mid A_1^c) = 0.7.$$

Hence

$$P(A_3 \mid A_1^c) = 0.8 \cdot 0.3 + 0.3 \cdot 0.7 = 0.24 + 0.21 = 0.45.$$

(b) Now use the law of total probability with respect to whether the first milestone is completed on time. Since $P(A_1) = 0.75$, we have $P(A_1^c) = 0.25$. Thus

$$P(A_3) = P(A_3 \mid A_1)P(A_1) + P(A_3 \mid A_1^c)P(A_1^c).$$

Using the values from part (a),

$$P(A_3) = 0.70 \cdot 0.75 + 0.45 \cdot 0.25 = 0.525 + 0.1125 = 0.6375.$$

Equivalently,

$$0.6375 = \frac{51}{80}.$$

Problem 4.

Source: Blitzstein and Hwang, Introduction to Probability, Chapter 2, Problem 20.

The Jack of Spades, Jack of Hearts, Queen of Spades, and Queen of Hearts are taken from a deck of cards. These four cards are shuffled, and then two are dealt. Literary references to cider, tarts, and winks do not need to be considered when solving this problem.

(a) Find the probability that both of these two cards are queens, given that the first card dealt is a queen.

(b) Find the probability that both are queens, given that at least one is a queen.

(c) Find the probability that both are queens, given that one is the Queen of Hearts.

Solution

(a) If the first card dealt is known to be a queen, then among the three remaining cards there is exactly one queen and two jacks. For both cards to be queens, the second card must be that one remaining queen. Therefore

$$P(\text{both queens} \mid \text{first card is a queen}) = \frac{1}{3}.$$

(b) Now condition only on the information that at least one of the two cards is a queen. The possible two-card hands from these four cards number

$$\binom{4}{2} = 6.$$

Only one of these hands contains no queen, namely the hand with the two jacks. Thus there are 5 hands with at least one queen. Among those 5 hands, only one hand contains both queens. Hence

$$P(\text{both queens} \mid \text{at least one queen}) = \frac{1}{5}.$$

(c) If we are told that one of the two dealt cards is the Queen of Hearts, then the other card is equally likely to be any one of the remaining three cards. Exactly one of those three cards is the Queen of Spades. Thus

$$P(\text{both queens} \mid \text{one card is the Queen of Hearts}) = \frac{1}{3}.$$

Problem 5.

Source: Blitzstein and Hwang, Introduction to Probability, Chapter 2, Problem 30.

A family has 3 children, creatively named A, B, and C.

(a) Discuss intuitively, but clearly, whether the event "A is older than B" is independent of the event "A is older than C."

(b) Find the probability that A is older than B, given that A is older than C.

Solution

(a) These two events are not independent. Knowing that A is older than C makes it more plausible that A is relatively old among the three children. That information also makes it more likely that A is older than B.

We can see this numerically by looking at the possible age orderings. Since all six orderings of A, B, C are equally likely, the event that A is older than both B and C has probability $1/3$, because A must be the oldest child. But

$$P(A \text{ older than } B) = \frac{1}{2}, \qquad P(A \text{ older than } C) = \frac{1}{2}.$$

If the two events were independent, the probability that both occur would be $1/4$, not $1/3$. Thus the events are not independent.

(b) Let

$$E = \{A \text{ is older than } B\}, \qquad F = \{A \text{ is older than } C\}.$$

Then

$$P(E \mid F) = \frac{P(E \cap F)}{P(F)}.$$

The event $E \cap F$ says that A is older than both B and C, so A is the oldest child. This has probability $1/3$. Also, $P(F) = 1/2$. Therefore

$$P(E \mid F) = \frac{1/3}{1/2} = \frac{2}{3}.$$

Problem 6.

Source: Blitzstein and Hwang, Introduction to Probability, Chapter 2, Problem 31.

Is it possible that an event is independent of itself? If so, when is this the case?

Solution

An event A is independent of itself exactly when

$$P(A \cap A) = P(A)P(A).$$

But $A \cap A = A$, so this condition becomes

$$P(A) = P(A)^2.$$

If we write $p = P(A)$, then this is

$$p = p^2,$$

or

$$p(1-p) = 0.$$

Thus $p = 0$ or $p = 1$. Therefore an event can be independent of itself, but only in the degenerate cases where the event is impossible or certain.

Problem 7.

Source: Blitzstein and Hwang, Introduction to Probability, Chapter 2, Problem 35.

You are going to play 2 games of chess with an opponent whom you have never played against before. Your opponent is equally likely to be a beginner, intermediate, or master. Depending on which, your chances of winning an individual game are 90%, 50%, or 30%, respectively.

(a) What is your probability of winning the first game?

(b) Congratulations: you won the first game! Given this information, what is the probability that you will also win the second game, assuming that, given the skill level of your opponent, the outcomes of the games are independent?

(c) Explain the distinction between assuming that the outcomes of the games are independent and assuming that they are conditionally independent given the opponent's skill level. Which of these assumptions seems more reasonable, and why?

Solution

(a) Before any games have been played, the opponent is equally likely to be a beginner, intermediate player, or master. Therefore the probability of winning the first game is the average of the three conditional win probabilities:

$$P(W_1) = \frac{1}{3}(0.90) + \frac{1}{3}(0.50) + \frac{1}{3}(0.30) = \frac{1.70}{3} = \frac{17}{30}.$$

(b) Winning the first game gives information about the opponent's likely strength. It makes it more likely that the opponent is a beginner and less likely that the opponent is a master.
Let the possible win probabilities be

$$0.90, \qquad 0.50, \qquad 0.30.$$

Since the three skill levels have equal prior probabilities, Bayes' rule says that after observing a first-game win, the posterior weights are proportional to these three numbers. Therefore the probability of winning the second game is

$$\frac{0.90^2 + 0.50^2 + 0.30^2}{0.90 + 0.50 + 0.30}.$$

Computing this gives

$$\frac{0.81 + 0.25 + 0.09}{1.70} = \frac{1.15}{1.70} = \frac{23}{34}.$$

(c) Assuming that the two game outcomes are independent would mean that the result of the first game gives no information about the result of the second game. That is not a very reasonable assumption here, because both games are played against the same opponent, whose skill level is initially unknown.
The more reasonable assumption is conditional independence given the opponent's skill level. Once we know whether the opponent is a beginner, intermediate player, or master, it is reasonable to model the two games as independent trials with the corresponding win probability. But before we know the opponent's skill level, the game outcomes are dependent, because winning the first game changes our belief about the opponent and therefore changes our predicted chance of winning the second game.