

# Mitzenmacher & Upfal (2nd Edition) Solutions

Leran Cai

Hayk Saribekyan

November 30, 2017



# Contents

<b>1</b>	<b>Events and Probability</b>	<b>5</b>
1.1	Comments on the main text . . . . .	5
1.2	Exercises . . . . .	5
<b>2</b>	<b>Discrete Random Variables and Expectation</b>	<b>13</b>
2.1	Comments on the main text . . . . .	13
2.2	Exercises . . . . .	13
<b>3</b>	<b>Moments and Deviations</b>	<b>25</b>
<b>4</b>	<b>Chernoff and Hoeffding Bounds</b>	<b>27</b>
<b>5</b>	<b>Balls, Bins and Random Graphs</b>	<b>29</b>
<b>6</b>	<b>The Probabilistic Method</b>	<b>31</b>
<b>7</b>	<b>Markov Chains and Random Walks</b>	<b>33</b>



# Chapter 1

## Events and Probability

### 1.1 Comments on the main text

### 1.2 Exercises

#### Exercise 1.1

Flip a fair coin ten times.

- (a)  $\mathbb{P}(\#heads = \#tails) = \frac{\binom{10}{5}}{2^{10}} = \frac{63}{256}$
- (b)  $\mathbb{P}(\text{more heads than tails}) = \left( \binom{10}{6} + \binom{10}{7} + \binom{10}{8} + \binom{10}{9} + \binom{10}{10} \right) / 2^{10}$
- (c)  $\mathbb{P}(\text{the } i\text{th flip and the } (11-i)\text{th flip are the same for } i = 1 \dots 5) = (1/2)^5$
- (d)  $\mathbb{P}(\text{at least 4 consecutive heads}) = \mathbb{P}(\text{consecutive 4}) + \mathbb{P}(\text{consecutive 5}) + \mathbb{P}(\text{consecutive 6}) + \mathbb{P}(\text{consecutive 7}) + \mathbb{P}(\text{consecutive 8}) + \mathbb{P}(\text{consecutive 9}) + \mathbb{P}(\text{consecutive 10}) = ((2 \times 2^5 + 5 \times 2^4 - 3 - 2) + (2 \times 2^4 + 4 \times 2^3) + (2 \times 2^3 + 3 \times 2^2) + (2 \times 2^2 + 2 \times 2^1) + (2 \times 2^1 + 1 \times 2^0) + 2 + 1) / 1024 = 251/1024$ .

Consecutive 5,6,7,8,9,10 are easy to understand. Consecutive 4 is special. There are 2 cases in consecutive 4 which are actually consecutive 5 (HHHHTHHHHH and HHHHHTHHHH). There are 3 cases in consecutive 4 which are duplicated (HHHHTTTHHHH, HHHHTTHHHHT, THHHHTHHHH).

Another method of doing this is to use recursion.

Let  $\mathbb{P}_4(n)$  be the probability that you have at least 4 consecutive heads after  $n$  flips. Clearly  $\mathbb{P}_4(0) = \mathbb{P}_4(1) = \mathbb{P}_4(2) = \mathbb{P}_4(3) = 0$  and  $\mathbb{P}_4(4) = 2^{-4}$ .

For more flips, either you have achieved 4 consecutive heads already or you have a string without 4 consecutive heads followed by THHHH. So for  $n > 4$ :  $\mathbb{P}_4(n) = \mathbb{P}_4(n-1) + 2^{-5}(1 - \mathbb{P}_4(n-5))$ . Because, if  $(n-1)$  already has 4 consecutive heads, the probability is just  $\mathbb{P}_4(n-1)$ , if it does not have, and we want  $n$  to have, then we know that the  $n$ th flip must be H and the last 5 flips out of  $n$  flips must be THHHH. Only in this way can we say that the first  $n-1$  does not achieve it but with the last flip it achieves.

#### Exercise 1.2

Roll two standard dice.

- (a) same number:  $1/6$
- (b) the number on the 1st dice  $>$  the number on the 2nd dice:  $15/36$ , (6 equal cases, the rest are symmetric)
- (c) the sum is even:  $1/2$
- (d) the product is a perfect square:  $(6+2)/36$ . (6 pairs of identical numbers,  $1 * 4$  and  $4 * 1$  are also two cases)

#### Exercise 1.3

Shuffle a standard deck of cards.

- (a) first two cards include at least one ace:  $1 - \frac{48}{52} \cdot \frac{47}{51}$ , use 1 minus the probability that neither of them is ace. Or we can view it as just draw two cards from a deck and they have at least one ace:  $1 - \frac{\binom{48}{2}}{\binom{52}{2}}$ .

- (b) first five cards include at least one ace:  $1 - \frac{48 \times 47 \times 46 \times 45 \times 44}{52 \times 51 \times 50 \times 49 \times 48}$   
 (c) first two cards are a pair of the same rank:  $3/51$   
 (d) first five cards are all diamonds:  $(13 \times 12 \times 11 \times 10 \times 9)/(52 \times 51 \times 50 \times 49 \times 48)$   
 (e) first five cards form a full house (three of one rank and two of another rank): in this problem, we see it as drawing 5 random cards from a deck. The sample space is  $\binom{52}{5}$ . To be a full house, we need to first pick two numbers  $\binom{13}{2}$ , then one of them has three cards and the other has two cards  $\binom{4}{3} \binom{4}{2} \times 2 = 3744$ . So the probability is  $3744/\binom{52}{5}$ .

### Exercise 1.4

Let  $E_k$  be the event that the loser has won  $k$  games, where  $0 \leq k \leq n-1$ .

If the game stops and the loser won  $k$  games, then it means there are  $n+k$  games in total. the  $k$  games won by the loser can happen at any place except the last one.

The answer for  $k$  is

$$\mathbb{P}[E_k] = \frac{\binom{n+k-1}{k}}{2^{n+k-1}}$$

We want to verify this result. Since  $k < n$ , we know that for a fixed  $n$  we need to prove,

$$\sum_{k=0}^{n-1} \mathbb{P}[E_k] = \sum_{k=0}^{n-1} \frac{\binom{n+k-1}{k}}{2^{n+k-1}} = 1$$

### Exercise 1.5

- (a) (b) (c) Just enumerate?

### Exercise 1.6

Show that the number of white balls is equally likely to be any number between 1 and  $n-1$ . Hint: Use mathematical induction.

Let  $W_n$  be the random variable representing the number of white balls when there are  $n$  balls in total.

Base case:  $n = 3$ .  $\mathbb{P}[W_3 = 1] = 1/2$  and  $\mathbb{P}[W_3 = 2] = 1/2$  because the third ball can be either white or black with equal probability.

Inductive step: Let  $X_n$  be the random variable representing the  $n$ th ball. It is  $b$  if it is black and  $w$  if white. Assuming  $\mathbb{P}[W_n = i] = \frac{1}{n-1} \forall i = 1, 2, \dots, n$ , then we can derive that

$$\begin{aligned} \mathbb{P}[W_{n+1} = i] &= \mathbb{P}[W_n = i-1] \cdot \mathbb{P}[X_n = w | W_n = i-1] \\ &\quad + \mathbb{P}[W_n = i] \cdot \mathbb{P}[X_n = b | W_n = i] \\ &= \frac{1}{n-1} \frac{i-1}{n} + \frac{1}{n-1} \frac{n-i}{n} \\ &= \frac{1}{n} \end{aligned}$$

### Exercise 1.7

- (a) [https://proofwiki.org/wiki/Inclusion-Exclusion\\_Principle](https://proofwiki.org/wiki/Inclusion-Exclusion_Principle)  
 (b)(c) Bonferroni Inequalities.

