Serving in tennis: deuce court and ad court differ*

Spieksma F.

KBI_1622

Serving in tennis: deuce court and ad court differ*

Frits C.R. Spieksma†

Abstract

We present statistical evidence that in professional tennis the probability of winning a point may depend on whether the serving player serves from the deuce court or from the ad court. Moreover, in this case of distinct win-probabilities, we show how to calculate the probability of winning a game, as well as winning a tiebreak.

Keywords: tennis, serving, iid assumption.

1 Introduction

"Try hitting a return from the stands with McEnroe standing at the net". This quote, from the book Winning Ugly by Brad Gilbert and Steve Jamison [2], describes the situation in tennis where the server (John McEnroe, a left hander) serves a diagonal service from the ad court to the receiver (Brad Gilbert, a right hander), and discusses the challenges arising from this situation from the receiver's point of view. Points in tennis can be partitioned into two sets, those served from the deuce court (where the server is required to stand right from the middle of the baseline), and those from the ad court (where the server is required to stand left from the middle of the baseline). The theme of this note, as illustrated by the first sentence, is that these two types of points differ. More specifically, using basic statistical tests we show in Section 2 that for some professional players (including many left-handed players), there is a statistically significant difference between the probability of winning the point when serving from the deuce court, as compared to serving from the ad court. Further, we show in Section 3, by extending known results, how to express the probability of winning a game, as well as the probability of winning a tiebreak, in terms of two probabilities: p_d and p_a, where p_d (p_a) stands for the probability for the server to win the point when serving from the deuce (ad) court.

2 The test

A very common assumption in statistical research on tennis is that the outcome of points (a win for the server, or a loss for the server) are independently and identically distributed

*This research is supported by the Interuniversity Attraction Poles Programme initiated by the Belgian Science Policy Office.

†KU Leuven, Faculty of Business and Economics. Email: frits.spieksma@kuleuven.be

(the *iid* assumption). This assumption is investigated in Klaassen and Magnus [3] who show that the iid assumption is violated, yet deviations from it are small. Knight and O'Donoghue [4] give evidence that break points are different from non-break points, in the sense that the receiver's probability of winning a break point is significantly larger than the probability of winning a non-break point, confirming the non-stationarity of points in tennis. Here, we give additional evidence that the iid assumption should be used with care. In fact, we argue that, for some professional players, the probability of winning a point depends upon the position of the server (deuce court or ad court).

There are two possible outcomes for a point played in tennis: it is either won by the server, or lost by the server. Assuming all points are created equal, we can model this outcome as the outcome of an experiment where we draw from a binomial distribution characterized by some probability p. Now consider, as a motivating example, the case of Angelique Kerber, a professional left-handed tennis player. Let us assume that she has played 3592 points in total, 1878 from the deuce court, and 1714 from the ad court. Out of these 1878 points, she won 1042 (55,48%), and out of these 1714 points, she won 1014 (59,16%). These numbers might raise the question whether indeed all points are created equal, or whether one should discriminate between points served from the deuce court and points served from the ad court. To do so, let us denote by p_d (p_a) the probability that Kerber wins a point when she serves from the deuce (ad) court, and let us use, as a traditional point estimate for these probabilities, the ratio of the number of points won by Kerber when serving from the deuce (ad) court and the number of points played by Kerber serving from the deuce (ad) court, denoted by $\hat{p_d}$ ($\hat{p_a}$). We proceed by setting up the null-hypothesis stating that p_d and p_a are identical, to be tested against the alternative hypothesis that p_d and p_a differ. Thus, we have the null hypothesis: $H_0 : p_d = p_a$ versus the alternative hypothesis $H_a : p_d \neq p_a$.

The appropriate test statistic (see Mood et al. [5]) equals

$$\frac{\hat{p_d} - \hat{p_a}}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_d} + \frac{1}{n_a})}}.$$

In this expression, n_d (n_a) stands for the number of points served from the deuce (ad) court, and $\hat{p}$ stands for $\frac{n_d\hat{p_d}+n_a\hat{p_a}}{n_d+n_a}$. We compute the test statistic, and find out whether this value lies within $[-1.96, 1.96]$. If so, we can accept the null-hypothesis, otherwise we can reject the null-hypothesis (using a 95% confidence).

Plugging in the numbers for Angelique Kerber gives a a value of the test statistic of $\approx -2,223858893$, meaning that we should reject the null hypothesis. Or, in other words, Kerber has a significant higher probability of winning the point when she serves from the ad court, than when she serves from the deuce court.

In Tables 1 and 2 listed below we give for the top 10 female and male players the corresponding data that were retrieved from www.tennisabstract.com on June 9, 2016. The first column contains the name of the player (with "(left)" indicating that the player is left-handed), the second column gives the number of matches for which the points were scored, the third (fifth) column gives n_d (n_a), the fourth (and sixth) column gives the number of points won, the seventh column gives the value of the test statistic, and the final column shows whether or not the null-hypothesis can be accepted for the corresponding player.

