COL202: Discrete Math	Tutorial 6
Instructor: Venkata Koppula	Release date: 28th September, 2024

The solutions for the ($\star$) marked problems must be submitted on Gradescope by 11:59am on 4th October, 2024.

This tutorial sheet requires basic probability, conditional probability and random variables. Here is a summary of the definitions, and theorems/results that we have seen in class (Lectures 16 and 17).

- A probability distribution is defined using a sample space Ω (which is finite or countably infinite) and a function $p : \Omega \to \mathbb{R}^{\geq 0}$ such that $\sum_{x \in \Omega} p(x) = 1$. The probability of any event $\mathcal{E} \subseteq \Omega$ is $\sum_{x \in \mathcal{E}} p(x)$.
- (Union Bound) For any k events $A_1, A_2, \ldots, A_k$,
$$\Pr\left[\bigcup_{i=1}^{k} A_i\right] \leq \sum_{i=1}^{k} \Pr[A_i].$$
- (Birthday bound) Suppose we sample t numbers from $\{1, 2, \ldots, n\}$, independently and uniformly at random. Let p_{coll} denote the probability that at least two of the sampled elements are equal. Then
$$1 - e^{-t(t-1)/2n} \leq p_{\mathsf{coll}} \leq \frac{t(t-1)}{2n}$$
- Given two events A, B such that $\Pr[B] > 0$, we define the conditional probability
$$\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]}.$$
- (Law of Total Probability) Let (Ω, p) denote a probability distribution over sample space Ω. Let $(\Omega_1, \ldots, \Omega_k)$ be any partitioning of Ω. Then for any event A,
$$\Pr[A] = \sum_{i=1}^{k} \Pr[A \cap \Omega_i] = \sum_{i=1}^{k} \Pr[A \mid \Omega_i] \cdot \Pr[\Omega_i].$$
- (Random variables and Expectation) Let (Ω, p) be a probability distribution, and $X : \Omega \to \mathbb{R}$ be any random variable. Let $Z = \{X(w) : w \in \Omega\}$, and for any $z \in Z$, let $X^{-1}(z) = \{w \in \Omega \ : \ X(w) = z\}$. The expectation of this random variable is
$$\mathbb{E}[X] = \sum_{w \in \Omega} X(w) \cdot p(w) = \sum_{z \in Z} w \cdot \Pr\left[X^{-1}(z)\right].$$
- (Linearity of Expectation) For any random variables $X_1, X_2, \ldots, X_k$ over the same probability distribution (Ω, p),
$$\mathbb{E}[X_1 + X_2 + \ldots + X_k] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \ldots + \mathbb{E}[X_k].$$

1 Tutorial Submission Problem ($\star$)

The following problem has a few parts. You only need to submit the ones that are ($\star$) marked.

Let $P = (\Omega, p)$ and $Q = (\Omega, q)$ be two discrete probability distributions. We define their Statistical Difference (SD) as:

$$\mathsf{SD}(P, Q) = \frac{1}{2} \sum_{x \in \Omega} |p(x) - q(x)| \,.$$

The statistical difference is between 0 and 1 and measures how "far apart" the two distributions are. Notice that $\mathsf{SD}(P, Q) = 0$ if and only if the distributions are *identical*, i.e.

$$p(x) = q(x) \;\; \forall x \in \Omega$$

Similarly, $\mathsf{SD}(P, Q) = 1$ if and only if the distributions have *disjoint supports*, i.e.

$$\{x \in \Omega \text{ s.t. } p(x) > 0\} \;\cap\; \{x \in \Omega \text{ s.t. } q(x) > 0\} = \varnothing$$

1.1. Prove that

$$\mathsf{SD}(P, Q) = \max_{S \subseteq \Omega} \Pr_P[S] - \Pr_Q[S]$$

where $\Pr_P[S] = \sum_{x \in S} p(x)$ and $\Pr_Q[S] = \sum_{x \in S} q(x)$.

Solution: First, show that

$$\begin{aligned} \mathsf{SD}(P, Q) &= \sum_{x \in \Omega \text{ s.t. } p(x) \geq q(x)} (p(x) - q(x)) \\ &= \sum_{x \in \Omega \text{ s.t. } q(x) \geq p(x)} (q(x) - p(x)) \end{aligned}$$

Take set $S^* = \{x \in \Omega \text{ s.t. } p(x) \geq q(x)\}$. Note that for any set $S \subseteq \Omega$,

$$\Pr_P\left[S^*\right] - \Pr_Q\left[S^*\right] \geq \Pr_P[S] - \Pr_Q[S].$$

Moreover, $\mathsf{SD}(P, Q) = \Pr_P[S^*] - \Pr_Q[S^*]$.

