The Enduring Appeal of the Probabilist's Urn

CÉCILE MAILLER

Pólya urns are classic objects in probability theory and have been extensively studied for more than 100 years due to their wide range of applications. We present classic results on these processes, their applications to large networks, and recent theoretical developments.

"Is it a contractual obligation for all probabilists to have a drawing of an urn of coloured balls on their blackboards?". This was the question of a non-mathematician friend of mine when they entered my office a few years ago. Drawing balls from urns appears in many classic exercises in a first course in probability (e.g. an urn contains two black balls and one red ball, I pick a ball uniformly at random from the urn, what is the probability that it is black?). The aim of this article is to explain why balls-in-urns is still an object of interest for probabilists (and hence why they may end up on their blackboards!) and what questions on this object remain open.

A Pólya urn is a stochastic process that depends on two parameters: the initial composition of the urn (how many balls of which colours there are in the urn at time zero), and a replacement rule, which is encoded by a $d \times d$ matrix $R = (R_{i,j})_{1 \le i,j \le d}$, where d is the number of different colours a ball can be. At every integer time step, a ball is picked uniformly at random from the urn and, if it is of colour i, it is put back in the urn together with $R_{i,j}$ additional balls of colour j, for all $1 \le j \le d$.

This article focuses on the following question: "What is the composition of the urn in the limit when time goes to infinity?". The difficulty in answering this question comes from the fact that the simple drawing-and-replacing procedure is repeated infinitely many times (at every integer time step). A classic example illustrating how complicated phenomena can arise from repeating a simple procedure infinitely many times is the Brownian motion, which can be defined through an infinite sequence of coin flips.

There exist two canonical cases from which the behaviour of most Pólya urns can be inferred: (a) the case when the replacement matrix is the identity, which was first studied by Markov in 1906, and (b) the case when it is "irreducible", for which landmark results were proved by Athreya and Karlin in 1968. These two cases lead to very different outcomes. This can be seen in the simulations of Figure 1: although in both cases, the composition of the urn seems to converge to a limiting value, this value seems to be random in the case of an identity replacement matrix, and deterministic in the irreducible case.

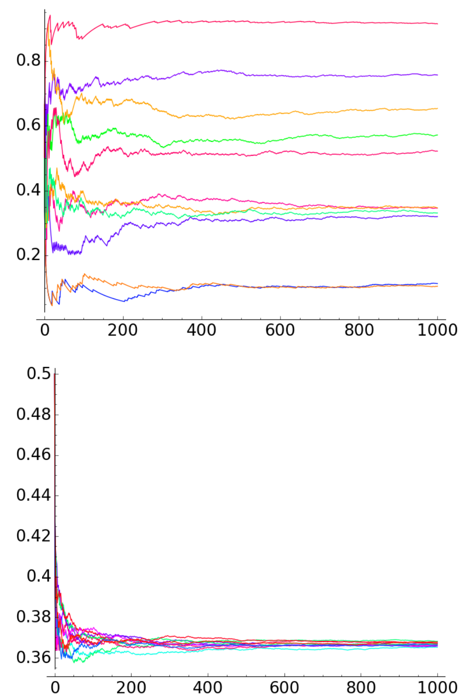


Figure 1. Ten realisations of a 2-colour Pólya urn whose initial composition is one ball of each colour, with identity replacement matrix (top) and with replacement matrix $\left(\begin{smallmatrix}2 & 1\\ 3 & 2\end{smallmatrix}\right)$ (bottom); the vertical axis is the proportion of balls of the first colour in the urn (equal to $1/2$ at time 0), the horizontal axis is time

The reason why Pólya urns have been of interest over the last century is because of their numerous applications. In fact, I often joke that any probability problem can be seen as an urn problem, which is, in essence, what Pólya wrote in 1957:

> “Any problem of probability appears comparable to a suitable problem about bags containing balls and any random mass phenomenon appears as similar in certain essential respects to successive drawings of balls from a system of suitably combined bags.”

