On a time-dependent Eggenberger-Pólya urn model

May-Ru Chen
National Sun Yat-sen University
Outline

1. Review
2. Multiple balls drawn
3. Time-Dependent
Pólya-Eggenberger Urn (1923)

- Suppose an urn initially contains $w$ white and $r$ red balls.
- One ball is drawn at random and then replaced together with $c$ balls of the same color. Repeat the procedure *ad infinitum*.
- Denote the added balls situation by the replacement matrix

$$M = \begin{cases} 
    (c & 0) & \text{the drawn ball is white} \\
    (0 & c) & \text{the drawn ball is red}
\end{cases}.$$
Pólya-Eggenberger Urn (1923)

- Suppose an urn initially contains $w$ white and $r$ red balls.
- One ball is drawn at random and then replaced together with $c$ balls of the same color. Repeat the procedure *ad infinitum*.
- Denote the added balls situation by the replacement matrix

$$M = \begin{cases} \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} & \text{the drawn ball is white} \\ \begin{pmatrix} 0 & c \\ c & 0 \end{pmatrix} & \text{the drawn ball is red} \end{cases}$$
After the $n$th action, let $W_n$ be the number of white balls and $T_n$ be the number of total balls. Also let $X_n = W_n / T_n$ and $\mathcal{F}_n = \sigma\{W_1, \ldots, W_n\}$. 
After the $n$th action, let $W_n$ be the number of white balls and $T_n$ be the number of total balls. Also let $X_n = W_n / T_n$ and $\mathcal{F}_n = \sigma\{W_1, \ldots, W_n\}$.

Then $T_{n+1} = T_n + c$ and $W_{n+1} \overset{(d)}{=} W_n + c\xi_{n+1}$, where $\xi_{n+1} \mid W_n \sim^d \text{Ber}(X_n)$. 

After the \( n \)th action, let \( W_n \) be the number of white balls and \( T_n \) be the number of total balls. Also let \\
\[ X_n = \frac{W_n}{T_n} \] and \\
\[ \mathcal{F}_n = \sigma\{W_1, \ldots, W_n\}. \]

Then \\
\[ T_{n+1} = T_n + c \] and \\
\[ W_{n+1} \overset{(d)}{=} W_n + c \xi_{n+1}, \] where \\
\[ \xi_{n+1} \mid W_n \sim^d \text{Ber}(X_n). \]

Thus \\
\[ X_{n+1} \overset{(d)}{=} \frac{W_n + c \xi_{n+1}}{T_{n+1}} \]
\[ = \frac{T_n}{T_{n+1}} X_n + \frac{c \xi_{n+1}}{T_{n+1}} \]
\[ = X_n + \frac{c}{T_{n+1}} (\xi_{n+1} - X_n). \]
Note that $E[\xi_{n+1} \mid \mathcal{F}_n] = E[\xi_{n+1} \mid \mathcal{W}_n] = X_n$ a.s.
Note that $E[\xi_{n+1} \mid \mathcal{F}_n] = E[\xi_{n+1} \mid \mathcal{W}_n] = X_n$ a.s.

Then

$$E[X_{n+1} \mid \mathcal{F}_n] = X_n + \frac{c}{T_{n+1}}(E[\xi_{n+1} \mid \mathcal{F}_n] - X_n) = X_n$$
• Note that $E[\xi_{n+1} \mid \mathcal{F}_n] = E[\xi_{n+1} \mid \mathcal{W}_n] = X_n$ a.s.

• Then

$$E[X_{n+1} \mid \mathcal{F}_n] = X_n + \frac{c}{T_{n+1}} (E[\xi_{n+1} \mid \mathcal{F}_n] - X_n) = X_n$$

• Hence $\{X_n\}$ is a bounded martingale.
Note that \( E[\xi_{n+1} \mid \mathcal{F}_n] = E[\xi_{n+1} \mid \mathcal{W}_n] = X_n \) a.s.

Then

\[
E[X_{n+1} \mid \mathcal{F}_n] = X_n + \frac{c}{T_{n+1}}(E[\xi_{n+1} \mid \mathcal{F}_n] - X_n) = X_n
\]

Hence \( \{X_n\} \) is a bounded martingale.

