Chapter 2

SURVEY OF PROCESSES WITH REINFORCEMENT

This chapter discusses Pólya's urn process and some of its generalizations. Since the main results of this thesis concern processes that generalize Pólya's urn, this chapter serves as historical background. Pólya's urn has been widely studied; this chapter singles out the generalizations and applications most relevant to this thesis. For want of an already established term, I call these *random processes with reinforcement.*

The term *random process with reinforcement* is intended to delimit a class of discrete-time processes, of which the Pólya urn process, described below, is prototypical. Which processes fall into this class? The question can best be answered by another question: why is the Pólya urn more interesting than the Ehrenfest urn or an urn scheme based on sampling without replacement, or any of a number of other urn schemes to be found in a general survey such as [JK]? One answer is that it is not easily reduced to a Markov chain which can then be completely understood. Perhaps I should say "fruitfully" instead of "easily", since any discrete-time process can be made into a Markov chain by expanding the state space to include the entire history of the process. In general, random processes with reinforcement can be described by their transition probabilities.

These are not constant as in the Markov case, but depend on the history of the process via a single (vector-valued) function. Typically, this function is just a normalized occupation vector, i.e. a vector indexed by the states of the process that says what fraction of the time the process has been in each state.

2.1 Pólya's urn

In 1923, Eggenberger and Pólya [EP] proposed the following urn scheme to model processes such as the spread of infectious disease. An urn contains R red balls and B black balls. A ball is drawn from the urn and replaced in the urn along with Δ balls of the same color. Here, Δ is any fixed constant. This process of drawing and replacing is repeated ad infinitum. At each draw, the chance of picking a red ball is equal to the fraction of balls in the urn that are red; we call any urn scheme satisfying this condition a linear scheme. Initially this chance is $R/(R+B)$, but it changes after each draw. The following is the basic theorem about this process; it is part of the mathematical folklore, i.e. I was too lazy to track down the first time it appeared in this form.

Theorem 2.1 (Pólya's urn) *The fraction of red balls in the urn under such a scheme converges almost surely to some random limit. The distribution of this limit is a beta distribution with parameters R/Δ and B/Δ.*

The randomness of this limit is surprising to most people (see [Co] for a discussion of this).

Proof: The successive fractions of red balls form a bounded martingale and hence converge almost surely. To verify the distribution of the limit is routine once you know that it is a beta with parameters R/Δ and B/Δ. See for example [Fe] v. 2 ch. VII sec. 4. $\square$

Remark: R, B and Δ may be taken to be arbitrary positive real numbers, although the process is usually stated for integral R, B and Δ because of the physical description in terms of balls in an urn. Similarly, the number of colors can be taken to be any $d \geq 2$ and the limiting distribution can be calculated as follows: consider all the colors but one to form a single supercolor and apply theorem 2.1 to this two-color problem;

Now since the distribution of colors within the supercolor is independent of the total fraction that have the supercolor, this procedure can then be inductively applied to the $d-1$ color problem that remains. Let the original number of balls of color i be w_i. It is easy to see that the joint distribution of the limiting fractions, $x_1, \ldots, x_d$, of balls of each color must have density

$$\frac{\Gamma((w_1 + \cdots + w_d)/\Delta)}{\Gamma(w_1/\Delta)\cdots\Gamma(w_d/\Delta)} x_1^{w_1/\Delta - 1} \cdots x_d^{w_d/\Delta - 1}$$

which is know as the Dirichlet distribution with parameters $w_1/\Delta, \ldots, w_d/\Delta$. This follows from the characteristic property of the Dirichlet distribution, namely that the marginal distribution of the sum $\sum_{i \in A} x_i$ is a beta with parameters $\sum_{i\in A} w_i/\Delta$ and $\sum_{i \notin A} w_i/\Delta$.

