Statistics and Probability Letters 157 (2020) 108625

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

On the distribution of the hitting time for the N–urn Ehrenfest model

Cheng Xin [a], Minzhi Zhao [a], Qiang Yao [b,*], Erjia Cui [a]

[a] School of Mathematical Sciences, Zhejiang University, Hangzhou 310027, China
[b] School of Statistics, East China Normal University, Shanghai 200241, China

ARTICLE INFO

Article history:
Received 9 April 2019
Accepted 9 September 2019
Available online 25 September 2019

Keywords:
Ehrenfest model
Markov chain
Hitting time
Laplace transform

ABSTRACT

In this paper, we consider the N–urn Ehrenfest model. By utilizing an auxiliary continuous-time Markov chain, we obtain the explicit formula for the Laplace transform of the hitting time from a single state to a set A of states where A satisfies some symmetric properties. After obtaining the Laplace transform, we are able to compute the high-order moments (especially, variance) for the hitting time.

1. Introduction

We consider the N–urn Ehrenfest model, where $N \geq 2$. In this model, there are N urns which are denoted by Urn 1, ... ,Urn N. In the beginning, we place M balls in the N urns in an arbitrary way. Then at each step, we choose a ball randomly and put it into another urn with equal probability. Formally, if we use $x = (x_1, \dots, x_M)$ to denote a state of the model, where $x_i \in \{1, \dots, N\}$ denotes the position of the ith ball, then the N–urn Ehrenfest model can be seen as a time-homogeneous Markov chain $\{X_n : \ n = 0, 1, 2, \dots\}$ on $E = \{1, \dots, N\}^M$. For $x = (x_1, \dots, x_M), y = (y_1, \dots, y_M) \in E$, denote by $s(x, y)$ the number of corresponding coordinates that are the same in x and y, that is,

$$s(x, y) := |\{1 \leq i \leq M : \ x_i = y_i\}| ,$$

where $|\cdot|$ denotes the cardinality of a set. Then the transition probability of the N-urn Ehrenfest model becomes

$$p_{xy} = \begin{cases} \frac{1}{M(N-1)}, & \text{if } s(x, y) = M - 1, \\ 0, & \text{otherwise.} \end{cases} \tag{1.1}$$

For nonempty $A \subseteq E$, denote the hitting time

$$T_A := \inf\{n \geq 0 : \ X_n \in A\},$$

and use T_x as an abbreviation for $T_{\{x\}}$, where $x \in E$. We follow the standard notation for the Markov chain to use P^x to denote the probability when the initial state is x, then use E^x and Var^x to denote the corresponding expectation and variance, respectively.

* Corresponding author.
E-mail addresses: xc1035514096@163.com (C. Xin), zhaomz@zju.edu.cn (M. Zhao), qyao@sfs.ecnu.edu.cn (Q. Yao), churchillcui@zju.edu.cn (E. Cui).

https://doi.org/10.1016/j.spl.2019.108625
0167-7152/

We are interested in the distribution of T_A for some special $A \subseteq E$. Denote by $\mathscr{A}$ the set containing all subsets $A = \{y_1, y_2, \ldots, y_k\} \subseteq E$ such that for each $y \in A$, the sequence $s(y, y_1), s(y, y_2), \ldots, s(y, y_k)$ is the same after being sorted by a monotonic increasing order. It is not difficult to see that the following special A_i $(i = 1, 2, 3, 4, 5)$ all belong to $\mathscr{A}$.

$$A_1 = \{y\}, \text{ where } y \in E;$$
$$A_2 = \{y, z\}, \text{ where } y, z \in E;$$
$$A_3 = \{(1, \ldots, 1), (2, \ldots, 2), \ldots, (N, \ldots, N)\};$$
$$A_4 = \{x \in E : s(x, (2, \ldots, 2)) = h\}, \text{ where } 0 \le h \le M;$$
$$A_5 = \{x = (x_1, \ldots, x_M) \in E : x_1, \ldots, x_M \text{ are all different}\}, \text{ where } M \le N.$$

Consider M permutation functions $\tau_1, \ldots, \tau_M$ from $\{1, \ldots, N\}$ to $\{1, \ldots, N\}$. For any $x = (x_1, \ldots, x_M)$, set $\tau(x) = (\tau_1(x_1), \ldots, \tau_M(x_M))$. Then τ is a one-to-one mapping from E to E. Obviously, $s(\tau(x), \tau(y)) = s(x, y)$ and hence $p_{\tau(x)\tau(y)} = p_{xy}$ for all $x, y \in E$. It follows that $E^{\tau(x)}(e^{-\lambda T_{\tau(A)}}) = E^x(e^{-\lambda T_A})$ for all $x \in E$, $A \subseteq E$ and $\lambda \ge 0$. Clearly, $A \in \mathscr{A}$ if and only if $\tau(A) \in \mathscr{A}$.

Our main result is as follows. It gives the Laplace transform and then illustrates the distribution of the hitting time T_A for $A \in \mathscr{A}$.

Theorem 1.1. *Suppose $x \in E$ and $A \in \mathscr{A}$. Then we have*

$$L(\lambda) := E^x\left(e^{-\lambda T_A}\right) = \begin{cases} \dfrac{\sum_{z\in A} f_{s(x,z)}(M(e^\lambda - 1))}{\sum_{z\in A} f_{s(y,z)}(M(e^\lambda - 1))}, & \text{if } \lambda > 0, \\ 1, & \text{if } \lambda = 0, \end{cases} \tag{1.2}$$

for any $y \in A$, where

$$f_k(u) := \sum_{0\le i\le k,\ 0\le j\le M-k} \frac{C_k^i C_{M-k}^j (N-1)^i (-1)^j}{N(i+j) + u(N-1)} \tag{1.3}$$

for $0 \le k \le M$ and $u > 0$. Here $C_n^m := \dfrac{n!}{m!(n-m)!}$ $(0 \le m \le n)$ denotes the combinatorial number.

