On the game of Memory

Paweł Hitczenko
(largely based on a joint work with H. Acan)

October 27, 2016

Description of the game

- A deck of n pairs of cards is shuffled and the cards are laid face down in a table (a row).
- In a *move* a player flips a card and then a second one. If they match the player collects them and continues.
- If the cards do not match they are flipped over again and play passes to the next player.
- The play ends when all pairs have been removed and the player who collects the most pairs is the winner.

Solitaire Version

- The game is played by a single person.
- The goal is to finish the game in the smallest possible number of moves.

Solitaire Version

- The game is played by a single person.
- The goal is to finish the game in the smallest possible number of moves.
- Under the assumption that the player has perfect memory, **Velleman and Warrington (2013)** studied the expected values of three characteristics of the game, namely:
  - the length of the game, G_n,
  - the waiting time till the first match, F_n, and
  - the number of *lucky moves*, L_n (a lucky move is a move in which the two cards, neither of which has been flipped before, match).

Velleman–Warrington results

Theorem (Velleman–Warrington (2013))

In a game played with n pairs of cards, as $n \to \infty$,

- $\mathbb{E}F_n = \frac{2^{2n}}{\binom{2n}{n}} \sim \sqrt{\pi n}$,
- $\mathbb{E}L_n \sim \ln 2$,
- $\mathbb{E}G_n \sim (3 - 2\ln 2)n$.

Our results

Theorem (Acan–H. (2016))

In a game played with n pairs of cards, as $n \to \infty$,

- $\frac{F_n}{2\sqrt{n}} \xrightarrow{d} W$,
  where W *is a standard Rayleigh random variable, i.e. a variable with density* $2xe^{-x^2}$ *if* $x > 0$ *and 0 otherwise.*
- $L_n \xrightarrow{d} Pois(\ln 2)$,
- $\frac{G_n-(3-2\ln 2)n}{\sqrt{n}} \xrightarrow{d} N(0,\sigma^2), \quad \sigma^2 = 4\ln^2 2 + 2\ln 2 - 13/4.$

Our results

In fact, we know more:

Theorem (Acan–H. (2016))

If $F_{n,i}$ is the number of moves between the $(i-1)^{st}$ and the i^{th} match, then for every fixed $k \geq 1$

$$\left(\frac{F_{n,1}}{2\sqrt{n}}, \frac{F_{n,2}}{2\sqrt{n}}, \dots, \frac{F_{n,k}}{2\sqrt{n}}\right) \xrightarrow{d} (W_1, \dots, W_k),$$

where $(W_1, \dots, W_k)$ has joint density given by

$$2^k x_1(x_1+x_2)\dots(x_1+\dots+x_k)e^{-(x_1+\dots+x_k)^2},$$

if $x_1, \dots, x_k \geq 0$ and 0 otherwise.

Methods

- The proof that $\frac{F_n}{2\sqrt{n}} \xrightarrow{d} W$ is straightforward:

$$\mathbb{P}(F_n > t) = \mathbb{P}(\text{first } t \text{ cards are distinct })$$
$$= \frac{\binom{n}{t} \cdot t! \cdot [(2n-t)!/2^{n-t}]}{(2n)!/2^n} = \frac{2^t \cdot (n)_t}{(2n)_t},$$

where $(x)_t = x(x-1)\dots(x-(t-1))$ is the falling factorial (but a more complicated proof will be sketched later).

Methods

- The proof that $\frac{F_n}{2\sqrt{n}} \xrightarrow{d} W$ is straightforward:

$$\mathbb{P}(F_n > t) = \mathbb{P}(\text{first } t \text{ cards are distinct })$$
$$= \frac{\binom{n}{t} \cdot t! \cdot [(2n-t)!/2^{n-t}]}{(2n)!/2^n} = \frac{2^t \cdot (n)_t}{(2n)_t},$$

  where $(x)_t = x(x-1)\ldots(x-(t-1))$ is the falling factorial (but a more complicated proof will be sketched later).

- That $L_n \xrightarrow{d} \mathrm{Pois}(\ln 2)$ is more involved, but still direct and relies on a method of (factorial) moments.

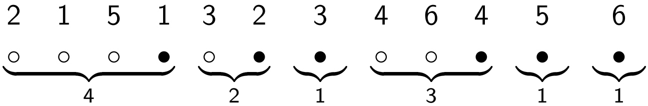
Length of the game

Length of the game:

- **Velleman– Warrington** showed that

$$G_n = \frac{3}{2}n + \frac{1}{2}\sum_i B_{n,2i} - L_n,$$

where $B_{n,j}$ is the number of blocks of size j.



