MAT 3172 Lecture Notes

Last Updated:

May 24, 2023

Contents

Chapter 1

Probability Measures

1.1 Review of Set Theory

Let Ω be an abstract set representing the sample space of a random experiment. The power set of Ω by $\mathcal{P}(\Omega)$ is defined to be the set of all subsets of Ω. Elements of Ω are outcomes and its subsets are events. Therefore,

$$\mathcal{P}(\Omega) = \{A : A \subseteq \Omega\}$$

For $A, B \in \mathcal{P}(\Omega)$, we define

$$A \cup B = \{x : x \in A \text{ or } x \in B\}$$

$$A \cap B = \{x : x \in A \text{ and } x \in B\}$$

$$\bar{A} = A^c = \{x : x \notin A\}$$

$$A \Delta B = (A \cup B) \setminus (A \cap B)$$

In terms of events $A \cup B$ occurs if and only if at least one of the two events A and B occurs. Also, $A \cap B$ occurs if both A and B occurs. The empty set is denoted by $\emptyset$.

Examples of Sample Spaces: When flipping a coin, we have two outcomes, so

$$\Omega = \{H, T\}$$

If we flip a coin and role a dice,

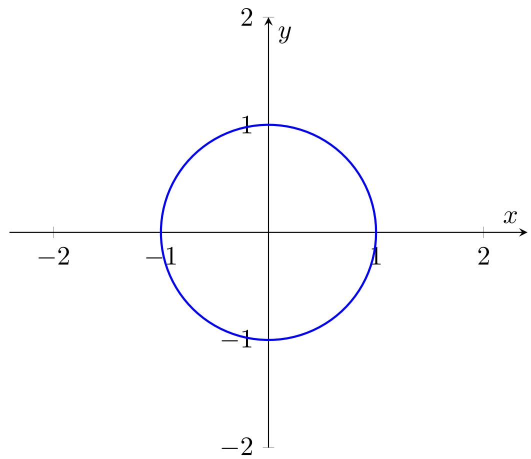
$$\Omega = \{1H, 2H, \ldots, 6H, 1T, \ldots 6T\}$$

If we flip a coin until we observe a head,

$$\Omega = \{H, TH, TTH, TTTH, \ldots\}$$

Here, there are infinite outcomes so the sample space is infinite, but it is countable since we can list all the possibilities.

If we pick a choose are sample space to be the points with distance one from the origin, we have the points in the unit circle,



The sample space is defined by

$$\Omega = \{(x, y) : d((x, y), (0, 0)) \leq 1\} = \{(x, y) : x^2 + y^2 \leq 1\}$$

In this example, the sample space omega is infinite as well, but it is uncountable.

Examples of Events: An event is a subset of the sample space. For example, in the case of rolling a dice and flipping a coin, we have

$$\Omega = \{1H, 2H, \ldots, 6H, 1T, \ldots 6T\}$$

And we can define an event E as

$$E = \{\text{Coin is heads and the dice is even}\} = \{2H, 4H, 6H\} \subset \Omega$$

If we flip a coin until the first head appears,

$$\Omega = \{H, TH, TTH, \ldots\}$$

And we can define an event E as

$$E = \{\text{First head appears before the 5th trial}\} = \{H, TH, TTH, TTTH, TTTTH\} \subset \Omega$$

Examples of Power Sets: Consider the sample space obtained by rolling a dice,

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

And let $E = \{2, 4, 6\}$ be the event we roll an even number. Then, the power set of E is

$$\mathcal{P}(E) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$$

The cardinality of the power set is

$$|\mathcal{P}(E)| = 2^{|E|} = 8$$

Examples of Set Operations:

$$\Omega = \{1, 2, \ldots, 6\}$$

$$A = \{1, 2, 3\}\ B = \{1, 2, 3\}$$

$$A \cup B = \{1, 2, 3, 4, 6\}$$

$$A \cap B = \{2\}$$

$$A^c = \{1, 3, 5\}$$

$$A \setminus B = \{x : x \in A, x \notin B\} = \{4, 6\}$$

$$A \Delta B = (A \cup B) \setminus (A \cap B)$$

Example of Empty Set: Is $\{\{\}\} = \{\}$? **No**. $\{\{\}\}$ is a set with one element, which is the empty set, so $\{\{\}\} = \{\emptyset\}$.

1.1.1 Properties of Sets

- $A \subset A$, $\emptyset \subset A$
- $A \subset B$ and $B \subset A$ implies $A = B$
- $A \subset C$ and $B \subset C$ implies $A \cup B \subset C$ and $A \cap B \subset C$.
- $A \subset B$ if and only if $B^c \subset A^c$
- $(A^c)^c = A$, ${}^c = \Omega$, $\Omega^c = \emptyset$
- $A \cup B = B \cup A$, $A \cap B = B \cap A$
- $A \cup A = A$, $A \cap \Omega = A$, $A \cup A^c = \Omega$, $A \cap A^c = \emptyset$.
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \cup (A \cap C)$
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- $(A \cup B)^c = A^c \cap B^c$, $(A \cap B)^c = A^c \cup B^c$