Since

$$E^x(T_A) = -L'(0) \text{ and } Var^x(T_A) = L''(0) - (L'(0))^2, \tag{1.4}$$

we can get the following corollary.

Corollary 1.2. *Suppose $x \in E$ and $A \in \mathscr{A}$. Then we have*

$$E^x(T_A) = \frac{M(N-1)}{|A|} \sum_{z\in A} [g_{s(y,z)}(0) - g_{s(x,z)}(0)] \tag{1.5}$$

and

$$Var^x(T_A) = \frac{2M(N-1)}{|A|} \sum_{z\in A} \left[M g'_{s(x,z)}(0) - M g'_{s(y,z)}(0) + E^x(T_A) g_{s(x,z)}(0)\right] + [E^x(T_A)]^2 - E^x(T_A) \tag{1.6}$$

for any $y \in A$, where

$$g_k(u) := \sum_{\substack{0\le i\le k,\ 0\le j\le M-k, \\ i+j\ne 0}} \frac{C_k^i C_{M-k}^j (N-1)^i (-1)^j}{N(i+j) + u(N-1)} \tag{1.7}$$

for $0 \le k \le M$ and $u \ge 0$.

Remark. (1) We can get the higher-order moments for T_A by taking higher-order derivatives of $L(\lambda)$. The notation will be more complicated, so we omit the details here.

(2) Fix $A \in \mathscr{A}$. Applying (1.5), we conclude that $E^x(T_A)$ is a decreasing function of $\sum_{z\in A} g_{s(x,z)}(0)$.

In history, the Ehrenfest model was first proposed in Ehrenfest and Ehrenfest (1907) as "a test bed of key concepts of statistical mechanics" (see Meerson and Zilber (2018)). There are many problems concerning this simple but insightful model. The study of the hitting time was first restricted in the 2–Urn case (when $N = 2$), see Blom (1989), Lathrop et al. (2016), Palacios (1994), etc. Recently, Chen et al. (2017) considered the 3–Urn case and computed $E^x(T_y)$ when $s(x, y) = 0$.

Then Song and Yao (2019) extended their result to all $N \geq 2$ and all $x, y \in E$. The N–Urn model has attracted attentions in application fields recently, see Aloisi and Nali (2018), for example. So it is worthwhile to investigate the model more deeply. The authors of all the above references did not consider the distribution of the hitting time. And they did not consider T_A for A other than a singleton. Furthermore, the methods in Chen et al. (2017) and Song and Yao (2019) cannot be used to prove Theorem 1.1. In this paper, we adopt a new method to solve the problem. The new method relies on an auxiliary continuous-time Markov chain, which is the main contribution of this paper.

The organization of this paper is as follows. In Section 2, we introduce an auxiliary continuous-time Markov chain and explore the relationship with the original discrete-time Markov chain. In Section 3, we prove Theorem 1.1 with the help of the above auxiliary chain, and then prove Corollary 1.2. In Section 4, we give some examples and use Corollary 1.2 to extend the results in Chen et al. (2017) and Song and Yao (2019).

2. An auxiliary continuous-time Markov chain

In this section, we introduce a continuous time Markov chain on $E = \{1, \ldots, N\}^M$. Let $\{Y_1(t) : t \geq 0\}, \ldots, \{Y_M(t) : t \geq 0\}$ be M independent continuous-time Markov chains on $\{1, \ldots, N\}$ with the same Q-matrix

$$Q = \begin{bmatrix} -1 & \frac{1}{N-1} & \cdots & \frac{1}{N-1} \\ \frac{1}{N-1} & -1 & \cdots & \frac{1}{N-1} \\ \vdots & \vdots & & \vdots \\ \frac{1}{N-1} & \frac{1}{N-1} & \cdots & -1 \end{bmatrix}. \tag{2.1}$$

Then define $Y(t) := (Y_1(t), \ldots, Y_M(t))$ for all $t \geq 0$. It follows that $\{Y(t) : t \geq 0\}$ is a continuous-time Markov chain with state space E. The next proposition gives the basic relationship between $\{X_n : n = 0, 1, 2, \ldots\}$ and $\{Y(t) : t \geq 0\}$.

Proposition 2.1. *If X_0 and $Y(0)$ have the same distribution, then $\{X_n : n = 0, 1, 2, \ldots\}$ and the embedded chain of $\{Y(t) : t \geq 0\}$ have the same finite-dimensional distributions.*

Proof. Set $\sigma_0 := 0$ and $\sigma_n := \inf\{t > \sigma_{n-1} : Y(t) \neq Y(\sigma_{n-1})\}$ for $n \geq 1$. Then for $n \geq 1$, denote $\xi_n := \sigma_n - \sigma_{n-1}$. Since $\xi_1, \xi_2, \ldots$ are i.i.d. exponential random variables with parameter M, and $\{\xi_1, \xi_2, \ldots\}$ are independent of $\{Y(\sigma_n) : n = 0, 1, 2 \cdots\}$, we can deduce that $\{Y(\sigma_n) : n = 0, 1, 2 \cdots\}$ is a time-homogeneous discrete-time Markov chain on E with transition probability

$$P(Y(\sigma_{n+1}) = y \mid Y(\sigma_n) = x) = \begin{cases} \frac{1}{M(N-1)}, & \text{if } s(x, y) = M - 1, \\ 0, & \text{otherwise}. \end{cases} \tag{2.2}$$

See Chapter 3 of Lawler (2006) for details. Comparing (2.2) with (1.1), we can get that $\{Y(\sigma_n)\}$ has the same finite-dimensional distribution with $\{X_n\}$, as desired. □

From Proposition 2.1, we can define $\{Y(t)\}$ and $\{X_n\}$ on the same probability space and treat $\{X_n\}$ as the embedded chain of $\{Y(t)\}$. So we can use the same notations P^x, E^x and Var^x when considering the both processes. Next, we use the superscript "Y" to denote the hitting time for $\{Y(t)\}$. That is, for nonempty $A \subseteq E$, denote the hitting time

$$T_A^Y := \inf\{t \geq 0 : Y(t) \in A\},$$

and use T_x^Y as an abbreviation for $T_{\{x\}}^Y$, where $x \in E$. The next proposition shows the relationship of the Laplace transforms between T_A and T_A^Y.

