D I D A C T I C S O F M A T H E M A T I C S

No. 10 (14) 2013

PENNEY'S GAME IN DIDACTICS

Andrzej Wilkowski

Abstract. This paper describes, at elementary level, Penney's game using the example of two players and a symmetric coin. It also provides a generalization for an unlimited number of players and coins, as an example, not an intuitive aspect of the teaching probability theory.

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

DOI: 10.15611/dm.2013.10.07

1. Introduction

Probability, in universities teaching economics, is usually discussed in the framework of the basic course of mathematics or statistics. Therefore relatively little time is devoted to this. Nevertheless it is still worth devoting some of the time to present problems that do not require too sophisticated theoretical apparatus, while their solutions are non-intuitive at the same time. It is also worth mentioning the paradox of Bertrand (Jakubowski, Sztencel 2000; Wilkowski 2007) when defining the probability space, or the concept of elementary events. While talking about the likelihood of the frequentist, one should mention specifically Chebyshev (Jagłom 1954; Nieznaj 2002), the problem of drawing a natural number (Adamaszek 2005). We also discuss the birthday paradox, when talking about the classic definition of probability (Adamaszek 2010; Nikodem 2010; Wilkowski 2010). The Bernoulli scheme, the concept of fair play, is naturally associated with the Penney game (Penney 1974; Gardner 1974; Nishiyama 2012), which this work is dedicated to.

Andrzej Wilkowski
Department of Mathematics and Cybernetics, Wrocław University of Economics, Komandorska Street 118/120, 53-345 Wrocław, Poland.
E-mail: andrzej.wilkowski@ue.wroc.pl

2. Penney's game between two players in fair coin case

This point is based on the work (Dniestrzański, Wilkowski 2008). We will now give an example of failing intuition in a game based on symmetric coin throwing. The game's description requires a short introduction.

Let X be a random variable accepting non-negative integers:

$$X : \Omega \to \{0,1,\ldots\}.$$

Further we continue to assume that X has a finite second moment:

$$E\left(X^2\right) < \infty .$$

When studying distributions of such type of random variables, it is convenient to use the **probability-generating function** of X, defined as a formal power series:

$$G_X(z) = \sum_{k=0}^{\infty} P\left(\omega : X(\omega) = k\right) z^k = E\left(z^X\right). \tag{1}$$

The following series of variable z contains all the information about the random variable X. One can see that:

$$G_X(1) = 1.$$

Conversely, each power series $G(z)$ of non-negative coefficients, satisfying the equation $G(1) = 1$ is the probability generating function of a particular random variable. An important feature of this function is that it simplifies the calculation of the mean and variance of the random variable X. To achieve this, it is suffice to set the first and second derivative of the series (1), for $z = 1$, and take a combination thereof. We get:

$$E(X) = G_X^{'}(1), \tag{2}$$

$$Var(X) = E\left(X^2\right) - \left(E(X)\right)^2 = G_X^{''}(1) + G_X^{'}(1) - \left(G_X^{'}(1)\right)^2 . \tag{3}$$

Example 1. Formulas (2) and (3) will be used in the case of a process that has only two results. When we toss a coin, the probability that we get heads (H), is p, and the probability of tails (T), is equal to q, where

$$p + q = 1, \quad p \text{ and } q > 0.$$

For a fair coin $p = q = \frac{1}{2}$. However, this is not always the case. As described in the book (Graham et al. 1989) in the case of a newly minted American one cent coin (penny) we get $p \approx 0{,}1$ (the weight distribution makes Lincoln fall to the bottom more often). Now let X_A be a random variable describing the number of independent coin tosses until the outcome of A = THHHH (string of heads and tails) is obtained for the first time. One should determine the mean and variance of this variable. We use the method given in the manual (Graham et al. 1989). Let S denote the sum of all possible outcomes which contain pattern A:

$$S = THHHH + HTHHHH + TTHHHH + \ldots$$

The N is the sum of all possible outcomes in which the pattern A does not appear:

$$N = 1 + H + T + HH + HT + TH + TT + \ldots$$

In view of the above, the relationship between S and N are true:

