Chapter 2

Probability Space: Sample Space, Events, Probability

2.1 Sample Space

In this chapter, we will establish the foundations which probability theory is built on, namely, the probability space, consisting of the three components: the sample space, events, and probability. We are interested in phenomena, whose outcomes are not deterministically predictable. We start out with some definitions.

Definition 2.1 A phenomenon whose outcome cannot be predicted with certainty is called an (random) experiment. Each performance of the experiment is called a trial, which results in an outcome of the experiment.

Definition 2.2 The set of all possible outcomes of an experiment is called the sample space, denoted by $S$ (sometimes denoted by $\Omega$).

Sample space corresponds to the universe set in set theory. Determining it is the first thing that needs to be done in setting up a probability problem. However, there may be some idealization or choices involved in specifying the sample space. Our probabilistic interest in the experiment determines which occurrences are to be specified as "outcomes" of the experiment. Certain physically possible occurrences may be left out of the sample space, if they are not considered probabilistically interesting. Therefore, different interests may lead to different specifications of the sample space for the same physical experiment. Although, often, proba-
bilistic interests associated with an experiment are common, leading to a common specification of the sample space. In any event, once it is specified at the outset, it remains fixed throughout the problem. Another important point to emphasize is that the outcomes of a random experiment, that is, the elements of the sample space $S$, are not necessarily real numbers. They can be any objects representing the outcomes of the experiment. Examples below will illustrate these points.

Example 2.1

1. **Coin toss.** The generally accepted sample space for this experiment is $S = \{h, t\}$, where $h$ and $t$ respectively denote the outcomes “heads” and “tails” showing in the toss. One can think of other “possible” occurrences in this experiment, such as the coin breaking up into two or standing on its edge and so on, but one would include these as outcomes only if they are “of interest”. Usually they are not of interest and they are left out. This is what is meant by the **idealization** in specifying the sample space.

2. **Die throw.** The usual sample space for this experiment is $S = \{f_1, f_2, f_3, f_4, f_5, f_6\}$, where $f_i$ denotes the outcome $i$-numbered face of the die showing.

3. **Coin tossed twice.** The usual sample space is $S = \{hh, ht, th, tt\}$, where $ht$ denotes a “heads” showing in the first toss and “tails” in the second and so on. However, if one were only interested in the “number of heads” showing, then $S = \{0, 1, 2\}$ would be a satisfactory sample space specification. This is an example of “our interest determining the sample space”. With the latter choice, given the the outcome is “1”, we cannot tell which of the tosses (first or second) resulted in heads. If this is of interest to us, then the second choice for the sample space would not be satisfactory. Note that the first choice for the sample space is finer or more detailed than the second one. Usually we prefer to work with a finer sample space, since it provides a better description of the outcomes of the experiment.

4. **Users on a computer system.** If the number of users on a computer system or customers in a facility at a given time is of interest, then $S = \{0, 1, 2, \ldots, N\}$, where $N$ denotes the maximum capacity, would be a satisfactory sample space.

5. **Radioactive disintegration.** If the number of particles emitted by a radioactive material in a certain time period is of interest, then $S = \{0, 1, 2, 3, \ldots\}$, having a countably infinite
2.1. SAMPLE SPACE

number of outcomes, would be a satisfactory sample space. We cannot limit the sample space to a finite number, however large, since there is always the possibility that there may be more emissions, unless one wishes to limit that as an idealization in the experiment.

6. Interarrival time. If the experiment is the emission of particles or customers arriving at a facility and we are interested in the time elapsed between two successive emissions or arrivals, then \( S = \{x \in \mathbb{R} \mid x > 0\} = \{x > 0\} = (0, \infty) \) would be the sample space. Such an uncountable sample space containing outcomes with a continuum of values is called a continuous sample space; in contrast, the sample spaces in the previous examples are called discrete sample spaces.

7. Communication network. A communication network consisting of five links, each of which may be operational or non-operational, is observed at a given time and its "state" is noted (see Fig. 2.1). An outcome of this experiment is the state of the network consisting of the state of each link. For example, if we let \( a_i \) for \( i = 1, \ldots, 5 \) denote the situation that link \( i \) is operational and \( \bar{a}_i \), that it is non-operational, then \( a_1 \bar{a}_2 a_3 a_4 a_5 \) is an outcome of this experiment that corresponds to the state of the network with links 1, 3, 5 operational and links 2, 4 non-operational. The sample space for this experiment can then be expressed as

\[
S = \{a_1a_2a_3a_4a_5, \bar{a}_1a_2a_3a_4a_5, a_1\bar{a}_2a_3a_4a_5, a_1a_2\bar{a}_3a_4a_5, \ldots, \bar{a}_1\bar{a}_2\bar{a}_3\bar{a}_4\bar{a}_5\}. \quad (2.1)
\]

Altogether \( S \) contains 32 elements corresponding to the 32 distinct situations or states that can be encountered in this 5-link network.

EOE

Frequently, we are interested in some characteristic of the outcomes of the random experiment. Specifying such a "characteristic" describes an "event" associated with the experiment. The set of all outcomes having this characteristic, that is a subset of \( S \), can be thought of as the event that represents that characteristic, thus establishing a correspondence between the "characteristic" and the "event".
2.2 Events

We have established in the previous section that we may be interested in certain subsets of the sample space $S$, whose elements are specified by some characteristic or other, and that we wish to call such subsets of $S$ events. From this it follows that, an event is a subset of $S$. The question then is: Are all subset of $S$ events? The short answer is: Not necessarily or it depends. The complete answer will emerge from the discussions below.

Let $\mathcal{F}$ denote the set of all events associated with the random experiment having the sample space $S$. Then the pair $(S, \mathcal{F})$ is called the event space. So the subsets of $S$ that are events are all elements of the set $\mathcal{F}$. (Note that, if all subsets of $S$ were/are events, then $\mathcal{F}$ would be the power set of $S$.) The set of events, that is the set $\mathcal{F}$, must have a structure reflecting certain intuitive requirements we have of events. This structure that $\mathcal{F}$ must satisfy and our interest in the problem together determine which subsets of $S$ are events.