Example. We have

$$\bigcup_{n=1}^{\infty}\left[0,\frac{n}{n+1}\right) = [0,1)$$

To show that these sets are equal, consider the limit of the sequence $\left(\frac{n}{n+1}\right)_{n=1}^{\infty}$. As n becomes large, the limit approaches 1. So, the union of all these sets will contain elements that become arbitrarly close to 1 but do not reach 1, so we have $[0,1)$.

$$\bigcap_{n=1}^{\infty}\left(0,\frac{1}{n}\right) = \emptyset$$

To show that these sets are equal, consider the sequence $\left(\frac{1}{n}\right)_{n=1}^{\infty}$, As this sequence approaches 0, the intersection of all these sets will contain elements that become arbitrarly close to 0 but do not reach 0, so we have the set $(0,0)$ which is empty. Therefore, when taking the intersection of all these sents with $\left(0,\frac{1}{n}\right)$ which become arbitrarly small, we have the empty set.

Example. Prove that $A\Delta B = A^c \Delta B^c$.

Proof. Note that $A \setminus B = A \cap B^c$.

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

□

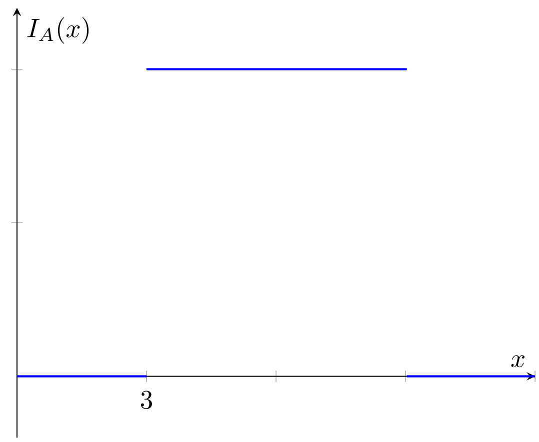
1.2 Indicator Function

Let $A \subset \Omega$. The indicator function of A is defined as

$$I(x \in A) = I_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

Example. $A = [1,3]$

$$I_A(x) = I_{[1,3]}(x)$$



1.2.1 Properties of Indicator Functions:

- $I_{A \cup B} = \max(I_A, I_B)$
- $I_{A \cap B} = I_A \cdot I_B$
- $I_{A \Delta B} = I_A + I_B \pmod 2$
- $A \subset B$ if and only if $I_A \leq I_B$
- $I_{\cup_i A_i} \leq \sum_i I_{A_i}$

1.3 Set Theoretic Limits

Exercise: Prove that

$$I_{\cup_{i=1}^{\infty} A_i} = 1 - \prod_{i=1}^{\infty}(1 - I_{A_i})$$

and

$$I_{A \Delta B} = (I_A - I_B)^2$$

Definition 1.3.1. *Let* $\{A_n\}$ *be a sequence of events. Then*

$$\liminf A_n = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m$$

and

$$\limsup A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$$

Example. Suppose we have a sequence of events $A_1, A_2, A_3, \ldots$, we can define

$$B_n = \bigcap_{m=n}^{\infty} A_m$$

So its sequence members are

$$B_n = A_n \cap A_{n+1} \cap \cdots$$
$$B_{n+1} = A_{n+1} \cap A_{n+2} \cap \cdots$$
$$B_{n+2} = A_{n+2} \cap A_{n+3} \cap \cdots$$

These sets are getting smaller because surely the intersection of more sets will be smaller since $|A \cap B| \leq \min(|A|, |B|)$. So as we take more intersections, the sets become smaller and smaller. Now we can look at the union of these sets,

$$\bigcup_{n=1}^{\infty} B_n = \liminf A_n$$

If instead we take B_n to be the union of all the sets,

$$B_n = \bigcup_{m=n}^{\infty} A_m$$

So the sequence members are

$$B_n = A_n \cup A_{n+1} \cup \cdots$$
$$B_{n+1} = A_{n+1} \cup A_{n+2} \cup \cdots$$
$$B_{n+2} = A_{n+2} \cup A_{n+3} \cup \cdots$$

These sets are getting larger since we are taking the union of more sets, then we can look at the intersection of these sets,

$$\bigcap_{m=n}^{\infty} = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_n = \limsup A_n$$

Lemma 1.3.1. *We have*

$$\limsup A_n = \left\{ \omega : \sum_{i=1}^{\infty} I_{A_i}(\omega) = \infty \right\}$$

and

$$\liminf A_n = \left\{ \omega : \sum_{i=1}^{\infty} I_{A_i^c}(\omega) < \infty \right\}$$

We can also express this as

$$\limsup_{n\to\infty} A_n = \left\{ x \in X : \limsup_{n\to\infty} I_{A_n}(x) = 1 \right\}$$

and

$$\liminf_{n\to\infty} A_n = \left\{ x \in X : \liminf_{n\to\infty} I_{A_n}(x) = 1 \right\}$$

Proof. If $\omega \in \limsup A_n$, then $\omega \in \bigcup_{m=n}^{\infty} A_m$ for all integers n. Therefore, for any integer n there exists an integer k_n such that $\omega \in A_{k_n}$, since

$$\sum_{i=1}^{\infty} A_{A_i}(\omega) \geq \sum_{i=1}^{\infty} I_{A_{k_i}}(\omega) = \infty$$

