# There Is More Than One Size of Infinity

January 12, 2017

I think there are a number of things that any good student of science or engineering should be aware of, not only for practical reasons but because they represent major milestones in human intellectual history. Among these things includes the notion of different sizes of infinity.

It isn't completely obvious how to compare the relative sizes of sets. One way to do it would be to declare the set $A$ to be at least as big as the set $B$ if $B$ is a subset of $A$. For example, if $A = \{x, y, z\}$ and $B = \{x, z\}$ then $B$ is smaller than $A$ in any reasonable sense. We will use this notation for comparing sizes of sets: $|B| < |A|$ means "$B$ is smaller than $A$".

Of course that way of comparing set sizes is limited only to pairs of sets in which one is a subset of another. So there needs to be a way to compare the size of any two sets. For finite sets the obvious thing to do is just to count the number of elements. But that won't work for infinite sets, so I'm going to introduce another way to compare the sizes of sets.

**Definition:** For sets $E$ and $F$ we say:

- $|E| \leq |F|$ if there is *any* 1-1 function $f : E \to F$.
- $|E| < |F|$ if there is *no* 1-1 function from $E$ into $F$ whose range is all of $F$.
- If $|E| \leq |F|$ and $|F| \leq |E|$, we say $|E| = |F|$.

This method obviously works for comparing any two finite sets, especially when the sets are written side-by-side in a table. For example, $C = \{x, y, 4, w\}$ is clearly bigger than $A$. Here are the sets in table form:

| $A$ | | $C$ |
|---|---|---|
| $x$ | $\leftrightarrow$ | $x$ |
| $y$ | $\leftrightarrow$ | $y$ |
| $z$ | $\leftrightarrow$ | 4 |
| | $\leftrightarrow$ | w |

The function that assigns $x$ to $x$, $y$ to $y$, and $z$ to 4 is clearly 1-1. So $|A| \leq |C|$. There is no 1-1 function (in fact no function at all) from $A$ to $C$ whose range is all of $C$. There aren't enough elements in $A$. So $|A| < |C|$, in accordance with our common sense.

If $D = \{r, q, 5, 23\}$, a table of the equally-sized sets $C$ and $D$ would be:

| $C$ | | $D$ |
|---|---|---|
| $x$ | $\leftrightarrow$ | $r$ |
| $y$ | $\leftrightarrow$ | $q$ |
| 4 | $\leftrightarrow$ | 5 |
| w | $\leftrightarrow$ | 23 |

The function that assigns elements in the left column to the elements in the same row in the right column is 1-1, so $|C| \leq |D|$. And the function that assigns elements in the right column to the elements in the same row in the left column is also 1-1, so $|D| \leq |C|$. Therefore $|E| = |F|$, again in accordance with our common sense

This all seems silly for finite sets. The real utility of this method is also useful because it can be extended to infinite sets.

The natural numbers themselves, $\mathbb{N}$, make up an infinite set. If you consider the set $\mathbb{E}$ of *even* natural numbers, you could argue that this is a smaller set than $\mathbb{N}$. That would be true in the sense that the even numbers are a proper subset of all the natural numbers. But we aren't using the subset method. We are using the 1-1 correspondence method. The function $f(n) = 2n$ is a 1-1 function from the natural numbers to the even numbers. Also, the function $g(m) = m/2$ is a 1-1 function from the even numbers to the natural numbers. So $|\mathbb{E}| = |\mathbb{N}|$. In table form:

| $\mathcal{N}$ | | $\mathcal{E}$ |
|---|---|---|
| 1 | $\leftrightarrow$ | 2 |
| 2 | $\leftrightarrow$ | 4 |
| 3 | $\leftrightarrow$ | 6 |
| 4 | $\leftrightarrow$ | 8 |
| $\vdots$ | | $\vdots$ |

If a set can be put in this kind of 1-1 correspondence with the natural numbers we way that set is *countable*. In class I've also said that countable means "the set can be written as a list".

The integers $\mathbb{Z}$ are countable. It is annoying to specify the two 1-1 functions explicitly, but here is the argument-via-table:

| $\mathcal{N}$ | | $\mathcal{I}$ |
|---|---|---|
| 1 | $\leftrightarrow$ | 0 |
| 2 | $\leftrightarrow$ | 1 |
| 3 | $\leftrightarrow$ | -1 |
| 4 | $\leftrightarrow$ | 2 |
| 5 | $\leftrightarrow$ | -2 |
| 6 | $\leftrightarrow$ | 3 |
| 7 | $\leftrightarrow$ | -3 |
| 8 | $\leftrightarrow$ | 4 |
| $\vdots$ | | $\vdots$ |

Are there infinite sets that aren't countable? Yes, lots! The real numbers are not countable. Consider just the reals in $(0,1)$. Suppose you *could* list all of them. Then you'd have a list as follows:

| $\mathcal{N}$ | | $\mathcal{R}$ |
|---|---|---|
| 1 | $\leftrightarrow$ | 0.2098438... |
| 2 | $\leftrightarrow$ | 0.8723456... |
| 3 | $\leftrightarrow$ | 0.2836350... |
| 4 | $\leftrightarrow$ | 0.9462734... |
| 5 | $\leftrightarrow$ | 0.4919983... |
| 6 | $\leftrightarrow$ | 0.0873545... |
| $\vdots$ | | $\vdots$ |

The important thing to note is that the decimal expansions for the real numbers could be non-terminating and non-repeating, as would be the case for any irrational number.

Anyway, the function that assigns the natural numbers in the left column to this so-called complete list of reals in the right column is certainly 1-1. So $|\mathbb{N}| \leq |\mathbb{R}|$.

