(Fall 2024) 515512 Probability **(Early Bird: 2024/9/25, 9pm; Normal: 2024/9/26, 9pm)**

Homework 1: Axioms, Sets, and Conditional Probability

Submission Guidelines: Please merge all your write-ups (photos/scanned copies are acceptable; please make sure that the electronic files are of good quality and reader-friendly) into one .pdf file and submit the file via E3.

Problem 1 (Probability Axioms) (8+10=18 points)

(a) Use the probability axioms to show that $P(A \cup B) + P(A \cap B) = P(A) + P(B)$.

(b) Anya Forger is a student currently taking an introductory Probability course. She recently got a homework problem about using the 3 Probability Axioms to prove the Union Bound, i.e., for any sequence of events $A_1, A_2, \cdots, A_n$, we have $P(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i)$. However, Anya does not know how to proceed and would like to ask ChatGPT for help. The response provided by ChatGPT is shown below in Figure 1 in Page 4. Could you point out the mathematical mistake(s) made by ChatGPT?

Problem 2 (Countable Set Operations) (8+6+8+8=30 points)

(a) Let $S_1, S_2, \cdots$ be an infinite sequence of sets. Prove that

$$\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} S_n = \{x | x \in S_n, \text{for infinitely many } n\}.$$

(Hint: To prove that $S = T$, we need to show $S \subseteq T$ and $T \subseteq S$)

(b) Based on (a), suppose we construct another set as $H = \lim_{m\to\infty} \bigcap_{k=1}^{m} \bigcup_{n=k}^{m} S_n$. Then, is the set H equivalent to $\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} S_n$? Please clearly explain your answer.

(c) Let Ω be the universal set and B, C be two sets that satisfy $B \subseteq \Omega$ and $C \subseteq \Omega$. Let $\{F_k\}_{k=1}^{\infty}$ denote the Fibonacci sequence, i.e., $F_1 = F_2 = 1$ and $F_{k+1} = F_k + F_{k-1}$, for $k \geq 2$. Define a countably infinite sequence of sets $A_1, A_2, A_3, \cdots$ as

$$A_n = \begin{cases} B - C, & \text{if } n \text{ is in the Fibonacci sequence } \{F_k\}, \\ C - B, & \text{otherwise.} \end{cases}$$

What are $\bigcap_{n=1}^{\infty} A_n$, $\bigcup_{n=1}^{\infty} A_n$, $\bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n$, and $\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n$? Please clearly explain your answer.

(d) Show that there are *uncountably* infinite many real numbers in the interval $(0, 1)$. (Hint: Prove this by contradiction. Specifically, (i) assume that there are countably infinite real numbers in $(0, 1)$ and denote them as $x_1, x_2, x_3, \cdots, x_i, \cdots$; (ii) express each real number x_i between 0 and 1 in decimal expansion; (iii) construct a number y whose digits are either 1 or 2. Can you find a way to choose 1 or 2 such that y is different from all the x_is?)

Problem 3 (Continuity of Probability Functions) (12+14=26 points)

(a) In this problem, we will learn a super useful property called *Borel-Cantelli Lemma* as follows: Let $A_1, A_2, A_3, \cdots$ be a countably infinite sequence of events.

Theorem 1 (Borel-Cantelli Lemma) *If $\sum_{n=1}^{\infty} P(A_n) < \infty$, then $P(\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n) = 0$.*

Please show the above property by using both the following 2 approaches:

- **Approach 1**: Start by considering the continuity of probability function for $\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n$. Then, apply the union bound to find an upper bound of $\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n$.

- **Approach 2**: Show the equivalent statement that "*If* $P(\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty} A_n) > 0$, *then* $\sum_{n=1}^{\infty} P(A_n) = \infty$".

(b) To get a better understanding of Borel-Cantelli Lemma, let's consider a random experiment of a countably infinite sequence of Sorting Hat decisions.

- For each $k \in \mathbb{N}$, define an event A_k of getting "Gryffindor" from the Sorting Hat at the k-th decision.
- Let $N > 0$ be a positive real number. The probability of A_k is p_k, where

$$p_k = \begin{cases} \frac{1}{10} \cdot k^{-N} & \text{, if } k \text{ is even} \\ 0 & \text{, if } k \text{ is odd} \end{cases}$$

- We use I to denote the event of observing an infinite number of "Gryffindor".

Please find out the value of $P(I)$ under different N. Please clearly explain your answer. (Hint: Leverage the result in (a). Consider two cases: $N > 1$ and $0 < N \leq 1$)

Problem 4 (The Generalized Monte Hall Problem) (4+8+8=20 points)

Let us consider the generalized version of the Monty Hall problem discussed in Lecture 3.

- Suppose Bill is given N doors ($N \geq 3$), where M of them are empty and the remaining $N - M$ doors come with a prize. Suppose $M \geq 2$.
- Then, Bill picks a door (say Option #1).
- Next, the moderator opens one empty door and asks Bill "*Do you want to stay or randomly switch to another door*"?