Proposition 2.2. *For any $x \in E$, $A \subseteq E$ and $\lambda \geq 0$, we have*

$$E^x\left(e^{-\lambda T_A}\right) = E^x\left(e^{-M(e^{\lambda}-1)T_A^Y}\right). \tag{2.3}$$

Proof. Note that $T_A^Y = \sum_{i=1}^{T_A} \xi_i$. Since T_A is independent of $\{\xi_1, \xi_2, \ldots\}$, and $\xi_1, \xi_2, \ldots$ are i.i.d exponential random variables with parameter M, we can deduce that for any $u \geq 0$,

$$\begin{aligned} E^x(e^{-uT_A^Y}) &= E^x\left(\exp\left\{-u\sum_{i=1}^{T_A}\xi_i\right\}\right) = E^x\left(\prod_{i=1}^{T_A} e^{-u\xi_i}\right) \\ &= E^x\left(E^x\left(\prod_{i=1}^{T_A} e^{-u\xi_i} \,\middle|\, T_A\right)\right) = E^x\left[\left(E^x(e^{-u\xi_1})\right)^{T_A}\right] \\ &= E^x\left[\left(\int_0^\infty e^{-ut} M e^{-Mt} dt\right)^{T_A}\right] = E^x\left[\left(\frac{M}{u+M}\right)^{T_A}\right]. \end{aligned} \tag{2.4}$$

Then for any $\lambda \geq 0$, take $u = M(e^{\lambda} - 1) \geq 0$, we have $\frac{M}{u+M} = e^{-\lambda}$. Substitute it into Eq. (2.4), we conclude that

$$E^x\left(e^{-\lambda T_A}\right) = E^x\left(e^{-M(e^{\lambda}-1)T_A^Y}\right),$$

as desired. □

As a corollary, we can get the following relationship of the expectations and variances between T_A and T_A^Y.

Corollary 2.3. *For any $x \in E$ and $A \subseteq E$, we have*

$$E^x(T_A) = ME^x(T_A^Y) \ \text{ and } \ Var^x(T_A) = M^2 Var^x(T_A^Y) - E^x(T_A).$$

Proof. Fix $x \in E$ and $A \subseteq E$. Define

$$L(\lambda) := E^x\left(e^{-\lambda T_A}\right), \quad L_Y(\lambda) := E^x\left(e^{-\lambda T_A^Y}\right)$$

for $\lambda \geq 0$. By Proposition 2.2, we have

$$L(\lambda) = L_Y\left(M(e^{\lambda} - 1)\right). \tag{2.5}$$

It is not difficult to get

$$L'(0) = ML_Y'(0) \ \text{ and } \ L''(0) = M^2 L_Y''(0) + L'(0).$$

Therefore, we can conclude that

$$E^x(T_A) = -L'(0) = -ML_Y'(0) = ME^x(T_A^Y),$$

and

$$\begin{aligned} Var^x(T_A) &= L''(0) - (L'(0))^2 \\ &= M^2[L_Y''(0) - (L_Y'(0))^2] + L'(0) = M^2 Var^x(T_A^Y) - E^x(T_A), \end{aligned}$$

as desired. □

Remark. By taking higher-order derivatives from Eq. (2.5), we can get the relationship of the higher-order moments between T_A and T_A^Y. Since the number of terms will increase in the higher-order moment case, we omit the detailed computation here.

3. Proofs of Theorem 1.1 and Corollary 1.2

By Proposition 2.2, we can see that we only need to calculate the Laplace transform of T_A^Y to obtain the Laplace transform of T_A. For a general $A \subseteq E$, the Laplace transform of T_A^Y cannot be easily calculated. However, for $A \in \mathscr{A}$ (the definition was given in Section 1), we can use the symmetric property to deal with it.

Denote by $\{p_t\}$ the transition semigroup of $\{Y_1(t)\}$ whose Q-matrix Q is given in (2.1). That is, $p_t(i,j) = P^i(Y_1(t) = j)$ for all $i, j \in \{1, \ldots, N\}$ and $t \geq 0$. The next lemma gives an explicit formula for $\{p_t\}$.

Lemma 3.1. *For any $i, j \in \{1, \ldots, N\}$ and $t \geq 0$, we have*

$$p_t(i,j) = \begin{cases} \frac{(N-1)e^{-\frac{N}{N-1}t}+1}{N}, & \text{if } i = j, \\ \frac{1-e^{-\frac{N}{N-1}t}}{N}, & \text{otherwise}. \end{cases}$$

Proof. By the symmetric property of the Q-matrix Q, we have for any $t \geq 0$,

$$p_t(i,j) = \begin{cases} p_t(1,1), & \text{if } i = j, \\ p_t(1,2), & \text{otherwise}. \end{cases}$$

Since $\sum_{j=1}^N p_t(i,j) = 1$, it follows that $p_t(1,2) = \dfrac{1 - p_t(1,1)}{N-1}$.

