AofA'12 DMTCS proc. **AQ**, 2012, 29–42

On death processes and urn models

Markus Kuba[1,2†] and Alois Panholzer[1‡]

[1] *Institut für Diskrete Mathematik und ,Geometrie, Technische Universität Wien, Austria*
[2] *Institut für Angewandte Mathematik und Naturwissenschaften, FH Technikum Wien, Austria*

We use death processes and embeddings into continuous time in order to analyze several urn models with a diminishing content. In particular we discuss generalizations of the pill's problem, originally introduced by Knuth and McCarthy, and generalizations of the well known sampling without replacement urn models, and OK Corral urn models.

Keywords: Urn models, Generating functions, Limiting distribution MSC2000 05A15,60F05,05C05

1 Introduction

1.1 Diminishing Pólya-Eggenberger urn models

In this work we are concerned with so-called Pólya-Eggenberger urn models, which in the simplest case of two colors can be described as follows. At the beginning, the urn contains n white and m black balls. At every step, we choose a ball at random from the urn, examine its color and put it back into the urn and then add/remove balls according to its color by the following rules: if the ball is white, then we put α white and β black balls into the urn, while if the ball is black, then γ white balls and δ black balls are put into the urn. The values $\alpha, \beta, \gamma, \delta \in \mathbb{Z}$ are fixed integer values and the urn model is specified by the transition matrix $M = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$. Models with r (≥ 2) types of colors can be described in an analogous way and are specified by an $r \times r$ transition matrix. Urn models are simple, useful mathematical tools for describing many evolutionary processes in diverse fields of application such as analysis of algorithms and data structures, and statistical genetics. Due to their importance in applications, there is a huge literature on the stochastic behavior of urn models; see for example [10, 11, 18]. Recently, a few different approaches have been proposed, which yield deep and far-reaching results [1, 4, 5, 8, 9, 19]. Most papers in the literature impose the so-called tenability condition on the transition matrix, so that the process of adding/removing balls can be continued ad infinitum. However, in some applications, examples given below, there are urn models with a very different nature, which we will refer to as "*diminishing urn models*." Such models have recently received some attention, see for example [21, 22, 2, 20, 4, 6]. For simplicity of presentation, we describe diminishing urn models in the case of balls with two types of colors, black and white. We

†E-mail: `kuba@dmg.tuwien.ac.at`
‡E-mail: `Alois.Panholzer@tuwien.ac.at` The second author was supported by the Austrian Science Foundation FWF, grant S9608-N13.

consider Pólya-Eggenberger urn models specified by a transition matrix $M = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$, and in addition we also specify a set of absorbing states $\mathcal{A} \subseteq \mathbb{N} \times \mathbb{N}$. The evolution of the urn takes place in the state space $\mathcal{S} \subseteq \mathbb{N} \times \mathbb{N}$. The urn contains m black balls and n white balls at the beginning, with $(m, n) \in \mathcal{S}$, and evolves by successive draws according to the transition matrix until an absorbing state in $\mathcal{A}$ is reached, and the process stops. Diminishing urn models with more than two type of balls can be considered similarly.

1.2 Plan of this note and notation

There are numerous examples of diminishing urns and related problems in literature. In the following we present three concrete problems, the pills problem urn model, the sampling without replacement urn model, and the OK Corral urn model, and summarize known results. For all three problems presented below, and suitable generalization of them, we will use stochastic processes and an embedding of the discrete time process of drawing and adding/removing balls in continuous time, in order to unify and extend the known results in the literature concerning exact distribution laws, generalizing some results of [2, 7, 15, 21, 12, 13, 17, 20, 4], trying to provide a lucid derivation. We will denote with $X \oplus Y$ the sum of independent random variable X and Y. Moreover, we use the notations $\mathbb{N} = \{1, 2, 3, \dots\}$ and $\mathbb{N}_0 = \{0, 1, 2, \dots\}$.

1.3 The pills problem and generalizations

Consider the diminishing urn problem with transition matrix given by $M = \left(\begin{smallmatrix} -1 & 0 \\ 1 & -1 \end{smallmatrix}\right)$, state space $\mathcal{S} = \mathbb{N}_0 \times \mathbb{N}$, and the absorbing axis $\mathcal{A} = \{(0, n) \mid n \in \mathbb{N}\}$. Following Knuth and McCarthy [14], a vivid interpretation is as follows: An urn has two types of pills in it, which are single-unit and double-unit pills, respectively. At every step, we pick a pill uniformly at random. If a single-unit pill is chosen, then we eat it up, and if the pill is of double unit, we break it into two halves—one half is eaten up and the other half is now considered of single unit and thrown back into the urn. The question is then, when starting with n single-unit pills and m double-unit pills, what is the probability that k single-unit pills remain in the urn when all double-unit pills are drawn? This problem was first posed in [14], where the authors asked for a formula for the expected number of remaining single-unit pills, when there are no double-unit pills in the urn. The solution appeared in [3]. A more refined study was given by Brennan and Prodinger [2], where they derive exact formulæ for the variance and the third moment of the number of remaining single-unit pills; furthermore, a few generalizations are proposed. The probability generating functions and limit laws for the pills problem and a variant of the problem have been derived in [7] using a generating functions approach. Furthermore, a study of the arising limiting distributions of a general class of related problems has been carried out in [15] using a recursive approach basically guessing the structure of the moments, together with an application of the so-called method of moments. However, some cases proved to be quite elusive using the techniques of [15]. Moreover, no simple explicit general formula for the probability mass function of the random variable of interest was obtain before. We will provide the solution for the general pills problem urn model of [15] with ball transition matrix given by $M = \left(\begin{smallmatrix} -\alpha & 0 \\ \gamma & -\delta \end{smallmatrix}\right)$, $\gamma = \alpha \cdot p$, $p \in \mathbb{N}_0$. Furthermore, we will discuss weighted generalizations of the pills problem, and also discuss extensions to higher dimensional pills problem urn models, with $r \times r$-transition matrix similar to

$$M = \begin{pmatrix} -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -1 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 1 & -1 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & -1 & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 1 & -1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & -1 \end{pmatrix}; \qquad (1)$$

Note that the matrix stated above is the natural generalization of the pills problem ball transition matrix to dimension $r \geq 2$.

