ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 162, 2002

Wiesław Pasewicz[*], *Wiesław Wagner*[**]

PROBABILITY MODEL OF WINNING TENNIS MATCH

ABSTRACT. Probability model on match involving two opposing players is discussed with particular emphasis on the relative probabity of a server in a play. It is assumed that player A has a constant probability p_A of winning any point while he is serving and that player B has a constant p_B of winning any point on his service. Tennis match consists of either the best 2 out of 3 sets or the best 3 out of 5 sets.

Expressions for the probability that a player wins a match are obtained. In order to simplify determination the probability of winning a match the special probability matrices are used. We present a simple numerical example for the illustration calculating the probability of winning a match.

Key words: Modulo 2 operation, probability matrices of winning a set, the probability of winning tennis match, service principle.

I. INDRODUCTION

The simplest model in analysing tennismatches is based on the assumption that two fixed probabilities govern a match: the probability of winning a service point for both players. Then, one can calculate the probability of winning a game or set (see: H s i and B u r y c h (1971) or C a r t e r and C r o w s (1974)). It seems to be natural calculating the probability of winning a match. Notice that the probability of winning a set is not equal the product of the probabilities of winning and losing games of the set by both players. Similarly, the probability of winning a match is not immediately stated as the product of the probabilities of winning and losing sets of the match.

[*] Dr., Department of Mathematics Agricultural University in Szczecin.
[**] Prof., Institute of Statistics, University School of Physical Education, Poznań.

In this paper we will discuss probability model of winning tennis match by one of two players. We assume that player A wins each point of his service with probability p_A and player B wins each point of his service with probability p_B. For illustrating our model we present a numerical example.

II. FORMAL DESCRIPTION STATES OF MATCH

In tennis there are two cases playing of match:

(a) the best 2 out of 3 sets (2/3),

(b) the best 3 out of 5 sets (3/5).

Let us assume that player A wins a match. We have to consider the following posibilites:

$\Rightarrow (a)$: player A wins two first sets with match 2 : 0, or loses the first set and wins the next two or loses the second one and wins the remaining two, with result of match 2 : 1.

$\Rightarrow (b)$: player A wins three first sets with match 3 : 0, or wins the fourth set and loses one of three played before with match 3 : 1 or wins the fifth set but loses exactly two of four played before with match 3 : 2.

Let S_A ($\overline{S}_A$) denote an event that player A wins (loses) a set. Denote by $M_A(2/3)$ or $M_A(3/5)$ events of winning the match by player A in case (a) and (b), respectively. At last let $M_A(m:k)$, $m = 2, 3$; $k = 0, 1, ..., m-1$ denote an event „player A won the match with result $m : k$". Now, we can write:

$$\Rightarrow (a): M_A(2:0) = S_A \cap S_A,$$

$$M_A(2:1) = (\overline{S}_A \cap S_A \cap S_A) \cup (S_A \cap \overline{S}_A \cap S_A),$$

$$M_A(2/3) = M_A(2:0) \cup M_A(2:1);$$

$$\Rightarrow (b): M_A(3:0) = S_A \cap S_A \cap S_A,$$

$$M_A(3:1) = (\overline{S}_A \cap S_A \cap S_A \cap S_A) \cup (S_A \cap \overline{S}_A \cap S_A \cap S_A) \cup \cup (S_A \cap S_A \cap \overline{S}_A \cap S_A),$$

$$M_A(3:2) = (\overline{S}_A \cap \overline{S}_A \cap S_A \cap S_A \cap S_A) \cup \cup (\overline{S}_A \cap S_A \cap \overline{S}_A \cap S_A \cap S_A) \cup \cup (\overline{S}_A \cap S_A \cap S_A \cap \overline{S}_A \cap S_A) \cup \cup (S_A \cap \overline{S}_A \cap \overline{S}_A \cap S_A \cap S_A) \cup \cup (S_A \cap \overline{S}_A \cap S_A \cap \overline{S}_A \cap S_A) \cup \cup (S_A \cap S_A \cap \overline{S}_A \cap \overline{S}_A \cap S_A),$$

$$M_A(3/5) = M_A(3:0) \cup M_A(3:1) \cup M_A(3:2).$$

Using the transcription: $S_A \to 1$ and $\overline{S}_A \to 0$ we have:

$\Rightarrow (a): M_A(2:0) = (11)$, $M_A(2:1) = (011) \cup (101)$;

$\Rightarrow (b): M_A(3:0) = (111)$, $M_A(3:1) = (0111) \cup (1011) \cup (1101)$,

$M_A(3:2) = (00111) \cup (01011) \cup (01101) \cup (10011) \cup$
$\cup (10101) \cup (11001)$.

