Chapter 1 Viewgraphs

Roy D. Yates & David J. Goodman
Set Theory
1.1 Comment: Sets

A set is a collection of things. We use capital letters to denote sets. The things that together make up the set are *elements*. When we use mathematical notation to refer to set elements, we usually use small letters. Thus we can have a set $A$ with elements $x$, $y$, and $z$. The symbol $\in$ denotes set inclusion. Thus $x \in A$ means “$x$ is an element of set $A$.” The symbol $\notin$ is the opposite of $\in$. Thus $c \notin A$ means “$c$ is not an element of set $A$.”
1.1 Comment: Set Union

The *union* of sets $A$ and $B$ is the set of all elements that are either in $A$ or in $B$, or in both. The union of $A$ and $B$ is denoted by $A \cup B$. In this Venn diagram, $A \cup B$ is the complete shaded area. Formally,

$$x \in A \cup B \text{ if and only if } x \in A \text{ or } x \in B.$$ 

The set operation union corresponds to the logical “or” operation.
1.1 Comment: Set Intersection

The *intersection* of two sets $A$ and $B$ is the set of all elements that are contained both in $A$ and $B$. The intersection is denoted by $A \cap B$. Another notation for intersection is $AB$. Formally, the definition is

$$x \in A \cap B \text{ if and only if } x \in A \text{ and } x \in B.$$ 

The set operation intersection corresponds to the logical “and” function.
1.1 Comment: Universal Set ($S$)

This is the set of all things that we could possibly consider in a given context.
1.1 Comment: Set Complement

The *complement* of a set $A$, denoted by $A^c$, is the set of all elements in $S$ that are not in $A$. The complement of $S$ is the null set $\emptyset$. Formally, $x \in A^c$ if and only if $x \notin A$. 

A collection of sets $A_1, \ldots, A_n$ is *mutually exclusive* if and only if

$$A_i \cap A_j = \emptyset, \quad i \neq j.$$  \hspace{1cm} (1)

The word *disjoint* is sometimes used as a synonym for mutually exclusive.
A collection of sets $A_1, \ldots, A_n$ is collectively exhaustive if and only if

$$A_1 \cup A_2 \cup \cdots \cup A_n = S.$$  \hfill (1)
A collection of sets $A_1, \ldots, A_n$ is a *partition* if it is both mutually exclusive and collectively exhaustive.
Theorem 1.1

De Morgan’s law relates all three basic operations:

\[(A \cup B)^c = A^c \cap B^c.\]
Proof: Theorem 1.1

There are two parts to the proof:

- To show \((A \cup B)^c \subseteq A^c \cap B^c\), suppose \(x \in (A \cup B)^c\). That implies \(x \not\in A \cup B\). Hence, \(x \not\in A\) and \(x \not\in B\), which together imply \(x \in A^c\) and \(x \in B^c\). That is, \(x \in A^c \cap B^c\).

- To show \(A^c \cap B^c \subseteq (A \cup B)^c\), suppose \(x \in A^c \cap B^c\). In this case, \(x \in A^c\) and \(x \in B^c\). Equivalently, \(x \not\in A\) and \(x \not\in B\) so that \(x \not\in A \cup B\). Hence, \(x \in (A \cup B)^c\).
Example 1.1 Problem

Phonesmart offers customers two kinds of smart phones, Apricot \((A)\) and Banana \((B)\). It is possible to buy a Banana phone with an optional external battery \(E\). Apricot customers can buy a phone with an external battery \((E)\) or an extra memory card \((C)\) or both. Draw a Venn diagram that shows the relationship among the items \(A,B,C\) and \(E\) available to Phonesmart customers.
Gerlandas offers customers two kinds of pizza crust, Tuscan \((T)\) and Neapolitan \((N)\). In addition, each pizza may have mushrooms \((M)\) or onions \((O)\) as described by the Venn diagram at right. For the sets specified below, shade the corresponding region of the Venn diagram.

(a) \(N\)

(b) \(N \cup M\)

(c) \(N \cap M\)

(d) \(T^c \cap M^c\)
Quiz 1.2

Monitor three consecutive packets going through a Internet router. Based on the packet header, each packet can be classified as either video (v) if it was sent from a Youtube server or as ordinary data (d). Your observation is a sequence of three letters (each letter is either v or d). For example, two video packets followed by one data packet corresponds to vvd. Write the elements of the following sets:

\[ A_1 = \{\text{second packet is video}\}, \quad B_1 = \{\text{second packet is data}\}, \]
\[ A_2 = \{\text{all packets are the same}\}, \quad B_2 = \{\text{video and data alternate}\}, \]
\[ A_3 = \{\text{one or more video packets}\}, \quad B_3 = \{\text{two or more data packets}\}. \]

For each pair of events \( A_1 \) and \( B_1 \), \( A_2 \) and \( B_2 \), and so on, identify whether the pair of events is either mutually exclusive or collectively exhaustive or both.
Section 1.2

Applying Set Theory to Probability
Example 1.2

An experiment consists of the following procedure, observation, and model:

- **Procedure:** Monitor activity at a Phonesmart store.

