Probability Theory

February 3, 2019

1 Probability space

Definition: The 3-tuple $(\Omega, \mathcal{F}, P)$ is called a probability space, where

1. Ω is a set called the sample space.
2. $\mathcal{F}$ is a σ-field.

   Definition of σ-field: $\mathcal{F}$ is a non-empty collection of subsets of Ω which satisfies

   (S1) $\Omega \in \mathcal{F}$.

   (S2) If $A \subseteq \Omega$ and $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$.

   (S3) If each set in the collection $\{A_n; n \in \mathbb{N}\}$ belongs to $\mathcal{F}$, *i.e.*, $A_n \in \mathcal{F}$, $\forall n \in \mathbb{N}$ (not necessarily disjoint), then $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

   Note that $A \subseteq \Omega$ is called $\mathcal{F}$-set if $A \in \mathcal{F}$.

3. P is a probability measure.

   Definition of probability measure: $P : \mathcal{F} \to [0,1]$ is called a probability measure if it satisfies:

   (M1) $P(\Omega) = 1$ and $P(\emptyset) = 0$.

   (M2) If $\{A_n\}_{n \in \mathbb{N}}$ is a disjoint collection of $\mathcal{F}$-sets, *i.e.*, $A_k \cap A_j = \emptyset$, for $k \neq j$, then

   $$P(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n). \tag{1}$$

   This property is called the *countable additivity of the probability measure.*

   Remark 1. *A similar concept to countable additivity is the finite additivity which is defined as follows: If $\{A_i; 1 \leq i \leq n\}$ is a finite collection of disjoint $\mathcal{F}$-sets, then $P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$. Note that countable* *additivity implies finite additivity. Indeed, by considering the countable collection $\{B_i; i \in \mathbb{N}\}$, where $B_1 = A_1, \ldots, B_n = A_n$, and $B_k = \emptyset$, for $k > n$, the claim follows.*

Lemma 1. *If A and B are $\mathcal{F}$-sets with $A \subseteq B$, then $P(A) \leq P(B)$. Also, $P(B \setminus A) = P(B) - P(A)$.*

Proof. Note that since $A \subseteq B$, we have $B = A \cup (B \setminus A)$ and, A and $B \setminus A$ are disjoint. Now, by the finite additivity of P, we have

$$P(B) = P(A) + \underbrace{P(B \setminus A)}_{\geq 0} \tag{2}$$
$$\Rightarrow P(B) \geq P(A).$$

This proves the first part. The second part follows from Eq. (2). □

Lemma 2. *If A is an $\mathcal{F}$-set, then $P(A^c) = 1 - P(A)$.*

Proof. Note that $A \cup A^c = \Omega$. Also, A and A^c are disjoint. Therefore by finite additivity property of P and M1, we have

$$1 = P(\Omega) = P(A \cup A^c) = P(A) + P(A^c).$$

Hence, the claim follows. □

1.1 Limit of sets

Definition: (Liminf of a sequence of sets) Given a sequence of sets $\{A_n\}_{n\in\mathbb{N}}$, where $A_n \subseteq \Omega$, we define

$$\liminf_n A_n = \cup_n \cap_{k\geq n} A_k. \tag{3}$$

Definition: (Limsup of a sequence of sets) Given a sequence of sets $\{A_n\}_{n\in\mathbb{N}}$, where $A_n \subseteq \Omega$, we define

$$\limsup_n A_n = \cap_n \cup_{k\geq n} A_k. \tag{4}$$

Definition: (Limit of a sequence of sets) We say the limit of the sequence of sets $\{A_n\}_{n\in\mathbb{N}}$ exists if $\liminf_n A_n = \limsup_n A_n$ and the $\lim_n A_n$ is that common set.

We will consider specific sequences here

1.1.1 Monotonically increasing sequence of sets

Definition: A sequence $\{A_n\}_{n\in\mathbb{N}}$ is called *monotonically increasing* sequence if $A_n \subseteq A_{n+1}$, $\forall n \in \mathbb{N}$.

