PERFORMANCE OF JUNGLER'S CHAMPIONS IN THE GAME LEGUE OF LEGENDS®

Victor Ferreira SILVA[1]
Édipo Menezes SILVA[1]
Kelly Pereira LIMA[1]
João Domingos SCALON[1]

■ ABSTRACT: Since the 1980s, there has been a widespread of high-speed internet connections, which is causing an expansion of the multiplayer electronic games industry. These games have gained the status of sport (called e-sports). Among these games, Lol (League of Legends®) deserves special mention. In this game, players form a team of five and assume the role of a "champion" (characters with unique abilities) generally varying around a type of class, and battle against a team of player (or computer-controlled champions). A recurring problem is to balance these champions, that is, to leave the game as honest as possible, so that only the player's ability will be a decisive factor for the victory. The aim of this work is to perform statistical analysis of the win rates of four champions of the tenth season of the Lol game mode "Wild Rift", doing control of the pachs. Using ANOVA and Tukey's HSD test, we showed that even when we take only the champions with the highest winning rates, one of them stands out from the others.

■ KEYWORDS: ANOVA; MOBA; e-sports; video games; Lol.

1 Introduction

The electronic games industry grows every year, having started its expansion in the nineties. Today, with the widespread use of the high-speed internet connections, multiplayer games have grown and received sport status. The practitioners of

[1]Universidade Federal de Lavras - UFLA, Instituto de Ciências Exatas e Tecnológicas, Departamento de Estatística, CEP: 37200-900, Lavras, MG, Brasil. E-mail: *victor.est.ufop@gmail.com; ediposvm01@gmail.com; kelly_limadm@hotmail.com; scalon@ufla.br*

these games (called e-sports) have training, salary, and prizes identical to those of traditional sports competitors.

In the current state of e-sports, the flagship of all success and profitability is the Lol (*League of Legends*®) (RIOT GAMES, 2009). Lol was created in South Korea in 2010 and nowadays has more than twenty million players. Even in Brazil, the game has a huge success, and the national championship has been broadcasting live on cable television.

Lol is a team-based strategy electronic game where two teams of five powerful champions (characters with unique abilities, generally varying around a type of class) face off to destroy the other's base. Choose from over 140 champions to make epic plays, secure kills, and take down towers as the players battle their way to victory. Champions are classified by classes and subclasses. Classes designate broad traits while subclasses denote offensive or protective capabilities. The major classes and their subclasses are: thanks, fighters, slayers, mages, controllers, and marksmen. Summoner's Rift is the flagship game mode of League of Legends and the only one featured in professional-level play. In this game mode, two teams of five players compete to destroy the enemy's Nexus, which is guarded by the enemy team and defensive structures known as turrets. (RIOT GAMES, 2009).

Wild Rift is the mobile version of Lol played on a smaller map compared to Summoner's Rift. Like other Multiplayer Online Battle Arena (MOBA), the game is a 5x5 battle, where your goal is to take down the other teams *Nexus* located at the heart of the base directly opposite yours. The team on the left is the blue team, while the team on the right is red. All matches are mirrored, so you will always play a game from the blue side even though you could be on the red, due to the nature of mobile controls where the movement wheel is on the left and abilities are on the right. Wild Rift has three lanes that run at the top, mid and bottom. In between lies the jungle that contains jungle monsters, or creeps. As such, champions will fit into different roles in-game in designated locations on the map. The most unique role in both LoL and Wild Rift is the Jungler. The Jungler is a player who thrives in the jungle part of the map (ONEESPORTS, 2020).

As in any sport like football, baseball, etc. which have a lot of fans and involve large amounts of money, naturally, the need for performance statistics arises. The open source website (OP.GG, 2020) is the most popular repository of builds, match results, weekly win rate variation, among other important statistics on how the game is developing over the seasons. The question that arises is: how can these data be used to better understand the sport? To the best of our knowledge, to date, there are no published papers that analyse Lol data. For exemple, a recurring problem of Lol, that can be understood using statistical analysis, is balancing these champions in order to make the game as honest as possible, so that the player's skill will be a decisive factor for the team's victory, and not the choice of a certain champion who may have a kit of superior powers than the others.

The aim of this work is to perform an analysis of variance on the available data of matches from "Wild Rift's Wheel of the Jungle" recorded in real time on the game server in order to understand the performance of the champions. More

specifically, we are testing the null hypothesis that there is no statistically significant difference between the mean winning rates of the four chosen champions.

2   Material and methods

2.1   The data set

We start by setting the first four champions in relation to their win rate, following: *Karthus, Olaf, Elise*, and *Zac*. Then, we define, as the study subject in this paper, the Jungler's position. This position is the one that most influences the results of "Solo Queue match", because it is the least dependent on the help of another player.