- **Observation:** Observe which type of phone (Apricot or Banana) the next customer purchases.

- **Model:** Apricots and Bananas are equally likely. The result of each purchase is unrelated to the results of previous purchases.
Example 1.3

Monitor the Phonesmart store until three customers purchase phones. Observe the sequence of Apricots and Bananas.
Example 1.4

Monitor the Phonesmart store until three customers purchase phones. Observe the number of Apricots.
Definition 1.1  Outcome

An outcome of an experiment is any possible observation of that experiment.
Definition 1.2  Sample Space

The sample space of an experiment is the finest-grain, mutually exclusive, collectively exhaustive set of all possible outcomes.
Example 1.5

- The sample space in Example 1.2 is $S = \{a, b\}$ where $a$ is the outcome “Apricot sold,” and $b$ is the outcome “Banana sold.”

- The sample space in Example 1.3 is

  $$S = \{aaa, aab, aba, abb, baa, bab, bba, bbb\}$$

(1)

- The sample space in Example 1.4 is $S = \{0, 1, 2, 3\}$. 
Example 1.6

Manufacture an integrated circuit and test it to determine whether it meets quality objectives. The possible outcomes are “accepted” \((a)\) and “rejected” \((r)\). The sample space is \(S = \{a, r\}\).
Definition 1.3  Event

An event is a set of outcomes of an experiment.
Example 1.7

Suppose we roll a six-sided die and observe the number of dots on the side facing upwards. We can label these outcomes $i = 1, \ldots, 6$ where $i$ denotes the outcome that $i$ dots appear on the up face. The sample space is $S = \{1, 2, \ldots, 6\}$. Each subset of $S$ is an event. Examples of events are

- The event $E_1 = \{\text{Roll 4 or higher}\} = \{4, 5, 6\}$.

- The event $E_2 = \{\text{The roll is even}\} = \{2, 4, 6\}$.

- $E_3 = \{\text{The roll is the square of an integer}\} = \{1, 4\}$. 
Example 1.8

Observe the number of minutes a customer spends in the Phonesmart store. An outcome $T$ is a nonnegative real number. The sample space is $S = \{T | T \geq 0\}$. The event “the customer stays longer than five minutes is $\{T | T > 5\}$.
Section 1.3

Probability Axioms
Definition 1.4  Axioms of Probability

A probability measure \( P[\cdot] \) is a function that maps events in the sample space to real numbers such that

**Axiom 1** For any event \( A \), \( P[A] \geq 0 \).

**Axiom 2** \( P[S] = 1 \).

**Axiom 3** For any countable collection \( A_1, A_2, \ldots \) of mutually exclusive events

\[
P [A_1 \cup A_2 \cup \cdots] = P [A_1] + P [A_2] + \cdots.
\]
Theorem 1.2

For mutually exclusive events $A_1$ and $A_2$,

$$P [A_1 \cup A_2] = P [A_1] + P [A_2].$$
Theorem 1.3

If $A = A_1 \cup A_2 \cup \cdots \cup A_m$ and $A_i \cap A_j = \emptyset$ for $i \neq j$, then

$$P[A] = \sum_{i=1}^{m} P[A_i].$$
Theorem 1.4

The probability measure $P[.]$ satisfies

(a) $P[\emptyset] = 0$.

(b) $P[A^c] = 1 - P[A]$.