Conversely, for any integer n, by definition of the limit superior,

$$\sum_{i=n}^{\infty} I_{A_i}(\omega) = \infty$$

This implies that $\omega \in \bigcup_{j=n}^{\infty} A_j$ for all integers n. Therefore,

$$\omega \in \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m = \limsup A_n$$

Then, we can notice that

$$\omega \in \liminf A_n = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m$$

implies that there exists an integer n_0 such that

$$\omega \in \bigcap_{k=n_0}^{\infty} A_k$$

Therefore,

$$\sum_{n=1}^{\infty} I_{A_n^c}(\omega) = \sum_{n=1}^{n_0-1} I_{A_n^c}(\omega) \leq n_0 < \infty$$

$\square$

Note: For this reason, sometimes we write $\limsup A_n = A_n$ infinitely often. If $\liminf A_n = \limsup A_n$, then

$$\lim A_n = \liminf A_n = \limsup A_n$$

Remark: The proof of the lemma above can be simplified by noticing the fact that

$$(\limsup A_n)^c = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m^c = \liminf A_n^c$$

$$(\liminf A_n)^c = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m^c = \limsup A_n^c$$

To summarize, with the limit superior we have infinitely many cases where $I_{A_i}(\omega) = 1$. So what it means for $\omega \in \limsup A_n$ is that ω is in infinitely many of the A_i's. For the limit inferior, it means that ω is in all but finitely many of the A_i's.

Example. Consider $\omega \in A_i$ when i is odd, so

$$\omega \in A_1, \omega \notin A_2, \omega \in A_3, \omega \notin A_4, \omega \in A_5, \ldots$$

ω is in infinitely many (but countable) number of the A_i's. So $\omega \in \limsup A_n$. Now consider $\omega \in A_i$ when $i \geq 10$, so

$$\omega \notin A_1, \omega \notin A_2, \omega \notin A_3, \ldots, \omega \notin A_9, \omega \in A_{10}, \omega \in A_{11}, \omega \in A_{12}, \ldots$$

So, ω is not in finitely many of the A_i's, therefore $\omega \in \liminf A_n$.

Example. Consider sample space of flipping a coin and infinite number of times, and the event $E = \{HTTHT\}$, so the event that we get $HTTHT$ in that order. Because this outcome is possible, it will occur an infinite number of times in the sequence of events, so E is in the limit superior, and the probability that E occurs infinitely often is 1.

Lemma 1.3.2. *Let $\{A_n\}$ be a sequence of events, then*

1. *If $A_n \subset A_{n+1}$ for any integer n, then*

$$\lim A_n = \bigcup_{n=1}^{\infty} A_n$$

2. *If $A_{n+1} \subset A_n$ for any integer n, then*

$$\lim A_n = \bigcap_{n=1}^{\infty} A_n$$

Proof. We can prove (1) and (2) similarly as follows, not that in this case,

$$\bigcap_{m=n}^{\infty} A_m = \bigcap_{m=1}^{\infty} A_m$$

for all integers n. If $A_n \subset A_{n+1}$ for any integer n, then we have that

$$A_1 \subset A_2 \subset A_3 \subset \cdots$$

So the set is getting bigger, now consider the limit superior

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m = \left(\bigcup_{m=1}^{\infty} A_m \right) \cap \left(\bigcup_{m=2}^{\infty} A_m \right) \cap \left(\bigcup_{m=3}^{\infty} A_m \right) \cap \cdots$$

These sets are equal in size since if $A_1 \subset A_2$, then $A_1 \cup A_2 = A_2$. Therefore, we get that the intersection of these sets is $\bigcup\limits_{m=1}^{\infty} A_m$, and thus

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m = \bigcup_{n=1}^{\infty} A_n$$

Furthermore,

$$\begin{aligned}
\liminf_{n\to\infty} A_n &= \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m \\
&= \left(\bigcap_{m=1}^{\infty} A_m\right) \cup \left(\bigcap_{m=2}^{\infty} A_m\right) \cup \left(\bigcap_{m=3}^{\infty} A_m\right) \cup \cdots \\
&= A_1 \cup A_2 \cup A_3 \cup \cdots \\
&= \bigcup_{n=1}^{\infty} A_n
\end{aligned}$$

Therefore

$$\limsup A_n = \liminf A_n \implies \limsup A_n = \liminf A_n = \lim A_n$$

The proof for (2) follows the same. $\square$

Example.

$$\lim_{n\to\infty} \left[0, 1 - \frac{1}{n}\right] = \lim_{n\to\infty} \left[0, 1 - \frac{1}{n}\right) = [0, 1)$$

To see this, we have that

$$A_1 = \{0\}, A_2 = \left[0, \frac{1}{2}\right], A_3 = \left[0, \frac{2}{3}\right], A_4 = \left[0, \frac{3}{4}\right], \ldots$$

So the set A_n is increasing, therefore

$$\limsup A_n = \liminf A_n = \lim A_n = \bigcup_{n=1}^{\infty} A_n = [0, 1]$$

Example.

