THE PREDICTION OF SECOND AND THIRD ROUND KNOCKOUTS USING SIGNIFICANT STRIKE METRICS ACROSS MIXED MARTIAL ARTS FIGHTS

David Deniziuk

Abstract

Mixed Martial Arts is a full contact sport that involves two combatants fighting each other inside an octagon. Previously published literature has delved into various questions relating to the performance of Mixed Martial Arts athletes. In this paper we study the relationship between various forms of significant strikes doled out by these fighters and the likelihood of them attaining a knockout or technical knockout. Specifically, we run binary logistic regressions to see if proportions of different strikes are significant in predicting which fighter achieves the knockout or technical knockout in subsequent rounds. We also make use of backwards selection, Akaike Information Criterion, and a validation set to select and test our models. We conclude that the proportion of first round head strikes is significant in predicting which fighter will achieve the knockout or technical knockout in the second round. We also conclude that the aggregate proportion of first and second round body strikes is significant in predicting third round knockouts. Additionally, our Akaike Information Criterion selected model for second round knockouts is predictively accurate, as proven by its validation set performance. These results support the idea that fighters should prioritize head and body strikes to maximize knockout probability.

Introduction

Mixed Martial Arts (MMA) is a sport that incorporates techniques from many different forms of martial arts. Techniques from Brazilian jiu-jitsu, Muay Thai, boxing, judo, wresting, along with others, are incorporated into MMA fighting styles. There are various ways a bout can end, and a winner decided, including knockouts (KO), technical knockouts (TKO), submissions, and decisions. Knockouts are very common in UFC competition, with more than half of all fights ending in this manner. They involve one fighter delivering a strike (punch, kick, elbow, knee, etc.) to his or her opponent that renders the opponent unable to continue, due to a loss of consciousness. Technical knockouts are very similar with the defendant remains conscious but unable to intelligently defend himself/herself. Knockouts and technical knockouts are seen as desirable ways to win a match due to there being no judge decision involved. Also, for most fighters, they are preferred over submissions. It is for these reasons that we are studying whether certain types of significant strikes result in an increased KO/TKO likelihood. Previous studies have looked at other metrics related to fight competency including the efficacy of perceived aggressiveness in MMA winner prediction, the use of matching law equations to determine strike selection, and a determination of fighter success based on physical characteristics. Although these studies provide insight into MMA success to some degree, they are all distinct from the hypothesis we are testing.

We are interested in the idea of a fighter being chronically weakened over the duration of a bout and how this impacts their susceptibility to fight ending strikes. Through this study, we test the impact of certain types of significant strikes: significant body strikes, significant clinch strikes, significant ground strikes, significant head strikes, and significant leg strikes on the likelihood of KO/TKO. Specifically, we will test whether the proportion of these strikes a certain fighter delivers in *previous* rounds influences the likelihood of that player attaining a KO in the *subsequent* round. Later in the paper, we will discuss this distinction, and its importance, further.

We are undertaking this study because we believe the application of quantitative methods to MMA will be beneficial to the sport as a whole. Statistical modeling and hypothesis testing have long been applied to other sports. The use of Sabermetrics in baseball is the quintessential example of statistical modeling being used to not only analyze the game from an outsider's perspective, but also from the inside. The advent of statistics in the game of baseball allowed the teams, themselves, to make more statistically informed decisions to maximize winnings. Although the scope of my study is much narrower than that of sabermetrics, the results are meaningful for the same reasons: inference is being utilized to improve athletic performance.

Statistical results from the hypothesis tests conducted will allow fighters to more intelligently select their strike type. This theoretically gives the fighter a higher probability of knocking out his/her opponent and therefore a higher probability of winning the bout. Additionally, the results from this study can aid athletes in defense, as they now know which strikes are most devastating from a knockout and technical knockout perspective. This can lead to increased training targeted to counter specific strikes. Overall, the results of this paper improve the sport, as a whole, as they help raise the bar for athletes and allow them to compete at higher levels.