Here, $G_n = \frac{3}{2} \cdot 6 + \frac{1}{2}(1+1) - 0 = 10$ (3 moves to remove 1's, 2 for 2's, 1 for 3's, 2 for 4's, 1 for 5's, and 1 for 6's).
(Lucky move: an even length block ending with a pair.)

Joint distribution of the lengths of blocks

Theorem

As $n \to \infty$, in $\mathbb{R}^\infty$,

$$\frac{1}{\sqrt{n}}\left(B_{n,i} - \frac{4n}{(i+2)_3}\right)_{i=1}^{\infty} \xrightarrow{d} (\Gamma_i)_{i=1}^{\infty},$$

where (Γ_i) are jointly Gaussian with mean zero and covariance matrix $\Sigma = [\sigma_{ij}]$ given by

$$\sigma_{ij} = \begin{cases} \frac{16}{(i+2)_3(j+2)_3} - \frac{24}{(i+j+2)_4}, & \text{if } i \neq j; \\ \frac{4}{(j+2)_3} + \frac{16}{(j+2)_3^2} - \frac{24}{(2j+2)_4}, & \text{if } i = j. \end{cases}$$

From that, obtaining the distribution of $\sum_j B_{2j}$ is straightforward.

Generalized Pólya Urn Model

This theorem is shown by interpreting the game in terms of generalized Pólya urn model as follows:

- WLOG assume that pairs are collected in an increasing order (such games are called *standard* by Velleman and Warrington).
- When a new pair of cards is added to an existing game, the second element of a pair is put at the end of the row and the first is put somewhere before it.
- If it is put in the existing block of size i it reduces the number of such blocks by 1, increases the number of blocks of length $i+1$ by 1, and the card inserted at the end creates a block of size 1.
- If both cards are placed at the end, this just creates one new block of size 2 (we need a model with immigration for that).

Generalized Pólya Urn Models

Theory of Pólya urn models is huge, in generality we need it was developed in

- **S.Janson** Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Processes and Applications* (2004).
- **L.-X. Zhang, F. Hu, S. H. Chung, W. S. Chan** Immigrated urn models – theoretical properties and applications. *Annals of Statistics* (2011).

Connections to Preferential Attachment Graphs

- Earlier picture: game of memory

$$\underbrace{\overset{2}{\circ}\quad\overset{1}{\circ}\quad\overset{5}{\circ}\quad\overset{1}{\bullet}}_{4}\quad\underbrace{\overset{3}{\circ}\quad\overset{2}{\bullet}}_{2}\quad\underbrace{\overset{3}{\bullet}}_{1}\quad\underbrace{\overset{4}{\circ}\quad\overset{6}{\circ}\quad\overset{4}{\bullet}}_{3}\quad\underbrace{\overset{5}{\bullet}}_{1}\quad\underbrace{\overset{6}{\bullet}}_{1}$$

- Preferential attachment graphs via chord diagrams (**Bollobás, Riordan, Spencer, and Tusnády (2001)**):

$$\underbrace{\overset{2}{\circ}\quad\overset{1}{\circ}\quad\overset{5}{\circ}\quad\overset{1}{\bullet}}_{v_1}\quad\underbrace{\overset{3}{\circ}\quad\overset{2}{\bullet}}_{v_2}\quad\underbrace{\overset{3}{\bullet}}_{v_3}\quad\underbrace{\overset{4}{\circ}\quad\overset{6}{\circ}\quad\overset{4}{\bullet}}_{v_4}\quad\underbrace{\overset{5}{\bullet}}_{v_5}\quad\underbrace{\overset{6}{\bullet}}_{v_6}$$

Draw an edge between v_i and v_j iff there is a pair whose one element is in v_i and the other in v_j.

In the language of preferential attachment graphs our results about the joint distribution of the first few matches in the game recover results by **Peköz, Röllin** and **Ross** about the joint degree distribution of the first few vertices in the preferential attachment graphs

- **E. Peköz, A. Röllin, N. Ross** Degree asymptotics with rates for preferential attachment random graphs, *Annals of Applied Probability* (2013).
- **E. Peköz, A. Röllin, N. Ross** Joint degree distributions of preferential attachment random graphs, http://arxiv.org/pdf/1402.4686v1.pdf.

The length of the game gives the asymptotic joint normality of the counts of the degrees of vertices.

