SPREADSHEETS in EDUCATION

Bond University

Volume 12 | Issue 1 | 2019

Tie-Break Tennis Matches: A Case Study for First Year Level Statistics Teaching

Graham D I Barr,
University of Cape Town, gdi@iafrica.com

Leanne Scott,
University of Cape Town, Leanne.Scott@uct.ac.za

Stephen J Sugden
Bond University, ssugden@bond.edu.au

Tie-Break Tennis Matches: A Case Study for First Year Level Statistics Teaching

Graham D I Barr,
University of Cape Town, gdi@iafrica.com

Leanne Scott,
University of Cape Town, Leanne.Scott@uct.ac.za

Stephen J Sugden
Bond University, ssugden@bond.edu.au

Abstract

We explore the effects of different tie-break scoring systems in tennis and how this can be used as a teaching case study to demonstrate the use of basic statistical concepts to contrast and compare features of different models or data sets. In particular, the effects of different tie-break scoring systems are compared in terms of how they impact on match length, as well as the chances of the "underdog" winning. This case study also provides an ideal opportunity to showcase some useful spreadsheet features such as array formulae and data tables.

Keywords: *first-year statistics, tie-break, spreadsheet, tennis*

1 Introduction

In this paper we explore the effects of different tie-break scoring systems in tennis and how this can be used as a teaching case study to demonstrate the use of basic statistical concepts to contrast and compare features of different models or data sets. In particular, the effects of different tie-break scoring systems are compared in terms of how they impact on match length, as well as the chances of the "underdog" winning. This case study also provides an ideal opportunity to showcase some useful spreadsheet features such as array formulae and data tables.

The modelling of tennis match outcomes and the discussion of different tennis scoring systems has attracted a large amount of research work in the statistical literature. The original work of Schutz (1970), extended by Riddle (1988) and Barnett and Clarke (2002) demonstrated the use of constant probability Markov chains to model tennis match outcomes. These models were particularly useful for determining probabilities of match outcome from a partially completed tennis match and had applicability in a sports-betting context. In fact, the paper of Barnett and Clarke (2002), interrogates the assumption that games, sets and matches are treated independently and shows how a series of interconnected sheets in Excel can be used to model and predict the outcome of tennis matches or partially completed tennis matches without this assumption.

Barnett, Brown and Pollard (2006) address the problem of match length directly and use the mathematical method of generating functions to show that the likelihood of long matches can be substantially reduced by using the tiebreak game in the fifth set, or more effectively by using a new type of game, the 50-40 game, throughout the match. The 50-40 game as outlined in Pollard and Noble (2004) is such that to win the game, the server requires four points and the receiver requires three points and there is, thus, at most six points played in this type of game. Pollard (1983) also did some of the original work on tennis match duration and calculated the mean and variance of the duration of a best-of-three sets, with and without a tiebreaker system.

Tennis match outcomes, and the associated scoring systems, has also provided a rich area for the **teaching** of mathematical and statistical concepts. A notable contributor has been the book of Stewart (1991) who has taken a playful approach to discussing and modelling the quirks of the tennis scoring system within a probability framework. Noubary Reza (2010) has treated the subject more comprehensively, and has used tennis, and the quirks of the tennis scoring system, as a platform to discuss a range of mathematical and statistical techniques including, inter alia, recursion, Markov chains, conditional probability, performance measures and even tournament organisation.

2 The Tiebreaker (TB) system

TB10 tennis tournaments, similar to other rapid result forms of spectator sport like Rugby 7s and T20 Cricket, have become increasingly popular. The inaugural TB10 tournament took place in London in 2015 and since then there have been regular TB10 tournaments, patronised by top professional players competing for significant prize money.

The TBX format consists of the following rules:

(i) First player to reach X points wins, provided they have a margin of at least 2 points on their opponent.
(ii) The match continues until one player has at least X points and a lead of at least 2 points on the opponent.
(iii) The first to serve is decided by a toss. Winner of the toss serves one serve, thereafter each player serves twice. Side change every six points.

Tennis provides a context which many students are familiar with, mainly due to the high-profile nature of international competitions which enjoy prime time TV coverage, as well as the celebrity status of some tennis champions. However, even students who are not familiar with the game of tennis can relate to the curious fact that vastly different scoring systems have been adopted across the spectrum of sporting codes. Consequently, as a teacher, one can draw on students' familiarity with a variety of different sporting codes and get them to contemplate the purpose behind the particular scoring system adopted by their own favourite sport. A discussion around the likely effects of the scoring system on the game itself across different sports will help to prompt students to consider which variables we might want to explore in our investigation.

The TB format appears to have developed in response to a demand to make the game of tennis a faster and more exciting spectator sport. It is worth noting, however, that the tie-break system has always been a part of tennis scoring, either implicitly or explicitly; we discuss this below and give a brief history of the introduction of the explicit tiebreak system.

2.1 The Tiebreak in Traditional Tennis

An explicit tiebreak system was first introduced into tennis at the US Open in 1970 and invoked at Wimbledon in 1979 to more rapidly reach a conclusion in evenly balanced matches. Although the scoring system already had a form of tiebreak at the **game** level, with players having to be 2 games ahead to win a set, this did not solve the problem of enduring sets, since each game could itself be very lengthy, and the set could continue for many games!

The introduction of the explicit tiebreak was precipitated by a number of long and extended matches played in major tournaments. One such example being the extraordinary first-round men's singles match in 1969 at Wimbledon between Pancho Gonzales and Charlie Pasarell. This was a 5-set match that lasted five hours and 20 minutes and took 2 days to complete. In the fifth set, the 41-year-old Gonzales managed to survive seven match points against him, twice coming back from 0–40 deficits but managing to win 22–24, 1–6, 16–14, 6–3, 11–9. However, even though the tie-break has been used in all Grand Slam tournaments for all sets except a (possible) 5th deciding set, the final set is not decided by a tie-break except for the US Open. The most recent example of the problematic absence of a tiebreak in the 5th deciding set at Wimbledon occurred in the 2018 men's semi-final at Wimbledon between Kevin Anderson and John Isner 7-6, 6-7, 6-7, 6-4, 26-24 which lasted 6hr and 35 minutes and pushed the 2nd semi-final match into an over-2-day affair!

On completion of the match, both players were supportive of a change in the current scoring system and were supportive of a TB7 tie-break system kicking it at 12-12 in games in the final set of Wimbledon. Anderson said. “If a match is 12-All in the fifth set, I don't think it needs to continue. The amount of times it gets to that point is pretty rare. I think it protects the players' health as well. Because being out there for this length can be pretty damaging from a health standpoint, too.” Isner also lent his support to that proposal. “I agree with Kevin,” Isner said. “I personally think a sensible option would be 12-All. If one person can't finish the other off before 12-All, then do a tiebreaker there. I think it's long overdue.”

The explicit tiebreak system (still currently in use in Grand Slam tournaments) attempts to circumvent this problem of extended sets and the referee will declare a points-based tiebreak in a set when the score is 6 games all. The players then play a series of points with the first player to reach at least 7 points, and be at least 2 points ahead, declared the winner of the **set**. Thus, a score of 10-8 points would win the set, but if both players reach 9 points each (9-9), the game must progress until one player is 2-points ahead. Hence a winning tiebreak score of 10-8 points is possible, but not a score of 10-9.

