# RATE OF CONVERGENCE OF POLYA'S URN TO THE BETA DISTRIBUTION

JOHN DRINANE

ADVISOR: YEVGENIY KOVCHEGOV
OREGON STATE UNIVERSITY

ABSTRACT. This paper will document the research of the Oregon State University 2008 Summer REU program. The convergence of the Urn to the Beta distribution will be shown. Different methods will be shown to arrive at specific rates. Also Mathematica 6 code will be included for generating the distribution of the urn so the reader can visualize the convergence.

## 1. INTRODUCTION

This project was motivated from the work of Kovchegov concerning the behavior of edge reinforced random walks. The particles move along the line, and the probability of the particles movement converges to the Beta Distribution. Two rates were derived for this process, one of order $1/n^2$ and one rate of the order of $1/\sqrt{n}$.

## 2. BACKGROUND INFORMATION

During the course of this paper the reader will need to be familiar with a couple of theorems, relationships, and definitions. This section is where these necessities are stated.

**Berry-Esseen 2.1.** [FW] *Let $X_1, X_2, \ldots$ be i.i.d with $E[X_i]=0, E[X_i^2]=\sigma^2$, and $E[|X_i|^3]=\rho < \infty$. If $F_n(x)$ is the distribution of $(X_1 + \cdots + X_n)/\sigma\sqrt{n}$ and $\mathcal{N}(x)$ is the standard normal distribution then*

$$|F_n(x) - \mathcal{N}(x)| \leq 3\rho/\sigma^3\sqrt{n}$$

**Cramér's 2.1.** [HF]
*Let $(X_i)$ be i.i.d. $\mathbb{R}$-valued random variables satisfying*

$$\varphi(t) = E[e^{tX_1}] < \infty \ \forall t \in \mathbb{R}$$

*Let $S_n = \sum_{i-1}^{n} X_i$. Then, for all $a > E[X_1]$,*

$$\lim_{n\to\infty} \frac{1}{n} \log P[S_n \geq an] = -I(a)$$

*where*

$$I(z) = \sup_{t\in\mathbb{R}} [zt - \log\varphi(t)]$$

---

*Date*: August 14, 2008.

This work was done during the Summer 2008 REU program in Mathematics at Oregon State University.

**Definition of Polya's Urn 2.1.** [RS]

*Suppose that an urn initially contains n red and m blue balls or the initial amounts $B_0, R_0$ of blue and red. At each stage a ball is randomly chosen, its color is noted, and it is then replaced along with another ball of the same color.*

**DeMoivre-Laplace Theorem 2.1.** [RS]

*If $S_n$ denotes the number of successes that occur when n independent trials, each resulting in a success with probability p, are performed then, for any $a < b$,*

$$P\left(a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right) \to \Phi(b) - \Phi(a)$$

**Beta Function 2.1.** [RS]

$$\beta(x,y) = \int_0^1 p^{x-1}(1-p)^{y-1}dp$$

**Beta Distribution 2.1.** [RS]

$$f(p) = \frac{1}{\beta(x,y)} p^{x-1}(1-p)^{y-1}$$

**Gamma-Beta Relationships 2.1.** [RS]

$$\Gamma(x+1) = x\,\Gamma(x)$$

$$\beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

**Sterling's Formula 2.1.** [FW]

$$\Gamma(z) \approx \sqrt{\frac{2\pi}{z}} \left(\frac{z}{e}\right)^z \left(1 + O\left(\frac{1}{z}\right)\right)$$

## 3. Beta Distribution and the Urn's Distribution

The distribution of the urn can be thought of the probabilities of reaching a certain proportion of red balls in the urn. You can calculate the distribution by hand by creating the tree of different possibilities. For example given the urn starts with 1 red and 1 blue ball our urn after one draw will either have 2 red and 1 blue or 1 red and 2 blue. So the probability that our proportion of red is 1/3 is 1/2 and the probability of the proportion being 2/3 is also 1/2. This is how the distribution is built.

The following section of graphs are the cumulative distribution graphs of the beta distribution combined with the cumulative distribution of the empirical distribution for the urn from left to right n=5,10,15. These graphs were generated in Mathematica 6.

