Problem Set 8

> **Due dates:** Electronic submission of this homework is due on **Friday 3/31/2023 before 11:59pm** on canvas.

Name: Ayushri Jain

Resources.

Cormen, Leiserson, Rivest, Stein: Introduction to Algorithms, 3rd edition, The MIT Press, 2009 (or 4th edition)

On my honor, as an Aggie, I have neither given nor received any unauthorized aid on any portion of the academic work included in this assignment. Furthermore, I have disclosed all resources (people, books, web sites, etc.) that have been used to prepare this homework.

Signature: Ajain

Problem 1 (20 points).

Solution. (a) We need to show that $\mathcal{F}$ is a σ-algebra, so it is enough to show that it follows the three properties discussed in lecture.
(i) The problem statement says that $\mathcal{F}$ is a family of subsets of Ω and $\mathcal{F}$ contains all finite sets. Thus $\mathcal{F}$ contains the empty set also. We can write Ω as the countable union of finite sets: $\Omega = \bigcup_{n=1}^{\infty} x_1, x_2, \ldots, x_n$, where $x_i \in \Omega$ for all i. Since $\mathcal{F}$ is closed under countable unions, Ω is in $\mathcal{F}$ ($\Omega \in \mathcal{F}$).
(ii) The problem statement says that $\mathcal{F}$ is closed under complements which means that if a set A is in $\mathcal{F}$, then its complement A^c is also in $\mathcal{F}$, i.e. if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$.
(iii) The problem statement says that $\mathcal{F}$ is closed under countable unions which means that if there are sets $E_1, E_2, \cdots$ present in $\mathcal{F}$, then their union $\bigcup_{k=1}^{\infty} E_k$ is also present in $\mathcal{F}$. So, $\mathcal{F}$ contains a countable union of sets also, i.e. $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$.
Since all three properties are satisfied, the given $\mathcal{F}$ is a σ-algebra.
(b) The power set of Ω will contain all the subsets of Ω. Any subset of countable set is countable. Subset of uncountable set can be countable or uncountable. When the sample space Ω is finite or countably finite, the power set $P(\Omega)$ will contain uncountable sets and countable sets both. But family $\mathcal{F}$ only contains the countable or finite subsets of sample space Ω. e.g. power set of natural numbers is uncountable but the sample space is finite or countably finite.
When the sample space Ω is uncountable, the power set $P(\Omega)$ will contain uncountable sets and their uncountable complements, countable sets and their countable complements. But the family $\mathcal{F}$ contains all finite sets according to the problem statement, so it contains subsets of sample space which are countable sets or their countable subsets. Hence $\mathcal{F}$ is σ-algebra but does not contain the uncountable sets and their uncountable complements which means it is not equal to the power set $P(\Omega)$.

Problem 2 (20 points).

Solution. (a) The problem statement says that $B_1, B_2, \cdots, B_t$ are the partitions of the sample space Ω. If A denotes any event of this sample space, we can denote the probability of event A ($Pr(A)$) with respect to the partitions of the sample space with the help of law of total probability as :

$$\Pr[A] = \sum_{k=1}^{t} \Pr[A \cap B_k]$$

Now, we can use the definition of conditional probability which is the probability of an event A given that event B occurs : $Pr(A \mid B) = \frac{Pr(A \cap B)}{Pr(B)}$. It can be re-written as $Pr(A \mid B) \cdot Pr(B) = Pr(A \cap B)$.
Here, event B is equivalent to the event that a partition $B_{k=[1..t]}$ is selected. So, $Pr(A \mid B_k) \cdot Pr(B_k) = Pr(A \cap B_k)$.

Let's substitute this value in the above law of total probability, then we get :

$$\Pr[A] = \sum_{k=1}^{t} \Pr[A \mid B_k] \Pr[B_k]$$

(b) Now, we need to show that probability of any event A is less than or equal to the maximum probability of event A given a partition $B_{k=[1..t]}$ of sample space Ω. We can use the formula that we derived above :

$$\Pr[A] = \sum_{k=1}^{t} \Pr[A \mid B_k] \Pr[B_k]$$

On the right hand side of this equation, if we replace all partition B_k with the partition which gives us the maximum conditional probability, then we will get an inequality.

$$\Pr[A] \leq \sum_{k=1}^{t} \max_{1 \leq j \leq t} (\Pr[A \mid B_j]) \Pr[B_k]$$

Now, we can take out the first term in summation outside because it is independent of k, so we get

$$\Pr[A] \leq \max_{1 \leq j \leq t} (\Pr[A \mid B_j]) \sum_{k=1}^{t} \Pr[B_k]$$

The summation actually represents the sum of all probabilities that partition $B_{k=[1..t]}$ will be selected which is equal to 1. So,

$$\Pr[A] \leq \max_{1 \leq j \leq t} (\Pr[A \mid B_j])$$

(In question, it is k, not j, variable name can be changed. Thus finally we have derived $\Pr[A] \leq \max_{1 \leq k \leq t} \Pr[A \mid B_k]$.)

