League of Legend: Destroy Towers Study

Application Ordinal Logistic Regression

Boram Shim

University of Minnesota

Abstract

Ordinal logistic regression is statistical method to apply the relation of variables with three or more response categories. The main purpose of this research is to analysis which factors influence on destroying towers in League of Legend game by using ordinal logistic regression. The results of this research describe that every 1 unit change in kills, the log odds of being accepted increased by 1.56. In other words, the more the number of kills, the faster the tower destroyed.

League of Legend: Destroy Towers Study

Application Ordinal Logistic Regression

Introduction and Background

League of Legends is a multiplayer online battle arena (MOBA) video game released by Riot Games in 2007. It has gained popularity by lowering the barriers to entry than previous MOBA games, and currently has more than 100 million users worldwide playing every month and holds the largest number of viewers of MOBA streams worldwide. The League of Legend is an example of e-Sports, and is hosted by Riot, and the League of Legends Championship Korea (LCK) is one of most popular e-Sports events in the world. Each team consists of five champions for each lane: Top, Mid, Bottom & Support, and Jungle lanes. There are total of 15 towers on the base of each team, and the rules of the game determine which team destroys 15 towers first. The main purpose of this paper is to examine what has the greatest impact on destroying towers, which is a winning condition using ordinal logistic regression. This paper consists of Literature Review, Method and Material, Conclusion and Discussion. In the Literature Review, there are three statistical papers that defined the ordinal logistic regression. In Method and Material, the data is described and a brief description of the statistical technique using collected data is provided. In Conclusion and Discussion, the summary of the analysis is provided, the meaning of conclusions is explained, and the potential issues and further research are discussed.

Literature Review

15.1 – Logistic Regression (n.d.) explains that Ordinal Logistic Regression is applied with three or more the category of the response with ordering to the levels. It also shows that ordinal logistic regression R output. The UCLA Institute for Digital Research and Education (n.d.) shows that ordinal logistic regression is called the proportional odds assumption or the parallel regression assumptions because between all group pairs are same and there is only one model. This model is based on the coefficients that describe the relationships between each category of the response variable. The outcome and the dependent variable are binary as 0 or 1. 8.4 - The Proportional-Odds Cumulative Logit Model. (n.d.) describes “Proportional-odds cumulative logit model is possibly the most popular model for ordinal data. This model uses cumulative probabilities up to a threshold, thereby making the whole range of ordinal categories binary at that threshold.” These articles show how to enter data to analyze logistic datasets.

Method and Materials

This research uses Ordinal Logistic Regression to analyze prediction between the number of kills (which is the continuous variable) and the number of destroyed towers (which is the ordered categorical response). There are many variables that can affect breaking a tower, but in this study, three of the most concentrated variables in a gaming competition were set up as the basis of the model. This research model has a sample size of 79 with the number of destroyed tower scaled for response and the number of kills, assists, and playing time for variables. The number of kills of a player refers to the number of other players they eliminate, the number of assists of a player means refers to the number of other players they help to eliminate, and playing time means that how long a player plays the game. The number of towers is counted as 1 to 15 was categorized in to “Low” if tower was destroyed 1 to 8, “Medium” for destroying 9 to 14, and “High” for 15. Independent variable is kills, assists, and playing time which are continuous.

$$r = \frac{\sum(x - m_x)(y - m_y)}{\sqrt{\sum(x - m_x)^2 \sum(y - m_y)^2}} \tag{1}$$

Table 1. Correlation with destroyed towers

	Kill	Assist	Time
Correlation	0.713	0.407	0.293

First, using the correlation equation (1), how related given variables to the response. Table 1 shows correlation result that the only destroyed towers and kills correlation is higher than 0.5, so there is a positive correlation between the number of destroyed towers and the number of kills.

Below equation (2a) and (2b) is cumulative logit form with ordering of the response where j= 1, …, j-1 and these two equations is used for analysis. On the right sides of the equal sign is a linear model with slope β and an intercept depends on j, αj (University of Virginia Library Research Data Services Sciences n.d).

$$\text{logit}(P(Y \leq j \mid X)) = \alpha j - \beta_1 \text{ Kill} - \beta_2 \text{ Assist} - \beta_3 \text{ Time} \tag{2a}$$

$$\text{logit}(P(Y \leq j \mid X)) = \alpha j - \beta_1 \text{Kill} \tag{2b}$$

Equations (2a) and (2b) are different models with or without Assist and Time variables. Below the equation (3) measures that log odds of destroyed towers level of “Low” versus “Medium” or “High” combined, and the equation (4) measures that log odds of destroyed towers level of “Low” or “Medium” combined versus “High”.

$$L_1 = \text{logit}(P(Y \leq 1 \mid X)) = \log\left(\frac{P(Y \leq 1 \mid X)}{P(Y > 1 \mid X)}\right) = \log\left(\frac{\pi(1)}{\pi(2) + \pi(3)}\right) \tag{3}$$

$$L_1 = \text{logit}(P(Y \leq 2 \mid X)) = \log\left(\frac{P(Y \leq 2 \mid X)}{P(Y > 2 \mid X)}\right) = \log\left(\frac{\pi(1) + \pi(2)}{\pi(3)}\right) \tag{4}$$

Below Table 2 and Table 3 are the results of the ordinal logistic regression model of (2a) and (2b). To select the best model for the analysis, it is checked AIC first. AIC is a constant estimator and the relative distance between the unknown true likelihood function and the fitted likelihood function of the model (AIC vs. BIC n.d). Therefore, small AIC model is desirable.