$$\lim_{n\to\infty} \left[0, 1 + \frac{1}{n}\right] = \lim_{n\to\infty} \left[0, 1 + \frac{1}{n}\right) = [0, 1)$$

Similarly,

$$A_1 = [0, 2], A_2 = \left[0, 1 + \frac{1}{2}\right], A_3 = \left[0, 1 + \frac{1}{3}\right], A_4 = \left[0, 1 + \frac{1}{4}\right], \ldots$$

The set A_n is decreasing, therefore

$$\limsup A_n = \liminf A_n = \bigcap_{n=1}^{\infty} A_n = \bigcap_{n=1}^{\infty} \left[0, 1 + \frac{1}{n}\right] = [0, 1]$$

Example. Let $B, C \subset \Omega$ and define the sequence

$$A_n = \begin{cases} B & \text{if } n \text{ is odd} \\ C & \text{if } n \text{ is even} \end{cases}$$

Then we have

$$\bigcup_{m=n}^{\infty} = B \cup C \implies \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m = B \cup C = \limsup A_n$$

Similarly for the limit inferior,

$$\bigcap_{m=n}^{\infty} A_m = B \cap C \implies \bigcup_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m = B \cup C = \liminf A_n$$

Therefore, we have

$$\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m = B \cup C \text{ and } \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m = B \cap C$$

If $B \cap C \neq B \cup C$, then $B \cap C = \liminf A_n \neq \limsup A_n = B \cup C$.

1.4 Fields and Algebras

Definition 1.4.1 (Fields (Algebras))**.** *A* field *(or* algebra*) is a class of subsets of Ω (called events) that contain Ω and are closed under finite union, finite intersection, and complementation. In otherwords, a family of subsets of Ω (say $\mathcal{A}$) is a field if*

- $\Omega \in \mathcal{A}$
- *If $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$*
- *If $A, B \in \mathcal{A}$, then $A \cup B \in \mathcal{A}$*

Remarks. If $A, B \in \mathcal{A}$ then $A \cap B \in \mathcal{A}$. This is true because

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

Definition 1.4.2 (σ-field)**.** *A σ-field (or σ-algebra) is a field that is closed under countable union (which implies that it is closed under countable intersection).*

Example. Let Ω be a set and $A, B \subset \Omega$. Then,

$$\mathcal{A} = \{\Omega, \emptyset, A, A^c, B, B^c, A \cup B, A \cap B, A \cap B^c, A^c \cap B, A^c \cup B^c, A^c \cap B^c\}$$

Examples of σ-fields.

- The power set $\mathcal{P}(\Omega)$
- $\mathcal{F} = \{\Omega, \emptyset\}$
- The family of subsets or $\mathbb{R}$ which are either countable or their complements are countable.
- Let $\mathcal{B}$ be the smallest σ-field containing all open sets. Then $\mathcal{B}$ is called the Borel σ-field.

Definition 1.4.3 (Probability Measure)**.** *Let Ω be a sample space and $\mathcal{F}$ be a σ-field on Ω. A probability measure P is defined on $\mathcal{F}$ such that*

(i) $P(\Omega) = 1$

(ii) If $A_1, A_2, \ldots \in \mathcal{F}$ are disjoint, then

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

1.4.1 Properties of Probability Measures

(i) Since $P(\Omega) = 1 = P(\Omega \cup \emptyset) = P(\Omega) + P(\emptyset)$, we have $P(\emptyset) = 0$

(ii) Since $(A \setminus B) \cup (A \cap B) = A$ and $(A \setminus B) \cap (A \cap B) = \emptyset$, we have

$$P(A \setminus B) = P(A) - P(A \cap B)$$

(iii) Similarly, $(A \setminus B) \cup B = A \cup B$ and $(A \setminus B) \cap B = \emptyset$, which implies

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

(iv) If $A \subset B$ then $A \cup (B \setminus A) = B$. Therefore,

$$P(A) + P(B \setminus A) = P(B)$$

and furthermore, $P(A) \leq P(B)$

1.4.2 Expectations

Definition 1.4.4. *Let $X : \Omega \to \mathbb{R}$, an expecation E be an operator with the following properties,*

(i) *If* $X \geq 0$ *then* $E(X) \geq 0$

(ii) *If* $c \in \mathbb{R}$ *is a constant, then* $E(cX) = cE(X)$

(iii) $E(X_1 + X_2) = E(X_1) + E(X_2)$

(iv) $E(1) = 1$

(v) *If* $X_n(\omega)$ *is monotonically increasing and* $X_n(\omega) \to X(\omega)$*, then*

$$\lim_{n\to\infty} E(X_n) = E(X)$$

Example. Flip 2 coins,

$$\Omega = \{HH, HT, TH, TT\}$$

Define $X : \Omega \to \mathbb{R}$, with

$$X(HH) = 2, X(HT) = X(TH) = 1, X(TT) = 0$$

So X is a random variable which represents the number of heads.

Example. Flip a coin until a head appears,

$$\Omega = \{H, TH, TTH, TTTH, \ldots\}$$

With

$$X(H) = 1, X(TH) = 2, X(TTH) = 3, X(TTTH) = 4, \ldots$$

So X is a random variable which represents the number of trials until a head appears.