$$1 + N(H + T) = N + S,$$

$$NTHHHH = S.$$

When H is replaced by a pz and T by qz and then, from the above presented relationships, we determine S, we obtain a function forming $G_A(z)$, of a random variable X_A:

$$G_A(z) = \frac{p^4 q z^5}{p^4 q z^5 - pz - qz + 1}.$$

Thus, on the basis of formulas (2), (3), we have:

$$E(X_A) = p^{-4} q^{-1},$$

$$Var(X_A) = p^{-8} q^{-2} - 9 p^{-4} q^{-1}.$$

When $p = q = \frac{1}{2}$, we get: $E(X_A) = 32$, $Var(X_A) = 736$.

The reasoning presented in this example can be generalized. The following theorem can be proved.

Theorem 1 (Graham et al. 1989). Let X_A be a random variable describing the number of individual tosses of a coin, till the first appearance of pattern A (a string of heads and tails) with the length of m = 1, 2, ... Let us assume that the probability of occurrence of heads (H) is p, the probability of tails (T), will be equal to q, where $p + q = 1$, $p > 0$, $q > 0$. Then

$$E\left(X_A\right) = \sum_{k=1}^{m} \bar{A}_{(k)}\left[A^{(k)} = A_{(k)}\right],$$

$$Var\left(X_A\right) = (E\left(X_A\right))^2 - \sum_{k=1}^{m}\left(2k-1\right)\bar{A}_{(k)}\left[A^{(k)} = A_{(k)}\right],$$

where $A^{(k)}$ and $A_{(k)}$ denote, respectively, the last k and first k elements of A pattern. $\bar{A}$ is the result of the substitution of p^{-1} for H and q^{-1} for T in pattern A, as for the square brackets [...] takes the value of 1, when the expression inside is true or 0 otherwise.

We assume again that the coin is balanced (fair), that is $p = q = \frac{1}{2}$. For a given pattern A of length l and pattern B consisting of m heads and tails let:

$$A : A = \sum_{k=1}^{l} 2^{k-1}[A^{(k)} = A_{(k)}], \tag{4}$$

$$A : B = \sum_{k=1}^{\min(l,m)} 2^{k-1}\left[A^{(k)} = B_{(k)}\right]. \tag{5}$$

We see that in general $A : B \neq B : A$.

With these values, based on Theorem 1, we have

$$E\ (X_A) = 2(A{:}A). \tag{6}$$

The formula (6) was shown for the first time in the work of (Soloviev 1966). This result seems at first sight paradoxical: patterns that do not overlap occur more often than the overlapping patterns!

Example 2. Let A = THHHH, B = HHHHH be two strings of heads and tails with independent coin tosses with a balanced (fair coin). Then, $E(X_A) = 32$, $E(X_B) = 62$ Waiting for the toss pattern B to occur takes almost twice as much time as waiting for the appearance of pattern A.

An interesting game associated with tossing a coin was proposed in 1969 by Walter Penney (Penney 1974). In Penney's Game there are two players involved. The first one selects the pattern A = HHT, the second player chooses pattern B = HTT. The winner is the player whose pattern appears as the first one, with independent fair coin tosses (it is known that at some point it will happen, and there will never be a tie because none of these patterns can occur inside the other). This game seems to be fair because patterns A and B when treated separately look very similar, and the functions generating the probability of random variables X_A and X_B equal:

$$G_A(z) = G_B(z) = \frac{z^3}{z^3 - 8(z-1)} .$$

It turns out, however, that when we analyze these two patterns simultaneously, one of them has the upper hand, the probability of the event that pattern A wins over B is different than the probability of the event that B wins over pattern A (Graham et al.1989). We have: $P(A \textit{ wins against } B) = 2/3$, $P(B \textit{ wins against } A) = 1/3$. The general formula for this type of problems was discovered by John Horton Conway (Gardner 1974).