Define

$$S_k := \sum_{1 \leq i_1 < \dots < i_k \leq n} \mathbb{P}(A_{i_1} \cap \dots \cap A_{i_k})$$

**Exercise 1.8**

Let  $E_1, E_2, E_3$  be the events that the number chosen is divisible by 4, 6, 9.

$$\begin{aligned}
 \mathbb{P}[E] &= \mathbb{P}[E_1] + \mathbb{P}[E_2] + \mathbb{P}[E_3] - \mathbb{P}[E_1 \cap E_2] - \mathbb{P}[E_2 \cap E_3] - \mathbb{P}[E_1 \cap E_3] \\
 &\quad + \mathbb{P}[E_1 \cap E_2 \cap E_3] \\
 &= 250000/10^6 + 166666/10^6 + 111111/10^6 - 83333/10^6 \\
 &\quad - 55555/10^6 - 27777/10^6 + 27777/10^6 \\
 &= 38889/1000000
 \end{aligned}$$

**Exercise 1.9**

Flip a fair coin  $n$  times. For  $k > 0$ , find an upper bound on the probability that there is a sequence of  $\log_2 n + k$  consecutive heads.

Let's begin by defining the probability space and events we will analyse. First, let  $H_i$  be the event that the  $i$ th coin comes up heads. Similarly, let  $S_i$  denote the event that  $\log_2 n + k$  consecutive coin flips are heads, starting with the  $i$ th flip. We can derive an upper bound on the probability  $p$  that there is a sequence of  $\log_2 n + k$  consecutive heads using the union bound. Note that only a single run of  $\log_2 n + k$  heads within  $n$  flips is required for success. As a result, we can apply the union bound to the sequence of events  $S_i$  to obtain

$$p = \mathbb{P}\left(\bigcup_{i \in I} S_i\right) \leq \sum_{i \in I} \mathbb{P}[S_i]$$

where  $I = \{1, 2, 3, \dots, n - \log_2 n - k + 1\}$ . Note that the limits of the summation have been selected to prevent indexing events which do not exist (e.g.  $S_0, S_{n - \log_2 n - k + 2}$ ).

At this point, we need to determine  $\mathbb{P}[S_i]$ . For any given sequence starting at flip  $i$ , each coin toss will be independent of the others (i.e.,  $\{H_i\}$  are mutually independent). As a result, we can express the desired probability

$$\mathbb{P}[S_1] = \mathbb{P}\left[\bigcap_{i=1}^{\log_2 n + k} H_i\right] = \prod_{i=1}^{\log_2 n + k} \mathbb{P}[H_i] = \left(\frac{1}{2}\right)^{\log_2 n + k} = \frac{1}{2^{\log_2 n + k}}$$

Similarly,  $\mathbb{P}[S_i]$  must also be  $1/2^{\log_2 n + k}$  because we only care about the coin sequence starting from the  $i$ th coin. Substituting into the union bound.

$$p \leq \sum_{i=1}^{n - \log_2 n - k + 1} \frac{1}{2^{\log_2 n + k}} = \frac{n - \log_2 n - k + 1}{2^{\log_2 n + k}} \leq 2^{-k}$$

*Remark.* This exercise shows us that to bound a probability we do not have to think about the exact probability very carefully. We can just simplify the procedure by omitting some details. For example, we know that when  $S_1$  happens, it can also mean that  $S_2$  may happen. However, we do not care about the duplicated cases. We just need to know that the actual probability is smaller than  $\mathbb{P}[S_1] + \mathbb{P}[S_2]$ . In practice, if this sum is small, then that is already enough for us.

**Exercise 1.10**

Let  $F$  represent the event that I flip a fair coin, let  $T$  represent the event that I flip a two-headed coin. Let  $X$  be the random variable representing the result of the flip.

$$\begin{aligned}
\mathbb{P}[T|X = H] &= \frac{\mathbb{P}[T \cap (X = H)]}{\mathbb{P}[X = H]} \\
&= \frac{1/2 \cdot 1}{\mathbb{P}[X = H|F]\mathbb{P}[F] + \mathbb{P}[X = H|T]\mathbb{P}[T]} \\
&= \frac{1/2}{1/2 \cdot 1/2 + 1 \cdot 1/2} \\
&= \frac{2}{3}
\end{aligned}$$

### Exercise 1.11

- (a) Obviously, only when the bit is flipped even times can we get a correct bit.  
(b) We study the probability that after the bit passes through the relays it is flipped.

$$\frac{1 - q_1}{2} \cdot \left(1 - \frac{1 - q_2}{2}\right) + \frac{1 - q_2}{2} \cdot \left(1 - \frac{1 - q_1}{2}\right) = \frac{1 - q_1 q_2}{2}$$

- (c) The proof is by induction. For the base case, let  $n = 1$ . The probability of receiving the correct bit is given by:

$$\frac{1 + (1 - 2p)}{2} = 1 - p$$

The base case now is verified. We want to prove that when  $n = k$  is valid, then  $n = k + 1$  is also valid. Let  $E_n$  be the event that we receive a correct bit after  $n$  relays.

$$\mathbb{P}[E_{n+1}] = \mathbb{P}[E_n] \cdot (1 - p) + (1 - \mathbb{P}[E_n]) \cdot p = \frac{1 + (1 - 2p)^{n+1}}{2}$$

□

### Exercise 1.12

Monty Hall Problem.

This is a classic problem in probability and the “counter-intuitive” result can best be seen by applying Bayes’ Law. To begin our analysis let’s enumerate the sample space. Let  $O_i$  correspond to the event where Monty opens door  $i$ . In addition, let  $C_i$  be the event that the car is behind door  $i$ . Without loss of generality we can assume that the contestant always initially chooses the first door and that Monty chooses the second (since we could always permute the door labels to achieve this condition). Subject to this condition, the sample space  $\Omega$  can be enumerated simply by the position of the car as  $\Omega = \{C_1, C_2, C_3\}$ .

We now know that the condition is that Monty opens the door 2. Therefore, all we need to do is to compare the probability  $\mathbb{P}[C_1|O_2]$  and  $\mathbb{P}[C_3|O_2]$ .

$$\mathbb{P}[C_1|O_2] = \frac{\mathbb{P}[O_2|C_1] \cdot \mathbb{P}[C_1]}{\sum_{i=1}^3 \mathbb{P}[O_2|C_i] \mathbb{P}[C_i]}, \mathbb{P}[C_3|O_2] = \frac{\mathbb{P}[O_2|C_3] \cdot \mathbb{P}[C_3]}{\sum_{i=1}^3 \mathbb{P}[O_2|C_i] \mathbb{P}[C_i]}$$

To calculate these two probabilities, we need to figure out all quantities here.  $\mathbb{P}[C_i] = 1/3$ .  $\mathbb{P}[O_2|C_1] = 1/2$ ,  $\mathbb{P}[O_2|C_2] = 0$ ,  $\mathbb{P}[O_2|C_3] = 1$ . Here, since the contestant already chose the door 1, if the car is behind the door 1, then Monty would open the door 2 with probability 1/2, if the car is behind the door 2, Monty would not open it, if the car is behind the door 3, Monty would definitely open the door 2 because there is no other doors to choose.

