Sets

Math 217 Probability and Statistics

Prof. D. Joyce, Fall 2014

Just a little bit about sets. We'll use the language of sets throughout the course, but we're not using much of set theory. This note just collects the background that you need to know about sets in one place

A set itself is just supposed to be something that has elements. It doesn't have to have any structure but just have elements. The elements can be anything, but usually they'll be things of the same kind.

If you've only got one set, however, there's no need to even mention sets. It's when several sets are under consideration that the language of sets becomes useful.

There are ways to construct new sets, too, and these constructions are important. The most important of these is a way to select some of the elements in a set to form another set, a subset of the first.

Examples. Let's start with sets of numbers. There are ways of constructing these sets, but let's not deal with that now. Let's assume that we already have these sets.

The natural numbers. These are the counting numbers, that is, whole positive numbers. 1 is the first natural number, 2 the second, 3, the third, etc. We'll use $\mathbf{N}$ to denote the set of all natural numbers. Some people like to include 0 in the natural numbers, but I follow Dedekind who started with 1. There is a structure on $\mathbf{N}$, namely there are operations of addition, subtraction, etc., but as a set, it's just the numbers. You'll often see $\mathbf{N}$ defined as

$$\mathbf{N} = \{1, 2, 3, \ldots\}$$

which is read as "$\mathbf{N}$ is the set whose elements are 1, 2, 3, and so forth." That's just an informal way of describing what $\mathbf{N}$ is. A complete description couldn't get away with "and so forth." If you want to see all of what "and so forth" entails, you can read Dedekind's 1888 paper *Was sind und was sollen die Zahlen?* and my comments on it. In that article he starts off developing set theory and ends up with the natural numbers.

The real numbers. These include all positive numbers, negative numbers, and 0. Besides the natural numbers, their negations and 0 are included, fractions like $\frac{22}{7}$, algebraic numbers like $\sqrt{5}$, and transcendental numbers like π and e. If a number can be named decimally with infinitely many digits, then it's a real number. We'll use $\mathbf{R}$ to denote the set of all real numbers. Like $\mathbf{N}$, $\mathbf{R}$ has lots of operations and functions associated with it, but treated as a set, all it has is its elements, the real numbers.

Note that $\mathbf{N}$ is a subset of $\mathbf{R}$ since every natural number is a real number.

Subsets. If you have a set and a language to talk about elements in that set, then you can form subsets of that set by properties of elements in that language.

For instance, we have arithmetic on $\mathbf{R}$, so solutions to equations are subsets of $\mathbf{R}$. The solutions to the equation $x^3 = x$ are 0, 1, and -1. We can describe its solution set using the notation

$$S = \{x \in \mathbf{R} \,|\, x^3 = x\}$$

which is read as "S is the set of x in $\mathbf{R}$ such that $x^3 = x$." We could also describe that set by listing its elements, $S = \{0, 1, -1\}$. When you name a set by listing its elements, the order that you name them doesn't matter. We could have also written $S = \{-1, 0, 1\}$ for the same set.

Open and closed intervals in $\mathbf{R}$ are also subsets of $\mathbf{R}$. For example,

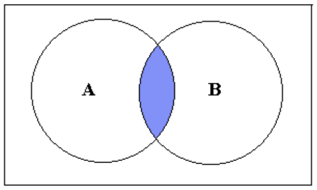
$$\begin{aligned} (3,5) &= \{x \in \mathbf{R} \,|\, 3 < x < 5\} \\ [3,5] &= \{x \in \mathbf{R} \,|\, 3 \leq x \leq 5\} \end{aligned}$$

There are a couple of notations for subsets. We'll use the notation $A \subseteq S$ to say that A is a subset of S. We allow $S \subseteq S$, that is, we consider a set S to be a subset of itself. If a subset A doesn't include all the elements of S, then A is called a *proper* subset of S. The only subset of S that's not a proper subset is S itself. We'll use the notation $A \subset S$ to indicate that A is a proper subset of S.