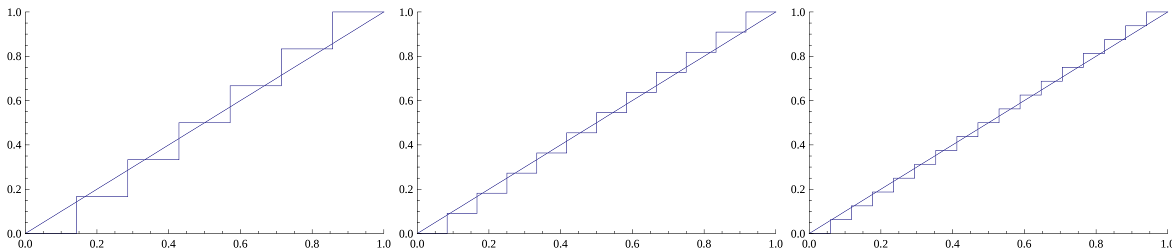


Figure 1. $R_0 = 1, B_0 = 1$

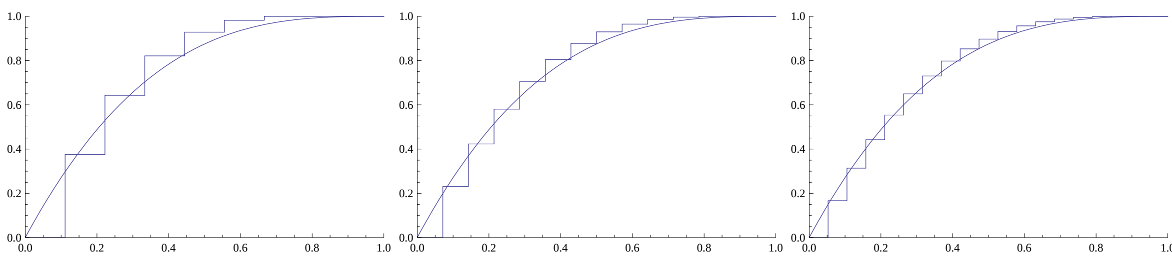


Figure 2. $R_0 = 1, B_0 = 3$

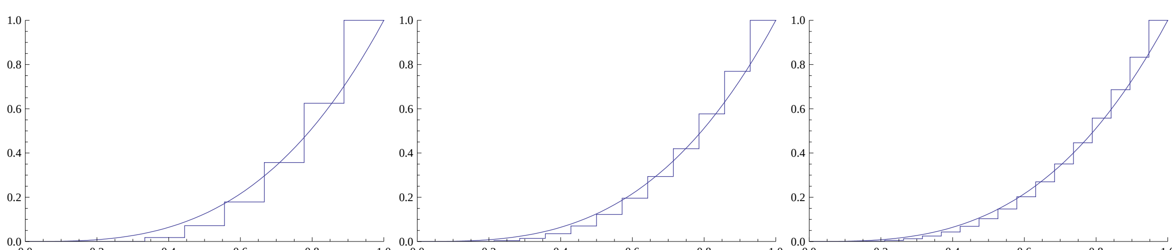


Figure 3. $R_0 = 3, B_0 = 1$

Now we can see the motivation for both showing this structure converges to the beta distribution and also finding the rate at which this occurs.

## 4. Convergence of Urn

Here we have that the probabilities of different rearrangements of draws from the urn as equal. We see this easily by comparing two different orders of a drawing k red and q blue.

$$P(\bullet\bullet\circ\bullet\circ\bullet) = P(\circ\circ\bullet\bullet\bullet) = \frac{(R_0)(R_0+1)...(R_0+k-1)(B_0)(B_0+1)...(B_0+q-1)}{(R_0+B_0)(R_0+B_0+1)...(R_0+B_0+n-1)}$$

Now we want to look at the probability of outcomes of the urn in terms of choosing k reds and the total probability of this which is found by multiplying by the binomial coefficient.

$$P(\text{k reds in n trials}) = \binom{n}{k} P(\underbrace{\bullet\bullet\cdots\bullet}_{k}\ \underbrace{\circ\circ\cdots\circ}_{n-k})$$

Now we are trying to say something about the long term behavior of this urn so we should define something that gives an idea of the urn at a known time step.

$$\rho_n = \frac{R_n}{R_n+B_n} \text{ then, } \rho_\infty = \lim_{n\to\infty} \frac{R_n}{R_n+B_n}$$

Now if we condition on a certain long term proportion of red balls in the urn then we can write the probability of drawing k reds in n trials in a new way.

$$P(\text{k reds in n trials}|\ \rho_\infty = p) = \binom{n}{k} p^k (1-p)^{n-k}$$

The term $\rho_\infty$ is a random variable and has a density function $f(p)$ and a cumulative distribution function $F(p)$.