Theorem 2. Let A and B be arbitrary patterns, not necessarily of equal length, of heads and tails, with independent tosses with a fair coin in Penney's game. Let us assume that pattern A is not contained in B, neither B is contained in pattern A. Then

$$\frac{P(A \, wins \, against \, B)}{P(B \, wins \, against \, A)} = \frac{B:B - B:A}{A:A - A:B}, \tag{7}$$

where the symbols on the right side of the equation are defined by the formulas (4) and (5).

Conclusion 1. For any pattern $A = a_1a_2 \ldots a_m$ and $B = (-a_2)a_1a_2 \ldots a_{m-1}$ we have:

$$P(A \textit{ wins against } B) < P(B \textit{ wins against } A),$$

where $m > 2$, and $(-a_2)$ is a heads-and-tails inversion of a_2.

Conclusion 2. For the given pattern $a_1a_2 \ldots a_m$, the biggest chances of a winning result from selecting one of two patterns: $Ha_1a_2 \ldots a_{m-1}$ or $Ta_1a_2 \ldots a_{m-1}$, $m > 2$ (Guibas, Odlyzko 1981).

Example 3. Let the patterns be: A = TTHH, B = TTT. From equation (7) it results that:

$$\frac{P(A\,wins\,against\;\;B)}{P(B\,wins\,against\;\;A)}=\frac{7}{5}.$$

In Penney's game it happens that the longer pattern wins against the shorter one.

The described game is another example of the unreliability of intuition in probabilistic issues. One can even talk here about unreliability on two levels: professional and amateur. We suspect that if a number of people, not professionally involved in the mathematics, was asked to determine which of the patterns, A=THHHH and B = HHHHH (Example 2), has a better chance of winning, the vast majority would state that it was pattern A, pattern B might not seem realistic and rare. So the answer would be correct. However, the "professionals" would most likely give the two patterns equal chances of occurrence. Their intuition (at this point unreliable) would be based on the knowledge that the likelihood of a four-coin-toss results of the two strings is the same. In turn, the ones not having this knowledge – mathematical laymen – would be misled by the intuition already at the four times coin toss. As in the case with Penney's game, they would tend to favor string A rather than B.

3. Penney's game in a general case

This section generalizes the considerations set out earlier. These considerations are based on the work of (Zajkowski 2012). Suppose that m players choose m patterns A_i ($1 \le i \le m$) of length l_i, respectively. Let us also assume that the coin is not a 'fair` one. Let p_{A_i} indicate the probability that pattern A_i will appear before other patterns with independent coin tosses ($p_{A_i} = P(A_i$ wins against the others)). The random variable X is the number of tosses till the end of the game i.e. until any pattern will appear for the first time (which is a certain occurrence). Let us note that $P\left(X=n\right)=p_n=\sum_{i=1}^{m}p_n^{A_i}$, where $p_n^{A_i}$ is the probability that the i-th player wins exactly in n-th toss. The generating functions of the strings (p_n) and $\left(p_n^{A_i}\right)$ will be marked G_X i $G_X^{A_i}$, respectively. Please note that:

$$G_X(z)=\sum_{j=1}^{m} G_X^{A_j}(z).$$

We shall define a polynomial $w_{A_i}^{A_j}$, as"

$$w_{A_i}^{A_j}(z)=\sum_{k=1}^{\min\{l_i\ ,\ l_j\}}\left[A_{i(k)}=A_j^{(k)}\right]P\left(A_i^{(l_i-k)}\right)z^{l_i-k}\ . \tag{8}$$

Let us recall that square brackets [...] take the value of 1 when the statement inside is true or 0 in the other case. Let us define the matrix now

$$\boldsymbol{A}(z)=\left(w_{A_i}^{A_j}(z)\right)_{1\le i,\,j\le m}, \tag{9}$$

where the polynomial is defined with the formula (8). The symbol $\boldsymbol{A}^j(z)$ denotes the matrix obtained from the matrix $\boldsymbol{A}(z)$ from the formula (9), after replacing its j-th column with a vector $(P\left(A_i\right)z^{l_i})_{1\le i\le m}$.