Literature Review

The first journal article, by Trebický et al., tested whether perceived aggressiveness predicts win rate in MMA fighters. Essentially, the study looked to determine whether or not certain facial characteristics were correlated with fighting success (Trebický et al., 2013). Individuals from the Czech Republic rated pictures of MMA fighters in both perceived fighting ability as well as perceived aggressiveness. These predictors were then tested against actual fighting success. Multiple linear regression was utilized to test the hypotheses. It was shown that perceived aggressiveness and perceived fighting ability were, in fact, positively correlated with win proportion (Trebický et al., 2013). It should be noted that both of these relationships were confounded with weight. For the correlation between perceived aggressiveness and win rate, the relationship remained statistically significant after controlling for weight. For the correlation between perceived fighting ability and win rate, however, the relationship only remained significant for fighters competing in the heavyweight division (Trebický et al., 2013).

The second article, by Seniuk et al., studied the application of the Generalized Matching Equation to strike selection in MMA competition. The Generalized Matching Equation is based on the matching law, a behavioral principle that deals with conditioning and response (Seniuk et al., 2019). It essentially states that a person's rate of response to something is related to the rate of reinforcement (Seniuk et al., 2019). This paper builds upon the preexisting literature on the application of matching law to team sports, however it was the first application of the principle to an individual sport. In this study, a set of UFC fighters were selected, and a subset of their previous fights were analyzed. In the setting of MMA, reinforcement is the landing of significant strikes, and response is the fighter's strike selection (Seniuk et al., 2019). Linear regression was used in assessing accuracy of the matching equation in predicting strike selection. It was found that the matching equation had high predictive accuracy with most of the fighters. Although more research is needed, it appears to be a good predictor of strike selection. A broad use of this newfound matching law relationship is with MMA coaching, as incorporating techniques to make a fighter's behavior less predictable would be advantageous for the fighter (Seniuk et al., 2019).

This last article, by James et al., deals with studying the physical characteristics which predict MMA fighters' success. Prior to this article's publication, little was known about the relationship between physical characteristics and mixed martial arts ability (James et al., 2016). This study was the first to shed light on this relationship. The article is a meta-analysis of previous studies which tracked the performance and characteristics of differing fighters within a given sport. These include ones with MMA fighters, as well as others with fighters from disciplines that are closely related to MMA. These disciplines include judo, Brazilian jiu-jitsu, kickboxing, etc. Data was then extracted from these studies and used in the meta-analysis. Some of the characteristics studied include strength, maximal neuromuscular power, aerobic capacity, and anaerobic capacity (James et al., 2016). There were multiple methods used in this study. Frequently, a confidence interval for a given metric related to strength and general physiology would be compared between elite competitors and amateur or less elite competitors within a given sport. Significant differences in a metric between differing levels of competition would imply that said metric may be important in predicting skill level. To extend further, if a certain type or expression of that metric is present (e.g. high anaerobic capacity), it may be more likely that the subject competes at an elite level (James et al., 2016). The results of this study are split between grappling and striking athletes, because both techniques are utilized in MMA competition, however, significantly differ in characteristics. For grappling athletes, dynamic strength and high-force neuromuscular power were significant indicators. For striking athletes, high-force efforts were less pronounced, however rapid expression of force proved more important (James et al., 2016).

These three articles all relate to performance in MMA fights. The Trebický and James articles study predictors of MMA success, while the Seniuk article looks to predict specific actions within a fight, as opposed to the result of the fight as a whole. The Seniuk article uses data from within MMA fights as a predictor, while the Trebický and James articles use data that relates to physical characteristics as predictors. Also, it should be noted that the Trebický article uses subjective interpretations from participants, while the other two use only physically measurable metrics. The use of subjective interpretation is necessary, however, as “perceived aggressiveness” and “perceived fighting ability” are purely subjective in nature.

Methodology

In this study, we attempt to determine whether certain types of significant strikes influence the likelihood of attaining a knockout or technical knockout in later rounds of UFC fights. To test this, we utilize multiple strike-related predictor variables. The null hypothesis, H_0, is that given a fight ends in KO/TKO, the proportion of significant strikes landed by a given fighter does not influence their likelihood of being the one who attains the KO/TKO. The alternative hypothesis, H_A, is that given the fight ends in KO/TKO, the proportion of at least one type of significant strike landed by a given fighter *does* significantly influence their likelihood of being the one who attains the KO/TKO.

