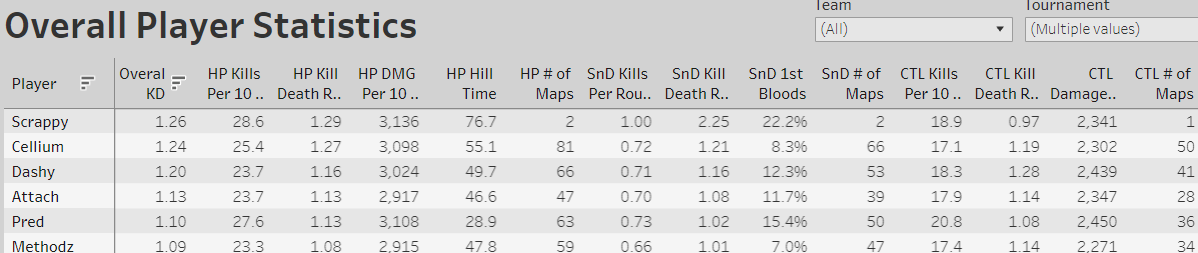
Kyle Antczak

Probability and Statistics using CDL 2022 data

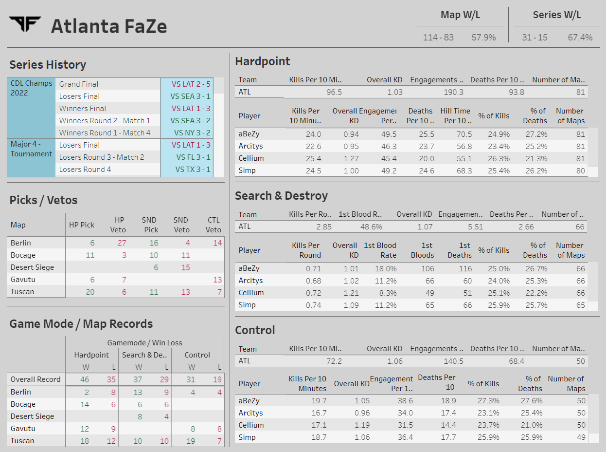
The data I choose to use was for the Call of duty league season of 2022. This season was for the game call of duty vanguard, which released in November of 2021. There is a total of 12 teams based around the country, with four players playing on each team. Throughout the season there was 4 regular season tournaments and qualifier rounds, which teams won cdl points based off tournament placement and winning qualifier matches. There then was a championship tournament with the top 8 teams. Each match is a best of 5 games, where 3 different game modes are played in the order of Hardpoint, Search and Destroy, Control, Hardpoint, Search and Destroy on 5 different maps of Berlin, Bocage, Desert Siege, Gavutu, and Tuscan. Maps are decided by each team banning one from the pool, (for Search Gavutu and for Hardpoint desert siege cannot be played) then each picking one of the 2. This is done twice, once for SnD and Hardpoint. Control is only out of a 3-map pool (Desert siege and bocage cannot be played), so each team bans one and the last is played.

For the specific data sets I will be looking at a few different things. First of which will be players overall stats (1). This table shows off every player and variables for their overall kill death ratio (KD), kills per 10 minutes for each game mode, KD for each game mode, number of games played for each game mode, and other game mode specific variables. I will also be using team specific data (2). The team statistics gives data like overall game win loss ratio, series win loss ratio, and the same variables as used for the players. It also gives per map win loss values and per map veto and picks.

In this report I will be using players variables to compare the players to one another. Player variables will also be used to answer questions about the probability of how they will perform based on certain conditions. I will be also breaking down the map choosing process to find out probabilities of playing on specific maps. I will also look at team specific games to see the break down of when they win and the probability that they will do it based off variable like map and mode.

Here is the table for 6 of the players: 

Here is the data for the Atlanta Faze team:



(1.2)

First thing I am showing is a histogram I made in excel for all the players overall kill death ratio. This shows all the players KD out of percent of grand total grouped by 0.05. With this I can see things like about 26% of players have a KD of 0.94-0.99 and about 8% have the highest KDs ranging from 1.09-1.24.

