They're Clutching up! Team Momentum in Round-Based Esports

A Preprint

Tony ElHabr

Georgia Institute of Technology

anthonyelhabr@gmail.com

August 13, 2022

Abstract

My research investigates patterns in round win percentages in professional Search and Destroy (SnD) matches of the popular first-person shooter game Call of Duty (CoD).

First, I find evidence in CoD defying the naive hypothesis that a series represents a sequence of independent events (rounds), with each team having a constant 50% probability of winning a given round.

Second, I examine post-streak round win probability. I find that teams perform significantly worse than expected after streaks of 2, 3, and 4 wins when series end up going to 9, 10, or 11 (maximum) rounds, even after accounting for the "hot-hand" phenomenon.

Third, I compare win percentages in round one versus all other rounds, hypothesizing that there may be some advantage on either side when there is no prior information about how the opponent intends to play a given map in either game. I find only one instance for which there seems to be a significant defensive advantage in round one.

Finally, I evaluate behavior when teams have two rounds left to win the series, observing a peak in COD offensive win percentages in the 4-4 state, and no such oddity in Valorant.

1 Introduction

1.1 Description of Call of Duty Search and Destroy

Call of Duty (CoD), first released in 2003, is one of the most popular first-person shooter (FPS) video game franchises of all-time. Tthe most popular mode in the competitive scene is "Search and Destroy" (SnD), which bears resemblance to "Bomb Defusal" in Counter-Strike and "Plant/Defuse" in Valorant, two other FPS games played in professional leagues. SnD is one-sided game mode in which one team, the offensive side, tries to destroy one of two designated bomb sites on the map.

In professional CoD, a team must win six rounds of SnD to win the match.[1] A round can end in one of four ways:

1. One team eliminates all members of the other team prior to a bomb plant. (Eliminating team wins.)
2. The offensive team eliminates all members of the defensive team after a bomb plant.[2] (Offense wins.)
3. The defensive team defuses the bomb after a bomb plant.[3] (Defense wins.)
4. The offensive team does not make a plant by the time the round timer ends. (Defense wins.)

[1] A maximum of 11 even rounds can be played. There is no "sudden death" or "win by two" rule like there are for SnD equivalent in professional Counter-Strike and Valorant matches.

[2] Merchanics regarding the bomb:

[3] Often the defensive team will try to eliminate all team members prior to making the defuse, but in some cases, they may try to "ninja" defuse.

- The bomb can be picked up by any member of the offensive team.
- The bomb carrier is not obstructed at all by carrying the bomb (i.e. movement is the same, weapon usage is the same).
- The defense does not get any visual indication for who is carrying the bomb.
- A bomb plant takes five seconds. The timer resets if the player stops planting site prior to completing it.
- A bomb defuse takes seven seconds. The timer resets if the player "drops" the bomb.
- The bomb takes 45 seconds to defuse after being planted.

Teams take turns playing offense and defense every round.

We adopt the terminology "series" to refer to what CoD SnD players typically call a "match", so as to emulate the terminology of playoff series in professional leagues like the National Basketball Association, National Hockey League, and Major League Baseball. A "game" or a "match" in such leagues is analogous to a "round" of CoD SnD.

1.2 Data

CoD has roughly gone through three eras of professional gaming: (1) Major League Gaming (MLG) tournaments prior to 2016; (2) the CoD World League (CWL), initiated in 2016; and (3) the 12-franchise CoD League (CDL), running since 2020 and completing three year-long "seasons" completed as of August 2022.[4] The data set consists of all SnD matches played in tournaments and qualifiers during the CDL era, totaling 15,584 rounds across 1,704 series. Data was collected in spreadsheets by community member "IOUTurtle".[5]

The empirical offensive round win percentage across all rounds is 47.8%.[6] To contextualize our subsequent discussion of series length and momentum, table 1 shows round win percentages by series "state" (i.e. the number of round wins by each team prior to an upcoming round). Offensive round win rate is not quite constant, although never veers more than 10% from this global average.

	Offensive team's round wins					
Defensive team's round wins	0	1	2	3	4	5
0	47.8% (852)	46.6% (408)	43.1% (216)	43.5% (115)	43.3% (67)	40.5% (37)
1	48.6% (444)	49.3% (418)	51.5% (309)	43.4% (205)	43.3% (120)	39.4% (99)
2	52.8% (218)	48.9% (305)	48.9% (315)	46.6% (262)	48.7% (189)	42.1% (133)
3	54.5% (123)	46.0% (200)	49.6% (250)	45.6% (248)	44.4% (214)	44.8% (174)
4	56.9% (65)	54.5% (145)	47.2% (193)	44.7% (228)	55.2% (221)	50.5% (208)
5	47.4% (38)	49.4% (83)	47.1% (136)	50.9% (175)	45.2% (177)	46.0% (202)