Our data are a sample of UFC fights from February 2014 through part of October 2018. The original dataset contained every UFC fight throughout this timespan, however not all fights contained the variables we want to test for, so we therefore removed a subset of observations, accordingly. Fights lasting five rounds, the format for championship bouts, were also removed due to the complexity they add to the analysis. The added two rounds of five round fights affect the "chronically weakened" characteristic that we are looking at, as the extra rounds will allow for more "chronic weakening" through additional significant strikes. For this reason, we believe it best to leave these out of the study. We also remove observations that do not result in a KO or TKO in the second or third rounds. This is done in order to better fit the models and isolate the variables we want to test for.

Each observation represents a UFC match that results in a round two or round three KO/TKO. There are roughly 161 observations in the combined dataset that fit the aforementioned criteria. 100 observations are of round two knockouts and 61 of round three knockouts. It should be noted that competitors who compete on multiple occasions could show up in multiple observations. This is allowed based on our definition of observation. Our population of interest is all MMA matches that result in KO/TKO, although the results of this study are especially relevant to those who compete within UFC.

The data were obtained from the website Kaggle.com. The creator of the dataset credits a github user for the "UFC api". This api likely scraped the data from a site with public fight data. Although the original dataset contained every observation within the specified time period, it is possible that as the observations with missing predictors were removed, the data became less random, but this appears unlikely.

Based on the quantity and apparent randomness of the removed observations, the data are very representative of the population. Upon downloading the data from Kaggle, we cleaned the data within excel. This included removing title bout observations, observations with incomplete predictor variables, and observations that did not result in round two or three KO/TKO. We used all of the remaining data and did not resample in any way. We did this in order to maximize our sample size, n, relative to our number of predictors, p. Along with removing certain observations, we also removed hundreds of unnecessary predictors that cluttered the dataset. These unnecessary variables ranged from attempted strikes, to ground time, to hometown. With this newly filtered dataset, we proceeded with our testing.

The response variable in our study is the likelihood of the fighter from the red corner attaining a KO/TKO in the given round. Before the fight, the underdog is assigned to the blue corner and the favorite to the red corner. We will always measure the probability for the red fighter winning for consistency. This is measured via a simple probability of the red fighter attaining the KO/TKO in the round of interest. Because all of our observations involve a KO/TKO, the sum of both fighters' individual probabilities is one. In our study, this knockout likelihood will be measured in both the second and third rounds, but not the first. This is because there are no rounds before the first for the fighter to land significant strikes. Our predictor variables are as follows:

- Significant Body Strike Proportion: Quantitative variable describing the proportion of significant strikes delivered by red fighter to opponent's torso region.
- Significant Clinch Strike Proportion: Quantitative variable describing the proportion of significant strikes delivered by red fighter while in the clinch position. The clinch position is defined as the position taken when both fighters grab ahold of one another and hold each other's body close to their own.
- Significant Ground Strike Proportion: Quantitative variable describing the proportion of significant strikes delivered by red fighter while both the combatant and the defender are on the ground (not standing on feet).
- Significant Head Strike Proportion: Quantitative variable describing the proportion of significant strikes delivered by red fighter to opponent's head.
- Significant Leg Strike Proportion: Quantitative variable describing the proportion of significant strikes delivered by red fighter to opponent's leg.

We use logistic regression to model the relationship between significant strikes and both round two and round three knockouts. Backward elimination is used to choose the best models with one through five predictors. Akaike Information Criterion, AIC, is then utilized to select the best model from the five candidate models. To further determine the generalizability error of the model, we split the data into train and test sets. With this new test/train split, we run the backwards elimination/AIC models to see how well they generalize.

In order to accurately predict the winner of a given bout, different types of significant strikes are aggregated. The response variable, *P*, represents the probability of the red corner fighter being the one who achieves the KO/TKO. The decision threshold is 0.5, meaning that if *P* is greater than 0.5, it is predicted that the red fighter wins. If *P* is less than 0.5, it is predicted that the blue fighter wins. The five predictors, X_1, X_2, X_3, X_4, X_5 , represent the proportion of total strikes that belong to red, for each category of strike. The full model for our research question regarding second round KO/TKO is as follows:

$$P = \frac{e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon}}{1 + e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon}}$$

$$logit(P) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon$$

Where the predictors, X_i, are defined as follows:

- X_1 is the proportion of total first round significant body strikes landed by Red
- X_2 is the proportion of total first round significant clinch strikes landed by Red
- X_3 is the proportion of total first round significant ground strikes landed by Red
- X_4 is the proportion of total first round significant head strikes landed by Red
- X_5 is the proportion of total first round significant legs strikes landed by Red

