Transactions of A. Razmadze
Mathematical Institute
Vol. **174** (2020), issue 2, 149–154

CURIOSITIES REGARDING WAITING TIMES IN PÓLYA'S URN MODEL

NORBERT HENZE[1] AND MARK HOLMES[2]

Abstract. Consider an urn, initially containing b black and w white balls. Select a ball at random and observe its colour. If it is black, stop. Otherwise, return the white ball together with another white ball to the urn. Continue selecting at random, each time adding a white ball, until a black ball is selected. Let $T_{b,w}$ denote the number of draws until this happens. Surprisingly, the expectation of $T_{b,w}$ is infinite for the "fair" initial scenario $b = w = 1$, but finite if $b = 2$ and $w = 10^9$. In fact, $\mathbb{E}[T_{b,w}]$ is finite if and only if $b \geq 2$, and the variance of $T_{b,w}$ is finite if and only if $b \geq 3$, regardless of the number w of white balls. These observations extend to higher moments.

1. Introduction

The classical Pólya–Eggenburger urn is an elegant model in probability theory that is often presented in a first course on martingales (typically in a graduate probability theory course). In its simplest case, the model can be described as follows. Starting with b black and w white balls in an urn, choose a ball uniformly at random from the urn, observe the colour, return the chosen ball to the urn together with another ball of the same colour, then repeat. The number B_n (say) of times a black ball is drawn after n drawings has the well-known Pólya distribution

$$\mathbb{P}(B_n = k) = \binom{n}{k} \frac{\prod_{i=0}^{k-1}(b+i) \prod_{j=0}^{n-k}(w+j)}{\prod_{\ell=0}^{n-1}(b+w+\ell)}, \qquad k = 0, \ldots, n, \tag{1}$$

where an empty product is defined to be one, see, e.g., [4, p. 177]. It is easy to see that the proportion $X_n = (b + B_n)/(b + w + n)$ of black balls at time n is a bounded martingale (with respect to the natural filtration), with $B_0 = b/(b + w)$, and thus X_n converges almost surely to a random variable X. Here, X has a beta $\beta(b, w)$ distribution, see, e.g., [7, Theorem 2.1]. In the special case $b = w = 1$, equation (1) reduces to the discrete uniform distribution $\mathbb{P}(B_n = k) = 1/(n+1)$, and the limit X has a standard uniform distribution.

For later purposes, it will be convenient to regard the distribution of B_n as a special case of a Beta-binomial distribution, see, e.g., [5, p. 242]. The latter distribution originates as follows: Let P have a Beta $\beta(u, v)$-distribution, where $u, v > 0$. Suppose that, conditionally on $P = p$, the random variable M has a binomial distribution $\mathrm{Bin}(n, p)$. Then, for $k \in \{0, 1, \ldots, n\}$, we have

$$\mathbb{P}(M = k) = \int_0^1 \binom{n}{k} p^k (1-p)^{n-k} \cdot \frac{1}{\mathrm{B}(u, v)} p^{u-1} (1-p)^{v-1} \, \mathrm{d}p \tag{2}$$

$$= \binom{n}{k} \frac{\mathrm{B}(u + k, v + n - k)}{\mathrm{B}(u, v)}, \tag{3}$$

where $\mathrm{B}(\cdot, \cdot)$ is the Beta function. The distribution of M is called the Beta-binomial distribution with parameters n, u and v. By using the relation $\mathrm{B}(u, v) = \Gamma(u)\Gamma(v)/\Gamma(u + v)$, where $\Gamma(\cdot)$ is the Gamma function, we see that the distribution of B_n is obtained from (3) by putting $u = b$ and $w = v$.

Inverse Pólya distributions originate if one asks for the number of drawings needed to observe a specified number of black balls under the above or more general replacement schedules, see, e.g., [4, p. 192]. Paper [3] considers waiting times for the first occurrence of a specified pattern in Pólya's urn scheme. A special case is the waiting time until the first occurrence of a black ball, which we will

2020 *Mathematics Subject Classification.* Primary 60E99, Secondary 00A08.
Key words and phrases. Pólya's urn model; Waiting time; Inverse Pólya distribution.

focus on in this note. For the recent work on the inverse Pólya distributions, see, e.g., [1,2], and [6]. In what follows, we consider some curiosities concerning the (random) time until we first draw a black ball, denoted by $T_{w,b}$, that evidently have not been highlighted before.

