The Implementation and Application of Statistics in Shotokan Karate

Olivia R. DiDonato[1]

Ryan Savitz[2]

[1]Neumann University, 1 Neumann Drive, Aston, PA. 19014
[2]Neumann University, 1 Neumann Drive, Aston, PA. 19014

Abstract

This paper introduces a model that addresses the way competitors in Shotokan karate martial arts competitions are evaluated. This new model is a combination of multiple variables that determine the winner of a sparring competition in Shotokan karate. The model we constructed predicts the winner of Shotokan karate matches with a near 100% success rate. The model was constructed using the forward LR variety of logistic regression. Due to some issues of multicollinearity, this model was then refined using the author's knowledge of karate. Finally, a single statistic was developed to capture the ability level of a karate participant.

Keywords: Sabermetrics, logistic regression, Shotokan karate

1. Introduction

The purpose of this paper is to use the different elements of a Shotokan karate sparring competition to create an unbiased, mathematical formula to predict the winner of these matches and, similarly, to assess the ability level of participants. The study of sports analytics has been observed and applied in various sports to determine the skill levels of each individual participant; however, little has been attempted in martial arts, especially Shotokan karate. The model created in this paper looks at a large number of karate skills to discern which competitor has a higher chance of winning a match. Additionally, the model looks at this group of variables and distinguishes which ones are of most importance for a competitor to win against their opponent. This paper starts off by discussing the history and backgrounds of Shotokan karate, including the organization and regulations. We then develop a model to predict the winner of Shotokan karate matches. This model is then evaluated, and the implications discussed. Finally, a single statistic is developed that will predict the winner of matches with high accuracy.

2. Literature Review

2.1 History/Origins of Shotokan Karate

Martial arts is a term that has been used to describe "traditional fighting styles" in different parts of the world all throughout history. Thus, it is essentially a very broad range of styles with different origin timelines. For example, the earliest kind of martial art that was taught through Europe was developed in ancient Greece during some of the very first Olympic games. Meanwhile, the earliest time period that historians found evidence of martial arts in India was in "the period of 2nd century BC to 2nd century AD," (Blue Martial arts Book). Eventually, a combination of Indian martial arts and Chinese martial

arts, which originated during the Xia Dynasty in approximately 2,000 BC, became the fundamental basis for the augmentation and advancement of martial arts all over Asia (Blue Martial Arts Book). While it can be very interesting discussing the differing origins of the diverse martial arts styles that have been developed, practiced, and refined over the centuries in various cultures, one style that originated in Okinawa, Japan, and founded by Gichin Funakoshi, that was brought over and introduced to America, is Shotokan karate.

As Funakoshi is famous for stating, "Karate ni sente nashi," or "There is no first attack in karate," Shotokan became a martial art of self-defense, rather than a method of violence against others. The two main parts of Shotokan are *kata* and *kumite*, and both are extremely integrated into the rules of the *dojo kun*, which is the etiquette on how to respect people in Japanese culture and also plays a big part in how martial artists are judged in competitions.

While martial artists have competitions of both their skill levels of kata and kumite, kata is the more artistic side of martial arts. *Kata* is mostly defined as "a set of sequences of karate moves that is organized into various directions as if the martial artist was in a fight with multiple opponents on all sides of them," (Shotokan Karate Club Calgary NW [SKC Calgary NW]). On the other hand, *kumite* more represents how someone could use their skills in a fight between two people in society. Although kata and kumite are split into both "individual" and "team" competitions, this research is mainly focused on how "individual kumite" competitions are scored and which techniques, tactics, etc. that might help a participant win a match.

