PROBABILITY AND STOCHASTIC PROCESS (MTL106)

ANANTA KUMAR MAJEE

1. Probability Theory:

Consider the experiment of tossing a coin. There are two possible outcome $\{H, T\}$ and we don't know whether H or T will come in advance for any performance of this experiment. From this example, we arrive at the following:

Definition 1.1 (**Random Experiment**)**.** A random experiment is an experiment in which

i) all outcomes of the experiment are known in advance
ii) any performance of the experiment results in an outcome that is not known in advance
iii) the experiment can be repeated under identical conditions.

Definition 1.2 (**Sample space**)**.** The set Ω of all possible outcomes of a random experiment is called sample space. An element $\omega \in \Omega$ is called a sample point.

A die is thrown once. We are interested in the possible occurrences of the 'events'

a) the outcome is the number 1
b) the outcome is even number
c) the outcome is even but does not exceed 3
d) the outcome is not even

Each of the above events can be specified as a subset A of the sample space $\Omega = \{1, 2, 3, 4, 5, 6\}$ as follows:

$$a)\;\; A = \{1\}, \quad b)\;\; A = \{2, 4, 6\}, \quad c)\;\; A = \{2, 4, 6\} \cap \{1, 2, 3\}, \quad d)\;\; A = \{2, 4, 6\}^{\complement}$$

Thus, events are subsets of Ω, but need all the subsets of Ω be events? The answer is NO. We need the collection of events be closed under *complement* and countable union.

Definition 1.3 (**σ-algebra**)**.** Let $\Omega \neq \emptyset$. A collection $\mathcal{F}$ of subsets of Ω is called a σ-algebra or σ-field over Ω if the following properties hold:

i) $\emptyset \in \mathcal{F}$
ii) If $A_1, A_2, \ldots \in \mathcal{F}$, then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$
iii) If $A \in \mathcal{F}$, then $A^{\complement}$.

The elements of $\mathcal{F}$ are called 'events'. One can easily show that if $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$.

Example 1.1. *i) The smallest σ-field associated to Ω is the collection $\mathcal{F} = \{\emptyset, \Omega\}$.*
ii) If A is any subset of Ω, then $\mathcal{F} = \{\emptyset, A, A^{\complement}, \Omega\}$ is a σ-field.
ii) $\mathcal{P}(\Omega) := \{A : A \subseteq \Omega\}$ is a σ-field, called the total σ-field over Ω.

Remark 1.1. In general, union of σ-algebras over Ω may not be a σ-algebra. For example, let $\Omega = \{1, 2, 3\}$ and $\mathcal{F}_1 = \{\emptyset, \{1\}, \{2, 3\}, \Omega\}$ and $\mathcal{F}_2 = \{\emptyset, \{1, 2\}, \{3\}, \Omega\}$. Then both $\mathcal{F}_1$ and $\mathcal{F}_2$ are σ-algebra, but $\mathcal{F}_1 \cup \mathcal{F}_2$ is NOT a σ-algebra.

Lemma 1.1. *Intersection of two σ-fields over Ω is a σ-field over Ω. More generally, if $\{\mathcal{F}_i : i \in I\}$ is a family of σ-fields over Ω, then $\mathcal{G} = \cap_{i \in I} \mathcal{F}_i$ is also a σ-field.*

Definition 1.4 (**Generated σ-field).** Let $\Omega \neq \emptyset$ and A be be collection of subsets of Ω. Let

$$\mathcal{M} := \{\mathcal{F} : \mathcal{F} \text{ is a } \sigma\text{-field over } \Omega \text{ containing } A\}.$$

Then $\sigma(A) := \cap_{\mathcal{F}\in\mathcal{M}}\mathcal{F}$ is the smallest σ-field over Ω containing A. $\sigma(A)$ is called the σ-field generated by A.