The scoring system in tennis also includes an *implicit* tiebreak system within each game. Each game in tennis is effectively a TB4 system, declared as “Deuce” when both players have 3 points. In the current scoring system, this is when the score is 40-40. The game continues from this point until one player is 2 points ahead. When “Deuce” is declared the game continues and then, if player A (say), scores the next point, it is scored as “Advantage A”. Then if player A wins the subsequent point (and is thus 2 points ahead, and has scored at least 4 points in total in that game), player A is awarded the game. If A loses the next point after having “Advantage”, the game score is reset to “Deuce”. The game continues until one of the players is two points ahead[1].

Thus, the traditional scoring system in tennis can be said to use points-based tiebreaking systems to decide games and games-based tiebreaking systems to decide sets.

In this paper we explore the effect of different TB scoring systems in terms of (i) probability of the weaker player winning (ii) expected length of each match and (iii) variability of match length. However, as an interesting aside we also explore whether it makes any material difference whether, within each game, we declare a tiebreak (“Deuce”) at 40-40 or at 30-30?

2.2 Deuce declared 30-30 is equivalent to declaring Deuce at 40-40

The issue of whether calling deuce at 30-30 rather than at 40-40 would influence the final outcome of a tennis game is well illustrated by a deterministic finite state automaton (DFA).

[1] . *It is worth noting that some tennis tournaments such as the World Team Championships use “no-ad” scoring system, in which the first person to win a point at deuce wins the game.*

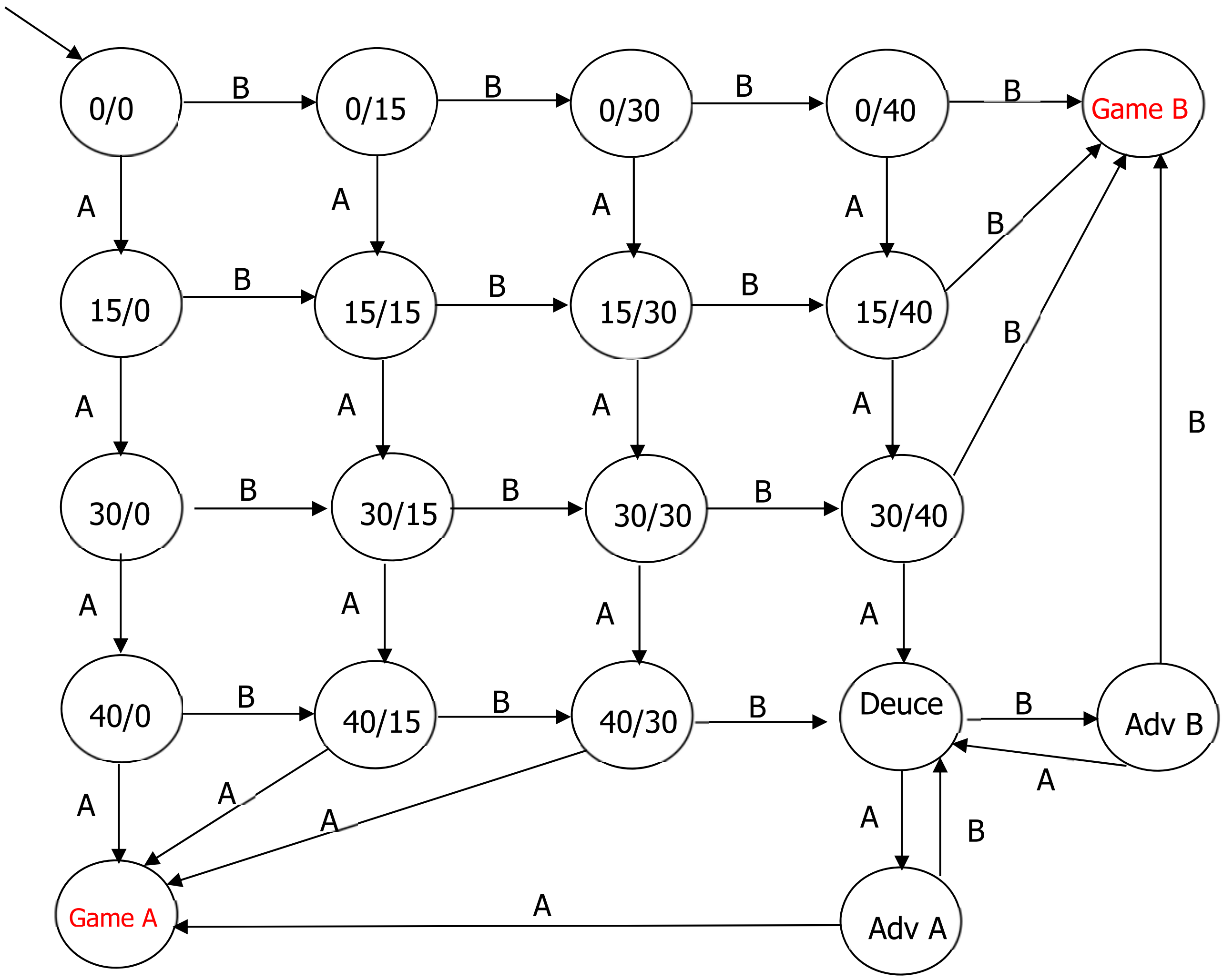


Figure 1 The Finite state Automaton of a Tennis Game

DFA are well-known in computer science and our instance here serves as a very useful diagrammatic illustration of simple tennis scoring. It is shown in Figure 1.

In the case of a tennis match, the initial state is 0-0 for both players; and we can use the diagram to track the changing score states until we reach the final state of the automaton which is either that A wins the game, or that player B wins the game.

The automaton of Figure 1 tracks the flow of possible score-states from the initial state of 0-0 to the two possible final games states. The symmetry of state possibilities in the two cases (30-30) or 40-40 is easily seen. In the case of 30-30, either A or B can win the game by scoring the next 2 points; similarly for the case of 40-40 (Deuce). If the score is 30-30 each and A and B both win one of the next 2 points, the state of Deuce is obtained.

Similarly, if one is already in the state of Deuce, and A and B both win one of the next 2 points, the score is returned to Deuce. Alternatively, if A and B win alternate points then, in the case where Deuce is the current state, the score simply returns to Deuce. In the case of 30-30 the score will reach Deuce and the keep returning to Deuce.

