“Luigi Bocconi” Commercial University in Milan
Department of Decision Sciences

XX PhD Course in Statistics

New Urn Approaches to Shock and Default Models

Prof. Dr. Jürg Hüsler
Institut für mathematische Statistik und Versicherungslehre
University of Bern (CH)

Prof. Pietro Muliere
Dipartimento di Scienze delle Decisioni
Bocconi University in Milan (IT)

Pasquale Cirillo
ID: 1003722

[Ego sum qui sum]

Summary

We introduce some new urn approaches to shock and default models, with a double aim: to reproduce and generalized some existing results and to present new ones.

The interest for shock and default models is due to their relevance for studying everyday life events, where the concept of "default", in its broader sense, is well-known.

Even if they are born for apparently different purposes, shock and default models can be considered as the two sides of the same coin.

Shock models are a particular class of models in which a system is randomly subject to different shocks of random magnitude. Sooner or later, these shocks make the system fail (think, for example, of an electrical system that breaks down after an energy jolt), so one is interested in studying the probability of default and the time to default.

Default models, on the contrary, analyze the so-called failing systems. A failing system is a system whose probability of failure is not negligible in a fixed time horizon. In this case, the stress is less on shocks and more on the system itself, but clearly the

final scope is in common with shock models. A failing system that everyone knows is, unfortunately, human life.

In this thesis, we present four different models, providing both the theoretical and the applicative aspects: a generalized extreme shock model via triangular urns, an urn-based shock model that incorporates obsolescence, a multidimensional default urn scheme and an urn chain model for studying interacting failures.

Acknowledgements

páthei máthos kyríos échein

First of all I am really grateful to Professor Dr. Jürg Hüsler from Bern University, Switzerland. His hospitality and helpfulness have been wonderful and fundamental for the development of this thesis. He has proposed me several interesting open problems, among which shock models. The first part of this thesis would not exist without his support.

In the same way I have to thank Prof. Pietro Muliere, the coordinator of my PhD Course at Bocconi University, for his continuous support and suggestions and for presenting me urn models and their usefulness. His supervision has been fundamental for the development of the second part of this work.

Then I want to express all my gratitude to Prof. Carlo Bianchi, from Pisa University, for being a constant presence in my academic life since I attended his econometric class more than 5 years ago. This year two of our papers have been published in important journals and I am really happy to be one of his coauthors, together with Prof. Pietro Vagliasindi, from the University of Parma, and Prof. Mauro Gallegati, from the Polytechnic University of Ancona.

I thank Prof. Philippe Flajolet, from INRIA France, for answering by mail to my questions about analytic combinatorics. His expert hints have been precious for my research.

I am very grateful to my PhD colleagues Gaia, Stefano P., Stefano F., Djil and Alberto for helping me, with their kindness, in facing this last three years of hard study.

Most of all, I thank and love my family for being always there when I need them and for supporting me in all my activities. My father Gennaro, my mother Elia and my Sister Elena are the main responsible of all my successes.

I cannot forget my Aunt Patrizia and my grandmas Elena and Noemi, for their love and trust. And my nephew Achille for being that “little lovely dwarf”.

Last but not least, I want to thank my girlfriend Viola. She is always present and her love and respect are the greatest things I have gained in the last years. How can I forget her funny inquisitive face, when she pretends to be interested about my tedious discussions on urn processes?

Contents

List of figures

Preface

Urn processes (or urn models or urn schemes) constitute a very large family of probabilistic models in which the probability of certain events is represented in terms of sampling, replacing and adding balls in one or more urns or boxes.

Urn problems have been an important part of the theory of probability since the publication of the *Ars conjectandi* by Jakob Bernoulli in 1713 [7]. Their most interesting characteristic is the possibility of simplifying complex probabilistic ideas, making them intuitive and concrete, and yet guaranteeing a good level of abstraction, that allows for general results.

As we will see in the next chapters, our choice of urn processes as a probabilistic tool is mainly due to the following reasons:

1. They are particularly suitable, thanks to their efficiency, for studying chance experiments, especially when these are characterized by countable spaces.

2. They represent an excellent way to describe the concept of "random choice", which can be easily tested a posteriori, but which is in principle "not accessible to an absolute mathematical definition" (see [7], pag. XXV).

3. Simple urns can be easily compounded into new ones in order to study more complex problems: for example, think about hyperurns and urn chains (see [50] for details), or about random partitions and other cutting-edge problems (see,

for example, [64]). This flexibility is probably one of their most interesting characteristics.

4. Urn schemes have as powerful as elegant combinatorial properties, that allow for general, complex results in a rather concise form.

5. There are many relationships and isomorphisms between urn models and other well-known mathematical objects. All this give the possibility to the researcher of switching from one approach to the other at her/his convenience. The isomorphism between special urn schemes and particular differential systems, as described in [34], is one of the milestones of the present thesis.

6. Urns are very useful objects in simulations, given their natural connections with sampling schemes. There are a lot of examples in this field and a very interesting one is presented in [46]. Moreover, as [6] puts it, "the term *simulation* can be interpreted as the statistical equivalent to the basic mathematical concept of *isomorphism* which is intrinsically associated with urn models". This explains the wide use of urns in sampling algorithms.

As discussed in [6] and [50], the prototype of urn processes is the well-known Polya urn, developed at the beginning of the last century to model the diffusion of infectious diseases. It represents one of the simplest ways to generate beta exchangeable random variables and it is based on the concept of reinforcement. Moreover, its multidimensional version, as shown in [14], is a very useful tool to obtain the fundamental Dirichlet distribution, an essential tool for Bayesian statistics.

Polya urn represents one of the basic pillars of the present thesis. In fact, the aim of this research work is to apply urn processes to give alternative modelization to well-known problems. In particular we will propose four new urn approaches to shock and default models.

The interest for shock and default models is due to their relevance for studying everyday life events, where the concept of “default”, in its broader sense, is common and important. A firm that goes bankrupt, a bridge that breaks down, a financial crisis are events that have a strong impact on human life. The same human life can be seen as a system subject to default.

Shock and default models can be considered as the two sides of the same coin. They simply use different perspectives to study the same subject.

Shock models are a particular class of models in which a system is randomly subject to different shocks of random magnitude. Sooner or later, these shocks make the system fail (think, for example, of an electrical system that breaks down after an energy jolt), so one is interested in studying the probability of default and the time to default.

Default models, on the contrary, analyze the so-called failing systems. A failing system is a system whose probability of failure is not negligible in a fixed time horizon. In this case, the stress is less on shocks and more on the system itself, but clearly the final scope is in common with shock models. An example of failing system are firms.

This thesis can be divided into two parts: the first one presents two new urn approaches to shock models; while the second is devoted to new urn-based default models, with a particular stress on failing systems.

The use of urns essentially has three important advantages in the analysis of shock and default models:

1. The modelization is rather intuitive and immediate, even if the underlying calculations are not;

2. Urns can be considered a first attempt to study shocks and defaults from a Bayesian nonparametrics point of view, in that they allow the researcher to introduce her/his prior knowledge into the analysis[1];

3. The flexibility of the adopted urn schemes is a very useful characteristic for simulations and empirical studies.

The division of the thesis into two parts is also reproduced if we look at the methodologies we have used for our studies.

---

1 We will come back to this point later in the thesis.

The first part is indeed basically treated by the means of the urn analytic approach introduced in [33], [34] and further developed in [67]. Analytic urns represent a very interesting way for dealing with balanced reinforcement matrices, since they allow to analytically solve them, on the basis of one fundamental theorem, that builds up an interesting isomorphism between urn processes and ordinary differential systems.

The second part, on the contrary, is developed starting from the more traditional Polya urn, even if we introduce extra structures in order to deal with default models.

In a nutshell, this can be considered a thesis in combinatorics, both in the classical (see [1]) and the new analytical (see [35]) sense.

Every part is made up of three chapters: the first two introduce new theoretical models, while the third is devoted to an application on actual data, in order to test the goodness of the proposed modelization.

To go in more detail we have:

Chapter 1 We introduce a new intuitive urn approach to generalized shock models. In particular, we aim to indirectly model the moving risky threshold of generalized extreme shock models introduced in [44]. The basic idea is to link the colors of the balls in the urn with the levels of risk a system can bear. Hence we model the evolution of the process using a special triangular reinforcement matrix, that we solve using the analytic approach;

Chapter 2 We describe an alternative modelization to the urn-based version of extreme shock models, using a ruinous urn, as we have called it. This new scheme has been inspired by the interesting cannibal urn described in [65]. The aim is to model, in a pessimistic way, a system subject to shocks of random magnitude, starting from the assumption that, however, the system will inevitably fail because of obsolescence;

Chapter 3 We here apply the model developed in Chapter 1 to study firms' default, assuming that the failure of a firm is given by a series of subsequent shocks. To prove the goodness of our results, we compare them with one of the actual market benchmark, Altman's Z-score;

Chapter 4 We introduce a multidimensional urn model for studying failing systems. The main purpose when studying a failing system is to calculate the probability of default and the distribution of the number of failures that may occur during the observation period. Here we introduce a new way for calculating the probabilities of joint defaults in k different homogeneous groups of defaulting systems, when each group is characterized by some sort of rating, that's some external information about its reliability.

Chapter 5 Starting from Doksum construction of neutral to the right processes, we develop a recursive model constructed by the means of interacting urns. The aim is to introduce a stronger dependence among defaults with respect to the multidimensional urn model of the previous chapter.

Chapter 6 We finally apply our default models to data and compare them with an effective benchmark, in order to study credit defaults.

In a word, this thesis can be considered a sort of puzzle, that tries to address different problems by proposing new possible roads.

It is now time to conclude this preface. In doing so, I would like to quote the famous Greek poet Challimacus (Antologia Palatina, XII 43) and remember to you, the Reader, that

méga biblíon méga kakón.

Part I

Urn-based Shock Models

Chapter 1
Generalized extreme shock models via triangular urns

In this chapter[2] we introduce a new intuitive approach to generalized shock models using urn processes. This allows us to indirectly model the moving risky threshold of generalized extreme shock models introduced in [46]. The basic idea is to link the colors of the balls in the urn with the levels of risk a system can face. Then we model the evolution of the process using a triangular reinforcement matrix.

Using the analytic approach proposed in [35], we explicitly derive probabilities, moments and limit laws for the different types of balls, that is for the distinct levels of risk. We also calculate the limit law for the urn composition. Our proposed model can be considered as a way to incorporate GESM's in a Bayesian framework. In fact, urn processes are an important tool of Bayesian nonparametrics, in the way they allow to compute posterior distributions without an explicit knowledge of priors. Moreover, we show that our process generates asymptotically exchangeable sequences of random variables, suggesting interesting links between our urn and the classical Polya.

1.1 Shock Models: an Introduction

Shock models are a particular class of models in which a system is randomly subject to different shocks of random magnitude.

In the literature (see [46] for example) there are essentially three distinct types of shock models: cumulative shock models, in which the failure of the system is due

[2] This chapter is based on the paper "An Urn Approach to Generalized Extreme Shock Models" by P.Cirillo and J.Hüsler, that has been presented at the EVA 2007 Conference in Bern as poster contribution.

to a cumulative effect; extreme shock models, whose default is caused by one single extreme shock; and mixed model, that combines the other two types.

In [46], the authors introduce several generalizations, among which we find generalized extreme shock models (GESM), that are the starting point of the present chapter. In this kind of models, one assumes that a system can be harmed by some large but non fatal shock, which jeopardizes the load the system can take without completely damaging it. This makes the system more sensitive to subsequent shocks, so that a less extreme shock can be sufficient for the system to break down.

Anyway, before introducing our new findings, let us now briefly review the main results concerning shock models.

1.1.1 Basic notation

Consider a family $\{Z_i, U_i\}_{i\geq 0}$ of nonnegative i.i.d. two-dimensional random vectors, such that Z_i represent the intensity of the i-th shock and U_i is the time between the $(i-1)$-th and the i-th shock. Let $S_0 = T_0 = 0$ and, $\forall i \geq 1$, define $S_n = \sum_{i=1}^{n} Z_i$ - the accumulated shock load at time n - and $T_n = \sum_{i=1}^{n} U_i$ - the amount of time that has elapsed after n shocks.

Now set $\mu_z = EZ_1$, $\mu_U = EU_1$, $\sigma_Z^2 = VarZ$ and $\sigma_U^2 = VarU$.

1.1.2 Cumulative Shock Models

In these models one defines the first passage time process $\{v(t)\}_{t \geq 0}$ by $v(t) = \min\{n : S_n > t\}$. The failure time is a random variable $T_{v(t)}$, that is shown to be a stopped random walk (see [43]).

Using renewal theory, we get the following immediate property:

$$\frac{v(t)}{t} \overset{a.s.}{\to} \frac{1}{\mu_Z} \text{ for } t \to \infty.$$

Moreover, as in [43], it is possible to state another interesting result:

Theorem 1 *If μ_Z and μ_U are finite, then*

$$\frac{T_{v(t)}}{t} \to_{a.s.} \frac{\mu_U}{\mu_Z} \textit{ as } t \to \infty.$$

Moreover, if $\sigma_Z^2 < +\infty$, $\sigma_U^2 < +\infty$ and $\sigma = Var(\mu_Z U - \mu_U Z) > 0$,

$$\frac{T_{v(t)} - \frac{\mu_U}{\mu_Z}}{\sqrt{\mu_Z^{-3} \sigma^2 t}} \overset{d}{\to} N(0,1) \textit{ as } t \to \infty.$$

1.1.3 Extreme Shock Models

As far as extreme shock models are concerned, the relevant stopping time is $\tau(t) = \min\{n : Z_n > t\}$.

A fundamental assumption for extreme shock models is that $p_t = P(Z_1 > t) = 1 - F(t) \to 0$ as $t \to z_F$, where $z_F = \{z : F(z) < 1\}$ is the right endpoint for F and F is the distribution function of Z.

Given that $\tau(t)$ is geometric with mean $\frac{1}{p_t}$, it can be easily shown that

$$p_t\tau(t) \xrightarrow{d} Exp(1) \text{ as } t \to z_F.$$

All this shows that the stopping times behave in two different ways for cumulative and extreme shock models. However, also in this case, $T_{\tau(t)}$ is a stopped random walk.

Theorem 2 (Gut and Hüsler - [45]**)** *If* $p_t = P(Z_1 > t) = 1 - F(t) \to 0$ *as* $t \to z_F$ *and* μ_U *is finite, then*

$$p_t T_{\tau(t)} \to^d \exp(\mu_U) \textit{ as } t \to z_F.$$

1.1.4 Mixed models

Imagine a system fails when one of the following situations appears: 1) the cumulative shocks reach some particular threshold level; 2) one big shock happens. In this case the number of shocks at default is equal to $\min\{v(t), \tau(t)\}$.

To avoid that one of the two stopping times dominates the other, it is crucial that p_t and $\frac{1}{t}$ have the same order of magnitude, that is $p_t \to 0$. In addition, assume μ_Z is finite[3].

Let q_t represent the $(1-\frac{\theta}{t})-$th quantile for the distribution of Z for some $\theta > 0$. It is evident that $q_t \to z_F$ when $t \to \infty$ and that $q_t = o(t)$ considering that the mean shock size is finite (see [44]).

[3] This is a necessary assumption in order to merge the two models.

Define the new stopping time

$$\tau_\lambda(t) = \min\{n : Z_n > \lambda_t\},\ t \geq 0.$$

Then, it is possible to get the appropriate number of shocks until failure, that is $\kappa(t) = \min\{\upsilon(t), \tau_\lambda(t)\}$. Thus, the failure time becomes $T_{\kappa(t)}$.

Theorem 3 *If $\mu_Z < +\infty$ and $\mu_U < +\infty$, then*

$$\frac{\kappa(t)}{t} \xrightarrow{d} V \textit{ as } t \to \infty,$$

where

$$f_V(x) = \theta e^{-\theta x} \ \ 0 < x < \frac{1}{\mu_Z}$$
$$P(V = \frac{1}{\mu_Z}) = e^{-\frac{\theta}{\mu_Z}}$$

or, similarly,

$$F_V(x) = \begin{cases} 1 - e^{-\theta x} \ 0 < x < \frac{1}{\mu_Z} \\ 1 \qquad\qquad\quad x \geq \frac{1}{\mu_Z} \end{cases}.$$

Additionally:

$$\frac{T_{\kappa(t)}}{t} \xrightarrow{d} \mu_U V \textit{ as } t \to \infty$$
$$\frac{S_{\kappa(t)}}{t} \xrightarrow{d} \mu_Z V \textit{ as } t \to \infty$$
$$\frac{Z_{\kappa(t)}}{t} \xrightarrow{p} 0 \textit{ and } \frac{\max_{1 \leq k \leq \kappa(t)} Z_k}{t} \xrightarrow{p} 0 \textit{ as } t \to \infty.$$

The three basic models we have presented have been generalized in several ways, using for example delayed sums or multiple critical shocks. In [56] and [46], one can find suitable references about these generalizations.

As already said, a generalization of the extreme shock model is the starting point of our urn process. Next section is just devoted to the theoretical description of generalized extreme shock models.

1.2 Generalized Extreme Shock Models

Following [46], we assume that large but not fatal shocks may effect system's tolerance to subsequent shocks. To be more exact, for a fixed t, a shock Z_i can damage the system if it is larger than a certain boundary value $\beta_t < t$. As long as $Z_i < t$ the system does not fail. The crucial hypothesis is the following: if a first nonfatal shock comes with values in $[\beta_t, t]$ the maximum load limit of the system is no more t, but decrease to $\alpha_t(1) \in [\beta_t, t]$, since the system has been damaged. At this point, if another large but not too strong shock occurs in $[\beta_t, \alpha_t(1)]$, the new crucial threshold is lowered again to $\alpha_t(2) \in [\beta_t, \alpha_t(1)]$ and so on until the system fails. We could call all this "risky threshold mechanism". It is obvious to say that, $\forall t$

$$t = \alpha_t(0) \geq \alpha_t(1) \geq \alpha_t(2) \geq ... \geq \beta_t. \tag{1.1}$$

We define the stopping time

$$\tau(t) = \min \{n : Z_n \geq \alpha_t(L_t(n-1))\}$$

with

$$L_n(t) = \sum_{i=1}^{n} 1_{\{Z_i \geq \beta_t\}} \text{ and } L_t(0) = 0.$$

When the $\alpha_t(k)$'s are all the same, a generalized extreme shock model simply behaves as an extreme shock one: it suffices to replace Z_i with $(Z_i - \beta_t)_+$.

More interesting is the case for different levels $\alpha_t(k)$.

Let $\alpha_t(k)$ be nonrandom: the distribution of $\tau(t)$ can be derived using the independence of the Z_i's. In particular, the event $\{\tau(t) > m\}_{m\geq 1}$ corresponds to "the union of the disjoint events that there are exactly j nonfatal shocks up to the index m" (see [46]), that's $L_t(m) = j$ for $0 \leq j \leq m$. Setting $\prod_{k=0}^{-1} = 1$, we have

$$P\left\{\tau(t) > m\right\} = \sum_{j=0}^{m} \binom{m}{j} F^{m-j}\left(\beta_t\right) \prod_{k=0}^{j-1} \left[F(\alpha_t(k) - F(\beta_{t)}\right].$$

As far as the asymptotic behavior of the model, it clearly depends on three fundamental elements: the distribution F and the two boundaries $\alpha_t(k)$ and β_t. The sequence $\{\alpha_t(k)\}_{k>0}$, for example, can be defined in different ways: it could be parametrized as $\alpha_t(k) = t(1 - \alpha k^{\delta}) \vee \beta_t$, which is linear for $\delta = 1$, or some other functional forms. However, it is possible to identify two major cases of interest.

Consider

$$\lim_{t\to\infty} \frac{\overline{F}(\alpha_t(\infty))}{\overline{F}(\beta_t)} = c \geq 0,$$

where $\overline{F}(x) = 1 - F(x)$. If $c > 0$, when a shock exceeds β_t, the chance of exceeding one of the $\alpha_t(k)$'s is not asymptotically negligible. In particular, it is possible to state the following result:

Theorem 4 (Gut and Hüsler - [46]**)** *Assume that* $\frac{\overline{F}(\alpha_t(\infty))}{\overline{F}(\beta_t)} \to c_k$ *for every* k *and* $\overline{F}(\beta_t) \to 0$. *Then*

$$P\left(\bar{F}(\beta_t)\tau(t) > m\right) \to \sum_{j\geq 0} e^{-m}\frac{m^j}{j!}\prod_{k=0}^{j-1}(1-c_k) \text{ as } t \to \infty.$$

If on the contrary $c = 0$, the boundary β_t has no influence on the extreme shocks exceeding the $\alpha_t(k)$ levels and the previous theorem does not hold, since $\tau(t)$ is not properly normalized with $\overline{F}(\beta_t)$. Anyway another theorem helps us in solving the problem:

Theorem 5 (Gut and Hüsler - [46]**)** *Let* $\overline{F}(\alpha_t(\infty)) \to 0$ *as* $t \to \infty$. *Assume that:* $\frac{\overline{F}(\alpha_t(k))}{\overline{F}(\alpha_t(\infty))} \to a_k$ *holds for every* k *with* $\frac{\overline{F}(\alpha_t(\infty))}{\overline{F}(\beta_t)} \to 0$. *If* $a_{k \to 1}$ *as* $k \to \infty$, *then*

$$P\left(\bar{F}(\alpha_t(\infty))\tau(t) > m\right) = e^{-m}.$$

At this point, the asymptotic behavior of the failure time $T_{\tau(t)}$ is described as follows:

Theorem 6 (Gut and Hüsler - [46]**)** *Assume that either the conditions of Theorem 4 or 5 hold. Let* μ_U *be defined. Then*

$$p_t T_{\tau(t)} \xrightarrow{d} \mu_U V \text{ as } t \to \infty,$$

where V *is the random variable with distribution given by Theorem 4 or 5 respectively.*

In this work our aim is to model the risky threshold mechanism without explicitly defining the moving threshold $\alpha_t(k)$. The idea is to create three different risk

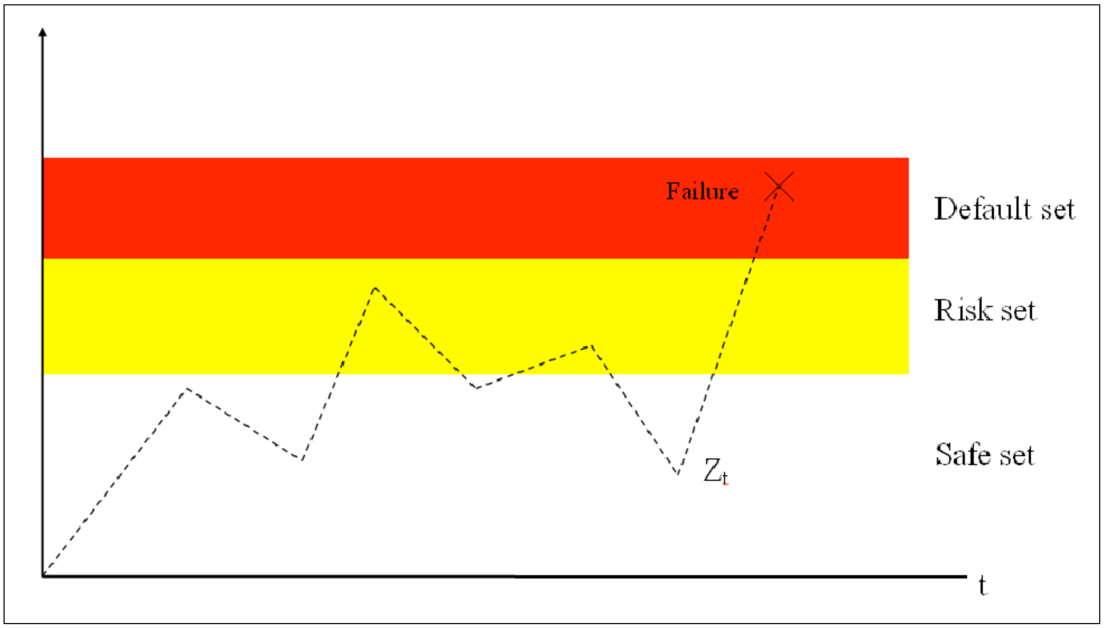

Figure 1.1: Example of the process with the three risk levels.

areas for the system - no risk or safe, risky and default - and to work with the probability for the process to enter each area. If every time the process enters the risky area the probability of failing increases, we can consider such a modelization a sort of intuitive approach to GESM's, getting around the definition of the moving threshold. In some sense, reinforcing the probability for the system to fail is like making the risky threshold move down. To do this we are going to use a special triangular urn process, discussed in the next section. This approach permits to use the good combinatorial properties of urns to study shock phenomena. Figure 1.1 gives an idea of the process.

