THE MATHEMATICS OF TENNIS

By

Tristan Barnett and Alan Brown

Strategic Games
ABN: 64663038362
www.strategicgames.com.au

Table of Contents

About the Authors

Tristan Barnett has lived tennis all of his life - as a club player for Roseville Park, Macquarie University and Hawthorn Tennis Club, as an academic with a PhD in tennis statistics from Swinburne University, as a performance analyst with Tennis Australia, and as a mathematician in predicting tennis outcomes for betting organizations Ladbrokes and Centrebet. Tristan also founded and is managing director for Strategic Games - an online site offering material on the mathematics of sport, gambling and conflicts. It was perhaps the realization that Tristan would not quite make it as a professional on the tennis circuit, and subsequently used his mathematical ability to become a "professional" tennis expert in the science of the game.

Alan Brown is not a tennis player. His wife, Marion, has played tennis for more than 50 years, and so Alan has an appreciation of the game. His training was in mathematics, education and statistics at the University of Melbourne. He had a long career in insurance working as a computer programmer, operations research scientist, and actuary specializing in life insurance, health insurance and investment valuation. His research interest is in the use of cumulants to measure risk in many practical contexts. Now retired, he has enjoyed using his computer to analyze risk in tennis.

Preface

The longest professional tennis match, in terms of both time and total games occurred at the first-round of Wimbledon 2010 between John Isner and Nicholas Mahut. It lasted 183 games, required 11 hours and 5 minutes of playing time, with Isner winning 70-68 in the advantage final set. Even with the introduction of a tiebreak set at Wimbledon in 1971 long matches still occur and records of long matches can still be broken. Was this long match predictable and what are the chances of this record being broken in the future? This book will provide insights to these questions by formulating a mathematical model that provides information such as chances of players winning the match, reaching the advantage final set and reaching 68-68 all in an advantage final set. Hence the mathematics of tennis is concerned with the chances of players winning the match (who is likely to win?) and match duration (when is the match going to finish?). These calculations are required prior to the start of the match, but also for the match in progress. For example, what are the chances of player A winning the match in 4 sets from 1 set all, 3 games all, 30-15 and player A serving? Whilst the mathematics of tennis could be of interest to tennis organizations, commentators, players, coaches and spectators; it could also be applied to teaching by using the well-defined scoring structure of tennis to teach concepts to students in probability and statistics. Such concepts include summing an infinite series, Binomial theorem, backward recursion, forward recursion, generating functions, Markov chain theory and distribution theory. The mathematics of tennis applied to teaching also allows students to build their own tennis calculator using spreadsheets, which in turn could assist in the understanding of probability and statistical concepts, and familiarize students with using spreadsheet software such as Excel.

The book is structured as follows. Chapters 1-7 define the mathematical model by obtaining calculations for chances of winning (chapters 1-3) and duration (chapters 4-7). The chances of winning is comprised firstly of chapter 1 on the chances of winning a game by demonstrating the results using mathematical techniques of counting paths, Binomial theorem, Markov Chain theory, backward recursion and forward recursion. This is followed by chapter 2 on the chances of winning a match using backward recursion and chapter 3 on the chances of winning a match using forward recursion. Duration is comprised firstly of chapter 4 on the duration of a game by demonstrating the results using mathematical techniques of the Binomial theorem, generating functions, forward recursion and backward recursion. Chapters 5-7 obtain calculations on the duration of a match using backward recursion. Chapter 5 obtains calculations for the mean and variance of the number of points in a game, the number of games in a set and the number of sets in match. This is extended in chapter 6 by obtaining calculations for the mean, variance, and coefficients of skewness and kurtosis of the number of points in a game, the number of games in a set and the number of sets in match. This is further extended in chapter 7 to obtain calculations for the number of points in a set and a match, followed by the time duration in a match. Chapter 8 outlines a prediction model by utilizing the mathematical model developed in chapters 1-7. Chapter 8 demonstrates the methodology by focusing on the 'long' mens singles match between John Isner and Nicholas Mahut at the 2010 Wimbledon Championships. The appeal of how predictions can be used in sports multimedia is also presented in chapter 8.

The authors would like to thank the following people for their support, encouragement, discussions and feedback in formulating this book.

Anthony Bedford - Associate Professor of Sports Statistics: RMIT University
Stephen Clarke - Professor of Sports Statistics: Swinburne University
Javier Maquirriain - President: Society for Tennis Medicine and Science
Stuart Miller - Executive Director: International Tennis Federation
Darren O'Shaughnessy - IT Consultant: Ranking Software
Babette Pluim - Former President: Society for Tennis Medicine and Science
Geoff Pollard - Vice President: International Tennis Federation
Graham Pollard - Professor of Applied Statistics: University of Canberra
Vladimir Ejov - Associate Professor of Applied Mathematics: Flinders University

Tristan Barnett would also like to thank the University of South Australia, School of Mathematics and Statistics for providing an office and workplace facilities to undertake this project. Tristan has been an adjunct research fellow with the university since 2011. Finally, Tristan would like to thank Kathleen Lumley College for providing an affordable and pleasant housing environment for academic visitors.

Chapter 1

Winning a game

1.1 Introduction

Chapter 1 obtains calculations for the probability of winning a game. This consists of winning the game from any point score within the game (i.e. winning the game from 30-0) and winning the game to a specific point score (i.e. winning the game to 0). The techniques to obtain these calculations consist of counting paths, Binomial theorem, Markov Chain theory, backward recursion and forward recursion.

1.2 Counting paths

Consider a game of tennis between the server and the receiver. The score starts at (0,0). If the server wins the point at (0,0) then the score progresses to (15,0). If the receiver wins the point at (0,0) then the score progresses to (0,15). Therefore the score progresses to either (15,0) or (0,15) from (0,0) as shown in figure 1.1.

$$
\begin{array}{ccc}
 & & (15,0) \\
 & \nearrow & \\
(0,0) & & \\
 & \searrow & \\
 & & (0,15)
\end{array}
$$

Figure 1.1: Graphical representation of the first point played in a game

If the server wins the point at (15,0) then the score progresses to (30,0). If the receiver wins the point at (15,0) then the score progresses to (15,15). If the server wins the point at (0,15) then the score progresses to (15,15) and if the receiver wins the point at (0,15) then the score progresses to (0,30). Therefore the score progresses to either (30,0) or (15,15) from (15,0), and the score progresses to either (0,30) or (15,15) from (0,15) as shown in figure 1.2.

$$\begin{array}{ccccc} & & & & (30,0)\\ & & & \nearrow & \\ & & (15,0) & & \\ & \nearrow & & \searrow & \\ (0,0) & & & & (15,15) \\ & \searrow & & \nearrow & \\ & & (0,15) & & \\ & & & \searrow & \\ & & & & (0,30)\end{array}$$

Figure 1.2: Graphical representation of the first two points played in a game

This process continues until either the server or receiver has won the game. To win a game requires winning four points. However if the scores are level after six points have been played (known as deuce) then play continues indefinitely until a player has a established a two point lead and wins the game. Hence the structure of a game can be broken up into two parts - winning a game prior to deuce and winning a game from deuce. For the server to win the game prior to deuce requires winning 4 points whilst the receiver can win 0, 1 or 2 points. Figure 1.3 represents winning a game prior to deuce. Note that game to server is represented by *sg* and game to receiver is represented by *rg*.

$$
\begin{array}{ccccccccccccc}
 & & & & & & & & sg & & & & \\
 & & & & & & & \nearrow & & & & & \\
 & & & & & & (40,0) & & & & sg & & \\
 & & & & & \nearrow & & \searrow & & \nearrow & & & \\
 & & & & (30,0) & & & & (40,15) & & & & sg \\
 & & & \nearrow & & \searrow & & \nearrow & & \searrow & & \nearrow & \\
 & & (15,0) & & & & (30,15) & & & & (40,30) & & \\
 & \nearrow & & \searrow & & \nearrow & & \searrow & & \nearrow & & \searrow & \\
(0,0) & & & & (15,15) & & & & (30,30) & & & & (40,40) \\
 & \searrow & & \nearrow & & \searrow & & \nearrow & & \searrow & & \nearrow & \\
 & & (0,15) & & & & (15,30) & & & & (30,40) & & \\
 & & & \searrow & & \nearrow & & \searrow & & \nearrow & & \searrow & \\
 & & & & (0,30) & & & & (15,40) & & & & rg \\
 & & & & & \searrow & & \nearrow & & \searrow & & & \\
 & & & & & & (0,40) & & & & rg & & \\
 & & & & & & & \searrow & & & & & \\
 & & & & & & & & rg & & & &
\end{array}
$$

Figure 1.3: Graphical representation of winning a game prior to deuce

A full game of tennis is represented in figure 1.4 below. Note that if the server wins the next point from deuce the score progresses to advantage server (abbreviated to ad-in) and if the receiver wins the next point from deuce the score progresses to advantage receiver (abbreviated to ad-out).

$$
\begin{array}{ccccccccccccccccc}
 & & & & & & & & sg & & & & & & & & \\
 & & & & & & & \nearrow & & & & & & & & & \\
 & & & & & & (40,0) & & & & sg & & & & & & \\
 & & & & & \nearrow & & \searrow & & \nearrow & & & & & & & \\
 & & & & (30,0) & & & & (40,15) & & & & sg & & & & sg \\
 & & & \nearrow & & \searrow & & \nearrow & & \searrow & & \nearrow & & & & \nearrow & \\
 & & (15,0) & & & & (30,15) & & & & (40,30) & & & & ad-in & & \\
 & \nearrow & & \searrow & & \nearrow & & \searrow & & \nearrow & & \searrow & & \nearrow\swarrow & & & \\
(0,0) & & & & (15,15) & & & & (30,30) & & & & (40,40) & & & & \\
 & \searrow & & \nearrow & & \searrow & & \nearrow & & \searrow & & \nearrow & & \searrow\nwarrow & & & \\
 & & (0,15) & & & & (15,30) & & & & (30,40) & & & & ad-out & & \\
 & & & \searrow & & \nearrow & & \searrow & & \nearrow & & \searrow & & & & \searrow & \\
 & & & & (0,30) & & & & (15,40) & & & & rg & & & & rg \\
 & & & & & \searrow & & \nearrow & & \searrow & & & & & & & \\
 & & & & & & (0,40) & & & & rg & & & & & & \\
 & & & & & & & \searrow & & & & & & & & & \\
 & & & & & & & & rg & & & & & & & &
\end{array}
$$

Figure 1.4: Graphical representation of a game

Suppose we assign a probability of the server winning a point. To keep the model simple, the probability is constant for every point of the game. Since the only two outcomes at each score within the game is either the server or receiver winning a point; the probability of the receiver winning a point is one minus the probability of the server winning a point. Figure 1.5 represents a game where the probability of the server winning a point is represented as p and the probability of the receiver winning a point is represented as $q = 1 - p$.

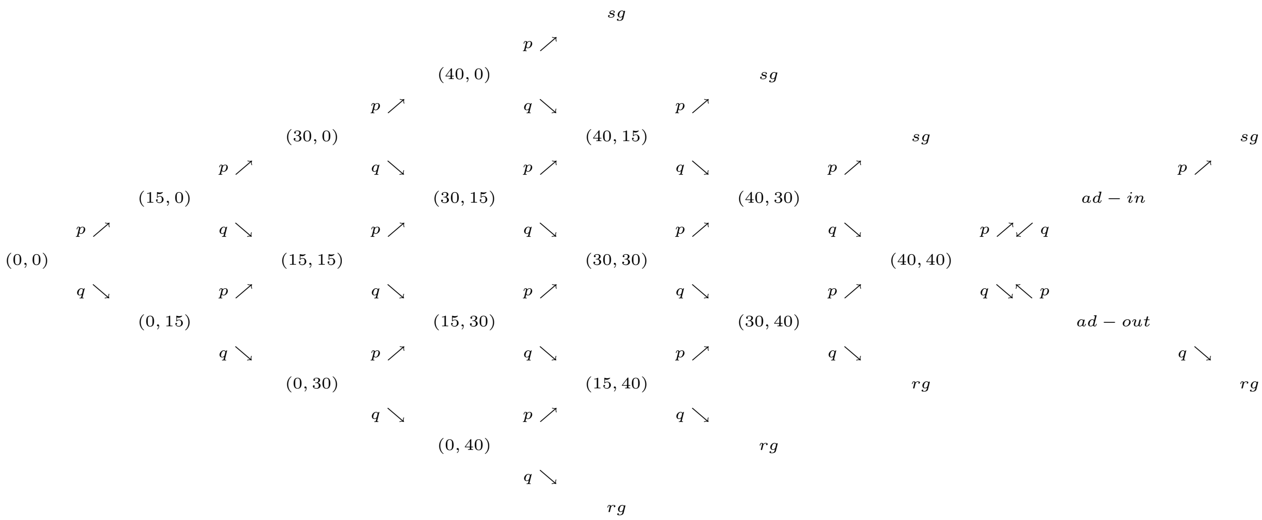


Figure 1.5: Graphical representation of a game where probabilities of winning a point are shown at each possible score within the game

An obvious question: What is the probability of the server or receiver winning the game given the server has a probability p of winning a point? The first observation is since the game can only end with the server winning or the receiver winning, then the two probabilities combined must sum to one.

The server can only win to 0 (known as love) by obtaining four points in a row from (0,0). This occurs with probability p^4 as given by the following path:

$(0, 0) \to^p (15, 0) \to^p (30, 0) \to^p (40, 0) \to^p sg$

The server can win to 15 from the following paths:

$(0, 0) \to^p (15, 0) \to^p (30, 0) \to^p (40, 0) \to^q (40, 15) \to^p sg$

$(0,0) \rightarrow^{p} (15,0) \rightarrow^{p} (30,0) \rightarrow^{q} (30,15) \rightarrow^{p} (40,15) \rightarrow^{p} sg$

$(0,0) \rightarrow^{p} (15,0) \rightarrow^{q} (15,15) \rightarrow^{p} (30,15) \rightarrow^{p} (40,15) \rightarrow^{p} sg$

$(0,0) \rightarrow^{q} (0,15) \rightarrow^{p} (15,15) \rightarrow^{p} (30,15) \rightarrow^{p} (40,15) \rightarrow^{p} sg$

Since there are 4 paths, and the server and receiver win 4 points and 1 point respectively; the server wins the game to 15 with probability $4p^4q$.

Similarly there are 10 paths for the server to win to 30 and this occurs with probability $10p^4q^2$.

Therefore the probability of the server winning the game without reaching deuce occurs with probability $p^4 + 4p^4q + 10p^4q^2 = p^4(1 + 4q + 10q^2)$.

The probability of the score reaching deuce occurs with probability $20p^3q^3$ (since there are 20 paths that can occur in reaching deuce and each player wins exactly 3 points).

The probability of the server winning the game from deuce is obtained as follows:

To win the game from deuce the server needs to win the next two points. This occurs with probability p^2. If the server wins the next point from deuce followed by the receiver winning a point, or the receiver wins the next point from deuce followed by the server winning a point, then the score returns to deuce. This occurs with probability $2pq$. If d is the probability that the server wins the game when the score is at deuce, then $d = p^2 + 2pqd$.

Solving for d gives $d = \frac{p^2}{1-2pq}$.

Using the fact that $p^2 + q^2 = (p+q)^2 - 2pq = 1 - 2pq$, we can also write $d = \frac{p^2}{p^2+q^2}$

Therefore the probability of the server winning the game from the outset (0,0) is given by:
$p^4(1 + 4q + 10q^2) + 20p^3q^3\frac{p^2}{p^2+q^2} = p^4(1 + 4q + \frac{10q^2}{p^2+q^2})$

The simplification arises since $(30, 30)$ and deuce are logically equivalent score lines (as will be proven in section 1.5).

Table 1.1 represents the probabilities of the server and receiver winning a game given the server has a 0.6 probability of winning a point on serve. It shows that the server and receiver have a 0.736 and 0.264 of winning the game respectively.

Win to	Server	Receiver
0	0.130	0.026
15	0.207	0.061
30	0.207	0.092
Deuce	0.191	0.085
Game	0.736	0.264

Table 1.1: The probabilities of the server and receiver winning a game given the server has a 0.6 probability of winning a point on serve

1.3 Binomial theorem

In elementary algebra, the Binomial theorem describes the algebraic expansion of powers of a binomial. According to the theorem, it is possible to expand the powers of a binomial $(p + q)^n$, for $n > 0$: integer.

The Binomial coefficients appear as the entries of Pascal's triangle. The diagram below gives the first 6 rows of Pascal's triangle where the rows are conventionally enumerated starting with row $n = 0$ at the top. The triangle is constructed such that the only element in row 0 is a 1 and the elements in each subsequent row are obtained by adding the number directly above and to the left with the number directly above and to the right. If either the number to the right or left is not present, substitute a zero in its place. For example the coefficients of $(p + q)^6$ are given by the 6th row.

$$\begin{array}{ccccccccccccc} & & & & & & 1 & & & & & & \\ & & & & & 1 & & 1 & & & & & \\ & & & & 1 & & 2 & & 1 & & & & \\ & & & 1 & & 3 & & 3 & & 1 & & & \\ & & 1 & & 4 & & 6 & & 4 & & 1 & & \\ & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\ 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 \end{array}$$

More formally, how many ways can n objects be chosen r at a time is equal to the Binomial coefficient ${}^nC_r = \frac{n!}{r!(n-r)!}$, where $k! = k \times (k-1) \times (k-2) \times ... \times 3 \times 2$ for any integer $k \geq 2$, and $1! = 0! = 1$.

Hence the Binomial theorem is given by:

$(p+q)^n = \sum_{r=0}^{n} {}^nC_r p^r q^{n-r}$

For example: $(p+q)^6$

$=^6 C_0 p^0 q^6 +^6 C_1 p^1 q^5 +^6 C_2 p^2 q^4 +^6 C_3 p^3 q^3 +^6 C_4 p^4 q^2 +^6 C_5 p^5 q^1 +^6 C_6 p^6 q^0$

$= q^6 + 6pq^5 + 15p^2q^4 + 20p^3q^3 + 15p^4q^2 + 6p^5q + p^6$

Notice how the coefficients agree with the 6th row in Pascal's triangle.

A Binomial experiment possesses the following properties:

1. The experiment consists of n identical trials
2. Each trial results in one of two outcomes; either a success S or a failure F
3. The probability of success on a single trial is equal to p and remains the same from trial to trial. The probability of a failure is equal to $q = 1 - p$.
4. The trials are independent.

5. The random variable of interest is R, the number of successes observed during n trials.

The Binomial theorem can be applied to a tennis game since there are two outcomes; either the server wins a point or receiver wins a point, the probability of success on a single trial p of the server winning a point is constant from trial to trial, the probability of a failure is equal to $q = 1 - p$, and the game consists of n identical and independent trials.

Suppose four points are played - then either there are zero successes by the receiver winning all four points and the game, one success by the server winning one point and the receiver winning three points for a score line of (15,40), two successes by the server and receiver winning two points each for a score line of (30,30), three successes by the the server winning three points and the receiver winning one point for a score line of (40,15), or four successes by the server winning all four points and the game. Using the Binomial theorem:

$(p+q)^4 = \sum_{r=0}^{4} {}^4C_r p^r q^{4-r}$
$={}^4 C_0 p^0 q^4 +{}^4 C_1 p^1 q^3 +{}^4 C_2 p^2 q^2 +{}^4 C_3 p^3 q^1 +{}^4 C_4 p^4 q^0$
$= q^4 + 4pq^3 + 6p^2q^2 + 4p^3q + p^4$

Table 1.2 represents the probability outcomes of a game of tennis from the Binomial theorem with $n = 4$. It shows the probability that the server wins the game to 0 is p^4. Alternatively, if the server wins to 0, then the game must last exactly four points and the server must win all four points. This is the situation of the number of ways 4 objects be chosen 4 at a time and occurs in ${}^4C_4 = 1$ way. Note that ${}^4C_4 = 1$ is given by the last entry in the 4th row in Pascal's triangle. Hence the the server wins the game to 0 with probability ${}^4C_4 p^4 q^{4-4} = p^4$. Suppose five points are played - then either there is one success by the server winning one point and the receiver winning four points and the game (which must include the 5th point), two successes by the server winning two points and the receiver winning three points for

Successes	Score line	Probability
0	Game receiver	q^4
1	(15,40)	$4pq^3$
2	(30,30)	$6p^2q^2$
3	(40,15)	$4p^3q$
4	Game server	p^4

Table 1.2: Probability outcomes of a game of tennis from the Binomial theorem with $n = 4$

a score line of (30,40), three successes by the server winning three points and the receiver winning two points for a score line of (40,30), or four successes by the server winning four points and the game (which must include the 5th point) and the receiver winning one point. Note that zero or five successes cannot occur since the server or receiver has already won the game after four consecutive points won. Using the Binomial theorem:

$$({}^5C_1 -{}^4 C_0)p^1q^4 +{}^5 C_2p^2q^3 +{}^5 C_3p^3q^2 + ({}^5C_4 -{}^4 C_4)p^4q^1$$
$$= 4pq^4 + 10p^2q^3 + 10p^3q^2 + 4p^4q$$

Therefore the probability that the server wins the game to 15 is $4p^4q$. Alternatively, if the server wins to 15, then the game must last exactly five points and the server must win four points (which must include the 5th point). This is the situation of the number of ways 5 objects be chosen 4 at a time minus the number of ways 4 objects be chosen 4 at a time (server winning to 0), and occurs in ${}^5C_4$-${}^4C_4$=5-1=4 ways. Hence the server wins the game to 15 with probability $({}^5C_1 -{}^4 C_4)p^4q^{5-4} = 4p^4q$.

If the server wins to 30, then the game must last exactly six points and the server must win 4 points (which must include the 6th point). This is the situation of the number of ways 6 objects be chosen 4 at a time minus the number of ways 5 objects be chosen 4 at a time (server winning to 15) minus the number of ways 4 objects be chosen 4 at a time (server winning to 0) and this occurs in ${}^6C_4$-${}^5C_4$-${}^4C_4$=15-4-1=10 ways. Hence the server wins the game to 30 with probability $({}^6C_4\text{-}{}^5C_4\text{-}{}^4C_4)p^4q^{6-4} = 10p^4q^2$.

For the score to reach deuce, the server must win three of the first six points played and this occurs in ${}^6C_3$=20 ways. Hence the probability of reaching deuce is ${}^6C_3p^3q^{6-3} = 20p^3q^3$.

1.4 Markov Chain theory

A Markov Chain, named for Andrey Markov, is a mathematical system that undergoes transitions from one state to another (from a finite or countable number of possible states) in a chainlike manner. It is a random process characterized as memoryless (i.e. exhibiting the Markov property): the next state depends only on the current state and not on the entire past. Markov chains have many applications as statistical models of real-world processes and can be applied to tennis.

Each score in a game of tennis is a state of the system as represented in table 1.3. Note that (40-30) and ad-in, (30,30) and deuce, and (30,40) and ad-out are equivalent states; which will be proven in section 1.5.

The transition matrix given by table 1.4 represents the probability $P_{i,j}$ that the process will, when in state i, next make a transition into state j. For example the probability $P_{0,1}$ of going from state 0 (score (0,0)) to state 1 (score (15,0)) is p. Similarly, the probability $P_{0,2}$ of going from state 0 (score (0,0)) to state 2 (score (0,15)) is q. Note that $\sum_i P_{i,j} = 1$.

We have already defined the one-step transition probabilities $P_{i,j}$. We now define the n-step transition probabilities $P^n_{i,j}$ to be the probability that a process in state i will be in state j after n additional transitions. The Chapman-Kolomogorov equations provide a method for computing these n-step transition probabilities and is shown that the n-step transition matrix may be obtained by multiplying the matrix by itself n times. Table 1.5 represents the transition matrix $P^2_{i,j}$. For example $P^2_{0,3}$ is the probability of being at score line $(30, 0)$ from score line $(0, 0)$ after two points have been played. From table 1.5 this is given as p^2

State	Score
0	(0,0)
1	(15,0)
2	(0,15)
3	(30,0)
4	(15,15)
5	(0,30)
6	(40,0)
7	(30,15)
8	(15,30)
9	(0,40)
10	(40,15)
11	(15,40)
12	game server
13	game receiver
14	(40,30) or ad-in
15	(30,30) or deuce
16	(30,40) or ad-out

Table 1.3: Scores in a game of tennis represented as states of a system

as expected. Note that $\sum_i P^n_{i,j} = 1$.

Using a four-step transition matrix:

Probability server winning to 0 $=P^4_{(0,12)} = p^4$

Probability receiver winning to 0 $=P^4_{(0,13)} = q^4$

Using a five-step transition matrix:

Probability server winning to 15 $=P^5_{(0,12)} - P^4_{(0,12)} = (4p^4q + p^4) - p^4 = 4p^4q$

Probability receiver winning to 15 $=P^5_{(0,13)} - P^4_{(0,13)} = (4q^4p + q^4) - q^4 = 4q^4p$

Using a six-step transition matrix:

Probability server winning to 30 $= P^6_{(0,12)} - (P^5_{(0,12)} - P^4_{(0,12)}) - P^4_{(0,12)} = P^6_{(0,12)} - P^5_{(0,12)} =$ $(10p^4q^2 + 4p^4q + p^4) - (4p^4q + p^4) = 10p^4q^2$

Probability receiver winning to 30 $= P^6_{(0,13)} - (P^5_{(0,13)} - P^4_{(0,13)}) - P^4_{(0,13)} = P^6_{(0,13)} - P^5_{(0,13)} =$

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	0	p	q	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	p	q	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	p	q	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	p	q	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	p	q	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	p	q	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	q	0	p	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	p	0	0	0	0	q	0
8	0	0	0	0	0	0	0	0	0	0	0	q	0	0	0	p	0
9	0	0	0	0	0	0	0	0	0	0	0	p	0	q	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	p	0	q	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	q	0	0	p
12	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	p	0	0	q	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	p	0	q
16	0	0	0	0	0	0	0	0	0	0	0	0	0	q	0	p	0

Table 1.4: Transition matrix for a game of tennis

$(10q^4p^2 + 4q^4p + q^4) - (4q^4p + q^4) = 10q^4p^2$

Probability reaching deuce $= P^6_{(0,15)} = 20p^3q^3$

Note the extensive matrix algebra computation to obtain these probabilities of winning prior to deuce. When the Markov Chain process is a Binomial experiment (as was demonstrated in section 1.3 for a game of tennis) the problem can be simplified by recursion formulas; which conveniently can be applied to spreadsheets (such as Excel) to obtain numerical results. This approach will be adopted for winning a game in section 1.5 using backward recursion and section 1.6 using forward recursion. Hence recursion formulas are developed in the remaining chapters for winning and the duration of a match. Note the probability of winning from deuce using Markov Chain theory is outlined in Kemeny and Snell[1].

[1]Kemeny, J. and Snell, J. Finite Markov Chains. Van Nostrand, Princeton, NJ. 1976, p163

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0	0	0	0	p^2	$2pq$	q^2	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	p^2	$2pq$	q^2	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	p^2	$2pq$	q^2	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	$2pq$	0	p^2	0	0	q^2	0
4	0	0	0	0	0	0	0	0	0	0	p^2	q^2	0	0	0	$2pq$	0
5	0	0	0	0	0	0	0	0	0	0	0	$2pq$	0	q^2	0	p^2	0
6	0	0	0	0	0	0	0	0	0	0	0	0	$pq+p$	0	q^2	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	p^2	0	$2pq$	0	q^2
8	0	0	0	0	0	0	0	0	0	0	0	0	0	q^2	p^2	0	$2pq$
9	0	0	0	0	0	0	0	0	0	0	0	0	0	$pq+q$	0	0	p^2
10	0	0	0	0	0	0	0	0	0	0	0	0	$pq+p$	0	0	q^2	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	$pq+q$	0	p^2	0
12	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	p	0	pq	0	q^2
15	0	0	0	0	0	0	0	0	0	0	0	0	p^2	q^2	0	$2pq$	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	q	p^2	0	pq

Table 1.5: Transition matrix P_{ij}^2 for a game of tennis

1.5 Backward recursion

The calculations presented in prior sections to obtain the probability of the server winning the game are all based on a path moving forwards i.e. starting from the value of $(0, 0)$, calculating the probabilities of winning to 0, 15, 30; the probability of reaching deuce; and lastly the probability of winning from deuce. This approach not only calculates the probability of the server winning the game, but also obtains probabilities of the server (and receiver) winning to a particular point score. For example the probability that the server wins the game to 15 was shown to be $4p^4q$. This can easily be extended to obtain the probability that the server (or receiver) wins the game after the first (or the second, third,...) deuce has been played. These calculations are given numerically using a forward recursion approach in section 1.6.

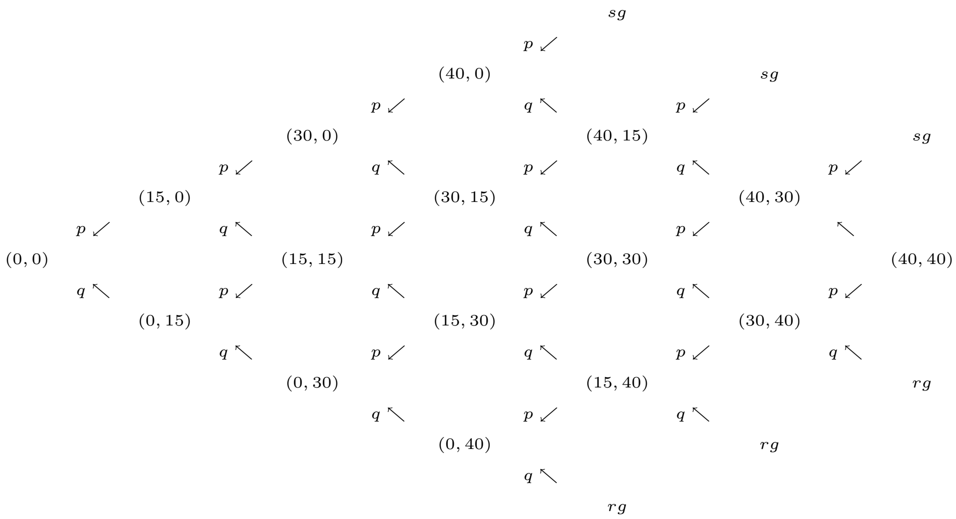


Figure 1.6: Graphical representation of a 'backward-path' approach to a game

Consider the diagram given in figure 1.6 which gives a graphical representation of a game using a 'backward-path' (in contrast to a 'forward-path' given by figure 1.5). The starting value is at deuce; where an explicit formula was obtained in section 1.2 for the probability of the server winning from deuce such that $d = \frac{p^2}{p^2+q^2}$.

Let $P(a, b)$ represent the probability that the server wins the game when the score is (a, b), where a is the server's score and b is the receiver's score. With simple logic, the probability of the server winning from (40,30) is then $p * 1 + q * P(40, 40) = p + qd$, since if the server wins the next point from (40,30) the server wins the game and if the receiver wins the next point from (40,30) the score is at deuce.

Similarly, the probability of the server winning from (30,40) is given by $p * P(40, 40) + q * 0 = pd$.

Likewise, the probability of the server winning from (30,30) is given by $p * P(40, 30) + q * P(30, 40)$. This backward recursive process could continue for the entire game to obtain the probability of the server winning from $(0, 0)$, and hence winning the game from the outset. Backward recursion simplifies calculations to obtain the probability of the server winning the game (compared to 'forward-path' approaches). However 'forward-path' approaches and in particular the forward recursion process presented in chapter 1.6 provides additional information on not just the probability of the server winning the game but also the probabilities of the server (and receiver) winning to a particular point score. Hence both backward and forward recursion methods will be used extensively throughout the remaining chapters depending on the context.

It is necessary in a tennis match to distinguish between which player is serving, since rotation of server occurs after each game. Therefore, the notation of a game will extend slightly to represent which player is currently serving.

More formally, we have two players; player A and player B

Let p_A represent a constant probability of player A winning a point on serve
Let p_B represent a constant probability of player B winning a point on serve

p_A and p_B become the two parameters for the model.

Let q_A represent a constant probability of player A losing a point on serve
Let q_B represent a constant probability of player B losing a point on serve

It follows that $q_A = 1 - p_A$ and $q_B = 1 - p_B$

Firstly, we set up a Markov Chain model of a single game for player A serving where the state of the game is the current score in points (thus $(40, 30)$ is $(3, 2)$).

With probability p_A the state changes from (a, b) to $(a + 1, b)$, and with probability $q_A = 1 - p_A$ the state changes from (a, b) to $(a, b + 1)$, where a and b represent the current point score for player A and player B respectively.

Let $P_A(a,b)$ represent the probability that player A wins the game on serve when the score is (a,b)

The backward recursion formula becomes:

$P_A(a,b) = p_A P_A(a+1,b) + q_A P_A(a,b+1)$

The boundary values are:

$P_A(a,b) = 1$, if $a = 4$ and $b \leq 2$

$P_A(a,b) = 0$, if $b = 4$ and $a \leq 2$

The boundary values and recursion formula can be entered on a simple spreadsheet (such as Excel). The problem of deuce can be handled in two ways. Since deuce is logically equivalent to (30-30), a formula for this can be entered in the deuce cell. This creates a circular cell reference, but the iterative function of Excel can be turned on, and Excel will iterate to a solution. In preference, an explicit formula was obtained in section 1.2, such that $P_A(3,3) = \frac{p_A^2}{p_A^2+q_A^2}$. An alternate way of calculating this explicit formula is by recognizing that the probability of winning from deuce is in the form of a geometric series

$P_A(3,3) = p_A^2 + p_A^2 2p_Aq_A + p_A^2(2p_Aq_A)^2 + p_A^2(2p_Aq_A)^3 + ...$

where the first term is p_A^2 and the common ratio is $2p_Aq_A$

The sum is given by $\frac{p_A^2}{1-2p_Aq_A}$ provided that $-1 < 2p_Aq_A < 1$. We know that $0 < 2p_Aq_A \leq 1/2$, for $0 < p_A < 1$ with $q_A = 1 - p_A$. Let $r(p_A) = 2p_Aq_A = 2p_A(1-p_A)$. Then $r(0) = r(1) = 0$ and $r(1/2) = 1/2$ is maximum in $0 < p_A < 1$, since $\frac{dr}{dp_A} = 0$ and $\frac{d^2r}{dp_A^2} < 0$ when $r = 1/2$.

Therefore the probability of winning from deuce is $\frac{p_A^2}{1-2p_Aq_A}$ and as obtained in section 1.2, this can be expressed as:

$P_A(3,3) = \frac{p_A^2}{p_A^2+q_A^2}$

A similar backward recursion formula with boundary conditions when player B is serving are as follows:

Let p_B represent a constant probability of player B winning a point on serve

With probability p_B the state changes from (a, b) to $(a, b+1)$, and with probability $q_B = 1 - p_B$ the state changes from (a, b) to $(a+1, b)$, where a and b represent the current point score for player A and player B respectively.

Let $P_B(a, b)$ represent the probability that player A wins the game when the score is (a, b) given player B is serving.

The backward recursion formula becomes:

$P_B(a, b) = q_B P_B(a+1, b) + p_B P_B(a, b+1)$

The boundary values are:

$P_B(a, b) = 1$, if $a = 4$ and $b \leq 2$

$P_B(a, b) = 0$, if $b = 4$ and $a \leq 2$

$P_B(3, 3) = \frac{q_B^2}{p_B^2 + q_B^2}$

Excel spreadsheet code to obtain the conditional probabilities of player A winning a game on serve is as follows:

Enter the text p_A in cell D1.

Enter the text q_A in cell D2

Enter 0.6 in cell E1

Enter =1-E1 in cell E2

Enter 1 in cells C11, D11 and E11

Enter 0 in cells G7, G8 and G9

Enter = E1ˆ2/(E1ˆ2+E2ˆ2) in cell F10

Enter =$E$1*C8+$E$2*D7 in cell C7

Copy and Paste cell C7 in cells D7, E7, F7, C8, D8, E8, F8, C9, D9, E9, F9, C10, D10 and E10

Notice the absolute and relative referencing used in the formula =$E$1*C8+$E$2*D7. By setting up an equation in this recursive format, the remaining conditional probabilities can easily and quickly be obtained by copying and pasting.

Table 1.6 represents the conditional probabilities of player A winning the game from various score lines for $p_A = 0.6$. It indicates that a player with a 60% chance of winning a point has a 73.6% chance of winning the game. Note that since advantage server is logically equivalent to 40-30, and advantage receiver is logically equivalent to 30-40, the required statistics can be found from these cells. Also worth noting is that the chances of winning from deuce and 30-30 are the same. The following theorems will prove these results to illustrate how arguments can be sustained about the probabilities of winning a game.

		B score				
		0	15	30	40	game
A score	0	0.736	0.576	0.369	0.150	0
	15	0.842	0.714	0.515	0.249	0
	30	0.927	0.847	0.692	0.415	0
	40	0.980	0.951	0.877	0.692	
	game	1	1	1		

Table 1.6: The conditional probabilities of player A winning the game on serve from various score lines for $p_A = 0.6$

Theorem 1.5.1. *A player has the same probability of winning a game from advantage server as they do from 40-30.*

Proof. In both cases, if the server wins the next point they win the game and if they lose the next point the score is back at deuce. $\square$

Theorem 1.5.2. *A player has the same probability of winning a game from advantage receiver as they do from 30-40.*

Proof. In both cases, if the server loses the next point they lose the game and if they win the next point the score is back at deuce. □

Theorem 1.5.3. *A player has the same probability of winning a game from deuce as they do from 30-30.*

Proof. At 30-30 if the server wins the next point the score goes to 40-30. At deuce if the server wins the next point the score goes to advantage server. From Theorem 1.5.1 advantage server is equivalent to 40-30. At 30-30 if the server loses the next point the score goes to 30-40. At deuce if the server loses the next point the score goes to advantage receiver. From Theorem 1.5.2 advantage receiver is equivalent to 30-40. □

1.6 Forward recursion

In section 1.5, backward recursion formulas were developed to obtain the probabilities of winning a game conditional on the point score. In section 1.6, spreadsheet calculations using forward recursion formulas are obtained for the probabilities of reaching a point score in a game (i.e. winning the game to 0). These calculations on reaching score lines can also be used to obtain probabilities of winning a game from the outset. The basic idea is that we start at the beginning of the game with an initial value and see what happens as the game progresses.

Let $N_A(g, h)$ be the probability of reaching a point score (g, h) in a game from the outset for player A serving, where g and h represent the projected point score for player A and player B respectively.

The initial value is $N_A(0, 0) = 1$

The forward recursion formulas are:

$N_A(g,h) = p_A N_A(g-1,h)$, if $g = 4$ and $0 \leq h \leq 2$; $h = 0$ and $1 \leq g \leq 4$; $g \geq 3$, $h \geq 3$ and $g = h+1$; $g \geq 3$, $h \geq 3$ and $g = h+2$

$N_A(g,h) = q_A N_A(g,h-1)$, if $h = 4$ and $0 \leq g \leq 2$; $g = 0$ and $1 \leq h \leq 4$; $g \geq 3$, $h \geq 3$ and $h = g+1$; $g \geq 3$, $h \geq 3$ and $h = g+2$

$N_A(g,h) = p_A N_A(g-1,h) + q_A N_A(g,h-1)$, if $1 \leq g \leq 3$ and $1 \leq h \leq 3$; $g \geq 4$, $h \geq 4$ and $g = h$

Table 1.7 lists the probability of reaching various score lines in a game from the outset with $p_A = 0.6$. It indicates that the probability of reaching deuce in such a game is 0.276.

		B score				
		0	15	30	40	game
	0	1	0.400	0.160	0.064	0.026
	15	0.600	0.480	0.288	0.154	0.061
A score	30	0.360	0.432	0.346	0.230	0.092
	40	0.216	0.346	0.346	0.276	
	game	0.130	0.207	0.207		

Table 1.7: The probability of reaching various score lines in a game from the outset with $p_A = 0.6$

Similar formulas can be obtained for when player B is serving such that $N_B(g,h)$ represents the probability of reaching a point score (g,h) in a game from the outset for player B serving, where g and h represent the projected point score for player A and player B respectively.

The probability of player A winning the game on serve from the outset can be obtained from:

$N_A(4,0) + N_A(4,1) + N_A(4,2) + N_A(3,3)P_A(3,3) = 0.130 + 0.207 + 0.207 + 0.276 * 0.692 = 0.736$ when $p_A = 0.6$.

Note that we have used here the value of $P_A(3,3) = 0.692$ obtained from a formula given in section 1.2.

After the point score of $(3,3)$ has been reached the recursive formulas tell us that

$N_A(4,3) = p_A N_A(3,3)$,

$N_A(3,4) = q_A N_A(3,3)$, and

$N_A(4,4) = p_A N_A(3,4) + q_A N_A(4,3)$.

These results can be summarized as

$N_A(4,4) = 2p_A q_A N_A(3,3)$.

This argument can be continued to establish that

$N_A(n,n) = (2p_A q_A)^{n-3} N_A(3,3)$ for all $n \geq 3$.

Now to win a game after any deuce either player must win two consecutive points. When player A is serving this occurs with probability $p_A^2 + q_A^2$. Thus the completion of the game after the first deuce whilst serving can be expressed in two equivalent ways:

$N_A(3,3) = (p_A^2 + q_A^2)(N_A(3,3) + N_A(4,4) + N_A(5,5) + ...)$

$= (p_A^2 + q_A^2) N_A(3,3)(1 + 2p_A q_A + (2p_A q_A)^2 + ...)$

If we substitute $r = 2p_A q_A$, then $0 < r \leq 1/2$ when $0 < p_A < 1$ and $q_A = 1 - p_A$ whilst $p_A^2 + q_A^2 = 1 - r$. By cancelling the term $N_A(3,3)$, which is not zero, we find that $1 = (1-r)(1 + r + r^2 + ...)$

This result can be expressed as the sum of an infinite geometric series by the following theorem:

Theorem 1.6.1. $1 + r + r^2 + ... = \frac{1}{1-r}$ *for* $0 < r \leq 1/2$.

Chapter 2

Winning a match: backward recursion

2.1 Introduction

Chapter 2 extends on chapter 1 (in section 1.5) by applying backward recursion to calculate the probability of winning a tiebreak game from any point score within the game, winning a tiebreak and advantage set from any point and game score within the set; and winning an all tiebreak set and final set advantage match from any point, game and set score within the match.

To extend the results of chapter 1 to so-called tiebreak games and beyond, we introduce some additional notation - after first noting that:

A tennis match consists of four levels - (points, games, sets, match). Games can be standard games (as presented in chapter 1) or tiebreak games, sets can be advantage or tiebreak, and matches can be the best-of-3 all tiebreak sets, best-of-3 final set advantage, best-of-5 all tiebreak sets or the best-of-5 final set advantage. In some circumstances we may be referring to points in a standard or tiebreak game and other circumstances points in a tiebreak or advantage set. Therefore, it becomes necessary to represent:
points in a game as pg,

points in a tiebreak game as pg_T,

points in an advantage set as ps,

points in a tiebreak set as ps_T,

points in a best-of-3 all tiebreak set match as pm_{3T},

points in a best-of-3 final set advantage match as pm_3,

points in a best-of-5 all tiebreak set match as pm_{5T},

points in a best-of-5 final set advantage match as pm_5,

games in an advantage set as gs,

games in a tiebreak set as gs_T,

sets in a best-of-3 all tiebreak set match as sm_{3T},

sets in a best-of-3 final set advantage match as sm_3.

sets in a best-of-5 all tiebreak set match as sm_{5T},

and sets in a best-of-5 final set advantage match as sm_5.

Using this additional notation:

Let $P_A^{pg}(a,b)$ and $P_B^{pg}(a,b)$ represent the probabilities that player A wins a game when the score is (a,b) given player A and player B are serving respectively.

At this stage the following notation on a game is introduced which is used throughout the book where applicable:

Let p_A^g and p_B^g represent the probabilities of player A and player B winning a game on serve from the outset.

Let q_A^g and q_B^g represent the probabilities of player A and player B losing a game on serve from the outset.

It follows that $p_A^g = P_A^{pg}(0,0)$, $p_B^g = 1 - P_B^{pg}(0,0)$, $q_A^g = 1 - p_A^g$ and $q_B^g = 1 - p_B^g$.

2.2 Winning a tiebreak game

The tiebreak game differs from a standard game, especially with its feature that the server changes several times within the game. The scoring structure within a tiebreak game of tennis is defined as follows: The first player to reach 7 points and be ahead by at least 2 points wins the game. If the point score reaches 6 points-all, then the game continues indefinitely until one player is two points ahead, and wins the game. One player serves the first point, and then the players alternate serving every two points.

Since the chance of a player winning a tiebreak game depends on who is serving, two interconnected sheets are required within the spreadsheet, one for when player A is serving and one for when player B is serving. The equations that follow for modelling a tiebreak game, tiebreak set and advantage set throughout the book are those for player A serving in the game. Similar formulas can be produced for player B serving in the game.

Let $P_A^{pg_T}(a,b)$ and $P_B^{pg_T}(a,b)$ represent the conditional probabilities of player A winning a tiebreak game from point score (a,b) given player A and player B are serving respectively, where a is the point score for player A and b is the point score for player B.

Recurrence formulas:
$P_A^{pg_T}(a,b) = p_A P_B^{pg_T}(a+1,b) + q_A P_B^{pg_T}(a,b+1)$, if $(a+b)$ is even
$P_A^{pg_T}(a,b) = p_A P_A^{pg_T}(a+1,b) + q_A P_A^{pg_T}(a,b+1)$, if $(a+b)$ is odd

Boundary values:
$P_A^{pg_T}(a,b) = 1$, if $a = 7$ and $0 \leq b \leq 5$
$P_A^{pg_T}(a,b) = 0$, if $b = 7$ and $0 \leq a \leq 5$

The formula for the probability of player A winning the tiebreak game from a score $(6,6)$ is calculated from the equation $P_A^{pg_T}(6,6) = p_A q_B + P_A^{pg_T}(6,6)(p_A p_B + q_A q_B)$, which simplifies to:

$P_A^{pg_T}(6,6) = \frac{p_A q_B}{p_A q_B + q_A p_B}$

It may be noted that we have used a short-cut in the argument used to develop this relationship. It is reasonably obvious that the probability of player A winning when serving at the point score $(6,6)$ is the same as when he is serving at the point score $(8,8)$. It is less obvious that this is same as the probability of player A winning when player B is serving at the point score of $(7,7)$. More detailed comment on this issue follows Theorem 2.2.1 below.

