UNIVERZA NA PRIMORSKEM
FAKULTETA ZA MATEMATIKO, NARAVOSLOVJE IN INFORMACIJSKE TEHNOLOGIJE

Zaključna naloga
(Final project paper)

Pólya–Eggenbergerjev model

(Pólya–Eggenberger model)

Ime in priimek: Elvir Merulić
Študijski program: Matematika
Mentor: doc. dr. Martin Raič

Koper, oktober 2018

Ključna dokumentacijska informacija

Ime in PRIIMEK: Elvir MERULIĆ

Naslov zaključne naloge: Pólya–Eggenbergerjev model

Kraj: Koper

Leto: 2018

Število listov: 39 Število slik: 8
Število prilog: 2 Število strani prilog: 5 Število referenc: 7

Mentor: doc. dr. Martin Raič

Ključne besede: verjetnost, Pólyeva žara, porazdelitev beta, šibka konvergenca

Math. Subj. Class. (2010): 60C05, 60F05

Izvleček:
Predstavimo Pólya–Eggenbergerjev model, v katerem je v žari sprva določeno število belih in modrih kroglic. Vsakič, ko izvlečemo eno kroglico, dodamo določeno število kroglic iste barve. V delu preučimo porazdelitev števila izvlečenih belih kroglic. Izpeljemo, da le-ta konvergira proti porazdelitvi beta.

Key words documentation

Name and SURNAME: Elvir MERULIĆ

Title of final project paper: Pólya–Eggenberger model

Place: Koper

Year: 2018

Number of pages: 39 Number of figures: 8
Number of appendices: 2 Number of appendix pages: 5 Number of references:7

Mentor: Assist. Prof. Martin Raič, PhD

Keywords: probability, Pólya urn, beta distribution, weak convergence

Math. Subj. Class. (2010): 60C05, 60F05

Abstract:
We will present the Pólya–Eggenberger model, in which we have a particular number of white and blue balls in the urn. Everytime, after we draw one ball, we add a fixed number of balls of the same color. In the paper we study the distribution of the number of white balls drawn. We obtain that it converges to the beta distribution.

Acknowledgement

I would like to thank my mentor, Professor Martin Raič, for a truly remarkable dedication to teaching students. He is very good at teaching and presenting the more complicated problems in easy to understand terms. He has a lot of patience and his lectures include a lot of examples. Writing with him is truly a joy.

Contents

List of Figures

1 Short introduction of Urn and balls models

The idea of presenting mathematical models via urns and balls is very old. The book of Johnson and Kotz [3] presents a historical perspective. The urns are traced to the times of the Old Testament. Plenty of noted mathematicians of the post-Renaissance era mention urns, such as Huygens, de Moivre, Laplace and Bernoulli. Some most common applications are listed below (for a more detailed description see Mahmoud [4]).
Models of urns can be used to present the Ballot problems that are concerned with the progress of a vote. Suppose we have two candidates running for A and B, with votes won represented by m and n, respectively. We would like to know the probabillity that A wins. We can view the progress of votes as white and blue balls placed in an urn. It was first used by W. Whitworth in 1817.
The Occupancy problem was first studied by de Moivre in 1713. There are n balls that will be thrown into k urns. Each ball is independently deposited in a uniformly chosen urn. One of the questions of interest is what is the probability that no urn will be empty.
The Gambler's ruin starts with two gamblers A and B who keep on betting until one of them goes bankrupt. A starts with m dollars, and B with n. They bet repeatedly on the outcome of a game of flipping a coin. A bets on heads, and B bets on tails. After the flip, the loser in that individual game gives a dollar to the winner. What is the probabillity that A wins? The problem can be presented via two urns, A and B, that start out with m and n balls, respectively. The problem is associated with a number of famous scientists and mathematicians including a few Bernoullis, de Moivre, Feller, Hermat, Huygens, Laplace, Lagrange, Pascal.
The Banach's Matchbox problem deals with a pipe smoker that has two matchboxes. He puts one box in each of the two pockets of his pants. Each complete new pack has n matches in it. The smoker reaches into a randomly selected pocket and takes a match from the box, whenever he needs a match. The problem can be presented by two urns with n balls in each. The problem is often attributed by mathematician Stefan Banach (1892–1945), who was a heavy smoker, but it may not be Banach who set the problem or provided an answer.

Urn Schemes can be applied to bioscience as well. Several models are used for population genetics in evolution of species, such as Wright–Fisher Allelic Urn models, Gene Miscopying, Mutations, Hoppe's Urn scheme, Ewens' Sampling Formula, Phylogenetic Trees. Some present the competitive exclusion. Others are used in Epidemiology, such as Kriz' Polytonomous Urn Scheme. Models also exist for the discriminatory monitoring of alternative medical treatments, such as Play-the-Winner Schemes.

2 The Pólya–Eggenberger Urn

Pólya urns are a model of urns that involve only one urn and general methods of replacement. The form we know today and its numerous derivatives seem to have first appeared in work of Markov around 1905–1907. A Pólya urn is an urn containing balls of up to k different colors. We observe the changes of the composition of the urn as it goes through discrete time steps. At each step we shake the urn well, and we sample uniformly one ball at random (with all balls being equally likely). We notice the color of the ball, and the ball is returned to the urn. We then add new balls to the urn in various numbers and of various colors, depending on color of the ball and on the chosen model. Primary interest is the long-term composition of the urn and in the stochastic path leading to it. The examples of the important parameters are the number of balls of each color and the number of times a ball of a particular color is drawn.

Pólya Urns can be applied in informatics in search trees such as Binary Search Trees, Fringe-Balanced Trees, m-ary Search Trees, 2-3 Trees, Paged Binary Trees, Bucket Quad Trees, Bucket k-d Trees. It can be used in The Recursive trees model such as Standard Recursive Trees, Pyramids, Plane-Oriented Recursive Trees, Bucket Recursive Trees, Sprouts, to name the most common.

In the present work, we focus on the basic model of Pólya and Eggenberger [2] which is one of the very first studies that focuses on balls of two different colors being present in the urn at the start. However it is reported that the model had been considered in Markov [5] and Tchuprov [7]. The model was introduced as a model for contagion. This is a fixed model, where we notice the color of the ball withdrawn and we add s balls of the same color to the urn (s is a positive integer). In this case, the probability distribution of the composition of the urn after a fixed number of draws can be expressed by a closed formula.

Let W_0 denote the number of white balls we have in the urn at the start, and B_0 the number of blue balls we have in the urn at the start. The total number of balls at the start is equal to the sum of the number of white balls at the start and the number of blue balls at the start, and thus $T_0 = W_0 + B_0$.

Let W_n and B_n denote the number of white and blue balls, respectively, after n draws. Let $\tilde{W}_n$ denote the number of white balls drawn from the urn after n draws, and similarly let $\tilde{B}_n$ denote the number of blue balls drawn from the urn after n draws. Since

we draw a total of n times, $\tilde{W}_n + \tilde{B}_n = n$. It must hold that the number of white balls in the urn after n draws is equal to the sum of the white balls we had at the start and the sum of the white balls added. Since we have drawn the white ball $\tilde{W}_n$ times, we have $W_n = W_0 + s\tilde{W}_n$. Similarly, we obtain $B_n = B_0 + s\tilde{B}_n$.
Let us denote by T_n the total number of balls in the urn after n draws. This is equal to the sum of the number of white balls after n draws and the number of blue balls after n draws. Thus, $T_n = W_n + B_n = W_0 + s\tilde{W}_n + B_0 + s\tilde{B}_n = W_0 + B_0 + s\tilde{W}_n + s\tilde{B}_n = W_0 + B_0 + sn$. The following assertion gives a closed formula for the distribution of $\tilde{W}_n$.

Theorem 2.1 (Eggenberger and Pólya, 1923)**.** *Let $\tilde{W}_n$ be the number of white balls drawn in the Pólya-Eggenberger urn after n draws. Then,*

$$\mathbb{P}(\tilde{W}_n = k) = \frac{W_0(W_0+s)\cdots(W_0+(k-1)s)\, B_0(B_0+s)\cdots(B_0+(n-k-1)s)}{T_0(T_0+s)\cdots(T_0+(n-1)s)}\binom{n}{k}.$$

Proof. We will use the notation introduced before. Following Mahmoud [4], suppose $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ are the time indices of the white ball draws. Thus, before time i_1, all drawn balls are blue. The probability that we will draw a blue ball in the particular draw is equal to the number of blue balls at that particular time divided by the number of total balls at that particular time.
In particular, the probability is equal to $\frac{B_0}{T_0}$. As we have drawn a blue ball, we add s new blue balls to the urn. The number of blue balls has now increased by s, giving us a $B_0 + s$ blue balls. The total number of balls has increased by s as well, giving us a total of $T_0 + s$.
Given that the first drawn ball is blue, the conditional probability that we draw a blue ball in the second draw is now, again, equal to the number of blue balls at that particular time divided by the number of total balls at that particular time. This probability equals to

$$\frac{B_0 + s}{T_1},$$

and the probability that we draw a blue ball in the first as well as in the second draw is now equal to the product:

$$\frac{B_0}{T_0} \cdot \frac{B_0 + s}{T_1}.$$

We continue the procedure until we have drawn a white ball for the first time, at time index i_1. We want to know the probability that we have drawn the blue ball at time index $i_1 - 1$. The number of draws before the $i_1 - 1$ time index is equal to $i_1 - 2$. The number of blue balls at time index $i_1 - 1$ is equal to $B_0 + (i_1 - 2)s$ because we have started with B_0 blue balls and after each draw we have added s new blue balls to the urn, giving us a total of $(i_1 - 2)s$ new blue balls. Given that all previous drawn balls are blue, the conditional probability that we have drawn a blue ball at time index $i_1 - 1$ is thus equal to

$$\frac{B_0 + (i_1 - 2)s}{T_{i_1-2}}.$$

We are now at time index i_1. The ball drawn must now be of white color. At that time index we have just as many white balls as we had at the start, W_0, because we were only drawing the blue balls in previous draws. The suitable conditional probability of this event is thus

$$\frac{W_0}{T_{i_1-1}}.$$

Thus, the probability that the first white ball is drawn at the time index i_1 equals

$$\frac{B_0}{T_0} \cdot \frac{B_0 + s}{T_1} \cdot \frac{B_0 + 2s}{T_2} \cdots \frac{B_0 + (i_1 - 2)s}{T_{i_1-2}} \cdot \frac{W_0}{T_{i_1-1}}$$