Player	Number of matches	n_d	points won	n_a	points won	test statistic	Is H_0 accepted (95 %)?
Serena Williams	96	3055	1994	2791	1768	1.534	yes
Garbine Muguruza	46	1714	1017	1587	933	0.318	yes
Agnieszka Radwanska	64	2307	1362	2132	1227	1.003	yes
Angelique Kerber (left)	51	1878	1042	1714	1014	-2.224	no
Simona Halep	153	5209	3106	4782	2751	2.128	no
Victoria Azarenka	54	1824	1066	1699	979	0.493	yes
Roberta Vinci	18	773	427	679	390	0.310	yes
Belinda Bencic	25	898	514	827	482	-0.439	yes
Venus Williams	36	1442	825	1334	773	-0.391	yes
Timea Bacsinszky	22	720	410	663	344	1.888	yes

Table 1: Results for female top 10 players

Player	Number of matches	n_d	points won	n_a	points won	test statistic	Is H_0 accepted (95 %)?
Novak Djokovic	144	6292	4169	5711	3691	1.875	yes
Andy Murray	77	3240	2078	2952	1835	1.609	yes
Roger Federer	170	7587	5257	6880	4670	1.827	yes
Rafael Nadal (left)	151	6459	4141	5892	3889	-2.203	no
Stanislas Wawrinka	48	2064	1361	1888	1177	2.358	no
Kei Nishikori	33	1300	819	1194	724	1.214	yes
Dominic Thiem	21	882	521	808	502	-1.285	yes
Tomas Berdych	40	1367	894	1236	812	-0.159	yes
Milos Raonic	35	1512	1052	1345	923	0.550	yes
Richard Gasquet	28	1035	613	952	571	-0.341	yes

Table 2: Results for male top 10 players

Out of these 20 players, there are four players for which the null-hypothesis is rejected: Angelique Kerber, Simona Halep, Rafael Nadal, and Stanislas Wawrinka. Intuition (see the quote starting this note) would predict that left-handed players may have a better percentage when serving from the ad court; the data support this intuition: in fact, for *all* left-handed players in the top 10, the null-hypothesis can be refuted. And apparently, Simona Halep, and Stanislas Wawrinka have a significant higher win-probability when serving from the deuce court, as compared to serving from the ad court. This allows us to conclude that, at least for some players, points serving from the deuce court are different from points serving from the ad court.

One might wonder whether each left-handed professional player has a significant higher win-percentage from the ad court than from the deuce court. We give the numbers in Table 3 below, where we list all left-handed male and female top 100 players with at least 10 matches scored.

And although it is true that the majority of left-handed players has a higher win probability when serving from the ad court than when serving from the deuce court, this

Player	Number of matches	n_d	points won	n_a	points won	test statistic	Is H_0 accepted (95 %)?
Angelique Kerber	51	1878	1042	1714	1014	-2.224	no
Petra Kvitova	47	1673	984	1545	949	-1.509	yes
Lucie Safarova	54	950	586	871	528	0.465	yes
Ekaterina Makarova	21	776	446	731	416	0.222	yes
Misaki Doi	10	422	236	397	223	-0.071	yes
Varvara Lepchenko	12	427	214	391	210	-1.027	yes
Rafael Nadal	151	6459	4141	5892	3889	-2.203	no
Feliciano Lopez	20	809	574	815	541	-0.818	yes
Albert Ramos-Vinolas	10	384	237	344	203	0.746	yes
Federico Delbonis	12	475	291	428	253	0.659	yes
Gilles Muller	11	544	363	497	348	-1.140	yes
Martin Klizan	10	374	233	340	213	-0.096	yes
Guido Pella	10	379	235	349	221	-0.367	yes
Fernando Verdasco	12	564	344	526	334	-0.852	yes
Adrian Mannarino	12	455	267	415	259	-1.123	yes
Thomaz Bellucci	15	619	388	570	353	0.267	yes
Jiri Vesely	11	428	262	375	263	-2.650	no

Table 3: Results for left-handed players with at least 10 matches

is only significant for 3 out of the 17 left-handed players.

Of course, one might discuss the set of matches that should be used for an analysis as carried out above. Here, we chose as a "universe" the set of matches played by a particular single player, but other possibilities exist. Indeed, one can well imagine to select the set of matches played between a specific pair of players. Or one could select all matches played at a particular venue, such as Wimbledon. Or one could select all matches played by left handers. Elaborating on this last suggestion, we summed all entries in Table 3, creating an "aggregate left handed player", whose results can be found in Table 4.

Player	Number of matches	n_d	points won	n_a	points won	test statistic	Is H_0 accepted (95 %)?
Aggr. left-h. player	469	17616	10843	16200	10217	-2.873	no

Table 4: Aggregated result for left-handed players with at least 10 matches

From Table 4, the conclusion seems justified that a professional left-handed tennis-player has a higher probability of winning the point when serving from the ad court compared to serving from the deuce court.

As a remark, notice that, for any player the number of points served from the deuce court (n_d) exceeds the number of points played served from the ad court (n_a). This is not a coincidence: each game features a number of points served form the deuce court (say, nod) and a number of points served from the ad court (say, noa). These numbers coincide in each game, except in those games which end when the server wins the point at a score of 40-15; only then $nod = noa + 1$.

3 The probability of winning a match

Newton and Keller [7] show how to derive the probability that player A defeats player B, assuming that p^A (the probability that player A wins a point when serving) and p^B (the probability that player B wins a point when serving) are given (and, of course, making the iid assumption). This derivation is based on first establishing the probability that the server wins a game (see (1)), and the probability that player A wins a tiebreak against player B. From this, the probability that player A wins a set, and subsequently, the probability that player A wins a match against player B is found. In order to extend this analysis to take into account distinct win-probabilities when serving from the deuce court versus serving from the ad court, we only need to show how to compute the probability of winning a game (Subsection 3.1), and a tiebreak (Subsection 3.2), in the situation of distinct win probabilities.