Example. Define

$$\Omega = \{(x, y) : x^2 + y^2 \leq 1\}$$

and

$$X(x, y) = \sqrt{x^2 + y^2} = \text{ Distance of } (x, y) \text{ from } (0, 0)$$

Example. Let X be a random variable, and

$$E(X) = \lim_{D\to\infty} \frac{\int_{-D}^{D} X(\omega)d\omega}{2D}$$

Check if E satisfies the definition for an expectation.

Solution. We have to check the 5 axioms,

1. Its clear that if $X \geq 0$ then $E(X) \geq 0$ since the integral of a non-negative function is non-negative.

2. $$E(cX) = \lim_{D\to\infty} \frac{\int_{-D}^{D} cX(\omega)d\omega}{2D} = c \lim_{D\to\infty} \frac{\int_{-D}^{D} X(\omega)d\omega}{2D} = cE(x)$$

3. $$\begin{aligned} E(X_1 + X_2) &= \lim_{D\to\infty} \frac{\int_{-D}^{D}(X_1(\omega) + X_2(\omega))d\omega}{2D} \\ &= \lim_{D\to\infty} \left(\frac{\int_{-D}^{D} X_1(\omega)d\omega}{2D} + \frac{\int_{-D}^{D} X_2(\omega)d\omega}{2D} \right) \\ &= \lim_{D\to\infty} \frac{\int_{-D}^{D} X_1(\omega)d\omega}{2D} + \lim_{D\to\infty} \frac{\int_{-D}^{D} X_2(\omega)d\omega}{2D} \\ &= E(X_1) + E(X_2) \end{aligned}$$

4. $$E(1) = \lim_{D\to\infty} \frac{\int_{-D}^{D} 1d\omega}{2D} = \lim_{D\to\infty} \frac{2D}{2D} = 1$$

5. The 5th axiom fails however. Take $\Omega = \mathbb{R}$ and $X_n(\omega) = I_{[-n,-n]}(\omega)$, then

$$\lim_{D\to\infty} \frac{\int_{-D}^{D} X(\omega)d\omega}{2D} = 0$$

But, $x_n(\omega) \to 1$. So the operator in this example is not a proper form of expectation.

1.5 Finding Probabilities Using Expectations

Definition 1.5.1. *For any event A, define*

$$P(A) = E(I_A(\omega))$$

For simplicity, we somtimes drop ω and write

$$P(A) = E(I_A)$$

1.5.1 Properties

1. $E\left(\sum\limits_{i=1}^{n} c_i X_i\right) = \sum\limits_{i=1}^{n} c_i E(X_i)$

2. If $X \leq Y \leq Z$, then $E(X) \leq E(Y) \leq E(Z)$

3. If $\{A_i\}$ is a sequence of events, then

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

The third property (known as Borel's inequality) is an important result and can be proved as follows.

Proof. Using the fact that

$$I_{\cup A_i} \leq \sum_{i=1}^{\infty} I_{A_i}$$

and the second property where if $X \leq Y$, then $E(X) \leq E(Y)$, then

$$E\left(I_{\cup A_i}\right) \leq \sum_{i=1}^{\infty} E(I_{A_i})$$

Then this is exactly

$$P(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i)$$

□

We can use these properties to prove the axioms of probability measures.

1. $P(\Omega) = E(I_\Omega)$, we have $I_\Omega(\omega) = 1 \ \forall \omega \in \Omega$, so $P(\Omega) = E(I_\Omega) = 1$
2. We want to show that the probability of disjoint events is the sum of their probablities, so we have that

   $$P\left(\bigcup_{i=1}^{\infty} A_i\right) = E\left(I_{\cup A_i}\right)$$

   We know from indicator functions that

   $$I_{\cup A_i} = \sum_{i=1}^{\infty} I_{A_i}$$

   Therefore,

   $$E\left(I_{\cup A_i}\right) = E\left(\sum_{i=1}^{\infty} I_{A_i}\right) = \sum_{i=1}^{\infty} E(I_{A_i}) = \sum_{i=1}^{\infty} P(A_i)$$

   as required.

Another useful result we can prove is that

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

and furthermore

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

Proof. We know that

$$I_{A\cup B\cup C} = 1 - (1 - I_A)(1 - I_B)(1 - I_C)$$

This can be easily shown by considering the cases when we have $\omega \in A\cup B\cup C$, then at least one of the terms $(1 - I_A)$, $(1 - I_B)$, or $(1 - I_c)$ will be zero, then

$$1 - (1 - I_A)(1 - I_B)(1 - I_C) = 1 - 0 = 1 = I_{A\cup B\cup C}$$

And if $\omega \notin A \cup B \cup C$, then $\omega \notin A$ and $\omega \notin B$ and $\omega \notin C$, so all their indicator functions will be zero, then we get

$$1 - (1 - 0)(1 - 0)(1 - 0) = 0 = I_{A\cup B\cup C}$$

So, we have the result that

$$I_{\cup A_i} = 1 - \prod_{i=1}^{\infty}(1 - A_i)$$

Now if we expand the equality, we get

$$\begin{aligned}
I_{A\cup B\cup C} &= 1 - (1 - I_A)(1 - I_B)(1 - I_C)\\
&= 1 - (1 - I_A - I_B + I_A I_B)(1 - I_C)\\
&= 1 - 1 + I_A + I_B + I_C - I_A I_B - I_A I_C - I_B I_C + I_A I_B I_C\\
&= I_A + I_B + I_C - I_A I_B - I_A I_C - I_B I_C + I_A I_B I_C
\end{aligned}$$