We continue until we draw the white ball again. In between the first and second white balls drawn, maybe some blue balls will be drawn. We are now at the time index $i_1 + 1$. It means that we have drawn i times already, we drew the white ball only once, and we drew the blue ball $i - 1$ times. The number of white balls is now $W_0 + s$. The number of blue balls has increased by s since the last time we drew the blue ball, and is thus equal to $B_0 + (i_1 - 2)s + s = B_0 + (i_1 - 1)s$. The probability of interest is

$$\frac{B_0 + (i_1 - 1)s}{T_{i_1}}.$$

We arrive to the time index i_2 when we draw the white ball the second time. We want to know what happened in the draw before, at time index $i_2 - 1$. Up to that point we drew $i_2 - 2$ times, out of which we only drew the white ball once. We drew the blue ball $i_2 - 3$ times. So the total amount of blue balls at time index $i_2 - 1$ is equal to $B_0 + (i_2 - 3)s$. The suitable conditional probability of drawing a blue ball at time index $i_2 - 1$ is then

$$\frac{B_0 + (i_2 - 3)s}{T_{i_2-2}}.$$

At time index i_2 we draw the white ball the second time. But, out of all the previous draws, we have drawn the white ball only once, at time index i_1. The total number of white balls before the (i_2)-th draw is then equal to $W_0 + s$, the total number of balls is T_{i_2-1}. The probability of drawing the white ball at time index i_2 is thus

$$\frac{W_0 + s}{T_{i_2-1}}.$$

We continue the procedure until we draw k white balls. That will happen at the time index i_k. Up to that point we have drawn $n-1$ times out of which we have drawn the white ball k times. The rest of the times we drew the blue ball. Thus, the number of blue balls after $n-1$ draws is $B_0 + (n-1-k)s$. The probability that we draw the blue ball at time index $i_n - 1$ is thus

$$\frac{B_0 + (n-k-1)s}{T_{n-1}}.$$

Therefore, the probability of drawing white balls exactly at time indices $i_1, i_2, \cdots i_k$ is then

$$\begin{aligned} &\frac{B_0}{T_0} \cdot \frac{B_0+s}{T_1} \cdot \frac{B_0+2s}{T_2} \cdots \frac{B_0+(i_1-2)s}{T_{i_1-2}} \cdot \frac{W_0}{T_{i_1-1}} \\ &\cdot \frac{B_0+(i_1-1)s}{T_{i_1}} \cdots \frac{B_0+(i_2-3)s}{T_{i_2-2}} \cdot \frac{W_0+s}{T_{i_2-1}} \\ &\cdots \\ &\cdot \frac{W_0+(k-1)s}{T_{i_k}-1} \cdot \frac{B_0+(i_k-k)s}{T_{i_k}} \cdots \frac{B_0+(n-k-1)s}{T_{n-1}}. \end{aligned}$$

Now observe that this expression does not depend on the indices, which can be chosen in exactly $\binom{n}{k}$ ways. This completes the proof. □

Remark 2.2. This standard proof bears an idea similar to the derivation of the binomial law on n independent trials. The difference is that in the binomial case the trials are identical, but here the probabilities change in time, because we add new balls after each draw. However, like in the binomial case, the probability of a given sequence of colors depends only of the number of white balls to be drawn, but not of the order of the colors.

The probabilities in Theorem 3.1 can be written a bit differently:

$$\begin{aligned} \mathbb{P}\big(\tilde{W}_n = k\big) &= \frac{s\left(\frac{W_0}{s}\right) s\left(\frac{W_0}{s}+1\right) \dots s\left(\frac{W_0}{s}+k-1\right) s\left(\frac{B_0}{s}\right) s\left(\frac{B_0}{s}+1\right) \dots s\left(\frac{B_0}{s}+(n-k-1)\right)}{s\left(\frac{T_0}{s}\right) s\left(\frac{T_0}{s}+1\right) \dots s\left(\frac{T_0}{s}+(n-1)\right)} \binom{n}{k} \\ &= \frac{s^k\left(\frac{W_0}{s}\right)\left(\frac{W_0}{s}+1\right) \dots \left(\frac{W_0}{s}+(k-1)\right) s^{n-k}\left(\frac{B_0}{s}\right)\left(\frac{B_0}{s}+1\right) \dots \left(\frac{B_0}{s}+(n-k-1)\right)}{s^n\left(\frac{T_0}{s}\right)\left(\frac{T_0}{s}+1\right) \dots \left(\frac{T_0}{s}+(n-1)\right)} \binom{n}{k} \\ &= \frac{\left(\frac{W_0}{s}\right)\left(\frac{W_0}{s}+1\right) \dots \left(\frac{W_0}{s}+(k-1)\right)\left(\frac{B_0}{s}\right)\left(\frac{B_0}{s}+1\right) \dots \left(\frac{B_0}{s}+(k-1)\right)}{\left(\frac{T_0}{s}\right)\left(\frac{T_0}{s}+1\right) \dots \left(\frac{T_0}{s}+(n-1)\right)} \binom{n}{k}. \end{aligned}$$

Introducing the rising factorial or the Pochhammer symbol $(x_0)_k$, as

$$(x_0)_k = x_0(x_0+1)(x_0+2)(x_0+3)\dots(x_0+k-1)$$

the probabillity can be rewritten as:

$$\mathbb{P}(\tilde{W}_n = k) = \binom{n}{k} \frac{\left(\frac{W_0}{s}\right)_k \left(\frac{B_0}{s}\right)_{n-k}}{\left(\frac{T_0}{s}\right)_n}.$$

Proposition 2.3 (Eggenberger and Pólya, 1923)**.** *Let $\tilde{W}_n$ be the number of white balls drawn in the Pólya-Eggenberger urn after n draws. Then,*

$$\mathbb{E}(\tilde{W}_n) = \frac{W_0}{T_0}\, n \tag{2.1}$$

$$\mathrm{Var}[\tilde{W}_n] = \frac{W_0\, B_0\, n\,(sn+T_0)}{T_0^2(T_0+s)}. \tag{2.2}$$

Proof. We will start with the classical definition of expected value:

$$\mathbb{E}[\tilde{W}_n] = \sum_{k=0}^{n} k\ \mathbb{P}(\tilde{W}_n = k) = \sum_{k=0}^{n} k \binom{n}{k} \frac{\left(\frac{W_0}{s}\right)_k \left(\frac{B_0}{s}\right)_{n-k}}{\left(\frac{T_0}{s}\right)_n}.$$

We can write

$$k\binom{n}{k} = k\,\frac{n!}{k!(n-k)!} = \frac{n!}{(k-1)!(n-k)!} = \frac{n!}{(k-1)!((n-1)-(k-1))!}$$

$$= n\,\frac{(n-1)!}{(k-1)!((n-1)-(k-1))} = n\binom{n-1}{k-1},$$

and use it in our main equation

$$\mathbb{E}[\tilde{W}_n] = \sum_{k=1}^{n} n \binom{n-1}{k-1} \frac{\left(\frac{W_0}{s}\right)_k \left(\frac{B_0}{s}\right)_{n-k}}{\left(\frac{T_0}{s}\right)_n}.$$

Recalling that $(W_0)_k$ is the rising factorial equal to

$$W_0(W_0+1)(W_0+2)(W_0+3)\dots(W_0+k-1),$$

and $(W_0)_{k-1}$ is the rising factorial equal to

$$W_0(W_0+1)(W_0+2)(W_0+3)\dots(W_0+k-2).$$

Then, we obtain that

$$\left(\frac{W_0}{s}\right)_k = \frac{W_0}{s}\left(\frac{W_0+s}{s}\right)_{k-1},$$

and similarly

$$\left(\frac{T_0}{s}\right)_k = \frac{T_0}{s}\left(\frac{T_0+s}{s}\right)_{k-1}.$$

We notice that n, $\frac{W_0}{s}, \frac{B_0}{s}$ are constants and can be written in front of the sum. Our equation is now:

$$\begin{aligned}\mathbb{E}[\tilde{W}_n] &= n\,\frac{\frac{W_0}{s}}{\frac{T_0}{s}}\sum_{k=1}^{n}\binom{n-1}{k-1}\frac{\left(\frac{W_0+s}{s}\right)_{k-1}\left(\frac{B_0}{s}\right)_{n-k}}{\left(\frac{T_0+s}{s}\right)_{n-1}}\\ &= n\,\frac{W_0}{T_0}\sum_{k=1}^{n}\binom{n-1}{k-1}\frac{\left(\frac{W_0+s}{s}\right)_{k-1}\left(\frac{B_0}{s}\right)_{n-k}}{\left(\frac{T_0+s}{s}\right)_{n-1}}.\end{aligned}$$

Here, we introduce the substitution $j = k-1$, and we get:

$$\mathbb{E}[\tilde{W}_n] = n\,\frac{W_0}{T_0}\sum_{j=0}^{n-1}\binom{n-1}{j}\frac{\left(\frac{W_0+s}{s}\right)_{j}\left(\frac{B_0}{s}\right)_{(n-1)-j}}{\left(\frac{T_0+s}{s}\right)_{n-1}}$$

Now observe that the summands in the right hand side are exactly the point probabilities for a Pólya-Eggenberger urn with W_0+s white and B_0 blue balls after $n-1$ draws, so that their sum equals to 1. This proves (2.1).
To prove the second statement we can use the formula

$$\mathrm{Var}(\tilde{W}_n) = \mathbb{E}(\tilde{W}_n^2) - \mathbb{E}(\tilde{W}_n)^2 = \mathbb{E}[\tilde{W}_n(\tilde{W}_n-1)] + \mathbb{E}(\tilde{W}_n) - \mathbb{E}(\tilde{W}_n)^2$$

Similarly as in the derivation of $\mathbb{E}(\tilde{W}_n)$, we observe that

$$\begin{aligned}\mathbb{E}[\tilde{W}_n(\tilde{W}_n-1)] &= \sum_{k=0}^{n} k(k-1)\,P(\tilde{W}_n = k)\\ &= \sum_{k=0}^{n} k(k-1)\binom{n}{k}\frac{\left(\frac{W_0}{s}\right)_k\left(\frac{B_0}{s}\right)_{n-k}}{\left(\frac{T_0}{s}\right)_n}\end{aligned}\tag{2.3}$$

Now write:

$$\begin{aligned}k(k-1)\binom{n}{k} &= k(k-1)\frac{n!}{k!(n-k)!} = \frac{n!}{(k-2)!(n-k)!}\\ &= n(n-1)\frac{(n-2)!}{(k-2)!((n-2)-(k-2))!} = n(n-1)\binom{n-2}{k-2}\end{aligned}$$

Substituting into (2.3), we obtain:

$$\begin{aligned}\mathbb{E}[\tilde{W}_n(\tilde{W}_n-1)] &= \sum_{k=2}^{n} n(n-1)\binom{n-2}{k-2}\frac{\left(\frac{W_0}{s}\right)_k\left(\frac{B_0}{s}\right)_{n-k}}{\left(\frac{T_0}{s}\right)_n}\\ &= n(n-1)\sum_{k=2}^{n}\binom{n-2}{k-2}\frac{\left(\frac{W_0}{s}\right)_k\left(\frac{B_0}{s}\right)_{n-k}}{\left(\frac{T_0}{s}\right)_n}\end{aligned}$$

Introducing a similar substitution as before, $j = k - 2$, we get:

$$\begin{aligned}\mathbb{E}[\tilde{W}_n(\tilde{W}_n - 1)] &= \sum_{j=0}^{n-2} n(n-1)\binom{n-2}{j}\frac{\left(\frac{W_0}{s}\right)_{j+2}\left(\frac{B_0}{s}\right)_{n-j-2}}{\left(\frac{T_0}{s}\right)_n} \\ &= n(n-1)\frac{\frac{W_0}{s}\left(\frac{W_0}{s}+1\right)}{\frac{B_0}{s}\left(\frac{B_0}{s}+1\right)}\sum_{j=0}^{n-2}\binom{n-2}{j}\frac{\left(\frac{W_0}{s}\right)_j\left(\frac{B_0}{s}\right)_{n-j-2}}{\left(\frac{T_0}{s}\right)_{n-2}}.\end{aligned}$$