Therefore,  $\mathbb{P}[C_1|O_2] = 1/3$  and  $\mathbb{P}[C_3|O_2] = 2/3$ . So we always change the door and we always get larger probability.

How to understand this intuitively? One idea is that we think about these two cases, the first case is that the contestant chose the right one with probability 1/3, and the wrong one with probability 2/3. Monty always opens a door behind which there is no car. If we fall in the first case, then it means changing would



give us a failure. If we fall in the second case, then it means that changing would give us a success. Since we have larger probability to fall in the second case, we like to choose to change our option.

### Exercise 1.13

We need to calculate the probability. Let the random variable  $R$  represent the result of the test. It is  $P$  if positive and  $N$  if negative. Let  $D$  be whether the patient has the disorder. It is  $T$  if true, and  $F$  if false.

$$\begin{aligned}\mathbb{P}[D = T | R = P] &= \frac{\mathbb{P}[R = P | D = T] \cdot \mathbb{P}[D = T]}{\mathbb{P}[R = P]} \\ &= \frac{\mathbb{P}[R = P | D = T] \cdot \mathbb{P}[D = T]}{\mathbb{P}[R = P | D = T] \cdot \mathbb{P}[D = T] + \mathbb{P}[R = P | D = F] \cdot \mathbb{P}[D = F]} \\ &= \frac{0.999 \cdot 0.02}{0.999 \cdot 0.02 + 0.005 \cdot 0.98} \\ &= 0.80305466237\end{aligned}$$

### Exercise 1.14

Let  $M$  represent the event that he wins 3 games and I win 1 game. Let  $E_1$  be the event that I am better,  $E_2$  equally good,  $E_3$  be the event that he is better.  $\mathbb{P}[E_i] = 1/3$ .

Then we want

$$\begin{aligned}\mathbb{P}[E_3 | M] &= \frac{\mathbb{P}[M | E_3] \mathbb{P}[E_3]}{\sum_{i=1}^3 \mathbb{P}[M | E_i] \mathbb{P}[E_i]} \\ &= \frac{0.0864}{0.0864 + 0.0625 + 0.0384} \\ &\approx 0.461\end{aligned}$$

### Exercise 1.15

We use the principle of deferred decisions.

It does not matter whether what outcomes are for the previous 9 dice. There is always one outcome of the 10th dice that can make the result divisible by 6. The probability that we can get this number is just  $1/6$ .

### Exercise 1.16

- (a) All three show the same number on the first roll:  $1/36$
- (b) Exactly two of them show the same number on the first roll:  $5/12$
- (c) When two of the three dice showing the same number on the first roll, the player can roll the different one once, if it succeeds, then the game ends, if it does not succeeds, the player can roll it one more time. The probability should be  $1 - (\frac{5}{6})^2 = 11/36$ .
- (d)

The probability that the player wins the game:

$$p = \frac{1}{36} + \frac{15}{36} \cdot \frac{11}{36} + \frac{20}{36} \cdot \left( \frac{1}{36} + \frac{15}{36} \cdot \frac{1}{6} + \frac{20}{36} \cdot \frac{1}{36} \right)$$

### Exercise 1.17

I don't think the analysis would change...

**Exercise 1.18**

(a) Choose  $x \in \{0, \dots, n-1\}$  uniformly and let  $y = z - x \pmod n$ . Then outputs  $F(z) = F(x) + F(y) \pmod m$ . The result is wrong if either of  $x$  and  $y$  is corrupted. By union bound, this is  $2/5$ .

(b) When we can do 3 times, we take the majority vote. If there is no such thing, then we return the first answer.

The algorithm is wrong when at least two outputs are wrong. Assuming the probability that one test gives us wrong answer is  $p_z$ . By (a), we know that  $p_z \leq 2/5$ . The error probability now is  $\binom{3}{2}p_z^2(1-p_z) + \binom{3}{3}p_z^3 \leq 0.352$ .

Note that the first term is actually  $2p_z^2(1-p_z)$  because if the first test is correct then the algorithm would also give us a correct answer.

**Exercise 1.19**

$$A|B < A, A|B = A, A|B > A$$

**Exercise 1.20**

If  $E_1, E_2, E_3, \dots, E_k$  are mutually independent then we know that for any subset  $I \subseteq [1, k]$ ,

$$\mathbb{P} \left[ \bigcap_{i \in I} E_i \right] = \prod_{i \in I} \mathbb{P}[E_i]$$

We now want to prove that  $\overline{E_1}, \overline{E_2}, \overline{E_3}, \dots, \overline{E_k}$  are mutually independent. So we want to prove that for any subset  $I \subseteq [1, k]$ ,

$$\mathbb{P} \left[ \bigcap_{i \in I} \bar{E}_i \right] = \prod_{i \in I} \mathbb{P}[\bar{E}_i]$$

By de morgan law, we know that  $\bigcap_{i \in I} \bar{E}_i = \overline{\bigcup_{i \in I} E_i}$ .

$$\begin{aligned} \mathbb{P} \left[ \bigcap_{i \in I} \bar{E}_i \right] &= 1 - \mathbb{P} \left[ \bigcup_{i \in I} E_i \right] \\ &= 1 - \left( \sum_{i \in I} \mathbb{P}(E_i) - \sum_{i < j \in I} \mathbb{P}(E_i \cap E_j) \right. \\ &\quad + \sum_{i < j < k \in I} \mathbb{P}(E_i \cap E_j \cap E_k) \\ &\quad \left. - \dots + (-1)^{l+1} \sum_{i_1 < i_2 < \dots < i_l \in I} \mathbb{P} \left( \bigcap_{r=1}^l E_{i_r} \right) \right) \\ &= 1 - \left( \sum_{i \in I} \mathbb{P}(E_i) - \sum_{i < j \in I} \mathbb{P}(E_i) \cdot \mathbb{P}(E_j) \right. \\ &\quad + \sum_{i < j < k \in I} \mathbb{P}(E_i) \cdot \mathbb{P}(E_j) \cdot \mathbb{P}(E_k) \\ &\quad \left. - \dots + (-1)^{l+1} \sum_{i_1 < i_2 < \dots < i_l \in I} \prod_{r=1}^l \mathbb{P}(E_{i_r}) \right) \\ &= \prod_{i \in I} (1 - \mathbb{P}[E_i]) \\ &= \prod_{i \in I} \mathbb{P}[\bar{E}_i] \end{aligned}$$

**Exercise 1.21**

Suppose that we throw a fair four-sided die (you may think of this as a square die thrown in a two-dimensional universe). We may take  $\Omega = \{1, 2, 3, 4\}$ , where each  $\omega \in \Omega$  is equally likely to occur. The events  $A = \{1, 2\}$ ,  $B = \{1, 3\}$ ,  $C = \{1, 4\}$  are pairwise independent but not independent.