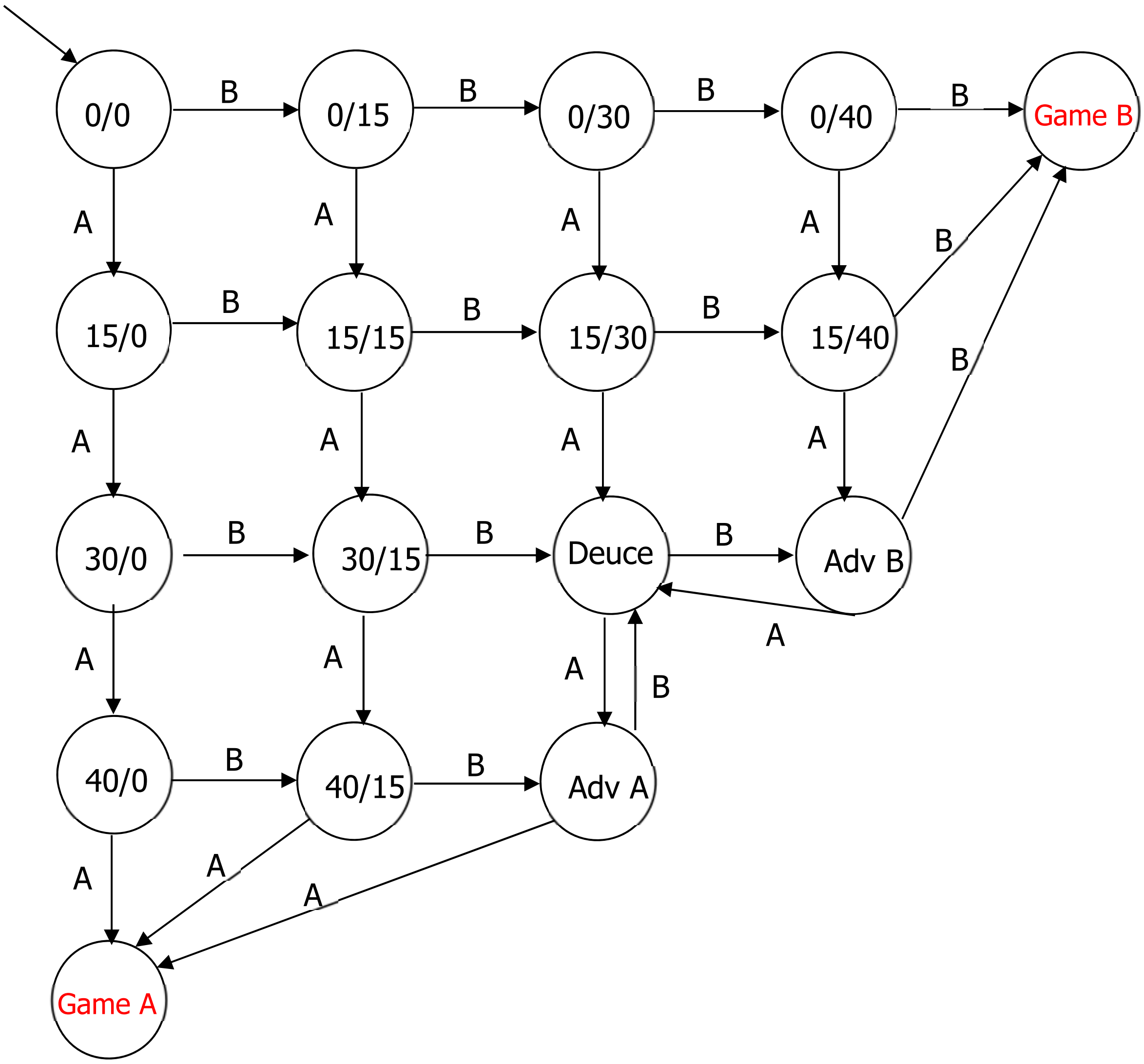


Figure 2 Reduced machine

Thus 30-30 all works as an effective Deuce and were Deuce called at 30-30, it would have no effect on the outcome of a game. Of course, experienced tennis players know that 30-30 is equivalent to Deuce. They are also aware that Advantage A is equivalent to 40/30 and that Advantage B is equivalent to 30/40.

The standard theory of deterministic finite state automata (DFA) includes a state-minimization algorithm (Aho et al, 1986). When applied to a given DFA, it will merge equivalent states and produce a "minimal equivalent machine". The concept of "equivalence" here is in the strict sense of equivalence relation: outcomes from all possible games will be identical to those of the original automaton. "Minimal" means that no equivalent machine with fewer states is possible. Our machine of Figure 1 has 20 states, with the two accept states indicated in red as Game A and Game B. Application of the standard minimization algorithm to the present machine yields an equivalent reduced machine of 17 states. Three states are lost as follows:

- 40/30 is merged with Advantage A,
- 30/40 is merged with Advantage B,

- 30/30 is merged with Deuce.

Our reduced machine appears in Figure 2.

We may explore the general principle of this "alternative Deuce" issue further. A normal tennis game, as mentioned above, is a T(4). Hence, we have concluded that calling "Deuce" in the standard T(4) tennis game at 3-points-all (40-40), OR at 2-points-all (30-30) does not affect the outcome of the tennis game. We can generalise this to say that in a hypothetical TB(n) game, calling "Deuce" at $n-1$ points-all OR at $n-2$ points-all will not affect the outcome of a TB(n) determined game. So, to take the example of a standard T(7) tiebreak, we could call Deuce at a score of 6-6 points OR, at a score of 5-5 and it would not affect the outcome of the tiebreak. Conventionally, of course, in a tennis tiebreak, "Deuce" is not explicitly called, and the game continues until one player has at least 7 points and is at least 2 points ahead.

2.3 Probability of player B (the "underdog") winning the match:

To construct the probability of the weaker player winning the match, we begin by defining the probabilities of each player winning an individual point. We let the probability of player A winning any point be p_A and that for player B be p_B ($= 1 - p_A$). Although it can be argued that players have a different probability of winning a point when they serve, this still results in an average probability of winning a point (taken over all points in a match). Table 1 below shows an example where player A has an average probability of 0.55 winning any point (regardless of who is serving) and thus the probability of B winning a point is 0.45. Clearly, in tennis, the player serving has an advantage. This could be (for example) that A has a 0.7 chance of winning points when serving, whereas player B only has a 0.6 probability of winning a point when serving. These probabilities are consistent with overall probabilities of winning points, $p_A = 0.55$ and $p_B = 0.45$ (because they meet the requirement that the probability of A winning a point while serving is 0.1 greater than that for player B).

Table 1 Probability of players winning (serving or not)

	$p_A = 0.55$	$p_B = 0.45$
	A serves	B serves
A wins	0.70	0.40
B wins	0.30	0.60

In order to win, player B needs to obtain TB points and to lead by at least 2 points. Thus, for example, in a TB7 match, player B can win in the following ways, as shown in Table 2.