One of the reasons for the wide applicability Pólya writes about is that Pólya urns are the simplest model for *reinforcement*, i.e. for the rich-get-richer phenomenon. Indeed, imagine a Pólya urn with black and red balls, with replacement matrix the identity (at every time step, we add one ball of the same colour as the drawn ball). If we start with many red balls and few black balls, then we are more likely to add even more red balls to the urn, thus *reinforcing* the higher proportion of red balls.

In this article, we focus on one of these many applications of Pólya urn results: namely, their application to the analysis of the degree distribution of two models of random “trees” that were originally introduced as models for large complex networks (such as the internet, social networks, etc.).

This application to random trees highlights the need for a generalisation of Pólya urns to infinitely many colours (i.e. $d = \infty$). The end of the article focuses on this generalisation, which, in the “irreducible” case, has only been achieved in the last decade, and for which important questions still remain open.

The classic Pólya urn

We start by looking at the classic Pólya urn, i.e. the urn with replacement matrix equal to the identity. We show that, in this case, the composition of the urn converges to a random variable (see the top of Figure 1) whose distribution depends on the initial composition of the urn. We start by looking at the 2-colour case with initial composition one ball of each colour, before looking at (a) different initial compositions, and (b) the d-colour generalisation.

The 2-colour case starting with one ball of each colour can be studied by explicit calculations: let $U_1(n)$ denote the number of red balls in the urn at time n. If $U_1(n) = k$ $(1 \le k \le n+1)$, then we have drawn $k-1$ times a red ball, and $n-k+1$ times a black ball from the urn. The probability that the first $k-1$ balls drawn from the urn are red and the next $n-k+1$ are black is

$$\frac{1}{2}\cdot\frac{2}{3}\cdots\frac{k-1}{2+k}\cdot\frac{1}{3+k}\cdot\frac{2}{4+k}\cdots\frac{n-k+1}{n+1} = \frac{(k-1)!(n-k+1)!}{(n+1)!}.$$

The same formula gives the probability of drawing the same number of red and black balls in any other order. Since there are $\binom{n}{k-1} = \frac{n!}{(k-1)!(n-k+1)!}$ different orders, we get

$$\mathbb{P}(U_1(n) = k) = \frac{1}{n+1}.$$

In other words, for all integer times n, $U_1(n)$ is uniformly distributed among all its possible values $\{1,\dots,n+1\}$, and thus, $U_1(n)/n$ converges in distribution to a uniform random variable on $[0,1]$ when n tends to infinity; this can be seen in the simulations of Figure 2.

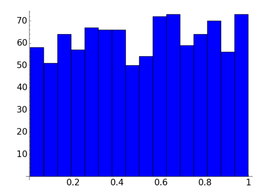


Figure 2. Histogram of the proportion of red balls at time 500 in 1000 realisations of a classic 2-colour Pólya urn

In fact, using “martingale theory”, one can prove that the proportion of red balls converges *almost surely* to this uniform limit, which is a stronger statement than convergence in distribution. Almost-sure convergence means that the probability of convergence of the proportion of red balls to a random uniformly-distributed value is equal to one, while convergence in distribution means that the distribution of the proportion of red balls converges to the uniform distribution. The almost-sure convergence of $U_1(n)/n$ is the reason why, on the top of Figure 1, each of the trajectories converges.

We now look at different initial compositions: let $U_1(0)$ and $U_2(0)$ be the number of red and black balls in the urn at time 0. A similar calculation to that above (using Stirling’s formula) implies that the proportion of red balls converges to a Beta distribution of parameter $(U_1(0), U_2(0))$. (A Beta distribution is a two-parameter distribution on $[0,1]$: Figure 3 shows how its density depends on the values of the parameters. The Beta distribution of parameter $(1,1)$ is the uniform distribution on $[0,1]$.) Consequently, the long-term behaviour of the urn depends on the initial composition of the urn.

