Chapter 1. Preliminaries.

1.1 The setting.

Probability theory is used to model the behavior of random experiments. In this context a random experiment is any process of observation or experimentation for which the particular outcome is not know with certainty in advance of performance of the experiment. For example we might consider the experiment of tossing a die, tossing a pair of dice, flipping a coin once, flipping a coin several times, drawing balls from a box of balls, placing several balls in several boxes, or selecting a person from a population and measuring some relevant characteristic of the person (*e.g.*, the person's sex, height, age, or weight).

Our goal is to formulate a theory which can be used to specify a formal model for the randomness in the outcomes of the experiment by assigning probabilities to events associated with the experiment. An event is a description of the outcome of the experiment. We need to distinguish between simple events and compound events. A simple event is an event which cannot be decomposed into simpler events. We will refer to simple events as elementary outcomes. A compound event is an event which can be decomposed into two or more events. For example, if we toss a die once, then the elementary outcomes, observe a 1, observe a 2, *etc.*, can be represented by the integers 1, 2, 3, 4, 5, and 6. The compound event "observe an even number" is the collection $\{2, 4, 6\}$ of three elementary outcomes.

The first step in forming a probability model for a particular experiment is the specification of a sample space consisting of the collection of all possible elementary outcomes of the experiment. Relevant events can then be viewed as subsets of the sample space. The elementary outcomes (simple events) are the singleton sets (sets containing a single elementary outcome) and compound events are sets containing two or more elementary outcomes.

1.2 Some representative examples.

Example 1. Placement of k balls into n boxes. Consider the experiment of placing k balls into n boxes. We will allow more than one ball to be placed in the same box. We can use a k–tuple $(b_1, \ldots, b_k)$, where the value of the i^{th} element b_i indicates the box into which the i^{th} ball is placed, to represent an elementary outcome for this experiment. Note that we have chosen to represent the elementary outcomes by indicating the locations of the k balls. We will show that there are n^k such elementary outcomes.

With $k = 3$ balls and $n = 2$ boxes (a and b) the $2^3 = 8$ elementary outcomes can be represented as: $aaa, aab, aba, abb, baa, bab, bba, bbb$,
where, for example, aba indicates that the first and third ball are placed in box a and the second ball is placed in box b. The event "the first box contains exactly two balls" is

represented by the set $\{aab, aba, baa\}$; the event "the first box contains at least two balls" is represented by the set $\{aaa, aab, aba, baa\}$; and, the event "both boxes contain at least one ball" is represented by the set $\{aab, aba, abb, baa, bab, bba\}$.

With $k = 3$ balls and $n = 3$ boxes (a, b, and c) the $3^3 = 27$ elementary outcomes can be represented as:
$aaa, aab, aac, aba, abb, abc, aca, acb, acc,$
$baa, bab, bac, bba, bbb, bbc, bca, bcb, bcc,$
$caa, cab, cac, cba, cbb, cbc, cca, ccb, ccc.$

Example 1a. Tossing a coin: If we toss a coin three times, then, letting H denote "heads" and T denote "tails", the 8 elementary outcomes can be represented as:
$HHH, HHT, HTH, HTT, THH, THT, TTH, TTT.$
More generally, tossing the coin k times is equivalent to placing k balls into $n = 2$ boxes and the 2^k elementary outcomes for this experiment can be envisioned as ordered k–tuples of $H's$ and $T's$ analogous to those listed above for $k = 3$.

Example 1b. Tossing a die. Tossing a die k times is equivalent to placing k balls into $n = 6$ boxes and the 6^k elementary outcomes for this experiment can be envisioned as ordered k–tuples of the integers $1, \dots, 6$. For example, with $k = 4$ tosses $(1, 3, 4, 2)$ indicates that the tosses yield 1, 3, 4, and 2 in this order.