1.3 The Urn Model

To model GESM's we introduce a particular urn process, completely characterized by its initial composition and its reinforcement matrix. In particular we make use of a Polya-like urn, with which we aim to replicate and generalize the results in [46].

Assume we have an urn containing balls of three different colors: x, y, and w; each color represents a possible state of risk for the process. In details, x-balls are related to the safe state, y-balls to the risky state and w-balls embody the default state.

Therefore, the evolution of the process is described by the following steps:

1. At time n a ball is sampled from the urn. The probability of sampling a particular ball obviously depends on the urn composition at time $n-1$;

2. According to the color of the ball, the process enters (or remains in) one of the three states of risk. For example, if the sampled ball is of type x, the process is in a safe state, while it fails if the chosen ball is w;

3. The urn is then reinforced according to its reinforcement matrix, that fully describes the sampling scheme.

1.3.1 The reinforcement matrix

The reinforcement matrix is a fundamental component of the urn process, since it determines how the process behaves each time a ball is sampled from the urn. In order to model the positive dependence between the risky and the default states, we will make use of a particular reinforcement rule, represented by the following balanced matrix, which is constant over time:

$$RM = \begin{array}{c} \\ x \\ y \\ w \end{array} \begin{array}{c} \begin{array}{ccc} x & y & w \end{array} \\ \begin{bmatrix} \theta & 0 & 0 \\ 0 & \delta & \lambda \\ 0 & 0 & \theta \end{bmatrix} \end{array}, \text{ where } \lambda = \theta - \delta > 0 \tag{1.2}$$

This matrix says that (see also figure 1.2 for a graphical example):

1. each time an x-ball is sampled, θ balls of type x are reintroduced in the urn;

2. if the sampled ball is y, the urn is reinforced with δ y-balls and λ w-balls (and this models dependence);

3. if a w-ball is picked up, θ balls of the same color are put in the urn[4].

In particular we will also assume two simple conditions:

1 *Let $\theta\delta \neq 0$, not to have degenerate cases.*

2 *Let $\lambda = \theta - \delta \geq 0$, as an expression of the positive dependence between y and w balls.*

1.3.2 The analytic approach

The so-called analytic approach to urn models was first introduced in [35], [36] and then further developed in [70]. It represents a very interesting way for dealing with

4 In reality, the reinforcement θ is fictitious: it is only necessary for calculations. In fact when a $w-$ball is sampled the system fails and the process stops.

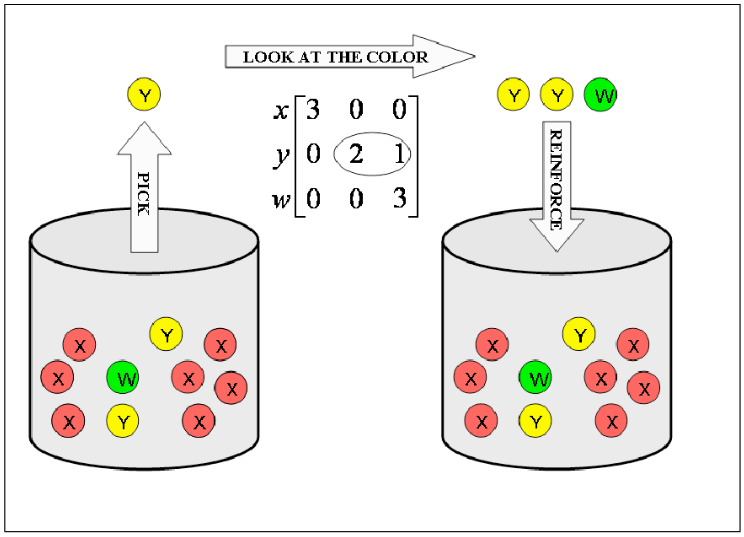

Figure 1.2: Example of triangular urn process for $a_0 = 6$, $b_0 = 3$, $c_0 = 1$, $\theta = 3$ and $\delta = 2$.

balanced urns (for another approach based on martingales see [50]), since it allows to analytically solve them, on the basis of one fundamental theorem, which is the *urn isomorphism theorem* presented in [36].

We here mention an adapted version of the main theorem in [36], which is based on the following definition.

Definition 1 *Given a* 3×3 *triangular urn model, whose reinforcement matrix is*

$$M = \begin{bmatrix} \alpha & \beta & \chi \\ 0 & \pi & \gamma \\ 0 & 0 & \sigma \end{bmatrix}, \quad \alpha, \beta, \sigma, \pi, \chi, \gamma \in \mathbb{Z},$$

the associated differential system is the ordinary differential system

$$\sum = \begin{cases} \dot{x} = x^{\alpha+1} y^{\beta} w^{\chi} \\ \dot{y} = y^{\pi+1} w^{\gamma} \\ \dot{w} = w^{\sigma+1} \end{cases},$$

where $\dot{x}$ *represents differentiation with respect to the independent variable time* z.

At this point we can state the following.

Theorem 7 (Flajolet, Gabarro, Pekari - [35]**)** *Consider a balanced urn of the form M, as above, which is initialized with a_0 balls of color x, b_0 balls of type y and c_0 for w. Let x_0, y_0 and w_0 be three complex numbers and assume that at least one of them is different from zero. Let $h_x(z; x_0, y_0, w_0)$, $h_y(z; x_0, y_0, w_0)$ and $h_w(z; x_0, y_0, w_0)$ be the solutions to the associated differential system $\sum$ with initial conditions x_0, y_0 and w_0. The generating function of urn histories satisfies (for z small enough)*

$$H(z; x_0, y_0, w_0) = h_x(z; x_0, y_0, w_0)^{a_0}\, h_y(z; x_0, y_0, w_0)^{b_0}\, h_w(z; x_0, y_0, w_0)^{c_0}\,.$$

So, following [36], it is possible to establish a fundamental isomorphism between discrete-time balanced urn processes and particular ordinary differential systems, that are nonlinear, autonomous and characterized by a simple monomial form. Such a relationship gives the possibility of analytically solving the urn process in finite terms, finding out the underlying generating function.

In our case, the reinforcement matrix RM determines the following ODE system:

$$\sum = \left\{ \begin{array}{l} \dot{x} = x^{\theta+1} \\ \dot{y} = y^{\delta+1} w^{\lambda} \\ \dot{w} = w^{\theta+1} \end{array} \right. \quad \text{with i.c.} \left\{ \begin{array}{l} x(0) = x_0 \\ y(0) = y_0 \\ w(0) = w_0 \end{array} \right., \qquad (1.3)$$

which is explicitly solvable.

In fact, the x and w components are immediately obtainable by simple integration:

$$x(t) = x_0(1 - \theta x_0^\theta t)^{-\frac{1}{\theta}} \tag{4}$$

$$w(t) = w_0(1 - \theta w_0^\theta t)^{-\frac{1}{\theta}}. \tag{5}$$

Then, noting that $\dot{y}y^{-\delta-1} = w^\lambda$, the solution is:

$$y(t) = y_0(1 - y_0^\delta \left(w_0^{-\delta} - \left[w_0(1 - \theta w_0^\theta t)^{-\frac{1}{\theta}} \right]^{-\delta} \right)^{-\frac{1}{\delta}}.$$

Given these solutions, we can then state the following proposition.

Proposition 8 *Consider an urn process with a reinforcement matrix RM as in 1.2, that satisfies conditions 1 and 2, and with an initial composition (a_0, b_0, c_0) of balls. The 4-variables generating function of urn histories is:*

$$H(z; x, y, w) = x^{a_0} y^{b_0} w^{c_0} (1 - \theta x^\theta z)^{-\frac{a_0}{\theta}} (1 - \theta w^\theta z)^{-\frac{c_0}{\theta}} \left(1 - y^\delta w^{-\delta} \left(1 - (1 - \theta w^\theta z)^{\frac{\delta}{\theta}} \right) \right)^{-\frac{b_0}{\delta}}.$$

Proof. The proof of the proposition is an immediate consequence of the multidimensional version of the urn isomorphism theorem, given the ODE system we have defined. ■

The possibility of having an explicit form for $H(z; x, y, w)$ is an useful result. It allows us to directly compute the closed-form expressions of the factorial moments of the number of x,y and w-balls in the urn by simply differentiating the generating function. Then, using the Stirling numbers of the second kind $\left\{ {k \atop j} \right\}$ (for a reference, see [37]), it is straightforward to recover the k-th power moment of N as $E\left[N^k\right] = \sum_j \left\{{k \atop j}\right\} E\left[(N)_j\right]$.

Proposition 9 *Let X_n, Y_n and W_n represent the number of x, y and w balls in the urn at time n. Their moments show to be hypergeometric functions, that is finite linear combinations of product and quotients of Euler Gamma functions. In particular:*

$$\begin{aligned}
E\left[X_n\right] &= \frac{a_0}{t_0}(t_0 + n\theta), \\
E\left[Y_n\right] &= b_0 \frac{\Gamma\left(\frac{t_0}{\theta}\right)}{\Gamma\left(\frac{t_0+\delta}{\theta}\right)} n^{\frac{\delta}{\theta}} + O(n^{\frac{\delta}{\theta}-1}), \\
E\left[W_n\right] &= \left[(t_0 - a_0)\frac{\lambda}{\theta}\right] \frac{\Gamma\left(\frac{t_0}{\theta}\right)}{\Gamma\left(\frac{t_0+\lambda}{\theta}\right)} n^{\frac{\delta}{\theta}} + O(n^{\frac{\delta}{\theta}-1}),
\end{aligned}$$

and, in general, for the moments of order l

$$\begin{aligned}
E\left[(X_n)_l\right] &= \theta^l \frac{\left(\frac{a_0}{\theta}\right)^{(l)}}{\left(\frac{t_0}{\theta}\right)^{(l)}} n^l + O(n^{l-1}), \\
E\left[(Y_n)_l\right] &= \delta^l \frac{\left(\frac{b_0}{\delta}\right)^{(l)} \Gamma\left(\frac{t_0}{\theta}\right)}{\Gamma\left(\frac{t_0+l\delta}{\theta}\right)} n^{l\frac{\delta}{\theta}} + O(n^{(l-1)\frac{\delta}{\theta}}), \\
E\left[(W_n)_l\right] &= \lambda^l \frac{\left(\frac{t_0-a_0}{\theta}\right)^{(l)} \Gamma\left(\frac{t_0}{\theta}\right)}{\Gamma\left(\frac{t_0+l\lambda}{\theta}\right)} n^{l\frac{\delta}{\theta}} + O(n^{(l-1)\frac{\delta}{\theta}}),
\end{aligned}$$

where $t_0 = a_0 + b_0 + c_0$, $\lambda = \theta - \delta$ and $(\cdot)^{(n)}$ represents the standard Pochhammer formula. Since this values are obtained through asymptotic approximations, their accuracy increases as $n \to \infty$.

Proof. We split the proof into two parts

1. We consider x balls. It quite easy to see that they are driven by a standard 3×3 Polya urn. In fact, since there is no interaction between the x-balls and the others, we can aggregate the reinforcement of y and w balls and then obtain a standard $\begin{bmatrix} \theta & 0 & 0 \\ 0 & \theta & 0 \\ 0 & 0 & \theta \end{bmatrix}$ reinforcement urn. Hence the result immediately follows from corollary 2.1 in [70], that's by using standard Lagrange inversion.

2. Since for y and w balls the methodology is the same, we will prove the result only for Y_n.

   Set $C_n = \frac{\Gamma(n+1)}{\left(\frac{t_0}{\theta}\right)^{(n)}}$. Taking derivatives of the multivariate generating function, one has

$$E\left[Y_n(Y_n-1)\cdots(Y_n-l+1)\right] = C_n\left[z^n\right]\frac{\partial^l H}{\partial y^l}|_{x=1,w=1},$$

   where $[z^n]$ represents the standard notation for the operation of coefficient extraction[5].

   Thanks to some simple but tedious manipulations and a small recurrence, it is possible to discover that

$$E\left[(Y_n)_l\right] = \left(\frac{b_0}{\delta}\right)^{(l)}(1-\theta z)^{-\frac{(t_0+l\theta)\delta}{\theta}} + \left(\frac{b_1}{\delta}\right)^{(l)}(1-\theta z)^{-\frac{(t_0+(l-1)\theta)\delta}{\theta}} + ...$$

   At this point, noting that for $\gamma_1 < \gamma_2$, $[z^n](1-z)^{-\gamma_1} = o([z^n](1-z)^{-\gamma_2})$ (for a reference, see [37]), we have that only the first term influences the asymptotic behavior. So, thanks to a coefficient extraction with respect to z, we get the desired result.

---

5 Let $f(X) = \sum_{i=0}^{\infty} u_i z^i$ be a generating function. One simply has $[z^m]f(z) = [z^m]\sum_{i=0}^{\infty} a_i z^i = a_m$.

■

An immediate consequence of this proposition is that, for example, the law leading the number of y balls in the urn is not Gaussian, since one can easily verify that expectation and standard deviation show the same order of growth. Similar results holds for the w balls. Finally, the number of x balls follows a standard Polya urn, whose behavior is different and well-known.

1.4 Urn's results and shock models

With our urn process we are trying to model shock models, so we are particularly interested in the dynamics of y and w balls, representing the risky and the default states. For example, it would be useful to know something about the limit distribution of Y_n (and analogously of W_n).

Looking at the moments of Y_n and W_n we have obtained in Proposition 9 and comparing them with those presented in [70] for a 2×2 triangular urn, we find that, up to scaling, they are the same. So, thanks to the standard moment convergence theorem (for a reference, see [12]) we can adapt some results in [70] to our case.

Proposition 10 *The random variable Y_n has a local limit law whose density is characterized by generalized Mittag-Leffler functions. In particular, for any compact set S of $\mathbb{R}^+$ and any $\gamma \in S$ such that $\gamma n^{\frac{\delta}{\theta}}$ is an integer, we have that*

$$P\left[Y_n = b_0 + \delta\gamma n^{\frac{\delta}{\theta}}\right] = n^{-\frac{\delta}{\theta}} g_y(\gamma) + O(n^{-2\frac{\delta}{\theta}})$$

where the error term holds uniformly with respect to $\gamma \in S$. Moreover, the function $g_y(\cdot)$ is defined on $\mathbb{R}^+$ according to the following series expansion:

$$g_y(\gamma) = \frac{\Gamma\left(\frac{t_0}{\theta}\right)}{\Gamma\left(\frac{b_0}{\delta}\right)}\gamma^{\frac{b_0}{\delta}-1}\sum_{k\geq 0}(-1)^k\frac{\gamma^k}{\Gamma(k+1)\Gamma\left(\frac{c_0-k\delta}{\theta}\right)}.$$

An analogous reasoning is valid for W_n with $g_w(\cdot)$

$$g_w(\gamma) = \frac{\Gamma\left(\frac{t_0}{\theta}\right)}{\Gamma\left(\frac{b_0}{\lambda}\right)\Gamma\left(\frac{c_0}{\theta}\right)}\gamma^{\frac{\delta}{\theta}-3}\sum_{k\geq 0}(-1)^k\frac{\gamma^k}{\Gamma(k+1)\Gamma\left(\frac{t_0-b_0-k\lambda}{\theta}\right)}$$

where $\gamma \in S$ and $\gamma n^{2-\frac{\delta}{\theta}}$ is an integer.

Proof. The result comes from some cumbersome calculations for extracting coefficients from the generating function of urn histories, with respect to y (or w) and then to z, using singularity analysis as in [28], [34] and [53].

In particular, we know that our function is analytic in a proper Camembert region. Then, defining a proper contour we approach the singularity at distance $\frac{1}{n}$. More specifically, we choose a quite standard semi-complete circle C (see figure 1.3 for further details[6]). Hence we apply Cauchy integral formula, according to which

$$[z^n]\, H(z) = \frac{1}{2\pi i}\int_C \frac{H(z)}{z^{n+1}}dz.$$

In the specific, from the generating function we have

$$[z^n]\,(1-\theta z)^{\frac{-b_0}{\theta}}\left(1-(1-\theta z)^{\frac{\delta}{\theta}}\right)^k = \frac{\theta^n}{2\pi i}\int_C (1-z)^{\frac{-b_0}{\theta}}\left(1-(1-z)^{\frac{\delta}{\theta}}\right)^k\frac{dz}{z^{n+1}},$$

where $k = \gamma n^{\frac{\delta}{\theta}}$.

6 This picture is an elaboration of Fig. 9 in [36].

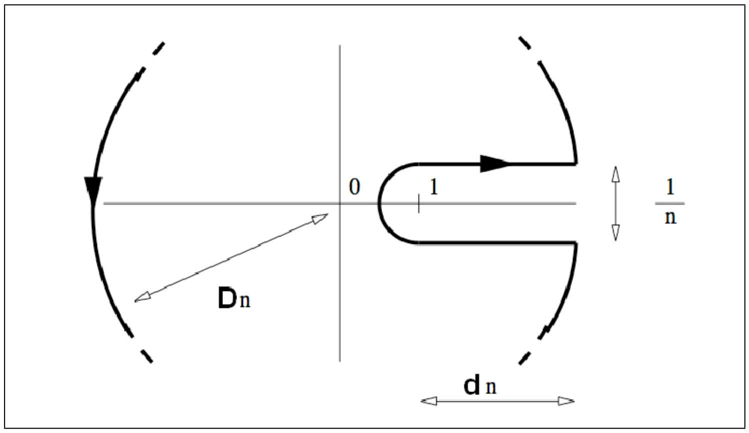

Figure 1.3: Integration contour for the extraction of coefficients, where, given $m \in \mathbb{R}^+$, $d_n = 1 + mn^{-\frac{\delta}{\theta}}$ and $D_n = \sqrt{d_n^2 + n^{-2}}$.

Then we choose $z = 1 - \frac{t}{n}$, obtaining

$$P\left[Y_n = b_0 + k\delta\right] \sim \frac{\Gamma\left(\frac{t_0}{\theta}\right)}{\Gamma\left(\frac{b_0}{\delta}\right)} \gamma^{\frac{b_0}{\delta}-1} \frac{1}{2\pi n^{\frac{\delta}{\theta}}} \int_Q t^{\frac{-b_0}{\theta}} e^{-\gamma t^{\frac{\delta}{\theta}}+t} dt,$$

where Q is a clockwise oriented negative loop around the negative real axis, as in [37].

At this point, we keep $t^{\frac{-b_0}{\theta}} e^t$ and, using Maple, expand the other terms.

The Hankel representation of Gamma functions (see [1]) allows to conclude the proof. ■

This last proposition is undoubtedly important, since it gives us very helpful information about our urn process. In fact, as known from complex analysis, the generalized Mittag-Leffler function $\sum_{k\geq 0} \frac{\gamma^k}{\Gamma(k+1)\Gamma\left(\frac{b_0-k\delta}{\theta}\right)}$ is an entire function and a special case of hyperbolic function. In particular, it interpolates between a purely exponential law and a power-law behavior (see [66]). This makes our modelization rather flexible. For example:

1. For $c_0 = 0$, that is if there are no w balls in the initial composition of the urn so that the process cannot fail immediately, the function $g_y(\gamma)$ represents a Paretian stable law of index $\frac{\delta}{\theta}$.

2. If $c_0 < \theta$, on the contrary, $g(\gamma)$ becomes a Gamma distribution (even an exponential for $b_0 = \delta$).

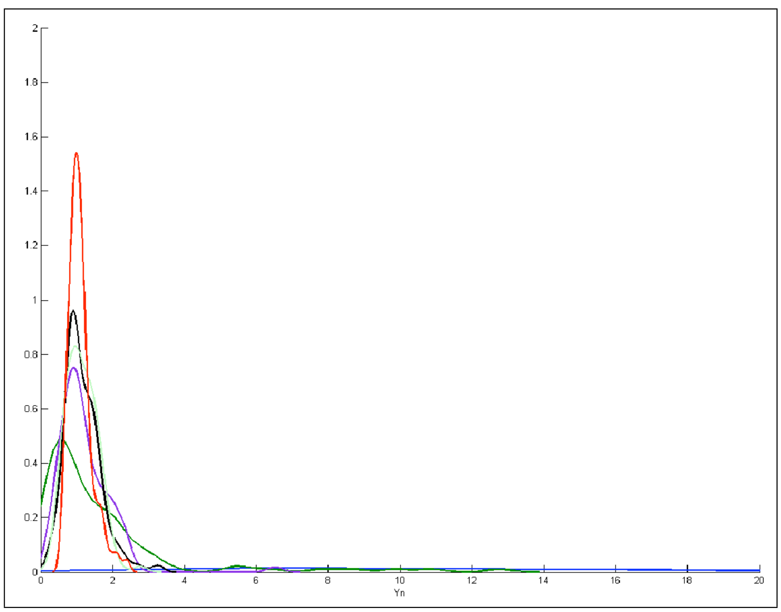

Figure 1.4: Limit law for Y_n for different values of θ and δ. Initial composition $(20, 5, 1)$.

The main consequence of this flexibility is that, for example, we are able to indirectly reproduce the results of Theorems 4 and 5, which was one of our first aims.

For example, consider the standard GESM in [46]. Computing the probability that $Y_{10} > b_0$ using 1.6 is like asking in [46] which is the probability for the system to overcome the risky threshold for the first time in $n = 10$. In the same way, $P[W_n > c_0]$ represents the probability for the model to fail at time n.

Figures 1.4 gives an idea of the limit laws for Y_n for particular values of θ and δ and given urn initial compositions.