We say that an event $A$ occurs if the outcome of the experiment is an element in $A$. The requirements we have of events can be stated as follows:

a) If $A$ is an event, so that we can talk about its occurrence, then $\overline{A}$ must also be an event, representing the event that $A$ did not occur.

b) If $A$ and $B$ are two events whose occurrences we can talk about, then we should be able to talk about the occurrence of $A$ or $B$. That is, $A \cup B$ must also be an event.

c) Similarly, if $A$ and $B$ are two events, then $A \cap B$ must be an event also. This third requirement is actually redundant, because it is implied by the first two. (Recall that due to De
2.2. EVENTS

Morgan's laws, the intersection operation can be expressed in terms of the complement and union operations.)

In set theory, there exists a concept describing the structure that corresponds to the requirements on events stated above, namely the field or algebra. Specifically, a field is defined as follows:

**Definition 2.3** For a set $S$, a non-empty set of subsets of $S$, denoted by $\mathcal{F}$, is called a field or algebra if and only if the following conditions are satisfied

i) If $A \in \mathcal{F}$, then $\overline{A} \in \mathcal{F}$.

ii) If $A, B \in \mathcal{F}$, then $A \cup B \in \mathcal{F}$.

(As noted above, that, if $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$, follows from the definition.) Thus, requiring $\mathcal{F}$, the set of events, to be a field defined on $S$, ensures the structure we require of events. Actually, for reasons beyond the scope of this presentation, a little more is required, namely, that the set of events must be a Borel field, (also called or $\sigma$-field, or $\sigma$-algebra). In the above definition, if condition (ii) is replaced by

ii') If $A_1, A_2, A_3, \ldots \in \mathcal{F}$, then $\bigcup_{n=1}^{\infty} A_i \in \mathcal{F}$,

then $\mathcal{F}$ is called a Borel field. Although it may appear like successive applications of condition (ii) will yield condition (ii'), hence the two are equivalent, they actually are not. Condition (ii') is slightly more stringent. It is not difficult to give examples of fields that are not Borel fields, thus showing that the two conditions are not equivalent. The distinction becomes important in more advanced probabilistic analyses. It is not important for our purposes, hence we will not dwell on it any further, except to say that (for the sake of correctness) we require the set of events, $\mathcal{F}$, to be a Borel field.

Thus we have determined the structure that the set of events must satisfy, but we still have not specified which subsets of $S$ are events. Before we do that, we note that it follows from the above definitions that, $S$ and $\phi$ are elements of any field or Borel field defined on $S$. In other words, $S$ and $\phi$ are always events. The sample space $S$ is the event which occurs always, that is, in every trial of the experiment. Hence, it is called the **universal event** (or the sure
event or certain event). And the null set $\phi$ is the event which never occurs; hence, it is called the impossible event (or the null event). The question we now address is: What else is in the set of events, $\mathcal{F}$?

One way to specify the set of events is as follows: Generally, there is a class of subsets of $S$ that we are interested in, say $C$. We “declare” the elements of $C$ as events and include them as elements of the set of events, $\mathcal{F}$. We then include in $\mathcal{F}$ all other subsets of $S$ which will ensure that $\mathcal{F}$ is a Borel field, that is, the conditions in the definition are satisfied. Whether as part of the subsets we are interested in or as the later additions, as mentioned earlier, $S$ and $\phi$ are always elements in $\mathcal{F}$. The set of events constructed this way contains all subsets that we are interested in as events and also it is a Borel field satisfying all the intuitive and mathematical requirements. The subsets of $S$ that are put in $\mathcal{F}$, so that it is a Borel field, may be too many and too unruly. But fortunately, we do not need to write down a list of all the elements in $\mathcal{F}$. It suffices to declare/specify $\mathcal{F}$ as the smallest Borel field that contains the class $C$ of subsets of $S$. This will give us the smallest Borel field that will do the job.

Another alternative in specifying $\mathcal{F}$, the set of events, is to take it to be $\mathcal{P}_S$, the power set of $S$. Since the power set is always a Borel field (trivial to show using the definition), this choice also satisfies all the requirements. It corresponds to the case where all subsets of $S$ are events, thus eliminating the problem of worrying about which subsets are events and which are not. So then, why don’t we simply take the power set to be the set of events in all cases? The answer we can give here is a partial one, hence it partly needs to be taken on faith. In many instances, such as in cases with continuous sample spaces, the power set contains some very “pathological” subsets of $S$ such that, the probability assignment to these potential events, which is the third step in the construction of the probability space, cannot be done in a convenient way. This difficulty, which will be briefly discussed again in the next section, deters us from selecting the power set as the set of events categorically in all cases. When there is no danger of difficulty in the probability assignment to events, such as in discrete sample space problems, then there is no disadvantage in taking the power set as the set of events, which is what is usually done. In cases where there may be difficulty, we will revert to the choice outlined in the previous paragraph and take the smallest Borel field that contains the class of subsets of $S$ that are of interest.

In summary, the set of events contains all subsets that are of interest, has the structure of
a Borel field and always contains $S$ and $\phi$ as two of its elements. With that long, but hopefully not confusing, discussion of issues concerning the set of events, we will now give some examples of subsets of $S$ which we may be interested in as events. For this, we will refer to some of the examples described in the previous section.

**Example 2.2**

1. Events, i.e., subsets of $S$, containing single elements are called *elementary events* (singleton, as a set). In the die throw experiment, if we are interested in “face ‘$t$’ showing” as an event, it is denoted as $\{f_t\}$. Note that this is not the samething as the outcome $f_t$. The event, “an even numbered face showing” is $\{f_2, f_4, f_6\}$, and the event, “a number less than 5 showing” is $\{f_1, f_2, f_3, f_4\}$.

2. In the ‘coin tossed twice’ experiment, the event “heads shows once” is $A = \{ht, th\}$ and the event “heads shows at least once” is $B = \{hh, ht, th\}$.

3. If the coin is tossed three times, the event “tails shows at most once” is $A = \{hht, hht, hth, thh\}$.

4. A system or a component is put in service at time 0. It operates until it fails, which is random. The sample space for this experiment would be $S = \{teR \mid t > 0\}$, where $t$ represents the outcome “the system failed at time $t$”. Then the event “the system is operational at time $t_o$ is $A = \{teR \mid t > t_o\} = (t_o, \infty)$.

5. In the communication network experiment, we might be interested in the event that “links 1, 2, and 3 are operational” which corresponds to $A = \{a_1a_2a_3a_4a_5, a_1a_2a_3\overline{a_4}a_5, a_1a_2a_3a_4\overline{a_5}, a_1a_2a_3\overline{a_4}\overline{a_5}\}$. Given the configuration of the network as in Fig. 2.1, we could be interested in the event that “there is communication (path with operational links) between points $x$ and $y$ in the network”.