(1.3)

With excel I was able to easily calculate the mean, variance, and standard deviation for the KD’s of all the players. The average KD was 0.978, the variance was 0.00774 and the standard deviation was 0.088686. With this information I can better understand the how all the players KDs compare. The mean shows how good or bad one player is compared to the average of all players. The standard deviation also shows how good or bad a player is by how far away from the mean they are, showing that players within 1 standard deviation of the mean fall between the KDs of 0.89 - 1.07. The Histogram very closely shows a bell shape curve, so using the empirical curve roughly 68% of players fall within 1 standard deviation. Also, the fact the variance value is so small shows how small the amount of variation within the set.

(2.3)

When it comes to a team the games played can be broken down into many ways. The overall set that contains them is total games played. The next most common way to break down games is by if they won or lost. These two sets are mutually exclusive. Games can be broken down even further into game modes and maps or combinations of these variables. Some examples include games played on hardpoint, games lost on control, hardpoint games played on Tuscan, and games lost on Bocage. Some of these sets have intersections like Search and destroy wins and games played on Berlin being Search games win on Berlin. Using this example for the Atlanta Faze team, they won 37 search games and played 40 games on Berlin. The intersection is 13 games of search won on Berlin.

(2.4)

Sticking with game break downs let’s say we wanted to find the probability of Atlanta Faze playing on a specific map. The sample space consists of 5 events being the different maps: = Berlin being played 40 times, =Bocage being played 32 times, =Desert Siege being played 12 times, =Gavutu being played 37 times, and =Tuscan being played 76 times. So the probability of each event would be the number of times played out of the total 197 being for each: P()= 0.20, P()= 0.16, P()= 0.06, P()= 0.19, P()=0.39.

(2.5)

Using the sample point method, we can see probabilities for occurrences of maps. Let’s say we wanted to see probabilities for maps of the first 2 games which are played in the order of Hard Point then Search and destroy. It is also important to remember certain maps cannot be played on certain gamemodes so the event: (Bocage,Gavutu) won’t appear because Gavutu can’t be played on Search. The Events are : (Berlin,Berlin) : (Berlin,Bocage) : (Berlin,Desert Siege) : (Berlin,Tuscan) : (Bocage,Bocage) : (Bocage,Berlin) : (Bocage,Desert Siege) : (Bocage,Tuscan) : (Gavutu,Berlin) : (Gavutu,Bocage) : (Gavutu,Desert Siege) : (Gavutu,Tuscan) : (Tuscan,Berlin) : (Tuscan,Bocage) : (Tuscan,Desert Siege) : (Tuscan,Tuscan)

By looking at these events we can see different things like P(playing Berlin at least once) = contains , *, , , , , =7/16*

P(not playing Tuscan) = , *, , , , , = 9/16*

P(playing Bocage and Tuscan)= , *= 2/16*

(2.6)

Placing in a tournament the first four teams are awarded different amounts of points being 65 for first, 50 for second, 40 for third, and 30 for fourth. I can see the total number of possible outcomes by doing which is 11880 ways.

(2.7)

I can use conditional probability to find out the probability if a team will win or loss based on what map or mode is being played. For example lets look at Atlanta Faze. Something I might want to look for is the probability that they win given they are playing Hardpoint or whats the probability they are playing on Berlin given that they win.

P(win|hardpoint) = = =

P(Berlin|win) = = =

(2.8)

Lets say I wanted to find the probability that Atlanta faze will win on hardpoint or will win on Tuscan.

P(winning on hardpoint) ꓴ P(winning on Tuscan) = P(winning on hardpoint) + P(winning on Tuscan) – P( winning on Tuscan hardpoint) = +- =

(2.10)

Before a game, each team gets to pick and veto a map. Let’s stick to Atlanta Faze. For the map Berlin they picked it 22/42 games = 0.524. When they picked it, they won 12/22 = 0.5454 and when they didn’t, they won 8/20 = 0.4. What is the probability they won?