If we want to find the probability of k reds in n trials without conditioning on a certain end proportion we can integrate over all p with the density function. Since p is a proportion then $0 \leq p \leq 1$.

$$P(\text{k reds in n trials}) = \int_0^1 P(\text{k reds in n} \mid \rho_\infty = p) f(p)dp = \binom{n}{k} \int_0^1 p^k (1-p)^{n-k} f(p)dp \tag{1}$$

For the next step of the derivation we need to examine this probability in a different way. Noticing there is a factorial type structure making up the probability of choosing k red in n trials we should try to simplify our expression using gamma functions.

$$P(\text{k reds in n trials}) = \binom{n}{k} P(\underbrace{\bullet\bullet\bullet}_{k}\ \underbrace{\circ\circ\cdots\circ}_{n-k})$$

$$= \binom{n}{k} \frac{R_0(R_0+1)\dots(R_0+k-1)B_0(B_0+1)\cdots(B_0+n-k-1)}{(R_0+B_0)(R_0+B_0+1)\cdots(R_0+B_0+n-1)}$$

$$= \binom{n}{k} \frac{\Gamma(R_0+k)\ \Gamma(B_0+n-k)\ \Gamma(R_0+B_0)}{\Gamma(R_0)\ \Gamma(B_0)\ \Gamma(R_0+B_0+n)}$$

$$= \binom{n}{k} \beta(R_0+k, B_0+n-k)\ \frac{1}{\beta(R_0,B_0)} \tag{2}$$

Thus setting equation (1) and equation (2) equal we have an interesting relationship for our probability of k red in n trials. Which holds for all $0 \leq k \leq n$.

$$\int_0^1 p^k(1-p)^{n-k} f(p)dp = \frac{\beta(R_0+k, B_0+n-k)}{\beta(R_0,B_0)} \tag{3}$$

$$\int_0^1 p^k(1-p)^{n-k} f(p)dp = \frac{1}{\beta(R_0,B_0)} \int_0^1 p^{R_0+k-1}(1-p)^{B_0+n-k-1}dp$$

$$\int_0^1 p^k(1-p)^{n-k} f(p)dp = \int_0^1 p^k(1-p)^{n-k} \frac{p^{R_0-1}(1-p)^{B_0-1}}{\beta(R_0,B_0)}dp$$

Since (3) is true for all $k \leq n$ then we can see that we have an expression for $f(p)$ which is the same for the beta distribution.

$$f(p) = \frac{1}{\beta(R_0,B_0)} p^{R_0-1}(1-p)^{B_0-1} \tag{4}$$

Now we can write $F_{\rho_\infty}(a)$ and $F_{\rho_n}(a)$ where $0 \leq a \leq 1$.

$$F_{\rho_\infty}(a) = P(\rho_\infty \leq a) = \frac{1}{\beta(R_0,B_0)} \int_0^a p^{R_0-1}(1-p)^{B_0-1}dp$$

$$F_{\rho_n}(a) = P(\rho_n \leq a) = P\left(\frac{R_n}{R_n+B_n} \leq a\right) = P\left(\frac{R_n}{R_0+B_0+n} \leq a\right) = P(R_n \leq a(R_0+B_0+n))$$

Where, $R_0 + k \leq a(R_0+B_0+n)$ and thus $k \leq (a-1)R_0 + aB_0 + an$.

$$= \sum_{k \leq (a-1)R_0+aB_0+an} P(\text{k reds in n trials}) = \sum_{k \leq (a-1)R_0+aB_0+an} \binom{n}{k} P(\underbrace{\bullet\bullet\bullet}_{k}\ \underbrace{\circ\circ\cdots\circ}_{n-k})$$

$$= \sum_{k \leq (a-1)R_0+aB_0+an} \binom{n}{k} \int_0^1 p^k(1-p)^{n-k} f(p)dp$$

Using Fubini's Theorem we can rewrite this with the summation inside the integral.