By Kolmogorov's backward equation, we have $p_t' = Qp_t$. So

$$\begin{aligned} p_t'(1,1) &= \sum_{i=1}^{N} q_{1i} p_t(i,1) \\ &= -p_t(1,1) + \sum_{i=2}^{N} \frac{1}{N-1} p_t(i,1) \\ &= -p_t(1,1) + p_t(1,2) \\ &= -p_t(1,1) + \frac{1-p_t(1,1)}{N-1} \\ &= -\frac{N}{N-1} p_t(1,1) + \frac{1}{N-1}. \end{aligned}$$

Since $p_0(1,1) = 1$, by solving the first-order linear ordinary differential equation, we get

$$p_t(1,1) = \frac{(N-1)e^{-\frac{N}{N-1}t} + 1}{N}.$$

And therefore,

$$p_t(1,2) = \frac{1-p_t(1,1)}{N-1} = \frac{1-e^{-\frac{N}{N-1}t}}{N}.$$

This completes the proof. □

Proof of Theorem 1.1. For $x \in E$, $A \in \mathscr{A}$ and $u > 0$, define

$$G_u(x,A) := E^x\left[\int_0^\infty e^{-ut} \mathbf{1}_A(Y(t))dt\right],$$

where

$$\mathbf{1}_A(Y(t)) := \begin{cases} 1, & \text{if } Y(t) \in A, \\ 0, & \text{otherwise.} \end{cases}$$

Denote by $\{P_t\}$ the transition semigroup of $\{Y(t)\}$. That is, $P_t(y,z) = P^y(Y(t) = z)$ for all $y, z \in E$ and $t \geq 0$. By Lemma 3.1, we have for any $x = (x_1, \ldots, x_M), z = (z_1, \ldots, z_M) \in E$,

$$P_t(x,z) = \prod_{i=1}^{M} p_t(x_i, z_i) = p_t^k(1,1) p_t^{M-k}(1,2) = \frac{1}{N^M}((N-1)e^{-\frac{N}{N-1}t} + 1)^k (1 - e^{-\frac{N}{N-1}t})^{M-k},$$

where $k = s(x,z)$. Thus for any $u > 0$,

$$\begin{aligned} G_u(x,\{z\}) &= \int_0^\infty e^{-ut} P_t(x,z)dt \\ &= \frac{1}{N^M}\int_0^\infty e^{-ut}\left((N-1)e^{-\frac{N}{N-1}t} + 1\right)^k \left(1 - e^{-\frac{N}{N-1}t}\right)^{M-k} dt \\ &\overset{s=e^{-\frac{N}{N-1}t}}{=\!=\!=} \frac{N-1}{N^{M+1}} \int_0^1 s^{\frac{N-1}{N}u-1}[(N-1)s+1]^k (1-s)^{M-k} ds \\ &= \frac{N-1}{N^{M+1}} \int_0^1 s^{\frac{N-1}{N}u-1} \sum_{i=0}^{k} C_k^i (N-1)^i s^i \sum_{j=0}^{M-k} C_{M-k}^j (-1)^j s^j ds \\ &= \frac{N-1}{N^{M+1}} \sum_{0\leq i\leq k,\ 0\leq j\leq M-k} C_k^i C_{M-k}^j (N-1)^i (-1)^j \int_0^1 s^{\frac{N-1}{N}u-1+i+j} ds \\ &= \frac{N-1}{N^M} \sum_{0\leq i\leq k,\ 0\leq j\leq M-k} \frac{C_k^i C_{M-k}^j (N-1)^i (-1)^j}{N(i+j)+u(N-1)} = \frac{N-1}{N^M} f_k(u). \end{aligned} \tag{3.1}$$

This, together with the fact that $A \in \mathscr{A}$, yields $G_u(y, A) = G_u(z, A)$ for all $y, z \in A$ and $u > 0$. Furthermore, by the strong Markov property, we obtain that for any $y \in A$ and $u > 0$,

$$G_u(x, A) = E^x \left[\int_0^\infty e^{-ut} \mathbf{1}_A(Y(t))dt \right] = E^x \left[\int_{T_A^Y}^\infty e^{-ut} \mathbf{1}_A(Y(t))dt \right] = E^x \left(e^{-uT_A^Y} \right) G_u(y, A)$$

and hence

$$E^x \left(e^{-uT_A^Y} \right) = \frac{G_u(x, A)}{G_u(y, A)} = \frac{\sum_{z\in A} G_u(x, \{z\})}{\sum_{z\in A} G_u(y, \{z\})} = \frac{\sum_{z\in A} f_{s(x,z)}(u)}{\sum_{z\in A} f_{s(y,z)}(u)}. \tag{3.2}$$

Therefore, the desired result (1.2) follows immediately from (2.3) and (3.2). □

Remark. From (3.1), we can get that

$$f_k(u) = \frac{1}{N} \int_0^1 s^{\frac{N-1}{N}u-1} [(N-1)s + 1]^k (1-s)^{M-k} ds \tag{3.3}$$

for any $0 \le k \le M$ and $u > 0$. We call (3.3) the "integral version" for $f_k(u)$. It will be useful for simplifying the formulas in Section 4.

Proof of Corollary 1.2. From (1.3) and (1.7), we can see that

$$g_k(u) = f_k(u) - \frac{1}{u(N-1)} \tag{3.4}$$

for any $0 \le k \le M$ and $u > 0$. Then by (1.2) and (3.4), we have for any $\lambda \ge 0$ and $y \in A$,

$$L(\lambda) = \frac{|A| + M(N-1)(e^\lambda - 1) \sum_{z\in A} g_{s(x,z)}(M(e^\lambda - 1))}{|A| + M(N-1)(e^\lambda - 1) \sum_{z\in A} g_{s(y,z)}(M(e^\lambda - 1))} =: \frac{L_1(\lambda)}{L_2(\lambda)},$$

where $L_1(\lambda)$ denotes the numerator and $L_2(\lambda)$ denotes the denominator. Hence $L_1(\lambda) = L(\lambda)L_2(\lambda)$. Therefore, by Leibnitz's formula we can get

$$\begin{cases} L_1'(0) = L'(0)L_2(0) + L(0)L_2'(0), \\ L_1''(0) = L''(0)L_2(0) + 2L'(0)L_2'(0) + L(0)L_2''(0). \end{cases} \tag{3.5}$$