The full model for our second research question, regarding third round KO/TKO, is similar to that of the second round but includes aggregate data from both the first and second rounds. It is as follows:

$$P = \frac{e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon}}{1 + e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon}}$$

$$logit(P) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon$$

Where the predictors, X_i, are defined as follows:

- X_1 is the proportion of total first and second round significant body strikes landed by Red
- X_2 is the proportion of total first and second round significant clinch strikes landed by Red
- X_3 is the proportion of total first and second round significant ground strikes landed by Red
- X_4 is the proportion of total first and second round significant head strikes landed by Red
- X_5 is the proportion of total first and second round significant legs strikes landed by Red

We use the software language R in the integrated development environment RStudio to process, tidy, and model the data. We make special use of the tidyverse, including the dplyr package, to alter the data via pipe operators and simple functions. This simplifies the process of altering the dataset and creating the exact variables needed to run analyses. Glmnet, a package consisting of generalized linear models, is used to run the logistic regression. The MASS package, and its function stepAIC, is used to run backwards elimination from the full model and then choose the best model via the Akaike Information Criterion.

Because the response is categorical, the performance of the models will be judged by their "misclassification rates". The misclassification rate is defined as one minus the accuracy of the model. We will use this metric instead of MSE or R^2, as would be used in a linear regression setting. The misclassification rate which we use is that from the model being fit to the validation set. Along with looking at measures of model predictive accuracy, we will also test the overall significance of each model, and the significance of each of the predictors via the Wald test. The Wald test essentially uses type III sums of squares to test if the addition of a given predictor is significant. It reports values from a standard normal distribution along with their respective p-values.

Results

In this study, we test the question of whether various types of significant strikes in MMA bouts influence and accurately predict knockouts and technical knockouts in later rounds. We use a sample consisting of various UFC fights from February 2014 through part of October 2018. Along with testing the significance of the various predictors, we also hope to build working models with high predictive accuracy. Our response variable represents a win or loss for the fighter from the red corner. It takes a value of either zero or one. One is arbitrarily assigned to be a win for red, while zero is arbitrarily assigned to be a loss. For our data dealing with round two knockouts, 51.0% of the bouts result in a red win, while 49.0% result in a loss. For our data dealing with round three knockouts, 50.8% of the bouts result in a red win, and 49.2% result in a loss. The splits are shown below. Note that this is the actual split of wins in the data, not a result from any sort of analysis being run.

Figure 1

Round Two KO/TKO

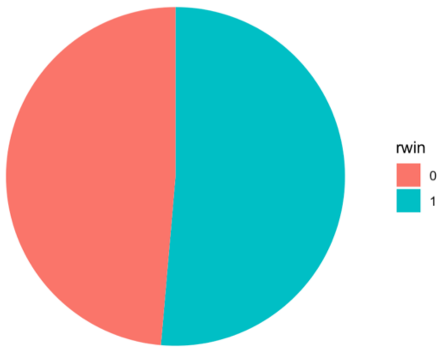
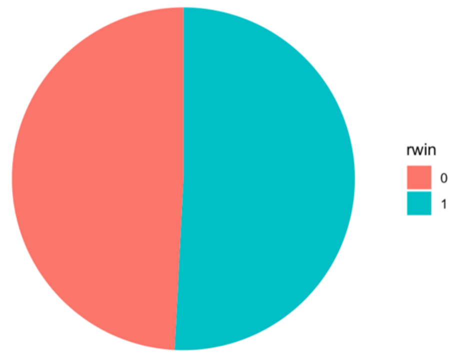


Figure 2

Round Three KO/TKO



Because the data has an almost perfectly even split between red and blue fighters winning, we do not have to worry about a class bias in the response variable. This should help the logistic regression to run well.