In this case, note that for $n \in \mathbb{N}$,

$$\cap_{k\geq n} A_k = A_n, \text{ since } A_n \subseteq A_{n+1} \subseteq A_{n+2} \ldots$$

Therefore,

$$\liminf_n A_n = \cup_n \cap_{k \geq n} A_k = \cup_n A_n. \tag{5}$$

Note that for $n > 1$, since $A_1 \subseteq A_2 \cdots \subseteq A_{n-1} \subseteq A_n$, we have

$$\cup_{k=1}^n A_k = A_n \Rightarrow \cup_{k \geq n} A_k = \cup_{k \geq 1} A_k \tag{6}$$

Therefore,

$$\limsup_n A_n = \cap_n \cup_{k \geq n} A_k = \cap_n \cup_{k \geq 1} A_k = \cup_{k \geq 1} A_k. \tag{7}$$

Therefore, by the definition of $\lim_n A_n$, we have

$$\lim_n A_n = \cup_n A_n. \tag{8}$$

The next question is what happens to the probability of the monotonically increasing sets A_n when each A_n is an $\mathcal{F}$-set. Indeed, we are considering the real sequence $\{P(A_n)\}_{n \in \mathbb{N}}$. The real sequence $\{P(A_n)\}_{n \in \mathbb{N}}$ is bounded since $0 \leq P(A_n) \leq 1$, $\forall n \in \mathbb{N}$. Also since the set sequence $\{A_n\}_{n \in \mathbb{N}}$ is monotonically increasing, we have, for $n \in \mathbb{N}$,

$$A_{n+1} \supseteq A_n \Rightarrow P(A_{n+1}) \geq P(A_n), \text{ (follows from Lemma 1)}$$

Therefore, the real sequence $\{P(A_n)\}_{n \in \mathbb{N}}$ is a monotonically increasing bounded sequence. Hence it should converge. But where does it converge to?

Theorem 1. *If $\{A_n\}_{n \in \mathbb{N}}$ is a monotonically increasing sequence of $\mathcal{F}-$sets, then*

$$\lim_{n \to \infty} P(A_n) = P(\lim_n A_n) = P(\cup_n A_n). \tag{9}$$

Proof. Let $A_0 = \emptyset$. Now set

$$\begin{aligned} B_1 &:= A_1 \setminus A_0; \\ B_2 &:= A_1 \setminus A_1; \\ &\vdots \\ B_n &:= A_n \setminus A_{n-1}; \\ &\vdots \end{aligned}$$

Now note that the set sequence $\{B_n\}_{n \in \mathbb{N}}$ is a disjoint sequence, *i.e.*, $B_i \cap B_j = \emptyset$, for $i \neq j$. Also,

$$\cup_n B_n = \cup_n A_n. \tag{10}$$

Therefore, from Eq. (10) and the fact that the set sequence $\{A_n\}_{n\in\mathbb{N}}$ is monotonically increasing, we have

$$\begin{aligned}
P(\lim_n A_n) = P(\cup_n A_n) &= P(\cup_n B_n) \\
&= \sum_{n\in\mathbb{N}} P(B_n) \text{ (follows from M2)} \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} P(A_i) - P(A_{i-1}) \text{ (follows from Lemma 1)} \\
&= \lim_{n\to\infty} P(A_n) - \underbrace{P(A_0)}_{=0} \\
&= \lim_{n\to\infty} P(A_n).
\end{aligned}$$

□

1.1.2 Monotonically decreasing sequence of sets

Definition: A sequence $\{A_n\}_{n\in\mathbb{N}}$ is called *monotonically decreasing* sequence if $A_{n+1} \subseteq A_n$, $\forall n \in \mathbb{N}$.

