Math 416

Worksheet 21. Intro to Probabilistic Algorithms

There are a lot of problems in this worksheet. Make sure you spend some time on to the Monte Carlo Algorithm section of the worksheet.

Part I. Crash course in basic probability

Events and probabilities

Definition. A **discrete probability space** is a finite or countable set Ω of **outcomes** together with a function (called a **probability measure**) P: Powerset$(\Omega) \to \mathbb{R}$ satisfying the following three conditions.

(i) for every **event** $E \subseteq \Omega$, $P(E) \in [0, 1]$;
(ii) $P(\Omega) = 1$;
(iii) P is **countably additive**: whenever $A_1, A_2, \ldots \subseteq \Omega$ is a (finite or countable) list of *pairwise disjoint* events ($j \neq k$ implies $A_j \cap A_k = \emptyset$), then

$$P\left(\bigcup_{k\geq 1} A_k\right) = \sum_{k\geq 1} P(A_k).$$

Problem 1. Use the definitions to prove the following:

(1) Prove that if $A \subseteq B$ then $P(A) \leq P(B)$.
(2) (Inclusion–Exclusion Principle):

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

(3) (Continuity lemma): If $A_1 \cup A_2 \cup \cdots$ is an increasing sequence of events, then

$$P\left(\bigcup_{k\geq 1} A_k\right) = \lim_{k\to\infty} P(A_k).$$

Corollary (Union Bound)**.** The following bound applies to *any* sequence $A_1, A_2, \ldots$ of events (countably infinite or finite).

$$P\left(\bigcup_{n\geq 1} A_n\right) \leq \sum_{n\geq 1} P(A_n).$$

Conditional probabilities and independence

Definition. If $P(B) > 0$ then we define the **conditional probability of A given B** as

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

(Sometimes it is useful to rearrange this: $P(A \cap B) = P(B) \cdot P(A \mid B)$.

Problem 2. In September 40% of days are warm and the rest are cool. On warm days there is a 60% chance of rain; on cool days there is a 30% chance of rain. Let R denote the event of rainy days, C the event of cool days, and W the event of warm days.

(a) Draw a picture. (Start with a rectangle.)
(b) Explain why $P(R) = P(R \mid W)P(W) + P(R \mid C)P(C)$. Find the probability that it is raining.
(c) Find the conditional probability $P(W \mid R)$. Find $P(C \mid R)$ too, while you're at it.

Proposition (Law of Total Probability)**.** Suppose that $B_1, \ldots, B_n$ are (disjoint) events that partition Ω. Then the following identity holds for any event A.

$$P(A) = \sum_{k=1}^{n} P(A \mid B_k)P(B_k).$$

Problem 3. Draw a picture illustrating the Law of Total Probability and explain why it follows from what we've already done.

Problem 4. Suppose that three urns each contain red and blue marbles, distributed as follows.

1^{st} urn:	3 red	1 blue
2^{nd} urn:	1 red	1 blue
3^{rd} urn:	2 red	3 blue

Suppose that you choose an urn at random and then choose a marble from that urn at random. What is the probability that your marble is red? (*Hint:* The answer is not $\frac{6}{11}$.)

What if we choose Urn 2 with probability $\frac{1}{2}$ and the other urns each with probability $\frac{1}{4}$?

Problem 5. Drastically change the number of red and/or blue marbles in the various urns and see what happens to the probability of drawing a red marble. Can you make it very close to 0? very close to 1?

Definition. Events A and B are **independent** if

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

More generally, events A_i, $i \in I$, are **pairwise independent** if for any different indices $i \neq j$ the events A_i and A_j are independent. Events A_i, $i \in I$ are **mutually independent** if for any finite $J \subseteq I$ we have $P(\bigcap_{j \in J} A_j) = \prod_{i \in J} P(A_i)$.

Problem 6. Suppose that $P(A) > 0$ and $P(B) > 0$. Prove that the following are equivalent.[1]

(a) A and B are independent events.
(b) $P(A \mid B) = P(A)$.
(c) $P(B \mid A) = P(B)$.

(Notice that the second and third conditions are not obviously symmetric in A and B, while the first one is!)

Problem 7. Independence is not the same as disjointness! In fact, neither implies the other.

(a) Give an example of two independent events that are not disjoint.
(b) Give an example of two disjoint events that are not independent.

Problem 8. One implication between 'pairwise independence' and 'mutual independence' holds; which is it?

Problem 9. Roll 2 (fair, 6-sided) dice. Consider the following three events.

A = the set of outcomes where the 1^{st} roll is a 3

B = the set of outcomes where the 2^{nd} roll is a 4

C = the set of outcomes where the sum of the two rolls is a 7

Show that A, B, and C are pairwise independent (so there are three things to do!) but not mutually independent.

---

[1] This problem gives a more intuitive way to think about independence: A and B are independent iff knowing one does not change the likelihood of the other.