(Warning. There's an alternate notational convention for subsets. In that notation $A \subset S$ means A is any subset of S, while $A \subsetneq S$ means A is a proper subset of S. I prefer the the notation we're using because it's analogous to the notations $\leq$ for less than or equal, and $<$ for less than.)

Operations on subsets. Frequently you deal with several subsets of a set, and there are operations of intersection, union, and difference that describe new subsets in terms of previously known subsets.

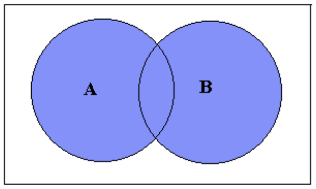
The intersection $A \cap B$ of two subsets A and B of a given set S is the subset of S that includes all the elements that are in both A and B:



$$A \cap B = \{x \in S \,|\, x \in A \text{ and } x \in B\}.$$

Our textbook uses the nonstandard notation AB instead of $A \cap B$ for intersection.

The union $A \cup B$ of two subsets A and B of a given set S is the subset of S that includes all the elements that are in A or in B or in both:



$$A \cup B = \{x \in S \,|\, x \in A \text{ or } x \in B\}.$$

As usual in mathematics, the word "or" means an inclusive or and implicitly includes "or both."

The difference $A - B$ of two subsets A and B of a given set S is the subset of S that includes all the elements that are in A but not in B:

$$A - B = \{x \in S \,|\, x \in A \text{ and } x \notin B\}$$

There's also the complement of a subset A of a set S. The complement is just $S - A$, all the elements of S that aren't in A. When the set S is understood, the complement of A often is denoted more simply as either A^{c}, $\overline{A}$, or A' rather than $S - A$. Set complement A^{c}.

These operations satisfy lots of identities. I'll just name a couple of important ones.

DeMorgan's laws describe a duality between intersection and union. They can be written as

$$\begin{aligned}(A \cap B)^{\mathsf{c}} &= A^{\mathsf{c}} \cup B^{\mathsf{c}} \\ (A \cup B)^{\mathsf{c}} &= A^{\mathsf{c}} \cap B^{\mathsf{c}}\end{aligned}$$

The distributivity laws say that intersection and union each distribute over the other

$$\begin{aligned}(A \cap B) \cup C &= (A \cup C) \cap (B \cup C) \\ (A \cup B) \cap C &= (A \cap C) \cup (B \cap C)\end{aligned}$$

Unions and intersections sometimes are taken of many subsets, even infinitely many. Suppose that $A_1, A_2, \ldots, A_n$ are subsets of S. The intersection of all of them can be written in an indexed notation as

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \cdots \cap A_n$$

and their union as

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n.$$

And when there are infinitely many, $A_1, A_2, \ldots, A_n, \ldots$, as

$$\bigcap_{i=1}^{\infty} A_i = \{x \in S \,|\, x \in A_i \text{ for all } i\}$$

and their union as

$$\bigcup_{i=1}^{\infty} A_i = \{x \in S \,|\, x \in A_i \text{ for at least one } i\}.$$

DeMorgan's laws and the distributivity laws also apply to indexed intersections and unions.

$$\begin{aligned}
\left(\bigcap_{i=1}^{n} A_i\right)^{\mathsf{c}} &= \bigcup_{i=1}^{n} A_i{}^{\mathsf{c}} \\
\left(\bigcup_{i=1}^{n} A_i\right)^{\mathsf{c}} &= \bigcap_{i=1}^{n} A_i{}^{\mathsf{c}} \\
\left(\bigcap_{i=1}^{n} A_i\right) \cup C &= \bigcap_{i=1}^{n} (A_i \cup C) \\
\left(\bigcup_{i=1}^{n} A_i\right) \cap C &= \bigcup_{i=1}^{n} (A_i \cap C)
\end{aligned}$$

Partitions. A set S is said to be *partitioned* into subsets $A_1, A_2, \ldots, A_n$ when each element of S belongs to exactly one of the subsets $A_1, A_2, \ldots, A_n$. That's logically equivalent to saying that S is the disjoint union of the $A_1, A_2, \ldots, A_n$.