Another way of looking at Pólya's urn is through exchangeability theory. An easy computation shows that the sequence of draws is an infinite exchangeable sequence of random variables. In other words, the probability of drawing a particular finite sequence of colors at the beginning depends only on the numbers of each color drawn and not on the specified order. An application of De Finetti's theorem [Fe v. 2 ch. VII sec. 4] now yields

Theorem 2.2 *Pick $p \in (0,1)$ randomly from a beta distribution with parameters R/Δ and B/Δ. Then generate a sequence of colors where each color is independently red with probability p and black with probability $1-p$. The sequences generated by this two-step process are identically distributed to the sequences of colors drawn from a Pólya urn with parameters R, B and Δ.* □

This theorem gives the Pólya urn process an interesting Bayesian interpretation. Suppose $p \in (0,1)$ is an unknown parameter corresponding to the probability of any ball being red, and put a prior on p that is beta with parameters R/Δ and B/Δ. This is a common prior, dating back to Bayes' original paper [Ba] where $R = B = \Delta$ and the prior is uniform. Then a Bayesian would guess that the probability of a red ball being drawn is just $\mathbf{E}(p) = R/(R+B)$. This is indeed just what happens on the first draw from the urn. After each draw, the Bayesian updates his information on p, while the urn changes its composition. These two things happen in such a way that the Bayesian's guess as to the probability of choosing a red ball next is always given by the fraction of red balls in the urn. In some sense, the urn is just an analog computer, carrying out the Bayesian's computations. Thus the effect of a red draw increasing the chance of a future red draw by actually altering the composition of the urn may be interpreted as a pseudo-effect where all that has changed is knowledge about a hidden parameter.

This interpretation makes the Pólya urn a natural model on which to base prediction of future probabilities. With a uniform prior, the process is equivalent to the Bayes-Laplace scheme. According to Feller ([Fe] v. 1 ch. V sec. 2) Laplace used this model to (jokingly?) estimate the probability that the sun will rise tomorrow to be $1,826,214$ to 1 because it has risen every day for the past $5,000$ years or $1,826,213$ days. More recently, Blackwell and Macqueen [BM] have proposed the following generalization. Let the index set of colors be arbitrary, and let μ be a probability measure on this set corresponding to the initial composition of the urn. Pick a random ball, say its color is i, and return it to the urn along with an extra ball of color i. Now the composition of the urn is $(\mu + \delta_i)/2$. Continuing in this way, the composition of the urn converges to a random measure. This random measure is almost surely atomic and is sometimes called a *Dirichlet process* or a *Ferguson's Dirichlet process* (after [Fer]) with parameter measure μ. It can be argued (see [Fer]) that the law of the Dirichlet process is a good prior for the distribution of a random measure, at least in situations where the measure is expected to be atomic, in which case this urn model has a useful Bayesian interpretation. At the least, the prior is mathematically convenient. A similar model has been proposed to predict and quantify clumping in the location of new industries [Ar2].

Various types of scientific data can be expected to fit the Pólya statistics, meaning that data ranging from 1 to n has the same distribution as the number of red draws among n draws from a Pólya urn of some fixed initial composition. This is because the scheme may be thought of revealing a hidden parameter that changes from one group of trials to the next. In [Ja] it is proposed that the number of male children in a family of a specified size is not binomially distributed, since the parents may have a predisposition toward producing one sex, but might fit the Pólya statistics better. The data actually presented in the paper are not convincing one way or the other. Mackerro and Lawson [ML] make a similar proposal for the number of days in a season that are suitable for crop spraying; the evidence in their case is somewhat more compelling. Several hypothetical examples are given in Cohen's friendly account of random limits in observed data [Co].

The inference process can also be reversed. The cross-section of the number of particles created in high-speed hadronic collisions is known experimentally to have a Greenwood-Yule distribution. This leads physicists to look for a mechanism responsible for the Pólya urn-like behaviour (see [Mi] and [YMN]).

Remark: Of historical interest is the following limiting case of the Pólya distribution which predates the general case by a few years. Suppose that the proportion of red balls is initially infinitessimally small, but that we draw often enough to expect an average of ρ red draws. More formally, suppose we let $R/\Delta = r$ remain fixed while $B/\Delta \to \infty$ and we look at the number ρ of red draws among the first N draws, where $N = \lambda(1 + B/\Delta R)$ so that $\mathbf{E}(\rho)$ remains constant at λ. Then the distribution of ρ converges to the Greenwood-Yule distribution, which is a generalization of the Poisson distribution that is sometimes also called the Eggenberger-Pólya distribution. The distribution is given by

$$prob(\rho = i) = \left(\frac{c}{1+c}\right)^r \frac{r(r+1)\cdots(r+i-1)}{i!(1+c)^i} \tag{1}$$

where $c = r/(r+\lambda)$. See [GY] for a wordy discussion of this distribution as it relates to their data on accidents in industry.