By martingale convergence theorem, \( \{X_n\} \) converges almost surely. Furthermore, the distribution of \( \lim_{n \to \infty} X_n \) follows a beta distribution with parameters \( b/c \) and \( r/c \).
Pemantle’s urn (1989)

- Suppose an urn initially contains \( w \) white and \( r \) red balls.
- At time \( n \), one ball is drawn at random and then replaced together with \( c_n \) balls of the same color, where \( c_n \) is a positive integer. Repeat the procedure \textit{ad infinitum}.
- Then after the \( n \)th drawn, the replacement matrix

\[
M_n = \begin{pmatrix} c_n & 0 \\ 0 & c_n \end{pmatrix}.
\]
After the $n$th action, let $W_n$, $T_n$ and $X_n$ defined as before.
After the $n$th action, let $W_n$, $T_n$ and $X_n$ defined as before.

Then $T_{n+1} = T_n + c_{n+1}$ and $W_{n+1} \overset{(d)}{=} W_n + c_{n+1} \xi_{n+1}$, where $\xi_{n+1} \mid W_n \sim^d \text{Ber}(X_n)$.
• After the $n$th action, let $W_n$, $T_n$ and $X_n$ defined as before.

• Then $T_{n+1} = T_n + c_{n+1}$ and $W_{n+1} \overset{(d)}{=} W_n + c_{n+1} \xi_{n+1}$, where $\xi_{n+1} \mid W_n \sim^d \text{Ber}(X_n)$.

• Thus

$$X_{n+1} \overset{(d)}{=} \frac{T_n}{T_{n+1}} X_n + \frac{c_{n+1} \xi_{n+1}}{T_{n+1}}$$

$$= X_n + \frac{c_{n+1}}{T_{n+1}} (\xi_{n+1} - X_n).$$

and so $E[X_{n+1} \mid F_n] = X_n$. 
After the $n$th action, let $W_n$, $T_n$ and $X_n$ defined as before.

Then $T_{n+1} = T_n + c_{n+1}$ and $W_{n+1} \overset{(d)}{=} W_n + c_{n+1} \xi_{n+1}$, where $\xi_{n+1} \mid W_n \sim^d \text{Ber}(X_n)$.

Thus

$$X_{n+1} \overset{(d)}{=} \frac{T_n}{T_{n+1}} X_n + \frac{c_{n+1} \xi_{n+1}}{T_{n+1}}$$

$$= X_n + \frac{c_{n+1}}{T_{n+1}} (\xi_{n+1} - X_n).$$

and so $E[X_{n+1} \mid \mathcal{F}_n] = X_n$.

Hence $\{X_n\}$ is a bounded martingale and so $X_n$ converges almost surely to a random variable $X$. 
Pemantle (1989) showed that

(i) the distribution of $X$ has no atoms on $(0, 1)$;
Pemantle (1989) showed that

(i) the distribution of $X$ has no atoms on $(0, 1)$;

(ii) $\sum_{n=1}^{\infty} \left( \frac{c_n}{T_{n-1}} \right)^2 = \infty$ if and only if $X \sim^d \text{Bernoulli}\left(\frac{w}{w+r}\right)$. 
Pemantle (1989) showed that

(i) the distribution of $X$ has no atoms on $(0, 1)$;

(ii) $\sum_{n=1}^{\infty} \left( \frac{c_n}{T_{n-1}} \right)^2 = \infty$ if and only if $X \sim^d \text{Bernoulli}(\frac{w}{w+r})$.

If $\sum_{n=1}^{\infty} \left( \frac{c_n}{T_{n-1}} \right)^2 < \infty$, what is the distribution of $X$?
\[ \sum_{n=1}^{\infty} \left( \frac{c_n}{T_{n-1}} \right)^2 < \infty \]

**Figure 1.** \( c_n = n \)

**Figure 2.** \( c_n = n^2 \)
\[ \sum_{n=1}^{\infty} \left( \frac{c_n}{T_{n-1}} \right)^2 < \infty \]

**Figure 3.** \( c_n = \ln(n) \)

**Figure 4.** \( c_n = \ln(\ln(n)) \)
Example

For Pemantle’s urn, if \( w = r = 1 \) and \( c_n = n \), then the probability that all drawn are of the same color is

\[
\frac{2}{3} \times \frac{6}{7} \times \cdots > 0.
\]

Thus the probability of \( X \in \{0,1\} \) is positive.
Example

For Pemantle’s urn, if $w = r = 1$ and $c_n = n$, then the probability that all drawn are of the same color is 
\[
\frac{2}{3} \times \frac{6}{7} \times \cdots > 0.
\]
Thus the probability of $X \in \{0, 1\}$ is positive.

