Two processes on compact spaces

David Aldous

5 October 2021

Most of my research nowadays is only a few steps away from data, with some plausible story, but this project is just Old School "intellectual curiosity".

Write (S, d) for a compact metric space and $\mathcal{P}(S)$ for the space of probability measures on S, with the weak topology. I want to **investigate** processes that can be defined on any S and be parametrized by any $\theta \in \mathcal{P}(S)$. One standard example is the i.i.d. empirical process (usually studied in greater generality). Are there other interesting examples?

One reason for studying processes defined on every S is that one can seek both general results and also sharper results for any given S – providing much scope for collaboration with students.

This talk discusses two unrelated such processes. The first has been studied (a little) in $\mathbb{R}^d$, the second is apparently novel. Can you think of any others?

1. A random coverage problem

Details in the arXiv preprint

Covering a compact space by fixed-radius or growing random balls .

We have general results (not deep); much scope for more precise analysis on particular S.

Consider a compact metric space (S, d), a probability measure μ on S, but now introduce two rates $0 < \lambda < \infty$ and $0 < v < \infty$. Write $0 < \tau_1 < \tau_2 < \ldots$ for the times of a rate-λ Poisson process, and write $\sigma_1, \sigma_2, \ldots$ for i.i.d. random points of S from distribution μ. The verbal description

> *seeds arrive at times of a Poisson process at i.i.d. random positions, and then create balls whose radius grows at rate v*

is formalized as the set-valued *growth process*

$$\mathcal{X}(t) := \cup_{i:\tau_i \le t} \mathrm{ball}\,(\sigma_i, v(t - \tau_i)). \tag{1}$$

We study the cover time

$$C := \min\{t : \ \mathcal{X}(t) = S\}$$

which is finite because $\mathbb{E}\tau_1 = 1/\lambda$ and so (for any μ)

$$1/\lambda \le \mathbb{E}C \le 1/\lambda + \Delta/v \tag{2}$$

where Δ is the diameter of S.

One observation: there is a *weak concentration* bound for the distribution of cover time C.
To obtain such a bound, it is natural to require that $\mathbb{E}C$ is large relative to the maximum expected time to cover any given single point, that is relative to

$$c^* := \max_{s \in S} \mathbb{E}C(s); \quad C(s) := \min\{t : \ s \in \mathcal{X}(t)\}.$$

It turns out this is the only requirement.

Proposition

In the growth model (1), $\mathrm{var}\left(\frac{C}{\mathbb{E}C}\right) \leq \frac{c^*}{\mathbb{E}C}$.

This is an (easy) consequence of (easy) general bounds for increasing set-valued processes, and these bounds are also useful for percolation-type processes (but that's a different talk).

We can "standardize" the model by choosing time and distance units to make $\lambda = v = 1$. This is "without loss of generality" as regards explicit inequalities, though does affects asymptotics for a sequence S_n. For the standardized model we can define

$$\chi(S) = \min_{\mu} \mathbb{E}_{\mu} C$$

which is just a number associated with S. One project (not done) would be to systematically compare with other numbers associated with compact spaces S. A related project (not done) is that there is no canonical notion of *uniform* distribution on S; to what extent can the minimizing μ play a role as proxy for uniform?

So what **is** done in the preprint? Because S is compact we have cover numbers

$$\mathrm{cov}(r) := \text{ minimum number of radius } r \text{ balls that cover } S$$

which are finite. It's natural to try to relate the one number $\chi(S)$ to the function $r \to \mathrm{cov}(r)$.

There are general upper and lower bounds for $\chi(S)$ in the standardized model.