2.2 Organization of Kumite Divisions

Kumite matches are not only organized by gender (male or female), but are also split into different age groups. The four main sections are Youth (7-17), Adults (18 - 44), Seniors (45 - 54), and Seniors 55+, aka "Super Seniors." For the most part, in regulated competitions, men and women are split into their own groups. However, the Youth division has smaller age groups within this category, and boys and girls are not split apart until about 11 years of age. After this, girls and boys are split apart, while still having the smaller age groups, for example, 12-13, 14-15, and 16-17. This is for safety purposes because even though self-control and discipline are two major aspects students learn during training, especially during kumite practice, kids of different ages have varying levels of self-control and strength. Thus, many senseis and adults do not expect kids to fully grasp how to control their strength against other kids and try to lessen the possibility of someone getting injured. On the other hand, by the time a student reaches 18 years of age, they are considered a young adult and are expected to have as much self-control as any other student in the Adult division. In addition to all of these categories, kumite matches are split by belt level/rank. In the Youth division, participants are divided into three different sections based on their rank level: Beginner (white - orange), Intermediate (green - 2nd degree purple), and Advanced (brown - black). Each section competes in a differing level of kumite that correlates with the rank level of the participants. However, in the Adult, Senior, and Senior 55+ divisions, participants from white - 2nd degree purple belt are mixed together, while the advanced belts (brown-black) are separated into their own mini-division. The differing levels of kumite are "Sambon" (three-step), "Ippon" (one-step), "Jiyū ippon" (semi-free sparring), and "Shobu-Ippon" (free sparring - for black belts only). Sambon, or, three-step, is mainly for martial artists of Beginner rank. Ippon, one-step, is for students of Intermediate rank. Despite brown and black belts being grouped in the same mini-division, Jiyū ippon is for brown belts while Shobu-Ippon is specifically *only* for black belts. This is another small gap seen in the organization of these events due to safety guidelines/precautions, as black belts, for the

most part, are expected to have an extremely high level of self-control and discipline over their strength and movements compared to lower ranks in Shotokan karate.

Below is a chart that represents all of the different elements that are involved to organize a kumite match in competitions.

Factor	Age Group	Gender	Rank Level/Belt Color	Attacks	Defense	Ways of Elimination
	Youth*	Male	Beginners (White, yellow, & orange belts)	Punch	Blocks	Run-out-of-ring
	Adults	Female	Intermediate (Green, purple)	Kicks	Counter-attacks	Use of Excessive Force
	Seniors		Advanced (Brown, black)	Other Strikes		
	Seniors 55+ (AKA Super Seniors)					
	Collegiate**					

These age groups are further split into smaller age groups
Collegiate = A separate event for 18-21 yr old college students only

2.3 Regulations/Rules of a Kumite Match

A kumite match is judged by five judges - one chief judge, called the referee, and four assistant judges, one in each corner of the ring. One match normally lasts approximately two minutes, or when one of the participants scores an *ippon*, which is one full point. However, there are multiple other actions that could cause a match to end prior to the full two-minute duration. If one of the participants receives a full point before the two minutes runs out, then they win and there is no need to further compete. Per contra, a participant can also become eliminated from a match, thus giving their opponent an automatic win. Each participant receives three warnings, which is a combination of being run out-of-bounds (out of the ring lines/past the corner judges) and use of excessive force in prohibited areas of the body. Although regulations allow light to medium contact to the person's midsection area (waist to chest), a participant is flagged and given a warning if they use excessive force during a strike (punch, kick, elbow strike, etc.) to their opponent's head or any area below their waist due to possible severe injury to their physical health. The first warning for out-of-bounds given is called *Jogai Keikoku*. The second warning for out-of-bounds is called *Jogai Chui*, however, this is also considered the first warning for usage of excessive force on an opponent due to how much more severe the impact can be to a person rather than being run out-of-bounds. The third and final warning that leads to elimination to a participant is *Jogai Hansoku*. To clarify the differences between a warning level for out-of-bounds and excessive force is an example situation where a participant gets run out-of bounds and receives the first warning, Jogai Keikoku, then proceeds to use excessive force against their opponent which then leads them to receive the second warning, Jogai Chui. On the other hand, a different example would be if a participant were to use excessive force against their opponent, then they would automatically receive the second warning, Jogai Chui, so if they were to get ran

out-of-bounds afterwards then they would receive the final warning, Jogai Hansoku, and become eliminated.