1.2. ($\star$) In this problem we will explore the problem of distinguishing between distributions. Let $Q_0 = (\Omega, q_0)$ and $Q_1 = (\Omega, q_1)$ be two probability distributions. Consider the following game: I sample a uniformly random bit $b \leftarrow \{0, 1\}$, then sample x from Q_b. You are given x (and you also know the two distributions Q_0, Q_1). You must guess b, and you win if your guess is correct.

- Suppose $\mathsf{SD}(Q_0, Q_1) = 1$. Give an algorithm that wins the game (that is, determines the distribution) with probability 1.
- Suppose $\mathsf{SD}(Q_0, Q_1) = \delta$. Give an algorithm that wins the game (that is, determines the distribution) with probability $(1 + \delta)/2$.

- Prove that no algorithm can determine the distribution with probability greater than $(1 + \delta)/2$.

1.3. Let S_n denote the set of all permutations of $\{1, 2, \ldots, n\}$.[1] Consider the following algorithm for sampling a permutation from S_n.

Algorithm 1 Fisher Yates Shuffling

$a = [1, 2, \ldots, n]$
for $i = 1$ to $n - 1$ **do**
  $j \leftarrow$ uniformly random integer such that $i \leq j \leq n$.
  Swap $a[i]$ and $a[j]$
end for
return a

Let $\mathcal{D}$ be the distribution obtained via Algorithm 1. Let $\mathcal{U}(S_n)$ be the uniform distribution on S_n. Show that $\mathsf{SD}(\mathcal{D}, \mathcal{U}(S_n)) = 0$.

Solution: Hint: Consider any permutation σ. Compute the probability of the above algorithm outputting σ. You may want to use conditional probability to compute the probability that, after the first i iterations, the first i elements in the array are $(\sigma(1), \sigma(2), \ldots, \sigma(i))$.

2 Problems - General Probability

2.1. A fair coin is flipped n times. What's the probability that all the heads occur at the end of the sequence? (If no heads occur, then "all the heads are at the end of the sequence" is vacuously true.)

Solution: Hint: Observe that once we select the number of heads in a sequence of coin flips, the permutation of the heads gets determined, as all the heads must occur at the end.

2.2. First one digit is chosen uniformly at random from $\{1, 2, 3, 4, 5\}$ and is removed from the set; then a second digit is chosen uniformly at random from the remaining set. What is the probability that an odd digit is picked the second time?

Solution: Hint: There are 4 possible cases (EE, OE, EO, OO). Calculate the required probabilities from these.

2.3. Consider a population where the probability of being born on any given day is given by the following (non-uniform) distribution: the probability of being born on day i, $p_i = \frac{C}{i^2}$ where $i \in [365]$ and C is a constant chosen such that $\sum_{i=1}^{365} p_i = 1$. What is the

[1] A permutation σ is a bijection from $\{1, 2, \ldots, n\}$ to itself.

probability that, in a room of 23 people, at least two people share the same birthday under this distribution? How does this compare to the uniform case?

Solution: The probability that all people have different birthdays will be $C^{23}23!S$ where

$$S = \sum_{1 \le i_1 < i_2 < \cdots < i_{23} \le 365} \frac{1}{(i_1 i_2 \ldots i_{23})^2}$$

The probability of at least two people having the same birthday will be $1 - C^{23} \cdot 23! \cdot S$. For comparing, think of what would happen for concentrated distributions - should the collisions increase or decrease?

2.4. (♦) Let S_n denote the set of all permutations of $\{1, 2, \ldots, n\}$. Given a permutation σ, let $\mathsf{inv}(\sigma)$ denote the number of pairs of indices (i, j) such that $i < j$ and $\sigma(i) > \sigma(j)$. Find the expected value of $\mathsf{inv}(\sigma)$, when σ sampled from the following distributions:

1. (Easy) Uniform distribution over S_n
2. A distribution $\mathcal{D}$ defined as follows: let $T_1 = \{1, 2, \ldots, n\}$ and $s_1 = \sum_{x \in T_1} x$. For $i = 1$ to n do the following:
   (a) for each $x \in T_i$, let $p_i(x) = x/s_i$. Sample an element y from (T_i, p_i).
   (b) Set $\sigma(i) = y$.
   (c) Let $T_{i+1} = T_i \setminus \{y\}$ and $s_{i+1} = s_i - y$.

Hint: in each of these two cases, decompose the random variable appropriately.