When you have a partition $A_1, A_2, \ldots, A_n$ of a set S like that, it induces a partition $E \cap A_1, E \cap A_2, \ldots, E \cap A_n$ on each subset E of S. Each element of E belongs to exactly one of its subsets $E \cap A_1, E \cap A_2, \ldots, E \cap A_n$.

Products of sets. So far we've looked at creating sets within set. There are some operations on sets that create bigger sets, the most important being creating products of sets. These depend on the concept of ordered pairs of elements. The notation for ordered pair (a, b) of two elements extends the usual notation we use for coordinates in the xy-plane. The important property of ordered pairs is that two ordered pairs are equal if and only if they have the same first and second coordinates:

$$(a, b) = (c, d) \text{ iff } a = c \text{ and } b = d.$$

The product of two sets S and T consists of all the ordered pairs where the first element comes from S and the second element comes from T:

$$S \times T = \{(a, b) \,|\, a \in S \text{ and } b \in T\}.$$

Thus, the usual xy-plane is $\mathbf{R} \times \mathbf{R}$, usually denoted $\mathbf{R}^2$.

Besides binary products $S \times T$, you can analogously define ternary products $S \times T \times U$ in terms of triples (a, b, c) where $a \in S$, $b \in T$, and $c \in U$, and higher products, too.

Sets of subsets; power sets. Another way to create bigger sets is to form sets of subsets. If you collect all the subsets of a given set S into a set, then the set of all those subsets is called the *power set* of S, denoted $\mathcal{P}(S)$ or sometimes 2^S.

For example, let S be a set with 3 elements, $S = \{a, b, c\}$. Then S has eight subsets. There are three singleton subsets, that is, subsets having exactly one element, namely $\{a\}$, $\{b\}$, and $\{c\}$. There are three subsets having exactly two elements, namely $\{a, b\}$, $\{a, c\}$, and $\{b, c\}$. There's one subset having all three elements, namely S itself. And there's one subset that has no elements. You could denote it $\{\}$, but it's always denoted $\emptyset$ and called the *empty set* or *null set*. Thus, the power set of S has eight elements

$$\mathcal{P}(S) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, S\}.$$

Countable *versus* uncountable sets. As Georg Cantor (1845–1918) discovered, not all infinite sets have the same cardinality. Some infinite sets are bigger than others.

Two sets S and $T$are said to have the *same cardinality* if there is a one-to-one correspondence of their elements. That means that there is some function $f : S \to T$ which is injective (also called one-to-one) and surjective (also called onto). A function which is both injective and surjective is called a *bijection.* For a bijection $f : S \to T$, the inverse function $f^{-1} : T \to S$ is also a bijection.

The smallest size an infinite set can be is that of the natural numbers $\mathbf{N}$. A set that has the same cardinality as $\mathbf{N}$ is called a *countably infinite* set.

An infinite set that doesn't have the same cardinality as $\mathbf{N}$ is called an *uncountable* set. The set of real numbers $\mathbf{R}$ is uncountable.

Finite sets are also said to be countable. Thus, a set is *countable* if it's either finite or countably infinite.

Series. From calculus, you're familiar with adding countably many numbers. A *series* is a countably infinite sum.

The first series you saw were geometric series. A geometric series is one where the ratio of the next term to the present term is the same for all terms. The general form for a geometric series is

$$a + ar + ar^2 + \cdots + ar^n + \cdots$$

where a is the first term of the series and r is the ratio for the series. A geometric series converges if and only if $-1 < r < 1$, in which case its sum is $\dfrac{a}{1-r}$.

Although you can sometimes add countably infinitely many positive numbers (like that geometric series), it is always the case that an uncountable sum of positive numbers diverges to infinity.

Math 217 Home Page at `http://math.clarku.edu/~djoyce/ma217/`