Note that

$$\begin{cases} L_1(0) = L_2(0) = |A|, \\ L_1'(0) = M(N-1) \sum\limits_{z\in A} g_{s(x,z)}(0), \;\; L_2'(0) = M(N-1) \sum\limits_{z\in A} g_{s(y,z)}(0), \\ L_1''(0) = M(N-1) \left[\sum\limits_{z\in A} g_{s(x,z)}(0) + 2M \sum\limits_{z\in A} g'_{s(x,z)}(0) \right], \\ L_2''(0) = M(N-1) \left[\sum\limits_{z\in A} g_{s(y,z)}(0) + 2M \sum\limits_{z\in A} g'_{s(y,z)}(0) \right]. \end{cases} \tag{3.6}$$

The desired results (1.5) and (1.6) then follow from (1.4), (3.5) and (3.6). □

4. Some special examples

In this section, we consider several special examples. To simplify the notation, we use g_k as an abbreviation for $g_k(0)$ in this section ($0 \le k \le M$). Before discussing the examples, we first give the expressions for g_0, g_M and $g_{k+1} - g_k$, that will be used later.

Note that for any real number a,

$$\sum_{i=1}^M \frac{C_M^i a^i}{i} = \int_0^a \frac{(1+t)^M - 1}{t} dt = \sum_{i=1}^M \int_0^a (1+t)^{i-1} dt = \sum_{i=1}^M \frac{(1+a)^i - 1}{i}. \tag{4.1}$$

By (1.7) and (4.1), we obtain

$$g_0 = \frac{1}{N} \sum_{i=1}^M \frac{C_M^i (-1)^i}{i} = -\frac{1}{N} \sum_{i=1}^M \frac{1}{i} \tag{4.2}$$

and

$$g_M = \frac{1}{N}\sum_{i=1}^{M}\frac{C_M^i(N-1)^i}{i} = \frac{1}{N}\sum_{i=1}^{M}\frac{N^i-1}{i}. \tag{4.3}$$

For $0 \le k \le M-1$, by (1.7), (3.3) and (3.4), we have

$$\begin{aligned} g_{k+1} - g_k &= \lim_{u\downarrow 0}[f_{k+1}(u) - f_k(u)] \\ &= \int_0^1 [(N-1)s+1]^k(1-s)^{M-k-1}ds \\ &= \sum_{i=0}^{k} C_k^i(N-1)^{k-i}\int_0^1 s^{k-i}(1-s)^{M-k-1}ds \\ &= \sum_{i=0}^{k} C_k^i(N-1)^{k-i}\frac{(k-i)!(M-k-1)!}{(M-i)!} \\ &= \frac{(N-1)^k}{MC_{M-1}^k}\sum_{i=0}^{k}\frac{C_M^i}{(N-1)^i}. \end{aligned} \tag{4.4}$$

4.1. First hitting time to a fixed singleton

When $A = \{y\}$ where $y \in E$, we have $A \in \mathscr{A}$. Therefore, by Theorem 1.1, we get that for $x, y \in E$,

$$E^x\left(e^{-\lambda T_y}\right) = \begin{cases} \dfrac{f_{s(x,y)}\left(M(e^\lambda - 1)\right)}{f_M\left(M(e^\lambda-1)\right)}, & \text{if } \lambda > 0, \\ 1, & \text{if } \lambda = 0. \end{cases} \tag{4.5}$$

Furthermore, by (1.5) in Corollary 1.2, we get that for any $x, y \in E$,

$$E^x(T_y) = M(N-1)[g_M - g_{s(x,y)}]. \tag{4.6}$$

Especially, for any $x, y \in E$ such that $s(x, y) = 0$ (for example, when $x = (1, \ldots, 1)$ and $y = (2, \ldots, 2)$), we have

$$E^x(T_y) = M(N-1)[g_M - g_0]. \tag{4.7}$$

We now explain how (4.6) and (4.7) match the results of Theorem 1.1 and Corollary 1.2 in Song and Yao (2019) respectively. Combining (4.6) with (4.4) we obtain

$$E^x(T_y) = M(N-1)\sum_{j=k}^{M-1}(g_{j+1} - g_j) = \sum_{j=k}^{M-1}\frac{(N-1)^{j+1}}{C_{M-1}^j}\sum_{l=0}^{j}\frac{C_M^l}{(N-1)^l}, \tag{4.8}$$

where $k = s(x, y)$. This is exactly the result of Theorem 1.1 in Song and Yao (2019). Applying (4.2) and (4.3), we can rewrite (4.7) as

$$E^x(T_y) = \frac{M(N-1)}{N}\sum_{i=1}^{M}\frac{N^i}{i} \tag{4.9}$$

for any $x, y \in E$ satisfying $s(x, y) = 0$. This is exactly the result of Corollary 1.2 in Song and Yao (2019).

Next, we compute $Var^x(T_y)$ for $x, y \in E$. This was not done in Song and Yao (2019). By (1.6) in Corollary 1.2, we get that for any $x, y \in E$,

$$Var^x(T_y) = 2M(N-1)\left[Mg'_{s(x,y)}(0) - Mg'_M(0) + E^x(T_y)g_{s(x,y)}\right] + \left[E^x(T_y)\right]^2 - E^x(T_y). \tag{4.10}$$

Especially, when $s(x, y) = 0$, (4.10) becomes

$$Var^x(T_y) = 2M(N-1)\left[Mg'_0(0) - Mg'_M(0) + E^x(T_y)g_0\right] + \left[E^x(T_y)\right]^2 - E^x(T_y). \tag{4.11}$$

By (4.1), for any real number a,

$$\sum_{i=1}^{M}\frac{C_M^i a^i}{i^2} = \int_0^a\sum_{i=1}^{M}\frac{C_M^i t^{i-1}}{i}dt = \int_0^a\sum_{i=1}^{M}\frac{(1+t)^i-1}{it}dt = \sum_{i=1}^{M}\frac{1}{i}\sum_{j=1}^{i}\frac{(1+a)^j-1}{j}. \tag{4.12}$$