$$F_{\rho_n}(a) = \int_0^1 \sum_{k \leq (a-1)R_0+aB_0+an} \binom{n}{k} p^k(1-p)^{n-k} f(p)dp \tag{5}$$

Let $S_n$ be Binomial $(n,p)$ then $P(X_n = k) = \binom{n}{k} p^k(1-p)^{n-k}$. This gives us some idea of how to rewrite the integral so we can show it goes to the beta distribution. Also it is worth noting that a Binomial random variable is the sum of n bernoulli random variables.

$$= \int_0^1 \underbrace{P(S_n \leq ((a-1)R_0 + aB_0 + an)}_{\sum_{k \leq (a-1)R_0+aB_0+an} P(X_n = k)} f(p)dp$$

For the next step we will center our probability within the integral at 0 and standardize it so we can use the DeMoivre-Laplace Theorem.

$$P(S_n \leq ((a-1)R_0 + aB_0 + an) = P(S_n - np \leq (a-1)R_0 + aB_0 + an - np)$$

$$(6)\qquad = P\left(\frac{S_n - np}{\sqrt{np(1-p)}} \leq \frac{(a-1)R_0 + aB_0}{\sqrt{np(1-p)}} + \frac{(a-p)n}{\sqrt{np(1-p)}}\right) \to \begin{cases} P(\mathbb{Z} \leq -\infty) = 0 & p > a \\ P(\mathbb{Z} \leq \infty) = 1 & p < a \end{cases}$$

$$\text{Now we can define the RHS as } = \mathcal{X}_{(-\infty,a)}(p) = \begin{cases} 1 & p \leq a \\ 0 & p > a \end{cases}$$

$$F_{\rho_n}(a) \to \int_0^1 \mathcal{X}_{(-\infty,a)}(p) f(p)dp = \int_0^a \mathcal{X}_{(-\infty,a)}(p) f(p)dp + \int_a^1 \mathcal{X}_{(-\infty,a)}(p) f(p)dp = \int_0^a f(p)fp$$

Which is the c.d.f of the beta distribution.

## 5. A Rate From Beery-Esseen For $\|F_{\rho_n} - F_{\rho_\infty}\|_{L_1}$

In this section we will derive a rate of the convergence for the urn's distribution to the beta distribution using the Beery-Esseen Theorem.

$$|F_{\rho_n} - F_{\rho_\infty}| = |\int_0^1 P(S_n \leq ((a-1)R_0 + aB_0 + an) f(p)dp - \int_0^a f(p)dp\ |$$

$$= |\int_0^1 P(S_n \leq ((a-1)R_0 + aB_0 + an) f(p)dp - \int_0^1 \mathcal{X}_{(-\infty,a)}(p) f(p)dp\ |$$

$$\leq \int_0^1 |P(S_n \leq ((a-1)R_0 + aB_0 + an) - \mathcal{X}_{(-\infty,a)}(p)|\ f(p)dp$$

Where $\mathcal{X}_{(-\infty,a)}(p)$ is equivalent to the cdf of the Normal so write $\mathcal{X}_{(-\infty,a)}(p) = \mathcal{N}[((a-1)R_0 + aB_0 + an)]$. Also $S_n$ is a Binomial random variable which is a sum of n Bernoulli trials. We want to change our random variable to work with Beery-Esseen so we subtract by $np$ and divide by $\sigma\sqrt{n}$.

$$= \int_0^1 |P(\frac{S_n - np}{\sigma\sqrt{n}} \leq \frac{((a-1)R_0 + aB_0 + an) - np}{\sigma\sqrt{n}} - \mathcal{N}(\frac{((a-1)R_0 + aB_0 + an) - np}{\sigma\sqrt{n}})|\ f(p)dp$$

$$(7)\qquad \text{Then, using Berry-Esseen } \leq \int_0^1 \frac{3 * p^*}{\sqrt{n}\sigma^3} f(p)dp = \Psi(n) \text{ or our rate function}$$

Where $p^* = E[|\frac{X_i - np}{\sigma\sqrt{n}}|^3] = E[\frac{X_i - np}{\sigma\sqrt{n}}^3] = \left(\frac{1-p}{\sigma\sqrt{n}}\right)^3 p + \left(\frac{0-p}{\sigma\sqrt{n}}\right)^3 (1-p) = \frac{p(1-p)(1-2p)}{\sigma^3 n^{3/2}}$

