PAPER 30 - MATHEMATICS IN TENNIS

S R Clarke

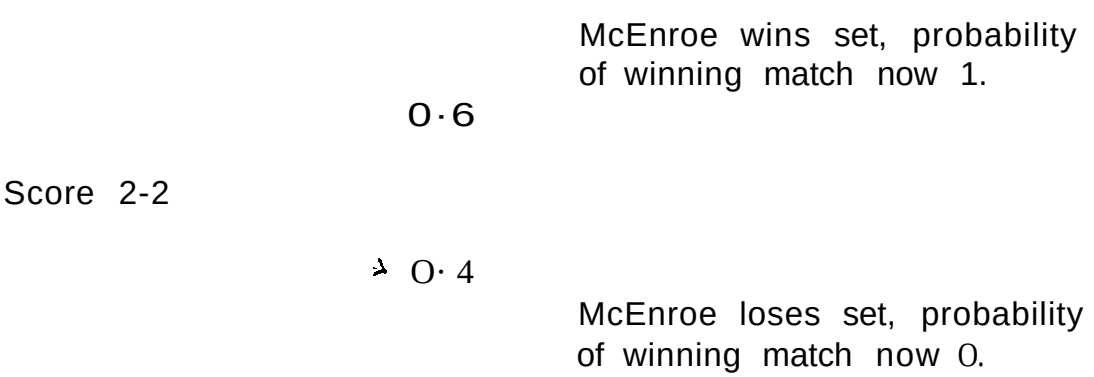
In recent years there has been an lncreasinq interest in the application of mathematics to problems in sport. Students interested in a range of mathematical applications in sports such as baseball, American football, basketball, hockey, cricket, tennis, golf, athletics and rowing should read Ladany and Machol (1977), Optimal Strategies in Sports, North Holland Pub!. Co. As an example, we will look at applying some probability to investigate the effects of different scoring systems in tennis.

Consider a 5-set tennis match between McEnroe and Cash. Suppose McEnroe has a 60% chance of winning any set. The following matrix can be set up, where the body of the matrix shows that probability of McEnroe winning the match when the score is as shown.

		Cash's Score			
McEnroe's Score	3	x	1	1	1
	2	0	*		
	1	0			
	0	0			

Thus the cell rnarked " would contain McEnroe's chance of winning when the score is 2 all. The first row is all 1's since McEnroe wins if he reaches 3 sets, and the first column is all O's since he loses if Cash reaches 3.

A tree diagram now allows us to calculate the number in the cell marked •



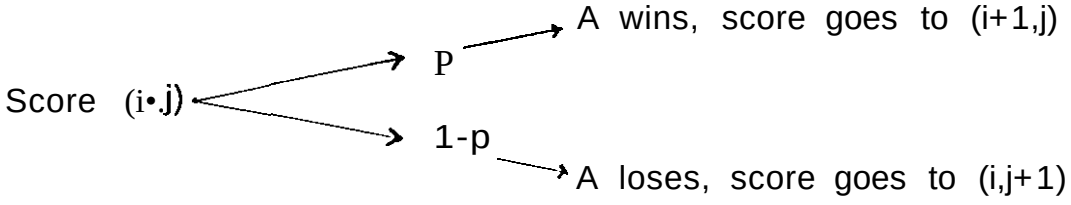
Hence the number in cell at 2-2 is 0·6 x 1 + 0·4 x 0 = 0·6. In a similar way the number in any cell is made up of 60% of the number above it plus 40% of the number to its left. We can progressively fill in the table using a calculator.

The results are shown below

			Cash's Score		
			2	1	
McEnroe's	3	x	1	1	1
Score	2	0	•60	•84	·94
	1	0	·36	·65	·82
	0	0	·22	·48	·68

Thus we find McEnroe's chance of winning from any position, and in particular that he has a 68% chance of winning a 5-set match, but only a 65% chance of winning a 3-set match. This is easily extended to any series of matches, sets, games, or points for any head to head contest in any sport.

In general, if P(i,j) is the probability that player A wins a match up to n points when A's score is i, and B's is j, and p is the probability that A wins any point, a tree diagram gives



Then P(i,j) = p.P(i+1,j) + (1-p).P(i,j+1).