1.5 The Bayesian perspective

We mentioned that the urn approach to GESM's can be seen as a Bayesian perspective on shock models. In fact, thanks to the urn modelization, the researcher is able to introduce in the process its prior knowledge by acting on the initial composition of the urn or on the reinforcement quantities θ and δ.

For example, as we have seen in the previous section, changing the values of a_0, b_0 and c_0 influences the limit law of Y_n. In the same way, modifying the number of balls that reinforce the process, we obtain different values for the moments and the probabilities of the three distinct colors in the urn. So, if we have observed a certain behavior for an actual system and we want to model it, we can intervene on the parameters of the urn processes, incorporating information. In fact, while in general we expect b_0 and c_0 to be quite small, since shocks and default are generally considered extreme events, there could be a particular application in which the system is constantly affected by shocks, so that b_0 can be greater. Therefore, modifying the initial values of the system, we are able to make prediction and compute posterior distributions for the process without an explicit knowledge of priors, and this surely simplifies the analysis.

Another interesting characteristic of our urn process is that, asymptotically, its composition - that is the proportions of different colors in the urn - behaves like that of a standard Polya urn, following a Dirichlet distribution (see [14] for an appealing perspective about Dirichlet distribution).

A similar result can be obtained looking at the basic definition of our reinforcement matrix. In fact, using the terminology of [69], if we standardize our reinforcement matrix by dividing each term by θ, that is the balance of the urn, we can say that our urn is *essentially Polya*, since 1 is a multiple (multiplicity is 2) eigenvalue for $\frac{1}{a}RM$. Moreover if $\delta > \frac{\theta}{2}$ (that is y-balls reinforce themselves more than they do for w-balls), we also have that our urn process is *large* and *semisimple*, according to the definition of [69].

Proposition 11 *Consider the urn process we have defined above. Let U_n represent the composition of the urn at time n, so that $U_0 = (a_0, b_0, c_0)$. Then the quantity $\frac{1}{\theta n}U_n$ converges almost surely and in any L^p, $p \geq 1$, to a random vector $V_1 r_1 + V_2 r_2 + V_3 r_3$ where the vectors r_1, r_2 and r_3 are the dual basis of $\frac{1}{\theta}RM^T$ and the random vector (V_1, V_2, V_3) has a Dirichlet distribution, whose density on the simplex*

$(u_x \geq 0, u_y \geq 0, u_z \geq 0, u_x + u_y + u_w = 1)$ of $\mathbb{R}^3$ is given by

$$\Gamma(\frac{t_0}{\theta})\frac{u_x^{a_0+c_0}}{\Gamma\left(a_0+c_0\right)}\frac{u_y^{b_0-c_0}}{\Gamma\left(b_0-c_0\right)}\frac{u_w^{\frac{1}{a}[\lambda c_0-\delta a_0]}}{\Gamma\left(\frac{1}{\theta}\left[\lambda c_0-\delta a_0\right]\right)},$$

where u_x represents the proportion of x balls in the urn.

Proof. First one computes the Jordan basis for the process, starting from its reinforcement matrix. As known a Jordan basis is a basis satisfying

$$\begin{aligned} RMb_{i,1} &= \lambda_i b_{i,1} \\ RMb_{i,j} &= \lambda_i b_{i,j} + b_{i,j-1}. \end{aligned}$$

Given the Jordan basis, the dual ones are calculated.

Then, using theorem 1 in [69], the result immediately follows. ■

In Bayesian statistics, an important form of dependence among random variables is exchangeability. The possibility of having exchangeable sequences is in fact very useful when studying the properties of as particular stochastic process.

Let $\overline{X}_n, \overline{Y}_{n,}, \overline{W}_n$ represent the proportions of x, y and w balls in the urn at time n. Unfortunately, it is easy to verify that, because of the lack of symmetry in the reinforcement matrix, the sequences $\{\overline{X}_n\}$, $\{\overline{Y}_n\}$ and $\{\overline{W}_n\}$ are not exchangeable in the sense of de Finetti (see [25]). However, they all show an useful property introduced in [8] and further investigated in [9].

Definition 2 ([8]**)** *A sequence $\{Z_n\}_{n\geq 1}$ defined on a probability space $(\Omega, \mathcal{L}, P)$, that take values in the measurable space (E, Σ) and is adapted to a filtration $\mathcal{F} = (\mathcal{F}_n)_{n\geq 0}$, is said to be conditionally identically distributed with respect to $\mathcal{F}$ ($\mathcal{F}$-c.i.d.)*

if

$$E\left[f\left(Z_k\right)|\mathcal{F}_n\right] = E\left[f\left(Z_{k+1}\right)|\mathcal{F}_n\right] \quad a.s.$$

for all $k > n \geq 0$ and all bounded measurable $f : E \to \mathbb{R}$.

Proposition 12 *Let $\left\{\overline{W}_n\right\}$ represent the sequence of the proportions of the balls of type w generated by the urn process defined in Section 3 up to time n. Set the filtration $\mathcal{F}$ such that $\mathcal{F}_0 = \{\varnothing, \Omega\}$ and $\mathcal{F}_n = \sigma\left(W_1, W_2, ..., W_n\right)$ for $n \geq 1$. Then the sequence $\left\{\overline{W}_n\right\}$ is $\mathcal{F}$-c.i.d. Similar results hold for $\left\{\overline{X}_n\right\}$ and $\left\{\overline{Y}_n\right\}$.*

Proof. To prove the proposition it is sufficient to show that $\left(E\left[\overline{W}_{n+1}|\mathcal{F}_n\right]\right)_{n\geq 0}$ is a $\mathcal{F}$-martingale.

First, note that

$$E\left[\overline{W}_{n+1}|\mathcal{F}_n\right] = \frac{c_0 + \sum_{i=1}^{n} \lambda 1_{\{R_i=y\}} + \sum_{i=1}^{n} \theta 1_{\{R_i=w\}}}{t_0 + n\theta} \quad a.s. \text{ for all } n \geq 1,$$

where R_n represents the color of the ball sampled at time n from the urn and $1_{\{\cdot\}}$ is the indicator function.

For $n = 0$ we have $E\left[E\left[\overline{W}_2|\mathcal{F}_1\right]|\mathcal{F}_0\right] = E\left[\overline{W}_2\right] = E\left[\overline{W}_1\right] = E\left[\overline{W}_1|\mathcal{F}_0\right]$ a.s. For $n \geq 1$, $E\left[\overline{W}_{n+1}|\mathcal{F}_n\right] = E\left[\overline{W}_{n+1}|\overline{W}_1, \overline{W}_2, ..., \overline{W}_n\right]$ a.s. Then

$$\begin{aligned} E\left[E\left[\overline{W}_{n+2}|\mathcal{F}_{n+1}\right]|\mathcal{F}_n\right] &= \frac{c_0 + \sum_{i=1}^{n} \lambda 1_{\{R_i=y\}} + \sum_{i=1}^{n} \theta 1_{\{R_i=w\}}}{t_0 + (n+1)\theta} + \\ &+ \frac{E\left[\lambda 1_{\{R_{n+1}=y\}} + \theta 1_{\{R_{n+1}=w\}}|\mathcal{F}_n\right]}{t_0 + (n+1)\theta} \\ &= E\left[\overline{W}_{n+1}|\mathcal{F}_n\right] \ a.s. \end{aligned}$$

Similar proofs hold for $\left\{\overline{X}_n\right\}$ and $\left\{\overline{Y}_n\right\}$. ■

This is a very interesting results. In fact, it has an immediate consequence, that gives us useful information about the dependence structure induced by our urn process.

Corollary 13 *The sequence* $\{\overline{W}_n\}$ *is asymptotically exchangeable in the sense of* [2]. *The same for* $\{\overline{X}_n\}$ *and* $\{\overline{Y}_n\}$.

Proof. The proof is a direct application of Lemma 2.4 and Theorem 2.5 in [8]. ■

This last corollary, together with Proposition 11, shows that, asymptotically, our urn process, even introducing some dependence among the balls, does not differ substantially from the classical Polya urn.

1.6 Conclusions

In this chapter we have proposed a new way for analyzing generalized extreme shock models using a special triangular urn process. The reinforcement matrix builds a positive dependence between the risky and the default states, indirectly modelling the moving threshold mechanism first introduced in [46]. The choice of an analytic approach to the urn has given us the possibility of explicitly derive probabilities, moments and limit laws for the three different types of balls, that is for the three distinct levels of risk. Using the algebraic approach proposed in [69], we have also calculated the limit law for the urn composition.

Moreover, the flexibility of the urn approach seems to be quite interesting, since it allows to incorporate prior knowledge into the process, opening GESM's to the Bayesian perspective. In fact, modifying the initial composition of the urn and the reinforcement quantities, the researcher can actively introduce his knowledge of the phenomenon into the modellisation.

In the future, it could be worth to introduce other forms of dependence among the states or let the reinforcement matrix evolve over time, as suggested in [64]. Moreover, it would be interesting to verify if other kinds of shock models could be modelled using urns, as we propose, for example, in the next chapter.

Chapter 2
An Alternative: The Ruinous Urn

This chapter introduces an alternative modelization to the urn-based version of generalized extreme shock models, following a recent research direction we have undertaken.

The idea of a generalized ruinous urn, as we have called it, has been inspired by the interesting cannibal urn described in [68]. The main characteristic of the ruinous urn is that it is not tenable, according to the definition of [36].

The aim is to model, in a pessimistic way, a system subject to shocks of random magnitude, starting from the assumption that, however, the system will inevitably fail. In fact, we will assume that it may default both for an extreme shock and for natural obsolescence, since the passage of time inescapably deteriorates the system.

We will show that this modelization produces interesting results.

2.1 Introduction

As we have seen in the previous chapter, shock models are a class of models that study the behavior of a system subject to shocks of random magnitude. In particular extreme shock models assume that the system fails after one big extreme shock that overcomes a certain endurance threshold.

Anyway, it is not ludicrous to reckon that obsolescence plays an important role in the life of the system. The passage of time has an unavoidable effect on system resilience to shocks and, even if no extreme shock occurs, sooner or later the system will break down for the usage.

The Ruinous Urn we are about to present is our modelization for such a situation and it can be considered an alternative to the triangular urn we have already described.

As we will see, the new urn can be considered a combination of three different urn schemes:

- a generalized Polya triangular urn, as the one of the previous chapter;
- a sacrificial urn, in the sense of [35];
- Pittel's cannibal urn (see [68]).

Hence, before analyzing the Ruinous Urn in detail, it can be useful to briefly revise some basics aspects of sacrificial and Pittel's urns, view that triangular urns have been already treated.

2.1.1 Sacrificial urns

Sacrificial urns are a class of urns in which one or more colors sacrifice themselves all to the good of the remaining ones. Famous examples of sacrificial urns are the two-chamber Ehrenfest urn (presented in [31]), whose reinforcement matrix is

$$A = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix},$$

and the peak-in-perms Mahmoud urn (see [55]), with

$$B = \begin{bmatrix} -1 & 2 \\ 1 & 0 . \end{bmatrix}$$

Suitable references on the subject can be [35] and [36]. The second one will also represent the canvas of this short introduction on sacrificial urns.

The general form for a 2×2 sacrificial urn is given by the initial composition (a_0, b_0) and the following balanced reinforcement matrix:

$$M = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix},$$

where $\alpha \leq 1$, $\gamma \cdot \beta \neq 0$ and $\sigma = \alpha + \beta = \gamma + \delta$. Moreover the urn is assumed tenable, as per the following definition.

Definition 3 *An sacrificial urn with reinforcement matrix M is said tenable when the diagonal entry $\alpha \leq 1$ is such that $|\alpha|$ divides both a_0 and γ.*

For convenience, set $p = \gamma - \alpha$, $s_0 = a_0 + b_0$, $s_n = s_0 + n\sigma$ and remember the definition of the following hypergoniometric functions (see [54] and [37]).

Definition 4 *Let $\phi \in \mathbb{Z}^+$ and $\xi \in \mathbb{Q}_0^+$ be two parameters. The fundamental integral $J = J_{\phi,\xi}$ is defined as*

$$J_{\phi,\xi}(u) = \int_0^u \frac{dy}{(1 + y^\phi)^\xi}. \tag{2.1}$$

The pseudo-sine function $S = S_{\phi,\xi}$ is defined in a complex neighbourhood of 0 as the inverse of J,

$$S(J(u)) = J(S(u)) = u.$$

The pseudo-cosine function $C = C_{\phi,\xi}$ is defined in a complex neighbourhood of 0 as

$$C(z) = (1 + S(z)^\phi)^{\frac{1}{k}},$$

where the principal determination of the $k-$th root is adopted.

S and C can be considered higher degree analogues of standard trigonometric or hyperbolic functions[7] (see [37]). For this reason it is quite easy to show that they are analytic[8] at 0, and in particular $S(z) = z + O(z^{\phi+1})$ and $C(z) = 1 + O(z^{\phi})$.

Moreover it can be shown that the fundamental integral of equation 2.1 can be written in a hypergeometric form and the same for the pseudo-sine and the pseudo-cosine functions. Following [1], the hypergeometric function is defined as a combination of products and quotients of Gamma functions, or

$$\begin{aligned} H\left[\begin{array}{c} a, b \\ c \end{array} \middle| x \right] &= 1 + \frac{ab}{c}\frac{z}{1!} + \frac{a(a+1)b(b+1)}{c(c+1)}\frac{z^2}{2!} + \ldots \\ &= \sum_{i=0}^{\infty} \frac{a^{(i)}b^{(i)}}{c^{(i)}} \frac{z^i}{\Gamma(i+1)}. \end{aligned}$$

So we have

$$J_{\phi,\xi}(z) = zH\left[\begin{array}{c} \xi, \frac{1}{\phi} \\ 1+\frac{1}{\phi} \end{array} \middle| -z^{\phi} \right].$$

This last property is very useful for calculations, since it dramatically simplifies the work of symbolic languages such as Maple11[9].

We are now ready to present the main theorem related to sacrificial urns.

---

7 For example, if $\phi = 2$ and $\xi = \frac{1}{2}$, we have $J(u) = \operatorname{arcsinh}(u)$, $S(u) = \sinh(u)$ and $C(u) = \cosh(u)$.

8 Here, an analytic function is a function that is locally given by a convergent power series. This clarification is meant to avoid the common superimposition between analytic and holomorphic functions.
In this thesis we refer to the definitions of [37].

9 As said before, Maple 11 has been the main symbolic language we have used for performing or controlling all our calculations.

Theorem 14 (Flajolet et al. [36]**)** *Consider a balanced sacrificial urn, whose reinforcement matrix* $M = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$ *is such that* $\alpha \leq 1$ *and* $\gamma \cdot \beta \cdot \delta \neq 0$. *Then the bivariate generating function of urn histories, for* (x, z) *in a neighbourhood of* $(0, 0)$, *is given by*

$$H(z; x, 1) = \Delta^{s_0} S(-\alpha z \Delta^{\sigma} + J(x^{-\alpha}\Delta^{\alpha}))^{-\frac{a_0}{\alpha}} C(-\alpha z \Delta^{\sigma} + J(x^{-\alpha}\Delta^{\alpha}))^{-\frac{b_0}{\delta}},$$

where $\Delta = (1 - x^p)^{\frac{1}{p}}$, J *is the fundamental integral,* S *is the pseudo-sine function and* C *is the pseudo-cosine function with* $k = -\frac{p}{\delta}$.

A special case of sacrificial urns are the so-called semi-sacrificial ones. The main feature is that, given the reinforcement matrix M, $\alpha \leq 1$ and $\delta > 0$, that is there is only one color that sacrifices itself, while the other is a hanger-on.

As we will see, at a first glance, our ruinous urn and a standard semi-sacrificial urn can seem a lot alike. Notwithstanding they are quite different, since our urn violates two fundamental assumptions of sacrificial urns: it is not tenable and $\gamma = 0$ so that $\gamma \cdot \beta = 0$.

2.1.2 A sketch on Pittel's cannibal urn

In 1960, R.F. Greene introduced, in an unpublished and nowhere-to-be-found paper, a simple urn model to study the behavior of cannibal tribes. That model has been renewed and deeply analyzed by Pittel in [68], that is why it is now known as Pittel's cannibal urn.

Imagine we have a population made up of cannibal and non-cannibal individuals. At every time step, a non-cannibal is selected as victim and removed. Then another member in the remaining population is selected at random. If the selected individual is a cannibal s/he remains the same, but if s/he is not a cannibal, s/he then becomes an antropophagous.

The reinforcement matrix of this urn process is simply given by

$$P = \begin{bmatrix} 0 & -1 \\ 1 & -2 \end{bmatrix} \begin{array}{l} cannibal \\ non-cannibal. \end{array}$$

This matrix is interesting since it is clearly not tenable. Moreover it is characterized by the presence of a set of absorbing states, that is to say states characterized by only one kind of balls - in this case those representing cannibals. To be more exact, the urn process continues until one the elements of $\aleph = \{(i,n) : n \geq 0\}$, with $i = 0, 1$, is reached.

Starting from a given initial composition (a_0, b_0), the urn evolves to $(a_t = a_{t-1},$ $b_t = b_{t-1} - 1)$ with probability $\frac{a_{t-1}}{b_{t-1}-1+a_{t-1}}$ and to $(a_t = a_{t-1} + 1, b_t = b_{t-1} - 2)$ with probability $\frac{b_{t-1}-1}{b_{t-1}-1+a_{t-1}}$, where a_t and b_t represent the number of cannibals and non-cannibals in the urn at time t. Figure 2.1 gives an idea of the urn behavior starting from the initial composition of 8 non-cannibals and 5 cannibals.

In his paper [68], Pittel has shown that the number of cannibals remaining when there are no more non-cannibals in the population is asymptotically normal as long as $b_0 \neq o(a_0)$. In this last case he has demonstrated the presence of a Poisson limit law.

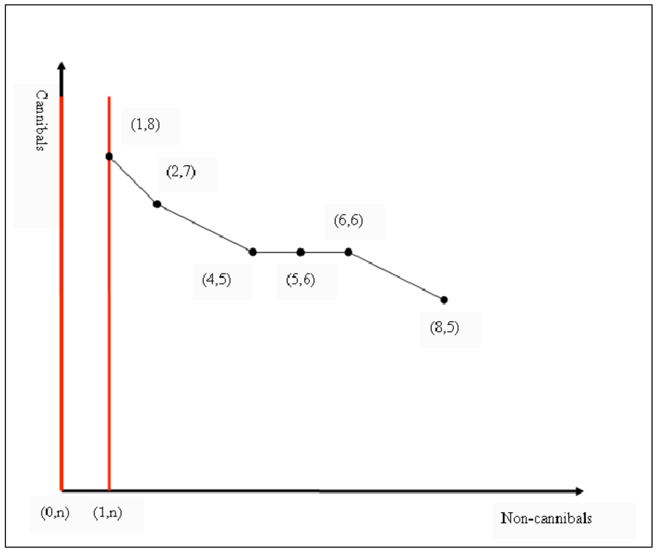

Figure 2.1: Example of Pittel's cannibal urn with initial composition $(8,5)$. In red the absorbing set.

2.2 The Ruinous Urn

It is now time to introduce the 2×2 ruinous urn by describing its reinforcement matrix and properties.

Consider the following reinforcement matrix

$$M = \begin{array}{c} x \\ y \end{array} \begin{bmatrix} \alpha & \sigma - \alpha \\ 0 & \sigma \end{bmatrix}, \tag{2.2}$$

where $\alpha \leq -1$ and $\sigma \geq 0$.

We can then do some do some obvious considerations:

1. the urn is balanced, i.e. at every time step we add σ balls in the urn;

2. the reinforcement matrix is triangular, but $\alpha \leq -1$ so that we cannot apply the same results of the previous chapter;

3. the tenability property is clearly violated, view that $\gamma = 0$;

4. the matrix M does not characterize a sacrificial (or semi-sacrificial) urn, since it presents a null term on the anti-diagonal;

5. As Pittel's cannibal urn, the ruinous urn possesses a set of absorbing states $S = \{(0, k) : k \geq 0\}$.

Before analyzing in detail the urn behavior, it can be useful to briefly describe how it works and the connections with shock models.

The urn is initialized with a_0 balls of type x, representing the safe state[10], and b_0 balls of type y, symbolizing default. When a ball is sampled from the urn, its color is observed and the urn is then reinforced according to the matrix M. In particular, if a $x-$ball is chosen, we remove $|\alpha|$ $x-$balls from the urn and we add $\beta = \sigma - \alpha$ $y-$balls. If, on the contrary, a $y - ball$ is picked up, we simply reintroduce the ball together with other σ balls of the same type.

Two clear consequences of this modelization are:

1. $x-$balls are destined to disappear from the urn (and this produces the absorbing set S);

2. the probability of picking a $y-$ball inescapably increases over time.

---

[10] See the previous chapter for a more detailed description about the relationships between urn colors and risk states.

In this way, we obtain that the passage of time increases the probability of default of the system, both by direct reinforcement[11] and by decrease of the probability of surviving. In a nutshell, the ruinous urn makes the system fail not only when a $y-ball$ is sampled, but also when there are no more $x-$balls available.

Set $X_n \geq 0$ and $Y_n \geq 0$ to represent the number of x and $y-$balls in the urn at time n, while C_n is the color of the $n-$th ball sampled. We are interested in calculating

$$P\left[X_n = 0\right],$$

that represents the probability of a default due to obsolescence, and

$$P\left[C_n = y\right],$$

that embodies the probability of a default due to an extreme shock. Finally it is worth to consider the stopping time $\tau_n^y = \min\left\{n : C_n = y\right\}.$

In order to derive the above probabilities and the other properties of the ruinous urn we need the following tedious but useful lemmas.

Lemma 15 *Let M be a reinforcement matrix as in equation 2.2 . We know that the associated differential system is*

$$\sum = \left\{ \begin{array}{l} \dot{x} = x^{\alpha+1}y^{\sigma-\alpha} \\ \dot{y} = y^{\sigma+1} \end{array} \right. \quad with\ i.c. \left\{ \begin{array}{l} x(0) = x_0 \\ y(0) = y_0 \end{array} \right. . \tag{2.3}$$

[11] We remember that the reinforcement $\alpha + \beta$ is fictitious, since it is only necessary for doing calculations. In fact, as already said, when a $y-$ball is sampled the system fails and the process stops.

Then $x^{-\alpha} - y^{-\alpha}$ is a first integral for the system.

Proof. We simply need to solve the system and verify that

$$\frac{d}{dt}\left(x^{-\alpha} - y^{-\alpha}\right) = -\alpha x^{-\alpha-1}\dot{x} + \alpha y^{-\alpha-1}\dot{y} = 0.$$

This guarantees that $x^{-a} - y^{-a}$ is constant. ■

Lemma 16 *Consider the differential equation*

$$\dot{x} = x^{\alpha+1}y^{\sigma-\alpha}.$$

We can eliminate the variable y simply using the first integral defined in the previous lemma.