3.1 The probability of winning a game

In case there is a single probability that reflects the win probability for the server, say p, (and making the iid assumption), there is a well-known expression for the probability that the server wins the game:

$$p^4(1 + 4(1-p) + 10(1-p)^2) + \frac{20p^5(1-p)^3}{1-2p(1-p)}. \tag{1}$$

The analysis leading to (1) can be found in Carter and Crews [1], see also Newton and Keller [7]. This expression has been generalized by Knight and O'Donoghue [4] taking breakpoints into account. We now extend this analysis for the case where there is a win probability for the deuce court (p_d), and a win probability for the ad court (p_a). Thus, we answer the following question: given that the server wins the point when serving from the deuce (ad) court with probability p_d (p_a), what is the probability that the server wins a game?

Clearly, a game in tennis may arrive at deuce, or it may not. Let p^D be the probability for the server to win the game given that the score in the game arrives at deuce. We have:

$$p^D = p_d p_a + p_d(1-p_a)p^D + (1-p_d)p_a p^D.$$

From this we arrive at an explicit expression for p^D:

$$p^D = \frac{p_d p_a}{1 - p_d - p_a + 2p_d p_a}. \tag{2}$$

If a game does not arrive at deuce, there are three ways for the server to win the game: the opponent makes either 0, or 1, or 2 points. The corresponding probabilities are denoted by respectively $p0$, $p1$, and $p2$. We have:

$$p0 = p_d p_a p_d p_a = p_d^2 p_a^2. \tag{3}$$

If the receiver wins a single point, it is either the first, the second, the third, or the fourth. Hence:

$$p1 = (1-p_d)p_a p_d p_a p_d + p_d(1-p_a)p_d p_a p_d + p_d p_a(1-p_d)p_a p_d + p_d p_a p_d(1-p_a)p_d.$$

It follows that:

$$p1 = 2p_d^2 p_a[(1-p_d)p_a + p_d(1-p_a)] = 2p_d^2 p_a[p_d + p_a - 2p_d p_a]. \tag{4}$$

Similarly, we enumerate all possibilities for the receiver to win two points. This leads to:

$$\begin{aligned} p2 &= (1-p_d)(1-p_a)p_d p_a p_d p_a + (1-p_d)p_a(1-p_d)p_a p_d p_a + \\ &+ (1-p_d)p_a p_d(1-p_a)p_d p_a + (1-p_d)p_a p_d p_a(1-p_d)p_a + \\ &+ p_d(1-p_a)(1-p_d)p_a p_d p_a + p_d(1-p_a)p_d(1-p_a)p_d p_a + \\ &+ p_d(1-p_a)p_d p_a(1-p_d)p_a + p_d p_a(1-p_d)(1-p_a)p_d p_a + \\ &+ p_d p_a(1-p_d)p_a(1-p_d)p_a + p_d p_a p_d(1-p_a)(1-p_d)p_a. \end{aligned}$$

Thus:

$$\begin{aligned} p2 &= 6(1-p_d)(1-p_a)p_d^2 p_a^2 + 3(1-p_d)^2 p_d p_a^3 + (1-p_a)^2 p_d^3 p_a = \\ &= p_d p_a \left(p_d^2 + 3p_a^2 - 8p_d^2 p_a - 12p_d p_a^2 + 10p_d^2 p_a^2 + 6p_d p_a\right). \end{aligned} \tag{5}$$

What is the probability that a game arrives at the deuce score, denoted by $p33$? We derive:

$$\begin{aligned} p33 &= 9(1-p_d)^2(1-p_a)p_d p_a^2 + 9(1-p_d)(1-p_a)^2 p_d^2 p_a + (1-p_d)^3 p_a^3 + (1-p_a)^3 p_d^3 = \\ &= 9p_d p_a(1-p_d)(1-p_a)[p_d + p_a - 2p_d p_a] + (1-p_d)^3 p_a^3 + (1-p_a)^3 p_d^3. \end{aligned} \tag{6}$$

Observe that the probability of winning a game equals:

$$p0 + p1 + p2 + p33 \times p^D. \tag{7}$$

Thus, plugging (2), (3), (4), (5) and (6) into (7) gives us the probability that a serving player wins the game assuming distinct win-probabilities p_d and p_a for serving from the deuce court and ad court respectively:

$$\begin{aligned} &p_d^2 p_a^2 + 2p_d^2 p_a[p_d + p_a - 2p_d p_a] + p_d p_a\left(p_d^2 + 3p_a^2 - 8p_d^2 p_a - 12p_d p_a^2 + 10p_d^2 p_a^2 + 6p_d p_a\right) + \\ &\left(9p_d p_a(1-p_d)(1-p_a)[p_d + p_a - 2p_d p_a] + (1-p_d)^3 p_a^3 + (1-p_a)^3 p_d^3\right) \times \frac{p_d p_a}{1 - p_d - p_a + 2p_d p_a}. \end{aligned} \tag{8}$$

Of course, when setting $p_d = p_a$ in (8), the resulting expression boils down to (1). Admittedly, the numerical difference between (1) and (8) in real-life situations is small. For instance, using the data in Table 1 to provide point estimates for p, p_d and p_a in the case of Angelique Kerber, leads to a value of 0,675416869 in case of (1), and 0,677483019 in case of (8).

3.2 The probability of winning a tiebreak

Let us assume that player A starts the tiebreak by serving first. The rules of the tiebreak say that player A then starts serving from the deuce court; next, player B serves twice, first once from the ad court, and then from the deuce court, followed again by player A serving twice in a similar fashion (first from the ad court, then from the deuce court), and so on. This continues until a player has made 7 points, while the opponent has made five points or less. If that does not happen, the score in the tiebreak is, at some point, 6-6, and the first player who has made two more points then the other player wins the tiebreak.

