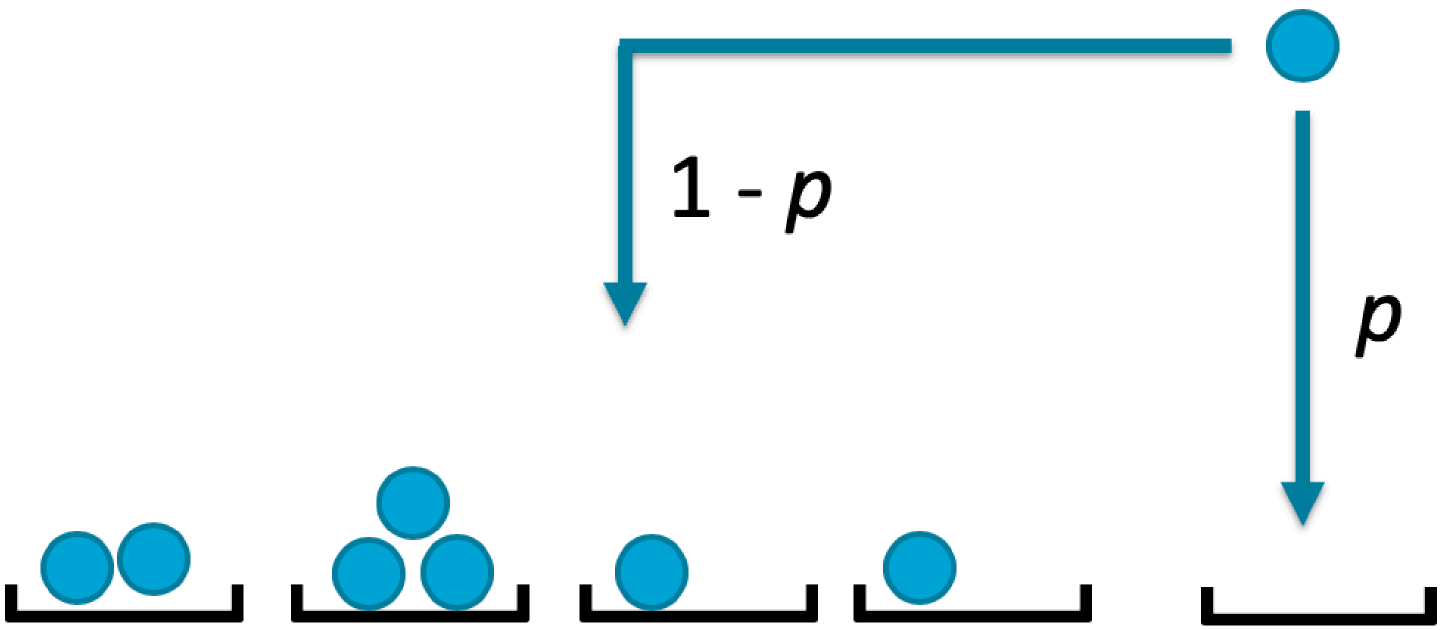
Finite & Infinite Polya Processes

Arian Joyandeh

Bachelor End Project



TU Delft

Finite & Infinite Polya Processes

by

Arian Joyandeh

to obtain the degree of Bachelor of Science
at the Delft University of Technology,
to be defended publicly on Friday July 19, 2019 at 12:00 AM.

Student number:	4568745	
Project duration:	April 22, 2019 – July 19, 2019	
Thesis committee:	Dr. W. M. Ruszel,	TU Delft, supervisor
	Dr. J. Söhl,	TU Delft
	Dr. J. G. Spandaw ,	TU Delft

TU Delft

Abstract

In this thesis we shall consider a generalization on Pólya Processes as have been described by Chung et al. [7]. Given finitely many bins, containing an initial configuration of balls, additional balls arrive one at a time. For each new ball, a new bin is created with probability p, or with probability $1-p$ this new ball shall be placed in an existing bin such that the probability of this ball ending in a specific bin, is proportional to $f(m)$ where m is the number of balls currently in that bin and f is some feedback function.

We shall show that for $p = 0$, which will be defined as Finite Pólya Processes, the behaviour of the process can be classified into one of three mutual exclusive regimes: Monopolistic Regime, Eventual Leadership Regime or Almost-Balanced Regime. This behaviour solely depends on the convergence of the following sums: $\sum_{n\geq 1} f(n)^{-1}$ and $\sum_{n\geq 1} f(n)^{-2}$. We shall explore the limiting distribution of fractions of balls in bins when $f(x) = x$, which is a known result for classical multi-coloured Pólya Urn problems.Using a similar method, we find a limiting distribution for Finite Pólya Processes with general positive linear feedback functions, which has not previously been researched.

We then consider the case where $p > 0$, which are defined as Infinite Pólya Processes and restrict the feedback function to be of the form $f(x) = x^{\gamma}$ where $\gamma \in \mathbb{R}$. We shall show that if $\gamma > 1$, almost surely one bin will dominate or a new bin will be created. We shall show that for $\gamma = 1$ a preferential attachment scheme arises. We consider $\gamma < 1$ under the assumptions that some limits exist and show that the fraction of bins having m balls shrinks exponentially as a function of m. Finally, we reflect on our results and discuss interesting future research subjects.

Preface

This thesis has been written to fulfill the graduation requirements for the Bachelor Applied Mathematics under the supervision of Dr. Ruszel on behalf of the department of Applied Probability of the faculty EEMCS at the Delft University of Technology.

A very interesting concept that I learned of during my research, is path-dependence, which refers to how a resulting state of a random process depends on the path it has taken to achieve that resulting state. This idea resonated with me. Every choice I have made during my research has lead up to this final thesis.

There have been days where I was stuck on a single sentence of a proof, that I could just not follow for some reason. Fortunately, eventually I understood even those parts. I believe that I could not have understood those parts if I had not initially struggled with them for so long: All of my past experiences have been crucial to having the understanding of the processes that I have now. In this sense, research is a path-dependence process as well.

One of the first choices I made that have brought me onto this path, is following the Advanced Probability course by Dr. Ruszel. I was intrigued by how mathematics can be used to gain a grasp of such, prior to me, ungraspable phenomenons as random processes. That made it clear to me that the only thesis subject for me was one with regard to Applied Probability.

I want to sincerely thank Dr. Ruszel for introducing me to this wonderful world of Probability and for her guidance during this project. I enjoyed a lot of freedom during the course of my research, but whenever there was something that I thought I had to discuss with someone, you were there with good advice and support.

Finally, I would like to thank J. Söhl and J. Spandaw for taking seat in my assessment committee.

Arian Joyandeh
Delft, July 2019

Contents

1 Introduction

Pólya Processes are random processes that can be used as models with a wide range of applications. This wide range of applications includes, but is not limited to: The development of neurons into either axons or dendrites[14]; The random limiting market share for selectively neutral products[2]; the growth of random networks, such as the world-wide web[4] and finding a generalized probability distribution for the Bose-Einstein statistic[4].

A Pólya Process is a random process such that: given finitely many bins k, containing an initial configuration of balls $(b_1^0, \dots, b_k^0)$, additional balls arrive one at a time. For each new ball, a new bin is created with probability p, or with probability $1 - p$ this new ball shall be placed in an existing bin such that the probability of this ball ending in a specific bin, is proportional to $f(m)$ where m is the number of balls currently in that bin and f is some feedback function. The only restriction we put on f, shall be that f is a positive function with respect to the natural numbers: $f : \mathbb{N} \to (0, \infty)$. With this definition of a Pólya Process we differ from the definition as proposed by Chung et al.[7], who restrict the feedback functions to be of the form $f(x) = x^\gamma$ where $\gamma \in \mathbb{R}$. Another aspect in which we differ from their definition, is that we allow an arbitrary initial configuration of balls, whereas Chung et al. only consider a standard initial configuration, where every bin contains one ball.

Pólya Processes can be split up in two categories, Finite Pólya Processes, where $p = 0$ and Infinite Pólya Processes, where $p > 0$. A great amount of literature is available for Finite Pólya Processes [16] [20]. However, with regard to Infinite Pólya Processes, there is a lot less literature available, the most notable of these works is the paper by Chung et al.[7]. That lack of literature is not due to Infinite Pólya Processes not being useful, in fact, they are equivalent to the preferential attachment scheme as has been described by Barabási et al.[4], which has found great use in studies of random networks, such as social media networks.

Noting this discrepancy between availability in literature for both processes, we wish to use the available methods of analyzing Finite Pólya Process to analyze Infinite Pólya Process and gain more knowledge with regard to these Infinite Pólya Processes.

In chapter 2 we shall speak shortly about the mathematical background that is crucial for understanding the ideas presented in this thesis.

Following this, in chapter 3 we shall discuss results considering Finite Pólya Processes. First of all we will consider the limiting behaviour under the standard feedback function $f(x) = x$. Note that under the standard feedback function, the Finite Pólya Process is equivalent to the Pólya-Eggenberger Urn Model[10]. We shall proof that the limiting distribution of a 2-bin Finite Pólya Process under the standard feedback function and initial configuration (b_1^0, b_2^0) will be a Beta distribution with parameters b_1^0 and b_2^0. This proof shall be done using De Finetti's theorem, which is one of the standard methods of analyzing Finite Pólya Processes. We also note that for a k-bin Finite Pólya Process a similar result holds with regard to the Dirichlet distribution.

Then we shall consider the limiting behaviour under general feedback function in the sense of different kind of regimes using an Exponential Embedding, inspired by the work of Oliveira[19][18]. A very interesting result emerges in the fact that any Finite Pólya Process can be classified in one of three distinct regimes: Monopolistic Regime, Eventual Leadership Regime or Almost-Balanced Regime. To which regime one Finite Pólya Process belongs, is solely dependent on the convergences of $\sum_{n\geq 1} f(n)^{-1}$ and $\sum_{n\geq 1} f(n)^{-2}$, where $f(n)$ is the feedback function of the Finite Pólya Process.

After having done this, we shall explicitly consider the class of linear feedback functions and we shall find explicit limiting distribution, for which there has not been any previous literature. We find that under feedback function $f(x) = b \cdot x + a$ where $a, b \in (0, \infty)$, for a 2-bin Finite Pólya Process with initial configuration (b_1^0, b_2^0) the limiting distribution shall be Beta distributed with parameters $b_1^0 + \frac{a}{b}, b_2^0 + \frac{a}{b}$ and we propose a conjecture for the k-bin scenario: a Dirichlet distribution shall then be the limiting distribution.

In chapter 4 we shall consider Infinite Pólya Processes. First we will discuss why the methods discussed in chapter 3 will not be applicable to Infinite Pólya Processes. Then we will restrict our feedback functions to be of the form $f(x) = x^\gamma$ where $\gamma \in \mathbb{R}$.

This chapter will largely be a revision of the proofs in Chung et al.[7] sections 3 and 4. We shall, however, add a great deal of probabilistic rigour that was lacking in their paper.

Lastly, in chapter 5 we shall conclude our results and reflect onto our results and we shall discuss some interesting future research subjects with regard to Pólya Processes.

2

Mathematical Background

2.1. Probability Theory

Any work regarding random processes, will have to deal with probability theory. We shall note the most important probabilistic theorems and definitions in this section. We will assume a basic knowledge of probability theory is known by the reader and we shall not be concerned with specific proofs of theorems. We refer the reader to 'Theory of Probability and Random Processes' by Kolarov and Sinai[15].

We shall denote the probabilistic space as $(\Omega, F, \mathbb{P})$, where Ω is the state space, which is the set of all possible outcomes. F is a σ-algebra on this state space, which one can interpret as all the collection of all events we would like to consider. $\mathbb{P}$ is the probability measure on this probability space, which is a function on every element of F, such that $\mathbb{P} : F \to [0,1]$ and $\mathbb{P}(\Omega) = 1$.

We note that there are events that happen with probability 0. In some cases we wish to ignore these events, that is, we are not concerned with the events that happen with probability 0. For that purpose, the idea of an event happening almost surely will be introduced.

Definition 2.1. *We say that an event $A \subset \Omega$ happens almost surely (a.s.) if $\mathbb{P}(A) = 1$.*

It is useful to consider a set of k events, $\{A_j\}_{j \in \{1,\ldots k\}}$. These events might be independent of each other, that means that the outcome of one of these events does not influence the outcome of another event. We shall define this concisely:

Definition 2.2. *A set $\{A_j\}_{j \in \{1,\ldots k\}}$ is said to be independent if and only if for any $1 \leq j \leq k$ the following equality holds:*

$$\mathbb{P}\Big(\bigcap_{i=1}^{j} A_{m_i}\Big) = \prod_{i=1}^{j} \mathbb{P}\big(A_{m_i}\big),$$

with $m_1, \ldots, m_k \in 1, \ldots n$.

If we consider a sequence of events $(A_n)_{n \geq 1}$, the limiting behaviour of this sequence might be of interest.

Definition 2.3. *For a sequence of events $(A_n)_{n \geq 1}$ we define:*

$$\liminf_{n \to \infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_n = \{A_n \text{ a.a.}\},$$

$$\limsup_{n \to \infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_n = \{A_n \text{ i.o.}\}.$$

Consider $\{A_n \text{ a.a.}\}$, this means intuitively that eventually, all events A_n occur, that is, there are just finitely many events A_n that do not occur. We note that almost always and eventually mean the same thing in this setting. Now consider $\{A_n \text{ i.o.}\}$, which can be interpreted as infinitely many of the A_n events occur. Note that this can also mean that infinitely many events A_n do not occur.

Theorem 2.1. *Consider a sequence of events $(A_n)_{n\geq 1}$, then $\mathbb{P}(\{A_n \text{ i.o.}\}) = 1 - \mathbb{P}(\{A_N \text{ a.a.}\})$.*

A great deal of knowledge regarding random processes can be gained if we understand its behaviour in the limiting sense. There is a very useful theorem regarding these probabilities of limiting events:

Theorem 2.2 (Borel-Cantelli Lemma)**.** *Consider $(A_n)_{n\geq 1}$ a sequence of events in some probability space.*

1. *If $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$, then $\mathbb{P}(\{A_n \text{ a.a.}\}) = 0$.*

2. *If $(A_n)_{n\geq 1}$ is an independent sequence of events and $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$, then $\mathbb{P}(\{A_n \text{ a.a.}\}) = 1$.*

Theorem 2.2 is a very useful theorem, however, the restriction of independence in the second part of the theorem proofs to be quite a harsh restriction. We shall introduce another theorem with regard to limiting events that will be more lenient with its restriction on the sequence of events. This theorem was found in page 110 of 'Probability: a Graduate Course' by Gut[11].

Theorem 2.3. *Let $(A_n)_{n\geq 1}$ be an arbitrary sequence of events such that:*

$$\liminf_{n\to\infty} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{P}(A_i \cap A_j)}{\left(\sum_{j=1}^{n} \mathbb{P}(A_j)\right)^2} = 1, \tag{2.1}$$

then

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty \Rightarrow \mathbb{P}(\{A_n \text{ i.o.}\}) = 1.$$

Whenever researching random processes, the interesting aspects are usually functions from the probability space to some other measurable space. These functions are called random variables. It is known that any random variable X has an expectation, which we shall note as $\mathbb{E}(X)$ and a variance, which we shall note as $\mathrm{Var}(X)$.

We will note that any random process can be seen as as some sequence of random variable, that is why we shall speak on known results with regard to convergence of these random variables.

Definition 2.4. *Let $(X_n)_{n\geq 1}$ be a sequence of random variables. We say that X_n converges almost surely (a.s.) to X if and only if:*

$$\mathbb{P}(X_n \xrightarrow{n\to\infty} X) = 1. \tag{2.2}$$

In practice, we will not use definition 2.4 for proofing almost sure convergence, but rather we shall use theorems that imply almost sure convergence, which will be noted shortly:

Theorem 2.4. *Let $(X_n)_{n\geq 1}$ be a sequence of random variables. X_n converges almost surely to X if and only if for all $\epsilon > 0$:*

$$\mathbb{P}(\{|X_n - X| < \epsilon\} \text{ a.a.}) = 1. \tag{2.3}$$

Theorem 2.5 (Strong Law of Large Numbers)**.** *Let $(X_n)_{n\geq 1}$ be a sequence of pairwise independent, random variables, then we know:*

$$\frac{\sum_{i=1}^{n}(X_i - \mathbb{E}(X_i))}{n} \xrightarrow{a.s.} 0.$$

Theorem 2.6 (Kolmogorov's Three Series Theorem). *Consider $(X_n)_{n\geq 1}$ a sequence of independent random variables, then $\sum_{n=1}^{\infty} X_n$ converges almost surely if and only if there is an $A > 0$ such that:*

1. $\sum_{n=1}^{\infty} \mathbb{P}(|X_n| \geq A) < \infty$,
2. $\sum_{n=1}^{\infty} \mathbb{E}(X_n \mathbb{1}_{\{|X_n| \leq A\}}) < \infty$,
3. $\sum_{n=1}^{\infty} Var(X_n \mathbb{1}_{\{|X_n| \leq A\}}) < \infty$.

Sometimes we know that some sequence $(X_n)_{n\geq 1}$ converges almost surely to X, but then we are interested in what happens with $g(X_n)$ where g is a continuous function. For such a scenario, there is the continuous mapping theorem.

Theorem 2.7 (Continuous Mapping Theorem). *Suppose some sequence $(X_n)_{n\geq 1}$ converges almost surely to X, and g is some continuous function, then $g(X_n)$ converges almost surely to $g(X)$.*

There are some probabilistic inequalities that will make our life a lot easier when we are trying to proof probabilistic statements.

Theorem 2.8 (Markov's inequality). *Let X be a non-negative random variable and $a > 0$, then we know:*

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

Theorem 2.9 (Chebyshev's inequality). *Let X be a random variable with finite mean $\mathbb{E}(X)$ and finite, non-zero variance $Var(X)$, then for any $k > 0$:*

$$\mathbb{P}(\left|X - \mathbb{E}(X)\right| \geq k) \leq \frac{Var(X)}{k^2}.$$

2.2. Gamma function and Beta function

The Gamma function and the Beta function are functions that we will see a lot when considering Pólya Processes. That is why we will discuss them in this section.

Definition 2.5. *The Gamma function for any complex z with real positive part, is defined as*

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt.$$

Theorem 2.10. *The Gamma function for any positive integer n is:*

$$\Gamma(n) = (n-1)!,$$

and for any value z with real positive part:

$$\Gamma(z+1) = z\Gamma(z).$$

Definition 2.6. *The Beta function for any complex x, y with positive real parts, is defined as*

$$\mathrm{B}(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1} dt.$$

Theorem 2.11. *For any complex x, y with positive real parts:*

$$\mathrm{B}(x,y) = \frac{\Gamma(x)\cdot\Gamma(y)}{\Gamma(x+y)}.$$

Let us remark that we will not consider complex-valued inputs in the Beta function or the Gamma function.

Definition 2.7. *The Beta distribution for any real parameters α, β is defined as:*

$$f(x;\alpha,\beta) = \frac{1}{\mathrm{B}(\alpha,\beta)} x^{\alpha-1} \cdot (1-x)^{\beta-1} \mathbb{1}_{x\in(0,1)}.$$

Definition 2.8 (Pochhammer Symbol)**.** *We shall denote $[a]_n^+$ as the following product:*

$$[a]_n^+ = (a)(a+1)\cdots(a+n-1) = \prod_{j=0}^{n-1}(a+j).$$

Theorem 2.12. *Let $[a]_n^+$ be defined as in definition 2.8, then we know:*

$$[a]_n^+ = \frac{\Gamma(a+n)}{\Gamma(a)}.$$

Proof.

$$\begin{aligned}\frac{\Gamma(a+n)}{\Gamma(a)} &= \frac{\Gamma(a+n-1+1)}{\Gamma(a)} = \frac{(a+n-1)\Gamma(a+n-1)}{\Gamma(a)},\\ &= \frac{(a+n-1)(a+n-2)\cdots(a)\Gamma(a)}{\Gamma(a)} = \prod_{j=0}^{n-1}(a+j),\\ &= [a]_n^+.\end{aligned}$$

□

We note that the power of this theorem comes from the fact that a does not necessarily need to be a natural number and thus one can see this as an extension of the factorial property of the Gamma function for natural input, which has been described in theorem 2.10.

2.3. Dirichlet Distribution

The Dirichlet distribution is a family of continuous multivariate distributions parametrized by a vector $\theta = (\theta_1, \ldots, \theta_K)$. It can be interpreted as the multivariate generalization of the Beta Distribution.

Definition 2.9. *A Dirichlet distribution of order $K \geq 2$ and parameters $\theta = (\theta_1, \ldots, \theta_K)$ has probability density:*

$$f(x_1, \ldots, x_K \mid \theta) = \frac{1}{\mathrm{B}(\theta)} \prod_{i=1}^{K} x_i^{\theta_i - 1},$$

where $\forall i \in \{1 \ldots K\} : x_1, \ldots, x_K \geq 0$ and $\sum_{i=1}^K x_i = 1$. $\mathrm{B}(\theta)$ is defined as:

$$\mathrm{B}(\theta) = \frac{\prod_{i=1}^K \Gamma(\theta_i)}{\Gamma(\sum_{i=1}^K \theta_i)},$$

which is called the multivariate Beta function.

Note that thus a realization $\{x_1, \ldots, x_k\}$ of the Dirichlet distribution belongs to the $K-1$-standard simplex in space $\mathbb{R}^K$.

2.4. Exchangeability

Knowing that a sequence is exchangeable, is a very powerful tool for analyzing that sequence. We shall see that there are sequences with regard to Pólya Processes that are exchangeable sequences. The information in this section is largely inspired by Heath and Sudderth[12].