2.2 Generalizations of Pólya's urn

Bernard Friedman [Fri] considers the following generalization of Pólya's urn. It is perhaps the simplest generalization, and it is quite instructive. Let the urn begin as before with R red balls and B black balls, but now when a ball is drawn, replace it along with α extra balls of the same color and β extra balls of the opposite color. The introduction of the parameter β radically changes the behaviour of the process. It will turn out that the *vertex-reinforced random walk* of chapter five is essentially like a Friedman urn and bears little resemblance to the *edge-reinforced random walk* of chapter four which is a Pólya urn type process. In the Friedman urn-like examples such as the Friedman urn itself, the successive fractions of red balls no longer form a martingale, but tend instead toward 1/2. This causes the set of possible limits to shrink from the whole interval $[0,1]$ to the set of points p such that the expected fraction of red balls after the next draw is p whenever the current fraction is p. This set is often a singleton (as in the above example) or a discrete set.

David Freedman [Fre] uses a moment calculation to obtain the precise rate of convergence of the successive fractions of red balls, $v_1, v_2, \ldots$ to 1/2.

Theorem 2.3 ([Fre]) *Let* $\rho = (\alpha-\beta)/(\alpha+\beta)$ *and* $\sigma^2 = (\alpha-\beta)^2/(1-2\rho)$. *Then*

$$\begin{aligned} (v_n - 1/2)\sqrt{n} &\to N(0,\sigma^2) \quad \text{if} \quad \rho < 1/2 \\ (v_n - 1/2)\sqrt{n\log(n)} &\to N(0,(\alpha-\beta)^2) \quad \text{if} \quad \rho = 1/2 \\ (v_n - 1/2)n^{1-\rho} &\to Z \quad \text{if} \quad \rho > 1/2 \end{aligned} \tag{2}$$

where Z is a non-degenerate random variable and the convergence is in distribution. □

Another way of generalizing Pólya's urn is to allow the probability of drawing a red ball to be some function of the fraction of red balls other than the identity. This will be called a nonlinear urn scheme. Hill, Lane and Sudderth [HLS] consider a simple nonlinear urn scheme which adheres to Pólya's original practice of always putting back one extra ball of the color drawn. Letting $f(x)$ denote the probability of drawing a red ball when the fraction of red balls is x, they show that the successive fractions of red balls, $v_1, v_2, \ldots$ must converge whenever the f is not too discontinuous: the set of points p such that $f(p) \neq p$ but every neighborhood of p contains some x_1 and x_2 for which $f(x_1) < x_1$ and $f(x_2) > x_2$ must be nowhere dense in $[0,1]$. For continuous f it is not hard to show that the limiting value v of the fractions $v_1, v_2, \ldots$ must satisfy $f(v) = v$ with probability 1. To see this, let v be any point with $f(v) \neq v$. Assume without loss of generality that $f(x) > v + \epsilon$ for all x in some neighborhood of v. If $v_n \to v$ then the probability of choosing a red ball at each step is eventually greater than $v + \epsilon$, so convergence of the v_n to v violates the strong law of large numbers. Hill, Lane and Sudderth also obtain a second order result about the behaviour of f near points of convergence of v_n. Say that a point v where the graph of f meets the diagonal is an upcrossing if $f(x) < x$ for $x < v$ and $f(x) > x$ for $x > v$. Say it is a downcrossing if $f(x) < x$ for $x > v$ and $f(x) > x$ for $x < v$. If f is differentiable at v then these will hold when $f'(v) > 0$ or $f'(v) < 0$ respectively.

Theorem 2.4 ([HLS]) *In the above urn scheme, suppose* $0 < f(x) < 1$ *for all* $x \in [0,1]$ *and let* $v \in (0,1)$ *be a fixed point for* f. *Then*

$$(i) \quad prob(v_n \to v) = 0 \text{ if } v \text{ is an upcrossing for } f$$
$$(ii) \quad prob(v_n \to v) > 0 \text{ if } v \text{ is a downcrossing for } f.$$

Remark: The conclusion of part (i) is only significant when the set of v for which $f(v) = v$ is discrete; in this case, with probability one, the process does not converge to any point outside the prescribed set. In the case of Pólya's urn, $f(v)$ is always equal to v and the probability of convergence to any point is zero though in fact the fraction always does converge somewhere.

