Miscellaneous results involving countability

Alex Fu

2023-02-04

I'm back from a two-week hiatus.

"It is better to have loved and lost than never to have loved at all." The former certainly isn't better in *every* sense, and I sincerely doubt most who use this proverb mean "better in every sense." But what *is* better than having never loved is the experience and memories of having once loved.

So, in the same manner, I think it is better to have written something mediocre than never to have written at all. It's *far* from better in every sense, but at least I'll have something to show for it, some work I can point to and say I haven't wasted my years. Better done than perfect.

Well, enough about birds in hands and birds in bushes. We have some countability to discuss.

1 Countably additive

The problem of *measurement* is one of the most basic problems in all of mathematics. The original purposes of numbers are to count, label, and *measure*; one of the oldest problems in classical geometry is to measure the length, area, or volume of a figure in space.

Measure theory is the formal study of *measures*, a generalization of the familiar measurements of quantity, length, area, volume, duration, and mass. Leaving the issue of uncertainty in measurement to science, the subject begins by answering two questions:

1. What are the rules of measuring?
2. What objects can be measured?

The first question has a rather simple answer: The measures of unrelated objects are unrelated. For example, knowing the length of a stick or the mass of a stone tells us nothing about the measure of any other object. In some sense, measurements are a bit arbitrary, so we should think of measures as "values we assign" as the measurer, which is a more flexible framework than "measures are values objectively determined."

(For example, while length seems like it is an empirical measure, choosing its unit or "what exactly has length 1" is an entirely arbitrary choice. We are also allowed to "weight" the measures of objects as we please, e.g. in currency, where we might declare that two coins with the same mass have different *value*, because one is made out of gold.)

In particular, to measure the "combination" of unrelated objects, it is enough to measure each object separately, then combine the separate measurements. This motivates the property of **additivity**: the measure of a disjoint union is the sum of individual measures.

The second question also has a straightforward answer that floats to the surface at this point: What is measured are *collections*. We normally measure collections of pieces, parts, components, atoms, or points — like shapes, solids, rocks, apples, etc. In particular, if we wanted to measure a single point, like an atom or an instant, we can treat it as a singleton collection, though usually these have negligible measure.

We describe collections in math using sets, and combining collections, like stacking blocks or putting two apples next to each other, using the operation of disjoint union. Sets also have boolean operations: unions, intersections, complements, so if certain collections can be measured, then their unions, intersections, and complements should also be measurable. That is, **measurability** should be preserved under set operations.

At this point, we have a fairly physically meaningful and intuitive definition of a measure as a *finitely additive set function on an algebra*. However, it is here that measure theory makes a decisive design choice: Measures should be **countably** additive, and measurability should be closed under **countable** set operations.

Why? It seems as if we sacrificed a more natural condition for a stranger requirement, but this concession has a key advantage: we can work with sequences and **limits**, the main object of study of mathematical analysis. Our measures are already idealized constructions that ignore the physical uncertainties present in measurement, so this additional assumption does not hurt as much as it gifts. We need not look further than Zeno's paradoxes of motion for a reason to consider countable infinities.

(Eventually, along a different path of development from proportions, ratios, percentages, and (relative) frequencies, the field historically rejected as "not mathematics" called probability theory would gain a solid foundation by adopting the language of measure theory. Kolmogorov's axioms would also bring about many analogies between probability and the physical measures of mass, length, area, density, etc., which come up in mass functions, Venn diagrams, balance equations, moments, and many other places we will explore in other days.)

2 Countably generated

Previously, in *2023-01-14*, we saw that the Borel σ-algebra $\mathcal{B}$ on $\mathbb{R}$ is generated by any or all intervals of the form $(a, b]$, $[a, b)$, (a, b), $[a, b]$, $(-\infty, b)$, $(-\infty, b]$, (a, ∞), and $[a, \infty)$. But, we can be even stingier, leaner and meaner, more aggressive in reducing the generating set — it is enough to have intervals with rational endpoints.

If we recall the proof from last time, closed endpoints are "created" by countable intersections of open sets (as unions of open sets remain open), and open endpoints by countable unions of closed sets. The same argument applies when we restrict the prelimiting intervals to ones with rational endpoints.

Topologically speaking, we use the fact that $\mathbb{Q}$ is a countable dense subset of $\mathbb{R}$, i.e. $\mathbb{R}$ is separable; $\mathbb{R}$ is a metric space, in which closedness is equivalent to sequential closedness; and the combined fact that every point in $\mathbb{R}$ is the limit of some sequence in $\mathbb{Q}$. Moreover, because every sequence in $\mathbb{R}$ has a monotone subsequence, every $x \in \mathbb{R}$ is a monotone limit of $\mathbb{Q}$.

But, a metric space is separable iff it is *second-countable*, i.e. has a countable basis. This equivalent condition gives a simpler description for when the Borel σ-algebra has a countable generating set.

Proposition 1.

The Borel σ-algebra for a second-countable topological space is countably generated.

Proof. Let $\mathcal{C}$ be the countable basis for τ. Then $\sigma(\mathcal{C}) \subseteq$ the Borel σ-algebra $\sigma(\tau)$, and every open set $O \in \tau$ is a countable union in $\mathcal{C}$, so $\tau \subseteq \sigma(\mathcal{C})$ in the other direction, and we win. $\square$

What's more, $\{(-\infty, x] : x \in \mathbb{Q}\}$ is a generating π-system! Probability measures are certainly σ-finite, so this means that distributions on $\mathbb{R}$ are uniquely determined by the values of their distribution functions on $\mathbb{Q}$.