Example 1.2 (**Borel σ-field).** *The smallest σ-algebra over $\mathbb{R}$ containing all intervals of the form $(-\infty, a]$ with $a \in \mathbb{R}$ is called the Borel σ-algebra and denoted by $\mathcal{B}(\mathbb{R})$.*

- *Any $A \in \mathcal{B}(\mathbb{R})$ is called a Borel subset of $\mathbb{R}$.*
- *The followings are Borel subsets of $\mathbb{R}$: for $a \le b$*

$$(a,\infty) = \mathbb{R}\setminus(-\infty,a], \quad (a,b] = (-\infty,b]\cap(a,\infty), \quad (\infty,a) = \cup_{n=1}^{\infty}(-\infty, a-\frac{1}{n}],$$
$$[a,\infty) = \mathbb{R}\setminus(-\infty,a), \quad (a,b) = (-\infty,b)\cap(a,\infty), \quad [a,b] = \mathbb{R}\setminus\Big((-\infty,a)\cup(b,\infty)\Big),$$
$$\{a\} = [a,a], \quad \mathbb{N} = \cup_{n=0}^{\infty}\{n\}, \quad \mathbb{Q} = \cup_{m,n\in\mathbb{Z},n\neq 0}\{\frac{m}{n}\}, \quad \mathbb{Q}^{\complement} = \mathbb{R}\setminus\mathbb{Q}.$$

Example 1.3 (**Borel σ-algebra on $\mathbb{R}^n$).** *It is σ-algebra generated by the n-dimensional rectangles of the form $\prod_{i=1}^n (a_i, b_i]$*

Definition 1.5 (**Measurable space).** Let $\Omega \neq \emptyset$ and $\mathcal{F}$ be a *sigma*-algebra over Ω. The pair $(\Omega, \mathcal{F})$ is called a measurable space.

Our goal now is to assign to each event A, a non-negative real number which indicates its chance of happening.

Definition 1.6 (**Probability measure).** Let $(\Omega, \mathcal{F})$ is a measurable space. A function $\mathbb{P} : \mathcal{F} \to [0,1]$ is called probability measure if

i) $\mathbb{P}(A) \ge 0 \quad \forall\, A \in \mathcal{F}$
ii) $\mathbb{P}(\emptyset) = 0$ and $\mathbb{P}(\Omega) = 1$
iii) If $\{A_i\} \subset \mathcal{F}$ and A_i's are pairwise disjoint, then

$$\mathbb{P}\big(\cup_{i=1}^{\infty} A_i\big) = \sum_{i=1}^{\infty}\mathbb{P}(A_i).$$

The triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space.

Example 1.4. *Let $\Omega = \{1,2,3\}$ and $\mathcal{F} = \{\emptyset, \Omega, \{1\}, \{2,3\}\}$. Define $\mathbb{P} : \mathcal{F} \to [0,1]$ via*

$$\mathbb{P}(A) = \begin{cases} 1, & \text{if } 3 \in A \\ 0, & \text{if } 3 \notin A. \end{cases}$$

Then $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$.

Example 1.5. *Let $\Omega = (0,\infty)$ and $\mathcal{F}$ is the Borel σ-field on Ω. Define, for each interval $A \subset \Omega$*

$$\mathbb{P}(A) = \int_A e^{-x}\,dx.$$

Observe that $\mathbb{P}(A) \ge 0$ for all $A \in \mathcal{F}$ and $\int_\Omega e^{-x}\,dx = 1$. Let $\{A_i\} \subset \mathcal{B}$ and pairwise disjoint. Then $\mathbb{P}\big(\cup_{i=1}^{\infty} A_i\big) = \int_{\mathbb{R}} e^{-x}\mathbf{1}_{\cup_{i=1}^{\infty} A_i}(x)\,dx$. We need to show that

$$\int_{\mathbb{R}} e^{-x}\mathbf{1}_{\cup_{i=1}^{\infty} A_i}(x)\,dx = \sum_{i=1}^{\infty}\int_{A_i} e^{-x}\,dx = \sum_{i=1}^{\infty}\mathbb{P}(A_i).$$

One can show the first equality by using monotone convergence theorem on the sequence of functions $f_n(x) = e^{-x}\mathbf{1}_{\cup_{i=1}^{n} A_i}(x)$.