**Exercise 1.22**

(a) Each possible subset has the same probability to be chosen,  $1/2^n$ . To generate this  $X$ , we know that for all elements in  $X$ , the probability that they got chosen is  $1/2^{|X|}$ . The probability that the others in  $\{1, \dots, n\}$  are not chosen is  $1/2^{(n-|X|)}$ . Then the probability that we generate this  $X$  is just  $1/2^n$ .

(b)

We generate two sets  $X$  and  $Y$ .

$$\mathbb{P}[X \subseteq Y] = \frac{\sum_{i=0}^n \binom{n}{i} \cdot 2^i}{2^n \cdot 2^n} = \frac{(2+1)^n}{4^n} = \left(\frac{3}{4}\right)^n$$

Another perspective is that, for each element, both  $X$  and  $Y$  have a probability  $1/2$  to choose it. We also know that to make  $X \subseteq Y$ , either  $Y$  chooses it or they both do not choose it. The probability is therefore  $1/2 + 1/4 = 3/4$ . There are  $n$  elements. So the probability is  $(3/4)^n$ .

Next, we look at the probability that  $X$  and  $Y$  include all the elements.

$$\mathbb{P}[X \cup Y = \{1, \dots, n\}] = (3/4)^n$$

For one element, it is in either  $X$  or  $Y$  with probability  $3/4$ . We require that all elements are either in  $X$  or in  $Y$ .

$$\mathbb{P} \left[ \bigcap_{\omega \in \{1, \dots, n\}} (\omega \in X \cup \omega \in Y) \right] = (3/4)^n$$

Another version of the answer can be found here: <http://oak.conncoll.edu/cchung/coursework/hw1.pdf>.

**Exercise 1.23**

By using the randomized min-cut algorithm, we can reduce the graph from  $n$  vertices to just 2 vertices. We take the edges between these two nodes as our min-cut set. There are at most  $\binom{n}{2} = n(n-1)/2$  such pairs of vertices. According to our analysis, some of them may be the min-cut sets we need while the others are not. Even if all of them are distinct min-cut sets, the number of min-cut sets would not exceed the number of such pairs. Hence, there are at most  $n(n-1)/2$  distinct min-cut sets.

**Exercise 1.24**

We know that to cut the graph into two connected components, we contract edges until there are only two vertices left. Hence, to cut the graph into  $r$  connected components, we contract edges until there are only  $r$  vertices left.

We use the similar way as we did in finding a 2-way cut set. Assuming the size of the  $r$ -way cut set is  $k$ . Then the average degree should be at least  $k/(r-1)$  (would be interesting to discuss, but should be easy). The number of edges should be at least  $nk/2(r-1)$ .

$$\mathbb{P}[F_{n-r}] \geq \prod_{i=1}^{n-r} \left( 1 - \frac{2(r-1)}{n-i+1} \right)$$

This should stop at some point. Another similar answer can be found at <https://hkn.eecs.berkeley.edu/~chris/CS174/mtireviewsoln.pdf>.

**Exercise 1.25**

(a) The number of edge contractions should be  $2(n-2)$ . The probability that it finds a min-cut is  $(1-p)^2$  where  $p = 2/n(n-1)$ .

(b) The number of edge contractions should be  $(n-k) + l(k-2)$ .

$$p = \frac{k(k-1)}{n(n-1)} \cdot \left(1 - \left(1 - \frac{2}{k(k-1)}\right)^l\right) \approx \frac{k^2}{n^2} (1 - e^{-\frac{2l}{k^2}})$$

(c) Find optimal values of  $k$  and  $l$  for the variation in (b) that maximize the probability of finding a minimum cut while using the same number of edge contractions as running the original algorithm twice.

We know the condition is that the number of edge contractions is  $(n-k) + l(k-2) = 2(n-2)$ . Hence,  $l = \frac{n+k-4}{k-2}$ , and the above expression becomes roughly  $\frac{k^2}{n^2} (1 - e^{-\frac{n+k}{k^3}})$ . We claim that, up to constant factors, this is maximized by taking  $k = n^{1/3}$ . To see this, note that if  $k \geq n^{1/3}$ , since  $1 - e^{-x} \leq x$  ( $x \geq 0$ ), this expression is at most  $\frac{k^2}{n^2} \frac{n+k}{k^3} = \frac{n+k}{kn^2}$ . Similarly, if  $k \leq n^{1/3}$ , then  $(1 - e^{-\frac{n+k}{k^3}})$  is at least a constant bounded away from zero, and hence, up to constant factors, the probability of success is  $\Omega(n^{2/3}/n^2) = \Omega(n^{-4/3})$ , which is significantly better than the  $O(1/n^2)$  in (a).

The solution to (c) is from here:

[theory.stanford.edu/~valiant/teaching/CS265\\_2015/ps1\\_sols.pdf](http://theory.stanford.edu/~valiant/teaching/CS265_2015/ps1_sols.pdf)

**Exercise 1.26**

Programming...

## Chapter 2

# Discrete Random Variables and Expectation

### 2.1 Comments on the main text

### 2.2 Exercises

#### Exercise 2.1

$$(k+1)/2$$

#### Exercise 2.2

The probability that typing “proof” is  $\frac{1}{26^5}$ . Let  $X_i$  be the random variable that is 1 if the word “proof” starts at the position  $i$ . Therefore, by linearity of expectation

$$\mathbb{E}[X] = \sum_{i=1}^{999996} \mathbb{E}[X_i] = \frac{999996}{26^5}$$

#### Exercise 2.3

By Jensen’s inequality, if  $f$  is convex, then we know that  $\mathbb{E}[f(X)] \geq f(\mathbb{E}[X])$ . If it is a linear function then we know that  $\mathbb{E}[f(X)] = f(\mathbb{E}[X])$ . If it is concave then  $\mathbb{E}[f(X)] \leq f(\mathbb{E}[X])$ .

#### Exercise 2.4

By Jensen’s inequality,  $\mathbb{E}[f(X)] \geq f(\mathbb{E}[X])$  for any convex functions  $f$ . If  $f$  is twice differentiable and its second derivative is non-negative, then  $f$  is convex. For  $f(x) = x^k$ , the second derivative is  $f''(x) = k(k-1)x^{k-2}$ . Since  $k$  is an even number greater than 1, we know that the second derivative is non-negative. Hence, it is convex.

#### Exercise 2.5

The probability that  $X$  is even is that

$$\begin{aligned}\mathbb{P}[X \text{ is even}] &= \sum_{i=0,2,4,\dots}^n \binom{n}{i} \left(\frac{1}{2}\right)^n \\ &= \left(\frac{1}{2}\right)^n \sum_{i=0,2,4,\dots}^n \binom{n}{i}\end{aligned}$$

We know that

$$\begin{aligned}\binom{n}{0} &= \binom{n-1}{0} \\ \binom{n}{1} &= \binom{n-1}{1} + \binom{n-1}{0} \\ \binom{n}{2} &= \binom{n-1}{2} + \binom{n-1}{1} \\ \binom{n}{3} &= \binom{n-1}{3} + \binom{n-1}{2} \\ &\dots \\ \binom{n}{n-2} &= \binom{n-1}{n-2} + \binom{n-1}{n-3} \\ \binom{n}{n-1} &= \binom{n-1}{n-1} + \binom{n-1}{n-2} \\ \binom{n}{n} &= \binom{n-1}{n-1}\end{aligned}$$

