(Fall 2022) 515502 Probability **(Early Bird: 2022/10/14, 9pm; Normal: 2022/10/17, 9pm)**

Homework 1: Axioms, Sets, Conditioning, and Random Variables

Submission Guidelines: Please compress 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 compressed file via E3.

Problem 1 (Countable Set Operations) (8+8=16 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) Let Ω be the universal set an let $\{A_n\}$ and $\{B_n\}$ be two countable sequences of subsets of Ω. Show that

$$\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} (A_n \cup B_n) = \Big(\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n\Big) \cup \Big(\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} B_n\Big).$$

(Hint: Again, to prove that the two sets are identical, we need to show that each is a subset of the other.)

Problem 2 (Probability Axioms) (8+8=16 points)

(a) Use the probability axioms to show the following three useful properties:

- $P(A^c) = 1 - P(A)$
- $P(A) = P(A - B) + P(A \cap B)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(b) Consider a random experiment with a sample space $\Omega = \{1, 2, 3, 4, 5\}$. Suppose we know $P(\{1, 2\}) = 0.2$, $P(\{2, 3\}) = 0.35$, $P(\{3, 4\}) = 0.45$, $P(\{4, 5\}) = 0.6$, and $P(\{1, 5\}) = 0.4$. Please write down one possible valid probability assignment. Moreover, does there exist one unique valid probability assignment that satisfies all the above conditions? Please clearly explain your answer. (Hint: We mentioned in the class that verifying the validity of a probability assignment can be done by solving a system of linear equations.)

Problem 3 (Continuity of Probability Functions) (10+10+8+8=36 points)

(a) Let $A_1, A_2, A_3, \cdots$ be a countably infinite sequence of events. Prove that if $\sum_{n=1}^{\infty} P(A_n) < \infty$, then $P(\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n) = 0$. This property is known as the *Borel-Cantelli Lemma*. (Hint: Consider the continuity of probability function for $\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n$ and then apply the union bound)

(b) Let $A_1, A_2, A_3, \cdots$ be a countably infinite sequence of *independent* events. Show that

$$P\Big(\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n\Big) = \begin{cases} 0, & \text{if } \sum_{n=1}^{\infty} P(A_n) < \infty, \\ 1, & \text{if } \sum_{n=1}^{\infty} P(A_n) = \infty. \end{cases}$$

This property is known as the *Borel Zero-One Law*. (Hint: For the first case, use the result in (a). For the second case, apply the "continuity of probability" to $P(\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n)$ and use the independence property.)

(c) Consider a countably infinite sequence of tosses of moon blocks. Recall that the possible outcomes of each toss are "Yes", "Laughing", and "No." Suppose that the outcomes of all the tosses are *independent*. The probability of having a "Yes" at the k-th toss is p_k, with $p_k = \frac{1}{100} \cdot k^{-N}$, where $N > 0$. We use I to denote the event of observing an infinite number of "Yes". Show that $P(I) = 1$ if $N \leq 1$. Under what values of N do we have $P(I) = 0$? Please clearly explain your answer. (Hint: Leverage the result in (a)-(b))

(d) Similar to the setting in (c), let's still consider a countably infinite sequence of tosses of moon blocks. However, suppose the different tosses are NOT necessarily independent. Again, the probability of having a "Yes" at the k-th toss is p_k, with $p_k = \frac{1}{100} \cdot k^{-N}$, where $N > 0$. We use I to denote the event of observing an infinite number of heads.

- Show that $P(I) = 0$ if $N > 1$.
- Is it possible to have $0 < P(I) < 1$? If so, could you construct an example of $\{A_n\}$ that gives us $0 < P(I) < 1$?

Please clearly explain your answer. (Hint: Leverage the result in (a)-(b))

Problem 4 (Inference via Bayes' Rule) (8+8+8=24 points)

Suppose we are given a special pair of moon blocks with unknown characteristics. Let $\theta_Y, \theta_L, \theta_N$ denote the unknown probabilities of getting a "Yes" (Y), "Laughing" (L), and "No" (N) at each toss, respectively. Moreover, suppose that the tuple of the unknown parameters $\theta \equiv (\theta_Y, \theta_L, \theta_N)$ can only be one of the following three possibilities: $(\theta_Y, \theta_L, \theta_N) \in \{(0.1, 0.3, 0.6), (0.3, 0.6, 0.1), (0.6, 0.3, 0.1)\}$. In order to infer the values $(\theta_Y, \theta_L, \theta_N)$, we experiment with the moon blocks and consider Bayesian inference as follows: Define events $A_1 = \{\theta_Y = 0.1, \theta_L = 0.3, \theta_N = 0.6\}$, $A_2 = \{\theta_Y = 0.3, \theta_L = 0.6, \theta_N = 0.1\}$, $A_3 = \{\theta_Y = 0.6, \theta_L = 0.3, \theta_N = 0.1\}$. Since initially we have no further information about $(\theta_Y, \theta_L, \theta_N)$, we simply consider the prior probability assignment to be $P(A_1) = P(A_2) = P(A_3) = 1/3$.

(a) Suppose we toss the pair of moon blocks once and observe an "L" (for ease of notation, we define the event $B = \{\text{the first toss is an L}\}$). What is the evidence $P(B)$? What is the posterior probability $P(A_1|B)$? How about $P(A_2|B)$ and $P(A_3|B)$? (Hint: use the Bayes' rule)

(b) Suppose we toss the pair of moon blocks for 10 times and observe YLNLYLLYLL (for ease of notation, we define the event $C = \{\text{YLNLYLLYLL}\}$). Moreover, all the tosses 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?

(c) Given the same setting as (b), suppose we instead choose to use a different prior probability assignment $P(A_1) = \frac{1}{3}, P(A_2) = \alpha, P(A_3) = \frac{2}{3} - \alpha$. Given the experimental results as in (b), under what value of α would $(\theta_Y, \theta_L, \theta_N) = (0.3, 0.6, 0.1)$ be the most probable value for θ?

Problem 5 (Random Variables and CDF) (8 points)

A random variable X is said to be "symmetric about 0" if for all $x \in \mathbb{R}$,

$$P(X \geq x) = P(X \leq -x).$$

Show that if X is symmetric about 0, then its CDF $F_X(t)$ must satisfy the following properties:

- $P(|X| \leq t) = 2 \cdot F_X(t) - 1$, for all $t \geq 0$
- $P(X = t) = F(t) + F(-t) - 1$