Definition 2.10. *An infinite sequence $(X_n)_{n\geq 1}$ is said to be exchangeable whenever all permutations of a finite subsequences have the same joint probability function, regardless of their order. That is for any sequence of arbitrary length k, $t_1, \dots, t_k$ and a random permutation of this sequence, $d_1, \dots, d_k$, $X_{t_1}, \dots, X_{t_k}$ has the same joint probability function as $X_{d_1}, \dots, X_{d_k}$.*

Theorem 2.13 (de Finetti's Representation Theorem)**.** *An infinite exchangeable sequence is distributed as a mixture of i.i.d. random variables. A binary sequence $\{X_n\}_{n\geq 1}$ is exchangeable if and only if there exists a distributing function $F(p)$ such that for any $n \geq 1$,*

$$\mathbb{P}(X_1 = x_1, \dots X_n = x_n) = \int_0^1 p^{S_n} \cdot (1-p)^{n-S_n} dF(p),$$

where $S_n = \sum_{j=1}^{n} x_j$, which is the number of realizations of $X_1, \dots, X_n$ that were equal to 1.

We note that this means that any exchangeable sequence can be seen as a mixture of independent, identically distributed random variables, and that the distribution F may be regarded as the prior of the random parameter p.

Remark 2.1. *Since $(X_j)_{j\in\{1\dots n\}}$ is a binary sequence of k successes in n trials, the probability that one observes exactly k success in n trials is equal to*

$$\mathbb{P}(S_n = k) = \binom{n}{k} \int_0^1 p^k \cdot (1-p)^{n-k} dF(p).$$

2.5. Exponential Distribution

The exponential distribution plays an essential role to the technique of the exponential embedding which shall be introduced in 3.2.1. In this section some aspects of this distribution will be discussed.

When X is a random variable such that $X \geq 0$ and, for some $\lambda > 0$, the following equation holds:

$$\mathbb{P}(X > t) = e^{-\lambda t} \ (t \geq 0). \tag{2.4}$$

We say that X is a random variable with an exponential distribution with parameter λ. We note this as $X \stackrel{d}{=} \text{Exp}(\lambda)$.

Lemma 2.1. *For a random variable X that is distributed exponentially with parameter $\lambda > 0$:*

$$\mathbb{E}(X) = \frac{1}{\lambda}, \ Var(X) = \frac{1}{\lambda^2}.$$

Proof.

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} t \cdot f_X(t) dt = \int_0^{\infty} \lambda t e^{-\lambda t} dt = \frac{1}{\lambda}.$$

$$\mathbb{E}(X^2) = \int_0^{\infty} t^2 e^{-\lambda t} dt = \frac{2}{\lambda^2}.$$

$$\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

□

Lemma 2.2 (Lack of memory)**.** *If $X \stackrel{d}{=} Exp(\lambda)$ and $Z \geq 0$ is a random variable independent from X, then*

$$\mathbb{P}(X - Z > t \mid X > Z) = \mathbb{P}(X > t).$$

Proof.

$$\begin{aligned}\mathbb{P}(X-Z>t \mid X>Z) &= \frac{\mathbb{P}(\{X-Z>t\}\cap\{X>Z\})}{\mathbb{P}(X>Z)} = \frac{\mathbb{P}(X-Z>t)}{\mathbb{P}(X>Z)},\\ &= \frac{\int_0^\infty \mathbb{P}(X>s+t)\cdot f_Z(s)ds}{\int_0^\infty \mathbb{P}(X>s)\cdot f_Z(s)ds} = \frac{\int_0^\infty e^{-\lambda(t+s)}\cdot f_Z(s)ds}{\int_0^\infty e^{-\lambda s}\cdot f_Z(s)ds},\\ &= e^{-\lambda t}\frac{\int_0^\infty e^{-\lambda(s)}\cdot f_Z(s)ds}{\int_0^\infty e^{-\lambda s}\cdot f_Z(s)ds} = \mathbb{P}(X>t)\cdot 1.\end{aligned}$$

□

Lemma 2.3 (Minimum property). *If $\{X_i \overset{d}{=} \text{Exp}(\lambda_i)\}_{i\in\{1,\dots,k\}}$ is a set of independent random variables, then*

$$X_{\min} = \min_{i\in\{1,\dots,k\}} X_i \overset{d}{=} \text{Exp}(\sum_{i=1}^k \lambda_i), \tag{2.5}$$

and for all $i \in \{1,\dots,k\}$

$$\mathbb{P}(X_i = X_{\min}) = \frac{\lambda_i}{\sum_{j=1}^k \lambda_j}. \tag{2.6}$$

Proof. To show that equation 2.5 holds, we need to proof that $\mathbb{P}(X_{\min} > t) = e^{-t\cdot(\sum_{i=1}^k \lambda_i)}$.

$$\begin{aligned}\mathbb{P}(X_{\min}>t) = \mathbb{P}(X_1>t,\dots,X_k>t) &= \prod_{i=1}^k \mathbb{P}(X_i>t)\\ &= \prod_{i=1}^k e^{-t\lambda_i} = e^{\sum_{i=1}^k -t\lambda_i} = e^{-t\cdot\sum_{i=1}^k \lambda_i}.\end{aligned}$$

Now we shall focus on proofing the statement in equation 2.6. We introduce $K_i = \{1,\dots,k\}\setminus\{i\}$.

$$\begin{aligned}\mathbb{P}(X_i = X_{\min}) &= \mathbb{P}(\forall j\in K_i : X_i < X_j) = \int_0^\infty \mathbb{P}(X_i < X_j \mid X_i = t)\cdot f_{X_i}(t)dt\\ &= \int_0^\infty \mathbb{P}(\forall j\in K_i : X_j > t)\cdot \lambda_i e^{-\lambda_i t}dt = \int_0^\infty \mathbb{P}(\forall j\in K_i : X_j>t)\cdot\lambda_i e^{-\lambda_i t}dt,\\ &= \int_0^\infty \prod_{j\in K_i}\mathbb{P}(X_j>t)\cdot\lambda_i e^{-\lambda_i t}dt = \lambda_i\int_0^\infty e^{-\lambda_i t}\cdot\prod_{j\in K_i} e^{-\lambda_j t}dt = \lambda_i\int_0^\infty\prod_{j=1}^k e^{-\lambda_j t}dt,\\ &= \lambda_i\int_0^\infty e^{-t\cdot(\sum_{j=1}^k\lambda_j)}dt = \frac{\lambda_i}{-\sum_{j=1}^k\lambda_j}e^{-t\cdot(\sum_{j=1}^k\lambda_j)}\Big|_{t=0}^{t=\infty} = \frac{\lambda_i}{-\sum_{j=1}^k\lambda_j}(0-1) = \frac{\lambda_i}{\sum_{j=1}^k\lambda_j}.\end{aligned}$$

□

Definition 2.11. *A continuous random variable X is called symmetric, if there exists an $x_0 \in \mathbb{R}$ such that for all $t > 0$:*

$$\mathbb{P}(X \le x_0 + t) = \mathbb{P}(X \ge x_0 - t).$$

Lemma 2.4. *If X_1, X_2 are two independent random variables with parameter λ, then the random variable $X_1 - X_2$ is a centered, symmetric random variable.*

Proof. The fact that this random variable is centered, follows directly from the linear property of expectations and the fact that X_1 and X_2 are identically distributed. We shall show that for any $t > 0$ $\mathbb{P}(X_1 - X_2 \le t) = \mathbb{P}(X_1 - X_2 \ge -t)$ using definition 2.11. Since X_1 and X_2 are identically distributed, the probability density function $f_{X_1}(s) = f_{X_2}(s)$. And thus the cumulative distribution function is also the same and can be exchanged.

$$\mathbb{P}(X_1 - X_2 \leq t) = \int_0^\infty \mathbb{P}(X_1 \leq t + s) \cdot f_{X_2}(s)ds = \int_0^\infty \mathbb{P}(X_2 \leq t + s) \cdot f_{X_1}(s)ds$$
$$= \mathbb{P}(X_2 \leq t + X_1) = \mathbb{P}(-X_2 \geq -t - X_1) = \mathbb{P}(X_1 - X_2 \geq -t).$$

□

2.6. Martingales

Another useful way of showing convergence in the almost sure sense, is by using martingales.

A sequence $\{\mathcal{F}_n : n \geq 0\}$ is a filtration on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ if it is an increasing sequence of σ-algebras. In this thesis we will only consider the natural filtration: $\mathcal{F}_n = \sigma(W_0, W_1, \dots, W_n)$, for a process $W = \{W_n : n \in \mathbb{N} \cup 0\}$. What this effectively means, is that for a process W at time n, $\mathcal{F}_n$ is just all the events that are determined by the first n time steps.

Definition 2.12. *A process $\{X_n\}$ is called a martingale relative to $(\{\mathcal{F}_n\}, \mathbb{P})$ if*

1. *X_n is $\mathcal{F}_n$-measurable for all n,*
2. *$\mathbb{E}(|X_n|) < \infty$ for all n,*
3. *$\mathbb{E}(X_n \mid \mathcal{F}_{n-1}) = X_{n-1}$ almost surely for all $n \geq 1$.*

We note that if the equality condition 3. from definition 2.12 is replaced by a $\leq$-sign or a $\geq$-sign, X is called a supermartingale or a submartingale, respectively.

Theorem 2.14 (Doob's forward convergence theorem)**.** *Suppose X_n is a supermartingale which satisfies $\sup_n E(|X_n|) < \infty$, then $X := \lim_{n\to\infty} X_n$ exists almost surely and is finite.*

Proof. A proof of this theorem can be found in 'Probability with Martingales' by Williams[22]. □

3
Finite Pólya Processes

3.1. Limiting behaviour under standard feedback function $f(x) = x$

In this section we will consider the standard feedback function $f(x) = x$. We introduce b_i^t to be the number of balls in bin i at time t. Similarly, we introduce x_i^t to be the fraction of balls of the total number of balls in bin i at time t. First, we shall proof that for any Finite Pólya Process this fraction of balls converges almost surely.

Theorem 3.1. *Under the standard feedback function, $f(x) = x$ for any Finite Pólya Process and any bin i, almost surely the following limit exists: $X_i = \lim_{t\to\infty} x_i^t$.*

Proof. We shall proof that x_i^t is a martingale by using definition 2.12 then, by theorem 2.14 x_i^t converges almost surely to some X_i. We will introduce ℓ to be the total number of balls in the bins at time t. We will introduce Y_n to be 1 whenever at time n a ball is put in bin i and 0 else. We shall proof that x_i^t is a martingale with respect to the filtration $\mathcal{F}_t = \sigma(Y_j : 1 \leq j \leq t)$.[1]

1. Note that we can write $x_i^t = \frac{1}{\ell}(b_i^0 + \sum_{j=1}^t Y_j)$, thus x_i^t is $\mathcal{F}_t$ measurable,
2. $\mathbb{E}(|x_i^t|) = \mathbb{E}(x_i^t) \leq 1 < \infty$,
3. $\mathbb{E}(x_i^{t+1} \mid \mathcal{F}_t) = \mathbb{E}(x_i^{t+1} \mid Y_t) = \mathbb{E}(x_i^{t+1} \mid Y_t = 1) \cdot \mathbb{P}(Y_t = 1) + \mathbb{E}(x_i^{t+1} \mid Y_t = 0) \cdot \mathbb{P}(Y_t = 0) = \frac{\ell \cdot x_i^t + 1}{\ell+1} \cdot x_i^t + \frac{\ell \cdot x_i^t}{\ell+1}(1 - x_i^t) = x_i^t \frac{\ell \cdot x_i^t}{\ell+1} - x_i^t \frac{\ell \cdot x_i^t}{\ell+1} + x_i^t \frac{1}{\ell+1} + \frac{\ell \cdot x_i^t}{\ell+1} = \frac{x_i^t \cdot (\ell+1)}{\ell+1} = x_i^t$.

Thus, we know that x_i^t converges almost surely to some X_i. □

A question that arises naturally when knowing that X_i exists, is: What will the value of X_i be? We shall see that rather than a deterministic value, the limit will be random. Meaning that if we were to run two different Finite Pólya Processes next to each other, even with the same initial configuration, they will have different limiting fractions. Then we ask ourselves: How will this limiting distribution be distributed? In the following sections 3.1.1 and 3.1.2 we shall find these limiting distributions for respectively 2-bin Finite Pólya Processes and k-bin Finite Pólya Processes.

3.1.1. 2-bin limit

In this section we will make use of theorem 2.13 to find the limiting distribution of 2-bin Finite Pólya Processes.

We will define Y_n to be 1 whenever ball n gets put into bin 1 and to be 0 whenever it gets put into bin 2.

Lemma 3.1. *The sequence $(Y_n)_{n\geq 1}$ is exchangeable.*

[1] $\mathcal{F}_0$ can be chosen to be the trivial σ-algebra.

Proof. We will show that for any arbitrary sequence of length n, $t_1, \ldots, t_n$ the probability $\mathbb{P}(Y_{t_1} = y_{t_1}, \ldots, Y_{t_n} = y_{t_n})$ is independent of the exact sequence of observations, as long as $\sum_{j=1}^{n} y_{t_j} = k$. In practice, this means that for any of the $\binom{n}{k}$ possible permutations of $t_1, \ldots, t_k$, one will have the same joint probability mass function.

We will consider the case where the first k elements y_{t_j} are equal to 1, that is, we place k balls back-to-back into bin 1, and afterwards we place $n-k$ balls in bin 2.

$$\mathbb{P}(Y_{t_1} = 1, \ldots, Y_{t_k} = 1, Y_{t_{k+1}} = 0 \ldots Y_{t_n} = 0), \tag{3.1}$$

$$= \frac{b_1^0}{b_1^0 + b_2^0} \cdot \frac{b_1^0 + 1}{b_1^0 + b_2^0 + 1} \cdots \frac{b_1^0 + k - 1}{b_1^0 + b_2^0 + k - 1} \cdot \frac{b_2^0}{b_1^0 + b_2^0 + k} \cdots \frac{b_2^0 + n - k - 1}{b_1^0 + b_2^0 + n - 1}, \tag{3.2}$$

$$= \frac{\prod_{i=0}^{k-1}(b_1^0 + i) \prod_{i=0}^{n-k-1}(b_2^0 + i)}{\prod_{i=0}^{n-1}(b_1^0 + b_2^0 + i)}. \tag{3.3}$$

Note how in equation 3.3 there is no dependence on the specific order of the observations.

We will argue that even if we were to switch two balls around, y_{t_j} and y_{t_i} with $i \neq j$, the joint density mass function would stay the same as in equation 3.3. Without loss of generality, we shall assume $y_{t_j} \neq y_{t_k}$[2] and we assume that $y_{t_j} = 1$ and thus $y_{t_k} = 0$. The only thing changing in 3.2, would be the exact order of the numerator in the product, but by the associative property, the joint probability would stay the same as in equation 3.3.

Any of the $\binom{n}{k}$ possible permutations can be made as a result of switching events in equation 3.1, and thus probability 3.1 is independent of the exact sequence of observations, knowing that $\sum_{j=1}^{n} y_{t_j} = k$. □

Knowing that $(Y_n)_{n \geq 1}$ is a binary, exchangeable sequence, one can find its mixing distribution function by using remark 2.1.

Theorem 3.2. *Given a 2-bin Finite Pólya Process under the standard feedback function, $f(x) = x$ and given initial condition (b_1^0, b_2^0), then the limiting fraction X_1 will be Beta distributed with parameters b_1^0 and b_2^0, that is:*

$$x_1^t \xrightarrow{D} \mathit{Beta}(b_1^0, b_2^0).$$

Proof. We shall introduce S_n to be the number of times a ball is put into bin 1 during the first n time steps. Using equation 3.3 and the fact that there are $\binom{n}{k}$ ways of putting k balls into bin 1 during n time steps, we get the following result:

$$\mathbb{P}(S_n = k) = \binom{n}{k} \frac{\prod_{i=0}^{k-1}(b_1^0 + i) \prod_{i=0}^{n-k-1}(b_2^0 + i)}{\prod_{i=0}^{n-1}(b_1^0 + b_2^0 + i)}. \tag{3.4}$$

[2]If they would have the same value, it would not change anything. e.g. if both values were 1, then you would add a ball to bin 1 at time t_j and time t_k, but by switching these events, you are not changing anything.

By using remark 2.1, in combination with equation 3.4, we can explicitly find the distribution $F(p)$.

$$\begin{aligned}
\int_0^1 p^k \cdot (1-p)^{n-k} dF(p) &= \frac{\prod_{i=0}^{k-1}(b_1^0 + i) \prod_{i=0}^{n-k-1}(b_2^0 + i)}{\prod_{i=0}^{n-1}(b_1^0 + b_2^0 + i)}, \\
&= \frac{[b_1^0]_k^+ [b_2^0]_{n-k}^+}{[b_1^0 + b_2^0]_n^+}, \\
&= \frac{\Gamma(b_1^0 + k)}{\Gamma(b_1^0)} \frac{\Gamma(b_2^0 + n - k)}{\Gamma(b_2^0)} \frac{\Gamma(b_1^0 + b_2^0)}{\Gamma(b_1^0 + b_2^0 + n)}, \\
&= \frac{\Gamma(b_1^0 + k) \cdot \Gamma(b_2^0 + n - k)}{\Gamma(b_1^0 + b_2^0 + n)} \cdot \frac{\Gamma(b_1^0 + b_2^0)}{\Gamma(b_1^0) \cdot \Gamma(b_2^0)}, \\
&= \mathrm{B}(b_1^0 + k, b_2^0 + n - k) \cdot \frac{1}{\mathrm{B}(b_1^0, b_2^0)}, \\
&= \frac{1}{\mathrm{B}(b_1^0, b_2^0)} \int_0^1 p^{b_1^0 + k - 1} (1-p)^{b_2^0 + n - k - 1} dp, \\
&= \int_0^1 p^k (1-p)^{n-k} \frac{p^{b_1^0 - 1} \cdot (1-p)^{b_2^0 - 1}}{\mathrm{B}(b_1^0, b_2^0)} dp.
\end{aligned}$$

Thus, one can conclude that $f(p) = \frac{1}{\mathrm{B}(b_1^0, b_2^0)} p^{b_1^0 - 1} \cdot (1-p)^{b_2^0 - 1} \mathbb{1}_{p \in (0,1)}$. Using definition 2.7, we can say that the limiting fraction of X_1 will be Beta distributed with parameters b_1^0 and b_2^0. □

Remark 3.1. *The limiting distribution of x_2^t is exactly Beta(b_2^0, b_1^0). By the same argument as in the proof theorem 3.2, but with the roles of bin 1 and bin 2 switched around.*

Remark 3.2. *We can find the limiting distribution of bin 2, X_2 by using the limiting distribution X_1 by looking at $X_2 = 1 - X_1$, since these are fractions, their sum should be 1.*

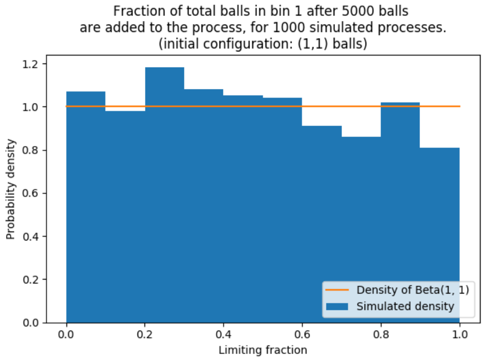
Validation

In this section we shall validate and reflect onto our theoretical results regarding 2-bin Finite Pólya Processes by means of simulations.

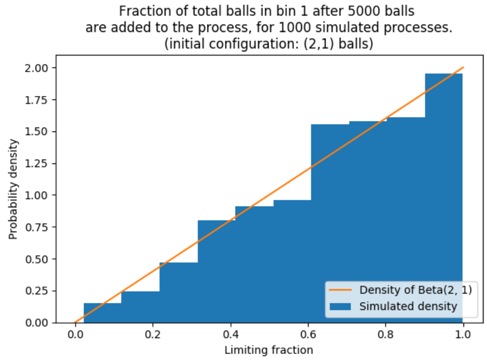
We shall start some Finite Pólya Process with a given initial number of balls, for example one ball in both bins, then we shall add 5000 balls to this process with respect to the feedback function. After doing this, we shall note the empirical fraction of balls in bin 1. We run 1000 processes and using the obtained empirical fractions, we make a histogram of these empirical fractions. Following this, we compare this resulting empirical distribution to the limiting distribution according to theorem 3.2. The results of these simulations have been put in figure 3.1.

It seems reasonable to assume that theorem 3.2 holds, since the simulated density is close to the theoretical density. Some very interesting observations can be made using figure 3.1. Note that if we initially start with one ball in both bins, we get a uniformly distributed limit. We remark that starting a process with initial configuration $(1, 1)$ and adding the first ball to $(2, 1)$, is the same as starting a process with initial configuration $(2, 1)$. With this idea, 3.1a and 3.1b can be seen as how the limiting distribution of the process changes if one knows that the first ball gets added to the first bin. The first couple of balls have a lot of influence on the limiting distribution of the process. Then again, it might still occur that even though one has received the first ball, the other bin will still gain the larger limiting fraction. Another interesting result is how centered the limiting distributions are if we start with five balls in both bins. Lastly we note how just having eight balls in one bin makes it very unlikely for the other bin to gain the limiting fraction.