It follows that there are seven mutually exclusive ways in which a player can win a tiebreak: with a score of 7-0, 7-1, 7-2, 7-3, 7-4, 7-5, and with a tiebreak in which the score after 12 points equals 6-6. We denote the probability of these first six options by $p70, p71, p72, p73, p74, p75$. The probability that the tiebreak reaches 6-6 is denoted by $p66$.

In the latter case, we claim that the probability that player A wins when the score is 6-6 is equal to the probability that player A wins when the score equals 8-8, 10-10, ... (this follows from the fact that after four points in the tiebreak, we return to an identical situation); we denote this probability by $p_{w@66}$.

The probability that player A wins the tiebreak equals:

$$p70 + p71 + p72 + p73 + p74 + p75 + (p66 \times p_{w@66}). \tag{9}$$

We now show how to express the terms $p_{w@66}$ and $p73$ in the probabilities $p_d^A, p_a^A, p_d^B, p_a^B$, where p_d^A (p_d^B) denotes the probability that player A (B) wins a point when serving from the deuce court, and p_a^A (p_a^B) denotes the probability that player A (B) wins a point when serving from the ad court. The expression of all other terms in (9) in $p_d^A, p_a^A, p_d^B, p_a^B$ can be found in the Appendix.

Expressing $p_{w@66}$

Let us first derive the probability that player A wins the tiebreak when the score has reached 6-6. The following holds:

$$\begin{aligned} p_{w@66} &= p_d^A(1-p_a^B) + p_d^A p_a^B(1-p_d^B)p_a^A + (1-p_d^A)(1-p_a^B)(1-p_d^B)p_a^A + \\ &+ p_{w@66}\left(p_d^A p_a^B(1-p_d^B)(1-p_a^A)\right) + p_{w@66}\left(p_d^A p_a^B p_d^B p_a^A\right) + \\ &+ p_{w@66}\left((1-p_d^A)(1-p_a^B)(1-p_d^B)(1-p_a^A)\right) + p_{w@66}\left((1-p_d^A)(1-p_a^B)p_d^B p_a^A\right). \end{aligned}$$

The first three terms reflect the three distinct ways in which the tiebreak can be won when playing at most four points from the score $x-x$ ($x \geq 6, x$ even); the last four terms reflect the four distinct ways in which the score reaches $(x+2)-(x+2)$ starting from the score $x-x$ ($x \geq 6, x$ even). This leads to:

$$p_{w@66} = \frac{p_d^A(1-p_a^B) + p_d^A p_a^B(1-p_d^B)p_a^A + (1-p_d^A)(1-p_a^B)(1-p_d^B)p_a^A}{p_d^A p_a^B(1-p_d^B)(1-p_a^A) + p_d^A p_a^B p_d^B p_a^A + (1-p_d^A)(1-p_a^B)(1-p_d^B)(1-p_a^A) + (1-p_d^A)(1-p_a^B)p_d^B p_a^A}.$$

Expressing $p73$

Let us now show, as an illustration, how $p73$ can be expressed in terms of p_d^A, p_a^A, p_d^B and p_a^B; for all other probabilities we refer to the Appendix.

There are four types of points: those where player A serves from the deuce court (type 1), those where player A serves from the ad court (type 2), those where player B serves from the deuce court (type 3), and those where player B serves from the ad court (type 4).

Clearly, in case player A wins the tiebreak with 7-3, the last (i.e., tenth) point of the tiebreak was won by player A. Since the tenth point features player B serving from the ad court, the probability that player A wins this point is given by $(1 - p_a^B)$. Further, it is a fact that from the first 9 points in the tiebreak, three go to player B. These 9 points consist of 3 points of type 1, 2 points of type 2, 2 points of type 3, and 2 points of type 4. Elementary analysis shows that there are 17 distinct ways in which the three points won by player B can be distributed over the four types of points: (3000), (2100), (2010), (2001), (1200), (1110), (1101), (1020), (1011), (1002), (0210), (0201), (0120), (0111), (0102), (0021), (0012). (Notice that (xyuv) indicates that player B won x points of type 1, y points of type 2, u points of type 3, and v points of type 4). For each of these ways, we can simply compute the associated probability, sum these probabilities, and since these are mutually exclusive, arrive at $p73$. In particular:

$$
\begin{aligned}
\tfrac{p73}{(1-p_a^B)} &= (1-p_d^A)^3(p_a^A)^2(1-p_d^B)^2(1-p_a^B)^2+ && (3000)\\
&+ (1-p_d^A)^2 p_d^A p_a^A(1-p_a^A)(1-p_d^B)^2(1-p_a^B)^2+ && (2100)\\
&+ (1-p_d^A)^2 p_d^A(p_a^A)^2(1-p_d^B)p_d^B(1-p_a^B)^2+ && (2010)\\
&+ (1-p_d^A)^2 p_d^A(p_a^A)^2(1-p_d^B)^2(1-p_a^B)p_a^B+ && (2001)\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)^2(1-p_d^B)^2(1-p_a^B)^2+ && (1200)\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)p_a^A(1-p_d^B)p_d^B(1-p_a^B)^2+ && (1110)\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)p_a^A(1-p_d^B)^2(1-p_a^B)p_a^B+ && (1101)\\
&+ (1-p_d^A)(p_d^A)^2(p_a^A)^2(p_d^B)^2(1-p_a^B)^2+ && (1020)\\
&+ (1-p_d^A)(p_d^A)^2(p_a^A)^2(1-p_d^B)p_d^B(1-p_a^B)p_a^B+ && (1011)\\
&+ (1-p_d^A)(p_d^A)^2(p_a^A)^2(1-p_d^B)^2(p_a^B)^2+ && (1002)\\
&+ (p_d^A)^3(1-p_a^A)^2(1-p_d^B)(p_d^B)(1-p_a^B)^2+ && (0210)\\
&+ (p_d^A)^3(1-p_a^A)^2(1-p_d^B)^2(1-p_a^B)p_a^B+ && (0201)\\
&+ (p_d^A)^3(1-p_a^A)p_a^A(p_d^B)^2(1-p_a^B)^2+ && (0120)\\
&+ (p_d^A)^3(1-p_a^A)p_a^A(1-p_d^B)(p_d^B)(1-p_a^B)p_a^B+ && (0111)\\
&+ (p_d^A)^3(1-p_a^A)p_a^A(1-p_d^B)^2(p_a^B)^2+ && (0102)\\
&+ (p_d^A)^3(p_a^A)^2(p_d^B)^2(1-p_a^B)p_a^B+ && (0021)\\
&+ (p_d^A)^3(p_a^A)^2(1-p_d^B)(p_d^B)(p_a^B)^2. && (0012)
\end{aligned}
$$