Proof. Let $X(t) = X(t\,; x_0, y_0)$ and $Y(t) = Y(t\,; x_0, y_0)$ represent the solution of the differential system 2.3. If $x_0 y_0 \neq 0$ then, for t in a neighborhood of 0, we have that $X(t)^{-\alpha} - Y(t)^{-\alpha} = x_0^{-\alpha} - y_0^{-\alpha}$.

Thus

$$\dot{x} = x^{\alpha+1}y^{\sigma-\alpha} = x^{\alpha+1}(x^{-\alpha} - x_0^{-\alpha} + y_0^{-\alpha})^{\frac{\alpha-\sigma}{\alpha}} = x^{\alpha+1}(x^{-\alpha} - D^{-\alpha})^{\frac{\alpha-\sigma}{\alpha}},$$

where $D = (y_0^{-\alpha} - x_0^{-\alpha})^{-\frac{1}{\alpha}}$. ■

Let us introduce a change of variable that will help us in calculations. Set

$$\begin{aligned} X(t) &= X(t\,; x_0, y_0) = stD^{\sigma+1}, \\ Y(t) &= Y(t\,; x_0, y_0) = wtD^{\sigma+1}. \end{aligned}$$

Then the system 2.3 can be rewritten as

$$\sum = \begin{cases} \dot{s} = s^{\alpha+1}(s^{-\alpha}+1)^{\frac{\alpha-\sigma}{\alpha}} \\ \dot{w} = w^{\sigma+1} \end{cases} \quad with\ i.c. \begin{cases} z(0) = x_0 D^{-1} \\ w(0) = y_0 D^{-1} \end{cases} .$$

From this new formulation we obtain

$$1 = \frac{\dot{s}}{s^{\alpha+1}(s^{-\alpha}+1)^{\frac{\alpha-\sigma}{\alpha}}}$$

and integrating

$$t = \int_{x_0 D^{-1}}^{s(t)} \frac{u^{-\alpha-1}}{(1+u^{-\alpha})^{\frac{\alpha-\sigma}{\alpha}}} du.$$

Now set $\psi = u^{-\alpha}$, so that we obtain

$$-\alpha t = \int_{x_0^{-\alpha} D^{\alpha}}^{s(t)^{-\alpha}} \frac{d\psi}{(1+\psi)^{\frac{\alpha-\sigma}{\alpha}}}.$$

This last expression should be familiar view that it can be expressed by the means of the fundamental integral introduced in Definition 4. In fact

$$-\alpha t = J_{\phi,\xi}(s(t)^{-\alpha}) - J_{\phi,\xi}(x_0^{-\alpha} D^{\alpha}),$$

where $\phi = 1$ and $\xi = \frac{\alpha}{\alpha-\sigma}$.

We are now ready to state the following fundamental result.

Proposition 17 *Consider a balanced ruinous urn with reinforcement matrix* $M = \begin{bmatrix} \alpha & \sigma-\alpha \\ 0 & \sigma \end{bmatrix}$, $\alpha \leq -1$ *and* $\sigma \geq 0$. *Initialized the urn with* $a_0 > b_0 > 0$ *balls (remember* $s_0 = a_0 + b_0$*) and let* S *be the pseudo-sine function. Then its bivariate generating function of urn histories is given by*

$$H(t;x,y) = (x_0^{-\alpha} - y_0^{-\alpha})^{\frac{s_0}{\alpha}} S(-\alpha t + J(x^{-\alpha} D^{\alpha}))^{\frac{a_0}{\alpha}} (1-\sigma y^{\sigma} t)^{-\frac{b_0}{\sigma}}. \tag{2.4}$$

Proof. Once again (see the previous chapter), the generating function is simply given by the application of theorem 7, once we have noticed the following elements:

1. The solution of $\dot{y} = y^{\sigma+1}$ is easily derived as $y_0(1-\sigma y_0^{\sigma} t)^{-\frac{1}{\sigma}}$;

2. Since $-\alpha t = J_{\phi,\xi}(s(t)^{-\alpha}) - J_{\phi,\xi}(x_0^{-\alpha} D^{\alpha})$, we have that

$$s^{-\alpha} = S(-\alpha t + J_{\phi,\xi}(x_0^{-\alpha} D^{\alpha})).$$

   But we known that $X_t = stD^{\sigma+1}$, so $X(t)^{-\alpha}$ is immediately obtained.

■

It is now time to study the probability of default of the system.

Proposition 18 *Given a ruinous urn as described above, we have that*

$$P[X_n = a_0 + \alpha l] = \binom{l + \frac{a_0}{\alpha} - 1}{l} \frac{n!}{\left(\frac{s_0}{\sigma}\right)^{(n)}} \sum_{j=1}^{l} (-1)^j \binom{l}{j} \binom{n(\sigma-1) + b_0 - \alpha j}{n},$$

and, in particular

$$P[X_n = 0] = pn^q U^n (1 + O(n^{-1}),$$

where $p > 0$ is a constant, $q = (-\frac{\alpha+\sigma}{\alpha\sigma})$ and $U = Beta(1, -\frac{\sigma}{\alpha})$

Proof. By construction we know that $x-$balls can only decrease in measure of α. So X_n can only be equal to $\max(a_0 + \alpha l, 0)$ with $l \in \mathbb{Z}^+$.

Then, we set $y = 1$ and we perform coefficient extraction with respect to x, noting the presence of the group $(1 - x^{-\alpha})^{\frac{1}{\alpha}}$, whose expansion is well-known, being a slight modification of Stirling formula (see [74]). At this point, we expand[12] the $l-$th power upon performing coefficient extraction with respect to t.

The probability $P[X_n = 0]$ can be easily computed from $H(t;0,1) = S(-\alpha t)^{\frac{a_0}{\alpha}}(1-\sigma t)^{-\frac{b_0}{\sigma}}$ using, as usual the $[t^n]$ operation. ■

On the contrary, given the simplicity of the result, no proposition is needed to study the probability $P\left[C_n = y\right]$.

Moreover, as far as the stopping time $\tau_n^y = \min\left\{n : C_n = y\right\}$ is concerned, we trivially have

$$P\left[\tau_n^y > n\right] = \prod_{i=1}^{n-1} \frac{X_i}{X_i + Y_i}.$$

Proposition 19 *The moment of a ruinous urn have hypergeometric form, i.e. they are $\mathbb{C}-$linear combinations of hypergeometric terms.*

Proof. In order to prove the proposition, first notice that

$$\begin{aligned} E\left[(X_n)_l\right] &= \frac{[t^n]l!\left[(x-1)^l\right]H(t;x,1)}{[t^n]H(t;x,1)}, \\ E\left[(Y_n)_l\right] &= \frac{[t^n]l!\left[(y-1)^l\right]H(t;1,y)}{[t^n]H(t;1,y)}, \end{aligned}$$

where $[a^n]$ is the usual operation of coefficient extraction.

Then notice that J and S are hypergoniometric functions (see [54]) and the result follows. ■

---

[12] Even in this case, calculations have been performed using Maple 11.

In this case, without defining the parameters in M, it is not possible to explicitly derive the moments of order l. For example, Maple 11 stops processing while searching the general solution.

Finally, we can give the following last result, that represents a clear link between the ruinous urn and sacrificial urns.

Proposition 20 *Let $a_0 > b_0 > 0$ and $\sigma \simeq s_0$. The number X_n of $x-$balls at time n is such that*

$$P\left[\frac{X_n - \mu n}{\nu\sqrt{n}} \leq x\right] = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt + O(n^{-\frac{1}{2}}),$$

where $\mu = \lim_{n\to\infty} \frac{E[X_n]}{n}$. and $\nu = \lim_{n\to\infty} \frac{Var(X_n)}{n}$.

Proof. Consider the following. For $x \longrightarrow 1$ we have

$$[t^n]H(t;x,1) \sim \frac{1}{\sigma} r(x)^{-n-1},$$

where $r(x) = x^{-\sigma} \sum_{j\geq 0} \binom{-\frac{\alpha}{\sigma-\alpha}}{j} \frac{1}{\sigma - j\alpha} \left[\frac{x^{-\alpha}}{(1-x^{-\alpha})}\right]^{-j}$ is the radius of convergence of $H(t;x,1)$.

Now, from the previous proposition, we have (set $s = x$ not to confuse notation)

$$E(s^{X_n}) = \frac{[t^n]l!\left[(n-1)^l\right] H(t;s,1)}{[t^n]H(t;1,1)} = \left(\frac{r(1)}{r(s)}\right)^{n+1} \left[1 + O(n^{\frac{\alpha}{\sigma}})\right].$$

An analysis of $E(s^{X_n})$ shows that it can be seen as the $(n+1)-$th power of a fixed function. A consequence of this is that X_n automatically corresponds to the sum of $n+1$ independent identically distributed random variables. It is rather intuitive to expect a Gaussian law in asymptotics. Fortunately the quasi-powers central limit theorem of [47] helps us in proving all this.

At this point, the coefficients can be obtained by extraction from $r(x)$. ■

2.3 A brief sketch about higher dimensional versions

What we have written for the 2×2 ruinous urn can be obviously adapted to higher dimensional urns.

In particular, this is a research direction we are still involved in. No need to underline the amounts of messy calculations that are required in order to get useful results. Anyway, some observations are always possible.

Consider the following 3×3 cases:

$$
\begin{aligned}
M_1 &= \begin{array}{c} x \\ y \\ z \end{array} \begin{bmatrix} a & b & f \\ 0 & c & d \\ 0 & 0 & e \end{bmatrix}, \ a, c \leq -1 \text{ and } b, d, e, f > 0 \\
M_2 &= \begin{array}{c} x \\ y \\ z \end{array} \begin{bmatrix} a & b & 0 \\ 0 & c & d \\ 0 & 0 & e \end{bmatrix} . a, c \leq -1 \text{ and } b, d, e > 0,
\end{aligned}
$$

where the balance of each row is equal to s.

The first one, M_1, is the simplest one. In fact, the reinforcement matrix can be divided into two smaller ruinous matrices: the first one, $\begin{bmatrix} a & s-a \\ 0 & s \end{bmatrix}$, describes the behavior of $x-$balls; while the second, $\begin{bmatrix} c & d \\ 0 & e \end{bmatrix}$, is related to y and $w-$balls.

The second case, M_2, can be divided into two matrices, but in this situation some problems about $y-$balls occur. In particular one has to model the dependence among the two smaller matrices that both influence $x-$balls.

2.4 Conclusions

In this chapter we have introduced the so-called Ruinous Urn, in order to model shock models that incorporate obsolescence.

The idea is to make the probability of default increase not only by direct reinforcement, but also by decreasing the probability of surviving.

Further analysis is needed for higher dimensional versions of the urn. This is a first step in order to produce useful applications.

Chapter 3
Applying Urn-based GESM's to Firms defaults

In the literature, there are several methods to estimate the probability of firms' default, from simple judgement-based methods to more complicated artificial intelligence systems and statistical regression models.
In this chapter[13] we propose a first simple application of the urn-based generalized extreme shock model we have presented.

3.1 Introduction

Modelling the probability of firms' default is a crucial issue for banks and financial institutions: assessing credit reliability of current and prospective counterparts is essential in defining loans operations in the banking business.

In the literature, one can find several methods to estimate the probability of firms' default, from simple judgement-based methods (see, for example, [72]) to more complicated artificial intelligence systems (a review in [77]) and statistical regression models ([24], [29]). The implementation of a reliable model to forecast the probability of default involves the determination of a period of observation. In order to identify default events, one has to control each debtor over time, identifying each transition from a non-default to a default state.

[13] This chapter is based on the paper "Modelling Credit Default with Generalized Extreme Shock Models" by P. Cirillo and J. Hüsler.
The paper has been presented at the Economic Lunch Seminar Series of the University of Pisa, Department of Economics.

In this chapter we propose a new stochastic model for the probability of firms' default starting from the results in [18], presented in chapter 1, and in [19]. As seen, this constructive model is based on an urn process, characterized by a special essentially Polya[14] triangular reinforcement matrix. In particular, assuming that a firm can experience three different levels of risk (safe, risk and default), we introduce a dependence among the last two levels, so that the probability of default increases every time the firm enters the risky state, while it decreases (but does not disappear) the more the firm spends in the non-risky one. The levels of risk are determined on the basis of aggregate balance indices. Figure 3.1 shows an example of the process.

As far as the results are concerned, we are both able to predict firms' default probabilities with a good degree of approximation (using Altman's Z-score as a benchmark) and to obtain limit distributions that nicely reproduce the empirical results one can find in the literature (see [38]).

3.2 A Simple Financial Application

We now present a first simple application of the urn-based generalized extreme shock model.

Despite its simplicity, the results are comparable with those of the actual market benchmark, that is Altman's Z-score.

[14] An urn is defined essentially Polya if, after normalisation, the maximum eigenvalue is equal to the unity with a multiplicity greater or equal to 2.

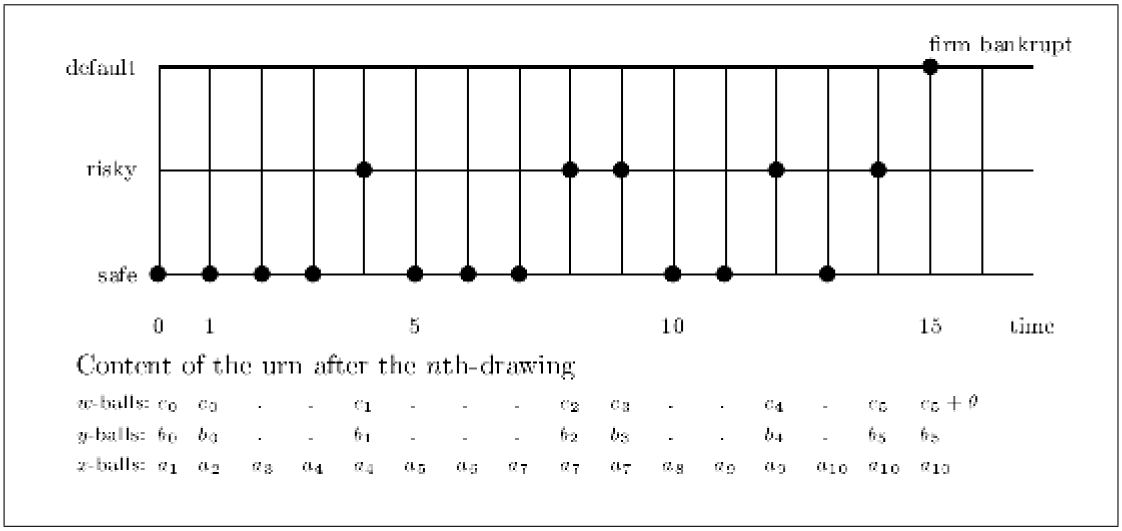

Figure 3.1: Example of the process with $a_k = a_0 + k\theta$, $b_k = b_0 + k\delta$ and $c_k = c_0 + k\lambda$.

The possibility of having asymptotic approximations for moments and probabilities can be seen as a useful tool for the analysis. In fact, it allows to explicitly compute the desired quantities without performing long and tedious simulations.

3.2.1 The Data

For our analysis, we have used data from the CEBI database. CEBI is a comprehensive database first developed by the Bank of Italy and now maintained by Centrale dei Bilanci Srl. It represents the biggest Italian industrial dataset and it contains firm-level observations and balance sheets of thousands of firms.

Our experiment has been performed with a subset of 380 manufacturing firms that respect the following conditions:

1. All firms' data are active in the period 1982-2000;

2. For every firm we have reliable information about capital and financial ratios and each has more than 100 employees;

3. They have been under bank control for possible insolvency at least once.

As far as the microstructure is concerned, the selected firms are comparable with those originally used by Altman in his famous paper [4], that we use as the benchmark in this work.

For every firm we have considered the following standard balance ratios: working capital / total assets (r_1), retained earnings / total assets (r_2), EBIT / total assets (r_3), market value of equity / book value of total liabilities (r_4), sales / total assets (r_5), equity ratio (r_6) and debt ratio (r_7).

3.2.2 Initialization of the process

In order to use our urn model in applications it is fundamental to decide the way the process is initialized and reinforced.

As far as reinforcement, we have simply set $\theta = 3$ and $\delta = 2$ for every firm. These values are the best we have obtained using a simple grid method. In particular, we have run our model several times, making the parameters change, but always respecting conditions 1 and 2, until we have found the best combination of parameters, that is the one that minimizes the distance between the actual distribution of defaults and the one given by the model, using the supremum metrics. The grid method is obviously a very naive way to calibrate a model's parameters. An alternative efficient and robust method could be represented by indirect inference, a simulation-based estimation technique introduced in [42]. Anyway, view that this is a first application and considering that this approach is quite common in urn models (see [52] or [6]), we prefer not to perplex our first computations.

This combination also assures that our reinforcement matrix RM is semi-simple and essentially Polya.

As far as the initial composition of firms' urns, we have developed the following heuristic method, based on well-known stylized facts of industrial economics. Suitable references can be found in [10], [11], [38], [39] and [65].

We first consider the equity ratio as a proxy of the proportion of $x-$balls. In fact, it is known that a firm with an equity ratio greater than 0.5 can be considered financially robust (see [38]). Hence, for every firm we trivially have $a_0 = round(r_6 * 100)$.

As far as the y and $w-$balls, expressing the risky and the default states, we combine the debt ratio with the complement of the equity ratio. Unlike the equity ratio, a higher debt ratio can be read as a signal of danger for firms' reliability (see [38], [39] and [65]), because of the excessive financial exposure. In particular, we have set $b_0 = round\left[r_7 * (1 - r_6) * 100\right]$ and $c_0 = round\left[(1 - r_6)(1 - r_7) * 100\right]$.

firm code	**year**	**equity ratio** r_6	**debt ratio** r_7	$\mathbf{a}_0$	$\mathbf{b}_0$	$\mathbf{c}_0$	**default in** $t+1$
$IM223A$	1982	0.42	0.71	42	41	17	0
$IM298A$	1982	0.62	0.52	62	20	18	0
$IM567B$	1982	0.68	0.37	68	12	20	1
$IM1031B$	1982	0.39	0.66	39	40	21	1
$IM1988A$	1982	0.72	0.57	72	16	12	0

Table 3.1: Example of initial urn composition for some firms of the dataset

It is easy to verify that, for every firm, the initial number of balls in the urn is equal to 100.

Let us analyze some firms, in order to better understand the data and the procedure. Table 3.1 summarizes the fundamental data for 5 different firms: using the equity ratio and the debt ratio, it is simple to derive $a_{0,}$ b_0 and c_0, by simply applying the formula we have proposed. The last column represents a dummy variable we have created to control for defaults (0 no default, 1 default).

If we consider all firms together, the average numbers of initialized x, y and $w-$balls in the urns are 48, 39 and 13 respectively. Figure 3.2 shows the distributions for $a_{0,}$ b_0 and c_0 all over the firms.

3.2.3 Results

Given the initial composition and the reinforcement matrix, for every firm we are able to compute all the probabilities we are interested in.

Using the results of Sections 2 and 3, it is possible to derive the probability of having a certain amount of x-balls or the probability of picking a w-ball (that's actual default) at time n. In a nutshell, we can say if a firm is in a safe, risky or default

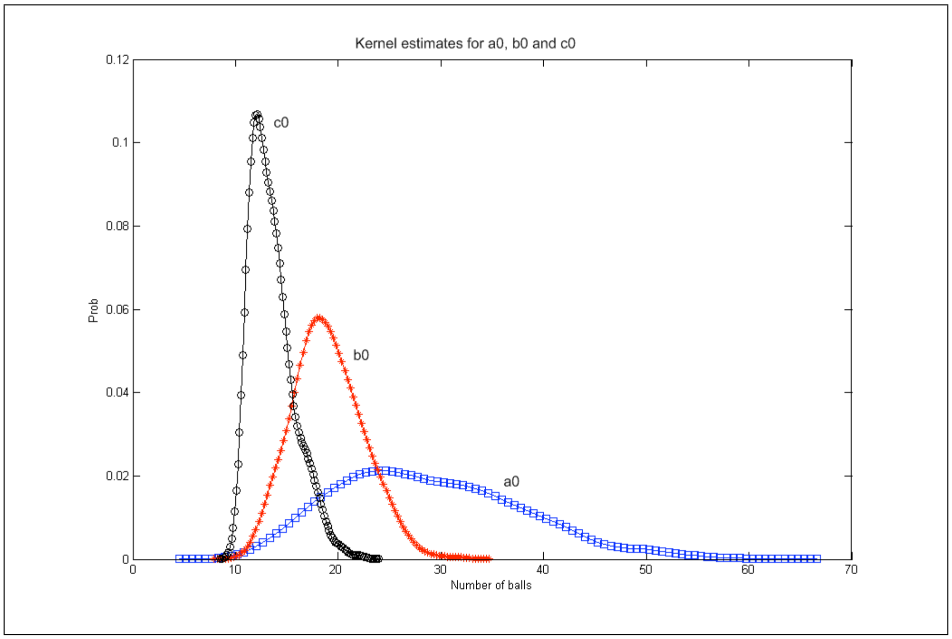

Figure 3.2: Kernel estimates for the distributions of a_0, b_0 and c_0.

level. In particular, in this experiment, we assume that a firm fails at time $n+1$ if the probability of extracting a w-ball is equal to or greater than 0.20 at time n, that represents a very common threshold value in studying firms' dynamics (see [39]).

For every firm, in every period, we have compared our predictions about failures with actual data and simple Altman's ones.

Altman's Z-score [4] is a popular measure, based on discriminant analysis, to classify firms' riskiness. In particular, using Altman's original formulation we have

$$Z = 0.012r_1 + 0.014r_2 + 0.033r_3 + 0.006r_4 + 0.999r_5.$$

According to this score, a firm is likely to default if $Z < 1.8$, it is safe if $Z > 3$, while it is in a "gray" condition otherwise.

Obviously we have also directly estimated the Z-score on CEBI data set to understand if its general formulation can be used on Italian data. The result we have obtained using standard regression techniques is the following

$$Z^* = \underset{(0.0062)}{0.014}\, r_1 + \underset{(0.0057)}{0.013}\, r_2 + \underset{(0.039)}{0.052} r_3 + \underset{(0.0028)}{0.007}\, r_4 + \underset{(0.3413)}{0.955}\, r_5,$$

where in brackets one can read the standard errors. As one can easily see, in our estimation exercise the sensitivity of r_3 is not significant.

In order to verify if there are substantial differences among our estimates and Altman's, we have imposed the standard restrictions and constraints on the two equations. The results are a little ambiguous and it is not possible to state that the differences are clear but, at the same time, we cannot completely reject them. For these reasons, in order to simplify the analysis, we have preferred to use the original Z-score formulation, as commonly done by practitioners. The subsequent results do not significantly change even if we use our estimates of Altman's Z.