Solution:

1. Let $\Omega = S_n$ and $\mathsf{inv}(\sigma)$ be a random variable that denotes the number of inversions in σ. We will decompose this random variable into a sum of simpler random variables. There can be multiple ways to represent $\mathsf{inv}(\sigma)$ as a sum of random variables. For instance, for each $i \in [n]$, we can define $Y_i(\sigma)$ to be the number of indices $j > i$ such that $\sigma(i) > \sigma(j)$. Clearly, $\mathsf{inv}(\sigma) = \sum_i Y_i(\sigma)$, and therefore $\mathbb{E}[\mathsf{inv}(\sigma)] = \sum_{i=1}^{n} \mathbb{E}[Y_i(\sigma)]$. However, computing $\mathbb{E}[Y_i]$ can be difficult.

   Let us consider the following decomposition. For every (i, j) pair where $i < j$, let X_{ij} be a random variable defined as follows:

   $$X_{ij}(\sigma) = \begin{cases} 1 & \text{if } \sigma(i) > \sigma(j) \\ 0 & \text{otherwise} \end{cases}$$

   Then $\mathsf{inv}(\sigma) = \sum_{1 \le i < j \le n} X_{ij}(\sigma)$ so that

   $$\mathbb{E}[\mathsf{inv}(\sigma)] = \sum_{1 \le i < j \le n} \mathbb{E}[X_{ij}(\sigma)] = \sum_{1 \le i < j \le n} \Pr[\sigma(i) > \sigma(j)]$$

Observation 1. *For any fixed $i < j$,*

$$\Pr[\sigma(i) > \sigma(j)] = \Pr[\sigma(i) < \sigma(j)] = \frac{1}{2}.$$

Proof. This 'obvious-looking' observation requires a proof. Since we are working with a uniform distribution, it suffices to just look at the set Γ of all permutations where $\sigma(i) > \sigma(j)$, and $\Omega \setminus \Gamma$. Observe that $|\Gamma| = |\Omega \setminus \Gamma|$, and therefore the two probabilities are equal.

For distributions that are not uniform, such a counting-based argument will not work. In such cases, you will need to use the law of total probabilities, and partition the sample space appropriately. $\square$

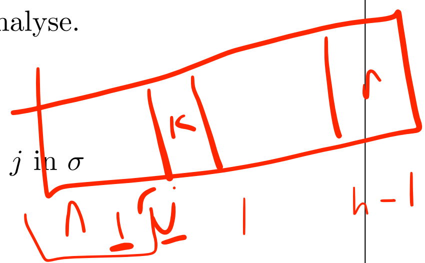
Thus, $\mathbb{E}[\mathsf{inv}(\sigma)] = \frac{n(n-1)}{4}$.

For this problem, we could have also considered the following decomposition: for every (i, j) pair where $i < j$, let X_{ij} be a random variable defined as follows:

$$X_{ij}(\sigma) = \begin{cases} 1 & \text{if } i \text{ appears after } j \text{ in } \sigma \\ 0 & \text{otherwise} \end{cases}$$

Convince yourself that this decomposition also works. For the easier version, both these approaches are equally easy to analyse. For the harder version of this problem, one of the approaches is easier to analyse.

2. Suppose we use the random variable as



$$X_{ij}(\sigma) = \begin{cases} 1 & \text{if } i \text{ occurs after } j \text{ in } \sigma \\ 0 & \text{Otherwise} \end{cases}$$

Let $\mathcal{E}_{ij}(\sigma)$ be the event that i occurs after j in σ. Then $\mathsf{inv}(\sigma) = \sum_{1 \le i < j \le n} X_{ij}(\sigma)$ so that

$$\mathbb{E}[\mathsf{inv}(\sigma)] = \sum_{1 \le i < j \le n} \mathbb{E}[X_{ij}(\sigma)] = \sum_{1 \le i < j \le n} \Pr[\mathcal{E}_{ij}]$$

Claim 1. *Fix i and j such that $i < j$. Then* $\boxed{\Pr\left[\mathcal{E}_{ij}\right] = \frac{j}{i+j}}$.