Example 1.6. *Let $\Omega = \{1, 2, 3, \dots\}$ and $\mathcal{F}$ be the set of all subsets of Ω. Define $\mathbb{P} : \mathcal{F} \to [0, 1]$ via*

$$\mathbb{P}(\{i\}) = \frac{1}{2^i}.$$

Note that since $\sum_{i=1}^{\infty} \frac{1}{2^i} = 1$, we have $\mathbb{P}(\Omega) = 1$. Check that $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$.

Theorem 1.2. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Then the followings hold:*

1) $\mathbb{P}(A^{\complement}) = 1 - \mathbb{P}(A)$ for any $A \in \mathcal{F}$.
2) If $A \subseteq B$, then $\mathbb{P}(A) \le \mathbb{P}(A) + \mathbb{P}(B \setminus A) = \mathbb{P}(B)$.
3) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$. More generally, if $A_1, A_2, \dots, A_n$ are events, then

$$\mathbb{P}(\cup_{i=1}^n A_i) = \sum_{i=1}^n \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i \cap A_j) + \sum_{i<j<k} \mathbb{P}(A_i \cap A_j \cap A_k) - \ldots + (-1)^{n+1} \mathbb{P}(\cap_{i=1}^n A_i).$$

4) Let $\{A_n\}$ be an increasing sequence of elements in $\mathcal{F}$ i.e., $A_n \in \mathcal{F}$ and $A_n \subseteq A_{n+1}$ for all $n = 1, 2, \dots$. Then

$$\mathbb{P}(\cup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mathbb{P}(A_n). \tag{1.1}$$

5) Let $\{A_n\}$ be a decreasing sequence of elements in $\mathcal{F}$ i.e., $A_n \in \mathcal{F}$ and $A_{n+1} \subseteq A_n$ for all $n = 1, 2, \dots$. Then

$$\mathbb{P}(\cap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mathbb{P}(A_n). \tag{1.2}$$

The propery (1.1) is called continuity of $\mathbb{P}$ from below and property (1.2) is called continuity of $\mathbb{P}$ from above.

Proof. Proof of 1): Since $A \cup A^{\complement} = \Omega$, we see that $\mathbb{P}(A^{\complement}) = 1 - \mathbb{P}(A)$ for any $A \in \mathcal{F}$.
Proof of 2): Observe that

$$B = A \sqcup (B \setminus A) \implies \mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \ge \mathbb{P}(A).$$

Proof of 3): We have

$$A \cup B = A \sqcup (B \setminus A) \implies \mathbb{P}(A \cup B) = \mathbb{A} + \mathbb{P}(B \setminus A). \tag{1.3}$$

Observe that $B \setminus A \subseteq B$, and hence by 2), we get that

$$\begin{aligned} &\mathbb{P}(B) = \mathbb{P}(B \setminus A) + \mathbb{P}\big(B \setminus (B \setminus A)\big) = \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B) \\ &\implies \mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B). \end{aligned} \tag{1.4}$$

Combining (1.3) and (1.4), we obtain

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

Proof of 4): Let $A = \cup_{n=1}^{\infty} A_i$. Define

$$C_1 = A_1, \quad C_2 = A_2 \setminus A_1, \quad C_3 = A_3 \setminus A_2, \dots, C_n = A_n \setminus A_{n-1}.$$

Then $\{C_n\}$ are pairwise disjoint and $A = \cup_{n=1}^{\infty} C_n$ with $A_n = \sqcup_{i=1}^n C_i$. Hence we have

$$\mathbb{P}(A) = \mathbb{P}(\cup_{n=1}^{\infty} C_n) = \sum_{n=1}^{\infty} \mathbb{P}(C_n) = \lim_{n\to\infty} \sum_{k=1}^n \mathbb{P}(C_k) = \lim_{n\to\infty} \mathbb{P}(\cup_{k=1}^n C_k) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