This procedure can be applied to all other probabilities $p70, p71, p72, p74, p75$ and $p66$ - we refer to the appendix for the precise expressions. Thus, plugging in all these probabilities into (9) gives us an expression for the probability that player A wins the tiebreak.

4 Conclusion

We presented statistical evidence that - for some professional tennisplayers, in particular left-handed players - win-probabilities depend on their position of serving. Moreover, we showed how to calculate the probability of winning a game assuming distinct win-probabilities when serving from the deuce court and serving from the ad court.

Acknowledgements: I would like to thank Martina Vandebroek for her suggestions, and Thomas Spieksma for collecting the data.

5 Appendix

We now give the expressions for all remaining probabilities.

Expressing $p70$

$$p70 = p_d^A(1-p_a^B)(1-p_d^B)p_a^A p_d^A(1-p_a^B)(1-p_d^B).$$

Expressing $p71$

$$\begin{aligned}
\frac{p71}{p_a^A} &= (1-p_d^A)(1-p_a^B)(1-p_d^B)p_a^A p_d^A(1-p_a^B)(1-p_d^B)+ \\
&+ p_d^A p_a^B(1-p_d^B)p_a^A p_d^A(1-p_a^B)(1-p_d^B)+ \\
&+ p_d^A(1-p_a^B)p_d^B p_a^A p_d^A(1-p_a^B)(1-p_d^B)+ \\
&+ p_d^A(1-p_a^B)(1-p_d^B)(1-p_a^A)p_d^A(1-p_a^B)(1-p_d^B)+ \\
&+ p_d^A(1-p_a^B)(1-p_d^B)p_a^A(1-p_d^A)(1-p_a^B)(1-p_d^B)+ \\
&+ p_d^A(1-p_a^B)(1-p_d^B)p_a^A p_d^A p_a^B(1-p_d^B)+ \\
&+ p_d^A(1-p_a^B)(1-p_d^B)p_a^A p_d^A(1-p_a^B)p_d^B.
\end{aligned}$$

Expressing $p72$
To compute $p72$, we note that there are 10 possibilities to distribute 2 points over 4 types

of points each with multiplicity 2.

$$
\begin{aligned}
\frac{p72}{p_d^A} &= (1-p_d^A)^2(p_a^A)^2(1-p_d^B)^2(1-p_a^B)^2+ && (2000)\\
&+ (1-p_d^A)p_d^A(1-p_a^A)p_a^A(1-p_d^B)^2(1-p_a^B)^2+ && (1100)\\
&+ (1-p_d^A)p_d^A(p_a^A)^2(1-p_d^B)(p_d^B)(1-p_a^B)^2+ && (1010)\\
&+ (1-p_d^A)p_d^A(p_a^A)^2(1-p_d^B)^2(1-p_a^B)p_a^B+ && (1001)\\
&+ (p_d^A)^2(1-p_a^A)^2(1-p_d^B)^2(1-p_a^B)^2+ && (0200)\\
&+ (p_d^A)^2(1-p_a^A)p_a^A(1-p_d^B)p_d^B(1-p_a^B)^2+ && (0110)\\
&+ (p_d^A)^2(1-p_a^A)p_a^A(1-p_d^B)^2(1-p_a^B)p_a^B+ && (0101)\\
&+ (p_d^A)^2(p_a^A)^2(p_d^B)^2(1-p_a^B)^2+ && (0020)\\
&+ (p_d^A)^2(p_a^A)^2p_d^B(1-p_d^B)(1-p_a^B)p_a^B+ && (0011)\\
&+ (p_d^A)^2(p_a^A)^2(1-p_d^B)^2(p_a^B)^2. && (0002)
\end{aligned}
$$

Expressing $p74$

To compute $p74$, we note that there are 25 possibilities to distribute 4 points over 4 types of points with multiplicities 3, 2, 2, and 3 respectively.

