Serving Strategy in Men's Professional Tennis

Introduction

I have played tennis for a number of years, and although I am not a particularly good player, I take interest in the workings of the game. One notable convention of the game is the concept of having two serves with differing strategies, a first serve and a second serve.

At the beginning of a tennis point, a player has two serves. Usually, a player's first serve is a hard, high velocity serve. In the event of missing the first serve, a player's second serve is typically slower and more measured. The obvious advantage of the fast first serve is that it can be difficult for the opponent to return. The faster the serve, the less time the opponent has to react to the ball; as a result, first serve returns will often be weaker, more defensive, and therefore easier for the server to remain on the offensive. The risk associated with this fast first serve is the higher chance of missing the serve when compared to conventional second serve, which is slower, accurate, and more reliable. The downside to a slower second serve is that as opposed to a fast first serve, it gives the opponent much more time to react, which often results in very powerful returns.

I will from now on refer to the specific technique of hitting a high velocity serve as hitting a "*first serve*" and I will refer to the technique of hitting a slower serve as hitting a "*second serve*" regardless of whether I am speaking of the first or second serve.

On the surface, it might appear logical to first attempt a "*first serve*" and in the event of missing the first serve follow it with a safer "*second serve*" to lower the possibility of missing both serves, which results in a double-fault, in which case the server loses the point. This is what the vast majority of top players do. However, there are two possible unconventional strategies that one might try. The first unconventional strategy is that of serving two "*first serves*" and unconventional strategy number two is to serve two "*second serves*."

There are a few notable professional tennis players who use this first unconventional strategy. Pete Sampras, Boris Becker, and Goran Ivanisevic are three well known players who often utilize a "*first serve*" on both serves. Nick Kyrgios - in particular - and Alexander Zverev - on occasion - are two current players who also utilize this strategy.

However, the second unconventional strategy I mentioned I find fascinating, as it is one which I have never seen explored or mentioned anywhere before. Could it be a new unexplored strategy that has promise, for certain players in some situations? From now onwards I will refer to the typical strategy of serving a "*first serve*" followed by a "*second serve*" as the "conventional" serving strategy. I will refer to the strategy of serving two "*first serves*" as unconventional strategy #1 (or the "bold" strategy) and I will refer to the strategy of serving two "*second serves*" as unconventional strategy #2 (or the "cautious" strategy). I will explore both of these unconventional strategies and how they compare with the conventional strategy.

Aim

My aim is to create a model with which which to determine which of three previously defined serving strategies, conventional strategy, unconventional strategy #1, and unconventional strategy #2, maximizes the probability of winning a point.

Rationale

My rationale for determining this is to gather data from the current top ten male tennis players, and apply the data to the model I create. I choose male tennis players purely out of simplicity, in order to avoid adding the confounding variable of gender.

The data for each player that I required for my calculations, along with the notation I gave each one, are as follows:

p1	probability that the 1st serve is in
w1	conditional probability of winning the point, given that the 1st serve is in
1-p1	probability of a 1st serve fault
1-w1	conditional probability of winning the point, given that the 1st serve is in
p2	probability that the 2nd serve is in
w2	conditional probability of winning the point, given that the 2nd serve is in
1-p2	probability of a 2nd serve fault
1-w2	conditional probability of winning the point, given that the 2nd serve is in

These are all statistics that are readily available online for professional tennis players, or can be easily derived from provided statistics. Recreational players would have to find out these statistics for themselves, by noting the relevant numbers down when they play a game, over the span of several sets.

Additional definitions:

- Service point: a point on which a player serves. The total number of service points for a given player is equal to the number of first serves in + the number of second serves in + the number of double faults. I use this term later when I calculate p2 from available statistics.

Results and concepts from probability theory

These are definitions of basic results and concepts from probability theory that I use in my model.

(1) Mutually exclusive and collectively exhaustive events

Three events E1, E2, and E3 are mutually exclusive and collectively exhaustive if the occurrence of any one event implies the non-occurrence of the other two events, and if one of the three events is certain to occur.

(2) The Law of Total Probability

If A, B, and C are 3 mutually exclusive and collectively exhaustive events; in other words, P(A or B or C) = 100%

And if E represents any other event, then:

P(E/A) is the conditional probability of E, given that A occurs

P(E/B) is the conditional probability of E, given that B occurs

P(E/C) is the conditional probability of E, given that C occurs

Then **P(E) = [P(E/A) P(A)] + [P(E/B) P(B)] + [P(E/C) P(C)].**

This formula for P(E) is the Law of Total Probability.

(3) Independent events

Two events A and B are independent if P(A and B) = P(A) P(B).

Creating the Model

(A) Conventional Strategy

The conventional strategy is to serve a hard first serve, followed by a slower second serve if the first serve is out.