A recurrence for generating polynomials

If

$$A_n(x) = \sum_{j=2}^{n+1} a_{n,j} x^j$$

is the generating polynomial of the number of standard games with n pairs of cards with first match at j then one has

$$A_n(x) = (2n-1)A_{n-1}(x) + x(x-1)A'_{n-1}(x),$$

which is an example of a recurrence of the form

$$P_n(x) = f_n(x)P_{n-1}(x) + g_n(x)P'_{n-1}(x)$$

for some fixed sequences of polynomials (f_n), (g_n) with

$$g_n(1) = 0.$$

A few more examples

- **H., Janson (2014)**: for $a, b > 0$ (but could be complex)

$$\begin{aligned} P_{n,a,b}(x) &= ((n-1+b)x+a)P_{n-1,a,b}(x) + x(1-x)P'_{n-1,a,b}(x) \\ P_{0,a,b}(x) &= 1. \end{aligned}$$

- $b = 0, a = 1$ give classical Eulerian polynomials. That is:

$$P_{n,1,0}(x) = E_n(x) = \sum_{k=0}^{n} \left\langle {n \atop k} \right\rangle x^k,$$

where $\left\langle {n \atop k} \right\rangle$ is the number of permutations of $[n]$ with exactly k ascents and the recurrence is:

$$E_n(x) = ((n-1)x+1)E_{n-1}(x) + x(1-x)E'_{n-1}(x).$$

A few more examples

- **Aval, Boussicault, Nadeau (2011)**:

$$B_n(x) = nx(x+1)B_{n-1}(x) + x(1-x^2)B'_{n-1}(x),$$
$$B_0(x) = x.$$

A few more examples

- **Aval, Boussicault, Nadeau (2011)**:

$$
\begin{aligned}
B_n(x) &= nx(x+1)B_{n-1}(x) + x(1-x^2)B'_{n-1}(x), \\
B_0(x) &= x.
\end{aligned}
$$

Question: Can we describe the limits of sequences of random variables (X_n) described by

$$\frac{P_n(x)}{P_n(1)} = \sum_{k\geq 0} \frac{p_{n,k}}{P_n(1)} x^k = \mathbb{E}x^{X_n}, \quad p_{n,k} \geq 0,$$

$$i.e. \quad \mathbb{P}(X_n = k) = \frac{p_{n,k}}{P_n(1)} = \frac{p_{n,k}}{\sum_j p_{n,j}}, \quad k \geq 0,$$

by looking at the sequences (f_n) and (g_n)?

Method of Moments

Often,

$$X_n \xrightarrow{d} X, \quad \text{as } n \to \infty$$

is implied by

$$\mathbb{E}(X_n)_r \longrightarrow \mathbb{E}(X)_r, \quad r = 1, 2 \dots$$

where

$$\mathbb{E}(X)_r := \mathbb{E}X(X-1)\dots(X-(r-1)).$$

In terms of polynomials

$$\mathbb{E}(X_n)_r = \frac{P_n^{(r)}(1)}{P_n(1)}$$

and so we are interested in limiting behavior of $P_n^{(r)}(1)$.

Method of Moments

Leibniz formula gives

$$\begin{aligned} P_n^{(r)}(x) &= (f_n(x)P_{n-1}(x))^{(r)} + (g_n(x)P'_{n-1}(x))^{(r)} \\ &= \sum_{k=0}^{r} \binom{r}{k} f_n^{(k)}(x) P_{n-1}^{(r-k)}(x) \\ &+ \sum_{k=0}^{r} \binom{r}{k} g_n^{(k)}(x) P_{n-1}^{(r+1-k)}(x). \end{aligned}$$

So, if f_n and g_n are low–degree polynomials then one obtains a reasonably simple recurrence for $P_n^{(r)}(1)$.

V–W recurrence

In the case of $A_n(x)$, $f_n = 2n-1$ has degree 0, $g_n(x) = x(x-1)$ has degree 2 and we get, after evaluating at $x=1$,

$$A_n(1)^{(r)} =: A_n^{(r)} = (2n-1+r)A_{n-1}^{(r)} + r(r-1)A_{n-1}^{(r-1)}.$$

In particular, if $r=0$ this is

$$A_n = A_n^{(0)} = (2n-1)A_{n-1} = \cdots = (2n-1)!!,$$

so that if $Q_n(x) = A_n(x)/A_n$ then

$$Q_n^{(r)} = \left(1 + \frac{r}{2n-1}\right) Q_{n-1}^{(r)} + \frac{r(r-1)}{2n-1} Q_{n-1}^{(r-1)}.$$