Theorem 3 (Zajkowski 2012). If m players chose m strings of heads and tails A_i $(1\le i\le m)$ such that any A_i is not a substring of other A_j then the probability-generating function $G_X^{A_i}$ of winning of the i-th player is given by the formula

$$G_X^{A_i}(z)=\frac{\det \boldsymbol{A}^i(z)}{\sum_{j=1}^{m}\det \boldsymbol{A}^j(z)+(1-z)\det \boldsymbol{A}(z)},$$

where $\boldsymbol{A}(z)$ is the matrix defined by (9).

Conclusion 3. The probability that string A_i occurs first is equal to

$$p_{A_i}=G_X^{A_i}(1)=\frac{\det \boldsymbol{A}^i(1)}{\sum_{j=1}^{m}\det \boldsymbol{A}^j(1)}\ .$$

Define a number A_j: A_i (generalization (5) in the "unfair" coin case) as

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

Define now a matrix $\boldsymbol{B}=(A_j:A_i)_{1\le i\ ,\ j\le m}$, $\boldsymbol{B}^j$ is the matrix formed by replacing the j-th column of $\boldsymbol{B}$ by the column vector $(1)_{1\le i\le m}$.

Corollary 1 (Zajkowski 2012). The probability that the i-th player wins is equal to

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

Example 4. Take three strings of heads (H) and tails (T): A_1 = HTT, A_2 = THT, A_3 = TTH . In this case

$$\boldsymbol{A}(z) = (w_{A_i}^{A_j}(z))_{1 \le i\ ,j \le 3} = \begin{pmatrix} 1 & qz & q^2 z^2 \\ pqz^2 & pqz^2 + 1 & qz \\ pqz^2 + pz & pqz^2 & 1 \end{pmatrix}.$$

By Theorem 3 one can obtain the probability-generating functions for the winnings of i-th player (the number of coin tosses until the i-th player wins).

Matrix $\boldsymbol{B}$ is equal to

$$\boldsymbol{B} = \left(\frac{w_{A_i}^{A_j}(1)}{P(A_i)} \right)_{1 \le i\ ,j \le 3} = \begin{pmatrix} \frac{1}{pq^2} & \frac{1}{pq} & \frac{1}{p} \\ \frac{1}{q} & \frac{pq+1}{pq^2} & \frac{1}{pq} \\ \frac{q+1}{q^2} & \frac{1}{q} & \frac{1}{pq^2} \end{pmatrix}.$$

By corollary 4 we can calculate the probability that the i-th player wins:

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

Example 5. Take three strings of heads (H) and tails (T): A_1 = HHTH, A_2 = HTHH, A_3 = THHH. We assume that $p = q = \frac{1}{2}$ (Graham at al. 1989). By corollary 4 we can calculate the probability that the i-th player wins:

$$p_{A_1} = \frac{16}{52},\ p_{A_2} = \frac{17}{52},\ p_{A_3} = \frac{19}{52}.$$

Example 6. Suppose that two players choose strings of heads (H) and tails (T): A_1 = TTH, A_2 = THT, a coin does not have to be symmetrical, $p, q \in (0, 1), p + q = 1$. On the basis of proposal 4, we have:

$$p_{A_1} = \frac{q^2}{1 - p - p^2 q}.$$

Note that $p_{A_1} = \frac{1}{2} \leftrightarrow (p-1)^3 = 0$ iff $p = 1$. For every coin string A_1 is more likely to occur than A_2.

The following table is based on the value of p (the probability of ejecting heads (H)), for which a two-person Penney`s game length of 3 becomes a fair game (the probability of one model against another is equal to 1/2). NP means not possible.

Table 1. The probability of ejecting heads (H))

	TTT	*TTH*	*THT*	*HTT*	*HHT*	*HTH*	*THH*	*HHH*
TTT	*x*	*0.5*	*0.28*	*0.21*	*0.4*	*0.39*	*0.44*	*0.5*
TTH	*0.5*	*x*	*np.*	*0.29*	*0.5*	*0.6*	*0.62*	*0.6*
THT	*0.28*	*NP*	*X*	*0.5*	*0.4*	*0.5*	*0.5*	*0.6*
HTT	*0.2*	*0.29*	*0.5*	*x*	*0.38*	*0.5*	*0.5*	*0.55*
HHT	*0.4*	*0.5*	*0.4*	*0.38*	*x*	*NP*	*0.7*	*0.5*
HTH	*0.39*	*0.6*	*0.5*	*0.5*	*NP*	*x*	*0.5*	*0.72*
THH	*0.44*	*0.62*	*0.5*	*0.5*	*0.7*	*0.5*	*x*	*0.79*
HHH	*0.5*	*0.6*	*0.61*	*0.55*	*0.5*	*0.72*	*0.79*	*x*

Source: own study.

