ℕ as a probability space

Saúl Rodríguez

This document is mostly based on the paper 'Sets of recurrence of $\mathbb{Z}^m$-actions and properties of sets of differences in $\mathbb{Z}^m$', by V. Bergelson.

1 Densities in ℕ

There are several ways in which we can compare the sizes of infinite sets of natural numbers. Today we will focus on densities:

Definition 1.1. The *natural density* of a set E of natural numbers is

$$d(E) = \lim_{N\to\infty} \frac{|E \cap \{1,\dots,N\}|}{N} \in [0,1],$$

if this limit is defined.

Density is somehow similar to a probability measure in the set of natural numbers, as it satisfies the following properties:

- $d(\varnothing) = 0, d(\mathbb{N}) = 1$.
- $d(A \cup B) = d(A) + d(B)$ if A, B are disjoint and have density.
- Translation invariance: $d(A+1) = d(A)$,[1] if $d(A)$ is defined.

Example 1.2. $d(3\mathbb{N}) = \frac{1}{3}$, $d(\text{prime numbers}) = 0$, $d(\text{squarefree numbers}) = \frac{6}{\pi^2}$.

Not all sets of natural numbers have density, as we will check in a moment. What is always defined is *upper density* and *lower density*:

$$\overline{d}(E) = \overline{\lim}_{N\to\infty} \frac{|E \cap \{1,\dots,N\}|}{N}$$
$$\underline{d}(E) = \underline{\lim}_{N\to\infty} \frac{|E \cap \{1,\dots,N\}|}{N}.$$

Thus, a set E has density $d(E) = a$ iff $\overline{d}(E) = \underline{d}(E) = a$.

Example 1.3. The set $E = \cup_{N\in\mathbb{N}}[(2N)!, (2N+1)!]$ satisfies $\overline{d}(E) = 1, \underline{d}(E) = 0$, so it does not have density. Draw picture of the set in blackboard.

For the set E of the previous example, we have $\overline{d}(E) = \overline{d}(E^c) = 1$, so that

$$\overline{d}(E \cup E^c) = 1 \neq \overline{d}(E) + \overline{d}(E^c).$$

So $\overline{d}$ does not exactly behave like a probability measure.

[1] By $A+1$ we mean $\{A+1; a \in A\}$

2 A translation-invariant density in $\mathcal{P}(\mathbb{N})$

Using the axiom of choice, we can turn $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ into a finitely additive probability measure space. Let's see how. We will need the following as a black box:

Theorem 2.1 (Existence of non-principal ultrafilters)**.** *There exists a finitely additive probability measure ω in $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ such that $\omega(A) \in \{0,1\}$ for all $A \subseteq \mathbb{N}$, and $\omega(A) = 0$ if A is finite.*

It turns out you can use ultrafilters to define the limit of an arbitrary bounded sequence of real numbers (even if it is not convergent in the usual sense). Let $L^\infty(\mathbb{N})$ be the family of bounded sequences of real numbers.

Theorem 2.2 (Taking limits of bounded sequences)**.** *There is a operator*

$$\begin{aligned}\lim_{n\to\omega} : L^\infty(\mathbb{N}) &\to \mathbb{R}\\ (a_n)_{n\in\mathbb{N}} &\mapsto \lim_{n\to\omega} a_n,\end{aligned}$$

satisfying the following properties:

1. $\lim_{n\to\omega} a_n + b_n = \lim_{n\to\omega} a_n + \lim_{n\to\omega} b_n$.
2. $\lim_{n\to\omega} ka_n = k \lim_{n\to\omega} a_n$.
3. $\underline{\lim}_n a_n \leq \lim_{n\to\omega} a_n \leq \overline{\lim} a_n$ *(so* $\lim_{n\to\omega} a_n = \lim_n a_n$ *if the limit is defined).*

Proof sketch. One can check that for any bounded sequence (a_n) there is a unique value L such that $\omega(\{n \in \mathbb{N}; a_n \in (L-\varepsilon, L+\varepsilon)\}) = 1$ for all $\varepsilon > 0$. So we define $\lim_{n\to\omega} a_n = L$.