Table 2.1 represents the conditional probabilities of player A winning the tiebreak game on serve from various score lines for $p_A = 0.62$ and $p_B = 0.60$. Table 2.2 is represented similarly with player A winning a tiebreak game from various score lines given player B is serving. It indicates that player A has a 0.533 probability of winning the tiebreak game from the outset for player A or player B serving. Note how the calculations are obtained by the interconnection of both sheets. For example $P_A^{pg_T}(0,0) = p_A P_B^{pg_T}(1,0) + q_A P_B^{pg_T}(0,1) = 0.62 * 0.620 + 0.38 * 0.389 = 0.533$. We now state some theorems on probabilities of winning a tiebreak game.

Theorem 2.2.1. *A player has the same probability of winning a tiebreak game from all points (n,n), $n \geq 5$.*

Proof. From (n,n), $n \geq 5$, a player always has to win the next two points to win the game, and one of the two points is on his own serve and the other point is on his opponent's serve. □

Theorem 2.2.2. *If player A is serving, he has the same probability of winning a tiebreak game from all points $(n+1,n)$, $n \geq 5$.*

Proof. If the server A wins the next point from $(n+1,n)$, $n \geq 5$, he wins the game. If the server A loses the next point from $(n+1,n)$, $n \geq 5$, the score is $(n+1,n+1)$. From

		B score							
		0	1	2	3	4	5	6	7
	0	0.533	0.441	0.295	0.205	0.096	0.044	0.008	0
	1	0.670	0.530	0.431	0.271	0.174	0.066	0.020	0
	2	0.755	0.680	0.528	0.417	0.240	0.134	0.032	0
A score	3	0.868	0.773	0.695	0.526	0.399	0.197	0.080	0
	4	0.926	0.892	0.798	0.716	0.523	0.372	0.129	0
	5	0.977	0.949	0.921	0.834	0.750	0.521	0.323	0
	6	0.994	0.991	0.975	0.959	0.891	0.818	0.521	
	7	1	1	1	1	1	1		

Table 2.1: The conditional probabilities of player A winning the tiebreak game on serve from various score lines for $p_A = 0.62$ and $p_B = 0.60$

		B score							
		0	1	2	3	4	5	6	7
	0	0.533	0.389	0.295	0.166	0.096	0.031	0.008	0
	1	0.620	0.530	0.374	0.271	0.135	0.066	0.013	0
	2	0.755	0.626	0.528	0.355	0.240	0.098	0.032	0
A score	3	0.834	0.773	0.635	0.526	0.328	0.197	0.052	0
	4	0.926	0.858	0.798	0.647	0.523	0.286	0.129	0
	5	0.967	0.949	0.890	0.834	0.669	0.521	0.208	0
	6	0.994	0.985	0.975	0.934	0.891	0.713	0.521	
	7	1	1	1	1	1	1		

Table 2.2: The conditional probabilities of player A winning the tiebreak game from various score lines given player B is serving for $p_A = 0.62$ and $p_B = 0.60$

Theorem 2.2.1, a player has the same probability of winning a tiebreaker game from all points (n, n), $n \geq 5$, or equivalently $(n + 1, n + 1)$, $n \geq 4$. $\square$

Theorem 2.2.3. *If player A is serving, he has the same probability of winning a tiebreak from all points $(n, n + 1)$, $n \geq 5$.*

Proof. The proof is obtained similarly to Theorem 2.2.2 $\square$

Let p_A^{gT} and p_B^{gT} represent the probabilities of player A and player B respectively winning a tiebreak game on serve from the outset.

It follows that $p_A^{gT} = P_A^{pgT}(0, 0)$ and $p_B^{gT} = 1 - P_B^{pgT}(0, 0)$.

Theorem 2.2.4. *There is no advantage in serving first in a tiebreak game. That is:*

$p_A^{g_T} = 1 - p_B^{g_T}$

Proof. The proof can be obtained combinatorially, by first observing that $P_A^{pg_T}(6,6) = P_B^{pg_T}(6,6)$, and more generally $P_A^{pg_T}(a,b) = P_B^{pg_T}(a,b)$, for $(a+b)$ even $\square$

Based on Theorem 2.2.4 it becomes convenient to let p^{g_T} and $q^{g_T} = 1 - p^{g_T}$ represent the respective probabilities of player A and player B winning a tiebreak game from the outset.

2.3 Winning a tiebreak set

The scoring structure of a tiebreak set of tennis is defined as follows. The first player to reach 6 standard games and be ahead by at least 2 standard games wins the set. If the game score reaches 6 games-all, then a tiebreak game is played to decide the set. Players alternate service each game. At 6 games-all, the player receiving in the prior game, serves the first point of the tiebreak game.

Let $P_A^{gs_T}(c,d)$ and $P_B^{gs_T}(c,d)$ represent the probabilities of player A winning a tiebreak set from game score (c,d) given player A and player B are serving respectively, where c is the game score for player A and d is the game score for player B.

Recurrence formula:

$P_A^{gs_T}(c,d) = p_A^g P_B^{gs_T}(c+1,d) + q_A^g P_B^{gs_T}(c,d+1)$

Boundary values:

$P_A^{gs_T}(c,d) = 1$, if $c = 6$ and $0 \leq d \leq 4$; $(7,5)$

$P_A^{gs_T}(c,d) = 0$, if $d = 6$ and $0 \leq c \leq 4$; $(5,7)$

$P_A^{gs_T}(6,6) = p^{g_T}$

Tables 2.3 and 2.4 show the probabilities of player A winning the tiebreak set, given $p_A = 0.62$ and $p_B = 0.60$. It indicates that player A has a 0.568 probability of winning the set from the outset for player A or player B serving. Notice how the cells $p_A^g = P_A^{pg}(0,0)$ and $q_A^g = 1 - P_A^{pg}(0,0)$, which represent the probability of player A winning and losing a game when serving respectively, is used in the recurrence formula for a tiebreak set.

		B score							
		0	1	2	3	4	5	6	7
	0	0.568	0.497	0.269	0.191	0.052	0.018	0	
	1	0.766	0.563	0.484	0.231	0.146	0.023	0	
	2	0.824	0.781	0.557	0.469	0.182	0.086	0	
A score	3	0.944	0.846	0.804	0.552	0.448	0.111	0	
	4	0.972	0.963	0.877	0.839	0.546	0.419	0	
	5	0.997	0.988	0.983	0.924	0.897	0.541	0.413	0
	6	1	1	1	1	1	0.895	0.533	
	7						1		

Table 2.3: The conditional probabilities of player A winning the tiebreak set on serve from various score lines for $p_A = 0.62$ and $p_B = 0.60$

		B score							
		0	1	2	3	4	5	6	7
	0	0.568	0.346	0.269	0.099	0.052	0.006	0	
	1	0.632	0.563	0.317	0.231	0.065	0.023	0	
	2	0.824	0.634	0.557	0.279	0.182	0.029	0	
A score	3	0.879	0.846	0.638	0.552	0.226	0.111	0	
	4	0.972	0.906	0.877	0.646	0.546	0.143	0	
	5	0.991	0.988	0.944	0.924	0.662	0.541	0.141	0
	6	1	1	1	1	1	0.656	0.533	
	7						1		

Table 2.4: The conditional probabilities of player A winning the tiebreak set from various score lines for $p_A = 0.62$ and $p_B = 0.60$ given player B is serving

Let p_A^{sT} and p_B^{sT} represent the probabilities of player A and player B respectively winning a tiebreak set on serve from the outset.

It follows that $p_A^{s_T} = P_A^{gs_T}(0,0)$ and $p_B^{s_T} = 1 - P_B^{gs_T}(0,0)$.

Theorem 2.3.1. *There is no advantage in serving first in a tiebreak set. That is:*

$p_A^{s_T} = 1 - p_B^{s_T}$

Proof. The proof can be obtained similar to the proof given in Theorem 2.2.4 $\square$

Based on Theorem 2.3.1 it becomes convenient to let p^{s_T} and $q^{s_T} = 1 - p^{s_T}$ represent the respective probabilities of player A and player B winning a tiebreak set from the outset.

Using the formulas given for a game and a tiebreak game conditional on the point score and a tiebreak set conditional on the game score, calculations are now obtained for a tiebreak set conditional on both the point and game score as follows.

Let $P_A^{ps_T}(a, b : c, d)$ represent the probability of player A winning a tiebreak set on serve from (c, d) in games and (a, b) in points. This can be calculated by:

$P_A^{ps_T}(a,b:c,d) = P_A^{pg}(a,b)P_B^{gs_T}(c+1,d) + (1 - P_A^{pg}(a,b))P_B^{gs_T}(c,d+1)$, if $(c,d) \neq (6,6)$
$P_A^{ps_T}(a,b:c,d) = P_A^{pg_T}(a,b)$, if $(c,d) = (6,6)$

2.4 Winning an advantage set

The scoring structure of an advantage set of tennis is defined as follows. The first player to reach 6 standard games and be ahead by at least 2 standard games wins the set. If the set score reaches 5 games-all, then the set continues indefinitely until one player is two games ahead, and wins the set. Players alternate service each game.

Let $P_A^{gs}(c, d)$ and $P_B^{gs}(c, d)$ represent the probabilities of player A winning an advantage set from game score (c, d) given player A and player B are serving respectively, where c is the game score for player A and d is the game score for player B.

Recurrence formula:

$P_A^{gs}(c,d) = p_A^g P_B^{gs}(c+1,d) + q_A^g P_B^{gs}(c,d+1)$

Boundary values:

$P_A^{gs}(c,d) = 1$, if $c = 6$ and $0 \leq d \leq 4$

$P_A^{gs}(c,d) = 0$, if $d = 6$ and $0 \leq c \leq 4$

The formula for the probability of player A winning the advantage set from a game score of $(5,5)$ is calculated from the equation $P_A^{gs}(5,5) = p_A^g q_B^g + P_A^{gs}(6,6)(p_A^g p_B^g + q_A^g q_B^g)$, which simplifies, using $P_A^{gs}(5,5) = P_A^{gs}(6,6)$ to:

$P_A^{gs}(5,5) = \frac{p_A^g q_B^g}{p_A^g q_B^g + q_A^g p_B^g}$

Tables 2.5 and 2.6 show the conditional probabilities of player A winning the advantage set, given $p_A = 0.62$ and $p_B = 0.60$. It indicates that player A has a 0.572 probability of winning the set from the outset for player A or player B serving.

		B score						
		0	1	2	3	4	5	6
	0	0.572	0.501	0.272	0.194	0.053	0.018	0
	1	0.769	0.567	0.489	0.235	0.149	0.023	0
	2	0.827	0.785	0.563	0.474	0.185	0.088	0
A score	3	0.945	0.849	0.808	0.558	0.455	0.114	0
	4	0.973	0.964	0.880	0.843	0.554	0.430	0
	5	0.997	0.988	0.984	0.926	0.900	0.554	
	6	1	1	1	1	1		

Table 2.5: The conditional probabilities of player A winning the advantage set from various score lines for $p_A = 0.62$ and $p_B = 0.60$, and player A serving

Theorem 2.4.1. *A player has the same probability of winning an advantage set from all games (n,n), $n \geq 4$. If player A is serving, he has the same probability of winning an advantage set from all games $(n+1,n)$, $n \geq 4$. If player A is serving, he has the same probability of winning an advantage set from all games $(n,n+1)$, $n \geq 4$.*

		B score						
		0	1	2	3	4	5	6
	0	0.572	0.350	0.272	0.101	0.053	0.006	0
	1	0.636	0.567	0.322	0.235	0.066	0.023	0
	2	0.827	0.638	0.563	0.284	0.185	0.030	0
A score	3	0.882	0.849	0.643	0.558	0.230	0.114	0
	4	0.973	0.909	0.880	0.653	0.554	0.146	0
	5	0.991	0.988	0.946	0.926	0.672	0.554	
	6	1	1	1	1	1		

Table 2.6: The conditional probabilities of player A winning the advantage set from various score lines for $p_A = 0.62$ and $p_B = 0.60$, and player B serving

Proof. The proof is obtained similar to the proofs given in Theorems 2.2.1 and 2.2.2 □

Let p^s_A and p^s_B represent the probabilities of player A and player B respectively winning an advantage set on serve from the outset.

It follows that $p^s_A = P^{gs}_A(0,0)$ and $p^s_B = 1 - P^{gs}_B(0,0)$.

Theorem 2.4.2. *There is no advantage in serving first in an advantage set. That is:*

$p^s_A = 1 - p^s_B$

Proof. The proof can be obtained similar to the proof given in Theorem 2.2.4 □

Based on Theorem 2.4.2 it becomes convenient to let p^s and $q^s = 1 - p^s$ represent the respective probabilities of player A and player B winning an advantage set from the outset.

Theorem 2.4.3. *If $p_A > p_B$ then $p^s > p^{s_T}$ and if $p_A < p_B$ then $p^s < p^{s_T}$*

Proof. The proof can be obtained combinatorially, by first observing that if $p_A > p_B$ then $P^{gs}_A(5,5) > p^{g_T}$ and if $p_A < p_B$ then $P^{gs}_A(5,5) < p^{g_T}$ □

Let $P_A^{ps}(a, b : c, d)$ represent the probability of player A winning an advantage set from (c, d) in games, (a, b) in points and player A serving in the set. This can be calculated by:

$$P_A^{ps}(a, b : c, d) = P_A^{pg}(a, b)P_B^{gs}(c + 1, d) + (1 - P_A^{pg}(a, b))P_B^{gs}(c, d + 1)$$

2.5 Winning an all tiebreak set match

The scoring structure for an all tiebreak set match of tennis is defined as follows. For a best-of-3 all tiebreak set match, the first player to reach 2 tiebreak sets wins the match. For a best-of-5 all tiebreak set match, the first player to reach 3 tiebreak sets wins the match. Usually the toss of a coin decides who will be serving the first game of the match. The server for the first game in the other sets will be the player who was receiving the last game in the prior set. If a set finishes with a tiebreak game, then the player that served first in that set, will be receiving for the first game in the next set.

Let $P_A^{sm_{3T}}(e, f)$ and $P_B^{sm_{3T}}(e, f)$ represent the probabilities of player A winning a best-of-3 all tiebreak set match from set score (e, f) given player A and player B are serving respectively, where e is the set score for player A and f is the set score for player B.

Similarly, let $P_A^{sm_{5T}}(e, f)$ and $P_B^{sm_{5T}}(e, f)$ represent the probabilities of player A winning a best-of-5 all tiebreak set match from set score (e, f) given player A and player B are serving respectively, where e is the set score for player A and f is the set score for player B.

Corollary 2.5.1. *$P_A^{sm_{3T}}(e, f) = P_B^{sm_{3T}}(e, f)$ and $P_A^{sm_{5T}}(e, f) = P_B^{sm_{5T}}(e, f)$, for all (e, f)*

Proof. This follows from Theorem 2.3.1 since there is no advantage in serving first in a tiebreak set $\square$

Based on Corollary 2.5.1 it becomes convenient to let $P^{sm_{3T}}(e, f)$ and $P^{sm_{5T}}(e, f)$ represent the probabilities of player A winning a best-of-3 and best-of-5 all tiebreak set match from set score (e, f) respectively and let $p^{m_{3T}}$ and $p^{m_{5T}}$ represent the probabilities of player A winning a best-of-3 and best-of-5 all tiebreak set match respectively from the outset.

It follows that $p^{m_{3T}} = P^{sm_{3T}}(0, 0)$ and $p^{m_{5T}} = P^{sm_{5T}}(0, 0)$.

Formulas are now given for a best-of-5 and best-of-3 all tiebreak set match conditional on set score (e, f).

Recurrence Formulas:

$P^{sm_{3T}}(e, f) = p^{s_T} P^{sm_{3T}}(e + 1, f) + q^{s_T} P^{sm_{3T}}(e, f + 1)$

$P^{sm_{5T}}(e, f) = p^{s_T} P^{sm_{5T}}(e + 1, f) + q^{s_T} P^{sm_{5T}}(e, f + 1)$

Boundary Values:

$P^{sm_{3T}}(e, f) = 1$, if $e = 2$ and $f \leq 1$

$P^{sm_{3T}}(e, f) = 0$, if $f = 2$ and $e \leq 1$

$P^{sm_{5T}}(e, f) = 1$, if $e = 3$ and $f \leq 2$

$P^{sm_{5T}}(e, f) = 0$, if $f = 3$ and $e \leq 2$

Table 2.7 shows the conditional probabilities of player A winning a best-of-5 all tiebreak set match, given $p_A = 0.62$ and $p_B = 0.60$. It indicates that player A has a 0.626 probability of winning the match from the outset. It also shows that a small increase on serve for the stronger player magnifies throughout the match. When $p_A = 0.62$ and $p_B = 0.60$, this 0.02 increase in probability on serve for player A, magnifies to a 0.07 increase in probability to win a set, and a 0.13 increase in probability to win the match.

Theorem 2.5.2. *A best-of-3 all tiebreak set match is identical to starting a best-of-5 all tiebreak set match at 1 set-all. That is:*

$P^{sm_{3T}}(e, f) = P^{sm_{5T}}(e + 1, f + 1)$

Proof. At 1 set-all in a best-of-5 all tiebreak set match the scores are level and 3 sets remain to be played. The equivalence to a best-of-3 set match is obvious. $\square$

		B score			
		0	1	2	3
	0	0.626	0.421	0.183	0
A score	1	0.782	0.601	0.323	0
	2	0.919	0.813	0.568	0
	3	1	1	1	

Table 2.7: The conditional probabilities of player A winning a best-of-5 all tiebreak set match from various score lines for $p_A = 0.62$ and $p_B = 0.60$

Using the formulas given for a tiebreak set conditional on the point and game score and a best-of-5 all tiebreak set match conditional on the set score, calculations are obtained for a best-of-5 all tiebreak set match conditional on the point, game and set score as follows.

Let $P_A^{pm_{5T}}(a,b:c,d:e,f)$ represent the probability of player A winning a best-of-5 all tiebreak set match on serve from (e,f) in sets, (c,d) in games and (a,b) in points. This can be calculated by:

$$P_A^{pm_{5T}}(a,b:c,d:e,f) = P_A^{ps_T}(a,b:c,d)P^{sm_{5T}}(e+1,f)+(1-P_A^{ps_T}(a,b:c,d))P^{sm_{5T}}(e,f+1)$$

When p_A=p_B, players are of equal strength and the probabilities of either player winning a set or match is 0.5. When $p_A = 1 - p_B$, there is no advantage in serving, since either player has the same probability of winning a point regardless of whether they are serving or receiving. Hence, this becomes a one parameter model, where a player has a constant probability of winning a point throughout the match.

2.6 Winning a final set advantage match

The scoring structure of a final set advantage match of tennis is defined as follows. For a best-of-5 final set advantage match, the first player to reach 3 sets wins the match. The first 4 sets are tiebreak sets and the 5th set is played as an advantage set. For a best-of-3 final set advantage match, the first player to reach 2 sets wins the match. The first 2 sets are tiebreak sets and the 3rd set is played as an advantage set. The serving is defined the same as an all tiebreak set match.

Let $P_A^{sm_3}(e, f)$ and $P_B^{sm_3}(e, f)$ represent the probabilities of player A winning a best-of-3 final set advantage match from set score (e, f) given player A and player B are serving respectively, where e is the set score for player A and f is the set score for player B.

Similarly, let $P_A^{sm_5}(e, f)$ and $P_B^{sm_5}(e, f)$ represent the probabilities of player A winning a best-of-5 final set advantage match from set score (e, f) given player A and player B are serving respectively, where e is the set score for player A and f is the set score for player B.

Corollary 2.6.1. *$P_A^{sm_3}(e, f) = P_B^{sm_3}(e, f)$ and $P_A^{sm_5}(e, f) = P_B^{sm_5}(e, f)$, for all (e, f)*

Proof. This follows from Theorems 2.3.1 and 2.4.2 since there is no advantage in serving first in a tiebreak or advantage set. □

Based on Corollary 2.6.1 it becomes convenient to let $P^{sm_3}(e, f)$ and $P^{sm_5}(e, f)$ represent the probabilities of player A winning a best-of-3 and best-of-5 final set advantage match from set score (e, f) respectively and let p^{m_3} and p^{m_5} represent the probabilities of player A winning a best-of-3 and best-of-5 final set advantage match respectively from the outset.

It follows that $p^{m_3} = P^{sm_3}(0, 0)$ and $p^{m_5} = P^{sm_5}(0, 0)$.

Formulas are now given for a best-of-5 and best-of-3 final set advantage match conditional on set score (e, f).

Recurrence Formulas:

$P^{sm_3}(e, f) = p^{s_T} P^{sm_3}(e + 1, f) + q^{s_T} P^{sm_3}(e, f + 1)$

$P^{sm_5}(e,f) = p^{s_T} P^{sm_5}(e+1,f) + q^{s_T} P^{sm_5}(e,f+1)$

Boundary Values:

$P^{sm_3}(e,f) = 1$, if $e = 2$ and $f = 0$

$P^{sm_3}(e,f) = 0$, if $f = 2$ and $e = 0$

$P^{sm_3}(1,1) = p^s$

$P^{sm_5}(e,f) = 1$, if $e = 3$ and $f \leq 1$

$P^{sm_5}(e,f) = 0$, if $f = 3$ and $e \leq 1$

$P^{sm_5}(2,2) = p^s$

Table 2.8 shows the conditional probabilities of player A winning a best-of-5 final set advantage match, given $p_A = 0.62$ and $p_B = 0.60$. It indicates that player A has a 0.627 probability of winning the match from the outset.

		B score			
		0	1	2	3
	0	0.627	0.422	0.184	0
A score	1	0.783	0.603	0.325	0
	2	0.920	0.815	0.572	
	3	1	1		

Table 2.8: The conditional probabilities of player A winning a best-of-5 final set advantage match from various score lines for $p_A = 0.62$ and $p_B = 0.60$

Theorem 2.6.2. *A best-of-3 final set advantage match is identical to starting a best-of-5 final set advantage match at 1 set-all. That is:*

$P^{sm_3}(e,f) = P^{sm_5}(e+1,f+1)$

Proof. At 1 set-all in a best-of-5 final set advantage match the scores are level and 3 sets remain to be played. The equivalence to a best-of-3 set match is obvious. $\square$

Theorem 2.6.3. *If* $p_A > p_B$ *then* $p^m > p^{m_T}$ *and if* $p_A < p_B$ *then* $p^m < p^{m_T}$

Proof. The proof can be obtained similarly to the proof given in Theorem 2.4.3 □

Using the formulas given for a tiebreak and advantage set conditional on the point and game score and a best-of-5 final set advantage match conditional on the set score, calculations are obtained for a best-of-5 final set advantage match conditional on the point, game and set score as follows.

Let $P_A^{pm_5}(a, b : c, d : e, f)$ represent the probability of player A winning a best-of-5 final set advantage match on serve from (e, f) in sets, (c, d) in games and (a, b) in points. This can be calculated by:

$P_A^{pm_5}(a, b : c, d : e, f) = P_A^{ps_T}(a, b : c, d)P^{sm_5}(e+1, f) + (1 - P_A^{ps_T}(a, b : c, d))P^{sm_5}(e, f+1)$, if $(e, f) \neq (2, 2)$

$P_A^{pm_5}(a, b : c, d : e, f) = P_A^{ps}(a, b : c, d)$, if $(e, f) = (2, 2)$

Chapter 3

Winning a match: forward recursion

3.1 Introduction

Chapter 3 extends on chapter 1 (in section 1.6) by applying forward recursion to calculate probabilities of reaching a specific point score from any point score within a game. Calculations are also obtained for reaching a specific point score from any point score within a tiebreak game, reaching a specific game score from any point and game score within a tiebreak and advantage set, and reaching a specific set score from any point, game and set score within an all tiebreak set and final set advantage match.

3.2 Winning a game

Let $N_A^{pg}(g,h|a,b)$ be the probability of reaching a point score (g,h) in a game from point score (a,b) for player A serving, where a and b represent the current point score for player A and player B respectively, and g and h represent the projected point score for player A and player B respectively.

The forward recursion formulas are:

If $a=g$ and $b=h$, then $N_A^{pg}(g,h|a,b)=1$, otherwise

$N_A^{pg}(g,h|a,b) = p_A N_A^{pg}(g-1,h|a,b)$, if $g=4$ and $0 \le h \le 2$; $h=0$ and $1 \le g \le 4$; $g \ge 3$, $h \ge 3$ and $g = h+1$; $g \ge 3$, $h \ge 3$ and $g = h+2$

$N_A^{pg}(g,h|a,b) = q_A N_A^{pg}(g,h-1|a,b)$, if $h=4$ and $0 \le g \le 2$; $g=0$ and $1 \le h \le 4$; $g \ge 3$, $h \ge 3$ and $h = g+1$; $g \ge 3$, $h \ge 3$ and $h = g+2$

$N_A^{pg}(g,h|a,b) = p_A N_A^{pg}(g-1,h|a,b) + q_A N_A^{pg}(g,h-1|a,b)$, if $1 \le g \le 3$ and $1 \le h \le 3$; $g \ge 4, h \ge 4$ and $g = h$

$N_A^{pg}(0,0|0,0) = 0$

Note that since each cell will either be an initial value (when $a = g$ and $b = h$) or a recursion formula, the use of 'IF' statements are required to implement these formulas in a spreadsheet. Also note that the formulas shown in section 1.6 are just the special case where $a = 0$ and $b = 0$.

3.3 Winning a tiebreak game

Let $N_A^{pgT}(g,h|a,b)$ represent the probabilities of reaching a point score (g,h) in a tiebreak game from point score (a,b) for player A serving at (a,b), where a and b represent the current point score for player A and player B respectively, and g and h represent the projected point score for player A and player B respectively.

Recurrence formulas:

If $a = g$ and $b = h$, then $N_A^{pgT}(g,h|a,b) = 1$, otherwise

$N_A^{pgT}(g,h|a,b)$

$=p_A N_A^{pgT}(g-1,h|a,b)$, if $(a+b)$ mod $4 = 0$ or 3

$=q_B N_A^{pgT}(g-1,h|a,b)$, if $(a+b)$ mod $4 = 1$ or 2

for $(g,h) = (1,0); (4,0); (5,0); (7,1); (7,2); (7,5)$; $g \ge 6, h \ge 6$, $g = h+1$ and $(g+h)$ mod $4 = 1$; $g \ge 6, h \ge 6$, $g = h+2$ and $(g+h)$ mod $4 = 0$

$N_A^{pg_T}(g, h|a, b)$

$= q_B N_A^{pg_T}(g-1, h|a, b)$, if $(a+b)$ mod $4 = 0$ or 3

$= p_A N_A^{pg_T}(g-1, h|a, b)$, if $(a+b)$ mod $4 = 1$ or 2

for $(g, h) = (2, 0)$; $(3, 0)$; $(6, 0)$; $(7, 0)$; $(7, 3)$; $(7, 4)$; $g \geq 6, h \geq 6$, $g = h+1$ and $(g+h)$ mod $4 = 3$; $g \geq 6, h \geq 6$, $g = h+2$ and $(g+h)$ mod $4 = 2$

$N_A^{pg_T}(g, h|a, b)$

$= q_A N_A^{pg_T}(g, h-1|a, b)$, if $(a+b)$ mod $4 = 0$ or 3

$= p_B N_A^{pg_T}(g, h-1|a, b)$, if $(a+b)$ mod $4 = 1$ or 2

for $(g, h) = (0, 1)$; $(0, 4)$; $(0, 5)$; $(1, 7)$; $(2, 7)$; $(5, 7)$; $g \geq 6, h \geq 6$, $h = g+1$ and $(g+h)$ mod $4 = 1$; $g \geq 6, h \geq 6$, $h = g+2$ and $(g+h)$ mod $4 = 0$

$N_A^{pg_T}(g, h|a, b)$

$= p_B N_A^{pg_T}(g, h-1|a, b)$, if $(a+b)$ mod $4 = 0$ or 3

$= q_A N_A^{pg_T}(g, h-1|a, b)$, if $(a+b)$ mod $4 = 1$ or 2

for $(g, h) = (0, 2)$; $(0, 3)$; $(0, 6)$; $(0, 7)$; $(3, 7)$; $(4, 7)$; $g \geq 6, h \geq 6$, $h = g+1$ and $(g+h)$ mod $4 = 3$; $g \geq 6, h \geq 6$, $h = g+2$ and $(g+h)$ mod $4 = 2$

$N_A^{pg_T}(g, h|a, b)$

$= q_B N_A^{pg_T}(g-1, h|a, b) + p_B N_A^{pg_T}(g, h-1|a, b)$, if $(a+b)$ mod $4 = 0$ or 3

$= p_A N_A^{pg_T}(g-1, h|a, b) + q_A N_A^{pg_T}(g, h-1|a, b)$, if $(a+b)$ mod $4 = 1$ or 2

for $(g, h) = (1, 1)$; $(2, 1)$; $(5, 1)$; $(6, 1)$; $(1, 2)$; $(4, 2)$; $(5, 2)$; $(3, 3)$; $(4, 3)$; $(2, 4)$; $(3, 4)$; $(6, 4)$; $(1, 5)$; $(2, 5)$; $(5, 5)$; $(6, 5)$; $(1, 6)$; $(4, 6)$; $(5, 6)$; $g \geq 6, h \geq 6$, $g = h$ and $(g+h)$ mod $4 = 2$

$N_A^{pg_T}(g, h|a, b)$

$= p_A N_A^{pg_T}(g-1, h|a, b) + q_A N_A^{pg_T}(g, h-1|a, b)$, if $(a+b)$ mod $4 = 0$ or 3

$= q_B N_A^{pg_T}(g-1, h|a, b) + p_B N_A^{pg_T}(g, h-1|a, b)$, if $(a+b)$ mod $4 = 1$ or 2

for $(g, h) = (3, 1)$; $(4, 1)$; $(2, 2)$; $(3, 2)$; $(6, 2)$; $(1, 3)$; $(2, 3)$; $(5, 3)$; $(6, 3)$; $(1, 4)$; $(4, 4)$; $(5, 4)$; $(3, 5)$; $(4, 5)$; $(2, 6)$; $(3, 6)$; $(6, 6)$; $g \geq 6, h \geq 6$, $g = h$ and $(g+h)$ mod $4 = 0$

$N_A^{pg_T}(0,0|0,0) = 0$

Tables 3.1 and 3.2 list the probability of reaching various score lines (g, h) in a tiebreak game from $(a = 0, b = 0)$ for player A and player B serving respectively at $(a = 0, b = 0)$, with $p_A = 0.62$ and $p_B = 0.60$. It indicates that the probability of reaching 6 points-all in such a tiebreak game is given by 0.231 for player A or player B serving from the outset.

		B score							
		0	1	2	3	4	5	6	7
	0	1	0.380	0.228	0.137	0.052	0.020	0.012	0.007
	1	0.620	0.524	0.406	0.239	0.123	0.082	0.054	0.020
	2	0.248	0.358	0.388	0.295	0.226	0.169	0.097	0.037
A score	3	0.099	0.260	0.339	0.322	0.284	0.212	0.141	0.085
	4	0.062	0.185	0.246	0.276	0.281	0.238	0.199	0.120
	5	0.038	0.097	0.157	0.231	0.262	0.252	0.231	0.088
	6	0.015	0.048	0.115	0.187	0.217	0.231	0.231	
	7	0.006	0.030	0.071	0.075	0.087	0.143		

Table 3.1: The probability of reaching various score lines (g, h) in a tiebreak game from $(a = 0, b = 0)$ with $p_A = 0.62$ and $p_B = 0.60$, for player A serving at $(a = 0, b = 0)$

		B score							
		0	1	2	3	4	5	6	7
	0	1	0.600	0.228	0.087	0.052	0.031	0.012	0.005
	1	0.400	0.524	0.340	0.239	0.164	0.082	0.038	0.023
	2	0.248	0.419	0.388	0.328	0.226	0.137	0.097	0.058
A score	3	0.154	0.260	0.311	0.322	0.263	0.212	0.166	0.063
	4	0.062	0.141	0.246	0.293	0.281	0.253	0.199	0.076
	5	0.025	0.097	0.189	0.231	0.251	0.252	0.220	0.132
	6	0.015	0.066	0.115	0.162	0.217	0.239	0.231	
	7	0.009	0.026	0.046	0.100	0.134	0.096		

Table 3.2: The probability of reaching various score lines (g, h) in a tiebreak game from $(a = 0, b = 0)$ with $p_A = 0.62$ and $p_B = 0.60$, for player B serving at $(a = 0, b = 0)$

Let $N_A^{pg_T}(g, h)$ represent the probabilities of reaching a point score (g, h) in a tiebreak game from the outset for player A serving at the outset of the game.

It follows that $N_A^{pg_T}(g,h) = N_A^{pg_T}(g,h|0,0)$.

The probability of player A winning the tiebreak game from the outset can be obtained from:

$N_A^{pg_T}(7,0)+N_A^{pg_T}(7,1)+N_A^{pg_T}(7,2)+N_A^{pg_T}(7,3)+N_A^{pg_T}(7,4)+N_A^{pg_T}(7,5)+N_A^{pg_T}(6,6)P_A^{pg_T}(6,6)$

3.4 Winning a tiebreak set

Let $N_A^{gs_T}(i,j|c,d)$ represent the probabilities of reaching a game score (i,j) in a tiebreak set from game score (c,d) for player A serving at (c,d), where c and d represent the current game score for player A and player B respectively, and i and j represent the projected game score for player A and player B respectively.

Recurrence formulas:

If $c=i$ and $d=j$, then $N_A^{gs_T}(i,j|c,d)=1$, otherwise

$N_A^{gs_T}(i,j|c,d)$
$= p_A^g N_A^{gs_T}(i-1,j|c,d)$, if $(c+d)$ is even
$= q_B^g N_A^{gs_T}(i-1,j|c,d)$, if $(c+d)$ is odd
for (i,j)=$(1,0)$; $(3,0)$; $(5,0)$; $(6,1)$; $(6,3)$; $(6,5)$

$N_A^{gs_T}(i,j|c,d)$
$= q_B^g N_A^{gs_T}(i-1,j|c,d)$, if $(c+d)$ is even
$= p_A^g N_A^{gs_T}(i-1,j|c,d)$, if $(c+d)$ is odd
for (i,j)=$(2,0)$; $(4,0)$; $(6,0)$; $(6,2)$; $(6,4)$; $(7,5)$

$N_A^{gs_T}(i,j|c,d)$

$= q_A^g N_A^{gs_T}(i, j-1|c, d)$, if $(c+d)$ is even

$= p_B^g N_A^{gs_T}(i, j-1|c, d)$, if $(c+d)$ is odd

for (i, j)=$(0, 1)$; $(0, 3)$; $(0, 5)$; $(1, 6)$; $(3, 6)$; $(5, 6)$

$N_A^{gs_T}(i, j|c, d)$

$= p_B^g N_A^{gs_T}(i, j-1|c, d)$, if $(c+d)$ is even

$= q_A^g N_A^{gs_T}(i, j-1|c, d)$, if $(c+d)$ is odd

for (i, j)=$(0, 2)$; $(0, 4)$; $(0, 6)$; $(2, 6)$; $(4, 6)$; $(5, 7)$

$N_A^{gs_T}(i, j|c, d)$

$= p_A^g N_A^{gs_T}(i-1, j|c, d) + q_A^g N_A^{gs_T}(i, j-1|c, d)$, if $(c+d)$ is even

$= q_B^g N_A^{gs_T}(i-1, j|c, d) + p_B^g N_A^{gs_T}(i, j-1|c, d)$, if $(c+d)$ is odd

for (i, j)=$(2, 1)$; $(4, 1)$; $(1, 2)$; $(3, 2)$; $(5, 2)$; $(2, 3)$; $(4, 3)$; $(1, 4)$; $(3, 4)$; $(5, 4)$; $(2, 5)$; $(4, 5)$

$N_A^{gs_T}(i, j|c, d)$

$= q_B^g N_A^{gs_T}(i-1, j|c, d) + p_B^g N_A^{gs_T}(i, j-1|c, d)$, if $(c+d)$ is even

$= p_A^g N_A^{gs_T}(i-1, j|c, d) + q_A^g N_A^{gs_T}(i, j-1|c, d)$, if $(c+d)$ is odd

for (i, j)=$(1, 1)$; $(3, 1)$; $(5, 1)$; $(2, 2)$; $(4, 2)$; $(1, 3)$; $(3, 3)$; $(5, 3)$; $(2, 4)$; $(4, 4)$; $(1, 5)$; $(3, 5)$; $(5, 5)$; $(6, 6)$

$N_A^{gs_T}(7, 6|c, d) = p^{g_T} N_A^{gs_T}(i-1, j|c, d)$

$N_A^{gs_T}(6, 7|c, d) = q^{g_T} N_A^{gs_T}(i, j-1|c, d)$

$N_A^{gs_T}(0, 0|0, 0) = 0$

Tables 3.3 and 3.4 list the probability of reaching various score lines (i, j) in a tiebreak set from $(c = 0, d = 0)$ for player A and player B serving respectively at $(c = 0, d = 0)$, with p_A=0.62 and p_B=0.60. It indicates that the probability of reaching a tiebreak game in such a tiebreak set is given by 0.183 for player A or player B serving from the outset.

		B score							
		0	1	2	3	4	5	6	7
	0	1	0.224	0.165	0.037	0.027	0.006	0.004	
	1	0.776	0.630	0.269	0.208	0.068	0.051	0.012	
	2	0.205	0.535	0.465	0.265	0.213	0.088	0.064	
A score	3	0.159	0.258	0.418	0.378	0.250	0.207	0.046	
	4	0.042	0.210	0.265	0.353	0.326	0.234	0.172	
	5	0.033	0.079	0.223	0.258	0.310	0.290	0.065	0.048
	6	0.009	0.062	0.059	0.200	0.082	0.225	0.183	0.085
	7						0.059	0.097	

Table 3.3: The probability of reaching various score lines (i, j) in a tiebreak set from $(c = 0, d = 0)$ with $p_A = 0.62$ and $p_B = 0.60$, for player A serving at $(c = 0, d = 0)$

		B score							
		0	1	2	3	4	5	6	7
	0	1	0.736	0.165	0.121	0.027	0.020	0.004	
	1	0.264	0.630	0.507	0.208	0.160	0.051	0.038	
	2	0.205	0.317	0.465	0.397	0.213	0.170	0.038	
A score	3	0.054	0.258	0.313	0.378	0.334	0.207	0.152	
	4	0.042	0.099	0.265	0.295	0.326	0.294	0.066	
	5	0.011	0.079	0.128	0.258	0.276	0.290	0.213	0.048
	6	0.009	0.021	0.100	0.068	0.214	0.077	0.183	0.085
	7						0.059	0.097	

Table 3.4: The probability of reaching various score lines (i, j) in a tiebreak set from $(c = 0, d = 0)$ with $p_A = 0.62$ and $p_B = 0.60$, for player B serving at $(c = 0, d = 0)$

Let $N_A^{gs_T}(g, h)$ represent the probabilities of reaching a game score (g, h) in a tiebreak set from the outset for player A serving at the outset of the set.

The probability of player A winning the tiebreak set from the outset can be obtained from:
$N_A^{gs_T}(6,0) + N_A^{gs_T}(6,1) + N_A^{gs_T}(6,2) + N_A^{gs_T}(6,3) + N_A^{gs_T}(6,4) + N_A^{gs_T}(7,5) + N_A^{gs_T}(7,6)$

Let $N_A^{gs_T}(i, j|a, b : c, d)$ represent the probabilities of reaching a game score (i, j) in a tiebreak set from point and game score $(a, b : c, d)$ for player A serving at (a, b).

$N_A^{gs_T}(i,j|a,b:c,d) = P_A^{pg}(a,b)N_B^{gs_T}(i,j|c+1,d)+(1-P_A^{pg}(a,b))N_B^{gs_T}(i,j|c,d+1)$, if $(c,d) \neq (6,6)$

$N_A^{gs_T}(i,j|a,b:c,d) = P_A^{pg_T}(a,b)N_B^{gs_T}(i,j|c+1,d) + (1-P_A^{pg_T}(a,b))N_B^{gs_T}(i,j|c,d+1)$, if $(c,d) = (6,6)$

For example with $p_A = 0.62$ and $p_B = 0.60$:

$N_A^{gs_T}(6,3|1,0:3,2) = P_A^{pg}(1,0)N_B^{gs_T}(6,3|4,2) + (1-P_A^{pg}(1,0))N_B^{gs_T}(6,3|3,3) = 0.870 * 0.167 + 0.130 * 0.054 = 0.152$

3.5 Winning an advantage set

Let $N_A^{gs}(i,j|c,d)$ represent the probabilities of reaching a game score (i,j) in an advantage set from game score (c,d) for player A serving at (c,d), where c and d represent the current game score for player A and player B respectively, and g and h represent the projected game score for player A and player B respectively.

Recurrence formulas:

If $c = i$ and $d = j$, then $N_A^{gs}(i,j|c,d) = 1$, otherwise

$N_A^{gs}(i,j|c,d)$
$= p_A^g N_A^{gs}(i-1,j|c,d)$, if $(c+d)$ is even
$= q_B^g N_A^{gs}(i-1,j|c,d)$, if $(c+d)$ is odd
for (i,j)=$(1,0)$; $(3,0)$; $(5,0)$; $(6,1)$; $(6,3)$; $(6,5)$; $i \geq 6$, $j \geq 6$ and $i = j+1$

$N_A^{gs}(i,j|c,d)$
$= q_B^g N_A^{gs}(i-1,j|c,d)$, if $(c+d)$ is even
$= p_A^g N_A^{gs}(i-1,j|c,d)$, if $(c+d)$ is odd
for (i,j)=$(2,0)$; $(4,0)$; $(6,0)$; $(6,2)$; $(6,4)$; $i \geq 5$, $j \geq 5$ and $i = j+2$

$N_A^{gs}(i,j|c,d)$

$= q_A^g N_A^{gs}(i,j-1|c,d)$, if $(c+d)$ is even

$= p_B^g N_A^{gs}(i,j-1|c,d)$, if $(c+d)$ is odd

for (i,j)=$(0,1)$; $(0,3)$; $(0,5)$; $(1,6)$; $(3,6)$; $(5,6)$; $i \geq 6$, $j \geq 6$ and $j = i+1$

$N_A^{gs}(i,j|c,d)$

$= p_B^g N_A^{gs}(i,j-1|c,d)$, if $(c+d)$ is even

$= q_A^g N_A^{gs}(i,j-1|c,d)$, if $(c+d)$ is odd

for (i,j)=$(0,2)$; $(0,4)$; $(0,6)$; $(2,6)$; $(4,6)$; $i \geq 5, j \geq 5$ and $j = i+2$

$N_A^{gs}(i,j|c,d)$

$= p_A^g N_A^{gs}(i-1,j|c,d) + q_A^g N_A^{gs}(i,j-1|c,d)$, if $(c+d)$ is even

$= q_B^g N_A^{gs}(i-1,j|c,d) + p_B^g N_A^{gs}(i,j-1|c,d)$, if $(c+d)$ is odd

for (i,j)=$(2,1)$; $(4,1)$; $(1,2)$; $(3,2)$; $(5,2)$; $(2,3)$; $(4,3)$; $(1,4)$; $(3,4)$; $(5,4)$; $(2,5)$; (4,5)

$N_A^{gs}(i,j|c,d)$

$= q_B^g N_A^{gs}(i-1,j|c,d) + p_B^g N_A^{gs}(i,j-1|c,d)$, if $(c+d)$ is even

$= p_A^g N_A^{gs}(i-1,j|c,d) + q_A^g N_A^{gs}(i,j-1|c,d)$, if $(c+d)$ is odd

for (i,j)=$(1,1)$; $(3,1)$; $(5,1)$; $(2,2)$; $(4,2)$; $(1,3)$; $(3,3)$; $(5,3)$; $(2,4)$; $(4,4)$; $(1,5)$; $(3,5)$; $(5,5)$; $(6,6)$; $i \geq 7$, $j \geq 7$ and $i = j$

$N_A^{gs}(0,0|0,0) = 0$

Let $N_A^{gs}(g,h)$ represent the probabilities of reaching a game score (g,h) in an advantage set from the outset for player A serving at the outset of the set.

The probability of player A winning the advantage set can be obtained from:

$N_A^{gs}(6,0) + N_A^{gs}(6,1) + N_A^{gs}(6,2) + N_A^{gs}(6,3) + N_A^{gs}(6,4) + N_A^{gs}(5,5)P_A^{gs}(5,5)$

Let $N_A^{gs}(i,j|a,b : c,d)$ represent the probabilities of reaching a game score (i,j) in an advantage set from point and game score $(a,b : c,d)$ for player A serving at (a,b).

$N_A^{gs}(i,j|a,b:c,d) = P_A^{pg}(a,b)N_B^{gs}(i,j|c+1,d) + (1-P_A^{pg}(a,b))N_B^{gs}(i,j|c,d+1)$

3.6 Winning an all tiebreak set match

Let $N^{sm_{5T}}(k,l|e,f)$ represent the probabilities of reaching a set score (k,l) in a best-of-5 all tiebreak set match from set score (e,f), where e and f represent the current set score for player A and player B respectively, and k and l represent the projected set score for player A and player B respectively.

Recurrence Formulas:

If $e = k$ and $f = l$, then $N^{sm_{5T}}(k,l|e,f) = 1$, otherwise

$N^{sm_{5T}}(k,l|e,f) = p^{sT}N^{sm_{5T}}(k-1,l|e,f)$, if $k = 3$ and $0 \leq l \leq 2$; $l = 0$ and $1 \leq k \leq 3$

$N^{sm_{5T}}(k,l|e,f) = q^{sT}N^{sm_{5T}}(k,l-1|e,f)$, if $l = 3$ and $0 \leq k \leq 2$; $k = 0$ and $1 \leq l \leq 3$

$N^{sm_{5T}}(k,l|e,f) = p^{sT}N^{sm_{5T}}(k-1,l|e,f)+q^{sT}N^{sm_{5T}}(k,l-1|e,f)$, if $1 \leq k \leq 2$ and $1 \leq l \leq 2$

$N^{sm_{5T}}(0,0|0,0) = 0$

Table 3.5 lists the probability of reaching various score lines (k,l) in a best-of-5 all tiebreak set match from $(e = 0, f = 0)$ with p_A=0.62 and p_B=0.60.

