

Recitation/Discussion 4

Midterm I Review

accompaniment to MATH 170E: Introduction to Probability and Statistics 1: Probability
at the University of California, Los Angeles

Allen Lin

In this recitation, we will review material such as

- laws of probability,
- combinatorics,
- conditional probability and independent events,
- Bayes' theorem, and
- discrete variables and expected value.

Exit Survey

Please fill out the exit survey.

This is so I can get feedback and know what to focus on next time!

Laws of Probability

Theorem 1. Let S be the sample space, and let $P: \mathcal{P}(S) \rightarrow [0, 1]$ denote the probability function. Then P satisfies the following properties.

- (i) $P(S) = 1$.
- (ii) For any event $A \subseteq S$, $P(A^c) = 1 - P(A)$.
- (iii) $P(\emptyset) = 0$.
- (iv) If $A \subseteq B$, then $P(A) \leq P(B)$.
- (v) If $\{A_k\}_{k \in \mathbb{N}}$ are **pairwise disjoint**, then

$$P\left(\bigcup_{n=1}^{\infty} A_k\right) = \sum_{n=1}^{\infty} P(A_k).$$

Theorem 2 (Inclusion-exclusion principle). Let A and B be events. Then

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

This is a good way to check your answer: When you calculate probability, your final answer should be a number between 0 and 1, inclusive.

Extremely useful trick, especially if calculating $P(A)$ is hard.

Make sure you check pairwise disjoint before simply summing.

Draw a Venn diagram to see why we need to subtract $P(A \cap B)$ to correct the overcount. What happens if A and B are disjoint?

Exercise 1. Let A , B , and C be events in a sample space S , where

$$P(A) = \frac{1}{2}, \quad P(B) = \frac{3}{5}, \quad P(C) = \frac{7}{10},$$

$$P(A \cap B) = \frac{3}{10}, \quad P(A \cap C) = \frac{2}{5}, \quad P(B \cap C) = \frac{1}{2}.$$

(i) Determine $P(A \cup B)$, $P(A \cup C)$, and $P(B \cup C)$.

(ii) Prove that $P(A \cap B \cap C) \geq 1/5$.

Hint. Consider $(A \cap B) \cup (A \cap C)$. Drawing a Venn diagram might be helpful.

(iii) Suppose that $P(A \cup B \cup C) = 9/10$. Determine $P(A \cap B \cap C)$.

Hint. Apply inclusion-exclusion twice.

Solution. (i) We use inclusion-exclusion and obtain

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{4}{5},$$

$$P(A \cup C) = P(A) + P(C) - P(A \cap C) = \frac{4}{5},$$

$$P(B \cup C) = P(B) + P(C) - P(B \cap C) = \frac{4}{5}.$$

(ii) We follow the hint and use inclusion-exclusion to get

$$\begin{aligned} P((A \cap B) \cup (A \cap C)) &= P(A \cap B) + P(A \cap C) - P((A \cap B) \cap (A \cap C)) \\ &= P(A \cap B) + P(A \cap C) - P(A \cap B \cap C). \end{aligned}$$

But $(A \cap B) \cup (A \cap C) \subseteq A$, so $P((A \cap B) \cup (A \cap C)) \leq P(A) = 1/2$.
Therefore,

$$\begin{aligned} P(A \cap B \cap C) &= P(A \cap B) + P(A \cap C) - P((A \cap B) \cup (A \cap C)) \\ &\geq \frac{3}{10} + \frac{2}{5} - \frac{1}{2} = \frac{1}{5}. \end{aligned}$$

(iii) We apply inclusion-exclusion to obtain

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

so $P((A \cup B) \cap C) = 4/5 + 7/10 - 9/10 = 3/5$. Then

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

so

$$\begin{aligned} P(A \cap B \cap C) &= P(A \cap C) + P(B \cap C) - P((A \cap C) \cup (B \cap C)) \\ &= 2/5 + 1/2 - 3/5 = 3/10. \end{aligned}$$

Exercise 2. We say that $A \supseteq B$ if and only if $B \subseteq A$. Suppose that $(A_n)_{n \in \mathbb{N}}$ is a sequence of events satisfying

$$A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots$$

and suppose that $P(A_n) = (1/10)^n$ for all $n \in \mathbb{N}$.

- (i) Let $B_n = A_n \cap A_{n+1}^c$ for all $n \in \mathbb{N}$. Justify that the B_n s are pairwise disjoint and calculate $P(B_n)$ for each $n \in \mathbb{N}$.

Hint. Drawing a picture might be helpful.

- (ii) Calculate $P(A_n^c)$.

- (iii) Calculate

$$P\left(\bigcup_{n=1}^{\infty} B_n\right).$$

Hint. Use the fact that the B_n s are pairwise disjoint.