(c) For any $A$ and $B$ (not necessarily mutually exclusive),

$$P[A \cup B] = P[A] + P[B] - P[A \cap B].$$

(d) If $A \subset B$, then $P[A] \leq P[B]$. 
Theorem 1.5

The probability of an event $B = \{s_1, s_2, \ldots, s_m\}$ is the sum of the probabilities of the outcomes contained in the event:

$$P[B] = \sum_{i=1}^{m} P[\{s_i\}].$$
Proof: **Theorem 1.5**

Each outcome $s_i$ is an event (a set) with the single element $s_i$. Since outcomes by definition are mutually exclusive, $B$ can be expressed as the union of $m$ mutually exclusive sets:

$$B = \{s_1\} \cup \{s_2\} \cup \cdots \cup \{s_m\}$$  \hspace{1cm} (1)

with $\{s_i\} \cap \{s_j\} = \emptyset$ for $i \neq j$. Applying Theorem 1.3 with $B_i = \{s_i\}$ yields

$$P[B] = \sum_{i=1}^{m} P[\{s_i\}] .$$  \hspace{1cm} (2)
Theorem 1.6

For an experiment with sample space $S = \{s_1, \ldots, s_n\}$ in which each outcome $s_i$ is equally likely,

$$P[s_i] = 1/n \quad 1 \leq i \leq n.$$
Proof: **Theorem 1.6**

Since all outcomes have equal probability, there exists \( p \) such that \( P[s_i] = p \) for \( i = 1, \ldots, n \). Theorem 1.5 implies

\[
P[S] = P[s_1] + \cdots + P[s_n] = np. \tag{1}
\]

Since Axiom 2 says \( P[S] = 1 \), \( p = 1/n \).
Problem 1.3.9

A student’s score on a 10-point quiz is equally likely to be any integer between 0 and 10. What is the probability of an $A$, which requires the student to get a score of 9 or more? What is the probability the student gets an $F$ by getting less than 4?
Section 1.4

Conditional Probability
Example 1.11

Consider an experiment that consists of testing two integrated circuits (IC chips) that come from the same silicon wafer and observing in each case whether a chip is accepted ($a$) or rejected ($r$). The sample space of the experiment is $S = \{rr, ra, ar, aa\}$. Let $B$ denote the event that the first chip tested is rejected. Mathematically, $B = \{rr, ra\}$. Similarly, let $A = \{rr, ar\}$ denote the event that the second chip is a failure.

The chips come from a high-quality production line. Therefore the prior probability $P[A]$ is very low. In advance, we are pretty certain that the second circuit will be accepted. However, some wafers become contaminated by dust, and these wafers have a high proportion of defective chips. When the first chip is a reject, the outcome of the experiment is in event $B$ and $P[A|B]$, the probability that the second chip will also be rejected, is higher than the $a$ priori probability $P[A]$ because of the likelihood that dust contaminated the entire wafer.
Definition 1.5  Conditional Probability

The conditional probability of the event $A$ given the occurrence of the event $B$ is

$$P[A|B] = \frac{P[AB]}{P[B]}.$$
Problem 1.4.8

Deer ticks can carry both Lyme disease and human granulocytic ehrlichiosis (HGE). In a study of ticks in the Midwest, it was found that 16% carried Lyme disease, 10% had HGE, and that 10% of the ticks that had either Lyme disease or HGE carried both diseases.

(a) What is the probability \( P[LH] \) that a tick carries both Lyme disease \((L)\) and HGE \((H)\)?

(b) What is the conditional probability that a tick has HGE given that it has Lyme disease?
Monitor three consecutive packets going through an Internet router. Classify each one as either video \((v)\) or data \((d)\). Your observation is a sequence of three letters (each one is either \(v\) or \(d\)). For example, three video packets corresponds to \(vvv\). The outcomes \(vvv\) and \(ddd\) each have probability 0.2 whereas each of the other outcomes \(vvd, vdv, vdd, dvv, dvd,\) and \(ddv\) has probability 0.1. Count the number of video packets \(N_V\) in the three packets you have observed. Describe in words and also calculate the following probabilities:

(a) \(P[N_V = 2]\)

(b) \(P[N_V \geq 1]\)

(c) \(P\{vvd\}|N_V = 2\)

(d) \(P\{ddv\}|N_V = 2\)

(e) \(P[N_V = 2|N_V \geq 1]\)

(f) \(P[N_V \geq 1|N_V = 2]\)
Partitions and the Law of Total Probability
In this example of Theorem 1.8, the partition is $B = \{B_1, B_2, B_3, B_4\}$ and $C_i = A \cap B_i$ for $i = 1, \ldots, 4$. It should be apparent that $A = C_1 \cup C_2 \cup C_3 \cup C_4$. 
Theorem 1.8

For a partition $B = \{B_1, B_2, \ldots\}$ and any event $A$ in the sample space, let $C_i = A \cap B_i$. For $i \neq j$, the events $C_i$ and $C_j$ are mutually exclusive and

$$A = C_1 \cup C_2 \cup \cdots.$$
Theorem 1.9

For any event $A$, and partition \( \{B_1, B_2, \ldots, B_m\} \),

\[
P[A] = \sum_{i=1}^{m} P[A \cap B_i].
\]
Problem 1.5.3

Suppose a cellular telephone is equally likely to make zero handoffs ($H_0$), one handoff ($H_1$), or more than one handoff ($H_2$). Also, a caller is either on foot ($F$) with probability $5/12$ or in a vehicle ($V$).

