Section 1.4. Set Theory

Note. In this section we introduce elementary set theory. In the next section we set up the axioms of probability using the language of set theory.

Definition 1.4.1. The collection (or "set") of all possible outcomes of an experiment is the *sample space* of the experiment.

Note. We will consider subsets of the sample space which are called "events." In all of the examples encountered in this chapter, the set of events will include all subsets of the sample space (that is, the set of events is the power set, $\mathcal{P}(S)$, of the sample space). For more general settings, though, we impose three conditions which the set of events must satisfy. We now state the first condition:

Condition 1 of the Set of Events. The sample space S itself must be an event.

Note. Instead of elements of the sample space, we speak of "outcomes" in the sample space. If outcomes a is in event A then we write $a \in A$.

Note. We now review some elementary properties of "naive" set theory. For a more formal treatment, see my online notes on Introductory Set Theory. Here, we mix the language of "sample space," "event," and "outcome" with the more traditional set theoretic terminology.

Definition 1.4.2. Set A is *contained in* set B if every element in A also is in B. This is denoted $A \subset B$ or $B \supset A$ and also stated as B *contains* A. The common set theoretic terminology is that A is a *subset* of B, or B is a *superset* of A.

Definition. Two sets A and B are *equal*, denoted $A = B$, if sets A and B contain exactly the same elements.

Theorem 1.4.1. Let A, B, and C sets. If $A \subset B$ and $B \subset A$, then $A = B$. If $A \subset B$ and $B \subset C$, then $A \subset C$.

Definition. The set consisting of no outcomes is the *empty set*, denoted $\varnothing$.

Note. Trivially, $\varnothing \subset A$ for any set A. DeGroot and Schervish state this (in terms of events) as Theorem 1.4.2.

Note. We now distinguish between two types of infinite sets.

Definition 1.4.4. An infinite set A is *countable* if there is a one-to-one correspondence (that is, a one to one and onto mapping) between the elements of A and the set of natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$. A set is *uncountable* is is is neither finite nor countable. A set has *at most countably many* elements if it is either finite or countable.

Note. It might come a surprise to you that there are "big" and "small" infinite sets. The standard Cantor diagonalization argument can be used to prove that the interval $(0, 1)$ of real numbers is *not* countable. DeGroot and Schervish deal with this on pages 13 and 14. For a more thorough presentation, see my online notes for Analysis 1 (MATH 4217/5217) on 1.3. The Completeness Axiom (see Theorem 1-20). Notice also Cantor's Theorem (Theorem 1-21) in these notes, which implies that there are "larger and larger" infinite sets and that there is not a "largest infinity."

Definition 1.4.5. The *complement* of a set A is the event that contains all elements in the universal set of discourse (the sample space S for our applications) which do not belong to A. This is denoted A^c.

Note. Trivially, $(A^c)^c = A$, $\varnothing^c = U$ (where U is the universal set of discourse, U is the sample space S for our applications), and $U^c = S^c = \varnothing$. DeGroot and Schervish state this (in terms of events) as Theorem 1.4.3.

Note. We are now ready to impose a second conditions on the set of events:

Condition 2 of the Set of Events. If A is an event, then A^c is also an event.

Notice that this implies, by Condition 1, that $\varnothing$ is an event.

Definition 1.4.6. If A and B are any two sets, the *union* of A and B is the set containing all outcomes that belong to A alone, to B alone, or to both A and B. This is denoted $A \cup B$.

Theorem 1.4.4. With S as the sample space (or the "universal set"), we have:

$$A \cup B = B \cup A,\ A \cup A = A,\ A \cup A^c = S,\ A \cup \varnothing = A, \text{ and } A \cup S = S.$$

Definition 1.4.7. The *union of n sets* $A_1, A_2, \ldots, A_n$ is defined to be the set that contains all elements that belong to at least one of these n sets. This is denoted $A_1 \cup A_2 \cup \cdots \cup A_n = \cup_{i=1}^n A_i$. Similarly, the *union of an infinite sequence of sets* $A_1, A_2, \ldots$ is the set that contains all elements that belong to at least one of the sets in the sequence. This is denoted $A_1 \cup A_2 \cup \cdots = \cup_{i=1}^\infty A_i$.

Note. We can now state the final of the three conditions on the set of events:

Condition 3 of the Set of Events. If $A_1, A_2, \ldots$ is a countable collection of events, then $\cup_{i=1}^\infty A_i$ is also an event.

Note. Notice that Conditions 1, 2, and 3 on the set of events requires that the universal set (sample set) S is an event, the complement of an event is an event, and the countable union of a collection of events is an event. In measure theory, a collection of sets satisfying these conditions is called a *σ-algebra of sets*; see my online notes for Real Analysis 1 (MATH 5210) on Section 1.4. Borel Sets. In

Real Analysis 1 we introduce Lebesgue measure on a collection of subsets of the real numbers which form a σ-algebra (the σ-algebra of Lebesgue measurable sets). Lebesgue measure is a generalization of the length of an interval and is used to set up Lebesgue integration, which is a generalization of Riemann integration. A class on modern probability theory requires a knowledge of Lebesgue measure and Lebesgue integration (given in Real Analysis 1, MATH 5210), abstract measure and integration (given in Real Analysis 2, MATH 5220), and Hilbert space, and linear operators on normed linear spaces (given in Fundamentals of Functional Analysis, MATH 5740). You can find my class notes for these classes, as well as notes on probability theory as follows:

- Real Analysis 1
- Real Analysis 2
- Fundamentals of Functional Analysis
- Measure Theory Based Probability

Note. An additional property of the set of events as as follows. The proof simply requires that we replace events $A_{n+1}, A_{n+2}, \ldots$ with $\varnothing$.