Similarly as before, observe that the summands in the right hand side are exactly the point probabilities for a Pólya-Eggenberger urn with $W_0 + 2s$ white and B_0 blue balls after $n - 2$ draws, so that their sum equals 1. Thus, we get:

$$\mathbb{E}[\tilde{W}_n(\tilde{W}_n - 1)] = n(n-1)\frac{\frac{W_0}{s}\left(\frac{W_0}{s}+1\right)}{\frac{B_0}{s}\left(\frac{B_0}{s}+1\right)} = n(n-1)\frac{W_0(W_0+s)}{T_0(T_0+s)},$$

and

$$\mathbb{E}[\tilde{W}_n^2] = \mathbb{E}[\tilde{W}_n(\tilde{W}_n - 1)] + \mathbb{E}(\tilde{W}_n) = n(n-1)\frac{W_0(W_0+s)}{T_0(T_0+s)} + n\frac{W_0}{T_0}.$$

Finally,

$$\begin{aligned}\operatorname{Var}(\tilde{W}_n) &= \mathbb{E}[\tilde{W}_n^2] - \mathbb{E}[\tilde{W}_n]^2 \\ &= n\frac{W_0}{T_0}\left((n-1)\frac{(W_0+s)}{(T_0+s)} + 1\right) - \frac{n^2W_0^2}{T_0^2} \\ &= n\frac{W_0}{T_0}\left(\frac{nW_0 + ns - W_0 - s + T_0 + s}{T_0 + s}\right) - \frac{n^2W_0^2}{T_0^2} \\ &= \frac{nW_0T_0\left(nW_0 + ns - W_0 + T_0\right)}{T_0^2(T_0+s)} - \frac{n^2W_0^2(T_0+s)}{T_0^2(T_0+s)} \\ &= \frac{nW_0T_0\left(nW_0 + ns - W_0 + T_0\right)}{T_0^2(T_0+s)} - \frac{n^2W_0^2T_0 + sn^2W_0^2}{T_0^2(T_0+s)} \\ &= \frac{nW_0\left(nW_0T_0 + nsT_0 - W_0T_0 + T_0^2\right) - n^2W_0^2T_0 - sn^2W_0^2}{T_0^2(T_0+s)} \\ &= \frac{nW_0\left(nW_0T_0 + nsT_0 - W_0T_0 + T_0^2 - nW_0T_0 - nW_0s\right)}{T_0^2(T_0+s)} \\ &= \frac{nW_0\left(nsT_0 - W_0T_0 + T_0^2 - nW_0s\right)}{T_0^2(T_0+s)} \\ &= \frac{nW_0\left(sn(T_0 - W_0) + T_0(T_0 - W_0)\right)}{T_0^2(T_0+s)} \\ &= \frac{nW_0(snB_0 + T_0B_0)}{T_0^2(T_0+s)} \\ &= \frac{nW_0B_0(sn + T_0)}{T_0^2(T_0+s)}\end{aligned}$$

completing the proof. □

3 Asymptotic behavior of the Pólya–Eggenberger urn distribution

We are interested to see what happens when the number of draws goes to infinity. We will display the distribution of $\tilde{W}_n$ as n increases to get an idea. Each time, we take $W_0 = 3, B_0 = 5, s = 2$.

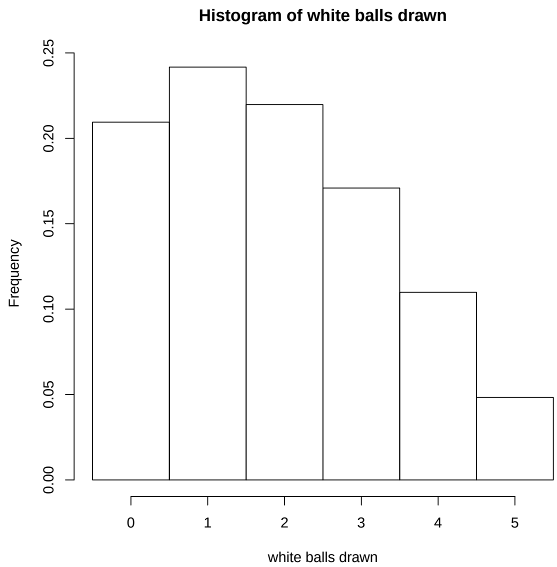

Figure 1: Histogram for $W_0 = 3, B_0 = 5, s = 2, n = 5$

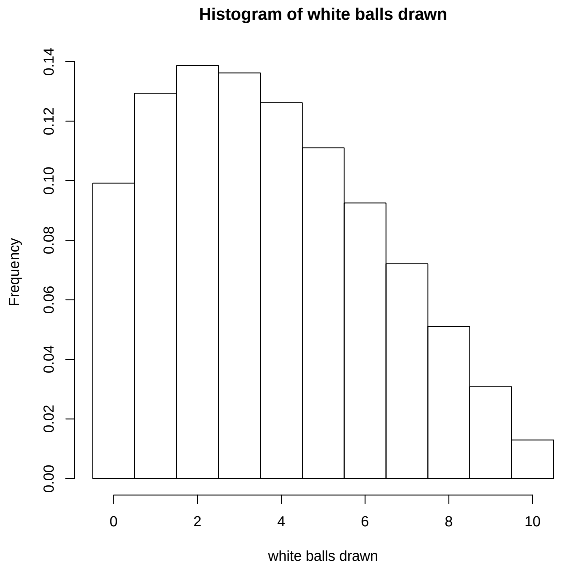

Figure 2: Histogram for $W_0 = 3, B_0 = 5, s = 2, n = 10$

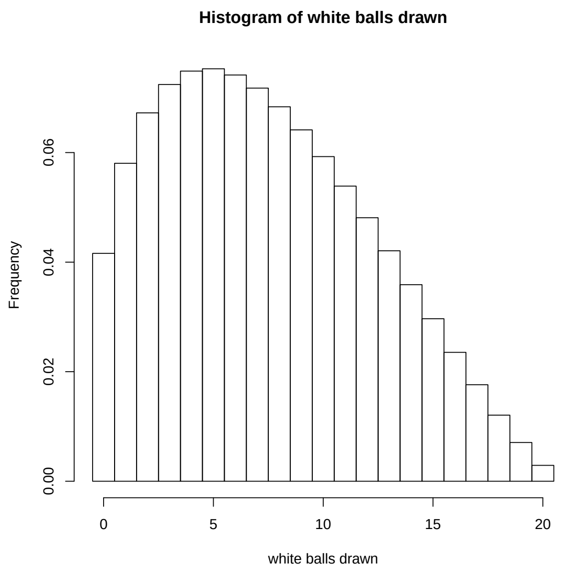

Figure 3: Histogram for $W_0 = 3, B_0 = 5, s = 2, n = 20$

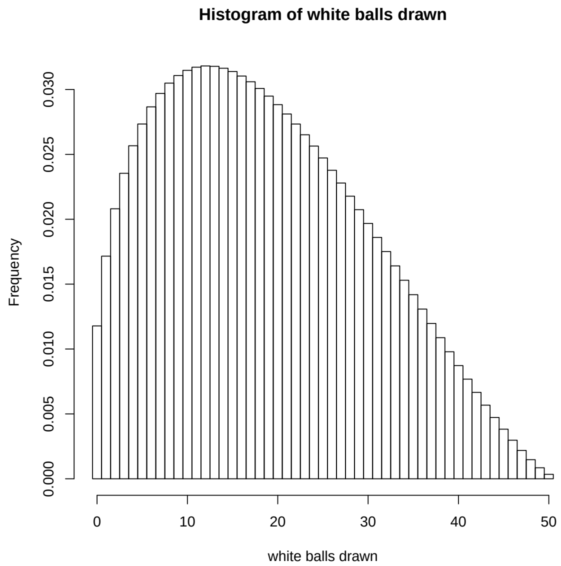

Figure 4: Histogram for $W_0 = 3, B_0 = 5, s = 2, n = 50$

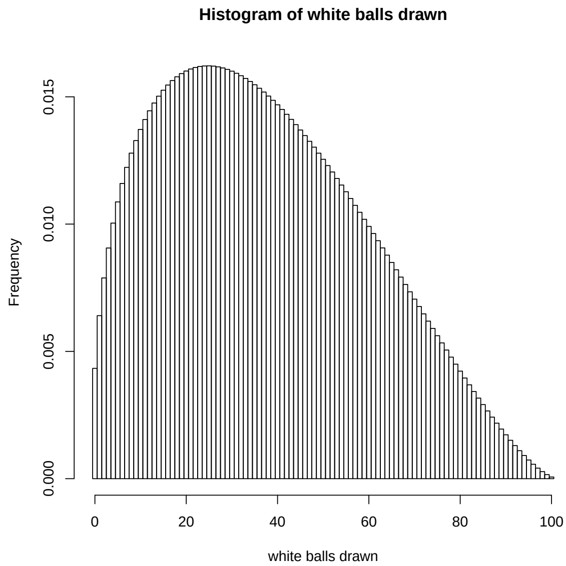

Figure 5: Histogram for $W_0 = 3, B_0 = 5, s = 2, n = 100$

It seems that $Z_n = \frac{\tilde{W}_n}{n}$ converges to a continuous distribution. This is also indicated by the expectation and variance, as, by Proposition 2.3, $\mathbb{E}(Z_n) = \frac{W_0}{T_0}$ and

$$\mathrm{Var}(Z_n) = \frac{W_0 B_0 \left(sn + T_0\right)}{n T_0^2 \left(T_0 + s\right)} \xrightarrow[n\to\infty]{} \frac{W_0 B_0 s}{T_0^2 \left(T_0 + s\right)}\,.$$

To find it, we examine the relative differences:

$$\frac{\mathbb{P}\big(\tilde{W}_n = k+1\big) - \mathbb{P}\big(\tilde{W}_n = k\big)}{\mathbb{P}\big(\tilde{W}_n = k\big)}.$$

Observe that:

$$\begin{aligned}
\mathbb{P}\big(\tilde{W}_n = k+1\big) &= \binom{n}{k+1} \frac{\left(\frac{W_0}{s}\right)_{k+1} \left(\frac{B_0}{s}\right)_{n-k-1}}{\left(\frac{T_0}{s}\right)_n} \\
&= \frac{n!}{(n-k-1)!\,(k+1)!} \frac{\left(\frac{W_0}{s}\right)_k \left(\frac{W_0}{s}+k\right) \left(\frac{B_0}{s}\right)_{n-k}}{\left(\frac{B_0}{s}+n-k-1\right) \left(\frac{T_0}{s}\right)_n} \\
&= \frac{n!\,(n-k)}{(n-k)!\,(k+1)\,k!} \frac{\left(\frac{W_0}{s}\right)_k \left(\frac{W_0}{s}+k\right) \left(\frac{B_0}{s}\right)_{n-k}}{\left(\frac{B_0}{s}+n-k-1\right) \left(\frac{T_0}{s}\right)_n}
\end{aligned}$$