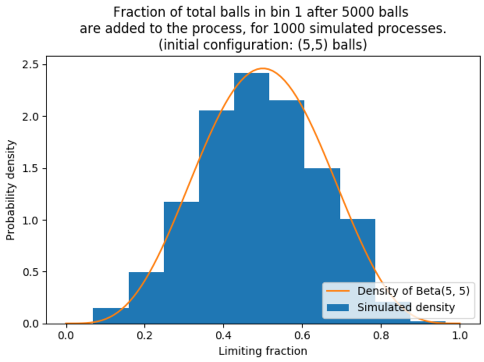
The fact that the first couple of balls being added holds such a large influence on the limiting distribution, is called path-dependence. It is one of the reasons to why Pólya Processes are interesting for modelling market phenomenons. One can see both bins as different



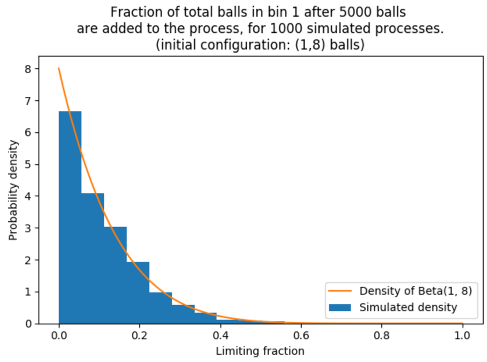
(a) Initial configuration $(1, 1)$



(b) Initial configuration $(2, 1)$



(c) Initial configuration $(5, 5)$



(d) Initial configuration $(1, 8)$

Figure 3.1: Plots comparing the results of theorem 3.2 to simulated 2-bin Finite Pólya Processes. Every plot has a different initial configuration. The Python code for these plots can be found in the Appendix.

players in some market. Since we have a standard feedback function, we assume that the products sold by the players are neutral products, that is, the consumer does not find one of the products inherently more attractive than the other. Note that whenever a player has sold more products, it is also more likely that the next customer buys their product. Thus, attracting early customers is vital for a player to gain a larger limiting market share, because of the very path-dependent nature of this process. Using Finite Pólya Processes we can model exactly how important these earlier customers are.

3.1.2. k-bin limit

In this section we will generalize the results of theorem 3.2 for an arbitrary, but finite, number of bins k.

Theorem 3.3. *Given a k-bin Finite Pólya Process under the standard feedback function, $f(x) = x$ and given initial condition $(b_1^0, \ldots, b_2^0)$, then the limiting fraction $X_1, \ldots, X_k$ will be Dirichlet distributed with parameters $b_1^0, \ldots, b_k^0$, that is:*

$$x_1^t, x_2^t, \ldots, x_k^t \xrightarrow{D} \text{Dirichlet}(b_1^0, \ldots, b_k^0).$$

Proof. A proof of this can be found in theorem 1 of the paper of Blackwell & MacQueen[6]. It can be proven in a similar way as we have done for the 2-bin case, by a De Finetti argument, note that if we were to introduce Y_t^j being the random variable that is 1 whenever bin j gets ball t, we can show that the sequence balls being added to different bins will be exchangeable. □

Note that a realization from the Dirichlet distribution contains $k - 1$ values, but that we need k fractions. This last value can be derived by using a similar argument as in remark 3.2

. We note that theorem 3.3 is a very logical conclusion from theorem 3.2, since the Dirichlet distribution is the multivariate generalization of the Beta distribution.

Validation

In this section we shall validate and reflect onto our theoretical results for k-bin Finite Pólya Process by means of simulations.

We note that for $k = 2$ theorem 3.3 just gives us theorem 3.2, which has already been validated. We shall focus our attention to the $k = 3$ case, since we van visualize this case. Note that the limiting fractions, X_1, X_2 and X_3 exist on the 3-dimensional standard 2-simplex, by nature of them being fractions, we know that $X_1 + X_2 + X_3 = 1$. What this means is, that any configuration of feasible limits, is some point on the simplex shown in figure 3.2. The Dirichlet distribution for $k = 3$ just assigns probabilities to points on this simplex, based on the Dirichlet parameter $\theta = (\theta_1, \theta_2, \theta_3)$.

We shall compare the empirical limiting fractions of a 3-bin Finite Pólya Process with initial configurations $(\theta_1, \theta_2, \theta_3)$ to a Dirichlet distribution with these parameters by means of looking at the simplex.

We shall start some Finite Pólya Process with a given initial number of balls, for example one ball in all three bins, then we shall add 5000 balls to this process according to the standard feedback function. After doing this, we shall note the empirical fraction of balls in all bins and mark this point on the simplex in 3.3. After having marked 500 points, we shall compare the result to the theoretical limiting distribution. Note that the Dirichlet distribution shall be plotted as a contour plot, where a more yellow hue corresponds with a greater probability density in that region. This simulation has been ran for different initial configurations, as can be seen in figure 3.3. It seems reasonable that theorem 3.3 holds based on these simulations. Once again we note the path-dependence. If every bin has just one ball, any point of the simplex will be equally likely to occur. However, if all bins have five balls, then it is much more likely that all bins will have a fraction around the same size. Note that once again one bin starting with a greater number of balls implies that this bin will have a larger limiting fraction.

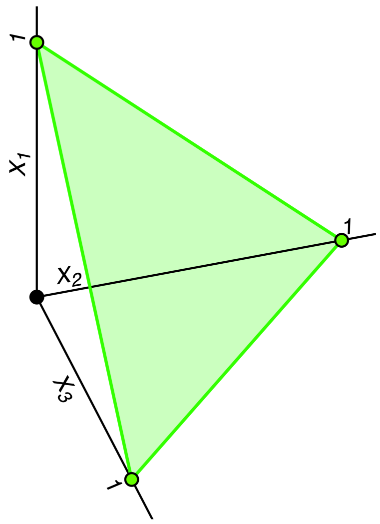


Figure 3.2: A 3-dimensional standard 2-simplex.

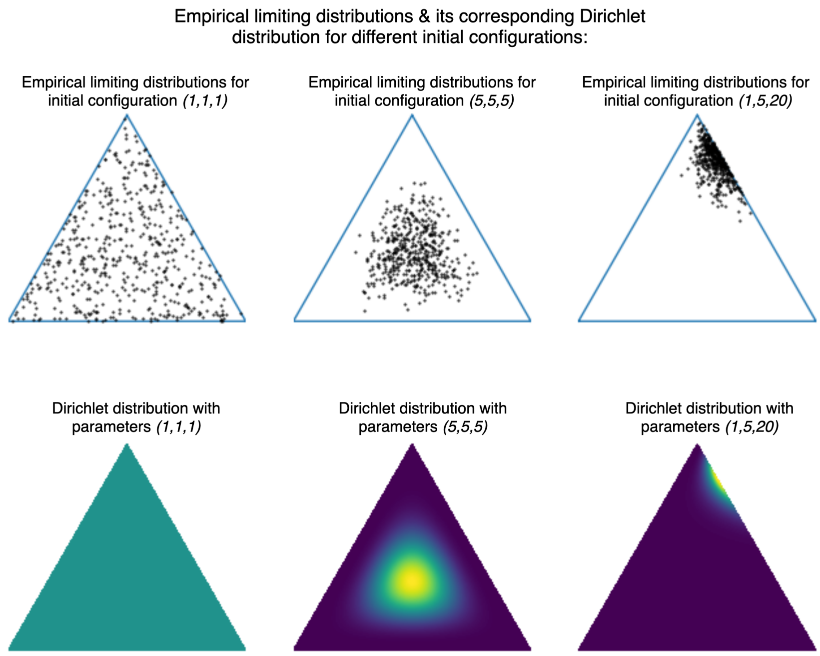


Figure 3.3: Plots comparing the results of theorem 3.3 to simulated 3-bin Finite Pólya Processes. We look perpendicularly at the simplex as described in figure 3.2. The left corner corresponds to bin 1, the right corner to bin 2 and the upper corner to bin 3. The Python code for these plots can be found in the Appendix.

3.2. Limiting behaviour under general feedback function $f(x)$

Until now, only the standard feedback function, $f(x) = x$ has been considered. However, any function of the form $f : \mathbb{N} \to (0, \infty)$ is a valid feedback function and in this section we shall discuss some characteristics regarding general feedback functions.

A well-known and well-researched class of feedback functions, is $f(x) = x^\gamma$ for some $\gamma \in \mathbb{R}$. It is known that for $\gamma > 1$ there is some bin j that will almost surely receive all the balls. And for $\gamma < 1$ all bins will almost surely have the same number of balls[7].

In this section we shall look at how different feedback functions behave in their almost sure limit. Questions that we will be concerned with, are: *For what kind of feedback functions will one bin receive all the balls almost surely? When will all the bins be equal in size almost surely? We know that for $f(x) = x$ neither of these scenarios happen, what makes this case so different, that this does not happen?* The results of this section are largely inspired by the work of Oliveira[19], [18].

We shall formally define two events that will be essential for our understanding of the limiting behaviour of Finite Pólya Processes under general feedback functions. For this purpose, we introduce i_m, which refers to bin i that gets the additional ball at time m.

Definition 3.1. *Mon_i refers to the event that bin i holds a monopoly. This means that eventually, bin i will receive all next balls. Thus, there is a $M \in \mathbb{N}$ such that after time M, all additional balls will be added to bin i.*

$$\text{Mon}_i = \{\exists M \in \mathbb{N} : \forall n \geq M : i_n = i\}.$$

Definition 3.2. *Lead_i refers to the event that bin i will eventually be the leader. This means that eventually bin i will always be larger than the other bins. Thus, there is a $M \in \mathbb{N}$ such that after time M, no bin will be larger than bin i.*

$$\text{Lead}_i = \{\exists M \in \mathbb{N} : \forall n \geq M : \forall j \in \{1, \dots, k\} \setminus \{i\} : b_i^n > b_j^n\}.$$

Remark 3.3. *Whenever there is a monopoly, there is an eventual leadership. The reverse does not hold. An example for this, is a Finite Pólya Process with 2 bins under the standard feedback function.*

To reach our desired conclusions, we will need to introduce an exponential embedding, which we shall do in the following section:

3.2.1. Proposed Embedding

We claim that there is an exponential embedding possible for the Finite Pólya Process and we shall show this embedding explicitly. *What we gain from introducing this exponential embedding, is that we can look at the growth of bins as separate, independent processes.* Before that is possible, however, one needs to consider what *makes* a Finite Pólya Process a Finite Pólya Process. We will start by denoting that one needs to know the feedback function $f : \mathbb{N} \to (0, \infty)$, the number of bins, k[3] and the initial configuration of balls in every bin, $(b_1^0, \dots, b_k^0)$.

A Finite Pólya Process, given feedback function f and some initial configuration $(b_1^0, \dots, b_k^0)$ is a process for which at every time t, there is exactly one bin i that gets an additional ball. That is, for some $j \in K$[4] $b_j^{t+1} = b_j^t$ and for all other $i \in K, i \neq j : b_i^{t+1} = b_i^t$.
And the probability that this bin, which gets the next ball, is exactly bin i, is related to the number of balls in bins at time n in the following way:

$$\mathbb{P}(b_i^{n+1} = b_i^n + 1 \mid b_1^n, \dots, b_k^n)^5 = \frac{f(b_i^n)}{\sum_{j=1}^k f(b_j^n)}. \tag{3.5}$$

[3] This value is actually not needed, since it can be retrieved from the initial configuration
[4] $K = \{1, \dots, k\}$
[5] We shall omit this conditioning statement in the rest of our work, for the sake of readability, but we note that it will always be assumed when we are working with these types of probabilities.

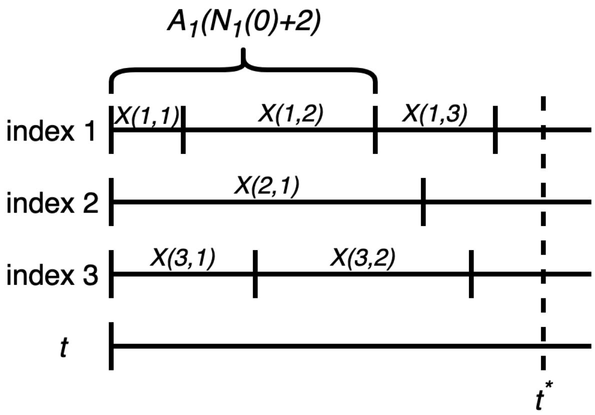


Figure 3.4: An example of some scenario of our proposed exponential embedding with $k = 3$. Note that by the definition $N_1(t^*) = N_1(0) + 3$, $N_2(t^*) = N_2(0) + 1$ and $N_3(t^*) = N_3(0) + 2$. Also note how $A_1(N_1(0) + 2) = X(1,1) + X(1,2)$.

Note that thus, a Finite Pólya Process can thus be seen as a sequence $(b_1^n, ..., b_k^n)_{n \in \mathbb{Z}^+}$ in which for every next time step n, there is one element b_i^n that is 1 greater in the next element of the sequence whereas the other elements stay the same size. The probability for which element gets selected to be the one that increases at a certain time step, is in accordance to equation 3.5.

We shall show that our proposed Exponential Embedding, given a feedback function and an initial configuration, gives rise to a process that is equivalent to the Finite Pólya Process.

For this purpose, we will require a set of independent random variables, $\{X(i,j) : i \in K, j \in \mathbb{N}\}$ such that for any element of that set, $X(i,j) \stackrel{d}{=} \text{Exp}(f(j))$. For any $i \in K$ and $t \geq 0$, we define:

$$N_i(t) = \sup_{n \in \mathbb{N}} \Big\{ \sum_{j=b_i^0}^{n-1} X(i,j) \leq t \Big\}. \tag{3.6}$$

Note that $X(i,j)$ is a positive, continuous random variable. Meaning that it is not possible for any $X(i,j)$ to be smaller or equal to 0, meaning that for $N_i(0)$, one will need to find the largest n, such that $\sum_{j=b_i^0}^{n-1} X(i,j) \leq 0$. Remembering that $\sum_i^k (...) = 0$ if $k > i$, we know that $N_i(0) = b_i^0$ for any $i \in K$.

We can also introduce the arrival time $A_i(k)$ for any process, which is the smallest value of t such that $N_i(t) = k$. A visualization of how one should interpret $A_i(k)$ and $N_i(t)$ can be found in figure 3.4.

The set of all arrival times, for all bins, can be put into a set A. Then, one can order these elements by their times, whilst taking into account the corresponding index of that specific arrival time. Then one can note A as a sequence $T_1 < T_2 < ....$ A visualization of how one should interpret this sequence can be found in figure 3.5. We claim that the process $\Big(N_1(T_m), N_2(T_m), ..., N_k(T_m)\Big)_{m \in \mathbb{Z}^+}$[6] is equivalent to a Finite Pólya Process. We have already shown that $N_i(0) = b_i^0$. Note that given any time T_t, at the next time T_{t+1}, there is exactly one indexed process that will increase by 1. Then the question arises, given all of these values at some time T_t, what is the probability that process N_i is that one process that will increase by 1?

[6]We define $T_0 = 0$.

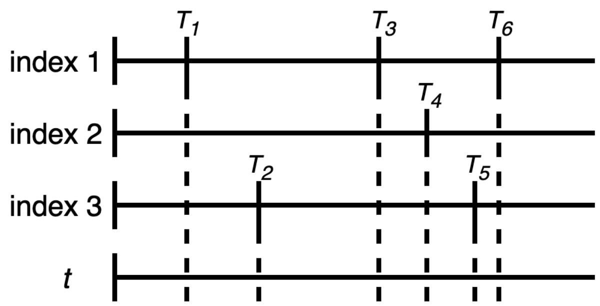


Figure 3.5: An example of some scenario of our proposed exponential embedding with $k = 3$, where one sees that by ordering the arrival times for all indices, one can get a global order of when the events happen.

At time T_t, the processes have exactly size $N_1(T_t), \dots, N_k(T_t)$. Which of these k processes will be the one that increases? We know that this question can be restated as: Which of the k processes will be the one with the first, next arrival? Using the lack of memory property, we know that for any process i, at time T_t, the time until the next arrival is exactly equal to $X(i, N_i(T_t))$. Thus the question becomes, for which $i \in K$ is $X(i, N_i(T_t))$ the smallest? We use the minimum property to derive:

$$\mathbb{P}\Big(X(i, N_i(T_t)) = \min_{j \in K} X(j, N_j(T_t))\Big) = \frac{f(N_i(T_t))}{\sum_{j=1}^{k} f(N_j(T_t))}. \tag{3.7}$$

It is no coincidence that equation 3.5 is similar to equation 3.7.

We shall recap our findings, $\big(N_1(T_m), N_2(T_m), \dots, N_k(T_m)\big)_{m \in \mathbb{Z}^+}$, our proposed exponential embedding, is a process that has initial configuration $(N_1(0), \dots, N_k(0)) = (b_1^0, \dots, b_k^0)$. For every next element in the sequence, there is exactly one process that increases with exactly 1, whereas the other processes stay the same. Which process will be the process that increases, is based on the current size of all processes. This relation is explicitly noted in equation 3.7. This exactly aligns with the definition of a Finite Pólya Process we have given earlier. Thus, the Finite Pólya Process and our proposed exponential embedding, are equivalent processes.

3.2.2. The three regimes

Now that we have our exponential embedding, we can start working towards our desired result. We shall first consider 2-bin Finite Pólya Processes.

Theorem 3.4. *For a 2-Bin Finite Pólya Process with feedback function $f(x)$ such that $\sum_{x=1}^{\infty} \frac{1}{f(x)} < \infty$ almost surely, one bin will dominate. $\mathbb{P}(\bigcup_{i=1}^{2} Mon_i) = 1 (= \mathbb{P}(\bigcup_{i=1}^{2} Lead_i))$.*

Proof. For this proof, we shall use the proposed exponential embedding as explained in section 3.2.1. Consider $F_1 = \sum_{j=b_1^0}^{\infty} X(1, j)$[7], which can be interpreted as the arrival time until bin 1 has infinitely many balls.

$$\mathbb{E}\Big[\sum_{j=b_1^0}^{\infty} X(1, j)\Big] = \mathbb{E}(F_1) = \sum_{j=b_1^0}^{\infty} \frac{1}{f(j)}. \tag{3.8}$$

This follows from the linearity of the expectation in combination with the fact that any $X(1, j)$ is distributed exponentially with parameter $f(j)$. By assumption, we know therefore that the expectation of the time until bin 1 has infinitely many balls, is finite. This also holds for bin 2. Thus, by the strong law of large numbers, $F_1 < \infty$ and $F_2 < \infty$ almost surely. Since these series are independent and they are continuous random variables, the probability that $F_1 = F_2$, is 0.

[7] We note that $F_1 = A(1, \infty)$ with $A(i, j)$ as has been introduced in 3.2.1

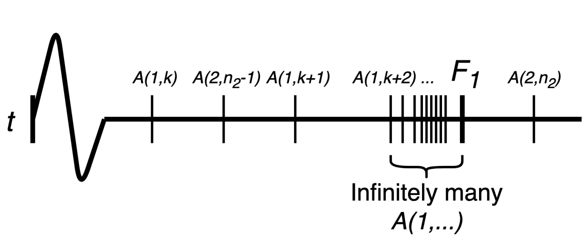


Figure 3.6: A visualization of how the existence of F_1 that is smaller than F_2 implies that there will be infinitely many arrival times that happen before $A(2, n_2)$ for some $n_2 \in \mathbb{N}$.

With probability 1, either $F_1 < F_2$ or $F_1 > F_2$. We shall assume the former. We know that for any arrival time $A(1, j)$, this arrival time is smaller than the arrival time of infinite balls, F_1. Other than that, we know that $F_2 > F_1$, thus for some $n_2 \in \mathbb{N}$, the arrival time $A(2, n_2)$ has to be larger than F_1, otherwise, F_2 can't be larger than F_1.

$$\exists n_2 \in \mathbb{N} : \forall n \geq n_2 : A(1, n) < F_1 < A(2, n_2).$$

Our exponential embedding was defined as $(N_1(T_m), N_2(T_m))_{m\in\mathbb{Z}^+}$. We claim that for the process as described above, after some point in this process, infinitely many times the next arrival time will be from the first bin, meaning that infinitely many times N_1 increases. Thus the next arrival time for bin 2, will never be represented in our sequence $\{N_1(T_m), N_2(T_m)\}_{m\in\mathbb{Z}^+}$. Figure 3.6 has been provided to give the reader an intuition for why this happens. Note that bin 2 will thus never have more than n_2 elements, and that eventually bin 1 will get all the balls. A same argument could be used if we had assumed $F_1 < F_2$, the roles of the bins would just have been reversed. Since either $F_1 > F_2$ or $F_1 < F_2$ with probability 1, thus with probability 1 there will be one bin that after some time $n_2 \in \mathbb{N}$, for all time $n > n_2$ that bin will receive the next ball, which means that $\mathbb{P}(\cup_{i=1}^{2} \text{Mon}_i) = 1$. □

Remark 3.4. *Note thus that there are quite some arrivals in our embedding that do not actually happen in the discrete process, these events are called ghost events. This is very similar to the continuation of a Galton-Watson process beyond its extinction time[1].*

Lemma 3.2. *If $\sum_{x=1}^{\infty} \frac{1}{f(x)} = \infty$, then:*

1. *For any $n \in \mathbb{N}$, $\sum_{x=n}^{\infty} \frac{1}{f(x)} = \infty$.*

2. *For any $n \in \mathbb{N}$, $\sum_{x=1}^{\infty} \frac{1}{n+f(x)} = \infty$.*

Proof. We shall first proof 1. Let n be an arbitrary element of $\mathbb{N}$.

$$\infty = \sum_{x=1}^{\infty} \frac{1}{f(x)} = \sum_{x=1}^{n-1} \frac{1}{f(x)} + \sum_{x=n}^{\infty} \frac{1}{f(x)} \leq (n-1) \max_{1\leq j\leq n-1} \frac{1}{f(j)} + \sum_{x=n}^{\infty} \frac{1}{f(x)},$$

$$= M + \sum_{x=n}^{\infty} \frac{1}{f(x)} \quad (M \in (0, \infty)).$$

This implies that $\sum_{x=n}^{\infty} \frac{1}{f(x)} = \infty$.