Aside: Playing cards, bridge, and poker. *A standard deck of playing cards contains 52 cards. The cards appear in four suits, hearts* $\heartsuit$*, diamonds* $\diamondsuit$*, clubs* $\clubsuit$*, and spades* $\spadesuit$*, of these the heart and diamond cards are colored red and the club and spade cards are colored black. The 13 cards of each suit have face values of* $2, 3, \dots, 10$*, jack, queen, king, and ace. Cards with the same face value are said to be of the same kind. For our purposes bridge means dividing the 52 cards into four hands (sets) of 13 cards and a poker hand is a collection of five cards.*

Example 1c. Aces in bridge. The distribution of the four aces among the four players (hands) in a bridge game is equivalent to the placement of $k = 4$ balls (the four aces) into $n = 4$ boxes (the four players). For example, if we list the aces in the order $\heartsuit\ \diamondsuit\ \clubsuit\ \spadesuit$, then $(1, 1, 2, 4)$ indicates that player 1 receives the $\heartsuit$ ace and the $\diamondsuit$ ace, player 2 receives the $\clubsuit$ ace, player 3 receives no ace, and player 4 receives the $\spadesuit$ ace.

Example 1d. Birthdays. The distribution of birthdays among a group of k people is equivalent to the placement of k balls (the people) into $n = 365$ boxes (the birthdays, assuming for simplicity that the year always has 365 days).

Example 1e. Eye color. The distribution of eye colors among a population of k individuals is equivalent to the placement of k balls (the k individuals) into n boxes (the n possible eye colors).

Example 1f. Elevators. The distribution of the floor at which a passenger exits an elevator, which stops at n floors, among k passengers is equivalent to the placement of k balls (the passengers) into n boxes (the floors).

Example 2. Tossing a coin or die repeatedly. Suppose that a coin is tossed repeatedly until a head occurs. In this case there is a countably infinite collection of elementary outcomes which are of the form: $H, TH, TTH, TTTH, \ldots$
Similarly, if a die is tossed repeatedly until a 1 occurs there is a countably infinite collection of elementary outcomes which are of this same form where H indicates that a 1 occurs and T indicates that a 2, 3, 4, 5, or 6 occurs.

Example 3. Sampling for one attribute. Suppose that a sample of 100 individuals (adults) is selected from a population with the goal of estimating the proportion of adults in the population who possess a 4-year college degree. For this purpose it suffices to use the collection of integers $0, 1, \ldots, 100$ (with the value x indicating that exactly x of the 100 adults possess a degree) to represent the elementary outcomes of this experiment. The event "the majority of the 100 adults possess a degree" can be represented by the set $\{51, \ldots, 100\}$.

Example 4. Sampling for two attributes. Returning to the preceding example consider the relationship between the possession of a college degree and the sex of the adult among the 100 adults in the sample. In this case we can represent an elementary outcome as a quadruple of integers of the form (F_1, F_2, M_1, M_2), with each element taking a value in $\{0, 1, \ldots, 100\}$, where F_1 is the number of females, F_2 is the number of females possessing a degree, M_1 is the number of males, and M_2 is the number of males possessing a degree. Note that we only need to consider quadruples (F_1, F_2, M_1, M_2) with $F_2 \le F_1$ and $M_2 \le M_1$. The event "the proportion of females in the sample who possess a degree is larger than the proportion of males in the sample who possess a degree" corresponds to the collection of these quadruples for which $F_2/F_1 > M_2/M_1$.

Example 5. Sampling for a numerical value or values. Suppose that we are interested in the distribution of weights among individuals in a population. Further suppose that we choose one individual from this population and determine the weight of this individual. We can use a positive real number x to represent an elementary outcome (the weight of an individual) and the positive part of the number line to represent the sample space. Relevant events such as "the individual weighs at least 200 pounds" correspond to intervals,

$[200, \infty)$ for this event, on the number line. If we are interested in the relationship between the weights of husbands and their wives, then we can represent an elementary outcome, corresponding to a particular married couple, by an ordered pair (x, y) where $x > 0$ is the husband's weight and $y > 0$ is the wife's weight. In this case we can use the upper right quadrant of the plane to represent the sample space. In this situation events correspond to regions in this quadrant, *e.g.*, the event "the husband weighs more than his wife" is the triangular region below the line $y = x$.