EOE

Finally, we note that, since events are simply sets, most of the definitions in set theory have counterparts in the world of events. For example, the definition of disjoint event is given below; collectively exhaustive events are defined analogous to collectively exhaustive sets.
2.8 \textit{CHAPTER 2. PROBABILITY SPACE: SAMPLE SPACE, EVENTS, PROBABILITY}

\textbf{Definition 2.4} Two events $A$ and $B$ are said to be disjoint (or mutually exclusive) if they do not have a common outcome, i.e., if $A \cap B = \emptyset$.

2.3 Probability

The third component of the probability space, which constitutes the foundation of the probability framework, is the \textit{probability} (function). To each event defined on $S$, a non-negative number called \textit{probability}, denoted by $P(A)$, is assigned. It is meant to represent the “relative likelihood” of a trial of the experiment will result in the occurrence of that event. This assignment of probability, however, has to be according to certain rules, most of which have strong intuitive justifications. The following definition states those rules, which are known as the \textit{axioms of probability}.

\textbf{Definition 2.5} Let $S$ be the sample space and $\mathcal{F}$ be the set of events (a Borel field defined on $S$) in a random experiment. Then a \textit{probability} (function or measure) is a (non-negative) real-valued set function defined on the set of all events, $\mathcal{F}$, which satisfies the following axioms (rules):

\begin{enumerate}
  \item \textbf{A1.} $P(A) \geq 0$, for all $A \in \mathcal{F}$.
  \item \textbf{A2.} $P(S) = 1$.
  \item \textbf{A3.} If $A$ and $B$ are two disjoint events, then $P(A \cup B) = P(A) + P(B)$.
\end{enumerate}

In other words, probability is a function defined on events, elements of $\mathcal{F}$, taking values in $[0, 1]$, that is, $P(\cdot) : \mathcal{F} \rightarrow [0, 1]$, which satisfies axioms \textbf{A1, A2, and A3}. Actually, again for some finer mathematical reason, we require the probability function to satisfy an axiom that is slightly more stringent than \textbf{A3}. The reasons for this are analogous to those requiring the set of events be a Borel field and not just a field. The axiom that needs to be satisfied is:

\begin{enumerate}
  \item \textbf{A3'.} If $A_1, A_2, A_3, \ldots$ are events such that any two are disjoint, then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.
\end{enumerate}

Although, again it may appear like successive application of \textbf{A3} implies \textbf{A3'}, hence the two would be equivalent, they are not. We will simply require $P(\cdot)$ to satisfy \textbf{A3'}, without dwelling
on the reasons. We now have all three components of the probability space and consequently the probability space defined:

Definition 2.6 The triplet \((S, \mathcal{F}, P)\), where \(S\) is the sample space, \(\mathcal{F}\) is the set of events and \(P\) is a probability function defined on events, is called the probability space.

The probability space embodies all the probabilistic characteristics or information of the random experiment. Note that the axioms of probability, aside from constituting a consistent and complete set of rules for probabilities, have strong intuitive justifications. In our daily untrained understanding of probability, these three rules are part of the properties we have gotten to expect probabilities to satisfy. Thus it is reaffirming that they constitute the set of rules probability is built on. The interesting point is that they are a sufficient set of rules and that all other intuitive or otherwise properties of probability are consequences of these three. Other intuitively motivated properties of probability are obtained as theorems that follow from the axioms. Below are a few:

Theorem 2.1 \(P(\emptyset) = 0\).

Proof: \(A = A \cup \emptyset\), for any event \(A\). Therefore, \(P(A) = P(A \cup \emptyset)\). Now since \(A\) and \(\emptyset\) are disjoint sets (events) (because \(A \cap \emptyset = \emptyset\)), by axiom \(A3'\), \(P(A \cup \emptyset) = P(A) + P(\emptyset)\). It follows then, that \(P(A) = P(A) + P(\emptyset)\). Therefore, \(P(\emptyset) = 0\) \(\hspace{1cm} EOP\)

Theorem 2.2 \(P(\overline{A}) = 1 - P(A)\), for any event \(A\).

Proof: Since \(A\) and \(\overline{A}\) are disjoint, by \(A3'\), \(P(A) + P(\overline{A}) = P(A \cup \overline{A})\). But \(A \cup \overline{A} = S\) and by \(A2\), \(P(S) = 1\). Therefore, \(P(A) + P(\overline{A}) = 1\) and the result follows. \(\hspace{1cm} EOP\)

Venn diagrams are also useful in representing events (i.e., subsets of \(S\)) and even their probabilities. Upon a conceptual normalization, for example, it is possible to think of probabilities of events as their "areas" in the Venn diagram representations. This can be helpful in visualizing certain probability operations. Although visualizations through Venn diagrams are not accepted as formal proofs, they can be very useful in formulating them. Hence, it is perfectly alright to use them as visual aid, as we do in the following theorems.
Figure 2.2: Some set operations

**Theorem 2.3** For $A$ and $B$ any two events,

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

**Proof:** We start with some set equalities,

$$A = A \cap S = A \cap (B \cup \overline{B}) = (A \cap B) \cup (A \cap \overline{B}).$$

$(A \cap B)$ and $(A \cap \overline{B})$ are disjoint, because $(A \cap B) \cap (A \cap \overline{B}) = A \cap B \cap \overline{B} = \emptyset$. Therefore, it follows from $A3'$ that

$$P(A) = P(A \cap B) + P(A \cap \overline{B}).$$

(2.3)

Note that this equation can be motivated using the corresponding Venn diagram, Fig. 2.2.a, by noting that the set (event) $A$ consists of two disjoint portions $A \cap B$ and $A \cap \overline{B}$, hence $P(A)$ should be the sum of the probabilities of these two portions, which is what eq.(2.3) is. Similarly, note the following set equalities,

$$B \cup A = (B \cup A) \cap S = (B \cup A) \cap (B \cup \overline{B}) = B \cup (A \cap \overline{B}).$$

(2.4)

$B$ and $(A \cap \overline{B})$ are disjoint, because $B \cap (A \cap \overline{B}) = A \cap (B \cap \overline{B}) = A \cap \emptyset = \emptyset$. That $A \cup B$ consists of two disjoint portions $B$ and $(A \cap \overline{B})$ is easily seen from the Venn diagram in Fig. 2.2.b. Then by $A3'$, it follows that

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

(2.5)

Now substituting eq.(2.3) into eq.(2.5) and rearranging, we obtain

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

(2.6)
2.3. PROBABILITY

which is what was to be proved.