Table 2 Probability of B winning TB7 match

A points	B points	Probability
0	7	p_B^7
1	7	$p_A^1 p_B^7 \binom{7}{1}$
2	7	$p_A^2 p_B^7 \binom{8}{2}$
3	7	$p_A^3 p_B^7 \binom{9}{3}$
4	7	$p_A^4 p_B^7 \binom{10}{4}$
5	7	$p_A^5 p_B^7 \binom{11}{6}$
6	8	$p_A^6 p_B^8 \binom{11}{6}\binom{2}{1}$
7	9	$p_A^7 p_B^9 \binom{11}{6}\binom{2}{1}\binom{2}{1}$
8	10	$p_A^8 p_B^{10} \binom{11}{6}\binom{2}{1}\binom{2}{1}\binom{2}{1}$
9	11	$p_A^9 p_B^{11} \binom{11}{6}\binom{2}{1}\binom{2}{1}\binom{2}{1}\binom{2}{1}$
etc		

The probability that player B wins a TB match thus comprises an infinite series. It may be decomposed into 2 parts:

1. *TB-2* terms of a (truncated) Negative Binomial, NB(*TB*, p_A)

$$\sum_{k=1}^{TB-2} \binom{TB+k-2}{k-1} p_A^{k-1} p_B^{TB}$$

2. Subsequent terms: *TB-1, TB, TB+1* … are of the form,

$$\binom{2TB-3}{TB-1} p_A^{TB+i-2} p_B^{TB+i} 2^{i-1} \quad , \qquad i = 1,2,\ldots$$

and decline geometrically by a factor of $2p_A p_B$.

This remaining infinite series thus sums to,

$$\frac{1}{1-2p_A p_B}\binom{2TB-3}{TB-1} p_A^{TB-1} p_B^{TB+1}$$

The probability that B wins the match is thus,

$$\sum_{k=1}^{TB-2} \binom{TB+k-2}{k-1} p_A^k p_B^{TB} + \frac{1}{1-2p_Ap_B} \binom{2TB-3}{TB-1} p_A^{TB-1} p_B^{TB+1}$$

In the case of a TB7, we would thus obtain:

$$p_B^7 + 7p_A p_B^7 + 28p_A^2 p^7{}_B + 84p_A^3 p^7{}_B + 210p_A^4 p_B^7 + \frac{462p_A^5 p_B^7}{1-2p_Ap_B}$$

3 Demonstrating the probabilities of winning in a TB using a spreadsheet

We can easily set up a spreadsheet to calculate these probabilities for TB matches of different lengths eg TB7, TB10, TB15, etc. and for players of different relative strengths. Table 3, below, shows some of the output that has been summarised from the spreadsheet.

Table 3 Probability of B winning TB7, TB10, TB15 or TB15 match

		TB7	TB10	TB15	TB20
p_A	p_B	P(B wins TB7)	P(B wins TB10)	P(B wins TB15)	P(B wins TB20)
0.50	0.50	0.500	0.500	0.500	0.500
0.55	0.45	0.346	0.321	0.287	0.259
0.60	0.40	0.213	0.174	0.128	0.097
0.65	0.35	0.113	0.078	0.043	0.024
0.70	0.30	0.051	0.027	0.010	0.004
0.75	0.25	0.018	0.007	0.001	0.000
0.80	0.20	0.005	0.001	0.000	0.000
0.85	0.15	0.001	0.000	0.000	0.000
0.90	0.10	0.000	0.000	0.000	0.000
0.95	0.05	0.000	0.000	0.000	0.000
1.00	0.00	0.000	0.000	0.000	0.000

The information in Table 3 is depicted graphically in Figure 1.

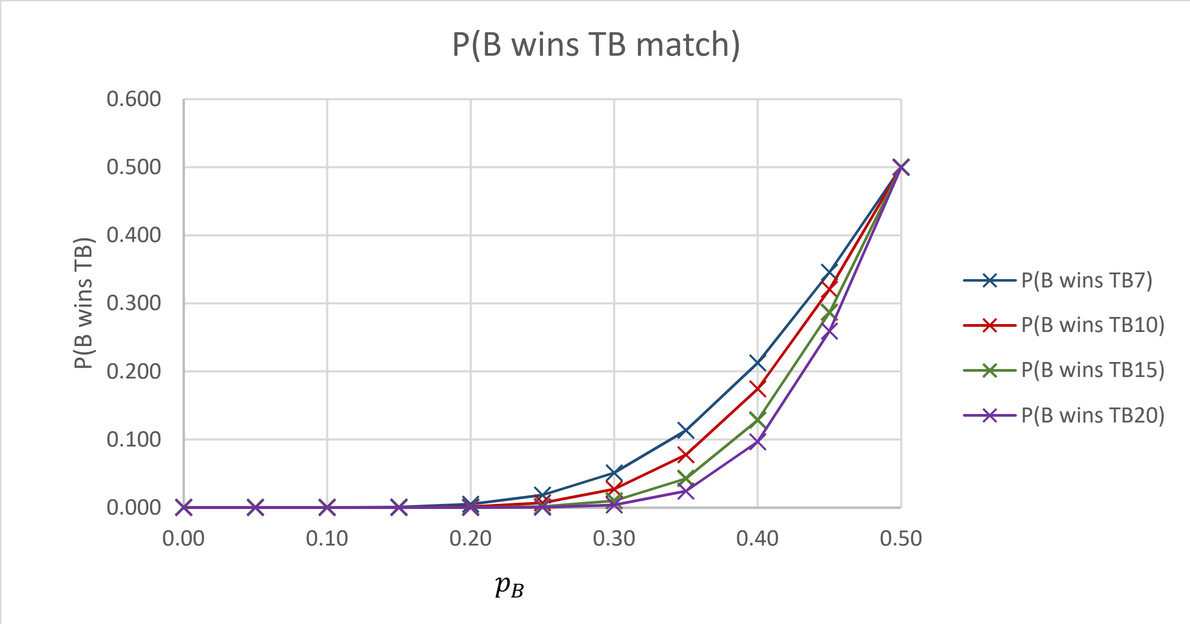


Figure 1 Probability that B wins a TB7, TB10, TB15 or TB20 match

We can see from Table 3 and Figure 1 that the shorter the TB format, the greater the chance of an underdog creating an upset and beating a stronger player. Thus a longer game is clearly a better way of teasing out the better player, but clearly the outcome is more predictable and hence less exciting than for a TB with a shorter format.

This raises another issue, namely the expected length of a tennis match which has important scheduling implications for both match officials and television producers. In theory, any TB match can continue for an infinite length of time, and the spreadsheet is an ideal tool to explore the variation in match length given differing relative player strengths and different TB formats.

3.1 The Expected Length of a TB match

To find the expected length of a match (L) we need to find

$E(L) = \sum_{i=1}^{\infty} L_i \ p(L_i) \qquad for\ L_i \in \{possible\ TB\ match\ lengths\}$

So, for example, E(L) for TB7 would be as shown in Table 4 (for $p_A = p_B = 0.5$).

These computations are easily done on a spreadsheet using **array formulas** for succinctness. To demonstrate we use the following spreadsheet (screenshot below), which gives the probabilities of a TB7 for outcomes up to (the very unlikely 20, 18) case. The spreadsheet is set up to give the probabilities for each outcome and the overall Expected Match Length and Standard Deviation (Match Length). See page ***TB7 using Array Formulae*** of attached spreadsheet.

We give a screenshot below of the "*TB7 using Array Formulae*" spreadsheet page.