Then, we can take the expected value of both sides,

$$\begin{aligned}
E(I_{A\cup B\cup C}) &= E(I_A + I_B + I_C - I_A I_B - I_A I_C - I_B I_C + I_A I_B I_C)\\
&= E(I_A) + E(I_B) + E(I_C) - E(I_A I_B) - E(I_A I_C) - E(I_B I_C) + E(I_A I_B I_C)\\
&= P(A) + P(B) + P(C) - P(A\cap B) - P(A\cap C) - P(B\cap C) + P(A\cap B\cap C)\\
&= P(A\cup B\cup C)
\end{aligned}$$

Therefore,

$$P(A\cup B\cup C) = P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-P(B\cap C)+P(A\cap B\cap C)$$

Showing that $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ follows the same. $\square$

Then we get that the general case for the union of n events as

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{1\leq i<j\leq n} P(A_i\cap A_j) + \sum_{i\leq i<j<k\leq n} P(A_i\cap A_j\cap A_k) + (-1)^{n+1} P(A_1\cap A_2\cap \cdots \cap A_n)$$

And this is again proved using the fact that

$$I_{\cup_{i=1}^{n} A_i} = 1 - \prod_{i=1}^{n}(1 - A_i)$$

Example. (Confused Secretary Problem) Suppose we have 100 distinct letters to be sent to 100 different people. The Secretary confuses the addresses, and sends 100 letters at random to these 100 people. What is the probability that at least one letter is sent to the correct address?

Solution. Let A_i denote the event that the ith letter goes to the right person. We want to find the probability that one of these events occurs, so $P\left(\bigcup_{i=1}^{100} A_i\right)$. From the formulas previously discussed,

$$\begin{aligned}
P\left(\bigcup_{i=1}^{100} A_i\right) &= P(A_1) + \cdots P(A_100) - P(A_1 \cap A_2) - \cdots - P(A_{99} \cap A_{100}) \\
&+ P(A_1 \cap A_2 \cap A_3) + \cdots + P(A_{98} \cap A_{99} \cap A_{100}) \\
&+ P(A_1 \cap A_2 \cap A_3 \cap A_4) - \cdots - P(A_{97} \cap A_{98} \cap A_{99} \cap A_{100}) \\
&+ \cdots - P(A_1 \cap A_2 \cap \cdots \cap A_{100})
\end{aligned}$$

Then we can evaluate each probability,

$$P(A_1) = \frac{1}{100},\ P(A_2) = \frac{1}{100}, \ldots, P(A_{100}) = \frac{1}{100}$$

$$P(A_1 \cap A_2) = \frac{1}{100} \cdot \frac{1}{99} = \cdots = P(A_{99} \cap A_{100})$$

$$P(A_1 \cap A_2 \cap A_3) = \frac{1}{100} \cdot \frac{1}{99} \cdot \frac{1}{98} = \cdots$$

Then continuing this sequence we have

$$\begin{aligned}
P\left(\bigcup_{i=1}^{100}\right) &= 100\left(\frac{1}{100}\right) - \binom{100}{2}\left(\frac{1}{100 \cdot 99}\right) + \binom{100}{3}\left(\frac{1}{100 \cdot 99 \cdot 98}\right) - \cdots - \binom{100}{100}\left(\frac{1}{100 \cdot 99 \cdots 1}\right) \\
&= 1 - \frac{100 \cdot 99}{2!} \cdot \frac{1}{100 \cdot 99} + \frac{100 \cdot 99 \cdot 98}{3!} \cdot \frac{1}{100 \cdot 99 \cdot 98} - \cdots - \frac{1}{100!} \\
&= 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \frac{1}{5!} - \cdots - \frac{1}{100!}
\end{aligned}$$

Now recall that e^x can be written as the infinite series

$$e^x = \sum_{n=1}^{\infty} \frac{x^n}{n!}$$

So,

$$\begin{aligned}
e^{-1} &= 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots \\
1 - e^{-1} &= 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots
\end{aligned}$$

Now since 100! is incredibly large, it is sufficiently large for this infinite sum. So we get that

$$P(\text{At least one letter goes to the right address}) = 1 - \frac{1}{e}$$

Lemma 1.5.1 (Fatou's Lemma)**.** *If $\{A_n\}$ is a family of events, then*

1. $P(\liminf A_n) \leq \liminf P(A_n) \leq \limsup P(A_n) \leq P(\limsup A_n)$
2. *If* $\lim A_n = A$*, then* $\lim P(A_n) = P(A)$

To prove this lemma, we first need to prove the following lemma.