We found that when  $n$  is even

$$\begin{aligned}\sum_{i=0,2,4,\dots}^n \binom{n}{i} &= \binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \dots + \binom{n}{n} \\ &= \binom{n-1}{0} + \binom{n-1}{1} + \binom{n-1}{2} + \dots + \binom{n-1}{n-1} \\ &= \binom{n}{1} + \binom{n}{3} + \binom{n}{5} + \dots + \binom{n}{n-1}\end{aligned}$$

and when  $n$  is odd

$$\begin{aligned}\sum_{i=0,2,4,\dots}^n \binom{n}{i} &= \binom{n}{0} + \binom{n}{2} + \binom{n}{4} + \dots + \binom{n}{n-1} \\ &= \binom{n-1}{0} + \binom{n-1}{1} + \binom{n-1}{2} + \dots + \binom{n-1}{n-1} \\ &= \binom{n}{1} + \binom{n}{3} + \binom{n}{5} + \dots + \binom{n}{n}\end{aligned}$$

Hence,  $\sum_{i=0,2,4,\dots}^n \binom{n}{i} = 2^{n-1}$ .  $\mathbb{P}[X \text{ is even}] = 1/2$ .

### Exercise 2.6

Two six-sided dice.  $X = X_1 + X_2$ .

- (a)  $\mathbb{E}[X|X_1 \text{ is even}] = \frac{3}{18} + \frac{4}{18} + \frac{10}{18} + \frac{12}{18} + \frac{21}{18} + \frac{24}{18} + \frac{18}{18} + \frac{20}{18} + \frac{11}{18} + \frac{12}{18}$ .  
 (b)  $\mathbb{E}[X|X_1 = X_2] = 7$ .  
 (c)  $\mathbb{E}[X_1|X = 9] = 9/2$ .  
 (d)  $\mathbb{E}[X_1 - X_2|X = k] = 0$  for  $k \in [2, 12]$ .

### Exercise 2.7

$X$  and  $Y$  are independent geometric random variables, where  $X$  has parameter  $p$  and  $Y$  has parameter  $q$ .

(a) What is  $\mathbb{P}[X = Y]$ ?

$$\mathbb{P}[X = Y = n] = (1-p)^{n-1}p(1-q)^{n-1}q.$$

$$\mathbb{P}[X = Y] = \sum_{n=1}^{\infty} (1-p)^{n-1}p(1-q)^{n-1}q = \sum_{n=1}^{\infty} pq[(1-p)(1-q)]^{n-1}.$$

Since we know that for geometric random variables  $\mathbb{P}[X \geq i] = \sum_{n=i}^{\infty} (1-p)^{n-1}p = (1-p)^{i-1}$ . So we have,

$$\begin{aligned} S = \mathbb{P}[X = Y] &= \sum_{n=1}^{\infty} pq[(1-p)(1-q)]^{n-1} \\ \frac{(p+q-pq)S}{pq} &= \sum_{n=1}^{\infty} (1-p-q+pq)^{n-1}(p+q-pq) \\ &= (1-p-q+pq)^{1-1} = 1 \\ S &= \frac{pq}{p+q-pq} \end{aligned}$$

Hence,  $\mathbb{P}[X = Y] = \frac{pq}{p+q-pq}$ .

(b) What is  $\mathbb{E}[\max(X, Y)]$ ?

From **MU 2.9** we know that  $\mathbb{E}[\max(X, Y)] = \mathbb{E}[X] + \mathbb{E}[Y] - \mathbb{E}[\min(X, Y)]$ . From part (c) we will know that  $\min(X, Y)$  is a geometric random variable mean  $p+q-pq$ . Therefore,  $\mathbb{E}[\min(X, Y)] = \frac{1}{p+q-pq}$ , and we get

$$\mathbb{E}[\max(X, Y)] = \frac{1}{p} + \frac{1}{q} - \frac{1}{p+q-pq}$$

(c) What is  $\mathbb{P}[\min(X, Y) = k]$ ?

We split this event into two disjoint events.

$$\begin{aligned} \mathbb{P}[\min(X, Y) = k] &= \mathbb{P}[X = k, Y \geq k] + \mathbb{P}[X > k, Y = k] \\ &= \mathbb{P}[X = k] \cdot \mathbb{P}[Y \geq k] + \mathbb{P}[X > k] \cdot \mathbb{P}[Y = k] \\ &= (1-p)^{k-1}p \cdot \sum_{n=k}^{\infty} (1-q)^{n-1}q + \left( \sum_{m=k+1}^{\infty} (1-p)^{m-1}p \right) \cdot (1-q)^{k-1}q \\ &= (1-p)^{k-1}p \cdot (1-q)^{k-1} + (1-p)^k \cdot (1-q)^{k-1}q \\ &= (1-p)^{k-1}(1-q)^{k-1}(p + (1-p)q) \\ &= [(1-p-q+pq)]^{k-1}(p+q-pq) \end{aligned}$$

(d) What is  $\mathbb{E}[X|X \leq Y]$ ?

$$\mathbb{E}[X|X \leq Y] = \sum_x x \frac{\mathbb{P}[X = x \cap x \leq Y]}{\mathbb{P}[X \leq Y]}$$

We investigate our denominator,

$$\begin{aligned}
\mathbb{P}[X \leq Y] &= \sum_{x=1}^{\infty} \mathbb{P}[Y \geq x] \mathbb{P}[X = x] \\
&= \sum_{x=1}^{\infty} (1-q)^{x-1} (1-p)^{x-1} p \\
&= p \sum_{x=1}^{\infty} (1-p-q+pq)^{x-1} \\
&= \frac{p}{p+q-pq}
\end{aligned}$$

Now the whole equation is

$$\begin{aligned}
\mathbb{E}[X|X \leq Y] &= \sum_x x \frac{\mathbb{P}[X = x \cap x \leq Y]}{\mathbb{P}[X \leq Y]} \\
&= \frac{p+q-pq}{p} \cdot \sum_x x \mathbb{P}[X = x \cap x \leq Y] \\
&= \frac{p+q-pq}{p} \cdot \sum_x x \mathbb{P}[X = x] \cdot \mathbb{P}[x \leq Y] \\
&= \frac{p+q-pq}{p} \cdot \sum_x x (1-p)^{x-1} p \cdot (1-q)^{x-1} \\
&= \sum_x x (1-p-q+pq)^{x-1} (p+q-pq)
\end{aligned}$$

This is equivalent to the expectation of a geometric random variable with mean  $p+q-pq$ . Hence,  $\mathbb{E}[X|X \leq Y] = \frac{1}{p+q-pq}$ .

### Exercise 2.8

Let  $G$  be the random variable that represents the number of girls they have and let  $B$  be the number of girls they have.

$$(a) \mathbb{E}[G] = 0 \cdot \left(\frac{1}{2}\right)^k + 1 \cdot \left(\frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \dots + \frac{1}{2^k}\right) = \frac{2^k-1}{2^k}.$$

$$\mathbb{E}[B] = \sum_{n=0}^{k-1} \frac{n}{2^{n+1}} + \frac{k}{2^k} = \frac{2^k-1}{2^k}.$$

(b) We know that the expectation of the number of their children is 2 since now it is a standard geometric random variable. The expected number of girls they can have is

$$\mathbb{E}[G] = \sum_{i=1}^{\infty} \frac{1}{2^i} = 1$$

Therefore,  $\mathbb{E}[B] = 1$ .