In 1989, Pemantle showed that if $\{c_n\}_{n\geq 1}$ is a bounded sequence, then $P(X = 0) = P(X = 1) = 0$, that is, $X$ has no atoms on $[0, 1]$. 
The urn of Johnson, Kotz, and Mahmoud (2004)

- In their model, the drawn, say \( m \geq 1 \), can be with or without replacement and the replacement matrix is

\[
\begin{pmatrix}
\text{#white balls drawn} \\
m \\
m - 1 \\
\vdots \\
1 \\
0
\end{pmatrix}
\begin{pmatrix}
-(m - 1) & m \\
-(m - 2) & m - 1 \\
\vdots & \vdots & \vdots \\
0 & 1 \\
1 & 0 \\
0 & 1 & 0
\end{pmatrix}.
\]
They gave an recursion formula for the distribution of white balls.

They also gave the expectation and the variance of the number of white balls.
Suppose an urn initially contains \( w \) white and \( r \) red balls. Chen-Wei considered that at each step, \( m \geq 1 \) balls are randomly drawn and then note their colors, say \( k \) white and \( m - k \) red balls. Replace the drawn balls together with \( ck \) white and \( c(m - k) \) red balls. Repeat the procedure \textit{ad infinitum}. 

---

\textbf{Chen-Wei Urn (2005)}
The replacement matrix is $M = \begin{pmatrix}
cm & 0 \\
c(m - 1) & c \\
\vdots & \vdots \\
c & c(m - 1) \\
0 & cm
\end{pmatrix}$. 
After the $n$th action, let $W_n$ be the number of white balls and $T_n$ be the number of total balls in the urn. Also let $X_n = W_n / T_n$. 
After the $n$th action, let $W_n$ be the number of white balls and $T_n$ be the number of total balls in the urn. Also let $X_n = W_n / T_n$.

Then $T_{n+1} = T_n + cm$ and $W_{n+1} \overset{d}{=} W_n + c\xi_{n+1}$, where $\xi_{n+1} \mid W_n \sim^d \text{Hypgeo}(W_n, T_n - W_n, m)$, that is,

$$\Pr\{\xi_{n+1} = k \mid W_n\} = \frac{\binom{W_n}{k}\binom{T_n-W_n}{m-k}}{\binom{T_n}{m}} = \frac{\binom{T_nX_n}{k}\binom{T_n(1-X_n)}{m-k}}{\binom{T_n}{m}},$$

where $0 \leq k \leq m$.
Note that $E[\xi_{n+1} \mid \mathcal{F}_n] = \sum_{k=0}^{m} \frac{k ( \binom{T_n X_n}{k} \binom{T_n (1-X_n)}{m-k})}{\binom{T_n}{m}} = mX_n$. 

Hence $\{X_n\}$ is a bounded martingale. Furthermore, as $n \to \infty$, $X_n$ converges almost surely to an absolutely continuous random variable.
Note that \( E[\xi_{n+1} \mid F_n] = \sum_{k=0}^{m} \frac{k(T_nX_n)(T_n(1-X_n))}{m-k} = mX_n. \)

Then

\[
X_{n+1} = \frac{T_n}{T_{n+1}} X_n + \frac{c}{T_{n+1}} \xi_{n+1}
\]

\[
= X_n + \frac{c}{T_{n+1}}(\xi_{n+1} - mX_n)
\]

and so \( E[X_{n+1} \mid F_n] = X_n. \)
Note that $E[\xi_{n+1} \mid \mathcal{F}_n] = \sum_{k=0}^{m} \frac{k(T_n X_n)(T_n(1-X_n))}{(T_n)} = mX_n$.