Proof of 5): Since A_n's are deceasing, $A_n^{\complement}$ are increasing. Hence by 4), we get

$$\mathbb{P}(\cup_{n=1}^{\infty} A_n^{\complement}) = \lim_{n\to\infty} \mathbb{P}(A_n^{\complement}) \implies \mathbb{P}\big(\cup_{n=1}^{\infty} A_n^{\complement})^{\complement}\big) = \lim_{n\to\infty} \mathbb{P}(A_n) \implies \mathbb{P}(\cap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

$\square$

Example 1.7. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $A, B \in \mathcal{F}$ such that $\mathbb{P}(A) = \frac{3}{4}$ and $\mathbb{P}(B) = \frac{1}{3}$. Show that*

$$\frac{1}{12} \le \mathbb{P}(A \cap B) \le \frac{1}{3}, \quad \frac{3}{4} \le \mathbb{P}(A \cup B) \le 1.$$

Solution: *Since $\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B)$, and $\mathbb{P}(A \cup B) \le 1$, we have*

$$\mathbb{P}(A \cap B) \ge -1 + \frac{1}{3} + \frac{3}{4} = \frac{1}{12}.$$

Also, since $A \cap B \subseteq A$ and $A \cap B \subseteq B$, we have

$$\mathbb{P}(A \cap B) \le \min\{\mathbb{P}(A), \mathbb{P}(B)\} = \frac{1}{3}.$$

Since $A \subseteq A \cup B$ and $B \subseteq A \cup B$, and fact that $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$, we get

$$\max\{\mathbb{P}(A), \mathbb{P}(B)\} \le \mathbb{P}(A \cup B) \le \min\{\mathbb{P}(A) + \mathbb{P}(B), 1\}$$
$$\implies \frac{3}{4} \le \mathbb{P}(A \cup B) \le 1.$$

Let $\Omega \subseteq \mathbb{R}^n$ be a given set, and A be a subset of Ω with $A \in \mathcal{F}$. We are interested in the probability that a 'randomly chosen point' in Ω which falls in A. We define

$$\mathbb{P}(A) = \frac{\text{measure}(A)}{\text{measure}(\Omega)}.$$

For example, if a point is chosen at random from the interval (a, b), the probability that it lies in the interval (c, d), $a \le c < d \le b$ is $\frac{d-c}{b-a}$. Moreover, the probability that the randomly selected point lies in any interval of length $(d - c)$ is the same.

Example 1.8. *A point is picked 'at random' from a unit square. Find the probability that the point falls in the set of all points which are at most a unit distance from the origin.*
Solution: *Here $\Omega = \{(x, y) \in \mathbb{R}^2 : 0 \le x \le 1,\ 0 \le y \le 1\}$ and $A = \{(x, y) \in \Omega : x^2 + y^2 \le 1\}$. Note that*

$$\text{Area}(\Omega) = 1, \quad \text{Area}(A) = \frac{\pi 1^2}{4}.$$

Thus, the probabilty that the point falls in A is given by

$$\mathbb{P}(A) = \frac{\text{measure}(A)}{\text{measure}(\Omega)} = \frac{\pi}{4}.$$

1.1. **Conditional Probability & its properties.** Suppose we know that an event B has happened. Then using this information, we want to measure the chances of happening another event A i.e., to find $\mathbb{P}(A|B)$. Given that the event B occurred, for if it is the case that A occurs if and only iff $A \cap B$ occurs. Thus, $\mathbb{P}(A|B) \sim \mathbb{P}(A \cap B)$ i.e., $\mathbb{P}(A|B) = \alpha \mathbb{P}(A \cap B)$ for some constant $\alpha = \alpha(B)$.