$$
\begin{aligned}
\frac{p74}{1-p_d^B} &= (1-p_d^A)^3(p_a^A)(1-p_a^A)(1-p_d^B)^2(1-p_a^B)^3+ && (3100)\\
&+ (1-p_d^A)^3(p_a^A)^2(1-p_d^B)p_d^B(1-p_a^B)^3+ && (3010)\\
&+ (1-p_d^A)^3(p_a^A)^2(1-p_d^B)^2(1-p_a^B)^2p_a^B+ && (3001)\\
&+ (1-p_d^A)^2p_d^A(1-p_a^A)^2(1-p_d^B)^2(1-p_a^B)^3+ && (2200)\\
&+ (1-p_d^A)^2p_d^A(1-p_a^A)p_a^A(1-p_d^B)p_d^B(1-p_a^B)^3+ && (2110)\\
&+ (1-p_d^A)^2p_d^A(1-p_a^A)p_a^A(1-p_d^B)^2(1-p_a^B)^2p_a^B+ && (2101)\\
&+ (1-p_d^A)^2p_d^A(p_a^A)^2(p_d^B)^2(1-p_a^B)^3+ && (2020)\\
&+ (1-p_d^A)^2p_d^A(p_a^A)^2(1-p_d^B)p_d^B(1-p_a^B)^2p_a^B+ && (2011)\\
&+ (1-p_d^A)^2p_d^A(p_a^A)^2(1-p_d^B)^2(1-p_a^B)^2(p_a^B)+ && (2002)\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)^2(1-p_d^B)p_d^B(1-p_a^B)^3+ && (1210)\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)^2(1-p_d^B)^2p_a^B(1-p_a^B)^2+ && (1201)\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)p_a^A(p_d^B)^2(1-p_a^B)^3+ && (1120)\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)p_a^A(p_d^B)(1-p_d^B)p_a^B(1-p_a^B)^2+ && (1111)\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)p_a^A(1-p_d^B)^2(p_a^B)^2(1-p_a^B)+ && (1102)\\
&+ (1-p_d^A)(p_d^A)^2(p_a^A)^2(p_d^B)^2(p_a^B)(1-p_a^B)^2+ && (1021)\\
&+ (1-p_d^A)(p_d^A)^2(p_a^A)^2(p_d^B)(1-p_d^B)(p_a^B)^2(1-p_a^B)+ && (1012)\\
&+ (1-p_d^A)(p_d^A)^2(p_a^A)^2(1-p_d^B)^2(p_a^B)^3+ && (1003)\\
&+ (p_d^A)^3(1-p_a^A)^2(p_d^B)^2(1-p_a^B)^3+ && (0220)\\
&+ (p_d^A)^3(1-p_a^A)^2(p_d^B)(1-p_d^B)(1-p_a^B)^2p_a^B+ && (0211)\\
&+ (p_d^A)^3(1-p_a^A)^2(1-p_d^B)^2(1-p_a^B)(p_a^B)^2+ && (0202)\\
&+ (p_d^A)^3(1-p_a^A)p_a^A(p_d^B)^2(1-p_a^B)^2(p_a^B)+ && (0121)\\
&+ (p_d^A)^3(1-p_a^A)p_a^A(p_d^B)(1-p_d^B)(1-p_a^B)(p_a^B)^2+ && (0112)\\
&+ (p_d^A)^3(1-p_a^A)p_a^A(1-p_d^B)^2(p_a^B)^3+ && (0103)\\
&+ (p_d^A)^3(p_a^A)^2(p_d^B)^2(1-p_a^B)(p_a^B)^2+ && (0022)\\
&+ (p_d^A)^3(p_a^A)^2(p_d^B)(1-p_d^B)(p_a^B)^3. && (0013)
\end{aligned}
$$

Expressing $p75$

To compute $p75$, we note that there are 34 possibilities to distribute 5 points over 4 types of points with multiplicities 3, 2, 3, and 3 respectively.