- (iv) Calculate

$$P\left(\bigcap_{n=1}^{\infty} A_n\right).$$

Hint. Note that $\bigcap_{i=1}^{\infty} A_i \subseteq A_n$ for every $n \in \mathbb{N}$. Then how can you bound the probability as a function of n ?

Solution. (i) If $m > n$, then $A_m \subseteq A_{n+1}$, so

$$B_n \cap B_m = A_n \cap (A_{n+1}^c \cap A_m) \cap A_{m+1}^c = A_n \cap \emptyset \cap A_{m+1}^c = \emptyset.$$

In particular, $A_n = A_{n+1} \cup B_n$ is a disjoint union since

$$A_{n+1} \cap B_n = A_{n+1} \cap A_n \cap A_{n+1}^c = \emptyset,$$

so $P(A_n) = P(A_{n+1}) + P(B_n)$. In particular,

$$P(B_n) = P(A_n) - P(A_{n+1}) = \frac{1}{10^n} - \frac{1}{10^{n+1}} = \frac{9}{10^{n+1}}.$$

- (ii) We have $P(A_n^c) = 1 - P(A_n) = 1 - (1/10)^n$.

- (iii) Since the B_n s are pairwise disjoint, we can just add them, so we get

$$P\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} P(B_n) = \sum_{n=1}^{\infty} \frac{9}{10^{n+1}} = \frac{1}{10}.$$

- (iv) Since $\bigcap_{i=1}^{\infty} A_i \subseteq A_n$ for every $n \in \mathbb{N}$ we have

$$P\left(\bigcap_{i=1}^{\infty} A_i\right) \leq P(A_n) = \frac{1}{10^n}$$

for every $n \in \mathbb{N}$. In particular, $1/10^n \rightarrow 0$ as $n \rightarrow \infty$, and since probability is always between 0 and 1, the only possible value is 0.

Combinatorics

Theorem 3. If there are n_1 choices for the first position, n_2 choices for the second position, and in general n_k choices for the k th position, then there are a total of $n_1 \cdot \dots \cdot n_k$ possible arrangements for k positions.

Theorem 4. There are ${}_nP_r$ permutations of n objects taken r at a time, and there are ${}_nC_r$ combinations of n objects taken r at a time.

$${}_nC_r = \binom{n}{r}.$$

Exercise 3. A password consists of 6 characters. Each character is either an uppercase English letter (26 choices) or a digit from 0 to 9. Repetition is allowed.

- (i) How many passwords are possible if there are no restrictions?
- (ii) How many passwords are possible if the first character must be a letter and the password must contain at least one digit?
- (iii) How many passwords are possible if all six characters must be distinct and the first character must be a letter?

Solution. (i) There are $26 + 10 = 36$ choices for each of the six characters, so there are 36^6 possible choices.

(ii) There are 26 choices for the first character and 36^5 for the last five characters. This is an overcount, so we subtract all the possibilities where the last five are letters, which is 26^5 . Hence there are $26(36^5 - 26^5)$ choices.

(iii) There are 26 choices for the first letter, then 35 for the second, 34 for the third, and so on, so we get $26 \cdot 35 \cdot 34 \cdot 33 \cdot 32 \cdot 31$ choices.

Exercise 4. A club has 10 members, including Alice and Bob. The club must select a committee of 4 members and then appoint a president and a secretary from among those selected. The president and secretary must be different people.

- (i) How many different committees of 4 members can be selected?
- (ii) For a given committee, how many ways can the president and secretary be appointed?
- (iii) How many possible outcomes are there if Alice and Bob cannot both be on the committee?

Solution. (i) We pick 4 from 10 where order does not matter, so there are $\binom{10}{4} = 210$ possibilities.

(ii) For a fixed committee, there are 4 ways the president can be selected and 3 for the secretary, so we get $4 \cdot 3 = 12$.

(iii) There are 12 ways to pick a president and secretary for each committee, so there are

$$\binom{10}{4} \cdot 12$$

ways. But this is an overcount because Alice and Bob cannot be on the same committee. If both are selected, then the other two members must be selected from the remaining 8, so there are

$$\binom{8}{2} = 28$$

ways to choose a committee. Again, there are 12 ways to pick a president and secretary, so the number of bad committees is

$$\binom{8}{2} \cdot 12$$

so the total is

$$\binom{10}{4} \cdot 12 - \binom{8}{2} \cdot 12.$$

Conditional Probability and Independent Events

Definition 5. The *conditional probability of an event B given A* is

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

Definition 6. Two events A and B are *independent* if $P(A \cap B) = P(A)P(B)$; otherwise, they are *dependent*.

Exercise 5. Let A and B be two events.

- (i) If A and B are mutually exclusive, are A and B always independent? If not, can they ever be independent?
- (ii) If $A \subseteq B$, are A and B always independent? If not, can they ever be independent?