The number of components for $M_A(m:k)$ is equal $\binom{m+k-1}{k}$, for $m = 2, 3$; $k = 0,1, ..., m-1$. Analogous formulas can be given for player B. We would like to emphasize that the probability of winning a set depends on the number played games – odd or even (see e.g. Pollard 1983, Riddle 1988, Wagner and Majewska 1996, Pasewicz and Wagner 2000), where in particular was considered the tie-breaker set if the games score reached 6 games each. It is very important who of players is serving first. Therefore, we have to consider both the aspects calculating the probability of winning tennis match.

III. PROBABILITY MODEL OF WINNING A MATCH

Consider the following events for an individual set:

$T_1(A)$ – „player A serves first",

$T_2(A)$ – „player A wins in an even number of games",

$T_3(A)$ – „player A wins in an odd number of games".

Similar events can be given for player B. Using the concatenation principle we can write:

$A_{00} = T_1(A) \,\&\, T_2(A)$, $A_{01} = T_1(A) \,\&\, T_3(A)$,

$A_{10} = T_1(B) \,\&\, T_2(A)$, $A_{11} = T_1(B) \,\&\, T_3(A)$,

$B_{00} = T_1(A) \,\&\, T_2(B)$, $B_{01} = T_1(A) \,\&\, T_3(B)$,

$B_{10} = T_1(B) \,\&\, T_2(B)$, $B_{11} = T_1(B) \,\&\, T_3(B)$,

where $A_{ij}(B_{ij})$, $i = 0,1; j = 0,1$ denote events that player A (player B) wins a set. The first subscript indicates which player served the first game of the set (0 for A and 1 for B), while the second subscript indicates the modulo 2 $((i + j) \bmod 2)$ that is (Riddle 1988):

$$(0+0) \bmod 2 = 0,\ (0+1) \bmod 2 = 1,\ (1+0) \bmod 2 = 1,\ (1+1) \bmod 2 = 0.$$

Let $a_{ij} = P(A_{ij})$ and $b_{ij} = P(B_{ij})$ denote the probabilities of events A_{ij} and B_{ij}, respectively. First we calculate the probability that player A wins a match in case (a) with result 2 : 0. In order to consider possible events A_{ij} in the case we present the service principle in tennis (tab. 1).

Table 1

The events for a match with 2:0

The events	Set I			Set II			Product of events
	f	e	o	f	e	o	
1	A	*		A	*		$A_{00}A_{00}$
2	A	*		A		*	$A_{00}A_{01}$
3	A		*	B	*		$A_{01}A_{10}$
4	A		*	B		*	$A_{01}A_{11}$

In table 1 the letters f, e and o denote a player serves first, an even number of games and an odd number of games in a set, respectively. Because of

$$M_A(2:0) = (A_{00} \cap A_{00}) \cup (A_{00} \cap A_{01}) \cup (A_{01} \cap A_{10}) \cup (A_{01} \cap A_{11}),$$

the probability p_1 that A wins the match with 2 : 0 has the following form:

$$p_1 = P(M_A(2:0)) = a_{00}^2 + a_{00}a_{01} + a_{01}a_{10} + a_{01}a_{11}.$$

Now, we calculate the probability p_2 that player A wins a match with 2 : 1 and loses the first set. We give the second table of service principle (tab. 2).

Table 2

The events for a match with 2:1 (A loses the first set)

The events	$\overline{S}_A$			S_A			S_A			Probabilities
	f	e	o	f	e	o	f	e	o	
1	A	*		A	*		A	*		$b_{00}a_{00}a_{00}$
2	A	*		A	*		A		*	$b_{00}a_{00}a_{01}$
3	A	*		A		*	B	*		$b_{00}a_{01}a_{10}$
4	A	*		A		*	B		*	$b_{00}a_{01}a_{11}$
5	A		*	B	*		B	*		$b_{01}a_{10}a_{10}$
6	A		*	B	*		B		*	$b_{01}a_{10}a_{11}$
7	A		*	B		*	A	*		$b_{01}a_{11}a_{00}$
8	A		*	B		*	A		*	$b_{01}a_{11}a_{01}$

Using results from Table 2, we have

$$p_2 = P(\bar{S}_A \cap S_A \cap S_A) = b_{00}(a_{00}^2 + a_{00}a_{01} + a_{01}a_{10} + a_{01}a_{11}) + \\ + b_{01}(a_{10}^2 + a_{10}a_{11} + a_{11}a_{00} + a_{11}a_{01}),$$

Similarly, the probability p_3 that player A wins a match with 2 : 1 and he loses the second set is given by

$$p_3 = P(S_A \cap \bar{S}_A \cap S_A) = a_{00}(b_{00}a_{00} + b_{00}a_{01} + b_{01}a_{10} + b_{01}a_{11}) \\ + a_{01}(b_{11}a_{00} + b_{10}a_{10} + b_{11}a_{01} + b_{10}a_{11}).$$