Lemma 1.5.2. *If $A_n \subset A_{n+1}$ for any $n \in \mathbb{N}$, then $\lim P(A_n) = P(A)$. Similarly, if $A_{n+1} \subset A_n$ for any $n \in \mathbb{N}$, then $\lim P(A_n) = P(A)$ where in both cases $\lim A_n = A$.*

Proof. If A_n is increasing (i.e $A_n \subset A_{n+1}$), start by defining $B_n = A_n \setminus A_{n-1}$ with $B_1 = A_1$. Then,

$$\bigcup_{i=1}^{n} B_i = \bigcup_{i=1}^{n} A_i = A_n$$

The point of this is now we have that the B_i's are disjoint. Now if we take the limit we get

$$\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i = A$$

Taking the probability now,

$$P(A) = P\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} P(B_i) = \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) = \lim_{n\to\infty} P(A_n)$$

If A_n is decreasing ($A_{n+1} \subset A_n$), then A_n^c is increasing. So,

$$\lim_{n\to\infty} P(A_n^c) = \lim(1 - P(A_n)) = P(A^c) = 1 - P(A)$$

Therefore,

$$\lim_{n\to\infty} P(A_n) = A$$

$\square$

To summarize, when A_n is increasing, we have

$$\lim P(A_n) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = P\left(\lim_{n\to\infty} A_n\right)$$

when A_n is decreasing,

$$\lim P(A_n) = P\left(\bigcap_{i=1}^{\infty} A_i\right) = P\left(\lim_{n\to\infty} A_n\right)$$

Now we can prove Fatou's Lemma.

Proof. To prove (1), notice that from the first part of lemma 1.5.2, we can write

$$P(\liminf A_n) = P\left(\lim_{n\to\infty} \cap_{i=n}^{\infty} A_i\right) = \lim_{n\to\infty} P\left(\cap_{i=n}^{\infty} A_i\right) \leq \liminf P(A_n)$$

since $\bigcap\limits_{i=n}^{\infty} A_i \subset A_n$. Likewise,

$$P(\limsup A_n) = P\left(\lim_{n\to\infty} \bigcup_{i=n}^{\infty} A_i\right) = \lim_{n\to\infty} P\left(\bigcup_{i=1}^{\infty} A_i\right) \geq \limsup P(A_n)$$

since $A_n \subset \bigcup\limits_{i=n}^{\infty} A_i$. For part (2), notice that if

$$A = \limsup A_n = \liminf A_n$$

we have

$$P(A) = P(\liminf A_n) \leq \liminf P(A_n) \leq \limsup P(A_n) \leq P(\limsup A_n) = P(A)$$

Then this implies that

$$\lim_{n\to\infty} P(A_n) = P(A)$$

□

1.6 Independence

Definition 1.6.1. *Let A and B be events. We say that A and B are independent if*

$$P(A \cap B) = P(A)P(B)$$

Example. Flip a coin and roll a die,

$$\Omega = \{1H, 2H, \ldots, 6H, 1T, 2T, \ldots, 6T\}$$

Let $A = \{1H, 1T\}$ and $B = \{1H, 2H, 3H, 4H, 5H, 6H\}$. Are A, B independent?

$$A \cap B = \{1H\} \neq \emptyset$$

$$P(A \cap B) = \frac{1}{12}$$

$$P(A) = \frac{1}{6},\ P(B) = \frac{1}{2}$$

$$P(A)P(B) = \frac{1}{6} \cdot \frac{1}{2} = \frac{1}{12} = P(A \cap B)$$

Therefore, A and B are independent.

1.6.1 Properties of Independence

1. If A and B are independent, then A and B^c are independent, A^c and B are independent, and A^c and B^c are independent.
2. If A, B, C are indepdent then A and $B \cup C$ are independent. similarly A and $B \cap C$ are independent.
3. An event A is independent of itself if and only if $A = \emptyset$ or $A = \Omega$.
4. Any event A is independent of Ω.

Proof. 1. To prove that A^c and B are independent, recall that

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

So,

$$\begin{aligned} P(A^c \cap B) &= P(B) - P(A \cap B) \\ &= P(B) - P(A)P(B) \\ &= P(B)(1 - P(A)) \\ &= P(B)P(A^c) \end{aligned}$$

The proof for A and B^c follows the same. To prove A^c and B^c are independent,

$$\begin{aligned} P(A^c \cap B^c) &= P((A \cup B)^c) \\ &= 1 - P(A \cup B) \\ &= 1 - [P(A) + P(B) - P(A \cap B)] \\ &= 1 - P(A) - P(B) + P(A)P(B) \\ &= 1 - P(A) - P(B)(1 - P(A)) \\ &= (1 - P(A))(1 - P(B)) \\ &= P(A^c)P(B^c) \end{aligned}$$

2. Given 3 independent events A, B, C, then any operations between the sets is independent. So we can show A and $B \cup C$ is independent since

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

Similarly for A and $B \cap C$,

$$\begin{aligned} P(A \cap (B \cap C)) &= P((A \cap B) \cap (A \cap C)) \\ &= P(A \cap B \cap A \cap C) \\ &= P(A)P(B)P(C) \\ &= P(A)P(B \cap C) \end{aligned}$$