Could such a hypothetical list in the right column possibly be complete? No. It is easy to constuct a number $x^*$ not on the list. Just make the $k$th digit of $x^*$ anything different from the $k$th digit of the $k$th number in the list. For example, the number

$$x^* = 0.384305\ldots$$

differs in (at least) the $k$th digit of the 6 real numbers in the table. In fact, I just took each digit in bold type from the table reproduced below and added 1 to it.

| $\mathcal{N}$ | | $\mathcal{R}$ |
|---|---|---|
| 1 | $\leftrightarrow$ | $0.\mathbf{2}098438\ldots$ |
| 2 | $\leftrightarrow$ | $0.8\mathbf{7}23456\ldots$ |
| 3 | $\leftrightarrow$ | $0.28\mathbf{3}6350\ldots$ |
| 4 | $\leftrightarrow$ | $0.946\mathbf{2}734\ldots$ |
| 5 | $\leftrightarrow$ | $0.4919\mathbf{9}83\ldots$ |
| 6 | $\leftrightarrow$ | $0.08735\mathbf{4}5\ldots$ |
| $\vdots$ | | $\vdots$ |

In other words, there is *no 1-1 function* from $\mathcal{N}$ to $\mathcal{R}$ whose range includes all of $\mathcal{R}$.

Conclusion: $(0, 1)$ is uncountable. By the same argument, any interval of real numbers is uncountable.

So in fact, there are different sizes of infinite sets. This fact has a substantial impact on a course in probability and statistics. We have to deal with finite and countable sample spaces differently than we deal with uncountable sample spaces.

These ideas were first elucidated by Georg Cantor in the late 1800s, who developed what is now a standard mathematical technique used in my demonstration of the uncountability of the real numbers. It's called Cantor's Diagonal Argument. For more details, you should be able to figure out what to Google, or you can ask me. All this theory was enormously controversial when it was first published. Now we teach it to undergraduates.

A fact that might alarm you is that the rational numbers, $\mathbb{Q}$, are countable! In other words, $\mathbb{Q}$ can be written as a list. Here's how. The following array contains all possible rational numbers (in fact it contains all of them many times, but that doesn't matter):

$$\begin{array}{ccccccc}
\frac{1}{1} & -\frac{1}{1} & \frac{2}{1} & -\frac{2}{1} & \frac{3}{1} & -\frac{3}{1} & \cdots \\
\frac{1}{2} & -\frac{1}{2} & \frac{2}{2} & -\frac{2}{2} & \frac{3}{2} & -\frac{3}{2} & \cdots \\
\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} & \frac{3}{3} & -\frac{3}{3} & \cdots \\
\frac{1}{4} & -\frac{1}{4} & \frac{2}{4} & -\frac{2}{4} & \frac{3}{4} & -\frac{3}{4} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

Now, to write a "list" of all rational numbers, follow this pattern:

$$\begin{array}{ccccccc}
\frac{1}{1} & -\frac{1}{1} & \frac{2}{1} & -\frac{2}{1} & \frac{3}{1} & -\frac{3}{1} & \cdots \\
\frac{1}{2} & -\frac{1}{2} & \frac{2}{2} & -\frac{2}{2} & \frac{3}{2} & -\frac{3}{2} & \cdots \\
\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} & \frac{3}{3} & -\frac{3}{3} & \cdots \\
\frac{1}{4} & -\frac{1}{4} & \frac{2}{4} & -\frac{2}{4} & \frac{3}{4} & -\frac{3}{4} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

Here is the table:

| $\mathcal{N}$ | | $\mathcal{Q}$ |
|---|---|---|
| 1 | $\leftrightarrow$ | $\frac{1}{1}$ |
| 2 | $\leftrightarrow$ | $-\frac{1}{1}$ |
| 3 | $\leftrightarrow$ | $\frac{1}{2}$ |
| 4 | $\leftrightarrow$ | $\frac{2}{1}$ |
| 5 | $\leftrightarrow$ | $-\frac{1}{2}$ |
| 6 | $\leftrightarrow$ | $\frac{1}{3}$ |
| 7 | $\leftrightarrow$ | $-\frac{2}{1}$ |
| 8 | $\leftrightarrow$ | $\frac{2}{2}$ |
| $\vdots$ | | $\vdots$ |

Amazingly, even the set of all *algebraic* numbers is countable too! What's an algebraic number? An algebraic number is a root of a polynomial $a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$ (where the $a_i$ are all integers) . For example, $\sqrt{2}$ is algebraic, since it is a root of $x^2 - 2$. Now, there are lots of polynomials, so there are lots and lots of algebraic numbers (most of them irrational like $\sqrt{2}$. But the set of algebraic numbers is still countable. Some fabulous numbers like $\pi$ and $e$ are not algebraic.

To see what's going on here, note that any rational number can be expressed in terms of two integers (the numerator and the denominator). The rationals are countable. Similarly, any algebraic number can represented by a finite number of integers (the $a_i$ in the relevant polynomial). The algebraic numbers are countable. But a real number cannot (in general) be represented by a finite number of integers or natural numbers. It turns out these non-algebraic real numbers are the most plentiful of all.

An obvious question (whose answer turned out to be quite shocking) is: Is there a set that is bigger than the natural numbers but smaller than the real numbers? There is no easy way to answer that question, except to say that the existence of such a set can neither be proved nor disproved using the accepted methods of mathematics. That doesn't mean it's too hard to prove. It means you actually can't do it—it has been proven that you can't!