### Exercise 2.9

(a)

$$\begin{aligned}
\mathbb{E}[\max(X_1, X_2)] &= \sum_{x_1} \sum_{x_2} \max(x_1, x_2) (1/k)(1/k) \\
&= 1/k^2 \sum_{x_1} \sum_{x_2 \leq x_1} x_1 + \sum_{x_2 > x_1} x_2
\end{aligned}$$



$$\begin{aligned}
\mathbb{E}[\min(X_1, X_2)] &= \sum_{x_1} \sum_{x_2} \min(x_1, x_2) (1/k)(1/k) \\
&= 1/k^2 \sum_{x_1} \sum_{x_2 \leq x_1} x_2 + \sum_{x_2 > x_1} x_1
\end{aligned}$$

(b)

$$\begin{aligned}
\mathbb{E}[\max(X_1, X_2)] + \mathbb{E}[\min(X_1, X_2)] &= 1/k^2 \sum_{x_1} \sum_{x_2 \leq x_1} x_1 + \sum_{x_2 > x_1} x_2 + 1/k^2 \sum_{x_1} \sum_{x_2 \leq x_1} x_2 + \sum_{x_2 > x_1} x_1 \\
&= 1/k^2 \sum_{x_1} \left( \sum_{x_2} x_2 + \sum_{x_2} x_1 \right) \\
&= 1/k^2 \sum_{x_1} \sum_{x_2} (x_2 + x_1) \\
&= \mathbb{E}[X_1] + \mathbb{E}[X_2]
\end{aligned}$$

(c)

By linearity of expectation,

$$\begin{aligned}
\mathbb{E}[\max(X_1, X_2)] + \mathbb{E}[\min(X_1, X_2)] &= \mathbb{E}[\max(X_1, X_2) + \min(X_1, X_2)] \\
&= \mathbb{E}[X_1 + X_2] \\
&= \mathbb{E}[X_1] + \mathbb{E}[X_2]
\end{aligned}$$

**Exercise 2.10**

(a)

When  $n = 1$ , there is only one  $\lambda = 1$ .

$$f(\lambda x) \leq \lambda f(x)$$

Inductive step, assuming for  $n = k$ ,  $f(\sum_{i=1}^k \lambda_i x_i) \leq \sum_{i=1}^k \lambda_i f(x_i)$ . Then we have,

$$\begin{aligned}
f\left(\sum_{i=1}^{k+1} \lambda_i x_i\right) &= f\left(\sum_{i=1}^{k-1} \lambda_i x_i + \lambda_k x_k + \lambda_{k+1} x_{k+1}\right) \\
&= f\left(\sum_{i=1}^{k-1} \lambda_i x_i + (\lambda_k + \lambda_{k+1})\left(\frac{\lambda_k x_k}{\lambda_k + \lambda_{k+1}} + \frac{\lambda_{k+1} x_{k+1}}{\lambda_k + \lambda_{k+1}}\right)\right) \\
&= f\left(\sum_{i=1}^{k-1} \lambda_i x_i + \lambda'_k x'_k\right) \\
&\leq \sum_{i=1}^k \lambda_i f(x_i) \\
&= \sum_{i=1}^{k-1} \lambda_i f(x_i) + \lambda'_k f(x'_k) \\
&= \sum_{i=1}^{k-1} \lambda_i f(x_i) + (\lambda_k + \lambda_{k+1}) f\left(\frac{\lambda_k x_k}{\lambda_k + \lambda_{k+1}} + \frac{\lambda_{k+1} x_{k+1}}{\lambda_k + \lambda_{k+1}}\right) \\
&\leq \sum_{i=1}^{k-1} \lambda_i f(x_i) + (\lambda_k + \lambda_{k+1}) \left(\frac{\lambda_k}{\lambda_k + \lambda_{k+1}} f(x_k) + \frac{\lambda_{k+1}}{\lambda_k + \lambda_{k+1}} f(x_{k+1})\right) \\
&= \sum_{i=1}^{k-1} \lambda_i f(x_i) + \lambda_k f(x_k) + \lambda_{k+1} f(x_{k+1}) \\
&= \sum_{i=1}^{k+1} \lambda_i f(x_i)
\end{aligned}$$

(b)

If  $X$  can only take finitely many values, then we can write its probability distribution as  $\sum_{i=1}^n \lambda_i = 1$  where  $\lambda_i$  is the probability that  $X$  takes the  $i$ th value in  $\text{Im}X$ .

### Exercise 2.11

Prove Lemma 2.6.

### Exercise 2.12

The expected number of cards we must draw to see all cards is

$$\mathbb{E}[X] = \sum \mathbb{E}[X_i] = n \sum_{i=1}^n \frac{1}{i}$$

If we draw  $2n$  cards, what is the expected number of cards in the deck that are not chosen at all?

The idea is to use indicator random variables. The probability that the  $i$ th card is not chosen is  $(\frac{n-1}{n})^{2n}$ .

$$\mathbb{E}[X] = \sum \mathbb{E}[X_i] = n \left(\frac{n-1}{n}\right)^{2n}$$

Chosen exactly once?

The probability that the  $i$ th card is chosen only once is  $\binom{2n}{1} \left(\frac{1}{n}\right) \left(\frac{n-1}{n}\right)^{2n-1}$ .

$$\mathbb{E}[X] = \sum \mathbb{E}[X_i] = n \binom{2n}{1} \left(\frac{1}{n}\right) \left(\frac{n-1}{n}\right)^{2n-1} = 2n \left(\frac{n-1}{n}\right)^{2n-1}$$

**Exercise 2.13**

(a)

This problem is exactly equivalent to the coupon collector's problem we did; assume that the relevant pairs of coupons form a label.

(b) No matter what  $k$  is, we can always get  $n \ln n + \Theta(n)$ .

**Exercise 2.14**

The last toss must be head, so for the previous  $n - 1$  tosses we know there are  $k - 1$  heads. Hence, the probability is  $\binom{n-1}{k-1} p^k (1-p)^{n-k}$ .

**Exercise 2.15**

Let  $X_i$  be the number of coin flips for the next head.

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}\left[\sum_{i=1}^k X_i\right] \\ &= \sum_{i=1}^k \mathbb{E}[X_i] \\ &= \sum_{i=1}^k 1/p = k/p \end{aligned}$$

**Exercise 2.16**

(a)

Let  $X_i$  be an indicator random variable which gets 1 if there is a streak of  $\log n + 1$  starting from the  $i$ th flip. Its expectation is  $(1/2)^{\log n} = 1/n$  because if a streak starts from the  $i$ th flip, the next  $\log n$  flips must be the same as the  $i$ th one. Hence the expectation of the number of streaks of length  $\log n + 1$  is  $(n - \log n)(1/n) = 1 - o(1)$ .

(b)

In other words, we want to prove that with high probability there is at least one streak with length at least  $\lfloor \log n - 2 \log \log n \rfloor$ .

Let  $k = \lfloor \log n - 2 \log \log n \rfloor$ . We break the sequence into disjoint blocks of  $k$  consecutive flips. There are  $\lfloor n/k \rfloor$  such blocks. For the sequence of  $n$  flips to not contain a streak of  $k$  flips (denote this event by  $A$ ) it is necessary that none of the blocks contains a streak of length  $k$  (denote this event by  $B$ ). Thus we have  $\mathbb{P}[A] \leq \mathbb{P}[B]$ .