Solution. (i) Suppose that A and B are mutually exclusive. By definition, $P(A \cap B) = 0$. If A and B are independent, then

$$0 = P(A \cap B) = P(A)P(B).$$

This only happens when $P(A) = 0$ or $P(B) = 0$. So yes, A and B can be independent, but not always.

- (ii) Suppose $A \subseteq B$. Then $A \cap B = A$. If A and B are independent, then

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

so $P(A)(1 - P(B)) = 0$. This happens when $P(A) = 0$ or $P(B) = 1$. So yes, A and B can be independent, but not always.

Exercise 6. A randomized algorithm performs three stages. Let A , B , and C denote the events that the algorithm passes stages 1, 2, and 3, respectively. Suppose

An application of probability in randomized algorithms!

$$P(A) = \frac{3}{4}, \quad P(B | A) = \frac{2}{3}, \quad P(C | A \cap B) = \frac{4}{5}.$$

- (i) Explain why the probability that the algorithm passes all three stages is $P(A \cap B \cap C)$.
- (ii) Determine the value of $P(B)$ so that A and B are independent.
- (iii) Compute $P(A \cap B \cap C)$.
- (iv) Determine the value of $P(C)$ so that $A \cap B$ and C are independent.
- (v) Disregard the value of $P(C)$ calculated in the last part. Suppose that

$$P(C | A \cap B^c) = \frac{1}{2}.$$

Compute $P(C | A)$ and $P(A \cap C)$.

Solution. (i) Passing all three stages means all three events occur, namely $A \cap B \cap C$.

- (ii) We have $P(A \cap B) = P(A)P(B | A) = 3/4 \cdot 2/3 = 1/2$. For A and B to be independent we must have $1/2 = P(A \cap B) = P(A)P(B) = 3P(B)/4$, so $P(B) = 2/3$.

- (iii) We have

$$P(A \cap B \cap C) = P(A)P(B | A)P(C | A \cap B) = \frac{3}{4} \cdot \frac{2}{3} \cdot \frac{4}{5} = \frac{2}{5}.$$

- (iv) We see that

$$\frac{2}{5} = P(A \cap B \cap C) = P(A \cap B)P(C) = \frac{1}{2}P(C),$$

so $P(C) = 4/5$.

- (v) We have $P(B^c | A) = 1 - P(B | A) = 1/3$. Using the law of total probability,

$$\begin{aligned} P(C | A) &= P(C \cap B | A) + P(C \cap B^c | A) \\ &= \frac{P(A \cap B \cap C)}{P(A)} + P(C | A \cap B^c) P(B^c | A) \\ &= \frac{2}{5} \cdot \frac{4}{3} + \frac{1}{2} \cdot \frac{1}{3} \\ &= \frac{7}{10}. \end{aligned}$$

Moreover, $P(A \cap C) = P(C | A)P(A) = 3/4 \cdot 7/10 = 21/40$.

Bayes' Theorem

Theorem 7 (Bayes' theorem). Let $B_1, \dots, B_m$ be events that partition the sample space S , so $S = B_1 \cup \dots \cup B_m$ and the events are pairwise disjoint ($B_i \cap B_j = \emptyset$ for all $i \neq j$). Then for all $i \in \{1, \dots, m\}$,

$$P(B_i | A) = \frac{P(B_i)P(A | B_i)}{\sum_{i=1}^m P(B_i)P(A | B_i)}.$$

Note that the denominator is just $P(A)$ by the law of total probability.

Exercise 7. You are a contestant on a game show. There are three doors. Behind one of the door sits a nice car, and behind the other two there are two goats. You choose a door. The host opens one of the other doors, which shows a goat. The host gives you the opportunity of switching to the remaining door. Should you switch or keep your original choice?

Hint. Let B be the event that your door has the car, and let A be the event that the host opens one of the other two doors and reveals a goat.

Solution. You pick a door. Let B be the event that your door has the car, so B^c is the event that your door does not have the car. Before the host does anything, your belief is

This is your *prior* belief.

$$P(B) = \frac{1}{3} \quad \text{and} \quad P(B^c) = \frac{2}{3}.$$

Now let A be the event that the host opens one of the other two doors and reveals a goat. This is the observed information. Note that

$$P(A | B) = 1 \quad \text{and} \quad P(A | B^c) = 1$$

because given B or B^c , the host always chooses one of the other doors that has a goat behind it. You want to calculate $P(B | A)$. By Bayes' theorem,

This is your *posterior* belief.

$$\begin{aligned} P(B | A) &= \frac{P(B)P(A | B)}{P(B)P(A | B) + P(B^c)P(A | B^c)} \\ &= \frac{1/3}{1/3 + 2/3} \\ &= \frac{1}{3}. \end{aligned}$$

On the other hand,

$$\begin{aligned} P(B^c | A) &= \frac{P(B^c)P(A | B^c)}{P(B)P(A | B) + P(B^c)P(A | B^c)} \\ &= \frac{2/3}{1/3 + 2/3} \\ &= \frac{2}{3}. \end{aligned}$$

So now your posterior belief is that the door you originally chose has the car with probability $1/3$ and does not have the car with probability $2/3$. Thus you should switch!