First, by considering the uniform distribution on the set of $\mathrm{cov}(r)$ points, we find (cf. coupon-collector)

$$\chi(S) \leq \min_{r>0}[r + \mathrm{cov}(r) \cdot (1 + \log \mathrm{cov}(r))].$$

Second, some μ attains $\chi(S)$, so consider the seeds of that process as a set to upper bound $\mathrm{cov}(r)$. We find

$$\chi(S) \geq \sup\{r : \mathrm{cov}(3r) > 9r\}.$$

How good are these general bounds? There is a notion I'll call *rough dimension*: A space like $[0, L]^d$ has rough dimension d characterized by

$$\mathrm{cov}(r) \asymp (L/r)^d \text{ for } r << L.$$

Here the general lower and upper bounds, for such a space, are of orders $L^{\frac{d}{d+1}}$ and $L^{\frac{d}{d+1}} \log L$.

For the actual torus $[0, L]^d$ we know sharp asymptotics as $L \to \infty$ as part of extensive historical "applied probability" work on coverage processes in Euclidean space.

Future project: study infinite-dimensional examples in more detail.

2. A Markov chain and a mapping $\mathcal{P}(S) \to \mathcal{P}(S)$

Take a pair (j, k) with $k \geq 2$ and $1 \leq j \leq k$. For any probability distribution $\theta \in \mathcal{P}(S)$, define a Markov chain on S by:

- from state s, take k i.i.d. (θ) samples, and jump to the j'th closest.

By considering the natural coupling, it is not hard to prove (a good homework problem in a course discussing coupling?) that

Theorem

Every *such chain converges in distribution (and variation distance) to some unique stationary distribution.*

Comment: Model apparently not studied. We mentally envisage S and θ as continuous, but a metric space might have only finitely many points. For this and other reasons, we explicitly specify "break ties uniform randomly". If ties are possible, the chain may not be Feller.

Call the stationary distribution $\pi_{j,k}(\theta)$. This defines a mapping $\pi_{j,k} : \mathcal{P}(S) \to \mathcal{P}(S)$. What happens when we iterate this mapping? In particular, what are the fixed points of this mapping? Fixed points would have a kind of "self-similarity under sampling" property and might provide interesting examples of specific non-uniform distributions on compact spaces S.

[Ongoing work with undergraduates Madelyn Cruz and Shi Feng: seeking more collaborators – will share extensive working notes]

The coupling proof tells us nothing explicit about the relation between θ and $\pi_{j,k}(\theta)$. By considering one step of the stationary chain we have, for $\pi = \pi_{j,k}(\theta)$

$$\theta^k(A) \le \pi(A) \le k\theta(A), \ A \subseteq S$$

and so π and θ are mutually absolutely continuous.

We study the **iterative process** which iterates the map $\pi_{j,k} : \mathcal{P}(S) \to \mathcal{P}(S)$. This does not have a simple "process" interpretation. And this project is **maybe crazy** because we don't know explicitly what the map $\pi_{j,k}$ actually is. However, for any given S and (j, k) there is an explicit equation determining fixed points θ so (in principle) one can try to solve to find all the fixed points.

The bottom line is:

- Simulations and conjectures reveal very counter-intuitive behavior.
- We have only some fragments of rigorous proofs.
- Proving anything substantial seems beyond the authors' capabilities ......

so my talk is also rather fragmentary.

First minor observations

Consider $\phi \in \mathcal{P}(S)$ which is invariant (that is, a fixed point) under $\pi_{j,k}$ for given (j, k). If the support of ϕ is smaller than S then it is more natural to consider ϕ as an invariant measure on the support. So our basic question can better be phrased as

- Given S and (j, k), what are all the invariant measures **with full support** on S?

On every compact metric space S we have an obvious "preservation of symmetry" result for the action of $\pi_{j,k}$.

Lemma

If $\theta \in \mathcal{P}(S)$ is invariant under an isometry ι of S then $\pi_{j,k}(\theta)$ is also invariant under ι.

Fragment 1: Fixed points existing by symmetry

In some cases there are distributions $\phi \in \mathcal{P}(S)$ which are invariant (that is, fixed points) "by symmetry" for all $\pi_{j,k}$. In particular

(i) The distribution δ_s degenerate at one point s;

(ii) The uniform two-point distribution $\delta_{s_1,s_2} = \frac{1}{2}(\delta_{s_1} + \delta_{s_2})$;

(iii) The Haar probability measure on a compact group S with a metric invariant under the group action.

