

## Statistical approach for table tennis athletes' success

Selcuk Goren, Ibrahim Tumay Gulbahar, and Muhammed Safak Pinar

Citation: *AIP Conference Proceedings* **1978**, 410003 (2018); doi: 10.1063/1.5044003

View online: <https://doi.org/10.1063/1.5044003>

View Table of Contents: <http://aip.scitation.org/toc/apc/1978/1>

Published by the *American Institute of Physics*

---

---

**AIP** | Conference Proceedings

Get **30% off** all  
print proceedings!

Enter Promotion Code **PDF30** at checkout



# Statistical Approach for Table Tennis Athletes' Success

Selcuk Goren<sup>a)</sup>, Ibrahim Tumay Gulbahar, Muhammed Safak Pinar

<sup>1</sup> *Engineering Faculty, Industrial Engineering Department, Abdullah Gül University, Kayseri, Turkey.*

<sup>a)</sup> Corresponding author: selcuk.goren@agu.edu.tr

**Abstract.** This report summarizes the statistical modeling and analysis results associated with the athletes' success and athletes' features. Main purpose of this report is to find any relation between athletes' success and their features. As a tool of creating correlation regression is used with SPSS.

## Introduction

Linear regression relates to the creation of models that investigate the linear dependence between variables. Two types of variables are included as the dependent (response) variable and the independent (regressor) variables. The purpose is to describe a linear model that expresses the dependent variables based on independent one. It can be used to estimate the unknown or unobservable values of the response variable based on the values of the regressor variables [1-4]. In this study, linear regression is used as it is defined above. On account of the fact that, there is a dependent variable, which is data column that refers to success points. The others are naturally independent (regressor) variables such as hand used, grip, height, weight, age; and forehand, backhand and blade brands.

For better understanding, Anova table has been used and checked during the study. Anova stands for Analysis of variance. According to K. Molugram and G.S. Rao, Anova is a hypothesis testing technique. It is also a very useful tool to get estimates of various parameters or to get the functions of the parameters in two models. The basic principle lying underneath the technique is that the total variation in dependent variance is known as a multiplication attributed to certain causes, variations between samples and variations between examples attributed to chance [5].

For an athlete training methods are important and it is taking the big portion for table tennis athletes. They generally support themselves with coaches for technical and tactical trainings. Also, to improve performance it is required athlete to decide simultaneously the movement and how should it be [6]. But nowadays, training is not enough for success, athlete or coach need to know other aspects of athlete to improve game performance. In this study, it is tried to show effects of some variances of athletes, such as age, hand used, grip type, height, weight, nationality, and brands used for bat, on success. To achieve this objective, data from table tennis players has been analyzed and visualized, then regression models fitted to understand which of the athletes' feature and their equipment affects their success and secondly the best regression model found.

## Data Description

The data consist of world-class table tennis players, their features and equipment. There are 41 rows and each row consists of 9 different variables. The part of data that consist of bat brands is taken from websites of brands and the rest is from International Table Tennis Federation (ITTF). A part of the data is shown and explained below.

TABLE 1. Sample of Data

Ranking	Name	Nationality	Pts.	Hand used	Grip	Height	Weight	Fore-hand	Back-hand	Blade	Age
1	MA Long	CHN	3376	Right	shakehand	175	72	DHS	DHS	DHS	28
2	FAN Zhendong	CHN	3272	Right	shakehand	173	77	DHS	BTFL	STG	20
3	XU Xin	CHN	3177	Left	Penhold	181	75	STG	STG	STG	27
4	ZHANG Jike	CHN	3046	Right	shakehand	178	73	DHS	BTFL	BTFL	28
5	MIZUTANI Jun	JPN	2982	Left	shakehand	172	68	BTFL	BTFL	BTFL	27
6	OVTCHA ROV Dimitrij	GER	2883	Right	shakehand	186	78	DON IC	BTFL	BTFL	28
7	WONG Chun Ting	HKG	2810	Right	Penhold	170	70	BTFL	BTFL	STG	25
8	JEOUNG Youngsik	KOR	2711	Right	shakehand	180	76	BTFL	BTFL	BTFL	25
8	SAMSON OV Vladimir	BLR	2700	Right	shakehand	189	83	TBR	TBR	TBR	40
10	CHUANG Chih Yuan	TPE	2630	Right	shakehand	169	60	BTFL	BTFL	BTFL	35
11	BOLL Timo	GER	2604	Left	shakehand	181	74	BTFL	BTFL	BTFL	35
12	FAN Bo	CHN	2577	Right	shakehand	170	67	DHS	BTFL	DHS	25
13	YAN An	CHN	2548	Right	shakehand	173	63	DHS	BTFL	BTFL	24
14	JOO Sachyuk	KOR	2546	Right	shakehand	180	68	BTFL	BTFL	BTFL	37
15	FREITAS Marcos	POR	2545	Left	shakehand	182	74	BTFL	BTFL	BTFL	28

Ranking shows world rank of the players. Pts. is points of the players collected. Forehand, backhand and blade show the brands of bat the players used. Other columns are features of players.