Now 2 shall be considered. Let $n \in \mathbb{N}$ be arbitrarily chosen and assume that $\sum_{x=1}^{\infty} \frac{1}{n+f(x)} < \infty$. We will show that this leads to a contradiction.
If $\sum_{x=1}^{\infty} \frac{1}{n+f(x)} < \infty$, then we know that $\lim_{x\to\infty} \frac{1}{n+f(x)} = 0$. That implies that $\lim_{x\to\infty} n + f(x) = \infty$. Note that n is a fixed element of $\mathbb{N}$, thus if we disregard it, the limit should still be infinite. $\lim_{x\to\infty} f(x) = \infty$.

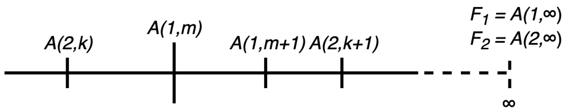


Figure 3.7: A visualization of how F_1 and F_2 being infinite warrants that for any $m \in \mathbb{N}$, the arrival time $A(1,m)$ will always have an arrival of the form $A(2,k+1)$ that comes after it for some $k \in \mathbb{N}$.

This means, that for any $m \in \mathbb{N}$, after some point N, for all $n^* \geq N$: $f(n^*) \geq m$. This also holds for our chosen n.

$$\infty > \sum_{x=N}^{\infty} \frac{1}{n+f(x)} \geq \sum_{x=N}^{\infty} \frac{1}{f(x)+f(x)} = \frac{1}{2} \sum_{x=N}^{\infty} \frac{1}{f(x)} = \infty.$$

This gives us our desired contradiction, and thus we know that $\sum_{x=1}^{\infty} \frac{1}{n+f(x)} = \infty$ for any $n \in \mathbb{N}$, since n was arbitrarily chosen. □

Theorem 3.5. *For a 2-bin Finite Pólya Process with feedback function $f(x)$ such that $\sum_{x=1}^{\infty} \frac{1}{f(x)} = \infty$, the occurrence of a monopoly has probability 0. $\mathbb{P}(\bigcup_{i=1}^{2} Mon_i) = 0$.*

Proof. We shall show that if $\sum_{x=1}^{\infty} \frac{1}{f(x)} = \infty$, then $F_1 = \sum_{j=b_1^0}^{\infty} X(1,j)$ will almost surely be infinite. Similarly F_2 will be infinite. We shall use the continuous mapping theorem 2.7, in combination with the function $g(x) = e^{-x}$.

For the sake of convenience, we introduce $Y_m = \sum_{j=b_1^0}^{m} X(1,j)$ and $Y = \lim_{m\to\infty} Y_m = \sum_{j=b_1^0}^{\infty} X(1,j)$. We shall show that $g(Y_m)$ converges almost surely to 0. We shall do this by means of the strong law of large numbers ,theorem 2.5. Thus, we need to show that $\mathbb{E}(g(Y)) = 0$. Note that Y is a positive random variable and $g(x)$ is a positive function, thus the expectation $\mathbb{E}(g(Y))$ is lower bounded by 0.

$$\begin{aligned} 0 \leq \mathbb{E}(Y) = \mathbb{E}(e^{-\sum_{j=b_1^0}^{\infty} X(1,j)}) &= \mathbb{E}(\prod_{j=b_1^0}^{\infty} e^{-X(1,j)}) = \prod_{j=b_1^0}^{\infty} \mathbb{E}(e^{-X(1j)}), \\ &= \prod_{j=b_1^0}^{\infty} \int_0^{\infty} e^{-t} \cdot f(j) e^{-tf(j)} dt = \prod_{j=b_1^0}^{\infty} f(j) \int_0^{\infty} e^{-t(1+f(j))} dt, \\ &= \prod_{j=b_1^0}^{\infty} \frac{f(j)}{1+f(j)} \int_0^{\infty} (1+f(j)) \cdot e^{-t(1+f(j))} dt = \prod_{j=b_1^0}^{\infty} \frac{f(j)}{1+f(j)} \cdot 1, \\ &= \prod_{j=b_1^0}^{\infty} 1 - \frac{1}{1+f(j)} \leq \prod_{j=b_1^0}^{\infty} e^{-\frac{1}{1+f(j)}} = e^{-\sum_{j=b_1^0}^{\infty} \frac{1}{1+f(j)}} = e^{-\infty} = 0. \end{aligned}$$

Thus, $Y_m = \sum_{j=b_1^0}^{m} X(1,j)$ converges almost surely to ∞. This implies therefore that F_1 and F_2 are almost surely infinite. Note that for any arrival time of ball m in bin 1, $A(1,m)$, it is not possible that there have been infinitely many arrivals prior to this arrival, since there can only have been infinitely many arrivals if either $F_1 < \infty$ or $F_2 < \infty$. That means that at time $A(1,m)$ bin 2 has also had finitely many arrivals, this number of arrivals will be noted as k. Thus, we know that before F_1 happens, $A(2,k+1)$ will happen. Since m was arbitrary, we note that there is no arrival for bin 1 such that all following arrivals will also be for bin 1. The same argument can be used in the same way for bin 2. A visualization of this exact scenario has been provided in figure 3.7. This means that $\mathbb{P}(\bigcup_{i=1}^{2} \text{Mon}_i) = 0$. □

Theorem 3.6. *For a 2-bin Finite Pólya Process with feedback function $f(x)$ such that $\sum_{x=1}^{\infty} \frac{1}{f(x)^2} < \infty$, there will almost surely be an eventual leader, that is, $\mathbb{P}(\bigcup_{i=1}^{2} Lead_i) = 1$.*

Proof. For the sake of notation, we introduce $x > y$. We also introduce $Z_j = X(1,j) - X(2,j)$. Then the series $\sum_{j=x}^{\infty} Z_j$ is a series of independent random variables. Note that for any j, Z_j is centered and symmetric by lemma 2.4. By our assumption we know that:

$$\sum_{j=x}^{\infty} \text{Var}(Z_j) = \sum_{j=x}^{\infty} \text{Var}(X(1,j)) - \text{Var}(X(2,j)) = \sum_{j=x}^{\infty} \frac{1}{f(j)^2} + \frac{1}{f(j)^2} = \sum_{j=x}^{\infty} \frac{2}{f(j)^2} < \infty.$$

Using Kolmogorov's Three Series Theorem 2.6, we shall show that $\sum_{j=x}^{m} Z_j$ converges almost surely to $\sum_{j=x}^{\infty} Z_j$[8]. Let $A > 0$ be arbitrary:

1. To proof this, Chebyshev's inequality (Theorem 2.9) will be used.

$$\sum_{j=1}^{\infty} \mathbb{P}(|Z_j| \geq A) = \sum_{j=1}^{\infty} \mathbb{P}(|Z_j - \mathbb{E}(Z_j)| \geq A) \leq \sum_{j=1}^{\infty} \frac{\text{Var}(Z_j)}{A^2} < \infty.$$

2. Since Z_j is symmetric and centered around 0 for any j, we know that $\mathbb{E}(Z_j \mathbb{1}_{|Z_j| \leq A}) = 0$, thus:

$$\sum_{j=1}^{\infty} \mathbb{E}(Z_j \mathbb{1}_{|Z_j| \leq A}) = \sum_{j=1}^{\infty} 0 = 0 < \infty.$$

3. We note that for any j: $\text{Var}(Z_j \mathbb{1}_{|Z_j| \leq A}) \leq \text{Var}(Z_j)$. If one were to consider a larger range of possible values a random variable can take, then the variance can only grow larger.

$$\sum_{j=1}^{\infty} \text{Var}(Z_j \mathbb{1}_{|Z_j| \leq A}) \leq \sum_{j=1}^{\infty} \text{Var}(Z_j) < \infty.$$

Thus we know that $\sum_{j=x}^{\infty} Z_j$ converges almost surely. Disregarding point masses, we can say that either of the following events happens almost surely:

$$\sum_{j=x}^{\infty} Z_j - \sum_{j=y}^{x-1} X(2,j) < 0 \text{ or } \sum_{j=x}^{\infty} Z_j - \sum_{j=y}^{x-1} X(2,j) > 0. \tag{3.9}$$

Assuming the former, we can get that:

$$\sum_{j=x}^{\infty} Z_j - \sum_{j=y}^{x-1} X(2,j) < 0,$$

$$\sum_{j=x}^{\infty} (X(1,j) - X(2,j)) - \sum_{j=y}^{x-1} X(2,j) < 0,$$

$$\sum_{j=x}^{\infty} X(1,j) - \sum_{j=x}^{\infty} X(2,j) - \sum_{j=y}^{x-1} X(2,j) < 0,$$

$$\sum_{j=x}^{\infty} X(1,j) - \sum_{j=y}^{\infty} X(2,j) < 0,$$

$$\sum_{j=x}^{\infty} X(1,j) < \sum_{j=y}^{\infty} X(2,j).$$

[8]We note that for convergence, you can disregard a finite number of elements in the series

This can only happen if for all large enough M the following equation holds:

$$\sum_{j=x}^{M-1} X(1,j) < \sum_{j=y}^{M-1} X(2,j).$$

Meaning, that for *all* large enough M, bin 1 reaches size M before bin 2. Thus bin 1 eventually becomes the leader. If we had assumed the other case from equation 3.9, then using the same logic, but the roles of bin 1 and bin 2 reversed, we can conclude that bin 2 eventually becomes the leader. We note that equation 3.9 hold almost surely and thus there almost surely will be an eventual leader. □

Remark 3.5. *In theorem 3.6, we have not put a restriction on* $\sum_{j=1}^{\infty} \frac{1}{f(j)}$. *It is remarkable that there are Finite Pólya Processes that are not monopolies, but that do have eventual leaders. An example of this, are Finite Pólya Processes under the standard feedback function* $f(x) = x$.

Theorem 3.7. *For a* 2*-bin Finite Pólya Process with feedback function* $f(x)$ *such that* $\sum_{x=1}^{\infty} \frac{1}{f(x)^2} = \infty$, *there will almost surely be no eventual leader. That is,* $\mathbb{P}(\bigcup_{i=1}^{2} \text{Lead}_i) = 0$.

Proof. If $\sum_{x=1}^{\infty} \frac{1}{f(x)^2} = \infty$, one can use Donsker's Invariance Principle and characeteristics of Brownian Motions, as has been done by Oliveira[18], to show that this means that both $A(1,m) < A(2,m)$ and $A(1,m) > A(2,m)$ occur infinitely often with probability 1. Meaning that there are infinitely many m such that bin 1 reaches level m before bin 2. The same holds for bin 2, thus neither bin will be the eventual leader, that is $\mathbb{P}(\bigcup_{i=1}^{2} \text{Lead}_i) = 0$. □

Theorem 3.8. *Let* $k \in \mathbb{N}, k \geq 2$. *For a* k*-bin Finite Pólya Process with feedback function* $f : \mathbb{N} \to (0, \infty)$, *then there are three mutually exclusive regimes that can occur.*

1. *If* $\sum_{n\geq 1} \frac{1}{f(n)} < \infty$, *then* $\mathbb{P}(\bigcup_{i=1}^{k} \text{Mon}_i) = \mathbb{P}(\bigcup_{i=1}^{k} \text{Lead}_i) = 1$ *This is called the monopolistic regime.*

2. *If* $\sum_{n\geq 1} \frac{1}{f(n)} = \infty$, *but* $\sum_{n\geq 1} \frac{1}{f(n)^2} < \infty$, *then* $\mathbb{P}(\bigcup_{i=1}^{k} \text{Mon}_i) = 0$, *but* $\mathbb{P}(\bigcup_{i=1}^{k} \text{Lead}_i) = 1$. *This is called the eventual leadership regime.*

3. *If* $\sum_{n\geq 1} \frac{1}{f(n)^2} = \infty$, *then* $\mathbb{P}(\bigcup_{i=1}^{k} \text{Mon}_i) = \mathbb{P}(\bigcup_{i=1}^{k} \text{Lead}_i) = 0$. *This is called the almost-balanced regime.*

This holds for any initial configuration of the process. A visual representation of this theorem has been provided in figure 3.8.

Proof. We shall show that our results in theorems 3.4, 3.6 and 3.7 hold for any $k \geq 2$,

First, we consider the case of $\sum_{n\geq 1} \frac{1}{f(n)} < \infty$. If we were to look at just pairs of bins, bin j and bin k, we would know by theorem 3.4 that there is one bin that will have the lowest expected time until infinitely many balls arrive, either $F_i < F_j$ or $F_i > F_j$. Knowing this, for any pair of bins, will give us that there is some index m such that $F_m \leq F_i$ for all $i \in \{1, ..., k\}$. This will be the bin that will have the monopoly. Since all of the arrival times coming after time F_m will be ghost events as described in remark 3.4, thus the probability in a k-bin Finite Pólya Process of some bin having a monopoly, is 1, given $\sum_{n\geq 1} \frac{1}{f(n)} < \infty$.

A similar argument can be used to show that if $\sum_{n\geq 1} \frac{1}{f(n)} = \infty$ and $\sum_{n\geq 1} \frac{1}{f(n)^2} < \infty$, then there will be some bin that will be the eventual leader. Note that once again one can consider pairs of bins i and j, and by theorem 3.6 we know that almost surely one of these two bins will gain eventual leadership over the other. Using this fact, we can find exactly one bin m such that there is no other bin that will gain eventual leadership over bin m. This means that for any other bin l, bin m will have eventual leadership over bin l, thus the probability in a k-bin Finite Pólya Process of some bin having eventual leadership, is 1, given $\sum_{n\geq 1} \frac{1}{f(n)} = \infty$ and $\sum_{n\geq 1} \frac{1}{f(n)^2} < \infty$.

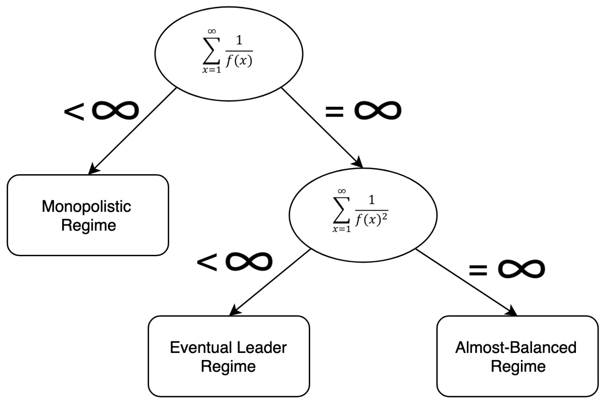


Figure 3.8: The different kind of regimes of a Finite Pólya Process can fall into depending on the feedback function $f(x) : \mathbb{N} \to (0, \infty)$.

Table 3.1: Table with examples of feedback functions $f(x)$ under their corresponding regime by theorem 3.8.

Monopolistic Regime	Eventual Leader Regime	Almost Balanced Regime
x^a when $a > 1$	x^a when $\frac{1}{2} < a \leq 1$	x^a when $a \leq \frac{1}{2}$
a^x when $a > 1$	$bx + a$ where $a, b \in \mathbb{R}^+$	a when $a \in \mathbb{R}^+$

If $\sum_{n\geq 1} \frac{1}{f(n)^2} = \infty$, then we know by theorem 3.7 that if we were to consider pairs of bins bin i and bin j, infinitely often bin i bins will reach some size m earlier than bin j and vice-versa. Fix i arbitrarily, if one considers bin i and bin j with j randomly in $\{1, \ldots, k\}$ we see that bin i is infinitely often larger than any other bin, but also infinitely often smaller than any other bin, thus it cannot ever be the leader. Since i was chosen randomly, no bin can ever gain eventual leadership, thus the probability of having eventual leadership when $\sum_{n\geq 1} \frac{1}{f(n)^2} = \infty$, is 0. □

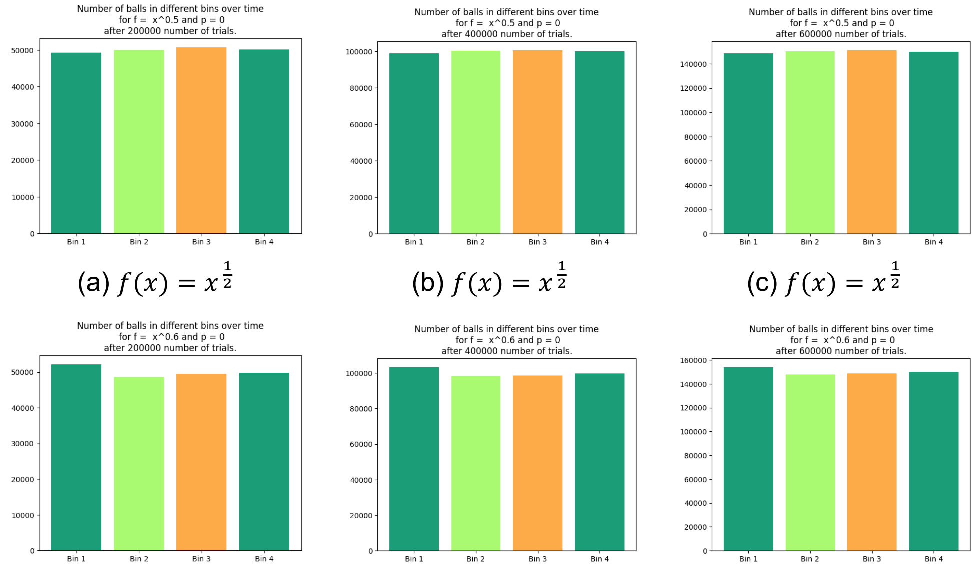
3.2.3. Implications of theorem 3.8

We note that theorem 3.8 gives us a very strong result, which warrants some discussion. Using theorem 3.8, we can divide different groups of feedback functions under the regime they will almost surely fall under. We have classified some specific feedback functions to their according regime in table 3.1. It is notable that $f(x) = x^a$ for $\frac{1}{2} < a < 1$ falls under the eventual leader regime. It is known that for feedback function $f(x) = x^a$ where $a < 1$, almost surely all bins will receive the same number of balls[7]. These seem to be two contradictory statements on first glance, but by means of simulations we can find an answer to our problem. We have added a plot in which the number of balls in every bin at a certain point in time are shown. It can be seen as a snapshot of a process. There have been made multiple of these snapshots for one process, as to see how they evolve over time, these results have been put in figure 3.9.

Note that for $f(x) = x^{\frac{1}{2}}$, the number of balls in every bin consistently stays around a quarter of the total number of balls. However, for $f(x) = x^{\frac{3}{5}}$, bin 1 stays the greatest, but is losing its lead over time. Thus, we have found a very interesting case of the eventual leadership regime, where the eventual leader loses his lead.

3.3. Limiting distributions under linear feedback functions

In the previous section, limiting behaviour in the sense of the three different regimes has been considered for general feedback functions. Even though we have a very powerful theorem in theorem 3.8, there is still a lot to be explored. We note that even though the standard feedback function falls under the eventual leadership regime by theorem 3.8, we can find an explicit distribution for its limit.



(a) $f(x) = x^{\frac{1}{2}}$ (b) $f(x) = x^{\frac{1}{2}}$ (c) $f(x) = x^{\frac{1}{2}}$

(d) $f(x) = x^{\frac{3}{5}}$ (e) $f(x) = x^{\frac{3}{5}}$ (f) $f(x) = x^{\frac{3}{5}}$

Figure 3.9: Plots comparing different snapshots for two 4-bin Finite Pólya Processes with different feedback functions after $200,000$, $400,000$ and $600,000$ balls have been added. The Python code for these plots can be found in the Appendix.

These interesting limits only happen for the eventual leadership regime of the k-bin Finite Pólya Process. This is because of the very fact that if there is a monopoly, the distribution of the fractions of balls in the bins will converge to 0 for all bins except for one bin, whose fraction will converge to 1. Similarly, if there is an almost balanced regime, all fractions will converge to $\frac{1}{k}$. Only for the eventual leadership regime there are non-trivial limiting distribution.

That explicit distribution, is a very interesting phenomenon. It would be of great interest to find an explicit distribution for a broader range of feedback functions. That is what this section will be concerned with. We shall find the explicit limiting distribution for Finite Pólya Processes with positive, linear feedback functions $f(x) = bx + a$ for $a, b \in \mathbb{R}^+$.

3.3.1. A result regarding scalar multiples of feedback functions

Consider some feedback function $f(x) : \mathbb{N} \to (0, \infty)$. How would a k-bin Finite Pólya Process differ in behaviour if one were to look at this process using feedback function $g(x) : x \mapsto a \cdot f(x)$ for some $a \in (0, \infty)$?

Theorem 3.9. *A k-bin Finite Pólya Process under some feedback function $f(x) : \mathbb{N} \to (0, \infty)$ and some initial configuration $(b_1^0, \ldots, b_k^0)$ will be equivalent to a k-bin Finite Pólya Process under feedback function $g(x) : x \mapsto a \cdot f(x)$ for any $a \in (0, \infty)$ with the same initial configuration $(b_1^0, \ldots, b_k^0)$.*

Proof. We shall refer to the process with feedback function $f(x)$ as the first process and to the process with feedback function $g(x)$ as the second process. If we can show that for some arbitrary time step t, the probability that some bin i gets the next ball is the same for both processes, then we are done.