3. ($\Longrightarrow$) If A is independent of itself, then

   $P(A \cap A) = P(A)P(A) \implies P(A) = P(A)^2 \implies P(A) = 0, 1 \implies A = \emptyset, A = \Omega$

   ($\Longleftarrow$) If $A = \emptyset$, then

   $$P(A \cap A) = P(A) = 0 = P(\emptyset)P(\emptyset)$$

   If $A = \Omega$, then

   $$P(A \cap A) = P(A) = 1 = P(\Omega)P(\Omega)$$

4. Every event is independent of Ω since

   $$P(A \cap \Omega) = P(A)P(\Omega) = P(A) \cdot 1$$

□

Lemma 1.6.1 (Borel Cantelli Lemma)**.** *Let* $(\Omega, \mathcal{F}, P)$ *be a probability space and let* $\{E_i\}$ *be a sequence of events. Then,*

(i) If $\sum\limits_{i=1}^{\infty} P(E_i) < \infty$*, then* $P(\limsup E_n) = 0$

(ii) If $\{E_i\}$ *is a sequence of independent events, then* $P(\limsup E_n) = 0$ *or 1 according to to whether the series* $\sum\limits_{i=1}^{\infty} P(E_i)$ *diverges or converges respectively.*

Proof. Set $E = \limsup E_n$. We have $E = \bigcap\limits_{n=1}^{\infty} F_n$ where $F_n = \bigcup\limits_{m=n}^{\infty} E_m$. For every positive integer n,

$$0 \leq P(F_n) \leq \sum_{m=n}^{\infty} P(E_m)$$

Since $\sum\limits_{i=1}^{\infty} P(E_i) < \infty$, then $\lim\limits_{n\to\infty} P(E_n) = 0$. Since $F_n \downarrow E$, from Fatou's lemma we can write

$$0 = \lim_{n\to\infty} P(F_n) = P\left(\lim_{n\to\infty} F_n\right) = P(\limsup E_n)$$

Thus (i) is proved. For (ii), suppose $E_1, E_2, \ldots$ are independent. From (ii), we know that $P(\limsup E_n) = 0$ if the sum $\sum_{n=1}^{\infty} E_n$ is finite. It remains to show that if the sum is infinite, then E_n occurs infinitely often. Let $E = \limsup E_n$. Then

$$E^c = \liminf E_n^c$$

The sequence of events $\{E_n^c\}$ are also independent, so we have

$$P\left(\bigcap_{m=n}^{\infty} E_m^c\right) \le P\left(\bigcap_{m=n}^{N} E_m^c\right) = \prod_{m=n}^{N}(1 - P(E_m)) \le \exp\left(-\sum_{m=n}^{N} P(E_m)\right)$$

This inequality comes from the Talor Series expansion of e^x, so

$$e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots$$
$$e^{-x} \ge 1 - x$$

As $N \to \infty$, we get $\bigcap\limits_{m=n}^{N} E_m^C \downarrow E^c$. Thus,

$$P\left(\bigcap_{m=n}^{N} E_m^c\right) \le \exp\left(-\sum_{m=n}^{N} P(E_m)\right) \le \exp\left(-\sum_{m=n}^{N} P(E_m)\right) \to 0$$

This implies that $P(E^c) = 0$, thus $P(E) = 1$. $\square$

Corollary 1.6.1. *If $\{E_i\}$ is a sequence of independent events then*

$$P(\limsup E_n) = 0 \iff \sum_{i=1}^{\infty} P(E_i) < \infty$$

Proof. ($\Longleftarrow$) We know from the Borel Cantelli lemma that if $\sum\limits_{i=1}^{\infty} P(E_i) < \infty$, then $P(\limsup E_n) = 0$.

($\Longrightarrow$) Suppose $P(\limsup E_n) = 0$. Then from part (2), $P(\limsup E_n) = 1$ when $\sum\limits_{i=1}^{\infty} P(E_i) = \infty$. So it must be that $\sum\limits_{i=1}^{\infty} P(E_i) < \infty$ as required. $\square$

Remark. Notice that indepdence is required by Corollary 1.6.1. To see this, let $(\Omega = [0,1], B, P)$ be a probability space with

$$P(A) = \int_A dx$$

for a Borel set A. It's easy to show that P is a probability measure on $[0,1]$. Now define $E = \left(0, \frac{1}{n}\right)$ and notice that $E_n \downarrow \emptyset$. Therefore,

$$P(\limsup E_n) = P(\lim E_n) = 0$$

Since

$$P(E_n) = \int_0^{\frac{1}{n}} dx = \frac{1}{n}$$

we have

$$\sum_{n=1}^{\infty} P(E_n) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

This does not violate Corollary 1.6.1 since the events E_n are not independent. For example. consider $E_2 = (0, 1/2)$, and $E_3 = (0, 1/3)$. Then

$$P(E_2) = \frac{1}{2}, \quad P(E_3) = \frac{1}{3}$$

$$P(E_2 \cap E_3) = P(E_3) = \frac{1}{3} \neq P(E_2)P(E_3) = \frac{1}{6}$$