Writing (7) out fully we see that this is a nice integral to work with and the rest follows easily.

$$\frac{3}{n^2\beta(R_0,B_0)} \int_0^1 \frac{(p)(1-p)(1-2p)}{\sigma^3} f(p)dp = \frac{3}{n^2\beta(R_0,B_0)} \left(\int_0^1 \frac{p^{R_0}(1-p)^{B_0+1}dp}{(p(1-p))^3} - \int_0^1 \frac{p^{R_0+1}(1-p)^{B_0}dp}{(p(1-p))^3}\right)$$

$$= \frac{3}{n^2\beta(R_0,B_0)} (\beta(R_0-2, B_0-1) - \beta(R_0-1, B_0-2))$$

$$= \frac{3}{n^2}\left(\frac{\Gamma(R_0+B_0)}{\Gamma(R_0)\Gamma(B_0)}\right)\left(\frac{\Gamma(R_0-2)\Gamma(B_0-1)}{\Gamma(R_0+B_0-3)} - \frac{\Gamma(R_0-1)\Gamma(B_0-2)}{\Gamma(R_0+B_0-3)}\right)$$

$$= \frac{3}{n^2}\left(\frac{(R_0+B_0-1)(R_0+B_0-2)(R_0+B_0-3)}{(R_0-1)(R_0-2)(B_0-1)} - \frac{(R_0+B_0-1)(R_0+B_0-2)(R_0+B_0-3)}{(R_0-1)(B_0-1)(B_0-2)}\right)$$

$$= \frac{3}{n^2}\frac{((R_0+B_0-1)(R_0+B_0-2)(R_0+B_0-3))(B_0-R_0)}{((R_0-1)(R_0-2)(B_0-1)(B_0-2)}$$

## 6. A Better Rate Using Cramér's Theorem

From the previous derivation of the cumulative distribution function we have:

$$F_{\rho_n}(a) = \int_0^1 P(X_n \leq ((a-1)R_0 + aB_0 + an)f(p)dp$$

Again if we notice that the random variable here is a Binomial random variable which is the sum of n Bernoulli random variables we can rewrite this integral in a similar form that is of the same order due to the fact that the difference between this integral and the new integral is minor as n ranges to infinity.

$$\text{So we have, } F_{\rho_n}(a) \sim \int_0^1 P(X_n \leq an)f(p)dp$$

Now consider the difference between the previous integral and the CDF of the beta distribution.

$$|F_{\rho_\infty}(a) - F_{\rho_n}(a)| \cong |\int_0^a f(p)dp - \int_0^1 P(S_n \leq an)f(p)dp|$$

$$= |\int_0^a f(p)dp - \int_0^a P(S_n \leq an)f(p)dp - \int_a^1 P(S_n \leq an)f(p)dp|$$

$$= |\int_0^a (1 - P(s_n \leq an))f(p)dp - \int_a^1 P(S_n \leq an)f(p)dp|$$

$$= |\int_0^a P(S_n \geq an)f(p)dp - \int_a^1 P(-S_n \geq -an)f(p)dp|$$

Notice from Cramérs Theorem that $\lim_{n\to\infty} \frac{1}{n}\log P[S_n \geq an] = -I(a)$ thus $P(S_n \geq an) \cong e^{-nI(a)}$. Continuing from the previous step we have the following expression where $I(a) = \sup_{t\in\mathbb{R}}[at - \log\varphi(t)]$.

(8)

$$|\int_0^a P(S_n \geq an)f(p)dp - \int_a^1 P(-S_n \geq -an)f(p)dp| \sim |\int_0^a e^{-nI(a)}f(p)dp - \int_a^1 e^{-nI(-a)}f(p)dp|$$

Thus $\varphi(t) = pe^t + 1 - p$ and to get I(a) solve for the supremum.

$$0 = \frac{d}{dt}\left[at - \log(pe^t + 1 - p)\right] = a - \frac{pe^t}{pe^t + 1 - p}, \text{ Solve for } t = \log\frac{a(1-p)}{(1-a)p}$$

$$I(a) = a\log\left(\frac{a(1-p)}{(1-a)p}\right) - \log\left(\frac{a(1-p)}{(1-a)} + p - 1\right)$$

$$= a\log a + (1-a)\log(1-a) + (a-1)\log(1-p) - a\log p$$

For the second integral: $\varphi(t) = pe^{-t} + 1 - p$ and to get I(a) solve for the supremum which turns out to be the same as that for $I(a)$. Writing the full difference of integrals from (8) we have the following.

