Limit distributions for Urn models with multiple drawings

Markus Kuba

Institut für Angewandte Mathematik und Naturwissenschaften



Joint work with Hosam Mahmoud; Basile Morcrette

AofA'15

Strobl, Austria - 11.06.2015



2 Urn models with multiple drawings - linear affine class



Analysis using Analytic Combinatorics

Urn models

- Introduction

Urn models

- Introduction

Pólya-Eggenberger urns: Setup

Urn contains W = n white and B = m black balls.

- Every discrete time steps a ball is drawn at random: $p_{\text{white}} = \frac{n}{n+m}, p_{\text{black}} = \frac{m}{n+m}.$
- Color inspection:

White - α white and b black balls are added/removed; Black - c white and d black balls are added/removed; α , b, c, d $\in \mathbb{Z}$.

2 × 2 ball replacement matrix.

$$\mathsf{M} = \begin{pmatrix} \mathsf{a} & \mathsf{b} \\ \mathsf{c} & \mathsf{d} \end{pmatrix}.$$

- Every discrete time steps a ball is drawn at random: $p_{\text{white}} = \frac{n}{n+m}, p_{\text{black}} = \frac{m}{n+m}.$
- Color inspection:

White - α white and b black balls are added/removed; Black - c white and d black balls are added/removed; α , b, c, d $\in \mathbb{Z}$.

• 2 × 2 ball replacement matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

- Every discrete time steps a ball is drawn at random: $p_{\text{white}} = \frac{n}{n+m}, p_{\text{black}} = \frac{m}{n+m}.$
- Color inspection:

White - a white and b black balls are added/removed; Black - c white and d black balls are added/removed; $a, b, c, d \in \mathbb{Z}$.

2 × 2 ball replacement matrix

$$\mathbf{l} = \begin{pmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{pmatrix}.$$

- Every discrete time steps a ball is drawn at random: $p_{\text{white}} = \frac{n}{n+m}, p_{\text{black}} = \frac{m}{n+m}.$
- Color inspection:

White - a white and b black balls are added/removed; Black - c white and d black balls are added/removed; $a, b, c, d \in \mathbb{Z}$.

• 2×2 ball replacement matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

- Every discrete time steps a ball is drawn at random: $p_{\text{white}} = \frac{n}{n+m}, p_{\text{black}} = \frac{m}{n+m}.$
- Color inspection:

White - a white and b black balls are added/removed; Black - c white and d black balls are added/removed; $a, b, c, d \in \mathbb{Z}$.

• 2 × 2 ball replacement matrix

$$M = \begin{cases} W & B \\ \{W\} & \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \end{cases}$$

Pólya urn:
$$M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
.

Beta limit law; Normal limit for martingale tail sum Hall and Heyde; **Grübel**.

Triangular urns: $M = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$

with $a, d, \in \mathbb{N}$ and $b \in \mathbb{N}_0$.

Generalized Mittag-Leffler limit law Janson; **Goldschmidt; Mahmoud**.

Pólya urn:
$$M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
.

Beta limit law; Normal limit for martingale tail sum Hall and Heyde; **Grübel**.

Triangular urns: $M = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$

with $a, d, \in \mathbb{N}$ and $b \in \mathbb{N}_0$.

Generalized Mittag-Leffler limit law Janson; **Goldschmidt; Mahmoud**.

Pólya urn:
$$M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
.

Beta limit law; Normal limit for martingale tail sum Hall and Heyde; **Grübel**.

Triangular urns:
$$M = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

with $a, d, \in \mathbb{N}$ and $b \in \mathbb{N}_0$.

Generalized Mittag-Leffler limit law Janson; **Goldschmidt; Mahmoud**.

Ball replacement matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$



Ball replacement matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$



Ball replacement matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$



Ball replacement matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$



Ball replacement matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$



Ball replacement matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$



Ball replacement matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$



Ball replacement matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$



Ball replacement matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$



We make the following assumptions:

- Tenable urns: process of drawing and adding/removing balls can be continued ad infinitum.
- Balance condition:

$$\mathsf{M} = \begin{pmatrix} \mathsf{a} & \mathsf{b} \\ \mathsf{c} & \mathsf{d} \end{pmatrix}$$

with

$$\sigma = a + b = c + d \geqslant 1.$$

Consequently: total number of balls T_n is **non-random**: $T_n = W_n + B_n = W_0 + T_0 + n \cdot \sigma$.

Two colors: white and black balls.

Pólya-Eggenberger urns

Main question: Given $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, and start with $W_0 \in \mathbb{N}_0$ white and $B_0 \in \mathbb{N}_0$ black balls: configuration W_n (B_n) after n draws?

Main task: Analyse

$$W_{n} \stackrel{(d)}{=} W_{n-1} + \mathfrak{a} \cdot \mathbb{I}_{n}(\{W\}) + \mathfrak{c} \cdot \mathbb{I}_{n}(\{B\}), \quad n \ge 1,$$

where

$$\mathbb{P}\{\mathbb{I}_{n}(\{W\}) = 1 | \mathcal{F}_{n-1}\} = \frac{W_{n-1}}{W_{n-1} + B_{n-1}} = \frac{W_{n-1}}{T_{n-1}}$$

and

$$\mathbb{P}\{\mathbb{I}_{n}(\{B\}) = 1 | \mathcal{F}_{n-1}\} = \frac{B_{n-1}}{W_{n-1} + B_{n-1}} = \frac{B_{n-1}}{T_{n-1}}.$$

Analytic combinatorics/Symbolic methods: generating functions, method of

moments, etc.