$$
\begin{aligned}
\frac{p75}{p_a^A} &= (1-p_d^A)^3(1-p_a^A)^2(1-p_d^B)^3(1-p_a^B)^3+ && (3200)\\
&+ (1-p_d^A)^3(p_a^A)(1-p_a^A)(1-p_d^B)^2p_d^B(1-p_a^B)^3+ && (3110)\\
&+ (1-p_d^A)^3(p_a^A)(1-p_a^A)(1-p_d^B)^3(1-p_a^B)^2p_a^B+ && (3101)\\
&+ (1-p_d^A)^3(p_a^A)^2(p_d^B)^2(1-p_d^B)(1-p_a^B)^3+ && (3020)\\
&+ (1-p_d^A)^3(p_a^A)^2(1-p_d^B)^2p_d^B(1-p_a^B)^2p_a^B+ && (3011)\\
&+ (1-p_d^A)^3(p_a^A)^2(1-p_d^B)^3(1-p_a^B)(p_a^B)^2+ && (3002)\\
&+ (1-p_d^A)^2p_d^A(1-p_a^A)^2(1-p_d^B)^2p_d^B(1-p_a^B)^3+ && (2210)\\
&+ (1-p_d^A)^2p_d^A(1-p_a^A)^2(1-p_d^B)^3(1-p_a^B)^2p_a^B+ && (2201)\\
&+ (1-p_d^A)^2(p_d^A)(1-p_a^A)p_a^A(p_d^B)^2(1-p_d^B)(1-p_a^B)^3+ && (2120)\\
&+ (1-p_d^A)^2(p_d^A)(1-p_a^A)p_a^Ap_d^B(1-p_d^B)^2p_a^B(1-p_a^B)^2+ && (2111)\\
&+ (1-p_d^A)^2(p_d^A)(1-p_a^A)p_a^A(1-p_d^B)^3(p_a^B)^2(1-p_a^B)+ && (2102)\\
&+ (1-p_d^A)^2p_d^A(p_a^A)^2(p_d^B)^3(1-p_a^B)^3+ && (2030)\\
&+ (1-p_d^A)^2p_d^A(p_a^A)^2(p_d^B)^2(1-p_d^B)p_a^B(1-p_a^B)^2+ && (2021)\\
&+ (1-p_d^A)^2p_d^A(p_a^A)^2p_d^B(1-p_d^B)^2(p_a^B)^2(1-p_a^B)+ && (2012)\\
&+ (1-p_d^A)^2p_d^A(p_a^A)^2(1-p_d^B)^3(p_a^B)^3+ && (2003)\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)^2(1-p_d^B)(p_d^B)^2(1-p_a^B)^3+ && (1220)\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)^2(p_d^B)(1-p_d^B)^2p_a^B(1-p_a^B)^2+ && (1211)\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)^2(1-p_d^B)^3(p_a^B)^2(1-p_a^B)+ && (1202)\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)p_a^A(p_d^B)^3(1-p_a^B)^3+ && (1130)\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)p_a^A(p_d^B)^2(1-p_d^B)p_a^B(1-p_a^B)^2+ && (1121)\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)p_a^A(1-p_d^B)^2p_d^B(p_a^B)^2(1-p_a^B)+ && (1112)\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)p_a^A(1-p_d^B)^3(p_a^B)^3+ && (1103)\\
&+ (1-p_d^A)(p_d^A)^2(p_a^A)^2(p_d^B)^3p_a^B(1-p_a^B)^2+ && (1031)\\
&+ (1-p_d^A)(p_d^A)^2(p_a^A)^2(p_d^B)^2(1-p_d^B)(p_a^B)^2(1-p_a^B)+ && (1022)\\
&+ (1-p_d^A)(p_d^A)^2(p_a^A)^2(1-p_d^B)^2p_d^B(p_a^B)^3+ && (1013)\\
&+ (p_d^A)^3(1-p_a^A)^2(p_d^B)^3(1-p_a^B)^3+ && (0230)\\
&+ (p_d^A)^3(1-p_a^A)^2(p_d^B)^2(1-p_d^B)(1-p_a^B)^2p_a^B+ && (0221)\\
&+ (p_d^A)^3(1-p_a^A)^2(1-p_d^B)^2p_d^B(1-p_a^B)(p_a^B)^2+ && (0212)\\
&+ (p_d^A)^3(1-p_a^A)^2(1-p_d^B)^3(p_a^B)^3+ && (0203)\\
&+ (p_d^A)^3(1-p_a^A)p_a^A(p_d^B)^3(1-p_a^B)^2p_a^B+ && (0131)\\
&+ (p_d^A)^3(1-p_a^A)p_a^A(p_d^B)^2(1-p_d^B)(1-p_a^B)(p_a^B)^2+ && (0122)\\
&+ (p_d^A)^3(1-p_a^A)p_a^Ap_d^B(1-p_d^B)^2(p_a^B)^3+ && (0113)\\
&+ (p_d^A)^3(p_a^A)^2(p_d^B)^3(1-p_a^B)(p_a^B)^2+ && (0032)\\
&+ (p_d^A)^3(p_a^A)^2(p_d^B)^2(1-p_d^B)(p_a^B)^3. && (0023)
\end{aligned}
$$

Expressing $p66$

To compute $p66$, we note that there are 44 possibilities to distribute 6 points over 4 types of points with multiplicity 3.