Rearranging the factors, we find that:

$$\begin{aligned}
\mathbb{P}\big(\tilde{W}_n = k+1\big) =& \frac{(n-k)}{(k+1)} \frac{\left(\frac{W_0}{s}+k\right)}{\left(\frac{B_0}{s}+n-k-1\right)} \frac{n!}{k!\,(n-k)!} \frac{\left(\frac{W_0}{s}\right)_k \left(\frac{B_0}{s}\right)_{n-k}}{\left(\frac{T_0}{s}\right)_n} \\
=& \frac{(n-k)}{(k+1)} \frac{\left(\frac{W_0}{s}+k\right)}{\left(\frac{B_0}{s}+n-k-1\right)} \binom{n}{k} \frac{\left(\frac{W_0}{s}\right)_k \left(\frac{B_0}{s}\right)_{n-k}}{\left(\frac{T_0}{s}\right)_n} \\
=& \frac{(n-k)}{(k+1)} \frac{\left(\frac{W_0}{s}+k\right)}{\left(\frac{B_0}{s}+n-k-1\right)} \, \mathbb{P}\left(\tilde{W}_n = k\right)
\end{aligned}$$

We now turn to the relative difference:

$$\begin{aligned}
&\frac{\mathbb{P}\big(\tilde{W}_n = k+1\big) - \mathbb{P}\big(\tilde{W}_n = k\big)}{\mathbb{P}\big(\tilde{W}_n = k\big)} \\
&= \frac{(n-k)\left(\frac{W_0}{s}+k\right) - (k+1)\left(\frac{B_0}{s}+n-k-1\right)}{(k+1)\left(\frac{B_0}{s}+n-k-1\right)} \\
&= \frac{n\frac{W_0}{s} + nk - k\frac{W_0}{s} - k^2 - \left(k\frac{B_0}{s} + kn - k^2 - k + \frac{B_0}{s} + n - k - 1\right)}{(k+1)\left(\frac{B_0}{s}+n-k-1\right)} \\
&= \frac{n\frac{W_0}{s} + nk - k\frac{W_0}{s} - k^2 - k\frac{B_0}{s} - kn + k^2 + k - \frac{B_0}{s} - n + k + 1}{(k+1)\left(\frac{B_0}{s}+n-k-1\right)} \\
&= \frac{(n-k)\frac{W_0}{s} - (k+1)\frac{B_0}{s} - n + 2k + 1}{(k+1)\left(\frac{B_0}{s}+n-k-1\right)}
\end{aligned}$$

Now express this in terms of $Z_n = \frac{\tilde{W}_n}{n}$, which takes its values in $\{0, \frac{1}{n}, \frac{2}{n}, \cdots, 1\}$. Clearly, $\tilde{W}_n = nZ_n$. With $k = nx$, where $x \in \{0, \frac{1}{n}, \frac{2}{n}, \cdots, 1\}$, we have

$$\mathbb{P}(\tilde{W}_n = k) = \mathbb{P}\left(Z_n = x\right)$$

and

$$\begin{aligned}
\frac{\mathbb{P}\big(\tilde{W}_n = k+1\big) - \mathbb{P}\big(\tilde{W}_n = k\big)}{\mathbb{P}\big(\tilde{W}_n = k\big)} &= \frac{\mathbb{P}\big(Z_n = x + \frac{1}{n}\big) - \mathbb{P}(Z_n = x)}{P(Z_n = x)} \\
&= \frac{(n-nx)\frac{W_0}{s} - (nx+1)\frac{B_0}{s} - n + 2k + 1}{(nx+1)\left(\frac{B_0}{s}+n-k-1\right)}
\end{aligned}$$

As seen from the histogram, it is reasonable to conjecture that the distribution of Z_n approaches a probability distribution with some sufficiently smooth density f. In this case,

$$\mathbb{P}\left(Z_n = x\right) = \mathbb{P}\left(x - \frac{1}{2n} < Z_n < x + \frac{1}{2n}\right) \approx \int_{x-\frac{1}{2n}}^{x+\frac{1}{2n}} f\left(t\right) dt \approx f(x)\frac{1}{n} \tag{3.1}$$

The error in the second approximation can be estimated by means of the following assertion:

Lemma 3.1. *For each function f, which is twice continuously differentiable on $(x - \frac{\delta}{2}, x + \frac{\delta}{2})$, we have*

$$\Big| \int_{x-\frac{\delta}{2}}^{x+\frac{\delta}{2}} f(t)\, dt - \delta f(x) \Big| \leq \frac{\delta^3}{24} \max_{x-\frac{\delta}{2}<t<x+\frac{\delta}{2}} |f''(t)| \,. \tag{3.2}$$

Proof. Write

$$\int_{x-\frac{\delta}{2}}^{x+\frac{\delta}{2}} f(t)\, dt = \int_{-\frac{\delta}{2}}^{\frac{\delta}{2}} f(x+s)\, ds.$$

We continue with the Taylor expansion:

$$f(x+s) = f(x) + f'(x)s + R(s)\,,$$

where $R(s) = \frac{f''(x+\theta_s s)}{2} s^2$ and $0 \leq \theta_s \leq 1$. Integrating, we obtain

$$\int_{-\frac{\delta}{2}}^{\frac{\delta}{2}} f(x+s)\, ds = f(x) \int_{-\frac{\delta}{2}}^{\frac{\delta}{2}} ds + f'(x) \int_{-\frac{\delta}{2}}^{\frac{\delta}{2}} s\, ds + \int_{-\frac{\delta}{2}}^{\frac{\delta}{2}} R(s)\, ds = \delta f(x) + \int_{-\frac{\delta}{2}}^{\frac{\delta}{2}} R(s)\, ds$$

Therefore,

$$\begin{aligned}
\left| \int_{-\frac{\delta}{2}}^{\frac{\delta}{2}} f(x+s)\, ds - \delta f(x) \right| &\leq \int_{-\frac{\delta}{2}}^{\frac{\delta}{2}} \big|R(s)\big|\, ds \\
&\leq \frac{1}{2} \sup_{x-\frac{\delta}{2}\leq t \leq x+\frac{\delta}{2}} |f''(t)| \cdot \int_{-\frac{\delta}{2}}^{\frac{\delta}{2}} s^2\, ds \\
&= \frac{1}{2} \sup_{x-\frac{\delta}{2}\leq t \leq x+\frac{\delta}{2}} |f''(t)| \cdot 2 \cdot \frac{1}{3} \left(\frac{\delta}{2}\right)^3 \\
&= \frac{\delta^3}{24}
\end{aligned}$$

□

From the preceding lemma, we saw that the error in the second approximation in (3.1) is of order $\frac{1}{n^3}$, typically much smaller than the difference

$$\mathbb{P}\left(Z_n = x + \frac{1}{n}\right) - \mathbb{P}\left(Z_n = x\right) \approx \int_{x-\frac{1}{2n}}^{x+\frac{1}{2n}} f\left(t\right) dt$$

which is typically of order $\frac{1}{n}$ .The left hand side of 3.1 can be approximated as:

$$\frac{\mathbb{P}\left(Z_n = x + \frac{1}{n}\right) - \mathbb{P}\left(Z_n = x\right)}{\mathbb{P}\left(Z_n = x\right)} \approx \frac{f\left(x + \frac{1}{n}\right)\frac{1}{n} - f\left(x\right)\frac{1}{n}}{f\left(x\right)\frac{1}{n}} \approx \frac{f'\left(x\right)\frac{1}{n}}{f\left(x\right)}$$

The right hand side of 3.1 can be approximated as:

$$\begin{aligned}
&\frac{\mathbb{P}\left(Z_n = x + \frac{1}{n}\right) - \mathbb{P}\left(Z_n = x\right)}{\mathbb{P}\left(Z_n = x\right)} \\
&\quad = \frac{(n - nx)\frac{W_0}{s} - (nx+1)\frac{B_0}{s} - n + 2k + 1}{(nx+1)\left(\frac{B_0}{s} + n - k - 1\right)} \\
&\quad = \frac{n}{n^2}\frac{(1-x)\frac{W_0}{s} - \left(x + \frac{1}{n}\right)\frac{B_0}{s} - 1 + 2x + \frac{1}{n}}{(nx+1)\left(\frac{B_0}{s} + n - k - 1\right)} \\
&\quad \approx \frac{1}{n}\frac{(1-x)\frac{W_0}{s} - x\frac{B_0}{s} - 1 + 2x}{x(1-x)}
\end{aligned}$$

It is reasonable to conjecture that in the approximation of both sides of (3.1), the relative error tends to zero as n tends to infinity. In this case, we have:

$$\frac{f'(x)}{f(x)} = \frac{(1-x)\frac{W_0}{s} - x\frac{B_0}{s} - 1 + 2x}{x(1-x)}$$

To solve this differential equation, we write $y = f(x)$ and express the derivative in terms of differentials:

$$\frac{dy}{dx\ y} = \frac{(1-x)\frac{W_0}{s} - x\frac{B_0}{s} - 1 + 2x}{x(1-x)}$$

Separating the variables, we obtain:

$$\begin{aligned}
\frac{dy}{y} &= \frac{\left((1-x)\frac{W_0}{s} - x\frac{B_0}{s} - 1 + 2x\right)\,dx}{x(1-x)} \\
\ln|y| &= \int\left(\frac{W_0}{s}\frac{1}{x} - \frac{B_0}{s}\frac{1}{1-x} - \frac{1}{x(1-x)} + \frac{2}{1-x}\right)dx \\
\ln|y| &= \frac{W_0}{s}\ln|x| + \frac{B_0}{s}\ln|1-x| - \ln|x| + \ln|1-x| - 2\ln|1-x| + C
\end{aligned}$$

No absolute values are necessary because $x \in [0,1]$, so that

$$\begin{aligned}
\ln y &= \left(\frac{W_0}{s} - 1\right)\ln x + \left(\frac{B_0}{s} - 1\right)\ln(1-x) + C \\
y &= x^{\frac{W_0}{s}-1}(1-x)^{\frac{B_0}{s}-1}e^C.
\end{aligned}$$

With

$$e^C = \frac{1}{\beta(\frac{W_0}{s}, \frac{B_0}{s})},$$

this is the probability density function of the beta distribution $\beta(\frac{W_0}{s}, \frac{B_0}{s})$, which should be the limiting distribution. Recall that $\beta(a,b)$ is a continuous distribution with density

$$p_{a,b}(t) = \begin{cases} \frac{1}{\beta(a,b)}\,t^{a-1}(1-t)^{b-1} & ;\ 0 < t < 1 \\ 0 & ;\ \text{otherwise} \end{cases} \tag{3.3}$$

Theorem 3.2 (Eggenberger and Pólya, 1923)**.** *Let $\tilde{W}_n$ be the number of white ball drawings in the Pólya-Eggenberger urn after n draws. Then, $Z_n = \frac{\tilde{W}_n}{n}$ converges in the Kolmogorov metric (see Appendix B) to the beta distribution $\beta(a,b)$, where $a = \frac{W_0}{s}$ and $b = \frac{B_0}{s}$. That is,*

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}(Z_n \le x) - F_{a,b}(x)\right| \xrightarrow[n\to\infty]{} 0\,,$$

where $F_{a,b}(x) = \int_{-\infty}^{x} p_{a,b}(t)\,dt$ is the cumulative distribution function of $\beta(a,b)$. Consequently, $\frac{\tilde{W}_n}{n}$ converges weakly to $\beta(a,b)$.