The intuitive idea behind this theorem is that at an upcrossing, v_{n+1} is expected to be further away from v than v_n, while at a downcrossing there is a restoring force that makes v_{n+1} closer to v than v_n, at least in expectation. To prove the non-convergence of v_n to v when v is an upcrossing, they use an elegant gambling-optimization argument. They allow the player to vary f after seeing the outcome of the draws up to time n, but always subject to the upcrossing condition: $(f(v_n) - v_n)(v_n - v) \geq 0$. They show that the best way for the player to maximize a payoff that is a function of the limit of the v_n, namely a beta density with peak at v, is always to choose $f(v_n) = v_n$. For this choice of f the distribution of the limit is a beta, and in particular has no atom at v.

Some independent work by Blum and Brennan [BB] implies the second part of this theorem under slightly stronger assumptions on f than are used in [HLS]. Their technique is to treat $n\, v_n$ as a sum of almost independent random variables. In chapter three, we shall generalize part (i) of this theorem to multicolor urns; the idea for the proof is present in [HLS], although we cannot duplicate their elegant methods. The proof uses instead an analysis of the increments of the process that is similar to the one in [BB], but more detailed.

Athreya generalizes a linear multicolor Pólya urn in two ways [At]. As before, the extra balls added at each step are always of the color drawn, but Athreya allows this number to be a random non-negative integer. Athreya also allows the distribution of this random number to be different for each color. As might be expected, the only colors that persist as a positive fraction in the limit are those with the greatest mean reinforcement.

Suppose there are d colors and that the generating functions for the distributions of the random numbers of additions for each color are $f_1, \ldots, f_d$ with finite means $f_i'(1) = \lambda_i$. Suppose in addition that the random number of additions a_i of each color are never zero and satisfy $\mathbf{E}|a_i \log a_i| < \infty$. Then letting the fraction of balls of color i at time n be denoted $v_i(n)$, we have

Theorem 2.5 (Athreya) *The vector $\vec{v}(n) = (v_1(n), \ldots, v_d(n))$ converges almost surely to a random vector $\vec{v}$ with $v_1 + \cdots + v_d = 1$. Furthermore, if $\lambda_1 = \cdots = \lambda_r > \lambda_{r+1} \geq \cdots \geq \lambda_d$, then with probability 1, $v_i = 0$ for $i > r$.*

Athreya's idea for proving this is to map the urn process into a continuous time branching process. Suppose each ball in the urn at any time waits for an exponential amount of time and then vanishes and gets replaced by $1 + y$ balls, where y is a random non-negative integer distributed appropriately for whatever color just vanished. Although this alters the time scale, the probability that the next ball to vanish has color i is always just the fraction of balls in the urn of color i, since the exponential waiting time has no memory. The details of mapping the urn process to a branching process are carried out in [AK]; The theorem then follows from the standard results about branching processes. This method of proof, unlike all the others cited above, actually does require the number of balls to be integral, since the results on branching processes are only applicable in this case (at least as they are usually stated).

2.3 Other processes with reinforcement

In this section we consider some random processes with reinforcement that are not urn schemes. There are many in the literature and I will not pretend to discuss them all; instead I will mention two of the larger categories of such processes, namely stochastic approximation and learning models. These are chosen because they relate loosely to the processes in chapter five below.

