# DIDACTICS

## OF

# MATHEMATICS

## 11(15)

# TABLE OF CONTENTS

# PENNEY'S GAME
# BETWEEN MANY PLAYERS

## Krzysztof Zajkowski

**Abstract.** We recall a combinatorial derivation of the functions generating the probability of winning for each of many participants of the Penney game and show a generalization of the Conway formula for this case.

**Keywords:** Penney's game, probability-generating functions, Conway's formula.

## 1. Introduction

In (2012), Wilkowski drew attention to the role of Penney's game in teaching probability at an elementary stage in economics. The purpose of the paper is to show a precise combinatorial approach to this subject. We think that questions that deal with Penney's game are also a good introduction to (discrete) stochastic processes.

Let us toss an 'unfair' coin with probabilities $p$ for heads ($H$) and $q = 1 - p$ for tails ($T$) and wait for the appearance of some chosen string of heads and tails. What is the expected number of tosses until this string occurs?

Let $A = a_1 a_2 ... a_l$ be a given pattern (a string of heads and tails) of the length $l$. By $P(A)$ we will denote a value $P(a_1)P(a_2) \cdot ... \cdot P(a_l)$. More precisely $P(A)$ is the probability of a cylindric set with some fixed coordinates: $a_1$, $a_2$,... and $a_l$, respectively. We flip a coin until we get $A$ as a run in the sequence of our trials. So we define the stopping time of the process in the following form

$$\tau_A = \min\{n \geq l : \xi_1 \xi_2 ... \xi_n, \xi_i \in \{H, T\} \text{ and } \xi_{n-l+1} = a_1, \xi_{n-l+2} = a_2, ..., \xi_n = a_l\},$$

**Krzysztof Zajkowski**
  Institute of Mathematics, University of Bialystok, Akademicka 2, 15-267 Bialystok, Poland.
  E-mail: kryza@math.uwb.edu.pl

if this minimum exists and $\infty$ if not. Now a more precise formulation of our question is: what is the expected value of $\tau_A$? An answer was first given by Solov'ev in (1966). In the paper presented we show some combinatorial solutions (compare Graham et. all, VIII.8.4), introducing at the same time the notations required and presenting a model reasoning.

Let $A_n$ denote the set of sequences in which the pattern $A$ appears exactly in the $n$-th toss, i.e. $A_n = \{\tau_A = n\}$, and $p_n$ the probability of $A_n$; $p_n = P(A_n)$. Let $B_n$ denote a set of sequences in which $A$ does not appear in the first $n$ tosses, i.e. $B_n = \{\tau_A > n\}$, and its probability by $q_n = P(B_n)$. Let us consider now a set of sequences in which $A$ does not appear in the first $n$ tosses and appears in the next $l$ trials, i.e. the set

$$\{(\xi_k) \in \{H,T\}^{\mathbb{N}} : (\xi_k) \in B_n \text{ and } \xi_{n+1} = a_1, \xi_{n+2} = a_2,...,\xi_{n+l} = a_l\}.$$

It seems that the probability of the set amounting to $q_n P(A)$ is equal to $p_{n+l}$ but we must check whether $A$ does not occur earlier in the trials from $n+1$ to $n+l-1$.

Let $A_{(k)}$ and $A^{(k)}$ denote strings of $k$-first and $k$-last terms of $A$ ($1 \le k \le l$), respectively. Note that $A_{(l)} = A^{(l)} = A$. Let $[A_{(k)} = A^{(k)}]$ equal 1 if $A_{(k)} = A^{(k)}$ or 0 if not. Additionally, let us assume that $P(A^{(0)}) = 1$. Now we can write the formula on $q_n P(A)$ as follows:

$$q_n P(A) = \sum_{k=1}^{l} [A_{(k)} = A^{(k)}] P(A^{(l-k)}) p_{n+k}. \tag{1}$$

Observe that the $l$-th summand in the above is equal to $p_{n+l}$. Remembering that $p_0 = p_1 = ... = p_{l-1} = 0$, multiplying the above equation by $s^{n+l}$ and summing from $n = 0$ to infinity we get