With the help of (1.7), it follows that

$$g_0'(0) - g_M'(0) = \frac{N-1}{N^2} \sum_{i=1}^{M} \frac{C_M^i[(N-1)^i - (-1)^i]}{i^2} = \frac{N-1}{N^2} \sum_{i=1}^{M} \frac{1}{i} \sum_{j=1}^{i} \frac{N^j}{j}. \tag{4.13}$$

Putting (4.9), (4.2) and (4.13) into (4.11), we have

$$Var^x(T_y) = \frac{M^2(N-1)^2}{N^2} \left[\left(\sum_{i=1}^{M} \frac{N^i}{i} \right)^2 - 2 \sum_{i=1}^{M} \frac{1}{i} \sum_{j=i+1}^{M} \frac{N^j}{j} \right] - \frac{M(N-1)}{N} \sum_{i=1}^{M} \frac{N^i}{i}$$

for any $x, y \in E$ such that $s(x,y) = 0$.

4.2. First hitting time to a fixed two-points set

Suppose x, y, z are three different points in E. Let $A = \{y, z\}$. We next compute $E^x(T_A)$ and $P^x(X_{T_A} = y)$. Since $A \in \mathscr{A}$. By (1.5) and (4.6), we have

$$\begin{cases} E^x(T_A) = \frac{M(N-1)}{2} \left[g_M + g_{s(y,z)} - g_{s(x,y)} - g_{s(x,z)} \right], \\ E^x(T_z) = M(N-1) \left[g_M - g_{s(x,z)} \right], \\ E^y(T_z) = M(N-1) \left[g_M - g_{s(y,z)} \right]. \end{cases} \tag{4.14}$$

On the other hand, by strong Markov property, we get

$$E^x(T_z) = E^x(T_A) + P^x(X_{T_A} = y) E^y(T_z). \tag{4.15}$$

From (4.14) and (4.15), we obtain

$$P^x(X_{T_A} = y) = \frac{E^x(T_z) - E^x(T_A)}{E^y(T_z)} = \frac{g_M + g_{s(x,y)} - g_{s(x,z)} - g_{s(y,z)}}{2\left[g_M - g_{s(y,z)} \right]}.$$

4.3. First time that all balls are in the same urn

Proposition 4.1. *Set $x = (x_1, x_2, \ldots, x_M) \in E$ and $A = \{(i, \ldots, i) : i = 1, \ldots, N\}$. Then we have*

$$E^x(T_A) = \frac{M(N-1)}{N} \left[g_M + (N-1) g_0 - \sum_{k=1}^{N} g_{n_k} \right],$$

and for $1 \le i \le N$,

$$P^x(X_{T_A} = (i, \ldots, i)) = \frac{1}{N} + \frac{g_{n_i} - \frac{1}{N} \sum_{k=1}^{N} g_{n_k}}{g_M - g_0},$$

where $n_i = s(x, (i, \ldots, i))$ for $i = 1, \ldots, N$.

Proof. Set $y = (1, \ldots, 1)$, then $y \in A \in \mathscr{A}$. By (1.5), we have

$$E^x(T_A) = \frac{M(N-1)}{|A|} \sum_{z \in A} \left[g_{s(y,z)} - g_{s(x,z)} \right] = \frac{M(N-1)}{N} \left[g_M + (N-1) g_0 - \sum_{k=1}^{N} g_{n_k} \right].$$

Next, set $z = (i, \ldots, i)$. By the strong Markov property, we obtain that

$$E^x(T_z) = E^x(T_A) + \sum_{j \neq i} P^x(X_{T_A} = (j, \ldots, j)) E^{(j,\ldots,j)}(T_z). \tag{4.16}$$

By (4.6), $E^x(T_z) = M(N-1)(g_M - g_{n_i})$ and $E^{(j,\ldots,j)}(T_z) = M(N-1)(g_M - g_0)$ for $j \neq i$. Hence (4.16) can be rewritten as

$$g_M - g_{n_i} = \frac{1}{N} \left[g_M + (N-1) g_0 - \sum_{k=1}^{N} g_{n_k} \right] + [1 - P^x(X_{T_A} = (i, \ldots, i))](g_M - g_0).$$

Therefore,

$$P^x(X_{T_A} = (i, \ldots, i)) = \frac{1}{N} + \frac{g_{n_i} - \frac{1}{N} \sum_{k=1}^{N} g_{n_k}}{g_M - g_0}$$

for $1 \le i \le N$, as desired. □

Corollary 4.2. *If $M \leq N$ and $x = (1, 2, \ldots, M)$, then we have*

$$E^x(T_A) = \frac{M(N-1)}{N^2} \sum_{i=2}^{M} \frac{N^i}{i}$$

and

$$P^x(X_{T_A} = (i, \ldots, i)) = \begin{cases} \frac{1}{N} + \frac{N-M}{M \sum_{i=1}^{M} \frac{N^i}{i}}, & \text{if } i \leq M, \\ \frac{1}{N} - \frac{1}{\sum_{i=1}^{M} \frac{N^i}{i}}, & \text{if } M < i \leq N. \end{cases}$$

Proof. For $1 \leq i \leq N$,

$$n_i = s(x, (i, \ldots, i)) = \begin{cases} 1, & \text{if } i \leq M, \\ 0, & \text{otherwise.} \end{cases}$$

Putting $k = 0$ into (4.4) yields

$$g_1 = g_0 + \frac{1}{M}. \tag{4.17}$$

By putting (4.2), (4.3) and (4.17) into Proposition 4.1, we get the desired results. □

4.4. First time that all balls are in different urns

Proposition 4.3. *Suppose that $M = N \geq 2$. Set $x = (1, \ldots, 1)$ and*

$$B = \{(i_1, \ldots, i_M) : (i_1, \ldots, i_M) \text{ is a permutation of } \{1, \ldots, M\}\}.$$