1.4 Sampling without replacement and generalizations

This classical example, often serving as a toy model, corresponds to the urn with transition matrix $M = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, state space $\mathcal{S} = \mathbb{N}_0 \times \mathbb{N}$, and absorbing axis $\mathcal{A} = \{(0, n) \mid n \in \mathbb{N}_0\}$. In this model, balls are drawn one after another from an urn containing balls of two different colors and not replaced. What is the probability that k white balls remain when all black balls have been removed, starting with n white and m black balls? This simple urn model has been discussed in [7] using generating functions. Moreover, generalizations of the sampling urn model have been discussed in [15, 16]. We will show that this urn model can be considered as a degenerate case of the pills problem, and that the solution for the generalized pills problem urn model also covers the sampling without replacement urn model.

1.5 The OK Corral urn model

The so-called OK Corral urn serves as a mathematical model of the historical gun fight at the OK Corral. This problem was introduced by Williams and McIlroy in [22] and studied recently by several authors using different approaches, leading to very deep and interesting results; see [21, 12, 13, 17, 20, 4]. Also the urn corresponding to the OK corral problem can be viewed as a basic model in the mathematical theory of warfare and conflicts; see [13, 17].

In the diminishing urn setting the OK corral problem corresponds to the urn with transition matrix $M = \left(\begin{smallmatrix} 0 & -1 \\ -1 & 0 \end{smallmatrix}\right)$, state space $\mathcal{S} = \mathbb{N} \times \mathbb{N}$, and two absorbing axes: $\mathcal{A} = \{(0, n) \mid n \in \mathbb{N}\} \cup \{(m, 0) \mid m \in \mathbb{N}\}$. An interpretation is as follows. Two groups of gunmen, group A and group B (with n and m gunmen, respectively), face each other. At every discrete time step, one gunman is chosen uniformly at random who then shoots and kills exactly one gunman of the other group. The gunfight ends when one group gets completely "eliminated". Several questions are of interest: what is the probability that group A (group B) survives, and what is the probability that the gunfight ends with k survivors of group A (group B)? This model was analyzed by Williams and McIlroy [22], who obtained a result for the expected value of the number of survivors. Using martingale arguments and the method of moments Kingman [12] gave limiting distribution results for the OK Corral urn model for the total number of survivors. Moreover, Kingman [13] obtained further results in a very general setting of Lanchester's theory of warfare. Kingman and Volkov [17] gave a more detailed analysis of the so-called balanced OK Corral urn model using a connection to the famous Friedman urn model; amongst others, they derived an explicit result for the number of survivors and even local limit laws. Puyhaubert [20] extended in his Ph. D. thesis the results of [12, 17] on the balanced OK Corral urn model using analytic combinatoric methods concerning the number of survivors of a certain group. His study is based on the connection to the Friedman urn showed in [17]. He obtained explicit expression for the probability distribution, the moments, and also reobtained (and refined) most of the limiting distribution results reported earlier. Some results of [20] where reported in the work of Flajolet et al. [4]. Apparently unknown to the previously stated authors was the earlier work of Stadje [21], who obtained several limiting distribution results for the generalized OK Corral urn, as introduced below, and also for related urn models with more general transition probabilities. In [21] the probability distributions for the most general transition probabilities are determined by a complex integral, but without any proof. The results of Stadje were then discussed in [16], and their connection to sampling without replacement urn models with general weight sequences uncovered, and a duality relation proved. Despite all the mentioned works and results, no transparent probabilistic derivation of the general results of [21, 16] were given before.

2 Probabilistic analysis of the pill's problem urn models

We are interested in a generalized pill's problem with ball transition matrix given by

$$M = \left(\begin{smallmatrix} -\alpha & 0 \\ \gamma & -\delta \end{smallmatrix}\right), \tag{2}$$

where $\alpha, \beta \in \mathbb{N}$, and $\gamma = \alpha \cdot p$, $p \in \mathbb{N}_0$. Let $X_{n,m}$ denote the random variable counting the number of remaining white balls (divided by α) when all black balls have been drawn. The probability generating function $h_{n,m}(v) = \mathbb{E}(v^{X_{n,m}})$ satisfies the recurrence relation

$$h_{n,m}(v) = \frac{\alpha n}{\alpha n + \delta m} h_{n-1,m}(v) + \frac{\delta m}{\alpha n + \delta m} h_{n+p,m-1}(v), \tag{3}$$

with $h_{n,0} = v^n$, $n \geq 1$. We analyze $X_{n,m}$ using a continuous time embedding. We start at time zero with n white balls and m black balls, and use two independent linear processes. The first one consists of n independent ordinary death processes (white balls) with death rate α. Let $W_n(t)$ denote the random variable counting the number of living white balls at time t, with $W_n(0) = n$. The second one (black balls) consists of m independent modified death processes, with rate δ, where each black ball gives at his death birth to p new white balls with death rate α, independent of all other balls, and $p \in \mathbb{N}_0$. We denote with $B_m(t)$ the random variable counting the number of living black balls at time t, with $B_m(0) = m$. Finally, let $C_m(t)$ denote the random variable counting the number of surviving white balls up to time t, which are children of black balls. Let $\tau = \inf_{t>0}\{B_m(t) = 0\}$ be the time when the black balls die out. Then