By definition, we know that for the first process the probability that bin i gets the next ball is:

$$\mathbb{P}(b_i^{t+1} = b_i^t + 1) = \frac{f(b_i^t)}{\sum_{j=1}^k f(b_j^t)}.$$

Applying that same definition to the second process, the probability that bin i gets the next ball is:

$$\mathbb{P}(b_i^t + 1 = b_i^t + 1) = \frac{g(b_i^t)}{\sum_{j=1}^k g(b_j^t)} = \frac{a \cdot f(b_i^t)}{a \cdot \sum_{j=1}^k f(b_j^t)} = \frac{f(b_i^t)}{\sum_{j=1}^k f(b_j^t)}.$$

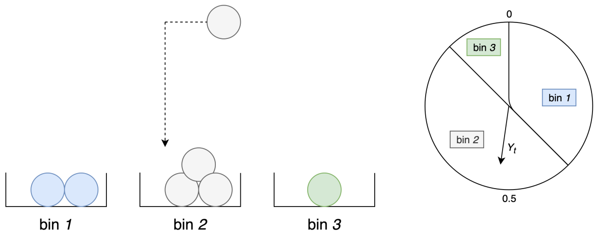


Figure 3.10: Some step in a 3-bin Finite Pólya Process under the standard feedback function. The probabilities of the next ball being added to any bin are represented in the pie chart on the right. Note that Y_t is some value that falls in the bin 2 section of the pie chart, thus the next ball will be added to bin 2.

Since both processes start with the same initial configuration and they have the same distribution for selecting the next bin to get the following ball, these two processes are equivalent. □

Even though theorem 3.9 might seem very trivial, this theorem gives a very interesting result in the sense that whenever we want to research some Finite Pólya Process with a certain feedback function, any positive scalar multiple of that feedback function will produce the same results as that feedback function.

Validation and Coupling Processes

We have stated that Finite Pólya Processes are equivalent under scalar multiples, but what does that exactly mean? If one were to run these processes next to each other, one might get very different outcomes. That is due to the inherent randomness of the processes. To circumvent this randomness, we could take a look at the distribution of fractions after a large number of steps for the different type of feedback functions and do statistical tests. We propose another method to compare these inherent random processes in such a way that statistical tests are not needed. We shall call this alternative method the coupling method.

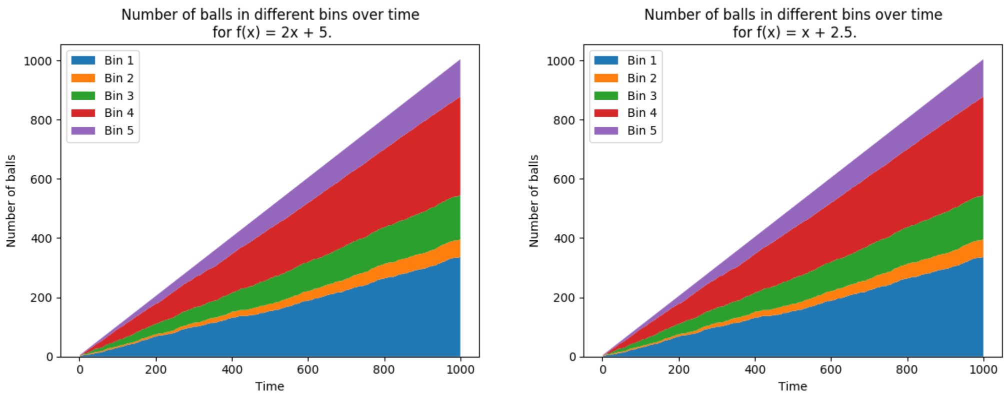
To understand the coupling method, a very important observation to be made is the fact that all of the randomness in a Finite Pólya Process can be put into one sequence of uniformly distributed random variables on the interval $(0, 1)$, $(Y_n)_{n\in\mathbb{N}}$. For any additional ball entering the process at time t, one can note the probabilities for any bin i to get the next ball, which we shall refer to as p_i. Note that $\sum_{j=1}^{k} p_j = 1$. Using our source of randomness for this time step, Y_t, we search for the smallest m such that $\sum_{j=1}^{m} p_j > Y_t$. The corresponding bin m shall get the next ball.

A visualization of this process is provided in figure 3.11. Computer Scientists have been using this method for Fitness Proportionate Selection, also known as, the Roulette Wheel Method in the field of Genetic Algorithms. The coupling method is based on this idea. The idea is to use the same sequence $(Y_n)_{n\in\mathbb{N}}$ for Finite Pólya Processes with different feedback functions and then look at their behaviour. If indeed theorem 3.9 holds, then for any scalar multiple of a feedback function, the process should add the same ball to the same bin.

We see in figure 3.11 that for a coupled Finite Pólya Processes where the feedback functions are scalar multiples of each other, the exact same evolution of bins over time occurs, which is what theorem 3.9 states.

3.3.2. Limiting distribution under feedback function $f(x) = x + a$

Using theorem 3.9, we can deduce the question with regards to the limiting distribution of any positive linear feedback function $f(x) = bx + a^*$ with $a^*, b \in (0, \infty)$ to finding the limiting distribution of any linear feedback function with $f(x) = x + a^*$ such that $a \in (0, \infty)$, since one can select $a = \frac{a^*}{b}$.



(a) $f(x) = 2x + 5$ (b) $f(x) = x + 2\frac{1}{2}$

Figure 3.11: Stackplots of two coupled Finite Pólya Processes with different feedback functions, that are scalar multiples of each other.

Using a similar strategy as in section 3.1, we shall find this limiting distribution. Once again, we shall consider the 2-bin limit and the k-bin limit separately. We note that there has not been recorded literature with regards to this limiting distribution as of yet, making us the first to dive into this specific topic.

2-bin limit

Once again, we shall define Y_n to be 1 whenever a ball n gets put into bin 1 and to be 0 whenever it gets put into bin 2. Since the proofs are very similar to the proofs from section 3.1, we shall omit some details.

Lemma 3.3. *The sequence $(Y_n)_{n\geq 1}$ is exchangeable.*

Proof. We shall show that for any sequence $t_1, \ldots, t_n$ of arbitrary length k, the probability $\mathbb{P}(Y_{t_1} = y_{t_1}, \ldots, Y_{t_1} = y_{t_k})$ is independent of the exact sequence of observations, as long as $\sum_{j=1}^{k} y_{t_j} = k$ for some k.

$$\mathbb{P}(Y_{t_1} = 1, Y_{t_k} = 1, Y_{t_{k+1}} = 0, Y_{t_n} = 0) \tag{3.10}$$

$$= \frac{b_1^0 + a}{b_1^0 + b_2^0 + 2a} \cdot \frac{b_1^0 + a + 1}{b_1^0 + b_2^0 + 2a + 1} \cdots \frac{b_1^0 + a + k - 1}{b_1^0 + b_2^0 + 2a + k - 1} \cdot \frac{b_2^0 + a}{b_1^0 + b_2^0 + 2a + k} \cdots \frac{b_2^0 + a + n - k - 1}{b_1^0 + b_2^0 + 2a + n - 1}, \tag{3.11}$$

$$= \frac{\prod_{j=0}^{k-1}(b_1^0 + a + j) \cdot \prod_{j=0}^{n-k-1}(b_2^0 + a + j)}{\prod_{j=0}^{n-1}(b_1^0 + b_2^0 + 2a + j)}. \tag{3.12}$$

Note that equation 3.12 is independent of the exact sequence of the values of $(Y_n)_{n\geq 1}$. Thus, $(Y_n)_{n\geq 1}$ is exchangeable. □

We shall use remark 2.1 to find the explicit limiting distribution of the Finite Pólya Process.

Theorem 3.10. *Given a 2-bin Finite Pólya Process under the feedback function $f(x) = x + a$ with $a \in \mathbb{R}^+$, then the limiting fraction X_1 will be Beta distributed with parameters $b_1^0 + a$ and $b_2^0 + a$, that is:*

$$x_1^t \xrightarrow{D} \textit{Beta}(b_1^0 + a, b_2^0 + a). \tag{3.13}$$

Proof. Let S_n be the number of times a ball is put into bin 1 during the first n time steps. Using equation 3.12 we get the following result:

$$\mathbb{P}(S_n = k) = \binom{n}{k} \frac{\prod_{j=0}^{k-1}(b_1^0 + a + j) \cdot \prod_{j=0}^{n-k-1}(b_2^0 + a + j)}{\prod_{j=0}^{n-1}(b_1^0 + b_2^0 + 2a + j)}. \tag{3.14}$$

Using remark 2.1 in combination with equation 3.14, we can explicitly find the distribution $F(p)$.

$$\begin{aligned}
\int_0^1 p^k \cdot (1-p)^{n-k} dF(p) &= \frac{\prod_{j=0}^{k-1}(b_1^0 + a + j) \cdot \prod_{j=0}^{n-k-1}(b_2^0 + a + j)}{\prod_{j=0}^{n-1}(b_1^0 + b_2^0 + 2a + j)}, \\
&= \frac{[b_1^0 + a]_k^+ \cdot [b_2^0 + a]_{n-k}^+}{[b_1^0 + b_2^0 + 2a]_n^+}, \\
&= \frac{\Gamma(b_1^0 + a + k)}{\Gamma(b_1^0 + a)} \frac{\Gamma(b_2^0 + a + n - k)}{\Gamma(b_2^0 + a)} \frac{\Gamma(b_1^0 + b_2^0 + 2a)}{\Gamma(b_1^0 + b_2^0 + 2a + n)}, \\
&= \frac{\Gamma(b_1^0 + a + k) \cdot \Gamma(b_2^0 + a + n - k)}{\Gamma(b_1^0 + b_2^0 + 2a + n)} \cdot \frac{\Gamma(b_1^0 + b_2^0 + 2a)}{\Gamma(b_1^0 + a)\Gamma(b_2^0 + a)}, \\
&= \mathrm{B}(b_1^0 + a + k, b_2^0 + a + n - k) \cdot \frac{1}{\mathrm{B}(b_1^0 + a, b_2^0 + a)}, \\
&= \frac{1}{\mathrm{B}(b_1^0 + a, b_2^0 + a)} \int_0^1 p^{b_1^0 + a + k - 1}(1-p)^{b_2^0 + a + n - k - 1} dp, \\
&= \int_0^1 p^k (1-p)^{n-k} \frac{p^{b_1^0 + a - 1}(1-p)^{b_2^0 + a - 1}}{\mathrm{B}(b_1^0 + a, b_2^0 + a)} dp.
\end{aligned}$$

Thus one can conclude that the density function of $F(p)$, $f^*(p)$, is exactly $f^*(p) = \frac{p^{b_1^0+a-1}(1-p)^{b_2^0+a-1}}{\mathrm{B}(b_1^0+a, b_2^0+a)}$, thus the limiting fraction of X_1 will be distributed $\mathrm{B}(b_1^0 + a, b_2^0 + a)$ for any $a \in \mathbb{Z}^+$. □

Having used the simplistic, yet elegant De Finetti argument for the limiting distributions of two classes of feedback functions, namely the standard feedback function and linear feedback functions, we wonder whether this argument can be used for even more limiting distributions. The answer to that question, is no. The De Finetti argument heavily relies on the following characteristic of feedback functions: $f(a + b) = f(a) + f(b)$. This shall be shown in remark 3.6.

Remark 3.6. *Consider a 2-bin Finite Pólya Process with feedback function $f(x)$, if $f(a + b) \neq f(a) + f(b)$, then the sequence $(Y_n)_{n\geq 1}$ which is the sequence indicating at every time step t whether bin 1 gets ball t, will not be exchangeable.*

Proof. We shall consider a $\{Y_1 = 1, Y_2 = 1, Y_3 = 0\}$ and $\{Y_1 = 1, Y_2 = 0, Y_3 = 1\}$. For a sequence to be considered exchangeable, these two events should be equal in probability. We shall show that if $f(a + b) \neq f(a) + f(b)$, the probabilities of these sequences happening will be different. And thus $(Y_n)_{n\geq 1}$ will not be exchangeable.

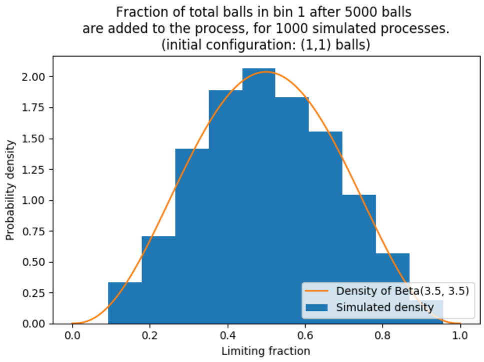
$$\mathbb{P}(Y_1 = 1, Y_2 = 1, Y_3 = 0) = \frac{f(b_1^0)}{f(b_1^0) + f(b_2^0)} \cdot \frac{f(b_1^0 + 1)}{f(b_1^0 + 1) + f(b_2^0)} \cdot \frac{f(b_2^0)}{f(b_1^0 + 2) + f(b_2^0)}. \tag{3.15}$$

$$\mathbb{P}(Y_1 = 1, Y_2 = 0, Y_3 = 1) = \frac{f(b_1^0)}{f(b_1^0) + f(b_2^0)} \cdot \frac{f(b_2^0)}{f(b_1^0 + 1) + f(b_2^0)} \cdot \frac{f(b_1^0 + 1)}{f(b_1^0 + 1) + f(b_2^0 + 1)}. \tag{3.16}$$

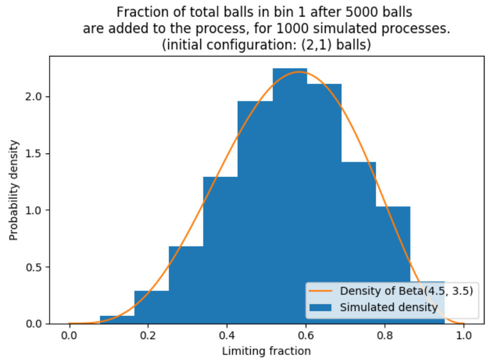
If the sequence were to be exchangeable, then equations 3.15 and 3.16 should be equal to each other. Note that the first fraction of both equations, is the same thus it can be ignored. that is:

$$\begin{aligned}
\frac{f(b_1^0 + 1)}{f(b_1^0 + 1) + f(b_2^0)} \cdot \frac{f(b_2^0)}{f(b_1^0 + 2) + f(b_2^0)} &= \frac{f(b_2^0)}{f(b_1^0 + 1) + f(b_2^0)} \cdot \frac{f(b_1^0 + 1)}{f(b_1^0 + 1) + f(b_2^0 + 1)}, \\
\frac{1}{f(b_1^0 + 2) + f(b_2^0)} &= \frac{1}{f(b_1^0 + 1) + f(b_2^0 + 1)}.
\end{aligned}$$

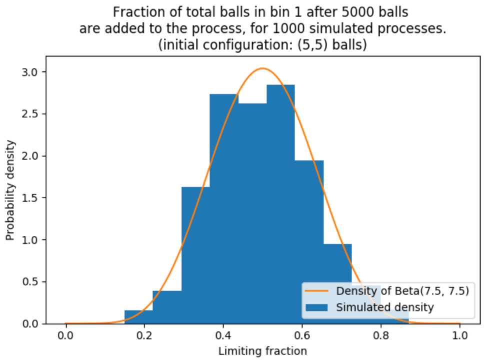
Which can only be equal if $f(a + b) = f(a) + f(b)$, thus if $f(a + b) \neq f(a) + f(b)$, then $(Y_n)_{n\geq 1}$ cannot be exchangeable. □



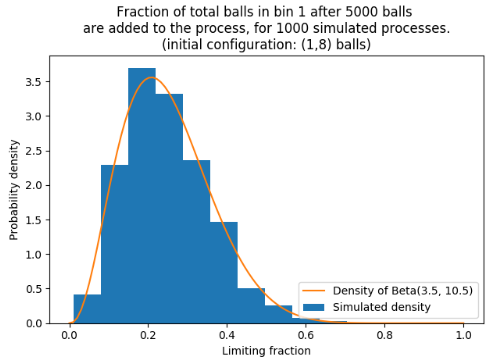
(a) Initial configuration $(1, 1)$



(b) Initial configuration $(2, 1)$

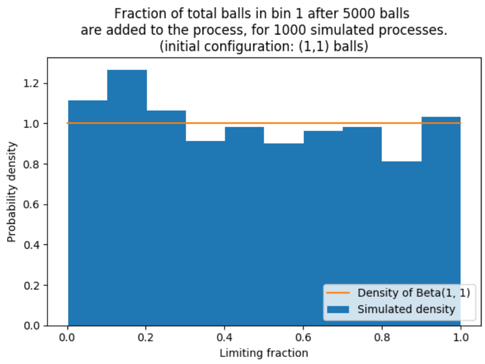


(c) Initial configuration $(5, 5)$

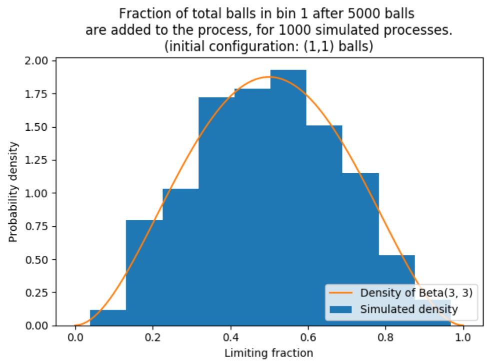


(d) Initial configuration $(1, 8)$

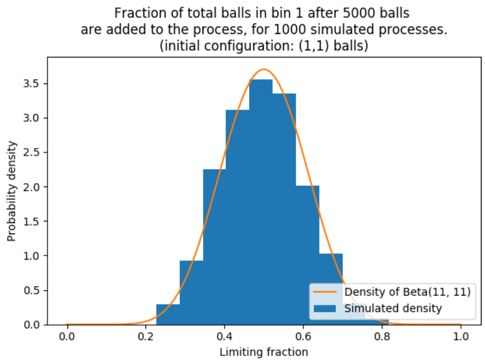
Figure 3.12: Plots comparing the results of theorem 3.10 to simulated 2-bin Finite Pólya Processes. In this plot the value of a, is constant, namely $a = 2\frac{1}{2}$ Every plot has a different initial configuration. The Python code for these plots can be found in the Appendix.



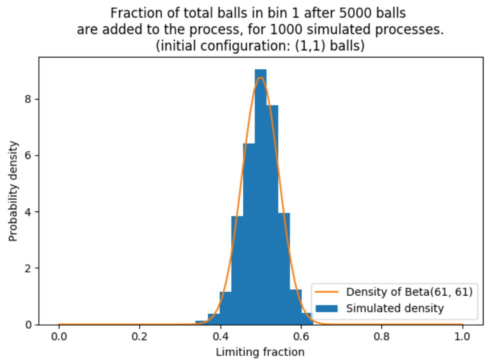
(a) $f(x) = x$, $a = 0$



(b) $f(x) = x + 2$, $a = 2$



(c) $f(x) = x + 10$, $a = 10$



(d) $f(x) = x + 60$, $a = 06$

Figure 3.13: Plots comparing simulated 2-bin Finite Pólya Processes. with different feedback functions of the form $f(x) = x + a$, where we vary a, but the same initial condition $(1, 1)$. The Python code for these plots can be found in the Appendix.

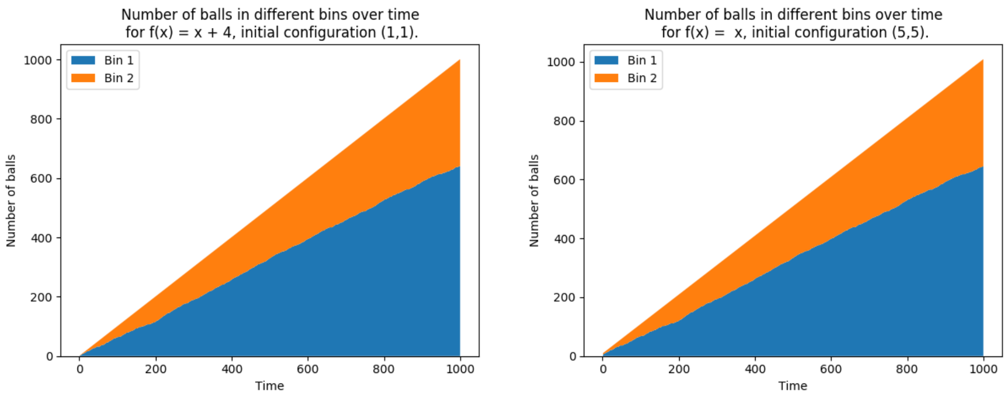
Validation and Simulations

We shall validate theorem 3.10 in a similar fashion as we have done for theorem 3.2 in section 3.1. Our simulations van be found in figure 3.12. We note that it seems reasonable to assume that theorem 3.10 holds, since the empirical probability density is close to the theoretical probability density.

The interesting findings from figure 3.12 can be found by comparing these results to the results of figure 3.1. Note that both figures have the same initial configuration, but are different with respect to their feedback function. Note that we had seen in figure 3.1 starting from $(1, 1)$ and adding the first ball to bin 1 changed the limiting distribution completely. This effect is a great deal more subtle when comparing it to the change in limiting distribution from figure 3.12a to figure 3.12b. Thus, it seems as if the first couple of balls have a smaller effect on the limiting distribution if $a > 0$, that is, the process is less path-dependent.

We test this hypothesis in figure 3.13, where we simulate different Finite Pólya Processes all with initial configuration $(1, 1)$, but with increasing a in their feedback function, we see that indeed, whenever a is larger, the limiting distributions seem to be less varied.

Another interesting observation is that if we had one Finite Pólya Process with initial configuration $(b_1^0 + a, b_2^0 + a)$ and the standard feedback function and another Finite Pólya Process with initial configuration (b_1^0, b_2^0), their limiting distribution is the same. Is it then true that these two processes are equal? We shall use the proposed coupling method of section 3.3.1 to make a conclusion regarding this fact. Figure 3.14 lets us believe that these processes are equivalent in the choices they make with respect to which bin gets the additional ball, however, since there initially are more balls in figure 3.14b, the total number of balls is greater that the total number of balls in figure 3.14a.



(a) $f(x) = x + 4$, initial configuration $(1, 1)$ (b) $f(x) = x$, initial configuration $(5, 5)$

Figure 3.14: Stackplots of two coupled Finite Pólya Processes with different feedback functions and different initial configurations, but the same limiting distribution.