One can give a Venn diagram interpretation of this result by noting that the (normalized) 'area' of the set $A \cup B$ (representing the probability of that event) can be viewed as the sum of the areas (probabilities) of $A$ and $B$, minus the area (probability) of the set $A \cap B$ as shown in Fig. 2.2.c. Since the portion $A \cap B$ is accounted for twice when the areas of $A$ and $B$ are added, it is subtracted once, so that each portion that needs to be considered is done so only once.

Corollary 2.1 $P(A \cup B) \leq P(A) + P(B)$.

Proof: The result follows from the previous theorem upon noting that $P(A \cap B) \geq 0$ by axiom $A1$.

Generalizing the result of the previous theorem to three events, we have:

Theorem 2.4 For any three events $A$, $B$, and $C$,

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(C \cap A) + P(A \cap B \cap C). \quad (2.7)$$

Proof: The proof follows similar lines as the proof of the previous theorem. It is left as an exercise.

One can give a graphical interpretation of this result using the Venn diagram shown in Fig. 2.3: The (normalized) 'area' (probability) of $A \cup B \cup C$ can be obtained by adding the areas (probabilities) of $A$, $B$, and $C$, then subtracting the areas (probabilities) of the portions considered twice, and then adding the area (probability) of the portion $A \cap B \cap C$ again, since it was counted three times but then subtracted three times. This way, all portions of $A \cup B \cup C$ are considered only once, as it should be. As noted earlier and evidenced here, Venn diagrams can be very helpful in visualizing certain set and probability relationships. The following is an extension of the above result to $N$ arbitrary events.

Theorem 2.5 Inclusion-exclusion principle. Let $A_1, A_2, \ldots, A_N$ be $N$ arbitrary events. Then

$$P(\bigcup_{i=1}^N A_i) = \sum_{i=1}^N P(A_i) - \sum_{i<j=2}^N P(A_i \cap A_j) + \sum_{i<j<k=3}^N P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{N-1} P(\bigcap_{i=1}^N A_i). \quad (2.8)$$
The proof for this can be obtained by mathematical induction. A Venn diagram interpretation of this is not helpful, since it is not practical to depict more than three arbitrary sets with all possible intersections in a Venn diagram.

**Remarks 2.1** The above theorem provides a way for calculating \( P(\bigcup_{i=1}^{N} A_i) \), which corresponds to the probability of at least one of the \( A_i \)'s occurring. If this way is too complicated, then there is an alternate way, using the De Morgan's laws:

\[
P(\bigcup_{i=1}^{N} A_i) = 1 - P(\bigcup_{i=1}^{N} A_i),
\]

\[
= 1 - P(\bigcap_{i=1}^{N} A_i),
\]

which may be simpler, since \( \bigcap_{i=1}^{N} A_i \) corresponds to the event that none of the \( A_i \)'s occur. Later we will give examples where using one or the other of these two methods is more convenient than the other.

Another intuitively motivated result which readily follows from the axioms is:

**Theorem 2.6** If \( A \subseteq B \), then \( P(A) \leq P(B) \).

**Proof:** The proof follows from expressing event \( B \) as the union of two disjoint events \( A \) and \( B - A \), using \( A \delta^i \), and noting that \( P(B - A) \geq 0 \) by \( A1 \). \textit{EOP
2.4 Probability Assignment

The axioms of probability and the resulting theorems provide a consistent theory and they also agree with our intuitive notions of probability. But, the question remains: How do we specify the probability function? One way of doing it would be to specify the probabilities of all events (in a manner consistent with the axioms, that is, such that the axioms are not violated). This, however, is not a practical way of doing it, because there are so many events to specify probabilities for. A more efficient way of specifying the probability function is to specify the “basic probabilities” or a “rule”, which, together with the axioms, enables us to compute the probabilities of all events. Below we will describe this probability assignment procedure for the discrete and continuous sample space cases separately.

2.4.1 Discrete Sample Spaces

Finite Sample Spaces. Suppose the sample space is finite, say $S = \{a_1, a_2, \ldots, a_N\}$, and the event space $\mathcal{F}$ is the power set of $S$. The probability function for such a random experiment is defined by specifying probabilities of all elementary events, i.e., events with single outcomes in them. In other words, we specify

$$P(\{a_i\}) = p_i, \quad \text{for } i = 1, 2, \ldots, N$$

(2.10)

such that

i) $p_i \geq 0$, for all $i \in \{1, 2, \ldots, N\}$.

ii) $\sum_{i=1}^{N} p_i = 1$.

Specifying the $p_i$'s such that the two conditions are satisfied, corresponds to specifying the basic probabilities. The two conditions that must be satisfied correspond to the probability assignment being complete and consistent (alluded to in Section 2.3). The probability of any other event can then be determined using the basic probabilities and the axioms. For example, $P(\{a_1, a_5, a_6\}) = P(\{a_1\}) + P(\{a_5\}) + P(\{a_6\}) = p_1 + p_5 + p_6$. Any probability assignment, that is, a set of $p_i$'s, satisfying conditions (i) and (ii) is a valid one. If either one of these conditions does not hold, then the axioms would be violated.
Example 2.3 Consider the die throwing experiment with $S = \{f_1, f_2, f_3, f_4, f_5, f_6\}$ and $\mathcal{F}$, the power set of $S$. The basic probabilities are specified as $P(\{f_i\}) = p_i$ for $i = 1, 2, \ldots, 6$, such that $p_i \geq 0$ for $i = 1, 2, \ldots, 6$ and $\sum_{i=1}^{6} p_i = 1$. Then the probability of, say $A = \text{"an odd numbered face showing"}$ is obtained as

$$A = \{f_1, f_3, f_5\} = \{f_1\} \cup \{f_3\} \cup \{f_5\},$$

implying

$$P(A) = P(\{f_1\}) + P(\{f_3\}) + P(\{f_5\}) = p_1 + p_3 + p_5.$$

Similarly, the probability of the event $B = \text{"a non-six showing"}$, would be given by $P(B) = p_1 + p_2 + p_3 + p_4 + p_5$ or by $P(B) = 1 - P(\overline{B}) = 1 - p_6$, which are of course equal due to the second condition.