The probability that a single block does not contain a streak is  $1 - (1/2)^{k-1}$ . Since the blocks are disjoint and independent, the probability that none of the blocks contains a streak is

$$\mathbb{P}[B] = \left(1 - \left(\frac{1}{2}\right)^{k-1}\right)^{\lfloor n/k \rfloor}$$

Now, since  $k - 1 \leq \log n - 2 \log \log n$  and  $\lfloor n/k \rfloor \leq n/\log n - 1$ , we get

$$\begin{aligned}
\mathbb{P}[B] &\leq \left(1 - \left(\frac{1}{2}\right)^{\log n - 2 \log \log n}\right)^{n / \log n - 1} \\
&= \left(1 - \frac{\log^2 n}{n}\right)^{n / \log n - 1} \\
&\leq \left(\exp\left(-\frac{\log^2 n}{n}\right)\right)^{n / \log n - 1} \\
&= \exp\left(-\frac{\log^2 n}{n} \left(\frac{n}{\log n} - 1\right)\right) \\
&= \exp\left(-\log n \left(1 - \frac{\log n}{n}\right)\right)
\end{aligned}$$

where the second inequality is based on the fact that  $1 - x \leq e^{-x}$ . Let  $g(n) = 1 - \log n/n$ . Since  $g(4) = 3/4 > 1/\log e$  and since  $g(n)$  is increasing for  $n \geq 4$  as the derivative  $g'(n) = n^{-2}(\log n - 1/(\ln 2)) > 0$  when  $n \geq 4$ , for  $n \geq 4$  we have that  $g(n) \geq 1/\log e$  and therefore

$$\mathbb{P}[B] \leq \exp\left(-\frac{\log n}{\log e}\right) = \exp(-\ln n) = \frac{1}{n}$$

Note that here we do not try to find out the exact probability of the event described in the problem(A). Instead, we find another event  $B$  which has higher probability than  $A$ . The probability of  $B$  is easier to bound.

The answer can be also found in <https://www.cs.helsinki.fi/u/mkhkoivi/teaching/RA-I/solutions1.pdf>

### Exercise 2.17

We see that  $\mathbb{E}[Y_0] = 1$  and  $\mathbb{E}[Y_1] = 2p$ . Next for  $i \geq 1$ , we have

$$\mathbb{E}[Y_i | Y_{i-1} = j] = 2pj$$

so that

$$\mathbb{E}[Y_i] = \mathbb{E}[\mathbb{E}[Y_i | Y_{i-1}]] = \sum_j \mathbb{P}[Y_{i-1} = j] 2pj = 2p\mathbb{E}[Y_{i-1}]$$

Thus, we have  $\mathbb{E}[Y_i] = (2p)^i$ . When  $p < 1/2$ , this probability is bounded.

### Exercise 2.18

Use Induction.

Let  $b_1, b_2, \dots, b_n$  be the values of the items observed at time  $b_t$ . We will prove this by induction. Let  $M_t$  be a random variable that takes the value of the item in memory at time  $t$ . We need to show that at time  $t$ ,  $\mathbb{P}[M_t = b_i] = 1/t$  for all  $1 \leq i \leq t$ .

The base case is when  $t = 1$ , which is trivially true since  $M_t = b_1$  with probability 1. Assume that at time  $t$ ,  $\mathbb{P}[M_t = b_i] = 1/t$  for all  $1 \leq i \leq t$ . Now we prove that this property holds for time  $t + 1$ . At time  $t + 1$ , we set  $M_{t+1} = b_{t+1}$  with probability  $1/(t + 1)$ . Therefore,  $\mathbb{P}[M_{t+1} = b_{t+1}] = 1/(t + 1)$ . For the rest,  $1 \leq i \leq t$ ,

$$\begin{aligned}
\mathbb{P}[M_{t+1} = b_i] &= \mathbb{P}[\text{no swap at time } t \text{ and } M_t = b_i] \\
&= \mathbb{P}[\text{no swap at time } t] \mathbb{P}[M_t = b_i] \\
&= \frac{t}{t+1} \frac{1}{t} \\
&= \frac{1}{t+1}
\end{aligned}$$

The answer can be found at <https://inst.eecs.berkeley.edu/~cs174/fa10/sol2.pdf>.

### Exercise 2.19

If in 2.18, when the  $k$ th item appears, it replaces the item in memory with probability  $1/2$ . Describe the distribution of the item in memory.

In this version of the algorithm, let  $X$  be random variable which is equal to  $i$  if and only if the item in memory is the  $i$ th item.

Then the probability distribution should be: for  $i = 2, \dots, n$ ,  $\mathbb{P}[X = i] = (\frac{1}{2})^{n-i+1}$ . For the first item, the probability is  $\mathbb{P}[X = 1] = (\frac{1}{2})^{n-1}$ .

### Exercise 2.20

There are  $n$  positions, each position can either be a fixed point or not.

Let  $X_i$  be an indicator showing whether the  $i$ th position is a fixed point. The permutation may put  $i$  still at  $i$  with probability  $1/n$  since it has  $n$  choices. So we have  $\sum_{i=1}^n \mathbb{E}[X_i] = n \cdot 1/n = 1$ .

The solution can also be found at <http://www.cs.nthu.edu.tw/~wkhon/random/assignment/assignment1ans.pdf>.

### Exercise 2.21

$$\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^n |a_i - i|\right] &= \sum_{i=1}^n \mathbb{E}[|a_i - i|] \\
&= \sum_{i=1}^n \sum_{j=1}^n \frac{1}{n} |j - i| \\
&= \sum_{i=1}^n \frac{1}{n} \left( \sum_{p=1}^{i-1} p + \sum_{q=1}^{n-i} q \right) \\
&= \sum_{i=1}^n \frac{2i^2 + n^2 + n - 2i - 2ni}{2n} \\
&= \sum_{i=1}^n \left( \frac{i^2 - i}{n} + \frac{n+1}{2} - i \right) \\
&= \sum_{i=1}^n \frac{i^2 - i}{n} \\
&= \frac{n^2 - 1}{3}
\end{aligned}$$

### Exercise 2.22

If  $i < j$  and  $a_i > a_j$ , let  $Y_{ij}$  be an indicator random variable. Let  $Y = \sum_{i < j} Y_{ij}$ . What we want is just the expectation of  $Y$ .

$$\begin{aligned}
\mathbb{E}[Y] &= \sum_{i < j} \mathbb{E}[X_{ij}] \\
&= \sum_{i < j} \mathbb{P}[a_i > a_j] \\
&= \sum_{i < j} \frac{1}{2} \\
&= \frac{(n-1)n}{4}
\end{aligned}$$

A similar solution can be seen at [http://www.cs.ox.ac.uk/people/varun.kanade/teaching/CS174-Fall2012/HW/HW2\\_sol.pdf](http://www.cs.ox.ac.uk/people/varun.kanade/teaching/CS174-Fall2012/HW/HW2_sol.pdf).

### Exercise 2.23

Let  $X_i$  be the number of swaps needed when the  $i$ th element is inserted.