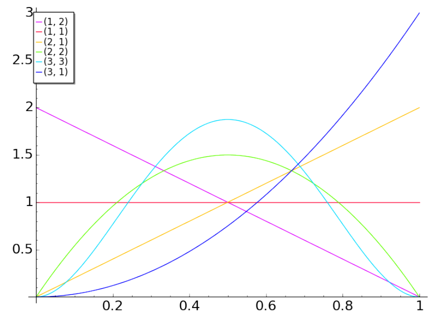


Figure 3. Different Beta distribution densities

Similar results hold when there are $d > 2$ colours in the urn. To study this case, we define $Z(n)$ as the vector whose d coordinates are the proportions of balls of each colour in the urn at time n, i.e.

$$Z(n) := \Big(\frac{U_1(n)}{n+d}, \dots, \frac{U_d(n)}{n+d}\Big), \tag{1}$$

where $U_i(n)$ is the number of balls of colour i in the urn at time n (note that $n + d$ is the total number of balls in the urn at time n since we start with d balls and add one ball at each time step). Using similar methods as in the 2-colour case, one can show that $Z(n)$ converges almost surely (i.e. with probability one) to a limiting value Z_∞, and this limiting value is uniformly distributed on the simplex, i.e. on the set

$$\Big\{(x_1, \dots, x_d) \in [0,1]^d : \sum_{i=1}^{d} x_i = 1\Big\},$$

where the vector $Z(n)$ lives.

The uniform distribution on the simplex is called a Dirichlet distribution of parameter $(1, \dots, 1)$. If we change the initial composition of the urn, then the limiting distribution is a Dirichlet distribution of parameter the initial composition vector:

Theorem 1. *Let $U_i(n)$ be the number of balls of colour i in the urn at time n, in the case when the replacement matrix equals the identity. Almost surely,*

$$Z(n) \to \mathrm{Dirichlet}(U_1(0), \dots, U_d(0)), \text{ as } n \to \infty.$$

The irreducible case

In this section, we look at the case of an irreducible replacement matrix. In contrast to the identity case, the composition of the urn in the irreducible case has a deterministic limit (see the bottom of Figure 1), which does not depend on the initial composition.

A $d \times d$ matrix R is irreducible if for all $1 \le i, j \le d$, there exists n such that $(R^n)_{i,j} \neq 0$; in particular, a matrix with all entries positive is irreducible. In our urn context, we also assume that all coefficients of the replacement matrix are non-negative (allowing -1s on the diagonal, meaning that we discard the ball that was drawn, is a straightforward generalisation).

To state the main result of this section, we need a linear algebra theorem, due to Perron and Frobenius:

Theorem 2. *If R is an irreducible matrix with non-negative coefficients, then the spectral radius Λ of R is also a simple eigenvalue of R. (That is, Λ is a real eigenvalue of R, has multiplicity one, and all other eigenvalues of R have absolute value less than Λ.)*

Moreover, there is a unit left-eigenvector v associated to eigenvalue Λ whose coefficients are all non-negative.

Sampling a classic Pólya urn via biased coin flips

A classic exercise in probability theory (see Williams's book Probability with Martingales, Exercise E10.8) is to prove the following result: Start with an urn containing one red and one black ball, sample a uniform random variable Θ on $[0,1]$, and then, at every time step, add to the urn either a red ball, with probability Θ, or a black ball, with probability $1 - \Theta$.

Lemma 1. *If $\hat{U}_1(n)$ is the number of red balls in this urn at time n, then the process $(\hat{U}_1(n))_{n\geq 0}$ has the same distribution as $(U_1(n))_{n\geq 0}$ in the classic Pólya urn case with initial composition $(1, 1)$.*

Conditionally on the random variable Θ, $\hat{U}_1(n)/n \to \Theta$ almost surely when n tends to infinity. Therefore, we can rephrase Lemma 1 as follows: one way to sample a Pólya urn with identity replacement matrix and initial composition one ball of each colour is to first sample its uniform limit Θ, and then perform coin flips with bias Θ to decide what ball to add to the urn at every time step.

To prove this result, you need basic notions of conditional expectation and martingale theory; if you have this background, I strongly encourage you to try!