$$
\begin{aligned}
&|\int_0^a exp[-n\left(a\log\left(\frac{a(1-p)}{(1-a)p}\right) - \log\left(\frac{a(1-p)}{(1-a)} + 1 - p\right)\right)]f(p)dp \\
&-\int_a^1 exp[-n\left(a\log\left(\frac{a(1-p)}{(1-a)p}\right) - \log\left(\frac{a(1-p)}{(1-a)} + 1 - p\right)\right)]f(p)dp| \\
&\leq \int_0^1 exp\left[-n\left(\log\left(\frac{a(1-p)}{(1-a)p}\right)^a - \log\left(\frac{1-p}{1-a}\right)\right)\right]f(p)dp \\
&= \int_0^1 e^{-n(a\log a+(1-a)\log(1-a)}e^{-n((a-1)\log(1-p)-a\log p)}f(p)dp \\
&= e^{-n(a\log a+(1-a)\log(1-a))}\int_0^1 (1-p)^{B_0-n(a-1)-1}p^{R_0+an-1}\frac{1}{\beta(R_0,B_0)}dp \\
&= e^{-n(a\log a+(1-a)\log(1-a))}\frac{\beta(R_0+an, B_0+n(1-a))}{\beta(R_0,B_0)}
\end{aligned}
$$

$$
= \frac{e^{-n(a\log a+(1-a)\log(1-a))}\Gamma(R_0+na)\Gamma(B_0+n(1-a))}{\beta(R_0,B_0)\Gamma(R_0+B_0+n)} \tag{9}
$$

Equation (9) is a good expression to study how the urn goes to the beta distribution. First we will rewrite it in terms of the Sterling formula to approximate the gamma function. Since this expression is so large it will be studied in separate terms. The following three equations are approximations of the gamma functions in equation (9).

$$
\begin{aligned}
\Gamma(R_0+na) &\approx \sqrt{\frac{2\pi}{R_0+na}}\left(\frac{R_0+na}{e}\right)^{R_0+na}(1+O(\frac{1}{n}) \\
\Gamma(B_0+n(1-a)) &\approx \sqrt{\frac{2\pi}{B_0+n(1-a)}}\left(\frac{B_0+n(1-a)}{e}\right)^{B_0+n(1-a)}(1+O(\frac{1}{n})) \\
\Gamma(R_0+B_0+n) &\approx \sqrt{\frac{2\pi}{R_0+B_0+n}}\left(\frac{R_0+B_0+n}{e}\right)^{R_0+B_0+n}(1+O(\frac{1}{n}))
\end{aligned}
$$

Rewriting terms with the approximations and then to base $e$ yields the equation we need to break apart and study.

$$(9) \approx exp[-n(a\log a+(1-a)\log(1-a))+(R_0+na)\log(R_0+na)+(B_0+n(1-a))\log(B_0+n(1-a))$$

$$-(R_0+B_0+n)\log(R_0+B_0+n)](1+O(\frac{1}{n}))\sqrt{\frac{2\pi}{R_0+na}}\sqrt{\frac{2\pi}{B_0+n(1-a)}}\sqrt{\frac{2\pi}{R_0+B_0+n}}^{-1}\beta(R_0,B_0)^{-1}$$

(10) The terms of square roots. $L_{\sqrt{2\pi}} = \dfrac{\sqrt{\frac{2\pi}{B_0+n(1-a)}}\sqrt{\frac{2\pi}{R_0+na)}}}{\sqrt{\frac{2\pi}{R_0+B_0+n}}} \cong \dfrac{\sqrt{2\pi}}{\sqrt{n}(a-a^2)}$

(11)

$exp[(R_0+na)\log(R_0+na)+(B_0+n(1-a))\log(B_0+n(1-a))-(R_0+B_0+n)\log(R_0+B_0+n)]$

$$= \overbrace{exp[(R_0)\log(R_0+na)+(B_0)\log(B_0+n(1-a))-(R_0+B_0)\log(R_0+B_0+n)]}^{\#1}$$
$$\cdot \underbrace{exp[(na)\log(R_0+na)+(n(1-a))\log(B_0+n(1-a))-(n)\log(R_0+B_0+n)]}_{\#2}$$