		B score			
		0	1	2	3
	0	1	0.432	0.187	0.081
A score	1	0.568	0.491	0.318	0.137
	2	0.323	0.418	0.361	0.156
	3	0.183	0.237	0.205	

Table 3.5: The probability of reaching various score lines (k,l) in a best-of-5 all tiebreak set match from $(e = 0, f = 0)$, with $p_A = 0.62$ and $p_B = 0.60$

Let $N^{sm_{5T}}(k,l)$ represent the probabilities of reaching a set score (k,l) in a best-of-5 all tiebreak set match from the outset.

The probability of player A winning a best-of-5 all tiebreak set match from the outset can be obtained from:

$N^{sm_{5T}}(3,0) + N^{sm_{5T}}(3,1) + N^{sm_{5T}}(3,2)$

Let $N_A^{sm_{5T}}(k,l|a,b:c,d:e,f)$ represent the probabilities of reaching a set score (k,l) in a best-of-5 all tiebreak set match from point, game and set score $(a,b:c,d:e,f)$ for player A serving at (a,b).

$N_A^{sm_{5T}}(k,l|a,b:c,d:e,f)$
$= P_A^{ps_T}(a,b:c,d)N^{sm_{5T}}(k,l|e+1,f) + (1 - P_A^{ps_T}(a,b:c,d))N^{sm_{5T}}(k,l|e,f+1)$

For example with $p_A = 0.62$ and $p_B = 0.60$:
$N_A^{sm_{5T}}(3,2|1,0:3,2:2,1)$
$= P_A^{ps_T}(1,0:3,2)N^{sm_{5T}}(3,2|3,1) + (1 - P_A^{ps_T}(1,0:3,2))N^{sm_{5T}}(3,2|2,2))$
=0.835*0+(1-0.835)*0.568
=0.094

3.7 Winning a final set advantage match

Let $N^{sm_5}(k,l|e,f)$ represent the probabilities of reaching a set score (k,l) in a best-of-5 final set advantage match from set score (e,f), where e and f represent the current set score for player A and player B respectively, and k and l represent the projected set score for player A and player B respectively.

Recurrence Formulas:

If $e = k$ and $f = l$, then $N^{sm_5}(k,l|e,f) = 1$, otherwise

$N^{sm_5}(k,l|e,f) = p^{s_T}N^{sm_5}(k-1,l|e,f)$, if $1 \leq k \leq 3$ and $l = 0$; $k = 3, l = 1$
$N^{sm_5}(k,l|e,f) = q^{s_T}N^{sm_5}(k,l-1|e,f)$, if $1 \leq l \leq 3$ and $k = 0$; $k = 1, l = 3$

$N^{sm_5}(k,l|e,f) = p^{s_T} N^{sm_5}(k-1,l|e,f) + q^{s_T} N^{sm_5}(k,l-1|e,f)$, if $1 \leq k \leq 2$ and $1 \leq l \leq 2$

$N^{sm_5}(k,l|e,f) = p^s N^{sm_5}(k-1,l|e,f)$, if $k=3, l=2$

$N^{sm_5}(k,l|e,f) = q^s N^{sm_5}(k,l-1|e,f)$, if $l=3, k=2$

$N^{sm_5}(0,0|0,0) = 0$

Table 3.6 lists the probability of reaching various score lines (k,l) in a best-of-5 final set advantage match from $(e=0, f=0)$, with p_A=0.62 and p_B=0.60.

		B score			
		0	1	2	3
	0	1	0.432	0.187	0.081
A score	1	0.568	0.491	0.318	0.137
	2	0.323	0.418	0.361	0.155
	3	0.183	0.237	0.207	

Table 3.6: The probability of reaching various score lines (k,l) in a best-of-5 final set advantage match from $(e=0, f=0)$, with $p_A=0.62$ and $p_B=0.60$

Let $N^{sm_5}(k,l)$ represent the probabilities of reaching a set score (k,l) in a best-of-5 final set advantage match from the outset.

The probability of player A winning a best-of-5 final set advantage match from the outset can be obtained from:

$N^{sm_5}(3,0) + N^{sm_5}(3,1) + N^{sm_5}(3,2)$

Let $N_A^{sm_5}(k,l|a,b:c,d:e,f)$ represent the probabilities of reaching a set score (k,l) in a best-of-5 final set advantage match from point, game and set score $(a,b:c,d:e,f)$ for player A serving at (a,b).

$N_A^{sm_5}(k,l|a,b:c,d:e,f)$

$= P_A^{ps_T}(a,b:c,d) N^{sm_5}(k,l|e+1,f) + (1-P_A^{ps_T}(a,b:c,d)) N^{sm_5}(k,l|e,f+1)$, if $(e,f) \neq (2,2)$

$$N_A^{sm_5}(k,l|a,b:c,d:e,f)$$
$$= P_A^{ps}(a,b:c,d)N^{sm_5}(k,l|e+1,f) + (1-P_A^{ps}(a,b:c,d))N^{sm_5}(k,l|e,f+1), \text{ if } (e,f)=(2,2)$$

Chapter 4

Duration of a game

4.1 Introduction

Chapter 4 obtains calculations for the duration of a game. This consists of two distributions; namely the distribution of the total number of points played in a game from any point score within the game and the distribution of the number of points remaining in the game from any point score within the game. The results obtained in chapter 3 (in section 3.2) on the probability of reaching a specific point score from any point score within the game are used to obtain the results for both distributions, and hence the connection between chances of winning and duration. Chapter 4 also obtains calculations for the parameters of distribution. Four parameters of distribution (mean, variance, coefficient of skewness and coefficient of kurtosis) are obtained in section 4.3 from the outset of the game, and two parameters of distribution (mean and variance) are obtained in section 4.5 from any point score within the game. The techniques to obtain these calculations consist of the Binomial theorem, generating functions, forward recursion and backward recursion.

4.2 Binomial theorem

Suppose we wish to determine the distribution of the total number of points played in a game from the outset; that is: “What is the probability of playing 4,5,6,8,10.....points?”. It was established in section 1.3 using the Binomial theorem, that the probability of player A winning the game to 0 on serve is given by p_A^4. Similarly the probability of player B winning the game to 0 given player A is serving is q_A^4. This implies that the probability of playing exactly 4 points in the game given player A is serving is $p_A^4 + q_A^4$ (since 4 points are required for either player A or player B to win the game to 0). Similarly, the probability of playing exactly 5 points in the game is $4p_A^4 q_A + 4q_A^4 p_A$. Similarly the probability of playing exactly 6 points in the game is $10p_A^4 q_A^2 + 10q_A^4 p_A^2$. The probability of reaching deuce was obtained as $20p_A^3 q_A^3$. Therefore the probability of playing exactly 8 points is obtained as $20p_A^3 q_A^3(p_A^2 + q_A^2)$. Similarly the probability of playing exactly 10 points is obtained as $20p_A^3 q_A^3(p_A^2 + q_A^2)(2p_A q_A)$. Similarly the probability of playing exactly 12 points is obtained as $20p_A^3 q_A^3(p_A^2 + q_A^2)(2p_A q_A)^2$. Therefore the probability of playing x points (x=8,10,12,14...) is obtained as $20p_A^3 q_A^3(p_A^2 + q_A^2)(2p_A q_A)^{\frac{x-8}{2}}$.

More formally, let X_A^{pg} be a random variable of the total number of points played in a game from the outset for player A serving. Let $f(X_A^{pg} = x_A^{pg})$ represent the distribution of the total number of points played in the game from the outset for player A serving. This distribution is given as follows.

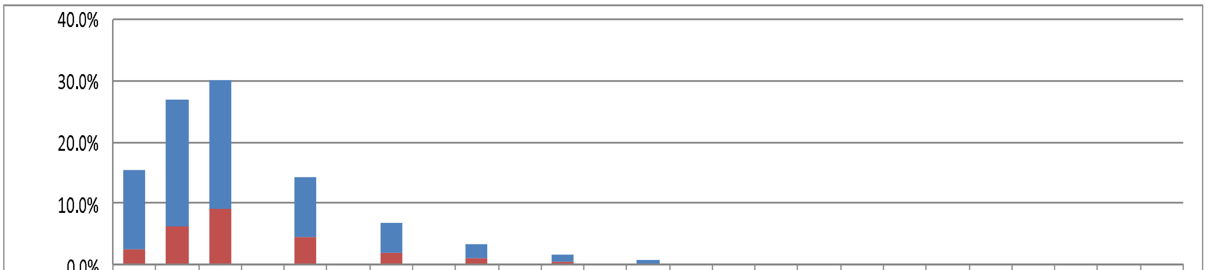
$f(X_A^{pg} = 4) = p_A^4 + q_A^4$

$f(X_A^{pg} = 5) = 4p_A^4 q_A + 4q_A^4 p_A$

$f(X_A^{pg} = 6) = 10p_A^4 q_A^2 + 10q_A^4 p_A^2$

$f(X_A^{pg} = x_A^{pg}) = 20p_A^3 q_A^3(p_A^2 + q_A^2)(2p_A q_A)^{\frac{x_A^{pg}-8}{2}}$, if $x_A^{pg} = 8, 10, 12, ...$

Figure 4.1 represents the distribution graphically of the total number of points played in a game from the outset for $p_A = 0.6$. Notice how the blue colour is the chances of player



	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
Player A	13.0%	20.7%	20.7%	0.0%	10.0%	0.0%	4.8%	0.0%	2.3%	0.0%	1.1%	0.0%	0.5%	0.0%	0.3%	0.0%	0.1%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Player B	2.6%	6.1%	9.2%	0.0%	4.4%	0.0%	2.1%	0.0%	1.0%	0.0%	0.5%	0.0%	0.2%	0.0%	0.1%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Figure 4.1: The distribution of the total number of points played in a game from the outset for $p_A = 0.6$

A winning the game and the maroon colour is the chances of player B winning the game. For example, the chances of player A winning the game to 15 is given by the frequency distribution of blue for 5 total points played. This numerical value is 20.74%. Similarly, the chances of player B winning the game to 15 is given by the frequency distribution of maroon for 5 total points played. This numerical value is 6.14%. Therefore, the game finishing with either player winning to 15 (or 5 total points played) is given by 20.74% + 6.14% = 26.9%.

4.3 Generating functions

Suppose we wish to calculate the mean (average value) of the total number of points played in a game. Using a standard formula for calculating the mean value of a discrete distribution, this can be calculated by $\mu = E(X) = \sum_{x_A^{pg}} x_A^{pg} f(x_A^{pg})$. Similarly the variance (standard deviation squared) of the total number of points played in a game can be calculated by $\sigma^2 = E(X^2) - E(X)^2 = \sum_{x_A^{pg}} (x_A^{pg})^2 f(x_A^{pg}) - (\sum_{x_A^{pg}} x_A^{pg} f(x_A^{pg}))^2$; which is recognized as a measure of the dispersion of a set of data from its mean. Both the mean and the variance contain important information to describe the shape of the distribution and these characteristics could be used to compare one distribution to another. For example, comparing the mean and variance of a tiebreak set to an advantage set to identify why 'long' matches can occur. However, if a distribution is not symmetric (as typically occurs in a game and an advantage set) the mean and variance do not 'adequately' describe the shape of the distribution. Two other characteristics that are used to describe the distribution and measure risk are skewness and kurtosis. Skewness is a measure of symmetry, or more precisely, the lack of symmetry. A distribution, or data set, is symmetric if it looks the same to the left and right of the center point. Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution. In order to calculate the skewness and kurtosis, it becomes convenient to work with generating functions, as the four characteristics (mean, variance, skewness and kurtosis) can readily be obtained.

The expectation of a random variable X is calculated by $E(X) = \sum_x xf(x)$.

The expectation of the n^{th} power of a random variable X is calculated by $E(X^n) = \sum_x x^n f(x)$. This is also known as the n^{th} moment of the random variable X and represented by m_{nX}, such that:
$m_{nX} = E(X^n)$ for $n = 1, 2, 3, 4, ...$

The moment generating function for a random variable X is defined by $M_X(t) = E(e^{tX}) = \sum_x e^{tx} f(x)$.

The n^{th} moment for a random variable X can be recovered from the moment generating function by differentiating n times and setting $t = 0$. Thus $E(X^n) = M_X^{(n)}(0) = \frac{d^n}{dt^n} M_X(0)$.

The moment generating function can be expanded in terms of the moments as an infinite series in t to obtain
$M_X(t) = 1 + E(X)t + E(X^2)\frac{t^2}{2!} + E(X^3)\frac{t^3}{3!} + E(X^4)\frac{t^4}{4!} + ...$

The moment generating function of X is said to exist if moments of X are finite. One way of showing that a moment generating function exists is to prove that the series converges to a finite value over an interval for t that includes 0.

For a geometric distribution the probabilities are given by

$P(X = k) = p_k = p(1-p)^k$ for $k = 0, 1, 2, ...$ with $0 < p < 1$.

Let $q = 1 - p$, so $P(Y = k) = p_k = pq^k$ for $k = 0, 1, 2, ...$ with $0 < q < 1$.

$M_X(t)$

$= pe^0 + pqe^t + pq^2e^{2t} + pq^3e^{3t} + pq^4e^{4t}...$

$= p(1 + qe^t + q^2e^{2t} + q^3e^{3t} + q^4e^{4t}...)$

$= p/(1 - qe^t)$ provided $0 \leq qe^t \leq 1$, i.e. $-\infty < t < \ln(1/q)$

The moments can be recovered from this expression by successive differentiation with respect to t, and putting $t = 0$. Thus

$M_X^{(1)}(t) = pqe^t/(1-qe^t)^2$, giving $E(X) = M_X^{(1)}(0) = q/p$

$M_X^{(2)}(t) = pqe^t(1+qe^t)/(1-qe^t)^3$, giving $E(X^2) = M_X^{(2)}(0) = q(1+q)/p^2$

The cumulant generating function $K_X(t)$ for a random variable X is given by $K_X(t) = log_e M_X(t)$. The mean, variance, coefficient of skewness and coefficient of excess kurtosis of X are given by:

$\mu(X) = K_X^{(1)}(0) = M_X^{(1)}(0)$

$\sigma^2(X) = K_X^{(2)}(0) = M_X^{(2)}(0) - M_X^{(1)}(0)^2$

$\gamma_1(X) = \frac{K_X^{(3)}(0)}{K_X^{(2)}(0)^{\frac{3}{2}}} = \frac{M_X^{(3)}(0) - 3M_X^{(2)}(0)M_X^{(1)}(0) + 2M_X^{(1)}(0)^3}{(M_X^{(2)}(0) - M_X^{(1)}(0)^2)^{\frac{3}{2}}}$

$\gamma_2(X) = \frac{K_X^{(4)}(0)}{K_X^{(2)}(0)^2} = \frac{M_X^{(4)}(0) - 4M_X^{(3)}(0)M_X^{(1)}(0) - 3M_X^{(2)}(0)^2 + 12M_X^{(2)}(0)M_X^{(1)}(0)^2 - 6M_X^{(1)}(0)^4}{(M_X^{(2)}(0) - M_X^{(1)}(0)^2)^2}$

where:

$\mu(X)$ = mean

$\sigma^2(X)$ = variance

$\gamma_1(X)$ = coefficient of skewness

$\gamma_2(X)$ = coefficient of excess kurtosis

$M_X^{(n)}(0)$ = the n^{th} derivative of the moment generating function evaluated at $t = 0$

$K_X^{(n)}(0)$ = the n^{th} derivative of the cumulant generating function evaluated at $t = 0$

In general, excess kurtosis = kurtosis - 3, so the normal distribution has an excess kurtosis of 0, and therefore a kurtosis of 3.

It can be observed that working with the cumulant generating function as oppose to using the moment generating function, simplifies calculations for the coefficients of skewness and excess kurtosis. This makes the calculations easier to implement on a mathematics software package such as *Mathematica* or a graphics calculator.

The moment generating function for the total number of points played in a game from the outset for player A serving, $M_{X_A^{pg}}(t)$, becomes:

$$\sum_{x_A^{pg}} e^{tx_A^{pg}} f(x_A^{pg}) = e^{4t}f(4) + e^{5t}f(5) + e^{6t}f(6) + \frac{N_A^{pg}(3,3)(1-N_A^{pg}(1,1))e^{8t}}{1-N_A^{pg}(1,1)e^{2t}}$$

In this example we have a few cases of positive probability before the point score of $(3,3)$, the first deuce, and then a shifted geometric distribution over the even point scores that follow. By comparing this situation with the geometric distribution example above, it is easy to see that the moment generating function will exist since the series will converge to a finite value over an interval of t that includes 0.

The cumulant generating function for the total number of points played in a game from the outset for player A serving, $K_{X_A^{pg}}(t)$, becomes:

$$log_e(e^{4t}f(4) + e^{5t}f(5) + e^{6t}f(6) + \frac{N_A^{pg}(3,3)(1-N_A^{pg}(1,1))e^{8t}}{1-N_A^{pg}(1,1)e^{2t}})$$

The first derivative of the cumulant generating function evaluated at $t = 0$, $K_{X_A^{pg}}^{(1)}(0)$, is equivalent to the mean of the total number of points played in a game, $\mu(X_A^{pg})$. It follows that:

$\mu(X_A^{pg}) = \frac{4(p_A q_A (6p_A^2 q_A^2 - 1) - 1)}{1 - 2p_A q_A}$

The second derivative of the cumulant generating function evaluated at $t = 0$, $K^{(2)}_{X_A^{pg}}(0)$, is equivalent to the variance of the total number of points played in a game, $\sigma^2(X_A^{pg})$. It follows that:

$\sigma^2(X_A^{pg}) = \frac{4p_A q_A (1 - p_A q_A (1 - 12 p_A q_A (3 - p_A q_A (5 + 12 p_A^2 q_A^2))))}{(1 - 2p_A q_A)^2}$

Let $\gamma_1(X_A^{pg})$ represent the coefficient of skewness of the total number of points played in a game for player A serving.

The third derivative of the cumulant generating function evaluated at $t = 0$, becomes:

$K^{(3)}_{X_A^{pg}}(0) = 4p_A q_A (q_A + 187p_A^2 - 840p_A^3 + 2118p_A^4 - 6108p_A^5 + 20916p_A^6 - 53952p_A^7 + 98160p_A^8 - 154656p_A^9 + 260928p_A^{10} - 412992p_A^{11} + 488160p_A^{12} - 387072p_A^{13} + 193536p_A^{14} - 55296p_A^{15} + 6912p_A^{16})/(1 - 2p_A + 2p_A^2)^3$

The coefficient of skewness of the total number of points played in a game can be calculated by:

$\gamma_1(X_A^{pg}) = \frac{\kappa^{(3)}_{X_A^{pg}}(0)}{\kappa^{(2)}_{X_A^{pg}}(0)^{\frac{3}{2}}}$

Let $\gamma_2(X_A^{pg})$ represent the coefficient of excess kurtosis of the total number of points played in a game for player A serving.

The fourth derivative of the cumulant generating function evaluated at $t = 0$, becomes:

$K^{(4)}_{X_A^{pg}}(0) = 4p_A q_A (1 + p_A + 871p_A^2 - 4004p_A^3 + 13364p_A^4 - 67596p_A^5 + 323140p_A^6 - 1077024p_A^7 + 2742960p_A^8 - 6502224p_A^9 + 15475344p_A^{10} - 33228864p_A^{11} + 59797440p_A^{12} - 94218048p_A^{13}$

p_A	$\mu(X_A^{pg})$	$\sigma(X_A^{pg})$	$c_v(X_A^{pg})$	$\gamma_1(X_A^{pg})$	$\gamma_2(X_A^{pg})$
0.50	6.75	2.77	0.41	2.16	6.95
0.55	6.68	2.73	0.41	2.17	7.01
0.60	6.48	2.59	0.40	2.20	7.21
0.65	6.19	2.37	0.38	2.25	7.59
0.70	5.83	2.10	0.36	2.34	8.25
0.75	5.45	1.78	0.33	2.46	9.27
0.80	5.09	1.44	0.28	2.61	10.71
0.85	4.75	1.10	0.23	2.74	12.47
0.90	4.46	0.79	0.18	2.81	13.91
0.95	4.21	0.49	0.12	2.92	13.93

Table 4.1: The parameters of the distributions of the number of points played in a game for different values of p_A

$+141430464p_A^{14}-201056256p_A^{15}+245099520p_A^{16}-233653248p_A^{17}+164643840p_A^{18}-82114560p_A^{19}+27371520p_A^{20}-5474304p_A^{21}+497664p_A^{22})/(1-2p_A+2p_A^2)^4$

The coefficient of excess kurtosis of the total number of points played in a game can be calculated by:

$$\gamma_2(X_A^{pg}) = \frac{K_{X_A^{pg}}^{(4)}(0)}{K_{X_A^{pg}}^{(2)}(0)^2}$$

Let $\sigma(X_A^{pg})$ represent the standard deviation of the total number of points played in a game for player A serving. Let $c_v(X_A^{pg})$ represent the coefficient of variation of the total number of points played in a game for player A serving. It follows that $\sigma(X_A^{pg}) = \sqrt{\sigma^2(X_A^{pg})}$ and $c_v(X_A^{pg}) = \frac{\sigma(X_A^{pg})}{\mu(X_A^{pg})}$.

Table 4.1 represents $\mu(X_A^{pg})$, $\sigma(X_A^{pg})$, $c_v(X_A^{pg})$, $\gamma_1(X_A^{pg})$ and $\gamma_2(X_A^{pg})$ for different values of p_A. The calculations were performed using *Mathematica*. For example when $p_A = 0.60$, $\mu(X_A^{pg}) = 6.48$ and $\sigma(X_A^{pg}) = 2.59$.

4.4 Forward recursion

Let $X_A^{pg}(a,b)$ be a random variable of the total number of points played in a game at point score (a,b) for player A serving. Let $f(X_A^{pg}(a,b) = x_A^{pg}(a,b))$ represent the distribution of the total number of points played in the game at point score (a,b) for player A serving.

Using forward recursion notation, the distribution of the total number of points played in a game at point score (a,b) for player A serving is represented by:

$f(X_A^{pg}(a,b) = 4) = N_A^{pg}(4,0|a,b) + N_A^{pg}(0,4|a,b)$
$f(X_A^{pg}(a,b) = 5) = N_A^{pg}(4,1|a,b) + N_A^{pg}(1,4|a,b)$
$f(X_A^{pg}(a,b) = 6) = N_A^{pg}(4,2|a,b) + N_A^{pg}(2,4|a,b)$
$f(X_A^{pg}(a,b) = x_A^{pg}(a,b)) = N_A^{pg}(\frac{x_A^{pg}(a,b)+2}{2}, \frac{x_A^{pg}(a,b)-2}{2}|a,b) + N_A^{pg}(\frac{x_A^{pg}(a,b)-2}{2}, \frac{x_A^{pg}(a,b)+2}{2}|a,b)$, for $x_A^{pg}(a,b) = 8, 10, 12, ...$

Let $Y_A^{pg}(a,b)$ be a random variable of the number of points remaining in a game at point score (a,b) for player A serving. Let $f(Y_A^{pg}(a,b) = y_A^{pg}(a,b))$ represent the distribution of the number of points remaining in the game at point score (a,b) for player A serving.

$f(Y_A^{pg}(a,b) = 4 - a - b) = N_A^{pg}(4,0|a,b) + N_A^{pg}(0,4|a,b)$
$f(Y_A^{pg}(a,b) = 5 - a - b) = N_A^{pg}(4,1|a,b) + N_A^{pg}(1,4|a,b)$
$f(Y_A^{pg}(a,b) = 6 - a - b) = N_A^{pg}(4,2|a,b) + N_A^{pg}(2,4|a,b)$
$f(Y_A^{pg}(a,b) = y_A^{pg}(a,b) - a - b) = N_A^{pg}(\frac{y_A^{pg}(a,b)+2}{2}, \frac{y_A^{pg}(a,b)-2}{2}|a,b) + N_A^{pg}(\frac{y_A^{pg}(a,b)-2}{2}, \frac{y_A^{pg}(a,b)+2}{2}|a,b)$, for $y_A^{pg}(a,b) = 8, 10, 12, ...$

Note the relation $f(X_A^{pg}(0,0) = x_A^{pg}(0,0)) = f(Y_A^{pg}(0,0) = y_A^{pg}(0,0))$, for all $x_A^{pg}(0,0) = y_A^{pg}(0,0)$. This implies that from the outset the distribution of the total number of points played in the game is equivalent to the distribution of the number of points remaining in the game.

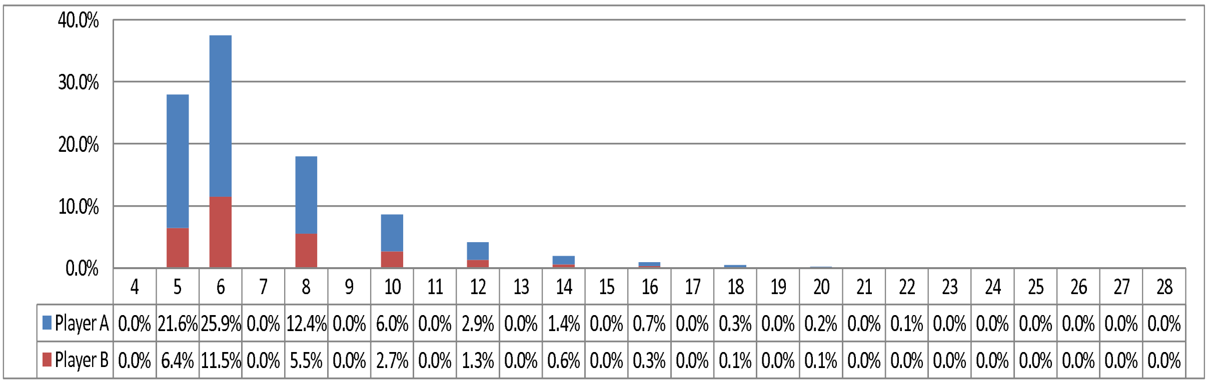


Figure 4.2: The distribution of the total number of points played in a game from $a = 1, b = 1$ for player A serving with $p_A = 0.6$

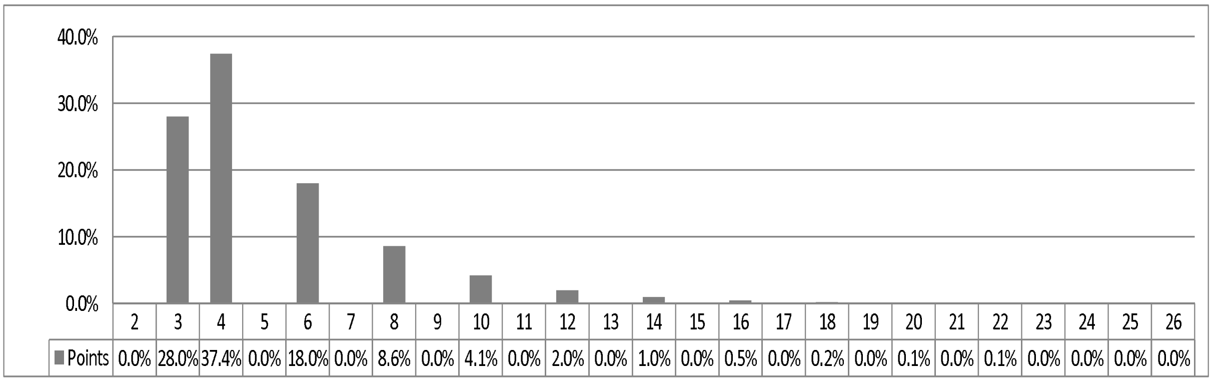


Figure 4.3: The distribution of the number of points remaining in a game from $a = 1, b = 1$ for player A serving with $p_A = 0.6$

Figures 4.2 and 4.3 represent the distributions of the total number of points played in a game and the number of points remaining in a game respectively from $a = 1, b = 1$ for player A serving with $p_A = 0.6$. Note that the shape of both distributions are the same. In other words the variance, and coefficients of skewness and excess kurtosis remain unchanged by adding a constant to all values of the variable. This is widely known as an invariant property in variance such that $V(X + a) = V(X)$. This result will be proven more formally at the conclusion of this chapter.

4.5 Backward recursion

Let $\mu(Y_A^{pg}(a,b))$ represent the mean number of points remaining in the game at point score (a,b) for player A serving. The following derivation is used to obtain the recurrence formula.

Consider the random variable $X_A^{pg}(a,b)$ of the total number of points played in a game at point score (a,b) for player A serving. If the game has not reached its completion then the next point must be contested, and one of two results must occur. If player A wins the point then the score progresses to $(a+1,b)$; otherwise the score progresses to $(a,b+1)$. It follows that:

$X_A^{pg}(a,b) = X_A^{pg}(a+1,b)$ with probability p_A, and

$X_A^{pg}(a,b) = X_A^{pg}(a,b+1)$ with probability q_A.

Taking expectations we obtain a backwards recurrence formula

$E(X_A^{pg}(a,b)) = p_A E(X_A^{pg}(a+1,b)) + q_A E(X_A^{pg}(a,b+1))$

The mixture law above applies at each and every score before the game is completed.

Consider the random variable $Y_A^{pg}(a,b)$ of the number of points remaining in a game at point score (a,b) for player A serving. Then

$X_A^{pg}(a,b) = a + b + Y_A^{pg}(a,b)$ for all $a \geq 0, b \geq 0$

As the score progresses we now have

$Y_A^{pg}(a,b) = 1 + Y_A^{pg}(a+1,b)$ with probability p_A, and

$Y_A^{pg}(a,b) = 1 + Y_A^{pg}(a,b+1)$ with probability q_A.

Taking expectations we obtain a backwards recurrence formula

$E(Y_A^{pg}(a,b)) = 1 + p_A E(Y_A^{pg}(a+1,b)) + q_A E(Y_A^{pg}(a,b+1))$

The mixture law above applies at each and every score before the game is completed.

Therefore the recurrence formula for the mean number of points remaining in a game at point score (a,b) for player A serving is:

$\mu(Y_A^{pg}(a,b)) = 1 + p_A\mu(Y_A^{pg}(a+1,b)) + q_A\mu(Y_A^{pg}(a,b+1))$

The boundary value $\mu(Y_A^{pg}(3,3))$ is obtained as follows.

Let $M_{Y_A^{pg}(a,b)}(t)$ represent the moment generating function for the number of points remaining in a game from point score (a,b) with player A serving. Therefore: $M_{Y_A^{pg}(3,3)}(t) = \frac{(p_A^2+q_A^2)e^{2t}}{1-2p_Aq_Ae^{2t}}$

The first moment $E(Y_A^{pg}(a,b))$ is obtained as $M^{(1)}_{Y_A^{pg}(3,3)}(0) = \frac{2(p_A^2+q_A^2)}{(1-2p_Aq_A)^2} = \frac{2}{p_A^2+q_A^2}$. Therefore $\mu(Y_A^{pg}(3,3)) = E(Y_A^{pg}(a,b)) = \frac{2}{p_A^2+q_A^2}$

Therefore the boundary values are obtained as

$\mu(Y_A^{pg}(a,b)) = 0$, if $b=4$ and $a \leq 2$; $a=4$ and $b \leq 2$

$\mu(Y_A^{pg}(3,3)) = \frac{2}{p_A^2+q_A^2}$

Table 4.2 lists the mean number of points remaining in a game from point score (a,b) for player A serving with $p_A = 0.6$. It indicates that the mean number of points remaining in such a game is 6.5.

		B score				
		0	15	30	40	game
	0	6.5	6.0	4.8	2.8	0
	15	5.2	5.0	4.5	3.0	0
A score	30	3.6	3.7	3.8	3.3	0
	40	1.8	2.0	2.5	3.8	
	game	0	0	0		

Table 4.2: The mean number of points remaining in a game from various score lines for player A serving with $p_A = 0.6$

The following analysis is used to obtain $\sigma^2(Y_A^{pg}(a,b))$, the variance of the number of points remaining in the game at point score (a,b) for player A serving.

Clarke and Norman[1] used recurrence relations to calculate probabilities of winning, mean and variance of lengths to squash. In particular they showed for a random variable Z which takes the value Z_1 with probability Π and the value Z_2 with probability $1-\Pi$, that

$E(Z) = \Pi E(Z_1) + (1-\Pi)E(Z_2)$

$\sigma^2(Z) = \Pi\sigma^2(Z_1) + (1-\Pi)\sigma^2(Z_2) + \Pi(1-\Pi)(E(Z_1) - E(Z_2))^2$

Since the equation representing $E(Z)$ is in the same format as $E(X_A^{pg}(a,b))$, then it follows that $\sigma^2(X_A^{pg}(a,b))$, the variance of the total number of points played in the game at point score (a,b) for player A serving is given by:

$$\sigma^2(X_A^{pg}(a,b))$$
$$= p_A\sigma^2(X_A^{pg}(a+1,b)) + q_A\sigma^2(X_A^{pg}(a,b+1)) + p_Aq_A(\mu(X_A^{pg}(a+1,b)) - \mu(X_A^{pg}(a,b+1)))^2$$

Let $\mu(X_A^{pg}(a,b))$ represent the mean of the total number of points played in a game at point score (a,b) for player A serving.

Given $X_A^{pg}(a,b) = a + b + Y_A^{pg}(a,b)$, with a and b fixed, it follows that:

$$\mu(X_A^{pg}(a,b))$$
$$= \mu(a + b + Y_A^{pg}(a,b))$$
$$= \mu(Y_A^{pg}(a,b)) + a + b$$

Let $\sigma^2(X_A^{pg}(a,b))$ represent the variance of the total number of points played in a game at point score (a,b) for player A serving.

Since $E(X_A^{pg}(a,b)) = \mu(X_A^{pg}(a,b))$, it follows that:

$E(X_A^{pg}(a,b)) = a + b + E(Y_A^{pg}(a,b))$

[1] Clarke, S and Norman, J. Comparison of North American and international squash scoring systems: analytical results, Research Quarterly 50(4) (1979), 723-728.

$E(X_A^{2_{pg}}(a,b)) = (a+b)^2 + 2(a+b)E(Y_A^{pg}(a,b)) + E(Y_A^{2_{pg}}(a,b))$

$\sigma^2(X_A^{pg}(a,b)) = E(X_A^{2_{pg}}(a,b)) - E(X_A^{pg}(a,b))^2 = \sigma^2(Y_A^{pg}(a,b))$

Therefore:

$\sigma^2(Y_A^{pg}(a,b))$

$= p_A\sigma^2(Y_A^{pg}(a+1,b)) + q_A\sigma^2(Y_A^{pg}(a,b+1)) + p_Aq_A(\mu(Y_A^{pg}(a+1,b)) - \mu(Y_A^{pg}(a,b+1)))^2$

The boundary value $\sigma^2(Y_A^{pg}(3,3))$ is obtained as follows.

Using the analysis above to obtain the moment generating function for the number of points remaining in a game from point score (a,b) with player A serving, the second moment $E(Y_A^{2_{pg}}(a,b))$ is obtained as $M^{(2)}_{Y_A^{pg}(3,3)}(0) = \frac{4(1+2p_Aq_A)}{(p_A^2+q_A^2)^2}$.

Therefore $\sigma^2(Y_A^{pg}(3,3)) = E(Y_A^{2_{pg}}(a,b)) - E(Y_A^{pg}(a,b))^2 = \frac{8p_Aq_A}{(p_A^2+q_A^2)^2}$

Therefore the boundary values are obtained as

$\sigma^2(Y_A^{pg}(a,b)) = 0$, if $b=4$ and $a \leq 2$; $a=4$ and $b \leq 2$

$\sigma^2(Y_A^{pg}(3,3)) = \frac{8p_Aq_A}{(p_A^2+q_A^2)^2}$

Table 4.3 lists the variance of the number of points remaining in a game from point score (a,b) for player A serving with $p_A = 0.6$. It indicates that the variance of the number of points remaining in such a game is 6.7.

		B score				
		0	15	30	40	game
	0	6.7	7.2	7.7	6.5	0
	15	6.2	6.7	7.4	7.3	0
A score	30	4.9	6.1	7.1	7.8	0
	40	2.6	4.1	6.4	7.1	
	game	0	0	0		

Table 4.3: The variance of the number of points remaining in a game from various score lines for player A serving with $p_A = 0.6$

Let $\gamma_1(Y_A^{pg}(a,b))$ and $\gamma_2(Y_A^{pg}(a,b))$ represent the coefficients of skewness and excess kurtosis of the number of points remaining in a game at point score (a,b) for player A serving respectively. Let $\gamma_1(X_A^{pg}(a,b))$ and $\gamma_2(X_A^{pg}(a,b))$ represent the coefficients of skewness and excess kurtosis of the total number of points played in a game at point score (a,b) for player A serving respectively. Using similar derivation to above, it follows that:

$\gamma_1(X_A^{pg}(a,b)) = \gamma_1(Y_A^{pg}(a,b))$

$\gamma_2(X_A^{pg}(a,b)) = \gamma_2(Y_A^{pg}(a,b))$

Chapter 5

Duration of a match: 1/2/

5.1 Introduction

Chapter 5 extends on chapter 4 (in section 4.4) by obtaining distributions of the total number of points played and the number of points remaining from any point score within a tiebreak game, distributions of the total number of games played and the number of games remaining from any game score within a tiebreak and advantage set, and distributions of the total number of sets played and the number of sets remaining from any set score within an all tiebreak set and final set advantage match. Chapter 5 also extends on chapter 4 (in section 4.5) by obtaining the mean and variance of the number of points remaining in a tiebreak game from any point score within the game, the mean and variance of the number of games remaining in a tiebreak and advantage set from any point and game score within the set, and the mean and variance of the number of sets remaining in an all tiebreak set and final set advantage match from any point, game and set score within the match.

Given that chapters 5-7 obtain calculations on the duration of a match using backward recursion, it therefore becomes necessary to distinguish the differences between each chapter by introducing the notation $n_1/n_2/$ such that n_1 represents the number of levels of nesting and n_2 represents the number of parameters of distribution. In particular $n_1 = 1$ (chapters 5 and 6) applies to obtaining calculations on the number of points in a game, the number of games in a set and the number of sets in a match; $n_1 = 4$ (chapter 7) applies to obtaining calculations on the time duration in a match; $n_2 = 2$ (chapter 5) applies to obtaining calculations on the mean and variance; and $n_2 = 4$ (chapters 6 and 7) applies to obtaining calculations on the mean, variance, skewness and kurtosis. Hence calculations on the duration of a match are obtained in chapter 5 given $n_1 = 1, n_2 = 2$ (1/2/), in chapter 6 given $n_1 = 1, n_2 = 4$ (1/4/), and chapter 7 given $n_1 = 4, n_2 = 4$ (4/4/).

5.2 Number of points in a tiebreak game

Let $X_A^{pgT}(a,b)$ be a random variable of the total number of points played in a tiebreak game at point score (a,b) for player A currently serving. Let $f(X_A^{pgT}(a,b) = x_A^{pgT}(a,b))$ represent the distribution of the total number of points played in a tiebreak game at point score (a,b) for player A currently serving.

$f(X_A^{pgT}(a,b) = 7) = N_A^{pgT}(7,0|a,b) + N_A^{pgT}(0,7|a,b)$
$f(X_A^{pgT}(a,b) = 8) = N_A^{pgT}(7,1|a,b) + N_A^{pgT}(1,7|a,b)$
$f(X_A^{pgT}(a,b) = 9) = N_A^{pgT}(7,2|a,b) + N_A^{pgT}(2,7|a,b)$
$f(X_A^{pgT}(a,b) = 10) = N_A^{pgT}(7,3|a,b) + N_A^{pgT}(3,7|a,b)$
$f(X_A^{pgT}(a,b) = 11) = N_A^{pgT}(7,4|a,b) + N_A^{pgT}(4,7|a,b)$
$f(X_A^{pgT}(a,b) = 12) = N_A^{pgT}(7,5|a,b) + N_A^{pgT}(5,7|a,b)$
$f(X_A^{pgT}(a,b) = x_A^{pgT}(a,b)) = N_A^{pgT}(\frac{x_A^{pgT}(a,b)+2}{2}, \frac{x_A^{pgT}(a,b)-2}{2}|a,b) + N_A^{pgT}(\frac{x_A^{pgT}(a,b)-2}{2}, \frac{x_A^{pgT}(a,b)+2}{2}|a,b)$,
for $x_A^{pgT}(a,b) = 14, 16, 18, ...$

Figure 5.1 represents the distribution graphically of the total number of points played in a tiebreak game from the outset for player A currently serving with $p_A = 0.62$ and $p_B = 0.60$.

Let $Y_A^{pgT}(a,b)$ be a random variable of the number of points remaining in a tiebreak game at point score (a,b) for player A currently serving. Let $f(Y_A^{pgT}(a,b) = y_A^{pgT}(a,b))$ represent

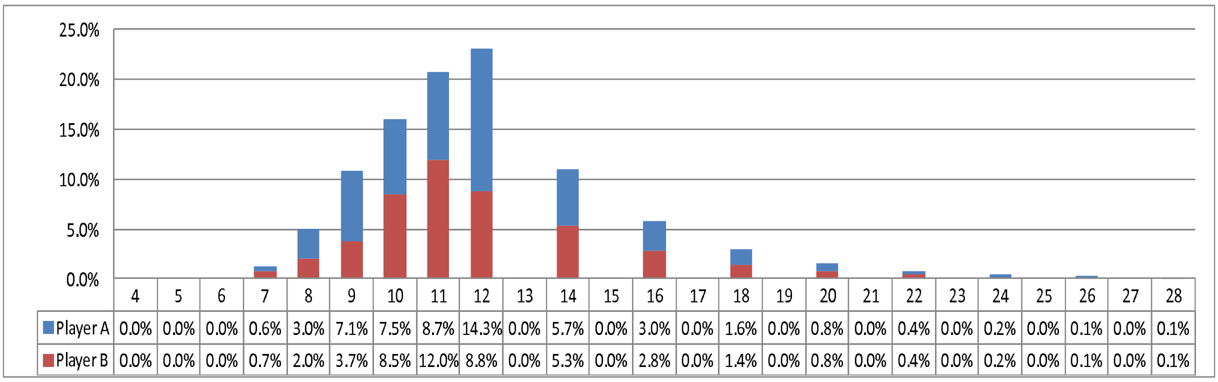


Figure 5.1: The distribution of the total number of points played in a tiebreak game from the outset for player A currently serving with $p_A = 0.62$ and $p_B = 0.60$

the distribution of the number of points remaining in a tiebreak game at point score (a, b) for player A currently serving.

$$f(Y_A^{pgT}(a,b) = 7 - a - b) = N_A^{pgT}(7,0|a,b) + N_A^{pgT}(0,7|a,b)$$
$$f(Y_A^{pgT}(a,b) = 8 - a - b) = N_A^{pgT}(7,1|a,b) + N_A^{pgT}(1,7|a,b)$$
$$f(Y_A^{pgT}(a,b) = 9 - a - b) = N_A^{pgT}(7,2|a,b) + N_A^{pgT}(2,7|a,b)$$
$$f(Y_A^{pgT}(a,b) = 10 - a - b) = N_A^{pgT}(7,3|a,b) + N_A^{pgT}(3,7|a,b)$$
$$f(Y_A^{pgT}(a,b) = 11 - a - b) = N_A^{pgT}(7,4|a,b) + N_A^{pgT}(4,7|a,b)$$
$$f(Y_A^{pgT}(a,b) = 12 - a - b) = N_A^{pgT}(7,5|a,b) + N_A^{pgT}(5,7|a,b)$$
$$f(Y_A^{pgT}(a,b) = y_A^{pgT}(a,b) - a - b) = N_A^{pgT}\left(\tfrac{y_A^{pgT}(a,b)+2}{2}, \tfrac{y_A^{pgT}(a,b)-2}{2}|a,b\right) + N_A^{pgT}\left(\tfrac{y_A^{pgT}(a,b)-2}{2}, \tfrac{y_A^{pgT}(a,b)+2}{2}|a,b\right),$$
for $y_A^{pgT}(a,b) = 14, 16, 18, \ldots$.

Let $\mu(Y_A^{pgT}(a,b))$ represent the mean number of points remaining in a tiebreak game from point score (a, b) with player A currently serving.

Recurrence Formulas:

$$\mu(Y_A^{pgT}(a,b)) = 1 + p_A \mu(Y_B^{pgT}(a+1,b)) + q_A \mu(Y_B^{pgT}(a,b+1)), \text{ if } (a+b) \text{ is even}$$
$$\mu(Y_A^{pgT}(a,b)) = 1 + p_A \mu(Y_A^{pgT}(a+1,b)) + q_A \mu(Y_A^{pgT}(a,b+1)), \text{ if } (a+b) \text{ is odd}$$

Boundary Values:

$$\mu(Y_A^{pgT}(a,b)) = 0, \text{ if } a = 7 \text{ and } 0 \le b \le 5;\ b = 7 \text{ and } 0 \le a \le 5$$

$\mu(Y_A^{pg_T}(6,6)) = \frac{2}{p_A q_B + q_A p_B}$

Table 5.1 represents the mean number of points remaining in a tiebreak game from various score lines with player A currently serving with $p_A = 0.62$ and $p_B = 0.60$. It indicates that the mean number of points remaining from the outset in such a game is 11.9.

		B score							
		0	1	2	3	4	5	6	7
	0	11.9	11.0	9.6	8.2	6.2	4.4	2.1	0
	1	10.7	10.1	9.2	7.9	6.4	4.2	2.4	0
	2	9.4	9.0	8.5	7.6	6.1	4.5	2.2	0
A score	3	7.6	7.7	7.3	6.9	6.0	4.3	2.7	0
	4	6.0	5.8	5.9	5.7	5.4	4.6	2.7	0
	5	3.7	4.1	3.9	4.3	4.0	4.2	3.6	0
	6	1.9	1.7	2.0	1.9	2.3	2.6	4.2	
	7	0	0	0	0	0	0		

Table 5.1: The mean number of points remaining in a tiebreak game from various score lines with player A currently serving with $p_A = 0.62$ and $p_B = 0.60$

Let $\sigma^2(Y_A^{pg_T}(a,b))$ represent the variance of the number of points remaining in a tiebreak game from point score (a,b) with player A serving.