3 Countably determined

How many probability distributions on $\mathbb{R}$ are there? We have seen that distributions are in bijection with cumulative distribution functions, and we have also seen that $\mathcal{B}$ admits a countable generating π-system. So, let us check the property we expect to hold right now:

Proposition 2.

Let F, G be CDFs that agree on the rationals: $F(q) = G(q)$ for all $q \in \mathbb{Q}$. Then $F(x) = G(x)$ everywhere.

Proof. It suffices to use the denseness of $\mathbb{Q}$ in $\mathbb{R}$ and the right-continuity of CDFs. If $F(x) \neq G(x)$ for some $x \in \mathbb{R}$, then consider any descending rational sequence $q_n \downarrow x$. Then $F(q_n) \downarrow F(x)$ and $G(q_n) \downarrow G(x)$, but $F(q_n) = G(q_n)$ for all n, so their limits must agree as well, $F(x) = G(x)$. $\square$

This provides an injection from the space of distributions on $\mathbb{R}$ to $[0,1]^{\mathbb{Q}}$, which has cardinality $\mathfrak{c}$, the same as the continuum. In fact, $\{F(q)\}_{q \in \mathbb{Q}}$ must be a nondecreasing sequence. We might ask the natural converse question: given any nondecreasing sequence in $[0,1]^{\mathbb{Q}}$, does there exist a CDF "extension" of that sequence to $\mathbb{R}$? If such an extension exists, then it must be unique. If not, then what additional properties does the sequence need?

Apparently there are some difficulties with the natural construction of $F(x) \coloneqq \inf\{F(q) : q > x\}$, which is monotone but not necessarily right-continuous, as we recall from the proof of Helly's selection theorem. For example, $F = \mathbb{1}_{q>0}$ on the rationals is certainly a nondecreasing sequence, but this construction leads to $F = \mathbb{1}_{x>0}$ on $\mathbb{R}$, which is not right-continuous at 0.

One possible resolution is to overwrite the original sequence, i.e. take $\bar{F}(x) \coloneqq \inf\{F(q) : q > x\}$ for all $x \in \mathbb{R}$, not just $x \notin \mathbb{Q}$. I believe $\bar{F}$ will be right-continuous, although it no longer agrees on $\mathbb{Q}$ with the original sequence. It seems likely that the bijection is really between CDFs and nondecreasing sequences $\{F(q)\}_{q \in \mathbb{Q}} \subseteq [0,1]^{\mathbb{Q}}$ such that $F(r) = \inf\{F(q) : q > r\}$ for all $r \in \mathbb{Q}$.

(Proposed proof: In the forward direction, we simply restrict F to $\mathbb{Q}$, where $F(r) = \inf\{F(x) : x > r\} = \inf\{F(q) : q > r \wedge q \in \mathbb{Q}\}$ for all $r \in \mathbb{Q}$ by the right-continuity of F. In the reverse direction, we extend the given sequence by $\bar{F}(x) \coloneqq \inf\{F(q) : q > x\}$, which should give a unique CDF in $\bar{F} \colon \mathbb{R} \to [0,1]$.)

And, the cardinality of both is not just $\leq \mathfrak{c}$, but equal to $\mathfrak{c}$; there are at least uncountably many CDFs $\mathbb{1}_{x \geq x_*}$, given by the point mass distribution, one for every point in $\mathbb{R}$.

4 Countably supported

The support of a random variable whose image is (surely) countable is the support of its probability mass function, the set of points where $\mathbb{P}(X = x)$ is nonzero. Likewise, the support of a continuous random variable — one which admits a density, whose distribution is absolutely continuous with respect to the Lebesgue measure, whose CDF is differentiable — is the support of its probability density function.

What is the support of a general random variable? This definition should allow us to identify "discrete" with "countably supported," even if the converse of "continuous $\to$ uncountable support" does not hold. Two equivalent definitions: the support of X is the set of points $x \in \mathbb{R}$ where Ball(x, r) has positive measure for every $r > 0$; and the support of X is the minimal closed set C such that X lies in C almost surely.

Some points of confusion: taking the limit as $r \downarrow 0$, does Ball$(x, r) \downarrow \{x\}$ not imply that $\mathbb{P}(X = x) > 0$ by continuity from above? Not necessarily, as the limit of a positive sequence of real numbers may still be 0, e.g. $(\frac{1}{n})_{n \geq 1}$. And, if X is $\mathbb{Q}$-valued, e.g. $X(\omega) = \omega \cdot \mathbb{1}_{\mathbb{Q}}$, would the second definition not mean that support$(X) =$ cl$(\mathbb{Q}) = \mathbb{R}$? Perhaps this is intended, but it would be nice to be able to identify discreteness with countability, even if this identification paints an oversimplified picture.

What *is* desirable in the definition of the support is the ability to consider random variables only up to almost sure equivalence — if X is almost surely constant or $\mathbb{N}$-valued, we should not identify X as being "not discrete" based on the naïve consideration of its (pointwise) image.

Some other results on countability — a CDF admits at most countably many (jump) discontinuities, or a measure has at most countably many atoms — will be left to another time. For now, good night.

■