$$X_{n,m} = W_n(\tau) \oplus C_m(\tau),$$

both random variables are independent, due to the construction. One readily obtains the recurrence relation for the probability generating function $h_{n,m}(v)$ by looking at the time when the first particle dies. It is well known that the index of the variable achieving the minimum out of r independent exponential distributed random variables $X_1, \dots, X_r$ with parameters $\lambda_1, \dots, \lambda_r$, is given by

$$\mathbb{P}\{X_k = \min\{X_1, \dots, X_r\}\} = \int_0^\infty \Big(\frac{d}{dt}(1 - e^{-\lambda_k t}\Big) \prod_{\substack{j=1 \\ j \neq k}} e^{-\lambda_j t} dt = \frac{\lambda_k}{\lambda_1 + \dots + \lambda_k}.$$

Hence, we note that the probability that any of the type one balls dies first is given by $\frac{\alpha n}{\alpha n + \delta m}$, and the opposite case happens with probability $\frac{\delta m}{\alpha n + \delta m}$. Moreover, if p new white balls with death rate α are being born, they can be grouped with the already existing white balls, due to the memorylessness of exponential distributions. This leads directly to (3).

Due to the construction of the two processes the random variables $W_n(t)$ and $C_m(t)$ can be decomposed themselves into sums of i. i. d. random variables,

$$W_n(t) = \bigoplus_{k=1}^{n} X_k(t), \qquad C_m(t) = \bigoplus_{k=1}^{m} Y_k(t)$$

where $X_k(t)$ denotes the indicator variable of the k-th white ball living at time t, $1 \leq k \leq n$, and $Y_k(t)$ denote the random variable counting the number of surviving white balls up to time t, which are children of the k-th black ball, $1 \leq k \leq m$.

The probability generating function of a single white ball at time t with rate δ is given by

$$\mathbb{E}(v^{X_k(t)}) = \mathbb{P}\{X_k(t) = 0\} + v\mathbb{P}\{X_k(t) = 1\} = 1 - e^{-\alpha t} + ve^{-\alpha t}.$$

so the probability generating function of the total number of white balls living at time t is given by

$$\mathbb{E}(v^{W_n(t)}) = \prod_{k=1}^{n} \mathbb{E}(v^{X_k(t)}) = \big(1 + (v-1)e^{-\alpha t}\big)^n,$$

due to the independence assumption. The probability generating function of $Y_k(t)$, assuming that the black ball dies before time t, is given by

$$\int_0^t \Big(\frac{d}{du}(1 - e^{-\delta u})\Big) \cdot (1 + (v-1)e^{-\alpha(t-u)})^p du = \int_0^t \delta e^{-\delta u} \cdot (1 + (v-1)e^{-\alpha(t-u)})^p du,$$

due to the fact that p independent white balls are being born at the death of the black ball. Consequently, the probability generating function of the children of the $m-1$ black balls, dying before time t, and the corresponding number of surviving child balls up to time t is given by

$$\Big(\int_0^t \delta e^{-\delta u} \cdot (1 + (v-1)e^{-\alpha(t-u)})^p du\Big)^{m-1}.$$

Furthermore, the density of the last remaining black ball is given by $\delta \cdot e^{-\delta t}$, giving birth to p more surviving white balls. Moreover, this final ball can be any one out the m balls. Alltogether, considering all possible final death times t, or more precisely by conditioning on the stopping time τ we obtain for the probability generating function $h_{n,m}(v) = \mathbb{E}(v^{X_{n,m}})$ the following result.

Theorem 1 *For arbitrary $\alpha, \delta \in \mathbb{N}$ and $p \in \mathbb{N}_0$, the probability generating function of the random variable $X_{n,m}$ counting the number of remaining white balls (divided by α) when all black balls have been drawn, $M = \left(\begin{smallmatrix} -\alpha & 0 \\ \gamma & -\delta \end{smallmatrix}\right)$, $\gamma = \alpha \cdot p$, is given by*

$$h_{n,m}(v) = \int_0^\infty \big(1 + (v-1)e^{-\alpha t}\big)^n \cdot \Big(\int_0^t \delta e^{-\delta u} \cdot (1 + (v-1)e^{-\alpha(t-u)})^p du\Big)^{m-1} \cdot v^p m \delta e^{-\delta t} dt.$$

The results above unify and extend the known results of [7]. Moreover, it allows to largely extend the results of [15] concerning the structure of the moments, as stated below, and also to give a complete analysis of the limit laws. Note that by setting $p = 0$ one also gets the probability generating function for a certain generalized sampling without replacement urn model. From the result above we will derive a closed formula for the s-th factorial moment $\mathbb{E}(\tilde{X}_{n,m}^{\underline{s}})$ of $\tilde{X}_{n,m} = X_{n,m} - p$, for $\alpha \neq \delta$ such that $\alpha\ell - \delta \neq 0$; the special case $\alpha = \delta$ has already been treated in [7]. Note that the factorial moments $\mathbb{E}(X_{n,m}^{\underline{s}})$ of $X_{n,m}$ are recovered using the binomial theorem for the falling factorials

$$\mathbb{E}(X_{n,m}^{\underline{s}}) = \mathbb{E}\big((\tilde{X}_{n,m} + p)^{\underline{s}}\big) = \sum_{\ell=0}^{s} \binom{s}{\ell} \mathbb{E}(\tilde{X}_{n,m}^{\underline{\ell}}) p^{\underline{s-\ell}}.$$