Recurrence Formulas:

$\sigma^2(Y_A^{pg_T}(a,b)) = p_A\sigma^2(Y_B^{pg_T}(a+1,b)) + q_A\sigma^2(Y_B^{pg_T}(a,b+1)) + p_Aq_A(\mu(Y_B^{pg_T}(a+1,b)) - \mu(Y_B^{pg_T}(a,b+1)))^2$, if $(a+b)$ is even

$\sigma^2(Y_A^{pg_T}(a,b)) = p_A\sigma^2(Y_A^{pg_T}(a+1,b)) + q_A\sigma^2(Y_A^{pg_T}(a,b+1)) + p_Aq_A(\mu(Y_A^{pg_T}(a+1,b)) - \mu(Y_A^{pg_T}(a,b+1)))^2$, if $(a+b)$ is odd

Boundary Values:

$\sigma^2(Y_A^{pg_T}(a,b)) = 0$, if $a = 7$ and $0 \leq b \leq 5$; $b = 7$ and $0 \leq a \leq 5$

$\sigma^2(Y_A^{pg_T}(6,6)) = \frac{4(p_Ap_B + q_Aq_B)}{(p_Aq_B + q_Ap_B)^2}$

Table 5.2 represents the variance of the number of points remaining in a tiebreak game from various score lines with player A currently serving with $p_A = 0.62$ and $p_B = 0.60$. It indicates that the variance of the number of points remaining from the outset in such a game is 9.2.

		B score							
		0	1	2	3	4	5	6	7
A score	0	9.2	9.1	9.4	9.2	8.0	5.9	2.4	0
	1	9.1	8.9	8.8	8.8	8.7	6.4	3.4	0
	2	9.1	8.9	8.7	8.7	8.6	7.4	3.6	0
	3	8.6	8.8	8.5	8.6	8.8	8.1	5.8	0
	4	7.3	7.3	8.1	8.3	8.7	8.8	6.6	0
	5	4.7	5.5	6.3	7.5	8.5	9.3	9.9	0
	6	2.4	2.2	3.5	3.8	6.6	7.7	9.3	
	7	0	0	0	0	0	0		

Table 5.2: The variance of the number of points remaining in a tiebreak game from various score lines with player A currently serving with $p_A = 0.62$ and $p_B = 0.60$

5.3 Number of games in a tiebreak set

Let $X_A^{gs_T}(a, b : c, d)$ be a random variable of the total number of games played in a tiebreak set at point and game score $(a, b : c, d)$ for player A currently serving in the set. Let $f(X_A^{gs_T}(a, b : c, d) = x_A^{gs_T}(c, d))$ represent the distribution of the total number of games played in a tiebreak set at point and game score $(a, b : c, d)$ for player A currently serving in the set.

$$f(X_A^{gs_T}(a, b : c, d) = 6) = N_A^{gs_T}(6, 0|a, b : c, d) + N_A^{gs_T}(0, 6|a, b : c, d)$$

$$f(X_A^{gs_T}(a, b : c, d) = 7) = N_A^{gs_T}(6, 1|a, b : c, d) + N_A^{gs_T}(1, 6|a, b : c, d)$$

$$f(X_A^{gs_T}(a, b : c, d) = 8) = N_A^{gs_T}(6, 2|a, b : c, d) + N_A^{gs_T}(2, 6|a, b : c, d)$$

$$f(X_A^{gs_T}(a, b : c, d) = 9) = N_A^{gs_T}(6, 3|a, b : c, d) + N_A^{gs_T}(3, 6|a, b : c, d)$$

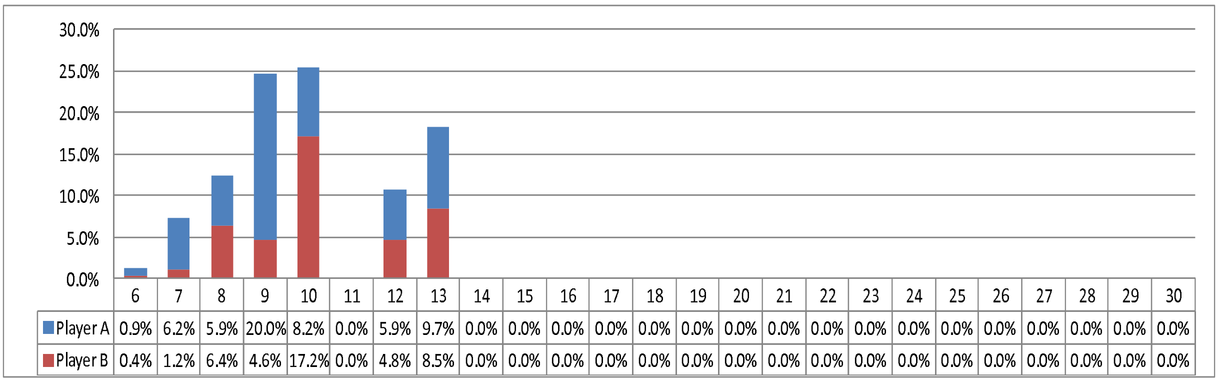


Figure 5.2: The distribution of the total number of games played in a tiebreak set from the outset for player A serving with $p_A = 0.62$ and $p_B = 0.60$

$$f(X_A^{gsT}(a,b:c,d) = 10) = N_A^{gsT}(6,4|a,b:c,d) + N_A^{gsT}(4,6|a,b:c,d)$$

$$f(X_A^{gsT}(a,b:c,d) = 12) = N_A^{gsT}(7,5|a,b:c,d) + N_A^{gsT}(5,7|a,b:c,d)$$

$$f(X_A^{gsT}(a,b:c,d) = 13) = N_A^{gsT}(7,6|a,b:c,d) + N_A^{gsT}(6,7|a,b:c,d)$$

Figure 5.2 represents the distribution graphically of the total number of games played in a tiebreak set from the outset for player A serving with $p_A = 0.62$ and $p_B = 0.60$.

Let $Y_A^{gsT}(a,b:c,d)$ be a random variable of the number of games remaining in a tiebreak set at point and game score $(a,b:c,d)$ for player A currently serving in the set. Let $f(Y_A^{gsT}(a,b:c,d) = y_A^{gsT}(c,d))$ represent the distribution of the number of games remaining in a tiebreak set at point and game score $(a,b:c,d)$ for player A currently serving.

$$f(Y_A^{gsT}(a,b:c,d) = 6-c-d) = N_A^{gsT}(6,0|a,b:c,d) + N_A^{gsT}(0,6|a,b:c,d)$$

$$f(Y_A^{gsT}(a,b:c,d) = 7-c-d) = N_A^{gsT}(6,1|a,b:c,d) + N_A^{gsT}(1,6|a,b:c,d)$$

$$f(Y_A^{gsT}(a,b:c,d) = 8-c-d) = N_A^{gsT}(6,2|a,b:c,d) + N_A^{gsT}(2,6|a,b:c,d)$$

$$f(Y_A^{gsT}(a,b:c,d) = 9-c-d) = N_A^{gsT}(6,3|a,b:c,d) + N_A^{gsT}(3,6|a,b:c,d)$$

$$f(Y_A^{gsT}(a,b:c,d) = 10-c-d) = N_A^{gsT}(6,4|a,b:c,d) + N_A^{gsT}(4,6|a,b:c,d)$$

$$f(Y_A^{gsT}(a,b:c,d) = 12-c-d) = N_A^{gsT}(7,5|a,b:c,d) + N_A^{gsT}(5,7|a,b:c,d)$$

$$f(Y_A^{gsT}(a,b:c,d) = 13-c-d) = N_A^{gsT}(7,6|a,b:c,d) + N_A^{gsT}(6,7|a,b:c,d)$$

Let $\mu(Y_A^{gs_T}(c,d))$ represent the mean number of games remaining in a tiebreak set from game score (c,d) for player A serving.

Recurrence formula:

$\mu(Y_A^{gs_T}(c,d)) = 1 + p_A^g \mu(Y_B^{gs_T}(c+1,d)) + q_A^g \mu(Y_B^{gs_T}(c,d+1))$

Boundary Values:

$\mu(Y_A^{gs_T}(c,d)) = 0$, if $c = 6$ and $0 \leq d \leq 4$; $d = 6$ and $0 \leq c \leq 4$; $(7,5)$; $(5,7)$

$\mu(Y_A^{gs_T}(6,6)) = 1$

Table 5.3 represents the mean number of games remaining in a tiebreak set from various score lines with player A currently serving with $p_A = 0.62$ and $p_B = 0.60$. It indicates that the mean number of games remaining from the outset in such a set is 10.0.

		B score							
		0	1	2	3	4	5	6	7
A score	0	10.0	9.2	7.9	6.3	4.3	2.3	0	
	1	8.7	8.3	7.4	6.1	4.4	2.3	0	
	2	7.2	7.0	6.6	5.8	4.2	2.4	0	
	3	5.4	5.4	5.2	5.0	4.2	2.3	0	
	4	3.5	3.5	3.6	3.5	3.7	3.0	0	
	5	1.5	1.5	1.5	1.7	1.6	2.6	1.8	0
	6	0	0	0	0	0	1.2	1	
	7						0		

Table 5.3: The mean number of games remaining in a tiebreak set from various score lines with player A currently serving with $p_A = 0.62$ and $p_B = 0.60$

Let $\mu(Y_A^{gs_T}(a,b:c,d))$ represent the mean number of games remaining in a tiebreak set from point and game score $(a,b:c,d)$ for player A serving.

$\mu(Y_A^{gs_T}(a,b:c,d)) = 1 + P_A^{pg}(a,b)\mu(Y_B^{gs_T}(c+1,d)) + (1 - P_A^{pg}(a,b))\mu(Y_B^{gs_T}(c,d+1))$, if $(c,d) \neq (6,6)$

$\mu(Y_A^{gs_T}(a,b:c,d)) = 1$, if $(c,d) = (6,6)$

Let $\sigma^2(Y_A^{gs_T}(c,d))$ represent the variance of the number of games remaining in a tiebreak set from game score (c,d) for player A serving.

Recurrence formula:

$\sigma^2(Y_A^{gs_T}(c,d)) = p_A^g\sigma^2(Y_B^{gs_T}(c+1,d)) + q_A^g\sigma^2(Y_B^{gs_T}(c,d+1)) + p_A^g q_A^g(\mu(Y_B^{gs_T}(c+1,d)) - \mu(Y_B^{gs_T}(c,d+1)))^2$

Boundary Values:

$\sigma^2(Y_A^{gs_T}(c,d)) = 0$, if $c = 6$ and $0 \leq d \leq 4$; $d = 6$ and $0 \leq c \leq 4$; $(7,5)$; $(5,7)$

$\sigma^2(Y_A^{gs_T}(6,6)) = 0$

Table 5.4 represents the variance of the number of games remaining in a tiebreak set from various score lines with player A currently serving with $p_A = 0.62$ and $p_B = 0.60$. It indicates that the variance of the number of games remaining from the outset in such a set is 3.7.

		B score							
		0	1	2	3	4	5	6	7
	0	3.7	3.5	3.6	4.2	3.2	1.8	0	
	1	3.7	3.3	3.1	3.2	3.6	1.6	0	
	2	4.1	3.2	2.9	2.8	2.6	2.2	0	
A score	3	3.3	3.6	2.7	2.5	2.4	1.6	0	
	4	2.5	2.1	2.7	1.9	1.8	1.4	0	
	5	1.0	1.3	1.1	1.8	1.3	0.2	0.2	0
	6	0	0	0	0	0	0.2	0	
	7						0		

Table 5.4: The variance of the number of games remaining in a tiebreak set from various score lines with player A currently serving with $p_A = 0.62$ and $p_B = 0.60$

Let $\sigma^2(Y_A^{gs_T}(a,b : c,d))$ represent the variance of the number of games remaining in a tiebreak set from point and game score $(a,b : c,d)$ for player A serving.

$\sigma^2(Y_A^{gs_T}(a,b\ :\ c,d))\ =\ P_A^{pg}(a,b)\sigma^2(Y_B^{gs_T}(c+1,d)) + (1 - P_A^{pg}(a,b))\sigma^2(Y_B^{gs_T}(c,d+1)) + P_A^{pg}(a,b)(1-P_A^{pg}(a,b))(\mu(Y_B^{gs_T}(c+1,d)) - \mu(Y_B^{gs_T}(c,d+1)))^2$, if $(c,d) \neq (6,6)$

$\sigma^2(Y_A^{gs_T}(a,b:c,d)) = 0$, if $(c,d) = (6,6)$

5.4 Number of games in an advantage set

Let $X_A^{gs}(a,b:c,d)$ be a random variable of the total number of games played in an advantage set at point and game score $(a,b\ :\ c,d)$ for player A currently serving in the set. Let $f(X_A^{gs}(a,b:c,d) = x_A^{gs}(c,d))$ represent the distribution of the total number of games played in an advantage set at point and game score $(a,b:c,d)$ for player A currently serving in the set.

$f(X_A^{gs}(a,b:c,d) = 6) = N_A^{gs}(6,0|a,b:c,d) + N_A^{gs}(0,6|a,b:c,d)$

$f(X_A^{gs}(a,b:c,d) = 7) = N_A^{gs}(6,1|a,b:c,d) + N_A^{gs}(1,6|a,b:c,d)$

$f(X_A^{gs}(a,b:c,d) = 8) = N_A^{gs}(6,2|a,b:c,d) + N_A^{gs}(2,6|a,b:c,d)$

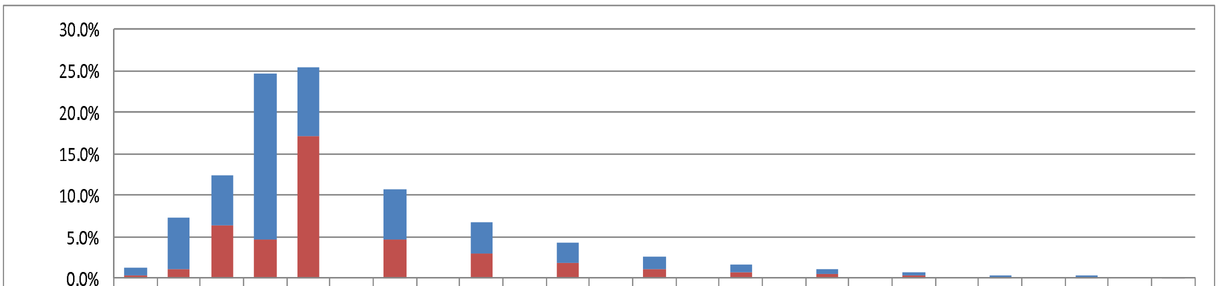
$f(X_A^{gs}(a,b:c,d) = 9) = N_A^{gs}(6,3|a,b:c,d) + N_A^{gs}(3,6|a,b:c,d)$

$f(X_A^{gs}(a,b:c,d) = 10) = N_A^{gs}(6,4|a,b:c,d) + N_A^{gs}(4,6|a,b:c,d)$

$f(X_A^{gs}(a,b:c,d) = x_A^{gs}(c,d)) = N_A^{gs}(\frac{x_A^{gs}(c,d)+2}{2}, \frac{x_A^{gs}(c,d)-2}{2}|a,b:c,d) + N_A^{gs}(\frac{x_A^{gs}(c,d)-2}{2}, \frac{x_A^{gs}(c,d)+2}{2}|a,b:c,d)$, for $x_A^{gs}(c,d) = 12, 14, 16, ...$

Figure 5.3 represents the distribution graphically of the total number of games played in an advantage set from the outset for player A serving with $p_A = 0.62$ and $p_B = 0.60$.

Let $Y_A^{gs}(a,b:c,d)$ be a random variable of the number of games remaining in an advantage set at point and game score $(a,b\ :\ c,d)$ for player A currently serving in the set. Let $f(Y_A^{gs}(a,b:c,d) = y_A^{gs}(c,d))$ represent the distribution of the number of games remaining in an advantage set at point and game score $(a,b:c,d)$ for player A currently serving in the set.



	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Player A	0.9%	6.2%	5.9%	20.0%	8.2%	0.0%	5.9%	0.0%	3.7%	0.0%	2.4%	0.0%	1.5%	0.0%	0.9%	0.0%	0.6%	0.0%	0.4%	0.0%	0.2%	0.0%	0.1%	0.0%	0.1%
Player B	0.4%	1.2%	6.4%	4.6%	17.2%	0.0%	4.8%	0.0%	3.0%	0.0%	1.9%	0.0%	1.2%	0.0%	0.8%	0.0%	0.5%	0.0%	0.3%	0.0%	0.2%	0.0%	0.1%	0.0%	0.1%

Figure 5.3: The distribution of the total number of games played in an advantage set from the outset for player A serving with $p_A = 0.62$ and $p_B = 0.60$

$f(Y_A^{gs}(a, b : c, d) = 6 - c - d) = N_A^{gs}(6, 0|a, b : c, d) + N_A^{gs}(0, 6|a, b : c, d)$

$f(Y_A^{gs}(a, b : c, d) = 7 - c - d) = N_A^{gs}(6, 1|a, b : c, d) + N_A^{gs}(1, 6|a, b : c, d)$

$f(Y_A^{gs}(a, b : c, d) = 8 - c - d) = N_A^{gs}(6, 2|a, b : c, d) + N_A^{gs}(2, 6|a, b : c, d)$

$f(Y_A^{gs}(a, b : c, d) = 9 - c - d) = N_A^{gs}(6, 3|a, b : c, d) + N_A^{gs}(3, 6|a, b : c, d)$

$f(Y_A^{gs}(a, b : c, d) = 10 - c - d) = N_A^{gs}(6, 4|a, b : c, d) + N_A^{gs}(4, 6|a, b : c, d)$

$f(Y_A^{gs}(a, b : c, d) = y_A^{gs}(c, d) - c - d) = N_A^{gs}(\frac{y_A^{gs}(c,d)+2}{2}, \frac{y_A^{gs}(c,d)-2}{2}|a, b : c, d) + N_A^{gs}(\frac{y_A^{gs}(c,d)-2}{2}, \frac{y_A^{gs}(c,d)+2}{2}|a, b : c, d)$, for $y_A^{gs}(c, d) = 12, 14, 16, ...$

Let $\mu(Y_A^{gs}(c, d))$ represent the mean number of games remaining in an advantage set from game score (c, d) for player A serving.

Recurrence formula:

$\mu(Y_A^{gs}(c, d)) = 1 + p_A^g \mu(Y_B^{gs}(c + 1, d)) + q_A^g \mu(Y_B^{gs}(c, d + 1))$

Boundary values:

$\mu(Y_A^{gs}(c, d)) = 0$, if $c = 6$ and $0 \leq d \leq 4$; $d = 6$ and $0 \leq c \leq 4$

$\mu(Y_A^{gs}(5, 5)) = \frac{2}{p_A^g q_B^g + q_A^g p_B^g}$

Table 5.5 represents the mean number of games remaining in an advantage set from various score lines with player A currently serving with $p_A = 0.62$ and $p_B = 0.60$. It indicates that the mean number of games remaining from the outset in such a set is 10.8.

		B score						
		0	1	2	3	4	5	6
A score	0	10.8	10.0	8.6	6.9	4.5	2.4	0
	1	9.4	9.2	8.4	6.8	5.0	2.4	0
	2	7.8	7.7	7.7	6.9	4.9	2.8	0
	3	5.6	6.0	5.9	6.3	5.7	2.9	0
	4	3.7	3.7	4.2	4.3	5.4	5.2	0
	5	1.5	1.6	1.6	2.1	2.2	5.4	
	6	0	0	0	0	0		

Table 5.5: The mean number of games remaining in an advantage set from various score lines with player A currently serving with $p_A = 0.62$ and $p_B = 0.60$

Let $\mu(Y_A^{gs}(a, b : c, d))$ represent the mean number of games remaining in an advantage set from point and game score $(a, b : c, d)$ for player A serving.

$$\mu(Y_A^{gs}(a, b : c, d)) = 1 + P_A^{pg}(a, b)\mu(Y_B^{gs}(c + 1, d)) + (1 - P_A^{pg}(a, b))\mu(Y_B^{gs}(c, d + 1))$$

Let $\sigma^2(Y_A^{gs}(c, d))$ represent the variance of the number of games remaining in an advantage set from game score (c, d) for player A serving.

Recurrence formula:

$$\sigma^2(Y_A^{gs}(c, d)) = p_A^g\sigma^2(Y_B^{gs}(c+1, d)) + q_A^g\sigma^2(Y_B^{gs}(c, d+1)) + p_A^g q_A^g(\mu(Y_B^{gs}(c+1, d)) - \mu(Y_B^{gs}(c, d+1)))^2$$

Boundary Values:

$\sigma^2(Y_A^{gs}(c, d)) = 0$, if $c = 6$ and $0 \leq d \leq 4$; $d = 6$ and $0 \leq c \leq 4$

$\sigma^2(Y_A^{gs}(5, 5)) = \frac{4(p_A^g p_B^g + q_A^g q_B^g)}{(p_A^g q_B^g + q_A^g p_B^g)^2}$

Table 5.6 represents the variance of the number of games remaining in an advantage set from various score lines with player A currently serving with $p_A = 0.62$ and $p_B = 0.60$. It indicates that the variance of the number of games remaining from the outset in such a set is 14.8.

		B score						
		0	1	2	3	4	5	6
	0	14.8	15.1	13.7	13.7	7.1	3.6	0
	1	13.1	15.1	15.6	13.3	12.4	3.7	0
	2	13.1	12.9	15.8	16.5	12.3	9.0	0
A score	3	7.6	12.5	12.5	17.0	18.2	9.2	0
	4	5.2	5.4	11.2	11.5	18.4	19.4	0
	5	1.4	2.8	2.9	8.6	9.2	18.4	
	6	0	0	0	0	0		

Table 5.6: The variance of the number of games remaining in an advantage set from various score lines with player A currently serving with $p_A = 0.62$ and $p_B = 0.60$

Let $\sigma^2(Y_A^{gs}(a, b : c, d))$ represent the variance of the number of games remaining in a tiebreak set from point and game score $(a, b : c, d)$ for player A serving.

$$\sigma^2(Y_A^{gs}(a, b : c, d)) = P_A^{pg}(a, b)\sigma^2(Y_B^{gs}(c+1, d)) + (1 - P_A^{pg}(a, b))\sigma^2(Y_B^{gs}(c, d+1)) + P_A^{pg}(a, b)(1 - P_A^{pg}(a, b))(\mu(Y_B^{gs}(c + 1, d)) - \mu(Y_B^{gs}(c, d + 1)))^2$$

5.5 Number of sets in a match

Let $X^{sm_5}(a, b : c, d : e, f)$ and $X^{sm_{5T}}(a, b : c, d : e, f)$ be random variables of the total number of sets played in a best-of-5 final set advantage and best-of-5 all tiebreak set match respectively at point, game and set score $(a, b : c, d : e, f)$. Let $f(X^{sm_5}(a, b : c, d : e, f) = x^{sm_5}(e, f))$ and $f(X^{sm_{5T}}(a, b : c, d : e, f) = x^{sm_{5T}}(e, f))$ represent the distribution of the total number of sets played in a best-of-5 final set advantage and best-of-5 all tiebreak set match respectively at set score $(a, b : c, d : e, f)$.

$$f(X^{sm_5}(a, b : c, d : e, f) = 3) = N^{sm_5}(3, 0|a, b : c, d : e, f) + N^{sm_5}(0, 3|a, b : c, d : e, f)$$
$$f(X^{sm_5}(a, b : c, d : e, f) = 4) = N^{sm_5}(3, 1|a, b : c, d : e, f) + N^{sm_5}(1, 3|a, b : c, d : e, f)$$
$$f(X^{sm_5}(a, b : c, d : e, f) = 5) = N^{sm_5}(3, 2|a, b : c, d : e, f) + N^{sm_5}(2, 3|a, b : c, d : e, f)$$

$$f(X^{sm_{5T}}(a, b : c, d : e, f) = 3) = N^{sm_{5T}}(3, 0|a, b : c, d : e, f) + N^{sm_{5T}}(0, 3|a, b : c, d : e, f)$$

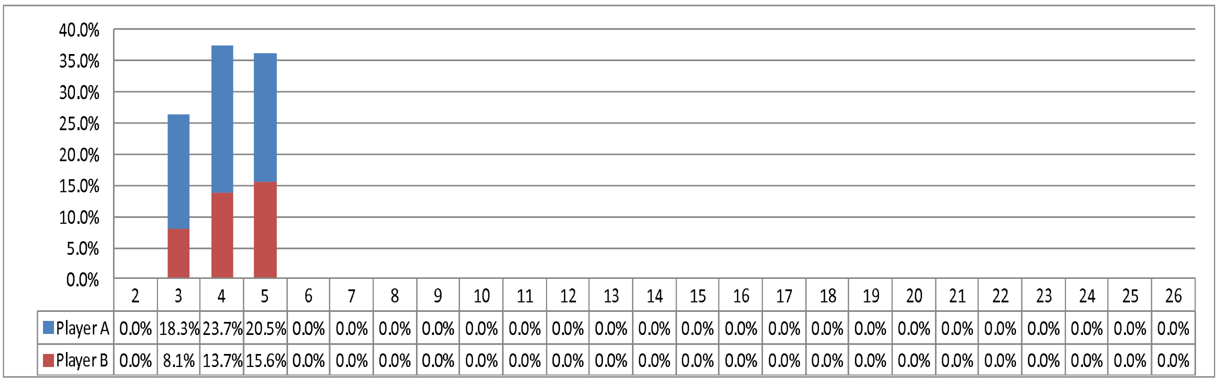


Figure 5.4: The distribution of the total number of sets played in a best-of-5 all tiebreak set match from the outset with $p_A = 0.62$ and $p_B = 0.60$

$$f(X^{sm_{5T}}(a,b:c,d:e,f) = 4) = N^{sm_{5T}}(3,1|a,b:c,d:e,f) + N^{sm_{5T}}(1,3|a,b:c,d:e,f)$$
$$f(X^{sm_{5T}}(a,b:c,d:e,f) = 5) = N^{sm_{5T}}(3,2|a,b:c,d:e,f) + N^{sm_{5T}}(2,3|a,b:c,d:e,f)$$

It can be shown that $f(X^{sm_5}(a,b:c,d:e,f) = x^{sm_5}(e,f)) = f(X^{sm_{5T}}(a,b:c,d:e,f) = x^{sm_{5T}}(e,f))$ for all $x^{sm_5}(e,f) = x^{sm_{5T}}(e,f)$. However the probability of playing 5 sets is comprised of the probability of player A winning in 5 sets and the probability of player B winning in 5 sets. These probabilities can differ depending on whether the final set is advantage of tiebreak. Figure 5.4 represents the distribution graphically of the total number of sets played in a best-of-5 all tiebreak set match from the outset with $p_A = 0.62$ and $p_B = 0.60$.

Let $Y^{sm_5}(a,b:c,d:e,f)$ be a random variable of the number of sets remaining in a best-of-5 final set advantage match at point, game and set score $(a,b:c,d:e,f)$. Let $f(Y^{sm_5}(a,b:c,d:e,f) = y^{sm_5}(e,f))$ represent the distribution of the number of sets remaining in a best-of-5 final set advantage match at point, game and set score $(a,b:c,d:e,f)$.

$$f(Y^{sm_5}(a,b:c,d:e,f) = 3-e-f) = N^{sm_5}(3,0|a,b:c,d:e,f) + N^{sm_5}(0,3|a,b:c,d:e,f)$$
$$f(Y^{sm_5}(a,b:c,d:e,f) = 4-e-f) = N^{sm_5}(3,1|a,b:c,d:e,f) + N^{sm_5}(1,3|a,b:c,d:e,f)$$
$$f(Y^{sm_5}(a,b:c,d:e,f) = 5-e-f) = N^{sm_5}(3,2|a,b:c,d:e,f) + N^{sm_5}(2,3|a,b:c,d:e,f)$$

It can be shown that $f(Y^{sm_5}(a,b:c,d:e,f) = y^{sm_5}(e,f)) = f(Y^{sm_{5T}}(a,b:c,d:e,f) =$

$y^{sm_{5T}}(e,f))$ for all $y^{sm_5}(e,f) = y^{sm_{5T}}(e,f)$.

Let $\mu(Y^{sm_5}(e,f))$ represent the mean number of sets remaining in a best-of-5 final set advantage match from set score (e,f). It follows from above that:
$\mu(Y^{sm_{5T}}(e,f)) = \mu(Y^{sm_5}(e,f))$, for all (e,f)

Therefore $\mu(Y^{sm_5}(e,f))$ represents the mean number of sets remaining in a best-of-5 set match from set score (e,f). This notation will be adopted for the higher order moments and other parameters of distribution for sets played in a match.

Recurrence formula:
$\mu(Y^{sm_5}(e,f)) = 1 + p^{sT}\mu(Y^{sm_5}(e+1,f)) + q^{sT}\mu(Y^{sm_5}(e,f+1))$

Boundary values:
$\mu(Y^{sm_5}(e,f)) = 0$, if $e = 3$ and $f \leq 1$; $f = 3$ and $e \leq 1$
$\mu(Y^{sm_5}(2,2)) = 1$

Table 5.7 represents the mean number of sets remaining in a best-of-5 set match from various score lines with $p_A = 0.62$ and $p_B = 0.60$. It indicates that the mean number of sets remaining from the outset in such a match is 4.1.

		B score			
		0	1	2	3
	0	4.1	3.2	1.9	0
A score	1	3.0	2.5	1.6	0
	2	1.6	1.4	1	
	3	0	0		

Table 5.7: The mean number of sets remaining in a best-of-5 set match from various score lines with $p_A = 0.62$ and $p_B = 0.60$

Let $\mu(Y_A^{sm_5}(a,b:c,d:e,f))$ represent the mean number of sets remaining in a best-of-5 set match from point, game and set score $(a,b:c,d:e,f)$ for player A serving. It follows from above that:

$\mu(Y_A^{sm_5}(a,b:c,d:e,f)) = 1 + P_A^{ps_T}(a,b:c,d)\mu(Y^{sm_5}(e+1,f)) + (1 - P_A^{ps_T}(a,b:c,d))\mu(Y^{sm_5}(e,f+1))$, if $(e,f) \neq (2,2)$

$\mu(Y_A^{sm_5}(a,b:c,d:e,f)) = 1$, if $(e,f) = (2,2)$

Let $\sigma^2(Y^{sm_5}(e,f))$ represent the variance of the number of sets remaining in a best-of-5 set match from set score (e,f).

Recurrence formula:

$\sigma^2(Y^{sm_5}(e,f)) = p^{s_T}\sigma^2(Y^{sm_5}(e+1,f)) + q^{s_T}\sigma^2(Y^{sm_5}(e,f+1)) + p^{s_T}q^{s_T}(\mu(Y^{sm_5}(e+1,f)) - \mu(Y^{sm_5}(e,f+1)))^2$

Boundary values:

$\sigma^2(Y^{sm_5}(e,f)) = 0$, if $e = 3$ and $f \leq 1$; $f = 3$ and $e \leq 1$

$\sigma^2(Y^{sm_5}(2,2)) = 0$

Table 5.8 represents the variance of the number of sets remaining in a best-of-5 set match from various score lines with $p_A = 0.62$ and $p_B = 0.60$. It indicates that the variance of the number of sets remaining from the outset in such a match is 0.62.

Let $\sigma^2(Y_A^{sm_5}(a,b:c,d:e,f))$ represent the variance of the number of sets remaining in a best-of-5 set match from point, game and set score $(a,b:c,d:e,f)$ for player A serving.

$\sigma^2(Y_A^{sm_5}(a,b:c,d:e,f)) = P_A^{ps_T}(a,b:c,d)\sigma^2(Y^{sm_5}(e+1,f)) + (1 - P_A^{ps_T}(a,b:c,d))$
$\sigma^2(Y^{sm_5}(e,f+1)) + P_A^{ps_T}(a,b:c,d)(1 - P_A^{ps_T}(a,b:c,d))(\mu(Y^{sm_5}(e+1,f))$
$-\mu(Y^{sm_5}(e,f+1)))^2$, if $(e,f) \neq (2,2)$

$\sigma^2(Y_A^{sm_5}(a,b:c,d:e,f)) = 0$, if $(e,f) = (2,2)$

		B score			
		0	1	2	3
	0	0.62	0.55	0.74	0
A score	1	0.64	0.25	0.25	0
	2	0.61	0.25	0	
	3	0	0		

Table 5.8: The variance of the number of sets remaining in a best-of-5 set match from various score lines with $p_A = 0.62$ and $p_B = 0.60$

Chapter 6

Duration of a match: 1/4/

6.1 Introduction

Chapter 6 extends on chapter 4 (section 4.5) by obtaining the coefficients of skewness and excess kurtosis of the number of points remaining in a game from any point score within the game. Chapter 6 also extends on chapter 5 by obtaining the coefficients of skewness and excess kurtosis of the number of points remaining in a tiebreak game from any point score within the game, the coefficients of skewness and excess kurtosis of the number of games remaining in a tiebreak and advantage set from any point and game score within the set, and the coefficients of skewness and excess kurtosis of the number of sets remaining in an all tiebreak set and final set advantage match from any point, game and set score within the match.

6.2 Number of points in a game

The following backward recurrence formulas were established in chapter 4 for the random variables $X_A^{pg}(a,b)$ (the total number of points played in a game at point score (a,b) for player A serving) and $Y_A^{pg}(a,b)$ (the number of points remaining in a game at point score (a,b) for player A serving).

$E(X_A^{pg}(a,b)) = p_A E(X_A^{pg}(a+1,b)) + q_A E(X_A^{pg}(a,b+1))$

$E(Y_A^{pg}(a,b)) = 1 + p_A E(Y_A^{pg}(a+1,b)) + q_A E(Y_A^{pg}(a,b+1))$

Let $X_A^{n_{pg}}(a,b)$ represent the n^{th} power of the random variable $X_A^{pg}(a,b)$ for each $n > 0$.

Then $E(X_A^{n_{pg}}(a,b))$ represents the n^{th} moment with the following important relation

$X_A^{n_{pg}}(a,b) = (a + b + Y_A^{n_{pg}}(a,b))$

which, when expanded involves various powers of $Y_A^{pg}(a,b)$. Thus calculation must proceed recursively, i.e. first moment, second moment, and so on. These higher moments can then be used to calculate other statistics such as variance, skewness and kurtosis.

Taking expectations gives the following recurrence formula:

$E(X_A^{n_{pg}}(a,b)) = p_A E(X_A^{n_{pg}}(a+1,b)) + q_A E(X_A^{n_{pg}}(a,b+1))$

The boundary values for $X_A^{n_{pg}}(a,b)$ are obtained as:

$E(X_A^{n_{pg}}(4,0))$ and $E(X_A^{n_{pg}}(0,4)) = 4^n$,

$E(X_A^{n_{pg}}(4,1))$ and $E(X_A^{n_{pg}}(1,4)) = 5^n$,

$E(X_A^{n_{pg}}(4,2))$ and $E(X_A^{n_{pg}}(2,4)) = 6^n$

The boundary values at $E(X_A^{n_{pg}}(3,3))$ are obtained as follows:

The moment generating function for the total number of points played in a game from $(3,3)$ with player A serving is given by:

$M_{X_A^{pg}(3,3)}(t) = \frac{(p_A^2+q_A^2)e^{8t}}{1-2p_Aq_Ae^{2t}}$

Therefore:

$E(X_A^{pg}(3,3)) = M^{(1)}_{X_A^{pg}(3,3)}(0) = \frac{4(3p_Aq_A-2)}{2p_Aq_A-1}$

$E(X_A^{2_{pg}}(3,3)) = M^{(2)}_{X_A^{pg}(3,3)}(0) = \frac{8(18p_A^2q_A^2-23p_Aq_A+8)}{(2p_Aq_A-1)^2}$

$E(X_A^{3_{pg}}(3,3)) = M^{(3)}_{X_A^{pg}(3,3)}(0) = \frac{16(108p_A^3q_A^3-200p_A^2q_A^2+131p_Aq_A-32)}{(2p_Aq_A-1)^3}$

$E(X_A^{4_{pg}}(3,3)) = M^{(4)}_{X_A^{pg}(3,3)}(0) = \frac{32(648p_A^4q_A^4 - 1556p_A^3q_A^3 + 1462p_A^2q_A^2 - 655p_Aq_A + 128)}{(2p_Aq_A - 1)^4}$

The following standard results are used to obtain $\mu(X_A^{pg}(a,b))$, $\sigma^2(X_A^{pg}(a,b))$, $\gamma_1(X_A^{pg}(a,b))$ and $\gamma_2(X_A^{pg}(a,b))$.

$E(X) = \mu(X)$

$E(X^2) = \sigma^2(X) + E(X)^2$

$E(X^3) = \gamma_1(X)\sigma^2(X)^{\frac{3}{2}} + 3E(X^2)E(X) - 2E(X)^3$

$E(X^4) = \gamma_2(X)\sigma^2(X)^2 + 4E(X^3)E(X) + 3E(X^2)^2 - 12E(X^2)E(X)^2 + 6E(X)^4$

Finally, the following results obtained in section 4.5 are used to obtain $\mu(Y_A^{pg}(a,b))$, $\sigma^2(Y_A^{pg}(a,b))$, $\gamma_1(Y_A^{pg}(a,b))$ and $\gamma_2(Y_A^{pg}(a,b))$.

$\mu(Y_A^{pg}(a,b)) = \mu(X_A^{pg}(a,b)) - a - b$

$\sigma^2(Y_A^{pg}(a,b)) = \sigma^2(X_A^{pg}(a,b))$

$\gamma_1(Y_A^{pg}(a,b)) = \gamma_1(X_A^{pg}(a,b))$

$\gamma_2(Y_A^{pg}(a,b)) = \gamma_2(X_A^{pg}(a,b))$

Let $E(X_A^{pg}(a,b))$ represent the first moment of the total number of points played in a game at point score (a,b) for player A serving.

Recurrence formula:

$E(X_A^{pg}(a,b)) = p_AE(X_A^{pg}(a+1,b)) + q_AE(X_A^{pg}(a,b+1))$

Boundary values:

$E(X_A^{pg}(4,0))$ and $E(X_A^{pg}(0,4)) = 4$,

$E(X_A^{pg}(4,1))$ and $E(X_A^{pg}(1,4)) = 5$,

$E(X_A^{pg}(4,2))$ and $E(X_A^{pg}(2,4)) = 6$,

$E(X_A^{pg}(3,3)) = \frac{4(3p_Aq_A-2)}{2p_Aq_A-1}$

Let $\mu(Y_A^{pg}(a,b))$ represent the mean number of points remaining in a game at point score (a,b) for player A serving.

$\mu(Y_A^{pg}(a,b)) = E(X_A^{pg}(a,b)) - a - b$, for $0 \leq a \leq 3$ and $0 \leq b \leq 3$

Let $E(X_A^{2_{pg}}(a,b))$ represent the second moment of the total number of points played in a game at point score (a,b) for player A serving.

Recurrence formula:
$E(X_A^{2_{pg}}(a,b)) = p_A E(X_A^{2_{pg}}(a+1,b)) + q_A E(X_A^{2_{pg}}(a,b+1))$

Boundary values:
$E(X_A^{2_{pg}}(4,0))$ and $E(X_A^{2_{pg}}(0,4)) = 16$,
$E(X_A^{2_{pg}}(4,1))$ and $E(X_A^{2_{pg}}(1,4)) = 25$,
$E(X_A^{2_{pg}}(4,2))$ and $E(X_A^{2_{pg}}(2,4)) = 36$,
$E(X_A^{2_{pg}}(3,3)) = \frac{8(18p_A^2q_A^2-23p_Aq_A+8)}{(2p_Aq_A-1)^2}$

Let $\sigma^2(Y_A^{pg}(a,b))$ represent the variance of the number of points remaining in a game at point score (a,b) for player A serving.

$\sigma^2(Y_A^{pg}(a,b)) = E(X_A^{2_{pg}}(a,b)) - E(X_A^{pg}(a,b))^2$, for $0 \leq a \leq 3$ and $0 \leq b \leq 3$

Let $E(X_A^{3_{pg}}(a,b))$ represent the third moment of the total number of points played in a game at point score (a,b) for player A serving.

Recurrence formula:
$E(X_A^{3_{pg}}(a,b)) = p_A E(X_A^{3_{pg}}(a+1,b)) + q_A E(X_A^{3_{pg}}(a,b+1))$

Boundary values:
$E(X_A^{3_{pg}}(4,0))$ and $E(X_A^{3_{pg}}(0,4)) = 64$,
$E(X_A^{3_{pg}}(4,1))$ and $E(X_A^{3_{pg}}(1,4)) = 125$,

$E(X_A^{3_{pg}}(4,2))$ and $E(X_A^{3_{pg}}(2,4)) = 216$,

$E(X_A^{3_{pg}}(3,3)) = \frac{16(108p_A^3q_A^3-200p_A^2q_A^2+131p_Aq_A-32)}{(2p_Aq_A-1)^3}$

Table 6.1 represents the third moment of the total number of points played in a game at various score lines for player A serving given $p_A = 0.6$.

		B score				
		0	15	30	40	game
	0	441.3	525.7	515.9	346.9	64
	15	385.1	532.2	628.6	535.5	125
A score	30	286.9	468.0	690.6	809.2	216
	40	166.2	319.6	611.5	1204.7	
	game	64	125	216		

Table 6.1: The third moment of the total number of points played in a game at various score lines for player A serving given $p_A = 0.6$

Let $\gamma_1(Y_A^{pg}(a,b))$ represent the coefficient of skewness of the number of points remaining in a game at point score (a,b) for player A serving.

$\gamma_1(Y_A^{pg}(a,b)) = \frac{E(X_A^{3_{pg}}(a,b))-3E(X_A^{2_{pg}}(a,b))E(X_A^{pg}(a,b))+2E(X_A^{pg}(a,b))^3}{\sigma^2(Y_A^{pg}(a,b))^{\frac{3}{2}}}$, for $0 \leq a \leq 3$ and $0 \leq b \leq 3$

Table 6.2 represents the coefficient of skewness of the total number of points played in a game at various score lines for player A serving given $p_A = 0.6$. It shows that the coefficient of skewness from the outset in such a game is 2.2.

		B score			
		0	15	30	40
	0	2.2	2.1	1.9	2.4
A score	15	2.4	2.2	2.0	2.1
	30	2.8	2.5	2.1	1.9
	40	4.1	3.4	2.4	2.1

Table 6.2: The coefficient of skewness of the total number of points played in a game at various score lines for player A serving given $p_A = 0.6$

Let $E(X_A^{4pg}(a,b))$ represent the fourth moment of the total number of points played in a game at point score (a,b) for player A serving.

Recurrence formula:

$E(X_A^{4pg}(a,b)) = p_A E(X_A^{4pg}(a+1,b)) + q_A E(X_A^{4pg}(a,b+1))$

Boundary values:

$E(X_A^{4pg}(4,0))$ and $E(X_A^{4pg}(0,4)) = 256$,

$E(X_A^{4pg}(4,1))$ and $E(X_A^{4pg}(1,4)) = 625$,

$E(X_A^{4pg}(4,2))$ and $E(X_A^{4pg}(2,4)) = 1296$,

$E(X_A^{4pg}(3,3)) = \frac{32(648p_A^4q_A^4 - 1556p_A^3q_A^3 + 1462p_A^2q_A^2 - 655p_Aq_A + 128)}{(2p_Aq_A - 1)^4}$

Table 6.3 represents the fourth moment of the total number of points played in a game at various score lines for player A serving given $p_A = 0.6$.

		B score				
		0	15	30	40	game
A score	0	4909.8	6009.7	5945.3	3809.3	256
	15	4176.6	6052.6	7369.3	6178.1	625
	30	2925.9	5174.9	8163.4	9880.2	1296
	40	1426.6	3182.5	7018.8	15603.0	
	game	256	625	1296		

Table 6.3: The fourth moment of the total number of points played in a game at various score lines for player A serving given $p_A = 0.6$

Let $\gamma_2(Y_A^{pg}(a,b))$ represent the coefficient of excess kurtosis of the number of points remaining in a game at point score (a,b) for player A serving.

$\gamma_2(Y_A^{pg}(a,b))$

$= \frac{E(X_A^{4pg}(a,b)) - 4E(X_A^{3pg}(a,b))E(X_A^{pg}(a,b)) - 3E(X_A^{2pg}(a,b))^2 + 12E(X_A^{2pg}(a,b))E(X_A^{pg}(a,b))^2 - 6E(X_A^{pg}(a,b))^4}{\sigma^2(Y_A^{pg}(a,b)^2}$,

for $0 \le a \le 3$ and $0 \le b \le 3$

Table 6.4 represents the coefficient of excess kurtosis of the total number of points played in a game at various score lines for player A serving given $p_A = 0.6$. It shows that the coefficient of excess kurtosis from the outset in such a game is 7.2.

		B score			
		0	15	30	40
	0	7.2	6.5	5.6	7.8
A score	15	8.3	7.2	6.1	6.3
	30	11.6	8.6	6.6	5.4
	40	25.4	15.5	8.0	6.6

Table 6.4: The coefficient of excess kurtosis of the total number of points played in a game at various score lines for player A serving given $p_A = 0.6$

6.3 Number of points in a tiebreak game

The moment generating function for the total number of points played in a tiebreak game from $(6,6)$ with player A serving is given by $M_{X_A^{pgT}(6,6)}(t) = \frac{(1-p_Ap_B-q_Aq_B)e^{14t}}{1-(p_Ap_B+q_Aq_B)e^{2t}}$

Let $E(X_A^{pgT}(a,b))$ represent the first moment of the total number of points played in a tiebreak game at point score (a,b) for player A serving.

Recurrence formulas:

$E(X_A^{pgT}(a,b)) = p_A E(X_B^{pgT}(a+1,b)) + q_A E(X_B^{pgT}(a,b+1))$, if $(a+b)$ is even

$E(X_A^{pgT}(a,b)) = p_A E(X_A^{pgT}(a+1,b)) + q_A E(X_A^{pgT}(a,b+1))$, if $(a+b)$ is odd

Boundary values:

$E(X_A^{pgT}(7,0))$ and $E(X_A^{pgT}(0,7)) = 7$,

$E(X_A^{pgT}(7,1))$ and $E(X_A^{pgT}(1,7)) = 8$,

$E(X_A^{pgT}(7,2))$ and $E(X_A^{pgT}(2,7)) = 9$,

$E(X_A^{pgT}(7,3))$ and $E(X_A^{pgT}(3,7)) = 10$,

$E(X_A^{pgT}(7,4))$ and $E(X_A^{pgT}(4,7)) = 11$,

$E(X_A^{pgT}(7,5))$ and $E(X_A^{pgT}(5,7)) = 12$,

$E(X_A^{pgT}(6,6)) = \frac{2(6q_Aq_B+6p_Ap_B-7)}{q_Aq_B+p_Ap_B-1}$

Let $\mu(Y_A^{pgT}(a,b))$ represent the mean number of points remaining in a tiebreak game at point score (a,b) for player A serving.