2 Literature review

There have been only a handful of studies of the distribution of games played in a series, most of which assume a constant probability p of a given team winning a game in the series, regardless of the series state. (Mosteller 1952) observed that the American League had dominated the National League in Major League Baseball's (MLB) World Series matchups, implying that matchups should not modeled with $p=0.5$. Mosteller proposed three approaches for identifying the optimal constant probability value of the stronger team in the World Series, finding $p \quad 0.65.$ in each case: (1) a "method of moments" approach in which one solves for $p.$ from the empirical average number of games won by the loser of the series; (2) maximizing the likelihood that the sample would have been drawn from a population in which the probability of a team winning a game is constant across the series, and (3) minimizing the chi-square goodness of fit estimate for p.

(Chance 2020) re-examines the constant probability notion in Major League Baseball's World Series (1923–2018), the National Basketball Association's Finals (1951–2018), and the National Hockey League's Stanley Cup (1939–2018). Chance applies finds strong evidence against the null hypothesis of $p = 0.5$ in the MLB and NHL championship series

[4] CoD is fairly unique compared to other esports in that it runs on an annual lifecycle (released coming in the late fall), where a new game is published every year under the same title. Each new game bears resemblance to past ones, often introducing relatively small variations ("improvements") to graphics, game modes, and other facets of gameplay. During the CDL era, the games released have been Modern Warfare (2020), Cold War (2021) and Vanguard (2022).

[5] Data: https://linktr.ee/CDLArchive. Author: https://twitter.com/IOUTurtle

[6] Offensive round win percentage has been nearly constant across the three games during the CDL era: MW (2020) 47.2%, Cold War (2021) 47.9%, Vanguard (2022) 48.1%

when applying Mosteller's first and second methods.[7] Chance's work is closely related to ours and, in fact, provides a guide for the first part of our investigation.

Momentum, one of most discussed topics in sports analytics, goes hand-in-hand with a discussion of the nature of series outcomes.[8] Two opposing fallacies are observed in the context of momentum: the "gambler's fallacy" (negative recency) and "hot hand fallacy" (positive recency). Per (Ayton and Fischer 2004), negative recency is "the belief that, for random events, runs of a particular outcome … will be balanced by a tendency for the opposite outcome", while positive recency is the expectation of observing future results that match recent results.

Studying both player streaks and team streaks in basketball, in both observational and controlled settings. (Gilovich, Vallone, and Tversky 1985) do not find evidence for the hot hand phenomenon. Recently , (Miller and Sanjurjo 2018) refute (Gilovich, Vallone, and Tversky 1985)'s conclusions, finding mathematical evidence that seems to support negative recency. Specifically, they find that a "bias exists in a common measure of the conditional dependence of present outcomes on streaks of past outcomes in sequential data" (streak selection bias) that imply that, under i.i.d. conditions, "the proportion of successes among the outcomes that immediately follow a streak of consecutive successes is expected to be strictly less than the underlying (conditional) probability of success". We agree with Miller's findings, accounting for the streak selection bias in our study of momentum.

Despite the plethora of existing research on games played in a series and momentum in sports, these topics have yet to be investigated heavily in esports. Work has been done to examine in-round win probability in other FPS titles such as Counter-Strike ((Xenopoulos, Freeman, and Silva 2022)) and Valorant ((DeRover 2021)), both of which are round-based like CoD SnD. However, research on round-level trends is sparse, perhaps because games like Counter-Strike and Valorant both have economic aspects that can create clear advantages on side in a given round, given how prior rounds played out.[9]

3 Methodology, results, and discussion

First, we investigate the constant probability assumption and the distribution of rounds played in a series. Afterwards, we investigate momentum, building on our learnings from the constant probability assumption analysis.

3.1 Distribution of rounds played

The general formula for the probability $P_E(i)$ that a best-of-s series lasts i rounds given constant probability p of one team[10] winning each round is

$$q = 1 \;\; p, s_1 = \frac{s \;\; 1}{2}, s_2 = \frac{s+1}{2}, s_2 = 1 \;\; s_2$$

$$P_E(i) = \frac{(i \;\; 1)!}{s_1!(i \;\; s_2)!}(p^{s_2}q^{s_2} + p^{s_2}q^{s_2}). \tag{1}$$

Table 1 shows the expected proportion for $s = 11$ and the naive assumption that $p = 0.5$, along with the observed proportion $P_O(i)$ for CoD SnD.

