Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2014, 6(3):926-930

Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

Research on regression model-based swimmer body shape features affect performance

Qinglin Hong

Sports Department, Xianyang Normal University, Xianyang, China

---

ABSTRACT

Excellent swimmers' competition performance is comprehensive reflection by multiple factors effects, body shape features as one of important factors to predict human body maximum movement potential, is an important guarantee for swimmers getting excellent performance. Exploring excellent swimmers body shape special features not only provides scientific guidance for swimmers high level selection, but also plays remote orientation roles in swimmers' basic selection, it can also makes certain references for swimmers' special physical training. This paper first through consulting documents, finds out swimming performance correlated each body shape indicator 4 large types that accounts for larger weight, by mathematical statistics method, with China swimmers four body shapes and performance through year 2005, 2009 and 2013 investigation as values basis, use MATLAB fitting, it respectively gets four body shape and swimming performance relationships, then use average method to solve weight algorithm, finally it gets four body shapes and swimming performance whole relations. At last, use swimmer Sun Yang swimming performance and his four body shapes data to make verifying, achieved results basically conform to fact, finally gets the conclusion that the model has feasibility.

Key words: Regression model, body shape, numerical fitting, swimmer

---

INTRODUCTION

Due to athletes' figures exist regional differences, China athlete figure compares with other areas athlete, it is generally not ideal. In order to shorten final competition performance gap as much as possible, it regards scientific talent selection as swimming circle important topics. Swimmers' performance is affected by multiple factors; body shape feature is regarded as one of important factors of human body movement potentials, meanwhile is also the guarantee of athletes getting excellent performance. Exploring excellent swimmers' body shape feature and swimming performance relationships, not only provides theoretical basis for selecting high level swimmers, but also plays crucial roles in swimmers basic talent selection [1-3].

Throughout previous researches documents, it is clear that most of researches only targeted for special age grades, special regions objects to carry out research, it hasn't get more involved in swimmers themselves constitutional requests. For example: Bai Xue, Li Ning, Liu Li-Kun published paper " Teenagers swimmers physical qualities training influences on future performance", through investigation, researched on teenagers swimmers' physical qualities, body shapes and functions, research results were periodical physical qualities training with purpose and planning can promote teenagers swimmers sports abilities comprehensive development that helpful for teenagers swimmers sports abilities improvement. And take another example, in Xu Hong-Qi published "China excellent swimmers' body shape simple mode evaluation suggestions"; he researched on swimming overall and each special athlete's body shape simple evaluation model [4-7]. It's mainly investigated on each special swimmers body shapes, values, and carried out simple evaluation.

It is clear that in previous research information, no intuitional research on swimming body shapes and swimming performances, here, this research through consulting documents information, it finds out swimmers body shapes four body shapes that of great important on swimming performance influences are height, foot area, span, shank plus foot height. Establish the four body shapes and performance relationships, so that provides better theoretical basis for athletes' performance.

BODY SHAPE AND PERFORMANCE RELATIONSHIP MATHEMATICAL MODEL

Least square method and unary regression model establishment

At first, according to sample values defining equation $\hat{y} = a + bx$ coefficient a, b. Assume in one experiment, it takes n pairs of data $(x_i, y_i)(i = 1,2,\cdots n)$, the n pairs of data (x_i, y_i) is a sequence of sample values, according to the sequence of sample values, it can look for a pair of coefficient a, b but due to y is a random variable, by another group of experiments, it can also get a pair of a, b values [8, 9]. That is to say, what we get from a group of data is coefficient a, b estimation values, record $\hat{a}, \hat{b}$ solved regression equation by a group of experiment data is:

$$\hat{y} = \hat{a} + \hat{b}x \tag{1}$$

It is called empirical equation of regression. It is also called empirical formula, $\hat{a}, \hat{b}$ as empirical regression coefficient.

In order to solve a, b estimation value, we common used method $\hat{a}, \hat{b}$ is least square method.

Assume in one experiment, it gets n pairs of number (x_i, y_i), from which y_i is random variable y corresponding x_i experiment values. Every experiment value y_i and regression value $\hat{y}_i$ difference $y_i - \hat{y}_i$, it shows as two vertical coordinates differences in figure, the differences have positive value and also negative value, its absolute value is $|y_i - \hat{y}_i|$.Obviously, the straight line we look for should be the straight line that lets all these distance sums be smallest, that $\sum_{i=1}^{n} |y_i - \hat{y}_i|$ is minimum. But due to absolute value is relative complicated in handling; replace it with sum of squares $Q = Q(\hat{a}, \hat{b}) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} \left[y_i - (\hat{a} + \hat{b}x_i)\right]^2$.