FLAJOLET, GABARRÓ AND PEKARI; BRENNAN AND PRODINGER; STADJE; DUMAS, FLAJOLET AND PUYHAUBERT; HWANG, K. AND PANHOLZER; MORCRETTE; MAHMOUD AND MORCRETTE;

Probabilistic methods: stochastic processes, martingales, contraction method, etc.

BAGCI AND PAL, NEININGER AND KNAPE; KINGMAN; KINGMAN AND VOLKOV; MAHMOUD; JOHNSON, KOTZ AND MAHMOUD; PITTEL; JANSON; POUYANNE; CHAUVIN, MAILLER AND POUYANNE; MAILLER; CHAUVIN, GARDY, POUYANNE AND TON-THAT; GRÜBEL; HIGUERAS, MOLER, PLO AND SAN MIGUEL; TSUKIJI AND MAHMOUD; CHEN AND WEI; RENLUND; ...

Analytic combinatorics/Symbolic methods: generating functions, method of

moments, etc.

FLAJOLET, GABARRÓ AND PEKARI; BRENNAN AND PRODINGER; STADJE; DUMAS, FLAJOLET AND PUYHAUBERT; HWANG, K. AND PANHOLZER; MORCRETTE; MAHMOUD AND MORCRETTE; ...

Probabilistic methods: stochastic processes, martingales, contraction method, etc.

BAGCI AND PAL, NEININGER AND KNAPE; KINGMAN; KINGMAN AND VOLKOV; MAHMOUD; JOHNSON, KOTZ AND MAHMOUD; PITTEL; JANSON; POUYANNE; CHAUVIN, MAILLER AND POUYANNE; MAILLER; CHAUVIN, GARDY, POUYANNE AND TON-THAT; GRÜBEL; HIGUERAS, MOLER, PLO AND SAN MIGUEL; TSUKIJI AND MAHMOUD; CHEN AND WEI; RENLUND; ...

Balanced urn models: partial differential equations; History counting approach

Analytic combinatorics/Symbolic methods: generating functions, method of

moments, etc.

FLAJOLET, GABARRÓ AND PEKARI; BRENNAN AND PRODINGER; STADJE; DUMAS, FLAJOLET AND PUYHAUBERT; HWANG, K. AND PANHOLZER; MORCRETTE; MAHMOUD AND MORCRETTE;

Probabilistic methods: stochastic processes, martingales, contraction method, etc.

BAGCI AND PAL, NEININGER AND KNAPE; KINGMAN; KINGMAN AND VOLKOV; MAHMOUD; JOHNSON, KOTZ AND MAHMOUD; PITTEL; JANSON; POUYANNE; CHAUVIN, MAILLER AND POUYANNE; MAILLER; CHAUVIN, GARDY, POUYANNE AND TON-THAT; GRÜBEL; HIGUERAS, MOLER, PLO AND SAN MIGUEL; TSUKIJI AND MAHMOUD; CHEN AND WEI; RENLUND; ...

Analytic combinatorics/Symbolic methods: generating functions, method of

moments, etc.

FLAJOLET, GABARRÓ AND PEKARI; BRENNAN AND PRODINGER; STADJE; DUMAS, FLAJOLET AND PUYHAUBERT; HWANG, K. AND PANHOLZER; MORCRETTE; MAHMOUD AND MORCRETTE; ...

Probabilistic methods: stochastic processes, martingales, contraction method, etc.

BAGCI AND PAL, NEININGER AND KNAPE; KINGMAN; KINGMAN AND VOLKOV; MAHMOUD; JOHNSON, KOTZ AND MAHMOUD; PITTEL; JANSON; POUYANNE; CHAUVIN, MAILLER AND POUYANNE; MAILLER; CHAUVIN, GARDY, POUYANNE AND TON-THAT; GRÜBEL; HIGUERAS, MOLER, PLO AND SAN MIGUEL; TSUKIJI AND MAHMOUD; CHEN AND WEI; RENLUND; ...

unbalanced r-color urns: general results; also covers 2-color triangular urn models.

Analytic combinatorics/Symbolic methods: generating functions, method of

moments, etc.

FLAJOLET, GABARRÓ AND PEKARI; BRENNAN AND PRODINGER; STADJE; DUMAS, FLAJOLET AND PUYHAUBERT; HWANG, K. AND PANHOLZER; MORCRETTE; MAHMOUD AND MORCRETTE;

Probabilistic methods: stochastic processes, martingales, contraction method, etc.

BAGCI AND PAL, NEININGER AND KNAPE; KINGMAN; KINGMAN AND VOLKOV; MAHMOUD; JOHNSON, KOTZ AND MAHMOUD; PITTEL; JANSON; POUYANNE; CHAUVIN, MAILLER AND POUYANNE; MAILLER; CHAUVIN, GARDY, POUYANNE AND TON-THAT; GRÜBEL; HIGUERAS, MOLER, PLO AND SAN MIGUEL; TSUKIJI AND MAHMOUD; CHEN AND WEI; RENLUND; ...

Analytic combinatorics/Symbolic methods: generating functions, method of

moments, etc.

FLAJOLET, GABARRÓ AND PEKARI; BRENNAN AND PRODINGER; STADJE; DUMAS, FLAJOLET AND PUYHAUBERT; HWANG, K. AND PANHOLZER; MORCRETTE; MAHMOUD AND MORCRETTE; ...

Probabilistic methods: stochastic processes, martingales, contraction method, etc.

BAGCI AND PAL, NEININGER AND KNAPE; KINGMAN; KINGMAN AND VOLKOV; MAHMOUD; JOHNSON, KOTZ AND MAHMOUD; PITTEL; JANSON; POUYANNE; CHAUVIN, MAILLER AND POUYANNE; MAILLER; CHAUVIN, GARDY, POUYANNE AND TON-THAT; GRÜBEL; HIGUERAS, MOLER, PLO AND SAN MIGUEL; TSUKIJI AND MAHMOUD; CHEN AND WEI; RENLUND; ...

contraction method for balanced urns; smoothing equations equations for large urns.

