First Server Advantage in Tennis

Michelle Okereke

Overview

- Background
- Research Question
- Methodology
- Results
- Conclusion

Background

- **Scoring**
  - Advantage Set: First player to win 6 games by a margin of 2 wins the set
  - Tie-Break Set: If score reaches 6-6, players play a tie-break. First to win 7 points by margin of 2 wins the set
- **First server selection**
  - First Set: Coin toss determines first server
  - Other Sets: Last server in set (or first server in tie-break) serves second the next set

Research Question

- **Question**: Is the “win by 2” rule enough to mitigate the advantage of serving first?
- **Hypothesis**: Given a set where the chances of winning on serve or return are equal, the first server will still be more likely to win
- **Relevance**:
  - Deciding set at Grand Slams
  - Consider further randomization (coin flips before each set and/or tie-breaks)

Methodology

- Analyze probability first server wins under "Advantage Set" rules (no tie-breaks)
- P(Win) = P(Win in 10 Games or Less) + P(Win in Over 10 Games)
  - **10 or Less**: Find equation for probability of each set score outcome, using Excel
  - **Over 10**: Find equation for winning by 2 recursively, given that the score is 5-5
- For First Server
  - Define "S" as P(1^{st} Server Holds a Serve) = P(Win Odd Game)
  - Define "B" as P(1^{st} Servers Breaks a Return) = P(Win Even Game)
- For Second Server
  - Same calculations, but replace "S" with "1-S" and "B" with "1-B"

10 Games or Less

- P(Win in 10 or Less) = P(6-0) + P(6-1) + P(6-2) + P(6-3) + P(6-4)
- Find all the "ways" score could happen and use combinatorics to solve
- Winner must win last game, so that position is fixed
  - Odd number of total games – First server must have won on a hold of serve
  - Even number of total games – First server must have won on a break of serve
- Example: 6-1
  - 4 Service Holds (0 failures to hold), 2 Breaks on Return (1 failure to break)
    - $S^4(1-S)^0B^2(1-B)^1$
  - Player 1 has 4 chances to hold 4 times (or 3 chances to hold 3 times when you consider last game fixed) and 3 chances to break, so this can occur ${}_{4-1}C_{4-1} \times {}_3C_2 = 3$ ways
    - P(Winning with 4 Holds, 2 Breaks) = $3(S^4(1-S)^0B^2(1-B)^1)$

Over 10 Games

- Can only occur after score has hit 5-5, at which point you must win 2 games in a row and avoid losing 2 in a row
  - **P(Win in over 10) = P(5-5)P(Win by 2)**
- Find P(5-5) using same method as before
  - (P(5 Holds, 0 Breaks) + P(4 Holds, 1 Break) … + P(0 Holds, 5 Breaks))
- Find P(Win by 2) recursively
  - W = P(Win by 2)
  - W = SB + S(1-B)W + (1-S)BW
    - **W(First Server) = (SB)/(1 – S – B + 2SB**)

Results

- Write equations for each combination probability as formulas in Excel and sum
- **For S > 0.5 and S+B ≥ 1**
  - **W > 50% (equal whenever S+B = 1)**
  - **P(First Server Wins) > 50% (even when S+B = 1)**
- **If S > B**
  - **P(Win as First Server) > P (Win as Second Server)**
- If both players are equally good and serving and receiving, then there is no advantage to serving first
- Given that 5-5 has been reached, the player who is better overall (higher combined S and B) is most likely to win
- Prior to this, the first server has an advantage as long as they are more likely to hold serve than to be broken. Even if both players are equally good at holding serve (S+B=1), the first server is still more likely to win

Limitations

- **Model**
  - When choosing S and B, do we consider stats in a particular Head-to-Head, or in general
  - Psychological effects of "momentum"
- **Relevance**
  - Only notable events to still use Advantage Sets are the 5^{th} set at Australian Open, French Open, and Wimbledon
- **Real World Data**
  - Tennis statistics do not report serving order
    - Through articles, able to find that 3 of the last 4 five-set Wimbledon men's finals have gone to the player who served first in final set

Conclusion

- Hypothesis partially confirmed
  - If first server is no better at winning on his game or the opponent, then no advantage
  - If first server is even slightly better on serve than than random (which is almost always the case), then serving first is more advantageous than serving second.
- Further inquiry needed to make models more accurate and applicable
  - First server advantage in Tie-Break Sets
  - Effect of introducing additional coin tosses
  - Incorporating other stats to determine mean probability and variance of winning a game (i.e. first serve percentage, break point conversions, etc.)
  - Finding a way to track first server for each set