Then

$$X_{n+1} = \frac{T_n}{T_{n+1}} X_n + \frac{c}{T_{n+1}} \xi_{n+1}$$

$$= X_n + \frac{c}{T_{n+1}} (\xi_{n+1} - mX_n)$$

and so $E[X_{n+1} \mid \mathcal{F}_n] = X_n$.

Hence $\{X_n\}$ is a bounded martingale.
• Note that \( E[\xi_{n+1} \mid \mathcal{F}_n] = \sum_{k=0}^{m} \frac{k(T_nX_n)(T_n(1-X_n))}{(T_n)^m} = mX_n. \)

• Then

\[
X_{n+1} = \frac{T_n}{T_{n+1}}X_n + \frac{c}{T_{n+1}}\xi_{n+1}
\]

\[
= X_n + \frac{c}{T_{n+1}}(\xi_{n+1} - mX_n)
\]

and so \( E[X_{n+1} \mid \mathcal{F}_n] = X_n. \)

• Hence \( \{X_n\} \) is a bounded martingale.

• Furthermore, as \( n \to \infty \), \( X_n \) converges almost surely to an absolutely continuous random variable.
The urn of Aoudia and Perron (2012)

- Aoudia and Perron (2012) proposed a new model which at time $n$, $M_n$ balls are sampled and a multiple of $C_n$ of the drawn balls are added, where $M_n$ and $C_n$ are random variables.

- They showed that $\{X_n\}$ is a bounded martingale and converges almost surely.

- They also showed that $X \sim^d \text{Bernoulli}\left(\frac{w}{w+r}\right)$ if and only if

$$
\sum_{n=1}^{\infty} E\left[ \frac{C_{n+1}^2 M_{n+1} X_n (1-X_n)(T_n-M_{n+1})}{T_{n+1}^2 (T_n-1)} \right] = \frac{wr}{(w+r)^2}.
$$
Assume an urn initially contains \( w \) white and \( r \) red balls.

After the \( n \)th adding balls, suppose \( m \geq 1 \) balls are randomly drawn and then note their colors, say \( k \) white and \( m - k \) red balls.

Replace the drawn balls together with \( c_{n+1}k \) white and \( c_{n+1}(m - k) \) red balls. Repeat the procedure \textit{ad infinitum}.

If \( m = 1 \), then the above model is Pemantle's urn.

If \( c_1 = c_2 = \cdots = c \), then the above model is Chen-Wei urn.
The replacement matrix at time $n$ is

$$M_n = \begin{pmatrix}
mc_n & 0 \\
(m-1)c_n & c_n \\
\vdots & \vdots \\
c_n & (m-1)c_n \\
0 & mc_n
\end{pmatrix}.$$
After the $n$th action, let $W_n$, $T_n$ and $X_n$ defined as before.

Then $\{X_n\}$ is a bounded martingale and so $X_n$ converges almost surely to a random variable, say $X$.

Let $\rho_n = c_n/T_{n-1}$, $n \in \mathbb{N}$. 

Theorem 1. (i) If $\sum_{j=1}^{\infty} \rho_{2j} + 1 = \infty$, then $X$ follows a Bernoulli distribution with parameter $w/(w+r)$.

(ii) If $\{c_n\}_{n \geq 1}$ is a bounded sequence by $c$, then $X$ is absolutely continuous.
After the $n$th action, let $W_n$, $T_n$ and $X_n$ defined as before.
Then $\{X_n\}$ is a bounded martingale and so $X_n$ converges almost surely to a random variable, say $X$.

Let $\rho_n = c_n / T_{n-1}$, $n \in \mathbb{N}$.

**Theorem 1.**

(i) If $\sum_{j=1}^{\infty} \rho_j^2 = \infty$, then $X$ follows a Bernoulli distribution with parameter $w/(w + r)$.

(ii) If $\{c_n\}_{n \geq 1}$ is a bounded sequence by $c$, then $X$ is absolutely continuous.
Proposition 1. (Chen and Wei(2005))

Let \((\Omega, \mathcal{F}, \mathbb{P})\) be the probability space and let \((\Omega_n)_{n \geq 1}\) be a sequence of increasing events such that \(\mathbb{P}\{\bigcup_{n=1}^{\infty} \Omega_n\} = 1\).