Theorem 2 *The factorial moments of the random variable $\tilde{X}_{n,m} = X_{n,m} - p$ are given in terms of a generalized beta integral,*

$$\mathbb{E}(\tilde{X}_{n,m}^{\underline{s}}) = \delta^{m-1} s! \sum_{j=0}^{s} \binom{n}{j} \sum_{\substack{\sum_{\ell=0}^{p} k_\ell = m-1,\, \sum_{\ell=0}^{p} \ell k_\ell = s-j \\ k_\ell \geq 0}} \binom{m-1}{k_0, \dots, k_\ell} \frac{\prod_{\ell=0}^{p} \binom{p}{\ell}^{k_\ell}}{\prod_{\ell=0}^{p} (\ell\alpha - \delta)^{k_\ell}}$$
$$\times \int_0^1 q^{\frac{\alpha j}{\delta} + m - 1} \prod_{\ell=0}^{p} (1 - q^{\frac{\alpha\ell}{\delta} - 1})^{k_\ell} dq.$$

In particular, we obtain for $p = 1$ the simple expression

$$\mathbb{E}(\tilde{X}_{n,m}^{\underline{s}}) = s! \sum_{\ell=0}^{s} \frac{\binom{n}{\ell}\binom{m}{s-\ell}}{\left(\frac{\alpha}{\delta} - 1\right)^{s-\ell}} \sum_{i=0}^{s-\ell} (-1)^{s-\ell-i} \frac{\binom{s-\ell}{i}}{\binom{m-s+\ell+i+\frac{\alpha}{\delta}(s-i)}{m-s+\ell}}.$$

Our starting point is the following expression for $\mathbb{E}(\tilde{X}_{n,m}^{\underline{s}})$:

$$\mathbb{E}(\tilde{X}_{n,m}^{\underline{s}}) = E_v D_v^s \frac{h_{n,m}(v)}{v^p},$$

where E_v denotes the operator which evaluates at $v = 1$, and D_v the differentiation operator. By the binomial theorem we have

$$\int_0^t \delta e^{-\delta u} \cdot (1 + (v-1)e^{-\alpha(t-u)})^p du = \delta \sum_{\ell=0}^{p} \binom{p}{\ell} (v-1)^\ell \frac{e^{-t\delta} - e^{-\alpha\ell t}}{\ell\alpha - \delta}.$$

Consequently, using the multinomial theorem, we obtain

$$\Big(\int_0^t \delta e^{-\delta u} \cdot (1 + (v-1)e^{-\alpha(t-u)})^p du \Big)^{m-1}$$
$$= \delta^{m-1} \sum_{\substack{k_0 + \dots + k_p = m-1 \\ k_\ell \geq 0}} \binom{m-1}{k_0, \dots, k_p} \prod_{\ell=0}^{p} \binom{p}{\ell}^{k_\ell} (v-1)^{\ell k_\ell} \frac{(e^{-t\delta} - e^{-\alpha\ell t})^{k_\ell}}{(\ell\alpha - \delta)^{k_\ell}}.$$

Using

$$\big(1 + (v-1)e^{-\alpha t}\big)^n e^{-\delta t} = \sum_{j=0}^{n} \binom{n}{j} (v-1)^j e^{-(\alpha j + \delta)t},$$

we get

$$\mathbb{E}(\tilde{X}_{n,m}^{\underline{s}}) = \delta^m \sum_{j=0}^{n} \binom{n}{j} \sum_{\substack{k_0 + \dots + k_p = m-1 \\ k_\ell \geq 0}} \binom{m-1}{k_0, \dots, k_p} \frac{\prod_{\ell=0}^{p} \binom{p}{\ell}^{k_\ell}}{\prod_{\ell=0}^{p} (\ell\alpha - \delta)^{k_\ell}} E_v D_v^s (v-1)^{j + \sum_{\ell=0}^{p} \ell k_\ell}$$
$$\times \int_{t=0}^{\infty} e^{-(\alpha j + \delta)t} \prod_{\ell=0}^{p} (e^{-t\delta} - e^{-\alpha\ell t})^{k_\ell} dt.$$

Since

$$E_v D_v^s (v-1)^{j+\sum_{\ell=0}^{p}\ell k_\ell} = \begin{cases} s!, & j+\sum_{\ell=0}^{p}\ell k_\ell = s,\\ 0, & j+\sum_{\ell=0}^{p}\ell k_\ell \neq s,\end{cases}$$

we get the simpler expression

$$\mathbb{E}(\tilde{X}_{n,m}^{\underline{s}}) = \delta^m s! \sum_{j=0}^{s}\binom{n}{j} \sum_{\substack{\sum_{\ell=0}^{p} k_\ell = m-1,\, \sum_{\ell=0}^{p}\ell k_\ell = s-j\\ k_\ell\ge 0}} \binom{m-1}{k_0,\dots,k_p} \frac{\prod_{\ell=0}^{p}\binom{p}{\ell}^{k_\ell}}{\prod_{\ell=0}^{p}(\ell\alpha-\delta)^{k_\ell}}$$
$$\times \int_{t=0}^{\infty} e^{-(\alpha j+\delta)t}\prod_{\ell=0}^{p}(e^{-t\delta}-e^{-\alpha\ell t})^{k_\ell}\,dt.$$

Now we use the substitution $q = e^{-\delta t}$ in order to convert the integral above into a beta-function type integral, which proves our result.