In this case, note that for $n \in \mathbb{N}$,

$$\cup_{k\geq n} A_k = A_n, \text{ since } A_n \supseteq A_{n+1} \supseteq A_{n+2} \ldots$$

Therefore,

$$\limsup_n A_n = \cap_n \cup_{k\geq n} A_k = \cap_n A_n. \tag{11}$$

Note that for $n > 1$, since $A_1 \supseteq A_2 \cdots \supseteq A_{n-1} \supseteq A_n$, we have

$$\cap_{k=1}^{n} A_k = A_n \Rightarrow \cap_{k\geq n} A_k = \cap_{k\geq 1} A_k \tag{12}$$

Therefore,

$$\liminf_n A_n = \cup_n \cap_{k\geq n} A_k = \cup_n \cap_{k\geq 1} A_k = \cap_{k\geq 1} A_k. \tag{13}$$

Therefore, by the definition of $\lim_n A_n$, we have

$$\lim_n A_n = \cap_n A_n. \tag{14}$$

What happens to the probability of the monotonically decreasing sets A_n when each A_n is an $\mathcal{F}$-set. Here also, the real sequence $\{P(A_n)\}_{n\in\mathbb{N}}$ is bounded since $0 \leq P(A_n) \leq 1$, $\forall n \in \mathbb{N}$. Also since the set sequence $\{A_n\}_{n\in\mathbb{N}}$ is monotonically decreasing, we have, for $n \in \mathbb{N}$,

$$A_{n+1} \subseteq A_n \Rightarrow P(A_{n+1}) \leq P(A_n), \text{ (follows from Lemma 1)}$$

Therefore, the real sequence $\{P(A_n)\}_{n\in\mathbb{N}}$ is a monotonically decreasing bounded sequence. Hence it should converge.

Theorem 2. *If $\{A_n\}_{n\in\mathbb{N}}$ is a monotonically decreasing sequence of $\mathcal{F}-$sets, then*

$$\lim_{n\to\infty} P(A_n) = P(\lim_n A_n) = P(\cap_n A_n). \tag{15}$$

Proof. Since $\{A_n\}_{n\in\mathbb{N}}$ is a monotonically decreasing sequence of $\mathcal{F}-$sets, we have $\{A_n^c\}_{n\in\mathbb{N}}$ to be a monotonically increasing sequence of $\mathcal{F}-$sets. This follows from S2.

Now from Theorem 1, we know that

$$\lim_{n\to\infty} P(A_n^c) = P(\lim_n A_n^c) = P(\cup_n A_n^c) \tag{16}$$

However, note that $\cup_n A_n^c = (\cap_n A_n)^c$. Therefore from Lemma 2 and Eq. (16), we have

$$\begin{aligned}
&\lim_{n\to\infty} 1 - P(A_n) = 1 - P(\cap_n A_n)\\
&\Leftrightarrow 1 - \lim_{n\to\infty} P(A_n) = 1 - P(\cap_n A_n)\\
&\Leftrightarrow \lim_{n\to\infty} P(A_n) = P(\cap_n A_n).
\end{aligned}$$

□

2 Random Variables

Notation: $\mathbb{R}^*$ is the extended real line, *i.e.*,

$$\mathbb{R}^* = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}. \tag{17}$$

Definition: (Borel σ-field) The smallest σ-field on $\mathbb{R}^*$ containing intervals. Recall that intervals are of the form $(a,b), [a,b], [a,b), (a,b]$, where $a,b \in \mathbb{R}^*$ and $a \le b$.

Remark 2. *The definition is indeed well-defined. Note that given a collection C of subsets of $\mathbb{R}^*$, one can ask what is the smallest σ-field containing C. We denote such a sigma field as $\sigma(C)$. Indeed, one can obtain $\sigma(C)$ as follows. Consider the new collection $\mathcal{G} := \{\mathcal{H}$ s.t. $\mathcal{H}$ is a σ-field and $C \subseteq \mathcal{H}\}$. Note that this is a collection of σ-fields. Is $\mathcal{G}$ non-empty? YES - since the power set of $\mathbb{R}^*$ itself is a σ-field and it contains C. Hence the power set belongs to $\mathcal{G}$. Now it is easy to verify that*

$$\sigma(C) = \cap_{\mathcal{H}\in\mathcal{G}} \mathcal{H}. \tag{18}$$

Definition: (Random variable) A function $X : \Omega \to \mathbb{R}^*$ is called a *random variable (r.v.)* if $X^{-1}(B) \in \mathcal{F}$, for every $B \in \mathcal{B}$. Here, $X^{-}(B)$ is defined as follows: for $B \subseteq \mathbb{R}^*$,

$$X^{-1}(B) := \{\omega \in \Omega : X(w) \in B\}. \tag{19}$$