If there exist nonnegative Borel measurable functions \((f_n)_{n \geq 1}\) such that \(\mathbb{P}(\Omega_n \cap X^{-1}(B)) = \int_B f_n(x) \, dx\) for all Borel sets \(B\), then \(f = \lim_{n \to \infty} f_n\) exists almost everywhere, and \(f\) is the density of \(X\).
Proposition 2.

For \( n \geq 1 \), let

\[
\Omega_n = \{ \omega : cm \leq W_n(\omega) \leq T_n - cm \}.
\]

Then \( \Omega_{n+1} \supset \Omega_n \) and \( P(\bigcup_{n=1}^{\infty} \Omega_n) = 1 \).
The proof of Theorem 1.(ii) with $m = 1$

- By Propositions 1 and 2, it is sufficient to show that the restriction of $X$ to $\Omega_\ell$ has a density for all positive integer $\ell \geq c$. 
The proof of Theorem 1.(ii) with $m = 1$

- By Propositions 1 and 2, it is sufficient to show that the restriction of $X$ to $\Omega_\ell$ has a density for all positive integer $\ell \geq c$.
- Recall $\rho_n = \frac{c_n}{T_{n-1}}$, $n \in \mathbb{N}$. Since $\{c_n\}_{n \geq 1}$ is bounded by $c$, $\sum_{j=1}^{\infty} \rho_j^2 < \infty$. 


The proof of Theorem 1.(ii) with \( m = 1 \)

- By Propositions 1 and 2, it is sufficient to show that the restriction of \( X \) to \( \Omega_\ell \) has a density for all positive integer \( \ell \geq c \).
- Recall \( \rho_n = c_n / T_{n-1}, \ n \in \mathbb{N} \). Since \( \{c_n\}_{n \geq 1} \) is bounded by \( c \), \( \sum_{j=1}^{\infty} \rho_j^2 < \infty \).
- For any given \( \epsilon > 0 \), choose
  \[ \delta = \epsilon / \left( T_{\ell-1} \exp \left\{ - \sum_{j=\ell}^{\infty} \rho_j^2 \right\} \right) > 0. \]
The proof of Theorem 1.(ii) with \( m = 1 \)

- By Propositions 1 and 2, it is sufficient to show that the restriction of \( X \) to \( \Omega_\ell \) has a density for all positive integer \( \ell \geq c \).
- Recall \( \rho_n = c_n / T_{n-1}, \ n \in \mathbb{N} \). Since \( \{c_n\}_{n \geq 1} \) is bounded by \( c \), \( \sum_{j=1}^{\infty} \rho_j^2 < \infty \).
- For any given \( \epsilon > 0 \), choose \( \delta = \epsilon / \left( T_{\ell-1} \exp \left\{ - \sum_{j=\ell}^{\infty} \rho_j^2 \right\} \right) > 0 \).
- Let \( x_1 < x'_1 \leq x_2 < x'_2 \leq \cdots < x_s < x'_s \) and \( \sum_{i=1}^{s}(x'_i - x_i) < \delta \).
Then by Fatou’s lemma,

$$\sum_{i=1}^{s} \Pr(\{x_i < X < x'_i\} \mid \Omega_\ell) = \sum_{i=1}^{s} E[1_{\{x_i < X < x'_i\}} \mid \Omega_\ell]$$

$$\leq \sum_{i=1}^{s} \liminf_{n \to \infty} E[1_{\{x_i < X_n < x'_i\}} \mid \Omega_\ell]$$

$$= \sum_{i=1}^{s} \liminf_{n \to \infty} P(x_i < X_n < x'_i \mid \Omega_\ell).$$
Since $X_n = W_n / T_n$,

$$
\sum_{i=1}^{s} \lim_{n \to \infty} \inf P(x_i < X_n < x_i' \mid \Omega_{\ell})
$$

$$
= \sum_{i=1}^{s} \lim_{n \to \infty} \inf P(T_n x_i < W_n < T_n x_i' \mid \Omega_{\ell})
$$