The outstanding question now is: How are $p_1, p_2, \ldots, p_6$ specified? They are either given or implied or assumed or estimated. For example, if we know that the die is "fair", then we may assume $p_i = \frac{1}{6}$ for $i = 1, 2, \ldots, 6$. If we cannot assume that the die is fair and $p_i$’s are not given, then we may want to throw the die $n$ times and count $n_i$, the number of times face $f_i$ shows in $n$ throws, for $i = 1, 2, \ldots, 6$. Of course, $n = n_1 + n_2 + \cdots + n_6$ would be true. Then, we may assign the basic probabilities as $p_i = \frac{n_i}{n}$, for $i = 1, 2, \ldots, 6$, that is equal to the corresponding relative frequencies. The resulting probability assignment would be valid, since the two conditions are satisfied. In any event, specifying $p_i$’s this way or any other way entails an assumption. As also discussed in Chapter 1, in this context, the concern is not on the accuracy or the origin of the assumption on the basic probabilities, as long as the two conditions are satisfied, that is, the $p_i$’s specified are complete and consistent. This is similar to, for example, in circuit theory, when a resistance value is given, we do not ask: Is it really that value? Or, how do you know? We assume that it is so and proceed with it.

Countably Infinite Sample Spaces This case is handled in a way similar to the finite sample spaces case. Suppose the experiment has a countably infinite sample space, say $S = \{a_1, a_2, a_3, \ldots\}$, and the set of events $\mathcal{F}$ is taken as the power set of $S$. Then the probability specification can be done by assigning probabilities to the elementary events:

$$P(\{a_i\}) = p_i, \quad \text{for } i = 1, 2, 3, \ldots \quad (2.11)$$

such that
2.4. PROBABILITY ASSIGNMENT

1. i) \( p_i \geq 0 \), for all \( i \in \{1, 2, 3, \ldots\} \).

2. ii) \( \sum_{i=1}^{\infty} p_i = 1 \).

Again the two conditions ensure the completeness and consistency of the assignment with the axioms and again the specification of the probabilities to the elementary events, i.e., the basic probabilities, is done according to some assumption based on, reasoning, experimentation, or subjective judgement on the random phenomenon.

Example 2.4 Consider the experiment: A fair coin is tossed until “heads” shows. The sample space for this experiment can be taken as

\[ S = \{h, th, tth, ttth, tttth, \ldots\} \]

where, for example, the outcome tttth represents the case where “tails” showed in the first four tosses and “heads” on the fifth toss. Note that, for example, ttth is not an outcome in this experiment, because the trial of the experiment ends as soon as a heads shows. Again we can take the power set of \( S \) as the set of events \( F \).

Given that it is a fair coin and the tosses are “independent”, the probability assignment to the elementary events can be made as:

\[
P(\{h\}) = \frac{1}{2}, \quad P(\{th\}) = \frac{1}{4}, \quad \ldots \quad P(\{tt\ldots \underbrace{t}_{k-1} h_{kth}\}) = \left(\frac{1}{2}\right)^k, \quad \ldots
\]

These are the “basic probabilities”, from which other probabilities can be determined using the axioms. For example, the probability of the event \( A \), that the experiment ends before the fifth toss can be computed as

\[
A = \{h, th, tth, ttth\}
\]

\[
P(A) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{15}{16}.
\]

EOE
The above described procedures are the most common ways of specifying the probability function for discrete (finite or countable) sample space experiments, but they may not be the only way. For example, if the set of events is not the power set of \( S \) and that some singleton subsets of \( S \) are not events, then specifying the "basic probabilities" can be done by assigning probabilities to the "smallest" events (some possibly not elementary events) of the event space, but again making sure that the corresponding two conditions are satisfied. Then again the probabilities of all other events can be determined from the basic probabilities by using the axioms. The "smallest event" here can be defined as a subset of \( S \) which is an event and no subset of it, except for \( \phi \) and the set itself, is an event.

2.4.2 Continuous Sample Spaces

In this case, the sample space is uncountable and the outcomes (generally) take on a continuum of values. The most common cases are those where the sample space \( S \) is some interval of (or union of intervals or all of) the real line. For example, \( S = [a, b] \), \( S = [0, \infty) \), and \( S = (-\infty, \infty) \) are such cases. It could also be that, \( S \) is a subset of a higher dimensional Euclidean space, such as \( R^3 \). But for the sake of this discussion, let us take \( S \) to be some interval or all of the real line. Clearly, there are a very large number (uncountably infinite) of outcomes and consequently a very large number (possibly also uncountably infinite) of events. Because of this, the probability specification has to be done in a different way than the discrete sample space case.

In continuous sample space random experiments, we are generally, although not exclusively, interested in "intervals" (subintervals of the sample space) being events. So then, the smallest Borel field containing the "intervals" (as the class of subsets of \( S \) that is of interest) would be an appropriate choice for the set of events \( \mathcal{F} \). We then have to make the probability specification in such a way that the probability of any event in \( \mathcal{F} \) is uniquely and unambiguously specified or determinable. This can be done by specifying a "rule" consistent with the axioms, which can be used to determine the probabilities of all events. Specifically, we define the probability function through

\[
P([a, b]) = \int_a^b w(\tau) \, d\tau, \tag{2.12}
\]

in terms of a \( w(\tau) \), called the "weighting function", which satisfies the conditions:
2.4. PROBABILITY ASSIGNMENT

i) \( w(\tau) \geq 0 \) for all \( \tau \in S \),

ii) \( \int_S w(\tau) \; d\tau = 1 \).

Note that \( w(\tau) \) is defined only on \( S \). Also note that these two conditions ensure that the axioms of probability are not violated. Moreover, using the axioms and this "rule", we can compute the probability of any event in the event space specified above, including the intervals and finite or countable unions of intervals. Thus, specifying the weighting function \( w(\tau) \) specifies the probability function.

Example 2.5 Consider the random experiment: A person arrives at the bus stop at random between 0 and \( T \) secs. Every point in \([0, T]\) is an outcome of this experiment. As a general convention, by "at random" it is generally implied that all outcomes (in this case, all points in \([0, T]\)) are equally likely. This implies that the weighting function \( w(\tau) \) is constant for all \( \tau \in [0, T] \). Then by the second condition, which corresponds to axiom A2, it follows that

\[
w(\tau) = \frac{1}{T},
\]

as shown in Fig. 2.4.a. It follows from this that \( P([a, b]) \), the probability that the person arrives at the bus stop sometime between \( a \) and \( b \), is given by

\[
P([a, b]) = \int_a^b \frac{1}{T} \; d\tau = \frac{b-a}{T}.
\]

In other words, \( P([a, b]) \) is the ratio of the length of the interval \([a, b]\) to the length of the sample space \( S = [0, T] \), as one would expect due to the assumption that all outcomes are equally likely.