Stochastic approximation was first introduced in 1951 by Robbins and Monro [RM]. They consider the following problem, which may be thought of as an attempt to modify Newton's method for root-finding when the values of the function have noise added. Suppose $y = y(x, \omega)$ is a random function of x with mean $M(x)$, where we do not necessarily know $M(x)$. We are allowed to sample $y(x_1), y(x_2), \ldots$ at points $x_1, x_2, \ldots$ of our choosing, and we want to approximate a solution of $M(x) = 0$ in the sense that $x_n \to x$ in probability for some x such that $M(x) = 0$. For example, suppose $M(x) = x - \theta$ for some unknown θ and $y(x_n) = x_n - \theta + \xi_n$ where $\{\xi_n\}$ are independent, identically distributed normal random variables with zero mean. Then the problem asks for an algorithm for generating x_{n+1} as a function of $x_1, \ldots, x_n, y_1, \ldots, y_n$ in such a way that $x_n \to \theta$ in probability. Robbins and Monro consider the case where M is monotone; we will assume hereafter that M is non-increasing. They propose the recursive estimate

$$x_{n+1} = x_n + a_n y_n \tag{3}$$

where $\{a_n\}$ is a sequence of type $1/n$, i.e. there are constants c_1, c_2 such that

$$c_1/n \leq a_n \leq c_2/n \text{ for all } n.$$

The idea is that the addition of $a_n y_n$ is a restoring force toward θ, since its expected value is positive when $x_n < \theta$ and negative when $x_n > \theta$. The reason for the particular choice of asymptotic magnitude of a_n is that $\sum a_n$ must be infinite in order to be able to compensate for any initial errors, but $\sum {a_n}^2$ must be finite if x_n is to converge.

Theorem 2.6 ([RM] theorem 2) *Suppose $\{a_n\}$ is of type $1/n$ and M is a nonincreasing function of x satisfying*

$$M(\theta) = 0 \tag{4}$$
$$M'(\theta) < 0 \tag{5}$$

for some θ. Suppose also that $y(x,\omega)$ is bounded almost surely. Then $\mathrm{E}(x_n - \theta)^2 \to 0$.
□

This result specializes to a version of part (*ii*) of Hill, Lane and Sudderth's theorem. For each x the random function $y(x,\omega)$ returns either $-x$ or $1-x$ with respective probabilities $f(x)$ and $1-f(x)$. For each x the mean return is therefore $\mathrm{E}(y(x)) = f(x) - x$. The Robbins-Monro algorithm finds a "root" of y, i.e. an x for which $\mathrm{E}(y(x)) = 0$. The algorithm uses only the sample values $y(x,\omega)$; it does not know f. Implementing the Robbins-Monro algorithm with $a_n = 1/n$ gives

$$x_{n+1} = \frac{nx + \xi}{n+1} \tag{6}$$

where ξ is a random variable that is 1 with probability $f(x)$ and 0 otherwise. The rule governing the sequence $x_1, x_2, \ldots$ of successive fractions in a Hill-Lane-Sudderth nonlinear urn whose urn function is f is precisely equation (6). The processes therefore have identical laws and since conditions (4) and (5) above are just the conditions for a differentiable downcrossing, part (*ii*) of Hill, Lane and Sudderths theorem (theorem 2.4) follows in the differentiable case from theorem 2.6 and the fact that θ is the unique fixed point for f because M is non-increasing.

The Robbins-Monro process hass been widely studied and generalized since 1951 (see [NH] for a survey of this field). A natural way to generalize it is to raise the dimension, so $\vec{y}$ is a random vector $\vec{y}(\vec{v},\omega)$. Immediately, it becomes more difficult to establish the convergence of $\vec{v}_n$ under the scheme $\vec{v}_{n+1} = \vec{v}_n + a_n\vec{y}_n$. In fact there are reasonable choices for $\vec{y}$ for which $\vec{v}_n$ does not converge, but moves in (ever slower) cycles. One approach to determining convergence is the method of Liapunov functions. This means finding, for the given approximation algorithm, a non-negative potential (scalar function) that decreases in expectation with each iteration. The following result underlies most convergence theorems.

Theorem 2.7 ([NH] ch. 2 thm. 5.1) *Let $\{v_n\}$ be as above, let H be a non-negative function and suppose $\{a_n\}$ is any positive sequence with $\sum a_n = \infty$. Let G be any measurable set and let τ be the first exit time from G. If*

$$\mathbf{E}(H(\vec{v}_{n+1}) \mid n, \vec{v}_n) \leq H(\vec{v}_n) - a_n$$

whenever $\vec{v}_n \in G$, then prob$(\tau < \infty) = 1$. □

This is essentially the reasoning used in chapter five below, lemma 5.7 and theorem 5.8.