We illustrate the convergence with a few histograms of the Pólya Eggenberger distributions along with the underlying beta distributions.

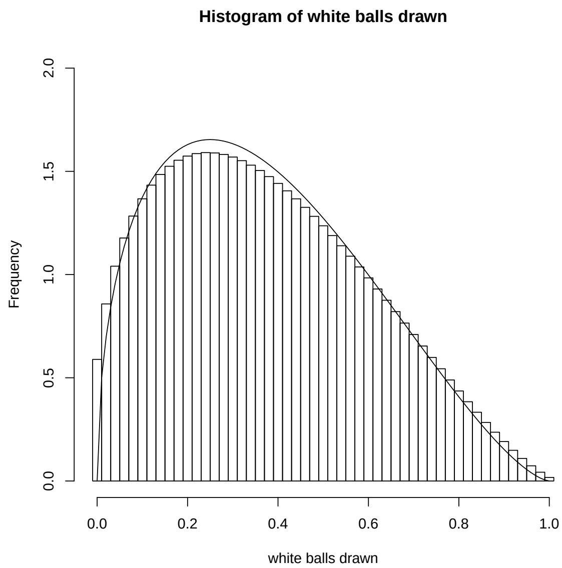

Figure 6: Histogram with beta density for $W_0 = 3, B_0 = 5, s = 2, n = 50$

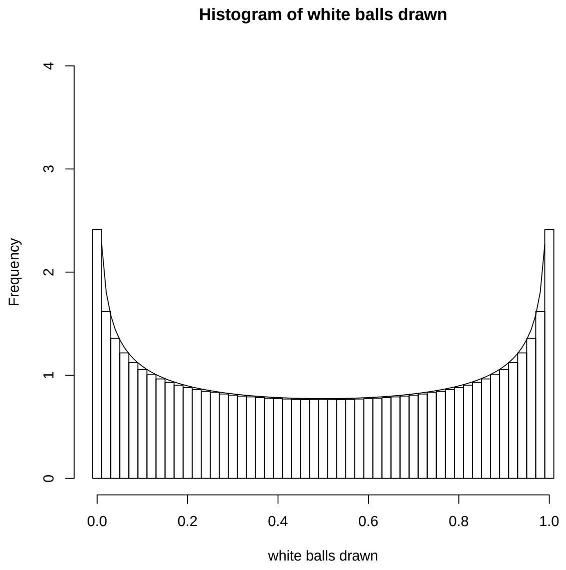

Figure 7: Histogram with beta density for $W_0 = 2, B_0 = 2, s = 3, n = 50$

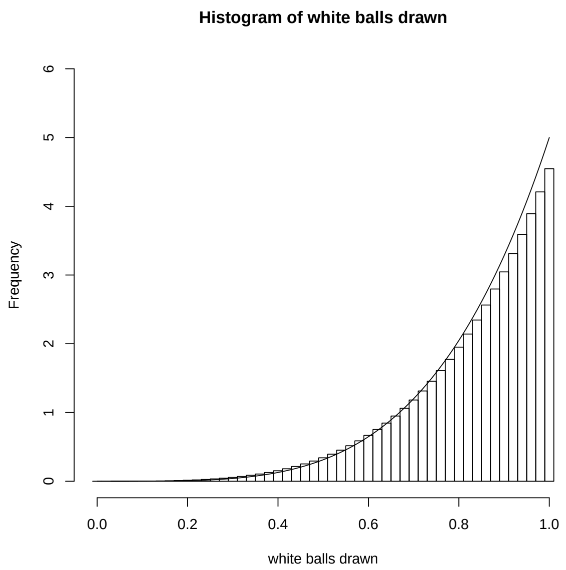

Figure 8: Histogram with beta density for $W_0 = 5, B_0 = 1, s = 1, n = 50$

Proof of Theorem 3.2. We can start by rewriting the equation

$$\mathbb{P}\left(\tilde{W}_n = k\right) = \binom{n}{k} \frac{\left(\frac{W_0}{s}\right)_k \left(\frac{B_0}{s}\right)_{n-k}}{\left(\frac{T_0}{s}\right)_n},$$

a bit differently. The rising factorial can be expressed in terms of the gamma function, which is defined in Appendix A, as follows:

$$(x)_n = \frac{\Gamma(x+n)}{\Gamma(x)}.$$

Furthermore, the binomial coefficient can be expressed as:

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} = \frac{\Gamma(n+1)}{\Gamma(k+1)\,\Gamma(n-k+1)}.$$

Putting it all together, we find that:

$$\mathbb{P}\big(\tilde{W}_n = k\big) = \frac{\Gamma(n+1)}{\Gamma(k+1)\,\Gamma(n-k+1)} \frac{\Gamma\left(k+\frac{W_0}{s}\right)\,\Gamma\left(n-k+\frac{B_0}{s}\right)}{\Gamma\left(\frac{W_0}{s}\right)\,\Gamma\left(\frac{B_0}{s}\right)\,\frac{\Gamma\left(n+\frac{T_0}{s}\right)}{\Gamma\left(\frac{T_0}{s}\right)}}$$

Now recall Theorem A.1, which can be alternatively written as:

$$\frac{\Gamma(x+a)}{\Gamma(x)} = x^a\, R(x,a),$$

where $\lim_{x\to\infty} R(x,a) = 1$ for all a. Thus, we can rewrite $\mathbb{P}(\tilde{W}_n = k)$, as:

$$\begin{aligned}
P(\tilde{W}_n = k) &= \frac{\Gamma(\frac{W_0}{s}+k)}{\Gamma(\frac{W_0}{s})}\,\frac{\Gamma(\frac{B_0}{s}+n-k)}{\Gamma(\frac{B_0}{s})}\,\frac{\Gamma(\frac{T_0}{s})}{\Gamma(\frac{T_0}{s}+n)}\,\frac{\Gamma(n+1)}{\Gamma(k+1)\Gamma(n-k+1)}\\
&= \frac{\Gamma(k+\frac{W_0}{s})}{\Gamma(k+1)}\,\frac{\Gamma(n-k+\frac{B_0}{s})}{\Gamma(n-k+1)}\,\frac{\Gamma(n+1)}{\Gamma(n+\frac{T_0}{s})}\,\frac{\Gamma(\frac{T_0}{s})}{\Gamma(\frac{W_0}{s})\Gamma(\frac{B_0}{s})}\\
&= \frac{\Gamma(k+\frac{W_0}{s})}{k\,\Gamma(k)}\,\frac{\Gamma(n-k+\frac{B_0}{s})}{(n-k)\,\Gamma(n-k)}\,\frac{n\,\Gamma(n)}{\Gamma(n+\frac{T_0}{s})}\,\frac{\Gamma(\frac{T_0}{s})}{\Gamma(\frac{W_0}{s})\,\Gamma(\frac{B_0}{s})}\\
&= k^{\frac{W_0}{s}-1} R\left(k, \frac{W_0}{s}\right)(n-k)^{\frac{B_0}{s}-1}\, R\left(n-k, \frac{B_0}{s}\right) n^{1-\frac{T_0}{s}}\,\frac{1}{R\left(n, \frac{T_0}{s}\right)}\,\frac{1}{B(\frac{W_0}{s}, \frac{B_0}{s})}\\
&= \frac{1}{n}\, p_{\frac{W_0}{s}, \frac{B_0}{s}}\left(\frac{k}{n}\right) S(n,k)
\end{aligned}$$

where

$$S(n,k) = \frac{R(k, \frac{W_0}{s})\, R(n-k, \frac{B_0}{s})}{R(n, \frac{T_0}{s})}.$$

Now we interrupt the proof of Theorem 3.2 by a few assertions.

Proposition 3.3.

$$\lim_{m\to\infty} \sup_{\substack{k\geq m\\ n-k\geq m}} |S(n,k)-1| = 0\,.$$

Proof. It is equivalent to prove:

$$\lim_{m\to\infty} \sup_{\substack{k\geq m\\ n-k\geq m}} |\ln S(n,k)| = 0$$

Observe that

$$|\ln S(n,k)| \le |\ln R(k, \frac{W_0}{s})| + |\ln R(n-k, \frac{B_0}{s})| + |\ln R(n, \frac{T_0}{s})|$$

The proof is now completed by the observation that: $\lim_{x\to\infty} \ln R(x,a) = 0$ implies $\lim_{m\to\infty} \sup_{x\ge m} |\ln R(x,k)| = 0$. □

Lemma 3.4.

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} \left| p_{a,b}\left(\frac{k}{n}\right) - \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} p_{a,b}\left(\frac{x}{n}\right) dx \right| = 0$$

Proof. Let:

$$D_k = \left| p_{a,b}\left(\frac{k}{n}\right) - \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} p_{a,b}\left(\frac{x}{n}\right) dx \right|$$

D_k can be estimated in two ways. First,

$$D_k \le p_{a,b}\left(\frac{x}{n}\right) + \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} p_{a,b}\left(\frac{x}{n}\right) dx \tag{3.4}$$

Second, by Lemma 3.1,

$$D_k \le \frac{1}{24} \max_{k-\frac{1}{2}<x<k+\frac{1}{2}} |p''_{a,b,n}(x)|\,, \tag{3.5}$$

where $p_{a,b,n}(x) = p_{a,b}(\frac{x}{n})$; notice that $p''_{a,b,n}(x) = \frac{1}{n^2} p''_{a,b}(\frac{x}{n})$, so that

$$D_k \le \frac{1}{24n^2} \max_{k-\frac{1}{2}\le x\le k+\frac{1}{2}} \left| p''_{a,b}\left(\frac{x}{n}\right) \right| = \frac{1}{24n^2} \max_{\frac{k}{n}-\frac{1}{2n}\le t\le \frac{k}{n}+\frac{1}{2n}} \left| p''_{a,b}(t) \right|.$$

Take $\epsilon > 0$. There exist $u, v > 0$, such that $u \le 1-v$,

$$\int_0^u p_{a,b}(x)\,dx < \frac{\epsilon}{8} \quad \text{and} \quad \int_{1-v}^1 p_{a,b}(x)\,dx < \frac{\epsilon}{8}.$$

Moreover, as $p_{a,b}$ has at most one stationary point, we can in addition assume that $p_{a,b}$ is monotone on $[0,u]$ and $[1-v,1]$. Take $r, s \in \{1, 2, \cdots, n-1\}$ and write:

$$\Sigma_1 = \sum_{k=0}^{r-1} D_k, \qquad \Sigma_2 = \sum_{k=r}^{s} D_k, \qquad \Sigma_3 = \sum_{k=s+1}^{n} D_k\,.$$