I define the following events:

E1 = 1st serve in

E2 = 2nd serve in (after a 1st serve fault)

E3 = Double fault (both 1st serve and 2nd serve faults)

I define the following probabilities:

P(E1) = p1

P(E2) = p2

P(Win Point/E1) = w1

P(Win Point/E2) = w2

w1 and w2 are conditional probabilities.

The events E1, E2, and E3 are mutually exclusive and collectively exhaustive events, so the Law of Total Probability applies to them.

I use the law of total probability,

P(Win point) = P(win point/1st serve in)P(1st serve in) +

P(win point/2nd serve in)P(2nd serve in) +

P(win point/2nd serve out)P(2nd serve out)

= P(win point/E1)P(E1) + P(win point/E2)P(E2) + 0

= w1p1 + w2 P(E2) + 0

I get a 0 in the 3rd term above, because the probability of winning a point when the 2nd serve is out is 0.

P(E2) = P(1st serve out and 2nd serve in)

= P(1st serve out) P(2nd serve in)

assuming that the event that the 2nd serve is in or out is independent of the event that the 1st serve is in or out.

This simplifies to

P(E2) = (1-p1)p2.

Plugging this back into the formula for P(Win Point), I get

P(Win Point) = w1p1 + (1-p1)w2p2

(B) Unconventional Strategy 1 (“Bold”)

The first unconventional strategy I analyze is the “bold” strategy of using a fast “*first serve*” for both the first and second serve.

P(Win point) = P(win point/E1)P(E1) + P(win point/E2)P(E2) + 0

= w1p1 + P(win point/E2)P(E2)

Again using the independence of the 1st serve and 2nd serve outcomes (in terms of whether or not the serves were in or out), I get

P(win point/E2) = P(1st serve fault and win point with second serve)

= P(1st serve fault)P(win point with second serve)

= (1-p1)w1

Plugging this into the formula for P(Win Point), I get

P(Win Point) = w1p1 + (1-p1)w1p1

(C) Unconventional Strategy 2 (“Cautious”)

The second unconventional strategy I analyze is the “cautious” strategy of using a slower “*second serve*” for both serves.

P(Win point) = P(win point/E1)P(E1) + P(win point/E2)P(E2) + 0

= w2p2 + P(win point/E2)P(E2)

Again using the independence of the 1st serve and 2nd serve outcomes (in terms of whether or not the serves were in or out), I get

P(win point/E2) = P(1st serve fault and win point with second serve)

= P(1st serve fault)P(win point with second serve)

= (1-p2)w2

Plugging this back into the formula for P(Win Point), I get

P(Win Point) = w2p2 + (1-p2)w2p2.

(D) Comparison

For the comparison I have to assume that p1 < p2, which reflects the real-world fact that the chances of a conventional 1st serve landing in are lower than the chances of a conventional 2nd serve landing in. This is because for a "*first serve*", in an effort to increase speed and power, a tennis player sacrifices some control and accuracy.

These are the results I get using the formulas for win probability from before:

Result 1, Bold strategy

For a player, if w1p1 > w2p2,

Then the player should use the "bold" strategy instead of the conventional strategy or the "cautious" strategy to maximize the probability of winning the point, as far as the serve affects it.

Proof: P(Win Point) with the conventional strategy is w1p1 + (1-p1)w2p2 and with the "bold" strategy it is w1p1 + (1-p1)w1p1.

If w1p1 > w2p2, then w1p1 + (1-p1)w1p1 > w1p1 + (1-p1)w2p2, which proves that the "bold" strategy gives a higher "win probability" than the conventional strategy.

now P(Win Point) with the "cautious" strategy is w2p2 + (1-p2)w2p2.

If $w1p1 > w2p2$, then together with the fact that

$p1 < p2$, which is equivalent to $(1-p1) > (1-p2)$,

this shows that

$w1p1 + (1-p1)w2p2 > w2p2 + (1-p2)w2p2$.

This shows if $w1p1 > w2p2$, then the probability of winning a point on serve is higher with the conventional strategy than with the "cautious" strategy. Thus the probability of winning a point on serve is maximized with the "bold" strategy, assuming that $w1p1 > w2p2$.

Result 2, Conventional strategy

For a player, if $w2p2 > w1p1$

Then the player should use the conventional strategy instead of the "bold" strategy to maximize the probability of winning a point.

Proof: If $w1p1 < w2p2$, then $w1p1 + (1-p1)w1p1 < w1p1 + (1-p1)w2p2$, which proves the result

Result 3, Cautious strategy

For a player, if $w2p2\ (1+p1-p2) > w1p1$

Then the player should use the "cautious" strategy instead of the conventional strategy or the "bold" strategy to maximize the probability of winning a point on serve.