## Preprocesses

Data pre-processing is an important step in the model fitting process. In this paper, two main pre-processing techniques are used which are data transformation and data cleaning. In data cleaning part, players with lack of information are deleted. In data transformation part, to use non-numerical data in nationality, hand-used, grip, forehand, backhand and blade columns, they are expressed in a different manner by data transformation.

To use hand-used and grip in regression, for each one, a binary variable is created. To illustrate, for a right-handed user the value 1 is given to the linked variable and 0 otherwise.

For rest of the variables that need to transform, for each observation of each column a new binary variable is created. For example, to express forehand column, 8 variables are created to explain 8 brands. If in a specific observation brand, A is used, the related variable got the value of 1 and 0 otherwise. An example variable set is given below.

**TABLE 2.** Sample of Forehand Brands

DHS	STG	BTFL	DNC	TBR	CRN	JLA	ANDR
1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0
0	0	0	1	0	0	0	0
0	0	1	0	0	0	0	0
0	0	1	0	0	0	0	0
0	0	0	0	1	0	0	0
0	0	1	0	0	0	0	0
0	0	1	0	0	0	0	0
1	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0
0	0	1	0	0	0	0	0

## Model Fitting

In this part of the paper, the techniques applied to fit predictive model will be explained. Regression models are used for the objective of explaining player success. As a measure of success points that players got from licensed tournaments by ITTF, are chosen. As independent variables transformed version of nationality, grip, hand-used, forehand, backhand and blade, and height, weight and age of players are used. First, stepwise regression model is used to learn which of these independent variables have a significant effect on the success rate which is response variable. As a result of this model, it is clarified that nationality is not good predictor variable for this model. Before releasing the nationality variable, a regression model is applied for only this variable without the model to be certain. The results of this analysis are given below.

### Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.627 <sup>a</sup>	.393	-.046	273.43538

a. Predictors: (Constant), CRO, NGR, SWE, AUT, BRA, BLR, FRA, POR, TPE, HKG, KOR, GER, JPN

### ANOVA<sup>a</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	869623.924	13	66894.148	.895	.573 <sup>b</sup>
	Residual	1345804.295	18	74766.905		
	Total	2215428.219	31			

a. Dependent Variable: sco

b. Predictors: (Constant), CRO, NGR, SWE, AUT, BRA, BLR, FRA, POR, TPE, HKG, KOR, GER, JPN

**FIGURE 1.** Result of Regression for Nationality

The significance level, which is the p-value, shows that the null hypothesis could not be rejected. Thus, nationality cannot explain player success. Considering this analysis, nationality is deleted from the model and a new model is fitted. Similarly, as seen from the ANOVA table below, brand columns, which are forehand, backhand and blade are not important for the success of the player.

		ANOVA <sup>a</sup>				
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	329.349	11	29.941	11.244	.016 <sup>b</sup>
	Residual	10.651	4	2.663		
	Total	340.000	15			

a. Dependent Variable: rank

b. Predictors: (Constant), x4bl, x1b, weight, age, xf2, x2bl, x1bl, xf3, x2b, x3bl, height

**FIGURE 2.** ANOVA for Brands

As seen above, adjusted most of the brand columns like x2bl, x2bl, x3bl for blade brand is deleted from the model. For models that divide a variable into more others we require all the components in the model, otherwise it is said that the model is not good to explain the dependent variable. Thus, this model shows that it is not possible to explain athletes' success with brand they use.

From the same ANOVA table above, it is easy to see that other independent variables, which are weight, age and height, are well explainer for rank variable. This means that if we know player feature we can estimate if the player will be succeeded or not.

## Conclusion

This study shows that the feature of nationality and the brand of blade that player use does not explain success of a player well. The other features are good to explain the success of player. What it means by this statement is that, the grip type, hand used, height, weight and rubbers for forehand and backhand affects the success of the athlete positively. Thus, managers can use the explained features and fit the model found in the paper to predict the success ratio of player.

## References

1. A. Rencher , G. Schaalje , Linear Models in Statistics, Wiley, 2008 .
2. D. Montgomery , E. Peck , G. Vining , Introduction to Linear Regression Analysis, 3rd Edition, Wiley Series in Probability and Statistics, Wiley, New York, NY [u.a.], 2001 .
3. N. Draper , H. Smith , Applied Regression Analysis, 2nd Edition, Applied Regression Analysis, Wiley, 1981 .
4. G. Seber , Linear Regression Analysis, Wiley Series in Probability and Statistics, Wiley, 1977 .
5. K. Molugram, G.S. Rao, Statistical Technique for Transportation Engineering, pages 451-462, 2017
6. M. Raab, R.S.W Masters, J.P. Maxwell, Improving The ‘How’ and ‘What’ Decisions Of Elite Table Tennis Players

### References for data taken

(All data belong to February 2017. It may show alteration between brands especially.)

<http://www.ittf.com/>

<http://www.old.ittf.com/>

<http://stigatabletennis.com/en/team-stiga/>

<http://www.cms.donic.de/cms/index.php/en/>

<http://www.joola.de/en/>

<https://www.butterfly.tt/info>

<http://www.cornilleau.com/en/dealer-locator-2/team/>

<http://www.dhs-tt.com/en>

<http://www.andro.de/inside-en-gb/players/overview/>

<http://www.tibhar.com/en/>