In Σ_1, we can estimate the summands according to (3.4), leading to:

$$\Sigma_1 \le \sum_{k=0}^{r-1} \left[p_{a,b}\left(\frac{k}{n}\right) + \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} p_{a,b}\left(\frac{x}{n}\right) dx \right].$$

Letting $r = \lfloor nu \rfloor, p_{a,b}$ is monotone on $[0, \frac{r}{n}]$, and we can estimate:

$$\sum_{k=1}^{r-1} p_{a,b}\left(\frac{k}{n}\right) \le \int_0^r p_{a,b}\left(\frac{x}{n}\right) dx\,,$$

and

$$\Sigma_1 \leq 2\int_0^r p_{a,b}\left(\frac{x}{n}\right)\,dx = 2n\int_0^{\frac{r}{n}} p_{a,b}(t)\,dt \leq 2n\int_0^n p_{a,b}(t)\,dt \leq \frac{n\epsilon}{4}.$$

Similarly, letting $s = \lceil n(1-v)\rceil, p_{a,b}$ is monotone on $[\frac{\epsilon}{n}, 1]$, and

$$\Sigma_3 \leq 2\int_s^n p_{a,b}\left(\frac{x}{n}\right)\,dx = 2n\int_{\frac{s}{n}}^1 p_{a,b}(t)\,dt \leq 2n\int_{1-v}^1 p_{a,b}(t)\,dt \leq \frac{n\epsilon}{4}.$$

In Σ_2, we estimate the $s-r-1$ summands according to (3.5) leading to:

$$\begin{aligned}\Sigma_2 &\leq \frac{s-r-1}{24\,n^2}\max_{r-\frac{1}{2}\leq x\leq s+\frac{1}{2}}|p''_{a,b}(x)|\\ &\leq \frac{1}{24\,n}\max_{nu-\frac{3}{2}\leq x\leq n(1-v)+\frac{3}{2}}|p''_{a,b}(x)|\\ &= \frac{1}{24\,n}\max_{u-\frac{3}{2n}\leq t\leq 1-v+\frac{3}{2n}}|p''_{a,b}(t)|\,.\end{aligned}$$

For sufficiently large n, we have $u-\frac{3}{2n}\geq\frac{u}{2}$ and $1-v+\frac{3}{3n}\leq 1-\frac{v}{2}$. As $p_{a,b}$ is infinitely smooth on $[\frac{u}{2}, 1-\frac{v}{2}]$, we have

$$M = \sup_{\frac{u}{2}\leq t\leq\frac{v}{2}}|p''_{a,b}(t)| < \infty.$$

Collecting all together, we obtain:

$$\frac{1}{n}\sum_{k=0}^n D_k = \frac{1}{n}\big(\Sigma_1+\Sigma_2+\Sigma_3\big) \leq \frac{\epsilon}{2} + \frac{M}{24\,n^2}.$$

For sufficiently large n, $\frac{M}{24\,n^2} < \frac{\epsilon}{2}$ and the proof is complete. □

Proposition 3.5. *For fixed W_0, B_0 and s we have $\mathbb{P}\big(\tilde{W}_n = 0\big) \xrightarrow[n\to\infty]{} 0$ and $\mathbb{P}\big(\tilde{W}_n = n\big) \xrightarrow[n\to\infty]{} 0$.*

Proof. It suffices to prove the first assertion, as the second one follows by symmetry. By Theorem 2.1, we have:

$$\mathbb{P}\big(\tilde{W}_n = 0\big) = \frac{B_0(B_0+s)\cdots(B_0+(n-1)s)}{T_0(T_0+s)\cdots(T_0+(n-1)s)}\,. \tag{3.6}$$

Observe that:

$$\frac{B_0+x}{T_0+x} = 1 - \frac{T_0-B_0}{T_0+x} \leq e^{-\frac{T_0-B_0}{T_0+x}} \tag{3.7}$$

The latter inequality follows from Taylor's expansion: $e^a = 1 + a + e^{\theta a}\frac{a^2}{2}$ where $0\leq\theta\leq 1$, so that $e^a \geq 1+a$ for all $a\in\mathbb{R}$. Applying (3.7) to (3.6), we obtain:

$$\mathbb{P}\big(\tilde{W}_n = 0\big) \leq e^{-\frac{B_0(B_0+s)\cdots(B_0+(n-1)s)}{T_0(T_0+s)\cdots(T_0+(n-1)s)}}$$

Estimating

$$\frac{T_0 - B_0}{T_0 + ks} \geq \int_k^{k+1} \frac{T_0 - B_0}{T_0 + ts}\, dt$$
$$\sum_{k=0}^{n-1} \frac{T_0 - B_0}{T_0 + ks} \geq \int_0^{n} \frac{T_0 - B_0}{T_0 + ts}\, dt$$

and integrating

$$\int_0^n \frac{T_0 - B_0}{T_0 + ts}\, dt = (T_0 - B_0) \int_0^n \frac{1}{T_0 + ts}\, dt = \frac{(T_0 - B_0)}{s} \ln \frac{T_0 + ns}{T_0}$$

we conclude that

$$\mathbb{P}\big(\tilde{W}_n = 0\big) \leq e^{-\frac{(T_0 - B_0)}{s} \ln \frac{T_0 + ns}{T_0}} \xrightarrow[n\to\infty]{} 0\,.$$

□

Proposition 3.6.

$$\lim_{n\to\infty} \sum_{k=0}^{n} \left| P(\tilde{W}_n = k) - \frac{1}{n}\, p_{a,b} \left(\frac{k}{n}\right) \right| = 0$$

Proof. Let

$$D_k = \left| P(\tilde{W}_n = k) - \frac{1}{n}\, p_{a,b} \left(\frac{k}{n}\right) \right|.$$

Observe that, by the definition of $p_{a,b}$ (Formula (3.3)), $D_0 = \mathbb{P}\big(\tilde{W}_n = 0\big)$ and $D_n = \mathbb{P}\big(\tilde{W}_n = n\big)$. For, $k = 1, 2, \cdots, n-1$ we can estimate

$$D_k \leq \frac{1}{n} \big| S(n,k) - 1 \big|\, p_{a,b} \left(\frac{k}{n}\right).$$

Let $\epsilon > 0$. By Proposition 3.3, there exists $m \in \mathbb{N}$, such that

$$\max_{m \leq k \leq n-m} \big| S(n,k) - 1 \big| < \frac{\epsilon}{8}.$$

Put

$$\Sigma_1 = \sum_{k=1}^{m-1} D_k, \qquad \Sigma_2 = \sum_{k=m}^{n-m} D_k, \qquad \Sigma_3 = \sum_{k=n-m+1}^{n} D_k\,.$$

We can estimate

$$\Sigma_2 \leq \frac{1}{n} \frac{\epsilon}{8} \sum_{k=0}^{n} p_{a,b} \left(\frac{k}{n}\right).$$

By Lemma (3.5), we have

$$\frac{1}{n} \sum_{k=0}^{n} \left| p_{a,b} \left(\frac{k}{n}\right) - \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} p_{a,b} \left(\frac{x}{n}\right) dx \right| < 1$$

for sufficiently large n. Then

$$\begin{aligned}\frac{1}{n}\sum_{k=0}^{n} p_{a,b}\left(\frac{k}{n}\right) &\leq \frac{1}{n}\sum_{k=0}^{n}\int_{k-\frac{1}{2}}^{k+\frac{1}{2}} p_{a,b}\left(\frac{x}{n}\right)\,dx+1\\ &= \frac{1}{n}\int_{-\frac{1}{2}\text{ or }0}^{n+\frac{1}{2}\text{ or }n} p_{a,b}\left(\frac{x}{n}\right)\,dx+1\\ &= \int_0^1 p_{a,b}\,dt+1\\ &= 2.\end{aligned}$$

As a result, $\Sigma_2 \leq \frac{\epsilon}{4}$ for sufficiently large n. Since $R(x,\frac{W_0}{s}), R(x,\frac{B_0}{s})$ and $R(x,\frac{T_0}{s})$ all converge to 1 as $x\to\infty$, $R(x,\frac{W_0}{s}), R(x,\frac{B_0}{s})$ and $R(x,\frac{T_0}{s})$ are all bounded in $x \geq 1$. Therefore,

$$M = \sup_{\substack{n\in\mathbb{N}\\ 1\leq k\leq n-1}} |S(n,k)-1| < \infty.$$

There exist $u>0$, $v>0$, such that $u \leq 1-v$, and

$$\int_0^u p_{a,b}(t)\,dt < \frac{\epsilon}{4M} \quad\text{and}\quad \int_{1-v}^1 p_{a,b}(t)\,dt < \frac{\epsilon}{4M}.$$

Again, we can assume that $p_{a,b}$ is monotone on $(0,u]$ and $[v,1)$. Now if $\frac{m}{n}\leq n$, $p_{a,b}$ is monotone on $(0,\frac{m}{n}]$ and

$$\Sigma_1 \leq \int_{k=1}^{m-1}\frac{1}{n}Mp_{a,b}\left(\frac{k}{n}\right)\,dx \leq \frac{M}{n}\int_0^m p_{a,b}\left(\frac{x}{n}\right)\,dx \leq M\int_0^{\frac{m}{n}} p_{a,b}(t)\,dt < \frac{\epsilon}{4}$$

Similarly, if $\frac{m}{n}\leq v$, $\frac{n-m}{n}\geq 1-v, p_{a,b}$ is monotone on $[\frac{n-m}{n},1)$ and

$$\Sigma_3 \leq \int_{k=n-m+1}^{n}\frac{1}{n}Mp_{a,b}\left(\frac{k}{n}\right)\,dx \leq \frac{M}{n}\int_{n-m}^n p_{a,b}\left(\frac{x}{n}\right)\,dx \leq M\int_{\frac{n-m}{n}}^{\frac{n}{n}} p_{a,b}(t)\,dt < \frac{\epsilon}{4}$$

By Proposition 3.5, $D_0 = \mathbb{P}\big(\tilde{W}_n=0\big) < \frac{\epsilon}{8}$ and $D_n = \mathbb{P}\big(\tilde{W}_n=0\big) < \frac{\epsilon}{8}$ for sufficiently large n. Summing up, we obtain that

$$\sum_{k=0}^{n}\left|\,\mathbb{P}\big(\tilde{W}_n=k\big)-\frac{1}{n}\int_{k-\frac{1}{2}}^{k+\frac{1}{2}} p_{a,b}\left(\frac{x}{n}\right)\,dx\right| < \epsilon\,,$$

for sufficiently large n. This completes the proof. □

Combining Lemma 3.4 and Proposition 3.6, we obtain the following

Corollary 3.7.

$$\lim_{n\to\infty}\sum_{k=0}^{n}\left|\,\mathbb{P}\big(\tilde{W}_n=k\big)-\frac{1}{n}\int_{k-\frac{1}{2}}^{k+\frac{1}{2}} p_{a,b}\left(\frac{x}{n}\right)\,dx\right| = 0$$

The preceding corollary can be equivalently written as:

$$\lim_{n\to\infty}\sum_{k=0}^{n}\left|\mathbb{P}\left(Z_n=\frac{k}{n}\right)-\frac{1}{n}\int_{k-\frac{1}{2n}}^{k+\frac{1}{2n}}p_{a,b}(t)\,dt\right|=0,$$

Defining a discretized beta distribution as the distribution of a $\{0,\frac{1}{n},\cdots,1\}$-valued random variable $B_{a,b,n}$ with

$$\mathbb{P}\left(B_{a,b,n}=\frac{k}{n}\right)-\frac{1}{n}\int_{k-\frac{1}{2n}}^{k+\frac{1}{2n}}p_{a,b}(t)\,dt=0,\quad k=0,1,\cdots,n$$

Corollary 3.7 says that

$$d_{TV}\big(\mathcal{L}(Z_n),\mathcal{L}(B_{a,b,n})\big)\xrightarrow[n\to\infty]{}0\,.$$

(See Proposition B.5).