In the round two knockout data, the mean and standard deviation of each of the predictors over the first round is given in the table below:

Table 1

Quantitative Predictors Round One

Predictors	X_1	X_2	X_3	X_4	X_5
Mean	0.516	0.499	0.530	0.514	0.471
Sd	0.245	0.306	0.367	0.223	0.314

For the round three knockout data, the mean and standard deviation of each predictor aggregated over the first two rounds is given:

Table 2

Quantitative Predictors Rounds One and Two

Predictors	X_1	X_2	X_3	X_4	X_5
Mean	0.521	0.536	0.512	0.511	0.455
Sd	0.259	0.303	0.318	0.197	0.290

We first run analyses relating to our hypothesis about round two knockouts. The first model is the logistic regression with all predictor variables for the data concerning round two knockouts. After performing a test for overall significance in this model, we conclude that it is not significant, overall. The test statistic is ($\chi^2_5 = 9.28$, $p = 0.098$) while the chi-square critical value, selected at 95% confidence is 11.07, so we fail to reject the null hypothesis. This failure to reject is largely because of there being five predictors in the model, with none of them independently significant (besides the intercept) according to Wald test p-values. The output of Wald test values can be seen in the table below.

Table 3

Wald Values Full Model (Round Two KO/TKO)

Predictor	*Coefficient*	*Z-value*	*p-value*
Intercept	-1.481	-2.372	0.018
X_1	-0.072	-0.064	0.949
X_2	0.205	0.234	0.815
X_3	-0.186	-0.262	0.794
X_4	2.149	1.523	0.128
X_5	0.961	1.263	0.207

In order to improve this model, both in terms of predictive accuracy as well as significance in the predictors, we run backwards elimination to choose the best model at each predictor level, and then use AIC as the criterion to pick the best model. These methods arrive at the one variable model with round one significant head strikes proportion, X_4, being the only significant predictor. The process of the backwards elimination via AIC can be seen in the table below:

Table 4

Backward Elimination Via AIC (Round 2 KO/TKO)

Step	*Predictor Removed*	*Overall Model AIC*
1	Full Model	141.31
2	X_1 (Body)	139.31
3	X_2 (Clinch)	137.36
4	X_3 (Ground)	135.46
5	X_5 (Leg)	135.36

Now, in this reduced model, both the overall model, as well as the sole predictor, are significant. The test statistic for the overall model is ($\chi^2_1 = 7.23$, $p = 0.007$) while the chi-square critical value, selected at 95% confidence, is 3.84, so we reject the null hypothesis. We conclude that the model containing round one significant head strikes percentage is significant in predicting the winner of the match.

We then split the data into a training set and a validation set to attain a "generalizability error" of the model. This allows us to understand how the model will perform in predicting future matches. This validation error rate is more important than that of the training data. We do a 75/25 split (75% of data to training set, 25% of data to test set). When we build the model on the training set and then validate it on the test set, we arrive at an accuracy of 60.0%. This corresponds to a misclassification rate of 40%. This matrix of results can be seen below.

Table 5

Validation Set Accuracy Table (Round Two KO/TKO)

Winner	*Actually Blue (0)*	*Actually Red (1)*
Predicted Blue (0)	7	4
Predicted Red (1)	6	8

Next, we run models concerning our next hypothesis, round three knockouts. First is the logistic regression comprised of all of the predictor variables. After performing a test for overall significance in this model, we conclude that it is not significant, overall. The test statistic is (χ^2_5 = 7.47, p = 0.188), while the chi-square critical value, selected at 95% confidence is 11.07, so we fail to reject the null hypothesis. This failure to reject is largely because of there being five predictors in the model with only one having a significant Wald test p-value. The output can be seen in the table below.

Table 6

Wald Values Full Model (Round Three KO/TKO)

Predictor	*Coefficient*	*Z-value*	*p-value*
Intercept	-1.033	-1.274	0.203
X_1	3.411	2.044	0.041
X_2	-1.280	-1.014	0.311
X_3	0.072	0.071	0.943
X_4	-0.494	-0.268	0.789

X_5	0.434	0.348	0.728

Like we previously did with the round two knockouts, we run a backward elimination with selection criterion AIC on the round three knockout data to arrive at a better, more significant model. These methods arrive at the one variable model with round one and two significant body strikes percentage being the only significant predictor. The process of the backwards elimination via AIC can be seen in the table below:

Table 7

Backward Elimination Via AIC (Round Three KO/TKO)

Step	*Predictor Removed*	*Overall Model AIC*
1	Full Model	89.08
2	X_3 (Ground)	87.09
3	X_4 (Head)	85.15
4	X_5 (Leg)	83.25
5	X_2 (Clinch)	82.53

Now, the overall AIC selected model is significant. The test statistic for the model is (χ^2_1 = 6.02, p = 0.014) while the chi-square critical value, selected at 95% confidence, is 3.84, so we can reject the null hypothesis. We conclude that the model containing proportion of round one and two significant body strikes is significant in predicting the winner of the match.