As in Equation (1), we let $Z(n)$ denote the vector whose coordinates are the proportions of balls of each colour in the urn at time n, i.e.,

$$Z(n) = \left(\frac{U_1(n)}{\|U(n)\|_1}, \dots, \frac{U_d(n)}{\|U(n)\|_1}\right),$$

where $\|U(n)\|_1 = U_1(n) + \cdots + U_d(n)$ is the total number of balls in the urn at time n. The following result is due to Athreya and Karlin in 1968:

Theorem 3. *If the replacement matrix R is irreducible, then, for all non-empty initial compositions, $Z(n) \to v$ almost surely when n tends to infinity, where v is the unit left-eigenvector with non-negative coordinates associated to the spectral radius of R.*

Note that this behaviour is drastically different from the behaviour of the classic Pólya urn (see Theorem 1) in the following two ways: (a) the limit in the irreducible case is deterministic and not random as in the identity case, and (b) the limit in the irreducible case does not depend on the initial composition of the urn, while it does in the identity case. I personally find point (b) surprising: say you have an urn with red and black balls, and replacement matrix $\binom{2\ 1}{3\ 2}$, the renormalised composition vector converges to $v = (1/2, \sqrt{3}/2)$ (see the right-hand side of Figure 1: the proportion of red balls indeed converges to $1/(\sqrt{3}+1) \approx 0.37$) whether we started with one black and one red ball or with a million red balls and one black ball!

Fluctuations

From Theorems 1 and 3, we know that the renormalised composition vector converges almost surely to a limit (both in the identity and the irreducible cases). It is natural to ask about its speed of convergence to this limit. Theorems about these “fluctuations” can be proved using martingale theory.

Interestingly, in the irreducible case the outcome depends on the spectral gap of the matrix R, i.e. the ratio σ between the largest of all other real parts of the eigenvalues of R and its spectral radius. If $\sigma < 1/2$, then the fluctuations of $U(n)$ around nv are Gaussian and of order $n^{1/2}$, while if $\sigma > 1/2$, they are non-Gaussian and of order n^{σ} (see, e.g, [3]).

Application to the random recursive tree

As mentioned in the introduction, results on Pólya urns can be applied to the study of more complicated stochastic processes, and among them, several models for complex networks. In this section, we apply Theorem 3 to prove convergence of the “degree distribution” of a model called the random recursive tree (RRT).

Embedding into continuous time

To prove Theorem 3, Athreya and Karlin embed the urn process into continuous time: a technique that is now standard for the analysis of Pólya urns. In particular, Janson uses it in [3] to generalise Athreya and Karlin’s result to a much wider class of Pólya urns: he allows balls of different colours to have different weights, the replacement matrix to be random, and relaxes the irreducibility assumption.

The idea is the following. Imagine that each ball in the urn is equipped with a clock that rings after a random time of exponential distribution with parameter 1, independently from all other clocks. When a clock rings, the corresponding ball, if it is of colour i, splits into $R_{i,j} + \mathbf{1}_{i=j}$ balls of colour j (for all $1 \le j \le d$).

Now imagine that every time a clock rings, we take a picture of the urn: using the standard properties of the exponential distribution, one can check that the ordered sequence of these pictures (which is now a discrete time process) is distributed like the Pólya urn of replacement matrix R.

The advantage of the continuous time process is that each ball is now independent from the other balls; the drawback is that we added some randomness by making the split times random. The continuous time process we get by embedding a Pólya urn into continuous time is called a multi-type Galton–Watson process.

The RRT is defined recursively as follows. At time zero, the tree is two nodes linked by one edge, and at every time step, we add a node to the tree and create an edge between this new node and a node chosen uniformly at random among the nodes that are already in the tree.

This model was introduced as a model for networks by Na and Rapaport in 1970. Two typical features of real-life complex networks are the "small-world" and the "scale-free" properties (see [2] where typical properties of complex networks are discussed). It is thus natural to ask whether the RRT has these two properties.

A graph has the small-world property if the distance between two nodes chosen uniformly at random among all n nodes of the graph is of order at most $\log n$ when n tends to infinity; the RRT has the small-world property (see, e.g., Dobrow 1996).