2. One Black Ball

We first consider the standard "fair" case where the urn contains one black and one white ball at the outset. We then have

$$\mathbb{P}(T_{1,1} > n) = \frac{1}{2} \cdot \frac{2}{3} \cdot \dots \cdot \frac{n-1}{n} \cdot \frac{n}{n+1} = \frac{1}{n+1}$$

and thus $\mathbb{P}(T_{1,1} < \infty) = 1$. Hence, the black ball will be drawn with probability one in finite time. However, since $\sum_{n=0}^{\infty} \mathbb{P}(T_{1,1} > n) = \infty$, *the expectation of $T_{1,1}$ is infinite.*

In view of $\mathbb{P}(T_{1,1} = j) = \mathbb{P}(T_{1,1} > j-1) - \mathbb{P}(T_{1,1} > j) = 1/(j(j+1))$, notice that the conditional expectation of $T_{1,1}$, given $T_{1,1} \le k$, is

$$\mathbb{E}[T_{1,1} | T_{1,1} \le k] = \frac{1}{\mathbb{P}(T_{1,1} \le k)} \sum_{j=1}^{k} j\, \mathbb{P}(T_{1,1} = j) = \frac{(k+1)}{k} \sum_{j=1}^{k} \frac{1}{j+1}.$$

Using $\sum_{j=1}^{n} \frac{1}{j} = \log n + \gamma + o(1)$, where $\gamma = 0.57721\ldots$ is the Euler–Mascheroni constant, it follows that

$$\mathbb{E}[T_{1,1} | T_{1,1} \le k] = \log k + \gamma - 1 + o(1), \qquad \text{as } k \to \infty.$$

In other words, given that you have selected a black ball by time k, on average you first picked one at a relatively early time of $\log(k)$. This is intuitively reasonable because it is much easier to choose a black ball for the first time at an early time, before white balls have been reinforced too much. Indeed, for large k, we find that $\mathbb{P}(T_{1,1} > k/2 | T_{1,1} \le k)$ is of order $1/k$.

We incidentally note that the probability that $T_{1,1}$ takes an odd value equals $\log 2$, since

$$\begin{aligned}\sum_{\ell=0}^{\infty} \mathbb{P}(T_{1,1} = 2\ell+1) &= \sum_{\ell=0}^{\infty} \frac{1}{(2\ell+1)(2\ell+2)} = \sum_{\ell=0}^{\infty} \left(\frac{1}{2\ell+1} - \frac{1}{2\ell+2} \right) \\ &= \sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{j}.\end{aligned}$$

Continue to set $b = 1$, but now allow w to be arbitrarily large. Since

$$\mathbb{P}(T_{1,w} > n) = \frac{w}{w+1} \cdot \frac{w+1}{w+2} \cdot \dots \cdot \frac{w+n-2}{w+n-1} \cdot \frac{w+n-1}{w+n} = \frac{w}{w+n},$$

it follows that $\mathbb{P}(T_{1,w} < \infty) = 1$, regardless of the number of white balls. If, for example, $w = 10^9$, drawing the only black ball seems to be like finding a needle in a haystack, but you have time beyond all limits, and the situation of having one black and 10^9 white balls in the urn could have happended in the course of the stochastic process involving over time under the initial scenario $b = w = 1$ after $10^9 - 1$ draws.

3. A Second Black Ball Works Wonders

Suppose now that at the beginning there are $b = 2$ black and w white balls in the urn. We now have

$$\mathbb{P}(T_{2,w} > n) = \frac{w}{w+2} \cdot \frac{w+1}{w+3} \cdot \frac{w+2}{w+4} \cdot \dots \cdot \frac{w+n-1}{w+n+1} = \frac{w(w+1)}{(w+n)(w+n+1)}.$$

Since $\sum_{n=1}^{\infty} \mathbb{P}(T_{2,w} > n) < \infty$, we do not only have $\mathbb{P}(T_{2,w} < \infty) = 1$, but, in addition, *the expectation of $T_{2,w}$ is finite, irrespective of the number of white balls.* More specifically, we have

$$\mathbb{E}[T_{2,w}] = \sum_{k=0}^{\infty} \mathbb{P}(T_{2,w} > k) = w(w+1) \sum_{k=0}^{\infty} \frac{1}{(w+k)(w+k+1)} = w+1.$$

Here, the last equality follows because the series is telescoping.