If all time instants (outcomes) are not equally likely, then the weighting function can be determined according to what is known or assumed about the experiment. For example, if it is known that an arrival in \([0, \frac{T}{4}]\) is three times more likely as one in \((\frac{T}{4}, T]\), then invoking the second conditions, we find that

\[
w(\tau) = \begin{cases} \frac{3}{4} & \text{if } \tau \in [0, \frac{T}{4}] \\ \frac{2}{3T} & \text{if } \tau \in (\frac{T}{4}, T] \end{cases}
\]

as shown in Fig. 2.4.b. Clearly, the two conditions are satisfied by these weighting functions.

Now, with this weighting function, for example, \( P([0.2T, 0.3T]) \) can be found as

\[
P([0.2T, 0.3T]) = \int_{0.2T}^{0.3T} \frac{2}{T} \; d\tau + \int_{0.2T}^{0.3T} \frac{2}{3T} \; d\tau = \frac{1}{10} + \frac{1}{30} = \frac{2}{15}.
\]

EOE
Figure 2.4: Some weighting functions for $S = [0, T]$.

Figure 2.5: Weighting function examples for $S = [0, \infty)$ and $S = (-\infty, \infty)$.

In the examples above, the weighting function is a step-function, but of course it can also be a continuous function defined on $S$, such as the one shown in Fig. 2.4.c. If the sample space $S$ is semi-infinite or infinite, such as $S = [0, \infty)$ or $S = (-\infty, \infty)$, then the weighting function is defined over the full range specified by $S$. For example, for $S = [0, \infty)$, $w(\tau) = ae^{-at}$, for $a > 0$, shown in Fig. 2.5.a, is a valid weighting function; an example of a weighting function for $S = (-\infty, \infty)$ is shown in Fig. 2.5.b.

Below we discuss some additional facts and issues regarding the probability specification through the use of a weighting function.

Remarks 2.2

1. Where is the weighting function $w(\tau)$ obtained from? Similar to the specification of "basic probabilities" $p_i$'s in the discrete sample space problems, $w(\tau)$ is determined by
an assumption or reasoning or possibly by experimentation, such as from past data on
the random phenomenon. Again we are not very concerned about how it is obtained, as
long as it satisfies the two conditions which in turn ensure that the axioms of probability
are not violated. The weighting function can be considered as the “basic probabilities”
associated with a continuous sample space random experiment.

2. We note that the weighting function method for specifying probabilities is restricted to
random experiments with sample spaces that are intervals (i.e., subsets) of the real num-
bers. This of course does not cover all of the continuous sample space cases, but a great
many of them fall into this category. We also note that the weighting function concept
is very much akin to, and is in some sense superseded by, the concept of the probability
density function for random variables, which will be studied in Chapter 5. Therefore,
the weighting function concept is often omitted in most probability books, despite the
fact that it is the most common method for specifying probabilities in continuous sample
space experiments.

3. We mentioned that we are generally interested in “intervals” as events. The endpoints of
an interval may or may not be included in the interval. We use the round parentheses “(”
and/or “)” if the endpoint is not included in the interval (called open) and use the square
bracket “[” and/or “]” if it is included (called closed), as we have been doing throughout.
Since the integral under a single point is zero, the probabilities of intervals open or closed
on either or both ends are equal, unless the weighting function has an impulse (delta)
function exactly at the endpoint. In that case, of course, it matters whether or not the
point is included in the interval. If the endpoint is included and there is an impulse of
strength c at the endpoint, then c is added to (or is part of) the probability. And it is
not unusual to have situations where the weighting function has some impulse function
components at various \( \tau \) values. In those cases, we need to keep an account of the
endpoints of the intervals.

4. We now discuss a consequence of the previous remark, using the “a person arriving at
the bus stop” example presented above. We ask: What is the probability that the person
arrives at the bus stop at \( t = t_0 \) exactly? Supposing that the weighting function \( w(\tau) \)
does not have an impulse at \( t_0 \), that probability is given by

\[
P([t_0, t_o]) = \int_{t_o}^{t_o} w(\tau) \, d\tau = 0.
\]

That is, the probability of each point (outcome) is zero, although one such outcome occurs at each trial of the experiment. The reason for this seeming contradiction, which really isn't one, is that there are infinitely many points, so each one has to have zero probability; otherwise the "sum" (i.e., the integral) cannot be finite, whereas it has to be 1. Of course, if there is an impulse with strength \( c \) in \( w(\tau) \) at \( \tau = t_o \), then \( P([t_0, t_o]) = c \).

5. Based on the discussion in the previous remark, we can state that, \( P(A) = 0 \) does not imply that \( A = \phi \), although the converse is true, as proved earlier. This point may appear counterintuitive, but upon a little reflection (as in the previous remark) one can see that it is not a contradiction.

6. For emphasis, we repeat a general convention that was alluded to earlier in an example. By at random, in discrete and continuous sample space experiments, it is implied that all outcomes (in \( S \) or in the declared range) are equally likely. In the discrete sample space experiments, this means all \( p_i \)'s are equal; and, in continuous sample space experiments, it means \( w(\tau) \) is constant over the range of \( S \).

7. It is not difficult to think of random experiments which require higher dimensional sample spaces. For example, consider the random experiment where a person arrives at the bus stop at a time instant \( t \) in \([0, T]\) and the bus arrives at \( s \) also in \([0, T]\). Therefore, an outcome of this experiment is \((t, s)\), where both \( t \) and \( s \) are elements of \([0, T]\). Thus, the sample space for this experiment is

\[
S = \{(t, s) \mid t \in [0, T], \, s \in [0, T]\}
= [0, T] \times [0, T],
\]

which is the square shown in Fig. 2.6. The event space for this experiment can be taken as the smallest Borel field that includes rectangles which are subsets of \( S \). And the probability function for this experiment can be specified through a two-dimensional weighting function \( w(t, s) \) defined on \( S \) such that it is non-negative everywhere and it integrates to unity over the square \( S \). Any event \( A \) regarding this random experiment
2.4. PROBABILITY ASSIGNMENT

Figure 2.6: An example of a two-dimensional sample space.