Theorem 1.4.5. The union of a finite number of events $A_1, A_2, \ldots, A_n$ is an event.

Theorem 1.4.6. The Associative Property. For sets A, B, C we have $(A \cup B) \cup C = A \cup (B \cup C)$.

Note. Theorem 1.4.6 makes it unambiguous when we write $A \cup B \cup C$ as notation for either $(A \cup B) \cup C$ or $\cup(B \cup C)$. Similarly we denote the union of n events $A_1, A_2, \ldots, A_n$ as $A_1 \cup A_2 \cup \cdots \cup A_n = \cup_{i=1}^n A_i$.

Definition 1.4.8. For sets A and B, the *intersection* of A and B is the set that contains all elements which belong to both A and B. This is denoted $A \cap B$.

Theorem 1.4.7. With S as the sample space (or the "universal set"), we have:

$$A \cap B = B \cap A,\ A \cap A = A,\ A \cap A^c = \varnothing,\ A \cap \varnothing = \varnothing, \text{ and } A \cap S = A.$$

Definition 1.4.9. The *intersection of n sets* $A_1, A_2, \ldots, A_n$ is defined to be the set that contains all elements that belong to all of these n sets. This is denoted $A_1 \cap A_2 \cap \cdots \cap A_n = \cap_{i=1}^n A_i$. Similarly, the *intersection of an infinite sequence of sets* $A_1, A_2, \ldots$ is the set that contains all elements that belong to all of the sets in the sequence. This is denoted $A_1 \cap A_2 \cap \cdots = \cap_{i=1}^\infty A_i$.

Theorem 1.4.8. For sets A, B, C we have $(A \cap B) \cap C = A \cap (B \cap C)$.

Note. Theorem 1.4.8 makes it unambiguous when we write $A \cap B \cap C$ as notation for either $(A \cap B) \cap C$ or $\cap(B \cap C)$. Similarly we denote the union of n events $A_1, A_2, \ldots, A_n$ as $A_1 \cap A_2 \cap \cdots \cap A_n = \cap_{i=1}^n A_i$.

Definition. If sets A and B have no elements in common (that is, $A \cap B = \varnothing$) then A and B are *disjoint* or *mutually exclusive*. The n sets $A_1, A_2, \ldots, A_n$ are *disjoint* if for every $i \neq j$ we have that A_i and A_j are disjoint. An arbitrary collection of sets is *disjoint* if no two of the sets in the collection have any elements in common.

Example 1.4.4. Suppose a coin is tossed three times. Then there are eight possible outcomes (in order, in terms of heads H and tails T) in the sample space: $s_1 : HHH$, $s_2 = THH$, $s_3 : HTH$, $s_4 : HHT$, $s_5 : HTT$, $s_6 : LTHT$, $s_7 : TTH$, $s_8 : TTT$. Let event A be the event that at least one head is obtained. Then $A = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7\}$. Let B be the event that a head is obtained on the second toss. Then $B = \{s_1, s_2, s_4, s_6\}$. Let C be the event that a tail is obtained on the third toss. Then $C = \{s_4, s_5, s_6, s_8\}$. Let D be the event that no heads are obtained. Then $D = \{s_8\}$. Notice that $B \subset A$, $A^c = D$, $B \cap D = \varnothing$, $A \cup C = S$, $B \cap C = \{s_4, s_6\}$, $(B \cup C)^c = \{s_3, s_7\}$, and $A \cap (B \cup C) = \{s_1, s_2, s_4, s_5, s_6\}$.

Theorem 1.4.9. De Morgan's Laws. For every two sets A and B we have

$$(A \cup B)^c = A^c \cap B^c \text{ and } (A \cap B)^c = A^c \cup B^c.$$

Note. The proof of Theorem 1.4.9 is to be given in Exercise 1.4.3. In fact, a more general result holds. For i an indexing set with events A_i for $i \in I$, we can define

$$\cup_{i \in I} A_i = \{a \mid a \in A_i \text{ for all } i \in I\} \text{ and } \cap_{i \in I} A_i = \{a \mid a \in A_i \text{ for all } i \in I\}$$

and then we have $(\cup_{i \in I} A_i)^c = \cap_{i \in I} A_i^c$ and $(\cap_{i \in I} A_i)^c = \cup_{i \in I} A_i^c$. This is to be shown in Exercise 1.4.5.

Note. Notice by Condition 2 and Condition 3 on the set of events, the generalized De Morgan's Law implies that if $A_1, A_2, \ldots, A_n$ are n events then $\cap_{i=1}^{n} A_i$ is an event. Also, if $A_1, A_2 \ldots$ is a sequence of events then $\cap_{i=1}^{\infty} A_i$ is an event. So the collection of events is a collection of sets closed under complements, finite and countable unions, and finite and countable intersections. By the way, this shows that the set of events form a σ-algebra of sets.

Note. The following two results are to be shown in Exercises 1.4.2 and 1.4.4. These will be useful in Section 1.5, "The Definition of Probability."

Theorem 1.4.10. Distributive Properties. For every three sets A, B, C we have

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \text{ and } A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

Theorem 1.4.11. Partitioning a Set. For every two sets A and B, $A \cap B$ and $A \cap B^c$ are disjoint and $A = (A \cap B) \mathbin{\dot\cup} (A \cap B^c)$. In addition, B and $A \cap B^c$ are disjoint and $A \cup B = B \mathbin{\dot\cup} (A \cap B^c)$.

Revised: 7/18/2019