Then, we set the response variable as being the win rate of each champion, and the nuisance factors as being each patch (update) of the tenth season, from 10.1 to 10.4. We observe that the updates, in a simple way to explain, are patches installed to do adjustments that are made periodically in the game's functionality in order to balance, and allow a more fluid game play without dominance.

The data regarding the winning rates of the four champions and runs in each patch were taken directly from the open source website OP.GG (2020). which contains updated statistics in real time from all servers. the first 2000 matches of each week were used to take each champion's win rate.

2.2   The randomized complete block design

Since we have one factor (or variable) that is of primary interest (champion's win rate), and there are also several other nuisance factors (patch) that must be controlled, we have a Randomized Complete Block Design (RCBD) also known as the two-way ANOVA without interaction. One useful way to look at a RCBD is to consider each block as a collection of completely randomized experiments. A key assumption in the analysis is that the effect of each level of the treatment factor is the same for each level of the blocking factor.

The model for an RCBD (or two-way ANOVA without interactions) is:

$$Y_{ij} = \mu + T_i + B_j + \epsilon$$

where:

- Y_{ij}: is any observation for the treatment i and block j;

- μ: is the general location parameter (that is, the overall mean of all experimental units);

- T_i: is the effect of treatment i ;

- B_j: is the effect of block j;

- ϵ: are random errors that are normally distributed with mean zero and unknown standard deviation, independent of the values for all other errors.

There are two hypothesis tests in an RCBD, and they are always the same:

- H_0: The means of all treatments are equal, versus

- H_a: At least one of the treatments has a different mean

  and

- H_0: The means of all blocks are equal, versus

- H_a: At least one of the blocks has a different mean

In the case of an RCBD, we divide the total sum of squares into the part attributable to differences between the treatment means, the part attributable to difference between the blocks, and the part attributable to differences within block-treatment groups, or the sum of squares due to pure error.

The test statistics for an RCBD look at the standardized ratio of the between-group sum of squares to the within-group sum of squares for the treatment and block effects separately.

Once we have calculated the test statistics, we refer them to F distributions. For the test of equal treatment effects, the numerator is $(i-1)$ degrees of freedom (df). For the test of equal block effects, the numerator is $(j-1)$ df. Both tests have the same denominator $(i-1)(j-1)$ df. If one or both are large (small p-values), then it suggest that some of the treatment or block effects are not equal.

In order to simplify the organization of the calculations in an RCBD, it is customary to write things in a usual analysis of variance (ANOVA) table. For more details on RCBD see Addelman (1969) and Box et al. (1978).

Regarding random errors, homogeneity between win rate variances was verified using the Oneill-Mathews tests (O'NEILL; MATHEWS, 2000), while normality of distribution and error randomness was assessed by the Shapiro-Wilk test (CANTELMO; FERREIRA, 2000).

After running an ANOVA and found significant results, then we can run Tukey's HSD to find out which specific groups?s means (compared with each other) are different. The test compares all possible pairs of means (BOX et al., 1978)

We performed the ANOVA of the data from the RCBD using the EspDes.pt (2014) available at the statistical software R (R CORE TEAM, 2020). It was used a significance level of 5% in all analyzes.

3   Results

We start our analysis by testing the random error assumptions of normality and homogeneity of the ANOVA model. We have got p-values equals to 0.6051 and

0.4399 for the Shapiro-Wilk and Oneill-Mathews tests, respectively, and therefore, we accept those assumptions. Thus, we can procede with the ANOVA to determine whether or not the four different champions produce different winning rates on the Lol game. Table 1 shows the output of ANOVA of the RCDB.

Table 1 - ANOVA of the RCDB

FV	DF	SS	MS	Fc	P-Value
Champions	3	134.858	44.953	81.680	0.00616
Patch	3	0.4707	0.1569	0.2851	0.83504
Residuals	9	49.532	0.5504		
Total	15	189.097			
CV = 1.43 %					

Table 1 shows that it would appear that there are significant differences between the the mean win rates of the four champions (p=0.00616) but not between the patches (p=0.83504) at the 0.05 significance level.

We observe that a non-significant block variable doesn't necessarily mean that the blocking was ineffective. It just wasn't significant. Given the nature of randomized complete block design, i.e. all factor levels are replicated randomly in each block, the blocking variable (patchs) may still account for potential variation present in the experiment (BOX et al., 1978). As we have pointed out before, the block was used to remove uncontrolled variation (patch) from our investigation. In other words, we have used blocking to permit a better understanding of the effect of the variables under consideration independent of the effect of those terms confounded with blocks.