Calculating the chi-square goodness of fit statistic

$${}^2 = {}^{11}_{i=6} \frac{(P_O(i) \;\; P_E(i))^2}{P_E(i)}, i \;\; R = [6, 7, 8, 9, 10, 11]$$

as 16.0 (p-value of 0.0068), we can comfortably reject the constant probability null hypothesis, event at a confidence level of $= 0.01$.

Table 4 shows the alternate values for the constant round win probability p for winning a given round in a CoD SnD that we find when applying the three methods suggested by (Mosteller 1952). Each is approximately or equal to 0.575, and each results in a chi^2 value for which we cannot reject constant probability null hypothesis.

[7]Chance goes on to outline a conditional probability framework (likelihood of winning a game given the series state) which can exactly explain the distribution of the number of games played.

[8]We often use use "streaks" and momentum interchangeably, but as (Steeger, Dulin, and Gonzalez 2021) note, momentum implies dependence between events, whereas streaking does not.

[9]Additionally, both Counter-Strike and Valorant have overtime rules and blocked offensive/defensive roles (i.e. playing either offense or defense for many consecutive rounds).

[10]If $p > 0.5$, then we might say that this team is the better team (presumably known in hindsight)

: The probabilities that a best-of-11 series lasts i rounds ($P(i)$, where i $R = [6, 7, 8, 9, 10, 11]$) under the assumption that each team has a 50% probability ($p = 0.5$) of winning each game {#tbl-prob-series-lasting-i-rounds} along with the observed frequencies of series lasting i rounds, expressed as a percentage $P_O(i)$ and as a count (N_O)

(a)

		Observed	
Series lasts i rounds	$P_E(i)$	$P_O(i)$	$N_O(i)$
6	3.1%	4.7%	40
7	9.4%	11.9%	101
8	16.4%	16.5%	141
9	21.9%	21.7%	185
10	24.6%	21.5%	183
11	24.6%	23.7%	202

Alternate estimates of the constant probability p for winning a given round in a CoD SnD, applying the three methods suggested by (Mosteller 1952), in addition to the naive $p = 0.5$.

Method	p	^{2} (p-value)
0. Naive	0.5000	16.0 (<=0.01)
1. Method of moments	0.5725	3.6 (0.6)
2. Maximum likelihood	0.5750	3.5 (0.62)
3. Minimum ^{2}	0.5775	3.5 (0.62)

Table 5 shows the new $P_E(i)$ when re-applying Equation 1 for each new p. We observe that $P_E(i)$ is notably smaller for $p_{1,2,3}$ when i $[10, 11]$ and higher for i $[6, 7]$, more closely matching $P_O(i)$.

Caption

Series lasts i rounds	0. p= 0.5	1. p = 0.5725	2. p = 0.575	3. p = 0.5775	$P_O(i)$
6	3.1%	4.1%	4.2%	4.3%	4.7%
7	9.4%	11.1%	11.2%	11.4%	11.9%
8	16.4%	17.7%	17.8%	17.9%	16.5%
9	21.9%	21.8%	21.8%	21.8%	21.7%
10	24.6%	23.1%	23.0%	22.9%	21.5%
11	24.6%	22.1%	22.0%	21.8%	23.7%

References

Ayton, Peter, and Ilan Fischer. 2004. "The Hot Hand Fallacy and the Gambler's Fallacy: Two Faces of Subjective Randomness?" *Memory & Cognition* 32 (8): 13691378.

Chance, Don. 2020. "Conditional Probability and the Length of a Championship Series in Baseball, Basketball, and Hockey." *Journal of Sports Analytics* 6 (2): 111–27. https://doi.org/10.3233/JSA-200422.

DeRover, DeMars. 2021. "Round Win Probabilities Based on Who's Alive + Time." https://www.reddit.com/r/VALORANT/comments/n3lpoo/round_win_probabilities_based_on_whos_alive_time/.

Gilovich, Thomas, Robert Vallone, and Amos Tversky. 1985. "The Hot Hand in Basketball: On the Misperception of Random Sequences." *Cognitive Psychology* 17 (3): 295314.

Miller, Joshua B., and Adam Sanjurjo. 2018. "Surprised by the Hot Hand Fallacy? A Truth in the Law of Small Numbers." *Econometrica* 86 (6): 20192047.

Mosteller, Frederick. 1952. "The World Series Competition." *Journal of the American Statistical Association* 47 (259): 355380.

Steeger, Gregory M., Johnathon L. Dulin, and Gerardo O. Gonzalez. 2021. "Winning and Losing Streaks in the National Hockey League: Are Teams Experiencing Momentum or Are Games a Sequence of Random Events?" *Journal of Quantitative Analysis in Sports* 17 (3): 155170.

Xenopoulos, Peter, William Robert Freeman, and Claudio Silva. 2022. "Analyzing the Differences Between Professional and Amateur Esports Through Win Probability." In, 34183427.