Analytic combinatorics/Symbolic methods: generating functions, method of

moments, etc.

FLAJOLET, GABARRÓ AND PEKARI; BRENNAN AND PRODINGER; STADJE; DUMAS, FLAJOLET AND PUYHAUBERT; HWANG, K. AND PANHOLZER; MORCRETTE; MAHMOUD AND MORCRETTE;

Probabilistic methods: stochastic processes, martingales, contraction method, etc.

BAGCI AND PAL, NEININGER AND KNAPE; KINGMAN; KINGMAN AND VOLKOV; MAHMOUD; JOHNSON, KOTZ AND MAHMOUD; PITTEL; JANSON; POUYANNE; CHAUVIN, MAILLER AND POUYANNE; MAILLER; CHAUVIN, GARDY, POUYANNE AND TON-THAT; GRÜBEL; HIGUERAS, MOLER, PLO AND SAN MIGUEL; TSUKIJI AND MAHMOUD; CHEN AND WEI; RENLUND; ...

Fundamental result

Theorem (Trichotomy of limit laws - Janson)

For a balanced two-color urn model $M = \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \end{pmatrix}$ let $\Lambda = \frac{\lambda_1}{\lambda_2}$ denote the ratio of the two eigenvalues of M.

• for
$$\Lambda \leq \frac{1}{2}$$
: $\frac{W_n - \mathbb{E}(W_n)}{\sqrt{\mathbb{V}(W_n)}} \to \mathbb{N}(0, 1)$.
• for $\Lambda > \frac{1}{2}$: $\frac{W_n - \mathbb{E}(W_n)}{n^{\Lambda}} \to L$.
• for $b_0 \cdot a_1 = 0$ the urn is triangular and $\frac{W_n}{n^{\Lambda}} \to T$.

... New problems.

SVANTE JANSON Functional limit theorems for multitype branching processes and generalized Pólya urns. Stoch. Proc. Appl. 110 (2004)

Remark 4.5: "Mahmoud has initiated the study of urn models where several, say 2, balls are drawn at the same time, and balls are added de- pending on the drawn combination of types. It may be possible to study such models too by the methods of this paper, first considering the corresponding continuous time model, but we have not pursued this. (This case is substantially more complicated than the standard case treated here; for example, the continuous time model will explode in finite time.")

Pólya-Eggenberger urns

Remark

PHILIPPE FLAJOLET, Talk by Basile Morcrette and Nicolas Pouyanne, 2011 Conference - Philippe Flajolet and Analytic Combinatorics, slides;



Pólya-Eggenberger urns

Remark

PHILIPPE FLAJOLET, Talk by Basile Morcrette and Nicolas Pouyanne, 2011 Conference - Philippe Flajolet and Analytic Combinatorics, slides;



Urn models

Multiple drawings - linear affine class

Urn models

Multiple drawings - linear affine class
Previously: Urn contains *w* white and b black balls. $M = \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \end{pmatrix}$ Draw at random a **single** ball: $p_{\text{white}} = \frac{w}{w+b}$, $p_{\text{black}} = \frac{b}{w+b}$.

Multiple drawings: We draw $m \ge 1$ balls in order to obtain our sample. Depending on the drawn multiset of white/black balls we add/remove balls according to the $(m + 1) \times 2$ ball replacement matrix

$$M = \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \\ \cdots & \cdots \\ a_{m-1} & b_{m-1} \\ a_m & b_m \end{pmatrix}$$

Balance: $\sigma = a_0 + b_0 = \cdots = a_m + b_m$.

$$\begin{split} & W & B \\ \{WW\} & \begin{pmatrix} a_0 & b_0 \\ \{WB\} & \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}. \end{split}$$

Urn models with multiple drawings: Example

$$M = \begin{cases} \textbf{WW} \\ \textbf{WW} \\ \textbf{WB} \\ \textbf{BB} \end{cases} \begin{pmatrix} \textbf{a}_0 & \textbf{b}_0 \\ \textbf{a}_1 & \textbf{b}_1 \\ \textbf{a}_2 & \textbf{b}_2 \end{pmatrix}.$$

$$\begin{split} & W & B \\ \{WW\} & \begin{pmatrix} a_0 & b_0 \\ \{WB\} & \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}. \end{split}$$

Urn models with multiple drawings: Example

$$M = \begin{cases} WW \\ \{WW\} \\ \{WB\} \\ \{BB\} \end{cases} \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}.$$

$$\begin{split} & W & B \\ \{WW\} & \begin{pmatrix} a_0 & b_0 \\ \{WB\} & \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}. \end{split}$$

Urn models with multiple drawings: Example

$$M = \begin{cases} WW \\ WW \\ WB \\ BB \end{cases} \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}.$$

$$\begin{split} & W & B \\ \{WW\} & \begin{pmatrix} a_0 & b_0 \\ \{WB\} & \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}. \end{split}$$

Sampling schemes - unordered samples:

Urn contains w white and b black balls.

Sampling without replacement

$$p_{\{k \text{ times white, } (m-k) \text{ times black }\}} = \frac{1}{(b+w)\underline{m}} \binom{m}{k} w^{\underline{k}} b^{\underline{m-k}},$$

with
$$x^{\underline{s}} = x(x-1)...(x-s+1)$$
.