Remark 3.1. Starting from $b = 1, w = 1$, we may continue observing Pólya's urn after $T_{1,1}$ until the time $T_{1,1}^{(2)}$ at which we draw a second black ball. At the time $T_{1,1}$ that we first draw a black ball, we return it and add another, so there are then 2 black balls and $T_{1,1}$ white balls. Since $\mathbb{E}[T_{1,1}^{(2)} - T_{1,1} \,|\, T_{1,1} = w] = \mathbb{E}[T_{2,w}] = w+1$, we know that this expectation is finite for every w. We can interpret this as $\mathbb{E}[T_{1,1}^{(2)} - T_{1,1} \,|\, T_{1,1}] = T_{1,1}+1$, or "given the value of $T_{1,1}$, the expected additional time required to draw a second black ball is finite" (a.s.). Nevertheless, $\mathbb{E}[T_{1,1}^{(2)} - T_{1,1}] = \mathbb{E}[T_{1,1} + 1] = \infty$.

4. The General Case

We now assume that the initial configuration is b black and w white balls. The event that each of the first n draws yields a white ball has probability

$$\begin{aligned}\mathbb{P}(T_{b,w} > n) &= \prod_{i=0}^{n-1} \frac{w+i}{b+w+i} \\ &= \frac{(b+w-1)!}{(w-1)!} \cdot \frac{(w-1+n)!}{(b+w-1+n)!}, \qquad n \geq 1.\end{aligned}$$

The first ratio does not depend on n, and the second is equal to

$$\frac{1}{(w+n) \cdot \cdots \cdot (b+w-1+n)}. \tag{4}$$

It immediately follows that $\mathbb{P}(T_{b,w} < \infty) = 1$, but we can infer more from (4). To this end, notice that this expression is bounded from below by $(b+w+n)^{-b}$ and from above by n^{-b}, which, for each integer r, shows that

$$\begin{aligned}\mathbb{E}\big[T_{b,w}^r\big] &= \sum_{n=1}^{\infty} n^r \mathbb{P}(T_{b,w} = n) \\ &= \sum_{n=1}^{\infty} n^r \frac{(b+w-1)!}{(w-1)!} \frac{(w+n-2)!}{(b+w+n-2)!} \frac{b}{b+w+n-1} \\ &= \sum_{n=1}^{\infty} n^r O(n^{-(b+1)}).\end{aligned}$$

Hence, $\mathbb{E}\big[T_{b,w}^r\big] < \infty$ if and only if $b > r$. Surprisingly, this moment condition does not depend on the number w of white balls. In particular, the variance of $T_{b,w}$ exists if and only if there are at least 3 black balls in the urn at the beginning. In the case $b = 3$, straightforward calculations involving telescoping series yield $\mathbb{E}[T_{3,w}] = (w+2)/2$, and, using the fact that $\mathbb{E}[L^2] = \sum_{n=0}^{\infty}(2n+1)\mathbb{P}(L > n)$ for a nonnegative integer-valued random variable L, we have $\mathbb{E}[T_{3,w}^2] = (w+2)(2w+1)/2$, and thus the variance is $\mathbb{V}(T_{3,w}) = 3w(w+2)/4$.

Remark 4.1. In [8], one finds the general formula

$$\mathbb{E}[T_{b,w}] = \frac{b+w-1}{b-1} \tag{5}$$

if $b \geq 2$, which was obtained from a hypergeometric series. As remarked in [9], (5) follows readily from (2), since, conditionally on $P = p$, drawings are according to an independent and identically distributed Bernoulli sequence with probability of success given by p, where success means drawing a black ball. Since, conditionally on $P = p$, the distribution of $T_{b,w}$ is geometric, we have $\mathbb{E}[T_{b,w}|P = p] = 1/p$ and thus

$$\begin{aligned}\mathbb{E}[T_{b,w}] &= \int_0^1 \mathbb{E}[T_{b,w}|P = p] \frac{1}{\mathrm{B}(b,w)} p^{b-1}(1-p)^{w-1}\, \mathrm{d}p = \frac{\mathrm{B}(b-1,w)}{\mathrm{B}(b,w)} \\ &= \frac{b+w-1}{b-1}.\end{aligned}$$