Exercise 8. A company uses a cryptographic system to protect its communications. On any given day, the company's secret key has probability $1/1000$ of having been compromised by an attacker.

A security monitoring system checks for signs of a compromised key. If the key has been compromised, the system raises an alert with probability $99/100$. If the key has not been compromised, the system nevertheless raises a false alert with probability $2/100$.

Let C be the event that the key has been compromised, and let A be the event that the monitoring system raises an alert.

- (i) Find $P(A)$.
- (ii) Given that the system raises an alert, calculate the probability that the key has actually been compromised $P(C | A)$.
- (iii) Suppose the system raises no alert. Calculate the probability that the key has nevertheless been compromised $P(C | A^c)$.
- (iv) The company wants the probability that a key is compromised given an alert to be at least $1/2$. Keeping $P(A | C) = 99/100$ and $P(C) = 1/1000$ fixed, determine the largest possible value of $P(A | C^c)$ for which $P(C | A) \geq 1/2$.

Solution. (i) By the law of total probability,

$$\begin{aligned} P(A) &= P(A | C)P(C) + P(A | C^c)P(C^c) \\ &= \frac{99}{100} \cdot \frac{1}{1000} + \frac{2}{100} \cdot \frac{999}{1000} \\ &= \frac{2097}{100000}. \end{aligned}$$

(ii) By Bayes' theorem,

$$P(C | A) = \frac{P(A | C)P(C)}{P(A)} = \frac{99/100 \cdot 1/1000}{2097/100000} = \frac{99}{2097}.$$

(iii) By Bayes' theorem,

$$P(C | A^c) = \frac{P(A^c | C)P(C)}{P(A^c)} = \frac{1/100 \cdot 1/1000}{1 - 2097/100000} = \frac{1}{97903}.$$

(iv) By Bayes' theorem,

$$\begin{aligned} P(C | A) &= \frac{P(A | C)P(C)}{P(A | C)P(C) + P(A | C^c)P(C^c)} \\ &= \frac{99/100 \cdot 1/1000}{99/100 \cdot 1/1000 + P(A | C^c) \cdot 999/1000} \geq 1/2. \end{aligned}$$

Solving for $P(A | C^c)$ we get $P(A | C^c) \leq 99/99900$.

Discrete Variables and Expected Value

Definition 8. A function $X: S \rightarrow \mathbb{R}$ is a *random variable* that assigns a real number $X(s) = x$ to each element $s \in S$.

Definition 9. An *indicator random variable* of an event A , denoted by $\mathbb{I}_A$, is a random variable that equals 1 if A occurs and 0 otherwise.

Definition 10. Let X be a discrete random variable with pmf $f(x)$, then the *expected value* of $u(X)$ is

$$\mathbb{E}[u(X)] = \sum_{x \in S} u(x)f(x).$$

In essence, this is just a weighted average, where the weights are determined by the probabilities.

Definition 11. The *variance* of a discrete random variable is

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

Theorem 12. If $\mathbb{I}_A$ is an indicator variable, then $\mathbb{E}[\mathbb{I}_A] = P(A)$.

Theorem 13 (Linearity of expectation). Let X be a discrete random variable with pmf $f(x)$, and let c and d be constants. Then

Literally the most useful theorem in all of probability theory.

- (i) $\mathbb{E}[c] = c$;
- (ii) $\mathbb{E}[cu(X)] = c\mathbb{E}[u(X)]$; and
- (iii) $\mathbb{E}[cu_1(X) + du_2(X)] = c\mathbb{E}[u_1(X)] + d\mathbb{E}[u_2(X)]$.

In fact, this extends to any finite sum using induction:

$$\mathbb{E}\left[\sum_{i=1}^n c_i u_i(X)\right] = \sum_{i=1}^n c_i \mathbb{E}[u_i(X)].$$

Exercise 9. A fair six-sided die is rolled once.

- (i) Suppose we care about the number shown on the die. Define the random variable X .
- (ii) Find $\mathbb{E}[X]$ and $\mathbb{E}[X^2]$.
- (iii) Compute $\text{Var}(X)$.
- (iv) We now care about if the outcome is prime. Define a random variable Y and calculate its expected value.
- (v) Suppose there are n independent rolls. We care about the number of rolls whose outcome is prime. Define a random variable Z and calculate the expected number of rolls that result in a prime number.

Solution. (i) $X(1) = 1, X(2) = 2, X(3) = 3, X(4) = 4, X(5) = 5, X(6) = 6$.

(ii) $\mathbb{E}[X] = 1/6 \cdot (1 + 2 + \dots + 6) = 7/2$ and $\mathbb{E}[X^2] = 1/6 \cdot (1^2 + 2^2 + \dots + 6^2) = 91/6$.