(iv) On a finite space S, a sufficient condition for the uniform distribution to be invariant is that S is *transitive*, that is if for each pair s, s' there is an isometry taking s to s'. This is equivalent to the finite case of Haar measure. But for finite S a weaker condition suffices, because all that matters is the *rank matrix* – see later.

In those cases the distribution is invariant for all $\pi_{j,k}$. So the question becomes:

for a particular S and (j, k), are there invariant distributions with full support, other than those "forced by symmetry" as above?

Fragment 2: The case $S = \{a, b\}$ is not trivial

One might suppose that the case of a 2-element set $S = \{a, b\}$ would be trivial, but it is not. Parametrizing a distribution θ on S by $p := \theta(a)$, we view the mapping $\pi_{j,k} : \mathcal{P}(S) \to \mathcal{P}(S)$ as a mapping $\pi_{j,k} : [0, 1] \to [0, 1]$ defined as follows. In the associated 2-state Markov chain, the transition probabilities are

$$\text{prob}(a \to b) = \mathbb{P}(\mathrm{Bin}(k, p) < j); \quad \text{prob}(b \to a) = \mathbb{P}(\mathrm{Bin}(k, p) > k - j)$$

for Binomial random variables. From the stationary distribution we find

$$\pi_{j,k}(p) = \frac{\mathbb{P}(\mathrm{Bin}(k, p) > k - j)}{\mathbb{P}(\mathrm{Bin}(k, p) > k - j) + \mathbb{P}(\mathrm{Bin}(k, p) < j)}.$$

So a fixed point is a solution of the equation

$$\pi_{j,k}(p) = p. \qquad (3)$$

We know by symmetry that $p = 0, p = 1/2, p = 1$ are fixed points; are there others? By symmetry it is enough to consider $0 < p < 1/2$.

We have not tried to find solutions analytically, but we will show results of numerical calculations of the iterates $\pi^n_{j,k}(p), n = 1, 2, 3, \ldots$. For a given (k, j), we observe three possible types of qualitative behavior:

1. $\pi^n_{j,k}(p) \to 0$ as $n \to \infty$, for all $0 < p < 1/2$.
2. $\pi^n_{j,k}(p) \to 1/2$ as $n \to \infty$, for all $0 < p < 1/2$.
3. There exists a critical value $p_{crit} \in (0, 1/2)$ such that p_{crit} is invariant : $\pi_{j,k}(p_{crit}) = p_{crit}$ and $\pi^n_{j,k}(p) \to 0$ as $n \to \infty$, for all $0 < p < p_{crit}$ and $\pi^n_{j,k}(p) \to 1/2$ as $n \to \infty$, for all $p_{crit} < p < 1/2$.

For us, (3) is the interesting case: there is a non-obvious fixed point, but it is unstable. It first arises with $k = 5, j = 4$, as shown in the Figure. We see the critical value $p_{crit} = 0.17267\ldots$.

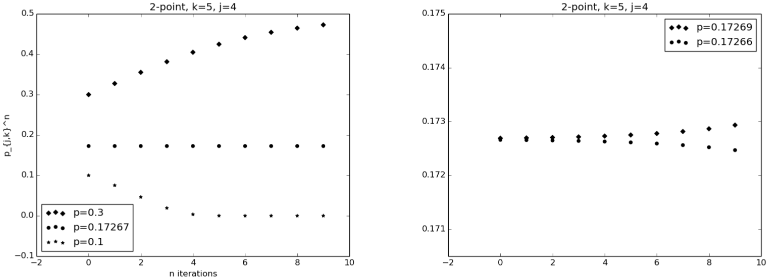


Figure: $S = \{a, b\}$; $k = 5, j = 4$. Iterates $n = 0, 1, 2, \ldots, 10$. Left panel shows type (3) behavior, Right panel shows the unstable fixed point at 0.17267.