For 245 firms from 380, our model seems to perform as good as Altman's Z-score. In particular, both methods correctly predict firms' default. However, for 72 firms from 380, our model seems to behave better. As far as the remaining 63 firms, both Altman's Z (apart from 8 correctly predicted firms) and the Ugesm are not able to predict default in the right way. However, while Altman's method usually underestimates the possibility of a failure, our model seems to be more pessimistic: for 40 firms from 63 we generally predict failure 2-3 periods before actual default. This is not correct, but it can be considered as a good result if we take the point of

	Ugesm	**Z-score**
correct	83%	66%
no correct	17%	34%

Table 3.2: Comparison of the number of correctly predicted defaults for Ugesm and Altman's Z-score

	Ugesm	**Z-score**
underestimated	63% (2.7)	28% (1.4)
overestimated	37% (3.2)	72% (2.3)

Table 3.3: Analysis of the 63 not correctly predicted firms. In brackets the average number of periods of difference between actual defaults and predictions

view of a bank or of another financial institution: a more prudent behavior is required for this kind of companies (consider Basel II, for example).

Table 3.2 compares our findings (the Ugesm) with the ones we have obtained using the Z-score. In particular we show:

1. The percentage of correctly reproduced firms (correct), which means those firms whose behavior is rightly predicted;

2. The percentage of incorrectly reproduced firms (not correct), that indicates those firms whose probability of failing is not correctly estimated and that, according to our calculations, have gone bankrupt before or after the actual time.

Table 3.3, on the contrary, focuses on our 63 not correctly predicted firms, analyzing the number of underestimated and overestimated defaults. In particular, we show how many defaults are wrongly expected before actual failure and how many after, also giving information about the average period of error.

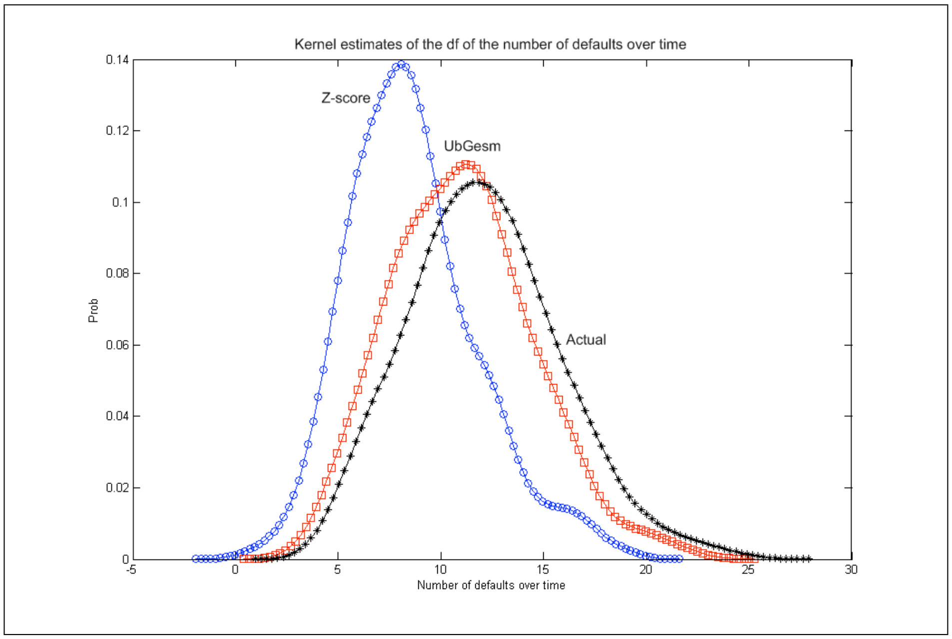

Figure 3.3: Kernel estimates of the number of defaults over time using the Ugesm and Altman Z-score and comparison with actual data.

Finally, in Figures 3.3 and 3.4, we give some results about the time pooled distribution of the number of defaults in the panel. It is easy to verify that the Ugesm seems to better reproduce the actual defaults' dynamics, showing a more similar shape. Kernel estimates have been obtained using the Epanechnikov kernel.

3.3 Conclusions

In this chapter we have proposed a new way for analyzing firms' default probabilities using a special triangular urn process. The reinforcement matrix models a positive dependence between the risky and the default states. The choice of the analytic approach to the urn has given us the possibility of explicitly deriving probabilities, mo-

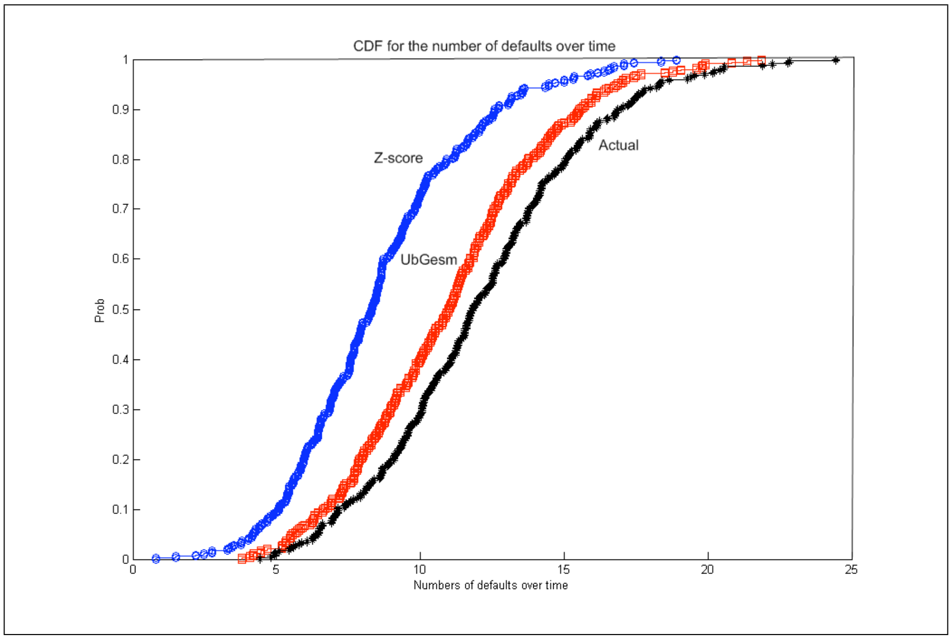

Figure 3.4: Cumulative distribution functions for the number of defaults over time in actual data and the two models.

ments and limit laws for the three different types of balls, that is for the three distinct levels of risk.

This first application has shown the good capabilities of the urn-based generalized extreme shock model in applications. In particular, we can be rather pleased with the obtained results, if we consider that our model behaves better than the one of the standard market benchmarks.

Moreover, the flexibility of the urn approach seems to be quite interesting, since it allows to incorporate prior knowledge into the process, opening the analysis to the Bayesian perspective. In fact, modifying the initial composition of the urn and the reinforcement quantities, the researcher can actively introduce his knowledge of the phenomenon into the modeling.

Obviously, a lot of improvements are needed. In particular, we need to improve the calibration of the reinforcement matrix parameters, using more statistically robust techniques. As already said, we might use indirect inference to estimate θ and δ, as we will show in a forthcoming paper.

Finally, as far as the data, it could be interesting to extend our studies to more general firms. In fact, considering firms that have been under bank control surely simplifies our analysis, since their probability of defaults is surely higher than the average.

Part II

Urn-based Default Models

Chapter 4
A Multidimensional Default Model for Failing Systems

In this chapter we introduce two new models for studying failing systems. A failing system (FS) is a system whose probability of failure is not negligible in a fixed time horizon (for example one year). The interest for such a topic is due to its diffusion in real-life problems: financial portfolios, electrical and mechanical systems, firms, the world wide web (and so on) can all be considered failing systems.

The main purpose when studying a FS is to calculate the probability of default and the distribution of the number of failures that may occur during the observation period.

Here we introduce a new way for calculating the probabilities of joint defaults in k different homogeneous groups of defaulting systems, when each group is characterized by some sort of rating, that's some external information about its reliability.

In our modelization, we will make use of a multidimensional urn scheme. We finally give an idea of one possible generalization.

4.1 The Multidimensional Urn Model

We now introduce a simple default model based on the Polya urn mechanism[15]. The properties of this well-known sampling scheme will help us in calculations, even if the use of a particular superstructure introduces some minor differences. The results we achieve are interesting for two different reasons: they serve as starting point for future research and they show to be useful in applications, as we will see in the last

---

[15] This chapter represents the evolution of a basic model [16], I developed when I was a student of Sant'Anna School of Advanced Studies in Pisa.

chapter of this thesis. Finally, even in this case, the urn modelization represents an intuitive tool to model dependence.

Consider N failing systems divided into k homogeneous groups. Each group will consist of n_j elements, such that $\sum_{j=1}^{k} n_j = N$. Assume that each group is characterized by a rating γ_j, $j = 1, ..., k$, i.e. some sort of information about the reliability of its components. A fundamental hypothesis of our analysis is that the set $G = \{\gamma_j : j = 1, ..., k\}$ is a poset. In other words, there exists a relation $\precsim$ on G which is reflexive, antisymmetric and transitive or, formally,

$$\gamma_r \precsim \gamma_r$$

$$\text{if } \gamma_r \precsim \gamma_s \text{ and } \gamma_s \precsim \gamma_r \text{ then } \gamma_r = \gamma_s$$

$$\text{if } \gamma_r \precsim \gamma_s \text{ and } \gamma_s \precsim \gamma_t \text{ then } \gamma_r \precsim \gamma_t.$$

Without any loss of generality, we will assume that the k group are completely ordered according to their ratings and, in particular, $\gamma_1 \succsim \gamma_2 \succsim ... \succsim \gamma_{k-1} \succsim \gamma_k$. We will read the relation $\succsim$ as "better than" so, for example, $\gamma_1 \succsim \gamma_2$ means that group 1 is "better than" group 2 and, as a consequence of this, the $k-$th group is the worst one, since it is characterized by the lowest reliability.

Let U represent a multicolored Polya urn with $k+1$ different types of balls and reinforcement matrix

$$RM = (\delta + 1) \cdot I_{k \times k},$$

where I is the identity matrix.

For each color $i = 1, ..., k, k+1$, there are α_i balls into the urn, so that the initial composition A is equal to $\sum_{i=1}^{k+1} \alpha_i$.

Starting from this well-known urn scheme, we will now describe a model for the analysis of joint default probabilities. For the sake of simplicity, we will assume that each FS can be characterized by only two states: safety and default.

The mechanism is as follows:

1. Look at all the failing systems and choose one of them. Obviously, it is part of one of the k groups and thus it is associated to a rating γ_j, $j = 1, ..., k$.

2. Now consider the multicolor Polya urn: pick a ball and look at its color i;

3. If the color is such that $i \in \{1, ..., j\}$ then the system fails and must be removed, otherwise if $i \in \{j+1, ..., k+1\}$ the systems is safe;

4. Return the ball in the urn and add δ extra balls of the same color.

5. Choose another failing system (even the same) and return to step 1.

In this way we are clearly introducing a positive dependence among the defaults.

Let $\mathbf{Y}_n$ be a random unit vector with $k+1$ rows representing the color of the $n-$th sampled ball, i.e.

$$\mathbf{Y}_n = \left[\underset{1}{0}, \underset{2}{0}, ..., \underset{j-1}{0}, \underset{j}{1}, \underset{j+1}{0}, ..., \underset{k+1}{0} \right]' \text{ if the } n-\text{th sampled ball has color } j.$$

Let $\mathbf{C}_n = [C_{n,1}, ..., C_{n,k+1}]$ be the vector of the conditional probabilities that the $n-$th sampled ball shows a particular color, with

$$C_{n,i} = P\left(\mathbf{Y}_n = \left[\underset{1}{0}, ..., \underset{i}{1}, ..., \underset{k+1}{0}\right]' | \mathbf{Y}_1, ..., \mathbf{Y}_{n-1}\right).$$

For convenience, in what follows, set $u_i = \left[\underset{1}{0}, ..., \underset{i}{1}, ..., \underset{k+1}{0}\right]'$. Thanks to the underlying Polya urn scheme, it is possible to state the following proposition, whose proof is already known (see [6] for details).

Proposition 21 *The random vectors $(\mathbf{Y}_n)_{n\geq 1}$ are exchangeable. In particular:*

$$P[\mathbf{Y}_1 = u_1, ..., \mathbf{Y}_n = u_n] = \frac{\prod_{j=1}^{k+1} \prod_{i=0}^{\epsilon_{n,j}-1} (\alpha_j + i\delta)}{\left(\frac{A}{\delta}\right)^{(n)}},$$

where $\epsilon_n = \sum_{i=1}^{n} u_{j_i}$ and $(\cdot)^{(l)}$ is the standard Pochammer formula.

Moreover, the random vector $\mathbf{C}_n$ of the conditional probabilities shows interesting properties.

Proposition 22 *The sequence $(C_{n,j})_{n\geq 1}$ is bounded martingale for every $1 \leq j \leq k+1$. Almost surely, it converges to the limit*

$$C_j \overset{a.s.}{=} \lim_{n\to\infty} C_{n,j} = \lim_{n\to\infty} \frac{\alpha_j + \delta\left(\sum_{i=1}^{n-1} Y_{i,j}\right)}{A + \delta(n-1)} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} Y_{i,j}.$$

Hence $(\mathbf{C}_n)_{n\geq 1}$ is convergent martingale according to the filtration $\mathcal{F}_n = \sigma(\mathbf{Y}_1, ..., \mathbf{Y}_n)$ with $\mathcal{F}_1 = \{\emptyset, \Omega\}$.

Proof. By construction we have that $0 \leq P_{n,j} \leq 1$ is $\mathcal{F}_n-$adapted. For this reason, we only have to prove that $C_{n,j}$ is a martingale:

$$\begin{aligned} E\left[C_{n,j}|\mathcal{F}_{n-1}\right] &= \left[\left.\frac{\alpha_j + \delta\left(\sum_{i=1}^{n-1} Y_{i,j}\right)}{A+\delta(n-1)}\right| \mathcal{F}_{n-1}\right] \\ &= \frac{\alpha_j + \delta\left(\sum_{i=1}^{n-2} Y_{i,j}\right) + \delta E\left[Y_{n-1,j}|\mathcal{F}_{n-1}\right]}{A+\delta(n-1)} \\ &= \frac{\left(A+\delta(n-2)\right) C_{n-1,j} + \delta C_{n-1,j}}{A+\delta(n-1)} \\ &= C_{n-1,j}. \end{aligned}$$

■

Proposition 23 *Set $k \geq 1$ fixed. Let $\left(\mathbf{Y}_n\right)_{n\geq 1}$ be an exchangeable sequence of vectors with values in $\{u_1, ..., u_{k+1}\}$. Then there exist a random vector $(C_1, ..., C_{k+1})$ with values in $\left\{(c_1, ..., c_{k+1}) \in [0,1]^{k+1} \mid \sum_{i=1}^{k+1} p_i = 1\right\}$ such that:*

$$P\left[\left.\sum_{i=1}^{n} \mathbf{Y}_i = \mathbf{E}\right| C_1, ..., C_{k+1}\right] \overset{a.s.}{=} \Gamma(n+1) \prod_{j=1}^{k+1} \frac{C_j^{\epsilon_j}}{\epsilon_j!}$$

and

$$C_j \overset{a.s.}{=} \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} Y_{i,j} \quad 1 \leq j \leq k+1,$$

where $\sum_{j=1}^{k+1} \epsilon_j = n$ and $\mathbf{E} \in \mathbb{N}_0^{k+1}$.

Proof. The proof is a simple and direct application of de Finetti's theorem (see [25]) for Polya urn schemes (see [41]). ■

4.1.1 Number of defaults and joint probabilities

The previous proposition is very useful if we want to calculate the joint distribution of the numbers of defaults within the k different groups. Let $(D_1, ..., D_k)'$ be the vector of interest, where D_j represents the number of failures within the $j-$th group of FS's.

Now remember that $N = \sum_j n_j$, $\sigma(C_1, ..., C_{k+1}) \subseteq \mathcal{F}_\infty$ and that w.r.t $(C_1, ..., C_k)$ the Y_n's are i.i.d.

It is straightforward to see that

$$P[D_1 = \epsilon_1, ..., D_k = \epsilon_k] =$$
$$= P\left[\sum_{i=1}^{n_1} Y_{i,1} = \epsilon_1, \sum_{i=n_1+1}^{n_1+n_2}\sum_{j=1}^{2} Y_{i,j} = \epsilon_2, ..., \sum_{i=N-n_k+1}^{N}\sum_{j=1}^{k} Y_{i,j} = \epsilon_k\right]$$
$$= E\prod_{j=1}^{k}\left[\binom{n_j}{\epsilon_j}\left(\sum_{i=1}^{j} C_i\right)^{\epsilon_j}\left(1-\sum_{i=1}^{j} C_i\right)^{n_j-\epsilon_j}\right].$$

Hence, looking at the distribution of $(C_1, ..., C_{k+1})$, we have

$$\begin{aligned} E\left[C_1^{\epsilon_1}\cdots C_{k+1}^{\epsilon_{k+1}}\right] &= E\left[P\left[\bigcap_{j=1}^{k+1}\left(\bigcap_{i=\epsilon_1+...+\epsilon_{j-1}+1}^{\epsilon_1+...+\epsilon_j}\{Y_i = u_j\}\right)\middle| C_1, ..., C_{k+1}\right]\right] \\ &= \frac{\prod_{j=1}^{k+1}\left(\frac{\alpha_j}{\delta}\right)^{(\epsilon_j)}}{\left(\frac{A}{\delta}\right)^{(n)}}. \end{aligned}$$

The random vector $(C_1, ..., C_{k+1})$ shows a bounded support, so its moments characterize its distribution function. In particular, as demonstrated in [14], we have that

$$f_{k+1}(c_1, ..., c_{k+1}) = \frac{\Gamma(\sum_{j=i}^{k+1}\frac{\alpha_j}{\delta})}{\prod_{j=1}^{k+1}\Gamma\left(\frac{\alpha_j}{\delta}\right)}\prod_{j=1}^{k+1} c_j^{\frac{\alpha_j}{\delta}-1},$$

where $\frac{\alpha_j}{\delta} > 0$, $0 \leq j \leq k+1$ and $\sum_{j=1}^{k+1} c_j = 1$. So $(C_1, ..., C_{k+1})$ is Dirichlet distributed $\mathcal{L}_{k+1}(\frac{\alpha_1}{\delta}, ..., \frac{\alpha_{k+1}}{\delta})$.

This last result is very useful when calculating the marginal distribution of D_i, that is the number of defaults in the $i-$th group. In fact, using the amalgamation property of the Dirichlet distribution (see [41]), we have that $\sum_{i=1}^{j} C_i \sim Beta\left(\widetilde{a}_j, \widetilde{b}_j\right)$, where $\widetilde{a}_j = \sum_{i=1}^{j} \frac{\alpha_j}{\delta}$ and $\widetilde{b}_j = \sum_{i=j+1}^{k+1} \frac{\alpha_j}{\delta}$. Hence

$$\begin{aligned} P\left[D_j = \epsilon_j\right] &= E\left[\binom{n_j}{\epsilon_j}\left(\sum_{i=1}^{j} C_i\right)^{\epsilon_j}\left(1-\sum_{i=1}^{j} C_i\right)^{n_j-\epsilon_j}\right] \qquad (4.1) \\ &= \binom{n_j}{\epsilon_j}\frac{B(\widetilde{a}_j+\epsilon_j;\widetilde{b}_j+n_j-\epsilon_j)}{B(\widetilde{a}_j;\widetilde{b}_j)} \end{aligned}$$

and we find the so-called beta-binomial (or compound binomial) distribution, which is very common in risk analysis. The interested reader can find more details in [33] and [41].

At this point we are ready to state the following:

Proposition 24 *For the multidimensional urn model we have constructed, the joint default probability $P\left[D_1 = \epsilon_1, ..., D_k = \epsilon_k\right]$, that is the joint probability of the numbers of failures in the different rating groups, is distributed according to a mixture of beta-binomial distributions.*

Proof. We have seen that

$$\begin{aligned} &P\left[D_1 = \epsilon_1, ..., D_k = \epsilon_k\right] = \\ &= E\prod_{j=1}^{k}\left[\binom{n_j}{\epsilon_j}\left(\sum_{i=1}^{j} C_i\right)^{\epsilon_j}\left(1-\sum_{i=1}^{j} C_i\right)^{n_j-\epsilon_j}\right]. \end{aligned}$$

Then using simple manipulations and the relationships we have underlined, we can continue showing that

$$E\prod_{j=1}^{k}\left[\binom{n_j}{\epsilon_j}\left(\sum_{i=1}^{j}C_i\right)^{\epsilon_j}\left(1-\sum_{i=1}^{j}C_i\right)^{n_j-\epsilon_j}\right]=$$

$$=\prod_{j=1}^{k}\binom{n_j}{\epsilon_j}\cdot\int_0^1\int_0^{1-p_1}\cdots\int_0^{1-\sum_{i=1}^{k-1}p_i}\left[\left(\left(\sum_{j=1}^{k}c_j\right)^{\epsilon_j}\left(1-\sum_{j=1}^{k}c_j\right)^{n_j-\epsilon_j}\right)\cdot\right.$$
$$\left.\cdot\frac{\Gamma(\sum_{h=i}^{k+1}\frac{\alpha_h}{\delta})}{\prod_{h=1}^{k+1}\Gamma\left(\frac{\alpha_h}{\delta}\right)}\prod_{h=1}^{k}c_h^{\frac{\alpha_h}{\delta}-1}\left(1-\sum_{h=1}^{k}c_h\right)^{\frac{\alpha_{k+1}}{\delta}-1}\right]dc_kdc_{k-1}\cdots dc_1=...$$

Now, remembering equation 4.1 and the properties of the beta function B, we get

$$\begin{aligned}... \quad = \quad & \prod_{j=1}^{k}\binom{n_j}{\epsilon_j}\frac{\Gamma(\sum_{h=i}^{k+1}\frac{\alpha_h}{\delta})}{\prod_{h=1}^{k+1}\Gamma\left(\frac{\alpha_h}{\delta}\right)}\sum_{s_1=0}^{\epsilon_k}\binom{\epsilon_k}{s_1}B\left(\tau_k+s_1;\tau_{k+1}+n_k-\epsilon_k\right)\cdot \\ & \cdot\sum_{s_2=0}^{\epsilon_k+\epsilon_{k-1}-s_1}\binom{\epsilon_k+\epsilon_{k-1}-s_1}{s_2}B\left(\tau_{k-1}+s_2;\tau_{k+1}+\tau_k+n_{k-1}+n_k-\epsilon_k-\epsilon_{k-1}+s_1\right)\cdot \\ & \cdot\sum_{s_{k-1}=0}^{\widetilde{\epsilon}_{2,k}-\widetilde{s}_{1,k-2}}\binom{\widetilde{\epsilon}_{2,k}-\widetilde{s}_{1,k-2}}{s_{k-1}}B\left(\tau_2+s_{k-1};\widetilde{\tau}_{3,k+1}+\widetilde{n}_{2,k}-\widetilde{\epsilon}_{2,k}+\widetilde{s}_{1,k-2}\right)\cdot \\ & \cdot B\left(\tau_1+\widetilde{\epsilon}_{1,k}-\widetilde{s}_{1,k-1};\widetilde{\tau}_{2,k+1}+\widetilde{n}_{1,k}-\widetilde{\epsilon}_{1,k}+\widetilde{s}_{1,k-1}\right)\end{aligned} \tag{4.2}$$

where $s_{-1}=s_0=0$, $\tau_j=\frac{\alpha_j}{\delta}$ and, in general, $\widetilde{x}_{a,b}=\sum_{i=a}^{b}x_i$. ■

4.1.2 Some more results about the number of defaults

We have seen that the random vector $(C_1,...,C_{k+1})$ is Dirichlet distributed $\mathcal{L}_{k+1}(\tau_1,...,\tau_{k+1})$. This result, that is one of the consequences of using Polia-like urns, can be used to deeply analyze the descriptive statistics of the number of defaults.