$$
= \sum_{i=1}^{s} \lim_{n \to \infty} \inf \left[ \sum_{T_n x_i < k < T_n x_i'} \Pr(W_n = k \mid \Omega_{\ell}) \right]
$$

$$
\leq \left[ \sum_{i=1}^{s} (x_i' - x_i) \right] \left[ \lim_{n \to \infty} \inf T_n \left( \max_{k} \Pr(W_n = k \mid \Omega_{\ell}) \right) \right]
$$

$$
\leq \delta \lim_{n \to \infty} \inf T_n \left( \max_{k} \Pr(W_n = k \mid \Omega_{\ell}) \right).
$$
Since \( \rho_n = c_n / T_{n-1} \),

\[
T_n = T_{n-1} + c_n = T_{n-1}(1 + \rho_n) = \cdots = T_{\ell-1} \prod_{s=\ell}^{n}(1 + \rho_s).
\]
Since $\rho_n = c_n / T_{n-1}$,

$$T_n = T_{n-1} + c_n = T_{n-1}(1 + \rho_n) = \cdots = T_{\ell-1} \prod_{s=\ell}^{n} (1 + \rho_s).$$

Thus

$$\sum_{i=1}^{s} \Pr\left( \{x_i < X < x'_i\} \mid \Omega_{\ell} \right) \leq \delta T_{\ell-1} \liminf_{n \to \infty} \left[ \prod_{j=1}^{n} (1 + \rho_j) \right] \left[ \max_{k} \Pr(W_n = k \mid \Omega_{\ell}) \right].$$
Observe that

\[
\max_k \Pr(W_n = k \mid \Omega_\ell) \\
= \max_k \{ \Pr(W_n = k \mid W_{n-1} = k) \Pr(W_{n-1} = k \mid \Omega_\ell) \\
+ \Pr(W_n = k \mid W_{n-1} = k - c_n) \Pr(W_{n-1} = k - c_n \mid \Omega_\ell) \} \\
\leq \left( 1 - \frac{k}{T_{n-1}} + \frac{k - c_n}{T_{n-1}} \right) \left( \max_k \Pr(W_{n-1} = k \mid \Omega_\ell) \right) \\
= (1 - \rho_n) \max_k \Pr(W_{n-1} = k \mid \Omega_\ell) \\
\vdots \\
\leq \prod_{j=\ell}^{n} (1 - \rho_j).
\]
Therefore,

\[
\sum_{i=1}^{s} \Pr(x_i < X < x'_i \mid \Omega_\ell)
\leq \delta T_{\ell-1} \liminf_{n \to \infty} \left[ \prod_{j=\ell}^{n} (1 + \rho_j) \right] \left[ \prod_{j=\ell}^{n} (1 - \rho_j) \right]
\leq \delta T_{\ell-1} \exp \left\{ - \sum_{j=\ell}^{\infty} \rho_j^2 \right\} \quad (\text{since } 1 - x \leq e^{-x})
\leq \epsilon.
\]

Hence by Theorem 31.7 of Billingsley (1995), the restriction of \(X\) to \(\Omega_\ell\) has a density.
References


Thanks for Your Attention
\[
\max_k \Pr(W_n = k \mid \Omega_\ell)
\]

\[
= \max_k \left\{ \sum_{i=0}^{m} \Pr(W_n = k \mid W_{n-1} = k - ic_n) \Pr(W_{n-1} = k - ic_n \mid \Omega_\ell) \right\}
\]

\[
\leq \left( \max_k \Pr(W_{n-1} = k \mid \Omega_\ell) \right) \max_k \left\{ \sum_{i=0}^{m} \frac{\binom{k-ic_n}{i} \binom{T_{n-1-k+ic_n}}{m-i}}{\binom{T_{n-1}}{m}} \right\}
\]

\[
\leq \max_k \Pr(W_{n-1} = k \mid \Omega_\ell) \left( 1 - \rho_n + \frac{\rho_n c_n (m - 1)}{T_{n-1} - 1} \right)
\]

\[
\vdots
\]

\[
\leq \prod_{j=\ell}^{n} \left( 1 - \rho_j + \frac{\rho_j c_j (m - 1)}{T_{j-1} - 1} \right).
\]