Maybe excessive to claim 0.17267... is *interesting* but encouraging that there exist non-obvious fixed points (for certain (j, k)).

Table: $S = \{a, b\}$ and $2 \leq k \leq 9$. The values of j with each type of behavior, and (critical values) of critical points.

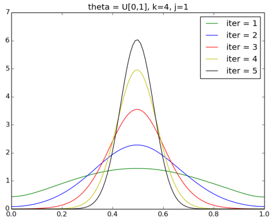
k	$(0 \leftarrow)$	$(critical)$	$(\rightarrow 1/2)$
2	1		2
3	$[1, 2]$		3
4	$[1-3]$		4
5	$[1-3]$	4 (0.17267)	5
6	$[1-4]$	5 (0.09558)	6
7	$[1-5]$	6 (0.06276)	7
8	$[1-5]$	6 (0.26405)	$[7, 8]$
9	$[1-6]$	7 (0.18884); 8 (0.03364)	9

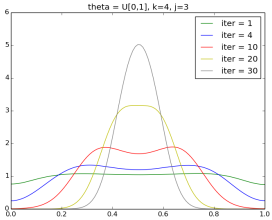
The Table shows the type of behavior – types (i) or (ii) or (iii) above – for all pairs (j, k) with $k \leq 9$. One take-away message is that for $S = \{a, b\}$ there exist some (j, k) for which $\pi_{j,k}$ has fixed points in addition to those existing by symmetry, but these fixed points are unstable.

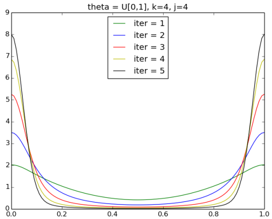
Of course the 2-point space may be very special. What properties extend to other S? Let's look at the unit interval.

Fragment 3: The case $S = [0, 1]$

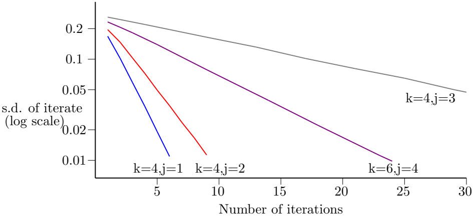
We have studied, by simulation, iterates starting from the uniform distribution $U[0, 1]$. Because $U[0, 1]$ is symmetric about $1/2$, all iterates must be symmetric about $1/2$.







The Figure above shows the case $k = 4$ and the first few iterates of $U[0, 1]$ for $j = 1, 2, 3$. Note that, here and throughout, the vertical scale and the numbers of iterations shown may not be the same from one panel to the next. What we see strongly suggests that the iterates are converging, quickly for $j = 1$ but rather slowly for $j = 3$, toward the degenerate distribution $\delta_{1/2}$. This is strongly supported by examining the standard deviations of the iterates, shown on log scale in the Figure below, and suggesting a scaling limit distribution.



In contrast, the Figure for $j = 4$ strongly suggests that the iterates are converging quickly toward the mixture $\delta_{0,1}$. These two "extreme" behaviors – convergence to $\delta_{1/2}$ for smaller j or to $\delta_{0,1}$ for larger j – appear to hold for all k. The Table shows which behavior appears to hold in simulations for each pair (j, k) with $k \leq 9$.

Table: Conjectured limits of iterates from $U[0, 1]$; the values of j with each type of behavior.

k	$\to \delta_{1/2}$	$\to \delta_{0,1}$
2	1	2
3	$[1, 2]$	3
4	$[1 - 3]$	4
5	$[1 - 4]$	5
6	$[1 - 4]$	$[5, 6]$
7	$[1 - 5]$	$[6, 7]$
8	$[1 - 6]$	$[7, 8]$
9	$[1 - 7(?)]$	$[8, 9]$

Though some cases are unclear from simulations. Need better simulations!