First note that, trivially,

$$E\left[Y_{i,j}\right] = E\left[C_j\right] = \frac{\tau_j}{\sum_{i=1}^{k+1} \tau_i} \tag{4.3}$$

and

$$Var\left(Y_{i,j}\right) = E\left[C_j\right]\left(1 - E\left[C_j\right]\right). \tag{4.4}$$

Hence, it is straightforward to see that

$$E\left[D_s\right] = E\left[\sum_{j=1}^{n_s}\sum_{i=1}^{s} Y_{j,i}\right] = n_s \sum_{i=1}^{s} E\left[C_i\right] = n_s \frac{\sum_{i=1}^{s} \tau_i}{\sum_{j=1}^{k+1} \tau_j}.$$

Analogously, as far as the variance of the number of defaults is concerned, remembering that

$$Cov(C_i, C_j) = -\frac{\tau_i \tau_j}{\left(\frac{A}{\delta}\right)^2 \left(\frac{A}{\delta} + 1\right)},$$

we have

$$\begin{aligned} Var\left(D_s\right) &= Var\left(\sum_{j=1}^{n_s}\sum_{i=1}^{s} Y_{j,i}\right) \\ = \left(n_s^2 - n_s\right)\left(\sum_{i=1}^{s}\sum_{l=1}^{s} Cov(C_i, C_l)\right) &+ n_s \left(\sum_{i=1}^{s}\sum_{l=1}^{s} Cov(Y_{j,i}, Y_{j,l})\right) \\ = \frac{n_s\left(n_s + \frac{A}{\delta}\right)}{\left(\frac{A}{\delta}\right)^2 \left(\frac{A}{\delta} + 1\right)} &\left(\sum_{i=1}^{s} \tau_i \left(\frac{A}{\delta} - \tau_i\right) - \sum_{l=1, l \neq i}^{s} \tau_i \tau_l\right). \end{aligned}$$

Using similar calculations, it is then possible to compute the covariance and all the other descriptive statistics one may be interested in. All this can be very useful in applications, as we will see in the last chapter.

4.2 A Possible Generalization

We now give a sketch about a generalized multidimensional urn scheme, starting from the reinforced urn process approach proposed in [60].

The main idea is to merge the positive aspects of both constructions, creating an interesting Bayesian approach to default analysis. The study of the model is still going on and this section represents only a possible development of the basic approach we have introduced.

4.2.1 A brief introduction to RUP's

Reinforced urn processes (RUP's) are a particular class of stochastic processes introduced by Muliere, Secchi and Walker in [60]. Having in mind the definition of [23], they can be seen as a reinforced random walk on a state space of urns.

Their usefulness is strictly linked to their characteristics of reinforcement and partial exchangeability, which are fundamental for constructing prior distributions and perform Bayesian nonparametric analysis. As a matter of fact, reinforced urn processes can be seen as a new approach to Bayesian analysis, as underlined in [60] and [76].

The definition of a RUP is essentially based on four distinct elements, that fully describe its behavior.

Definition 5 (Muliere et al. [60]**)** *Let*

- *S be a countable state space;*

- *$E = \{c_1, ..., c_k\}$ $k \geq 1$ a finite set of colors;*

- *$U(x) = (n_x(c_1), ..., n_x(c_k))$ an unr composition function that maps S into the set of k-tuples of nonnegative numbers whose sum is strictly positive;*

- *$q : S \times E \to S$ a law of motion such that for every $x, y \in S$, there is at most one color $c(x, y) \in E$ such that $q(x, c(x, y)) = y$.*

Fix $X_0 = x_0 \in S$. If for $n \geq 1$ $X_{n-1} = x$, a ball is picked from the urn associated to x and its color c discovered. Then $X_n = q(x, c)$ and the urn is Polya reinforced.

At this point, the sequence $X = \{X_n,\ n \geq 0\}$ is defined Reinforced Urn Process with initial state x_0 and parameters S, E, U and q, or

$$X \sim RUP(x_0, S, E, U, q).$$

Now, for every $x \in S$, set

$$R_x = \{y \in S : n_x(c(x, y)) > 0\} \text{ and } R^{[0]} = \{x_0\}$$

to be the set of all the states one can reach from state x in one step.

Moreover define, for every $n \geq 1$,

$$R^{[n]} = \bigcup_{x \in R^{[n-1]}} R_x$$

and

$$R = \bigcup_{n=0}^{\infty} R^{[n]},$$

the latter representing the set of all the states visited by the Reinforced Urn Process.

Using the approach proposed in [26], it is possible to show that a Reinforced Urn Process is partially exchangeable in the sense of Diaconis and Freedman (see [26] and [60]).

Moreover, if $X \sim RUP(x_0, S, E, U, q)$ is recurrent, that is there is a positive probability of visiting a particular state in a finite number of steps, it can be demonstrated that the reinforced urn process is a mixture of Markov chains. In particular, there exist a probability measure μ on the set $\wp$ of stochastic matrix on $R \times R$ such that for every $n \geq 1$ and $(x_1, ..., x_n)$

$$P\left[X_0 = x_0, ..., X_k = x_k\right] = \int_{\wp} \prod_{j=0}^{n-1} \pi(x_j, x_{j+1}) \mu(d\pi).$$

So, given a random matrix Π of $\wp$ with distribution μ, set $\Pi(x)$ to be the $x-$th row of Π for $x \in R$, and $\alpha(x)$ the measure on R that assigns a mass $n_x(c)$ to $q(x, c)$ for every color c such that $n_x(c) > 0$ and 0 otherwise. Hence the following holds:

Theorem 25 *If the Reinforced Urn Process $X = \{X_n, n \geq 1\}$ is recurrent, the rows of Π are mutually independent random probability distributions on R and, for all $x \in R$, the law of $\Pi(x)$ is that of a Dirichlet process with parameter $\alpha(x)$.*

4.2.2 Back to defaults

We can now construct the so-called Reinforced Urn Process for Defaults or Rating RUP (RRUP), by combining the standard RUP approach with the multidimensional default scheme we have introduced in the previous section. The idea can be considered a generalization of the approach proposed in [5].

Consider the following elements:

- a countable state space S representing time (even in a fictitious sense);
- a set of $k+1$ colors $E = \{c_1, ..., c_{k+1}\}$, $k \geq 1$, as required by the multidimensional default urn scheme;
- an urn composition function $U(x) = (n_x(c_1), ..., n_x(c_k))$, whose Polya reinforcement is δ, that maps S into the set of $k+1$-tuples of nonnegative numbers whose sum is strictly positive;
- a sampling mechanism as the one described in the previous section, with the only difference that now, after sampling, we move to the next step urn;
- a law of motion such that

$$q(x, c_i) = \begin{cases} 0 \text{ if } i = 1 \\ x+1 \text{ if } i = 2, ..., k+1 \end{cases} .$$

This construction entails two important consequences:

1. At every time step we consider a different multidimensional default urn, whose behavior is equal to the one we have already described before. This is important in performing calculations. In particular, for every urn,

$$P\left[D_j = d_j\right] = \binom{n_j}{d_j} \frac{B(\widetilde{a}_j + d_j; \widetilde{b}_j + n_j - d_j)}{B(\widetilde{a}_j; \widetilde{b}_j)},$$

where $\widetilde{a}_j = \sum_{i=1}^{j} \frac{n_x(c_j)}{\delta}$ and $\widetilde{b}_j = \sum_{i=j+1}^{k+1} \frac{n_x(c_j)}{\delta}$.
For $P\left[D_1 = \epsilon_1, ..., D_k = \epsilon_k\right]$, equation 4.2 still holds, once we have simply changed the parameters according to the new construction;

2. Time is set to 0 and the process is re-initialized every time a ball of color c_1 is sampled. This choice is simply due to the construction of the process. In fact, thanks to the rating scheme, if a c_1 ball is picked up we necessarily have a default, view that it is associated to the best rating group.

As usual, we are interested in the joint number of defaults for the k risk groups. Embedding the multidimensional scheme into the reinforced urn process framework gives the possibility of easily obtaining these quantities. Some considerations can be made:

1. In every time period we can have at most one default;

2. If we consider the first rating group, that is the one associated with the re-initialization of the RRUP, we have that for every cycle/block of the process

(in the terminology of [60]) we can observe at most one default. So, in general, after m cycles we cannot have more than m defaults in the first rating group;

3. For the other risk groups we can have a greater number of defaults d but, for every block, they cannot exceed the number x of elapsed periods (i.e. $d < x$) before re-initialization.

It is not difficult to see that, when considering a RRUP, equation 4.2 must be modified introducing Stirling numbers of the first kind, in order to consider permutation cycles.

As we will see in the last chapter, there are no substantial problems in simulating the RRUP. The implementation in Fortran code, indeed, is a little messy, but not in the least impossible. At the moment we are working on a more efficient and time-saving algorithm. Most of the problems, while studying the RRUP, are due to the research of a compact combinatorial form for the results.

4.3 Conclusions

We have introduced a multidimensional default scheme based on a multicolored version of Polya urn. The sampling mechanism allows for building dependence among defaults, showing interesting results as far as probabilities are concerned.

We have also proposed a first generalization of the scheme using the reinforced urn process framework. Here we have presented some preliminary results, view that our research is still going on.

Chapter 5
The Interacting Urn Chain Model

This chapter is dedicated to a recursive model constructed by the means of interacting urns. The aim is to introduce a stronger dependence among defaults with respect to the multidimensional urn model we have previously presented. Some seminal ideas for such a construction have been introduced in [57] and [63], where systems of interacting Polya urns are discussed. Anyway, the main references for our interacting urn chain, which is based on an iterative framework of interacting urns, are [27], [61] and [76].
In particular, the basic ideas in [61] can be considered the very starting point of this work.
We first present the general model, with few assumptions and very general results. Then we impose some more conditions on urns' behavior obtaining more complete and applicable results.

5.1 General scheme

Suppose we have N failing systems divided into k different homogeneous groups of risk with n_i elements $i = 1, ..., k$.

Once again, assume that, thanks to some external information, there is the possibility of ordering the distint groups. In particular, we require that the first group is the best one and the last the worst, where the label "best" and "worst" are related to the probability of default C_i of the different groups. In general - and this is not a strong neither a ludicrous assumption - we want that $C_i < C_{i+1}$ for $i = 1, ..., k$.

Hence, the aim of this chapter is to model the dependence among failures both within and between the k groups, introducing a specific scheme based - what a surprise - on urns.

Having in mind the construction of [27], we consider $C_1^*, ..., C_k^*$ independent random variables such that $C_i^* \in (0,1)$. We then construct the probabilities of failure of the k groups as

$$C_1 = C_1^*$$

$$C_2 = C_1 + (1 - C_1)C_2^*$$

$$\vdots$$

$$C_k = C_{k-1} + (1 - C_{k-1})C_k^* = 1 - \prod_{i=1}^{k} (1 - C_i^*) \, .$$

It is easy to verify that:

1. this construction respects all our assumptions, so that the better groups of FS's show a lower probability of default;

2. the probabilities of default of the different groups are strictly linked together by the means of the recursive scheme.

So, thanks to this simple and intuitive iterative modelization, we obtain the probabilities of default for the different risk groups and, for every FS, we are able to say in which measure it is likely to fail.

We now introduce a particular class of stochastic processes, that will be very useful for our analysis.

Definition 6 (Doksum Doksum [27]) *The random distribution function F is said to be neutral to the right if for each $h > 1$ and $t_1 < ... < t_h$, there exist nonnegative independent random variables $V_1, ..., V_h$ such that*

$$(F(t_1), F(t_2), ..., F(t_h)) =_{\mathcal{L}} \left(V_1, 1 - (1 - V_1)(1 - V_2), ..., 1 - \prod_{i=1}^{h} (1 - V_i) \right).$$

The equations

$$F(t_j) = 1 - \prod_{i=1}^{j} (1 - V_i) \quad j = 1, ..., h$$

yield

$$F(t_j) - F(t_{j-1}) = V_j \prod_{i=1}^{j-1} (1 - V_i)$$

and

$$V_j = \frac{(F(t_j) - F(t_{j-1}))}{(1 - F(t_{j-1}))} \quad j = 1, ..., h \text{ and } t_0 = -\infty.$$

Thus "F is neutral to the right" mainly means that the normalized increments

$$F(t_1), \frac{(F(t_2) - F(t_1))}{(1 - F(t_1))}, ..., \frac{(F(t_h) - F(t_{h-1}))}{(1 - F(t_{h-1}))}$$

are independent for all the $t_1 < ... < t_h$.

Neutral to the right processes have been introduced to generalize Ferguson's Dirichlet processes. For their conjugacy property and construction, neutral to the right processes have become an essential tool of Bayesian statistics, especially for studying censored data.

Proposition 26 *The process that governs the probabilities of failure C_i $i = 1, ..., k$ is a neutral to the right process.*

Proof. First of all let $E_i = C_i - C_{i-1}$.with $i = 1, ..., k$ and $E_1 = C_1$. It is easy to verify that

$$(E_1, E_2, ..., E_k) \stackrel{d}{=} \left(C_1^*, C_2^* \left(1 - C_1^*\right), ..., C_k^* \prod_{i=1}^{k-1} \left(1 - C_i^*\right) \right). \tag{5.1}$$

This, as shown in remark $3.1b$ in [27], assures that the process governing $(C_1, C_2, ..., C_k)$ is neutral to the right. ■

Let D_i, $i = 1, ..., k$, be the number of failures in the $i-$th group with n_i elements. The (marginal) probability of having $D_i = d_i$ failures in the $i-$th group is equal to

$$P\left[D_i = d_i\right] = E\left[\binom{n_i}{d_i} C_i^{d_i} (1 - C_i)^{n_i - d_i}\right] = E\left[P\left[D_i = d_i | C_{i-1}\right]\right].$$

Then, using standard combinatorial considerations, the joint default probabilities can easily be computed as

$$\begin{aligned} P\left[D_1 = d_1, ..., D_k = d_k\right] &= E\left[\binom{n_1}{d_1} C_1^{d_1} (1 - C_1)^{n_1 - d_1} \binom{n_2}{d_2} C_2^{d_2} (1 - C_2)^{n_2 - d_2} \cdots \right. \\ &\qquad \left. \cdots \binom{n_k}{d_k} C_k^{d_k} (1 - c_k)^{n_k - d_k} \right] \\ &= E\left[\prod_{i=1}^{k} \left(\binom{n_i}{d_i} \widetilde{E}_i^{d_i} \left(1 - \widetilde{E}_i\right)^{n_i - d_i} \right)\right], \end{aligned}$$

where $\widetilde{E}_j = \sum_{i=1}^{j} E_i$.

These results, even if interesting, are not very useful in practice. However, they can be further developed if we assume some more information about the random variables $C_1^*, ..., C_k^*$.

5.2 Polya Urn Chain

We now want to deepen what we have found in the previous section.

Consider k two-color Polya urns. We will make them evolve and interact as described above.

At time $t = 1$ each urn contains $w_{i,1}$ white balls and $b_{i,1}$ black balls $i = 1, ..., k$. When sampled, every urn is Polya reinforced with r_i balls of the same color that has been selected[16]. In this way, at time t, each urn is equal to itself in $t - 1$ plus reinforcement.

Hence, as known (see [33] and [52]), we have that, using Polya urns, we can obtain $C^*_{i,t} \sim Beta\left(\frac{w_{i,t}}{r_i}, \frac{b_{i,t}}{r_i}\right)$. At this point, following the recursive scheme, once again we have

$$C_{1,t} = C^*_{1,t}$$

$$C_{2,t} = C_{1,t} + (1 - C_{1,t})C^*_{2,t}$$

$$\vdots$$

$$C_{k,t} = C_{k-1,t} + (1 - C_{k-1,t})C^*_{k,t} = 1 - \prod_{i=1}^{k} \left(1 - C^*_{i,t}\right).$$

This construction, also shown in Figure 5.1, has three interesting consequences:

1. The Polya urn scheme guarantees the dependence among the failures within the same group of risk;

[16] No need to describe the Polya sampling scheme once again.

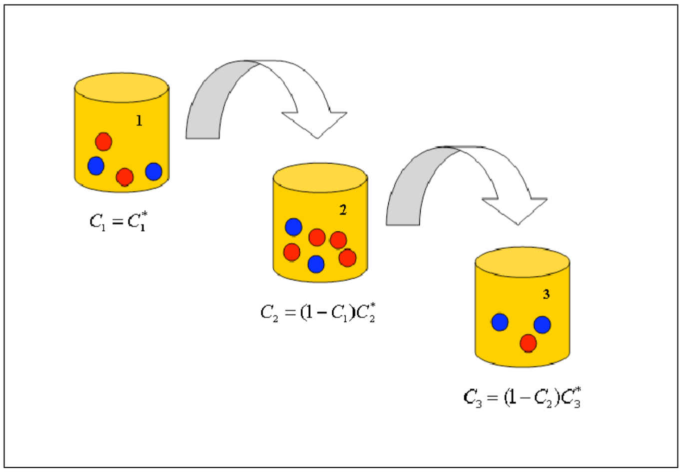

Figure 5.1: Graphical example of the urn chain mechanism.

2. The construction of the probabilities $C_1, ..., C_k$ induces dependence among the defaults of the different k groups. In fact, if the probability of default of a better group increases, automatically the probabilities of failure for the worse rating groups do the same.

3. The use of Polya urns allows for the incorporation of prior and empirical knowledge in the model, as we will see in the next chapter.

Given these details, we have that, at time $t = 1$, the number of default in the first group (the best one) follows, once again, the well-known beta-bernoulli distribution (see [14] and [33]), or

$$P(D_{1,1} = d_{1,1}) = \binom{n_{1,1}}{d_{1,1}} \frac{B(\frac{w_{1,1}}{r_1} + d_{1,1}; \frac{b_{1,1}}{r_1} + n_{1,1} - d_{1,1})}{B(\frac{w_{1,1}}{r_1}; \frac{b_{1,1}}{r_1})},$$

where - we remember - $n_{1,1}$ is the number of failing systems in group 1 at time $t = 1$.

Then, after some tedious calculations, it follows that the joint probability of failure, for $t = 1, 2, ...$, is given by[17]

$$P\left[D_1 = d_1, ..., D_k = d_k\right] = \binom{n_1}{d_1}\left[B(\tau_1;\varphi_1)\right]^{-1}\cdots\binom{n_k}{d_k}\left[B(\tau_k;\varphi_k)\right]^{-1}$$

$$\cdot\sum_{s_1=0}^{d_k}\binom{d_k}{s_1}B(\tau_k + d_k - s_1;\varphi_k + n_k - d_k)\cdot$$

$$\cdot\sum_{s_j=0}^{d_{k+1-j}+s_{j-1}}\binom{d_{k+1-j}+s_{j-1}}{s_j}B(\tau_{k+1-j} + d_{k+1-j} + s_{j-1} - s_j;\varphi_{k+1-j} + \widetilde{n}_{k+1-j,k} - d_{k+1-j} - s_{j-1})$$

$$\cdot\sum_{s_{k-1}=0}^{d_2+s_{k-2}}\binom{d_2+s_{k-2}}{s_{k-1}}B(\tau_2 + s_2 + s_{k-2} - s_{k-1};\varphi_2 + \widetilde{n}_{2,t} - d_{2t} - s_{k-2})$$

$$\cdot B(\tau_1 + d_1 + s_{k-1};\varphi_1 + n_{1,k} - d_1 - s_{k-1}),$$

where $\tau_j = \frac{w_{j,t}}{r_j}$, $\varphi_j = \frac{b_{j,t}}{r_j}$and $\widetilde{n}_{a,b} = \sum_{i=a}^{b} n_{i,t}$.

Proposition 27 *The process that governs the probabilities of failure C_i $i = 1, ..., k$ is a beta-Stacy process, as defined in* [76]*, characterized by the initial compositions of the k urns.*

Before giving the proof, let us spend some words about this new process.

The beta-Stacy process, introduced by [76], is a special case of neutral to the right process. The most interesting aspect of the beta-Stacy process is that it virtually encapsulates all the neutral to the right processes mentioned in the literature. In their work, Walker and Muliere [76] give both a discrete and continuous definition of the

---

[17] In what follows, in order not to abuse in notation, we will omit the time indication, so that $x_{1,t}$ is x_1.

process, whose name comes from the beta-Stacy distribution introduced in [21]. As far as the discrete case is concerned, it suffices to notice that our urn chain is nothing but the construction of a discrete beta-Stacy process by the means of Polya urns. For the continous case, on the contrary, we suggest [76] and [61] as suitable references.

Proof. Since the beta-Stacy process is a special case of neutral to the right process, when the independent variables are Betas, it is straightforward to prove the proposition. In fact, using equation 5.1 we known that

$$E_1 = C_1^*$$

$$E_2 = C_2^*(1 - C_1^*)$$

$$\vdots$$

$$E_k = C_k^* \prod_{i=1}^{k-1} (1 - C_i^*)\,.$$

An obvious consequence of this is that

$$\begin{aligned} E_1 &\sim \Im(\tau_1; \varphi_1; 1) \\ E_2|E_1 &\sim \Im(\tau_2; \varphi_2; 1 - E_1) \\ &\vdots \\ E_k|E_{k-1}, ..., E_1 &\sim \Im(\tau_k; \varphi_k; 1 - \sum_{j=1}^{k-1} E_j), \end{aligned}$$

where $\Im(a; b; c)$ is the so-called beta-Stacy distribution introduced by [58], whose density function is

$$\frac{1}{B(a,b)} x^{a-1} \frac{(c-x)^{b-1}}{c^{a+b-1}} I_{(o,c)}(x).$$

Hence the final result immediately follows. ∎

Corollary 28 *Given the construction above and* $E_i = C_i - C_{i-1}$ *we have that*

$$(E_1, E_2, ..., E_k) =_{\mathcal{L}} GD\left(\frac{w_1}{r_1}, ..., \frac{w_k}{r_k}; \frac{b_1}{r_1}, ..., \frac{b_k}{r_k}\right),$$

where GD *is the generalized Dirichlet distribution introduced in* [21].

Proof. This is a clear consequence of the beta-Stacy process. Remembering that this process generalizes the Dirichlet distribution, it suffices to apply the results in [76] and [61] for the proof. ■

Finally, it is possible to have some additional information about the behavior of the probabilities of default in the entire model, simply using the following immediate corollary.

Corollary 29 *For every* $k \geq 1$, $E_k < 1 - \sum_{j=1}^{k-1} E_j$. *Moreover*

$$E\left[\sum_{j=1}^{k} E_j\right] = \sigma_k + (1-\sigma_k)E\left[\sum_{j=1}^{k-1} E_j\right],$$

where $E_0 = 0$ *and* $\sigma_k = \frac{\tau_k}{\tau_h + \varphi_k}$.