We know that  $\mathbb{E}[X_i] = \frac{i-1}{2}$ .

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{i=1}^n \mathbb{E}[X_i] \\
&= \sum_{i=1}^n \frac{i-1}{2} \\
&= \frac{n^2 - n}{4}
\end{aligned}$$

### Exercise 2.24

We use the memoryless property. This is a very important trick. There are two cases for us to discuss:

1. if the first is not 6, then we forget it and start from the next. The expected number of flips should be  $1 + \mathbb{E}[X]$ . The corresponding probability for this case is  $5/6$ .
2. if the first is 6, then we look at the second one. If it is not 6, then we forget it and keep going. If it is 6, we count it as 2.

Let  $A_i$  means the event that  $i$ th flip is six.  $\mathbb{P}[A_i] = 1/6$ .

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{P}[\overline{A_1}] \mathbb{E}[X | \overline{A_1}] + \mathbb{P}[A_1] \mathbb{E}[X | A_1] \\
&= \mathbb{P}[\overline{A_1}] \mathbb{E}[X | \overline{A_1}] + \mathbb{P}[A_1] (\mathbb{P}[\overline{A_2} | A_1] \mathbb{E}[X | A_1, \overline{A_2}] + \mathbb{P}[A_2 | A_1] \mathbb{E}[X | A_1, A_2]) \\
&= \mathbb{P}[\overline{A_1}] \mathbb{E}[X | \overline{A_1}] + \mathbb{P}[A_1] (\mathbb{P}[\overline{A_2}] \mathbb{E}[X | A_1, \overline{A_2}] + \mathbb{P}[A_2] \mathbb{E}[X | A_1, A_2]) \\
&= \frac{5}{6} \mathbb{E}[X | \overline{A_1}] + \frac{1}{6} \left( \frac{5}{6} \mathbb{E}[X | A_1, \overline{A_2}] + \frac{1}{6} \mathbb{E}[X | A_1, A_2] \right)
\end{aligned}$$

Hence we know that,

$$\mathbb{E}[X] = \frac{5}{6}(1 + \mathbb{E}[X]) + \frac{1}{6} \left( \frac{5}{6}(2 + \mathbb{E}[X]) + \frac{1}{6} \cdot 2 \right) = \frac{35}{36} \mathbb{E}[X] + \frac{42}{36}.$$

### Exercise 2.25

### Exercise 2.26

These directed arcs will generate disjoint cycles because a node can only be an end point for two edges.

Let  $X_i^k$  be an indicator variable which is 1 if vertex  $i$  belongs to a cycle of length  $k$ . Let  $Y_k = \sum_{i=1}^n X_i^k$  be the total number of nodes belonging to a  $k$ -cycle and let  $N_k$  be the number of  $k$ -cycles in the graph. By definition we must have  $N_k = Y_k/k$  because these cycles are disjoint. Finally let  $N$  be the total number of cycles in the graph, which is  $N = \sum_{k=1}^n N_k$ .

We want to know how many permutations  $\pi$  there are such that  $X_i^k = 1$ . If vertex  $i$  belongs to a  $k$ -cycle, then the next vertex  $j = \pi(i)$  can not be vertex  $i$ , vertex  $\pi(j)$  can be neither  $i$  nor  $j$ , and so on, until the  $k$ th vertex is again  $i$ . Thus, we can choose the consecutive  $k-1$  vertices on the same cycle in  $(n-1)(n-2)\dots(n-k+1)$  ways. The remaining  $n-k$  vertices can be in any order in  $\pi$ , so the number of possibilities is  $(n-k)(n-k-1)\dots 2 \cdot 1$ . The total number of permutations such that  $X_i^k = 1$  therefore  $(n-1)!$ . As there are  $n!$  permutations in total, the probability of such event is  $\mathbb{P}[X_i^k = 1] = (n-1)!/n! = 1/n$ .

By using the linearity of expectations twice, we can get the expected number of cycles.

$$\mathbb{E}[N] = \sum_{k=1}^n \mathbb{E}[N_k] = \sum_{k=1}^n \frac{1}{k} \mathbb{E}[Y_k] = \sum_{k=1}^n \frac{1}{k} \sum_{i=1}^n \mathbb{E}[X_i^k] = \sum_{k=1}^n \frac{1}{k} \sum_{i=1}^n \frac{1}{n} = \sum_{k=1}^n \frac{1}{k} = H(n)$$

### Exercise 2.27

$$\begin{aligned} \mathbb{E}[X] &= \sum_{x=1}^{\infty} x \mathbb{P}[X = x] \\ &= \sum_{x=1}^{\infty} \frac{6}{\pi^2} \frac{1}{x} \end{aligned}$$

This sum diverges since  $\sum \frac{1}{x}$  diverges, so the expectation is not finite.

### Exercise 2.28

### Exercise 2.29

### Exercise 2.30

### Exercise 2.31

### Exercise 2.32

(a)

$E_i$ s are disjoint events, therefore  $\mathbb{P}[E] = \sum_{i=1}^n \mathbb{P}[E_i]$ . For  $i \leq m$ ,  $\mathbb{P}[E_i] = 0$ , since none of them are selected. Now, for  $i > m$  two independent events make up  $E_i$ .

$$\begin{aligned} \mathbb{P}[E_i] &= \mathbb{P}[\text{the } i\text{th candidate is the best}] \cdot \mathbb{P}[\text{the } i\text{th candidate is chosen}] \\ &= \frac{1}{n} \cdot \mathbb{P}[\text{best of the first } i-1 \text{ candidates is in the first } m \text{ candidates}] \\ &= \frac{1}{n} \frac{m}{i-1} \end{aligned}$$

Now, putting this all together, we get

$$\mathbb{P}[E] = \sum_{i=m+1}^n \mathbb{P}[E_i] = \frac{m}{n} \sum_{i=m+1}^n \frac{1}{i-1}$$

(b)

Using Lemma 2.10 from the book,

$$\mathbb{P}[E] \geq \frac{m}{n} \int_{m+1}^{n+1} \frac{1}{x-1} dx = \ln(x-1)|_{m+1}^{n+1} = \frac{m}{n} (\ln n - \ln m)$$

and

$$\mathbb{P}[E] \leq \frac{m}{n} \int_m^n \frac{1}{x-1} dx = \ln(x-1)|_m^n = \frac{m}{n} (\ln(n-1) - \ln(m-1))$$

(c)

Since the bound from above is concave, we can take the derivative to find the  $m$  so that the function is maximized.

$$\frac{d}{dm} \frac{m}{n} (\ln n - \ln m) = \frac{\ln n}{n} - \frac{\ln m}{n} + \frac{1}{n} = 0$$

Then we get  $\ln m = \ln n - 1$ , which is

$$m = e^{\ln n - 1} = e^{\ln n} e^{-1} = n e^{-1} = \frac{n}{e}$$

Substituting this  $m$  back into the bound from part (b), we get

$$\mathbb{P}[E] \geq \frac{1}{e} (\ln n - \ln \frac{n}{e}) = 1/e$$



## Chapter 3

# Moments and Deviations



## Chapter 4

# Chernoff and Hoeffding Bounds



## Chapter 5

# Balls, Bins and Random Graphs



## Chapter 6

# The Probabilistic Method





## Chapter 7

# Markov Chains and Random Walks