V-W recurrence

By induction $Q_n^{(r)} = 0$ for $0 \le n < r - 1$ and for $n \ge r - 1$

$$Q_n^{(r)} = C_r \frac{\Gamma(n + \frac{r+1}{2})}{\Gamma(n + \frac{1}{2})} \sum_{1 \le k_1 < k_2 < \cdots < k_{r-1} \le n} \prod_{l=1}^{r-1} \frac{\Gamma(k_l + \frac{l-1}{2})}{\Gamma(k_l + \frac{l+2}{2})}$$

with

$$C_1 = \sqrt{\pi}, \quad C_r = \binom{r}{2} C_{r-1} = \sqrt{\pi} \frac{\Gamma(r+1)(r-1)!}{2^{r-1}}, \quad r \ge 2.$$

The multiple sum can be evaluated:

$$\sum_{1 \le k_1 < \cdots < k_{r-2} < k_{r-1}} \prod_{l=1}^{r-1} \frac{\Gamma(k_l + \frac{l-1}{2})}{\Gamma(k_l + \frac{l+2}{2})} = \frac{2^{r-1}}{(r-1)!\Gamma(\frac{r+1}{2})}.$$

V-W recurrence

This gives

$$Q_n^{(r)} \sim \frac{\Gamma(n+\frac{r+1}{2})}{\Gamma(n+\frac{1}{2})} C_r \frac{2^{r-1}}{(r-1)!\Gamma(\frac{r+1}{2})} \sim \sqrt{\pi}\frac{\Gamma(r+1)}{\Gamma(\frac{r+1}{2})} n^{r/2}.$$

By the duplication formula for Gamma function:

$$\frac{\Gamma(r+1)}{\Gamma(\frac{r+1}{2})}\sqrt{\pi} = 2^r\Gamma(1+\frac{r}{2}),$$

and thus

$$\frac{Q_n^{(r)}}{(2\sqrt{n})^r} \sim \Gamma(1+\frac{r}{2}),$$

which are the moments of the standard Rayleigh distribution.

Real–rootedness and convergence to normal

If all roots of P_n are real then

$$\frac{P_n(x)}{P_n(1)}$$

is a p.g.f. of a sum of independent indicators, so the asymptotic normality holds as long as variance of the sum goes to infinity (usually easy to check from the recurrence for $P_n(x)$).

Conditions for the real–rootedness

- **L. L. Liu and Y. Wang (2007)** consider

$$P_n(x) = f_n(x)P_{n-1}(x) + g_n(x)P'_{n-1}(x) + h_n(x)P_{n-2}(x),$$

under extra assumptions (but not on the degrees f_n, g_n, and h_n). The minimum assumption is that

$$g_n(x) \le 0, \quad h_n(x) \le 0 \quad \text{for} \quad x \le 0.$$

- **D. Dominici, K. Driver, K Jordaan (2011)** consider

$$P_n(x) = f_n(x)P_{n-1}(x) + g_n(x)P'_{n-1}(x),$$

where f_n's have degrees at most 1 and g_n's at most 2.

A–B–N recurrence

Aval, Boussicault, Nadeau (2011) recurrence

$$B_n(x) \quad = \quad nx(x+1)B_{n-1}(x) + x(1-x^2)B'_{n-1}(x)$$

does not fall in either of the cases:

- $x(1-x^2) \leq 0$ fails for $x \leq 0$ and
- Both $x(1+x)$ and $x(1-x^2)$ have too high degrees.

A–B–N recurrence

Aval, Boussicault, Nadeau (2011) recurrence

$$B_n(x) \quad = \quad nx(x+1)B_{n-1}(x) + x(1-x^2)B'_{n-1}(x)$$

does not fall in either of the cases:

- $x(1-x^2) \leq 0$ fails for $x \leq 0$ and
- Both $x(1+x)$ and $x(1-x^2)$ have too high degrees.
- Their method can be adapted since all B_n's have common roots at -1 and 0 (**H. & A. Lohss**).
- Alternative approach (based on shifting the mean trick) has been developed by **H.–K. Hwang** and can be used to show asymptotic normality (but not real rootedness, of course).

A–B–N recurrence

Aval, Boussicault, Nadeau (2011) recurrence

$$B_n(x) \quad = \quad nx(x+1)B_{n-1}(x) + x(1-x^2)B'_{n-1}(x)$$

does not fall in either of the cases:

- $x(1-x^2) \leq 0$ fails for $x \leq 0$ and
- Both $x(1+x)$ and $x(1-x^2)$ have too high degrees.
- Their method can be adapted since all B_n's have common roots at -1 and 0 (**H. & A. Lohss**).
- Alternative approach (based on shifting the mean trick) has been developed by **H.–K. Hwang** and can be used to show asymptotic normality (but not real rootedness, of course).

Thank you :-)