We can more generally define limits $\lim_{n\to\omega} a_n$ for any sequence $(a_n)_{n\in\mathbb{N}}$ defined in a compact metric space. □

And finally, we may use these limits $\lim_{n\to\omega}$ to define a translation-invariant mean (a density) in all of $\mathcal{P}(\mathbb{N})$: for $A \subseteq \mathbb{N}$, let

$$d_\omega(A) = \lim_{N\to\omega} \frac{|A \cap \{1, \ldots, N\}|}{N} \in [0,1]. \tag{1}$$

This density, d_ω, satisfies the following properties:

1. $d_\omega(\varnothing) = 0$, $d_\omega(\mathbb{N}) = 1$.
2. $d_\omega(A \cup B) = d_\omega(A) + d_\omega(B)$ if A, B are disjoint.
3. $d_\omega(A+1) = d_\omega(A)$ for all $A \in \mathcal{P}(\mathbb{N})$.
4. $\underline{d}(A) \leq d_\omega(A) \leq \overline{d}(A)$. So $d_\omega(A) = d(A)$, if $d(A)$ is defined.

d_ω is not quite a probability measure on $\mathbb{N}$; indeed, for all n we have $d_\omega(\{n\}) \leq \overline{d}(\{n\}) = 0$, so

$$d_\omega(\mathbb{N}) = 1 \neq \sum_n d_\omega(\{n\}).$$

The concept of translation-invariant means can be studied much more generally for topological groups, see Wikipedia - amenable group.

3 Using $\mathbb{N}$ as a probability space to prove a cool theorem

We now prove the following theorem proved by Bergelson in his 1985 article *Sets of Recurrence of $\mathbb{Z}^m$-Actions and Properties of Sets of Differences in $\mathbb{Z}^m$.*

Theorem 3.1. *If a set $A \subseteq \mathbb{N}$ has $d(A) > 0$, then there exists $B \subseteq \mathbb{N}$ such that $\overline{d}(B) > 0$ and $B + B \subseteq A - A$.*

Here, we used the notation $B + B = \{b + b'; b, b' \in B\}$ and $A - A = \{a - a'; a, a' \in A\}$.

First we need a nice, but elementary, measure theory lemma:

Lemma 3.2 (Intersectivity lemma)**.** *Let $(X, \mathcal{B}, \mu)$ be a probability space, and let (A_n) be a sequence of measurable subsets of X with $\mu(X) \geq a > 0$. Then there exists a set $B \subseteq \mathbb{N}$ such that $\overline{d}(B) \geq a$ and for all $F \subseteq B$ finite, $\mu\left(\bigcap_{b\in F} A_b\right) > 0$.*

We in fact prove the more general statement:

Lemma 3.3. *Let $(X, \mathcal{B}, \mu)$ be a probability space, and let (A_n) be a sequence of measurable subsets of X. Then there exists a set $B \subseteq \mathbb{N}$ such that for all $F \subseteq B$ finite $\mu\left(\bigcap_{b\in F} A_b\right) > 0$, and*

$$\overline{d}(B) \geq \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A_n).$$

In particular, for all $b, b' \in B$ we have $\mu(A_b \cap A_{b'}) = \varnothing$.

Proof. For $F \subseteq \mathbb{N}$ finite, let $A_F = \cap_{n\in F} A_n$. We define

$$X' = X \setminus \bigcup_{\mu(A_F)=0} A_F.$$

Thus, $\mu(X') = 1$. Now let $f_N = \frac{1}{N}\sum_{n=1}^{N} \chi_{A_n}$, and $f = \limsup_N f_N$. We have, by Fatou's lemma,

$$\int_{X'} f d\mu = \int_X f d\mu \geq \limsup_N \int_X f_N d\mu = \limsup_N \frac{1}{N}\sum_{n=1}^{N} \mu(A_n) = \limsup_N \frac{1}{N}\sum_{n=1}^{N} a = a.$$

Thus, there is some point $x_0 \in X'$ such that $f(x_0) \geq a$. Letting $B = \{n \in \mathbb{N}; x_0 \in A_n\}$, we have

$$\overline{d}(B) = \limsup_N \frac{\{n \leq N; x_0 \in A_n\} \cap \{1, \ldots, N\}}{N} = \limsup_N f_N(x_0) = f(x_0) \geq a.$$

Moreover, if $F \subseteq B$ is finite, then $x_0 \in A_F$, so as $x_0 \in X'$, $\mu(A_F) > 0$. □

Note that the intersectivity lemma applies to countably additive probability measure spaces, and our proof does not work for finitely additive probability measure spaces. But in any case, we can 'transfer it' to finitely additive probability measure spaces.

This is because we can, in some sense, 'imbed' any finitely additive probability measure space into a countably additive one.

Proposition 3.4. *Let $(X, \mathcal{B}, \mu)$ be a finitely additive probability measure space (so $\mathcal{B}$ is an algebra but not necessarily a σ-algebra). There exists a countably additive probability measure space $(\overline{X}, \overline{\mathcal{B}}, \overline{\mu})$ and an injective function $f : \mathcal{B} \to \overline{\mathcal{B}}; A \mapsto \overline{A}$ (in particular $f(X) = \overline{X}$, we are abusing notation here) such that*

$$\begin{aligned} \overline{\mu}(\overline{A}) &= \mu(A) \text{ for all } A \in \mathcal{B}. \\ \overline{\varnothing} &= \varnothing \\ \overline{A \cup B} &= \overline{A} \cup \overline{B} \\ \overline{A \cap B} &= \overline{A} \cap \overline{B}. \end{aligned}$$

I'll black box this, one can prove it using Loeb measures, if anyone wants a proof I can send my notes about Loeb measures/limits of a sequence of finitely additive probability measure spaces.