The sum of squares Q is changing with regression coefficient $\hat{a}, \hat{b}$ change, therefore it is one binary function of $\hat{a}, \hat{b}$, from which x_i, y_i is constant.

According to binary function solving extreme value method, solves partial derivatives and it gets:

$$\begin{cases} \dfrac{\partial Q}{\partial \hat{a}} = -2\sum_{i=1}^{n} \left[y_i - (\hat{a} + \hat{b}_i)\right], \\ \dfrac{\partial Q}{\partial \hat{b}} = -2\sum_{i=1}^{n} \left[y_i - (\hat{a} + \hat{b}x_i)\right]x_i. \end{cases}$$

Let $\dfrac{\partial Q}{\partial \hat{a}} = 0, \dfrac{\partial Q}{\partial \hat{b}} = 0$, It gets:

$$\begin{cases} \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \hat{a} - \hat{b}\sum_{i=1}^{n} x_i = 0 \\ \sum_{i=1}^{n} x_i y_i - \hat{a}\sum_{i=1}^{n} x_i - \hat{b}\sum_{i=1}^{n} x_i^2 = 0 \end{cases}$$

Solve regression coefficient $\hat{a}, \hat{b}$ as:

$$y\begin{cases} \hat{b} = \dfrac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i^{\ 2} - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2} \\ \hat{a} = \bar{y} - \hat{b}\bar{x} \end{cases} \tag{2}$$

From which: $\bar{x} = \sum_{i=1}^{n} x_i$, $\bar{y} = \frac{1}{n}\sum_{i=2}^{n} y_i$

Formula (3) or written as:

$$\begin{cases} \hat{b} = \dfrac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^{\ 2} - n\bar{x}^2} \\ \hat{a} = \bar{y} - \hat{b}\bar{x} \end{cases} \tag{3}$$

In formula $\hat{a}, \hat{b}$, that are points of Q minimum values, makes $Q = Q(\hat{a}, \hat{b}) = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ arrive at minimum. Linear equation with $\hat{a}, \hat{b}$ as regression coefficient is our required regression equation. It can best represent these points scatter states, due to when solving coefficient $\hat{a}, \hat{b}$, Q let sum of squares to be minimum, the method is least square method.

Curve goodness of fit testing

The merits of curve fitting can be measured by errors sum of squares. Assume that actual tested value is Y, its average value is Y_1, according to fit curve solved theoretical value is Y_2, and then it can get its error sum of square $\sum(Y - Y_2)^2$, mean variance is $\sum(Y - Y_1)^2$, if error sum of squares and mean variance ratios be smaller, then it indicates that actual observation value and estimation value get more closer, curve fitting be better. Therefore, it can define correlation coefficient R^2 , its computation formula is: $R^2 = 1 - \frac{\sum(Y - Y_2)^2}{\sum(Y - Y_1)^2}$

BODY FEATURE AND PERFORMANCE RELATIONSHIP MODEL SOLUTION

At first through consulting document information, it finds out swimming performance related each body shape indicator larger weight four types, by mathematical statistics method, with China swimmers four body shapes and performance through year 2005, 2009 and 2013 investigation as values basis, refer to Table 1.

Table 1: Body shape feature value

Year	Height\cm	Foot area\ cm^2 cm	Span\cm	Shank plus foot height\cm	Average performance
Year 05	172.26	236.76	179.24	42.38	3’59”29
Year 09	181.17	257.67	190.41	46.74	3’49”11
Year 13	184.62	252.63	193.83	48.86	3’47’32

Use MATLAB software fitting, it can get following result:
b = 406.3528 -0.9789
bint =15.0368 797.6689 -2.6658 0.7081
stats = 0.9819 54.3580 0.0858 1.4696

Table 2: R^2 test result table

Item	R^2 test result
Height	0.9819
Foot area	0.8678
Span	0.9925
Shank plus foot height	0.9652

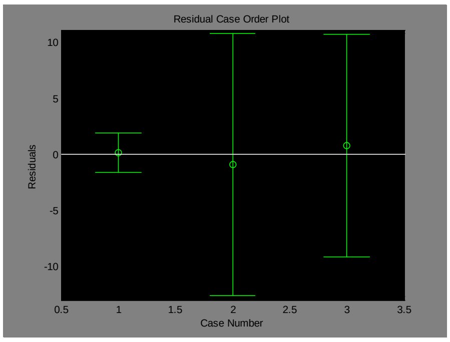

Figure 1: Residual figure

Through Table 2 and Figure 1 result, it indicates that Table 3 all regression model expressions are at work.