Although ANOVA is a powerful and useful parametric approach to analyzing approximately normally distributed data with more than two groups, it does not provide any deeper insights into patterns or comparisons between specific groups. Thus, After conducting an ANOVA, it is often desired to know more about the specific groups to find out if they are significantly different. This step after ANOVA is referred to as 'post-hoc analysis'.

One common and popular method of post-hoc analysis is Tukey's Test. The test is known by several different names such as the Tukey HSD ("honestly significant difference" or "honest significant difference"). Tukey's test compares the means of all treatments to the mean of every other treatment, and is considered the best available method in cases when confidence intervals are desired or if sample sizes are unequal.

Table 2 shows the output of the Tukey HSD of the ANOVA presented in table 1.

Table 2 - Tukey HSD

Group	Champion	Win rate
a	*Karthus*	53.5
b	*Olaf*	51.53
b	*Zac*	51.32
b	*Elsie*	50.93

The output presented in table 2 gives the win rate means for each champion, and the significance of between-group differences. Different letters mean that the treatment means are statistically different at the 0.05. In other words, the champion *Karthus* is in a cluster all on its own because it was significantly different from the other three. *Olaf*, *Zac*, and *Elsie* are in a cluster together because their difference in win rate means was not large enough to be significant.

Conclusions

The ANOVA and the Tukey's HSD test of the RCDB seem to be plausible tools for detecting a possible imbalance in the Lol game. The statistical methods used have shown that there is statistically significant difference between the winning rate means of the four chosen champions. Our approach allow us to advocate that even when we take only the champions with the highest winning rate, it is still possible that one of them stands out from the others.

Acknowledgments

SILVA, V. F.; SILVA, E. M.; LIMA, K. P.; SCALON,J . D. Análise da variabilidade do desempenho dos campeões da selva no jogo Legue of Legends® . *Rev. Bras. Biom.,* Lavras, v.39, n.1, p.114-121, 2021.

■ *RESUMO: Desde os anos 80, observa-se uma grande difusão de conexões de internet de alta velocidade, o que está acarretando uma expansão da indústria de jogos eletrônicos multiplayers. Esses jogos ganharam o status de esporte (chamados de e-sports). Dentre esses jogos, merece destaque o Lol (League of Legends®). Neste jogo, os jogadores assumem o papel de "invocadores", que controlam campeões com habilidades únicas, que formam um time e lutam contra o time adversário de outros invocadores ou controlados pelo computador. Um problema recorrente é balancear esses campeões, ou seja, deixar o jogo o mais honesto possível, de forma que apenas a habilidade do jogador venha ser um fator decisivo para a vitória. O objetivo deste trabalho é analisar estatisticamente as taxas de vitória de quatro campeões da décima temporada fazendo controle dos pachs. Usando ANOVA e o teste HSD de Tukey pode-se mostrar que mesmo quando tomamos apenas os campeões com as maiores taxas de vitórias, um deles se destaca dos demais.*

■ *PALAVRAS-CHAVE: ANOVA; MOBA; e-sports; video games; Lol.*

References

ADDELMAN, S. The Generalized Randomized Block Design. *The American Statistician*, v.23, n.4, p.35-36, 1969.

BOX, G. E. P.; HUNTER, W. G.; HUNTER, J. S. *Statistics for experimenters: an introduction to design, data analysis, and model building.* New York: Wiley, 1978. 653p.

CANTELMO, N. F.; FERREIRA, D. F. Desempenho de testes de normalidade multivariados avaliado por simulação Monte Carlo. Ciência e Agrotecnologia, v.31, n.6, p.1630-1636, 2007.

FERREIRA, E.; CAVALCANTI, P.; NOGUEIRA, D. ExpDes: An R Package for ANOVA and Experimental Designs. *Applied Mathematics*, v.5, p.2952-2958, 2014.

RIOT GAMES. League of legends ® [Video game]. *Riot Games*, Los Angeles, CA, 2009.

R CORE TEAM *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

ONEESPORTS. Available at: *https://www.oneesports.gg/wild-rift/everything-you-should-know-before-you-start-playing-wild-rift/.* Access in: December 2020.

O'NEILL, M. E.; MATHEWS, K. L. A weighted least squares approach to Levene test of homogeneity of variance. *Australian e New Zealand Journal Statistical*, v.42, n.1, p.81-100, 2000.

OPP.GG. Statistics. Available at: *https://br.op.gg/ranking/ladder/page=7858*. Access in : March 28-2020.

VIEIRA, V. A. Experimental designs using ANOVA. *Revista de Administração Contemporânea*, v.15, n.2, p.363-365, 2011.