2.4 Scoring System of a Kumite Match

To win a kumite match, a participant needs to receive an ippon, or one full point, and there are two possible ways to gain an ippon. The first and much more common way that happens in many competitions is where a participant gains two *Waza-Ari*, or half-points, that add up to a total of one full point, or ippon. The second way that is more rare to occur in a kumite match is when a participant gains a full Ippon in one attack against their opponent. The primary target to gain either a Waza-Ari or a full Ippon is the body's midsection area (waist to chest). For the judges to award a participant with any points at all, a participant is supposed to aim to land a strike within a distance of two inches from their opponent's midsection. To gain one Waza-Ari, a participant has to land a strike within the required distance to their opponent's body, but "is executed with some imperfection or is not strong enough for a full point" ("Explanation of Karate Contest."). To gain a full Ippon, a participant needs to have all elements of a strike - perfect posture, correct distance to body, strong focus and extremely good timing. The only other way(s) to gain a full Ippon is if a participant "executes a blow that beats the opponent to the attack, attacks in a series of continuous techniques, delivers a blow after unbalancing the opponent, or strikes when the opponent is defenseless," ("Explanation of Karate Contest."). If someone were to attempt a strike that ends up landing below the belt/waistline but did not use excessive force against their opponent, it will not lead to elimination, they are just considered "off-target" due to being away from the midsection area, thus not receiving any points. If neither participant has gained a full Ippon in total by the end of the two-minute duration of the match, but one of them has one Waza-Ari, then the participant with the one half-point is normally declared the winner by the judges. However, if both participants have not gained any points at all after two minutes or the judges believe that both participants are of equal skill against each other, then they may declare *Hikiwaka* - which means a "draw" - and an additional two-minute overtime match will be held.

2.5 History of Current Scoring System

With the exception of the one-point system and the various regulations previously discussed, there is currently no quantitative system available to determine the skill levels of a martial artist in Shotokan karate. Additionally, there is no kind of mathematically based formula or system that judges use in competitions to determine whether a participant landed a strike against their opponent with enough skill to gain a half-point, full point, or none at all.

2.6 Biases Present in the Population Ratios of Kumite Matches

For a very long time throughout history, the majority of sports, activities, and events were male dominated in society, while many youth and women were unable to participate. Due to this, most of the videos and data available for observation and analysis involve male participants. As a result, the main focal point of observation for this research will specifically be on advanced male adults.

3. Methodology

To determine how to confirm the validity of this model that predicts which competitor has a higher chance of winning, a list of 40 ISKF (International Shotokan Karate Federation) sparring matches was gathered and assigned a number. With this data, a random number generator was used to randomly select 30 matches. Since both competitors need to be analyzed, this gave a larger sample size of 60 participants to individually observe for this model. For each individual participant, fifteen variables were observed: Punches, Kicks, Other Strikes, Blocks, Counter-attacks, Runs-out-of-ring, Usage of Excessive Force, Waza-Ari (½ point), Ippon (1 point), Winner, Unsuccessful Punches, Unsuccessful Kicks, Unsuccessful Other Strikes, Unsuccessful Blocks, and Unsuccessful Counter-Attacks. Additionally, each person's belt rank was recorded and the success percentage of each participant's Punches, Kicks, Other Strikes, Blocks, and Counter-Attacks, however, since every participant in the study were advanced black belts, this variable is not included in the model.

Once all the videos were watched and all the data was collected, the data was entered into the SPSS statistical software program. Due to having a large number of independent variables, we used the Forward LR (likelihood ratio) logistic regression method in SPSS to construct a parsimonious model. This model is used to help determine which variables are the most relevant for the model.