$\mu(Y_A^{pgT}(a,b)) = E(X_A^{pgT}(a,b)) - a - b$, for $0 \le a \le 6$ and $0 \le b \le 6$

Let $E(X_A^{2_{pgT}}(a,b))$ represent the second moment of the total number of points played in a tiebreak game at point score (a,b) for player A serving.

Recurrence formulas:

$E(X_A^{2_{pgT}}(a,b)) = p_AE(X_B^{2_{pgT}}(a+1,b)) + q_AE(X_B^{2_{pgT}}(a,b+1))$, if $(a+b)$ is even

$E(X_A^{2_{pgT}}(a,b)) = p_AE(X_A^{2_{pgT}}(a+1,b)) + q_AE(X_A^{2_{pgT}}(a,b+1))$, if $(a+b)$ is odd

Boundary values:

$E(X_A^{2_{pgT}}(7,0))$ and $E(X_A^{2_{pgT}}(0,7)) = 49$,

$E(X_A^{2_{pgT}}(7,1))$ and $E(X_A^{2_{pgT}}(1,7)) = 64$,

$E(X_A^{2_{pgT}}(7,2))$ and $E(X_A^{2_{pgT}}(2,7)) = 81$,

$E(X_A^{2_{pgT}}(7,3))$ and $E(X_A^{2_{pgT}}(3,7)) = 100$,

$E(X_A^{2_{pgT}}(7,4))$ and $E(X_A^{2_{pgT}}(4,7)) = 121$,

$E(X_A^{2_{pgT}}(7,5))$ and $E(X_A^{2_{pgT}}(5,7)) = 144$,

$E(X_A^{2_{pgT}}(6,6)) = \frac{4(36q_A^2q_B^2+72p_Ap_Bq_Aq_B-83q_Aq_B+36p_A^2p_B^2-83p_Ap_B+49)}{(q_Aq_B+p_Ap_B-1)^2}$

Let $\sigma^2(Y_A^{pgT}(a,b))$ represent the variance of the number of points remaining in a tiebreak game at point score (a,b) for player A serving.

$\sigma^2(Y_A^{pgT}(a,b)) = E(X_A^{2_{pgT}}(a,b)) - E(X_A^{pgT}(a,b))^2$, for $0 \le a \le 6$ and $0 \le b \le 6$

Let $E(X_A^{3_{pgT}}(a,b))$ represent the third moment of the total number of points played in a tiebreak game at point score (a,b) for player A serving.

Recurrence formulas:

$E(X_A^{3_{pgT}}(a,b)) = p_A E(X_B^{3_{pgT}}(a+1,b)) + q_A E(X_B^{3_{pgT}}(a,b+1))$, if $(a+b)$ is even

$E(X_A^{3_{pgT}}(a,b)) = p_A E(X_A^{3_{pgT}}(a+1,b)) + q_A E(X_A^{3_{pgT}}(a,b+1))$, if $(a+b)$ is odd

Boundary values:

$E(X_A^{3_{pgT}}(7,0))$ and $E(X_A^{3_{pgT}}(0,7)) = 343$,

$E(X_A^{3_{pgT}}(7,1))$ and $E(X_A^{3_{pgT}}(1,7)) = 512$,

$E(X_A^{3_{pgT}}(7,2))$ and $E(X_A^{3_{pgT}}(2,7)) = 729$,

$E(X_A^{3_{pgT}}(7,3))$ and $E(X_A^{3_{pgT}}(3,7)) = 1000$,

$E(X_A^{3_{pgT}}(7,4))$ and $E(X_A^{3_{pgT}}(4,7)) = 1331$,

$E(X_A^{3_{pgT}}(7,5))$ and $E(X_A^{3_{pgT}}(5,7)) = 1728$,

$E(X_A^{3_{pgT}}(6,6)) = M^{(3)}_{X_A^{pgT}(6,6)}(0)$

Table 6.5 represents the third moment of the total number of points played in a tiebreak game at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$.

		B score							
		0	1	2	3	4	5	6	7
A score	0	2043.9	2087.1	1956.3	1777.2	1344.3	1032.7	610.2	343
	1	1977.6	2167.2	2215.6	2042.4	1825.6	1291.6	941.7	512
	2	1861.4	2096.4	2321.8	2381.2	2152.9	1837.8	1205.1	729
	3	1519.7	1958.3	2229.6	2521.2	2588.2	2225.7	1824.3	1000
	4	1250.9	1537.7	2051.0	2425.2	2810.4	2942.2	2329.5	1331
	5	808.7	1223.1	1525.9	2189.1	2608.3	3317.7	3608.9	1728
	6	554.8	742.8	1119.5	1445.9	2173.5	2880.8	4761.7	
	7	343	512	729	1000	1331	1728		

Table 6.5: The third moment of the total number of points played in a tiebreak game at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$

Let $\gamma_1(Y_A^{pg_T}(a,b))$ represent the coefficient of skewness of the number of points remaining in a tiebreak game at point score (a,b) for player A serving.

$$\gamma_1(Y_A^{pg_T}(a,b)) = \frac{E(X_A^{3pg_T}(a,b)) - 3E(X_A^{2pg_T}(a,b))E(X_A^{pg_T}(a,b)) + 2E(X_A^{pg_T}(a,b))^3}{\sigma^2(Y_A^{pg_T}(a,b))^{\frac{3}{2}}},$$

for $0 \leq a \leq 6$ and $0 \leq b \leq 6$

Table 6.6 represents the coefficient of skewness of the total number of points played in a tiebreak game at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$. It shows that the coefficient of skewness from the outset in such a game is 1.9.

		B score						
		0	1	2	3	4	5	6
A score	0	1.9	1.9	1.8	1.8	2.0	2.5	3.8
	1	1.9	2.0	2.1	2.0	2.0	2.5	3.5
	2	1.9	2.0	2.1	2.1	2.1	2.5	3.9
	3	1.9	2.0	2.2	2.2	2.2	2.4	3.0
	4	2.2	2.3	2.3	2.3	2.2	2.2	2.8
	5	2.6	2.8	2.6	2.5	2.3	2.1	1.9
	6	3.4	4.4	3.8	4.1	2.9	2.5	2.1

Table 6.6: The coefficient of skewness of the total number of points played in a tiebreak game at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$

Let $E(X_A^{4pg_T}(a,b))$ represent the fourth moment of the total number of points played in a tiebreak game at point score (a,b) for player A serving.

Recurrence formulas:

$E(X_A^{4pg_T}(a,b)) = p_A E(X_B^{4pg_T}(a+1,b)) + q_A E(X_B^{4pg_T}(a,b+1))$, if $(a+b)$ is even

$E(X_A^{4pg_T}(a,b)) = p_A E(X_A^{4pg_T}(a+1,b)) + q_A E(X_A^{4pg_T}(a,b+1))$, if $(a+b)$ is odd

Boundary values:

$E(X_A^{4pg_T}(7,0))$ and $E(X_A^{4pg_T}(0,7)) = 2401$,

$E(X_A^{4pg_T}(7,1))$ and $E(X_A^{4pg_T}(1,7)) = 4096$,

$E(X_A^{4_{pg_T}}(7,2))$ and $E(X_A^{4_{pg_T}}(2,7)) = 6561$,

$E(X_A^{4_{pg_T}}(7,3))$ and $E(X_A^{4_{pg_T}}(3,7)) = 10000$,

$E(X_A^{4_{pg_T}}(7,4))$ and $E(X_A^{4_{pg_T}}(4,7)) = 14641$,

$E(X_A^{4_{pg_T}}(7,5))$ and $E(X_A^{4_{pg_T}}(5,7)) = 20736$,

$E(X_A^{4_{pg_T}}(6,6)) = M^{(4)}_{X_A^{pg_T}(6,6)}(0)$

Table 6.7 represents the fourth moment of the total number of points played in a tiebreak game at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$.

		B score							
		0	1	2	3	4	5	6	7
	0	30733.3	31527.0	29188.6	25934.2	18217.3	12761.5	5950.0	2401
	1	29514.6	32960.2	33865.9	30663.9	26655.1	16936.2	10604.2	4096
	2	27402.7	31635.7	35828.5	36979.2	32611.9	26554.1	14593.1	6561
A score	3	21282.1	29066.0	34053.8	39655.9	41036.4	33885.1	25581.2	10000
	4	16511.3	21271.7	30620.2	37655.4	45419.6	48097.5	35130.9	14641
	5	9153.4	15541.9	20723.6	32896.7	41335.7	56044.7	62513.5	20736
	6	5237.9	7579.3	13262.6	18481.8	32320.4	46341.6	88119.1	
	7	2401	4096	6561	10000	14641	20736		

Table 6.7: The fourth moment of the total number of points played in a tiebreak game at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$

Let $\gamma_2(Y_A^{pg_T}(a,b))$ represent the coefficient of excess kurtosis of the number of points remaining in a tiebreak game at point score (a,b) for player A serving.

$$\gamma_2(Y_A^{pg_T}(a,b)) = \frac{E(X_A^{4_{pg_T}}(a,b)) - 4E(X_A^{3_{pg_T}}(a,b))E(X_A^{pg_T}(a,b)) - 3E(X_A^{2_{pg_T}}(a,b))^2 + 12E(X_A^{2_{pg_T}}(a,b))E(X_A^{pg_T}(a,b))^2 - 6E(X_A^{pg_T}(a,b))^4}{\sigma^2(Y_A^{pg_T}(a,b))^2},$$

for $0 \leq a \leq 6$ and $0 \leq b \leq 6$

Table 6.8 represents the coefficient of excess kurtosis of the total number of points played in a tiebreak game at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$. It shows that the coefficient of excess kurtosis from the outset in such a game is 6.3.

		B score						
		0	1	2	3	4	5	6
	0	6.3	6.3	5.9	6.1	7.4	10.9	26.2
	1	6.3	6.8	6.8	6.8	6.7	10.3	21.3
	2	6.3	6.6	7.1	7.0	7.0	9.2	23.7
A score	3	6.6	6.6	7.4	7.3	7.1	8.1	13.1
	4	8.2	8.9	7.9	7.8	7.2	7.0	11.2
	5	12.9	12.8	10.9	8.9	7.6	6.4	5.6
	6	22.9	32.1	22.0	23.5	11.4	9.0	6.4

Table 6.8: The coefficient of excess kurtosis of the total number of points played in a tiebreak game at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$

6.4 Number of games in a tiebreak set

Let $E(X_A^{gs_T}(c,d))$ represent the first moment of the total number of games played in a tiebreak set at game score (c,d) for player A serving.

Recurrence formula:

$E(X_A^{gs_T}(c,d)) = p_A^g E(X_B^{gs_T}(c+1,d)) + q_A^g E(X_B^{gs_T}(c,d+1))$

Boundary values:

$E(X_A^{gs_T}(6,0))$ and $E(X_A^{gs_T}(0,6)) = 6$,

$E(X_A^{gs_T}(6,1))$ and $E(X_A^{gs_T}(1,6)) = 7$,

$E(X_A^{gs_T}(6,2))$ and $E(X_A^{gs_T}(2,6)) = 8$,

$E(X_A^{gs_T}(6,3))$ and $E(X_A^{gs_T}(3,6)) = 9$,

$E(X_A^{gs_T}(6,4))$ and $E(X_A^{gs_T}(4,6)) = 10$,

$E(X_A^{gs_T}(7,5))$ and $E(X_A^{gs_T}(5,7)) = 12$,

$E(X_A^{gs_T}(6,6))$=13

Let $\mu(Y_A^{gs_T}(c,d))$ represent the mean number of games remaining in a tiebreak set at game score (c,d) for player A serving.

$\mu(Y_A^{gs_T}(c,d)) = E(X_A^{gs_T}(c,d)) - c - d$, for $0 \leq c \leq 5$ and $0 \leq d \leq 5$; $(6,5)$; $(5,6)$; $(6,6)$

Let $E(X_A^{gs_T}(a,b:c,d))$ represent the first moment of the total number of games played in a tiebreak set at point and game score $(a,b:c,d)$ for player A serving.

$E(X_A^{gs_T}(a,b:c,d)) = P_A^{pg}(a,b)E(X_B^{gs_T}(c+1,d)) + (1 - P_A^{pg}(a,b))E(X_B^{pg_T}(c,d+1))$

Let $\mu(Y_A^{gs_T}(a,b:c,d))$ represent the mean number of games remaining in a tiebreak set at point and game score $(a,b:c,d)$ for player A serving.

$\mu(Y_A^{gs_T}(a,b:c,d)) = E(X_A^{gs_T}(a,b:c,d)) - c - d$

Let $E(X_A^{2_{gs_T}}(c,d))$ represent the second moment of the total number of games played in a tiebreak set at game score (c,d) for player A serving.

Recurrence formula:
$E(X_A^{2_{gs_T}}(c,d)) = p_A^g E(X_B^{2_{gs_T}}(c+1,d)) + q_A^g E(X_B^{2_{gs_T}}(c,d+1))$

Boundary values:
$E(X_A^{2_{gs_T}}(6,0))$ and $E(X_A^{2_{gs_T}}(0,6)) = 36$,
$E(X_A^{2_{gs_T}}(6,1))$ and $E(X_A^{2_{gs_T}}(1,6)) = 49$,
$E(X_A^{2_{gs_T}}(6,2))$ and $E(X_A^{2_{gs_T}}(2,6)) = 64$,
$E(X_A^{2_{gs_T}}(6,3))$ and $E(X_A^{2_{gs_T}}(3,6)) = 81$,
$E(X_A^{2_{gs_T}}(6,4))$ and $E(X_A^{2_{gs_T}}(4,6)) = 100$,
$E(X_A^{2_{gs_T}}(7,5))$ and $E(X_A^{2_{gs_T}}(5,7)) = 144$,
$E(X_A^{2_{gs_T}}(6,6))$=169

Let $\sigma^2(Y_A^{gs_T}(c,d))$ represent the variance of the number of games remaining in a tiebreak set at game score (c,d) for player A serving.

$\sigma^2(Y_A^{gs_T}(c,d)) = E(X_A^{2_{gs_T}}(c,d)) - E(X_A^{gs_T}(c,d))^2$, for $0 \leq c \leq 5$ and $0 \leq d \leq 5$; $(6,5)$; $(5,6)$; $(6,6)$

Let $E(X_A^{2_{gs_T}}(a,b:c,d))$ represent the second moment of the total number of games played in a tiebreak set at point and game score $(a,b:c,d)$ for player A serving.

$$E(X_A^{2_{gs_T}}(a,b:c,d)) = P_A^{pg}(a,b)E(X_B^{2_{gs_T}}(c+1,d)) + (1-P_A^{pg}(a,b))E(X_B^{2_{gs_T}}(c,d+1))$$

Let $\sigma^2(Y_A^{gs_T}(a,b:c,d))$ represent the variance of the number of games remaining in a tiebreak set at point and game score $(a,b:c,d)$ for player A serving.

$$\sigma^2(Y_A^{gs_T}(a,b:c,d)) = E(X_A^{2_{gs_T}}(a,b:c,d)) - E(X_A^{gs_T}(a,b:c,d))^2$$

Let $E(X_A^{3_{gs_T}}(c,d))$ represent the third moment of the total number of games played in a tiebreak set at game score (c,d) for player A serving.

Recurrence formula:
$$E(X_A^{3_{gs_T}}(c,d)) = p_A^g E(X_B^{3_{gs_T}}(c+1,d)) + q_A^g E(X_B^{3_{gs_T}}(c,d+1))$$

Boundary values:
$E(X_A^{3_{gs_T}}(6,0))$ and $E(X_A^{3_{gs_T}}(0,6)) = 216$,
$E(X_A^{3_{gs_T}}(6,1))$ and $E(X_A^{3_{gs_T}}(1,6)) = 343$,
$E(X_A^{3_{gs_T}}(6,2))$ and $E(X_A^{3_{gs_T}}(2,6)) = 512$,
$E(X_A^{3_{gs_T}}(6,3))$ and $E(X_A^{3_{gs_T}}(3,6)) = 729$,
$E(X_A^{3_{gs_T}}(6,4))$ and $E(X_A^{3_{gs_T}}(4,6)) = 1000$,
$E(X_A^{3_{gs_T}}(7,5))$ and $E(X_A^{3_{gs_T}}(5,7)) = 1728$,
$E(X_A^{3_{gs_T}}(6,6))$=2197

Table 6.9 represents the third moment of the total number of games played in a tiebreak set at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$.

		B score							
		0	1	2	3	4	5	6	7
	0	1111.5	1155.2	1072.5	933.0	651.6	428.5	216	
	1	1025.5	1187.7	1236.5	1113.2	930.9	604.8	343	
	2	908.6	1084.1	1286.1	1345.1	1145.7	898.7	512	
A score	3	676.0	945.3	1148.3	1422.4	1502.7	1149.1	729	
	4	492.3	676.7	980.1	1219.9	1644.9	1794.1	1000	
	5	294.0	454.5	643.2	958.5	1229.4	2023.5	2091.9	1728
	6	216	343	512	729	1000	1833.1	2197	
	7						1728		

Table 6.9: The third moment of the total number of games played in a tiebreak set at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$

Let $\gamma_1(Y_A^{gs_T}(c,d))$ represent the coefficient of skewness of the number of games remaining in a tiebreak set at game score (c,d) for player A serving.

$\gamma_1(Y_A^{gs_T}(c,d)) = \frac{E(X_A^{3gs_T}(c,d)) - 3E(X_A^{2gs_T}(c,d))E(X_A^{gs_T}(c,d)) + 2E(X_A^{gs_T}(c,d))^3}{\sigma^2(Y_A^{gs_T}(c,d))^{\frac{3}{2}}}$,
for $0 \leq c \leq 5$ and $0 \leq d \leq 5$; $(6,5)$; $(5,6)$

$\gamma_1(Y_A^{gs_T}(6,6)) = 0$

Table 6.10 represents the coefficient of skewness of the total number of games played in a tiebreak set at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$. It shows that the coefficient of skewness from the outset in such a set is 0.3.

Let $E(X_A^{3gs_T}(a,b:c,d))$ represent the third moment of the total number of games played in a tiebreak set at point and game score $(a,b:c,d)$ for player A serving.

$$E(X_A^{3gs_T}(a,b:c,d)) = P_A^{pg}(a,b)E(X_B^{3gs_T}(c+1,d)) + (1 - P_A^{pg}(a,b))E(X_B^{3gs_T}(c,d+1))$$

Let $\gamma_1(Y_A^{gs_T}(a,b:c,d))$ represent the coefficient of skewness of the number of games remaining in a tiebreak set at point and game score $(a,b:c,d)$ for player A serving.

$$\gamma_1(Y_A^{gs_T}(a,b:c,d)) = \frac{E(X_A^{3gs_T}(a,b:c,d)) - 3E(X_A^{2gs_T}(a,b:c,d))E(X_A^{gs_T}(a,b:c,d)) + 2E(X_A^{gs_T}(a,b:c,d))^3}{\sigma^2(Y_A^{gs_T}(a,b:c,d))^{\frac{3}{2}}}$$

		B score						
		0	1	2	3	4	5	6
A score	0	0.3	0.2	0.3	0.5	1.0	2.3	
	1	0.3	0.3	0.2	0.4	0.8	1.9	
	2	0.6	0.4	0.3	0.2	0.5	1.5	
	3	0.9	0.7	0.5	0.1	-0.1	1.2	
	4	1.8	1.5	1.1	0.9	-0.3	-0.9	
	5	2.7	3.0	2.6	1.8	1.5	-0.5	-1.3
	6						1.3	0

Table 6.10: The coefficient of skewness of the total number of games played in a tiebreak set at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$

Let $E(X_A^{4_{gs_T}}(c,d))$ represent the fourth moment of the total number of games played in a tiebreak set at game score (c,d) for player A serving.

Recurrence formula:

$E(X_A^{4_{gs_T}}(c,d)) = p_A^g E(X_B^{4_{gs_T}}(c+1,d)) + q_A^g E(X_B^{4_{gs_T}}(c,d+1))$

Boundary values:

$E(X_A^{4_{gs_T}}(6,0))$ and $E(X_A^{4_{gs_T}}(0,6)) = 1296$,

$E(X_A^{4_{gs_T}}(6,1))$ and $E(X_A^{4_{gs_T}}(1,6)) = 2401$,

$E(X_A^{4_{gs_T}}(6,2))$ and $E(X_A^{4_{gs_T}}(2,6)) = 4096$,

$E(X_A^{4_{gs_T}}(6,3))$ and $E(X_A^{4_{gs_T}}(3,6)) = 6561$,

$E(X_A^{4_{gs_T}}(6,4))$ and $E(X_A^{4_{gs_T}}(4,6)) = 10000$,

$E(X_A^{4_{gs_T}}(7,5))$ and $E(X_A^{4_{gs_T}}(5,7)) = 20736$,

$E(X_A^{4_{gs_T}}(6,6))$=28561

Table 6.11 represents the fourth moment of the total number of games played in a tiebreak set at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$.

Let $\gamma_2(Y_A^{gs_T}(c,d))$ represent the coefficient of excess kurtosis of the number of games remaining in a tiebreak set at game score (c,d) for player A serving.

		B score							
		0	1	2	3	4	5	6	7
A score	0	12297.5	12882.3	11737.3	9970.3	6214.3	3548.6	1296	
	1	11097.2	13312.7	13987.8	12230.9	9848.1	5427.2	2401	
	2	9628.6	11827.0	14662.5	15510.5	12588.4	9205.8	4096	
	3	6529.7	10034.8	12638.4	16604.3	17811.1	12441.4	6561	
	4	4331.5	6371.4	10363.8	13511.6	19870.6	22154.8	10000	
	5	2103.3	3740.5	5794.6	9915.2	13511.4	25666.2	26807.1	20736
	6	1296	2401	4096	6561	10000	22489.9	28561	
	7						20736		

Table 6.11: The fourth moment of the total number of games played in a tiebreak set at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$

$\gamma_2(Y_A^{gs_T}(c,d))$

$= \frac{E(X_A^{4gs_T}(c,d)) - 4E(X_A^{3gs_T}(c,d))E(X_A^{gs_T}(c,d)) - 3E(X_A^{2gs_T}(c,d))^2 + 12E(X_A^{2gs_T}(c,d))E(X_A^{gs_T}(c,d))^2 - 6E(X_A^{gs_T}(c,d))^4}{\sigma^2(Y_A^{gs_T}(c,d))^2},$

for $0 \leq c \leq 5$ and $0 \leq d \leq 5$; $(6,5)$; $(5,6)$

$\gamma_2(Y_A^{gs_T}(6,6)) = 0$

Table 6.12 represents the coefficient of excess kurtosis of the total number of games played in a tiebreak set at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$. It shows that the coefficient of excess kurtosis from the outset in such a set is -0.9.

		B score						
		0	1	2	3	4	5	6
A score	0	-0.9	-1.0	-0.8	-0.7	0.7	6.5	
	1	-0.7	-1.1	-1.2	-0.9	-0.5	4.2	
	2	-0.6	-0.8	-1.4	-1.5	-0.8	1.1	
	3	0.4	-0.5	-0.8	-1.6	-1.6	0.2	
	4	3.4	2.0	-0.1	-0.7	-1.7	-0.8	
	5	9.4	9.8	6.7	1.4	0.3	-1.7	-0.2
	6						-0.2	0

Table 6.12: The coefficient of excess kurtosis of the total number of games played in a tiebreak set at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$

Let $E(X_A^{4gs_T}(a,b:c,d))$ represent the fourth moment of the total number of games played in a tiebreak set at point and game score $(a, b : c, d)$ for player A serving.

$$E(X_A^{4_{gs_T}}(a, b : c, d)) = P_A^{pg}(a, b)E(X_B^{4_{gs_T}}(c + 1, d)) + (1 - P_A^{pg}(a, b))E(X_B^{4_{gs_T}}(c, d + 1))$$

Let $\gamma_2(Y_A^{gs_T}(a, b : c, d))$ represent the coefficient of excess kurtosis of the number of games remaining in a tiebreak set at point and game score $(a, b : c, d)$ for player A serving.

$$\gamma_2(Y_A^{gs_T}(a, b : c, d))$$
$$= (E(X_A^{4_{gs_T}}(a, b : c, d)) - 4E(X_A^{3_{gs_T}}(a, b : c, d))E(X_A^{gs_T}(a, b : c, d)) - 3E(X_A^{2_{gs_T}}(a, b : c, d))^2 + 12E(X_A^{2_{gs_T}}(a, b : c, d))E(X_A^{gs_T}(a, b : c, d))^2 - 6E(X_A^{gs_T}(a, b : c, d))^4)/\sigma^2(Y_A^{gs_T}(a, b : c, d))^2$$

6.5 Number of games in an advantage set

The moment generating function for the total number of games played in an advantage set from $(5, 5)$ with player A serving is given by $M_{X_A^{gs}(5,5)}(t) = \frac{(p_A^g q_B^g + q_A^g p_B^g)e^{12t}}{1-(p_A^g p_B^g + q_A^g q_B^g)e^{2t}}$

Let $E(X_A^{gs}(c, d))$ represent the first moment of the total number of games played in an advantage set at game score (c, d) for player A serving.

Recurrence formula:

$$E(X_A^{gs}(c, d)) = p_A^g E(X_B^{gs}(c + 1, d)) + q_A^g E(X_B^{gs}(c, d + 1))$$

Boundary values:

$E(X_A^{gs}(6, 0))$ and $E(X_A^{gs}(0, 6)) = 6$,

$E(X_A^{gs}(6, 1))$ and $E(X_A^{gs}(1, 6)) = 7$,

$E(X_A^{gs}(6, 2))$ and $E(X_A^{gs}(2, 6)) = 8$,

$E(X_A^{gs}(6, 3))$ and $E(X_A^{gs}(3, 6)) = 9$,

$E(X_A^{gs}(6, 4))$ and $E(X_A^{gs}(4, 6)) = 10$,

$E(X_A^{gs}(5, 5)) = \frac{2(5q_A^g q_B^g + 5p_A^g p_B^g - 6)}{q_A^g q_B^g + p_A^g p_B^g - 1}$

Let $\mu(Y_A^{gs}(c,d))$ represent the mean number of games remaining in an advantage set at game score (c,d) for player A serving.

$\mu(Y_A^{gs}(c,d)) = E(X_A^{gs}(c,d)) - c - d$, for $0 \leq c \leq 5$ and $0 \leq d \leq 5$

Let $E(X_A^{gs}(a,b:c,d))$ represent the first moment of the total number of games played in an advantage set at point and game score $(a,b:c,d)$ for player A serving.

$E(X_A^{gs}(a,b:c,d)) = P_A^{pg}(a,b)E(X_B^{gs}(c+1,d)) + (1-P_A^{pg}(a,b))E(X_B^{gs}(c,d+1))$

Let $\mu(Y_A^{gs}(a,b:c,d))$ represent the mean number of games remaining in an advantage set at point and game score $(a,b:c,d)$ for player A serving.

$\mu(Y_A^{gs}(a,b:c,d)) = E(X_A^{gs}(a,b:c,d)) - c - d$

Let $E(X_A^{2_{gs}}(c,d))$ represent the second moment of the total number of games played in an advantage set at game score (c,d) for player A serving.

Recurrence formula:
$E(X_A^{2_{gs}}(c,d)) = p_A^g E(X_B^{2_{gs}}(c+1,d)) + q_A^g E(X_B^{2_{gs}}(c,d+1))$

Boundary values:
$E(X_A^{2_{gs}}(6,0))$ and $E(X_A^{2_{gs}}(0,6)) = 36$,
$E(X_A^{2_{gs}}(6,1))$ and $E(X_A^{2_{gs}}(1,6)) = 49$,
$E(X_A^{2_{gs}}(6,2))$ and $E(X_A^{2_{gs}}(2,6)) = 64$,
$E(X_A^{2_{gs}}(6,3))$ and $E(X_A^{2_{gs}}(3,6)) = 81$,
$E(X_A^{2_{gs}}(6,4))$ and $E(X_A^{2_{gs}}(4,6)) = 100$,
$E(X_A^{2_{gs}}(5,5)) = \frac{4(25(q_A^g)^2(q_B^g)^2+50p_A^g p_B^g q_A^g q_B^g-59q_A^g q_B^g+25(p_A^g)^2(p_B^g)^2-59p_A^g p_B^g+36)}{(q_A^g q_B^g+p_A^g p_B^g-1)^2}$

Let $\sigma^2(Y_A^{gs}(c,d))$ represent the variance of the number of games remaining in an advantage set at game score (c,d) for player A serving.

$\sigma^2(Y_A^{gs}(c,d)) = E(X_A^{2_{gs}}(c,d)) - E(X_A^{gs}(c,d))^2$, for $0 \leq c \leq 5$ and $0 \leq d \leq 5$

Let $E(X_A^{2_{gs}}(a,b:c,d))$ represent the second moment of the total number of games played in an advantage set at point and game score $(a,b:c,d)$ for player A serving.

$$E(X_A^{2_{gs}}(a,b:c,d)) = P_A^{pg}(a,b)E(X_B^{2_{gs}}(c+1,d)) + (1-P_A^{pg}(a,b))E(X_B^{2_{gs}}(c,d+1))$$

Let $\sigma^2(Y_A^{gs}(a,b:c,d))$ represent the variance of the number of games remaining in an advantage set at point and game score $(a,b:c,d)$ for player A serving.

$$\sigma^2(Y_A^{gs}(a,b:c,d)) = E(X_A^{2_{gs}}(a,b:c,d)) - E(X_A^{gs}(a,b:c,d))^2$$

Let $E(X_A^{3_{gs}}(c,d))$ represent the third moment of the total number of games played in an advantage set at game score (c,d) for player A serving.

Recurrence formula:

$$E(X_A^{3_{gs}}(c,d)) = p_A^g E(X_B^{3_{gs}}(c+1,d)) + q_A^g E(X_B^{3_{gs}}(c,d+1))$$

Boundary values:

$E(X_A^{3_{gs}}(6,0))$ and $E(X_A^{3_{gs}}(0,6)) = 216$,

$E(X_A^{3_{gs}}(6,1))$ and $E(X_A^{3_{gs}}(1,6)) = 343$,

$E(X_A^{3_{gs}}(6,2))$ and $E(X_A^{3_{gs}}(2,6)) = 512$,

$E(X_A^{3_{gs}}(6,3))$ and $E(X_A^{3_{gs}}(3,6)) = 729$,

$E(X_A^{3_{gs}}(6,4))$ and $E(X_A^{3_{gs}}(4,6)) = 1000$,

$E(X_A^{3_{gs}}(5,5)) = M^{(3)}_{X_A^{gs}(5,5)}(0)$

Table 6.13 represents the third moment of the total number of games played in an advantage set at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$.

Let $\gamma_1(Y_A^{gs}(c,d))$ represent the coefficient of skewness of the number of games remaining in an advantage set at game score (c,d) for player A serving.

		B score						
		0	1	2	3	4	5	6
	0	1879.1	1976.4	1754.2	1524.3	861.9	514.8	216
	1	1644.0	2049.2	2169.7	1814.5	1486.4	716.0	343
	2	1456.8	1745.9	2286.3	2452.4	1829.5	1319.8	512
A score	3	908.0	1509.3	1850.6	2652.0	2918.1	1691.8	729
	4	628.3	855.8	1530.1	1932.3	3312.4	3847.5	1000
	5	310.1	526.5	741.0	1395.0	1822.6	4670.1	
	6	216	343	512	729	1000		

Table 6.13: The third moment of the total number of games played in an advantage set at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$

$\gamma_1(Y_A^{gs}(c,d)) = \frac{E(X_A^{3gs}(c,d)) - 3E(X_A^{2gs}(c,d))E(X_A^{gs}(c,d)) + 2E(X_A^{gs}(c,d))^3}{\sigma^2(Y_A^{gs}(c,d))^{\frac{3}{2}}}$, for $0 \le c \le 5$ and $0 \le d \le 5$

Table 6.14 represents the coefficient of skewness of the total number of games played in an advantage set at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$. It shows that the coefficient of skewness from the outset in such a set is 2.5.

		B score					
		0	1	2	3	4	5
	0	2.5	2.4	2.6	2.6	3.5	5.4
	1	2.6	2.4	2.4	2.7	2.9	5.3
A score	2	2.7	2.7	2.4	2.3	2.9	3.7
	3	3.4	2.8	2.8	2.2	2.1	3.6
	4	4.5	4.4	3.2	3.1	2.1	1.9
	5	5.8	6.5	6.3	3.9	3.7	2.1

Table 6.14: The coefficient of skewness of the total number of games played in an advantage set at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$

Let $E(X_A^{3_{gs}}(a,b:c,d))$ represent the third moment of the total number of games played in an advantage set at point and game score $(a,b:c,d)$ for player A serving.

$E(X_A^{3_{gs}}(a,b:c,d)) = P_A^{pg}(a,b)E(X_B^{3_{gs}}(c+1,d)) + (1-P_A^{pg}(a,b))E(X_B^{3_{gs}}(c,d+1))$

Let $\gamma_1(Y_A^{gs}(a,b:c,d))$ represent the coefficient of skewness of the number of games remaining in an advantage set at point and game score $(a,b:c,d)$ for player A serving.

$$\gamma_1(Y_A^{gs}(a,b:c,d)) = \frac{E(X_A^{3gs}(a,b:c,d)) - 3E(X_A^{2gs}(a,b:c,d))E(X_A^{gs}(a,b:c,d)) + 2E(X_A^{gs}(a,b:c,d))^3}{\sigma^2(Y_A^{gs}(a,b:c,d))^{\frac{3}{2}}}$$

Let $E(X_A^{4_{gs}}(c,d))$ represent the fourth moment of the total number of games played in an advantage set at game score (c,d) for player A serving.

Recurrence formula:

$E(X_A^{4_{gs}}(c,d)) = p_A^g E(X_B^{4_{gs}}(c+1,d)) + q_A^g E(X_B^{4_{gs}}(c,d+1))$

Boundary values:

$E(X_A^{4_{gs}}(6,0))$ and $E(X_A^{4_{gs}}(0,6)) = 1296$,

$E(X_A^{4_{gs}}(6,1))$ and $E(X_A^{4_{gs}}(1,6)) = 2401$,

$E(X_A^{4_{gs}}(6,2))$ and $E(X_A^{4_{gs}}(2,6)) = 4096$,

$E(X_A^{4_{gs}}(6,3))$ and $E(X_A^{4_{gs}}(3,6)) = 6561$,

$E(X_A^{4_{gs}}(6,4))$ and $E(X_A^{4_{gs}}(4,6)) = 10000$,

$E(X_A^{4_{gs}}(5,5)) = M^{(4)}_{X_A^{gs}(5,5)}(0)$

Table 6.15 represents the fourth moment of the total number of games played in an advantage set at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$.

Let $\gamma_2(Y_A^{gs}(c,d))$ represent the coefficient of excess kurtosis of the number of games remaining in an advantage set at game score (c,d) for player A serving.

$$\gamma_2(Y_A^{gs}(c,d)) = \frac{E(X_A^{4gs}(c,d)) - 4E(X_A^{3gs}(c,d))E(X_A^{gs}(c,d)) - 3E(X_A^{2gs}(c,d))^2 + 12E(X_A^{2gs}(c,d))E(X_A^{gs}(c,d))^2 - 6E(X_A^{gs}(c,d))^4}{\sigma^2(Y_A^{gs}(c,d))^2},$$

for $0 \leq c \leq 5$ and $0 \leq d \leq 5$

Table 6.16 represents the coefficient of excess kurtosis of the total number of games played in an advantage set at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$.

		B score						
		0	1	2	3	4	5	6
	0	32606.2	34608.6	29771.3	25613.1	11778.5	5832.5	1296
	1	27459.1	36105.7	38678.7	30785.3	24544.5	8370.9	2401
	2	24132.6	29336.4	41126.1	44805.6	30680.0	20344.9	4096
A score	3	12666.9	24957.4	31219.0	49136.0	55258.3	26798.4	6561
	4	7930.6	11110.0	24914.1	32358.8	63988.2	76481.9	10000
	5	2530.1	5644.7	8382.7	21462.1	29205.8	95687.7	
	6	1296	2401	4096	6561	10000		

Table 6.15: The fourth moment of the total number of games played in an advantage set at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$

It shows that the coefficient of excess kurtosis from the outset in such a set is 9.0.

		B score					
		0	1	2	3	4	5
	0	9.0	8.7	9.9	10.0	21.2	46.7
	1	10.6	8.7	8.3	10.6	11.8	45.5
A score	2	10.7	11.1	8.2	7.6	12.0	18.8
	3	19.5	11.6	11.6	7.2	6.4	18.2
	4	31.5	30.5	14.0	13.5	6.2	5.6
	5	72.7	64.4	60.9	20.2	18.4	6.2

Table 6.16: The coefficient of excess kurtosis of the total number of games played in an advantage set at various score lines for player A serving given $p_A = 0.62$ and $p_B = 0.60$

Let $E(X_A^{4gs}(a, b : c, d))$ represent the fourth moment of the total number of games played in an advantage set at point and game score $(a, b : c, d)$ for player A serving.

$$E(X_A^{4gs}(a, b : c, d)) = P_A^{pg}(a, b)E(X_B^{4gs}(c + 1, d)) + (1 - P_A^{pg}(a, b))E(X_B^{4gs}(c, d + 1))$$

Let $\gamma_2(Y_A^{gs}(a, b : c, d))$ represent the coefficient of excess kurtosis of the number of games remaining in an advantage set at point and game score $(a, b : c, d)$ for player A serving.

$$\gamma_2(Y_A^{gs}(a, b : c, d)) = \frac{E(X_A^{4gs}(a,b{:}c,d)) - 4E(X_A^{3gs}(a,b{:}c,d))E(X_A^{gs}(a,b{:}c,d)) - 3E(X_A^{2gs}(a,b{:}c,d))^2 + 12E(X_A^{2gs}(a,b{:}c,d))E(X_A^{gs}(a,b{:}c,d))^2 - 6E(X_A^{gs}(a,b{:}c,d))^4}{\sigma^2(Y_A^{gs}(a,b{:}c,d))^2}$$

6.6 Number of sets in a match

Let $E(X^{sm_5}(e,f))$ represent the first moment of the total number of sets played in a best-of-5 set match at set score (e,f)

Recurrence formula:

$E(X^{sm_5}(e,f)) = p^{s_T}E(X^{sm_5}(e+1,f)) + q^{s_T}E(X^{sm_5}(e,f+1))$

Boundary values:

$E(X^{sm_5}(3,0))$ and $E(X^{sm_5}(0,3)) = 3$,

$E(X^{sm_5}(3,1))$ and $E(X^{sm_5}(1,3)) = 4$,

$E(X^{sm_5}(2,2)) = 5$

Let $\mu(Y^{sm_5}(e,f))$ represent the mean number of sets remaining in a best-of-5 set match at set score (e,f).

$\mu(Y^{sm_5}(e,f)) = E(X^{sm_5}(e,f)) - e - f$, for $0 \le e \le 2$ and $0 \le f \le 2$

Let $E(X_A^{sm_5}(a,b:c,d:e,f))$ represent the first moment of the total number of sets played in a best-of-5 set match at point, game and set score $(a,b:c,d:e,f)$ for player A serving.

$$E(X_A^{sm_5}(a,b:c,d:e,f))$$
$$= P_A^{ps_T}(a,b:c,d)E(X^{sm_5}(e+1,f)) + (1-P_A^{ps_T}(a,b:c,d))E(X^{sm_5}(e,f+1))$$

Let $\mu(Y_A^{sm_5}(a,b:c,d:e,f))$ represent the mean number of sets remaining in a best-of-5 set match at point, game and set score $(a,b:c,d:e,f)$ for player A serving.

$\mu(Y_A^{sm_5}(a,b:c,d:e,f)) = E(X_A^{sm_5}(a,b:c,d:e,f)) - e - f$

Let $E(X^{2_{sm_5}}(e,f))$ represent the second moment of the total number of sets played in a best-of-5 set match at set score (e,f).

Recurrence formula:

$E(X^{2_{sm_5}}(e,f)) = p^{s_T}E(X^{2_{sm_5}}(e+1,f)) + q^{s_T}E(X^{2_{sm_5}}(e,f+1))$

Boundary values:

$E(X^{2_{sm_5}}(3,0))$ and $E(X^{2_{sm_5}}(0,3)) = 9$,

$E(X^{2_{sm_5}}(3,1))$ and $E(X^{2_{sm_5}}(1,3)) = 16$,

$E(X^{2_{sm_5}}(2,2)) = 25$

Let $\sigma^2(Y^{sm_5}(e,f))$ represent the variance of the number of sets remaining in a best-of-5 set match at set score (e,f).

$\sigma^2(Y^{sm_5}(e,f)) = E(X^{2_{sm_5}}(e,f)) - E(X^{sm_5}(e,f))^2$, for $0 \leq e \leq 2$ and $0 \leq f \leq 2$

Let $E(X_A^{2_{sm_5}}(a,b:c,d:e,f))$ represent the second moment of the total number of sets played in a best-of-5 set match at point, game and set score $(a,b:c,d:e,f)$ for player A serving.

$$E(X_A^{2_{sm_5}}(a,b:c,d:e,f))$$
$$= P_A^{ps_T}(a,b:c,d)E(X^{2_{sm_5}}(e+1,f)) + (1 - P_A^{ps_T}(a,b:c,d))E(X^{2_{sm_5}}(e,f+1))$$

Let $\sigma^2(Y_A^{sm_5}(a,b:c,d:e,f))$ represent the variance of the number of sets remaining in a best-of-5 set match at point, game and set score $(a,b:c,d:e,f)$ for player A serving.

$\sigma^2(Y_A^{sm_5}(a,b:c,d:e,f)) = E(X_A^{2_{sm_5}}(a,b:c,d:e,f)) - E(X_A^{sm_5}(a,b:c,d:e,f))^2$

Let $E(X^{3_{sm_5}}(e,f))$ represent the third moment of the total number of sets played in a best-of-5 set match at set score (e,f).

Recurrence formula:

$E(X^{3_{sm_5}}(e,f)) = p^{s_T}E(X^{3_{sm_5}}(e+1,f)) + q^{s_T}E(X^{3_{sm_5}}(e,f+1))$

Boundary values:

$E(X^{3_{sm_5}}(3,0))$ and $E(X^{3_{sm_5}}(0,3)) = 27$,

$E(X^{3_{sm_5}}(3,1))$ and $E(X^{3_{sm_5}}(1,3)) = 64$,

$E(X^{3_{sm_5}}(2,2)) = 125$

Table 6.17 represents the third moment of the total number of sets played in a best-of-5 set match at various score lines given $p_A = 0.62$ and $p_B = 0.60$.

		B score			
		0	1	2	3
A score	0	76.3	82.6	67.7	27
	1	71.5	93.9	98.6	64
	2	54.4	90.4	125	
	3	27	64		

Table 6.17: The third moment of the total number of sets played in a best-of-5 set match at various score lines given $p_A = 0.62$ and $p_B = 0.60$

Let $\gamma_1(Y^{sm_5}(e,f))$ represent the coefficient of skewness of the number of sets remaining in a best-of-5 set match at set score (e,f).

$$\gamma_1(Y^{sm_5}(e,f)) = \frac{E(X^{3sm_5}(e,f)) - 3E(X^{2sm_5}(e,f))E(X^{sm_5}(e,f)) + 2E(X^{sm_5}(e,f))^3}{\sigma^2(Y^{sm_5}(e,f))^{\frac{3}{2}}},$$

for $0 \le e \le 1$ and $0 \le f \le 1$; $(2,0)$; $(2,1)$; $(0,2)$; $(1,2)$

$\gamma_1(Y^{sm_5}(2,2)) = 0$

Table 6.18 represents the coefficient of skewness of the total number of sets played in a best-of-5 set match at various score lines given $p_A = 0.62$ and $p_B = 0.60$. It shows that the coefficient of skewness from the outset in such a match is -0.17.

Let $E(X_A^{3_{sm_5}}(a,b:c,d:e,f))$ represent the third moment of the total number of sets played in a best-of-5 set match at point, game and set score $(a,b:c,d:e,f)$ for player A serving.

$$E(X_A^{3_{sm_5}}(a,b:c,d:e,f))$$
$$= P_A^{ps_T}(a,b:c,d)E(X^{3_{sm_5}}(e+1,f)) + (1 - P_A^{ps_T}(a,b:c,d))E(X^{3_{sm_5}}(e,f+1))$$

		B score		
		0	1	2
	0	-0.17	-0.40	0.21
A score	1	0.01	0.04	-0.27
	2	0.78	0.27	0

Table 6.18: The coefficient of skewness of the total number of sets played in a best-of-5 set match at various score lines given $p_A = 0.62$ and $p_B = 0.60$

Let $\gamma_1(Y_A^{sm_5}(a, b : c, d : e, f))$ represent the coefficient of skewness of the number of sets remaining in a best-of-5 set match at point, game and set score $(a, b : c, d : e, f)$ for player A serving.

$\gamma_1(Y_A^{sm_5}(a, b : c, d : e, f))$
$= \frac{E(X_A^{3sm_5}(a,b:c,d:e,f)) - 3E(X_A^{2sm_5}(a,b:c,d:e,f))E(X_A^{sm_5}(a,b:c,d:e,f)) + 2E(X_A^{sm_5}(a,b:c,d:e,f))^3}{\sigma^2(Y_A^{sm_5}(a,b:c,d:e,f))^{\frac{3}{2}}}$

Let $E(X^{4_{sm_5}}(e, f))$ represent the fourth moment of the total number of sets played in a best-of-5 set match at set score (e, f).

Recurrence formula:

$E(X^{4_{sm_5}}(e, f)) = p^{s_T} E(X^{4_{sm_5}}(e + 1, f)) + q^{s_T} E(X^{4_{sm_5}}(e, f + 1))$

Boundary values:

$E(X^{4_{sm_5}}(3, 0))$ and $E(X^{4_{sm_5}}(0, 3)) = 81$,

$E(X^{4_{sm_5}}(3, 1))$ and $E(X^{4_{sm_5}}(1, 3)) = 256$,

$E(X^{4_{sm_5}}(2, 2)) = 625$

Table 6.19 represents the fourth moment of the total number of sets played in a best-of-5 set match at various score lines given $p_A = 0.62$ and $p_B = 0.60$.