(iii) $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 91/6 - 49/4 = 35/12$.

(iv) Let $A = \{2, 3, 5\}$. Then $Y = \mathbb{I}_A$ and $\mathbb{E}[Y] = \mathbb{E}[\mathbb{I}_A] = P(A) = 1/2$.

(v) Let $Z = \mathbb{I}_{A,1} + \dots + \mathbb{I}_{A,n}$, where $\mathbb{I}_{A,i}$ is $\mathbb{I}_A$ but on the i th roll. By linearity,

$$\mathbb{E}[Z] = \mathbb{E}[\mathbb{I}_{A,1} + \dots + \mathbb{I}_{A,n}] = \mathbb{E}[\mathbb{I}_{A,1}] + \dots + \mathbb{E}[\mathbb{I}_{A,n}] = n/2.$$

Exercise 10. A permutation of $\{1, 2, \dots, n\}$ is chosen uniformly at random. Let X be the number of indices i for which the permutation maps i to itself; that is, the number of i satisfying $\pi(i) = i$.

A *permutation* is just an ordering that switches $1, 2, \dots, n$. There are $n!$ ways to order $1, 2, \dots, n$ so there are $n!$ permutations.

- (i) For each $i \in \{1, \dots, n\}$, define an indicator random variable $\mathbb{I}_i$ for the event $\pi(i) = i$. Express X as a sum of indicators.
- (ii) Compute $\mathbb{E}[X]$.
- (iii) Compute $\mathbb{E}[X^2]$.
- (iv) Use your answers to compute $\text{Var}(X)$.

Solution. (i) Define

$$\mathbb{I}_i = \begin{cases} 1 & \text{if } \pi(i) = i, \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$X = \sum_{i=1}^n \mathbb{I}_i.$$

- (ii) We have $\mathbb{E}[\mathbb{I}_i] = P(\pi(i) = i) = (n-1)!/n! = 1/n$ since exactly $(n-1)!$ of the $n!$ permutations fix i . By linearity of expectation,

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^n \mathbb{I}_i\right] = \sum_{i=1}^n \mathbb{E}[\mathbb{I}_i] = \sum_{i=1}^n \frac{1}{n} = 1.$$

- (iii) We have

$$X^2 = \left(\sum_{i=1}^n \mathbb{I}_i\right)^2 = \sum_{i=1}^n \mathbb{I}_i^2 + 2 \sum_{i=1}^n \sum_{j=i+1}^n \mathbb{I}_i \mathbb{I}_j = \sum_{i=1}^n \mathbb{I}_i + 2 \sum_{i=1}^n \sum_{j=i+1}^n \mathbb{I}_i \mathbb{I}_j.$$

We have

$$\mathbb{E}[\mathbb{I}_i \mathbb{I}_j] = P((\pi(i) = i) \cap (\pi(j) = j)) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$$

since fixing two coordinates we have $(n-2)!$ permutations. Then

$$\begin{aligned} \mathbb{E}[X^2] &= \mathbb{E}\left[\sum_{i=1}^n \mathbb{I}_i + 2 \sum_{i=1}^n \sum_{j=i+1}^n \mathbb{I}_i \mathbb{I}_j\right] \\ &= \sum_{i=1}^n \mathbb{E}[\mathbb{I}_i] + 2 \sum_{i=1}^n \sum_{j=i+1}^n \mathbb{E}[\mathbb{I}_i \mathbb{I}_j] \\ &= 1 + 2 \binom{n}{2} \frac{1}{n(n-1)} \\ &= 2 \end{aligned}$$

by linearity of expectation.

- (iv) $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 2 - 1 = 1.$

Binomial Distributions

Exercise 11. A basketball player makes each free throw with $9/10$ probability. Suppose the player attempts 20 free throws.

- (i) What is the probability that the player makes all 20 free throws?
- (ii) What is the probability that the player makes 15 free throws?
- (iii) What is the probability that the player makes at least 10 free throws?
- (iv) What is the number of free throws the player is expected to make?
- (v) What is the variance in the number of free throws the player makes?

Solution. Let's create a random variable

$$X = \text{\#successful free throws.}$$

- (i) This is $P(X = 20) = (9/10)^{20}$; each free throw has a $9/10$ chance of success, so 20 successes happen with probability $(9/10)^{20}$.
- (ii) This is $P(X = 15) = \binom{20}{15}(9/10)^{15}(1/10)^5$. For 15 free throws, there must be exactly 15 successes and 5 failures, and there are $\binom{20}{15}$ ways to arrange 15 successes in 20 attempts.
- (iii) This is $P(10 \leq X \leq 20) = \sum_{i=10}^{20} \binom{20}{i}(9/10)^i(1/10)^{20-i}$.
- (iv) When X is binomial, $\mathbb{E}[X] = np$, where n is the number of trials and p is the probability of success. Here $n = 20$ and $p = 9/10$, so $\mathbb{E}[X] = 20 \cdot 9/10 = 18$.
- (v) When X is binomial, $\text{Var}(X) = np(1 - p)$, where n is the number of trials and p is the probability of success. Here $n = 20$ and $p = 9/10$, so $\text{Var}(X) = 20 \cdot 9/10 \cdot 1/10 = 9/5$.