1.3 Sample space and events.

The examples of Section 1.2 illustrate the process of defining an elementary outcome, a sample space, and events for several specific experiments. Recall that for our purposes an experiment is a process of observation or experimentation which results in one of several possible elementary outcomes. In this section we begin a more formal treatment of these aspects of an experiment. We will assume that all of the possible elementary outcomes of the experiment are known but the actual outcome of the experiment will not be known with certainty until after the experiment is performed.

The sample space Ω of a particular experiment is the collection of all possible elementary outcomes for the experiment. We will assume that these elementary outcomes are mutually exclusive so that two distinct elementary outcomes cannot occur at the same time. A generic elementary outcome will be denoted ω.

An event A is a collection of elementary outcomes, *i.e.* a subset of Ω. For convenience we define the null event (empty set), $\emptyset$, as the event with no elements. When the experiment is conducted and the elementary outcome ω occurs: if $\omega \in A$ (ω is an element of A), then we say that event A has occurred; and if $\omega \notin A$ (ω is not an element of A), then we say that event A has not occurred.

Given two events A and B we write $A \subset B$ (A is contained in B or A is a subset of B) when every element of A is also an element of B. More formally, $A \subset B$ means that if $\omega \in A$, then $\omega \in B$. Note that if $A \subset B$, then the occurrence of A implies the occurrence of B. Also note that for every event A, $A \subset \Omega$, $A \subset A$, and $\emptyset \subset A$.

Two events A and B are said to be equivalent (equal), $A = B$, when they contain the same elementary outcomes. That is, every element of A is an element of B and every element of B is an element of A. Another way to say this is: $A = B$ if and only if $A \subset B$ and $B \subset A$.

Given an event A the complementary event A^c (the complement of A) is the event containing all elements of Ω which do not belong to A, *i.e.*,

$$A^c = \{\omega \in \Omega : \omega \notin A\}$$

Note that A and A^c are mutually exclusive in the sense that they cannot both occur at the same time. Furthermore, since they are complementary, exactly one of A and A^c must occur. Also note that $\Omega^c = \emptyset$ and $\emptyset^c = \Omega$.

The union of events A and B, $A \cup B$, is the collection of elementary outcomes which belong to A or B, *i.e.*,

$$A \cup B = \{\omega \in \Omega : \omega \in A \text{ or } \omega \in B\}.$$

Note that the *or* in this definition is the *logical or* meaning one or the other or both

The intersection of events A and B, AB, is the collection of elementary outcomes which belong to both A and B, *i.e.*,

$$AB = \{\omega \in \Omega : \omega \in A \text{ and } \omega \in B\}.$$

We will generally use the notation AB for the intersection, however, the alternate notation $A \cap B$ may be used for clarity in some contexts.

Example. Toss a coin four times. Letting H denote heads and T denote tails we can express an elementary outcome as a 4–tuple of $H's$ and $T's$. Thus we have the sample space

$$\begin{aligned}\Omega = \{&HHHH, HHHT, HHTH, HHTT, HTHH, HTHT, HTTH, HTTT,\\ &THHH, THHT, THTH, THTT, TTHH, TTHT, TTTH, TTTT\}.\end{aligned}$$

Some representative events are:

The event A "exactly 2 heads occur" is

$A = \{HHTT, HTHT, HTTH, THHT, THTH, TTHH\}$;

The event B "at least 2 heads occur" is

$$\begin{aligned}B = \{&HHHH, HHHT, HHTH, HHTT, HTHH, HTHT,\\ &HTTH, THHH, THHT, THTH, TTHH\};\end{aligned}$$

The event C "an odd number of heads occur" is

$C = \{HHHT, HHTH, HTHH, HTTT, THHH, THTT, TTHT, TTTH\}$.