2.1 Higher dimensional urn models

One can readily extend the 2×2 transition matrix (2) to higher dimensions,

$$M = \begin{pmatrix} -\alpha_1 & 0 & 0 & \cdots & 0 & 0 & 0\\ p_2\alpha_1 & -\alpha_2 & 0 & \ddots & \ddots & \ddots & 0\\ 0 & p_3\alpha_2 & -\alpha_3 & \ddots & \ddots & \ddots & 0\\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots\\ 0 & \ddots & \ddots & \ddots & -\alpha_{r-2} & 0 & 0\\ 0 & \ddots & \ddots & \ddots & p_{r-1}\alpha_{r-2} & -\alpha_{r-1} & 0\\ 0 & 0 & 0 & \cdots & 0 & p_r\alpha_{r-1} & -\alpha_r \end{pmatrix},$$

with $\alpha_i \in \mathbb{N}$ and $p_i \in \mathbb{N}_0$. We consider the distribution of the random vector $\mathbf{X_n} = (X_{\mathbf{n}}^{[1]},\dots,X_{\mathbf{n}}^{[r-1]})$, which counts the number of type 1 up to type $r-1$ pills when all pills of r units are all taken, starting with n_i pills of i units, $i=1,\dots,r$. One may use similar arguments to the 2×2 case to obtain the following result.

Theorem 3 *The probability generating function of $\mathbf{X_n}$ is given by*

$$h_{\mathbf{n}}(\mathbf{v}) = \int_0^\infty \Big(g_r(t,\mathbf{v})\Big)^{n_r-1} v_{r-1}^{p_r}\alpha_r n_r e^{-\alpha_r t}\prod_{\ell=1}^{r-1}\big(f_\ell(t,\mathbf{v})\big)^{n_\ell}dt.$$

Here $f_j(t,\mathbf{v})$ denotes a sequence of functions defined by $f_0(t,\mathbf{v})=1$, and

$$f_j(t,\mathbf{v}) = v_j e^{-\alpha_j t} + \int_0^t \alpha_j e^{-\alpha_j u_j}\big(f_{j-1}(t-u_j,\mathbf{v})\big)^{p_j}du_j, \quad j\ge 1,$$

with $g_j(t,\mathbf{v}) = f_j(t,\mathbf{v}) - v_j e^{-\alpha_j t}$.

Note that the result above can be further generalized to $r\times r$ ball transition matrices $M=(m_{i,j})$ with entries $-\alpha_i$, $1\le i\le r$, and entries $m_{i,j}=-\alpha_i$, for $1\le i=j\le r$, $m_{i,j}=p_{i,j}\cdot\alpha_j$ for $1\le i<j\le r$ and $p_{i,j}\in\mathbb{N}_0$, and $m_{i,j}=0$ for $1\le j<i\le r$.

2.2 General weight sequences

One may also obtain the result for $X_{n,m}$ using a slightly different probabilistic model. Our first process still consists of n independent ordinary death processes (white balls) with death rate α. However, concerning the second process, we consider a single modified death process $B(t)$ with death rates $\theta_m, \dots, \theta_1$, starting with $B(0) = m$. Note that for $\theta_k = \delta \cdot k$, we reobtain our earlier results. At each transition of $B(t)$ exactly p white balls are being born, modeled by p independent ordinary death processes (white balls) with death rate α. Consequently, one obtains the alternative description

$$h_{n,m}(v) = \int_0^\infty \left(1 + (v-1)e^{-\alpha t}\right)^n \cdot p_m(t,v)dt.$$

where $p_m(t,v)$ denotes density of $B(t)$ dying out before time t, with variable v marking the living white balls at time t,

$$\begin{aligned} p_m(t,v) = \frac{d}{dt}\bigg(\int_0^t \theta_m e^{-\theta_m u_m} &(1+(v-1)e^{-\alpha(t-u_m)})^p \\ &\int_0^{t-u_m} \theta_{m-1} e^{-\theta_{m-1}u_{m-1}} (1+(v-1)e^{-\alpha(t-u_m-u_{m-1})})^p \dots \\ &\dots \int_0^{t-\sum_{\ell=2}^m u_\ell} \theta_1 e^{-\theta_1 u_1} (1+(v-1)e^{-\alpha(t-\sum_{\ell=1}^m u_\ell)})^p du_1 \dots du_m \bigg). \end{aligned} \tag{4}$$

For $\theta_k = \delta \cdot k$, the nested integrals simplify and we reobtain our earlier result. Note that the general case corresponds to a biased pills problem urn model, first described in [16]: assuming that the urn contains n white and m black balls, the probability of choosing a white ball is given by $n/(n+\theta_m)$, whereas the probability of choosing a black ball is given by $\theta_m/(n+\theta_m)$. Similar to the standard model with transition matrix (2), a chosen white ball will be discarded, and a chosen black ball will be discarded, but p additional white balls are added to the urn. One gets the following recurrence for $\mathbb{P}\{X_{m,n} = k\}$:

$$\mathbb{P}\{X_{m,n} = k\} = \frac{n}{n+\theta_m}\mathbb{P}\{X_{m,n-1} = k\} + \frac{\theta_m}{n+\theta_m}\mathbb{P}\{X_{m-1,n+c} = k\},$$

with initial values $\mathbb{P}\{X_{0,n} = n\} = 1$, for $n \in \mathbb{N}_0$.

3 Probabilistic analysis of sampling without replacement and OK Corral type urn models

We will generalize the sampling without replacement urns, and OK Corral urn models by analyzing two urn models associated to sequences of positive real numbers $A = (\alpha_n)_{n\in\mathbb{N}}$ and $B = (\beta_m)_{m\in\mathbb{N}}$. The dynamics of the discrete time process of drawing and replacing balls is as follows: At every discrete time step, we draw a ball from the urn according to the number of white and black balls present in the urn, with respect to the sequences (A, B), subject to the two models defined below. The choosen ball is discarded and the sampling procedure continues until one type of balls is completely drawn.