$$Q_{\tau_A}(s) P(A) s^l = g_{\tau_A}(s) \sum_{k=1}^{l} [A_{(k)} = A^{(k)}] P(A^{(l-k)}) s^{l-k}, \tag{2}$$

where $g_{\tau_A}(s) = \sum_{n=0}^{\infty} p_n s^n$ is the probability generating function for a random variable $\tau_A$ of the number of tosses until $A$ occurs and $Q_{\tau_A}(s) = \sum_{n=0}^{\infty} q_n s^n$ is the generating function of tail probabilities $q_n$. Because one can bound $\tau_A$ by a random variable with the geometric

distribution then one can show that $E\tau_A < \infty$. Hence $q_n = \sum_{k=n+1}^{\infty} p_k$ so we can obtain the second equation that relates $g_{\tau_A}$ and $Q_{\tau_A}$:

$$Q_{\tau_A}(s) = \frac{1 - g_{\tau_A}(s)}{1 - s}. \tag{3}$$

Solving the above two equations we get

$$g_{\tau_A}(s) = \frac{P(A)s^l}{P(A)s^l + (1-s)\sum_{k=1}^{l}[A_{(k)} = A^{(k)}]P(A^{(l-k)})s^{l-k}}$$

and

$$Q_{\tau_A}(s) = \frac{\sum_{k=1}^{l}[A_{(k)} = A^{(k)}]Pr(A^{(l-k)})s^{l-k}}{P(A)s^l + (1-s)\sum_{k=1}^{l}[A_{(k)} = A^{(k)}]P(A^{(l-k)})s^{l-k}}.$$

Since $E\tau_A = Q_{\tau_A}(1)$, one can calculate the general formula for the expected number of tosses as follows:

$$E\tau_A = \frac{\sum_{k=1}^{l}[A_{(k)} = A^{(k)}]P(A^{(l-k)})}{P(A)} = \sum_{k=1}^{l}\frac{[A_{(k)} = A^{(k)}]}{P(A_{(k)})}.$$

This is the answer to the question posed at the beginning.

   In the classical Penney Ante game (see: Penney 1974), for a given string of fixed length we want to show a second one of the same length with a higher probability to be the first to occur. In (Chen, Zame 1979), Chen and Zame proved that for two-person games, public knowledge of the opponent's string leads to an advantage. Guibas and Odlyzko (1981), showed some optimal strategy for the second player. An algorithm for computing the odds of winning for the competing patterns was discovered by Conway and described by Gardner (1974). The Conway formula allows us to compare the probability of winning for two players.

   In this paper we show a generalization of the Conway formula in the case of many gamblers (Section 3). But first (Section 2), we present a derivation and solution of the system of equations proposed by Guibas and Odlyzko (1981, Th. 3.3).

## 2. Functions generating probability of winning

Let $m$ players choose $m$ strings $A_i$ ($1 \le i \le m$) of heads and tails of lengths $l_i$, respectively. We start to toss an 'unfair' coin and wait for the occurrence of some $A_i$. We ask about the chances of winning for each player, that is about the probability $p_{A_i}$ that the string $A_i$ will be the first to occur. We assume that any $A_i$ is not a substring of other $A_j$, in the opposite case $p_{A_j} = 0$ or for some sequences both players may win simultaneously.

Let $\tau$ denote the number of tosses to the end of the game, i.e. $\tau = \min\{\tau_{A_i} : 1 \le i \le m\}$, where $\tau_{A_i}$ is the stopping time until pattern $A_i$ occurs. Notice that $P(\tau = n) = p_n = \sum_{i=1}^{m} p_n^{A_i}$, where $p_n^{A_i} = P(\tau = \tau_{A_i} = n)$ is the probability that the $i$-th player wins exactly in the $n$-th toss. Let $g_\tau$ and $g_\tau^{A_i}$ denote the functions generating distributions of probability $(p_n)$ and $(p_n^{A_i})$, respectively, and $Q_\tau$ the generating function of tail distributions $q_n = P(\tau > n)$.