A graph is called scale-free if its degree distribution is a power-law, i.e. if there exists $\tau > 0$ such that the proportion of nodes of degree i in the graph is of order $i^{-\tau}$ when $i \to +\infty$. The degree of a node is the number of other nodes it is linked to by edges (e.g. in a friendship network, it is the number of friends of a node). Most real-life networks are scale-free with τ usually between 2 and 3 (e.g. for the internet, it is estimated that $\tau \approx 2.5$).

In the rest of this section, we prove that the RRT is *not* scale-free; to do so, we use Theorem 3, as originally done by Mahmoud and Smythe in 1992. For all $i \geq 1$, we let $U_i(n)$ denote the number of nodes of degree i in the n-node RRT.

The idea is to view the process $U(n) = (U_1(n), U_2(n), \dots)$ as a Pólya urn: the nodes of the tree are the balls in the urn, and their colour is their degree. At time zero, the urn contains two nodes of colour 1 (because the tree contains two nodes of degree 1), and at every time step, we pick a ball in the urn, say of colour i, remove it from the urn and add instead a ball of colour 1 and a ball of colour $i+1$. In other words, the process $U(n)$ is a Pólya urn of initial composition (2, 0, 0, ...) and replacement matrix

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & \cdots & \cdots \\ 1 & -1 & 1 & 0 & 0 & \cdots & \cdots \\ 1 & 0 & -1 & 1 & 0 & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}.$$

At this point, we could be worried for two reasons: (a) the matrix has negative coefficients and (b) it is infinite, i.e. there are infinitely many colours! As explained above, the -1s on the diagonal are not a problem and Theorem 3 still holds if the matrix is irreducible. But (b) is more worrying since Theorem 3 only holds for urns with finitely-many colours.

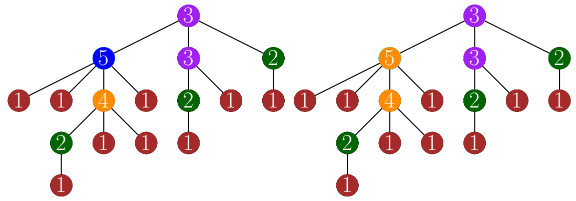


Figure 4. The degree distribution in the RRT seen as an urn, and its truncated version at $m = 4$ (in the truncated version, the node of degree 5 is considered as being of the same colour as the node of degree 4)

Luckily, there exists a trick to reduce to a finite number of colours: we decide to consider all colours above a threshold $m \geq 2$ as one colour. Remarkably, the m-colour process is also a Pólya urn process (this is not true for all infinitely-many-colour urns). For example, take $m = 4$ (see Figure 4), and look at the vector $(U_1(n), U_2(n), U_3(n), \sum_{i\geq 4} U_i(n))$: it is a Pólya urn with replacement matrix

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

This matrix is irreducible (its spectral radius is 1 and left-eigenvector $v = (1/2, 1/4, 1/8, 1/8)$); thus Theorem 3 applies and gives that, almost surely when n tends to infinity,

$$\frac{1}{n}(U_1(n), U_2(n), U_3(n)) \to \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}\right).$$

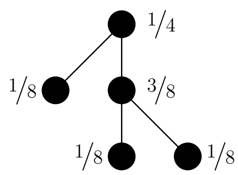
Here we chose $m = 4$ as a threshold, but one can choose an arbitrarily large threshold m, which gives the following result:

Theorem 4. *Let $U_i(n)$ be the number of nodes of degree i in the n-node RRT. For all $i \geq 1$, almost surely when n tends to infinity, $U_i(n)/n \to 2^{-i}$. Consequently, the RRT is not scale-free.*

Application to the preferential attachment tree

Although the RRT does not have the scale-free property (and is therefore not a realistic model for complex networks), the idea of Na and Rapaport to define a dynamical structure (i.e. nodes arriving one by one in the tree) led to the definition of more realistic models for networks. One of these models is the preferential attachment tree (PAT), which was originally defined by Yule in 1923, and popularised by Barabási and Albert in 1999. In this section, we show how one can prove that the PAT is scale-free (see Figure 5 where the typical shapes of the RRT and of the PAT are compared).