Sampling with replacement,

$$p_{\{k \text{ times white, } (m-k) \text{ times black }\}} = \frac{1}{(b+w)^m} \binom{m}{k} w^k b^{m-k}$$

Distributions: We consider balanced urn models and study the number of white balls W_n :

$$W_{n} \stackrel{(d)}{=} W_{n-1} + \sum_{k=0}^{m} a_{m-k} \cdot \mathbb{I}_{n}(W^{k}B^{m-k}), \quad n \ge 1.$$

 $\mathbb{I}_n(W^k B^{m-k})$, indicator variables k white and m-k black balls. \mathfrak{F}_{n-1} sigma-field generated by first n-1 draws For sampling without replacement $\mathbb{P}\{\mathbb{I}_n(W^k B^{m-k}) = 1 | \mathfrak{F}_{n-1}\}$ is given by

$$\frac{\binom{W_{n-1}}{k}\binom{B_{n-1}}{m-k}}{\binom{T_{n-1}}{m}} = \frac{\binom{W_{n-1}}{k}\binom{T_{n-1}-W_{n-1}}{m-k}}{\binom{T_{n-1}}{m}}$$

and for sampling with replacement

$$\binom{\mathfrak{m}}{k} \frac{W_{n-1}^{k} B_{n-1}^{\mathfrak{m}-k}}{T_{n-1}^{\mathfrak{m}}} = \binom{\mathfrak{m}}{k} \frac{W_{n-1}^{k} (T_{n-1} - W_{n-1})^{\mathfrak{m}-k}}{T_{n-1}^{\mathfrak{m}}}.$$

Distributions: We consider balanced urn models and study the number of white balls W_n :

$$W_{n} \stackrel{(d)}{=} W_{n-1} + \sum_{k=0}^{m} a_{m-k} \cdot \mathbb{I}_{n}(W^{k}B^{m-k}), \quad n \ge 1.$$

 $\mathbb{I}_n(W^k B^{m-k})$, indicator variables k white and m-k black balls. \mathfrak{F}_{n-1} sigma-field generated by first n-1 draws For sampling without replacement $\mathbb{P}\{\mathbb{I}_n(W^k B^{m-k}) = 1 | \mathfrak{F}_{n-1}\}$ is given by

$$\frac{\binom{W_{n-1}}{k}\binom{B_{n-1}}{m-k}}{\binom{T_{n-1}}{m}} = \frac{\binom{W_{n-1}}{k}\binom{T_{n-1}-W_{n-1}}{m-k}}{\binom{T_{n-1}}{m}}$$

and for sampling with replacement

$$\binom{m}{k} \frac{W_{n-1}^{k} B_{n-1}^{m-k}}{T_{n-1}^{m}} = \binom{m}{k} \frac{W_{n-1}^{k} (T_{n-1} - W_{n-1})^{m-k}}{T_{n-1}^{m}}.$$

Distributions: We consider balanced urn models and study the number of white balls W_n :

$$W_{n} \stackrel{(d)}{=} W_{n-1} + \sum_{k=0}^{m} a_{m-k} \cdot \mathbb{I}_{n}(W^{k}B^{m-k}), \quad n \ge 1.$$

 $\mathbb{I}_n(W^k B^{m-k})$, indicator variables k white and m-k black balls. \mathfrak{F}_{n-1} sigma-field generated by first n-1 draws For sampling without replacement $\mathbb{P}\{\mathbb{I}_n(W^k B^{m-k}) = 1 | \mathfrak{F}_{n-1}\}$ is given by

$$\frac{\binom{W_{n-1}}{k}\binom{B_{n-1}}{m-k}}{\binom{T_{n-1}}{m}} = \frac{\binom{W_{n-1}}{k}\binom{T_{n-1}-W_{n-1}}{m-k}}{\binom{T_{n-1}}{m}}$$

and for sampling with replacement

$$\binom{\mathfrak{m}}{k} \frac{W_{n-1}^{k} B_{n-1}^{\mathfrak{m}-k}}{T_{n-1}^{\mathfrak{m}}} = \binom{\mathfrak{m}}{k} \frac{W_{n-1}^{k} (T_{n-1} - W_{n-1})^{\mathfrak{m}-k}}{T_{n-1}^{\mathfrak{m}}}.$$

Pólya urn:
$$M = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix}$$
, $c \in \mathbb{N}$.

Generalization: If we draw $\{W^k S^{m-k}\}$ we add $k \cdot c$ white and $(m-k) \cdot c$ black balls, with $c \in \mathbb{N}$. $(m+1) \times 2$ -matrix

$$M = \begin{pmatrix} mc & 0 \\ (m-1)c & c \\ \dots & \dots \\ c & (m-1)c \\ 0 & mc \end{pmatrix}$$

Pólya urn:
$$M = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix}, c \in \mathbb{N}.$$

Generalization: If we draw $\{W^kS^{m-k}\}$ we add $k \cdot c$ white and $(m-k) \cdot c$ black balls, with $c \in \mathbb{N}$. $(m+1) \times 2$ -matrix

$$M = \begin{pmatrix} mc & 0 \\ (m-1)c & c \\ \dots & \dots \\ c & (m-1)c \\ 0 & mc \end{pmatrix}$$

Friedman urn:
$$M = \begin{pmatrix} 0 & c \\ c & 0 \end{pmatrix}$$
, $c \in \mathbb{N}$.