<u>Proof:</u> If $w2p2\ (1+p1-p2) > w1p1$, then

$w2p2\ (2-1+p1-p2) > w1p1$, which is the same as

$w2p2\ (2-p2) > w1p1 + w2p2(1-p1)$, which is equivalent to

$w2p2 + (1-p2)w2p2 > w1p1 + (1-p1)w2p2$.

The last inequality states that the “cautious” strategy results in a higher win probability than the conventional strategy.

Now $p1 - p2 < 0$ is always true. So,

$0 < 1+p1-p2 < 1$, which implies that $1/(1+p1-p2) > 1$. This inequality implies that if $w2p2\ (1+p1-p2) > w1p1$ is true, then $w2p2 > w1p1$ is true as well. But Result 2 states that if $w2p2 > w1p1$ is true, then the conventional strategy results in a higher win probability than the “bold” strategy. This implies that if $w2p2\ (1+p1-p2) > w1p1$ is true, then the “cautious” strategy, which results in a higher win probability than the conventional strategy under this condition, also results in a higher win probability than the “bold” strategy. This completes the proof of Result 3.

Result 4

For a player, if $w1p1 = w2p2$

Then the conventional strategy and the “bold” strategy will both give equal win probabilities, and the “cautious” strategy will have the lowest win probability.

Proof: If $w1p1 = w2p2$, then

$w1p1 + (1-p1)w1p1 = w1p1 + (1-p1)w2p2$,

which means that the conventional and “bold” strategies have the same win probability.

Since $p2 > p1$, which is equivalent to $(1-p1) > (1-p2)$, we also have the inequality

$w1p1 + (1-p1)w1p1 > w2p2 + (1-p2)w2p2$

if $w1p1 = w2p2$.

These show that either of the three strategies I looked at, conventional, “bold”, and “cautious”, could be the right strategy to adopt. A player’s personal values for w1, w2, p1, and p2 plugged into the model will determine which serving strategy they should use.

Assumptions and limitations for the model

There are some assumptions and limitations of the model in order for it to be valid:

1. The model assumes that the player using the model does not already use unconventional serve strategies, or uses it so infrequently that it would have a negligible impact on the player’s statistics. It assumes that the player uses conventional serve strategy most of the time. If not, then it would be impossible to calculate p1, p2, w1, and w2 from a player’s published serve statistics.
2. The model can’t account for fatigue during a tennis game. This could easily affect serve performance, and the model ignores the effect of fatigue.
3. The model assumes that the player’s performance on serves are statistically independent, from one serve or attempted serve to the next. In a real tennis game, there could be correlations between different serves. For example, if a player had a string of first serve faults, it may cause them to be timid or hesitant on their second serve, which might diminish their win probability on the second serve.
4. Last, the model does not account for the possible factor of an opponent’s surprise if they do make use of one of the unconventional strategies. The element of surprise could be an added advantage for the player.

Applying the model

In order to determine which of three previously defined serving strategies, conventional strategy, unconventional strategy #1, and unconventional strategy #2, maximizes the probability of winning a point, I will apply the model to the current top ten male tennis players. As previously mentioned, I only used the male top ten for simplicity and to avoid the confounding variable of gender.

The top ten ranked male tennis players in the world right now are, in order from first to tenth, Novak Djokovic, Rafael Nadal, Dominic Thiem, Roger Federer, Daniil Medvedev, Stefanos Tsitsipas, Alexander Zverev, Matteo Berrettini, Gael Monfils, and Roberto Bautista Agut.

There is a problem in the case of Alexander Zverev. As stated in the introduction, Zverev is known to already employ the two "*first serves*" strategy on occasion. As mentioned in the list of assumptions, for the model to be accurate, the player must use traditional serving strategy as opposed to the two "*first serves*" strategy. Because of this, it is reasonable to remove Alexander Zverev from my calculations and not use the model on his statistics. I also excluded Matteo Berrettini, as he is a relatively new player and does not not have some of the necessary statistics available.

I proceed by gathering p1, p2, w1, and w2, as defined previously, for each of the top ten players (with the exception of Alexander Zverev and Matteo Berrettini).

The probability that a second serve is in for a given player (p2) is unfortunately not directly available; however, the double fault percentage is available. The website I found calculates the double fault percentage as (the number of double faults) divided by (the number of first serves in + the number of second serves in + the number of double faults), and the result is multiplied by 100 to get the double fault percentage. I will use d2 to refer to this percentage.

Using d2, I calculated p2 for each of the players I wanted to use the model on:

p2 = (number of attempted second serves) / (number of attempted second serves + number of double faults) = (100 - 100p1 - d2) / (100 - 100p1 - d2 + d2).