Urn model I (Sampling without replacement with general weights). Assume that n white and m black balls are contained in the urn, with arbitrary $n, m \in \mathbb{N}$. A white ball is drawn with probability $\alpha_n/(\alpha_n + \beta_m)$, and a black ball is drawn with probability $\beta_m/(\alpha_n+\beta_m)$. Additionally, we assume for urn model I that $\alpha_0=\beta_0=0$.

Urn model II (OK Corral urn model with general weights). For arbitrary $n,m\in\mathbb{N}$ assume that n white and m black balls are contained in the urn. A white ball is drawn with probability $\beta_m/(\alpha_n+\beta_m)$, and a black ball is drawn with probability $\alpha_n/(\alpha_n+\beta_m)$.

The absorbing states, i.e. the points where the evolution of the urn models stop, are given for both urn models by the positive lattice points on the the coordinate axes $\{(0,n)\mid n\geq 1\}\cup\{(m,0)\mid n\geq 1\}$. These two urn models generalize two famous Pólya-Eggenberger urn models with two types of balls, namely the classical sampling without replacement (I), and the so-called OK-Corral urn model (II), described in detail below. We are interested in a probabilistic derivation of the distribution of the random variable $X_{n,m}$, counting the number of white balls, when all black balls have been drawn. In order to simplify the analysis we note that there only exists a single one urn model.

Lemma 1 ([16]) *Let $\mathbb{P}\{X_{n,m,[A,B,I]}=k\}$ denote the probability that k white balls remain when all black balls have been drawn in urn model I with weight sequences $A=(\alpha_n)_{n\in\mathbb{N}}$, $B=(\beta_m)_{m\in\mathbb{N}}$ and $\mathbb{P}\{X_{n,m,[\tilde{A},\tilde{B},II]}=k\}$ the corresponding probability in urn model II with weight sequences $\tilde{A}=(\tilde{\alpha}_n)_{n\in\mathbb{N}}$, $\tilde{B}=(\tilde{\beta}_m)_{m\in\mathbb{N}}$. The probabilities $\mathbb{P}\{X_{n,m,[A,B,I]}=k\}$ and $\mathbb{P}\{X_{n,m,[\tilde{A},\tilde{B},II]}=k\}$ are dual to each other, i.e. they are related in the following way.*

$$\mathbb{P}\{X_{n,m,[A,B,I]}=k\}=\mathbb{P}\{X_{n,m,[\tilde{A},\tilde{B},II]}=k\},$$

for $\alpha_n=\frac{1}{\tilde{\alpha}_n}$, $\beta_m=\frac{1}{\tilde{\beta}_m}$, $n,m\in\mathbb{N}$, and $k>0$.

Without loss of generality, we will restrict ourselves to the urn model I. Note that the recurrence relation for the probability generating function $h_{n,m}(v)=\mathbb{E}(v^{X_{n,m}})$ of $X_{n,m}$ is given by

$$h_{n,m}(v)=\frac{\alpha_n}{\alpha_n+\beta_m}h_{n-1,m}(v)+\frac{\beta_m}{\alpha_n+\beta_m}h_{n,m-1}(v),\quad n,m\geq 1, \tag{5}$$

with initial values $h_{n,0}(v)=v^n$, $h_{0,m}(v)=1$ $n,m\geq 0$. For the sake of simplicity will assume in the following that $\alpha_j\neq\alpha_\ell$ and $\beta_j\neq\beta_\ell$, $1\leq j<\ell<\infty$.

3.1 Probabilistic embedding

We use a probabilistic approach, embedding the discrete-time model into a continuous-time model. The basic idea is as follows. We consider two independent death processes $X(t)$, and $Y(t)$, which stop at 0. Their death rates are defined using the weight sequences, $A=(\alpha_n)_{n\in\mathbb{N}}$, $B=(\beta_m)_{m\in\mathbb{N}}$: the death rates of $X(t)$, starting with $X(0)=n$ are $\alpha_n,\dots,\alpha_1$, and the death rates of $Y(t)$, starting with $Y(0)=m$ are $\beta_m,\dots,\beta_1$. For the sake of convenience we set $\beta_0=0$. We can model the random variable $X_{n,m}$ of urn model I by looking at the distribution of $C_{n,m}=X(\tau)$, starting with $X(0)=n$, where τ denotes the time of the process $Y(t)$ dying out, $\tau=\inf\{t>0:Y(t)=0\}$. By conditioning on the first transition of the two processes one directly obtains the recurrence relation (5) for $\mathbb{E}(v^{C_{n,m}})$, which proves that $X_{n,m}$ and $C_{n,m}$ have the same distribution. Now things are simple. The probability that the process $X(t)=k$, is according to the definition given by the iterated integral

$$\mathbb{P}\{X(t)=k\} = \int_0^t \alpha_n e^{-\alpha_n u_n} \int_0^{t-u_n} \alpha_{n-1} e^{-\alpha_{n-1}u_{n-1}} \dots \\ \dots \int_0^{t-u_n-\dots-u_{k+2}} \alpha_{k+1} e^{-\alpha_{k+1}u_{k+1}} \cdot e^{\alpha_k(t-u_n-\dots-u_{k+2}-u_{k+1})} du_{k+1}\dots du_n.$$