Generalization: If we draw $\{W^kS^{m-k}\}$ we add $(m-k) \cdot c$ white and $k \cdot c$ black balls, with $c \in \mathbb{N}$. $(m+1) \times 2$ -matrix

$$M = \begin{pmatrix} 0 & mc \\ c & (m-1)c \\ \vdots & \vdots \\ (m-1)c & c \\ mc & 0 \end{pmatrix}$$

Friedman urn:
$$M = \begin{pmatrix} 0 & c \\ c & 0 \end{pmatrix}, c \in \mathbb{N}.$$

Generalization: If we draw $\{W^kS^{m-k}\}$ we add $(m-k) \cdot c$ white and $k \cdot c$ black balls, with $c \in \mathbb{N}$. $(m+1) \times 2$ -matrix

$$M = \begin{pmatrix} 0 & mc \\ c & (m-1)c \\ \vdots & \vdots \\ (m-1)c & c \\ mc & 0 \end{pmatrix}$$

Urn models with multiple drawings: linear affine models.

Given

$$M = \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \\ \dots & \dots \\ a_{m-1} & b_{m-1} \\ a_m & b_m \end{pmatrix}$$

one cannot easily obtain the expected value.

Difficulty:

$$\mathbb{E}(W_n) = \sum_{k=0}^m f_{n,k} \mathbb{E}(W_{n-1}^k),$$

for certain sequences $f_{n,k}$ depending only on n and m (and k). When is it possible to **calculate** the expected value of W_n (in a **simple** way)?

Urn models with multiple drawings: linear affine models.

Given

$$M = \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \\ \dots & \dots \\ a_{m-1} & b_{m-1} \\ a_m & b_m \end{pmatrix}$$

one cannot easily obtain the expected value.

Difficulty:

$$\mathbb{E}(W_n) = \sum_{k=0}^{m} f_{n,k} \mathbb{E}(W_{n-1}^k),$$

for certain sequences $f_{n,k}$ depending only on n and m (and k). When is it possible to **calculate** the expected value of W_n (in a **simple** way)?

Urn models with multiple drawings: linear affine models.

Given

$$M = \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \\ \dots & \dots \\ a_{m-1} & b_{m-1} \\ a_m & b_m \end{pmatrix}$$

one cannot easily obtain the expected value.

Difficulty:

$$\mathbb{E}(W_n) = \sum_{k=0}^m f_{n,k} \mathbb{E}(W_{n-1}^k),$$

for certain sequences $f_{n,k}$ depending only on n and m (and k). When is it possible to **calculate** the expected value of W_n (in a **simple** way)?

Proposition (Mahmoud and K.)

Given matrix M, the numbers a_{m-1} , a_m , and the balance factor $\sigma = a_k + b_k \ge 0$. W_n satisfies a linear affine relation of the form

$$\mathbb{E}\big[W_{n}|\mathcal{F}_{n-1}\big] = \alpha_{n}W_{n-1} + \beta_{n}, \qquad n \ge 1,$$

if and only if, for $0\leqslant k\leqslant m,$ the numbers a_k satisfy the condition

$$a_k = (m-k)a_{m-1} - (m-k-1)a_m.$$

(equivalently, $a_k = a_0 + hk$, with h an integer guaranteeing tenability.) Then

$$\alpha_n = \frac{T_{n-1} + m(\mathfrak{a}_{m-1} - \mathfrak{a}_m)}{T_{n-1}}, \qquad \textit{and} \qquad \beta_n = \mathfrak{a}_m, \qquad n \geqslant 1.$$

Remark

For m = 1 we reobtain the ordinary balanced urns. Markus Kuba - Limit laws for urns with multiple draws

Proof (sketch). For sampling with replacement:

$$\mathbb{E}\left[W_{n}|\mathcal{F}_{n-1}\right] = W_{n-1} + \sum_{k=0}^{m} a_{m-k}\mathbb{E}\left[\mathbb{I}_{n}\left(W^{k}B^{m-k}\right) = 1|\mathcal{F}_{n-1}\right]$$

$$= W_{n-1} + \sum_{k=0}^{m} a_{m-k}\binom{m}{k} \frac{W_{n-1}^{k}B_{n-1}^{m-k}}{T_{n-1}^{m}}$$

$$= W_{n-1} + \sum_{k=0}^{m} \left(k(a_{m-1} - a_{m}) + a_{m}\right)\binom{m}{k} \frac{W_{n-1}^{k}B_{n-1}^{m-k}}{T_{n-1}^{m}}$$

$$= W_{n-1} + \frac{1}{T_{n-1}^{m}} \left(m(a_{m-1} - a_{m})W_{n-1}(W_{n-1} + B_{n-1})^{m-1} + a_{m}(W_{n-1} + B_{n-1})^{m-1}\right)$$

Taking the expectation gives the "if-part" of the stated result.

Proof (sketch). For sampling with replacement:

$$\begin{split} & \mathbb{E}\left[W_{n}|\mathcal{F}_{n-1}\right] = W_{n-1} + \sum_{k=0}^{m} a_{m-k} \mathbb{E}\left[\mathbb{I}_{n}(W^{k}B^{m-k}) = 1|\mathcal{F}_{n-1}\right] \\ & = W_{n-1} + \sum_{k=0}^{m} a_{m-k} \binom{m}{k} \frac{W_{n-1}^{k}B_{n-1}^{m-k}}{T_{n-1}^{m}} \\ & = W_{n-1} + \sum_{k=0}^{m} (k(a_{m-1} - a_{m}) + a_{m}) \binom{m}{k} \frac{W_{n-1}^{k}B_{n-1}^{m-k}}{T_{n-1}^{m}} \\ & = W_{n-1} + \frac{1}{T_{n-1}^{m}} \left(m(a_{m-1} - a_{m})W_{n-1}(W_{n-1} + B_{n-1})^{m-1} + a_{m}(W_{n-1} + B_{n-1})^{m-1}\right) \\ & = W_{n-1} + \frac{1}{T_{n-1}^{m}} \left(m(a_{m-1} - a_{m})W_{n-1}T_{n-1}^{m-1} + a_{m}T_{n-1}^{m}\right). \end{split}$$

Taking the expectation gives the "if-part" of the stated result.