Let $\gamma_2(Y^{sm_5}(e, f))$ represent the coefficient of excess kurtosis of the number of sets remaining in a best-of-5 set match at set score (e, f).

		B score			
		0	1	2	3
	0	343.1	377.6	299.4	81
A score	1	316.9	437.1	465.6	256
	2	225.5	415.4	625	
	3	81	256		

Table 6.19: The fourth moment of the total number of sets played in a best-of-5 set match at various score lines given $p_A = 0.62$ and $p_B = 0.60$

$\gamma_2(Y^{sm_5}(e,f))$

$$= \frac{E(X^{4sm_5}(e,f)) - 4E(X^{3sm_5}(e,f))E(X^{sm_5}(e,f)) - 3E(X^{2sm_5}(e,f))^2 + 12E(X^{2sm_5}(e,f))E(X^{sm_5}(e,f))^2 - 6E(X^{sm_5}(e,f))^4}{\sigma^2(Y^{sm_5}(e,f))^2},$$

for $0 \leq e \leq 1$ and $0 \leq f \leq 1$; $(2,0)$; $(2,1)$; $(0,2)$; $(1,2)$

$\gamma_2(Y^{sm_5}(2,2)) = 0$

Table 6.20 represents the coefficient of excess kurtosis of the total number of sets played in a best-of-5 set match at various score lines given $p_A = 0.62$ and $p_B = 0.60$. It shows that the coefficient of excess kurtosis from the outset in such a match is -1.4.

		B score		
		0	1	2
	0	-1.4	-1.1	-1.6
A score	1	-1.4	-2.0	-1.9
	2	-0.9	-1.9	0

Table 6.20: The coefficient of excess kurtosis of the total number of sets played in a best-of-5 set match at various score lines given $p_A = 0.62$ and $p_B = 0.60$

Let $E(X_A^{4sm_5}(a,b:c,d:e,f))$ represent the fourth moment of the total number of sets played in a best-of-5 set match at point, game and set score $(a,b:c,d:e,f)$ for player A serving.

$E(X_A^{4sm_5}(a,b:c,d:e,f))$

$= P_A^{ps_T}(a,b:c,d)E(X^{4sm_5}(e+1,f)) + (1 - P_A^{ps_T}(a,b:c,d))E(X^{4sm_5}(e,f+1))$

Let $\gamma_2(Y_A^{sm_5}(a,b:c,d:e,f))$ represent the coefficient of excess kurtosis of the number of sets remaining in a best-of-5 set match at point, game and set score $(a,b:c,d:e,f)$ for player A serving.

$$\gamma_2(Y_A^{sm_5}(a,b:c,d:e,f))$$
$$= (E(X_A^{4_{sm_5}}(a,b:c,d:e,f)) - 4E(X_A^{3_{sm_5}}(a,b:c,d:e,f))E(X_A^{sm_5}(a,b:c,d:e,f)) - 3E(X_A^{2_{sm_5}}(a,b:c,d:e,f))^2 + 12E(X_A^{2_{sm_5}}(a,b:c,d:e,f))E(X_A^{sm_5}(a,b:c,d:e,f))^2 - 6E(X_A^{sm_5}(a,b:c,d:e,f))^4)/\sigma^2(Y_A^{sm_5}(a,b:c,d:e,f))^2$$

Chapter 7

Duration of a match: 4/4/

7.1 Introduction

Chapter 7 extends on chapter 6 by obtaining the parameters of distribution (mean, variance, skewness and kurtosis) of the number of points remaining in a match from any point, game and set score within the match. This is followed by the time duration of a match from any point, game and set score within the match.

In chapter 4 on the duration of a game a moment generating function was derived to obtain the parameters of distribution of the total number of points played in a game. Since prior to deuce, there are only three possible outcomes (the game finishes in either 4, 5 or 6 points); the probabilities were readily utilized from chapter 1, and therefore it was not necessary to condition on whether player A or player B won the game. Similar reasoning applies for obtaining moment generating functions for the total number of games played in a set and the total number of sets played in a match. However, to obtain a moment generating function for the total number of points played in a set requires conditioning on whether player A or player B won each game of the set, since there are a relatively 'large' number of possible outcomes to use the standard moment generating function formula (as was the case for the total number of points played in a game). Further, to obtain a moment generating function for the total number of points played in a match requires identifying the probabilities of player A and player B serving first at the start of each set, and therefore obtaining probabilities of each set finishing in an odd and even number of games. So to obtain the parameters of distribution for the total number of points played in a match (and likewise the number of points remaining in a match) conditional on the score line, requires some deeper mathematical ideas to what has been used in previous chapters.

Our use of the moment generating functions to add up the statistics on the number of points is based on the following theorem.

Theorem 7.1.1. *Let $Z = X + Y$, where X and Y are independent random variables, then $M_Z(t) = M_X(t)M_Y(t)$.*

Proof. $M_Z(t) = E(e^{Zt}) = E(e^{(X+Y)t}) = E(e^{Xt})E(e^{Yt}) = M_X(t)M_Y(t)$. □

The important step in this proof is the factorization of the joint distribution of X and Y to calculate two separate expectations. The main assumption in our model of tennis is that the probabilities of each server winning a point on serve are constant, but not necessarily the same. This assumption enables us to forget about the winner of the last point once we have recorded the result and updated the score. This loss of memory is called the Markov property. It enables us to claim that the winner of the next point is independent of the winner of the previous point. This independence is critical to our model; however the assumption of constant probability of the server winning a point could be weakened a little provided we still retained the Markov property.

The algorithm for extracting the moments of the sum of independent random variables is the following:

$m_{1Z} = m_{1Y} + m_{1X}$

$m_{2Z} = m_{2Y} + 2m_{1X}m_{1Y} + m_{2X}$

$$m_{3Z} = m_{3Y} + 3m_{1X}m_{2Y} + 3m_{2X}m_{1Y} + m_{3X}$$

$$m_{4Z} = m_{4Y} + 4m_{1X}m_{2Y} + 6m_{2X}m_{2Y} + 4m_{3X}m_{1Y} + m_{4X}$$

This is a simple corollary of Theorem 7.1, which can be obtained by successive differentiation with respect to t, and putting t to 0.

The outline of our strategy to solve the problems raised in this chapter is easy to state. For a partially completed match, we combine estimates of the remaining points played in the partially completed game, estimates of the remaining points played in the partially completed set at the end of the current game, and estimates of the remaining points played in the partially completed match at the end of the current set.

Our task is made more difficult by the rules relating to rotation of service between the players after each game, and for the rotation of service within tiebreak games. We develop a notation that requires us to specify the server at the start of each game, or each set, in an effort to reduce the complexity of our task. Thus an initial step, when given the current score and current server in a partially completed match, is the determination of the server at the start of the current set.

The score for a match in progress will be denoted by $(a, b : c, d : e, f)$, where (a, b) is the score in points, (c, d) is the score in games, and (e, f) is the score in sets, for player A and player B respectively. We will use a truncated form of this notation whenever it is convenient so to do.

Example: Best-of-5 final set advantage match, current score $(2, 1 : 2, 0 : 0, 2)$, current server player A. We can deduce from this information that the current set is a tiebreak set, with player A serving at the first point of this set.

The common notation for the score in this example might be 4-6, 4-6, 2-0 with player A serving at 30-15. The common notation is more than just a transcription and transposition of our preferred notation. It provides more information about the past history of the match, and as such it is possible to determine that player A served at the first point of the match, in addition to serving at the first point of the current set. However this extra information does not help in estimating the number of remaining points in the match.

Consider the scenario in the example above. There is a possibility that player A will win the third tiebreak set, and the number of remaining points will be large. However if player A loses his current lead in this set and player B succeeds in winning the set, then the number of remaining points played in the match will be modest, even if the outcome is decided by playing a tiebreak game. The key point here is that it is necessary to track separately the cases where player A is winning or losing a set. The same observation applies to a point, a game, a match, or even the server.

A tennis match consists of four levels - (points, games, sets, match). It becomes necessary to represent:
points in a point as pp,
points in a game as pg,
points in a tiebreak game as pg_T ,
points in an advantage set as ps,
points in a tiebreak set as ps_T
points in a best-of-5 all tiebreak set match as pm_{5T},
points in a best-of-5 final set advantage match as pm_5.

Let s_A, s_B represent the condition that player A and player B, respectively served first at the beginning of a set. Let c_A, c_B represent the condition that player A and player B, respectively are currently serving in the set at the score $(a, b : c, d)$. If (a, b) is not a boundary score for the current game then
$s_A = c_A$ and $s_B = c_B$, if $(c + d)$ mod 2=0

$s_A = c_B$ and $s_B = c_A$, if $(c + d)$ mod 2=1

except in the case of the tiebreak game of the tiebreak set, with $c = 6, d = 6$, when

$s_A = c_A$ and $s_B = c_B$, if $(a + b)$ mod 4=0 or 3

$s_A = c_A$ and $s_B = c_B$, if $(a + b)$ mod 4=1 or 2

Let $P^{ps}(a, b : c, d|s_A)$ represent the probability of player A winning an advantage set at this score, and player A serving first in the current set. Let $Y^{ps}(a, b : c, d|s_A)$ be the number of points remaining in the set at this score with player A serving first in the current set. This number is a random variable. Let $M_{Y^{ps}(a,b:c,d|s_A)}(t)$ be its moment generating function.

Similarly let $P^{ps}(a, b : c, d|w_A, s_A)$ represent the probability of player A winning an advantage set at this score, and player A serving first in the current set. Let $Y^{ps}(a, b : c, d|w_A, s_A)$ be a random variable of the number of points remaining in the set at this score conditional on player A both winning the set, and serving first in the current set. Let $M_{Y^{ps}(a,b:c,d|w_A,s_A)}(t)$ be its moment generating function conditional on player A both winning the set, and serving first in the current set.

Many variants of this notation will be used. The representation of the score will be restricted whenever it is not essential to display the full score. Other symbols include B for player B, l for the condition of losing, and n for the condition of serving next.

The next step is to introduce weighted moment generating functions. Let X be a conditional random variable. Let C be the condition that X occurs with probability p_X.

Then

$W_{X|C}(t) = p_X M_X(t)$

This product of a probability and its associated moment generating function is defined as a weighted moment generating function. The weight is the probability measure such that the conditions applied to the random variable are true.

Denote by w_{nX} the weighted n^{th} moment of the random variable X. Then

$w_{nX} = p_X m_{nX}$ for $n = 1, 2, 3, 4, ...$

The more important situation for us arises when the score does change. Let X and Y be independent random variable with conditional probabilities p_X and p_Y, respectively, of occurring. Let Z denote the random variable for their sum, $Z = X + Y$ when both X and Y occur. Then $p_Z = p_X p_Y$. It follows from Theorem 7.1 that the weighted moment generating functions satisfy

$W_{Z|C_1,C_2}(t) = W_{X|C1}(t) W_{Y|C2}(t)$.

The algorithm for extracting the weighted moments is the following:

$p_Z = p_X p_Y$

$w_{1Z} = p_X w_{1Y} + w_{1X} p_Y$

$w_{2Z} = p_X w_{2Y} + 2w_{1X} w_{1Y} + w_{2X} p_Y$

$w_{3Z} = p_X w_{3Y} + 3w_{1X} w_{2Y} + 3w_{2X} w_{1Y} + w_{3X} p_Y$

$w_{4Z} = p_X w_{4Y} + 4w_{1X} w_{2Y} + 6w_{2X} w_{2Y} + 4w_{3X} w_{1Y} + w_{4X} p_Y$

We now develop the algebra for weighted moment generating functions. We are able to add together two weighted moment generating functions whenever we encounter two mutually exclusive cases. Two simple examples where the score does not change are:

(a) Condition on initial server

$W_{Y^{ps}(a,b:c,d)}(t) = W_{Y^{ps}(a,b:c,d|s_A)}(t) + W_{Y^{ps}(a,b:c,d|s_B)}(t)$

(b) Condition on winning or losing

$W_{Y^{ps}(a,b:c,d|s_A)}(t) = W_{Y^{ps}(a,b:c,d|w_A,s_A)}(t) + W_{Y^{ps}(a,b:c,d|l_A,s_A)}(t)$

We now apply these ideas to the playing of a single point. In this case some of the notation appears to degenerate, so we must be careful. However this analysis will be used whenever the score changes as a point is played in a game, a set, or a match.

Each point played is a single point, irrespective of the score. For player A serving, the probability of winning the point is denoted by p_A irrespective of the score and $q_A = 1 - p_A$. Let $P^{pp}(()|c_A, w_A)$ and $P^{pp}(()|c_A, l_A)$ represent the probabilities of player A winning and losing a point on serve respectively from score line () within the point. It follows that:

$P^{pp}(()|c_A, w_A) = p_A$

$P^{pp}(()|c_A, l_A) = q_A$

Let $Y^{pp}(()|c_A)$ represent the number of points remaining in the point from score line () with player A serving. Each point played is a single point, so $Y^{pp}(()|c_A) = 1$. Let $Y^{pp}(()|c_A, w_A)$ and $Y^{pp}(()|c_A, l_A)$ represent the number of points remaining in the point from score line () given player A won and lost the point respectively with player A serving.

Therefore:

$M_{Y^{pp}(()|c_A)}(t) = E(e^{Y^{pp}(()|c_A)t}) = E(e^t) = e^t$

$W_{Y^{pp}(()|c_A, w_A)}(t) = P^{pp}(()|c_A, w_A) M_{Y^{pp}(()|c_A)}(t) = p_A e^t$

This is a fundamental brick in the model.

It is easy to check that

$w_n(Y^{pp}(()|c_A, w_A)) = p_A$ for $n = 0, 1, 2, 3, 4, ...$

Likewise $W_{Y^{pp}(()|c_A, l_A)}(t) = P^{pp}(()|c_A, l_A) M_{Y^{pp}(()|c_A)}(t) = q_A E(e^t)$

and

$w_n(Y^{pp}(()|c_A, l_A)) = q_A$ for $n = 0, 1, 2, 3, 4, ...$

In the following sections of this chapter we will describe the stages in the solution to our problems using backwards recursion. If we were using forward recursion it would be more natural to study the total number of points played to reach the current score line. The results for a completed game, set or match should be identical irrespective of the method used. Thus the two alternative methods provide us with a means of checking numerical calculations.

7.2 Number of points in a game

Whenever a point is played in a game, the score takes a branch in one of two directions, as illustrated in Figure 7.1. Weighted moment generating functions can be used to analyze the change in the number of remaining points in the game conditional on the server winning or losing the game.

$$\begin{array}{ccc} & & (a+1,b) \\ & \nearrow & \\ (a,b) & & \\ & \searrow & \\ & & (a,b+1) \end{array}$$

Figure 7.1: Graphical representation of the point played in a game at score (a,b)

Let $W_{Y^{pg}(a,b|c_A,w_A)}(t)$ and $W_{Y^{pg}(a,b|c_A,l_A)}(t)$ represent the weighted moment generating functions of the number of points remaining in the game from score line (a,b) given player A is serving and player A won and lost the game respectively.

Theorem 7.2.1. $W_{Y^{pg}(a,b|c_A,w_A)}(t)$
$= W_{Y^{pp}(()|c_A,w_A)}(t)W_{Y^{pg}(a+1,b|c_A,w_A)}(t) + W_{Y^{pp}(()|c_A,l_A)}(t)W_{Y^{pg}(a,b+1|c_A,w_A)}(t)$

Proof. $M_{Y^{pg}(a,b|c_A,w_A)}(t) = E(e^{tY^{pg}(a,b|c_A,w_A)})$ is an expectation that is calculated before the point at score (a,b) has been played. The point played is won and lost with probability p_A and q_A respectively, where $p_A + q_A = 1$ since there are only two possible outcomes. When we try to recalculate the original expectation after the point has been played we obtain the weighted sum of two expressions

$M_{Y^{pg}(a,b|c_A,w_A)}(t)$
$= p_A E(e^{t(1+Y^{pg}(a+1,b|c_A,w_A))}) P^{pg}(a+1,b|c_A,w_A)/P^{pg}(a,b|c_A,w_A)$
$+ q_A E(e^{t(1+Y^{pg}(a,b+1|c_A,w_A))}) P^{pg}(a,b+1|c_A,w_A)/P^{pg}(a,b|c_A,w_A)$

where the odds ratios
$P^{pg}(a+1,b|c_A,w_A)/P^{pg}(a,b|c_A,w_A)$ and $P^{pg}(a,b+1|c_A,w_A)/P^{pg}(a,b|c_A,w_A)$
reflect the changes in the chances of player A winning when the score is updated after winning or losing the point, respectively. The count of 1 for the point played is independent of the distribution of the remaining points after the point has been played, so, as in Theorem 7.1, we can factorize the expectations to obtain
$E(e^{t(1+Y^{pg}(a+1,b|c_A,w_A))}) = E(e^t)E(e^{t(Y^{pg}(a+1,b|c_A,w_A))})$ and
$E(e^{t(1+Y^{pg}(a,b+1|c_A,w_A))}) = E(e^t)E(e^{t(Y^{pg}(a,b+1|c_A,w_A))})$.

After some rearrangement we find that
$P^{pg}(a,b|c_A,w_A)M_{Y^{pg}(a,b|c_A,w_A)}(t)$
$= p_A E(e^t) P^{pg}(a+1,b|c_A,w_A) E(e^{t(Y^{pg}(a+1,b|c_A,w_A))})$
$+ q_A E(e^t) P^{pg}(a,b+1|c_A,w_A) E(e^{t(Y^{pg}(a,b+1|c_A,w_A))})$

The only step that is left is the identification of the various terms in this expression as weighted moment generating functions, to obtain
$W_{Y^{pg}(a,b|c_A,w_A)}(t) = W_{Y^{pp}(()|c_A,w_A)}(t)W_{Y^{pg}(a+1,b|c_A,w_A)}(t) + W_{Y^{pp}(()|c_A,l_A)}(t)W_{Y^{pg}(a,b+1|c_A,w_A)}(t)$

$\square$

Note carefully in Theorem 7.2.1 how first we are able to multiply the weighted moment generating functions on each path of this branching process which arises when scoring, because the steps on each branch are independent; and then add the results of this multiplication, because the paths are mutually exclusive.

It follows that $W_{Y^{pg}(a,b|c_A,w_A)}(t) = P^{pg}(a,b|c_A,w_A)M_{Y^{pg}(a,b|c_A,w_A)}(t)$

where $M_{Y^{pg}(a,b|c_A,w_A)}(t)$ is the moment generating function of the random variable $Y^{pg}(a,b|c_A,w_A)$.

By successive differentiation with respect to t from Theorem 7.2.1, and setting $t = 0$ we obtain the following recurrence formulas.

$w_1(Y^{pg}(a,b|c_A,w_A)) = p_A w_1(Y^{pg}(a+1,b|c_A,w_A)) + q_A w_1(Y^{pg}(a,b+1|c_A,w_A)) + p_A P^{pg}(a+1,b|c_A,w_A) + q_A P^{pg}(a,b+1|c_A,w_A)$

$w_2(Y^{2_{pg}}(a,b|c_A,w_A)) = p_A w_2(Y^{pg}(a+1,b|c_A,w_A)) + q_A w_2(Y^{pg}(a,b+1|c_A,w_A)) + 2p_A w_1(Y^{pg}(a+1,b|c_A,w_A)) + 2q_A w_1(Y^{pg}(a,b+1|c_A,w_A)) + p_A P^{pg}(a+1,b|c_A,w_A) + q_A P^{pg}(a,b+1|c_A,w_A)$

$w_3(Y^{pg}(a,b|c_A,w_A)) = p_A w_3(Y^{pg}(a+1,b|c_A,w_A)) + q_A w_3(Y^{pg}(a,b+1|c_A,w_A)) + 3p_A w_2(Y^{pg}(a+1,b|c_A,w_A)) + 3q_A w_2(Y^{pg}(a,b+1|c_A,w_A)) + 3p_A w_1(Y^{pg}(a+1,b|c_A,w_A)) + 3q_A w_1(Y^{pg}(a,b+1|c_A,w_A)) + p_A P^{pg}(a+1,b|c_A,w_A) + q_A P^{pg}(a,b+1|c_A,w_A)$

$w_4(Y^{pg}(a,b|c_A,w_A)) = p_A w_4(Y^{pg}(a+1,b|c_A,w_A)) + q_A w_4(Y^{pg}(a,b+1|c_A,w_A)) + 4p_A w_3(Y^{pg}(a+1,b|c_A,w_A)) + 4q_A w_3(Y^{pg}(a,b+1|c_A,w_A)) + 6p_A w_2(Y^{pg}(a+1,b|c_A,w_A)) + 6q_A w_2(Y^{pg}(a,b+1|c_A,w_A)) + 4p_A w_1(Y^{pg}(a+1,b|c_A,w_A)) + 4q_A w_1(Y^{pg}(a,b+1|c_A,w_A)) + p_A P^{pg}(a+1,b|c_A,w_A) + q_A P^{pg}(a,b+1|c_A,w_A)$

Boundary Values:

$w_n(Y^{pg}(a,b|c_A,w_A)) = 0$, if $a = 4$ and $0 \leq b \leq 2$; $b = 4$ and $0 \leq a \leq 2$

The boundary values for $w_n(Y^{pg}(3,3|c_A,w_A))$ are obtained as follows:

If the point score reaches $(3,3)$ in a game, the first 'deuce', then the game continues. At least two points must be played after deuce to complete the game. For the point played after $(3,3)$ we use a mixture law to obtain:

$W_{Y^{pg}(3,3|c_A,w_A)}(t) = W_{Y^{pp}(()|c_A,w_A)}(t) W_{Y^{pg}(4,3|c_A,w_A)}(t) + W_{Y^{pp}(()|c_A,l_A)}(t) W_{Y^{pg}(3,4|c_A,w_A)}(t)$

This can be simplified to:

$W_{Y^{pg}(3,3|c_A,w_A)}(t) = e^t(p_A W_{Y^{pg}(4,3|c_A,w_A)}(t) + q_A W_{Y^{pg}(3,4|c_A,w_A)}(t))$

Likewise for the following point played we obtain:

$W_{Y^{pg}(4,3|c_A,w_A)}(t) = e^t(p_A W_{Y^{pg}(5,3|c_A,w_A)}(t) + q_A W_{Y^{pg}(4,4|c_A,w_A)}(t))$

$W_{Y^{pg}(3,4|c_A,w_A)}(t) = e^t(p_A W_{Y^{pg}(4,4|c_A,w_A)}(t) + q_A W_{Y^{pg}(3,5|c_A,w_A)}(t))$

leading to:

$W_{Y^{pg}(3,3|c_A,w_A)}(t) = e^{2t}(p_A^2 W_{Y^{pg}(5,3|c_A,w_A)}(t) + 2p_A q_A W_{Y^{pg}(4,4|c_A,w_A)}(t) + q_A^2 W_{Y^{pg}(3,5|c_A,w_A)}(t))$

Now the scores $(5,3)$ and $(3,5)$ are boundary scores, so this relation simplifies further to:

$W_{Y^{pg}(3,3|c_A,w_A)}(t) = e^{2t}(p_A^2 + 2p_A q_A W_{Y^{pg}(4,4|c_A,w_A)}(t))$

There is an equivalence of the game situation at each deuce: player A continues to serve with a constant probability of winning a point, and the game only concludes when one of the players has an advantage of at least 2 points. This implies that there is a constant probability of the server winning the game after each deuce. The equivalence of the game situation at each deuce also implies that there is a common distribution in the number of points played after each deuce. We can express this equivalence in terms of the weighted moment generating functions, as:

$W_{Y^{pg}(n,n|c_A,w_A)}(t) = W_{Y^{pg}(n+1,n+1|c_A,w_A)}(t)$, for $n \geq 3$.

It follows that:

$W_{Y^{pg}(n,n|c_A,w_A)}(t) = e^{2t}(p_A^2 + 2p_A q_A W_{Y^{pg}(n,n|c_A,w_A)}(t))$, for $n \geq 3$.

which simplifies to:

$W_{Y^{pg}(n,n|c_A,w_A)}(t) = \frac{e^{2t}p_A^2}{1-2p_A q_A e^{2t}}$, for $n \geq 3$.

Therefore:

$w_1(Y^{pg}(a,b|c_A,w_A)) = \frac{2p_A^2}{(2p_A^2-2p_A+1)^2}$

$w_2(Y^{pg}(a,b|c_A,w_A)) = \frac{4p_A^2(1-2p_A^2+2p_A)}{(2p_A^2-2p_A+1)^3}$

$w_3(Y^{pg}(a,b|c_A,w_A)) = \frac{8p_A^2(4p_A^4-8p_A^3-4p_A^2+8p_A+1)}{(2p_A^2-2p_A+1)^4}$

$w_4(Y^{pg}(a,b|c_A,w_A)) = \frac{16p_A^2(1-2p_A^2+2p_A)(4p_A^4-8p_A^3-16p_A^2+20p_A+1)}{(2p_A^2-2p_A+1)^5}$

Tables 7.1, 7.2, 7.3 and 7.4 represent the weighted first, second, third and fourth moments respectively of the number of points remaining in a game at various score lines given player A is serving and wins the game for $p_A = 0.6$.

		B score				
		0	15	30	40	game
	0	4.7	3.7	2.4	1.0	0
	15	4.1	3.6	2.7	1.5	0
A score	30	3.1	2.9	2.7	2.0	0
	40	1.7	1.7	1.9	2.7	
	game	0	0	0		

Table 7.1: The weighted first moment of the number of points remaining in a game at various score lines given player A is serving and wins the game for $p_A = 0.6$

		B score				
		0	15	30	40	game
	0	34.6	27.7	18.5	8.1	0
	15	24.9	22.5	18.1	10.3	0
A score	30	14.1	14.8	15.2	12.7	0
	40	4.8	6.1	9.1	15.2	
	game	0	0	0		

Table 7.2: The weighted second moment of the number of points remaining in a game at various score lines given player A is serving and wins the game for $p_A = 0.6$

Similar recursion formulas with boundary values can be obtained for $w_n(Y^{pg}(a,b|c_A,l_A))$.

Let $M_{Y^{pg}(a,b|c_A)}(t)$ represent the moment generating function of the number of points remaining in a game at point score (a,b) for player A serving.

		B score				
		0	15	30	40	game
	0	309.6	254.6	173.5	75.8	0
	15	195.3	187.7	157.5	90.9	0
A score	30	95.2	111.8	124.1	106.9	0
	40	27.6	42.9	71.9	124.1	
	game	0	0	0		

Table 7.3: The weighted third moment of the number of points remaining in a game at various score lines given player A is serving and wins the game for $p_A = 0.6$

		B score				
		0	15	30	40	game
	0	3418.1	2883.2	1986.4	864.7	0
	15	2027.1	2037.0	1746.9	1009.8	0
A score	30	941.4	1182.1	1352.1	1170.5	0
	40	268.4	453.2	780.9	1352.1	
	game	0	0	0		

Table 7.4: The weighted fourth moment of the number of points remaining in a game at various score lines given player A is serving and wins the game for $p_A = 0.6$

Using the rule for combining weighted moment generating functions with mutually exclusive conditions we obtain

$M_{Y^{pg}(a,b|c_A)}(t) = W_{Y^{pg}(a,b|c_A,w_A)}(t) + W_{Y^{pg}(a,b|c_A,l_A)}(t)$

since the probability that a game will eventually end is 1.

Converting moments to parameters of distribution (mean, variance, coefficients of skewness and excess kurtosis) can readily be obtained by the formulas given in section 4.3.

Similar formulas and parameters of distribution can be obtained for when player B is serving such that $W_{Y^{pg}(a,b|c_B,w_B)}(t)$ and $W_{Y^{pg}(a,b|c_B,l_B)}(t)$ represent the weighted moment generating functions of the number of points remaining in the game from score line (a, b) given player B is serving and player B wins and loses the game respectively.

7.3 Number of points in a tiebreak game

The analysis of a tiebreak game is similar to that of a standard game except that it is necessary to allow for the rotation of service before each odd point in the tiebreak game.

Let $P^{pg_T}(a,b|s_A,w_A)$ represent the probability of player A winning a tiebreak game at point score (a,b) given player A served first in the game.

Recurrence Formulas:

$P^{pg_T}(a,b|s_A,w_A) = p_A P^{pg_T}(a+1,b|s_A,w_A) + q_A P^{pg_T}(a,b+1|s_A,w_A)$,

if $(a+b)$ mod $4 = 0$ or 3

$P^{pg_T}(a,b|s_A,w_A) = q_B P^{pg_T}(a+1,b|s_A,w_A) + p_B P^{pg_T}(a,b+1|s_A,w_A)$,

if $(a+b)$ mod $4 = 1$ or 2

Boundary Values:

$P^{pg_T}(a,b|s_A,w_A) = 1$, if $a = 7$ and $0 \le b \le 5$; $b = 7$ and $0 \le a \le 5$

$P^{pg_T}(6,6|s_A,w_A) = \frac{p_A(p_B-1)}{2p_Ap_B-p_B-p_A}$

Table 7.5 represents the probability of player A winning a tiebreak game at various score lines given player A serves first in the game, for $p_A = 0.62$ and $p_B = 0.60$.

Let $W_{Y^{pg_T}(a,b|s_A,w_A)}(t)$ represent the weighted moment generating function of the number of points remaining in the tiebreak game from score line (a,b) given player A served first and wins the game.

$W_{Y^{pg_T}(a,b|s_A,w_A)}(t) = W_{Y^{pp}(()|c_A,w_A)}(t)W_{Y^{pg_T}(a+1,b|s_A,w_A)}(t) + W_{Y^{pp}(()|c_A,l_A)}(t)W_{Y^{pg_T}(a,b+1|s_A,w_A)}(t)$,

for $(a+b)$ mod $4 = 0$ or 3

$W_{Y^{pg_T}(a,b|s_A,w_A)}(t) = W_{Y^{pp}(()|c_B,l_B)}(t)W_{Y^{pg_T}(a+1,b|s_A,w_A)}(t) + W_{Y^{pp}(()|c_B,w_B)}(t)W_{Y^{pg_T}(a,b+1|s_A,w_A)}(t)$,

for $(a+b)$ mod $4 = 1$ or 2

		B score							
		0	1	2	3	4	5	6	7
	0	0.533	0.389	0.295	0.205	0.096	0.031	0.008	0
	1	0.620	0.530	0.431	0.271	0.135	0.066	0.020	0
	2	0.755	0.680	0.528	0.355	0.240	0.134	0.032	0
A score	3	0.868	0.773	0.635	0.526	0.399	0.197	0.052	0
	4	0.926	0.858	0.798	0.716	0.523	0.286	0.129	0
	5	0.967	0.949	0.921	0.834	0.669	0.521	0.323	0
	6	0.994	0.991	0.975	0.934	0.891	0.818	0.521	
	7	1	1	1	1	1	1		

Table 7.5: The probability of player A winning a tiebreak game at various score lines given player A serves first in the game, for $p_A = 0.62$ and $p_B = 0.60$

Let $w_n(Y^{pg_T}(a,b|s_A,w_A))$ represent the weighted n^{th} moment of the number of points remaining in a tiebreak game at point score (a,b) given player A served first and wins the game.

Recurrence Formulas:

For $(a+b)$ mod $4 = 0$ or 3:

$w_1(Y^{pg_T}(a,b|s_A,w_A)) = p_A w_1(Y^{pg_T}(a+1,b|s_A,w_A)) + q_A w_1(Y^{pg_T}(a,b+1|s_A,w_A)) + p_A P^{pg_T}(a+1,b|s_A,w_A) + q_A P^{pg_T}(a,b+1|s_A,w_A)$

$w_2(Y^{pg_T}(a,b|s_A,w_A)) = p_A w_2(Y^{pg_T}(a+1,b|s_A,w_A)) + q_A w_2(Y^{pg_T}(a,b+1|s_A,w_A)) + 2p_A w_1(Y^{pg_T}(a+1,b|s_A,w_A)) + 2q_A w_1(Y^{pg_T}(a,b+1|s_A,w_A)) + p_A P^{pg_T}(a+1,b|s_A,w_A) + q_A P^{pg_T}(a,b+1|s_A,w_A)$

$w_3(Y^{pg_T}(a,b|s_A,w_A)) = p_A w_3(Y^{pg_T}(a+1,b|s_A,w_A)) + q_A w_3(Y^{pg_T}(a,b+1|s_A,w_A)) + 3p_A w_2(Y^{pg_T}(a+1,b|s_A,w_A)) + 3q_A w_2(Y^{pg_T}(a,b+1|s_A,w_A)) + 3p_A w_1(Y^{pg_T}(a+1,b|s_A,w_A)) + 3q_A w_1(Y^{pg_T}(a,b+1|s_A,w_A)) + p_A P^{pg_T}(a+1,b|s_A,w_A) + q_A P^{pg_T}(a,b+1|s_A,w_A)$

$w_4(Y^{pg_T}(a,b|s_A,w_A)) = p_A w_4(Y^{pg_T}(a+1,b|s_A,w_A)) + q_A w_4(Y^{pg_T}(a,b+1|s_A,w_A)) + 4p_A w_3(Y^{pg_T}(a+1,b|s_A,w_A)) + 4q_A w_3(Y^{pg_T}(a,b+1|s_A,w_A)) + 6p_A w_2(Y^{pg_T}(a+1,b|s_A,w_A)) + 6q_A w_2(Y^{pg_T}(a,b+$

$1|s_A, w_A)) + 4p_A w_1(Y^{pg_T}(a+1, b|s_A, w_A)) + 4q_A w_1(Y^{pg_T}(a, b+1|s_A, w_A)) + p_A P^{pg_T}(a+1, b|s_A, w_A) + q_A P^{pg_T}(a, b+1|s_A, w_A)$

For $(a+b) \bmod 4 = 1$ or 2:

$w_1(Y^{pg_T}(a, b|s_A, w_A)) = q_B w_1(Y^{pg_T}(a+1, b|s_A, w_A)) + p_B w_1(Y^{pg_T}(a, b+1|s_A, w_A)) + q_B P^{pg_T}(a+1, b|s_A, w_A) + p_B P^{pg_T}(a, b+1|s_A, w_A)$

$w_2(Y^{pg_T}(a, b|s_A, w_A)) = q_B w_2(Y^{pg_T}(a+1, b|s_A, w_A)) + p_B w_2(Y^{pg_T}(a, b+1|s_A, w_A)) + 2q_B w_1(Y^{pg_T}(a+1, b|s_A, w_A)) + 2p_B w_1(Y_P^{pg_T}(a, b+1|s_A, w_A)) + q_B P^{pg_T}(a+1, b|s_A, w_A) + p_B P^{pg_T}(a, b+1|s_A, w_A)$

$w_3(Y^{pg_T}(a, b|s_A, w_A)) = q_B w_3(Y^{pg_T}(a+1, b|s_A, w_A)) + p_B w_3(Y^{pg_T}(a, b+1|s_A, w_A)) + 3q_B w_2(Y^{pg_T}(a+1, b|s_A, w_A)) + 3p_B w_2(Y^{pg_T}(a, b+1|s_A, w_A)) + 3q_B w_1(Y^{pg_T}(a+1, b|s_A, w_A)) + 3p_B w_1(Y^{pg_T}(a, b+1|s_A, w_A)) + q_B P^{pg_T}(a+1, b|s_A, w_A) + p_B P^{pg_T}(a, b+1|s_A, w_A)$

$w_4(Y^{pg_T}(a, b|s_A, w_A)) = q_B w_4(Y^{pg_T}(a+1, b|s_A, w_A)) + p_B w_4(Y^{pg_T}(a, b+1|s_A, w_A)) + 4q_B w_3(Y^{pg_T}(a+1, b|s_A, w_A)) + 4p_B w_3(Y^{pg_T}(a, b+1|s_A, w_A)) + 6q_B w_2(Y^{pg_T}(a+1, b|s_A, w_A)) + 6p_B w_2(Y^{pg_T}(a, b+1|s_A, w_A)) + 4q_B w_1(Y^{pg_T}(a+1, b|s_A, w_A)) + 4p_B w_1(Y^{pg_T}(a, b+1|s_A, w_A)) + q_B P^{pg_T}(a+1, b|s_A, w_A) + p_B P^{pg_T}(a, b+1|s_A, w_A)$

Boundary Values:

$w_n(Y^{pg_T}(a, b|s_A, w_A)) = 0$, if $a = 7$ and $0 \leq b \leq 5$
$w_n(Y^{pg_T}(a, b|s_A, w_A)) = 0$, if $b = 7$ and $0 \leq a \leq 5$

To find the boundary values for $w_n(Y^{pg_T}(6, 6|s_A, w_A))$ we follow the same argument as used for deuce in the standard game, except that we must allow for the change in server after the odd point has been played. The order of serving is unimportant as each player serves just one of the two points. It can be shown that after playing two points that
$W_{Y^{pg_T}(6,6|s_A,w_A)}(t) = e^{2t} p_A q_B + (p_A p_B + q_A q_B) e^{2t} W_{Y^{pg_T}(7,7|s_A,w_A)}(t)$

From the equivalence of the situation at the scores $(6,6)$ and $(7,7)$ it follows that

$W_{Y^{pg_T}(6,6|s_A,w_A)}(t) = \frac{e^{2t}p_Aq_B}{1-(p_Ap_B+q_Aq_B)e^{2t}}$

Therefore:

$w_1(Y^{pg_T}(6,6|s_A,w_A)) = \frac{2p_A(1-p_B)}{(2p_Ap_B-p_B-p_A)^2}$

$w_2(Y^{pg_T}(6,6|s_A,w_A)) = \frac{4p_A(p_B-1)(2p_Ap_B-p_B-p_A+2)}{(2p_Ap_B-p_B-p_A)^3}$

$w_3(Y^{pg_T}(6,6|s_A,w_A)) = \frac{8p_A(1-p_B)(4p_A^2p_B^2-4p_Ap_B^2+p_B^2-4p_A^2p_B+14p_Ap_B-6p_B+p_A^2-6p_A+6)}{(2p_Ap_B-p_B-p_A)^4}$

$w_4(Y^{pg_T}(6,6|s_A,w_A)) = \frac{16p_A(p_B-1)(2p_Ap_B-p_B-p_A+2)(4p_A^2p_B^2-4p_Ap_B^2+p_B^2-4p_A^2p_B+26p_Ap_B-12p_B+p_A^2-12p_A+12)}{(2p_Ap_B-p_B-p_A)^5}$

Tables 7.6, 7.7, 7.8 and 7.9 represent the weighted first, second, third and fourth moments respectively of the number of points remaining in a tiebreak game at various score lines given player A is serving first and wins the game, given $p_A = 0.62$ and $p_B = 0.60$.

		B score							
		0	1	2	3	4	5	6	7
A score	0	6.3	4.5	3.3	2.2	1.0	0.3	0.1	0
	1	6.5	5.4	4.2	2.6	1.3	0.6	0.2	0
	2	6.7	5.8	4.5	3.0	2.0	1.1	0.3	0
	3	6.2	5.5	4.5	3.6	2.7	1.4	0.4	0
	4	5.2	4.9	4.3	3.8	2.8	1.6	0.8	0
	5	3.9	3.5	3.2	3.2	2.7	2.2	1.7	0
	6	2.0	1.7	1.8	2.1	1.9	1.6	2.2	
	7	0	0	0	0	0	0		

Table 7.6: The weighted first moment of the number of points remaining in a tiebreak game at various score lines for player A serving first and wins the game, given $p_A = 0.62$ and $p_B = 0.60$

Let $W_{Y^{pg_T}(a,b|s_A,l_A)}(t)$ represent the weighted moment generating function of the number of points remaining in a tiebreak game at point score (a,b) given player A served first and loses the game.

Let $M_{Y^{pg_T}(a,b|s_A)}(t)$ represent the moment generating function of the number of points

		B score							
		0	1	2	3	4	5	6	7
	0	79.3	56.1	39.7	25.8	11.7	3.7	0.9	0
	1	74.1	59.1	44.9	27.7	13.5	6.3	1.9	0
	2	65.5	55.0	42.6	28.4	18.3	10.0	2.5	0
A score	3	49.6	44.9	37.6	29.5	21.6	11.3	3.2	0
	4	33.9	32.9	28.6	24.5	19.9	12.3	6.2	0
	5	19.2	17.0	15.1	16.4	16.2	14.0	11.7	0
	6	5.6	4.4	5.5	7.7	7.6	7.8	14.0	
	7	0	0	0	0	0	0		

Table 7.7: The weighted second moment of the number of points remaining in a tiebreak game at various score lines for player A serving first and wins the game, given $p_A = 0.62$ and $p_B = 0.60$

remaining in a tiebreak game at point score (a, b) given player A served first. Since the game must end with probability 1 it follows that:

$M_{Y^{pgT}(a,b|s_A)}(t) = W_{Y^{pgT}(a,b|s_A,w_A)}(t) + W_{Y^{pgT}(a,b|s_A,l_A)}(t)$

Similar formulas can be obtained for when player B is serving such that $W_{Y^{pgT}(a,b|s_B,w_B)}(t)$ and $W_{Y^{pgT}(a,b|s_B,l_B)}(t)$ represent the weighted moment generating functions of the number of points remaining in the tiebreak game from score line (a, b) given player B is serving first and player B wins and loses the game respectively.

Let $M_{Y^{pgT}(a,b|c_A)}(t)$ represent the moment generating function of the number of points remaining in a tiebreak game at point score (a, b) for player A serving.

It follows that:

$M_{Y^{pgT}(a,b|c_A)}(t) = M_{Y^{pgT}(a,b|s_A)}(t)$, if $(a + b)$ mod $4 = 0$ or 3

$M_{Y^{pgT}(a,b|c_A)}(t) = M_{Y^{pgT}(a,b|s_B)}(t)$, if $(a + b)$ mod $4 = 1$ or 2

Converting moments to parameters of distribution (mean, variance, coefficients of skewness and excess kurtosis) can readily be obtained by the formulas given in section 4.3.

		B score							
		0	1	2	3	4	5	6	7
	0	1080.9	759.3	525.4	334.4	150.2	46.3	11.2	0
	1	923.8	722.0	538.0	332.7	162.1	74.0	21.7	0
	2	718.1	593.5	466.3	315.5	202.2	109.9	26.9	0
A score	3	462.4	432.5	373.8	293.9	217.3	117.5	32.7	0
	4	269.4	276.4	243.9	213.3	186.3	120.8	60.7	0
	5	124.9	113.2	105.5	128.1	143.0	130.3	111.2	0
	6	24.9	21.6	32.4	53.4	59.2	68.8	130.3	
	7	0	0	0	0	0	0		

Table 7.8: The weighted third moment of the number of points remaining in a tiebreak game at various score lines for player A serving first and wins the game, given $p_A = 0.62$ and $p_B = 0.60$

Similar formulas and parameters of distribution can be obtained for when player B is serving such that $M_{Y^{pgT}(a,b|c_B)}(t)$ represents the moment generating function of the number of points remaining in the tiebreak game from score line (a, b) given player B is serving.

7.4 Number of points in a tiebreak set

We study here the model for a tiebreak set. To account for the rotation of service in this type of set it is necessary to allow for the rotation of server at the beginning of each game. Using this convention, whenever a tiebreak game is required to resolve the winner of the set, this tiebreak game is marked to the server of the first point of the game, and hence to the server of the first point of the set when it comes to determining the first server of the next set. This rule applies irrespective of the outcome of the tiebreak game.

For player A serving in the first game of the set there are four cases to be dealt with separately. Consider the case where player A not only serves in the first game of the set, but wins the set, and serves in the first game of the next set.

		B score							
		0	1	2	3	4	5	6	7
A score	0	16210.0	11382.1	7768.3	4883.3	2185.1	666.8	157.6	0
	1	12923.5	10007.9	7405.6	4615.8	2253.2	1018.8	295.2	0
	2	9107.3	7526.1	6052.5	4169.1	2674.1	1452.0	352.8	0
	3	5214.9	5095.9	4581.9	3654.2	2741.3	1506.5	417.7	0
	4	2745.8	3021.4	2756.2	2492.5	2288.5	1517.2	761.1	0
	5	1134.3	1101.2	1101.1	1455.9	1734.6	1612.2	1380.5	0
	6	187.2	191.0	330.3	598.4	706.1	846.7	1612.2	
	7	0	0	0	0	0	0		

Table 7.9: The weighted fourth moment of the number of points remaining in a tiebreak game at various score lines for player A serving first and wins the game, given $p_A = 0.62$ and $p_B = 0.60$

Let $W_{Y^{psT}(0,0:c,d|s_A,w_A,n_A)}(t)$ represent the weighted moment generating function of the number of points remaining in a tiebreak set at point and game score $(0, 0 : c, d)$ given player A served first, wins the set and is serving first in the next set to be played. Then by considering a complete game being played at that score we obtain, for $c + d < 12$,

$W_{Y^{psT}(0,0:c,d|s_A,w_A,n_A)}(t)$

$= W_{Y^{pg}(0,0|c_A,w_A)}(t)W_{Y^{psT}(0,0:c+1,d|s_A,w_A,n_A)}(t)+W_{Y^{pg}(0,0|c_A,l_A)}(t)W_{Y^{psT}(0,0:c,d+1|s_A,w_A,n_A)}(t)$, for $(c + d)$ mod 2=0

$W_{Y^{psT}(0,0:c,d|s_A,w_A,n_A)}(t)$

$= W_{Y^{pg}(0,0|c_B,l_B)}(t)W_{Y^{psT}(0,0:c+1,d|s_A,w_A,n_A)}(t)+W_{Y^{pg}(0,0|c_B,w_B)}(t)W_{Y^{psT}(0,0:c,d+1|s_A,w_A,n_A)}(t)$, for $(c + d)$ mod 2=1

There is a special case for the tiebreak game, with $c = 6, d = 6$, where due to the rotation of serve player A cannot serve first in the next set, so

$W_{Y^{psT}(0,0:6,6|s_A,w_A,n_A)}(t) = 0$

However $W_{Y^{psT}(0,0:6,6|s_A,w_A,n_A)}(t)$

$= W_{Y^{pg_T}(0,0|c_A,w_A)}(t) W_{Y^{ps_T}(0,0:7,6|s_A,w_A,n_B)}(t) + W_{Y^{pg_T}(0,0|c_A,l_A)}(t) W_{Y^{ps_T}(0,0:6,7|s_A,w_A,n_A)}(t)$

which simplifies to

$W_{Y^{ps_T}(0,0:6,6|s_A,w_A,n_A)}(t) = W_{Y^{pg_T}(0,0|c_A,w_A)}(t)$

as expected since $W_{Y^{ps_T}(0,0:7,6|s_A,w_A,n_B)}(t) = 1$ and $W_{Y^{ps_T}(0,0:6,7|s_A,w_A,n_A)}(t) = 0$

Let $P^{gs_T}(c,d|s_A,w_A,n_A)$ represent the probability of player A winning a tiebreak set at game score (c,d) given player A served first and is serving first in the next set to be played. Then by setting $t = 0$ in the formulas above we obtain the following results.