You should know how to prove this. Set up indicator variables for each success, and apply linearity of expectation.

Exercise 12. A factory produces items independently, each of which has probability $1/4$ of being defective. A batch contains 20 items.

- (i) Find the probability that at least 3 items are defective and at most 6 are defective.
- (ii) Given that at least one item is defective, find the probability exactly two are defective. There should be no summation $\sum$ in your answer.
- (iii) Suppose the batch is accepted if at most 2 items are defective. Find the probability that the batch is accepted.
- (iv) The factory inspects items one at a time until it finds the first defective item. Let N denote the number of items inspected, including the first defective item. Find $P(N = 5)$. Explain why N does not follow a binomial distribution. What distribution does it follow?
- (v) Now suppose the defect probability p is unknown. You observe that exactly 4 of the 20 items are defective. Find the value of p that maximizes the likelihood of this observation.

Solution. Let's create a random variable

$$X = \text{\#defective items.}$$

(i) This is $P(3 \leq X \leq 6) = \sum_{i=3}^6 \binom{20}{i} (1/4)^i (3/4)^{20-i}$.

(ii) We calculate

$$P(X = 2 \mid X \geq 1) = \frac{P(X = 2 \text{ and } X \geq 1)}{P(X \geq 1)} = \frac{P(X = 2)}{P(X \geq 1)}.$$

But $P(X = 2) = \binom{20}{2} (1/4)^2 (3/4)^{18}$ and

$$P(X \geq 1) = 1 - P(X = 0) = 1 - (3/4)^{20},$$

so the final answer is

$$P(X = 2 \mid X \geq 1) = \frac{\binom{20}{2} (1/4)^2 (3/4)^{18}}{1 - (3/4)^{20}}.$$

(iii) This is $P(X \leq 2) = \sum_{i=0}^2 \binom{20}{i} (1/4)^i (3/4)^{20-i}$.

(iv) For $N = 5$, the first four items must be acceptable and the fifth is defective. Therefore, $P(N = 5) = (3/4)^4 (1/4)$. A binomial random variable counts the number of successes in a fixed number of independent trials. Here N follows a geometric distribution where the process stops as soon as a defective item is found.

(v) Here $P(X = 4) = \binom{20}{4} p^4 (1-p)^{16}$. To find a p that maximizes this, we can take the derivative to get

$$\binom{20}{4} (4p^3(1-p)^{16} - 16p^4(1-p)^{15}) = \binom{20}{4} p^3(1-p)^{15}(4-20p),$$

which is zero when $p = 0$, $p = 1$, or $p = 1/5$. One can check that the maximum probability is attained when $p = 1/5$.

The derivative is positive for $p < 1/5$ and negative for $p > 1/5$, or you can use your favorite second derivative test.

Geometric Distributions

Exercise 13. You flip a weighted coin that lands on heads with probability $1/5$ and on tails with probability $4/5$.

- (i) What is the probability that you see your first heads at the 6th flip?
- (ii) Find the probability that you flip the coin at least 5 times before getting heads.
- (iii) Find the expected number of attempts you need to flip your coin until you get heads.
- (iv) What is the variance in the number of attempts you need to flip until you get heads?
- (v) Given that the first two flips were tails, what is the probability that the first heads is on the 5th flip?

Solution. Let's create a random variable

$X = \text{\#of coin flips until first heads.}$

(i) Here $P(X = 6) = (4/5)^5(1/5)$.

(ii) Here

$$P(X \geq 5) = \sum_{i=5}^{\infty} \left(\frac{4}{5}\right)^i \left(\frac{1}{5}\right) = \frac{1}{5} \cdot \frac{(4/5)^5}{1 - 4/5} = \frac{(4/5)^5}{5 - 4} = \left(\frac{4}{5}\right)^5.$$

(iii) When X is geometric, $\mathbb{E}[X] = 1/p$, where p is the probability of success. Here $p = 1/5$, so $\mathbb{E}[X] = 1/(1/5) = 5$.

(iv) When X is geometric, $\text{Var}(X) = (1 - p)/p^2$, where p is the probability of success. Here $p = 1/5$, so $\text{Var}(X) = (1 - 1/5)/(1/5)^2 = 20$.