Then we have

$$E^x(T_B) = M(M-1)\left[-g_1 + \sum_{k=0}^{M-2} \frac{1}{k!}\left(\frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^{M-k}\frac{1}{(M-k)!}\right) g_k\right] + \frac{1}{(M-2)!} g_M.$$

Proof. Set $y = (1, 2, \ldots, M)$. Then $y \in B \in \mathscr{A}$. By (1.5), we have

$$E^x(T_B) = \frac{M(M-1)}{|B|} \sum_{z \in B} \left[g_{s(y,z)} - g_{s(x,z)}\right] = M(M-1)\left[\sum_{k=0}^{M} \frac{a_k}{M!} g_k - g_1\right],$$

where $a_k = |\{z \in B : s(y, z) = k\}|$. Pick an element of B with equal probability and denote it by $Z = (Z_1, Z_2, \ldots, Z_M)$. Let $\eta = s(Z, y)$. Then $\frac{a_k}{M!} = P(\eta = k)$ for $0 \leq k \leq M$. Since

$$P(\eta = k) = \frac{1}{k!}\left(\frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^{M-k}\frac{1}{(M-k)!}\right)$$

for $k \leq M - 2$, $P(\eta = M - 1) = 0$ and $P(\eta = M) = \frac{1}{M!}$, the proof is completed. □

4.5. First time that h balls are in a fixed urn (for example, Urn 2)

In this subsection, set $A_h = \{y \in E : s(y, (2, \ldots, 2)) = h\}$ and $s(x, (2, 2, \ldots, 2)) = k$.

Proposition 4.4. *If $k < h$, then*

$$E^x(T_{A_h}) = M(N-1)(g_h - g_k) = \sum_{i=k}^{h-1} \frac{(N-1)^{i+1}}{C_{M-1}^i} \sum_{j=0}^{i} \frac{C_M^j}{(N-1)^j}. \tag{4.18}$$

If $k > h$, then

$$E^x(T_{A_h}) = \sum_{i=h}^{k-1} \frac{(N-1)^{i+1}}{C_{M-1}^i} \sum_{j=i+1}^{M} \frac{C_M^j}{(N-1)^j}. \tag{4.19}$$

Proof. Firstly, suppose that $k<h$. By (4.4), (4.6) and the strong Markov property, we have

$$\begin{aligned}E^x(T_{A_h}) &= E^x(T_{(2,\dots,2)}) - E^x\left(E^{X_{T_{A_h}}}(T_{(2,2,\dots,2)})\right)\\ &= M(N-1)(g_M-g_k) - M(N-1)(g_M-g_h)\\ &= \sum_{i=k}^{h-1}\frac{(N-1)^{i+1}}{C_{M-1}^i}\sum_{j=0}^{i}\frac{C_M^j}{(N-1)^j}.\end{aligned}$$

Thus (4.18) holds.

Now we suppose that $k>h$. Let $y=(1,\dots,1)$. Then $y\in A_0\in\mathscr{A}$. By (1.5), we get

$$E^x(T_{A_0}) = \frac{M(N-1)}{|A_0|}\sum_{z\in A_0}[g_{s(y,z)}-g_{s(x,z)}].$$

Pick an element of A_0 with equal probability and denote it by Z. Let $\eta_1=s(y,Z)$ and $\eta_2=s(x,Z)$. Then we have

$$E^x(T_{A_0}) = M(N-1)[E(g_{\eta_1})-E(g_{\eta_2})].$$

It is easy to check that $\eta_1\sim\mathrm{Bin}\left(M,\frac{1}{N-1}\right)$ and $\eta_2\sim\mathrm{Bin}\left(M-k,\frac{1}{N-1}\right)$. Suppose that $\zeta_1,\zeta_2,\dots,\zeta_k$ are i.i.d., independent of η_2, and $\zeta_1\sim\mathrm{Bin}\left(1,\frac{1}{N-1}\right)$. For convenience, set $S_i=\eta_2+\sum_{l=1}^{i}\zeta_l$ for $0\le i\le k$. Then $S_i\sim\mathrm{Bin}\left(M-k+i,\frac{1}{N-1}\right)$. In particular, S_k has the same distribution as η_1. It follows that

$$E^x(T_{A_0}) = M(N-1)E(g_{S_k}-g_{S_0}) = M(N-1)\sum_{i=0}^{k-1}E(g_{S_{i+1}}-g_{S_i}).$$

For $0\le i\le k-1$,

$$E(g_{S_{i+1}}-g_{S_i}) = E(g_{S_i+\zeta_{i+1}}-g_{S_i}) = P(\zeta_{i+1}=1)E(g_{S_i+1}-g_{S_i}).$$

If we have showed that for any $\zeta\sim\mathrm{Bin}\left(m,\frac{1}{N-1}\right)$ with $0\le m\le M-1$,

$$E(g_{\zeta+1}-g_\zeta) = \frac{(N-1)^{M-m}}{MC_{M-1}^m}\sum_{i=M-m}^{M}\frac{C_M^i}{(N-1)^i}, \tag{4.20}$$

then

$$E^x(T_{A_0}) = \sum_{i=0}^{k-1}\frac{(N-1)^{k-i}}{C_{M-1}^{k-i-1}}\sum_{u=k-i}^{M}\frac{C_M^u}{(N-1)^u} = \sum_{j=0}^{k-1}\frac{(N-1)^{j+1}}{C_{M-1}^j}\sum_{u=j+1}^{M}\frac{C_M^u}{(N-1)^u}.$$

It follows that for $k>h$,

$$E^x(T_{A_h}) = E^x(T_{A_0}) - E^x\left(E^{X_{T_{A_h}}}(T_{A_0})\right) = \sum_{i=h}^{k-1}\frac{(N-1)^{i+1}}{C_{M-1}^i}\sum_{j=i+1}^{M}\frac{C_M^j}{(N-1)^j}.$$