To see how the notion of upcrossing in theorem 2.4 generalizes, consider any point $\vec{p}$ for which $M(\vec{p}) = 0$. In order for $a_n \vec{y}_n$ to be a restoring force, the linear approximation to $M(\vec{p}+\vec{v})$ for small $\vec{v}$ should be $T\vec{v}$ where T is a negative definite matrix. This ensures that $\vec{v}_{n+1}$ is always pushed back toward $\vec{p}$ in some appropriate metric. Requiring T to be negative definite is analogous to condition (5) in the one-dimensional case. Under certain conditions, this implies $\vec{v}_n \to \vec{p}$ just as in theorem 2.4 (ii). Conversely, if T is positive definite then a non-convergence theorem similar to theorem 2.4 (i) holds. For various versions of this theorem see [NH] ch. 5 sec. 3 and following.

Specializing once more to an urn scheme, let $\vec{F} : \Delta \to \Delta$ be any continuous function from the unit simplex $\Delta \subseteq \mathbf{R}^d$ to itself. Consider an urn with the fraction of balls of color i at step n given by $v_i(n)$ and let a ball of color i be drawn with probability $F_i(\vec{v}(n))$, being replaced along with one extra ball of the same color. This is the multicolor version of the simple nonlinear urn scheme studied in [HLS]. Arthur, Ermol′ev and Kaniovskii use stochastic approximation theorems such as the one above to derive the following theorems about nonlinear urns with more than 2 colors.

Theorem 2.8 ([Ar1] theorem 1) *Let a multicolor nonlinear urn be governed by a continuous urn function $\vec{F}$ as above. Suppose for some $\vec{p} \in \Delta$, $\vec{F}(\vec{p}) = \vec{p}$ and either*

$$\begin{aligned} &(i) \qquad C(\vec{v}-\vec{p}, \vec{F}(\vec{v})-\vec{p}) \;>\; 0 \text{ for all } x \notin Q \\ \text{or} \quad &(ii) \qquad C(\vec{v}-\vec{p}, \vec{F}(\vec{v})-\vec{p}) \;<\; 0 \text{ for all } x \notin Q \end{aligned}$$

where C is a symmetric positive definite bilinear form and Q is the set of fixed points of $\vec{F}$. Then $\vec{v}_n \to \vec{v}$ almost surely for some random variable $\vec{v}$ with $prob(\vec{v} \in Q) = 1$. □.

Theorem 2.9 ([Ar1] theorem 2) *Assume the hypotheses of the above theorem with case (ii) holding. Suppose also that $\vec{F}$ satisfies a Hölder condition with exponent $\mu \in (0,1]$ in a neighborhood of $\vec{p}$. Then $prob(\vec{v}_n \to \vec{p}) = 0$.* □

An important heuristic is pointed out in both [NH] and [Ar1]: the multidimensional stochastic approximation process

$$\vec{v}_{n+1} = \vec{v}_n + a_n(\vec{y}(\vec{v}_n, \omega) - \vec{v}_n) \quad (7)$$

is in some sense a discrete version of the differential equation

$$\frac{d}{dt}\vec{v}(t) = a(t)(\vec{F}(\vec{v}(t)) - \vec{v}(t)) \quad (8)$$

where $\vec{F}(\vec{x}) = \mathbf{E}(\vec{y}(\vec{x}))$. The differences are that the stochastic version (7) is discrete and also that is has noise added (which is mean zero by definition of $\vec{F}$. Because of the noise, the reader may prefer to think of (7) as analogous instead to a diffusion, for which (8) then gives the drift.

By a time change, assume $a(t) \equiv 1$. Then the flow for the differential equation has a critical point wherever $\vec{F}(\vec{v}) = \vec{v}$ and the linear approximation T to $\vec{F}(\vec{p} + \vec{v})$ being positive or negative definite corresponds to the flow having a repelling or attracting node respectively. Of course, one expects that the stochastic version will converge only to stable critical points, which are the attracting nodes. However, the case where T has both positive and negative eigenvalues and the critical point is thereby unstable – though only in some directions – is not dealt with. One reason may be that there are few theorems stating that the process converges in this case, so a theorem giving non-convergence to such a point is not worth very much. We consider such cases in chapter five and in the latter part of chapter three.