Table 2. Regression results for model (2a). se= standard error

Coefficient	j= 1	j=2	P value
Intercept	1.834 (se= 1.645)	3.665 (se= 1.734)	0.2648(j=1), 0.0345 (j=2)
Kill	0.373, (se= 0.091)		0.00004
Assist	0.124 (se= 0.06565)		0.059
Time	-0.044 (se= 0.04873)		0.362
Residual deviance	87.123		
AIC	97.123		

Table 3. Regression results for model (2b). se= standard error

Coefficient	j= 1	j=2	P value
Intercept	3.006 (se= 0.729)	4.728 (se= 0.893)	0.00001 (j=1), 0 (j=2)
Kill	0.445 (se= 0.0792)		0
Residual deviance	91.718		
AIC	97.718		

From the value of AIC, model (2b) AIC is 97.718 which is slightly higher than model (2a) AIC which is 97.123, so both model can be selected. However, above Table 2, model (2a) p-value (0.2648 > 0.05) is insignificant, so final model is (2b). Table2 and Table 3 show that only Kill p- value is less than 0.05, it is significant statistically. From Table 3, with using ordinal logistic regression equation, as kill increases by one unit, the number of destroyed towers is expected to increase by 0.445 in the log odd scale. The results of Table 3 apply to the formula (5) and (6) below, then odds ratio comes out by applying formula (7).

Cumulative probability: $(P(Y \leq j \mid X)) = \frac{\exp(\alpha j - \beta 1 \mathrm{Kill})}{1+\exp(\alpha j - \beta 1 \mathrm{Kill})}$ (5)

Odds: $\frac{\mathrm{P}(\mathrm{Y} \leq \mathrm{j} \mid \mathrm{X})}{\mathrm{P}(\mathrm{Y} > j \mid X)} = \exp(\alpha j - \beta_1 \text{Kill})$ (6)

Odds Ratio: $\frac{\frac{\mathrm{P}(\mathrm{Y} \leq Y_j \mid \text{less Kill})}{\mathrm{P}(\mathrm{Y} > Y_j \mid \text{less Kill})}}{\frac{\mathrm{P}(\mathrm{Y} \leq Y_j \mid \text{high Kill})}{\mathrm{P}(\mathrm{Y} > Y_j \mid \text{high Kill})}} = \frac{odds(\text{less Kill})}{odds(\text{high Kill})}$ (7)

Table 4. Odds ratio for each category and Kill

	j=1 (Low)		j=2 (Medium | High)	
	Less Kill	High Kill	Less Kill	High Kill
Odds Ratio	1.560 or exp (0.445)			

Table 4 represents odds ratio for each category. High kill means that the number of kills has high possibility to destroy towers than against teams and less kill means that the number of kills has less possibility to destroy towers than against teams. Using the equation (7), the odds ratio gets 1.56 which is same as the exponential of the coefficients of the model (2b). It represents that for every 1 unit change in kills, the log odds of being accepted increased by 1.56.

Conclusion and Discussion

This research analysis to examine the relationship between the level of destroyed towers and effected variables with the using ordinal logistic regression. The results of this research show that the odds of the number of high kills can destroy towers 1.56 more than the odds of the number of less kill. Because this research is an analysis of a game called League of Legend, there are a variety of unpredictable variables. There are two limitations. One limitation of this research is the fact that since it is a competitive game, the outcome of the game is determined by both sets of players. Thus, the data set is dynamically evolving every moment, making it difficult to perform statistical analysis. If there is an unaltered absolute figure associated with the champion's ability, it may be changed the approach of this research to get more accurate results. Further, since the game is highly strategic, players may choose to focus on simply breaking towers as opposed to getting kills. In this research pattern of analysis, a player following this strategy would skew the data set because of their low kills even though they achieved the objective. Thus, this model would not predict their victory. If it is possible to collect the data of strategy, the most significant variable may change to strategy data from the number of kills.

References

AIC vs. BIC. (n.d.). Retrieved from https://www.methodology.psu.edu/resources/aic-vs-bic/

Champions - League of Legends - League of Legends. (n.d.). Retrieved May 01, 2019, from https://na.op.gg/champion/statistics

South Korea / League of Legend Inventory. (n.d.). Retrieved May 01, 2019, from http://lol.inven.co.kr/

University of Virginia Library Research Data Services Sciences. (n.d). Retrieved from https://data.library.virginia.edu/fitting-and-interpreting-a-proportional-odds-model/

UCLA Institute for Digital Research and Education, n.d. (n.d.). Retrieved May 01, 2019, from https://stats.idre.ucla.edu/stata/dae/ordered-logistic-regression/

8.4 - The Proportional-Odds Cumulative Logit Model. (n.d.). Retrieved May 01, 2019, from https://onlinecourses.science.psu.edu/stat504/node/176

15.1 – Logistic Regression (n.d.). Retrieved May 01, 2019, from https://newonlinecourses.science.psu.edu/stat501/node/374/