The following is the rewriting of #2.

$$exp(na\log(na)+n(1-a)\log(n(1-a))-n\log n$$
$$+\log\left(1+\frac{B_0}{(n(1-a))}\right)^{n(1-a)}+\log\left(1+\frac{R_0}{na}\right)^{na}-\log\left(1+\frac{R_0+B_0}{n}\right)^{n})$$

The expression with the logarithms to powers is familiar and is useful to examine it's limit.

$$L_1 = \log\left(1+\frac{B_0}{(n(1-a))}\right)^{n(1-a)}+\log\left(1+\frac{R_0}{na}\right)^{na}-\log\left(1+\frac{R_0+B_0}{n}\right)^{n} \rightarrow \frac{B_0+R_0}{R_0+B_0} = 1$$

If we combine what is left from expression #2 with $exp[-n(a\log a+(1-a)\log(1-a))]$ from (9) the following results.

$$exp(na\log(na)+n(1-a)\log(n(1-a))-n\log n-na\log a-n(1-a)\log(1-a)))$$

=

$$exp[na\log n+n(1-a)\log n-n\log n] = \frac{n^{na}n^{n(1-a)}}{n^n} = 1$$

Rewriting the equation (9) with the simplifications we have and labeling the functions that we know converge to a specific limit we have.

$$(12) \qquad \frac{1}{\beta(R_0,B_0)}L_1L_{\sqrt{2\pi}}\left(1+O\left(\frac{1}{n}\right)\right) \cong \frac{1}{\beta(R_0,B_0)}\frac{\sqrt{2\pi}}{\sqrt{n}(a-a^2)}\left(1+O\left(\frac{1}{n}\right)\right)$$

## 7. Conclusion

The first rate I derived seems to be incorrect when I examine numerical examples of the distance between the urn's distribution and the actual. The second rate does however seem to match the numerical evidence. I believe one could derive a much better rate if they incorporated the relationship between $R_0$ and $B_0$ into their investigation. The reason for this is the great difference between distributions where $B_0 \geq R_0$ and the distributions where $B_0 \leq R_0$.

I would like to personally thank the advisors of the REU which are Nathan Gibson, Yevgeniy Kovchegov, Donald Solmon and Paul Cull, the organizer Dennis Garrity, and Oregon State University. I also would like to thank the NSF for their gracious funding. I had a wonderful time and enjoyed the hospitality that I was privileged to experience.

## References

[DR] Durrett, Richard. *Probability: Theory and Examples Second Edition*
[FW] Feller, William. *An Introduction to Probability Theory and Its Application*
[HF] Hollander, Frank den. *Large Deviations (Fields Institute Monographs)*
[RS] Ross, Sheldon. *A First Course in Probability Seventh Edition*