Please find out the following:

- What is the sample space?
- Let $E_{M,N}$ denote the event that Bill wins the prize when Bill chooses to uniformly randomly switches to another door under the parameters M, N. What's the probability $P(E_{M,N})$? Please write your answers in terms of N, M and clearly explain your answers. (Hint: You need to consider two scenarios: (i) Bill's initial choice has a prize and (ii) Bill's initial choice does not have a prize)
- Please write a short MATLAB or python program to draw a heatmap that visualizes the value of $P(E_{M,N})$ under all valid combinations of M, N for all $3 \leq N \leq 100$. Can you find any interesting trends in the heatmap?

Problem 5 (Inference via Bayes' Rule) (10+6=16 points)

Suppose we are given a customized Sorting Hat with unknown characteristics. Let $\theta_G, \theta_R, \theta_S, \theta_H$ denote the unknown probabilities of getting "Gryffindor" (G), "Ravenclaw" (R), "Slytherin" (S), and "Hufflepuff" (H) at each play, respectively. Moreover, suppose that the tuple of the unknown parameters $\theta \equiv (\theta_G, \theta_R, \theta_S, \theta_H)$ can only be one of the following three possibilities:

$$(\theta_G, \theta_R, \theta_S, \theta_H) \in \{(0.25, 0.25, 0.25, 0.25), (0.5, 0.2, 0.2, 0.1), (0.2, 0.5, 0.2, 0.1)\}.$$

In order to infer the values $(\theta_G, \theta_R, \theta_S, \theta_H)$, we experiment with this customized Sorting Hat and consider Bayesian inference as follows:

- Define events $A_1 := \{\theta = (0.25, 0.25, 0.25, 0.25)\}$, $A_2 := \{\theta = (0.5, 0.2, 0.2, 0.1)\}$, $A_3 := \{\theta = (0.2, 0.5, 0.2, 0.1)\}$.
- Moreover, Prof. Dumbledore told us some useful information about the prior probability assignment to be $P(A_1) = \frac{1}{2}$, $P(A_2) = \frac{1}{4}$, and $P(A_3) = \frac{1}{4}$.

Suppose we try the Sorting Hat for 6 times and observe SGRGGH (for ease of notation, we define the event $C = \{\text{SGRGGH}\}$). Moreover, all the trials are assumed to be independent.

- What is the posterior probability $P(A_1|C)$, $P(A_2|C)$, and $P(A_3|C)$?
- Given the experimental results, what is the most probable value for θ?

The union bound is a fundamental result in probability theory, stating that the probability of the union of two or more events is at most the sum of the individual probabilities of those events. It can be proven using the three probability axioms:

1. **Non-negativity:** For any event A, $P(A) \geq 0$.
2. **Normalization:** The probability of the entire sample space S is 1, i.e., $P(S) = 1$.
3. **Additivity:** For any two mutually exclusive events A and B (meaning they cannot both occur at the same time), $P(A \cup B) = P(A) + P(B)$.

Now, let's prove the union bound for two events A and B:

1. **Using Additivity:**
   $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

This equation comes from the inclusion-exclusion principle. The events A and B can both occur when A occurs and B occurs, but we counted this intersection ($A \cap B$) twice, so we subtract it once to correct for the overcounting.

1. **Using Non-negativity:**
   $P(A \cup B) \geq P(A) + P(B)$

From the non-negativity axiom, $P(A \cap B)$ is greater than or equal to 0. So, subtracting $P(A \cap B)$ from both sides of the equation does not change the inequality.

Now, the union bound can be extended to the union of n events $A_1, A_2, \ldots, A_n$ using mathematical induction:

Base Case: For $n = 2$, we have already proven the union bound for A and B.

Inductive Step: Assume the union bound holds for $n = k$, i.e.,
$P\left(\bigcup_{i=1}^{k} A_i\right) \leq \sum_{i=1}^{k} P(A_i)$

Now, consider $n = k + 1$. Using the base case for $A_1, A_2, \ldots, A_k$ and the additivity axiom, we have:
$P\left(\bigcup_{i=1}^{k+1} A_i\right) = P\left(\bigcup_{i=1}^{k} A_i\right) + P(A_{k+1}) - P\left(\left(\bigcup_{i=1}^{k} A_i\right) \cap A_{k+1}\right)$

Using the inductive assumption, this can be simplified to:
$\leq \sum_{i=1}^{k} P(A_i) + P(A_{k+1}) - P\left(\bigcup_{i=1}^{k} (A_i \cap A_{k+1})\right)$

Now, by the additivity axiom and the fact that intersections of events are also events:
$\leq \sum_{i=1}^{k} P(A_i) + P(A_{k+1}) - \sum_{i=1}^{k} P(A_i \cap A_{k+1})$

$= \sum_{i=1}^{k+1} P(A_i)$

Therefore, by mathematical induction, the union bound holds for any finite collection of events.

Figure 1: The response provided by ChatGPT.