Table 3: Body shape feature indicators and swimming performance relations expression

Body shape feature	m	b	Expression
Height	406.3528	-0.9789	$y_1 = 406.3528x_1 - 0.9789$
Foot area	614.1536	-1.5745	$y_2 = 614.1536x_2 - 1.5745$
Span	460.8610	-1.1773	$y_3 = 460.8610x_3 - 1.1773$
Shank plus foot height	162.5832	-0.5027	$y_4 = 162.5832x_4 - 0.5027$

Note:

x_i Four main test items values;

y_i Four main test items corresponding competition performance;

m_i x_i and y_i linear relations coefficient;

a_i Four main test items weight in entirety;

Solve weight with average method, $a_i = \dfrac{m_i}{\sum m_i}$

Handle with coefficient in expression of Table 3 each body shape feature value and swimming performance, it can get $a_i = 0.2472, 0.3736, 0.2803, 0.0989$ $i = 1,2,3,4$

That is getting performance y expression: $y = 0.2472x_1 + 0.3736x_2 + 0.2803x_3 + 0.0989x_4$

EMPIRICAL RESEARCHES

To verify that achieved expression has reliability, use swimmer Sun Yang four body shape feature values and Olympic Games performance in 2012 as data, carry out verifying with final got expression:

Beijing time, July.29th, 2012, Sun Yang in 400 meters free stroke final, broke through Olympic Games record with 3 minutes 40 seconds 14 performance, obtained China swimming team first gold medal with very excellent performance. Meanwhile, he also received first gold medal of Olympic Games in China men swimming team history.

Therefore, this paper carries out verifying and estimation with Sun Yang as an example, Sun Yang body correlation data as Table 4 shows.

Table 4: Sun Yang body data information

Item	Data
Height \ cm	198
Foot area \ cm^2	280
Span \ cm	211
Shank plus foot height \ cm	56

Input Table 4 correlation data correspondingly into height (x_1), foot area (x_2), span (x_3), shank plus foot height (x_4) and the item predicted performance (y) relations: $y = 0.2472x_1 + 0.3736x_2 + 0.2803x_3 + 0.0989x_4$

It gets through calculation $y = 218.235$

Through model prediction performance and actual performance comparing as Table 5:

Table 5: Predicted performance and actual performance comparative table

\	Best result by prediction	Actual best result
400 meters free stroke performance	3Minutes38seconds24	3Minutes40seconds 14
Relative error	0.86%	

The model is mainly established with China swimmers as research objects, the research can apply into national team swimmers selection and excellent swimmers best results prediction, so that it has better basis for planning more perfect training, diet and to other schemes.

CONCLUSION

Through fitting, it got each feature indicator and average performance relationships, used average method solving weights, it got each factor weight on average performance, which could get performance y expression, after that utilized master swimmer Sun Yang empirical verification such expression, the result was basically conforming to actual that indicated the modeling process was at work and reliable. The top advantage in the paper lied in its bold calculation and guess, exploring for problems without previous researching that started the trend.

There are promotion existing on research results, it mainly four body shape feature indicators from them is defined according to previous investigation China swimmers physical qualities feature data, only represents China athletes physiques, it can collect some foreign athletes' body shape feature values, expanding the research to international level, which produces effective promotion for improving international swimmers' performance.

Acknowledgment
Scientific research project in Shaanxi province department of education: (2013 jk0524)

REFERENCES

[1] ZHANG Ming-fei, CHEN Yan, ZHANG Yi-hua. *China Sport Science and Technology*, **2006**, 42(5):85-88.
[2] SUN Bing, XU Hongqi. *Journal of Tianjin Institute of Physical Education*, **2008**, 23(5):433-436.
[3] Li Yi-yu, Yang Fang and He Jiang-chuan. *Neural Regeneration Research*, **2009**, 13(37):7381-7384.
[4] Tang Qingwen, Chen Li. *Sichuan Journal of Anatomy*, **1997**, 5(2):93-96.
[5] JI Gang. *Journal of Guangzhou Physical Education Institute*, **2009**, 29(4).
[6] LUO Zhi. *Journal of Chengdu Physical Education Institute*, **2005**, 31(5):93-97.
[7] WANG Ju-tao, MIAO Miao. *Journal of Chengdu Physical Education Institute*, **2012**, 38(4):72-76.
[8] Zhang B.; Zhang S.; Lu G.. *Journal of Chemical and Pharmaceutical Research*, **2013**, 5(9), 256-262.
[9] Bing Zhang. *Journal of Chemical and Pharmaceutical Research*, **2014**, 5(2), 649-659.