This integral can be evaluated,

$$\mathbb{P}\{X(t)=k\} = \Big(\prod_{h=k+1}^{n} \alpha_h\Big) \sum_{h=k}^{n} \frac{e^{-\alpha_h t}}{\prod_{\substack{j=k\\ j\neq h}}^{n}(\alpha_j-\alpha_h)}, \tag{6}$$

which can easily be checked by induction; here we need the assumptions that $\alpha_j \neq \alpha_\ell$ and $\beta_j \neq \beta_\ell$, $1 \le j < \ell < \infty$. Note that a different integral representation holds true for general weight sequences. This result above is covered in standard textbooks or lecture notes, its derivation is usually based on the Kolomogorov equation and an application of the Laplace transform. The exact distribution of τ is given by

$$\mathbb{P}\{\tau < t\} = \int_0^t \beta_m e^{-\beta_m u_m} du_m \int_0^{t-u_m} \beta_{m-1} e^{-\beta_{m-1}u_{m-1}} \dots \int_0^{t-u_m-\dots-u_2} \beta_1 e^{-\beta_1 u_1} du_1 \dots du_m.$$

One obtains the closed formula

$$\mathbb{P}\{\tau < t\} = 1 + \Big(\prod_{\ell=1}^{m} \beta_\ell\Big) \sum_{\ell=1}^{m} \frac{e^{-\beta_\ell t}}{\prod_{\substack{i=0\\ i\neq \ell}}^{m}(\beta_i-\beta_\ell)},$$

using the convention $\beta_0 = 0$. Hence, then density function of the stopping time τ is given by

$$\frac{d}{dt}\mathbb{P}\{\tau < t\} = \Big(\prod_{\ell=1}^{m} \beta_\ell\Big) \sum_{\ell=1}^{m} \frac{e^{-\beta_\ell t}}{\prod_{\substack{i=1\\ i\neq \ell}}^{m}(\beta_i-\beta_\ell)}. \tag{7}$$

Considering all possible times when the second process dies out leads to the integral representation

$$\begin{aligned}\mathbb{P}\{X_{n,m}=k\} &= \int_0^\infty \frac{d}{dt}\mathbb{P}\{\tau < t\}\mathbb{P}\{X(t)=k\}dt \\ &= \Big(\prod_{\ell=1}^{m} \beta_\ell\Big)\Big(\prod_{h=k+1}^{n} \alpha_h\Big) \sum_{h=k}^{n}\sum_{\ell=1}^{m} \frac{1}{\prod_{\substack{i=1\\ i\neq \ell}}^{m}(\beta_i-\beta_\ell)\prod_{\substack{j=k\\ j\neq h}}^{n}(\alpha_j-\alpha_h)} \int_0^\infty e^{-(\beta_\ell+\alpha_h)t}dt \\ &= \Big(\prod_{\ell=1}^{m} \beta_\ell\Big)\Big(\prod_{h=k+1}^{n} \alpha_h\Big) \sum_{h=k}^{n}\sum_{\ell=1}^{m} \frac{1}{(\beta_\ell+\alpha_h)\prod_{\substack{i=1\\ i\neq \ell}}^{m}(\beta_i-\beta_\ell)\prod_{\substack{j=k\\ j\neq h}}^{n}(\alpha_j-\alpha_h)}.\end{aligned} \tag{8}$$

The result above can be simplified in two different ways using the partial fraction identities

$$\begin{aligned}\frac{1}{\prod_{j=k}^{n}(\alpha_j+x)} &= \sum_{h=k}^{n}\frac{1}{(x+\alpha_h)\prod_{\substack{j=k\\ j\neq h}}^{n}(\alpha_j-\alpha_h)},\\ \frac{1}{\prod_{i=1}^{m}(\beta_i+x)} &= \sum_{\ell=1}^{n}\frac{1}{(x+\beta_\ell)\prod_{\substack{i=1\\ i\neq \ell}}^{m}(\beta_i-\beta_\ell)}.\end{aligned} \tag{9}$$

Consequently, we obtain a transparent probabilistic proof of the following result.

Theorem 4 ([16]) *The probability mass function of the random variable $X_{n,m}$, counting the number of remaining white balls when all black balls have been drawn in urn model I with weight sequences $A=(\alpha_n)_{n\in\mathbb{N}}$, $B=(\beta_m)_{m\in\mathbb{N}}$, is for $n,m\geq 1$ and $n\geq k\geq 1$ given by the explicit formula*

$$\begin{aligned}\mathbb{P}\{X_{n,m}=k\} &= \Big(\prod_{h=1}^{m}\beta_h\Big)\Big(\prod_{h=k+1}^{n}\alpha_h\Big)\sum_{\ell=1}^{m}\frac{1}{\Big(\prod_{j=k}^{n}(\alpha_j+\beta_\ell)\Big)\Big(\prod_{\substack{i=1\\ i\neq\ell}}^{m}(\beta_i-\beta_\ell)\Big)}\\ &= \Big(\prod_{h=1}^{m}\beta_h\Big)\Big(\prod_{h=k+1}^{n}\alpha_h\Big)\sum_{\ell=k}^{n}\frac{1}{\Big(\prod_{\substack{j=k\\ j\neq\ell}}^{n}(\alpha_j-\alpha_\ell)\Big)\Big(\prod_{i=1}^{m}(\beta_i+\alpha_\ell)\Big)},\end{aligned}$$

assuming that $\alpha_j\neq\alpha_\ell$ and $\beta_j\neq\beta_\ell$, $1\leq j<\ell<\infty$, and that $\alpha_0=0$.

It can be shown that the result above is also valid for $k=0$. Moreover, by the duality of the two urn models, one also gets the corresponding result for the urn model II, OK-Corral type urn models, by switching to weight sequences, $\tilde{A}=(1/\alpha_n)_{n\in\mathbb{N}}$, $\tilde{B}=(1/\beta_m)_{m\in\mathbb{N}}$.