A	B	C	D	E
2	**TB7 in Tennis - use of Array formulas to Calculate Length of Matches**			
3	**TBn system**	**n**	**7**	
4	**Player a (Stronger)**	**P(a wins a point)**	**0.50**	
5	**Player b (Weaker)**	**P(b wins a point)**	**0.50**	
6				
7	**Prob b wins TB7**	**0.5000**	**E(Match Length)**	**11.743**
8	**Prob a wins TB8**	**0.5000**	**StdDev(Match Length**	**2.905**
9				
10	**7**	**TB7**	**Case b wins**	**Expected Match Length**
11	**pts for b**	**pts for a**	**Prob b wins**	**(for each case)**
12	7	0	0.0078	0.0547
13	7	1	0.0273	0.2188
14	7	2	0.0547	0.4922
15	7	3	0.0820	0.8203
16	7	4	0.1025	1.1279
17	7	5	0.1128	1.3535
18	8	6	0.0564	0.7896
19	9	7	0.0282	0.4512
20	10	8	0.0141	0.2538
21	11	9	0.0070	0.1410
22	12	10	0.0035	0.0775
23	13	11	0.0018	0.0423
24	14	12	0.0009	0.0229
25	15	13	0.0004	0.0123
26	16	14	0.0002	0.0066
27	17	15	0.0001	0.0035
28	18	16	0.0001	0.0019
29	19	17	0.0000	0.0010
30	20	18	0.0000	0.0005
31		**TB7**	**Case a wins**	
32	**pts for a**	**b**	**Prob a wins**	
33	7	0	0.0078	0.0547
34	7	1	0.0273	0.2188
35	7	2	0.0547	0.4922
36	7	3	0.0820	0.8203
37	7	4	0.1025	1.1279
38	7	5	0.1128	1.3535
39	8	6	0.0564	0.7896
40	9	7	0.0282	0.4512
41	10	8	0.0141	0.2538
42	11	9	0.0070	0.1410
43	12	10	0.0035	0.0775
44	13	11	0.0018	0.0423
45	14	12	0.0009	0.0229
46	15	13	0.0004	0.0123
47	16	14	0.0002	0.0066
48	17	15	0.0001	0.0035
49	18	16	0.0001	0.0019
50	19	17	0.0000	0.0010
51	20	18	0.0000	0.0005

Screenshot 1 Sheet to compute Probability that B wins, E(L) and Std. Dev. (L) for TB7 match

To compute the expected match length, we must multiply (Points for a + Points for b) by Probability (this configuration occurred).

As mentioned above, for convenient exposition on the spreadsheet for the TB7 case, we truncate at a score of (20, 18). Note that the probability of a higher score (21, 19 and above) is very small, even when the players are evenly matched (0.0000 to 4 decimal places).

The appropriate array formula for computing the match length is the following array formula:

E7= {SUM((B12:B30+C12:C30)*(D12:D30+D33:D51))}, where

B12:B30+C12:C30 represents the possible values for (Pts for a + Pts for b), and,

D12:D30+D33:D51 the associated probabilities, that is, the Prob (this configuration occurred).

The product of these quantities gives component match lengths (Expected Match Lengths for each case). To get the overall expected match length these quantities are summed and multiplied by 2, to cover the case that either a, or b may win. We see that for the case of TB7 in the case $p_A = p_B = 0.5$, the expected length of a match is 11.743

3.2 Expected Match Length as a function of p_A and p_B

One would expect the length of the match to depend on how evenly matched the players were. If we vary the relative probabilities of winning, we can see (Table 4 below) that for TB7, the maximum expected match length is when both players are exactly evenly matched (the case shown in Table 88 where $p_A = p_B = 0.5$). Table 4 below shows how match length varies with varying relative player abilities. As p_B decreases below 0.5, the match length will clearly reduce (the distribution of L is symmetric around 0.5) and so the maximum expected match length for a TB7 match is close to 12 points.

Table 4 Expected Length of TB7 Match for varying player strengths

Expected length of match for TB7 (E(L))		
p_A	p_B	E(L)
0.00	1.00	7.000
0.05	0.95	7.368
0.10	0.90	7.778
0.15	0.85	8.239
0.20	0.80	8.762
0.25	0.75	9.353
0.30	0.70	9.995
0.35	0.65	10.641
0.40	0.60	11.208
0.45	0.55	11.602
0.50	0.50	11.743

3.3 Using Data Tables to tabulate the Expected match length (and Standard Deviation) for a range of different values for p_A and p_B

One can generate a table of values as in Table 4 by laboriously changing the values of p_A and hence p_B and recording the associated values of expected match length (and standard deviation of length) from the spreadsheet. A much neater way of doing this is with the Data Tables feature of Excel which allows us to create a table of the expected length of a match (and standard deviation of length) for a range of different values for p_A (p_B) .

To demonstrate, we set up the spreadsheet table (see Screenshot1 below) on a sheet labelled Data Table with the range of values one might want to use for p_A (p_B) as substitutions into the Expected Match Length and the Standard Deviation (Match Length) formulas. Note that the array formulas will not work if we put the p_A (p_B) to be **exactly** 0 so we make the smallest value for p_A equal to 0.01.

A	B	C	D
2	**P(a wins)**	**E(Match Length)**	**StDev(Match L)**
3	0.01	7.0707	0.2672
4	0.05		
5	0.1		
6	0.15		
7	0.2		
8	0.25		
9	0.3		
10	0.35		
11	0.4		
12	0.45		
13	0.5		
14	0.55		
15	0.6		
16	0.65		
17	0.7		
18	0.75		
19	0.8		
20	0.85		
21	0.9		
22	0.95		
23	1		

Screenshot 1 Data Table Setup : Step1

We then link Cell C3 and cell D3 on the **Data Table** sheet to the appropriate cells on the spreadsheet. For demonstration purposes we copy the sheet **TB7 and using Array Formulas** and rename it **Data Table Input** sheet. Hence, we have

(C3)='Data Table Input'!E7

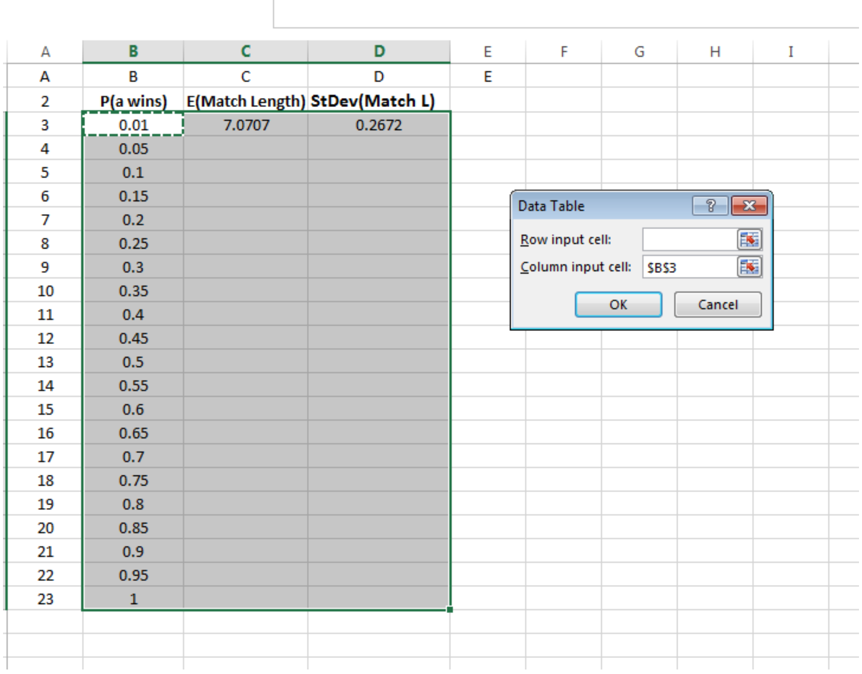
(D3)='Data Table Input'!E8

On the Data Table Input Sheet, we link the cell we want to vary (p_A) to the first cell of sheet **Data Table**.

