Classic Riddler

6 September 2019

Riddle:

Suppose you're playing a match at the U.S. Open, and you're slightly better than the competition: your chances of winning any given point are exactly 55 percent. (Yes, most players are more likely to win the points they serve, but we're simplifying things a bit.) What are your chances of winning a three-set match, as played by the women, or a five-set match, as played by the men? And what are your chances of winning the whole tournament (seven consecutive matches)?

If you're not familiar with the scoring system in tennis, you can read more about it here. In short, the first to 4 points wins a game (as long as they've won 2 more points than their opponent), the first to six games wins a set (as long as they've won two more games), and the first to two sets (for women) or three sets (for men) wins the match. If at any point a set is tied at six games apiece, that set is decided by a tiebreaker, in which the first to 7 points wins (with the same 2-point margin rule applying here).

You know, I thought tennis scoring was pretty straightforward until I had to write it down.

Solution:

I set up this problem as a tree with interconnected probabilities. The tree for a single game is shown on the next page. Each state is a score, and there is a probability for each state that is the probability to win starting from that state. The probabilities are determined by the two outcomes after each state: winning the point with probability p, and losing the point with probability $1 - p$. For this riddle $p = 0.55$. The probabity of winning at state W is 1, and the probability of winning at state L is 0. The other probabilities can be solved by a series of iterated equations.

The calculation starts with the three states at the bottom, $P(40-40)$, $P(\text{Ad in})$, and $P(\text{Ad out})$, which are described as follows:

$$P(40 - 40) = pP(\text{Ad in}) + (1-\text{p})\text{P}(\text{Ad out})$$
$$P(\text{Ad in}) = \text{pP(W)} + (1-\text{p})\text{P}(40-40)$$
$$P(\text{Ad out}) = \text{pP}(40-40) + (1-\text{p})\text{P(L)}$$

Solving gives the initial (approximate) solutions $P(40 - 40) = 0.599$, $P(\text{Ad in}) = 0.8196$, and $P(\text{Ad out}) = 0.3295$. The remaining probabilities are solved bottom-to-top with similar equations.

This kind of iterated equation solving is perfect for spreadsheet calculations, which I have done in the file `tennis.xlsx`. The first part of the spreadsheet solves both the tiebreak and the game based on winning individual points. The probability of winning a game is mathematically equivalent to the probability of winning a tiebreak starting from a score of 3-3. The spreadsheet shows that the probability of winning any game is (approximately) 0.6231, and for a tiebreak is (approximately) 0.6542.

These numbers are then used in the spreadsheet to make a similar calculation for winning a set. The overall probability of winning a set is therefore (approximately) 0.815.

The final portion of the spreadsheet then calculates the probability of winning a womens match and tournament, and a mens match and tournament, which are the final solutions. I list all the combinations of sets which lead to a match win in both cases, and add up the individual probabilities to get the final number. This is then raised to the seventh power to get the total probability of winning each tournament. The probability of winning a womens match and tournament, respectively, are (approximately) $\boxed{\textbf{91.0\% and 51.7\%}}$, and the probability of winning a mens match and tournament, respectively, are (approximately) $\boxed{\textbf{95.3\% and 71.4\%}}$.