From (2) and the fact that $\mathbb{V}(T_{b,w}) = \mathbb{E}[\mathbb{V}(T_{b,w}|P)] + \mathbb{V}(\mathbb{E}[T_{b,w}|P])$, we can also obtain a general formula for the variance of $T_{b,w}$ if $b \geq 3$. Since the conditional variance of $T_{b,w}$, given $P = p$, is the variance of a geometric distribution with parameter p and thus equal to $(1-p)/p^2$, a straightforward algebra gives

$$\mathbb{E}[\mathbb{V}(T_{b,w}|P)] = \int_0^1 \frac{1-p}{p^2} \frac{1}{\mathrm{B}(b,w)} p^{b-1}(1-p)^{w-1}\,\mathrm{d}p = \frac{w(b+w-1)}{(b-1)(b-2)}.$$

Furthermore, $\mathbb{E}[T_{b,w}|P] = 1/P$, and thus some algebra yields

$$\mathbb{V}(\mathbb{E}[T_{b,w}|P]) = \frac{w(b+w-1)}{(b-1)^2(b-2)}.$$

Summing up, we obtain

$$\mathbb{V}(T_{b,w}) = \frac{bw(b+w-1)}{(b-1)^2(b-2)}.$$

Notice that, in view of $\mathbb{E}[T_{b,w}^\ell] = \mathbb{E}\big[\mathbb{E}[T_{b,w}^\ell|P]\big]$, one can fairly easily even obtain closed-form expressions for higher moments of $T_{b,w}$.

5. A General Replacement Scheme

Suppose now that if a white ball shows up at time k, we return this ball and additionally a_k white balls, where $a_k \geq 1$. Notice that this flexible model includes the special case $a_k = 1$ that has been considered so far, but also the case that a constant number larger than one of white balls is returned to the urn together with the chosen ball. The following result gives a necessary and sufficient condition on the sequence (a_k) for the probability that a black ball shows up at a finite time.

Lemma 5.1. *Let $s_k = a_1 + \cdots + a_k$, $k \geq 1$. We then have*

$$\mathbb{P}(T_{b,w} < \infty) = 1 \Longleftrightarrow \sum_{j=1}^{\infty} \frac{1}{s_j} = \infty.$$

Proof. Putting $s_0 = 0$, we have

$$\mathbb{P}(T_{b,w} > n) = \prod_{j=0}^{n-1} \frac{w+s_j}{b+w+s_j}.$$

Using the inequalities $1 - 1/t \leq \log t \leq t-1$, $t > 0$, straightforward calculations yield

$$-b\sum_{j=0}^{n-1} \frac{1}{w+s_j} \leq \log \mathbb{P}(T_{b,w} > n) \leq -b\sum_{j=0}^{n-1} \frac{1}{b+w+s_j}.$$

Hence $\log \mathbb{P}(T_{b,w} > n) \to -\infty$ as $n \to \infty$ if and only if the series $\sum_{j=0}^{\infty} 1/s_j$ diverges, and the assertion follows. □

From this result it follows that $\mathbb{P}(T_{b,w} < \infty) = 1$ even if $b = 1$, w is arbitrarily large, and a fixed huge number of additional white balls is added to the urn after each draw of a white ball, but not if at the kth time we select a white ball we return it and add k extra white balls, for example.

In the case where we add a constant c additional number of white balls to the urn whenever we select a white ball, we can also consider the expected time to select a black ball.

Lemma 5.2. *In the case where we start with w white balls and b black balls in the urn, and add $c \geq 1$ additional white balls whenever white is selected from the urn, we find that $\mathbb{E}[T_{b,w}] < \infty$ if and only if $b > c$.*

Proof. In this context we can write

$$\begin{aligned}\mathbb{P}(T_{b,w} > n) &= \prod_{j=0}^{n-1} \frac{\frac{w}{c}+j}{\frac{b}{c}+\frac{w}{c}+j} \\ &= \frac{\frac{w}{c}}{\frac{b}{c}+\frac{w}{c}} \times \frac{\frac{w}{c}+1}{\frac{b}{c}+\frac{w}{c}+1} \times \cdots \times \frac{\frac{w}{c}+n-2}{\frac{b}{c}+\frac{w}{c}+n-2} \times \frac{\frac{w}{c}+n-1}{\frac{b}{c}+\frac{w}{c}+n-1}.\end{aligned}$$

If $b/c \leq 1$, then the numerator of the $j+1$st term in the product is greater than or equal to the denominator of the jth term and so this product is at least

$$\frac{\frac{w}{c}}{\frac{w}{c}+\frac{b}{c}+n-1},$$

which is not summable in n, so the expectation of $T_{b,n}$ is infinite.