Problem 3 (20 points)**.**

Solution. If H denotes that a coin shows head and T denotes that a coin shows tails then there are total four outcomes possible for 1 coins - {HH,HT,TH,TT}. This is our sample space (4 possible outcomes).
(a) We want to find two events A_1 and B_1 such that probability of event A_1 occurring is greater than the probability of event A_1 occurring given that event B_1 already occurred. Let A_1 be the event that both coins show heads and B_1 be the event that first coin shows tail. Let's calculate the individual probabilities $Pr[A_1]$ and $Pr[B_1]$ and the conditional probability $Pr[A_1 \mid B_1]$.

Event $A_1 = \{\text{HH}\}$. So, $Pr[A_1] = 1/4$

Event $B_1 = \{\text{T\_}\}$ (can be TH or TT). So, $Pr[B_1] = 2/4 = 1/2$

Now, if tail is shown in first coin (event B_1), then there is no chance that both coins can show heads (event A_1). So, $Pr[A_1 \mid B_1] = 0$.
Therefore, in this case $\Pr[A_1 \mid B_1] < \Pr[A_1]$ because $0 < 1/4$.

(b) We want to find two events A_2 and B_2 such that probability of event A_2 occurring is equal to the probability of event A_2 occurring given that event B_2 already occurred. Let A_2 be the event that second coin shows heads and B_2 be the event that first coin shows heads. Let's calculate the individual probabilities $Pr[A_2]$ and $Pr[B_2]$, intersection probability $Pr[A_2 \cap B_2]$ and the conditional probability $Pr[A_2 \mid B_2]$.

Event $A_2 = \{\text{\_H}\}$ (can be HH or TH). So, $Pr[A_2] = 2/4 = 1/2$

Event $B_2 = \{\text{H\_}\}$ (can be HT or HH). So, $Pr[B_2] = 2/4 = 1/2$

Events A_2 and B_2 both occur (intersection) = {HH}. So, $Pr[A_2 \cap B_2] = 1/4$

Let's use conditional probability formula now -
$\Pr[A_2 \mid B_2] = \frac{\Pr[A_2 \cap B_2]}{\Pr[B_2]} = \frac{(1/4)}{(1/2)} = 2/4 = 1/2$
Therefore, in this case $\Pr[A_2 \mid B_2] = \Pr[A_2]$ because $1/2 = 1/2$.

(c) We want to find two events A_3 and B_3 such that probability of event A_3 occurring is less than the probability of event A_3 occurring given that event B_3 already occurred. Let A_3 be the event that at least one coin shows heads and B_3 be the event that first coin shows heads. Let's calculate the individual probabilities $Pr[A_3]$ and $Pr[B_3]$, intersection probability $Pr[A_3 \cap B_3]$ and the conditional probability $Pr[A_3 \mid B_3]$.

Event A_3 can be HH or TH or HT. So, $Pr[A_3] = 3/4$

Event $B_3 = \{\text{H\_}\}$ (can be HT or HH). So, $Pr[B_3] = 2/4 = 1/2$

Events A_3 and B_3 both occur (intersection) = {HH, HT}. So, $Pr[A_3 \cap B_3] = 2/4 = 1/2$

Let's use conditional probability formula now -
$\Pr[A_3 \mid B_3] = \frac{\Pr[A_3 \cap B_3]}{\Pr[B_3]} = \frac{(1/2)}{(1/2)} = 2/2 = 1$
Therefore, in this case $\Pr[A_3 \mid B_3] > \Pr[A_3]$ because $1 > 3/4$.

Problem 4 (20 points)**.**

Solution. It is true that there may be several different min-cut sets in a graph and can be generated by the randomized min-cut algorithm. The probability of generating a specific min-cut set depends on the probability that the algorithm will contract all the edges in that set. Let S be a specific min-cut set, and let k be the number of edges in S. The probability that the first edge selected by the algorithm is not in S is $(n-k)/n$, since there are n vertices in total and $n-k$ vertices that are not in S. If this edge is contracted, then the remaining graph still has $n-1$ vertices, of which $n-k-1$ are not in S. The probability that the second edge selected by the algorithm is not in S, given that the first edge was contracted, is $(n-k-1)/(n-1)$. Continuing in this way, the probability