D4 (of 'Data Table Input) ='Data Table'!B3

We then highlight **B3:D23** on the Data Table sheet, and click on Menu - DATA, What-If Analysis, Data Table.

We enter **B3** for the Column Input Cell; see Screenshot 2 below.



Screenshot 2 Data Table Setup : Step2

On clicking OK, I obtain the E(Match Length) and StDev(Match L) for the range of p_A as listed, to give the following sheet of values on page **Data Table (see Screenshot 3, below)**.

A	B	C	D
2	**P(a wins)**	**E(Match Length)**	**StDev(Match L)**
3	0.01	7.0707	0.2672
4	0.05	7.3684	0.6229
5	0.1	7.7783	0.9333
6	0.15	8.2391	1.2280
7	0.2	8.7624	1.5461
8	0.25	9.3531	1.8907
9	0.3	9.9954	2.2329
10	0.35	10.6408	2.5289
11	0.4	11.2081	2.7436
12	0.45	11.6017	2.8662
13	0.5	11.7430	2.9049
14	0.55	11.6017	2.8662
15	0.6	11.2081	2.7436
16	0.65	10.6408	2.5289
17	0.7	9.9954	2.2329
18	0.75	9.3531	1.8907
19	0.8	8.7624	1.5461
20	0.85	8.2391	1.2280
21	0.9	7.7783	0.9333
22	0.95	7.3684	0.6229
23	1	7.0000	0.0000

Screenshot 3 Data Table Setup : Step3

We now have a table of Expected Match Lengths (and Standard Deviations of Match Length) for an appropriate range of p_A

3.4 Expected Match Length for other TB formats

If we consider different TB formats, we see, as expected, that the distribution of expected match length is always symmetric around $p_A = p_B = 0.5$. Table 5 below shows how E(L) varies with varying relative player strengths for TB10, TB15 and TB20.

Table 5 Expected match lengths for TB10, TB15 and TB20 for varying player strengths

p_A	p_B	E(L) TB10	E(L) TB15	E(L) TB20
0.00	1.00	10.000	15.000	20.000
0.05	0.95	10.526	15.789	21.053
0.10	0.90	11.111	16.667	22.222
0.15	0.85	11.765	17.647	23.529
0.20	0.80	12.502	18.750	25.000
0.25	0.75	13.337	20.000	26.667
0.30	0.70	14.267	21.416	28.565
0.35	0.65	15.242	22.968	30.697
0.40	0.60	16.142	24.496	32.894
0.45	0.55	16.793	25.668	34.671
0.50	**0.50**	**17.032**	**26.114**	**35.371**
0.55	0.45	16.793	25.668	34.671
0.60	0.40	16.142	24.496	32.894

We display these results graphically in Figure 2 below.

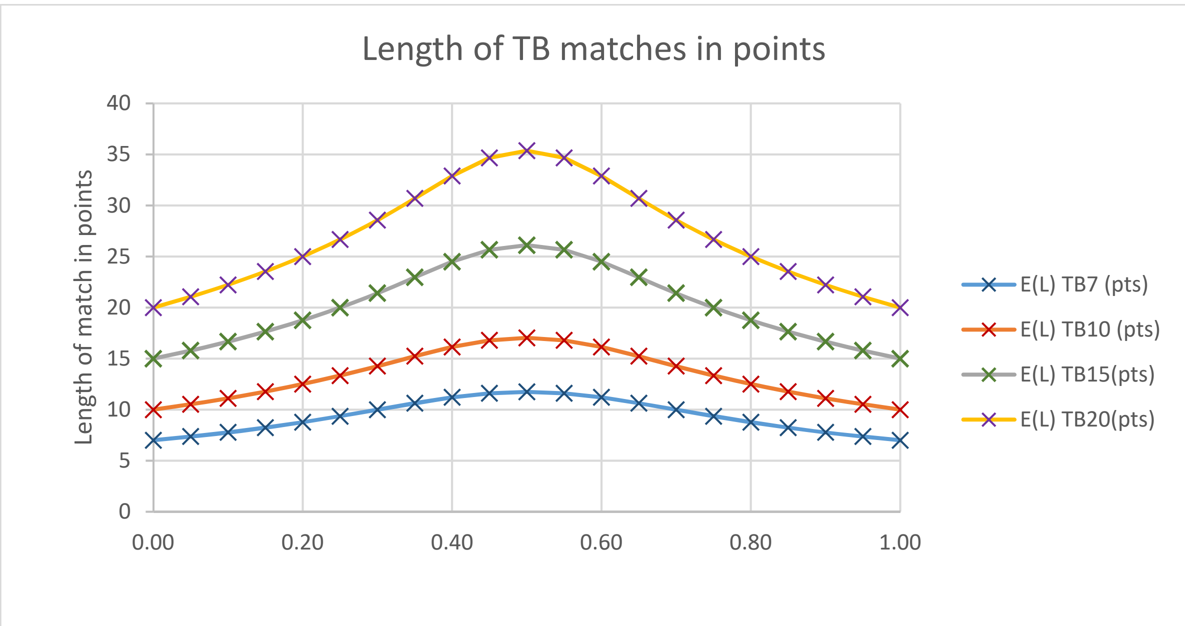


Figure 2 Expected Match Length of different TB formats.

It can be seen that, regardless of the minimum number of points required to win the match, the matches where players are evenly matched are expected to last the longest. Moreover, variability in match length raises a problem for match organisers who have to schedule tennis matches of unknown length into fixed TV schedules. There is thus a need to provide some sense of how great the uncertainty is around match length.

3.5 The Standard deviation of Match Length for TB formats

To assist with this we can calculate the theoretical standard deviations of the match lengths associated with values for $p_B$which will allow us to calculate appropriate confidence intervals of match length for any given p_B.

Thus, for each p_B, we compute the standard deviation of length of match (L) using the following formula:

$$Std.Dev.(L) = \sqrt{\sum_i L_i^2 p(L_i) - \left(\sum_i L_i p(L_i)\right)^2}$$

Where $p(L_i)$ is the probability that the match has length L_i.

The appropriate array formula for this is (see the Array Formulae page, noting the use of (<Ctrl><Shift><Enter>) to invoke the Array Formula):

(E8)={SQRT(SUM((B12:B30+C12:C30)^2*(D12:D30+D33:D51))-(SUM((B12:B30+C12:C30)*(D12:D30+D33:D51))^2))}, where

B12:B30+C12:C30 represents the possible values for (Pts for A + Pts for B); these are then squared. And,

D12:D30+D33:D51 contains the associated probabilities, that is, the probability that a particular match score was reached.