Calculating p2 for each player:

Djokovic: d2 = 3.0 and p1 = 0.65

p2 = (100 - 100(0.65) - 3.0) / (100 - 100(0.65) - 3.0 + 3.0) = 0.91

Nadal: d2 = 2.2 and p1 = 0.68

p2 = (100 - 100(0.68) - 2.2) / (100 - 100(0.68) - 2.2 + 2.2) = 0.93

Thiem: d2 = 3.5 and p1 = 0.60

p2 = (100 - 100(0.60) - 3.5) / (100 - 100(0.60) - 3.5 + 3.5) = 0.91

Federer: d2 = 2.4 and p1 = 0.62

p2 = (100 - 100(0.62) - 2.4) / (100 - 100(0.62) - 2.4 + 2.4) = 0.93

Medvedev: d2 = 4.4 and p1 = 0.59

p2 = (100 - 100(0.59) - 4.4) / (100 - 100(0.59) - 4.4 + 4.4) = 0.91

Tsitsipas: d2 = 2.9 and p1 = 0.61

p2 = (100 - 100(0.61) - 2.9) / (100 - 100(0.61) - 2.9 + 2.9) = 0.93

Monfils: d2 = 4.4 and p1 = 0.62

p2 = (100 - 100(0.62) - 4.4) / (100 - 100(0.62) - 4.4 + 4.4) = 0.88

Agut: d2 = 2.5 p1 = 0.66

p2 = (100 - 100(0.66) - 2.5) / (100 - 100(0.66) - 2.5 + 2.5) = 0.93

Player	**p1**	**p2**	**w1**	**w2**
Novak Djokovic	0.65	0.91	0.74	0.56
Rafael Nadal	0.68	0.93	0.72	0.57
Dominic Thiem	0.60	0.91	0.74	0.53
Roger Federer	0.62	0.93	0.77	0.57
Daniil Medvedev	0.59	0.91	0.74	0.52
Stefanos Tsitsipas	0.61	0.93	0.80	0.61
Gael Monfils	0.62	0.88	0.73	0.50
Roberto Bautista Agut	0.66	0.93	0.70	0.54

Calculations

I used the data I gathered, plugged it into the equations I got for the model, and calculated the probability and put it in this following table to compare the win probabilities for each of the three serving strategies.

Player	w1p1 + (1-p1)w2p2 (conventional strategy)	w1p1 + (1-p1)w1p1 (bold strategy)	w2p2 + (1-p2)w2p2 (cautious strategy)
Novak Djokovic	0.6594	0.6494	0.5555
Rafael Nadal	0.6592	0.6463	0.5672
Dominic Thiem	0.6369	0.6216	0.5257
Roger Federer	0.6788	0.6588	0.5672
Daniil Medvedev	0.6306	0.6156	0.5158
Stefanos Tsitsipas	0.7092	0.6783	0.6070
Gael Monfils	0.6198	0.6246	0.4928
Roberto Bautista Agut	0.6327	0.6191	0.537354

Conclusion

With the exception of Gael Monfils, for all of the top professional tennis players whose statistics I analyzed, the popular conventional strategy gives a higher probability of winning a point when compared with the "bold" strategy or the "cautious" strategy. For all the players, with the exception of Monfils, the condition $w2p2 > w1p1$ was true, which means that the "conventional" strategy results in the highest win probability. For Monfils, the condition $w1p1 > w2p2$ was true, and for him the "bold" strategy results in the highest win probability.

The strategy of serving two "*second serves*" is one that I have never seen discussed, and for good reason; the values for the cautious strategy were significantly lower than the other two serving strategies in all cases.

However, for many of the players, the difference in probability between the conventional strategy and the "bold" strategy comes down to the hundredths place. This suggests that the "bold" two "*first serves*" may be an option that players could experiment with. Even if this model does not show a higher probability of using it, the bold strategy may prove advantageous to use occasionally, given that the element of surprise is not mathematically factored into the equation. Players normally stand farther back when they expect a "*first serve*" and stand closer to the net when they expect a "*second serve*" so using the unconventional strategy would surprise the opponent and possibly leave them in an uncomfortable position to return the serve. So, even if the model outputs a lower win probability for the "bold" strategy, it may still be advantageous to the player if used occasionally.

Bibliography

"ATP Stats Leaderboards." *ATP Tour*, ATP Tour, 2020, www.atptour.com/en/stats.

Clarey, Christopher. "A Second Serve That's Just as Big as the First." *The New York Times*, The New York Times, 25 Oct. 2017, www.nytimes.com/2017/10/25/sports/tennis/second-serve-vs-first-serve.html.

Ultimate Tennis Statistics, www.ultimatetennisstatistics.com/.