Continuation of Proof of Theorem 3.2. Corollary 3.7 says that the distribution of Z_n and $B_{a,b,n}$ are close as $n\to\infty$. If $\mathcal{L}(B_{a,b,n})$ is close to $\beta(a,b)$, then Z_n is close to $\beta(a,b)$. However, since $B_{a,b,n}$ is discrete and $\beta(a,b)$ is continuous, the latter two distributions cannot be close in d_{TV}, but they are close in d_K. To show the latter, first observe that

$$\mathbb{P}\left(B_{a,b,n}\le\frac{k}{n}\right)=F_{a,b}\left(\frac{k}{n}+\frac{1}{2n}\right)=\int_0^{\frac{k}{n}+\frac{1}{2n}}p_{a,b}(t)\,dt$$

for all $k=0,1,\ldots,n$. Now let $t\in[0,1)$ and $k=\lfloor nt\rfloor$. Write

$$\mathbb{P}\big(B_{a,b,n}\le t\big)=\mathbb{P}\left(B_{a,b,n}\le\frac{k}{n}\right)=F_{a,b}\left(\frac{k}{n}+\frac{1}{2n}\right)\approx F_{a,b}(t)\tag{3.8}$$

Let $\epsilon>0$. Observe that $\beta(a,b)$ is continuous and uniformly continuous on $[0,1]$. Therefore, there exists δ, such that $|F_{a,b}(x)-F_{a,b}(y)|<\epsilon$ for all $x,y\in[0,1]$ with $|x-y|<\delta$. In particular, if $\left|\frac{k}{n}+\frac{1}{2n}-t\right|<\delta$, then

$$\left|F_{a,b}\left(\frac{k}{n}+\frac{1}{2n}\right)-F_{a,b}(t)\right|<\epsilon\,.$$

However, according to the choice of k, we have $\left|\frac{k}{n}+\frac{1}{2n}-t\right|\le\frac{1}{2n}$. Therefore, if $n\ge\frac{1}{2\delta}$, then, recalling (3.8), we have

$$\big|\mathbb{P}(B_{a,b,n}\le t)-F_{a,b}(t)\big|<\epsilon$$

for $t\in[0,1)$. As $\mathbb{P}(B_{a,b,n}\le t)=F_{a,b}(t)=0$ for $t<0$ and $\mathbb{P}(B_{a,b,n}\le t)=F_{a,b}(t)=1$ for $t\ge1$, we conclude that

$$\sup_{t\in\mathbb{R}}\big|\mathbb{P}(B_{a,b,n}\le t)-F_{a,b}(t)\big|<\epsilon\,.$$

Therefore, $d_K\big(\mathcal{L}(B_{a,b,n}),\beta(a,b)\big)\xrightarrow[n\to\infty]{}0$. Recalling that $d_K\big(\mathcal{L}(Z_n),\mathcal{L}(B_{a,b,n})\big)\xrightarrow[n\to\infty]{}0$, we obtain the desired result.

□

4 Conclusion

The rescaled Pólya Eggenberger distribution converges to the beta distribution rapidly as the number of draws increases. It is curious that the limiting properties of a Pólya Eggenberger urn depend critically on the initial conditions. However, the usefulness of this fact may be limited as the point probabilities of the Pólya Eggenberger distribution can be expressed by a closed formula. It would be interesting to examine convergence of more complicated urn models, but this is beyond the scope of this paper.

5 Povzetek naloge v slovenskem jeziku

Eni izmed najpogostejših modelov, s katerimi se srečamo v kombinatorični verjetnosti, so modeli v katerih iz posod (žar) jemljemo kroglice. Z njimi se da predstaviti veliko problemov, kot so potek volitev, problem zasedenosti, kockarjev bankrot, Banachov problem vžigalic in tudi kar nekaj modelov v bioznanosti.
V pričujočem delu predstavimo Pólya-Eggenbergerjev model, pri katerem imamo eno samo posodo, v kateri je najprej W_0 belih in B_0 modrih kroglic. Pri vsakem koraku na slepo izvlečemo eno kroglico, pogledamo njeno barvo, jo vrnemo v posodo in dodamo še s novih kroglic iste barve. Označimo z $\tilde{W}_n$ število izvlečenih belih kroglic po n korakih. Potem se porazdelitev te slučajne spremenljivke da izraziti z eksplicitno formulo:

$$\mathbb{P}(\tilde{W}_n = k) = \frac{W_0(W_0+s)\cdots(W_0+(k-1)s)\,B_0(B_0+s)\cdots(B_0+(n-k-1)s)}{T_0(T_0+s)\cdots(T_0+(n-1)s)}\binom{n}{k}.$$

To formulo izpeljemo korak za korakom, ko vlečemo kroglice. Najprej določimo, ob katerih trenutkih naj bodo izvlečene bele kroglice, nakar za vsako vlečenje izračunamo ustrezno pogojno verjetnost, da bomo izvlekli kroglico ustrezne barve. Zanimivo je, da je verjetnost posameznega zaporedja barv neodvisna od razporeditve belih kroglic, kar se zgodi tudi pri binomski porazdelitvi.
Na preprost način se izražata pričakovana vrednost in varianca:

$$\mathbb{E}(\tilde{W}_n) = \frac{W_0}{T_0}\,n$$
$$\mathrm{Var}[\tilde{W}_n] = \frac{W_0\,B_0\,n\,(sn+T_0)}{T_0^2(T_0+s)}.$$

Pričakovano vrednost izračunamo tako, da člene $k\,\mathbb{P}\big(\tilde{W}_n = k\big)$ izrazimo kot točkaste verjetnosti za neko drugo žaro, pomnožene s konstantnim faktorjem, nato pa uporabimo, da je vsota verjetnosti enaka 1. Pri varianci pa izhajamo iz $\mathbb{E}\big[\tilde{W}_n(\tilde{W}_n-1)\big]$, kar izračunamo na enak način kot $\mathbb{E}(\tilde{W}_n)$.
V nadaljevanju nas zanima obnašanje porazdelitve, ko gre n proti neskončno, W_0, B_0 in s pa so fiksni. Narišemo histograme za neko izbiro parametrov W_0, B_0 in s ter vedno večje vrednosti parametra n. Histogrami nakazujejo, da slučajne spremenljivke

$Z_n = \frac{\tilde{W}_n}{n}$ konvergirajo proti neki zvezni porazdelitvi. Da bi ugotovili, katera porazdelitev je to, studiramo relativne razlike

$$\frac{\mathbb{P}\big(\tilde{W}_n = k+1\big) - \mathbb{P}\big(\tilde{W}_n = k\big)}{\mathbb{P}\big(\tilde{W}_n = k\big)}$$

in jih primerjamo z logaritamskim odvodom gostote limitne porazdelitve. Izpeljemo diferencialno enačbo, ki ji mora ta gostota zadoščati, jo rešimo in ugotovimo da gre za porazdelitev beta s parametroma $a := \frac{W_0}{s}$ in $b := \frac{B_0}{s}$. To je porazdelitev z gostoto

$$p_{a,b}(t) = \begin{cases} \frac{1}{\beta(a,b)}\, t^{a-1}(1-t)^{b-1} & ;\ 0 < t < 1 \\ 0 & ;\ \text{sicer} \end{cases}$$

Konvergenco pokažemo v metriki Kolmogorova, kar pomeni, da gre

$$\sup_{x\in\mathbb{R}}\Big|\mathbb{P}(Z_n \le x) - F_{a,b}(x)\Big| \xrightarrow[n\to\infty]{} 0\,,$$

kjer je $F_{a,b}(x) = \int_{-\infty}^{x} p_{a,b}(t)\, dt$ kumulativna porazdelitvena funkcija porazdelitve $\beta(a,b)$. Pokažemo tudi, da gre razdalja v totalni variaciji med ustrezno skrčeno Pólya-Eggenbergerjevo porazdelitvijo in ustrezno diskretizirano porazdelitvijo beta proti 0. Natančneje, pokažemo, da je

$$\lim_{n\to\infty} \sum_{k=0}^{n} \left| \mathbb{P}\big(\tilde{W}_n = k\big) - \frac{1}{n} \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} p_{a,b}\left(\frac{x}{n}\right)\, dx \right| = 0$$

To naredimo z neposredno oceno razlik v slednji vsoti. Verjetnosti pri Pólya-Eggenbergerjevi porazdelitvi izrazimo s funkcijo gama in uporabimo da je logaritem funkcije gama pri velikih vrednostih približno linearen, natančneje, da za vse $a \in \mathbb{R}$ velja

$$\lim_{n\to\infty} \frac{\Gamma(x+a)}{x^a\, \Gamma(x)} = 1\,.$$

To dejstvo tudi eksplicitno izpeljemo iz Gaussove izražave funkcije gama

$$\Gamma(x) = \lim_{n\to\infty} \frac{n!\, n^x}{x(x+1)(x+2)\cdots(x+n)}\,.$$

6 Bibliography

[1] A.D. Barbour, L. Holst, and S. Janson, *Poisson Approximation.* Clarendon Press, 1992. *(Cited on page .)*

[2] F. Eggenberger and G. Pólya, Über die Statistik Verketteter Vorgänge. *Zeitschrift für Angewandte Mathematik und Mechanik* 3 (1923) 279–289. *(Cited on page 3.)*

[3] N. Johnson and S. Kotz, *Urn models and their applications:An approach to Modern Discrete Probability Theory.* Wiley, 1977. *(Cited on page 1.)*

[4] H.M. Mahmoud, *Pólya Urn Models*, CRC Press, 2009. *(Cited on pages 1 and 4.)*

[5] A. Markov, *Generalization of a problem on a sequential exchange of balls(in Russian)*, Collected works, Meeting of PhysicoMathematical Society of the Academy of Sciences, 1917. *(Cited on page 3.)*

[6] E.D. Rainville, *Special Functions*, Macmillan, 1960. *(Cited on page .)*

[7] A. Tchuprov, Ist die normale Stabilität empirisch nachweisbar?. *Nord Stat. Tidskr* B/I (1922) 373–378. *(Cited on page 3.)*