Extra practice

Events and probabilities

10. Two fair six-sided dice are painted: each have two sides red, two sides black, one side yellow, and one side white. If we roll both dice, what is the probability that the dice land with the same color facing up?
11. A market study determined that 60% of respondents drink coffee, 30% drink both coffee and tea, and 80% drink one or both of these beverages. What percentage drink tea?
12. Prove the following basic properties carefully, from the axioms of probability:
    (a) If A and B are events and $A \subseteq B$, then $P(A) \leq P(B)$.
    (b) For any event A, we have $P(A^c) = 1 - P(A)$. (In particular, $P(\emptyset) = 0$.)
    (c) The Union Bound for finitely many events:

$$P(A_1 \cup \cdots \cup A_n) \leq \sum_{k=1}^{n} P(A_k).$$

13. The more general version of the Inclusion–Exclusion Lemma can be proved for n events by induction. Illustrate this by deducing the version for 3 events from the version for 2.

Conditional probability and independence

14. I have three cards; each side of each card is either red or blue. The cards are RR, RB, and BB. I select a card randomly (uniformly at random) and then show one of its sides at random. Given that I see a red side, what is the probability that I'm holding the RR card? (*Hint:* It's not $\frac{1}{2}$!)
15. Verify (in the previous example) by conditioning on the probability of each card that the probability of seeing a red side is $\frac{1}{2}$.
16. I have two coins: one is a normal fair coin, and the other one always shows heads. Suppose that I choose one of the two coins at random and flip it twice. Given that I see two heads, what is the probability that I've chosen the biased coin?
17. Prove that $P(A \mid B) + P(A^c \mid B) = 1$ for all events A and B (with $P(B) > 0$).

More practice

18. In poker, a standard deck of 52 cards is used, and a hand consists of 5 cards. You are dealt a *flush* if all five cards that you receive have the same suit but are not in order. You are dealt a *straight* if all five cards are in sequential order, but at least two suits appear. Assume that an Ace can be either the highest or the lowest card.[2] Assume that you are randomly dealt a hand of 5 cards.
    (a) What is the probability that you are dealt a flush?
    (b) What is the probability that you are dealt a straight?
    (c) What is the probability that you are dealt a *straight flush*, i.e., five cards all of the same suit that are also in sequential order?
19. Prove (this special case of) Bayes' Theorem:

$$P(A \mid B) = \frac{P(A)P(B \mid A)}{P(A)P(B \mid A) + P(A^c)P(B \mid A^c)}.$$

[2]So e.g. 10♣ J♢ Q♠ K♡ A♡ and A♣ 2♣ 3♣ 4♣ 5♡ are each valid straights.

Part II. Simple Monte Carlo Algorithms

We introduce randomness to our analysis of algorithms in two ways:

1. How long do we expect a (deterministic) algorithm to run on an *average* input?[3]
2. Can we improve deterministic algorithms (especially their running time) with access to random numbers?

We will consider two types of randomized algorithms:

Monte Carlo algorithms: These always run in polynomial time, and they probably return the correct answer.

Las Vegas algorithms: These probably run in polynomial time, and always return a correct answer.

Problem 10. Explain how to turn a Las Vegas algorithm into a Monte Carlo algorithm. (NB. There is no known general way to go the other way, i.e., to turn a MC algorithm into a LV one.)

Finding a top-half element The problem is this: given an array $x[1 \cdots n]$ find k such that x_k is in the top half of the array, i.e., $x_k \geq$ median.

Problem 11. Explain how to solve this problem deterministically after only $n/2$ comparisons.

Problem 12. Suppose that we choose two elements x_i, x_j of the array independently at random.

(a) (This does not refer to x_j.) Explain why the probability that x_i is in the upper half of the array is $\geq 1/2$. Why $\geq$ and not $=$?

(b) Prove that, if we simply return $\max(x_i, x_j)$, then our probability of success is $\geq 3/4$.

A common trick for Monte Carlo algorithms is to repeat a selection or increase the number of random selections to increase the likelihood of correctness.

Here we choose k elements instead of 2 and return the max of those k.

Problem 13. Explain why this gives a probability $\geq 1 - \frac{1}{2^k}$ of success.

(E.g. for $k = 100$ this is basically 1.) So we have an algorithm that solves the problem with overwhelming probability using only 100 comparisons (i.e., in $O(1)$ time, independent of $n!$).

Verifying polynomial identities Polynomials might be given to you in many different forms. For example, one might be given as a product of monomials and the other in 'standard form', e.g.:

$$(x+1)(x-2)(x+3)(x-4)(x+5)(x-6) = x^6 - 7x^3 + 25?$$

Problem 14. Doing as little work as possible, show that the two polynomials above are not equal.

You could multiply the product in $O(d^2)$ time, where d is the degree, but let's do better. Here's the idea:

- Suppose that $\deg F(x) = \deg G(x) = d$.
- Choose $r \in \{1, 2, \ldots, 100d\}$ uniformly at random.
- If $F(r) = G(r)$ return `equal`; else return `not equal`.

Problem 15. How can the algorithm be incorrect? Can it produce a false positive? false negative?

Problem 16.

(a) Show that the probability that the algorithm returns an incorrect answer is $\leq \frac{1}{100}$.

(b) Explain how, by repeating the algorithm, we can produce an answer that is as accurate as we'd like.

[3] A more sophisticated question that follows up to 1: How confident can be be that the outcome will tend toward the expected value? (E.g. analyze quantities like $P(|X - \mu| > \varepsilon)$.