The expected value **{SUM((B12:B30+C12:C30)*(D12:D30+D33:D51))}** squared, is then subtracted.

The square root is then extracted to obtain the standard deviation.

Table 6 below shows how Std. Dev.(L) varies with varying relative player strengths for TB7, TB10, TB15 and TB20.

Table 6 Standard Deviation of Match Length (in points) for TB7, TB10, TB15 and TB20.

Standard Deviations of Match Length (points)					
p_A	p_B	T7	T10	T15	T20
0.00	1.00	0.000	0.000	0.000	0.000
0.05	0.95	0.623	0.744	0.912	1.053
0.10	0.90	0.933	1.111	1.361	1.571
0.15	0.85	1.228	1.444	1.765	2.038
0.20	0.80	1.546	1.782	2.166	2.500
0.25	0.75	1.891	2.145	2.587	2.982
0.30	0.70	2.233	2.514	3.021	3.489
0.35	0.65	2.529	2.832	3.396	3.951
0.40	0.60	2.744	3.041	3.486	3.974
0.45	0.55	2.866	3.133	3.555	4.033
0.50	0.50	2.905	3.153	3.571	4.041
0.55	0.45	2.866	3.133	3.555	4.033
0.60	0.40	2.744	3.041	3.486	3.974
0.65	0.35	2.529	2.832	3.396	3.951
0.70	0.30	2.233	2.514	3.021	3.489
0.75	0.25	1.891	2.145	2.587	2.982
0.80	0.20	1.546	1.782	2.166	2.500
0.85	0.15	1.228	1.444	1.765	2.038
0.90	0.10	0.933	1.111	1.361	1.571
0.95	0.05	0.623	0.744	0.912	1.053
1.00	0.00	0.000	0.000	0.000	0.000

Figure 3 displays the E(L) graphically with confidence intervals for the case of TB7 for different values of p_A.

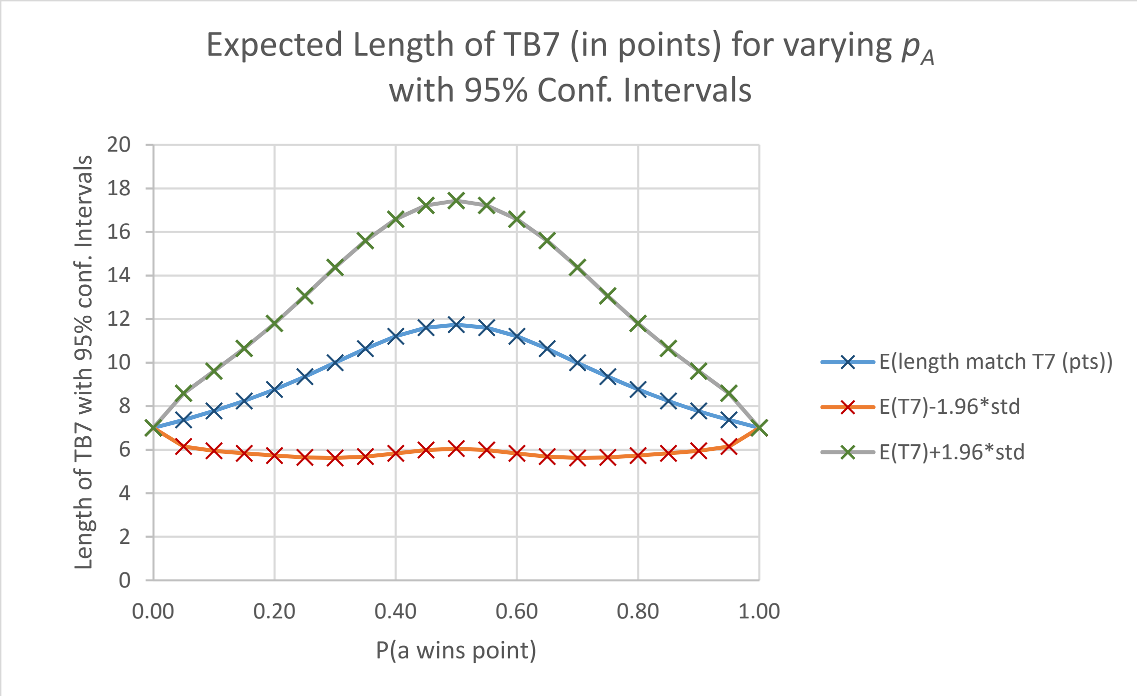


Figure 3 Expected Match Length with 95% confidence Intervals

4 Teaching aspects of this Case Study

Apart from the development of formulae for probability of a player winning a match and the expected length of a match, which are very appropriate exercises for first year students to apply the theory they have learnt, the TB case study allows us to demonstrate the power of using spreadsheets to explore variation across one or two dimensions. A typical tutorial would be structured as follows and the associated spreadsheet is attached:

(i) Students are stepped through how to set up a spreadsheet to work out the probability that the underdog wins for a TB7 match. (Cell **C7** on ***TB7 using Array Formulae*** page on spreadsheet). We start with the case where the players are evenly matched i.e. there is no real underdog and $p_A = p_B = 0.5$. For each possible winning score (resulting in player B winning), students calculate an associated probability (as shown in Table 2). This is a good exercise in basic probability using a Binomial scenario.

(ii) Expected length of TB7 calculation: Students employ their understanding of basic probability and expected value of a random variable.

(iii) In theory, TB matches can go on ad infinitum. We set up our TB7 tables (see page ***TB7 and using Array Formulas*** on the attached spreadsheet) to include all potential scores until the probabilities of a match continuing for that long become very small (eg the probability of a match score being 20-18 in favour of player B is 0.00001377 (to 8 decimal places)).
Using an array formula, students can calculate the Expected match length (L)
(E7)={SUM((B12:B30+C12:C30)*(D12:D30+D33:D51))}

Students can then, as an exercise, repeat these calculations for TB10, TB15 and TB20.

(iv) We ask the students to do a check that the probabilities of different scores for a match we have produced are reasonable. If the players are evenly matched ($p_A = p_B = 0.5$) then the overall probability that B wins a match must be 0.5. Summing over all probabilities of possible winning scores for B should give us 0.5 (we show students how we can do this check in cell **C7**).

(vi) In our initial exercise with the students, we work through all the calculations for probabilities of different final scores and expected match lengths for the case, $p_A = p_B = 0.5$. Instead of the students laboriously repeating this process for different values of p_A and p_B , we show the students how to make use of the Data Tables feature of Excel as demonstrated above. We discuss how students might do this in the section below. We spend some time talking about the features of data tables with the students and then get them to repeat the exercise for TB10, TB15 and TB20.

(v) The values in these tables can better be displayed as graphs and so we get the students to explore a variety of graphical options such as those in Figure 2 and Figure 3.