We note that therefore instead of looking at a Finite Pólya Process with linear feedback function with $f(x) = x+a$, where $a \in \mathbb{N}$, we could instead add a balls to every bin and we would get the exact same process, which would be nearly equivalent processes. The only difference comes in the total number of balls at time t, since the initial configuration is different. Note that this only holds for integer values of a, since we cannot add fractions of balls.

k-bin limit

Since the Dirichlet distribution is the multivariate generalization of the Beta distribution one would expect that the k-bin limit of a Finite Pólya Process with feedback function $f(x) = x+a$ would be Dirichlet distributed with parameters $b_1^0 + 1, \dots, b_k^0 + a$.

Conjecture 3.1. *Given a k-bin Finite Pólya Process under the feedback function, $f(x) = x+a$ with $a \in \mathbb{R}^+$ and given initial condition $(b_1^0, \dots, b_2^0)$, then the limiting fraction $X_1, \dots, X_k$ will be Dirichlet distributed with parameters $b_1^0, \dots, b_k^0$, that is:*

$$x_1^t, x_2^t, \dots, x_{k-1}^t \xrightarrow{D} \text{Dirichlet}(b_1^0 + a, \dots, b_k^0 + a).$$

We shall make this conjecture a more plausible by comparing this conjectured limiting distribution with an empirical limiting distribution, similar as has been done in figure 3.3. The results of this can be found in figure 3.15. We note that the conjecture seems quite likely, since the empirical limiting distribution and the conjectured limiting distribution seem to be similar.

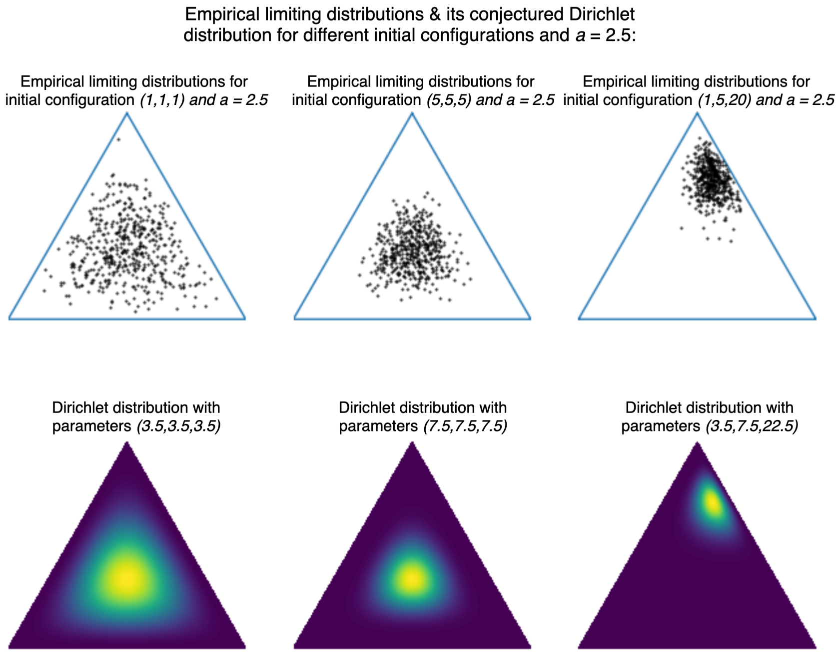


Figure 3.15: Plots comparing simulated 3-bin Finite Pólya Processes with feedback function $f(x) = x + 2\frac{1}{2}$ to the proposed limiting distribution for it by conjecture 3.1.We look perpendicularly at the simplex as described in figure 3.2. The left corner corresponds to bin 1, the right corner to bin 2 and the upper corner to bin 3.

4

Infinite Pólya Processes

Having discussed the results for Finite Pólya Processes, we wish to use these same results in the setting of Infinite Pólya Processes. In section 4.1 we shall speak on why our previous defined methods are not applicable. Then we shall proof the results for Infinite Pólya Processes as given by Chung et al.[7], our contribution shall be a certain level of probabilistic rigour that was previously not available in these cases. In this chapter, we shall only consider feedback functions of the form $f(x) = x^\gamma$ for $\gamma \in \mathbb{R}$.

4.1. Methods of the Finite Pólya Processes

Unfortunately, the methods that were of such great use to us for Finite Pólya Processes, namely the De Finetti method and the Exponential Embedding, do not work for Infinite Pólya Processes. This is because of the fact that these methods are heavily dependent on the fact that our state of bins is finite. For an Infinite Pólya Process the number of bins can be infinite, this makes it that our methods are not applicable.

One can generalize Pólya Urns in the following way: Whenever you choose some bin i for adding a ball, we only add 1 ball to that bin i, however one could also say that some bin j also gets an additional ball if bin i is selected, in fact, one can vary a lot with these replacement rules. Mahmoud has written extensively about these replacement rules [16]. Using the replacement rule for any bin i, we can build a replacement matrix R that contains all the replacement rules. If for any bin i the total number of balls being added to bins, is the same regardless of i, then this replacement matrix is called balanced and we can normalize these matrices such that their rows sum up to 1, by dividing over the total number of balls.

We note that for our Pólya Processes, these replacement matrices are identity matrices and therefore are balanced. Whenever such a replacement matrix is balanced, we can find a Markov Chain that has transition matrix R. This matrix is essential for the analysis of the limiting behaviour of Finite Pólya Process [3] [17]. Thus, when this matrix is infinite-dimensional, as happens in an Infinite Pólya Process, we cannot use methods reliant on the characteristics of finite replacement matrices. Both the Exponential Embedding method and the De Finetti method are dependent on this fact. We will have to resort to other methods of analyzing Infinite Pólya Processes.

4.2. Limiting behaviour under $f(x) = x^\gamma$, where $\gamma > 1$

In this section we shall work towards the following result:

Theorem 4.1. *Given an Infinite Pólya Process with feedback function $f(x) = x^\gamma$ with $\gamma > 1$, then there is a bin i such that the eventually all balls will either go into bin i, or go into a new bin. Next to that, for any $m \in \mathbb{Z}^+$ such that $m \le (m-1)\gamma$, only finitely many bins will ever reach size m.*

We will introduce $\max_t$ to be the largest number of balls in any bin at time t.

Lemma 4.1. *For any Infinite Pólya Process with feedback function $f(x) = x^\gamma$ such that $\gamma > 1$, the number of balls in the largest bin will eventually become infinite, that is: $\lim_{t\to\infty} \max_t = \infty$.*

Proof. We define A_t for any $t \in \mathbb{N}$ as the event that the largest bin at time t gets the following ball. If we can proof that eventually, A_t, will infinitely often be one, then we know that the largest bin will grow infinitely large almost surely, that is: $\lim_{t\to\infty} \max_t = \infty$. We shall proof this by using theorem 2.3.

We shall proof that equation 2.1 holds by proofing that the left hand side simultaneously is ≥ 1 and ≤ 1. First we will proof the ≥ 1 case. A very important observation will be that: $\mathbb{P}(A_i \mid A_j) \geq \mathbb{P}(A_i)$ for all $i > j$. Intuitively, if we know that the largest bin has increased at least once in the past, the largest bin will be larger than if we had not known this fact. Since the probability of a ball going to the largest bins is grows more than linear when an additional ball is added to the largest bin as a result of $\gamma > 1$, we know that $\mathbb{P}(A_i \mid A_j) \geq \mathbb{P}(A_i)$. Let $n \in \mathbb{N}$ be arbitrarily chosen:

$$\begin{aligned}
\frac{\sum_{i=1}^n \sum_{j=1}^n \mathbb{P}(A_i \cap A_j)}{\left(\sum_{j=1}^n \mathbb{P}(A_j)\right)^2} &= \frac{2\sum_{i=1}^n \sum_{j=1}^{i-1} \mathbb{P}(A_i \cap A_j) + \sum_{i=1}^n \mathbb{P}(A_i)}{\Big(\mathbb{P}(A_1) + \cdots + \mathbb{P}(A_n)\Big)\Big(\mathbb{P}(A_1) + \cdots + \mathbb{P}(A_n)\Big)}, \\
&= \frac{2\sum_{i=1}^n \sum_{j=1}^{i-1} \mathbb{P}(A_i \mid A_j)\mathbb{P}(A_j) + \sum_{i=1}^n \mathbb{P}(A_i)}{2\sum_{i=1}^n \sum_{j=1}^{i-1} \mathbb{P}(A_i)\mathbb{P}(A_j) + \sum_{i=1}^n \mathbb{P}(A_i)^2}, \\
&\geq \frac{2\sum_{i=1}^n \sum_{j=1}^{i-1} \mathbb{P}(A_i)\mathbb{P}(A_j) + \sum_{i=1}^n \mathbb{P}(A_i)^2}{2\sum_{i=1}^n \sum_{j=1}^{i-1} \mathbb{P}(A_i)\mathbb{P}(A_j) + \sum_{i=1}^n \mathbb{P}(A_i)^2} = 1.
\end{aligned}$$

Since n was arbitrarily chosen, this holds for all n and therefore it also holds for the $\liminf$. The ≤ 1 case can be proven by writing out the probability $\mathbb{P}(A_i \cap A_j)$. □

Lemma 4.2. *Given an Infinite Pólya Process with feedback function $f(x) = x^\gamma$ such that $\gamma > 1$, then for any $\epsilon > 0$, there is some $t \in \mathbb{N}$ such that the largest bin gets the next ball, ball $t+1$, with probability at least $\frac{\max_t}{t}(1-\epsilon)$.*

Proof. First, we will introduce K_t, which will be the random variable that is equal to the number of bins at time t. Without loss of generality, we shall assume $K_0 = 0$. We start with zero bins in our Infinite Pólya Process. This assumption is done for the sake of readability. Note that the number of bins only increases if some ball t creates a new bin. For this event, we introduce the random variable A_t, which shall be 1 whenever ball t creates a new bin and 0 else.

$$A_j = \begin{cases} 1, & \text{with probability } p, \\ 0, & \text{with probability } (1-p). \end{cases}$$

Note that $K_t = K_0 + \sum_{j=1}^t A_j = \sum_{j=1}^t A_j$. Using the fact that K_t is a series of t independent, identically distributed random variables, we know that $\frac{1}{t}K_t$ converges almost surely to its expectation $\mathbb{E}(\frac{1}{t}K_t)$ by theorem 2.5, that is:

$$\frac{1}{t}K_t \xrightarrow{a.s.} \mathbb{E}(\frac{1}{t}K_t) = \mathbb{E}(\frac{1}{t}\sum_{j=1}^t A_j) = \frac{1}{t}\sum_{j=1}^t \mathbb{E}(A_j) = \frac{1}{t}\cdot t \cdot (1\cdot p + 0 \cdot (1-p)) = p, \tag{4.1}$$

$$\frac{1}{t}K_t \xrightarrow{a.s.} p. \tag{4.2}$$

Using equation 4.2 in combination with the continuous mapping theorem 2.7 and continuous function $g(x) = \frac{1}{x}$, we can state the following:

$$\frac{t}{K_t} \xrightarrow{a.s.} \frac{1}{p}. \tag{4.3}$$

Which one can interpret as the following: The average number of balls per bin[1] will eventually almost surely be equal to $\frac{1}{p}$.

[1]We assume that there are t balls in total in the system, ignoring any initial balls

Select $\delta > 0$ such that $\frac{1}{1+\delta} > 1 - \epsilon$, note that this can be done since one can choose $\delta < \frac{\epsilon}{1-\epsilon}$.

We can therefore select $N > \frac{1}{p\delta}$. This implies that eventually the fraction of bins that have size less than N is at least $(1 - \delta)$. Note that this is equivalent to:

$$\frac{\sum_{j=1}^{K_t} \mathbb{1}_{\{b_j \leq N\}}}{K_t} \overset{a.a.}{\geq} 1 - \delta \Leftrightarrow \frac{\sum_{j=1}^{K_t} \mathbb{1}_{\{b_j > N\}}}{K_t} \overset{a.a.}{\leq} \delta. \tag{4.4}$$

We shall show that the right equation of equation 4.4 holds by using theorem 2.5. This can be applied, since for any bin j its size being larger than N, is independent of whether any other bins are larger than N. We also note that K_t is distributed binomially with parameters t and p, thus its probability mass function is explicitly known. We will disregard the event in which $K_t = 0$.

$$\begin{aligned}
\mathbb{E}\left(\frac{\sum_{j=1}^{K_t} \mathbb{1}_{\{b_j > N\}}}{K_t}\right) &= \sum_{m=1}^{t} \mathbb{E}\left(\frac{1}{m}\sum_{j=1}^{m} \mathbb{1}_{\{b_j > N\}}\right) \cdot \mathbb{P}(K_t = m) = \sum_{m=1}^{t} \frac{1}{m}\sum_{j=1}^{m} \mathbb{E}(\mathbb{1}_{\{b_j > N\}}) \cdot \mathbb{P}(K_t = m),\\
&= \sum_{m=1}^{t} \frac{1}{m}\sum_{j=1}^{m} \mathbb{P}(b_j > N) \cdot \mathbb{P}(K_t = m) \leq \sum_{m=1}^{t} \frac{1}{m}\sum_{j=1}^{m} \frac{\mathbb{E}(b_j)}{N} \cdot \mathbb{P}(K_t = m),\\
&= \sum_{m=1}^{t} \frac{\mathbb{E}(\frac{1}{m}\sum_{j=1}^{m} b_j)}{N} \cdot \mathbb{P}(K_t = m) = \sum_{m=1}^{t} \frac{\frac{1}{p}}{N} \cdot \mathbb{P}(K_t = m) = \frac{1}{p}\frac{1}{N} \cdot \sum_{m=1}^{t} \mathbb{P}(K_t = m),\\
&< \frac{1}{p} \cdot p\delta = \delta.
\end{aligned}$$

In this calculation we have assumed that $\mathbb{E}(\frac{1}{m}\sum_{j=1}^{m} b_j) = \frac{1}{p}$, this is exactly the expectation of the average number of balls per bin, for which we had found almost sure convergence in equation 4.3.

Another useful result is that eventually, the fraction of balls in bins of size less than N, is at least $(1-\delta)p$. We will denote all bins with a size smaller than N by $b_1, ..., b_{j_t}$. That makes the result equivalent to:

$$\frac{b_1 + ... + b_{j_t}}{b_1 + ... + b_{j_t} + ... + b_k} \overset{a.a.}{\geq} (1 - \delta)p. \tag{4.5}$$

This can be shown quite easily using the fact that equation 4.4 is equivalent to $\frac{j_t}{K_t} \overset{a.a.}{\geq} 1-\delta$ in combination with equation 4.2.

$$\begin{aligned}
\frac{b_1 + ... + b_{j_t}}{b_1 + ..., b_{j_t} + ... + b_k} &\geq \frac{j_t}{b_1 + ... + b_{j_t} + ... + b_k} = \frac{j_t}{K_t \cdot \frac{1}{K_t}(b_1 + ... + b_{j_t} + ... + b_k)},\\
&= \frac{j_t}{K_t} \cdot \frac{K_t}{t} \overset{a.a.}{\geq} (1 - \delta) \cdot p.
\end{aligned}$$

All of this work is needed so we can say the following: Eventually, the number of balls in bins that have size larger than N, are bounded above by $(1 - (1 - \delta)p)t$, that is:

$$\sum_{i=j+1}^{k} b_i \overset{a.a.}{\leq} (1 - (1 - \delta)p)t.$$

We can select M such that $M^{\gamma-1} > \frac{N^{\gamma-1}}{\delta}$. Using lemma 4.1, it follows that there is some t large such that $\max_t > M$. By definition of the feedback function, we know that the probability that the largest bin gets the $(t+1)^{\text{st}}$ ball is:

$$\begin{aligned}
\frac{(1-p)\max_t^{\gamma}}{\sum_{i=1}^{k} b_i^{\gamma}} &= \frac{(1-p)\max_t \max_t^{\gamma-1}}{\sum_{i=1}^{j} b_i b_i^{\gamma-1} + \sum_{i=j+1}^{k} b_i b_i^{\gamma-1}} \geq \frac{(1-p)\max_t \max_t^{\gamma-1}}{\sum_{i=1}^{j} b_i N^{\gamma-1} + \sum_{i=j+1}^{k} b_i \max_t^{\gamma-1}},\\
&= \frac{(1-p)\max_t}{\sum_{i=1}^{j} b_i (\frac{N}{\max_t})^{\gamma-1} + \sum_{i=j+1}^{k} b_i} \geq \frac{(1-p)\max_t}{\sum_{i=1}^{j} b_i \delta + \sum_{i=j+1}^{k} b_i},\\
&\geq \frac{(1-p)\max_t}{t\delta + (1-(1-\delta)p)t} = \frac{\max_t}{t}\cdot\frac{1-p}{\delta + 1 - (1-\delta)p},\\
&= \frac{\max_t}{t}\cdot\frac{1-p}{\delta(1-p) + (1-p)} = \frac{\max_t}{t}\cdot\frac{1}{\delta+1} > \frac{\max_t}{t}(1-\epsilon).
\end{aligned}$$

□

Lemma 4.3. *Suppose there are t independent balls being placed in bins, such that every ball has probability at least $\frac{m}{t}$ of landing in the first bin. Then the probability that the first bin receives fewer than $m-c$ balls, is less than $me^{\frac{-c^2}{2m}}$.*

Proof. Without loss of generality, we shall assume that each ball has probability $\frac{m}{t}$ to land in the first bin. We shall denote p_k as the probability that the first bin receives exactly k balls, then we can deduce the following:

$$\begin{aligned}
p_{m-c-1} &< \frac{p_{m-c-1}}{p_m} = \frac{\binom{t}{m-c-1}(\frac{m}{t})^{m-c-1}(\frac{t-m}{t})^{t-m+c+1}}{\binom{t}{m}(\frac{m}{t})^{m}(\frac{t-m}{m})^{t-m}} = \frac{\binom{t}{m-c-1}(t-m)^{c+1}}{\binom{t}{m}(m)^{c+1}},\\
&= \frac{\frac{(t)!}{(t-m+c+1)!(m-c-1)!}}{\frac{(t)!}{(t-m)!(m)!}}\cdot(\frac{t-m}{m})^{c+1} = \frac{(t-m)!}{(t-m+c+1)!}\frac{m!}{(m-c-1)!}(\frac{t-m}{m})^{c+1},\\
&= \frac{1}{(t-m+1)\cdots(t-m+c+1)}\frac{(m-c)(m-c+1)\ldots(m)}{1}(\frac{t-m}{m})^{c+1},\\
&= \prod_{i=1}^{c+1}\frac{(m-i+1)(t-m)}{(t-m+i)(m)} \leq \prod_{i=1}^{c+1}\frac{m-i+1}{m} = \prod_{i=1}^{c+1} 1+\frac{1-i}{m} \leq \prod_{i=1}^{c+1} e^{\frac{1-i}{m}} = \prod_{j=1}^{c} e^{-\frac{j}{m}},\\
&= e^{-\frac{1}{m}\sum_{j=1}^{c} j} \leq e^{\frac{-c^2}{2m}}.
\end{aligned}$$

We remark that $\frac{p_i}{p_{i+1}} < 1$ if $i+1<m$. This means that p_i is an increasing function as long as $i+1<m$. Note that this is equivalent to showing that a binomially distributed random variable has an increasing probability mass function for all values prior to its expectation.

$$\begin{aligned}
\frac{p_i}{p_{i+1}} &= \frac{\binom{t}{i}(\frac{m}{t})^{i}(\frac{t-m}{t})^{t-i}}{\binom{t}{i+1}(\frac{m}{t})^{i+1}(\frac{t-m}{t})^{t-i-1}} = \frac{(i+1)!\,(t-i-1)!}{i!\,(t-i)!}\frac{t-m}{m} = \frac{i+1}{t-i}\frac{t-m}{m},\\
&< \frac{i+1}{t-m}\frac{t-m}{m} = \frac{i+1}{m} < \frac{m}{m} = 1.
\end{aligned}$$

Since $m-c-1<m$ for any $c\in\mathbb{N}$, we know that $\sum_{i=0}^{m-c-1} p_i < mp_{m-c-1}$. Thus $\sum_{i=0}^{m-c-1} p_i < mp_{m-c-1} < me^{\frac{-c^2}{2m}}$. □

Lemma 4.4. *Given an Infinite Pólya Process with feedback function $f(x)=x^{\gamma}$ for some $\gamma>1$, for any $\epsilon>0$, almost surely $\max_t > t^{1-\epsilon}$.*

Proof. Choose $\delta_1,\delta_2,\delta_3>0$ sufficiently small such that: $\frac{(1-\delta_2)(1-\delta_3)}{(1+\delta_1)} > 1-\frac{\epsilon}{2}$ and $1+\delta_1(1-\frac{\epsilon}{2}) > (1+\delta_1)^{(1-\epsilon)}$. We define ρ as $\rho = \frac{(1-\delta_2)\delta_1}{(1+\delta_1)}$. Combining the former two conditions, we get: $(1-\delta_3)\rho > (1+\delta_1)^{(1-\epsilon)}$.

Introduce t_M to be the first time for which the largest bin contains M balls. We shall consider the interval from t_M until $(1+\delta_1)t_M$, which will be referred to as the M^{th} time interval. We will show that during this M^{th} time interval, the largest bin will almost surely grow more than factor $(1+\delta_1)^{(1-\epsilon)}$ for large enough M.