(v) This is

$$\begin{aligned} P(X = 5 \mid \text{first two flips are tails}) &= \frac{P(X = 5 \text{ and first two flips are tails})}{P(\text{first two flips are tails})} \\ &= \frac{(4/5)^4(1/5)}{(4/5)^2} = \left(\frac{4}{5}\right)^2 \left(\frac{1}{5}\right). \end{aligned}$$

Negative Binomial Distributions

Exercise 14. You flip a weighted coin that lands on heads with probability $1/3$ and on tails with probability $2/3$.

- (i) Find the probability that the third head appears on the 7th flip.
- (ii) Find the probability that at least 7 flips are needed to obtain the third head.
- (iii) Find the expected number of attempts you need to flip your coin until you get three heads.
- (iv) What is the variance in the number of attempts you need to flip until you get three heads.
- (v) Given that the first two flips were tails, what is the probability that the third heads is on the 7th flip?

Solution. Let's create a random variable

$X = \text{\#of coin flips until third heads.}$

- (i) For the third head to appear on the 7th flip, the first six flips must contain exactly two heads, and the seventh flip must be a head. Thus

$$P(X = 7) = \binom{6}{2} \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right).$$

- (ii) Here the event $N \geq 7$ means that the third head has not appeared in the first six flips. Equivalently, those 6 flips contain at most 2 heads.

$$P(X = 7) = P(\text{at most 2 heads in 6 flips}) = \sum_{i=0}^2 \binom{6}{i} \left(\frac{1}{3}\right)^i \left(\frac{2}{3}\right)^{6-i}.$$

- (iii) When X is negative binomial, $\mathbb{E}[X] = r/p$, where p is the probability of success and r is the number of required successes. Here $p = 1/3$ and $r = 3$, so $\mathbb{E}[X] = 3/(1/3) = 9$.
- (iv) When X is negative binomial, $\text{Var}(X) = r(1 - p)/p^2$, where p is the probability of success and r is the number of required successes. Here $p = 1/3$ and $r = 3$, so $\text{Var}(X) = 3(1 - 1/3)/(1/3)^2 = 18$.
- (v) This is

$$P(X = 7 \mid \text{first two flips are tails}) = \frac{P(X = 7 \text{ and first two flips are tails})}{P(\text{first two flips are tails})}.$$

This quantity is

$$\frac{P(\text{tail on flips 1 and 2})P(2 \text{ heads for flips 3, 4, 5, 6})P(\text{heads on flip 7})}{(2/3)^2}$$

$$= \frac{(2/3)^2 \binom{4}{2} (1/3)^2 (2/3)^2 (1/3)}{(2/3)^2} = \frac{8}{81}.$$

Poisson Distributions

Exercise 15. A university help desk receives an average of 4 support requests per hour. Assume requests arrive according to a Poisson process, with a constant arrival rate and independent numbers of arrivals in disjoint time intervals.

- (i) Find the probability that the help desk receives exactly 6 requests in one hour.
- (ii) Find the probability that it receives at most 2 requests in one hour.
- (iii) Given that 3 requests arrived in the first hour, find the probability that exactly 2 requests arrive in the next hour.
- (iv) Find the expectation and variance of the number of requests received in a chosen hour.

Solution. Let's create a random variable

$$X = \text{\#requests received in one hour.}$$

Since the average number of requests per hour is 4, here we choose $\lambda = 4$.

- (i) Here $P(X = 6) = e^{-4} \cdot 4^6 / 6!$.
- (ii) Here $P(X \leq 2) = e^{-4} (4^0 / 0! + 4^1 / 1! + 4^2 / 2!) = 13e^{-4}$.
- (iii) Since requests in disjoint time intervals are independent, the number of requests in the next hour is independent of the number in the first hour. So this is just $P(X = 2) = e^{-4} \cdot 4^2 / 2! = 8e^{-4}$.
- (iv) When X is Poisson, $\mathbb{E}[X] = \text{Var}(X) = \lambda = 4$.

Challenge Problem

This problem is a usage of linearity of expectation in randomized algorithms and leads into hash functions. You should really only attempt this if you are bored and feel comfortable with the previous problems.

Warning: This is supposed to be hard.

Exercise 16. Suppose n balls are placed uniformly at random in n bins. Find the expected number of triple of balls that all land in the same bin.

Hint. Let $X_{ijk} = 1$ if balls i, j, k are all placed in the same bin. Compute $\mathbb{E} [\sum X_{ijk}]$.