Recurrence Formulas:

$P^{gs_T}(c,d|s_A,w_A,n_A) = p_A^g P^{gs_T}(c+1,d|s_A,w_A,n_A) + q_A^g P^{gs_T}(c,d+1|s_A,w_A,n_A)$, if $(c+d)$ mod $2 = 0$

$P^{gs_T}(c,d|s_A,w_A,n_A) = q_B^g P^{gs_T}(c+1,d|s_A,w_A,n_A) + p_B^g P^{gs_T}(c,d+1|s_A,w_A,n_A)$, if $(c+d)$ mod $2 = 1$

Boundary Values:

$P^{gs_T}(c,d|s_A,w_A,n_A) = 1$, if $(6,0)$; $(6,2)$; $(6,4)$; $(7,5)$

$P^{gs_T}(c,d|s_A,w_A,n_A) = 0$, if $d = 6$ and $0 \leq c \leq 4$; $(5,7)$; $(6,1)$; $(6,3)$

$P^{gs_T}(6,6|s_A,w_A,n_A) = 0$

Table 7.10 represents the probability of player A winning a tiebreak set at various score lines given player A served first and is serving first in the next set to be played, for $p_A = 0.62$ and $p_B = 0.60$.

Let $w_n(Y^{ps_T}(0,0:c,d|s_A,w_A,n_A))$ represent the weighted n^{th} moment of the number of points remaining in a tiebreak set at point and game score $(0,0:c,d)$ given player A served first, wins the set and is serving first in the next set to be played. By differentiating n times and setting $t = 0$ we obtain the following results.

		B score							
		0	1	2	3	4	5	6	7
	0	0.209	0.146	0.120	0.047	0.025	0.002	0	
	1	0.227	0.215	0.142	0.110	0.031	0.009	0	
	2	0.261	0.237	0.230	0.132	0.095	0.011	0	
A score	3	0.268	0.255	0.258	0.237	0.119	0.042	0	
	4	0.305	0.254	0.319	0.270	0.334	0.054	0	
	5	0.319	0.075	0.333	0.093	0.415	0.205	0.000	0
	6	1	0	1	0	1	0.264	0	
	7						1		

Table 7.10: The probability of player A winning a tiebreak set at various score lines given player A served first and is serving first in the next set to be played, for $p_A = 0.62$ and $p_B = 0.60$

Recurrence Formulas:

For $(c + d)$ mod $2 = 0$:

$w_1(Y^{ps_T}(0, 0 : c, d|s_A, w_A, n_A))$

$=p^g_A w_1(Y^{ps_T}(0, 0 : c + 1, d|s_A, w_A, n_A)) + q^g_A w_1(Y^{ps_T}_P(0, 0 : c, d + 1|s_A, w_A, n_A))$

$+w_1(Y^{pg}(0, 0|c_A, w_A))P^{gs_T}(c+1, d|s_A, w_A, n_A)+w_1(Y^{pg}(0, 0|c_A, l_A))P^{gs_T}(c, d+1|s_A, w_A, n_A)$

$w_2(Y^{ps_T}(0, 0 : c, d|s_A, w_A, n_A))$

$= p^g_A w_2(Y^{ps_T}(0, 0 : c + 1, d|s_A, w_A, n_A)) + q^g_A w_2(Y^{ps_T}(0, 0 : c, d + 1|s_A, w_A, n_A))$

$+2w_1(Y^{pg}(0, 0|c_A, w_A))w_1(Y^{ps_T}(0, 0 : c + 1, d|s_A, w_A, n_A))$

$+2w_1(Y^{pg}(0, 0|c_A, l_A))w_1(Y^{ps_T}(0, 0 : c, d + 1|s_A, w_A, n_A))$

$+w_2(Y^{pg}(0, 0|c_A, w_A))P^{gs_T}(c+1, d|s_A, w_A, n_A)+w_2(Y^{pg}(0, 0|c_A, l_A))P^{gs_T}(c, d+1|s_A, w_A, n_A)$

$w_3(Y^{ps_T}(0, 0 : c, d|s_A, w_A, n_A))$

$= p^g_A w_3(Y^{ps_T}(0, 0 : c + 1, d|s_A, w_A, n_A)) + q^g_A w_3(Y^{ps_T}(0, 0 : c, d + 1|s_A, w_A, n_A))$

$+3w_1(Y^{pg}(0, 0|c_A, w_A))w_2(Y^{ps_T}(0, 0 : c + 1, d|s_A, w_A, n_A))$

$+3w_1(Y^{pg}(0, 0|c_A, l_A))w_2(Y^{ps_T}(0, 0 : c, d + 1|s_A, w_A, n_A))$

$+3w_2(Y^{pg}(0,0|c_A,w_A))w_1(Y^{ps_T}(0,0:c+1,d|s_A,w_A,n_A))$

$+3w_2(Y^{pg}(0,0|c_A,l_A))w_1(Y^{ps_T}(0,0:c,d+1|s_A,w_A,n_A))$

$+w_3(Y^{pg}(0,0|c_A,w_A))P^{gs_T}(c+1,d|s_A,w_A,n_A)+w_3(Y^{pg}(0,0|c_A,l_A))P^{gs_T}(c,d+1|s_A,w_A,n_A)$

$w_4(Y^{ps_T}(0,0:c,d|s_A,w_A,n_A))$

$= p_A^g w_4(Y^{ps_T}(0,0:c+1,d|s_A,w_A,n_A)) + q_A^g w_4(Y^{ps_T}(0,0:c,d+1|s_A,w_A,n_A))$

$+4w_1(Y^{pg}(0,0|c_A,w_A))w_3(Y^{ps_T}(0,0:c+1,d|s_A,w_A,n_A))$

$+4w_1(Y^{pg}(0,0|c_A,l_A))w_3(Y^{ps_T}(0,0:c,d+1|s_A,w_A,n_A))$

$+6w_2(Y^{pg}(0,0|c_A,w_A))w_2(Y^{ps_T}(0,0:c+1,d|s_A,w_A,n_A))$

$+6w_2(Y^{pg}(0,0|c_A,l_A))w_2(Y^{ps_T}(0,0:c,d+1|s_A,w_A,n_A))$

$+4w_3(Y^{pg}(0,0|c_A,w_A))w_1(Y^{ps_T}(0,0:c+1,d|s_A,w_A,n_A))$

$+4w_3(Y^{pg}(0,0|c_A,l_A))w_1(Y^{ps_T}(0,0:c,d+1|s_A,w_A,n_A))$

$+w_4(Y^{pg}(0,0|c_A,w_A))P^{gs_T}(c+1,d|s_A,w_A,n_A)+w_4(Y^{pg}(0,0|c_A,l_A))P^{gs_T}(c,d+1|s_A,w_A,n_A)$

For $(c+d) \bmod 2 = 1$:

$w_1(Y^{ps_T}(0,0:c,d|s_A,w_A,n_A))$

$= q_B^g w_1(Y^{ps_T}(0,0:c+1,d|s_A,w_A,n_A)) + p_B^g w_1(Y^{ps_T}(0,0:c,d+1|s_A,w_A,n_A))$

$+w_1(Y^{pg}(0,0|c_B,l_B))P^{gs_T}(c+1,d|s_A,w_A,n_A)+w_1(Y^{pg}(0,0|c_B,w_B))P^{gs_T}(c,d+1|s_A,w_A,n_A)$

$w_2(Y^{ps_T}(0,0:c,d|s_A,w_A,n_A))$

$= q_B^g w_2(Y^{ps_T}(0,0:c+1,d|s_A,w_A,n_A)) + p_B^g w_2(Y^{ps_T}(0,0:c,d+1|s_A,w_A,n_A))$

$+2w_1(Y^{pg}(0,0|c_B,l_B))w_1(Y^{ps_T}(0,0:c+1,d|s_A,w_A,n_A))$

$+2w_1(Y^{pg}(0,0|c_B,w_B))w_1(Y^{ps_T}(0,0:c,d+1|s_A,w_A,n_A))$

$+w_2(Y^{pg}(0,0|c_B,l_B))P^{gs_T}(c+1,d|s_A,w_A,n_A)+w_2(Y^{pg}(0,0|c_B,w_B))P^{gs_T}(c,d+1|s_A,w_A,n_A)$

$w_3(Y^{ps_T}(0,0:c,d|s_A,w_A,n_A))$

$= q_B^g w_3(Y^{ps_T}(0,0:c+1,d|s_A,w_A,n_A)) + p_B^g w_3(Y^{ps_T}(0,0:c,d+1|s_A,w_A,n_A))$

$+3w_1(Y^{pg}(0,0|c_B,l_B))w_2(Y^{ps_T}(0,0:c+1,d|s_A,w_A,n_A))$

$+3w_1(Y^{pg}(0,0|c_B,w_B))w_2(Y^{ps_T}(0,0:c,d+1|s_A,w_A,n_A))$

$+3w_2(Y^{pg}(0,0|c_B,l_B))w_1(Y^{ps_T}(0,0:c+1,d|s_A,w_A,n_A))$

$+3w_2(Y^{pg}(0,0|c_B,w_B))w_1(Y^{ps_T}(0,0:c,d+1|s_A,w_A,n_A))$

$+w_3(Y^{pg}(0,0|c_B,l_B))P^{gs_T}(c+1,d|s_A,w_A,n_A)+w_3(Y^{pg}(0,0|c_B,w_B))P^{gs_T}(c,d+1|s_A,w_A,n_A)$

$w_4(Y^{ps_T}(0,0:c,d|s_A,w_A,n_A))$

$=q_B^g w_4(Y^{ps_T}(0,0:c+1,d|s_A,w_A,n_A))+p_B^g w_4(Y^{ps_T}(0,0:c,d+1|s_A,w_A,n_A))$

$+4w_1(Y^{pg}(0,0|c_B,l_B))w_3(Y^{ps_T}(0,0:c+1,d|s_A,w_A,n_A))$

$+4w_1(Y^{pg}(0,0|c_B,w_B))w_3(Y^{ps_T}(0,0:c,d+1|s_A,w_A,n_A))$

$+6w_2(Y^{pg}(0,0|c_B,l_B))w_2(Y^{ps_T}(0,0:c+1,d|s_A,w_A,n_A))$

$+6w_2(Y^{pg}(0,0|c_B,w_B))w_2(Y^{ps_T}(0,0:c,d+1|s_A,w_A,n_A))$

$+4w_3(Y^{pg}(0,0|c_B,l_B))w_1(Y^{ps_T}(0,0:c+1,d|s_A,w_A,n_A))$

$+4w_3(Y^{pg}(0,0|c_B,w_B))w_1(Y^{ps_T}(0,0:c,d+1|s_A,w_A,n_A))$

$+w_4(Y^{pg}(0,0|c_B,l_B))P^{gs_T}(c+1,d|s_A,w_A,n_A)+w_4(Y^{pg}(0,0|c_B,w_B))P^{gs_T}(c,d+1|s_A,w_A,n_A)$

Boundary Values:

$w_n(Y^{ps_T}(0,0:c,d|s_A,w_A,n_A))=0$, if $c=6$ and $0\leq d\leq 4$; $d=6$ and $0\leq c\leq 4$; $(7,5)$; $(5,7)$

$w_n(Y^{ps_T}(0,0:6,6|s_A,w_A,n_A))=0$

Tables 7.11, 7.12, 7.13 and 7.14 represent the weighted first, second, third and fourth moments respectively of the number of points remaining in a tiebreak set at various score lines given player A is serving first, wins the game and is serving first in the next set to be played, for p_A=0.62 and p_B=0.60.

Similar formulations can be obtained for $w_n(Y^{ps_T}(0,0:c,d|s_A,w_A,n_B))$, $w_n(Y^{ps_T}(0,0:c,d|s_A,l_A,n_A))$, $w_n(Y^{ps_T}(0,0:c,d|s_A,l_A,n_B))$, $w_n(Y^{ps_T}(0,0:c,d|s_B,w_B,n_A))$, $w_n(Y^{ps_T}(0,0:c,d|s_B,w_B,n_B))$, $w_n(Y^{ps_T}(0,0:c,d|s_B,l_B,n_A))$ and $w_n(Y^{ps_T}(0,0:c,d|s_B,l_B,n_B))$.

		B score							
		0	1	2	3	4	5	6	7
A score	0	13.3	8.9	6.8	2.5	1.2	0.1	0	
	1	12.8	11.3	7.1	5.0	1.3	0.3	0	
	2	11.5	10.7	9.4	5.1	3.2	0.4	0	
	3	9.6	8.7	8.7	7.3	3.2	1.1	0	
	4	5.8	6.5	6.2	6.6	6.0	1.1	0	
	5	3.1	1.3	3.5	1.7	4.7	2.7	0.0	0
	6	0	0	0	0	0	1.8	0	
	7						0		

Table 7.11: The weighted first moment of the number of points remaining in a tiebreak set at various score lines for player A serving first, wins the set and serving first in the next set given $p_A = 0.62$ and $p_B = 0.60$

Let $W_{Y^{ps_T}(a,b:c,d|s_A,w_A,n_A)}(t)$ represent the weighted moment generating function of the number of points remaining in a tiebreak set at point and game score $(a, b : c, d)$ given player A served first, wins the set and is serving first in the next set to be played. Then by considering the current game being completed at that score we obtain, for $c + d < 12$,

$W_{Y^{ps_T}(a,b:c,d|s_A,w_A,n_A)}(t)$
$= W_{Y^{pg}(a,b|c_A,w_A)}(t)W_{Y^{ps_T}(0,0:c+1,d|s_A,w_A,n_A)}(t)+W_{Y^{pg}(a,b|c_A,l_A)}(t)W_{Y^{ps_T}(0,0:c,d+1|s_A,w_A,n_A)}(t)$, for $(c + d)$ mod 2=0

$W_{Y^{ps_T}(a,b:c,d|s_A,w_A,n_A)}(t)$
$= W_{Y^{pg}(a,b|c_B,l_B)}(t)W_{Y^{ps_T}(0,0:c+1,d|s_A,w_A,n_A)}(t)+W_{Y^{pg}(a,b|c_B,w_B)}(t)W_{Y^{ps_T}(0,0:c,d+1|s_A,w_A,n_A)}(t)$, for $(c + d)$ mod 2=1

The special case for the tiebreak game, with $c = 6, d = 6$, where due to the rotation of serve player A cannot serve first in the next set, so
$W_{Y^{ps_T}(a,b:6,6|s_A,w_A,n_A)}(t) = 0$

Let $w_n(Y^{ps_T}(a, b : c, d|s_A, w_A, n_A))$ represent the weighted n^{th} moment of the number of

		B score							
		0	1	2	3	4	5	6	7
A score	0	879.9	567.4	397.2	136.1	58.4	4.9	0	
	1	762.3	619.8	367.4	233.7	55.7	13.6	0	
	2	560.4	518.0	409.3	209.4	111.9	12.4	0	
	3	394.2	326.9	324.2	244.0	93.6	29.8	0	
	4	146.4	196.3	155.0	178.2	127.0	22.5	0	
	5	49.3	27.9	59.6	35.9	74.1	37.7	0.0	0
	6	0	0	0	0	0	14.1	0	
	7						0		

Table 7.12: The weighted second moment of the number of points remaining in a tiebreak set at various score lines for player A serving first, wins the set and serving first in the next set given $p_A = 0.62$ and $p_B = 0.60$

points remaining in a tiebreak set at point and game score $(a, b : c, d)$ given player A served first and wins the set, and player A is serving first in the next set to be played.

$w_1(Y^{ps_T}(a, b : c, d|s_A, w_A, n_A)) = P^{pg}(a, b|c_A, w_A)w_1(Y^{ps_T}(a, b : c + 1, d|s_A, w_A, n_A)) + P^{pg}(a, b|c_A, l_A)w_1(Y^{ps_T}(a, b : c, d+1|s_A, w_A, n_A))+w_1(Y^{pg}(a, b|c_A, w_A))P^{gs_T}(c+1, d|s_A, w_A, n_A)+ w_1(Y^{pg}(a, b|c_A, l_A))P^{gs_T}(c, d + 1|s_A, w_A, n_A)$, if $(c + d)$ mod 2=0 and $(c, d) \neq (6, 6)$

$w_1(Y^{ps_T}(a, b : c, d|s_A, w_A, n_A)) = P^{pg}(a, b|c_B, l_B)w_1(Y^{ps_T}(a, b : c + 1, d|s_A, w_A, n_A))+ P^{pg}(a, b|c_B, w_B)w_1(Y^{ps_T}(a, b : c, d+1|s_A, w_A, n_A))+w_1(Y^{pg}(a, b|c_B, l_B))P^{gs_T}(c+1, d|s_A, w_A, n_A)+ w_1(Y^{pg}(a, b|c_B, w_B))P^{gs_T}(c, d + 1|s_A, w_A, n_A)$, if $(c + d)$ mod 2=1

$w_1(Y^{ps_T}(a, b : c, d|s_A, w_A, n_A)) = 0$, if $(c, d) = (6, 6)$

However $w_1(Y^{ps_T}(a, b : c, d|s_A, w_A, n_A)) = w_1(Y^{pg_T}(a, b|c_A, w_A))$, if $(c, d) = (6, 6)$

Similar formulations can be obtained for $w_2(Y^{ps_T}(a, b : c, d|s_A, w_A, n_A))$, $w_3(Y^{ps_T}(a, b : c, d|s_A, w_A, n_A))$ and $w_4(Y^{ps_T}(a, b : c, d|s_A, w_A, n_A))$.

Let $W_{Y^{ps_T}(a,b:c,d|s_A,w_A)}(t)$ represent the weighted moment generating function of the number

		B score							
		0	1	2	3	4	5	6	7
A score	0	60739.5	37306.4	24029.3	7658.2	2979.3	237.2	0	
	1	47603.8	35761.6	19927.0	11460.4	2503.0	561.3	0	
	2	29504.8	26559.5	19009.3	8977.3	4203.8	436.1	0	
	3	17969.1	13666.9	13051.4	8687.7	2962.8	843.5	0	
	4	4792.0	6830.8	4747.3	5280.0	3071.6	497.0	0	
	5	1242.5	764.0	1482.1	871.5	1429.1	578.4	0.0	0
	6	0	0	0	0	0	131.7	0	
	7						0		

Table 7.13: The weighted third moment of the number of points remaining in a tiebreak set at various score lines for player A serving first, wins the set and serving first in the next set given $p_A = 0.62$ and $p_B = 0.60$

of points remaining in a tiebreak set at point and game score $(a, b : c, d)$ given player A served first and wins the set.

It follows that:

$$W_{Y^{ps_T}(a,b:c,d|s_A,w_A)}(t) = W_{Y^{ps_T}(a,b:c,d|s_A,w_A,n_A)}(t) + W_{Y^{ps_T}(a,b:c,d|s_A,w_A,n_B)}(t)$$

Let $M_{Y^{ps_T}(a,b:c,d|s_A)}(t)$ represent the moment generating function of the number of points remaining in a tiebreak set at point and game score $(a, b : c, d)$ given player A served first in the set.

Since the probability that the set is completed when player A serves first is 1, it follows that:

$$M_{Y^{ps_T}(a,b:c,d|s_A)} = W_{Y^{ps_T}(a,b:c,d|s_A,w_A)} + W_{Y^{ps_T}(a,b:c,d|s_A,l_A)}$$

Tables 7.15, 7.16, 7.17 and 7.18 represent the mean $\mu(Y^{ps_T}(0, 0 : c, d|s_A))$, variance $\sigma^2(Y^{ps_T}(0, 0 : c, d|s_A))$, coefficient of skewness $\gamma_1(Y^{ps_T}(0, 0 : c, d|s_A))$ and coefficient of excess kurtosis $\gamma_2(Y^{ps_T}(0, 0 : c, d|s_A))$ respectively of the number of points remaining in a

		B score							
		0	1	2	3	4	5	6	7
	0	4348830.1	2531649.9	1501380.0	444617.0	157208.0	11688.1	0	
	1	3104364.0	2153880.6	1122411.8	584095.8	116913.3	23855.7	0	
	2	1664140.0	1434402.9	930984.2	403331.0	166587.0	15831.8	0	
A score	3	895225.4	623596.8	561190.3	329545.6	100227.7	24938.1	0	
	4	190829.8	268631.6	167692.3	169887.7	82637.9	11657.6	0	
	5	42033.6	25258.0	45019.9	23546.0	31371.9	9766.5	0.0	0
	6	0	0	0	0	0	1491.8	0	
	7						0		

Table 7.14: The weighted fourth moment of the number of points remaining in a tiebreak set at various score lines for player A serving first, wins the set and serving first in the next set given $p_A = 0.62$ and $p_B = 0.60$

tiebreak set at point and game score $(0, 0 : c, d)$ given player A served first, and $p_A = 0.62$, $p_B = 0.60$.

		B score						
		0	1	2	3	4	5	6
	0	65.3	58.7	51.5	37.5	27.8	10.3	
	1	58.9	54.3	47.3	39.8	24.6	14.6	
	2	47.2	48.1	43.8	36.0	27.9	10.6	
A score	3	38.8	35.6	37.8	33.9	24.7	15.6	
	4	22.9	26.8	24.1	28.3	25.7	11.9	
	5	13.6	9.7	14.7	11.2	21.4	20.3	9.6
	6						15.2	11.9

Table 7.15: The mean number of points remaining in a tiebreak set at $(0, 0 : c, d)$ given player A served first, and $p_A = 0.62$, $p_B = 0.60$

7.5 Number of points in an advantage set

We study here the model for an advantage set. Most of the recurrence formulas and boundary conditions are similar to those for a tiebreak set. However, since the advantage

		B score						
		0	1	2	3	4	5	6
	0	256.6	244.7	238.2	220.7	175.8	74.5	
	1	251.9	229.6	215.6	206.0	154.8	90.1	
	2	256.5	225.6	202.6	186.3	169.7	85.4	
A score	3	235.5	220.7	198.9	177.3	144.6	109.0	
	4	140.8	189.1	170.0	176.6	136.7	103.6	
	5	76.6	72.5	115.2	117.1	119.9	49.4	38.5
	6						38.7	9.2

Table 7.16: The variance of the number of points remaining in a tiebreak set at $(0, 0 : c, d)$ given player A served first, and $p_A = 0.62$, $p_B = 0.60$

		B score						
		0	1	2	3	4	5	6
	0	0.54	0.56	0.59	0.95	1.23	2.54	
	1	0.55	0.57	0.63	0.70	1.44	2.15	
	2	0.79	0.57	0.57	0.70	0.86	2.38	
A score	3	0.97	0.93	0.55	0.46	0.94	1.46	
	4	1.87	1.21	1.28	0.35	0.21	1.41	
	5	2.55	2.98	1.97	1.91	-0.04	0.25	1.22
	6						-0.01	1.90

Table 7.17: The coefficient of skewness of the number of points remaining in a tiebreak set at $(0, 0 : c, d)$ given player A served first, and $p_A = 0.62$, $p_B = 0.60$

rule applies in an advantage set (compared to a tiebreak game at 6 games-all in a tiebreak set), this requires obtaining explicit expressions for the boundary values at 5 games-all.

For player A serving in the first game of the set there are four cases to be dealt with separately due to the rotation of serve in and between sets. Consider the case where player A not only serves in the first game of the advantage set, but wins the set, and serves in the first game of the next set.

Let $W_{Y^{ps}(a,b:c,d|s_A,w_A,n_A)}(t)$ represent the weighted moment generating function of the number of points remaining in a tiebreak set at point and game score $(a, b : c, d)$ given player

		B score						
		0	1	2	3	4	5	6
A score	0	-0.22	-0.17	-0.12	0.68	1.73	8.80	
	1	-0.26	-0.32	-0.22	-0.11	1.98	5.86	
	2	0.08	-0.38	-0.51	-0.23	0.09	6.27	
	3	0.48	0.24	-0.62	-0.77	-0.06	1.40	
	4	4.27	1.05	0.80	-0.95	-0.97	0.82	
	5	9.58	11.18	3.54	2.62	-0.91	-0.19	0.74
	6						0.03	6.28

Table 7.18: The coefficient of excess kurtosis of the number of points remaining in a tiebreak set at $(0, 0 : c, d)$ given player A served first, and $p_A = 0.62$, $p_B = 0.60$

A served first, wins the set and is serving first in the next set to be played.

Then by considering the completion of the game being played at that score we obtain, for $0 \leq c \leq 5$ and $0 \leq d \leq 5$,

$W_{Y^{ps}(a,b:c,d|s_A,w_A,n_A)}(t)$

$= W_{Y^{pg}(a,b|c_A,w_A)}(t)W_{Y^{ps}(0,0:c+1,d|s_A,w_A,n_A)}(t) + W_{Y^{pg}(a,b|c_A,l_A)}(t)W_{Y^{ps}(0,0:c,d+1|s_A,w_A,n_A)}(t)$, for $(c+d)$ mod 2=0

$W_{Y^{ps}(a,b:c,d|s_A,w_A,n_A)}(t)$

$= W_{Y^{pg}(a,b|c_B,l_B)}(t)W_{Y^{ps}(0,0:c+1,d|s_A,w_A,n_A)}(t) + W_{Y^{pg}(a,b|c_B,w_B)}(t)W_{Y^{ps}(0,0:c,d+1|s_A,w_A,n_A)}(t)$, for $(c+d)$ mod 2=1

Similar equations apply when $c \geq 5$ and $d \geq 5$, provided the additional restriction $|c-d| < 2$ holds, otherwise the set is completed. The boundary conditions can be readily stated as weighted moment generating functions, in the form

$W_{Y^{ps}(0,0:c,d|s_A,w_A,n_A)}(t) = 1$, for $c = 6$ and $d = 0$; $c = 6$ and $d = 2$; $c = 6$ and $d = 4$; $c > 6$ and $c - d = 2$

$W_{Y^{ps}(0,0:c,d|s_A,w_A,n_A)}(t) = 0$, for $c = 6$ and $d = 1$; $c = 6$ and $d = 3$; $0 \leq c \leq 4$, $d > 6$ and $d - c = 2$

However in this form there is a gap at the boundary conditions at infinity. We can close this gap if we find an explicit expression for $W_{Y^{ps}(0,0:5,5|s_A,w_A,n_A)}(t)$. It will turn out that this is more difficult than it seems at first glance, so we will examine several possible approaches to this problem.

The first approach is to follow the strategy which we have used more than once before, e.g. at the point score $(3,3)$ in the standard game, and at the point score $(6,6)$ in the tiebreaker game. We attempt to use the equivalence of the situation at the game score of $(5,5)$ with that at the game score of $(6,6)$. Using the formulas above we find that

$$W_{Y^{ps}(0,0:5,5|s_A,w_A,n_A)}(t)$$
$$= W_{Y^{pg}(0,0|c_A,w_A)}(t)W_{Y^{ps}(0,0:6,5|s_A,w_A,n_A)}(t) + W_{Y^{pg}(0,0|c_A,l_A)}(t)W_{Y^{ps}(0,0:5,6|s_A,w_A,n_A)}(t)$$
$$W_{Y^{ps}(0,0:6,5|s_A,w_A,n_A)}(t)$$
$$= W_{Y^{pg}(0,0|c_B,l_B)}(t)W_{Y^{ps}(0,0:7,5|s_A,w_A,n_A)}(t) + W_{Y^{pg}(0,0|c_B,w_B)}(t)W_{Y^{ps}(0,0:6,6|s_A,w_A,n_A)}(t)$$
$$W_{Y^{ps}(0,0:5,6|s_A,w_A,n_A)}(t)$$
$$= W_{Y^{pg}(0,0|c_B,l_B)}(t)W_{Y^{ps}(0,0:6,6|s_A,w_A,n_A)}(t) + W_{Y^{pg}(0,0|c_B,w_B)}(t)W_{Y^{ps}(0,0:5,7|s_A,w_A,n_A)}(t)$$

Using the boundary conditions

$W_{Y^{ps}(0,0:7,5|s_A,w_A,n_A)}(t) = 1$ and

$W_{Y^{ps}(0,0:5,7|s_A,w_A,n_A)}(t) = 0$

together with the equivalence

$$W_{Y^{ps}(0,0:5,5|s_A,w_A,n_A)}(t) = W_{Y^{ps}(0,0:6,6|s_A,w_A,n_A)}(t)$$

the expression simplifies to

$$W_{Y^{ps}(0,0:5,5|s_A,w_A,n_A)}(t) = W_{Y^{pg}(0,0|c_A,w_A)}(t)W_{Y^{pg}(0,0|c_B,l_B)}(t) + (W_{Y^{pg}(0,0|c_A,w_A)}(t)W_{Y^{pg}(0,0|c_B,w_B)}(t) + W_{Y^{pg}(0,0|c_A,l_A)}(t)W_{Y^{pg}(0,0|c_B,l_B)}(t))W_{Y^{ps}(0,0:5,5|s_A,w_A,n_A)}(t)$$

It seems that all that remains is to write $W_{Y^{ps}(0,0:5,5|s_A,w_A,n_A)}(t)$ in the form of a fraction and we are done. However there is a practical difficulty of extracting the probability and the first four weighted moments from this expression. When we differentiate a product of three functions of t four times we finish up with eighty-one terms, some of which are equal. There is a high risk of clerical error in manipulating this many terms unless a computer algebra package is employed. So we abandon this approach and contemplate other ways to solve the problem, which could be readily implemented in a spreadsheet.

We know that the weighted moment generating function for the remaining points when the game score is at $(5, 5)$ is the same as when the game score is at $(6, 6)$. If we guess the values of the probability and weighed moments at $(6, 6)$ we can use backward recursion to obtain the values at $(5, 5)$, and thus check that the guesses are correct or not. By a process of successive refinement, starting from the probability of winning first, followed by the lower moments, we can arrive at the correct numerical values in a particular instance. We can even check that the values at the game score $(4, 4)$ are the same. The weakness of this approach is that we still lack a general formula for the results obtained in this manner. To overcome this weakness we now seek to develop suitable formulas based on successive approximation.

Let $R(t)$ denote the weighted moment generating function for the number of points played in two games that lead to from level scores to a repetition of level scores. Let r denote the probability that this repetition occurs.

The order of service for these two games does not matter as each player must serve for just one of the games. The repetition occurs when either both players win whilst serving or both players lose whilst serving. Thus
$r = p_A^g p_B^g + q_A^g q_B^g$

where p_A^g and p_B^g are the probabilities of player A and player B winning a game respectively, with $q_A^g = 1 - p_A^g$ and $q_B^g = 1 - p_B^g$.

Furthermore the number of points played in the first game is independent of the second, and this enables us to multiply the weighted moment generating functions to add the number of points played in the various cases. So

$R(t) = W_{Y^{pg}(()|s_A,w_A)}(t)W_{Y^{pg}(()|s_B,w_B)}(t) + W_{Y^{pg}(()|s_A,l_A)}(t)W_{Y^{pg}(()|s_B,l_B)}(t)$

The probability of no repetition is $1 - r$. The weighted moment generating function for the points played with no repetition is $(1 - r)e^0 = 1 - r$. The weighted moment generating function for the points played with exactly n repetitions is $(1 - r)R^n(t)$, for $n > 0$. Let T denote the total number of points played with repetition of level scores. Then the moment generating function of T is given by

$M_T(t) = (1 - r)(1 + R(t) + R^2(t) + R^3(t) + R^4(t) + ...) = \frac{1-r}{1-R(t)}$

We can recover the n^{th} moment by differentiating n times and setting $t = 0$.

Perhaps more insight into the approximation can be obtained by algebraic manipulation. Let

$R(t) = r + r_1t + \frac{r_2t^2}{2!} + \frac{r_3t^3}{3!} + ...$

Then

$M_T(t)$

$= (1 - r)/(1 - r - r_1t - r_2t^2/2! - r_3t^3/3! + ...) = 1/(1 - r_1/(1 - r)t - r_2/(1 - r)t_2/2! - r_3/(1 - r)t_3/3! - ....)$

so as a first approximation, for small t, we obtain

$M_T(t) \approx 1 + r_1/(1 - r)t$

If we multiply both the numerator and denominator by this result we obtain

$M_T(t) = (1 + r_1/(1 - r)t)/(1 - (r_2/(1 - r) + 2r_1^2/(1 - r)^2)t_2/2! - (r_3/(1 - r) + 3r_1r^2/(1 - r)^2)t^3/3! - ....)$

This enables us to obtain a second approximation as

$M_T(t)$

$\approx (1 + r_1/(1-r)t)(1 + (r_2/(1-r) + 2r_1^2/(1-r)^2)t^2/2!)$

$\approx 1 + r_1/(1-r)t + (r_2/(1-r) + 2r_1^2/(1-r)^2)t^2/2!$

Proceeding in this way we can obtain an extra moment of T at each iteration.

An algorithm for this process is the following:

$m_{0T} = 1$

$m_{1T} = r_1 m_{0T}/(1-r)$

$m_{2T} = (r_2 m_{0T} + 2r_1 m_{1T})/(1-r)$

$m_{3T} = (r_3 m_{0T} + 3r_2 m_{1T} + 3r_1 m_{2T})/(1-r)$

$m_{4T} = (r_4 m_{0T} + 4r_3 m_{1T} + 6r_2 m_{2T} + 4r_1 m_{3T})/(1-r)$

The pattern for extending this algorithm is obvious when we recognize where the binomial coefficients appear.

Having determined the moment generating function for the points played whilst repetition is occurring, we now have to adjust these results to allow for the points played in the two extra games required to complete the set. If player A wins the set the probability of this player winning the two games is $p_A^g q_B^g$ and we obtain

$W_{Y^{gs}(n,n|s_A,w_A)}(t) = W_{Y^{pg}(()|s_A,w_A)}(t) W_{Y^{pg}(()|s_B,l_B)}(t) M_T(t)$ for $n \geq 5$
since the points played in each game are independent.

Likewise if player A loses the set

$W_{Y^{gs}(n,n|s_A,w_A)}(t) = W_{Y^{pg}(()|s_A,l_A)}(t) W_{Y^{pg}(()|s_B,w_B)}(t) M_T(t)$ for $n \geq 5$

Adding these two alternatives together we find that

$M_{Y^{ps}(n,n|s_A)}(t) = (W_{Y^{pg}(()|s_A,w_A)}(t) W_{Y^{pg}(()|s_B,l_B)}(t) + W_{Y^{pg}(()|s_A,l_A)}(t) W_{Y^{pg}(()|s_B,w_B)}(t)) M_T(t)$ for $n \geq 5$.

7.6 Number of points in a match

We study here the model for the best-of-5 final set advantage match. Similar calculations can be obtained for the best-of-5 all tiebreak set match. Because we have to take into account both the winner of the current set and the server at the start of the next set, the recurrence formulas have to allow for four-way branching rather than the two-way branching that we have previously met.

For player A winning the match and currently serving,

$$W_{Y^{pm_5}(0,0:0,0:e,f|c_A,w_A)}(t) = W_{Y^{ps_T}(0,0:0,0|s_A,w_A,n_A)}(t)W_{Y^{pm_5}(0,0:0,0:e+1,f|c_A,w_A)}(t)+$$
$$W_{Y^{ps_T}(0,0:0,0|s_A,l_A,n_A)}(t)W_{Y^{pm_5}(0,0:0,0:e,f+1|c_A,w_A)}(t)+$$
$$W_{Y^{ps_T}(0,0:0,0|s_A,w_A,n_B)}(t)W_{Y^{pm_5}(0,0:0,0:e+1,f|c_B,w_A)}(t)+$$
$$W_{Y^{ps_T}(0,0:0,0|s_A,l_A,n_B)}(t)W_{Y^{pm_5}(0,0:0,0:e,f+1|c_B,w_A)}(t), \text{ if } 0 \leq e+f < 4$$

$$W_{Y^{pm_5}(0,0:0,0:2,2|c_A,w_A)}(t) = W_{Y^{ps}(0,0:0,0|s_A,w_A)}(t)$$

Let $P^{sm_5}(e,f|c_A,w_A)$ represent the probability of player A winning a best-of-5 final set advantage match at set score (e,f) given player A wins the match and is currently serving.

Recurrence Formula:

$$P^{sm_5}(e,f|c_A,w_A) = P^{gs_T}(0,0|s_A,w_A,n_A)P^{sm_5}(e+1,f|c_A,w_A)+P^{gs_T}(0,0|s_A,l_A,n_A)P^{sm_5}(e,f+1|c_A,w_A)+P^{gs_T}(0,0|s_A,w_A,n_B)P^{sm_5}(e+1,f|c_B,w_A)+P^{gs_T}(0,0|s_A,l_A,n_B)P^{sm_5}(e,f+1|c_B,w_A)$$

Boundary Values:

$$P^{sm_5}(e,f|c_A,w_A) = 1, \text{ if } (3,0);\ (3,1)$$
$$P^{sm_5}(e,f|c_A,w_A) = 0, \text{ if } (0,3);\ (1,3)$$
$$P^{sm_5}(2,2|c_A,w_A) = P^{gs}(0,0|s_A,w_A)$$

Table 7.19 represents the probability of player A winning a best-of-5 final set advantage match at various score lines given player A wins the match and is currently serving, where $p_A = 0.62$ and $p_B = 0.60$.

		B score			
		0	1	2	3
	0	0.627	0.422	0.184	0
A score	1	0.783	0.603	0.325	0
	2	0.920	0.815	0.572	
	3	1	1		

Table 7.19: The probability of player A winning a best-of-5 final set advantage match at set score various score lines given player A wins the match and is currently serving, where $p_A = 0.62$ and $p_B = 0.60$

Let $w_n(Y^{pm_5}(0,0:0,0:e,f|c_A,w_A))$ represent the weighted n^{th} moment of the number of points remaining in a best-of-5 final set advantage match at point, game and set score $(0,0:0,0:e,f)$ given player A wins the match and is currently serving.

Recurrence Formulas:

$w_1(Y^{pm_5}(0,0:0,0:e,f|c_A,w_A)) =$

$P^{gs_T}(0,0|s_A,w_A,n_A)w_1(Y^{pm_5}(0,0:0,0:e+1,f|c_A,w_A))+$

$P^{gs_T}(0,0|s_A,l_A,n_A)w_1(Y^{pm_5}(0,0:0,0:e,f+1|c_A,w_A))+$

$P^{gs_T}(0,0|s_A,w_A,n_B)w_1(Y^{pm_5}(0,0:0,0:e+1,f|c_B,w_A))+$

$P^{gs_T}(0,0|s_A,l_A,n_B)w_1(Y^{pm_5}(0,0:0,0:e,f+1|c_B,w_A))+$

$w_1(Y^{ps_T}(0,0:0,0|s_A,w_A,n_A))P^{sm_5}(e+1,f|c_A,w_A)+$

$w_1(Y^{ps_T}(0,0:0,0|s_A,l_A,n_A))P^{sm_5}(e,f+1|c_A,w_A)+$

$w_1(Y^{ps_T}(0,0:0,0|s_A,w_A,n_B))P^{sm_5}(e+1,f|c_B,w_A)+$

$w_1(Y^{ps_T}(0,0:0,0|s_A,l_A,n_B))P^{sm_5}(e,f+1|c_B,w_A)$

$w_2(Y^{pm_5}(0,0:0,0:e,f|c_A,w_A)) =$

$P^{gs_T}(0,0|s_A,w_A,n_A)w_2(Y^{pm_5}(0,0:0,0:e+1,f|c_A,w_A))+$

$P^{gs_T}(0,0|s_A,l_A,n_A)w_2(Y^{pm_5}(0,0:0,0:e,f+1|c_A,w_A))+$

$P^{gs_T}(0,0|s_A,w_A,n_B)w_2(Y^{pm_5}(0,0:0,0:e+1,f|c_B,w_A))+$

$P^{gs_T}(0,0|s_A,l_A,n_B)w_2(Y^{pm_5}(0,0:0,0:e,f+1|c_B,w_A))+$

$2w_1(Y^{ps_T}(0,0:0,0|s_A,w_A,n_A))w_1(Y^{pm_5}(0,0:0,0:e+1,f|c_A,w_A))+$

$2w_1(Y^{ps_T}(0,0:0,0|s_A,l_A,n_A))w_1(Y^{pm_5}(0,0:0,0:e,f+1|c_A,w_A))+$

$2w_1(Y^{ps_T}(0,0:0,0|s_A,w_A,n_B))w_1(Y^{pm_5}(0,0:0,0:e+1,f|c_B,w_A))+$

$2w_1(Y^{ps_T}(0,0:0,0|s_A,l_A,n_B))w_1(Y^{pm_5}(0,0:0,0:e,f+1|c_B,w_A))+$

$w_2(Y^{ps_T}(0,0:0,0|s_A,w_A,n_A))P^{sm_5}(e+1,f|c_A,w_A)+$

$w_2(Y^{ps_T}(0,0:0,0|s_A,l_A,n_A))P^{sm_5}(e,f+1|c_A,w_A)+$

$w_2(Y^{ps_T}(0,0:0,0|s_A,w_A,n_B))P^{sm_5}(e+1,f|c_B,w_A)+$

$w_2(Y^{ps_T}(0,0:0,0|s_A,l_A,n_B))P^{sm_5}(e,f+1|c_B,w_A)$

$w_3(Y^{pm_5}(0,0:0,0:e,f|c_A,w_A)) =$

$P^{gs_T}(0,0|s_A,w_A,n_A)w_3(Y^{pm_5}(0,0:0,0:e+1,f|c_A,w_A))+$

$P^{gs_T}(0,0|s_A,l_A,n_A)w_3(Y^{pm_5}(0,0:0,0:e,f+1|c_A,w_A))+$

$P^{gs_T}(0,0|s_A,w_A,n_B)w_3(Y^{pm_5}(0,0:0,0:e+1,f|c_B,w_A))+$

$P^{gs_T}(0,0|s_A,l_A,n_B)w_3(Y^{pm_5}(0,0:0,0:e,f+1|c_B,w_A))+$

$3w_1(Y^{ps_T}(0,0:0,0|s_A,w_A,n_A))w_2(Y^{pm_5}(0,0:0,0:e+1,f|c_A,w_A))+$

$3w_1(Y^{ps_T}(0,0:0,0|s_A,l_A,n_A))w_2(Y^{pm_5}(0,0:0,0:e,f+1|c_A,w_A))+$

$3w_1(Y^{ps_T}(0,0:0,0|s_A,w_A,n_B))w_2(Y^{pm_5}(0,0:0,0:e+1,f|c_B,w_A))+$

$3w_1(Y^{ps_T}(0,0:0,0|s_A,l_A,n_B))w_2(Y^{pm_5}(0,0:0,0:e,f+1|c_B,w_A))+$

$3w_2(Y^{ps_T}(0,0:0,0|s_A,w_A,n_A))w_3(Y^{pm_5}(0,0:0,0:e+1,f|c_A,w_A))+$

$3w_2(Y^{ps_T}(0,0:0,0|s_A,l_A,n_A))w_3(Y^{pm_5}(0,0:0,0:e,f+1|c_A,w_A))+$

$3w_2(Y^{ps_T}(0,0:0,0|s_A,w_A,n_B))w_3(Y^{pm_5}(0,0:0,0:e+1,f|c_B,w_A))+$

$3w_2(Y^{ps_T}(0,0:0,0|s_A,l_A,n_B))w_3(Y^{pm_5}(0,0:0,0:e,f+1|c_B,w_A))+$

$w_3(Y^{ps_T}(0,0:0,0|s_A,w_A,n_A))P^{sm_5}(e+1,f|c_A,w_A)+$

$w_3(Y^{ps_T}(0,0:0,0|s_A,l_A,n_A))P^{sm_5}(e,f+1|c_A,w_A)+$

$w_3(Y^{ps_T}(0,0:0,0|s_A,w_A,n_B))P^{sm_5}(e+1,f|c_B,w_A)+$

$w_3(Y^{ps_T}(0,0:0,0|s_A,l_A,n_B))P^{sm_5}(e,f+1|c_B,w_A)$

$w_4(Y^{pm_5}(0,0:0,0:e,f|c_A,w_A)) =$

$P^{gs_T}(0,0|s_A,w_A,n_A)w_4(Y^{pm_5}(0,0:0,0:e+1,f|c_A,w_A))+$

$P^{gs_T}(0,0|s_A,l_A,n_A)w_4(Y^{pm_5}(0,0:0,0:e,f+1|c_A,w_A))+$

$P^{gs_T}(0,0|s_A,w_A,n_B)w_4(Y^{pm_5}(0,0:0,0:e+1,f|c_B,w_A))+$

$P^{gs_T}(0,0|s_A,l_A,n_B)w_4(Y^{pm_5}(0,0:0,0:e,f+1|c_B,w_A))+$

$4w_1(Y^{ps_T}(0,0:0,0|s_A,w_A,n_A))w_3(Y^{pm_5}(0,0:0,0:e+1,f|c_A,w_A))+$

$4w_1(Y^{ps_T}(0,0:0,0|s_A,l_A,n_A))w_3(Y^{pm_5}(0,0:0,0:e,f+1|c_A,w_A))+$

$4w_1(Y^{ps_T}(0,0:0,0|s_A,w_A,n_B))w_3(Y^{pm_5}(0,0:0,0:e+1,f|c_B,w_A))+$

$4w_1(Y^{ps_T}(0,0:0,0|s_A,l_A,n_B))w_3(Y^{pm_5}(0,0:0,0:e,f+1|c_B,w_A))+$

$6w_2(Y^{ps_T}(0,0:0,0|s_A,w_A,n_A))w_2(Y^{pm_5}(0,0:0,0:e+1,f|c_A,w_A))+$

$6w_2(Y^{ps_T}(0,0:0,0|s_A,l_A,n_A))w_2(Y^{pm_5}(0,0:0,0:e,f+1|c_A,w_A))+$

$6w_2(Y^{ps_T}(0,0:0,0|s_A,w_A,n_B))w_2(Y^{pm_5}(0,0:0,0:e+1,f|c_B,w_A))+$

$6w_2(Y^{ps_T}(0,0:0,0|s_A,l_A,n_B))w_2(Y^{pm_5}(0,0:0,0:e,f+1|c_B,w_A))+$

$4w_3(Y^{ps_T}(0,0:0,0|s_A,w_A,n_A))w_1(Y^{pm_5}(0,0:0,0:e+1,f|c_A,w_A))+$

$4w_3(Y^{ps_T}(0,0:0,0|s_A,l_A,n_A))w_1(Y^{pm_5}(0,0:0,0:e,f+1|c_A,w_A))+$

$4w_3(Y^{ps_T}(0,0:0,0|s_A,w_A,n_B))w_1(Y^{pm_5}(0,0:0,0:e+1,f|c_B,w_A))+$

$4w_3(Y^{ps_T}(0,0:0,0|s_A,l_A,n_B))w_1(Y^{pm_5}(0,0:0,0:e,f+1|c_B,w_A))+$

$w_4(Y^{ps_T}(0,0:0,0|s_A,w_A,n_A))P^{sm_5}(e+1,f|c_A,w_A)+$

$w_4(Y^{ps_T}(0,0:0,0|s_A,l_A,n_A))P^{sm_5}(e,f+1|c_A,w_A)+$

$w_4(Y^{ps_T}(0,0:0,0|s_A,w_A,n_B))P^{sm_5}(e+1,f|c_B,w_A)+$

$w_4(Y^{ps_T}(0,0:0,0|s_A,l_A,n_B))P^{sm_5}(e,f+1|c_B,w_A)$

Boundary Values:

$w_n(Y^{pm_5}(0,0:0,0:e,f|c_A,w_A)) = 0$, if $(3,0)$; $(3,1)$; $(0,3)$; $(1,3)$

$w_n(Y^{pm_5}(0,0:0,0:2,2|c_A,w_A)) = w_n(Y^{ps}(0,0:0,0|s_A,w_A))$

Tables 7.20, 7.21, 7.22 and 7.23 represent the weighted first, second, third and fourth moments respectively of the number of points remaining in a best-of-5 final set advantage match at various score lines given player A wins the match and is currently serving.