corresponds to a subset of the square $S$; one such event $A$ is depicted in Fig. 2.6. The probability of such an event then can be computed by

$$P(A) = \int \int_A w(t, s) \, dt \, ds.$$ 

Finally, we can further comment on the reasons for not taking the power set of $S$ as the event space $\mathcal{F}$ in continuous sample space experiments. The power set of $S$ has such pathological subsets of $S$ in it that, probability assignment to these sets through the integral of a weighting function as proposed in eq.(2.12) is not possible. In other words, these subsets of $S$ cannot be integrated over; hence probabilities cannot be assigned to these subsets through this preferred probability specification method. Consequently, these subsets cannot be allowed to be events. Thus, in order to be able to use this preferred method of probability assignment, we need to limit the extent of the event space to those subsets of $S$ which can be integrated over. The smallest Borel field which contains the intervals (rectangles for higher dimensions) is such an event space, where a probability assignment to each element of this space can be done through the integral over a weighting function. In summary, this is the reason for selecting a smaller Borel field as the event space in continuous sample space random experiments. In discrete sample space experiments, since the probability assignment is done in terms of the basic probabilities of elementary events and there are a only finite or countable number of them, taking the power set of $S$ as the event space does not cause any difficulty in specifying the probability function. Hence, in discrete sample space experiments, that is what is usually done.
2.5 Problems

1. Write down a sample space for the following random experiments:
   a) A coin is tossed four times.
   b) A die is thrown three times.
   c) A coin is tossed until two "heads" or "tails" show in two consecutive tosses.
   d) Each trial of an experiment results in a "success" (s) or a "failure" (f); the experiment is repeated until the first success. The probability of success in a single trial is given as $p$.
   e) A card is drawn from a deck of 52 cards.
   f) A large lot of IC chips contains a number of defective ones. Three chips are drawn at random and tested.
   g) A box contains 2 red and 4 white balls. 3 balls are drawn at random. The colors of the drawn balls are noted.
   h) Five items two of which are defective are tested in a random order; test results are observed.
   i) A customer arrives at a facility at a random instant between 0 and $T$ seconds; the arrival time is observed. Also, sketch the sample space.
   j) A customer arrives sometime between 0 and $T$ seconds, and the customer's service lasts a random period anywhere between 0 and $\tau$ seconds. A probabilist notes the arrival and departure times of the customer. Also, sketch the sample space.

2. For the random experiments described in Prob. 2.1,
   a) Specify an event space, based on the sample space specified.
   b) Specify a probability function, based on the event space specified and any other information provided.
   c) Specify two events of possible interest, and determine their probabilities.

3. Consider the following random experiment involving a coin and a die, both fair.
   a) The coin is tossed. If "heads" shows, the experiment ends; if "tails" shows, the die is thrown and the experiment ends. Construct the probability space for this experiment.
   b) The die is thrown. If an odd-numbered face shows, the coin is tossed once; if an even-numbered face shows, the coin is tossed twice, and the experiment ends. Construct the probability space for this experiment. Also find the probability that no "heads" is observed in the experiment.

4. A coin with $P(\{h\}) = p$ is tossed until both heads and tails show at least once.
   a) Construct the probability space for this random experiment.
   b) Find the probability (as a function of $k$) that after the $k$th toss the experiment is still going on.

5. Two players $A$ and $B$ roll a fair die taking turns; whoever rolls the first "six" wins and the game ends.
   a) Construct the probability space for this random experiment.
2.5. PROBLEMS

b) Find the probability that the player who starts the game wins, eventually.

6. Two chess players A and B play the final match in a chess tournament. They play game after game. The player who wins two games in a row wins the match. Suppose the outcomes of successive games are independent, the probability that player A wins any particular game is \( \frac{2}{3} \), and for the sake of this problem assume there are no draws.

a) Construct the sample space for this random experiment (write down at least 10 outcomes).

b) Find the probability that the match will be over on or before the fourth game.

c) Find the probability that A wins the match.

d) Write down the event "the match does not end" and find its probability.

7. Al, Bo and Cy take turns (in that order) in throwing a die until the first "six" shows. The person who throws the first six wins the game and the game ends. Suppose the probability of a "six" showing in each throw is \( \frac{1}{6} \).

a) Write down a sample space for this experiment.

b) Assign probabilities to the elementary events.

c) Find the respective probabilities of Al, Bo and Cy winning, eventually.

8. Consider the following experiment (see Fig. 2.7): Experiment starts at point A. A coin with \( P(\{h\}) = \frac{3}{4} \) is tossed. If the outcome is "heads", we take a step to the right; and if it is "tails", we take a step to the left. The experiment terminates when we reach either point \( T_1 \) or point \( T_2 \).

a) Write down the sample space for this experiment.

b) Define a probability function by specifying the probabilities of the elementary events in this experiment. Write down the probabilities for six elementary events.

c) Determine the probability that the experiment ends on or before the fourth toss of the coin.

d) Determine the probability that the experiment ends at point \( T_2 \).

9. Consider the random walk experiment depicted in Fig. 2.8. A man stands at point A and tosses a fair coin. If "heads" shows, he takes a step in the clockwise direction; if "tails" shows, he takes a step in the counterclockwise direction. He repeats tossing the coin and taking a step according to the outcome until he reaches back to the point A, when the experiment ends.

a) Write down the sample space for this experiment.

b) Develop a probability space for this experiment suitable to handle the questions that follow.
c) Write down explicitly the event "the experiment ends at an even number of steps" and calculate its probability.

d) Write down the event "the experiment never ends" and calculate its probability.

10. a) Given $P(A) = 0.6$, $P(B) = 0.7$ and $P(A \cap B) = p$, find the range of values $p$ can take.

   b) Given $P(A) = 0.5$, $P(B) = 0.6$ and $P(\overline{A} \cap \overline{B}) = 0.25$, find $P(A \cap B)$ and $P((A \cap \overline{B}) \cup (\overline{A} \cap B))$.

11. Suppose that $A$, $B$, and $C$ are events such that $P(A) = P(B) = P(C) = 0.35$, $P(A \cap B) = P(B \cap C) = 0$, and $P(A \cap C) = 0.15$. Evaluate the following probabilities:

   a) At least one of the events $A$, $B$, or $C$ occurs,

   b) Exactly one of the events occurs,

   c) At most one of the events occurs,

   d) $A$ occurs, but neither $B$ nor $C$ occurs.

