APMA1690: Homework # 1 (Due by 11 pm Sept 21)

1 Review

I would suggest you read the Review section before going to the problem set. - Mike

1.1 Three Building Blocks of Probability Theory

1.1.1 Sample Spaces

Definition 1.1. 1. *The* ***sample space*** *of an experiment is the collection of all possible outcomes of the experiment. A sample space is usually denoted by* Ω.

2. *Any subset* A *of* Ω *(allowed to be empty* $\emptyset$*) is called an* ***event****, and* $\emptyset$ *is called the/an* ***impossible event****. The sample space* Ω*, as a subset of itself, is called the* ***inevitable event****.*

1.1.2 Random Variables

Definition 1.2. *Let* Ω *be a sample space and* $\mathbb{R}^d$ *denote d-dimensional space.*

- *Any map* $X : \Omega \to \mathbb{R}^d$, $\omega \mapsto X(\omega)$ *is called a (*$\mathbb{R}^d$*-valued)* ***random variable****; when* $d \geq 2$*, the* $\mathbb{R}^d$*-valued random variable is also referred to as a* ***random vector****.*
- *If there exists a fixed* $x \in \mathbb{R}^d$ *such that* $X(\omega) = x$ *for all* $\omega \in \Omega$*, i.e.,* $X(\omega)$ *is a constant function of* ω*, we call* X ***deterministic****.*
- *If random variable* X *is not deterministic, we call* X ***truly random****.*

1.1.3 Probabilities

Definition 1.3. *Let* Ω *be a sample space. Suppose* $\mathbb{P}$ *is a real-valued function of subsets of* Ω*, i.e.,*

$$\begin{aligned} \mathbb{P}:\ & \{\textit{subsets of } \Omega\} \to \mathbb{R}, \\ & A \mapsto \mathbb{P}(A). \end{aligned}$$

If $\mathbb{P}$ *satisfies the following three axioms,* $\mathbb{P}$ *is called a* ***probability****, and the pair* $(\Omega, \mathbb{P})$ *is called a* ***probability space***

1. $\mathbb{P}(A) \geq 0$ *for any subset* $A \subseteq \Omega$*;*
2. $\mathbb{P}(\Omega) = 1$*;*
3. *For any infinitely long sequence of disjoint subsets* $\{A_i\}_{i=1}^{\infty}$*, i.e.,* $A_i \cap A_j = \emptyset$ *if* $i \neq j$*, we have*

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

1.2 Properties of Probabilities

Theorem 1.1. *Let $(\Omega, \mathbb{P})$ be a probability space. Then, we have the following*

1. $\mathbb{P}(\emptyset) = 0$, *i.e., the probability of the impossible event is zero;*
2. *if two events E_1 and E_2 satisfy $E_1 \cap E_2 = \emptyset$, we have $\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2)$;*
3. *suppose $A, B \subseteq \Omega$. If $A \subseteq B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$;*
4. $0 \leq \mathbb{P}(A) \leq 1$ *for any subsets $A \subseteq \Omega$;*
5. *for any $A, B \subseteq \Omega$, we have $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$;*
6. *for any sequence of subsets $\{A_n\}_{n=1}^{\infty}$, we have $\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n)$.*

1. Let $A_1 = \Omega$ and $A_n = \emptyset$ for all $n \geq 2$. Then, $\{A_n\}_{n=1}^{\infty}$ is a sequence of disjoint sets. We have

$$A_1 = \Omega = \Omega \cup \emptyset \cup \emptyset \cdots \cup \emptyset \cup \cdots = \bigcup_{n=1}^{\infty} A_n,$$

   which implies $\mathbb{P}(A_1) = \mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n) = \mathbb{P}(A_1) + \mathbb{P}(\emptyset) + \mathbb{P}(\emptyset) + \cdots + \mathbb{P}(\emptyset) + \cdots$. We cancel $\mathbb{P}(A_1)$ and get the following

$$0 = \mathbb{P}(\emptyset) + \mathbb{P}(\emptyset) + \cdots + \mathbb{P}(\emptyset) + \cdots .$$

   Since the definition of probability enforce $\mathbb{P}(\emptyset) \geq 0$, we have $\mathbb{P}(\emptyset) = 0$.