Proof. Let $k \in \{1, 2, \ldots, n-1\}$ and $S \subseteq \{1, 2 \ldots n\} \setminus \{i, j\}, |S| = k - 1$. Consider the following sets:

- $\mathcal{F}_{k,S}$: the set of permutations where the first $k - 1$ elements of the permutation are from the set S.
- $\mathcal{G}_{k,S}$: the set of permutations where i or j appear at the k^{th} position.
- $\Omega_{k,S} = \mathcal{F}_{k,S} \cap \mathcal{G}_{k,S}$

Observe that $\Omega_{k,S}$ forms a partition of Ω. Then by the law of total probability, we have

$$\Pr[\mathcal{E}_{ij}] = \sum_{k,s} \Pr[\mathcal{E}_{ij} | \Omega_{k,S}] \Pr[\Omega_{k,S}]$$

p(j/i)

The k^{th} element of $\Omega_{k,S}$ can be either i or j and $\Pr[\mathcal{E}_{ij}|\Omega_{k,S}]$ is same as the probability that it is j. Let $A = \sum_{t\in\{1,2\dots n\}\setminus S} t$. Then

$$\boxed{\Pr[\mathcal{E}_{ij}|\Omega_{k,S}]} = \frac{\boxed{\Pr[\mathcal{F}_{k,S}]}\frac{j}{A}}{\Pr[\mathcal{F}_{k,S}]\frac{i}{A} + \Pr[\mathcal{F}_{k,S}]\frac{j}{A}} = \frac{j}{i+j}$$

Thus

$$\Pr[\mathcal{E}_{ij}] = \frac{j}{i+j}\sum_{k,S}\Pr[\Omega_{k,S}] = \frac{j}{i+j}$$

□

Finally

$$\mathbb{E}[\mathsf{inv}(\sigma)] = \sum_{1\le i<j\le n}\frac{j}{i+j}$$

2.5. (♦) There are n seats in an aeroplane, and n passengers each having a unique ticket to a seat. Passengers come one at a time to occupy seats. The first passenger loses his/her ticket and thereby chooses a seat uniformly at random. For each successive passenger, if the seat corresponding to his/her ticket is vacant, he/she occupies that seat, else he/she chooses a vacant seat uniformly at random. What is the probability that the n^{th} passenger occupies the seat corresponding to his/her ticket?

Solution: Let P_n denote the probability that in the above situation with n seats, the n^{th} passenger occupies the correct seat. Then we have the following cases:

1. The first passenger occupies the correct seat. In this all other passengers also occupy the correct seat.
2. The first passenger occupies the i^{th} seat for $2 \le i \le n-1$. Thus the passengers having seats before seat i will occupy their correct seat. Essentially, the problem size now reduces to $n - i + 1$.

From the above cases we obtain the following recursion for $n \ge 2$:

$$P_n = \frac{1}{n} + \frac{1}{n}\sum_{i=2}^{n-1} P_i$$

which reduces to

$$nP_n = 1 + \sum_{i=2}^{n-1} P_i$$

$$\implies nP_n - (n-1)P_{n-1} = P_{n-1} \implies P_n = P_{n-1}$$

Thus $P_n = P_{n-1} = \cdots = P_2 = \frac{1}{2}$. So the probability that the n^{th} passenger gets the correct seat is $\frac{1}{2}$.

One can also prove this using strong induction.

2.6. We play a game with a deck of 52 regular playing cards, of which 26 are red and 26 are black. I randomly shuffle the cards and place the deck face down on a table. You have the option of "taking" or "skipping" the top card. If you skip the top card, then that card is revealed and we continue playing with the remaining deck. If you take the top card, then the game ends; you win if the card you took was revealed to be black, and you lose if it was red. If we get to a point where there is only one card left in the deck, you must take it. Prove that you have no better strategy than to take the top card—which means your probability of winning is 1/2.

Solution: First, let us consider deterministic strategies. Let $\mathcal{S}_i$ be a strategy in which we pick the ith card and A_i be the event that the i^{th} card is red in a random shuffle of cards. The number of permutations of the cards in which the i^{th} card is red is $\binom{26}{1} \times 51!$. Then probability of winning given strategy $\mathcal{S}_i$ is

$$\Pr[\text{Win}|\mathcal{S}_i] = \Pr[A_i] = \frac{\binom{26}{1} \times 51!}{52!} = \frac{26}{52} = \frac{1}{2}$$

Observe that the winning probability is independent of i. Hence, even if we use a randomized strategy (in which we choose strategy $\mathcal{S}_i$ with probability p_i) the winning probability remains the same: let $\mathcal{S} = (p_1, p_2 \ldots p_{56})$ be the randomized strategy, where p_i denotes the probability of choosing strategy $\mathcal{S}_i$. Then

$$\Pr[\text{Win}|\mathcal{S}] = \sum_{i=1}^{56} \Pr[\text{Win}|\mathcal{S}_i] \Pr[\mathcal{S}_i] = \sum_{i=1}^{56} \frac{1}{2} p_i = \frac{1}{2}$$

Thus, no strategy works better than picking the first card.