Definition 1.7 (Conditional Probability). Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. If $A, B \in \mathcal{F}$ with $\mathbb{P}(B) > 0$, then the conditional probability of A given B is defined as

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Example 1.9. *Two fair dice are thrown. Given that the first shows 3, what is the probability that total exceeds 6?*

Solution: *Here $\Omega = \{(i, j) : 1 \le i, j \le 6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, $\mathbb{P}(A) = \frac{|A|}{36}$ for any $A \subseteq \Omega$. Let B be the event that first die shows 3 and A be the event that total exceeds 6, i.e.,*

$$B := \{(3, b) : 1 \le b \le 6\}, \quad A := \{(a, b) \in \Omega : a + b \ge 6\}$$

Note that $\mathbb{P}(B) = \frac{6}{36} = \frac{1}{6}$.

$$A \cap B = \{(3,4),(3,5),(3,6)\} \implies \mathbb{P}(A\cap B) = \frac{3}{36} = \frac{1}{12}.$$

Thus, the required probability is given by

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

We now discuss some important properties of conditional probability.

Theorem 1.3. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space.*

1) *Let* $A \in \mathcal{F}$ *with* $\mathbb{P}(A) > 0$. *Then* $\mathbb{P}_A := \mathbb{P}(\cdot|A)$ *is a probability measure over* Ω *with* $\mathbb{P}_A(A) = 1$.
2) **Multiplication rule:** *If* $A_1, A_2, \ldots, A_n \in \mathcal{F}$ *with* $\mathbb{P}(\cap_{i=1}^{n-1} A_i) > 0$, *then*
$$\mathbb{P}(\cap_{i=1}^n A_i) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_2\cap A_1)\ldots\mathbb{P}(A_n|\cap_{i=1}^{n-1} A_i).$$
3. **Total probability rule:** *Let* $A_1, A_2, \ldots$ *be finite or countable partition of* Ω *i.e.,*
$$A_i \cap A_j = \emptyset \quad \forall i \neq j, \quad \text{and} \quad \cup_{i=1}^\infty A_i = \Omega$$
such that $\mathbb{P}(A_i) > 0$. *Then for any* $B \in \mathcal{F}$
$$\mathbb{P}(B) = \sum_{i=1}^\infty \mathbb{P}(B|A_i)\mathbb{P}(A_i)\,.$$
4. **Bayes' rule:** *Let* $A_1, A_2, \ldots$ *be finite or countable partition of* Ω *such that* $\mathbb{P}(A_i) > 0$. *Then for any* $B \in \mathcal{F}$ *with* $\mathbb{P}(B) > 0$,
$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(A_i)\mathbb{P}(B|A_i)}{\sum_{j=1}^\infty \mathbb{P}(B|A_j)\mathbb{P}(A_j)}\,.$$

Proof. Proof of 1): Observe that $0 \le \mathbb{P}_A(B) \le 1$ for any $B \in \mathcal{F}$ and $\mathbb{P}_A(\Omega) = 1$. Moreover, $\mathbb{P}_A(A) = 1$. Let $\{B_i\}$ are disjoint events. Then $C_i := B_i \cap A$ are disjoint. Hence we have

$$\mathbb{P}_A(\cup_{i=1}^\infty B_i) = \frac{\mathbb{P}\big(A \cap (\cup_{i=1}^\infty B_i)\big)}{\mathbb{P}(A)} = \frac{\mathbb{P}(\sqcup_{i=1}^\infty C_i)}{\mathbb{P}(A)} = \frac{\sum_{i=1}^\infty \mathbb{P}(C_i)}{\mathbb{P}(A)} = \sum_{i=1}^\infty \frac{\mathbb{P}(C_i)}{\mathbb{P}(A)} = \sum_{i=1}^\infty \mathbb{P}_A(B_i).$$

This show that $\mathbb{P}_A := \mathbb{P}(\cdot|A)$ is a probability measure over Ω.