2. Let $A_1 = E_1$, $A_2 = E_2$, and $A_n = \emptyset$ for $n \geq 3$. Then, $\{A_n\}_{n=1}^{\infty}$ is a sequence of disjoint sets. We have

$$E_1 \cup E_2 = A_1 \cup A_2 \cup \emptyset \cup \emptyset \cup \cdots \cup \emptyset \cup \cdots = \bigcup_{n=1}^{\infty} A_n,$$

   which implies

$$\begin{aligned}
\mathbb{P}(E_1 \cup E_2) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) + \sum_{n=1}^{\infty} \mathbb{P}(A_n) \\
&= \mathbb{P}(E_1) + \mathbb{P}(E_2) + \mathbb{P}(\emptyset) + \mathbb{P}(\emptyset) + \cdots + \mathbb{P}(\emptyset) + \cdots \\
&= \mathbb{P}(E_1) + \mathbb{P}(E_2).
\end{aligned}$$

3. $B = A \cup (B - A)$. Since $A \cap (B - A) = \emptyset$, we have $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B - A)$. Because $\mathbb{P}(B - A) \geq 0$, we have $\mathbb{P}(B) \geq \mathbb{P}(A)$.

The proofs of other results are left for homework.

1.3 Indicator Functions

Let A be a subset of $\mathbb{R}^d$. The **indicator function** $\mathbf{1}_A$ of A is defined as

$$\mathbf{1}_A(x) := \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{otherwise.} \end{cases} \tag{1.1}$$

The function $\mathbf{1}_A(x)$ is sometimes represented as $\mathbf{1}(x \in A)$.

1.4 Cumulative Distribution Functions (CDFs)

Definition 1.4. *Let X be an $\mathbb{R}$-valued random variable defined on an underlying probability space $(\Omega, \mathbb{P})$. The function F_X defined as follows*

$$F_X(t) := \mathbb{P}\left(\{\omega \in \Omega : X(\omega) \leq t\}\right) = \mathbb{P}\left(\{\omega \in \Omega : X(\omega) \in (-\infty, t]\}\right), \quad \textit{for all } t \in \mathbb{R}, \tag{1.2}$$

*is called the **cumulative distribution function** (CDF) of X, which is denoted as $X \sim F_X$. (F_X is sometimes briefly denoted by F.)*

Remark: $F_X(t)$ is defined for **all** real numbers $t \in \mathbb{R}$.

2 Problem Set

1. (2 points) Suppose $\Omega = \{1, 2, \cdots, n\}$ is the sample space of interest, where $n < +\infty$ is a positive integer. For any subset (i.e., event) $A \subseteq \Omega$, we define

$$\mathbb{P}(A) := \frac{\#A}{n}$$

where $\#A$ denotes the number of elements in A. Please verify that the $\mathbb{P}$ defined above is a probability.

To be a probability, $\mathbb{P}$ must satisfy three axioms:

(a) $\mathbb{P}(A) \geq 0 \quad A \subset \Omega$
To see that this is true, observe that $A \subset \Omega$ so $0 \leq \#A \leq n$. Hence,

$$\frac{0}{n} \leq \frac{\#A}{n} \leq \frac{n}{n}$$

By definition of $\mathbb{P}$,

$$0 \leq \mathbb{P}(A) \leq 1$$

which is a stronger condition than $\mathbb{P}(A) \geq 0$

(b) $\mathbb{P}(\Omega)$ By definition of $\mathbb{P}$,

$$\mathbb{P}(\Omega) = \frac{\#\Omega}{n} = \frac{n}{n} = 1$$

(c) For any sequence of disjoint subsets, $\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$
Let $A := \bigcup_{i=1}^{m} A_i$ be a sequence of m disjoint events in Ω. Then $\#A =$ so

$$\mathbb{P}\left(\bigcup_{i=1}^{m} A_i\right) = \mathbb{P}(A) = \frac{\#A}{n} = \frac{m}{n}$$

But as m is a positive integer,

$$\frac{m}{n} = \sum_{i=1}^{n} \frac{1}{n}$$

Then as each A_i is disjoint, $\mathbb{P}(A_i) = \frac{1}{n}$ so

$$\frac{m}{n} = \sum_{i=1}^{m} \mathbb{P}(A_i) = \mathbb{P}\left(\bigcup_{i=1}^{m} A_i\right)$$