4. Data Analysis

Using SPSS, multiple logistic regressions were executed. In every match that was analyzed, the win/loss was the dependent variable. Before conducting any logistic regressions to create a model, we computed the descriptive statistics (range, mean, and standard deviation) of each variable (including variables not used in the final model), as shown in Table 1.

Descriptive Statistics

	N	Range	Mean	Std. Deviation
Punches	60	2.00	.6833	.79173
Kicks	60	2	.07	.312
Other	60	1	.03	.181
Blocks	60	15	3.43	3.357
CounterAt	60	2	.27	.548
RunoutRing	60	2	.22	.524
ExcessForce	60	0	.00	.000
Unpunch	60	10.00	2.6000	2.58549
Unkick	60	6.00	1.0667	1.42456
Unother	60	3.00	.3167	.67627
Uncounter	60	7.00	1.3000	1.61875
UnBlock	60	2.00	.3333	.65527
PerPunch	53	1.00	.2855	.35560
Perkick	33	1.00	.0480	.18521
PerOther	14	1.00	.1071	.28947
PerBlock	53	1.00	.8543	.30397
PerCounter	40	1.00	.1917	.33022
Valid N (listwise)	9			

Table 1. Descriptive Statistics

For the first logistic regression, the data set contained the eleven core variables: Punches, Kicks, Other Strikes, Blocks, Runs-out-of-ring, Usage of Excessive Force, Counter-Attacks, Unsuccessful Punches, Unsuccessful Kicks, Unsuccessful Other

Strikes, Unsuccessful Blocks, and Unsuccessful Counter-Attacks. For this first attempt, we wanted to see the success rate of the model when including all of the main variables combined. As shown in Tables 1 and 2 below, the first model produced results of a near-100% success rate, with 100% prediction accuracy and a 1.00 for the Nagelkerke R^2.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
4	.000[a]	.747	1.000

Table 1. Nagelkerke R^2

Classification Table[a]

			Predicted		
			Winner		Percentage Correct
	Observed		0	1	
Step 4	Winner	0	33	0	100.0
		1	0	27	100.0
	Overall Percentage				100.0

Table 2. Classification Table

Although Model 1 seemed promising, Table 3 informs us that despite these high accuracy rates, the individual independent variables were statistically insignificant at the 0.05 level. This result appears to be due to the effect of multicollinearity among the independent variables. This obfuscates their individual effects.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B) Lower	Upper
Step 4[a]	Punches	38.333	6239.335	.000	1	.995	4.443E+16	.000	.
	Kicks	76.226	17499.092	.000	1	.997	1.272E+33	.000	.
	Other	38.860	24659.921	.000	1	.999	7.528E+16	.000	.
	CounterAt	38.799	7420.789	.000	1	.996	7.080E+16	.000	.
	Constant	-56.994	9000.107	.000	1	.995	.000		

Table 3. Model of all combined variables

Because of this multicollinearity problem, we attempted several small logistic regression models of each variable individually and numerous combinations of two or more core variables together to determine a model that had both high predictive ability and statistically significant independent variables. In the end, we obtained the model presented in Table 4 below, which uses the independent variables of punches and counter-attacks.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B) Lower	Upper
Step 2[a]	Punches	3.048	.782	15.175	1	.000	21.074	4.547	97.671
	CounterAt	3.457	1.174	8.678	1	.003	31.733	3.181	316.611
	Constant	-3.394	.905	14.076	1	.000	.034		

Table 4. Model with Punches and Counter-Attacks

As can be seen, both punches and counter-attacks are highly statistically significant. The Exp(B) column represents how much more likely a competitor is to win the match if they land a punch or counter-attack. For example, if a competitor lands a punch, their chance of winning a match increases by approximately 21 times. Likewise, if a competitor lands a counter-attack, their chances of winning increase by approximately 31.7 times.

