

League of Legends: Inferential Statistics

Intro

After going through these projects, my project has shifted more towards giving players more of an advantage based on data, inferential statistics and advanced algorithms.

Question

So my question was if and how much of a difference does it make if you are assigned to the blue side in contrast to the red side.

Background

Take note that the assignment is random but since there are objectives on the map that are fixed on a certain side, champions that do need to go a certain jungle route and this will be fixed based on which side you are assigned on, as well as one more teammate on the bottom side, there are several variables in my head that will be affected by which side you are assigned on.

Process

1. So I performed a two-sample Z-test. Even though I don't have the population proportion, we have a significant number of records in our sample that will give us a relatively normal distribution for our sampling distribution. It is a two sample test because we have the proportion of wins on the blue side and proportion of wins on the red side.
2. I checked for the conditions for the Central Limit Theorem so that we are able to use the following statistical techniques revolving around the normal distribution and our samples check out.
3. Our null hypothesis was that there was no difference between the proportion of wins on the blue side and proportion of wins on the red side. Our alternative hypothesis was that there was a difference between the two.
4. I assigned our significance level to .01 because I wanted to decrease the power and decrease the probability of making a type I error because if

players leave every time they see they're on a specific side that they think makes a difference but actually doesn't, they'll be wasting A LOT of time and rank points by leaving the game before the game starts whereas if they stay in even though there might be a slight chance of having a disadvantage, you will still be able to play it out.

Conclusion

Within the analysis in my jupyter notebook, I show that there is a difference and it is statistically significant relative to my significance level but my confidence interval is so miniscule that although it is statistically significant, it is not practically significant enough for me to suggest players to leave a game before the game starts to optimize their time and energy.

Final Thoughts

This was an extremely interesting insight that I've found and may look to further delve into this and maybe propose suggestions and different ways to balance the difference in winning percentages on the different sides so that right off the bat, the chances are equal for both teams going into a ranked match.