Winning a game of tennis: using backward recursion

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. Let p represent the probability of the server winning a point and $q=1-p$. Let d represent $P(40,40)$. With simple logic, the probability of the server winning from (40,30) is then $p \times 1 + q \times P(40,40) = p + qd$, 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 $p \times P(40,40) + q \times 0 = pd$.

Likewise, the probability of the server winning from (30,30) is given by $p \times P(40,30) + q \times P(30,40)$. 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 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.

We set up a Markov Chain model of a single game for the server where the state of the game is the current score in points (thus (40,30) is (3,2).

With probability p the state changes from (a,b) to $(a+1,b)$ and with probability $q=1-p$ the state changes from (a,b) to $(a,b+1)$, where a and b represent the current point score for server and receiver respectively.

Let $P(a,b)$ represent the probability that the server wins when the score is (a,b)

The backward recursion formula becomes:
$P(a,b)=pP(a+1,b)+qP(a,b+1)$

The boundary values are:
$P(a,b)=1$, if $a=4$ and $b\leq 2$
$P(a,b)=0$, 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[1], such that $P(3,3)=p^2/(p^2+q^2)$. 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

$P(3,3)=p^2+p^2pq+p^2(2pq)^2+p^2(2pq)^3+\ldots\ldots..$

where the first term is p^2 and the common ratio is $2pq$

The sum is given by $p^2/(1-2pq)$ provided that $-1<2pq<1$. We know that $0<2pq\leq 1/2$, for $0<p<1$ with $q=1-p$. Let $r(p)=2pq=2p(1-p)$. Then $r(0)=r(1)=0$ and $r(1/2)=1/2$ is maximum in $0<p<1$, since $dr/dp=0$ and $d^2r/dp^2<0$ when $r=1/2$.

Therefore the probability of winning from deuce is $p^2/(1-2pq)$ and as obtained in[1], this can be expressed as:
$P(3,3) = p^2/(p^2+q^2)$

Excel spreadsheet code to obtain the conditional probabilities of the server winning a game is as follows:
Enter the text *p* in cell D1.
Enter the text *q* 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 = E1^2/(E1^2+E2^2) 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 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 represents the conditional probabilities of the server winning the game from various score lines for $p=0.60$. 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
	0	0.736	0.576	0.369	0.150	0
	15	0.842	0.714	0.515	0.249	0
A score	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: The conditional probabilities of the server winning the game from various score lines for $p=0.60$

Theorem 1.1. *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.

Theorem 1.2. *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 1.3. *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 1.1 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 1.2 advantage receiver is equivalent to 30-40.

Reference

[1]Barnett T. Winning a game of tennis: using the Binomial theorem. To appear in JMST