Of course, one need not have $\overline{\mu}\left(\bigcup_{n\in\mathbb{N}} \overline{A_n}\right) = \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right)$ even if $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{B}$, as the measure μ need not be countably additive.

Thus, Theorem 3.3 holds for any finitely additive probability measure space $(X, \mathcal{B}, \mu)$, as it has to hold for the associated space $(\overline{X}, \overline{\mathcal{B}}, \overline{\mu})$.

The last ingredient we need is a measure theory fact:

Proposition 3.5. *Let $(X, \mathcal{B}, \mu)$ be a finitely additive probability space and $T : X \to X$ a measure-preserving transformation. For any set $A \in \mathcal{B}$, we have*

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-n}A) \geq \mu(A)^2$$

In fact the limsup is a limit and Theorem 3.5 holds for finitely additive probability spaces too, this is a basic ergodic theory fact. I will likely blackbox this in the talk due to time constraints, but let's give an elementary proof anyways, without von Neumann's ergodic theorem.

Proof. Proof by contradiction, suppose there is $\varepsilon > 0, N_0 \in \mathbb{N}$ such that, for all $N > N_0$, $\frac{1}{N}\sum_{n=1}^{N} \mu(A \cap T^{-n}A) < \mu(A)^2$, or, $\sum_{n=1}^{N} \mu(A \cap T^{-n}A) < N\mu(A)^2$.

But then, by the Cauchy-Schwartz' inequality and letting $K \in \mathbb{N}$ be a big natural number and $f_K = \frac{1}{K}\sum_{k=1}^{K} \chi_{T^{-k}A}$,

$$\begin{aligned} K^2\mu(A)^2 = \left(\int_X f_K d\mu\right)^2 \le \int_X f_K^2 d\mu &= \sum_{i,j=1}^{K} \mu(T^{-i}A \cap T^{-j}A) \\ &= K\mu(A) + 2\sum_{1\le i<j\le K} \mu(A \cap T^{i-j}A) \\ &= K\mu(A) + 2\sum_{k=1}^{K-1}\sum_{n=1}^{k} \mu(A \cap T^{-k}A) \\ &\le K\mu(A) + N_0^2 + 2\sum_{k=N+1}^{K-1}\sum_{n=1}^{k} \mu(A \cap T^{-k}A) \\ &\le K\mu(A) + N_0^2 + 2\sum_{k=N+1}^{K-1} k(\mu(A)^2 - \varepsilon) \\ &\le K\mu(A) + N_0^2 + K^2(\mu(A)^2 - \varepsilon), \end{aligned}$$

so that $K^2\varepsilon \le K\mu(A) + N_0^2$. This is a contradiction for big enough K. □

We can finally prove Theorem 3.1.

Proof. We consider the finitely additive probability measure d_ω from Equation (1) in $(\mathbb{Z}, \mathcal{P}(\mathbb{Z}))$ (say any set of negative numbers has measure 0). Note that $T : \mathbb{Z} \to \mathbb{Z}; T(n) = n + 1$ is a d_ω-preserving map, as $d_\omega(A) = d_\omega(A + 1)$ for all A. Moreover, $d_\omega(A) = d(A) > 0$. Now, for each $n \in \mathbb{N}$ define the set $A_n = (A - n) \cap (A + n) = T^n A \cap T^{-n} A$. Then, we have

$$\begin{aligned} \limsup_N \frac{1}{N}\sum_{n=1}^{N} d_\omega(A_n) &= \limsup_N \frac{1}{N}\sum_{n=1}^{N} d_\omega(A \cap T^{-2n}A) \\ &= \limsup_N \frac{1}{N}\sum_{n=1}^{N} d_\omega(A \cap (T^2)^{-n}A) \ge d_\omega(A)^2. \end{aligned}$$

So by Theorem 3.3, there is $B \subseteq \mathbb{N}$ such that $\overline{d}(B) \ge d_\omega(A)^2$ and $d_\omega(A_b \cap A_{b'}) > 0$ for all $b, b' \in B$. But, for all $b, b' \in B$,

$$(A - b) \cap (A + b') \supseteq (A - b) \cap (A + b) \cap (A - b') \cap (A + b') = A_b \cap A'_b.$$

So $d_\omega((A - b) \cap (A + b')) > 0$. In particular, $(A - b) \cap (A + b') \neq \varnothing$, so $b + b' \in A - A$ (as, letting $t \in (A - b) \cap (A + b')$, we have $t + b, t - b' \in A$, so $b + b' = (t + b) - (t - b') \in A - A$).

We conclude that $B + B \subseteq A - A$. □