While this gave confirmation that various multicollinearities exist in Model 1 and increased the significance levels of the variables, it slightly lowered the Nagelkerke R^2 from 1.000 down to 0.780 (Table 5). Additionally, the success rate of the prediction accuracy (as seen in the classification table slightly decreased from 100% to a 91.7% (Table 6).

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
2	30.129[a]	.583	.780

Table 5. Nagelkerke R^2

Classification Table[a]

			Predicted		
			Winner		Percentage Correct
	Observed		0	1	
Step 2	Winner	0	32	1	97.0
		1	4	23	85.2
	Overall Percentage				91.7

Table 6. Classification

Our next goal was to come up with a single statistic, based on the variables listed in Table 1, that would be able to successfully predict the outcome of matches. After some exploratory analysis, we ran a logistic regression with the independent variables of punches, counter-attacks, unsuccessful counter-attacks, and unsuccessful Blocks. From this, we computed the following statistic:

$$D = \sum_{i=1}^{4} [Bi / \sum Bi] \mathrm{X_i} ,$$

Where B_i is the logistic regression parameter estimate of the i^{th} variable, and X_i is the value of the i^{th} variable. (We order the variables from 1 through 4 as listed in the paragraph above). The B_i values were 2.415, 2.993, .419, and 1.664, respectively.

In order to evaluate the aforementioned statistic, D, we collected data from 30 additional Shotokan Karate matches. 5 of these matches resulted in draws, so they were excluded from the analysis. Using these data, we found that D was able to successfully predict the outcome of 94% of the matches. Finally, when D was used as the sole independent variable in a logistic regression to predict match outcome, its P-value was found to be 0.001 (highly significant) and the Nagelkerke R^2 for the model was 0.822. This information is presented in Tables 7 and 8 below. Note that if winner = 0, that participant lost the match, while if winner = 1, that participant won.

Variable	**B**	**Standard Error**	**P-value**	**Exp(B)**
D	9.907	2.888	0.001	20,078.469
Constant	-1.161	0.697	0.096	0.313
Nagelkerke R^2	0.822			

Table 7

		Predicted		**% Correct**
		Winner		
Observed		0	1	92.0
Winner	0	23	2	96.0
	1	1	24	94.0

Table 8

5. Discussion and Conclusions

In this paper, we presented two successful models for predicting the outcome of Shotokan Karate matches. The first model utilized the independent variables of punches and counter-attacks, while the second model utilized our newly constructed statistic, D.

Our statistic, D, was validated using a new data set and found to have excellent predictive power and high statistical significance. Based on this, we believe that D has the potential to be used within the Shotokan community for both predicting the outcome of matches and in the process of assigning matches that will produce competitors of similar ability. This second use of our statistic is especially important, as matches of competitors with widely disparate abilities have the potential to be physically dangerous to the competitors.

One future area for research involves employing similar analyses to predict the outcome of matches in other martial arts. A related area for future research involves developing similar statistics in other martial arts that can be used to assign competitors to matches that will be equitable.

References

Evans, B. (1998). *Karate Rules: A Player's Guide*. Ward Lock.

"Explanation of Karate Contest." 53rd Annual ECSKA Championships, 2016.

Funakoshi, G. (1981). *Karate-dō: my way of life*. Kodansha International.

IntroBooks. (2018). *Martial Arts Book*.

Lund, G. J. (2010). *The essential karate book: for white belts, black belts and all levels in between*. Tuttle Publishing.

Shotokan Karate Club Calgary NW. (n.d.). Kata. Shotokan Karate Kata. https://www.shotokankaratecalgary.com/kata.php.

Yuncza, E. (2011). *Sport karate point sparring: an essential guide to the point fighting method*. Indomitable Pub.

35th Annual ISKF/US National Karate Championships, 2013.

38th Annual ISKF/US National & Collegiate Karate Championship, 2016.

Acknowledgements

This research was made possible through the generosity of The Clare Boothe Luce Program of The Henry Luce Foundation.