However, this does not depend on the finiteness of m so

$$\sum_{i=1}^{\infty} \mathbb{P}(A_i) = \mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right)$$

Then, as $\mathbb{P}$ satisfies all three requirements, it is a probability. ■

2. Please prove the results iv), v), and vi) of Theorem 1.1 (see the Review section), i.e.,

- (1 point) $0 \leq \mathbb{P}(A) \leq 1$ for any subsets $A \subseteq \Omega$;
  As $(\mathbb{P}, \Omega)$ is a probability space, $\mathbb{P}(A) \geq 0$. But as $A \subseteq \Omega$, $\mathbb{P}(A) \leq \mathbb{P}(\Omega)$. Thus by the definition of a probability,

$$0 \leq A \leq \mathbb{P}(\Omega) = 1 \implies 0 \leq A \leq 1 \quad \blacksquare$$

- (1 point) for any $A, B \subseteq \Omega$, we have $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$;
  For any $A, B \subseteq \Omega$, $A \cap (B \cap A^c) = \emptyset$ so from Property 2,

$$\mathbb{P}(A \cup (B \cap A^c)) = \mathbb{P}(A) + \mathbb{P}(B \cap A^c)$$

  Additionally,

$$(B \cap A^c) \cap (B \cap A) = \emptyset$$

  and partition Ω. Therefore,

$$\mathbb{P}(B) = \mathbb{P}(B \cap A^c) + \mathbb{P}(B \cap A)$$

  Rearranging,

$$\mathbb{P}(B \cap A^c) = \mathbb{P}(B) - \mathbb{P}(B \cap A)$$

  so together with the first equation,

$$\begin{aligned} \mathbb{P}(A \cup (B \cap A^c)) &= \mathbb{P}(A) + \mathbb{P}(B \cap A^c) \\ &= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(B \cap A) \end{aligned}$$

  Finally, observe that

$$\begin{aligned} A \cup (B \cap A^c) &= (A \cup B) \cap (A \cup A^c) \\ &= (A \cup B) \cap \Omega \\ &= A \cup B \end{aligned}$$

  so

$$\mathbb{P}(A \cup (B \cap A^c)) = \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(B \cap A) \quad \blacksquare$$

- (1 point) for any sequence of subsets $\{A_n\}_{n=1}^{\infty}$, we have $\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n)$.
  In the case where $\{A_n\}_{n=1}^{\infty}$ is mutually disjoint, the equality follows trivially from Axiom 3. When the sequence is not mutually disjoint, we observe that

$$\mathbb{P}(A_1 \cup A_2) \leq \mathbb{P}(A) + \mathbb{P}(A_2)$$

  because $\mathbb{P}(A \cap A_2) \geq 0$.

Now to establish the inductive step, we see that

$$\begin{aligned}\mathbb{P}(A_1 \cup A_2 \cup A_3) &= \mathbb{P}(A_1 \cup A_2) + \mathbb{P}(A_3) - \mathbb{P}((A_1 \cup A_2) \cap A_3) \\ &= \mathbb{P}(A_1) + \mathbb{P}(B_2) + \mathbb{P}(A_3) - \mathbb{P}(A_1 \cap A_2) - \mathbb{P}((A_1 \cup A_2) \cap A_3)\end{aligned}$$

with $\mathbb{P}(A_1 \cap A_2) \geq 0$ and $\mathbb{P}((A_1 \cup A_2) \cap A_3) \geq 0$ so

$$\mathbb{P}(A_2 \cup A_2 \cup A_3) \leq \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3)$$

That is, for $n \geq 2$,

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = -\tilde{P} + \sum_{i=1}^{n} \mathbb{P}(A_i)$$

where $\tilde{P} \geq 0$ is the sequence of intersections of earlier n. Thus,

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mathbb{P}(A_i) \quad \blacksquare$$

Since the results *i)*, *ii)*, and *iii)* have been proved in the Review section, you can directly apply these results (i.e., results *i)*, *ii)*, and *iii)*) in your proofs.