The PAT is defined as follows. Start with two nodes linked by an edge, and at every time step, add a node to the tree, and link it to a node chosen at random among existing nodes with probability proportional to their degrees: this is the "preferential attachment". For example, on the picture, the next node to enter the tree will connect with one of the 5 existing nodes with the probabilities displayed.



Theorem 5. *Let $U_i(n)$ denote the number of nodes of degree $i \geq 1$ in the n-node PAT. Almost surely when n tends to infinity,*

$$\frac{U_i(n)}{n} \to \frac{2}{i(i+1)(i+2)},$$

and thus the PAT is scale-free with index $\tau = 3$.

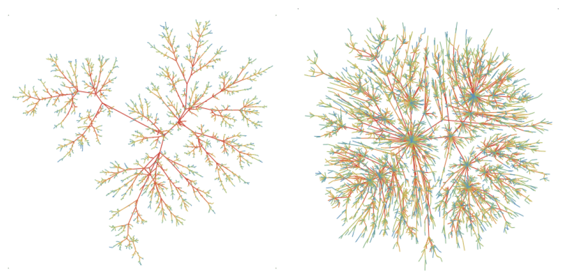


Figure 5. A realisation of the RRT (left) and the PAT (right) at large time (courtesy of Igor Kortchemski); the scale-free property of the PAT can be "seen" on this simulation: heuristically speaking, the PAT has large hubs while degrees in the RRT are much more homogeneous

This result was proved by Mahmoud, Smythe and Szymański in 1993; they use similar methods as for the RRT. The only additional subtlety is that the vector $(U_1(n), U_2(n), \dots)$ is not a Pólya urn, one has to consider the vector $(U_1(n), 2U_2(n), 3U_3(n), \dots)$ instead. As in the RRT case, one can truncate the number of colours to finitely-many, and apply Theorem 3 to conclude the proof of Theorem 5.

Infinitely-many colours

The Pólya urns arising from the application to the RRT and the PAT both have infinitely many colours. In these two examples, we considered all colours above a threshold to be the same, and used the remarkable fact that the finitely-many-colour urn obtained after this reduction was still a Pólya urn. However, this property does not hold for all infinitely-many-colour Pólya urns and consequently, as Janson wrote in [3],

> "These examples suggest the possibility of (and desire for) an extension of the results in this paper to infinite sets of types."

For the case when the replacement matrix equals the identity, the generalisation to infinitely many colours dates back to Blackwell and McQueen in 1973. Its equivalent for the irreducible case, however, only dates back to 2017 (with a preliminary particular case in a 2013 paper by Bandyopadhyay and Thacker) and is still the object of some open problems.

The main idea is to look at the composition of the urn, not as a vector as in the finitely-many colour case, but as a measure on a set $\mathscr{C}$ of colours. As in the finitely-many-colour case, we define a discrete time process that depends on two parameters: the initial composition, which is now a finite measure m_0 on $\mathscr{C}$, and the replacement kernel $(K_x)_{x\in\mathscr{C}}$, which is a set (indexed by $\mathscr{C}$) of finite measures on $\mathscr{C}$.

For all integers $n \geq 0$, we set

$$m_{n+1} = m_n + K_{\xi(n+1)},$$

where $\xi(n+1)$ is a $\mathscr{C}$-valued random variable of distribution $m_n/m_n(\mathscr{C})$ (the normalisation is so that the total mass of $m_n/m_n(\mathscr{C})$ equals 1). The random variable $\xi(n+1)$ can be interpreted as the colour of the ball drawn in the urn at time $n+1$.

The finitely-many-colour case fits in this framework: it corresponds to $\mathscr{C} = \{1, \dots, d\}$, $m_0 = \sum_{i=1}^d U_i(0)\delta_i$ and, for all $1 \leq i \leq d$, $K_i = \sum_{j=1}^d R_{i,j}\delta_j$, where δ_x is the Dirac mass at x.

However, this new model of "measure-valued Pólya processes" is much more general than the finitely-many colour case: the set of colours can be any measurable set. From our applications to random trees, it is natural to consider $\mathscr{C}$ to be $\mathbb{N}$ or $\mathbb{Z}$, but we can take $\mathscr{C}$ to be $\mathbb{R}$ or any other measurable set.