Once again, we split the data into a training set and a validation set (75/25) to attain the misclassification rate (generalizability error). When we build the model on the training set and then validate it on the test set, we arrive at an accuracy of 50.0%. This output is unfortunately no more accurate than random guessing by overall proportion. This accuracy corresponds to a misclassification rate of 50.0%. The matrix of predicted classes can be seen below.

Table 8

Validation Set Accuracy Table (Round Three KO/TKO)

Winner	*Actually Blue (0)*	*Actually Red (1)*
Predicted Blue (0)	2	1
Predicted Red (1)	7	6

In both the round two knockout model as well as the round three knockout model, one variable logistic regression models are chosen. Because of this, multicollinearity of predictors as a possible problem can be ruled out. Along with multicollinearity, we can also consider the condition of linearity of the predictors against the logit of the outcome. We assess this visually by looking at graphs of predictor values against corresponding logit values. The plots for both round two and three KO/TKOs are displayed below.

Figure 4

Linearity (Round Two KO/TKO)

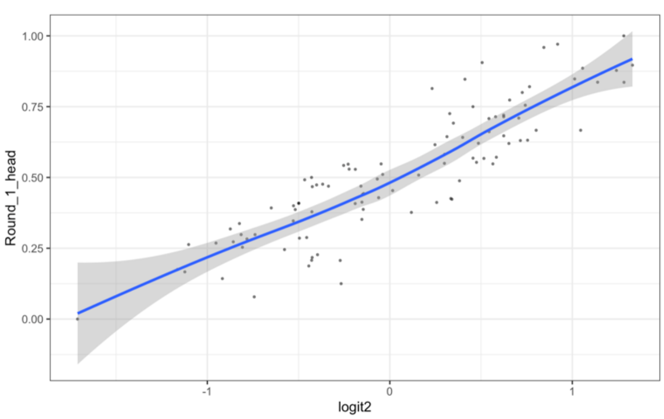
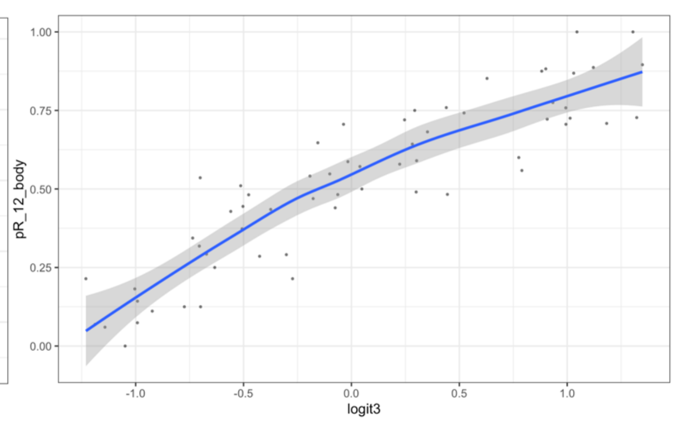


Figure 3

Linearity (Round Three KO/TKO)



As can be seen above, in the case of round two KO, the predictor, round one significant head strike proportion, appears to be linearly related to the logit of the response. Similarly, in the case of round three KO, the predictor, round one and two significant body strike proportion, appears linearly related to the logit of the response. Because these two predictors, the ones used in the final models, are linearly related with the logit of the response, we conclude that the linearity assumption is not violated.

Conclusion

In this article, we study the relationship between various kinds of significant strikes and KO/TKO probability in UFC fights. We first look at the relationship between significant strikes in the first round and the likelihood of KO/TKO in the second round. We then look at aggregate significant strike data from rounds one and two and the likelihood of KO/TKO in the third round.

In both cases, we start off by running a logistic regression on full models involving all of the five respective predictors. In both the second round KO/TKO case as well as the third round KO/TKO case, the models were not significant overall, and most of the individual predictor variables had very large p-values, as well. Because of this, we then refine both models via backward elimination and AIC to arrive at simple logistic regression models. Unlike the full models, these were both significant.