Proof (sketch). For sampling with replacement:

$$\begin{split} \mathbb{E}\left[W_{n}|\mathcal{F}_{n-1}\right] &= W_{n-1} + \sum_{k=0}^{m} a_{m-k} \mathbb{E}[\mathbb{I}_{n}(W^{k}B^{m-k}) = 1|\mathcal{F}_{n-1}] \\ &= W_{n-1} + \sum_{k=0}^{m} a_{m-k} \binom{m}{k} \frac{W_{n-1}^{k}B_{n-1}^{m-k}}{T_{n-1}^{m}} \\ &= W_{n-1} + \sum_{k=0}^{m} (k(a_{m-1} - a_{m}) + a_{m}) \binom{m}{k} \frac{W_{n-1}^{k}B_{n-1}^{m-k}}{T_{n-1}^{m}} \\ &= W_{n-1} + \frac{1}{T_{n-1}^{m}} \left(m(a_{m-1} - a_{m})W_{n-1}(W_{n-1} + B_{n-1})^{m-1} + a_{m}(W_{n-1} + W_{n-1}) + \frac{1}{T_{n-1}^{m}} \right) \\ &= W_{n-1} + \frac{1}{T_{n-1}^{m}} \left(m(a_{m-1} - a_{m})W_{n-1}T_{n-1}^{m-1} + a_{m}T_{n-1}^{m} \right). \end{split}$$

Taking the expectation gives the "if-part" of the stated result.

Proof (sketch). For sampling with replacement:

$$\begin{split} & \mathbb{E}\left[W_{n}|\mathcal{F}_{n-1}\right] = W_{n-1} + \sum_{k=0}^{m} a_{m-k} \mathbb{E}\left[\mathbb{I}_{n}(W^{k}B^{m-k}) = 1|\mathcal{F}_{n-1}\right] \\ & = W_{n-1} + \sum_{k=0}^{m} a_{m-k} \binom{m}{k} \frac{W_{n-1}^{k}B_{n-1}^{m-k}}{T_{n-1}^{m}} \\ & = W_{n-1} + \sum_{k=0}^{m} (k(a_{m-1} - a_{m}) + a_{m}) \binom{m}{k} \frac{W_{n-1}^{k}B_{n-1}^{m-k}}{T_{n-1}^{m}} \\ & = W_{n-1} + \frac{1}{T_{n-1}^{m}} \left(m(a_{m-1} - a_{m})W_{n-1}(W_{n-1} + B_{n-1})^{m-1} + a_{m}(W_{n-1} + W_{n-1})^{m-1}\right) \\ & = W_{n-1} + \frac{1}{T_{n-1}^{m}} \left(m(a_{m-1} - a_{m})W_{n-1}T_{n-1}^{m-1} + a_{m}T_{n-1}^{m}\right). \end{split}$$

Taking the expectation gives the "if-part" of the stated result.

Proof (sketch). For sampling with replacement:

$$\begin{split} & \mathbb{E}\left[W_{n}|\mathcal{F}_{n-1}\right] = W_{n-1} + \sum_{k=0}^{m} a_{m-k} \mathbb{E}\left[\mathbb{I}_{n}(W^{k}B^{m-k}) = 1|\mathcal{F}_{n-1}\right] \\ & = W_{n-1} + \sum_{k=0}^{m} a_{m-k} \binom{m}{k} \frac{W_{n-1}^{k}B_{n-1}^{m-k}}{T_{n-1}^{m}} \\ & = W_{n-1} + \sum_{k=0}^{m} (k(a_{m-1} - a_{m}) + a_{m}) \binom{m}{k} \frac{W_{n-1}^{k}B_{n-1}^{m-k}}{T_{n-1}^{m}} \\ & = W_{n-1} + \frac{1}{T_{n-1}^{m}} \left(m(a_{m-1} - a_{m})W_{n-1}(W_{n-1} + B_{n-1})^{m-1} + a_{m}(W_{n-1} + B_{n-1})^{m-1}\right) \\ & = W_{n-1} + \frac{1}{T_{n-1}^{m}} \left(m(a_{m-1} - a_{m})W_{n-1}T_{n-1}^{m-1} + a_{m}T_{n-1}^{m}\right). \end{split}$$

Taking the expectation gives the "if-part" of the stated result.

Although only three parameters a_{m-1} , a_m and σ , we **unify earlier** treated models:

Example

- Case m = 2: MAHMOUD's condition $a_0 = 2a_1 a_2$
- Case $a_m=mc,\,a_{m-1}=(m-1)c$ and $\sigma=mc:$ generalized Friedman urn model
- Case $a_m=$ 0, $a_{m-1}=c$ and $\sigma=mc:$ generalized Pólya urn model
- Case $a_m = 1$ and $a_{m-1} = 0$ and $\sigma = 1$ we obtain $a_k = -(m-k) + 1$, urn model for logical circuits.

Urn models

- A few results

Urn models

- A few results

Proposition (Mahmoud and K.)

For balanced affine urn schemes the expected value is given by

$$\mathbb{E}[W_n] = \frac{a_m(n + \frac{T_0}{\sigma})}{1 - \Lambda} + \left(W_0 - \frac{\frac{a_m T_0}{\sigma}}{1 - \Lambda}\right) \frac{\binom{n - 1 + \frac{T_0}{\sigma} + \Lambda}{n}}{\binom{n - 1 + \frac{T_0}{\sigma}}{n}},$$

as well as the asymptotic expansion

$$\mathbb{E}[W_n] = \frac{a_m}{1 - \Lambda} n + \left(W_0 - \frac{\frac{a_m T_0}{\sigma}}{1 - \Lambda}\right) \frac{\Gamma(\frac{T_0}{\sigma})}{\Gamma(\frac{T_0}{\sigma} + \Lambda)} n^{\Lambda} + O(1)$$

Moreover, for $\Lambda = 1$ we obtain $\mathbb{E}[W_n] = W_0 \frac{n\sigma + T_0}{T_0}$.