$$
\begin{align}
p66 &= (1-p_d^A)^3(1-p_a^A)^3(1-p_d^B)^3(1-p_a^B)^3+ \tag{3300}\\
&+ (1-p_d^A)^3p_a^A(1-p_a^A)^2(1-p_d^B)^2p_d^B(1-p_a^B)^3+ \tag{3210}\\
&+ (1-p_d^A)^3p_a^A(1-p_a^A)^2(1-p_d^B)^3(1-p_a^B)^2p_a^B+ \tag{3201}\\
&+ (1-p_d^A)^3(p_a^A)^2(1-p_a^A)(p_d^B)^2(1-p_d^B)(1-p_a^B)^3+ \tag{3120}\\
&+ (1-p_d^A)^3(p_a^A)^2(1-p_a^A)(1-p_d^B)^2p_d^B(1-p_a^B)^2p_a^B+ \tag{3111}\\
&+ (1-p_d^A)^3(p_a^A)^2(1-p_a^A)(1-p_d^B)^3(1-p_a^B)(p_a^B)^2+ \tag{3102}\\
&+ (1-p_d^A)^3(p_a^A)^3(p_d^B)^3(1-p_a^B)^3+ \tag{3030}\\
&+ (1-p_d^A)^3(p_a^A)^3(p_d^B)^2(1-p_d^B)(1-p_a^B)^2p_a^B+ \tag{3021}\\
&+ (1-p_d^A)^3(p_a^A)^3(1-p_d^B)^2p_d^B(1-p_a^B)(p_a^B)^2+ \tag{3012}\\
&+ (1-p_d^A)^3(p_a^A)^3(1-p_d^B)^3(p_a^B)^3+ \tag{3003}\\
&+ (1-p_d^A)^2p_d^A(1-p_a^A)^3(1-p_d^B)^2p_d^B(1-p_a^B)^3+ \tag{2310}\\
&+ (1-p_d^A)^2p_d^A(1-p_a^A)^3a(1-p_d^B)^3(1-p_a^B)^2p_a^B+ \tag{2301}\\
&+ (1-p_d^A)^2p_d^A(1-p_a^A)^2p_a^A(1-p_d^B)(p_d^B)^2(1-p_a^B)^3+ \tag{2220}\\
&+ (1-p_d^A)^2p_d^A(1-p_a^A)^2p_a^A(1-p_d^B)^2p_d^B(1-p_a^B)^2p_a^B+ \tag{2211}\\
&+ (1-p_d^A)^2p_d^A(1-p_a^A)^2(p_a^A)(1-p_d^B)^3(p_a^B)^2(1-p_a^B)+ \tag{2202}\\
&+ (1-p_d^A)^2p_d^A(1-p_a^A)(p_a^A)^2(p_d^B)^3(1-p_a^B)^3+ \tag{2130}\\
&+ (1-p_d^A)^2p_d^A(p_a^A)^2(1-p_a^A)(p_d^B)^2(1-p_d^B)(p_a^B)(1-p_a^B)^2+ \tag{2121}\\
&+ (1-p_d^A)^2p_d^A(1-p_a^A)(p_a^A)^2(p_d^B)(1-p_d^B)^2(p_a^B)^2(1-p_a^B)+ \tag{2112}\\
&+ (1-p_d^A)^2p_d^A(1-p_a^A)(p_a^A)^2(1-p_d^B)^3(p_a^B)^3+ \tag{2103}\\
&+ (1-p_d^A)^2p_d^A(p_a^A)^3(p_d^B)^3p_a^B(1-p_a^B)^2+ \tag{2031}\\
&+ (1-p_d^A)^2p_d^A(p_a^A)^3(p_d^B)^2(1-p_d^B)(p_a^B)^2(1-p_a^B)+ \tag{2022}\\
&+ (1-p_d^A)^2p_d^A(p_a^A)^3p_d^B(1-p_d^B)^2(p_a^B)^3+ \tag{2013}\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)^3(1-p_d^B)(p_d^B)^2(1-p_a^B)^3+ \tag{1320}\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)^3(p_d^B)(1-p_d^B)^2p_a^B(1-p_a^B)^2+ \tag{1311}\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)^3(1-p_d^B)^3(p_a^B)^2(1-p_a^B)+ \tag{1302}\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)^2p_a^A(p_d^B)^3(1-p_a^B)^3+ \tag{1230}\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)^2p_a^A(p_d^B)^2(1-p_d^B)p_a^B(1-p_a^B)^2+ \tag{1221}\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)^2p_a^A(1-p_d^B)^2p_d^B(p_a^B)^2(1-p_a^B)+ \tag{1212}\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)^2p_a^A(1-p_d^B)^3(p_a^B)^3+ \tag{1203}\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)(p_a^A)^2(p_d^B)^3(p_a^B)(1-p_a^B)^2+ \tag{1131}\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)(p_a^A)^2(p_d^B)^2(1-p_d^B)(p_a^B)^2(1-p_a^B)+ \tag{1122}\\
&+ (1-p_d^A)(p_d^A)^2(1-p_a^A)(p_a^A)^2(1-p_d^B)^2p_d^B(p_a^B)^3+ \tag{1113}\\
&+ (1-p_d^A)(p_d^A)^2(p_a^A)^3(p_d^B)^3(1-p_a^B)(p_a^B)^2+ \tag{1032}\\
&+ (1-p_a^A)(p_d^A)^2(p_a^A)^3(p_d^B)^2(1-p_d^B)(p_a^B)^3+ \tag{1023}\\
&+ (p_d^A)^3(1-p_a^A)^3(p_d^B)^3(1-p_a^B)^3+ \tag{0330}\\
&+ (p_d^A)^3(1-p_a^A)^3(p_d^B)^2(1-p_d^B)(1-p_a^B)^2(p_a^B)+ \tag{0321}\\
&+ (p_d^A)^3(1-p_a^A)^3(1-p_d^B)^2p_d^B(1-p_a^B)(p_a^B)^2+ \tag{0312}\\
&+ (p_d^A)^3(1-p_a^A)^3(1-p_d^B)^3(p_a^B)^3+ \tag{0303}\\
&+ (p_d^A)^3(1-p_a^A)^2p_a^A(p_d^B)^3(1-p_a^B)^2(p_a^B)+ \tag{0231}\\
&+ (p_d^A)^3(1-p_a^A)^2p_a^A(p_d^B)^2(1-p_d^B)(1-p_a^B)(p_a^B)^2+ \tag{0222}\\
&+ (p_d^A)^3(1-p_a^A)^2p_a^Ap_d^B(1-p_d^B)^2(p_a^B)^3+ \tag{0213}\\
&+ (p_d^A)^3(p_a^A)^2(1-p_a^A)(p_d^B)^3(1-p_a^B)(p_a^B)^2+ \tag{0132}\\
&+ (p_d^A)^3(p_a^A)^2(1-p_a^A)(p_d^B)^2(1-p_d^B)(p_a^B)^3+ \tag{0123}\\
&+ (p_d^A)^3(p_a^A)^3(p_d^B)^3(p_a^B)^3. \tag{0033}
\end{align}
$$

References

[1] Carter, W. and S. Crews (1974). An analysis of the game of tennis, *American Statistician* **28**, 130-134.

[2] Gilbert, B. and S. Jamison (1993). *Winning Ugly*, Simon and Schuster, New York.

[3] Klaassen, F. and J. Magnus (2001). Are points in tennis independent and identically distributed? Evidence from a binary panel data model, *Journal of the American Statistical Association* **96**, 500-509.

[4] Knight, G. and P. O'Donoghue (2012). The probability of winning break points in Grand Slam men's singles tennis, *European Journal of Sport Science* **12**, 462-468.

[5] Mood, A., F. Graybill and D. Boes (1974). *Introduction to the Theory of Statistics*, McGraw Hill, New York.

[6] Newton, P. and K. Aslam (2006). Monte Carlo Tennis, *SIAM Review* **48**, 722-742.

[7] Newton, P. and J. Keller (2005). Probability of Winning at Tennis I. Theory and Data, *Studies in Applied Mathematics* **114**, 241-269.

FACULTY OF ECONOMICS AND BUSINESS
Naamsestraat 69 bus 3500
3000 LEUVEN, BELGIË
tel. + 32 16 32 66 12
fax + 32 16 32 67 91
info@econ.kuleuven.be
www.econ.kuleuven.be