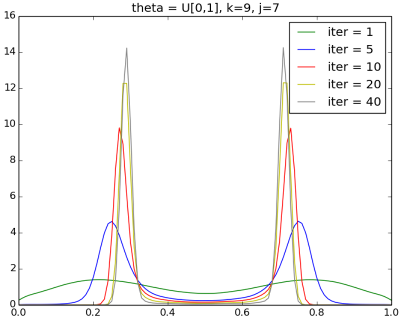


Figure: Iterates from $U[0,1]$; $k = 9, j = 7$.

As mentioned earlier, one can always write down an equation for a fixed point. On [0, 1] a density function $f(t)$ is a fixed point for $\pi_{j,k}$ iff

Proof. According to (26), invariant distribution must satisfies the following equation for all t in [0, 1]

$$1 = \binom{k}{j-1, 1, k-j} \cdot \Bigg\{ \Bigg[\int_0^{\frac{t}{2}} f_n(y) \left(\int_0^t f_n(x)\, dx \right)^{j-1} \cdot \left(\int_t^1 f_n(x)\, dx \right)^{k-j} dy \Bigg]$$
$$+ \Bigg[\int_{\frac{t}{2}}^t f_n(y) \left(\int_{2y-t}^t f_n(x)\, dx \right)^{j-1} \cdot \left(\int_0^{2y-t} f_n(x)\, dx + \int_t^1 f_n(x)\, dx \right)^{k-j} dy \Bigg]$$
$$+ \Bigg[\int_t^{\frac{1+t}{2}} f_n(y) \left(\int_t^{2y-t} f_n(x)\, dx \right)^{j-1} \cdot \left(\int_0^t f_n(x)\, dx + \int_{2y-t}^1 f_n(x)\, dx \right)^{k-j} dy \Bigg]$$
$$+ \Bigg[\int_{\frac{1+t}{2}}^1 f_n(y) \left(\int_t^1 f_n(x)\, dx \right)^{j-1} \cdot \left(\int_0^t f_n(x)\, dx \right)^{k-j} dy \Bigg] \Bigg\}$$

Shi Feng (undergrad) studied this by careful and elaborate calculus, initially in the case $j = 2, k = 2$. From the "$t = 0$" identity one can argue to a contradiction, and this can be made into a rigorous proof of

Theorem

There are no $\pi_{2,2}$-invariant distributions on $[0, 1]$ *other than those of the form δ_s or δ_{s_1,s_2}.*

The argument extends to some, but not all, pairs (j, k).

Simulations of the iterative process on $[0, 1]$ **starting with a non-uniform distribution** show analogous behavior: either convergence to $\delta_{0,1}$ or to δ_s for some s depending on the initial distribution.

At a rigorous level, the key open questions for $S = [0, 1]$ are

- Does there exist (for any (j, k)) any invariant distribution with full support?
- Does there exist (for any (j, k)) any distribution other than δ_s or $\delta_{0,1}$ that occurs as a limit of iterates from some initial distribution with full support?

We suspect the answer to each is "no". Of course, "no" to the second question would imply "no" to the first question.

Fragment 4: An unexpected equivalence

The mappings $\theta \to \pi_{1,2}(\theta)$ and $\theta \to \pi_{2,2}(\theta)$ behave quite differently, but Shi Feng discovered the following, which is easy to verify but which I would never have imagined.

Proposition

For every compact metric space S, the set of invariant distributions for $\pi_{1,2}$ is the same as the set of invariant distributions for $\pi_{2,2}$.

Maybe a useful starting point for some "proof by contradiction" that certain fixed points for $\pi_{1,2}$ do not exist?

Fragment 5: Finite S.