		B score			
		0	1	2	3
	0	165.5	98.7	36.5	0
A score	1	144.4	97.6	43.2	0
	2	90.1	69.8	38.9	
	3	0	0		

Table 7.20: The weighted first moment of the number of points remaining in a best-of-5 final set advantage match at various score lines given player A wins the match and is currently serving, where $p_A = 0.62$ and $p_B = 0.60$

		B score			
		0	1	2	3
	0	46250.8	24051.7	7444.3	0
A score	1	29527.5	17028.9	6039.5	0
	2	11336.6	7246.4	3031.5	
	3	0	0		

Table 7.21: The weighted second moment of the number of points remaining in a best-of-5 final set advantage match at various score lines given player A wins the match and is currently serving, where $p_A = 0.62$ and $p_B = 0.60$

Let $W_{Y^{pm_5}(a,b:c,d:e,f|c_A,w_A)}(t)$ represent the weighted moment generating function of the number of points remaining in a best-of-5 final set advantage match at point, game and set score $(a,b:c,d:e,f)$ given player A wins the match and is currently serving.

$W_{Y^{pm_5}(a,b:c,d:e,f|c_A,w_A)}(t) =$

		B score			
		0	1	2	3
	0	13605392.5	6107473.5	1566640.3	0
A score	1	6611541.6	3197334.7	895555.3	0
	2	1753389.2	899559.1	281444.4	
	3	0	0		

Table 7.22: The weighted third moment of the number of points remaining in a best-of-5 final set advantage match at various score lines given player A wins the match and is currently serving, where $p_A = 0.62$ and $p_B = 0.60$

		B score			
		0	1	2	3
	0	4191414520.7	1613213525.1	341426656.8	0
A score	1	1599056057.3	642798820.0	142067031.6	0
	2	315190870.1	129735828.6	31904969.8	
	3	0	0		

Table 7.23: The weighted fourth moment of the number of points remaining in a best-of-5 final set advantage match at various score lines given player A wins the match and is currently serving, where $p_A = 0.62$ and $p_B = 0.60$

$W_{Y^{psT}(a,b:c,d|s_A,w_A,n_A)}(t)W_{Y^{pm_5}(0,0:0,0:e+1,f|c_A,w_A)}(t)+W_{Y^{psT}(a,b:c,d|s_A,l_A,n_A)}(t)W_{Y^{pm_5}(0,0:0,0:e,f+1|c_A,w_A)}(t)+$

$W_{Y^{psT}(a,b:c,d|s_A,w_A,n_B)}(t)W_{Y^{pm_5}(0,0:0,0:e+1,f|c_B,w_A)}(t)+W_{Y^{psT}(a,b:c,d|s_A,l_A,n_B)}(t)W_{Y^{pm_5}(0,0:0,0:e,f+1|c_B,w_A)}(t),$

if $0 \leq e+f < 4$ and $(c+d)$ mod 2 =0

$W_{Y^{pm_5}(a,b:c,d:e,f|c_A,w_A)}(t) =$

$W_{Y^{psT}(a,b:c,d|s_B,l_B,n_A)}(t)W_{Y^{pm_5}(0,0:0,0:e+1,f|c_A,w_A)}(t)+W_{Y^{psT}(a,b:c,d|s_B,w_B,n_A)}(t)W_{Y^{pm_5}(0,0:0,0:e,f+1|c_A,w_A)}(t)+$

$W_{Y^{psT}(a,b:c,d|s_B,l_B,n_B)}(t)W_{Y^{pm_5}(0,0:0,0:e+1,f|c_B,w_A)}(t)+W_{Y^{psT}(a,b:c,d|s_B,w_B,n_B)}(t)W_{Y^{pm_5}(0,0:0,0:e,f+1|c_B,w_A)}(t),$

if $0 \leq e+f < 4$ and $(c+d)$ mod 2 =1

$W_{Y^{pm_5}(a,b:c,d:e,f|c_A,w_A)}(t) =$

$W_{Y^{ps}(a,b:c,d|s_A,w_A,n_A)}(t)W_{Y^{pm_5}(0,0:0,0:e+1,f|c_A,w_A)}(t)+W_{Y^{ps}(a,b:c,d|s_A,l_A,n_A)}(t)W_{Y^{pm_5}(0,0:0,0:e,f+1|c_A,w_A)}(t)+$

$W_{Y^{ps}(a,b:c,d|s_A,w_A,n_B)}(t)W_{Y^{pm_5}(0,0:0,0:e+1,f|c_B,w_A)}(t)+W_{Y^{ps}(a,b:c,d|s_A,l_A,n_B)}(t)W_{Y^{pm_5}(0,0:0,0:e,f+1|c_B,w_A)}(t),$

if $(e,f)=(2,2)$ and $(c+d)$ mod 2 =0

$W_{Y^{pm_5}(a,b:c,d:e,f|c_A,w_A)}(t) =$

$W_{Y^{ps}(a,b:c,d|s_B,l_B,n_A)}(t)W_{Y^{pm_5}(0,0:0,0:e+1,f|c_A,w_A)}(t)+W_{Y^{ps}(a,b:c,d|s_B,w_B,n_A)}(t)W_{Y^{pm_5}(0,0:0,0:e,f+1|c_A,w_A)}(t)+$

$W_{Y^{ps}(a,b:c,d|s_B,l_B,n_B)}(t)W_{Y^{pm_5}(0,0:0,0:e+1,f|c_B,w_A)}(t)+W_{Y^{ps}(a,b:c,d|s_B,w_B,n_B)}(t)W_{Y^{pm_5}(0,0:0,0:e,f+1|c_B,w_A)}(t),$

if $(e,f)=(2,2)$ and $(c+d)$ mod 2 =1

The formulations for $w_1(Y^{pm_5}(a,b:c,d:e,f|c_A,w_A))$, $w_2(Y^{pm_5}(a,b:c,d:e,f|c_A,w_A))$, $w_3(Y^{pm_5}(a,b:c,d:e,f|c_A,w_A))$ and $w_4(Y^{pm_5}(a,b:c,d:e,f|c_A,w_A))$ can be obtained in the usual manner.

Let $W_{Y^{pm_5}(a,b:c,d:e,f|c_A)}(t)$ represent the weighted moment generating function of the number of points remaining in a best-of-5 final set advantage match at point, game and set score $(a,b:c,d:e,f)$ given player A is currently serving.

It follows that:

$W_{Y^{pm_5}(a,b:c,d:e,f|c_A)}(t) = W_{Y^{pm_5}(a,b:c,d:e,f|c_A,w_A)}(t) + W_{Y^{pm_5}(a,b:c,d:e,f|c_A,w_B)}(t)$

Let $M_{Y^{pm_5}(a,b:c,d:e,f|c_A)}(t)$ represent the moment generating function of the number of points remaining in a best-of-5 final set advantage match at point, game and set score $(a,b:c,d:e,f)$ given player A is currently serving. The set must end with probability 1.

It follows that:

$M_{Y^{pm_5}(a,b:c,d:e,f|c_A)}(t) = W_{Y^{pm_5}(a,b:c,d:e,f|c_A)}(t)$

Tables 7.24, 7.25, 7.26 and 7.27 represent the mean $\mu(Y^{pm_5}(0,0:0,0:e,f|c_A))$, variance $\sigma^2(Y^{pm_5}(0,0:0,0:e,f|c_A))$, coefficient of skewness $\gamma_1(Y^{pm_5}(0,0:0,0:e,f|c_A))$ and coefficient of excess kurtosis $\gamma_2(Y^{pm_5}(0,0:0,0:e,f|c_A))$ respectively of the number of points remaining in a best-of-5 final set advantage match at point, game and set score $(0,0:0,0:e,f)$ given player A is currently serving, and $p_A = 0.62$, $p_B = 0.60$.

		B score		
		0	1	2
A score	0	269.4	213.0	124.9
	1	197.1	164.8	104.8
	2	106.5	95.3	69.4

Table 7.24: The mean number of points remaining in a best-of-5 final set advantage match at $(0,0:0,0:e,f)$ given player A is currently serving, and $p_A = 0.62$, $p_B = 0.60$

		B score		
		0	1	2
A score	0	4030.7	3405.5	3794.0
	1	3957.2	2053.0	1704.9
	2	3436.5	1853.4	672.2

Table 7.25: The variance of the number of points remaining in a best-of-5 final set advantage match at $(0,0:0,0:e,f)$ given player A is currently serving, and $p_A = 0.62$, $p_B = 0.60$

Calculating the distribution of the number of points remaining in a match using the forward recursion results obtained in chapter 5 is somewhat cumbersome within a spreadsheet, due to the large number of possible outcomes that can occur. However, by programming forward recursion formulas it is possible to obtain numerical results by identifying combinations of the number of points played in a game as well as a set finishing in a particular set score. Simulation techniques could also be used to obtain approximate results. An interesting

		B score		
		0	1	2
A score	0	0.25	0.21	0.55
	1	0.34	0.65	0.68
	2	0.96	0.83	2.20

Table 7.26: The coefficient of skewness of the number of points remaining in a best-of-5 final set advantage match at $(0,0:0,0:e,f)$ given player A is currently serving, and $p_A = 0.62$, $p_B = 0.60$

		B score		
		0	1	2
	0	-0.34	-0.22	-0.55
A score	1	-0.44	0.43	0.69
	2	0.15	0.54	7.65

Table 7.27: The coefficient of excess kurtosis of the number of points remaining in a best-of-5 final set advantage match at $(0,0:0,0:e,f)$ given player A is currently serving, and $p_A = 0.62$, $p_B = 0.60$

technique to obtain approximate results using the parameters of distribution of the number of points remaining in a match is as follows.

It has been demonstrated throughout chapter 7 that the total number of points played in a tennis match has a discrete distribution. The moments of this distribution can be calculated using a lattice model with the Markov property and a few other modest assumptions. The Normal distribution has been widely studied, and tables of the probabilities for this distribution are readily available. The basic idea of the Normal Power approximation is to use these tables to estimate the tail probabilities of other distributions. This method uses the first four moments and produces a continuous approximation to the cumulative distribution. The approximation to the frequency distribution can be recovered using differences.

Let X be a random variable with a cumulative distribution $F(x)$, so that $P(X \geq x) = F(x)$

Let μ, σ, γ_1, γ_2 be the mean, standard deviation, skewness and excess kurtosis of X. Let Z be a standardized random variable with mean 0 and standard deviation 1, with

$P(Z \geq z) = P(X \geq x)$

Denote the cumulative Normal distribution by $\phi(.)$. Then the Normal Power approximation can be written as

$F(x) \approx \phi(y)$

with

$z = \frac{x-\mu}{\sigma}$

and

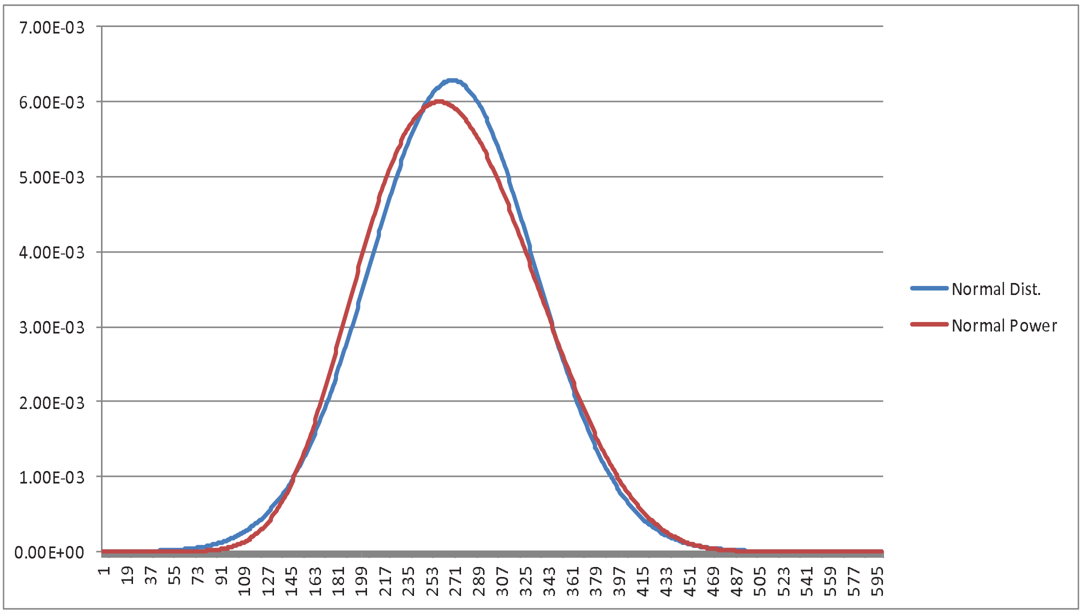
$y = z - \frac{1}{6}\gamma_1(z^2 - 1) - \frac{1}{24}\gamma_2(z^3 - 3z) + \frac{1}{36}\gamma_1^2(4z^3 - 7z)$

It is well known that the distribution of a random variable is not uniquely determined by its moments. In order for this approximation to be successful, the target distribution must be similar in some sense to the Normal distribution. The Normal distribution is continuous, whilst the distribution of points played in a match is discrete. So the best we can hope to achieve is a good approximation as an exact fit cannot be achieved with the Normal Power approximation. Figure 7.1 represents the frequency distribution of the total number of points played in a best-of-5 final set advantage match using the Normal distribution and the Normal Power as approximations for when (a) p_A=0.62 and $p_B = 0.60$ and for (b) $p_A = 0.72$ and $p_B = 0.70$. It shows visually that the tail probabilities are underestimated using the Normal distribution since the higher order moments of skewness and excess kurtosis and not present in obtaining frequency probabilities. These differences in the tail increase as p_A and p_B increase, since the skewness and excess kurtosis increase relatively.

Despite the improvement in the Normal Power approximation over the Normal distribution, there are still several weaknesses that are of some concern:

(a) The Normal Power approximation preserves the unimodal property of the Normal distribution and so fails to replicate the multimodal property of the distribution of points played.

(b) The Normal Power approximation produces a distribution with Normal tails, and fails to fit a distribution with exponential tails. Thus is inappropriate to use in conjunction



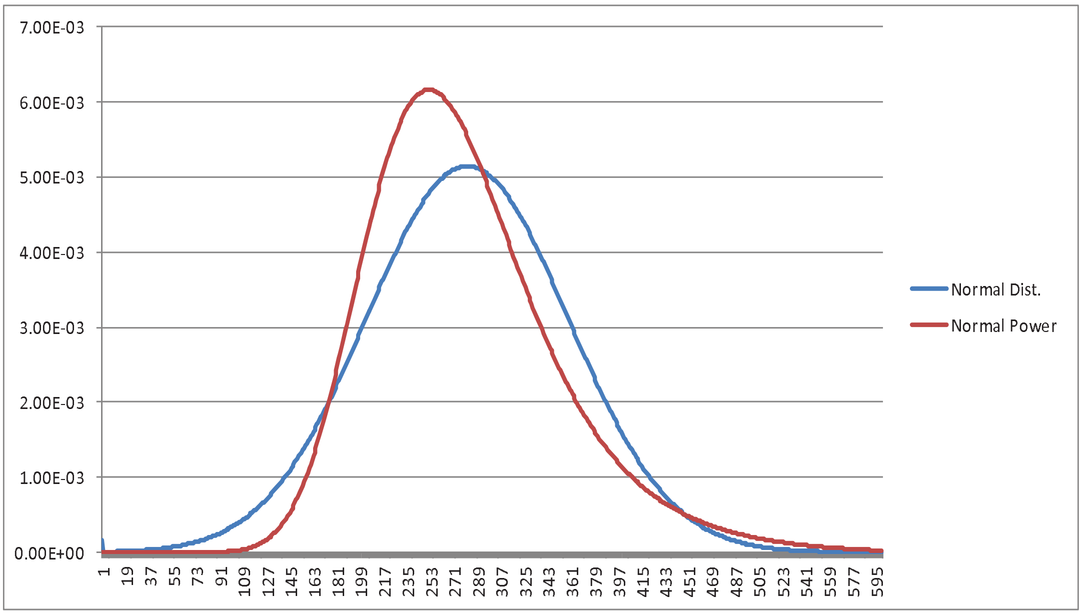


Figure 7.1: Frequency distribution of the total number of points played in a best-of-5 final set advantage match using the Normal distribution and the Normal Power as approximations for when Above: (a) $p_A = 0.62$ and $p_B = 0.60$ and Below: (b) $p_A = 0.72$ and $p_B = 0.70$

with the statistics for a tennis match where an advantage set might be played. This is due to the long tail of the distribution of the points in an advantage set when two good servers are opposed to each other, such as the John Isner vs Nicholas Mahut match played at the 2010 Wimbledon Championships.
(c) The Normal Power approximation is based on an asymptotic expansion. It can become numerically unstable when applied more than 4 standard deviations from the mean.

For information on steps towards improving the Normal Power approximation refer to Brown[1].

7.7 Time duration in a match

As documented in the ITF 2012 Rules of Tennis[2] "*Between points, a maximum of twenty (20) seconds is allowed. When the players change ends at the end of a game, a maximum of ninety (90) seconds are allowed. However, after the first game of each set and during a tie-break game, play shall be continuous and the players shall change ends without a rest. At the end of each set there shall be a set break of a maximum of one hundred and twenty (120) seconds*".

To simplify the analysis we will work with a constant rest time such that:
Average Rest Time = Total Rest Time / Number of Points Played

Note that this includes the time between points within a game as well as the time between change of ends at the completion of a game.

To simplify the analysis even further we will assume that the time to play a point is constant for either player serving such that:

[1]Brown, A. Better approximation to the distribution of points played in a tennis match. In proceedings of the 11th Australasian Conference on the Mathematics and Computers in Sport. 2012, pp99-103
[2]www.itftennis.com

Average Time of Point = Average Time of Point (during play) + Average Rest Time

Note that this is equivalent to:

Average Time of Point = Time Duration of Match / Number of Points Played

More formally, let $X^{tp}(a,b)$ be a constant random variable of the time to play a point at point score (a,b).

It follows that:

$\sigma^2(X^{tp}(a,b)) = 0$

$\gamma_1(X^{tp}(a,b)) = 0$

$\gamma_2(X^{tp}(a,b)) = 0$

Let $Y^{tm_5}(a,b:c,d:e,f|c_A)$ be a random variable of the amount of time remaining in a best-of-5 final set advantage match at point score $(a,b:c,d:e,f)$ given player A is currently serving. It follows that:

$\mu(Y^{tm_5}(a,b:c,d:e,f|c_A)) = \mu(X^{tp}(a,b))\mu(Y^{pm_5}(a,b:c,d:e,f|c_A))$

$\sigma^2(Y^{tm_5}(a,b:c,d:e,f|c_A)) = \mu(X^{tp}(a,b))^2\sigma^2(Y^{pm_5}(a,b:c,d:e,f|c_A))$

$\gamma_1(Y^{tm_5}(a,b:c,d:e,f|c_A)) = \gamma_1(Y^{pm_5}(a,b:c,d:e,f|c_A))$

$\gamma_2(Y^{tm_5}(a,b:c,d:e,f|c_A)) = \gamma_2(Y^{pm_5}(a,b:c,d:e,f|c_A))$

Note that the standard results of $\mu(aX) = a\mu(X)$ and $\sigma^2(aX) = a^2\sigma^2(X)$, for a: constant were used in obtaining the above.

Example: It is commonly known that grass is a fast surface with players winning a relatively high percentage of points on serve (and the time to play a point being relatively low in comparison to clay court surfaces). We will proceed to obtain the parameters of distribution for the amount of time remaining in a best-of-5 set final set advantage match from the outset for a) grass court match as typically occurs at Wimbledon and b)clay court match as typically occurs at the French Open.

a) $p_A = 0.72$, $p_B = 0.70$, $\mu(X^{tp}(a, b) = 38$ secs

b) $p_A = 0.62$, $p_B = 0.60$, $\mu(X^{tp}(a, b) = 47$ secs

These parameters of distribution are given in table 7.28. Note that the mean amount of time for a clay court match is 33.1 minutes longer on average than for a grass court match. This is somewhat expected due to the increased amount of time to play a point. However the coefficients of skewness and excess kurtosis are greater on grass due to the increased serving dominance which leads to a greater chance of a longer advantage deciding set. Interestingly, the standard deviation is greater on clay, and therefore demonstrates why the standard deviation can be insufficient information for measuring risk.

	Mean (min)	S. Deviation	Variance	Variation	Skewness	Kurtosis
a) Grass	177.9	49.1	2411.7	0.28	1.07	2.55
b) Clay	211.0	49.7	2473.3	0.24	0.25	-0.34

Table 7.28: The parameters of distribution for the amount of time remaining in a best-of-5 final set advantage match from the outset for a) grass and b) clay

Chapter 8

Predictions

8.1 Introduction

By assigning two parameters, the constant probabilities of player A and player B winning a point on serve; the probability of winning and duration can be determined using the methods outlined in chapters 1-7. Estimating these two parameters when two elite players meet on a particular surface is now obtained in chapter 8, and an updating rule is derived for the match in progress. The method is demonstrated by focusing on the 'long' men's singles match between John Isner and Nicholas Mahut at the 2010 Wimbledon Championships. The appeal of how predictions in sports multimedia can be used is also presented in chapter 8.

8.2 Data Analysis

OnCourt (www.oncourt.info) is a software package for all tennis fans, containing match results for men's and women's tennis, along with statistical information about players, tournaments or histories of the head-to-head matches between two players. Match statistics can be obtained for the majority of Association of Tennis Professionals (ATP) and

Women's Tennis Association (WTA) matches going back to 2003. The ATP is the governing body of men's professional tennis for the allocation of player rating points in matches to determine overall rankings and seedings for tournaments. The WTA is used similarly in women's professional tennis. Table 8.1 gives the match statistics broadcast from the Isner vs Mahut match at the 2010 Wimbledon Championships where John Isner defeated Nicholas Mahut 70-68 in the advantage fifth set; with a total match time of 11 hours and 5 minutes and total games played of 183. Notice that the Serving Points Won is not given directly in the table. This statistic can be derived from the Receiving Points Won such that Serving Points Won for Isner and Mahut are 1-117/491=76.2% and 1-104/489=78.7%, respectively. Alternatively, the Serving Points Won can be obtained from a combination of the 1st Serve %, Winning % on 1st Serve and Winning % on 2nd Serve such that Serving Points Won for Isner and Mahut are (361/491)(292/361)+(1-361/491)(82/130)=76.2% and (328/489)(284/328)+(1-328/489)(101/161)=78.7%, respectively. Note that the Winning % on 1st Serve is conditional on the 1st serve going in whereas the Winning % on the 2nd Serve is unconditional on the 2nd serve going in. Also worth observing is that Mahut won a higher percentage of points on serve even though Isner won the match, the relatively high proportion of aces served by each player (113/491=23.0% and 103/489=21.1% for Isner and Mahut respectively), and the the relatively low number of break point opportunities (14 and 3 for Isner and Mahut respectively). These calculations could be used as a teaching exercise in interpreting and analyzing data, and in conditional probabilities.

Combining player statistics is a common challenge in sport. Although we would expect a good server to win a higher proportion of serves than average, this proportion would be reduced somewhat if his opponent is a good receiver. We can calculate the percentage of points won on serve when player i meets player j on surface s (f_{ijs}) as follows:

Statistic	Isner	Mahut
1st Serve %	361 of 491=74%	328 of 489=67%
Aces	113	103
Double Faults	10	21
Unforced Errors	52	39
Winning % on 1st Serve	292 of 361=81%	284 of 328=87%
Winning % on 2nd Serve	82 of 130=63%	101 of 161=63%
Winners	246	244
Receiving Points Won	104 of 489=21%	117 of 491=24%
Break Point Conversions	2 of 14=14%	1 of 3=33%
Net Approaches	97 of 144=67%	111 of 155=72%

Table 8.1: Match statistics for the men's 2010 1st round Wimbledon Championships between John Isner and Nicholas Mahut

$$f_{ijs} = f_{is} - g_{js} + g_{avs} \tag{8.2.1}$$

where f_{is} is the percentage of points won on serve for player i on surface s; g_{is} is the percentage of points won on return of serve for player i on surface s; g_{avs} represents the average (across all ATP/WTA players) percentage of points won on return of serve on surface s.

The average percentage of points won on return of serve across all players on each of six different surfaces (grass, hard, indoor hard, clay, carpet and acrylic) was calculated from OnCourt and represented in table 8.2. Note that the serving averages for carpet and indoor hard are approximately the same and are therefore combined as the one surface. Similarly, hard and acrylic are combined as the one surface. Therefore, the surfaces are defined as: $s = 1$ for grass, $s = 2$ for carpet and indoor hard, $s = 3$ for hard and acrylic, and $s = 4$ for clay. For example, suppose male player i with $f_{i1} = 0.7$ and $g_{i1} = 0.4$ meets male player j with $f_{j1} = 0.68$ and $g_{j1} = 0.35$ on a grass court surface. Then the estimated percentage of points won on serve for player i and player j are given by $f_{ij1} = 0.7 - 0.35 + 0.347 = 69.7\%$

and $f_{ji1} = 0.68 - 0.4 + 0.347 = 62.7\%$, respectively.

Surface	Men	Women
Grass	0.347	0.420
Carpet-I.hard	0.358	0.430
Hard-Acrylic	0.375	0.448
Clay	0.400	0.464

Table 8.2: The average probabilities of points won on return of serve for men's and women's tennis

The general form for updating the rating of a player is:

New rating=Old rating+α(actual margin-predicted margin) for some α.

Using serving and receiving player statistics as ratings, we get

$$f_{is}^n = f_{is}^o + \alpha_s(f_{is}^a - f_{ijs}) \tag{8.2.2}$$

$$g_{is}^n = g_{is}^o + \alpha_s(g_{is}^a - g_{ijs}) \tag{8.2.3}$$

where f_{is}^n, f_{is}^o and f_{is}^a represent the new, old and actual percentage of points won on serve for player i on surface s; g_{is}^n, g_{is}^o and g_{is}^a represent the new, old and actual percentage of points won on return of serve for player i on surface s; a_s is the weighting parameter for surface s.

Experimental techniques show that $\alpha_s = 0.049$ is a suitable weighting parameter for all surfaces for both men and women. Further, every player is initialized with surface averages as given in table 8.2.

Equations 8.2.2 and 8.2.3 treat each surface independently. A more advanced approach is to update the serving and receiving statistics for each surface when playing on a particular surface. For example, if a match is played on grass, then how are the other surfaces of clay, carpet/i.hard and hard/acrylic updated on the basis of the player's performances on the grass?

This more complicated approach is given as follows:

$$f_{ist}^{n} = f_{is}^{o} + \alpha_{st}(f_{is}^{a} - f_{ijs}) \tag{8.2.4}$$

$$g_{ist}^{n} = g_{is}^{o} + \alpha_{st}(g_{is}^{a} - g_{ijs}) \tag{8.2.5}$$

where f_{ist}^{n} represents the new expected percentage of points won on serve for player i on surface t when the actual match is played on surface s; g_{ist}^{n} represents the new expected percentage of points won on return of serve for player i on surface t when the actual match is played on surface s; a_{st} is the weighting parameter for surface t when the actual match is played on surface s. Table 8.3 gives the weighting parameters based on experimental techniques.

Surface	t=Grass	t=Carpet/I.hard	t=Hard/Acrylic	t=Clay
s=Grass	0.049	0.02	0.015	0.01
s=Carpet/I.hard	0.02	0.049	0.02	0.015
s=Hard/Acrylic	0.015	0.02	0.049	0.02
s=Clay	0.01	0.015	0.02	0.049

Table 8.3: Surface weighting parameters as applicable to both men and women

8.3 Updating Rule

Although prior estimates of points won on serve may be reliable for the first few games or even the first set, it would be useful to update the prior estimates with what has actually occurred throughout the match. We will use an updating system of the form where for player i the proportion of initial serving statistics (X_i) is combined with actual serving statistics (Y_i) to give updated serving statistics (Z_i) at any point within the match.

$$Z_i = e^{-\frac{n}{c}} X_i + (1 - e^{-\frac{n}{c}}) Y_i \tag{8.3.1}$$

where n represents the total number of points played and c is a constant.

Experimental results reveal that $c = 200$ is a suitable constant for best-of-5 and best-of-3 set matches. Note that the updating process occurs after each point.

8.4 John Isner vs Nicholas Mahut

Serving and receiving rating statistics for Isner and Mahut prior to the match at the 2010 Wimbledon Championships are given in table 8.4 using the forecasting algorithm given by equations 8.2.4 and 8.2.5. Note that the serving statistics on grass for Isner (70.9%) and Mahut (68.8%) are above the average serving statistic on grass (65.3%). Also note that the receiving statistic on grass for Isner (32.7%) is below the average receiving statistic on grass (34.7%). Combining these serving and receiving rating statistics on grass using equation 8.2.1 gives predicted serving statistics for Isner and Mahut at the outset of the match as 69.3% and 70.7% respectively, and as expected both values are above the average serving statistic of 65.3%.

The predicted serving statistics of 69.3% and 70.7% for Isner and Mahut respectively are used as input parameters in the model to give various resultant statistics, as represented in table 8.5. Note that the mean number of games in a set/match with the associated standard deviations are calculated for each player serving first in the set/match.

Player(i)	f_{i1}	g_{i1}	f_{i2}	g_{i2}	f_{i3}	g_{i3}	f_{i4}	g_{i4}
Isner	0.709	0.327	0.722	0.335	0.748	0.349	0.698	0.356
Mahut	0.688	0.363	0.690	0.360	0.630	0.381	0.616	0.382

Table 8.4: Serving and receiving rating statistics for Isner and Mahut prior to the match played at the 2010 Wimbledon Championships

Parameter	Scoring unit	Isner	Mahut
Chance of winning	Point on serve	69.3%	70.7%
	Game on serve	89.2%	90.9%
	Tiebreak game	47.5%	52.5%
	Tiebreak set	45.7%	54.3%
	Advantage set	44.8%	55.2%
	Match	41.6%	58.4%
Mean number of games	Tiebreak set	11.0	10.9
	Advantage set	14.8	14.8
	Match	46.4	46.4
Standard deviation of number of games	Tiebreak set	1.8	1.9
	Advantage set	9.1	9.1
	Match	12.1	12.1

Table 8.5: Predicted parameters for the Isner vs Mahut match played at the 2010 Wimbledon Championships

Two questions were proposed in the preface. Firstly, was this ‘long’ match of 11 hours and 5 minutes predictable? In order to simplify the analysis, the ‘long’ match will be interpreted as the number of games played in the advantage final set as this typically reflects why long matches occur with a high serving dominance from both players, and hence the difficulty for either player to obtain a break of serve in order to finish the match. Incidentally, the 138 games for the final set was also a record for the longest set by games played.

Both players are above the ATP tour averages on grass for the percentage of points won on serve and Isner is below the ATP tour averages on grass for the percentage of points won returning serve. When the player’s statistics are combined together we find that both players are still above the ATP tour averages on grass for percentage of points won on serve. From table 8.5, Isner is expected to win 69.3% of points on serve and Mahut is expected to win 70.7% of points on serve. Isner is expected to win 89.2% of games on serve and Mahut 90.9%. This means that it will be difficult for either player to break serve and if the match does reach an advantage fifth set, there is a possibility it will go on for a long time. Table 8.6 gives the chances of reaching various score lines from the outset of an advantage set, where in column 2 the initial predicted serving statistics for Isner and Mahut are represented. There is a 37.9% chance the set will reach 6 games all, $0.379 \times (0.892 \times 0.909 + 0.108 \times 0.091) = 31.1\%$ chance it will reach 7 games all, $0.379 \times (0.892 \times 0.909 + 0.108 \times 0.091)^2 = 25.5\%$ chance of reaching 8 games all and so on (where 0.892 and 0.909 are the probabilities of Isner and Mahut winning games on serve respectively).

The Isner vs Mahut match stood out amongst the other men's singles matches played at the 2010 Wimbledon Championships, as this match had the highest predicted expected number of games for an advantage set (14.8) along with the highest predicted standard deviation on the number of games played in an advantage set (9.1). In the actual match both players served better than predicted, with Isner winning 76.2% and Mahut 78.7% of points on serve, both higher than the ATP tour average on grass of 65.3%.

The percentage of points won on serve for each player in each set are given in table 8.7. In particular, the percentage of points won on serve in the fifth set for both players are greater than their respective averages for the match. Using Equation 8.3.1 to update prior estimates with actual match statistics, shows that at the start of the advantage fifth set, Isner was predicted to win $0.261 \times 0.693 + (1 - 0.261) \times 0.732 = 72.2\%$ of points on serve and Mahut predicted to win $0.261 \times 0.707 + (1 - 0.261) \times 0.712 = 71.1\%$ of points on serve. This equates to Isner and Mahut expected to win 92.4% and 91.3% of games on serve respectively. The chances of reaching various score lines in an advantage set using these updated serving statistics from the start of the fifth set is given in column 3 in table 8.6. As expected, the chances of a 'long' match are greater when the match has reached the fifth deciding set, since both player's had won a higher percentage of points on serve throughout the match than predicted.

However, the updating formula is an estimate, which includes prior estimates and actual statistics. For example, at the start of the fifth set it is very possible that the serving statistics for both players reflect directly on the serving match statistics that have occurred prior to the fifth set. It is also possible that at the start of the fifth set that the serving statistics for both player's reflect directly on the serving statistics that have occurred on any particular set, an in particular the highest dominance of serving that occurred for both players in the 3rd set. Therefore, the chances of reaching various score lines in an advantage set using the actual serving statistics that occurred in sets 1-4 and set 3 only are given in columns 4 and 5 respectively in table 8.6. Using these results it could be argued that although it is someone difficult to predict a 'long' match from the outset, there is indication based on the serving performance by both players throughout the match, that a 'long' match was predictable prior to the start of the fifth advantage set.

The second question proposed in the preface: What are the chances of this record being broken in the future? The best scenario for a 'long' match to occur, is for both player's winning a 'high' percentage of points on serve in the deciding advantage set. Note that the US Open plays a tiebreak game at 6 games-all in the deciding set, whereas the other three grand slams play an advantage deciding set. Therefore, the Isner vs Mahut record of 138 games cannot occur at the US Open (under the current scoring system). From table 8.2, the average percentage of points won on serve is highest on grass for both men and women. The proportion of matches represented by serving percentages at men's and women's grand slam singles events in 2010 is given in tables 8.8 and 8.9 respectively. As

Score line	$p_A = 0.693$ $p_B = 0.707$	$p_A = 0.722$ $p_B = 0.711$	$p_A = 0.732$ $p_B = 0.712$	$p_A = 0.811$ $p_B = 0.775$	$p_A = 0.772$ $p_B = 0.819$
6-6	37.9%	43.8%	45.7%	74.2%	74.5%
7-7	31.1%	37.3%	39.3%	70.5%	70.8%
8-8	25.5%	31.7%	33.8%	66.9%	67.2%
9-9	21.0%	27.0%	29.1%	63.5%	63.8%
10-10	17.2%	22.9%	25.0%	60.3%	60.6%
20-20	2.4%	4.6%	5.5%	35.9%	36.2%
30-30	0.3%	0.9%	1.2%	21.4%	21.6%
40-40	0.0%	0.2%	0.3%	12.8%	12.9%
50-50	0.0%	0.0%	0.1%	7.6%	7.7%
60-60	0.0%	0.0%	0.0%	4.5%	4.6%
68-68	0.0%	0.0%	0.0%	3.0%	3.1%
69-69	0.0%	0.0%	0.0%	2.8%	2.9%

Table 8.6: Chances of reaching a score line in an advantage set for the Isner vs Mahut match at the 2010 Wimbledon Championships

expected, Wimbledon for both men and women have the highest proportion of matches played with the greatest serving dominance. In particular, both players were winning at least 70% of points on serve in 7.2% of matches played at Wimbledon for men (compared to 3.1% at the US Open, 2.4% at the Australian open and 0.8% at the French Open). Therefore, if the 138 game record was to be broken in the future it would most likely occur at the men's Wimbledon Championships. From table 8.6 it was shown that if the

Set	Isner	Mahut
1	76.9%	65.5%
2	57.1%	76.9%
3	81.1%	77.5%
4	71.8%	66.7%
5	77.2%	81.9%
All	76.2%	78.7%

Table 8.7: Percentage of points won on serve in each set for the Isner vs Mahut match at the 2010 Wimbledon Championships

percentage of points won on serve for both players prior to the fifth set reflected their serving performance in the third set, then there is a 2.8% chance that the match would reach 69-69 all in the deciding fifth set. It is also shown in table 8.6 (column 6) that if the percentage of points won on serve for both players prior to the fifth set reflected their serving performance in the fifth set, then there is a 2.9% chance that the match would reach 69-69 all in the deciding fifth set.

Serving percentages	Wimbledon	US Open	Australian Open	French Open
>70%	7.2%	3.1%	2.4%	0.8%
>65%	29.6%	7.9%	9.6%	7.1%
>60%	66.4%	39.4%	36.8%	30.7%
>55%	91.2%	71.7%	72.0%	55.1%
>50%	98.4%	89.0%	86.4%	81.1%

Table 8.8: Proportion of matches represented by serving percentages at men's grand slam singles events in 2010

Serving percentages	Wimbledon	US Open	Australian Open	French Open
>70%	0.8%	0.0%	0.0%	0.0%
>65%	3.2%	0.8%	0.8%	0.8%
>60%	19.0%	8.7%	7.1%	1.6%
>55%	48.4%	21.4%	31.5%	15.0%
>50%	79.4%	40.5%	63.8%	43.3%

Table 8.9: Proportion of matches represented by serving percentages at women's grand slam singles events in 2010

8.5 Sports Multimedia

Profiting from sports betting is an obvious application of predicting outcomes in sport. Although sports betting was originally restricted to betting before the start of the match, it is now possible and common to be betting throughout a match in progress. However, the appeal of predictions throughout a match may not just be restricted to punters. In-play sports predictions could be used in sports multimedia, and hence could be appealing to the spectator, coach or technology buff without involving actual betting. According to Wikipedia: "Multimedia is media and content that includes a combination of text, audio, still images, animation, video, and interactivity content forms. Multimedia is usually recorded and played, displayed or accessed by information content processing devices, such as computerized and electronic devices". There are many ways predictions through multimedia could be used. Spectators could engage with live predictions through multimedia for entertainment when watching a live match. If a spectator was to place a bet, then live predictions through multimedia could be used as a decision support tool as to when and how much to bet on a particular event, and hence the combination of betting and multimedia becomes a powerful form of entertainment. Another interesting application of sports predictions in multimedia is in teaching mathematical concepts, which students often relate to sporting events and hence may be stimulated in learning mathematics through an activity of personal interest. Another application is in using predictions as a coaching tool. For example, it is common for players and coaches to watch a replayed match to discuss strategies for upcoming matches. The graphical and visual aspects of the predictions could enhance the TV replay. A further application is in using the predictions for TV commentary. For example a time series plot of the probability of winning a tennis match in progress can be useful for TV commentary by supporting his/her discussion on the likely winner of the match. Commentators could also use the graph to evaluate the match after completion to identify turning points and key shifts in momentum.

Strategic Games specializes in delivering online sports content and currently has an interactive tennis calculator freely available in a Java applet[1]. The user interacts by firstly entering the probabilities of each player winning a point on serve followed by the current

[1] www.strategicgames.com.au

server and score line; and the calculator outputs the chances of winning the game, set and match. Note that the backward recursion formulas given in chapters 1 and 2 were used to obtain the results. A more extensive version of the calculator could include the chances of reaching a future score line from a particular score line (chapter 3), and the parameters of distribution and the distributions at different levels within a tennis match (chapters 4,5,6 and 7). Further, a predictive feature could be included that gives serving probabilities for any two given players on the men's and women's main tour on a particular surface (chapter 8).

Figure 8.1 represents a tennis calculator at the outset of the Isner versus Mahut match played at the 2010 Wimbledon Championships. It shows that Mahut has a 90.9% chance of winning a game on serve, 54.3% chance of winning the set and a 58.4% chance of winning the match. It also shows that there is 17.8% chance of reaching deuce on Mahut's serve, 37.9% chance of reaching a tiebreak game and a 36.9% chance of reaching a deciding advantage 5th set. From the graph of the total number of sets played in the match, it shows that Isner has a 9.5%, 15.5% and 16.6% chance of winning the match in 3, 4 and 5 sets respectively, and that Mahut has a 16.0%, 22.0% and 20.4% chance of winning the match in 3, 4 and 5 sets respectively. The parameters of distribution of the total number of sets played in the match are also displayed, and shows that the mean number of sets to be played is 4.1 with a corresponding standard deviation of 0.8. Similar results are also displayed for the total number of games played in a set and the total number of points played in a game.

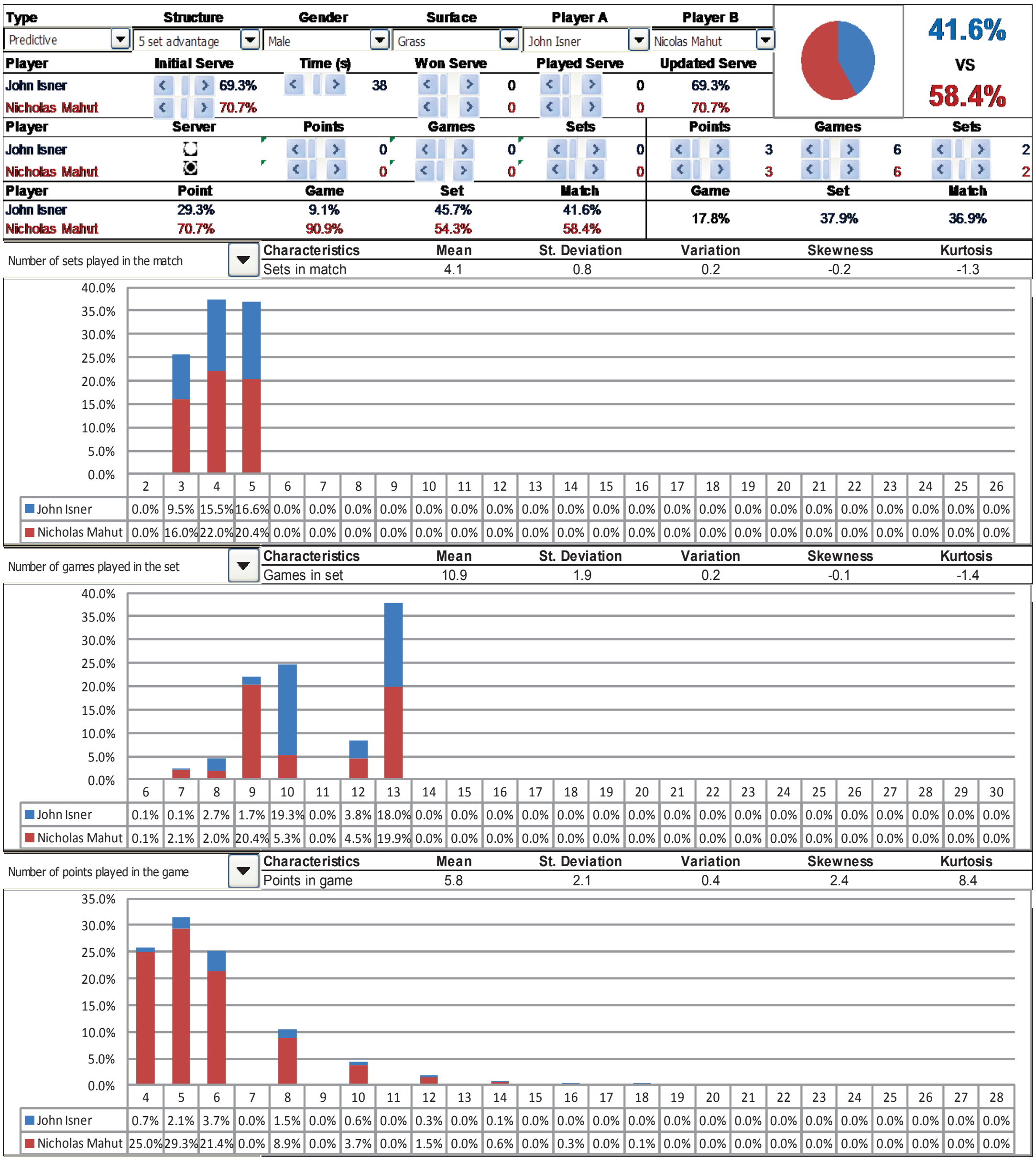


Figure 8.1: Screenshot of a tennis calculator for the Isner versus Mahut match at the 2010 Wimbledon Championships