12. In a certain population of 25 years or older males, it is known that 60% are married, 50% are college graduates, 30% smoke, 27% are single college graduates, and 9% are married college graduates who smoke.

   a) A male is selected at random from this population. What is the probability that he is either single or a married college graduate who smokes?

   b) Is the information given sufficient to calculate any probability regarding the three characteristics? If so, find the probability that the person is a single college graduate who smokes; if not, provide additional information to complete the probability specification and find the required probability.

   c) What are his chances for leading a long happy life? (Note the answer to this question is not unique.)

13. Consider the communication network between points $x$ and $y$, shown in Fig. 2.9. Let $A_i$ denote the event that link $a_i$ is operational (and $\overline{A}_i$, that it is not). Let $B$ be the event that there is communication between $x$ and $y$. Express $B$ in terms of $A_i$'s and $\overline{A}_i$'s. Simplify your expressions. Also obtain an expression for $P(B)$ in terms of probabilities of some sets involving $A_i$'s and $\overline{A}_i$'s.

14. a) Given the set (sample space) $S = \{1, 2, 3, 4, 5, 6\}$, find the smallest Borel field that contains $\{1, 2\}$, $\{3\}$, $\{3, 4, 5\}$.
b) Given the sample space $S = [0, 1]$, find the smallest Borel field that contains $[0, 0.2]$ and $[0, 0.5]$.

15. To illustrate the distinction between a field and Borel field, consider the following: Let

- $S$ = set of positive integers,
- $B$ = set of all finite subsets of $S$,
- $C$ = set of all subsets of $S$, whose complements are finite.

Note that $C$ is not the complement of $B$, nor is $B \cup C$ equal to the power set of $S$. Prove that $B \cup C$ is a field, but not a Borel field. The latter can be shown by finding a counterexample that violates the conditions for a Borel field.

16. Prove the following. State reasons for each step in your proof. You may use Venn diagrams, but only as a visual aid.

a) $P(A) \leq 1$ for any event $A$.

b) $P((A \cap \overline{B}) \cup (A \cap B)) = P(A) + P(B) - 2P(A \cap B)$, which corresponds to the probability of the exclusive or.

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

17. a) Given that $(B \cap C) \subseteq A$, prove that $P(A) \geq P(B) + P(C) - 1$.

b) Given that $(B \cap C \cap D) \subseteq A$, prove that $P(A) \geq P(B) + P(C) + P(D) - 2$.

18. Using mathematical induction, prove the inclusion-exclusion principle (Theorem 2.5).
19. Mr. Poorsoul (in the tradition of A.W. Drake, that is, the tradition of probabilistic real-life situations) each night after work before going home stops at a pub for a drink. He is absolutely opposed to drinking and driving, therefore he takes the train to go home. When he gets to the train station, he takes the first train that comes along, whether it is going toward his home or away from it. The trains run in both directions regularly at 10 min. intervals, although not necessarily at the same time. Mr. Poorsoul ends up going away from home 70% of the time, which makes no one happy. Assuming that he arrives at the train station at random, how can you explain this “strange” phenomenon which does not give him even an even chance? Short of buying a watch or looking at the destination of the trains, what can he do to increase his chances of going home?

20. The Random Express passes through Stochasticitytown once every day between 0 and 24 hours. Its time of arrival is of course random; let it be \( \tau \). It is twice more likely to arrive in \([12, 24]\) than in \([0, 12]\), otherwise equally likely in each interval.

   a) Find the weighting function in this random experiment.

   b) Write down the probability assignment for an interval that lies completely within \([0, 12]\); for one that lies completely within \([12, 24]\); and one that has parts in both intervals.

   c) For \( A = \{ \tau > 6 \} \), \( B = \{ \tau < 18 \} \) and \( C = \{ |\tau - 12| > 3 \} \), find the probabilities of the events \( A, B, C, A \cup C, A \cup B \cup C, A \cap B \cap C \).

21. Consider the random experiment with sample space \( S = [0, 1] \). The probability assignment is such that the outcome “0.5” occurs with probability 0.5 and all other outcomes are equally likely.

   a) Determine the weighting function corresponding to the given probability assignment.

   b) Find \( P([0.5, 0.8]), P([0, 0.5]), \) and \( P([0.3, 0.6]) \).

22. Suppose a computer starts processing a job at \( t = 0 \) and it completes the job at a random time \( \tau \). Let \([a, b]\) denote the event that the job is completed at some time in the interval \([a, b]\). Suppose the probability assignment is made such that the probability that the job is completed before \( t \) is given \( 1 - e^{-2t} \), for any \( t \geq 0 \).

   a) Find the weighting function for the job completion time, corresponding to the above probability assignment.

   b) Find the probability of the event that at \( t = 10 \), the job is still running.

   c) Find the probability of the event that the job is completed in \([5, 10]\).

23. A point is selected at random inside a square. Find the probability that the point is closer to the center of the square than to any of the sides.

24. A train and a man arrive at a station between 9:00 and 9:30 AM at random and independently of each other. The train makes a 6 min. stop. The man wants to board the train, but he is not willing to wait more than 5 min. for the train. If the train has come and gone before he arrives, of course, he will not be able to board.

   a) Sketch the sample space for this experiment and sketch the event (on the sketch of the sample space) that the man boards the train.

   b) Determine the probability that the man boards the train.

25. Within a specified 1 hour period \(([0,60])\), a customer arrives at a facility, is served and leaves the facility. Let \( t_1 \) denote the customer’s arrival time and \( t_2 \) his/her departure time.
a) Write down and sketch the sample space of the experiment, specifying both the arrival and departure times.

b) Specify and sketch the event $A$, that at $t = 20$ min., the customer is being served.

c) Specify and sketch the event $B$, that the customer's service takes longer than 15 mins.

d) Specify and sketch the event $C$, that at $t = 20$ mins. the customer is in the facility and his/her service takes (eventually) less than 15 mins.

e) If every outcome in the sample space is equally likely, then find the probabilities of the events described in parts b), c), and d).

26. A student takes an exam consisting of two questions $Q1$ and $Q2$. According to the instructions he/she can spend up to 40 mins. on each question and up to 60 mins. on the whole exam. The student notes the time he/she spends on each question. Assume that all possible outcomes of this experiment (exam) are equally likely.

a) Sketch and write down an expression for the sample space of this experiment (representing the time spent on each of the two questions).

b) Sketch and find the probability of the event that at minute 40, the student is not finished with the exam.

c) Sketch and find the probability of the event that the student answers $Q1$ in less than 10 mins.