= P(won|picked) P(picked) + P(won|didn’t pick) P(didn’t pick)

P(didn’t pick) = 1 – 0.524 = 0.476

=.5454 \* .524 + .4\* .476 = .2858 + .1904 = .476

(3.2)

Let’s look back to 2.5 and use that sample space to see the probability distribution of playing Tuscan:

P(0) = , *, , , , , = 9/16*

P(1) = , ,,,, = 6/16

P(2) = = 1/16

(3.3)

The expected number of times to play Tuscan is 0 (9/16) + 1 (6/16) + 2 (1/16) = 14/32 = 0.4375

The expected variance is

=(9/16) + (6/16) + (1/16) =

=.1077 + .1187 + .1526 = .379

(3.4)

The game mode Search and destroy is based on rounds first to 6. In a round the first player to get a kill gets first blood. One player stat is 1st blood percentage. For one play attach his is 11.7%. Let’s say he plays a Search game that lasts 11 rounds. What is the probability that he will get 2 first bloods?

Using binomial distribution, we can find this by = 55 \* .0137 \* .28814 = 0.2171 This means in an 11 round game attach has a 21.71% chance of getting 2 first bloods. His expected is 1.404 and expected variance is 1.24

(3.5)

Using geometric distribution, the probability that attach will get a first blood before or on the 5th round. Which would be = 0.537. The expected number of rounds it will take is 8.547 round. This was calculated by .

(3.7)

Lets say I had the videos for every game that Atlanta Faze played on Desert Siege(12). Lets say I wanted to review 6 of the videos so I randomly pick 6 to watch. What would be the probability that I choose 4 games that they win? This can be done using hypergeometric distribution. The equation would be:

== 0.4545

(3.8)

On hardpoint one player Cellium, averages 25.4 kills per 10mins. What if I wanted to find the probability that he would get 60 kills in 20 minutes? Using Poisson distribution, it would be =0.023. If I wanted to find out the probability of 27 kills in 10 min it would be = 0.100

(3.11)

Using the mean and Standard deviation from 1.3, what percentage of values fall between 0.886 and 1.07. 0.092 is the within number by 0.092 = 1.07 - 0.978 = 0.978 - 0.886. so, k = 0.092/0.088686 = .8119 And plugging that into tchebysheff it gets 1- = .07

(4.2)

Using the histogram from 1.2 a distribution function can be made.

F(y) =

The corresponding density function would be 0 at all points because the =0 Using this F(1.00) = .24

(4.3)

Another distribution function can be made from taking the players stats of time on hard point hill per game. In hardpoint there is a hill that moves around the map and while standing in that your team gets points. Every second on the hill is equivalent to 1 point. This make the distribution of time to points 1 to 1 so the distribution of amount of time to number of points is where time is measured in the interval of 2 minutes

F(y) =

The density function for this would be

f(y) =

With this the expected value can be calculated by =60( ) =120

(4.4)

In Search and destroy there is an attacking side and a defending side. The attacking side need to plan a bomb before the time runs out in order not to lose. Each round is 2 minutes. Using uniform distribution, we can figure out probabilities of when the bomb will be planted. Like the probability of planting the bomb in the last 45 seconds is = 0.625.

(5.2)

Using 4.2 and the probabilities of getting a kill in a search round. lets say we wanted to get the probability of a players KD being over or under one and they also get amount of kills in a search round. This distribution function would be

(5.3)

Here is a table for the join probability function for the problem of 2.5 of = Playing on Berlin and = Playing on Tuscan

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | Total |
| 0 | 4/16 | 4/16 | 1/16 | 9/16 |
| 1 | 4/16 | 2/16 | 0 | 6/16 |
| 2 | 1/16 | 0 | 0 | 1/16 |
| Total | 9/16 | 6/16 | 1/16 | 1 |

This table can be used to find P(Berlin =1|Tuscan = 1)== =

(5.4)

Using the density function for 5.2, testing f(x,y) = f(x)f(y) will show if the variables and independent or dependent.

Because f(x,y) ≠ 0.2456 these two variables are dependent on each other in the function.

1. Overall Player Statistics table: <https://breakingpoint.gg/player-stats/2022/advanced-stats/>
2. Team Statistics: <https://breakingpoint.gg/team-stats/2022/>