Similarly, as in the introduction, let $B_n = \{\tau > n\}$ be the set of sequences of tails and heads in which any string $A_i$ does not appear in the first $n$ tosses. In the system of $m$ patterns, if we add the string $A_i$ to the set $B_n$ then we must check if neither $A_i$ nor other patterns appear earlier. For this reason a system of equations

$$q_n P(A_i) = \sum_{j=1}^{m} \sum_{k=1}^{\min\{l_i, l_j\}} [A_{i(k)} = A_j^{(k)}] P(A_i^{(l_i - k)}) p_{n+k}^{A_j},$$

for each $1 \le i \le m$, where $[A_{i(k)} = A_j^{(k)}] = 1$ if $A_{i(k)} = A_j^{(k)}$ or $0$ if not, corresponds to equation (1).

Multiplying the above equation by $s^{n+l_i}$ and summing from $n = 0$ to infinity, we get the following recurrence equations

$$Q_\tau(s) P(A_i) s^{l_i} = \sum_{j=1}^{m} g_\tau^{A_j}(s) \sum_{k=1}^{\min\{l_i, l_j\}} [A_{i(k)} = A_j^{(k)}] P(A_i^{(l_i - k)}) s^{l_i - k}.$$

Let $w_{A_i}^{A_j}(s)$ denote the polynomial $\sum_{k=1}^{\min\{l_i,l_j\}}[A_{i(k)} = A_j^{(k)}]P(A_i^{(l_i-k)})s^{l_i-k}$ ; now we can rewrite the above system of $m$ equations as follows

$$Q_\tau(s)P(A_i)s^{l_i} = \sum_{j=1}^{m}g_\tau^{A_j}(s)w_{A_i}^{A_j}(s) \quad (1 \le i \le m). \tag{4}$$

Since $g_\tau = \sum_{j=1}^{m}g_\tau^{A_j}$ , by virtue of (3), we get

$$Q_\tau(s) = \frac{1 - \sum_{j=1}^{m}g_\tau^{A_j}(s)}{1-s}.$$

Inserting the form of $Q_\tau$ into (4) we obtain

$$P(A_i)s^{l_i} = \sum_{j=1}^{m}g_\tau^{A_j}(s)[P(A_i)s^{l_i} + (1-s)w_{A_i}^{A_j}(s)] \quad (1 \le i \le m).$$

To solve this system of functional equations we use Cramer's rule. Define now functional matrices

$$\mathsf{A}(s) = \left(P(A_i)s^{l_i} + (1-s)w_{A_i}^{A_j}(s)\right)_{1\le i,j\le m}$$

and

$$\mathsf{B}(s) = \left(w_{A_i}^{A_j}(s)\right)_{1\le i,j\le m}. \tag{5}$$

Notice that because $w_{A_i}^{A_i}(0) = 1$, and $w_{A_i}^{A_j}(0) = 0$ for $i \ne j$ then $\mathsf{A}(0)$ and $\mathsf{B}(0)$ are the identity matrices. Let $\mathsf{B}^j(s)$ denote the matrix formed by replacing the $j$-th column of $\mathsf{B}(s)$ by the column vector $[P(A_i)s^{l_i}]_{1\le i\le m}$. Because the determinant of matrices $m \times m$ is a $m$-linear functional with respect to columns (equivalently to rows), then one can check that

$$\det \mathsf{A}(s) = (1-s)^m \det \mathsf{B}(s) + (1-s)^{m-1}\sum_{j=1}^{m}\det \mathsf{B}^j(s).$$

The determinant $\det \mathsf{A}(s)$ is a polynomial of variable $s$ and $\det \mathsf{A}(0) = 1$. For these reasons $\det \mathsf{A}(s) \ne 0$ in some neighborhood of zero. This means that in this neighborhood there exists a solution of the system.

If now, similarly, $\mathsf{A}^j(s)$ denotes the matrix formed by replacing the $j$-th column of $\mathsf{A}(s)$ by the column vector $[P(A_i)s^{l_i}]_{1\le i\le m}$, then the determinant's calculus gives $\det \mathsf{A}^j(s) = (1-s)^{m-1}\det \mathsf{B}^j(s)$. Finally, by Cramer's rule we obtain:

$$g_\tau^{A_i}(s) = \frac{\det \mathsf{A}^i(s)}{\det \mathsf{A}(s)} = \frac{\det \mathsf{B}^i(s)}{\displaystyle\sum_{j=1}^{m}\det \mathsf{B}^j(s) + (1-s)\det \mathsf{B}(s)}$$