2.7. (♦) Consider a game played between Alice, Bob and the Referee where the Refree asks questions x to Alice and y to Bob. Alice and Bob reply with their respective answers a, b. Here $x, y, a, b \in \{0, 1\}$. Note that the questions may be sampled from any distribution by the Referee.

They win the game if

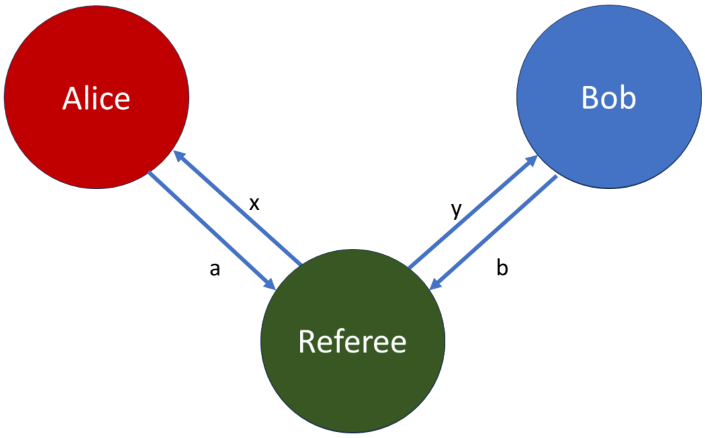
$$x \wedge y = a \oplus b$$

Alice and Bob are allowed to agree on a strategy before the game begins, but once it starts, they are *spatially separated* and cannot communicate.

Determine an optimal strategy for Alice and Bob to win the game. Provide a proof showing that this strategy is indeed optimal.

Solution: Below we discuss the case where the referee picks x and y uniformly at random. The general distribution case should be similar. The following table summarizes all the cases:

x	y	Win	ab
0	0	0	00 or 11
0	1	0	00 or 11
1	0	0	00 or 11
1	1	1	01 or 10



On analyzing the cases, we find that an optimal deterministic classical strategy is to always output 00 for a and b, which wins with a probability 3/4 (if the referee uses a uniform distribution over the inputs).

We claim that even with a probabilistic strategy, the best winning probability is 3/4. Let a_{ij} denote the probability that Alice outputs j if she is asked query i, i.e.

$$a_{ij} = \Pr[a = j | x = i]$$

$$b_{ij} = \Pr[b = j | y = i]$$

The probability of winning is:

$$\begin{aligned}\Pr[\mathrm{Win}] = &(a_{00}b_{00} + a_{01}b_{01})\Pr[x = 0, y = 0]\\ &+ (a_{00}b_{10} + a_{01}b_{11})\Pr[x = 0, y = 1]\\ &+ (a_{10}b_{00} + a_{11}b_{01})\Pr[x = 1, y = 0]\\ &+ (a_{10}b_{11} + a_{11}b_{10})\Pr[x = 1, y = 1]\end{aligned}$$

We also have $a_{00} + a_{01} = 1$. Consider the first term

$$a_{00}b_{00} + a_{01}b_{01} = a_{00}b_{00} + (1 - a_{00})(1 - b_{00}) = 1 - a_{00} - b_{00} + 2a_{00}b_{00} \leq \frac{3}{4}$$

Where the last inequality follows from the value of $\max_{x,y\in[0,1]}(1 - x - y + 2xy)$. So,

$$\Pr[\mathrm{Win}] \leq \frac{3}{4}(\Pr[x = 0, y = 0]+\Pr[x = 0, y = 1]+\Pr[x = 1, y = 0]+\Pr[x = 1, y = 1])$$

Hence,

$$\Pr[\mathrm{Win}] \leq \frac{3}{4}$$

This is version 1.0 of the tutorial sheet. Let me know if something is unclear. In case of any doubt or for help regarding writing proofs, feel free to contact me or TAs.

Venkata Koppula - kvenkata@iitd.ac.in
Ananya Mathur - cs5200416@iitd.ac.in
Anish Banerjee - cs1210134@cse.iitd.ac.in
Eshan Jain - cs5200424@cse.iitd.ac.in
Mihir Kaskhedikar - cs1210551@iitd.ac.in
Naman Nirwan - Naman.Nirwan.cs521@cse.iitd.ac.in
Pravar Kataria - Pravar.Kataria.cs121@cse.iitd.ac.in
Shashwat Agrawal - csz248012@cse.iitd.ac.in
Subhankar Jana - csz248009@iitd.ac.in