(a) Given the preceding information, find three ways to fill in the following probability table:

<table>
<thead>
<tr>
<th></th>
<th>$H_0$</th>
<th>$H_1$</th>
<th>$H_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$V$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) Suppose we also learn that $1/4$ of all callers are on foot making calls with no handoffs and that $1/6$ of all callers are vehicle users making calls with a single handoff. Given these additional facts, find all possible ways to fill in the table of probabilities.
Theorem 1.10  Law of Total Probability

For a partition \( \{B_1, B_2, \ldots, B_m\} \) with \( P[B_i] > 0 \) for all \( i \),

\[
P[A] = \sum_{i=1}^{m} P[A|B_i] P[B_i].
\]
Theorem 1.11 Bayes’ theorem

\[ P[B|A] = \frac{P[A|B]P[B]}{P[A]} \]
Quiz 1.5

Monitor customer behavior in the Phonesmart store. Classify the behavior as buying \((B)\) if a customer purchases a smartphone. Otherwise the behavior is no purchase \((N)\). Classify the time a customer is in the store as long \((L)\) if the customer stays more than three minutes; otherwise classify the amount of time as rapid \((R)\). Based on experience with many customers, we use the probability model \(P[N] = 0.7\), \(P[L] = 0.6\), \(P[NL] = 0.35\). Find the following probabilities:

(a) \(P[B \cup L]\)
(b) \(P[N \cup L]\)
(c) \(P[N \cup B]\)
(d) \(P[LR]\)
Section 1.6

Independence
Definition 1.6  Two Independent Events

Events $A$ and $B$ are independent if and only if

$$P[AB] = P[A]P[B].$$
Problem 1.6.5

In an experiment, $A$ and $B$ are mutually exclusive events with probabilities $\Pr[A] = 1/4$ and $\Pr[B] = 1/8$.
(a) Find $\Pr[A \cap B]$, $\Pr[A \cup B]$, $\Pr[A \cap B^c]$, and $\Pr[A \cup B^c]$.
(b) Are $A$ and $B$ independent?
Problem 1.6.6

In an experiment, \( C \) and \( D \) are independent events with probabilities \( P[C] = 5/8 \) and \( P[D] = 3/8 \).

(a) Determine the probabilities \( P[C \cap D] \), \( P[C \cap D^c] \), and \( P[C^c \cap D^c] \).

(b) Are \( C^c \) and \( D^c \) independent?
Problem 1.6.11

For independent events $A$ and $B$, prove that
(a) $A$ and $B^c$ are independent.
(b) $A^c$ and $B$ are independent.
(c) $A^c$ and $B^c$ are independent.
Definition 1.7  Three Independent Events

$A_1$, $A_2$, and $A_3$ are mutually independent if and only if

(a) $A_1$ and $A_2$ are independent,

(b) $A_2$ and $A_3$ are independent,

(c) $A_1$ and $A_3$ are independent,

More than Two Independent

**Definition 1.8 Events**

If \( n \geq 3 \), the events \( A_1, A_2, \ldots, A_n \) are mutually independent if and only if

(a) all collections of \( n - 1 \) events chosen from \( A_1, A_2, \ldots A_n \) are mutually independent,

(b) \( P[A_1 \cap A_2 \cap \cdots \cap A_n] = P[A_1] \cdot P[A_2] \cdots P[A_n] \).
Quiz 1.6

Monitor two consecutive packets going through a router. Classify each one as video (v) if it was sent from a Youtube server or as ordinary data (d) otherwise. Your observation is a sequence of two letters (either v or d). For example, two video packets corresponds to vv. The two packets are independent and the probability that any one of them is a video packet is 0.8. Denote the identity of packet $i$ by $C_i$. If packet $i$ is a video packet, then $C_i = v$; otherwise, $C_i = d$. Count the number $N_V$ of video packets in the two packets you have observed. Determine whether the following pairs of events are independent:

(a) $\{N_V = 2\}$ and $\{N_V \geq 1\}$

(b) $\{N_V \geq 1\}$ and $\{C_1 = v\}$

(c) $\{C_2 = v\}$ and $\{C_1 = d\}$

(d) $\{C_2 = v\}$ and $\{N_V \text{ is even}\}$