If $b/c \geq 2$, then the numerator of the $j+2$nd term in the product is no larger than the denominator of the jth term, so for some constant a we have $\mathbb{P}(T_{b,w} > n) \leq an^{-2}$ for all n sufficiently large. This is summable in n, so the expectation is finite when $b/c \geq 2$.

The case $b/c \in (1,2)$ can be handled by a slightly more elaborate (but standard) approach, which we now quickly present. We can write

$$\mathbb{P}(T_{b,w} > n) = \prod_{j=0}^{n-1} \left(1 - \frac{\frac{b}{c}}{\frac{b}{c}+\frac{w}{c}+j}\right) \leq \exp\left\{ -\frac{b}{c} \sum_{j=0}^{n-1} \frac{1}{\frac{b}{c}+\frac{w}{c}+j} \right\},$$

where we have used $1-x \leq e^{-x}$ and that the product of exponentials is the exponential of a sum. For $n \geq 1$, the sum is at least $\int_0^{n-1} \frac{1}{d+x} \mathrm{d}x = \log(d+n-1) - \log(d)$, where $d = (b+w)/c > 0$. Thus for $n \geq 1$,

$$\mathbb{P}(T_{b,w} > n) \leq d^{b/c} \exp\left\{ \log\left((n-1+d)^{-b/c} \right) \right\} = \frac{d^{b/c}}{(n-1+d)^{b/c}}.$$

Since $b/c > 1$, this is summable in n. $\square$

Acknowledgement

The work of MH is supported by Future Fellowship FT160100166 from the Australian Research Council. The authors thank a referee whose comments helped to improve the paper. The authors thank the organisers of the International Conference on Probability Theory and Statistics in Tbilisi, 2019, for bringing them together. Finally, the authors thank and congratulate Estate Khmaladze for many years of outstanding contributions to statistics and probability.

References

1. Ch. A. Charalambides, A q-Pólya urn model and the q-Pólya and inverse q-Pólya distributions. *J. Statist. Plann. Inference* **142** (2012), no. 1, 276–288.
2. T. Garg, P. N. Agrawal, A. Kajla, Jain-Durrmeyer operators involving inverse Pólya-Eggenberger distribution. *Proc. Nat. Acad. Sci. India Sect. A* **89** (2019), no. 3, 547–557.
3. K. Inoue, S. Aki, Generalized waiting time problems associated with pattern in Pólya's urn scheme. *Ann. Inst. Statist. Math.* **54** (2002), no. 3, 681–688.
4. N. L. Johnson, S. Kotz, *Urn Models and their Application.* An approach to modern discrete probability theory. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, New York-London-Sydney, 1977.
5. N. L. Johnson, S. Kotz, A. Kemp, *Univariate Discrete Distributions.* Second edition. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1992.
6. F. Makri, A. N. Philippou, Z. M. Psillakis, Pólya, inverse Polya, and circular Polya distributions of order k for l-overlapping success runs. *Comm. Statist. Theory Methods* **36** (2007), no. 1-4, 657–668.
7. R. Pemantle, A survey of random processes with reinforcement. *Probab. Surv.* **4** (2007), 1–79.
8. https://math.stackexchange.com/questions/2560846/polya-urn-mean-number-of-draws-to-get-a-specific-sequence-of-colors

9. https://mathoverflow.net/questions/289247/polya-urn-mean-number-of-draws-to-get-a-specific-sequence-of-colors

(Received 28.02.2020)

[1]Institute of Stochastics, Karlsruhe Institute of Technology (KIT), Englerstr. 2, D-76133 Karlsruhe
E-mail address: Norbert.Henze@kit.edu

[2]University of Melbourne, School of Mathematics and Statistics, Peter Hall Building, Vic 3010, Australia
E-mail address: holmes.m@unimelb.edu.au