that all k edges in S are not contracted is : $\frac{n-k}{n} \cdot \frac{n-k-1}{n-1} \cdot \frac{n-k-2}{n-2} \cdots \frac{2}{k+1}$ which is greater than or equal to $2/(n(n-1))$ (Using the fact (harmonic number) that $1/2 + 1/3 + \cdots + 1/n \geq \ln(n)/2$, this probability can be shown to be at least $2/(n(n-1))$). Thus, the probability of generating S as the min-cut set = $\Pr[E_S]$ is at least $2/(n(n-1))$. This is also denotes the minimum probability of generating any other min-cut set in general : $\Pr[E_i] \geq 2/(n(n-1))$.
Now, there are n vertices in the graph and so there are $n(n-1)/2$ pairs of vertices. Each pair of vertices could potentially be the two vertices remaining at the end of the algorithm and choosing these pairs are disjoint events. Let $\Pr[E_i]$ denote the probability to choose one such pair, so $\sum_{i=1}^{c} P[E_i] \leq 1$ (since there are c such disjoint events). Substituting the probability value derived previously, we the inequality :

$$c \cdot \frac{2}{n(n-1)} \leq 1 \implies 2c \leq n(n-1) \implies c \leq \frac{n(n-1)}{2}$$

Hence, the number of distinct min-cut sets are at most $n(n-1)/2$.

Problem 5 (20 points)**.**

Solution. Suppose that the median-of-three is the m-th smallest element of the array. Then it gives at worst an a-to-$(1-a)$ split if and only if $an \leq m \leq (1-a)n$.

Let us count the number of sets of three elements that lead to the pivot being the m-th smallest element. There are $\binom{n}{3} = \frac{n(n-1)(n-2)}{6}$ ways to choose any three elements from the array. To count the number of sets of three elements that lead to the pivot being the m-th smallest element, we need to consider the following cases:
Case 1: The pivot is the smallest element among the three. In this case, we need to choose two elements from the remaining elements that are larger than the pivot. There are $\binom{n-m}{2}$ ways to do this.
Case 2: The pivot is the largest element among the three. In this case, we need to choose two elements from the remaining elements that are smaller than the pivot. There are $\binom{m-1}{2}$ ways to do this.
Case 3: The pivot is the median of the three. In this case, we need to choose one element from the set of elements that are smaller than the pivot and one element from the set of elements that are larger than the pivot. There are $(m-1)(n-m)$ ways to do this.
Now, we are interested in case 3 when the pivot is the median of the 3 elements. The probability that this pivot gives us at worst an a-to-$(1-a)$ split is the sum over $an \leq m \leq (1-a)n$ of the ratio of the number of sets obtained in case 3 and the total number of sets of three elements :

$$\Pr = \sum_{m=an}^{(1-a)n} \frac{(m-1)(n-m)}{\frac{n(n-1)(n-2)}{6}}$$

$$= \sum_{m=an}^{(1-a)n} 6\frac{(m-1)(n-m)}{n(n-1)(n-2)}$$

$$= \frac{6}{n(n-1)(n-2)} \sum_{m=an}^{(1-a)n} (m-1)(n-m)$$

$$= \frac{6}{n(n-1)(n-2)} \sum_{m=an}^{(1-a)n} mn - m^2 - n + m$$

$$= \frac{6}{n(n-1)(n-2)} \sum_{m=an}^{(1-a)n} m(n+1) - m^2 - n$$

$$= \frac{6}{n(n-1)(n-2)} [\sum_{m=an}^{(1-a)n} m(n+1) - \sum_{m=an}^{(1-a)n} m^2 - \sum_{m=an}^{(1-a)n} n]$$

$$= \frac{6}{n(n-1)(n-2)} [(n+1) \sum_{m=an}^{(1-a)n} m - \sum_{m=an}^{(1-a)n} m^2 - n \sum_{m=an}^{(1-a)n} 1]$$

Trying to simplify individual terms :
Using the formula for the sum of integers from an to $(1-a)n$ -
$(n+1)\sum_{m=an}^{(1-a)n} m = [an + (an+1) + \cdots + (1-a)n] = \frac{(n+1)(1-a^2)n^2}{2}$
Using the formula for the sum of squares of integers from an to $(1-a)n$
$\sum_{m=an}^{(1-a)n} m^2 = \frac{(1-3a+3a^2)n^3+(2-3a)n^2}{3}$.
The sum of the constant function 1 over the range $an \leq m \leq (1-a)n$ is simply the number of terms in the sum
$n\sum_{m=an}^{(1-a)n} 1 = n - an^2 + (1-a)n$
Combining, we get -

$$\Pr = \frac{6}{n(n-1)(n-2)} [\frac{(n+1)(1-a^2)n^2}{2} - \frac{(1-3a+3a^2)n^3 + (2-3a)n^2}{3} - (n - an^2 + (1-a)n)]$$

$$= \frac{(1-9a^2+6a)n^2 + (-1+12a-3a^2)n + (6a-12)}{(n-1)(n-2)}$$