During this M^{th} time interval, the time will be smaller than $(1+\delta_1)t_M$ and the largest bin will contain at least M balls. By using lemma 4.2 with $\epsilon = \delta_2$, we can lower bound the probability for any ball entering the largest bin during time interval M.

$$(1-\delta_2)\frac{\max_t}{t} \geq \frac{(1-\delta_2)M}{(1+\delta_1)t_M}.$$

Since the time interval is of length $\delta_1 \cdot t_M$, the expected number of balls to enter the largest bin during this interval will at least be: $\frac{(1-\delta_2)M}{(1+\delta_1)t_M} \cdot (\delta_1 t_M) = \frac{(1-\delta_2)\delta_1 M}{1+\delta_1} = \rho M$. By lemma 4.3, noting that $m = \rho M$ and we select $c = \delta_3 \rho M$, we know that the probability that the largest bin receives fewer than $(1-\delta_3)\rho M$ balls is at most, $\rho M e^{-\frac{\delta_3^2 \rho M}{2}}$.

Note that:

$$\sum_{M=1}^{\infty} \rho M e^{-\frac{\delta_3^2 \rho M}{2}} < \infty,$$

so by theorem 2.2, there can only be finitely many times for which the largest bin fails to receive $(1-\delta_3)\rho M$ balls during the M^{th} time interval. Thus we know that eventually the time grows by a factor $(1+\delta_1)$, the largest bin grows by a factor of $1+(1-\delta_3)\rho > (1+\delta_1)^{1-\epsilon}$ and therefore, eventually $\max_t > t^{1-\epsilon}$. □

With all of this knowledge we can now work towards our original theorem 4.1.

Proof. We will first of all fix a $m \in \mathbb{Z}^+$ such that $m < (m-1)\gamma$. With this condition, one can select $\epsilon > 0$ such that $m < (m-1)\gamma(1-\epsilon)$. For the sake of readability, $\gamma(1-\epsilon)$ shall be denoted as γ^*. Using lemma 4.4, we know that eventually the largest bin has size $t^{(1-\epsilon)}$. This means that after some time t, for all times greater than t, $N > t$, one knows that the largest bin will have size at least $t^{(1-\epsilon)}$.

The probability that a ball goes to a specific bin of size i at time $N > t$, is then bounded above by $\frac{i^\gamma}{t^{\gamma^*}}$.

$$\mathbb{P}(\text{one bin of size } i \text{ gets next ball at time N}) = \frac{i^\gamma}{\sum_{j=1}^{k} b_j^\gamma} \leq \frac{i^\gamma}{\max_t^\gamma} < \frac{i^\gamma}{t^{(1-\epsilon)\gamma}} = \frac{i^\gamma}{t^{\gamma^*}}.$$

Suppose that a new bin is created at some time $t_1 > N$ The probability that it receives balls at times $t_2 < \cdots < t_m$, is upper bounded by $\prod_{j=1}^{m-1} \frac{j^\gamma}{t_{i+1}^{\gamma^*}}$. That is:

$$\mathbb{P}(\text{bin created at time } t_1 \text{ ever receives } m \text{ balls}) \leq \sum_{\{(t_2,\ldots,t_m)\,:\,t_1<\cdots<t_m\}} \prod_{i=1}^{m-1} \frac{i^\gamma}{t_{i+1}^{\gamma^*}}. \tag{4.6}$$

We shall show that there are just finitely many bins that will be created after time N and receive k balls, that is:

$$\mathbb{P}(\{\text{bin created at time } t > N \text{ ever receives } k \text{ balls}\} \text{ i.o.}) = 0. \tag{4.7}$$

We notice that we can use theorem 2.2 to proof equation 4.7. For that purpose, we will need a useful upper bound for 4.6, in terms of its time t_1.

$$\sum_{\{(t_2,\ldots,t_m)\,:\,t_1<\cdots<t_m\}} \prod_{i=1}^{m-1} \frac{i^\gamma}{t_{i+1}^{\gamma^*}} < (m-1)!^\gamma \int_{t_1}^{\infty} \cdots \int_{t_{m-1}}^{\infty} (t_2 \cdots t_m)^{-\gamma^*} dt_m \cdots dt_2, \tag{4.8}$$

$$= \frac{(m-1)!^\gamma}{(\gamma^*-1)(2\gamma^*-2)\ldots((m-1)\gamma^*-(m-1))} t_1^{-(m-1)\gamma^*+(m-1)}, \tag{4.9}$$

$$= \frac{(m-1)!^{\gamma-1}}{(\gamma^*-1)^{m-1}} t_1^{-(m-1)\gamma^*+(m-1)}. \tag{4.10}$$

By theorem 2.2, proving 4.7 is equivalent to proving:

$$\sum_{n=N}^{\infty} \mathbb{P}(\text{bin created at time } n \text{ ever receives } m \text{ balls}) < \infty \tag{4.11}$$

Which can be done quite easily using equations 4.6 in combination with 4.10.

$$\begin{aligned}
\sum_{n=N}^{\infty} \mathbb{P}(\text{bin created at time } n \text{ ever receives } m \text{ balls}) &\leq \sum_{n=N}^{\infty} \frac{(m-1)!^{\gamma-1}}{(\gamma^*-1)^{m-1}} n^{-(m-1)\gamma^*+(m-1)}, &(4.12)\\
&\leq \int_{n=N}^{\infty} \frac{(m-1)!^{\gamma-1}}{(\gamma^*-1)^{m-1}} n^{-(m-1)\gamma^*+(m-1)} dn, &(4.13)\\
&= \frac{(m-1)!^{\gamma-1}}{(\gamma^*-1)^{m-1}((m-1)\gamma^*-m} N^{-(m-1)\gamma^*+m}, &(4.14)\\
&< \infty. &(4.15)
\end{aligned}$$

Thus we know that there are only finitely many bins that will ever receive m balls. Note that even if all of the bins created prior to time N were to receive m balls, then there would still be finitely many bins that have received m balls.

Now we want to show that the probability that for some ball t, this ball enters an **existing** bin of size less than m, does not happen infinitely often, that is:

$$\mathbb{P}(\{\text{ball } t \text{ enters an existing bin of size less than } m\} \text{ i.o.}) = 0. \tag{4.16}$$

We remark that the complement of ball t entering an existing bin of size less than k is: ball t producing a new bin or ball t entering a bin of greater or equal to m, we note that showing that equation 4.16 is equivalent to showing:

$$\mathbb{P}(\{\text{ball } t \text{ produces a new bin or ball } t \text{ enters a bin of size greater than or equal to } m\} \text{ a.a.}) = 1, \tag{4.17}$$

which will give us exactly what we need. We know that there are finitely many bins of size greater than or equal to m. Therefore, using our result for Finite Pólya Process with feedback function $f(x) = x^\gamma$ with $\gamma > 1$, which will almost surely lead to the monopoly scenario, we know that one the bins greater than m will almost surely hold the monopoly.

All that remains to proof for our theorem, is equation 4.16. We shall denote the number of existing bins smaller than m at time t as j_t. We shall use theorem 2.2 to proof that equation 4.16 indeed holds.

$$\begin{aligned}
\sum_{t=1}^{\infty} \mathbb{P}(\{\text{ball } t \text{ enters an existing bin of size less than } m\}) &= \sum_{t=1}^{\infty} \frac{\sum_{i=1}^{j_t} (b_i^t)^\gamma}{\sum_{i=1}^{j_t} (b_i^t)^\gamma + \sum_{i=j_t+1}^{k} (b_i^t)^\gamma},\\
\leq \sum_{t=1}^{\infty} \frac{j_t m^\gamma}{j_t + (\max_t)^\gamma} = \sum_{t=1}^{\infty} \frac{m^\gamma}{1 + \frac{t^{\gamma^*}}{j_t}} \leq m^\gamma \sum_{t=1}^{\infty} \frac{1}{1+t^{\gamma^*}} < \infty.
\end{aligned}$$

□

4.3. Limiting behaviour under $f(x) = x^\gamma$, where $\gamma \leq 1$

Having discussed the limiting behaviour in Infinite Pólya Processes with feedback function $f(x) = x^\gamma$ with $\gamma > 1$ a very natural question that arises, is the following: What limiting behaviour can we expect in Infinite Pólya Processes with feedback function $f(x) = x^\gamma$ with $\gamma \leq 1$?

We shall introduce some terms to be able to be able to answer this question. Denote $f_{i,t}$ as the fraction of bins at time t that contain exactly i balls. An interesting question is whether for such an Infinite Pólya Process with negative feedback for all bin sizes i this limit exists. Chung et al.[7] state that it seems clear that this limit exists, but they could not prove it, except for the $\gamma = 1$ case. They did two assumptions and derived some limiting behaviour of the process under those assumptions.

We shall recap their findings and run some simulations. If our simulations were to show very different results than the limiting behaviour we were to expect under Chung et al.'s assumptions, there is ground to believe that perhaps these limits might not exist in the same sense that they have assumed.

Assumptions 4.1 (Chung's assumptions)**.**

1. *For each i, there exists $f_i \in \mathbb{R}^+$ such that, almost surely $\lim_{t\to\infty} f_{i,t}$ exists and is equal to f_i.*

2. *Almost surely $\sum_{j=1}^{\infty} f_{j,t} j^\gamma$ exists, is finite and is equal to $\sum_{j=1}^{\infty} f_j j^\gamma$.*

Theorem 4.2. *For an Infinite Pólya Process with feedback function $f(x) = x^\gamma$ such that $\gamma \leq 1$, if assumptions 4.1 hold, then there exists a constant $K > 0$ (which depends on p and γ) such that for $i \geq 2$ the following holds almost surely:*

$$f_i = \Big(\frac{(i-1)^\gamma}{K + i^\gamma}\Big) f_{i-1}. \tag{4.18}$$

Proof. Introduce K_t as the number of bins at time t. We note that for $i > 0$ the number of bins with exactly i balls at time t, is equal to $f_{i,t} \cdot K_t$

We shall denote $p_{i,t}$ as the probability that ball t is placed in a bin of size i. Let us remark that $p_{0,t} = p$. Then for $i > 0$ the following holds:

$$p_{i,t} = \frac{(1-p) f_{i,t} K_t \cdot i^\gamma}{K_t \sum_{j=1}^{\infty} f_{j,t} j^\gamma} = \frac{(1-p) f_{i,t} i^\gamma}{\sum_{j=1}^{\infty} f_{j,t} j^\gamma}. \tag{4.19}$$

We denote $e_{i,t}$ as the change of bins of size i at time t. and we introduce $E_{i,t} = \mathbb{E}(e_{i,t})$. There are two ways the number of bins of size i at time t can change. It can happen that a bin of size $i-1$ at time t receives the additional ball, this will result into a positive change of $e_{i,t}$, but it could also happen that one of the bins of size i gets the additional ball, resulting in a negative change of $e_{i,t}$. We introduce $B_{i,t}$, which is the random variable that is equal to 1 if some bin of size i gets the additional ball at time t and 0 else, note that $B_{i,t}$ is distributed Bernoulli with parameter $p_{i,t}$.

$$E_{i,t} = \mathbb{E}(e_{i,t}) = \mathbb{E}(B_{i-1,t} - B_{i,t}) = p_{i-1,t} - p_{i,t}.$$

By assumptions 4.1, $p_{i,t}$ for any i should converge, and thus $E_{i,t}$ should converge. We shall refer to these limits as p_i and E_i respectively. By definition, f_i should stay the same in the limit, meaning that there can be balls that get added to bins of size i, as long as the balls that get added to bins of size $i+1$ happens at the same rate. this means thus that $\frac{E_i}{f_i} = \frac{E_j}{f_j} = C$ for any $i, j \in \mathbb{N}$ and some constant $C \in \mathbb{R}^+$. Using this fact we can deduce:

$$\begin{aligned} C f_i = E_i = p_{i-1} - p_i &= \frac{(1-p)(f_{i-1}(i-1)^\gamma - f_i i^\gamma)}{\sum_{j=1}^{\infty} f_j j^\gamma}, \\ f_i &= \frac{1-p}{C \sum_{j=1}^{\infty} f_j j^\infty}(f_{i-1}(i-1)^\gamma - f_i i^\gamma), \\ f_i(1 + \frac{i^\gamma}{K}) &= \frac{1}{K} f_{i-1}(i-1)^\gamma, \\ f_i &= \Big(\frac{(i-1)^\gamma}{K + i^\gamma}\Big) f_{i-1}, \end{aligned}$$

for $i \geq 2$ and $K = C \sum_{j=1}^{\infty} f_j j^\gamma / (1-p)$. Thus we have found the relation we were looking for. □

Note that theorem 4.2 contains a recurrence. We shall use those recurrence equations to estimate values of f_i for large i, if assumptions 4.1 hold. We shall use $f_i \propto g(i)$ to denote $f_i = c(1 + o(1)(g(i))$ for some constant c.

Theorem 4.3. *For an Infinite Pólya Process with feedback function $f(x) = x^\gamma$ and $\gamma \leq 1$, suppose assumptions 4.1 hold, then the limit f_i of fractions of bins with i balls will almost surely satisfy the following:*

$$f_i \propto \begin{cases} i^{-(1+\frac{1}{1-p})}, & \text{if } \gamma = 1, \\ i^{-\gamma} e^{-\frac{K}{1-\gamma} i^{1-\gamma}}, & \text{if } 0 < \gamma < 1, \\ \frac{1}{(K+1)^i}, & \text{if } \gamma = 0, \\ O\Big(\frac{((i-1)!)^\gamma}{K^i}\Big), & \text{if } \gamma < 0. \end{cases}$$

Proof. We shall define $E_i, E_{i,t}, p_i, p_{i,t}, C$ and K as defined in the proof of theorem 4.2. We shall first show that $\sum_{i=1}^\infty f_i = 1$. Note that for any time t, $\sum_{i=1}^\infty f_{i,t} = 1$. We define $a_t = \sum_{i=1}^\infty f_{i,t} = 1$. note that $\lim_{t\to\infty} a_t = a = \sum_{i=1}^\infty f_i$. Note that we have a constant sequence $a_t = 1$ and we want to proof that the limit of this sequence is also equal to 1. This trivially holds, thus $\sum_{i=1}^\infty f_i = 1$.

We shall show that $\sum_{i=1}^\infty E_i = p$. At any time t, if we were to look at the expected number of changes in bins that have size smaller or equal to n, we know that this value would be equal to $\sum_{j=1}^n E_{i,t} = p - p_{n,t}$, thus $\sum_{j=1}^n E_i = p - p_n$. Note that $\sum_{n=1}^\infty p_n$ is bounded by assumptions 4.1, thus $\lim_{n\to\infty} p_n = 0$. Therefore, we know that $\sum_{i=1}^\infty E_i = \lim_{n\to\infty} p - p_n = p$.

Using these facts, we can deduce that $C = \frac{E_i}{f_i} = C$ for all $\gamma \leq 1$ and thus $K = \frac{p}{1-p} \sum_{i=1}^\infty f_i i^\gamma$.

Consider $\gamma = 1$, We know that $\sum_{i=1}^\infty f_i i$ is exactly the average bin size, which we know converges almost surely to $\frac{1}{p}$ as has been shown in equation 4.3[2]. Thus for $\gamma = 1$, $K = \frac{1}{1-p}$. Using the recurrence relation from theorem 4.2, we note the following:

$$f_i \propto \prod_{j=2}^{i} \frac{j-1}{j + \frac{1}{1-p}} \propto \frac{\Gamma(i)}{\Gamma(i + 1 + \frac{1}{1-p})} \propto i^{-(1+\frac{1}{1-p})}.$$

Thus the bin sizes obey a power-law distribution with power law exponent $1 + \frac{1}{1-p}$.

For the other values $\gamma < 1$ we shall approximate the asymptotic behavior of f_i when $\gamma < 1$. If $0 < \gamma < 1$, then for large i:

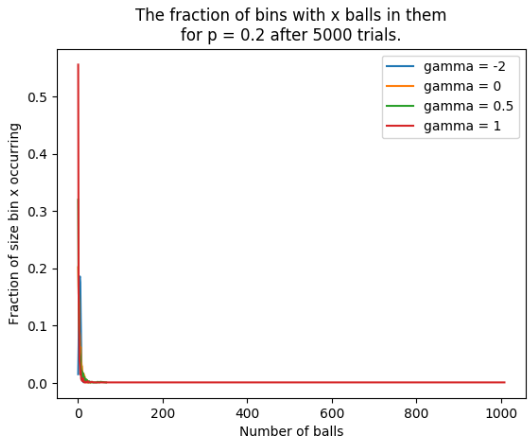
$$\begin{aligned} f_i &\propto \prod_{j=2}^{i} \frac{(j-1)^\gamma}{K + j^\gamma} \propto i^{-\gamma} \prod_{j=1}^{i} \frac{j^\gamma}{K + j^\gamma}, \\ &= i^{-\gamma} \prod_{j=1}^{i} \frac{1}{1 + \frac{K}{j^\gamma}} \propto i^{-\gamma} e^{-\sum_{j=1}^{i} \frac{K}{j^\gamma}}, \\ &\propto i^{-\gamma} e^{-\frac{K}{1-\gamma} i^{1-\gamma}}. \end{aligned}$$

If $\gamma = 0$, we note that $f_i \propto \frac{1}{(K+1)^i}$. Lastly, if $\gamma < 0$, then:

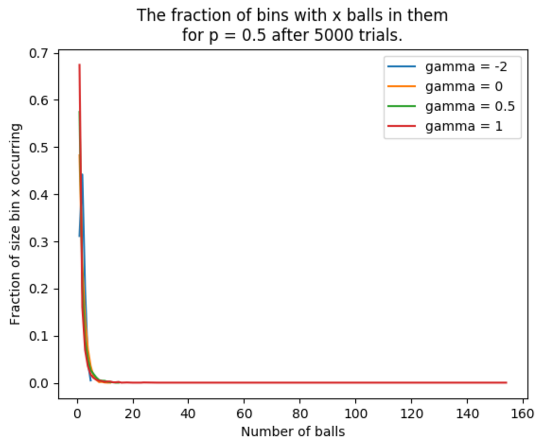
$$f_i \propto \prod_{j=2}^{i} \frac{(j-1)^\gamma}{K + j^\gamma} = O\Big(\prod_{j=2}^{i} \frac{(j-1)^\gamma}{K}\Big) = O\Big(\frac{((i-1)!)^\gamma}{K^i}\Big).$$

□

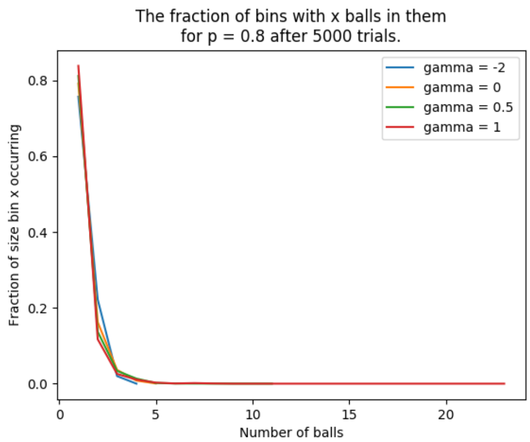
[2]Note that this specific result holds for all Infinite Pólya Processes.



(a) $p = 0.2$



(b) $p = 0.5$



(c) $p = 0.8$

Figure 4.1: Plots for three different values of p, such that every plot contains 4 simulations of an Infinite Pólya Process, each of these 4 simulations has a different feedback function.

We will take a look at some simulations of these Infinite Pólya Processes. Unfortunately, the results of theorem 4.3 are not very useful to us in these simulations. We note that for $\gamma < 1$, we are not sure what the value of K is. Other than that, theorem 4.3 only gives us a result with regards to the fractions of bins being proportional to a function, giving us another parameter that we have to estimate. For these reasons, we note that using simulations based on theorem 4.3 will not increase our knowledge with regards to assumptions 4.1.

We have provided some simulations with regard to the fractions of bins with size i at time t $f_{i,t}$ for different values of γ and p in the figure 4.1. In any of the three plots provided in figure 4.1, four different Infinite Pólya Processes have been ran, with different values for γ, corresponding to the different classes as proposed by theorem 4.3. Every plot has a different value for p, we note that it seems as if the different value for γ does not warrant a very different distribution with regards to the fractions of numbers of bins with some size. It is logical that there will be a lot of bins with size 1, but how larger the bin gets, how less likely it is to receive a next ball.

5 Conclusion and Recommendations

Our conclusions can be divided into three categories. First of all, there are the conclusions with regard to Finite Pólya Processes, secondly there are the conclusions with using the Finite Pólya Process methods for Infinite Pólya Processes, and lastly there are the conclusions with regard to Infinite Pólya Processes. After having discussed these results, we shall look at some recommended research subjects.

5.1. Finite Pólya Processes

In our analysis of Finite Pólya Processes, we have made use of two known methods, namely De Finetti method and the Exponential Embedding method. We have also introduced a new method of analyzing Finite Pólya Processin the 'Coupling Method'.

We have proven that for 2-bin Finite Pólya Processes under the standard feedback function $f(x) = x$ with initial configuration (b_1^0, b_2^0), the limiting distribution will be a Beta distribution with parameters (b_1^0, b_2^0) in theorem 3.2. We also note that this can be generalized to k-bin Finite Pólya Processes, which will have a Dirichlet distribution as its limiting distribution in. We have validated these results by means of simulations.

After having analyzed processes under the standard feedback function, we have analyzed processes under general feedback functions, we have done this by using the Exponential Embedding methods of Oliveira [19] [18]. Our main result is that, purely based on properties of the feedback function, any Finite Pólya Process can be put into one of three mutually exclusive regimes: Monopolistic Regime, Eventual Leader Regime or Almost-Balanced Regime. Our contribution was finding examples of functions that could be put into one of these three regimes.

An interesting result was that even though for $f(x) = x^\gamma$ for $\gamma < 1$ it is known to that the fractions of ball in every bin will converge to $\frac{1}{k}$ almost surely, we have found that for $\frac{1}{2} \leq \gamma < 1$ there will also almost surely be an eventual leader. We have gained insight in these seemingly contradictory statements by means of simulations. It appears that even though there indeed is an eventual leader, the lead of this eventual leader will converge to 0.