Player A wins when A's score is n and loses when B's score is n, so

$$P(n,j) = 1 \quad 0 \le j < n$$
$$P(i,n) = 0 \quad 0 \le j < n.$$

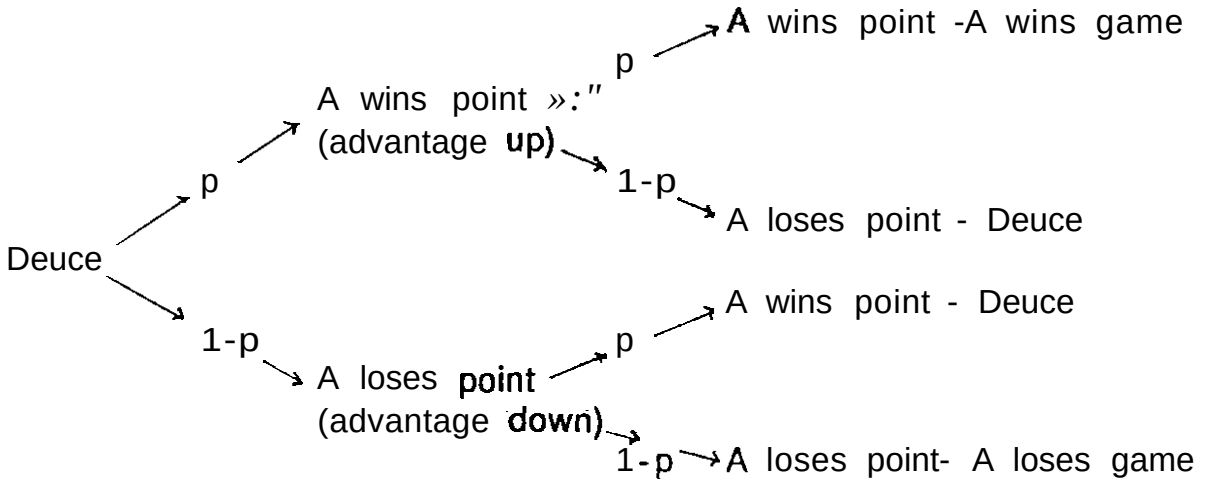
The short computer program below, written in Microsoft Basic, solves these equations progressively, for any match up to 25 points.·

```
 10  DIM P(25,25)
 20  INPUT"number of points in match",N
 30  INPUT"probability of player A winning point", P
 40  FOR 1=0 TO N:P(I,N)=0:P(N-1)=1:NEXT I
 50  FOR I=N-1 TO 0 STEP -1
 60  FOR J=N-1 TO 0 STEP-1
 70  P(I,J)=P*P(I+1,J)+(1-P)*P(I,J+1)
 80  PRINT P(I,J)
 90  NEXTJ
100  PRINT
110  NEXT I
120  END
```

In general we would be interested in P(O,O), the chance of winning at the beginning of the game, but the program gives us the chance of winning from any position. The previous results were obtained by running the program with n = 3 and p = 0-6.

When applying these equations to a set of tennis, where p is now the probability of winning a game, p may alter depending on whether A is serving or not. We would then need to input the two values of p at line 30, and incorporate a line at 65 which chooses between them depending on whether i+i is even or odd.

Situations such as deuce cause a slight problem since at advantage the score may revert to deuce. Thus we need to know the answer for the deuce cell before calculating the advantage cell, but we need the advantage cell before calculating the deuce cell. This can be handled either by iteration - letting your computer program repeatedly calculate each one until its answers are not changing, or by more simply considering 2 points ahead. Again a tree diagram helps.



So if X is the probability that player A wins game at deuce, we have

$$X = p2.1 + (1-p)2.0 + 2p(1-p)X.$$

Solving for X gives

$$X = \frac{p2}{1-2p+2p2} = \frac{p2}{(1-p)2+p^2}.$$

Also the chance of winning from advantage up is p.1 + (1-p)X and from advantage down is pX.

Since deuce usually occurs one point before usual winning score (e.g .at 40-all in a game of tennis, 20-all in table tennis etc) we can incorporate these equations by adding to our program:

```
45 X = P*P/(P*P + (1-P)*(1-P))
46 P(N,N)=X:P(N,N-1)=P+(1-P)*X:P(N-1,N)=P*X.
```

The scoring systems of most sports are nested - e.g. in tennis a certain number of points wins a game, a certain number of games wins a set, a certain number of sets wins a match..

This is most easily incorporated by working through a series of recurrence relations (see Function Vol 2, Part 5 and Vol 3, Part 2.). Starting with the probability of winning a point, we can calculate p(O,O), the probability of winning the game. This then becomes P in the next set to calculate the probability of winning the set, etc.

It is quite surprising how a small advantage in winning a point gives a large advantage in winning the match.

For example, suppose McEnroe wins 70% of points that he serves. Running our program with n = 4, P = 0*7 show that McEnroe will win 90% of his service games. Suppose Cash wins 65% of points that he serves. The program shows, with n = 4 and p = 0·35, that McEnroe will win 17% of Cash's service games. Using a slight approximation, we could avoid this with some extra work, the program with n = 7 and $P = \frac{0{\cdot}70 + 0{\cdot}35}{2} = 0{\cdot}525$ tells us that McEnroe will win approximately 58% of tiebreaker games. By inserting this in as lines 45, 46 and making other adjustments so that when i+j is even, p = 0·17 but when i+j is odd, P = 0·90 we find that McEnroe will win 65% of sets.

Now running the original program (without lines 45, 46) with n = 3 and p = 0·65 we get that McEnroe wins 76% of matches. Thus a very slight advantage in each point (winning 70% of serves against opponent winning 65%) means the better player will win over $\frac{3}{4}$ of best of 5-set matches.

In many cases we may be interested in the length of a match. If $\mu(i,j)$ is the mean number of points in the rest of the match when the score is (i,j), then it can be shown that the basic recurrence relations become

$$\mu(i,j) = 1 + P\mu(i+1,j) + (1-p)\mu(i,j+1)$$
$$\mu(n,j) = \mu(i,n) = 0. \quad 0 \leq i,j < n.$$

Some of the problems which can be investigated using the above techniques are:

- The effect of 3-set, 5-set, 6, 8 or 10-game advantage, no advantage, tie breaker sets on chances of better player, and length of game.
- Probabilities of winning from any position in tennis, table tennis etc.
- The effect of giving players starts, e.g. if you beat a player 21-15 at table tennis, is it fair that you give a 6 start on the next game?
- The most important points in games, Le. at which scores does the probability of winning/losing alter the most?
- The efficiency of scoring systems - which system gives the better player the most chance of winning in as few points as possible?

In Paper 31, we will look at squash and badminton, where you only score . points on your own serve.