3. Suppose the sample space of interest is $\Omega = [0,1] = \{$all the real numbers that are ≥ 0 and $\leq 1\}$. For any subset (i.e., event) $A \subseteq [0,1] = \Omega$, we define

$$\mathbb{P}(A) = \int_0^1 \mathbf{1}_A(x)\,dx,$$

where $\mathbf{1}_A(x)$ is the indicator function of A (see Eq. (1.1)). The $\mathbb{P}$ defined above is a probability (you do not need to prove this fact).

Let X be a random variable defined by

$$X(\omega) = \omega + 1,$$

for all $\omega \in \Omega = [0,1]$.

(a) (1 point) Consider the event

$$A = \{\omega \in \Omega \,|\, X(\omega) = 1.5\}. \tag{2.1}$$

Please calculate the probability of the event A defined in Eq. (2.1), i.e., $\mathbb{P}(A)$.

$$X = 1.5 \implies \omega = 0.5$$

$$\begin{aligned}
\mathbb{P}(A) &= \int_0^1 \mathbb{1}_A(x)\;dx \\
&= \int_0^{0.5} \mathbb{1}_A(x)\;dx + \int_{0.5}^{0.5} \mathbb{1}_A(x)\;dx + \int_{0.5}^{1} \mathbb{1}_A(x)\;dx \\
&= \int_0^{0.5} 0\;dx + \int_{0.5}^{0.5} 1\;dx + \int_{0.5}^{1} 0\;dx \\
&= 0 + 0 + 0 = \boxed{0}
\end{aligned}$$

(b) (0.5 points) Is the event A defined in Eq. (2.1) an impossible event?

Despite the fact that $\mathbb{P}(A) = 0$, the event $\boxed{\text{is not impossible}}$. $X(\omega) = 1.5 = 0.5 + 1$ occurs when $\omega = 0.5 \in [0,1]$ so the event can occur.

(c) (0.5 points) Consider the event

$$B = \{\omega \in \Omega \,|\, X(\omega) = 0.5\}. \tag{2.2}$$

Please calculate the probability of the event B defined in Eq. (2.2), i.e., $\mathbb{P}(B)$.

$$X = 0.5 \implies \omega = -0.5$$

But $-0.5 \notin [0,1]$ so

$$\mathbb{P}(A) = \int_0^1 \mathbb{1}_A(x)\;dx = \int_0^1 0\;dx = \boxed{0}$$

(d) (0.5 points) Is the event B defined in Eq. (2.2) an impossible event?

$\omega \notin [0,1] = \Omega$ so it is $\boxed{\text{impossible}}$.

(e) (0.5 points) Please calculate the CDF of X.

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(\omega + 1 \leq x) = \begin{cases} 0 & x < 1 \\ 1 & 1 \leq x \leq 2 \\ 0 & x > 2 \end{cases}$$

4. (2 points) Let $(\Omega, \mathbb{P})$ be a probability space and X a random variable defined on Ω. Please prove that the CDF $F_X(t)$ of X is a non-decreasing function, i.e., $F_X(t_1) \leq F_X(t_2)$ if $t_1 \leq t_2$.

   By definition,

$$F_X(t_1) = \mathbb{P}(X \leq t_1) = \mathbb{P}(X \in (-\infty, t_1])$$
$$F_X(t_2) = \mathbb{P}(X \leq t_2) = \mathbb{P}(X \in (-\infty, t_2])$$

   However, if $t_1 \leq t_2$,

$$\begin{aligned} F_X(t_2) &= \mathbb{P}(X \in (-\infty, t_1] \cup (t_1, t_2]) \\ &= \mathbb{P}(-\infty < X \leq t_1) + \mathbb{P}(t_1 < X \leq t_2) \\ &= F_X(t_1) + \mathbb{P}(t_1 < X \leq t_2) \end{aligned}$$

   and by the fact that $\mathbb{P}(A) \geq 0 \quad A \in \Omega$, $\mathbb{P}(t_1 < X \leq t_2) \geq 0$ so

$$F_X(t_2) - F_X(t_1) \geq 0$$

   and

$$F_X(t_2) \geq F_X(t_1) \quad \blacksquare$$