5.3 Conclusions

In this chapter we have proposed a new model for studying the behavior of failing systems, when their defaults show some sort of dependence both within and between groups of homogeneous characteristics. In this way the urn chain model tries to address some problems that are not solved by wide used modelizations such as

CreditRisk+© by Credit Suisse or RiskMetrics© by JP Morgan Chase (see [30] for a reference).

As shown, the model presents a very general framework that can be simply specialized by better defining the properties of random variables $C_1^*, ..., C_k^*$. In particular, if we assume that they come from the famous Polya urn scheme, we are able to explicitly derive all the information we need about the number of defaults and the probabilities of failures. Obviously, properly changing the generation mechanism of the $C_1^*, ..., C_k^*$ variables, it is possible to get different result. An idea is to use Ehrenfest-like urns. This would be the object of future studies.

Finally, as we will show in the next chapter and in [20], the Polya Chain Model, using CreditRisk+© as a benchmark, performs surprisingly well, suggesting useful developments and applications.

Chapter 6
An application to Credit Risk Modelling

This chapter is devoted to the application of the multidimensional and the urn chain model to credit risk modelling.

This cutting-edge topic is quite important for banks and financial institutions, mainly because of the Basel II Accord, that contains recommendations on banking laws and regulations issued by the Basel Committee on Banking Supervision.

Credit risk, that represents one of the major risk factors for banks, is the risk of loss due to a debtor's non-payment of a loan or other line of credit.

In the literature there are a lot of models concerning credit risk and some of them are very used by practitioners. In particular, CreditRisk+© by Credit Suisse is a fundamental tool for assessing credit risk. For these reasons, we have chosen to use it as a benchmark.

6.1 Introduction

Credit risk is risk due to uncertainty in a counterparty's ability to meet its obligations. Since there are many types of counterparties and many different types of obligations, credit risk takes many forms and institutions manage it in different ways.

In general in assessing credit risk from a single counterparty, an institution must consider three issues:

1. the default probability, that is the likelihood that the counterparty will default on its obligation either over the life of the obligation or over some specified horizon;

2. the credit exposure, that is to say how large the outstanding obligation will be when the default happens;

3. the recovery rate, which represents the fraction of the exposure that can be recovered through bankruptcy proceedings or other similar methods.

In order to place credit exposure and credit quality - the counterparty's ability to perform on that obligation - in perspective, one has to remember that every risk comprises two elements: exposure and uncertainty. For credit risk, credit exposure represents the former, and credit quality represents the latter.

Many banks, investment managers and insurance companies hire their own credit analysts who prepare credit ratings for internal use. Other firms, such as Moody's and Standard and Poor's, are in the business of developing credit ratings for use by investors or other third parties.

There are many ways credit risk can be managed or mitigated. The "first line of defense" is the using credit scoring or credit analysis to avoid extending credit to parties that entail excessive credit risk. For these reasons several models, such as CreditRisk+©, RiskMetrics© and Credit Monitor©, have been developed and an interesting survey can be found in [3].

Here we have chosen to use CreditRisk+© as a benchmark, for it represents the most used approach to credit scoring. In the next section we spend some words to introduce this model.

6.2 A sketch on the CreditRisk+© Model

The CreditRisk+© model, developed by Credit Suisse, is a proprietary model described in [22].

The main aim of the model is to estimate the unexpected loss in a portfolio of credit obligations. The basic assumptions are:

- credit obligations are evaluated at their face values, ignoring market values;
- counterparties can experience only two states: default and non default;
- the recovery rate is not computed in CreditRisk+©, but it is assumed it has been applied before running the model;
- debtors are grouped into exposure bands, in which defaults are modelled using a Poisson process, whose intensity measure is given by the sum of the probabilities of default of all the counterparties in the group;
- the default processes of the different groups can be uncorrelated (base model) or correlated (volatile default rates model).

Moreover, the CreditRisk+© model can be implemented in two different ways:

- Analytical approach (or closed form model), in which the solution of the model is derived analytically, by using the fundamental properties of Poisson Processes.

- Monte Carlo approach, that uses simulations to create scenarios from which one can obtain the distribution of losses and apply the VaR methodology.

In this chapter we will only deal with the analytical approach, in order to have closed form solutions, as in our modelizations. Moreover we will consider the base model because of some constraints on data availability.

The starting point of the CreditRisk+© model is the following equation for the random portfolio loss L over all obligors A, that is

$$L = \sum_A 1_A \nu_A,$$

where ν_A is the exposure net of recovery and

$$1_A = \begin{cases} 1 \text{ if } A \text{ fails} \\ 0 \text{ otherwise} \end{cases}.$$

L is considered in units of l, where l is used as numeraire (see [22]).

Set p_A to be the unconditional default probability of obligor A. The dependence between obligors is incorporated in the common risk factor R which is assumed to be Gamma distributed with mean $\mu = \sum_A p_A$ and variance σ. In particular, R represents the intensity of the number of defaults in the economy as a whole. At this point it is possible to derive the expected loss E as

$$E = \sum_A p_A \nu_A.$$

In order to get the loss distribution, the CreditRisk+© model assumes that the events 1_A, conditionally on R, are Poisson variables. Hence, it turns out that the probability generating function $G(z) = \sum_{n=0}^{\infty} p(n)z^n$ of the portfolio loss is such that

$$G(z) = \left(1 - \sigma^2\mu^{-1}\left(\mu^{-1}\sum_A p_A z^{\nu_A} - 1\right)\right)^{-\frac{\mu^2}{\sigma^2}}.$$

Standard considerations (see [22] for further details) on the probability generating function $G(z)$ show that the coefficients of the power expansion are given by

$$p(n) = \frac{1}{(n+\sigma^2\mu^{-1})^{-\frac{\mu^2}{\sigma^2}}}\sum_{j=1}^{\min(m,n)}\left(\sum_{A:\nu_A=j} p_A\right)\left[\sigma^2\mu^{-2}n + \left(1-\sigma^2\mu^{-2}\right)j\right)\right]p(n-j),$$

where m is the largest exposure in the portfolio in terms of l units and

$$p(0) = (1+\sigma^2\mu^{-1})^{-\frac{\mu^2}{\sigma^2}}.$$

These equations can be further extended to the case of several segments, with, for example, a set of systematic risk factors $R_1, ..., R_N$. In particular, the extension for the case of independent Gamma distributed segments is given in [22].

Finally, as far as the number of defaults D is concerned, because of its assumptions, the CreditRisk+© model requires that, in every group $i = 1, ..., k$

$$P(D_i = d_i) = \frac{e^{-\mu_i}\mu_i^{d_i}}{d_i!},$$

where μ_i is the expected number of defaults in group i. So, because of the assumed independence of defaults between groups,

$$P(D_1 = d_1, ..., D_k = d_k) = \prod_{i=1}^{k} P(D_i = d_i) = \frac{e^{-\mu}\mu_i^d}{d!},$$

where $d = \sum_{i=1}^{k} d_i$. Hence the total number of defaults trivially follows a Poisson distribution.

6.3 The Data

The subsequent analysis is based on Moody's data in [59], about defaults and recovery rates of European corporate bond issuers from 1989 to 2006. These data have the following characteristics:

- The traditional 21 risk classes of Moody's are here collapsed into 4 greater groups: AAA-AA, A-BAA, BA-B and C firms. This is due to the not commercial but academic nature of the dataset;
- For every rating group we have the total number of firms and the number of defaults from 1989 to 2006.

Tables 6.1 and 6.2 aggregate and summarize the available data about defaults. For every risk class and for every year, the first number (above) represents the total amount of bond issuers in the class, while the second (below) is the number of actual defaults[18].

These data are sufficient for implementing our three default models (the multidimensional urn, the RRUP and the Polya urn chain) and for comparing them with the basic CreditRisk+© model, that is the purpose of this chapter. In what follows

---

[18] Note the 09-11 effects on firms' defaults: 2002 is the worst year for every risk class.

	1989	1990	1991	1992	1993	1994	1995	1996	1997
AAA-AA	304 2	312 0	298 0	302 0	325 1	326 0	333 0	321 0	330 1
A-BAA	601 0	605 4	627 0	631 0	644 3	657 2	653 0	657 0	666 1
BA-B	322 2	343 0	341 3	345 3	356 7	367 0	362 0	371 3	372 1
C	33 9	26 6	24 3	22 3	35 14	25 10	22 7	39 5	45 5

Table 6.1: Moody's data. Part I.

	1998	1999	2000	2001	2002	2003	2004	2005	2006
AAA-AA	342 0	342 0	338 0	345 0	323 2	380 0	402 0	409 1	409 0
A-BAA	365 1	374 0	377 3	378 2	383 6	392 0	386 2	381 0	377 1
BA-B	403 3	405 1	398 1	407 0	396 9	403 6	404 5	411 0	409 2
C	37 11	44 4	51 6	56 7	63 20	74 17	77 11	79 8	72 6

Table 6.2: Moody's data. Part II.

we will consider the pooled case, summing the different risk classes over the years. Obviously, a fundamental step for the analysis is to estimate the different parameters of interest, as we discuss in the next section.

6.4 Implementing the models

In describing how we have implemented the different models, starting from the actual data in [59], we can consider three different initializations: the one of the two multidimensional models, the one of the Polya urn chain and the one related to CreditRisk+©. Since we do not have an in-depth knowledge of the phenomena we are describing - we are not economists or financial gurus, we cannot express any a priori representation of the behavior of firms and their defaults. For these reasons, we do prefer to use actual data to objectively initialize the models.

6.4.1 The Multidimensional Urn and the RRUP

The Multidimensional Urn and the RRUP models are initialized in the same way. After all, the RRUP is nothing but a generalization of the simpler multidimensional case. In fact, at every time point, the different urns are standard Polya urns and this makes implementation easier. The main difference between the two model is given by the computational algorithms: for the RRUP model, indeed, the amount of calculation is sensibly greater.

From equation 4.2 and remembering the amalgamation property of the Dirichlet distribution, we know that the frequency of defaults for D_r is $\sum_{j=1}^{r} C_j$ and that

$$\sum_{j=1}^{r} C_j \underset{n\to\infty}{\to} \frac{1}{m}\sum_{l=1}^{m}(Y_{l,1},...,Y_{l,r}).$$

At this point, following a standard approach (see [52]), we can consider $\frac{d_r}{n_r}$ as a proxy of $\sum_{j=1}^{r} C_j$, where d_r represents the number of observed defaults in the group r, as in tables 6.1 and 6.2. Then, it follows that, setting $\tau_i = \frac{d_i}{n_i}$ for $i = 1, ..., k$,

$$(\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_2, ..., \tau_k - \tau_{k-1})$$

can be considered a sample from

$$(C_1, ..., C_{k+1}) \sim \mathcal{L}_{k+1}(\frac{\alpha_1}{\delta}, ..., \frac{\alpha_{k+1}}{\delta}).$$

	$AAA-AA$	$A-BAA$	$BA-B$	C	IN
$\tau's$	0.1123	0.8445	4.2611	17.6161	25.4346

Table 6.3: Dirichlet parameters for multidimensional models

This last results in an optimal starting point for the generalized expectation maximization (GEM) algorithm approach to Polya urns proposed in [71] and [73]. The GEM algorithm is a well-known estimation technique and there is no need to describe it here. A very complete discussion of the topic can be easily found in [40] or in any other book about computational statistics.

The main idea in [71] and [73], that we have adapted to our case, is to use the GEM algorithm to maximize the log-likelihood function one can deduce from equation 4.2, once assumed that the available observations are incomplete. In fact, given equation 4.2, it suffices to estimate the theoretical moments of equations 4.3 and 4.4, using a standard estimation method (e.g. the method of moments (see [40])), and then to run the GEM algorithm to get the desired likelihood.

Table 6.3 contains the estimates[19] for the starting multidimensional urn. Column IN is the instrumental group we have created by merging the AAA-AA and the C groups. As a matter of fact, our model needs $k + 1$ groups, so, in order to have estimates for the four main risk classes, we require an extra fictitious group.

[19] All the codes for computations have been written using Fortan language.

	$AAA-AA$	$A-BAA$	$BA-B$	C
$\psi's$	1.5234	2.2312	3.5677	5.2113
$\rho's$	8239.23	5485.19	427.94	98.381

Table 6.4: MLE estimates for the Polya Urn Chain

6.4.2 The Polya Urn Chain

Even for the Polya Urn Chain the vector $(\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_2, ..., \tau_k - \tau_{k-1})$ can be considered a realization of

$$(E_1, ..., E_k) \sim GD\left(\frac{w_1}{r_1}, ..., \frac{w_k}{r_k}; \frac{b_1}{r_1}, ..., \frac{b_k}{r_k}\right).$$

In this case, in order to apply the method of moments, we need to use the theoretical moments derived in [21] for the Generalized Dirichlet distribution and to compare them with the empirical ones. In particular, trivial but tedious calculations show that, setting $\psi_i = \frac{w_i}{r_i}$ and $\rho_i = \frac{b_i}{r_i}$,

$$\begin{aligned} \psi_i &= \frac{M_i(M_i - L_{i-1}) + S_i}{M_i(L_{i-1} - P_{i-1}) - P_{i-1}S_iM_i^{-1}} \\ \rho_i &= \frac{[M_i(M_i - L_{i-1}) + S_i]\,(P_i - M_i)}{M_i^2(L_{i-1} - P_{i-1}) - P_{i-1}S_i}, \end{aligned}$$

where M_i and S_i are the sample mean and variance, $P_j = \prod_{s=1}^{j} \frac{\rho_s}{\psi_s + \rho_s}$ and $L_j = \prod_{s=1}^{j} \frac{\rho_s + 1}{\psi_s + \rho_s + 1}$

Table 6.4 contains the estimates for the Polya Urn Chain.

6.4.3 CreditRisk+©

As far as CreditRisk+© is concerned, simply notice that tables 6.1 and 6.2 contain all the information we need to estimate the model parameters. In particular, given the total number of firms in the groups and the related number of defaults, it is very simple to derive the intensity measures we need to perform calculations as $\mu_j = \frac{d_j}{n_j}$, $j = 1, ..., k$.

6.5 The Results

We are now ready to analyze the results of our very demanding simulations, that have required awhile, using an euphemism. In particular, the hardest calculations have been those of the RRUP model. Anyway, the use of Fortran codes for the algorithms has sensibly reduced the running time, with respect to other languages used for the initial trials.

Figure 6.1 shows the comparison of the pdf's for the number of defaults in the AAA-AA risk classes using the different models and their properties about marginal distributions. As one can clearly see, the differences are extremely moderate. We can only appreciate a substantial similarity between the multidimensional urn model and the RRUP. All this is also confirmed by Figure 6.2, in which we consider the different cdf's and we compare them with the empirical one calculated on the actual data in [59]. Here it is possible to see that the model which better reproduce actual data seems to be the Polya urn chain.

Figures 6.3 and 6.4 show the results for the A-BAA classes. Here the CreditRisk+© model and the Polya urn chain perform very similarly, while the worst one seems to be the RRUP.

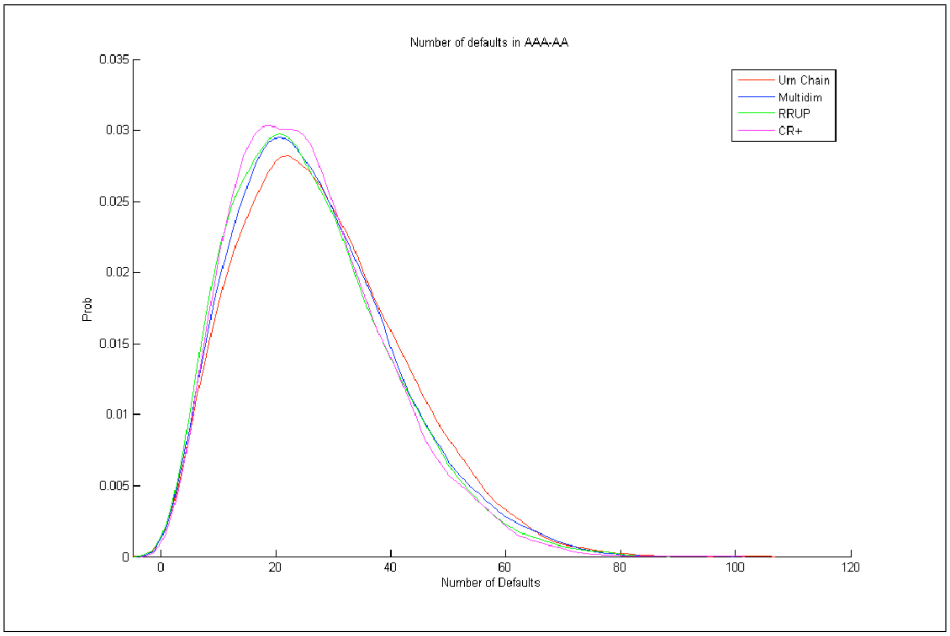

Figure 6.1: Comparison of the kernel density estimates for the number of defaults in the AAA-AA risk classes.

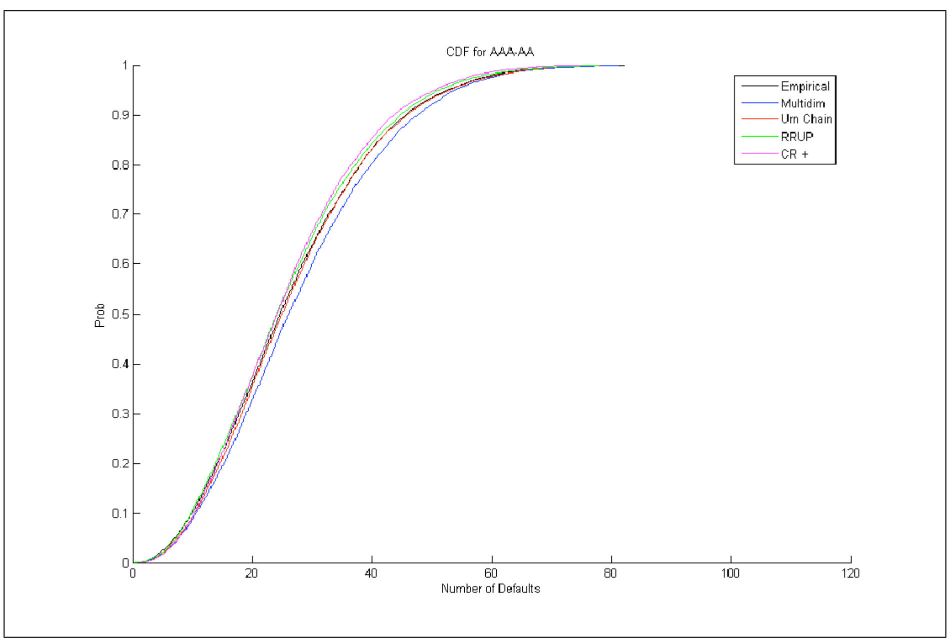

Figure 6.2: Comparison of the cumulative distribution functions for the number of defaults in the AAA-AA risk classes.

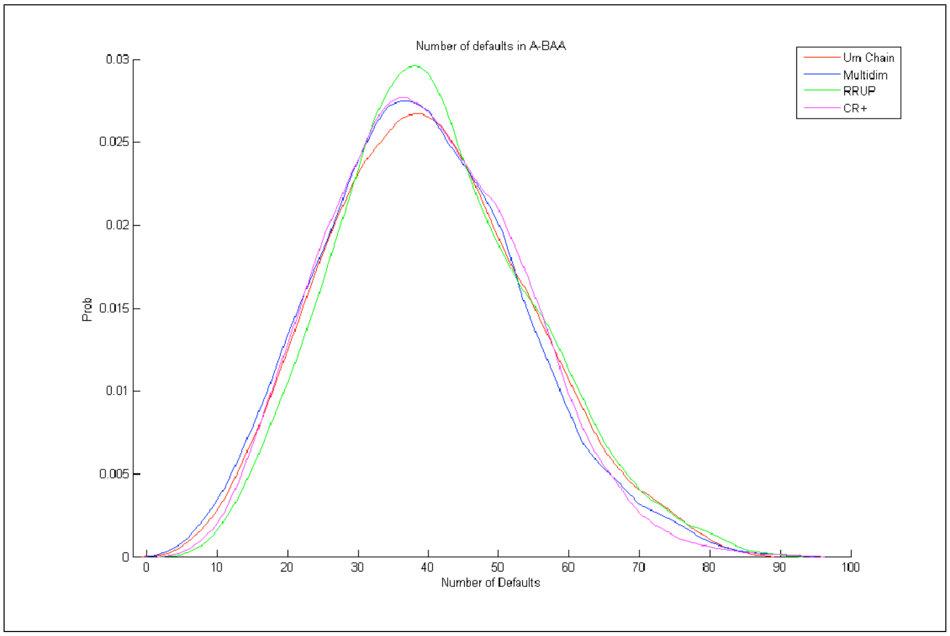

Figure 6.3: Comparison of the kernel density estimates for the number of defaults in the A-BAA risk classes.

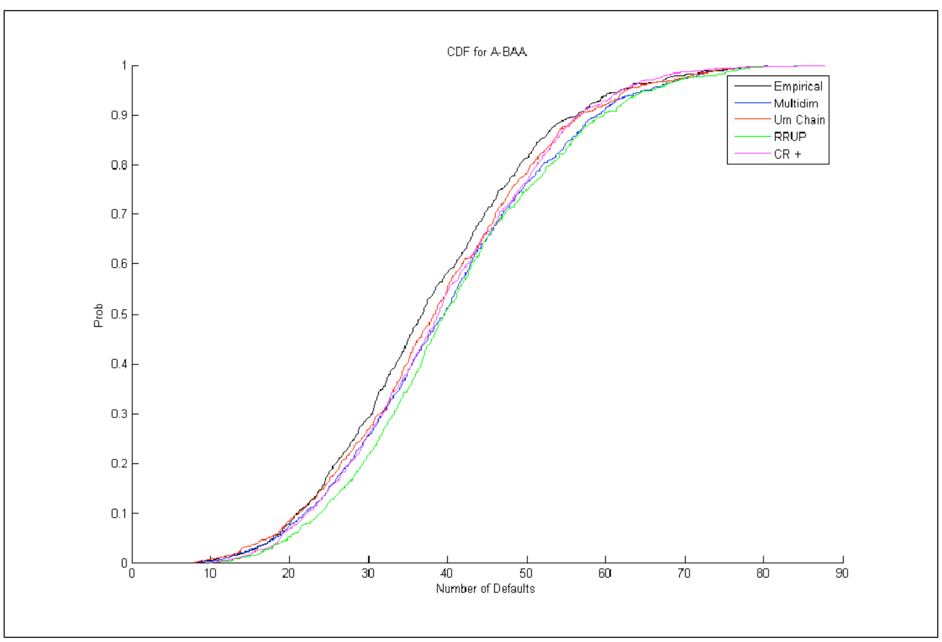

Figure 6.4: Comparison of the cumulative distribution functions for the number of defaults in the A-BAA risk classes.