Example 7. Suppose we have 2^n players and a fair coin $(p = q = \frac{1}{2})$. Each player chooses a different string of heads and tails of length n. The Penney game becomes a fair game, and we have:

$$p_{A_1} = \ldots = p_{A_{2^n}} = \frac{1}{2^n}.$$

4. Conclusions

Notice that the above results are true not only for binary strings but also for strings that take many values. On the basis of considerations within this article, one will find that even the classic and simple discrete probabilistic issues such as independent coin tosses, may lead to a situation not entirely consistent with intuition. Therefore it is worth to take up these types of problems when teaching probability and statistics. Penney`s game is a handy tool when discussing the theory of teaching countable probability space. It does not require too sophisticated mathematical apparatus (to keep it short, one can skip the last point and limit it to situations when the coins are symmetrical). Introducing this game within the statistics or probability theory courses provides some motivation to the process of creating and testing countable probability space. It also allows to analyze the concept of justice games and gives motivation to form both probabilistic and mathematical problems and tasks.

References

Adamaszek M. (2005). *Losowanie liczby naturalnej*. Delta. No 7. Warszawa.

Adamaszek M. (2010). *Potrójne urodziny*. Delta. No 4. Warszawa.

Dniestrzański P., Wilkowski A. (2008). *O paradoksie Halla i rzucaniu monetą*. Didactics of Mathematics. No. 5-6(9-10). Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu. Pp. 43-52.

Gardner M. (1974). *On the paradoxical situations that arise from non-transitive relations*. Scientific American 231. Vol. 4.

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

Guibas L., Odlyzko A. (1981). *String overlaps, pattern matching, and non-transitive games*. Journal of Combinatorial Theory. Series A. Vol. 30. Pp.183-208.

Jakubowski J., Sztencel R. (2000). *Wstęp do teorii prawdopodobieństwa*. Script. Warszawa.

Jagłom A.M., Jagłom I.M. (1954). *Nieelemientarnyje zadaczi w elemientarnym izłożenii*. Moskwa.

Nieznaj E. (2002). *Zadanie Czebyszewa*. Delta. No 2. Warszawa.

Nikodem T. (2010). *Paradoks dni urodzin i pokrewne, czyli o pewnych zagadnieniach związanych z rozmieszczeniem kul w komórkach*. Delta. No 4. Warszawa.

Nishiyama Y. (2012). Pattern matching probabilities *and* paradoxes – *A* new variation *on Penney's coin game*. Osaka Keidai Ronhsu, Vol. 63 No. 4. November 2012.

Penney W. (1974). *Problem 95: Penney-Ante,* Journal of Recreational Mathematics 7.

Smoluk A. (2002). *Co jest przedmiotem rachunku prawdopodobieństwa?* Ekonomia Matematyczna 6. Wydawnictwo Akademii Ekonomicznej we Wrocławiu. Pp. 27-48.

Sołowiew A. (1966). *A combinatorial identity and its application to the problem concerning the first occurrence of a rare even*. Theory of Probability and its Applications 11 (tłumaczenie z rosyjskiego).

Wilkowski A. (2007). *Uwagi o prawdopodobieństwie częstościowym i przestrzeni probabilistycznej*. Didactics of Mathematics No. 4(8). Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu. Pp. 79-84

Wilkowski A. (2010). *On changing money and the birthday paradox*. Didactics of Mathematics. No 7(11). Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu. Pp. 127-134.

Zajkowski K. (2012). *Penney's game between many players*. arXiv:1212.3973v1 [math.PR] 17 Dec 2012.