For the second round KO/TKO model, significant head strike percentage is the sole predictor chosen via AIC. This can be interpreted to mean that significant head strikes play the largest role in round two knockouts of the predictors considered. Landing a high percentage of these strikes is more indicative of round two KO/TKO success than any other type of significant strike. For the third round KO/TKO model, significant body strike percentage was the predictor chosen via AIC. This can be interpreted to mean that body strikes in previous rounds have the largest impact on achieving a KO in the last round of the fight. They play a larger role than any other type of significant strike. Both of these results support the idea that there is "chronic weakening" at play in UFC fights.

Despite both being significant, only the model involving round two KO/TKOs has predictive accuracy over 50.0% when utilizing a validation set. The other model, the one involving round three knockouts, has a validation set accuracy rate of 50.0%, so it is no better than random guessing based on proportion. In practice, this means that the model involving round two knockouts can theoretically be used to predict which fighter will achieve a knockout in future fights, based on first round performance in that given fight. Because of its 60.0% accuracy rate, we expect this model to generalize well to new observations. We cannot say the same about the round three KO/TKO case.

The round two knockout model likely performs better in the validation set approach because the predictor variables only involve data from one round. All of the predictors are different types of significant strikes solely from round one. There is less variability in the number of strikes that happen within one round than in two rounds, so it can be expected that this model will perform better than its round three KO/TKO counterpart. Further, when a fighter lands the majority of head strikes early on in a bout, it may be more indicative of an impending KO than if he/she merely has the majority of strikes in the first two rounds. First round strikes may set the stage for the rest of the match more so than those in the second round.

Possible implications for these conclusions are that fighters should attempt to land more significant head strikes in the first round if they desire to increase KO/TKO likelihood in round two. Additionally, fighters should attempt to land more significant body strikes in rounds one and two if they desire to achieve a KO/TKO in round three. Generally, to combine the results from both models, fighters may decide to throw more head and body strikes, as opposed to leg, clinch, and ground strikes, as both of these predictors prove significant in one of the two models.

Along with fighters choosing to throw these specific types of strikes in future fights, the results can also prove meaningful to other people, as well. Coaches at MMA gyms may choose to emphasize the importance of head and body strikes when leading training sessions. Further, sports bettors may now consider these metrics more heavily when choosing a fighter to bet on.

Despite the broad implications of our study, there are also various limitations. First, the sample size was relatively small, with only 100 observations in the round two KO/TKO data and 61 observations in the round three KO/TKO data. This creates problems in generalizability, as the sample may not be completely indicative of the entire population. Further, we narrowed our populations, in each of our models, to fights that specifically ended in KO/TKO during the round of interest. We did this in order to run binary logistic regression. This means that we may not be able to generalize our findings to fights in which knockouts are not achieved in rounds two or three. We choose to generalize anyway but must acknowledge the possible downfalls.

Although this article provides a good introduction to the study of various significant strikes and their implications on round two and round three knockouts, future research should concern itself with running models on larger sets of data, and also including observations in which no KO/TKO is achieved in the round of interest. This would better allow for generalizability to future fights, because in these future instances it is unknown whether a KO/TKO will occur, and therefore it does not make sense to limit the population of a study to only KO/TKO fights. Such a study will involve a response variable with multiple response categories and may employ either logistic regression for more than two classes, or possibly a non-parametric local model similar to K Nearest Neighbors (KNN).

Despite its limitations, we believe that our study provides a good look into the impact of various types of significant strikes on knockout likelihood. It is our hope that this study will be viewed holistically to best understand its results and implications.

References

Trebický, V., Havlícek, J., Roberts, S. C., Little, A. C., & Kleisner, K. (2013). Perceived aggressiveness predicts fighting performance in mixed-martial arts fighters. *Psychological Science, 24*(9), 1664-1672. https://doi.org/10.1177/0956797613477117

Seniuk, H. A., Vu, J. P., & Nosik, M. R. (2019). Application of the matching law to mixed martial arts. *Journal of Applied Behavior Analysis, 53*(2), 846-856. https://doi.org/10.1002/jaba.653

James, L. P., Haff, G. G., Kelly, V. G., & Beckman, E. M. (2016). Towards a determination of the physiological characteristics distinguishing successful mixed martial arts athletes: A systematic review of combat sport literature. *Sports Medicine (Auckland), 46*(10), 1525-1551. https://doi.org/10.1007/s40279-016-0493-1