Theorem (Mahmoud and K.)

For balanced affine urn schemes, the variance satisfies the following expansions. Small-index urns, the case $\Lambda < \frac{1}{2}$:

$$\mathbb{V}[W_n] = \frac{a_m b_0 \Lambda^2}{m(1-2\Lambda)(1-\Lambda)^2} n + o(n).$$

Critical-index urns, the case $\Lambda = \frac{1}{2}$:

$$\mathbb{V}[W_n] = rac{a_m b_0}{m} n \log n + \mathcal{O}(n),$$

Large-index urns, the case $\Lambda > \frac{1}{2}$:

$$\mathbb{V}[W_n] = Cn^{2\Lambda} + \mathcal{O}(n),$$

26/36

with the constant C being model-dependent given by an infinite AofA¹⁵ Markus Kuba - Limit laws for urns with multiple draws

Pólya-Eggenberger urns

Theorem (continued)

Constant C:

$$\begin{split} C &= \frac{W_0^2}{\psi_0} + \sum_{j=1}^{\infty} \frac{\beta_j \mathbb{E}[W_{j-1}] + a_m^2 - \psi_j 2a_m j^{-\Lambda} \Big(\frac{a_m}{1-\Lambda} j^{1-\Lambda} + (W_0 - \frac{a_m T_0}{1-\Lambda}) \frac{\Gamma(\frac{T_0}{\sigma})}{\Gamma(\frac{T_0}{\sigma} + \Lambda)} \Big)}{\psi_j} \\ &+ \frac{2a_m^2}{1-\Lambda} \zeta(2\Lambda - 1) + 2a_m \Big(W_0 - \frac{a_m T_0}{\sigma} \Big) \frac{\Gamma(\frac{T_0}{\sigma})}{\Gamma(\frac{T_0}{\sigma} + \Lambda)} \zeta(\Lambda) - \Big(W_0 - \frac{a_m T_0}{1-\Lambda} \Big)^2 \frac{\Gamma^2(\frac{T_0}{\sigma})}{\Gamma^2(\frac{T_0}{\sigma} + \Lambda)}, \end{split}$$

with $\zeta(z)$ denoting the Riemann zeta function and β_i , ψ_i certain sequences:

$$\psi_{n} = \begin{cases} \frac{\Gamma(n + \lambda_{1})\Gamma(n + \lambda_{2})}{\Gamma(n + \frac{T_{0}}{\sigma})\Gamma(n + \frac{T_{0} - 1}{\sigma})}, \text{ without replacement} \\ \frac{\Gamma(n + \mu_{1})\Gamma(n + \mu_{2})}{\Gamma(n + \frac{T_{0}}{\sigma})^{2}} \text{ with replacement,} \end{cases}$$

$$and \lambda_{1,2} = \frac{m(a_{m-1}-a_m)+T_0 - \frac{1}{2} \pm \frac{1}{2}\sqrt{1+4m(a_{m-1}-a_m)(a_{m-1}-a_m+1)}}{\sigma},$$

$$\beta_n = (a_{m-1}-a_m)^2 \left(\frac{m}{T_{n-1}} - \frac{m^2}{T_{n-1}^2}\right) + \frac{2ma_m(a_{m-1}-a_m)}{T_{n-1}} + 2a_m. \text{ for sampling without replacement,}$$

$$or \mu_{1,2} = \frac{m(a_{m-1}-a_m)+T_0 \pm (a_{m-1}-a_m)\sqrt{m}}{\sigma}, \text{ and}$$

$$\beta_n = \frac{(a_{m-1}-a_m)^2m}{T_{n-1}} + \frac{2ma_m(a_{m-1}-a_m)}{T_{n-1}} + 2a_m \text{ for sampling with replacement;}$$
AofA'15

Theorem (Trichotomy of limit laws for affine linear models - Mahmoud and K.)

For a balanced linear affine two-color urn model M with sample size $m \ge 1$ let $\Lambda = \frac{\lambda_1}{\lambda_2}$ denote the ratio of the two eigenvalues of M.

$$\begin{aligned} \bullet \quad & \text{for } \Lambda \leqslant \frac{1}{2} \colon \frac{W_n - \mathbb{E}(W_n)}{\sqrt{\mathbb{V}(W_n)}} \to \mathcal{N}(0, 1). \\ \bullet \quad & \text{for } \Lambda > \frac{1}{2} \colon \frac{W_n - \mathbb{E}(W_n)}{n^{\Lambda}} \to L. \\ \bullet \quad & \text{for } b_0 \cdot a_m = 0 \text{ the urn is triangular and } \frac{W_n}{n^{\Lambda}} \to T. \end{aligned}$$

There exists explicit expressions for all moments of $\mathbb{E}(L^s)$ and $\mathbb{E}(T^s)$ as nested infinite sums; almost sure convergence for large-index urns and triangular urns.