for $1\le i\le m$. In this way we have proved the following

**Theorem 2.1.** *If $m$ players choose $m$ strings of heads and tails $A_i$ $(1\le i\le m)$ such that any $A_i$ is not a substring of another $A_j$, then the function $g_\tau^{A_i}$ generating the probability of winning of the i-th player is given by the following formula:*

$$g_\tau^{A_i}(s) = \frac{\det \mathsf{B}^i(s)}{\displaystyle\sum_{j=1}^{m}\det \mathsf{B}^j(s) + (1-s)\det \mathsf{B}(s)}, \tag{6}$$

*where $\mathsf{B}(s)$ is the matrix defined by (5).*

Notice that the probability generating function $g_\tau^{A_i}(s) = \sum_{n=0}^{\infty} p_n^{A_i} s^n$ is undoubtedly well defined on the interval $[-1,1]$ (it is an analytic function on $(-1,1)$). The right hand side of (6) is a rational function equal to $g_\tau^{A_i}$ in the neighborhood of zero. By analytic extension we know that there exists the limit of the right hand side of (6) by $s\to 1^-$ which is equal to $g_\tau^{A_i}(1)$. Thus the probability $p_{A_i}$ that the string $A_i$ occurs first is given by the following formula

$$g_\tau^{A_i}(1) = \frac{\det \mathsf{B}^i(1)}{\displaystyle\sum_{j=1}^{m}\det \mathsf{B}^j(1)}, \tag{7}$$

where the right hand side of the above equation is understood as the limit of (6) with $s$ approaching the left-side to 1 ($s\to 1^-$).

## 3. A generalization of Conway's formula

Define a number $A_j : A_i$ as

$$\frac{w_{A_i}^{A_j}(1)}{P(A_i)} = \sum_{k=1}^{\min\{l_i, l_j\}} [A_{i(k)} = A_j^{(k)}] \frac{P(A_i^{(l_i - k)})}{P(A_i)} = \sum_{k=1}^{\min\{l_i, l_j\}} \frac{[A_{i(k)} = A_j^{(k)}]}{P(A_{i(k)})}.$$

For a balanced coin ( $p = q = \dfrac{1}{2}$ )

$$A_j : A_i = \sum_{k=1}^{\min\{l_i, l_j\}} [A_{i(k)} = A_j^{(k)}] 2^k$$

and it coincides (up to the scalar 2) with the notation introduced in (Graham, Knuth, Patashnik 1989).

Define now a matrix $\mathsf{C} = \left( (A_j : A_i) \right)_{1 \leq i, j \leq m}$. Observe that $\det \mathsf{B}(1) = \prod_{i=1}^{m} P(A_i) \det \mathsf{C}$ and $\det \mathsf{B}^j(1) = \prod_{i=1}^{m} P(A_i) \det \mathsf{C}^j$, where $\mathsf{C}^j$ is the matrix formed by replacing the $j$-th column of $\mathsf{C}$ by the column vector of units $(1)_{1 \leq i \leq m}$. Due to (7) and the above observations we can formulate the following

**Corollary 3.1.** *The probability that the $i$-th player wins is equal to*

$$p_{A_i} = \frac{\det \mathsf{C}^i}{\sum_{j=1}^{m} \det \mathsf{C}^j}.$$

Let us emphasize that the above Corollary is a generalization of Conway's formula. For two players we get

$$\frac{p_{A_1}}{p_{A_2}} = \frac{\det \mathsf{C}^1}{\det \mathsf{C}^2} = \det \begin{pmatrix} 1 & (A_2 : A_1) \\ 1 & (A_2 : A_2) \end{pmatrix} : \det \begin{pmatrix} (A_1 : A_1) & 1 \\ (A_1 : A_2) & 1 \end{pmatrix} = \frac{(A_2 : A_2) - (A_2 : A_1)}{(A_1 : A_1) - (A_1 : A_2)}.$$