Proof of 2): Let us take $A_0 = \Omega$ for convenient. Then, we have

$$\begin{aligned}&\mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_2\cap A_1)\ldots\mathbb{P}(A_n|\cap_{i=1}^{n-1} A_i)\\ &= \prod_{r=1}^n \mathbb{P}(A_r|\cap_{k=1}^{r-1} A_k) = \prod_{r=1}^n \frac{\mathbb{P}\big(A_r \cap (\cap_{k=1}^{r-1} A_k)\big)}{\mathbb{P}(\cap_{k=1}^{r-1} A_k)} = \prod_{r=1}^n \frac{\mathbb{P}(\cap_{k=1}^r A_k)}{\mathbb{P}(\cap_{k=1}^{r-1} A_k)} = \mathbb{P}(\cap_{i=1}^n A_i)\,.\end{aligned}$$

Proof of 3): Observe that, since $\{A_i\}$ is disjoint, $\{B \cap A_i\}$ is also disjoint. Moreover,

$$B = B \cap \Omega = B \cap (\cup_{i=1}^\infty A_i) = \sqcup_{i=1}^\infty (B \cap A_i).$$

Thus, by using the property of $\mathbb{P}$, we obtain

$$\mathbb{P}(B) = \mathbb{P}\big(\sqcup_{i=1}^\infty (B\cap A_i)\big) = \sum_{i=1}^\infty \mathbb{P}(B \cap A_i) = \sum_{i=1}^\infty \mathbb{P}(B|A_i)\mathbb{P}(A_i)\,.$$

Proof of 4): By using the definition of conditional probability and the total probability rule 2), we have

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(A_i\cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A_i)\mathbb{P}(B|A_i)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A_i)\mathbb{P}(B|A_i)}{\sum_{j=1}^\infty \mathbb{P}(B|A_j)\mathbb{P}(A_j)}.$$

□

Example 1.10. *Three teenagers want to get into a R-rated movie. At the box office, ther are asked to produce their IDs; after the clerk checks them and denies them to enter, he returns the IDs randomly. Find the probability that none of the teenagers get their own ID.*

Solution: *Let A be the event that none of the teenagers get their own ID and A_i denotes the event that i-th teenager gets his own ID. Our required probability is $\mathbb{P}(A)$. Note that since there are three possible cases and only one is favourable, $\mathbb{P}(A_i) = 1/3$ for $i = 1, 2, 3$. Note that after giving the right ID to i-th teenager, probability that j-th teenager will get his own ID is equal to $1/2$ as there are two possible ones in which one is favourable. Thus, for $i \neq j$, we see that*

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j|A_j) = \frac{1}{3} \times \frac{1}{2} = \frac{1}{6}.$$

In a similar fashion, we get

$$\mathbb{P}(\cap_{i=1}^3 A_i) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|(A_1 \cap A_2)) = 1/3 \times 1/2 \times 1 = 1/6.$$

Hence , we compute

$$\mathbb{P}(A) = \mathbb{P}(\cap_{i=1}^3 A_i^{\complement}) = 1 - \mathbb{P}(\cup_{i=1}^3 A_i) = 1 - \sum_{i=1}^3 \mathbb{P}(A_i) + \sum_{i<j} \mathbb{P}(A_i \cap A_j) - \mathbb{P}(\cap_{i=1}^3 A_i)$$
$$= 1 - (3 \times 1/3) + (3 \times 1/6) - 1/6 = 2/6 = 1/3.$$

Example 1.11. *An insurance company believes that people can be divided into two classes: those who are accident prone and those who are not. During any given year, an accident-prone person will have an accident with probability 0.4, whereas the corresponding figure for a person who is not prone to accidents is 0.2. If we assume that 30 percent of the population is accident prone*

a) *What is the probability that a new policyholder will have an accident within a year of purchasing a policy?*
b) *What is the probability that policyholder is accident prone given that he/she has an accident within a year of purchasing a policy?*
c) *What is the probability that a new policyholder will have accident in his/her second year of policy ownership given that the policyholder has had accident in the first year?*

Solution: *Let A be the event that the policyholder is accident prone and A_i be the event that the policyholder has had accident in i-th year. We need find $\mathbb{P}(A_1), \mathbb{P}(A|A_1)$ and $\mathbb{P}(A_2|A_1)$. From the given conditions, we have*

$$\mathbb{P}(A_1|A) = 0.4 = \mathbb{P}(A_2|A \cap A_1), \quad \mathbb{P}(A_1|A^{\complement}) = 0.2 = \mathbb{P}(A_2|A^{\complement} \cap A_1), \quad \mathbb{P}(A) = 0.3, \quad \mathbb{P}(A^{\complement}) = 0.7$$