3.2 Sums of independent exponential random variables

Of course, the formulas (6), (7) stated before do not come as a surprise, since one can take yet another viewpoint. The time τ until the second process $Y(t)$ dies out has the same distribution as the sum of m independent exponential distributed random variables ϵ_{β_i} with parameters $\beta_m,\dots,\beta_1$ stemming from the death rates of the process. Hence,

$$\tau=\bigoplus_{\ell=1}^{m}\epsilon_{\beta_\ell},$$

where ϵ_{β_i} denotes an exponential distribution with parameter β_i, and the density is simply the formula stated in (7). Furthermore, the distribution of $X(t)$ can also be modeled by k independent random variables: let

$$\theta=\bigoplus_{\ell=k+1}^{n}\epsilon_{\alpha_\ell}.$$

If $X(t)=k$, then the k transitions of the process X have occurred before t, and no more transition afterwards. Hence,

$$\mathbb{P}\{X(t)=k\}=\mathbb{P}\{\theta<t,\theta+\epsilon_{\alpha_k}>t\}=\int_0^t\left(\prod_{h=k+1}^{n}\alpha_h\right)\sum_{h=k+1}^{n}\frac{e^{-\alpha_h u}}{\prod_{\substack{j=k+1\\ j\neq h}}^{n}(\alpha_j-\alpha_h)}e^{-\alpha_k(t-u)}du$$

$$=\sum_{h=k+1}^{n}\frac{\left(\prod_{h=k+1}^{n}\alpha_h\right)e^{-\alpha_h t}}{\prod_{\substack{j=k\\ j\neq h}}^{n}(\alpha_j-\alpha_h)}+\sum_{h=k+1}^{n}\frac{\left(\prod_{h=k+1}^{n}\alpha_h\right)e^{-\alpha_k t}}{(\alpha_h-\alpha_k)\prod_{\substack{j=k+1\\ j\neq h}}^{n}(\alpha_j-\alpha_h)}$$

which simplifies to (6) after an application of (9).

References

[1] Z.-D. Bai, F. Hu, and L.-X. Zhang. Gaussian approximation theorems for urn models and their applications. *Annals of applied probability*, 12,(4):1149–1173, 2002.

[2] C. Brennan and H. Prodinger. The pills problem revisited. *Quaestiones Mathematicae*, 26:427–439, 2003.

[3] T. Hesterberg et al. Problems and solutions, e3429. *American Mathematical Monthly*, 99:684, 1992.

[4] P. Flajolet, P. Dumas, and V. Puyhaubert. Some exactly solvable models of urn process theory. *Discrete Mathematics and Theoretical Computer Science vol. AG*, pages 59–118, 2006.

[5] P. Flajolet, J. Gabarró, and H. Pekari. Analytic urns. *Annals of Probability*, 33:1200–1233, 2005.

[6] P. Flajolet and T. Huillet. Analytic combinatorics of the mabinogion urn. *Discrete mathematics and Theoretical Computer Science (DMTCS), Proceedings of Fifth Colloquium on Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probabilities*, 2008.

[7] H. K. Hwang, M. Kuba, and A. Panholzer. Analysis of some exactly solvable diminishing urn models. *Proceedings of the 19th International Conference on Formal Power Series and Algebraic Combinatorics, Nankai University, Tianjin*, 2007. available at http://dmg.tuwien.ac.at/kuba/urnsub.pdf.

[8] S. Janson. Functional limit theorems for multitype branching processes and generalized pólya urns. *Stochastic processes and applications*, 110:177–245, 2004.

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

[10] N. L. Johnson and S. Kotz. *Urn models and their application. An approach to modern discrete probability theory*. John Wiley, New York, 1977.

[11] S. Kotz and N. Balakrishnan. Advances in urn models during the past two decades. In *Advances in combinatorial methods and applications to probability and statistics, Stat. Ind. Technol.*, pages 203–257. Birkhauser, Boston, 1997.

[12] J. F. C. Kingman. Martingales in the ok corral. *Bulletin of the London Mathematical Society*, 31:601–606, 1999.

[13] J. F. C. Kingman. Stochastic aspects of lanchester's thoery of warfare. *J. Appl. Prob.*, 39:455465, 2002.

[14] D. E. Knuth and J. McCarthy. Problem e3429: Big pills and little pills. *American Mathematical Monthly*, 98:264, 1991.

[15] M. Kuba and A. Panholzer. Limiting distributions for a class of diminishing urn models. *Advances in Applied Probability*, pages 1–31, 2012.

[16] M. Kuba. On sampling without replacement and ok-corral urn models. available at `http://arxiv.org/abs/1003.1603`, 2010.

[17] J. F. C. Kingman and S. E. Volkov. Solution to the ok corral model via decoupling of friedman's urn. *Journal of Theoretical Probability*, 16:267–276, 2003.

[18] H. Mahmoud. Urn models and connections to random trees: a review. *Journal of the Iranian Mathematical Society*, 2:53–114, 2003.

[19] N. Pouyanne. An algebraic approach to pólya processes. *Annales de l'Institut Henri Poincaré*, 44, No. 2:293–323, 2008.

[20] V. Puyhaubert. *Modéles d'urnes et phénomènes de seuil en combinatoire analytique*. PhD thesis, Ecole Polytechnique, Palaiseau, France, 2005.

[21] W. Stadje. Asymptotic probabilities in a sequential urn scheme related to the matchbox problem. *Advances in Applied Probability*, 30:831–849, 1998.

[22] D. Williams and P. McIlroy. The ok corral and the power of the law (a curious poisson kernel formula for a parabolic equation). *Bulletin of the London Mathematical Society*, 30:166–170, 1998.