The teaching example above is quite demanding of Excel skills and although it might be used for a specialist (say Actuarial stream) first year Statistics course, we feel it is probably more appropriate for a second year course in Applied Statistics. We were, however, keen to use the vagaries of the tennis scoring system to introduce our first year students to a challenging problem and to introduce them to practical cases when the Binomial distribution was applicable. We thus set an assignment around the various tennis scoring systems for our first year Statistics course and discuss the assignment below.

4.1 The Tennis Assignment (for first year Statistics at the University of Cape Town (UCT))

At the University of Cape Town, we run a general first year Statistics course for 1600 students STA1000. This year we sketched out the tennis-scoring problem-backdrop and provided the students with the data from a simulation of 1 000 tennis matches, with results for different scoring systems. We then gave them the following assignment, with an associated set of questions. The data was provided in an Excel file, with different sheets for the match outcomes under the various different scoring systems, as discussed and described below.

Tennis is a game which has quite a strange scoring system. Neither time (like, for example, soccer) nor total points played (as, for example, in fencing) are fixed. This can result in unusually long matches such as at the 2018 Men's Wimbledon final! Variants of the tennis scoring system are now seriously being considered.

Let us look at two possible tennis scoring systems. Firstly we let two players A and B play a match in which exactly 11 points (alternating serves) are played. In each case, the player to win the most points wins. Note that these 11-point matches always continue until 11 points in total have been played (they wouldn't stop say at say 6-0 to A, because then A would be the guaranteed winner). This is because there is also a Grand Prize at the end of all the matches, for the player with the most points overall!

The Excel file "Tennis Data.xlsx" on sheet "11-point Matches" contains data on the outcome of 1 000 such tennis matches between player A and player B. We assume that the probability of A winning any point (p) is fixed over the 1000 matches. We are trying to work out who is the better player, and to estimate p. If we let X be the number of points won by A in any particular match, then we can see that X will follow a Binomial distribution with n equal to 11 and p unknown (p is the chance of A winning a single point).

1. As a starting point in trying to determine p , calculate the proportion of times A wins.
2. Compute also the average number of points won by A over the 1000 matches as an estimate of p.
3. Compute (from the data) an empirical estimate for the variance of X.
4. Using your best estimate for p (on the basis of questions 2and 3 above) calculate a value for the variance of X on the basis of n and p (and compare it to your answer above).

We now consider a second tennis scoring system, namely 7-point tie-break. The Excel sheet "Tiebreak Matches" contains data on the outcome of 1 000 tennis matches between player A and player B, when playing 7-point tie-break tennis. 7-point tie-break is a scoring system whereby the winner is the first player to obtain at least 7 points and be at least 2 points ahead of their opponent. We assume, again, that the probability of A winning any point (p) is fixed over the 1000 matches and that we are trying to estimate p.

Note the difference between this system and the former. Fixed-point matches can be modelled as a Binomial distribution as the total number of points played per match (trials) is fixed**. However, in tie-break matches, the total number of points played varies, since one player must be at least 2 points ahead to win.**

5. Using the Excel data, calculate the average chance of A winning a 7-point tiebreak match (i.e. the proportion of times A wins).
6. Now calculate the average proportion of the points won in each match by player A.
7. If player A is the better player, do you think their chance of winning a match is higher than their chance of winning a point? Or lower?
8. If one lengthened the number of points that a player was required to win, in order to win a tie-break match, from 7 to say 50, would this raise or lower the percentage of matches won by the better player?
9. Would the weaker player be better off playing a short or a long tie-break tennis match?

This assignment was well received and well executed and represents, we believe, an appropriate challenge for Statistics students at first year level. It forms part of our "modelling thrust" focus which we use when teaching first year Statistics. That is, rather than giving students a theoretical focus, often unnecessarily dense in mathematics, we give them problems which are thought provoking, and challenge them to think about solving these problems in a statistical modelling context.

5. Discussion and Conclusion

In this paper we have used a topical conundrum in professional tennis competitions as a teaching exercise for statistics students. We believe that this provides a practical and alluring example of how statistics can be employed to solve problems. For those students who are not likely to continue with the study of statistics and for whom the subject is a bit of a chore, it at least indicates that the subject can be relevant and applicable and provides them with an opportunity to apply some basic statistical concepts. For those students who find themselves drawn to pursue the subject, it showcases the ability of statistics to support applied research into everyday phenomena. The example of the tennis scoring dilemma and the teaching case study above provide a variety of hooks for student interest. The use of spreadsheets assists by making the calculations at once understandable (the different elements of calculation are broken down into bite-sized chunks!) and also relatively painless (repetition is made easy and seamless).

The students will also begin to be challenged to think about how to model an everyday situation of uncertainty (how do we begin to think about which player will win and how long it will take?). What's a good way to check that our thinking (about which player will win) is correct? What matters when it comes to choosing between different game formats (ie what will make us prefer one game format over another)? From this we hope to provoke the students to consider that equally important to finding the right answer is finding the right question. In this vein, spreadsheets are an extremely valuable tool to allow the students to experiment with different "questions". The "What If" tools provided in Excel (of which we have only introduced Data Tables in this exercise) are extremely useful, particularly in Statistics, in enabling students to explore data and investigate patterns, trends and hypotheses. In this exercise we were able to answer a number of practical questions, of the kind that tournament organisers and players alike might need answers to, with regard to different forms of TB Tennis and the uncertainty surrounding match length and probability of players of differing relative strength winning.

References

1. Aho, Sethi, Ullman, Compilers: Principles, Techniques, and Tools, Addison-Wesley, 1986. ISBN 0-201-10088-6
2. Barnett T, Brown A and Pollard G (2006). Reducing the likelihood of long tennis matches. J Sports Sci Med 5: 567-574.
3. Barnett TJ & Clarke, SR (2002) Using Microsoft Excel to model a tennis match. In Cohen G. and Langtry T. (eds). Proceedings of the Sixth Australian Conference on Mathematics and Computers in Sport. Bond University: Queensland, pp. 63-68.
4. Barnett TJ & Clarke, SR (2005) Combining player statistics to predict outcomes of tennis matches. IMA Journal of Management Mathematics, Vol 16, 2, 113-120.
5. Noubary Reza (2010) Teaching Mathematics and Statistics Using Tennis. In Joseph A. Gallian (ed) Mathematics and Sports, Mathematical Association of America (pub.).
6. Pollard GH (1983) An Analysis of Classical and Tie-breaker Tennis, Austral. J. Statist., 25, 496-505.
7. Pollard G and Noble K (2004). The benefits of a new scoring system in tennis: the 50-40 game. In R. Morton and S. Ganesalingam (eds.) Proceeding of the Seventh Australasian Conference on Mathematics and Computers in Sport. Massey University, Massey, New Zealand. 262-265.
8. Riddle LH (1988) probability models for tennis scoring systems Appl Stat 37: 63-75
9. Schutz RW (1970), A mathematical model for evaluating scoring systems with specific reference to tennis, Res. Quart. For Exercise and Sport, 41, 552-561.
10. Stewart I (1991) Game, Set and Math: Enigmas and Conundrums. Penguin: London.