For these events we have: $A \subset B$ thus $A \cup B = B$ and $A \cap B = A$;

$$\begin{aligned}A^c = \{&HHHH, HHHT, HHTH, HTHH, HTTT,\\ &THHH, THTT, TTHT, TTTH, TTTT\};\end{aligned}$$

$$\begin{aligned}A \cup C = \{&HHHT, HHTH, HHTT, HTHH, HTHT, HTTH, HTTT,\\ &THHH, THHT, THTH, THTT, TTHH, TTHT, TTTH, TTTT\};\end{aligned}$$

$AC = \emptyset$; and, $BC = \{HHHT, HHTH, HTHH, THHH\}$.

Commutative properties. *For any events A and B*
$A \cup B = B \cup A$ *and* $AB = BA$ $(A \cap B = B \cap A)$

Associative properties. *For any events A, B, and C*
(a) $(A \cup B) \cup C = A \cup (B \cup C)$ *(thus the notation $A \cup B \cup C$ is unambiguous)*
(b) $(AB)C = (A \cap B) \cap C = A \cap (B \cap C) = A(BC)$ *(thus the notation ABC or $A \cap B \cap C$ is unambiguous)*

Distributive properties. *For any events A, B, and C*
(a) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ *equivalently* $A(B \cup C) = AB \cup AC$
(b) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ *equivalently* $A \cup (BC) = (A \cup B)(A \cup C)$

Additional properties.
For any event A: $AA = A$, $A \cup A = A$, $A\Omega = A$, $A \cup \Omega = \Omega$, $A\emptyset = \emptyset$, $A \cup \emptyset = A$, $(A^c)^c = A$, $AA^c = \emptyset$, *and* $A \cup A^c = \Omega$.
If $A \subset B$, then $AB = A$ and $A \cup B = B$.

DeMorgan's laws. *For any events A and B:* $(A \cup B)^c = A^c \cap B^c$ *and* $(A \cap B)^c = A^c \cup B^c$.

If $AB = \emptyset$, then A and B are said to be mutually exclusive or disjoint events indicating that they cannot occur at the same time. Note that for any event A we have $A\emptyset = \emptyset$ thus the null event $\emptyset$ has the property that it is disjoint from every event. In fact, the null event is the only event with this property.

Given events A and B we will often find it useful to use event B to decompose (partition) A into disjoint subevents. The relevant decomposition (partition) of A is

$$A = AB \cup AB^c.$$

Note that this simply says that (1) the occurrence of A is equivalent to the occurrence of either A and B or the occurrence of A and the nonoccurrence of B and (2) these two subevents cannot occur at the same time. Also note that if we take $B = A^c$, then we get the partition $\Omega = A \cup A^c$ of Ω into disjoint events.
Given a collection of events $\{A_1, \ldots, A_n\}$:

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n \text{ denotes the union of these events}$$

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \cdots \cap A_n \text{ denotes the intersection of these events}$$

Given a sequence of events $\{A_1, A_2, \ldots\}$:

$$\bigcup_{i=1}^{\infty} A_i \text{ denotes the union of these events}$$

$$\bigcap_{i=1}^{\infty} A_i \text{ denotes the intersection of these events}$$

Given the sequence of events $\{A_1, A_2, \ldots\}$ and an index set I, *i.e.* a set of positive integers,

$$\bigcup_{i \in I} A_i \text{ denotes the union of the } A_i \text{ with subscripts } i \in I$$

$$\bigcap_{i \in I} A_i \text{ denotes the intersection of the } A_i \text{ with subscripts } i \in I$$

DeMorgan's laws. *Given a sequence of events* $\{A_1, A_2, \ldots\}$ *and an index set* I*:*

$$\left(\bigcup_{i \in I} A_i\right)^c = \bigcap_{i \in I} A_i^c \text{ and } \left(\bigcap_{i \in I} A_i\right)^c = \bigcup_{i \in I} A_i^c$$