In these continuous cases, for all Borel set $B \subseteq \mathscr{C}$, we interpret $m_n(B)$ as the mass of all balls in the urn at time n whose colour belongs to B, and $K_x(B)$ as the mass of balls with colour in B that we add in the urn when drawing a ball of colour x.

Note that $m_n(B)$ and $K_x(B)$ do not have to be integer valued, and in fact m_n and K_x can be absolutely continuous measures, in which case we should think of coloured dust in an urn instead of coloured balls (see Figure 7).

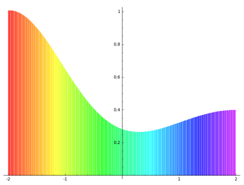


Figure 7. Composition measure of an urn with colour set $\mathbb{R}$

We have now defined a model that allows for infinitely many colours, but can we prove anything about it? Yes: Blackwell and McQueen in 1973 proved an analogue to Theorem 1 for the identity case ($K_x = \delta_x$ for all $x \in \mathscr{C}$). In the appropriate analogue of the irreducible case for infinitely many colours, Bandyopadhyay and Thacker, and Mailler and Marckert, simultaneously in 2017, proved an equivalent of Theorem 3.

More recent results on the infinitely-many-colour irreducible case induce a stronger statement than Theorem 4 for the degree distribution of the RRT (see Mailler and Villemonais 2020):

Theorem 6. *Let $U_i(n)$ be the number of nodes of degree i in the n-node RRT. For all $\varepsilon > 0$, for all functions $f : \mathbb{N} \to \mathbb{R}$ such that $f(n) = o((2-\varepsilon)^n)$ when $n \to \infty$, we have, almost surely when $n \to \infty$,*

$$\frac{1}{n}\sum_{i\geq 1} f(i)U_i(n) \to \sum_{i\geq 1} f(i)2^{-i}.$$

The difficulty of the infinitely-many-colour case comes from two main factors: (a) the embedding into continuous-time method, which was very successful in the finitely-many colour case seems not to be applicable to the infinitely-many-colour case and (b) all the proofs that rely on linear algebra in the finitely-many colour case need to be adapted to the infinite-dimensional setting using operator theory. For example, a question that remains open at the time of writing this article, and on which I am currently working, is to prove theorems for the fluctuations of these processes around their limits.

Summary

In this article, we focused on the convergence of the composition of Pólya urns in the case when the replacement matrix is the identity or when it is irreducible. We then looked at applications of these results to the analysis of random tree models for complex networks. Motivated by these applications, we looked at the definition a new framework for Pólya urns with infinitely many colours, a topic which is the object of ongoing research.

In the introduction, we stated a quote of Pólya from 1957 highlighting the wide applicability of Pólya urns and this wide applicability continues today. Indeed, in parallel universes, articles with the same title as this one could have told stories ending with applications of Pólya urns to adaptive clinical trials (e.g. Laruelle and Pagès 2013, Zhang 2016), or tissue growth modelling (Borovkov 2019), or Monte-Carlo approximation methods (Wang, Roberts and Steinsaltz 2018, Mailler and Villemonais 2020), or reinforcement learning (ants finding shortest paths in networks: Kious, Mailler and Schapira 2020), all of which are still active areas of research at the time of writing.

FURTHER READING

[1] K.B. Athreya, P.E. Ney. Branching Processes (Chapter V). Springer Berlin, 1972.
[2] R. van der Hofstad. Random Graphs and Complex Networks. Cambridge Series in Statistical and Probabilistic Mathematics, 2017.
[3] S. Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. Stoch. Proc. App. 110(2) (2004) 177–245.
[4] N.L. Johnson, S. Kotz. Urn models and their application; an approach to modern discrete probability theory. Wiley, 1977.

Cécile Mailler

Cécile is a lecturer in probability and an EPSRC postdoctoral fellow at the University of Bath. Her main research interests are on branching processes and reinforcement. She particularly enjoys working on problems that arise from applications, and even more so if the model is named after an animal (she has so far managed to work on the "monkey walk" and the "ant walk"!).