## 8. Appendix

The following is Mathematica 6 code used to generate the empirical distributions used for the diagrams in this paper.

```
TheoreticalDist[Red0_, Blue0_, DrawN_] :=

  Module[

   {RedInitial, BlueInitial, DepthN},

   RedInitial = Red0;
   BlueInitial = Blue0;
   DepthN = DrawN;

   NewTable = {{RedInitial, BlueInitial, 1, 1}};

   For
    [
    MainI = 0, MainI < DepthN, MainI++, OldTable = NewTable;

    LoopIterations = Length[NewTable];
    NewTable = Array[1 &, {2*Length[OldTable], 4}];

    For
     [
     i = 1, i <= LoopIterations, i++,
     (*Red*)
     NewTable[[(2*i - 1), 1]] = OldTable[[i, 1]] + 1;
     NewTable[[(2*i - 1), 2]] = OldTable[[i, 2]];
     NewTable[[(2*i - 1),
       3]] = (OldTable[[i,
          1]]/(OldTable[[i, 1]] + OldTable[[i, 2]])) // N;
     NewTable[[(2*i - 1), 4]] =
      OldTable[[i, 4]]*NewTable[[(2*i - 1), 3]];
     (*Blue*)
     NewTable[[(2*i), 1]] = OldTable[[i, 1]];
     NewTable[[(2*i), 2]] = OldTable[[i, 2]] + 1;
     NewTable[[(2*i),
       3]] = (OldTable[[i,
          2]]/(OldTable[[i, 1]] + OldTable[[i, 2]])) // N;
     NewTable[[(2*i), 4]] = OldTable[[i, 4]]*NewTable[[(2*i), 3]];
     ];
    ];
   (*Take NewTable and Build CDF Array*)
```

```
SortedTable = SortBy[NewTable, First];
NewIndex = SortedTable[[1, 1]];
CDFBuildTable = Array[0 &, {DepthN + 1, 2}];
CDFIndex = 1;
CDFBuildTable[[1, 1]] = RedInitial;

For
 [i = 1, i <= Length[SortedTable], i++,

 OldIndex = NewIndex;
 NewIndex = SortedTable[[i, 1]];

 If
  [
  OldIndex != NewIndex,

  CDFIndex = CDFIndex + 1;
  CDFBuildTable[[CDFIndex, 1]] = NewIndex;
  CDFBuildTable[[CDFIndex, 2]] = SortedTable[[i, 4]];,

  CDFBuildTable[[CDFIndex, 2]] =
    CDFBuildTable[[CDFIndex, 2]] + SortedTable[[i, 4]];

  ];
 ];
For[
 i = 1, i <= Length[CDFBuildTable], i++,
 CDFBuildTable[[i, 1]] =
  CDFBuildTable[[i, 1]]/(RedInitial + BlueInitial + DepthN) // N;
 If[i >= 2,
  CDFBuildTable[[i, 2]] =
    CDFBuildTable[[i, 2]] + CDFBuildTable[[i - 1, 2]];
  ];
 ];

CDFBuildTableFinal = Array[0 &, {2*Length[CDFBuildTable] + 2, 2}];
For[
 i = 1, i <= Length[CDFBuildTable] - 1, i++,
 If[i >= 1,
   CDFBuildTableFinal[[(2*i + 1), 1]] = CDFBuildTable[[i, 1]];
   CDFBuildTableFinal[[(2*i + 1), 2]] = CDFBuildTable[[i, 2]];

   CDFBuildTableFinal[[(2*i + 2), 1]] = CDFBuildTable[[i + 1, 1]];
```

```
      CDFBuildTableFinal[[(2*i + 2), 2]] = CDFBuildTable[[i, 2]];
      ];
    ];
   CDFBuildTableFinal[[2, 1]] = CDFBuildTable[[1, 1]];

   CDFBuildTableFinal[[Length[CDFBuildTableFinal] - 1, 1]] =
    CDFBuildTable[[Length[CDFBuildTable], 1]];
   CDFBuildTableFinal[[Length[CDFBuildTableFinal] - 1, 2]] = 1;

   CDFBuildTableFinal[[Length[CDFBuildTableFinal], 1]] = 1;
   CDFBuildTableFinal[[Length[CDFBuildTableFinal], 2]] = 1;

   (*Distance Stuff*)
   MaxAbDistance = 0;
   PositionAbDistance = 0;
   DistFunction = CDFBuildTableFinal;
   For[i = 1, i <= Length[CDFBuildTableFinal], i++,
    AbDistance =
     Abs[(CDFBuildTableFinal[[i, 2]] -
         CDF[BetaDistribution[RedInitial, BlueInitial],
          CDFBuildTableFinal[[i, 1]]]) // N];
    DistFunction[[i, 2]] = AbDistance;
    If[AbDistance > MaxAbDistance, MaxAbDistance = AbDistance;
     PositionAbDistance = DistFunction[[i, 1]];];
    ];

   ];
```

The next section of code will display the plot of the empirical distribution vs. the actual beta distribution. First execute the pervious section and then this section.

```
Manipulate[TheoreticalDist[R0, B0, n];
 Show[ListLinePlot[CDFBuildTableFinal, PlotRange -> {{0, 1}, {0, 1}},
   PlotStyle -> Directive[Thickness[Large], Red]],
  Plot[CDF[BetaDistribution[R0, B0], x], {x, 0, 1},
   PlotStyle -> Directive[Thickness[Large], Blue]]], {R0, 1, 10,
  1}, {B0, 1, 10, 1}, {n, 1, 15, 1}]
```

WINONA STATE UNIVERSITY
*E-mail address*: jdrinane4878@winona.edu