That is, (4.19) holds. Therefore it remains to prove (4.20). Applying (3.3) and (3.4), we obtain

$$\begin{aligned}E(g_{\zeta+1}-g_\zeta) &= E\left\{\lim_{u\downarrow 0}[f_{\zeta+1}(u)-f_\zeta(u)]\right\}\\ &= E\left[\int_0^1((N-1)s+1)^\zeta(1-s)^{M-\zeta-1}ds\right]\\ &= \int_0^1(1-s)^{M-1}E\left[\left(\frac{(N-1)s+1}{1-s}\right)^\zeta\right]ds\\ &= \int_0^1(1-s)^{M-1}\left[1-\frac{1}{N-1}+\frac{1}{N-1}\frac{(N-1)s+1}{1-s}\right]^m ds\end{aligned}$$

$$= (N-1)^{-m}\int_0^1 (N-1+s)^m(1-s)^{M-1-m}ds$$

$$= (N-1)^{-m}\sum_{j=0}^{m} C_m^j (N-1)^j \int_0^1 s^{m-j}(1-s)^{M-1-m}ds$$

$$= \frac{1}{MC_{M-1}^m}\sum_{j=0}^{m} C_M^j (N-1)^{j-m}$$

$$\overset{i=M-j}{=\!=\!=} \frac{(N-1)^{M-m}}{MC_{M-1}^m}\sum_{i=M-m}^{M}\frac{C_M^i}{(N-1)^i}$$

as desired. □

Remark. We can also use the method of electric networks to prove (4.19) after getting (4.18). Readers can refer to Doyle and Snell (1984) and Lyons and Peres (2017) for the detailed instruction of this method. For $n = 0, 1, \ldots$, let $\Phi_n = s(X_n, (2, \ldots, 2))$. Then Φ_n denotes the number of balls in Urn 2 at time n. We see that $\{\Phi_n : n = 0, 1, \ldots\}$ is the random walk on the electric network $\{0, 1, \ldots, M\}$ with

$$C_{i,i+1} = \frac{C_{M-1}^i}{(N-1)^{i+1}} \text{ and } C_{i,i} = (N-2)C_{i,i+1} \quad (i = 0, 1, \ldots, M).$$

It is easy to check that $C_i = \dfrac{C_M^i}{(N-1)^i}$ for $i = 0, 1, \ldots, M$. Use E_Φ^k to denote the expectation when $\Phi_0 = k$. For any $0 \le h \le M$, set $T_h^\Phi = \inf\{n \ge 0 : \Phi_n = h\}$. Clearly, for any $x \in E$ and $0 \le h \le M$, we have

$$E^x(T_{A_h}) = E_\Phi^{s(x,(2,\ldots,2))}(T_h^\Phi). \tag{4.21}$$

When $0 \le h < k \le M$, by Corollary 2.21 on Page 48 of Lyons and Peres (2017), we get

$$E_\Phi^k(T_h^\Phi) + E_\Phi^h(T_k^\Phi) = \sum_{i=0}^{M} C_i \mathcal{R}(h \leftrightarrow k)$$
$$= \sum_{i=0}^{M}\frac{C_M^i}{(N-1)^i}\sum_{j=h}^{k-1}\frac{1}{C_{j,j+1}} = \sum_{i=0}^{M}\frac{C_M^i}{(N-1)^i}\sum_{j=h}^{k-1}\frac{(N-1)^{j+1}}{C_{M-1}^j}. \tag{4.22}$$

(4.19) then follows immediately from (4.18), (4.21) and (4.22), as desired.

Acknowledgments

This work is partially supported by NSFC with grant numbers 11671145 and 11771286. The second author is also partially supported by Natural Science Foundation of Zhejiang Province (LQ18A010007). The third author is also partially supported by the program of China Scholarships Council (No. 201806145024) and the Natural Science Foundation of Shanghai (16ZR1409700).

References

Aloisi, A.M., Nali, P.F., 2018. Marbles and bottles–or boxes illustrate irreversibility and recurrence. Phys. Educ.(India) 34, 1–18.
Blom, G., 1989. Mean transition times for the Ehrenfest urn model. Adv. Appl. Probab. 21, 479–480.
Chen, Y.-P., Goldstein, I.H., Lathrop, E.D., Nelsen, R.B., 2017. Computing an expected hitting time for the 3–urn Ehrenfest model via electric networks. Stat. Prob. Lett. 127, 42–48.
Doyle, P.G., Snell, E.J., 1984. Random Walks and Electric Networks. In: Carus Math. Monographs, vol. 22, Math. Assoc. Amer., Washington, D.C.
Ehrenfest, P., Ehrenfest, T., 1907. Über zwei bekannte Einwände gegen das Boltzmannsche H–Theorem. Phys. Z. 8, 311–314.
Lathrop, E.D., Goldstein, I.H., Chen, Y.-P., 2016. A note on a generalized Ehrenfest urn model: another look at the mean transition times. J. Appl. Probab. 53, 630–632.
Lawler, G.F., 2006. Introduction to Stochastic Processes, second ed. Chapman and Hall/CRC, New York.
Lyons, R., Peres, Y., 2017. Probability on Trees and Networks. In: Cambridge Series in Statistical and Probabilistic Mathematics, vol. 42, Cambridge University Press, New York.
Meerson, B., Zilber, B., 2018. Large deviations of a long-time average in the Ehrenfest urn model. J. Stat. Mech. Theory Exp. 2018 (5), 053202.
Palacios, J.L., 1994. Another look at the Ehrenfest urn via electric networks. Adv. Appl. Probab. 26, 820–824.
Song, S., Yao, Q., 2019. A new method for computing the expected waiting time between arbitrary different configurations of the multiple–urn Ehrenfest model. Preprint, current version available at https://arxiv.org/pdf/1610.09745.pdf.