A finite metric space can be represented by the matrix D of distances $d(i,j)$. By taking all the non-zero distances to be between 1 and 2, the triangle inequality is automatically satisfied. Consider the example of a 5-element space with distance matrix

$$D = \begin{pmatrix} 0 & 1.714 & 1.341 & 1.656 & 1.74 \\ 1.714 & 0 & 1.298 & 1.794 & 1.03 \\ 1.341 & 1.298 & 0 & 1.715 & 1.844 \\ 1.656 & 1.794 & 1.715 & 0 & 1.524 \\ 1.74 & 1.03 & 1.844 & 1.524 & 0 \end{pmatrix}$$

What matters for our purposes, assuming as in this example that all non-zero distances are distinct, is the *rank matrix* R, where $r(i,j) = 4$ means that $d(i,j)$ is the 4'th smallest of $\{d(i,1), d(i,2), \ldots, d(i,|S|)\}$. For the distance matrix D above, the rank matrix is

$$R = \begin{pmatrix} 1 & 4 & 2 & 3 & 5 \\ 4 & 1 & 3 & 5 & 2 \\ 3 & 2 & 1 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \\ 4 & 2 & 5 & 3 & 1 \end{pmatrix}$$

By numerical calculation, for $\pi_{1,2}$ on this space there is an invariant distribution

$$\theta \approx (0.149\ 0.188\ 0.203\ 0.298\ 0.162)$$

for which the transition matrix is

$$K \approx \begin{pmatrix} 0.276 & 0.097 & 0.304 & 0.297 & 0.026 \\ 0.111 & 0.341 & 0.222 & 0.089 & 0.237 \\ 0.159 & 0.265 & 0.365 & 0.185 & 0.026 \\ 0.139 & 0.036 & 0.118 & 0.507 & 0.201 \\ 0.083 & 0.28 & 0.041 & 0.298 & 0.298 \end{pmatrix}$$

This example was found (by Shi Feng) by simulating random distance matrices D, obtaining the rank matrix R, and then numerically solving for invariant distributions θ until finding a solution with full support. Note this involved non-linear equations: we need to solve $\theta K = \theta$ but here K depends on θ, for instance for $\pi_{1,2}$

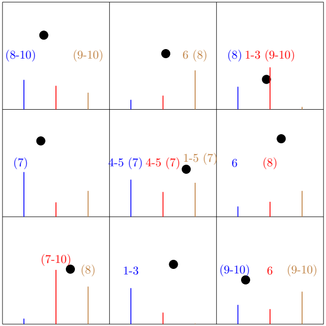
$$k(i,i) = 1 - (1 - \theta(i))^2$$

$$\text{if } r(i,j) = 5 \text{ then } k(i,j) = \theta^2(j).$$

Note also that for $|S| = 5$ there are only a finite number of possible rank matrices R, so this counter-example is not like a counter-example depending on a real parameter taking a specific value.

Fragment 6: Finite S embedded in $\mathbb{R}^2$.

Example; S consists of 9 points in the plane.
$k = 10$; each $1 \leq j \leq 10$.
Three different initial distributions (colors of bars); limits of iterative process are δ_s for marked s or δ_{s_1,s_2} for marked (s_1) and (s_2).



Simulations by Madelyn Cruz (undergrad).
What we observe (in each of 11 trials from different initial distributions) is that with $k = 10$

> *(*) for $j \leq 6$ the limit is always some δ_s whereas for $j \geq 7$ the limit is always some δ_{s_1,s_2}. But the precise limit – which s or s_1, s_2 – depends both on j and the initial distribution.*

Remarkably, simulations for an analogous perturbed 5×5 pattern S show exactly the same behavior described in (*), and so do simulations for an analogous perturbed $3 \times 3 \times 3$ pattern in three dimensions.

Suggests some form of "universality"?

State of this project

Disappointing: we have not found interesting distributions on particular S.

Instead we have a range of open problems about ways in which the behavior is non-interesting.

- For which S and (j, k) are there invariant distributions other than those "forced by symmetry"?
- True or false: For every S and every (j, k), every invariant distribution except δ_s and δ_{s_1, s_2} is unstable (to a generic perturbation).
- True or false: For every S, the iterative process for $\pi_{1,2}$ from almost all initial $\phi \in \mathcal{P}(S)$ converges to some δ_s (depending on ϕ).

If not true in general, is it true for $S \subset \mathbb{R}^d$?