By the total probability rule, we obtain

$$\mathbb{P}(A_1) = \mathbb{P}(A_1|A)\mathbb{P}(A) + \mathbb{P}(A_1|A^{\complement})\mathbb{P}(A^{\complement}) = (0.4)(0.3) + (0.2)(0.7) = 0.26.$$

Moreover, by using definition of conditional probability, we get

$$\mathbb{P}(A|A_1) = \frac{\mathbb{P}(A \cap A_1)}{\mathbb{P}(A_1)} = \frac{\mathbb{P}(A_1|A)\mathbb{P}(A)}{\mathbb{P}(A_1)} = \frac{(0.3)(0.4)}{0.26} = \frac{6}{13}.$$

Note that $\mathbb{P}_{A_1}(\cdot) := \mathbb{P}(\cdot|A_1)$ is a probability measure. Thus, by total probability rule, we have

$$\begin{aligned}\mathbb{P}(A_2|A_1) = \mathbb{P}_{A_1}(A_2) &= \mathbb{P}_{A_1}(A_2|A)\mathbb{P}_{A_1}(A) + \mathbb{P}_{A_1}(A_2|A^{\complement})\mathbb{P}_{A_1}(A^{\complement}) \\ &= \mathbb{P}(A_2|(A \cap A_1))\mathbb{P}(A|A_1) + \mathbb{P}(A_2|(A^{\complement} \cap A_1))\mathbb{P}(A^{\complement}|A_1) \\ &= \mathbb{P}(A_2|(A \cap A_1))\mathbb{P}(A|A_1) + \mathbb{P}(A_2|(A^{\complement} \cap A_1))\big[1 - \mathbb{P}(A|A_1)\big] \\ &= (0.4)\frac{6}{13} + (0.2)\frac{7}{13} \approx 0.29.\end{aligned}$$

1.2. **Independence of Events and its properties:** Sometimes the occurrence of an event B does not affect the probability of an event A i.e., $\mathbb{P}(A|B) = P(A)$. In that case, we say the event A is "independent" from event B.

Definition 1.8. We make the following definition regarding various notion of independent of events:

i) **Independent Events:** Two events A and B are said to be independent if
$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$
ii) **Pairwise Independent Events:** A family of events $\{A_i : i \in I\}$ is said to be pairwise independent if
$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j) \quad \forall\ i \neq j.$$
iii) **Independent family:** A family of events $\{A_i : i \in I\}$ is said to be independent / mutually independent if for every finite subset $J \neq \emptyset$ of I, there holds
$$\mathbb{P}(\cap_{j\in J} A_j) = \prod_{j \in J} \mathbb{P}(A_j).$$

Example 1.12. *Two fair dice are thrown. Consider the following events: let A be the events that first die shows 2, B be the event that second die shows 5 and C be the event that total is equal to 7, i.e.,*
$$A = \{(2,b) : 1 \leq b \leq 6\}, \quad B = \{(a,5) : 1 \leq a \leq 6\}, \quad C = \{(a,b) : a+b = 7, 1 \leq a, b \leq 6\}$$
Here $\Omega = \{(i,j) : 1 \leq i, j \leq 6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$, $\mathbb{P}(D) = \frac{|D|}{36}$ for any $D \subseteq \Omega$. It is easy to check that
$$\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{6}, \quad \mathbb{P}(A \cap B) = \mathbb{P}(B \cap C) = \mathbb{P}(C \cap A) = \frac{1}{36}, \mathbb{P}(A \cap B \cap C) = \frac{1}{36}.$$
Hence

a) *The family $\{A, B, C\}$ is pairwise independent.*
b) *The family $\{A, B, C\}$ is NOT mutually independent as $\mathbb{P}(A \cap B \cap C) = \frac{1}{36} \neq \frac{1}{216} = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$.*