Appendices

A Some Properties of the Gamma Function

The purpose of the gamma function is to extend the factorial to real numbers: for historical reasons, we set $\Gamma(x) = (x-1)!$ for $x \in \mathbb{N}$. As $x! = (x-1)!\,x$, it would be reasonable to assume that $\Gamma(x+1) = x\,\Gamma(x)$ for all x. However, the gamma function is not uniquely determined by these two properties. What would be another reasonable assumption that would determine Γ uniquely? Observe that for $a \in \mathbb{N}$, $\Gamma(x+1) = \Gamma(x)x(x-1)\cdots(x+a-1)$ and that $x(x+1)\cdots(x+a+1) \sim x^a$ if x is large. Therefore, it would be reasonable to assume $\lim_{x\to\infty} \frac{\Gamma(x+1)}{x^a\,\Gamma(x)} = 1$ for all $a \in \mathbb{R}$. According to Gauss, we define the gamma function as:

$$\Gamma(x) = \lim_{n\to\infty} \frac{n!\,n^x}{x(x+1)(x+2)\cdots(x+n)} \tag{A.1}$$

(see also [6], page 12, Equation (4)). This function is defined for all $x \in \mathbb{R}\backslash\{0,-1,-2,\cdots\}$ (see again [6]). We first show that this is an extension of the factorial. Observe that

$$\Gamma(1) = \lim_{n\to\infty} \frac{n!\,n}{1\cdot 2\cdots n(n+1)} = \lim_{n\to\infty} \frac{n}{n+1} = 1,$$

and

$$\begin{aligned}
\Gamma(x+1) &= \lim_{n\to\infty} \frac{n!\,n^{x+1}}{(x+1)(x+2)\cdots(x+n)(x+1+n)} \\
&= \lim_{n\to\infty} \frac{n}{x+1+n}\frac{n!\,n^x}{(x+1)(x+2)\cdots(x+n)} \\
&= \lim_{n\to\infty} \frac{n!\,n^x}{(x+1)(x+2)\cdots(x+n)} \\
&= \lim_{n\to\infty} \frac{x\,n!\,n^x}{x(x+1)(x+2)\cdots(x+n)} \\
&= x\,\Gamma(x)
\end{aligned}$$

It follows by induction that Γ is an extension of the factorial: $\Gamma(n) = (n-1)!$ for all $n \in N$.

Theorem A.1. *For all $a \in \mathbb{R}$*

$$\lim_{n\to\infty} \frac{\Gamma(x+a)}{x^a\,\Gamma(x)} = 1.$$

In order to prove the result, we need the following assertion:

Lemma A.2.

$$\lim_{t\to\infty} t\,[\ln(a+t) - \ln t] = a \tag{A.2}$$

Proof.

$$\begin{aligned}\lim_{t\to\infty} t\,[\ln(a+t) - \ln t] &= \lim_{t\to\infty} t\,\ln(\frac{a+t}{t}) \\ &= \lim_{t\to\infty} \frac{\ln(a+t) - \ln t}{\frac{1}{t}}\end{aligned}$$

As the numerator and the denominator both tend to 0 as $t \to \infty$, we can use L'Hôpital's rule to obtain:

$$\lim_{t\to\infty} \frac{\frac{1}{a+t} - \frac{1}{t}}{-\frac{1}{t^2}} = \lim_{t\to\infty} \frac{\frac{t-(a+t)}{(a+t)\,t}}{-\frac{1}{t^2}} = \lim_{t\to\infty} \frac{at^2}{(a+t)\,t} = \lim_{t\to\infty} \frac{at}{a+t} = a$$

□

Now we can prove Theorem A.1.

Proof of Theorem A.1. Observe that by (A.1)

$$\begin{aligned}\frac{\Gamma(x+a)}{\Gamma(x)} &= \lim_{n\to\infty} n^a \frac{x(x+1)\cdots(x+n)}{(x+a)(x+a+1)\cdots(x+a+n)} \\ \ln\frac{\Gamma(x+a)}{x^a\,\Gamma(x)} &= \lim_{n\to\infty} F(x,n),\end{aligned}$$

where

$$F(x,n) = a\ln n - a\ln x - \sum_{k=0}^{n}\big[\ln(x+a+k) - \ln(x+k)\big].$$

It suffices to prove that

$$\lim_{x\to\infty}\lim_{n\to\infty} F(x,n) = 0 \tag{A.3}$$

Write

$$\ln(x+a+k) - \ln(x+k) = \ln\left(\frac{x+a+k}{x+k}\right) = \ln\left(1 + \frac{a}{x+k}\right),$$

which is a decreasing function in k, so we can bound:

$$\begin{aligned}\int_0^{n+1}\big(\ln(x+a+t) - \ln(x+t)\big)\,dt &\le \sum_{k=0}^{n}\big(\ln(x+a+k) - \ln(x+k)\big) \\ &\le \int_{-1}^{n}\left(\ln(x+a+t) - \ln(x+t)\right)\,dt\end{aligned}$$

We will write the left hand side as $I^-(x,n)$, and the right hand side as $I^+(x,n)$. Here we introduce a substitution $s = t+1$, and continue:

$$\begin{aligned} I^+(x,n) &= \int_{-1}^{n} (\ln(x+a+t) - \ln(x+t))\ dt \\ &= \int_{0}^{n+1} (\ln(x+a+s-1) - \ln(x+s-1))\ ds \\ &= I^-(x-1,n) \end{aligned}$$

Furthermore, it holds that:

$$a\ln n - a\ln x - I^+(x,n) < F(x,n) < a\ln n - a\ln x - I^-(x,n)$$

In order to show (A.3), it suffices to show

$$\lim_{x\to\infty}\lim_{n\to\infty}\left[a\ln n - a\ln x - I^+(x,n)\right] = 0.$$

$$\lim_{x\to\infty}\lim_{n\to\infty}\left[a\ln n - a\ln x - I^-(x,n)\right] = 0.$$

Observe that:

$$\begin{aligned} &a\ln n - a\ln x - I^+(x,n) \\ &\quad= a\ln n - a\ln x - I^-(x-1,n) \\ &\quad= a\ln n - a\ln(x-1) - I^-(x-1,n) + a\ln(x-1) - a\ln x \\ &\quad= a\ln n - a\ln(x-1) - I^-(x-1,n) + a\ln\frac{x-1}{x} \end{aligned}$$

Clearly,

$$\lim_{x\to\infty}\lim_{n\to\infty} a\ln\frac{x-1}{x} = 0,$$

because $\lim_{x\to\infty}\frac{x-1}{x} = 1$.
So it suffices to prove:

$$\lim_{x\to\infty}\lim_{n\to\infty}[a\ln n - a\ln x - I^-(x,n)] = 0.$$

Computing:

$$\begin{aligned} I^-(x,n) &= \int_{0}^{n+1} (\ln(x+a+t) - \ln(x+t))\ dt \\ &= \int_{0}^{n+1} \ln(x+a+t)\,dt - \int_{0}^{n+1} \ln(x+t)\,dt \\ &= (a+x+t)\ln(a+x+t)\Big|_{0}^{n+1} - (x+t)\ln(x+t)\Big|_{0}^{n+1} \\ &= (a+x+n+1)\ln(a+x+n+1) - (a+x)\ln(a+x) \\ &\qquad - (x+n+1)\ln(x+n+1) + x\ln x \end{aligned}$$

we obtain:

$$\begin{aligned} a\ln n - a\ln x - I^{-}(x,n) =& a\ln\frac{n}{a+x+n+a} \\ &-(x+n+1)\big[\ln(a+x+n+1)-\ln(x+n+1)\big] \\ &-a\ln x+(a+x)\ln(a+x)-x\ln x \end{aligned}$$

Letting $n\to\infty$, the first term tends to 0, while, by lemma (3.3), the second term tends to a. Therefore,

$$\begin{aligned} &\lim_{n\to\infty}\big[a\ln n - a\ln x - I^{-}(x,n)\big] \\ &= -a\ln x - a\ln x + (a+x)\ln(a+x) - x\ln x \end{aligned}$$

Now let $x\to\infty$:

$$\lim_{x\to\infty}\lim_{n\to\infty}\big[a\ln n - a\ln x - I^{-}(x,n)\big] = \lim_{x\to\infty}\big[-a\ln x - a + a\ln(a+x) + x(\ln(a+x)-\ln x)\big]$$

Applying Lemma (3.2) once again we find that: $\lim_{x\to\infty} x(\ln(a+x)-\ln x) = a$. Finally, we get:

$$\begin{aligned} \lim_{x\to\infty}\lim_{n\to\infty}\big(a\ln n - a\ln x - I^{-}(x,n)\big) &= \lim_{x\to\infty}\lim_{n\to\infty}\big[-a\ln x - a + a\ln(a+x) + a\big] \\ &= \lim_{x\to\infty}\lim_{n\to\infty}\big[-a\ln x + a\ln(a+x)\big] \\ &= \lim_{x\to\infty}\lim_{n\to\infty} a\,\ln\frac{a+x}{x} = 0 \end{aligned}$$

$\square$

B Convergence of Probability Distributions

There are numerous concepts of convergence of probability distribtutions. Many of them are based on metrics. Let us denote with $\mathcal{L}(X)$ the distribution of X.

Definition B.1. The total variation distance between the distributions of X and Y is defined as:

$$d_{TV}(\mathcal{L}(X), \mathcal{L}(Y)) = \sup_{\text{A measurable}} |\,\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)|\,.$$

This metric is usually too strong: Let X be a discrete random variable which takes values in a countable set S, and let Y be a continuous random variable. Then, $\mathbb{P}(X \in S) = 1$ and $\mathbb{P}(Y \in S) = 0$. This implies that $d_{TV}(\mathcal{L}(X), \mathcal{L}(Y)) = 1$. So we shall also consider a weaker metric.

Definition B.2. The Kolmogorov distance is given as:

$$d_K(\mathcal{L}(X), \mathcal{L}(Y)) = \sup_{a \in \mathbb{R}} |\,\mathbb{P}(X \leq a) - \mathbb{P}(Y \leq a)|.$$

Clearly,

$$d_K(\mathcal{L}(X), \mathcal{L}(Y)) \leq d_{TV}(\mathcal{L}(X), \mathcal{L}(Y)).$$

Definition B.3. A sequence of the distributions of random variables X_n $(n = 1, 2, \dots)$ is said to converge weakly to the distribution of X if

$$\lim_{n \to \infty} \mathbb{P}(X_n \leq x) = \mathbb{P}(X \leq x)$$

at each x where $F_X(x) = \mathbb{P}(X \leq x)$ is continuous. We write

$$X_n \xrightarrow[n \to \infty]{w} X\,.$$

Remark B.4. For a sequence of random variables X_n and an additional random variable X, we have the following implications:

$$d_{TV}(\mathcal{L}(X), \mathcal{L}(Y)) \xrightarrow[n \to \infty]{} 0 \implies d_K(\mathcal{L}(X), \mathcal{L}(Y)) \implies X_n \xrightarrow[n \to \infty]{w} X$$

Proposition B.5. *For discrete random variables X and Y, both taking values on a countable set S, we have*

$$d_{TV}(\mathcal{L}(X), \mathcal{L}(Y)) = \frac{1}{2} \sum_{a \in S} |\,\mathbb{P}(X = a) - \mathbb{P}(Y = a)|\,.$$

See [1].