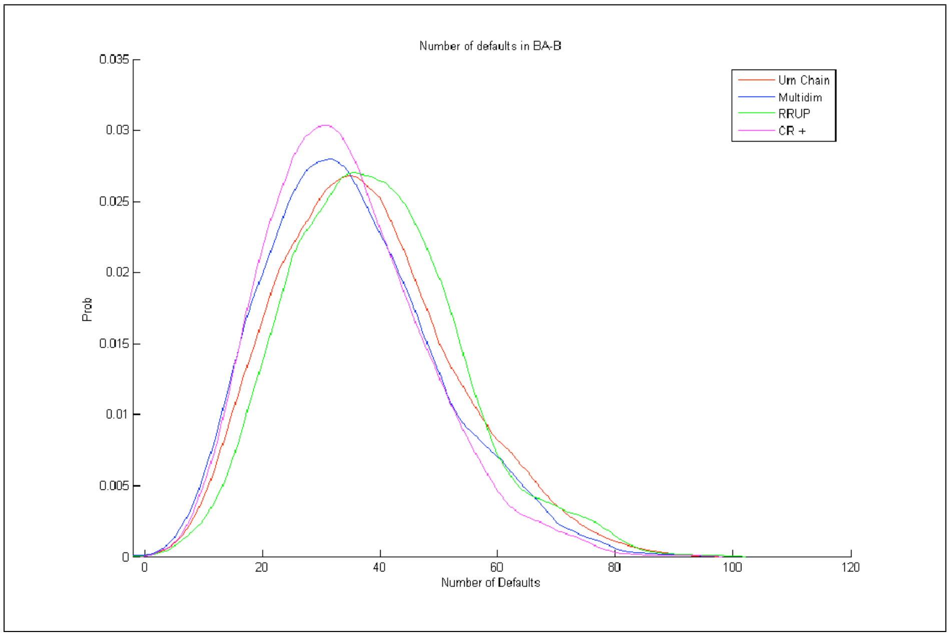

Figure 6.5: Comparison of the kernel density estimates for the number of defaults in the BA-B risk classes.

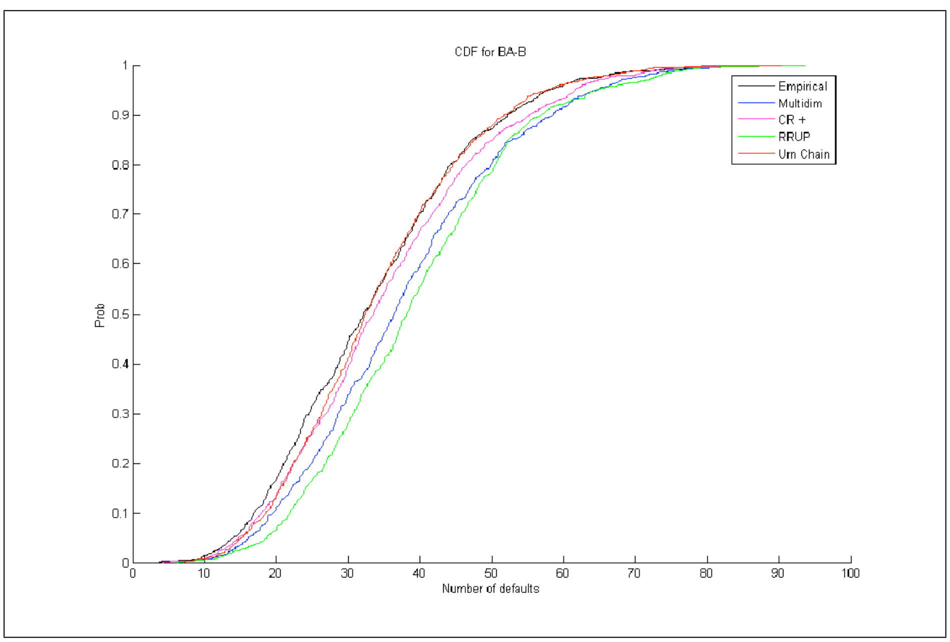

Figure 6.6: Comparison of the cumulative distribution functions for the number of defaults in the BA-B risk classes.

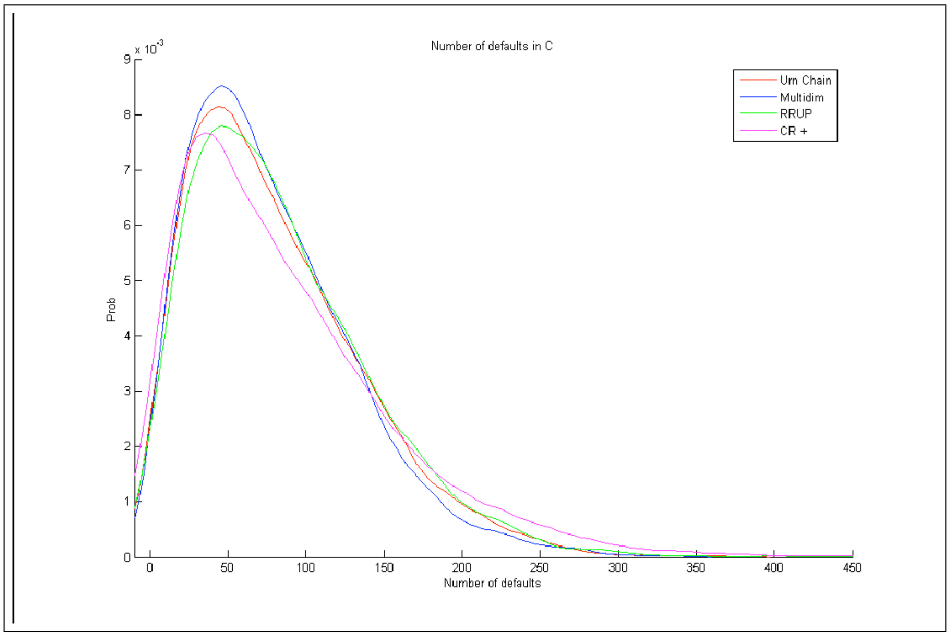

Figure 6.7: Comparison of the kernel density estimates for the number of defaults in the C risk class.

For the firms belonging to the BA-B risk group, Figures 6.5 and 6.6 display how the Polya chain gives the best results, if compared with the other models. The multidimensional models, on the contrary, are those that worse reproduce actual data.

In Figures 6.7 and 6.8, we analyze the most dangerous risk class, the one made up of C ranking bond issuers. In this case, that can be considered the most important one, view that we are interested in the number of defaults in the worst risk class, CreditRisk+© seems to give the less accurate estimates. This is probably due to the use of the basic model, however we cannot deny the very good behavior of the Polya urn chain, even in this case.

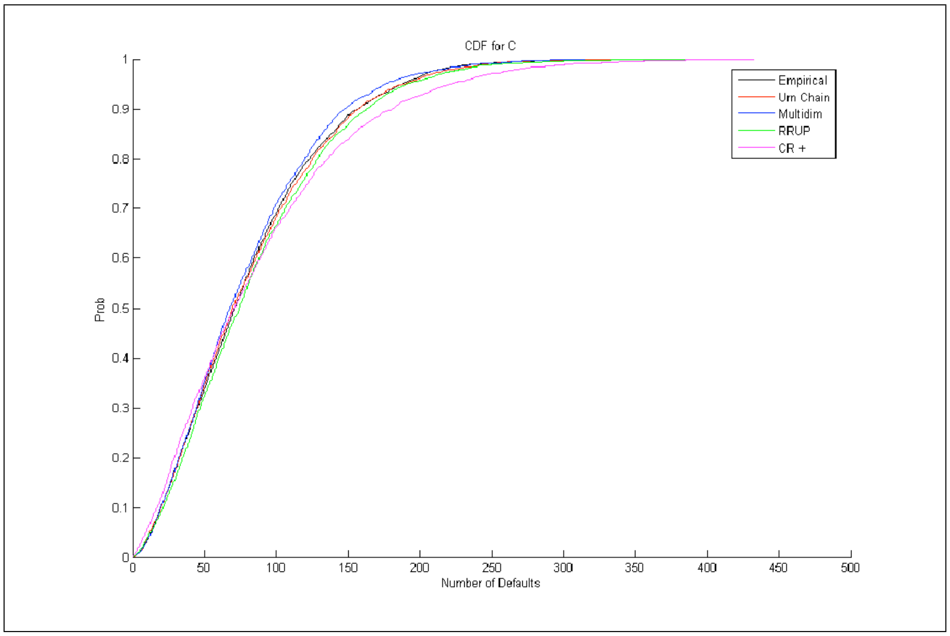

Figure 6.8: Comparison of the cumulative distribution functions for the number of defaults in the C risk class.

6.6 Conclusions

Finally we have applied our theoretical models to actual data and compared them with one of the standard model used by practitioners. The results, even if they could be further developed, give interesting information about the goodness of fit of the models.

In particular, the RRUP model, despite its demanding calculations, seems to be the less accurate in estimating the distribution of the number of defaults in the different risk classes. In practice, it simply reproduces the results of the multidimensional urn model, suggesting that, without improvements, it is better to use the simpler time-saving model.

On the contrary, the Polya urn chain seems to be the best model, with the most accurate estimates. It overreaches both the multidimensional urn models and CreditRisk+©. This is probably due to its construction, that strongly builds dependence among the risk groups, better reproducing the empirical evidence. All this suggests to better develop the chain mechanism, in order to improve the results.

Concluding Remarks

It is now time to summarize what we have done to have an useful *vue d'ensemble*.

Our aim was to propose some new urn approaches to shock and default models, to see if it was possible to reproduce and generalize existing results.

In addition, mainly if we think about the ruinous urn and the urn chain model, we have also introduced some new - and we hope interesting - results.

For our purposes we have made use of combinatorial methods, both in the classical and the analytical approach. This last one, in particular, has shown very good capabilities in solving problems that cannot be easily treated using the standard methodology. The isomorphism between urn schemes and differential systems have many fascinating properties and consequences and we aim to continue this line of research in the future.

Even if further work is needed, we have also seen that our urn-based shock and defaults models seem to have good applications, allowing for the study of real-life phenomena.

In general, the results we have obtained can be considered the very first part of a series of researches one could pursue. In particular, some ideas could be:

- To extend the class of urn-based shock models, also focusing on higher dimensional versions. One could use elliptic urns in order to model cumulative shock models;

Concluding Remarks

It is now time to summarize what we have done to have an useful *vue d'ensemble*.

Our aim was to propose some new urn approaches to shock and default models, to see if it was possible to reproduce and generalize existing results.

In addition, mainly if we think about the ruinous urn and the urn chain model, we have also introduced some new - and we hope interesting - results.

For our purposes we have made use of combinatorial methods, both in the classical and the analytical approach. This last one, in particular, has shown very good capabilities in solving problems that cannot be easily treated using the standard methodology. The isomorphism between urn schemes and differential systems have many fascinating properties and consequences and we aim to continue this line of research in the future.

Even if further work is needed, we have also seen that our urn-based shock and defaults models seem to have good applications, allowing for the study of real-life phenomena.

In general, the results we have obtained can be considered the very first part of a series of researches one could pursue. In particular, some ideas could be:

- To extend the class of urn-based shock models, also focusing on higher dimensional versions. One could use elliptic urns in order to model cumulative shock models;

- To study (generalized) extreme shock models when the reinforcement matrix is not constant over time. If the changing rule is deterministic, this only produces messier calculations, but what about the random case?
- To consider time varying ratings in default models, also modeling the case in which one failing systems migrates from one risk group to another;
- To develop new applications, also using more accurate and robust estimation techniques.

These are only a few possible roads we want to investigate in the immediate future. The rest is up to imagination.

Bibliography

[1] Aigner M., 1997. Combinatorial Theory. *Springer-Verlag*, Berlin.

[2] Aldous D.J., 1983. Exchangeability and related topics. In: Lecture Notes in Mathematics 1117, *Springer-Verlag,* Berlin.

[3] Allen L., 2004. Credit risk modelling of middle markets. *Scientific Report* 11, Zicklin School of Business.

[4] Altman, E.I., 1968. Financial ratio, discriminant analysis and prediction of corporate bankruptcy. Journal of Finance 4, 589-610.

[5] Amerio E., Muliere P., Secchi P., 2004. Reinforced urn processes for modeling credit default distributions. *International Journal of Theoretical and Applied Finance* 7, 407-423.

[6] Balakrishnan N. (Ed.), 1997. Advances in combinatorial methods and applications to probability and statistics. *Birkhauser*, Berlin.

[7] Bernoulli J., 1713. Ars Conjectandi. German version in Wahrscheinlichkeitsrechnung: Ars conjectandi. 1., 2., und 4. Theil, *Harry Deutsch*, 1998, Berlin.

[8] Berti P., Pratelli L., Rigo P., 2004. Limit theorems for a class of identically distributed random variables. *Annals of Probability,* 32, 2029-2052.

[9] Berti P., Rigo P., 2006. Finitely additive uniform limit theorems. *Sankyā,* 68, 24-44.

[10] Bianchi C., Cirillo P., Gallegati M., Vagliasindi P., 2007. Validation in ACE models: an investigation on the C@s model. *Journal of Economic Behavior and Organization*, doi:10.1016/j.jebo.2007.08.008.

[11] Bianchi C., Cirillo P., Gallegati M., Vagliasindi P., 2007. Validating and calibrating agent-based models: a case study. *Computational Economics* 30, 245-264.

[12] Billingsley P., 1995. Probability and Measure (3rd edition). *Wiley-Interscience,* New York.

[13] Beaver, W.H., 1966. Financial ratios as predictors of failure. Empirical Research in Accounting: Selected Studies. Supplement of the Journal of Accounting Research, 77-111.

[14] Blackwell D., MacQueen J.B., 1977. Ferguson distribution via Polya urn schemes. *Annals of Statistics,* 1, 353-355.

[15] Cifarelli D.M., Muliere P., Secchi P., 2000. Urn schemes for constructing priors. *Technical Report 429/P,* Dipartimento di Matematica, Politecnico di Milano.

[16] Cirillo P., 2004. Modelling firms' defaults with urns. *Unpublished Manuscript*.

[17] Cirillo P., 2007. Does Gibrat's law hold true? The Italian case. *Economics Letters*, forthcoming.

[18] Cirillo P., Hüsler J., 2007. An urn approach to generalized extreme shock models. Submitted to *Statistics and Probability Letters*.

[19] Cirillo P., Hüsler J., 2008. An urn-based generalized extreme shock models for modeling the probability of firms' defaults. Submitted to *Risk Analysis*.

[20] Cirillo P., Muliere P., 2008. An urn chain model for credit scoring. Submitted to *Statistical Modeling*.

[21] Connor R.J., Mosiman J.E., 1969. Concepts of independence for proportions with a generalization of the Dirichlet distribution. *Journal of the American Statistical Association* 64, 194-206.

[22] Credit Suisse Financial Products, 1997. CreditRisk+: a credit risk management framework. *Technical Report* available online.

[23] Davis B., 1990. Reinforced random walks. *Probability Theory and Related Fields* 84 203- 229.

[24] Davydenko S., 2005. When Do Firms Default? A Study of the Default Boundary. *EFA Moscow Meeting Paper*. Preprint at http://rotman.utoronto.ca/davydenko/Boundary.pdf

[25] De Finetti B., 1974. Theory of Probability (2 volumes). *Wiley-Interscience,* New York.

[26] Diaconis P., Freedman D., 1980. De Finetti's theorem for Markov chains. *Annals of Probability* 1, 115-130.

[27] Doksum K., 1974. Tailfree and neutral random probabilities and their posterior distributions. *Annals of Statistics* 2, 183-201.

[28] Drmota A., 1994. Asymptotic distributions and a multivariate Darboux method in enumeration problems. *Journal of Combinatorial Theory* A67, 169-184.

[29] Duffee G.R., 1999. Estimating the Price of Default Risk. *The Review of Financial Studies*, 12, 197-226.

[30] Duffie D., Singleton K.J., 2003. Credit Risk. *Princeton University Press*, Princeton NJ.

[31] Ehrenfest P., Ehrenfest T., 1907. Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem. *Physicalische Zeitschrift* 8, 311-314.

[32] Esary J.D., Marshall A.W., 1973. Shock models and wear processes. *Annals of Probability* 1, 627-649.

[33] Feller W., 1966. An introduction to probability theory and its applications. *Wiley*, New York.

[34] Flajolet P., Odlyzko A., 1990. A singularity analysis of generating functions. *SIAM Journal of Discrete Mathematics* 3, 216-240.

[35] Flajolet P., Gabarro J., Pekari H., 2005. Analytic Urns. *Annals of Probability,* 33, 1200-1233.

[36] Flajolet P., Dumas P., Puyhaubert V., 2006. Some exactly solvable models of urn process theory. *Fourth Colloquium on Mathematics and Computer Science,* 59-118.

[37] Flajolet P., Sedgewick R., 2008. Analytic Combinatorics. Forthcoming book for *Cambridge University Press*. Preliminary version downloadable at algo.inria.fr/flajolet/

[38] Fujiwara Y., 2004. Zipf Law in Firms Bankruptcy. *Physica A,* 337, 219-230.

[39] Delli Gatti D., Gaffeo E., Gallegati M., Giulioni G., Kirman A., Palestrini A., Russo A., 2007. Complex dynamics and empirical evidence. *Information Sciences* 177, 1204-1221.

[40] Gentle E., Härdle W., Mori Y., Gentle J.E., 2004. Handbook of computational statistics. *Springer-Verlag*, Berlin.

[41] Ghosh J.K., Ramamoorthi R.V., 2003. Bayesian Nonparametrics. *Springer-Verlag*, New York.

[42] Gouriéroux C., Monfort A., 1996. Simulation-Based Econometric Methods. *Oxford University Press*, Oxford.

[43] Gut A., 1990. Cumulative Shock Models. *Advances in Applied Probability,* 22, 504-507.

[44] Gut A., 2001. Mixed Shock Models. *Bernoulli,* 7, 541-555.

[45] Gut A., Hüsler J., 1999. Extreme Shock Models. *Extremes,* 2, 293-305.

[46] Gut A., Hüsler J., 2005. Realistic variation of shock models. *Statistics & Probability Letters,* 74, 187-204.

[47] Hwang Hsien-Huei. Théorèems limites pour les structures combinatoires et les functions arithmetiques. *PhD Thesis*, Ens Paris.

[48] Ishwaran H., Lancelot F.J., 2001. Gibbs sampling methods for stick-breaking priors. *Journal of the American Statistical Association* 96, 161-173.

[49] Kiyoshi I., Sigeo A., 2001. Polya urn models under general replacement schemes. *Journal of the Japanese Statistical Society,* 31, 193-205.

[50] Janson S., 2005. Limit theorems for triangular urn schemes. *Probability Theory and Related Fields*, 134, 417-452.

[51] Jeffrey A., 2003. Handbook of mathematical formulas and integrals, 3rd edition. *Academic Press*, Oxford.

[52] Johnson N.L., Kotz S., 1977. Urn models and their applications. *Wiley-Interscience*, New York.

[53] Lladser E.M., 2003. Asymptotic enumeration via singularity analysis. *PhD Thesis*, Ohio State University.

[54] Lundberg E., 1879. Om Hypergoniometriskafunctioner af komplexa variabla. *Manuscript*. The English translation "On hypergoniometric functions of complex variables" is available online at www.maths.lth.se/matematiklu/personal/jaak/hypergf.ps

[55] Mahmoud H., 1998. On rotations in fringe-balanced binary trees. *Information Processing Letters* 65, 41-46.

[56] Mallor F., Omey E., 2001. Shocks, runs and random sums. *Journal of Applied Probability*, 38, 438-448.

[57] Marsili M., Valleriani A., 1998. Self-organization of interacting Polya urns. *Preprint* at citeseer.ist.psu.edu/194060.html.

[58] Mihram G.A., Hultquist R.A., 1967. A bivariate warning-time/failure distribution. *Journal of the American Statistical Association* 62, 589-599.

[59] Moody's Corporation, 2007. Default and recovery rates of European corporate bonds issuers 1989-2006. *Technical Report* J0607.

[60] Muliere P., Secchi P., Walker S., 2000. Urn schemes and reinforced random walks. *Stochastic Processes and Their Applications* 88, 59-78.

[61] Muliere P., Walker S., 2000. Neutral to the right processes from a predictive perspective: a review and new developments. *Metron* 3-4, 13-30.

[62] Ohlson, J., 1980. Financial ratio and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* 18, 109-131.

[63] Paganoni A.M., Secchi P., 2004. Interacting reinforced-urn systems. *Advances in Applied Probability*, 36, 791-804.

[64] Pemantle R., 1990. A time-dependent version of Polya's urn. *Journal of Theoretical Probability*, 3, 627-637.

[65] Phlips L., 1998. Applied Industrial Economics. *Cambridge University Press*, Cambridge.

[66] Pillai R.N., 1990. On Mittag-Leffler functions and related distributions. *Annals of the Institute of Statistical Mathematics*, 42, 157-161.

[67] Pitman J., 1996. Some developments of the Blackwell-MacQueen urn scheme. In Ferguson et al., eds, Statistics, Probability and Game Theory, Papers in honor of David Blackwell, *Lecture Notes-Monograph Series* 30, Institute of Mathematical Statistics, 245-267.

[68] Pittel B., 1987. An urn model for cannibal behavior. *Journal of Applied Probability 24*, 522-526.

[69] Pouyanne N., 2005. Classification of large Polya-Eggenberger urns with regard to their asymptotics. *Discrete Mathematics and Theoretical Computer Science*, AD, 275-286.

[70] Puyhaubert V., 2005. Modèles d'urnes et phénomènes de seuils en combinatoire analytique. *PhD Thesis*, Ecole Polytechinque, INRIA. Downloadable at algo.inria.fr/puyhaubert

[71] Ray B., Singh R., Shashanka M., Smaragdis P., 2007. Bandwitch expansion with a Polya urn model. *Mitsubishi Research Labs* TR58.

[72] Sandström A., 2007. Political Risk and Firm Default Probability. *Working Paper,* Downloadable at project.hkkk.fi/gsf/seminar_papers/Political%20... ...risk%20and%20firm%20default%20probability.pdf.

[73] Sohn K-A., Xing E.P., 2006. Hidden Markov Dirichlet processes. Modelling genetic recombination in open ancestral space. *Advances in Neural Information Processing Systems* 19, 20-28.

[74] Stanley R.P., 1997. Enumerative Combinatorics. Cambridge Studies in Advanced Mathematics. *Cambridge University Press*, Cambridge.

[75] Taufer E., 2006. Modeling stylized features in default rates. *Alea Tech Report* 21.

[76] Walker S., Muliere P., 1997. Beta-stacy processes and a generalization of the Polya urn scheme. *Annals of Statistics* 25, 1762-1780.

[77] Young U.R., Wei T.Y., 2005. Firms Bankruptcy Prediction: Experimental Comparison of Isotonic Separation and Other Classification Approaches. *IEEE Transactions on Systems, Man and Cybernetics: Part A*, 35, 727-737.