We have found that the most interesting processes, are the processes that have an Eventual Leadership Regime. That is why we have investigated the limiting distribution of Finite Pólya Processes under general positive linear feedback functions, for which there has not been previous literature. We have found that Finite Pólya Processes are equivalent under positive scalar multiples of their feedback function. Using that fact we have found that for any 2-bin Finite Pólya Process with feedback function $f(x) = bx + a$ where $a, b \in (0, \infty)$ and initial configuration (b_1^0, b_2^0), the limiting distribution is a Beta distribution with parameters $b_1^0 + \frac{a}{b}, b_2^0 + \frac{a}{b}$. By means of simulations, we have found that Finite Pólya Processes under linear feedback functions are less path-dependent than Finite Pólya Processes under the standard feedback function. We have also found a transformation of Finite Pólya Processes under the standard feedback function and under linear positive feedback functions such that these processes are equivalent.

We have conjectured a limiting distribution for linear feedback functions in higher dimensions, by noting the similarities between this process and the processes under the standard feedback function. We have also shown that the De Finetti argument is not viable for non-linear feedback functions.

5.2. Analyzing Infinite Pólya Processes using methods for Finite Pólya Processes

We have found that two popular methods of analyzing Finite Pólya Processes, which are De Finetti methods and dding methods cannot be used to analyze Infinite Pólya Processes. This is due to the occurrence of infinitely many bins in Infinite Pólya Processes. We conclude that these processes are so fundamentally different, that we have to consider different ways of interpreting these processes than has been done as of yet.

5.3. Infinite Pólya Processes

For Infinite Pólya Processes, we have proven results for processes under feedback function $f(x) = x^{\gamma}$, wheter $\gamma \in \mathbb{R}$, which were largely inspired by Chung et al.[7]. Our contribution is adding a level of probabilistic rigour to these proofs that was previously not included. We note that especially in the proofs of lemma's 4.1 and 4.2 and theorem 4.1 we have found some mistakes in the work of Chung et al. , which we have corrected. We shall speak on the mistakes and how we have solved them in the upcoming three subsections, after which we shall dedicate a section to our simulations with respect to theorem 4.3.

5.3.1. Lemma 4.1

Chung et al. proof this lemma by showing that the events A_t as defined in the our proof of the lemma happen almost surely by means of using the second case of theorem 2.2. Note however that for this case to hold, $(A_t)_{t\geq 1}$ should be an independent sequence, which it is not. In our proof, we use theorem 2.3, which does not need this independent characteristic of the sequence, to show that A_t happens infinitely often, which is enough for our purposes.

5.3.2. Lemma 4.2

Chung et al. use a different restriction on δ, which they base on the fact that $\sum_{i=1}^{j} b_i \overset{a.a.}{\leq} (1-\delta)pt$, which they assume follows from equation 4.5, however this is not the case, since they mix up it happening almost always and almost surely. Our restriction on δ produces the same result as they have, but it does not rely on $\sum_{i=1}^{j} b_i \overset{a.a.}{\leq} (1-\delta)pt$.

5.3.3. Theorem 4.1

Chung et al. deduce that equation 4.16 holds because of the fact that the limit of this probability goes to 0. This is not correct. We use the first statement of theorem 2.2 to proof this statement. In doing this, we have found a tighter bound for the probability $\mathbb{P}(\{\text{ball } t \text{ enters an existing bin of size less than } m\})$, since their proposed bound for this probability would not have given us a converging series, which is required for theorem 2.2 to hold.

5.4. On simulation and theorem 4.3

We have seen that theorem 4.3 relies on assumptions 4.1 to hold. We have tried to get a better understanding of whether assumptions 4.1 holds by means of simulations, but we have failed to do so. This is because of the fact that theorem 4.3 is too broad in terms of parameters to be able to simulate it.

5.5. Recommendations

There are a lot of very interesting that are still to be explored in the field of Pólya Processes.

(Random Recursive Trees) Janson [13] has found a connection between infinitely many-coloured Pólya urns and random recursive trees, for which there are results available. The link between our Infinite Pólya Processand the growth of these random recursive trees could be investigated and could lead to new insights.

(Proofing conjecture 3.1) This conjecture remains to be unproven. Using a similar strategy as in the proof of theorem 3.10 and using the proof of theorem 3.3, it should follow straight forward. Stirling's Formula will probably play an important role in this part.

(Division of Regimes) Note that theorem 3.8 gives us that for feedback function $f(x) = x^\gamma$ where $\frac{1}{2} < \gamma \leq 1$, we know that any Finite Pólya Process will almost surely converge to an eventual leadership regime, but it is a well known result that for $\gamma < 1$ all fraction sizes converge to $\frac{1}{k}$, by means of simulation in figure 3.9, we note that these results are not contradictory, since it can be possible that there is an eventual leadership that loses its lead.Perhaps another division could be made in these regimes: Diminishing Eventual Leadership and Non-Diminishing Eventual Leadership, where in the first case the leaders lead will eventually go to 0. For what other kind of feedback functions does this Diminishing Eventual Leadership regime happen?

A
Python Code

A.1. Bin.py

```
class Bin:

    def __init__(self,numberOfBalls):
        self.numberOfBalls = numberOfBalls

    def addBall(self):
        self.numberOfBalls += 1

    def getBalls(self):
        return self.numberOfBalls
```

A.2. Urn.py

```
from Bin import Bin

class Urn:
    #Every urn model contains some parameters
    # 1. The parameter p which decides whether we add a new bin
    # 2. The parameter gamma which decides how you reinforce.
    # 3. The number of balls in every bin.

    def __init__(self,p,f,numberOfBins):
        self.p = p
        self.f = f
        self.numberOfBins = numberOfBins
        binList = []
        for i in range(0,numberOfBins):
            binList.append(Bin(1))
        self.binList = binList

    def addBin(self):
        self.numberOfBins += 1
        self.binList.append(Bin(1))

    # Note: The reinforcedBinSelection is the same as a roulette wheel
       selection
    # the random number has to be between 0 and 1.
    def reinforcedBinSelection(self,randomNumber):
```

```
        binSizeList = []
        for bin in self.binList:
            binSizeList.append(bin.getBalls())
        #print(binSizeList)
        currDensity = 0
        binCounter = 0
        totalDensity = 0
        for binSize in binSizeList:
            totalDensity += self.f(binSize)
        randomDensityPoint = totalDensity*randomNumber

        while(randomDensityPoint >= currDensity):
            currDensity += self.f(binSizeList[binCounter])
            binCounter += 1
        return self.binList[binCounter - 1]

    def addIteration(self,randomNumberOne,randomNumberTwo):
        if(randomNumberOne < self.p):
            self.addBin()
        else:
            selectedBin = self.reinforcedBinSelection(randomNumberTwo)
            selectedBin.addBall()

    def getBinSizes(self):
        binSizeList = []
        for bin in self.binList:
            binSizeList.append(bin.getBalls())
        return binSizeList
```

A.3. Process.py

```
from Urn import Urn
import random
import matplotlib.pyplot as plt
import numpy as np
from scipy.stats import beta

class Process:
    # Most important:
    # Number of trials, starting number of Bins
    # Urnmodel
    # Big matrix in which we have all the values of balls per bin


    def __init__(self,trialNumber,p,f,startingNumberOfBins):
        self.trialNumber = trialNumber
        self.p = p
        self.f = f
        self.urn = Urn(p,f,startingNumberOfBins)
        self.bigList = []

    def process(self):
        for i in range(0,self.trialNumber):
            #self.bigList[i] = self.urn.getBinSizes()
            #print(self.urn.getBinSizes())
            self.bigList.append(self.urn.getBinSizes())
            self.urn.addIteration(random.random(),random.random())
```

```
def coupledProcess(self,randomNumberOne,randomNumberTwo):
    for i in range(0,self.trialNumber):
        self.bigList.append(self.urn.getBinSizes())
        self.urn.addIteration(randomNumberOne[i],randomNumberTwo[i])


def plot(self):
    maxNumberOfBins = len(self.bigList[-1])
    listAllNumbers = np.zeros((maxNumberOfBins,self.trialNumber))
    for index in range(0,self.trialNumber):
        currentList = self.bigList[index]
        for indexTwo in range(0,len(currentList)):
            listAllNumbers[indexTwo][index] = currentList[indexTwo]

    for i in range (0,maxNumberOfBins):
        string = "Bin "+ str(i+1)
        plt.plot(listAllNumbers[i], label = string)
    plt.legend(loc='upper left')
    plt.xlabel("Time")
    plt.ylabel("Number of balls")
    plt.title("Number of balls in different bins over time for f(x) =
        "+str(self.f)+" and p = " +str(self.p))
    plt.show(block=True)
    #plt.show()

def stackplot(self,functionName="default"):
    maxNumberOfBins = len(self.bigList[-1])
    listAllNumbers = np.zeros((maxNumberOfBins, self.trialNumber))
    toPlot = []
    for index in range(0, self.trialNumber):
        currentList = self.bigList[index]
        for indexTwo in range(0, len(currentList)):
            listAllNumbers[indexTwo][index] = currentList[indexTwo]
    labelsList = []
    for i in range(0, maxNumberOfBins):
        string = "Bin " + str(i + 1)
        labelsList.append(string)
        toPlot.append(listAllNumbers[i])
    plt.stackplot(range(1,self.trialNumber+1),listAllNumbers,labels =
        labelsList)
    plt.legend(loc='upper left')
    plt.xlabel("Time")
    plt.ylabel("Number of balls")
    plt.title("Number of balls in different bins over time\n for f(x)
        = " + functionName)
    plt.savefig("stackplot"+functionName+".png",bbox_inches = "tight")
    plt.show(block=True)

def hist(self):
    maxNumberOfBins = len(self.bigList[-1])
    listAllNumbers = np.zeros((maxNumberOfBins, self.trialNumber))
    for index in range(0,self.trialNumber):
        currentList = self.bigList[index]
        for indexTwo in range(0,len(currentList)):
            listAllNumbers[indexTwo][index] = currentList[indexTwo]
```

```
        finalList = listAllNumbers[:,-1]
        names = []
        values = []
        for i in range(0,len(finalList)):
            string = "Bin "+str(i+1)
            names.append(string)
            values.append(finalList[i])

        plt.title(
            "Number of balls in different bins over time\n for f = " + str
               (self.f) + " and p = " + str(self.p)+"\nafter " +str(self.
               trialNumber)+" number of trials.")
        plt.bar(names,values,color = ['#1b9e77', '#a9f971', '#fdaa48'])
        plt.show(block=True)

def f(x,gamma=1):
    return x^gamma

def summary_plot_polya(n,p,f,k):
    thisProcess = Process(n,p,f,k,)
    thisProcess.process()

    thisProcess.hist()
    thisProcess.plot()
    thisProcess.stackplot()




def g(x,gamma=1,a=0.0,b=1):
    return (b*x+a)**gamma


def Beta_distribution_test(bin1,bin2=1,a=0,fileName= "BetaPlot.png",
   n_process = 1000, n_balls = 5000):
    fractions = []
    def gfunc(x):
        return x+a
    for i in range(0,n_process):
        thisProcess = Process(n_balls,0,gfunc,2)
        for i in range(0,bin1-1):
            thisProcess.urn.binList[0].addBall()
        for i in range(0,bin2-1):
            thisProcess.urn.binList[1].addBall()
        thisProcess.process()
        bin_one = thisProcess.bigList[-1][0]*1.0/(n_balls+bin1+bin2)

        fractions.append(bin_one)
    plt.hist(fractions,density=True,label="Simulated density")
    x = np.linspace(0,1,100)

    plt.plot(x,beta.pdf(x,bin1+a,bin2+a),label="Density of Beta("+str(bin1
       +a)+", "+str(bin2+a)+")")
    plt.title("Fraction of total balls in bin 1 after "+str(n_balls)+"
       balls \nare added to the process, for "+str(n_process)+" simulated
```

```
        processes."+"\n(initial configuration: ("+str(bin1)+","+str(bin2)+"
        ) balls)")

    plt.xlabel("Limiting fraction")
    plt.ylabel("Probability density")
    plt.legend(loc=4)
    plt.tight_layout()
    plt.savefig(fileName,bbox_inches = "tight")
    plt.show()



def multiplesave(arrayOfInit,arrayOfA):
    for i in range(0,len(arrayOfInit)):
        initOne = arrayOfInit[i][0]
        initTwo = arrayOfInit[i][1]
        aVal = arrayOfA[i]
        thisFile = "BetaplotA_"+str(aVal)+"_"+str(initOne)+","+str(initTwo
            )+".png"
        Beta_distribution_test(initOne,initTwo,a=aVal,fileName = thisFile)
```

A.4. SimplexPlots.py

```
'''Functions for drawing contours of Dirichlet distributions.'''

# Author: Thomas Boggs

import numpy as np
import matplotlib.pyplot as plt
import matplotlib.tri as tri
from Process import Process

_corners = np.array([[0, 0], [1, 0], [0.5, 0.75**0.5]])
_triangle = tri.Triangulation(_corners[:, 0], _corners[:, 1])
_midpoints = [(_corners[(i + 1) % 3] + _corners[(i + 2) % 3]) / 2.0 \
              for i in range(3)]

def xy2bc(xy, tol=1.e-3):
    '''Converts 2D Cartesian coordinates to barycentric.
    Arguments:
        `xy`: A length-2 sequence containing the x and y value.
    '''
    s = [(_corners[i] - _midpoints[i]).dot(xy - _midpoints[i]) / 0.75 \
         for i in range(3)]
    return np.clip(s, tol, 1.0 - tol)

class Dirichlet(object):
    def __init__(self, alpha):
        '''Creates Dirichlet distribution with parameter `alpha`.'''
        from math import gamma
        from operator import mul
        self._alpha = np.array(alpha)
        self._coef = gamma(np.sum(self._alpha)) / \
                     reduce(mul, [gamma(a) for a in self._alpha])
    def pdf(self, x):
        '''Returns pdf value for `x`.'''
```

```
        from operator import mul
        return self._coef * reduce(mul, [xx ** (aa - 1)
                                          for (xx, aa)in zip(x, self._alpha
                                              )])
    def sample(self, N):
        '''Generates a random sample of size 'N'.'''
        return np.random.dirichlet(self._alpha, N)

def draw_pdf_contours(dist, border=False, nlevels=200, subdiv=8, **kwargs)
    :
    '''Draws pdf contours over an equilateral triangle (2-simplex).
    Arguments:
        'dist': A distribution instance with a 'pdf' method.
        'border' (bool): If True, the simplex border is drawn.
        'nlevels' (int): Number of contours to draw.
        'subdiv' (int): Number of recursive mesh subdivisions to create.
        kwargs: Keyword args passed on to 'plt.triplot'.
    '''
    from matplotlib import ticker, cm
    import math

    refiner = tri.UniformTriRefiner(_triangle)
    trimesh = refiner.refine_triangulation(subdiv=subdiv)
    pvals = [dist.pdf(xy2bc(xy)) for xy in zip(trimesh.x, trimesh.y)]

    plt.tricontourf(trimesh, pvals, nlevels, **kwargs)
    plt.axis('equal')
    plt.xlim(0, 1)
    plt.ylim(0, 0.75**0.5)
    plt.axis('off')
    if border is True:
        plt.hold(1)
        plt.triplot(_triangle, linewidth=1)

def plot_points(X, barycentric=True, border=True, **kwargs):
    '''Plots a set of points in the simplex.
    Arguments:
        'X' (ndarray): A 2xN array (if in Cartesian coords) or 3xN array
                       (if in barycentric coords) of points to plot.
        'barycentric' (bool): Indicates if 'X' is in barycentric coords.
        'border' (bool): If True, the simplex border is drawn.
        kwargs: Keyword args passed on to 'plt.plot'.
    '''
    if barycentric is True:
        X = X.dot(_corners)
    plt.plot(X[:, 0], X[:, 1], 'k.', ms=1, **kwargs)
    plt.axis('equal')
    plt.xlim(0, 1)
    plt.ylim(0, 0.75**0.5)
    plt.axis('off')
    if border is True:
        plt.hold(1)
        plt.triplot(_triangle, linewidth=1)

def get_point(trials,startNumberBalls):
    res = []
```

```
    def f(x):
        return x+2.5
    thisProcess = Process(trials ,0,f,len(startNumberBalls))
    for index in range(0,len(startNumberBalls)):

        thisUrn = thisProcess.urn.binList[index]
        for i in range(0,startNumberBalls[index]−1):
            thisUrn.addBall()
    thisProcess.process()
    for index in range(0,len(thisProcess.urn.getBinSizes())):
        res.append(1.0*thisProcess.urn.getBinSizes()[index]/(trials+
            startNumberBalls[index]))
    return res

def get_points(points, trials, startnumberballs):
    allPoints = []
    for i in range(0,points):
        currPoint = get_point(trials, startnumberballs)
        allPoints.append(currPoint)
    return allPoints

if __name__ == '__main__':
    f = plt.figure(figsize=(8, 6))
    alphas = [[1+2.5] * 3,
              [5+2.5,5+2.5,5+2.5],
              [1+2.5, 5+2.5, 20+2.5]]
    balls = [[1,1,1],
             [5,5,5],
             [1,5,20]]
    for (i, alpha) in enumerate(alphas):
        plt.subplot(2, len(alphas), i + 1)
        dist = Dirichlet(alpha)
        plot_points(np.array(get_points(500, 5000, balls[i])), barycentric
            =True)
        plt.subplot(2, len(alphas), i + 1 + len(alphas))
        draw_pdf_contours(dist)
    plt.savefig('dirichlet_plots.png',bbox_inches = "tight")
    print 'Wrote plots to "dirichlet_plots.png".'
```

A.5. CoupledProcess.py

```
import numpy as np
from Process import Process

def g1(x,gamma=1,a=4,b=1):
    return (b*x+a)**gamma

def g2(x,gamma=1,a=0,b=1):
    return (b*x+a)**gamma


def coupleProcesses(bin1, bin2):
    randomNumbersOne = np.random.rand(10000)
    randomNumbersTwo = np.random.rand(10000)
    processOne = Process(1000,0,g1,3)
    processTwo = Process(1000,0,g2,3)
    for i in range(0,bin1−1):
```

```
        processOne.urn.binList[0].addBall()
        processTwo.urn.binList[0].addBall()
    for i in range(0,bin2-1):
        processOne.urn.binList[1].addBall()
        processTwo.urn.binList[1].addBall()
    processOne.coupledProcess(randomNumbersOne, randomNumbersTwo)
    processTwo.coupledProcess(randomNumbersOne, randomNumbersTwo)
    processOne.stackplot()
    processTwo.stackplot()
```

Bibliography

[1] Noga Alon and Joel H. Spencer. *The Probabilistic Method.* Wiley Publishing, 4th edition, 2016.

[2] W. Brian Arthur, Yu.M. Ermoliev, and Yu.M. Kaniovski. Path-dependent processes and the emergence of macro-structure. *European Journal of Operational Research*, 30(3):294 – 303, 1987.

[3] Antar Bandyopadhyay and Debleena Thacker. A New Approach to Pólya Urn Schemes and Its Infinite Color Generalization. *arXiv e-prints*, art. arXiv:1606.05317, 2016.

[4] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[5] José M. Bernardo. Modern bayesian inference: Foundations and objective methods. In Prasanta S. Bandyopadhyay and Malcolm R. Forster, editors, *Philosophy of Statistics*, volume 7 of *Handbook of the Philosophy of Science*, pages 263 – 306. 2011.

[6] David Blackwell and James B. MacQueen. Ferguson distributions via polya urn schemes. *The Annals of Statistics*, 1(2):353–355, 1973.

[7] Fan Chung, Shirin Handjani, and Doug Jungreis. Generalizations of polya's urn problem. *Annals of Combinatorics*, 7(2):141–153, 2003.

[8] Anirban DasGupta. *Markov Chains and Applications*, pages 343–378. Springer, 2010.

[9] Eleni Drinea, Alan Frieze, Michael Mitzenmacher, and Michael Mitzenmacher. Balls and bins models with feedback. In *Proceedings of the Thirteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '02, pages 308–315. Society for Industrial and Applied Mathematics, 2002.

[10] F. Eggenberger and G. Pólya. Über die statistik verketteter vorgänge. *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik*, 3(4):279–289, 1923.

[11] Allan Gut. *Probability: A Graduate Course.* Springer, 2005.

[12] David Heath and William Sudderth. De finetti's theorem on exchangeable variables. *The American Statistician*, 30(4):188–189, 1976.

[13] Svante Janson. A.s. convergence for infinite colour Pólya urns associated with random walks. *arXiv e-prints*, art. arXiv:1803.04207, 2018.

[14] Konstantin Khanin and Raya Khanin. A probabilistic model for the establishment of neuron polarity. *Journal of mathematical biology*, 42:26–40, 2001.

[15] Leonid Koralov and Yakov G. Sinai. *Theory of Probability and Random Processes.* Springer, 2012.

[16] Hosam Mahmoud. *Polya Urn Models.* Chapman & Hall/CRC, 1 edition, 2008.

[17] Cécile Mailler and Jean-François Marckert. Measure-valued pólya urn processes. *Electron. J. Probab.*, 22:33 pp., 2017.

[18] Roberto Oliveira. Balls-in-bins processes with feedback and brownian motion. *Combinatorics, Probability Computing*, 17:87–110, 2008.

[19] Roberto Oliveira and Joel Spencer. Avoiding defeat in a balls-in-bins process with feedback. *arXiv e-prints*, art. math/0510663, 2005.

[20] Robin Pemantle. A survey of random processes with reinforcement. *Probab. Surveys*, 4: 1–79, 2007.

[21] M.V. Simkin and V.P. Roychowdhury. Re-inventing willis. *Physics Reports*, 502(1):1 – 35, 2011.

[22] David Williams. *Probability with Martingales*. Cambridge University Press, 1991.