**Example 3.2.** Take three strings of heads and tails: $A_1 = THH$, $A_2 = HTH$ and $A_3 = HHT$. In this case

$$\mathsf{B}(s) = \left( w_{A_i}^{A_j}(s) \right)_{1 \leq i, j \leq 3} = \begin{pmatrix} 1 & ps & p^2 s^2 \\ pqs^2 & pqs^2 + 1 & ps \\ pqs^2 + qs & pqs^2 & 1 \end{pmatrix}.$$

By Theorem 2.1 one can obtain the probability generating functions for winnings of *i*-th player. The matrix

$$
\mathsf{C} = \left( \frac{w_{A_i}^{A_j}(1)}{P(A_i)} \right)_{1 \le i,j \le 3} =
\begin{pmatrix}
\dfrac{1}{p^2 q} & \dfrac{1}{pq} & \dfrac{1}{q} \\[2ex]
\dfrac{1}{p} & \dfrac{pq+1}{p^2 q} & \dfrac{1}{pq} \\[2ex]
\dfrac{p+1}{p^2} & \dfrac{1}{p} & \dfrac{1}{p^2 q}
\end{pmatrix}
$$

and

$$
\mathsf{C}^1 =
\begin{pmatrix}
1 & \dfrac{1}{pq} & \dfrac{1}{q} \\[2ex]
1 & \dfrac{pq+1}{p^2 q} & \dfrac{1}{pq} \\[2ex]
1 & \dfrac{1}{p} & \dfrac{1}{p^2 q}
\end{pmatrix},\quad
\mathsf{C}^2 =
\begin{pmatrix}
\dfrac{1}{p^2 q} & 1 & \dfrac{1}{q} \\[2ex]
\dfrac{1}{p} & 1 & \dfrac{1}{pq} \\[2ex]
\dfrac{p+1}{p^2} & 1 & \dfrac{1}{p^2 q}
\end{pmatrix},\quad
\mathsf{C}^3 =
\begin{pmatrix}
\dfrac{1}{p^2 q} & \dfrac{1}{pq} & 1 \\[2ex]
\dfrac{1}{p} & \dfrac{pq+1}{p^2 q} & 1 \\[2ex]
\dfrac{p+1}{p^2} & \dfrac{1}{p} & 1
\end{pmatrix}.
$$

On the basis of Corollary 3.1 we can calculate the probability that the *i*-th player wins:

$$
p_{A_1} = \frac{q(1+pq)}{1+q},\quad p_{A_2} = \frac{q}{1+q},\quad p_{A_3} = p^2.
$$

For a fair coin

$$
p_{A_1} = \frac{5}{12},\quad p_{A_2} = \frac{1}{3} \quad \text{and} \quad p_{A_3} = \frac{1}{4}.
$$

## 4. Conclusions

Sequences of Bernoulli trials are the first historical example of discrete stochastic processes. Questions dealing with the appearances of the chosen strings are effortlessly formulated. For instance, in teaching elementary probability, waiting up to the first success is one of the basic models of infinite probability spaces. The problems of Penney's game develop this approach to the subject and may serve as a good introduction to statistics and the theory of stochastic processes.

The presented combinatorial derivation of the formulas for the chances of winning of many players is an example of the applications of the determinant calculus and it shows how different techniques of mathematics penetrate each other and lead to the solutions of given problems.

## References

Chen R., Zame A. (1979). *On the fair coin-tossing games*. J. Multivariate Anal. Vol. 9, pp. 150-157.

Gardner M. (1974). *On the paradoxical situations that arise from nontransitive relations*. Scientific American 231 (4), pp. 120-124.

Guibas L.J., Odlyzko A.M. (1981). *String overlaps, pattern matching, and nontransitive games*. Journal of Combinatorial Theory (A) 30, pp. 183-208.

Graham R.L., Knuth D.E. and Patashnik O. (1989). *Concrete Mathematics: a Foundation for Computer Science*. Addison-Wesley Publishing Company.

Penney W. (1974). *Problem 95: Penney-Ante*. Journal of Recreational Mathematics. No 7, p. 321.

Solov'ev A.D.(1966). *A combinatorial identity and its application to the problem concerning the first occurrence of a rare event*. Theory of Probability and its Applications 11, pp. 313-320.

Wilkowski A. (2012). *Penney's game in didactics*, Didactics of Mathematics. No 10(14), pp. 77-86.