Solution. Since each ball is thrown independently, for any $\ell \in \{1, \dots, n\}$,

$$\Pr[i, j, k \text{ are in bin } \ell] = \frac{1}{n^3}.$$

By linearity of expectation, we have

$$\begin{aligned} \mathbb{E} \left[\sum_{i=1}^n \sum_{i < j \leq n} \sum_{j < k \leq n} X_{ijk} \right] &= \sum_{i=1}^n \sum_{i < j \leq n} \sum_{j < k \leq n} \mathbb{E}[X_{ijk}] \\ &= \sum_{i=1}^n \sum_{i < j \leq n} \sum_{j < k \leq n} \Pr[i, j, k \text{ are in same bin}] \\ &= \sum_{i=1}^n \sum_{i < j \leq n} \sum_{j < k \leq n} \sum_{\ell=1}^n \Pr[i, j, k \text{ are in bin } \ell] \\ &= \sum_{i=1}^n \sum_{i < j \leq n} \sum_{j < k \leq n} \sum_{\ell=1}^n \frac{1}{n^3} \\ &= \sum_{i=1}^n \sum_{i < j \leq n} \sum_{j < k \leq n} \frac{1}{n^2} \\ &= \binom{n}{3} \frac{1}{n^2}, \end{aligned}$$

as required.

Challenge Problem

This problem is a usage of linearity of expectation in randomized algorithms! I really like it, but you should really only attempt this problem if you are bored and feel comfortable with the previous problems.

A *bloom filter* is a probabilistic data structure that is space-efficient and is used to test whether an element is in a set. The way it works is this:

- An empty bloom filter is an array of size n , and each element of the array is initially set to 0.
- Every time you add an element in your set, it goes uniformly random into one of the n elements in the array a total of k times, and those elements are turned to 1.
- Once an element of the array is set to 1 it stays at 1.
- To test whether an element is in the set, you simulate throwing a ball uniformly at random k times, and you look at where the ball hits.
 - If any of the elements at these positions were 0, then the element is not in the set since if it were, then all the bits would have been set to 1 when it was inserted.
 - If all are 1, then either the element is in the set or the elements in the array have been set to 1 by accident by other elements, resulting in a false positive.

Kind of like throwing a ball uniformly at random into n elements, but you're throwing the ball k times.

Exercise 17. Bloom filters can be used to estimate the difference between two sets. Suppose that you have sets X and Y , each with m elements and with r elements in common. Create an n -bit Bloom filter for each, with k insertions for each element.

- (i) Determine the expected number of bits where the two Bloom filters differ, as a function of m , n , k , and r .
- (ii) Assuming that the number of differing bits is tightly concentrated about its expectation, estimate r , the number of elements in common.

Warning: This is supposed to be hard.

Solution. (i) Let B_X and B_Y denote the Bloom filters corresponding to X and Y , respectively. Let Z_i be the random variable that is 1 when the i th bit of B_X and B_Y (denoted by $B_X[i]$ and $B_Y[i]$, respectively) are different and 0 otherwise. Consider the sets

$$\begin{aligned} X' &= X \setminus (X \cap Y), \\ Y' &= Y \setminus (X \cap Y); \end{aligned}$$

The event $B_X[i] = 1$ and $B_Y[i] = 0$ occurs only when some element of X' sets $B_X[i] = 1$ and no element of Y sets $B_Y[i] = 1$; if some

$x \in X \cap Y$ sets $B_X[i] = 1$, then we must also have $B_Y[i] = 1$. We now compute the probabilities separately; by independence we have

$$\Pr[B_X[i] = 1 \text{ and } B_Y[i] = 0] = \left(1 - \left(1 - \frac{1}{n}\right)^{(m-r)k}\right) \left(1 - \frac{1}{n}\right)^{mk}.$$

By symmetry, we also have

$$\Pr[B_X[i] = 0 \text{ and } B_Y[i] = 1] = \left(1 - \left(1 - \frac{1}{n}\right)^{(m-r)k}\right) \left(1 - \frac{1}{n}\right)^{mk}.$$

Therefore, the probability that $B_X[i] \neq B_Y[i]$ is

$$\mathbb{E}[Z_i] = 2 \left(1 - \left(1 - \frac{1}{n}\right)^{(m-r)k}\right) \left(1 - \frac{1}{n}\right)^{mk},$$

so by linearity of expectation the expected number of bits that differ is

$$\begin{aligned} \mathbb{E}\left[\sum_{i=1}^n Z_i\right] &= \sum_{i=1}^n \mathbb{E}[Z_i] \\ &= 2n \left(1 - \left(1 - \frac{1}{n}\right)^{(m-r)k}\right) \left(1 - \frac{1}{n}\right)^{mk}, \end{aligned}$$

as required.

- (ii) From the two sets X and Y we obtain two Bloom filters B_X and B_Y . Let d be the number of differing bits between B_X and B_Y . Assuming that d is tightly concentrated around its expectation, we say

$$d \approx 2n \left(1 - \left(1 - \frac{1}{n}\right)^{(m-r)k}\right) \left(1 - \frac{1}{n}\right)^{mk},$$

from which we can estimate

$$r \approx \frac{2mk \log\left(1 - \frac{1}{n}\right) - \log\left(\frac{2(1-1/n)^{mk}n-d}{2n}\right)}{k \log\left(1 - \frac{1}{n}\right)},$$

as required.