Finally, the probability that player A wins a match in case (a) is the following:

$$P(M_A(2/3)) = p_1 + p_2 + p_3.$$

We can simplify determination probabilities p_1, p_2, p_3 using the probability matrices of winning a set of the form

$$P = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} \text{ and } \tilde{P} = \begin{bmatrix} a_{00} & a_{01} \\ a_{11} & a_{10} \end{bmatrix}.$$

Note that transformation $(i, j) \to (i, i+j)$ mod 2:

$$(0,0) \to (0,0+0) \bmod 2 = (0,0), \quad (0,1) \to (0,0+1) \bmod 2 = (0,1), \\ (1,0) \to (1,1+0) \bmod 2 = (1,1), \quad (1,1) \to (1,1+1) \bmod 2 = (1,0).$$

leads to $\tilde{P}_{i,j} = P_{i,i+j}$ for $i, j = 0, 1$. Let Q and $\tilde{Q}$ denote the corresponding matrices for the probabilities b_{ij} for player B, that is

$$Q = \begin{bmatrix} b_{00} & b_{01} \\ b_{10} & b_{11} \end{bmatrix} \text{ and } \tilde{Q} = \begin{bmatrix} b_{00} & b_{01} \\ b_{11} & b_{10} \end{bmatrix}.$$

Thus for a match consisting of the best 2 out of 3 sets the probability that A wins given that A serves first is

$$P(M_A(2/3)) = [PP + Q\tilde{P}P + P\tilde{Q}P]_* \quad (1)$$

where the subscript * denotes summation over the first row.

Note that, the order notation matrices $P, \tilde{P}, Q, \tilde{Q}$ ensue directly from the binary form of registration events of played sets during the match. We introduce the following principle of notation:

– a number of elements of a binary sequence is equal a number of matrices,

– the first and the last matrix is without of ~ while the remaining matrices are with ~,

– introduce the transcription: $0 \rightarrow Q$, $1 \rightarrow P$.

For instance, considering the binary sequence 01011, we have $Q\tilde{P}\tilde{Q}\tilde{P}P$. Using the transcription the probability that player A wins the match given that A serves first in case (b) is the following:

$$P(M_A(3/5)) = [P\tilde{P}P + (Q\tilde{P}\tilde{P}P + P\tilde{Q}\tilde{P}P + P\tilde{P}\tilde{Q}P) +$$
$$+ (Q\tilde{Q}\tilde{P}\tilde{P}P + Q\tilde{P}\tilde{Q}\tilde{P}P + Q\tilde{P}\tilde{P}\tilde{Q}P + P\tilde{Q}\tilde{Q}\tilde{P}P + P\tilde{Q}\tilde{P}\tilde{Q}P + P\tilde{P}\tilde{Q}\tilde{Q}P)]_* \qquad (2)$$

We can obtain similar evaluation for player B who wins a match. The formulas (1) and (2) were found by R i d d l e (1988). In order to use the formulas in practice we have to know a_{ij} and b_{ij}. Therefore it is necessary the formulas of winning a set by player A and B when:

(i) a set is finished by an even number of games, i.e. $6 : k$, $k = 0, 2, 4$;

(ii) a set is finished by an odd number of games, i.e. $6 : k$, $k = 1, 3$;

(iii) a set is finished by like $(k+2) : k$, $k = 5, 6, 7, ...$;

(iv) a set is finished by the tie-breaker. The tie-breaker game is played if the games' score reaches 6 games each. The first player to score at least 7 points and be at least 2 points ahead of his opponent is the winner of the tie-breaker game and set. If the points score reaches 6 points each in the tie-breaker one of players wins the tie-breaker by $(i + 2)$ points to i ($i = 6, 7, ...$). For these cases we give the formulas by authors mentioned in point 2.

Now, we give satisfactory formulas which are required in a numerical example when player A wins but A or B serves first in cases (i) and (ii):

(a) A serves first and A wins 6 : 0, 6 : 2 or 6 : 4

$$P(S_{(p)}) = (1 - g_B) \sum_{i=0}^{2} \sum_{j=1}^{2i+1} Bin(i+3, j-1.1-g_A) Bin(i+2.2i-j+1, g_B) \qquad (3)$$

(b) B serves first and A wins 6 : 0, 6 : 2 or 6 : 4

$$P(S_{(p)}) = g_A \sum_{i=0}^{2} \sum_{j=1}^{2i+1} Bin(i+2, j-1.1-g_A) Bin(i+3.2i-j+1, g_B) \qquad (4)$$