Urn models

- Multiple drawings Analytic combinatorics

Urn models

- Multiple drawings Analytic combinatorics Analytic combinatorial approach à la Flajolet:

- Symbolic description of the problem, replacement of balls modelled by differential operator
- Generating functions
- Translate description to higher order linear partial differential equation
- Use analytic machinery to analyze parameters of interest

Using the idea DUMAS, FLAJOLET AND PUYHAUBERT we study urn histories using differential operators:

 ∂_z : differential operator with respect to z: $\partial_z(z^n) = n \cdot z^{n-1}$ $\Theta_z = z \cdot \partial_z$, such that $\Theta_z(z^n) = n \cdot z^n$

Assume we have w white, b black balls \implies encoded by $x^w y^b$:

$$\frac{\binom{m}{k}}{(b+w)\underline{m}}y^{m-k}\partial_{y}^{m-k}x^{k}\partial_{x}^{k}(x^{w}y^{b}) = \frac{\binom{m}{k}w^{\underline{k}}b\underline{m-k}}{(b+w)\underline{m}}x^{w}y^{b},$$

$$\frac{\binom{\mathfrak{m}}{k}}{(\mathfrak{b}+w)^{\mathfrak{m}}}\Theta_{x}^{k}\Theta_{y}^{\mathfrak{m}-k}(x^{w}y^{\mathfrak{b}}) = \frac{\binom{\mathfrak{m}}{k}w^{k}\mathfrak{b}^{\mathfrak{m}-k}}{(\mathfrak{b}+w)^{\mathfrak{m}}}x^{w}y^{\mathfrak{b}}.$$
Analysis using Analytic combinatorics

Let $H_n(x, y) = \mathbb{E}(x^{W_n}y^{B_n})H_n$ and $H(x, y, z) = \sum_{n \ge 0} H_n(x, y) \frac{z^n}{(n!)^m}$ denote the complete history generating function

Proposition (Morcrette and K.)

Starting with W_0 white and B_0 black balls the generating function of all urn histories H(x, y; z) satisfies

$$\mathcal{D} * \mathsf{H}(\mathbf{x}, \mathbf{y}, z) = \frac{1}{z} \cdot \Theta_z^{\mathfrak{m}} * \mathsf{H}(\mathbf{x}, \mathbf{y}, z),$$

with

$$\mathcal{D}_{\mathsf{M}} = \sum_{k=0}^{m} \binom{m}{k} x^{a_{\mathfrak{m}-k}+k} y^{b_{\mathfrak{m}-k}+\mathfrak{m}-k} \vartheta_{x}^{k} \vartheta_{y}^{\mathfrak{m}-k},$$

and

$$\mathcal{D}_R = \sum_{k=0}^m \binom{m}{k} x^{\mathfrak{a}_{\mathfrak{m}-k}} y^{\mathfrak{b}_{\mathfrak{m}-k}} \Theta_x^k \Theta_y^{\mathfrak{m}-k}.$$

Proposition (Linear affine model - AC version)

The history generating function $H_x(1, 1, z)$ of the expected value $\mathbb{E}(W_n)$ satisfies an ordinary differential equation if the coefficients satisfy an affinity condition: $a_k = a_0 + hk$, with h an integer guaranteeing tenability. Then,

$$\mathsf{E}_{x}\mathsf{E}_{y}\mathfrak{d}_{x}\mathcal{D} = \begin{cases} \left(\mathfrak{m}(1-h)\cdot\mathcal{V}^{\underline{m}-1}+\mathcal{V}^{\underline{m}}\right)\mathsf{E}_{x}\mathsf{E}_{y}\mathfrak{d}_{x} + (\mathfrak{a}_{0}+h\mathfrak{m})\mathcal{V}^{\underline{m}}\mathsf{E}_{x}\mathsf{E}_{y},\\ \left(\mathfrak{m}(1-h)\cdot\mathcal{V}^{\underline{m}-1}+\mathcal{V}^{\underline{m}}\right)\mathsf{E}_{x}\mathsf{E}_{y}\mathfrak{d}_{x} + (\mathfrak{a}_{0}+h\mathfrak{m})\mathcal{V}^{\underline{m}}\mathsf{E}_{x}\mathsf{E}_{y}, \end{cases}$$

with $\mathcal{V} = \sigma \Theta_z + T_0$. Thus,

$$\left(\mathfrak{m}(1-\mathfrak{h})\cdot\mathcal{V}^{\underline{\mathfrak{m}}-1}+\mathcal{V}^{\underline{\mathfrak{m}}}-\frac{1}{z}\cdot\Theta_{z}^{\mathfrak{m}}\right)H_{x}(1,1,z)=-(\mathfrak{a}_{0}+\mathfrak{h}\mathfrak{m})\mathcal{V}^{\underline{\mathfrak{m}}}H(1,1,z)$$

for sampling without replacement and

$$\left(\mathfrak{m}(1-\mathfrak{h})\cdot\mathcal{V}^{\mathfrak{m}-1}+\mathcal{V}^{\mathfrak{m}}-\frac{1}{z}\cdot\Theta_{z}^{\mathfrak{m}}\right)H_{x}(1,1,z)=-(\mathfrak{a}_{0}+\mathfrak{h}\mathfrak{m})\mathcal{V}^{\mathfrak{m}}H(1,1,z)$$

for sampling with replacement (with H(1,1,z) explicitly given and $E_xf(x)=f(1)\mbox{.}$

AofA'15

Urn models

- Outlook

Urn models

- Outlook

Outlook and conclusion

Conclusion

- Characterization of linear affine urns
- Extension of the trichotomy small, large and triangular urns from m = 1 to $m \ge 1$ for linear affine urns.
- Analytic combinatorics PDEs and an alternative derivation of the linear affine urns.

Outlook

We obtain higher order PDEs using Morcrette's approach also for

- r-colors
- unbalanced models
- ordered samples

Linear affine models: We can characterize and analyze r-colors models (joint work with Hosam Mahmoud).

Thanks for your attention!

Danke für Ihre Aufmerksamkeit!

Merci de votre attention!