(c) A serves first and A wins 6 : 1 or 6 : 3

$$P(S_{(n)}) = g_A \sum_{i=0}^{1} \sum_{j=1}^{2i+2} Bin(i+3, j-1.1-g_A) Bin(i+3.2(i+1)-j, g_B) \qquad (5)$$

(d) B serves first and A wins 6 : 1 or 6 : 3

$$P(S_{(n)}) = (1-g_B) \sum_{i=0}^{1} \sum_{j=1}^{2i+2} Bin(i+3, j-1.1-g_A) Bin(i+3.2(i+1)-j, g_B) \qquad (6)$$

where

$$g_i = \begin{cases} \dfrac{p_i^4[1-16(1-p_i)^4]}{p_i^4-(1-p_i)^4}, & \text{if } p_i \neq \dfrac{1}{2}, \\ \dfrac{1}{2}, & \text{if } p_i = \dfrac{1}{2}, \end{cases}$$

for $i = A, B$ denotes the probability of winning a game by player A or B (see e.g. H s i and B u r y c h (1971)) and p_i ($i = A, B$) is the probability of winning a point by player A (or B) when he is serving. The formulas (4) and (5) are authors' propositions while the formulas (3) and (5) we can find in R i d d l e (1988) or in P a s e w i c z and W a g n e r (2000).

IV. A NUMERICAL EXAMPLE

An example is an illustration of calculation in case of (a), when player A serves first and wins the match with 2 : 0. There we assume that $p_A = 0.9$ and $p_B = 0.6$ and the set score reaches 6 : k for k = 0, 1, 2, 3, 4. A computer program written in Department of Statistics Academy of Physical Education in Poznań can be obtained from the authors upon requst. Using formulae (3) – (6) we have different possibilities of the probability that A wins a set:

(a) „6 : 0" – 0.01838, „6 : 2" – 0.11299, „6 : 4" – 0.07768,
(b) „6 : 0" – 0.01838, „6 : 2" – 0.22575, „6 : 4" – 0.38597,
(c) „6 : 1" – 0.15334, „6 : 3" – 0.41925,
(d) „6 : 1" – 0.04058, „6 : 3" – 0.11096.

Hence $a_{00} = 0.20905$, $a_{10} = 0.63010$, $a_{01} = 0.57259$ and $a_{11} = 0.15154$ and the probability that A wins the match is equal $p_1 = 0.61096$.

Now we calculate p_1 using matrix P. In the case matrices P and PP have the forms:

$$P = \begin{bmatrix} 0.20905 & 0.57259 \\ 0.63010 & 0.15154 \end{bmatrix}, \quad PP = \begin{bmatrix} 0.40449 & 0.20647 \\ 0.30453 & 0.38375 \end{bmatrix},$$

The sum of the elements over the first row of matrix PP is equal p_1.

REFERENCE

Carter W. H., Crews S. L. (1974), *An Analysis of the Game of Tennis*, „American Statistician", **28**, 4, 130–134.

Hsi B. P., Burych D. M. (1971), *Games of Two Players*, „J. of Royal Stat. Society Series C", **20**, 86–92.

Pasewicz W., Wagner W. (2000), *Charakterystyka modeli probabilistycznych w opisie gema i seta w tenisie ziemnym*, „Wyzwania i Dylematy Statystyki XXI wieku", AE Wrocław, 140–147.

Pollard G. H. (1983), *An Analysis of Classical and Tie-Breaker Tennis*, „Australian Journal of Statistics", **25**, 496–505.

Riddle L. H. (1988), *Probability Models for Tennis Scoring Systems*, „Appl. Statist.", **37**, 1, 63–75.

Wagner W., Majewska K. (1996), *Model probabilistyczny dwumianowy systemu punktacji seta typu 6 : k w tenisie ziemnym*, „Colloquium Biometryczne", AR Lublin, 365–375.

Wiesław Pasewicz, Wiesław Wagner

PROBABILISTYCZNY MODEL ZAKOŃCZENIA MECZU W TENISIE ZIEMNYM

W tenisie ziemnym mecz jest rozgrywany przez dwóch graczy i składa się z setów podzielonych na gemy. Przyjmując stałe prawdopodobieństwa wygrania własnego serwisu przez każdego z graczy w trakcie trwania meczu, można podać odpowiednie wzory na prawdopodobieństwa zakończenia gema oraz seta. Naturalny wydaje się być problem obliczania prawdopodobieństw zakończenia meczu. Wyprowadzone są wzory na wygranie meczu przez jednego z graczy. W celu uproszczenia wyprowadzenia wzorów stosowane są specjalne macierze probabilistyczne. Przedstawiony jest również prosty przykład numeryczny obliczania prawdopodobieństwa wygrania meczu.