ISSN 1750-9823 (print)
International Journal of Sports Science and Engineering
Vol. 01 (2007) No. 03, pp. 177-182

Gray Models for the Prediction of Results in Track and Field Events *

Yong Jiang [1], Jianliang Li [1], Hongguang Bao [2, +], Jiefu Wang [1], Fengyou Hua [1]

[1] Department of Mathematics, Nanjing University of Science and Technology, Nanjng, 210094, P. R. China
[2] Nanjing First Hospital, Nanjing University of Medicine, Nanjng, 210006, P. R. China

(*Received July 28 2007, accepted August 2 2007*)

Abstract. Considering the result record of a sport event regarding years as a time serial, it can be described by a model of time serial including certain model and uncertain model. In this paper, we created the gray model GM(1, 1) as differential equation for track and field events based on gray system theory. Based on historical record data of Olympic Games, parameters of the equation are solved by using the method of least squares. The validation is down. And the comparison with other methods shows that the gray model it is accepted. It can be used to predict result of track and field events of 2008 Bejing Olympic Games, etc.

Key Words: Gray Models, Prediction, Track and Field, Olympic Games, Mathematical Modelling.

1. Introduction

Formerly mainly used the time series analysis method to predict tendency of the Track and Field Events results which along with the year development. And directly uses the Track and Field Events results time series $Y_0(1)$, …, $Y_0(N)$ as data. Then chooses a mathematical model which has specific expression form to carryon the fitting, and obtains the approximate model which reflects track and field games results change rules. Using the model, obtains the forecast value.

However, regarding Prediction of Results in Track and Field Events, because the influence of track and field games results is extremely various, some known or clear, some unknown, inaccurate or is fuzzy. Therefore, even if thought the historical statistics in the results of track and field games has contained all-around information of exterior factor influence and action, also has the change rule insufficiently accurate and insufficient obvious and so on. So unavoidably has a bigger error. In fact, the results of track and field games sometimes also affected by social factors, such as the significant technological breakthrough, the appearance of remarkable persons, entrant country resists, contingency etc. and other accidental factors. We regard the results of track and field games and each kind known, unknown influence factors as a system, using the cybernetics gray system, and union the computational mathematics method to carry on the analysis model. This article take the gray system as a foundation, proposed the gray model method to predict the results in track and field events (or called gray time series analysis method). This method does not directly use the movement result sequence modeling, but uses AGO sequence modeling.

2. Gray model method (GM)

The gray system obtains success in the research of grain prediction, the talented person prediction and so on. The gray system believed, the objective system regardless of how complex, it all has the connection, has the rule, has the overall function. One of its main characteristics is not directly uses the primitive series modeling. Because the primitive data series' rule is disordered or insufficiently obvious and insufficiently accurate. So we process the primitive data series (namely production), make it more orderly or more obviously and accurate, then use the production to modeling. This article applied this theory and the characteristic to propose sports results prediction gray model method.

* This paper is presented at the International Conference on Mathematical Modelling in Sport, Manchester, June, 2007
+ Correspondence author, *E-mail address*: math@mail.njust.edu.cn

2.1. Accumulated generating operation (AGO)

First, we process the results of track and field games by AGO. Its goal is: Provides the middle information for the modeling, and make randomness of originally movement result random series to attenuate.

Definition 1. Let Y_0 be the original series, i.e.,

$$Y_0: Y_0(1), \ldots, Y_0(i),\ldots$$

Let Y_1 be the generating series, i.e,

$$Y_1: Y_1(1), \ldots, Y_1(i),\ldots$$

Where $Y_1(i)$ satisfies

$$Y_1(i) = \sum_{i=1}^{i} Y_0(i) \tag{1}$$

Definition 2 If we regard AGO series as some continual variable which correspond to the time serial number of separate series, then called this continual variable for the AGO value, and denoted by $Y^1(t)$.

Obviously, the AGO series as well as abstractly produced AGO value all is strict monotonic increase. And its randomness was attenuated. This obvious and accurate rule brings very big convenience for us to determine the model.

2.2. Determinate gray models type

The general gray model is the n rand, h variable, denoted by GM(n, h). Regarding the single item sports result, the corresponding AGO value is the single variable, thus takes h = 1. Furthermore, considering the AGO value is monotonic increasing, the sports result tendency basic stability and the actual computation work load, it is not difficult for us to ascertain n=1.So we took GM(1,1) as the model of sports results development law. And it is a definition differential equation (refer to reference 4):

$$\frac{dY_1(t)}{dt} - KY_1(t) = L \tag{2}$$

Or denoted by a equivalent type

$$\frac{\frac{dY_1(t)}{dt}}{Y_1(t)} = K + \frac{L}{Y_1(t)} \tag{3}$$

K, *L* are two parameters which decided by the single item sports result series. Considering (3) , it is obvious that K can be regard as the AGO value's relative " growth" rate.

Use $\overline{Y}(1) = Y_0(1)$ as the initial condition, from (2) or (3) we can get $Y(t)$:

$$\overline{Y_1}(t) = (Y_0(1) + \frac{L}{K})e^{K(t-1)} - \frac{L}{K} \tag{4}$$

2.3. Single item result gray model

Supposes $Y_0(i)(i = 1,2,..., N)$ is some sports result series (known). So directly from equation (2), we can use the computational mathematics method to obtain the normal equation about *K*, *L*. thus obtains *K*, *L*'s least square solution

$$\begin{pmatrix} \overline{K} \\ \overline{L} \end{pmatrix} = (B^T B)^{-1} B^T X \tag{5}$$

where

$$B=\begin{bmatrix} \frac{1}{2}[Y_1(1)+Y_1(2)] & 1 \\ \frac{1}{2}[Y_1(2)+Y_1(3)] & 1 \\ \cdots & \\ \frac{1}{2}[Y_1(N-1)+Y_1(N)] & 1 \end{bmatrix} \tag{6}$$

and

$$X=[Y_0(2),Y_0(3),...,Y_0(N)]^T$$

$$Y_1(i)=Y_1(i-1)+Y_0(i) \quad i = 2, 3, \ldots, N$$

When we obtain the answer of $\overline{K},\overline{L}$, the solution of (4) is:

$$\overline{Y_1}(i+1) = (Y_0(1)+\frac{\overline{L}}{\overline{K}})e^{\overline{K}(i-1)}-\frac{\overline{L}}{\overline{K}} \qquad \text{i=1,2,…} \tag{7}$$

2.4. Recover and predict

By (7) the formula, according to defination 1, $Y_0(i)$ can be recovered (using inverse accumulated generating operation). So we can obtain the sports result's model value which correspond to the time series:

$$\overline{Y_0}(i+1) = (Y_0(1)+\frac{\overline{L}}{\overline{K}})(e^{\overline{K}(i+1)}-e^{K(i-1)}) \qquad i = 1,2,\ldots$$

When $i \geq N$, $\overline{Y_0}(i+1)$ is the forecast value.

In fact, (7), (8) is a unification pattern to general sports. When practical application we can only according to the known single item sports result sequence and direct use (5), (6) to obtain the parameters $\overline{K},\overline{L}$ to that sports. Then obtain (7), (8) two types gray models of sports result, and use (8) formula to calculate sports result series' model value and forecast value.

2.5. Model examination

To the sports results gray model, we present the following the method of post-validation (refer to reference Wang, X., et al, (2001)).

Let $q(i)$ be the error (actual value subtract model value), i. e

$$q(i)=\frac{Y_0(i)-\overline{Y_0}(i)}{Y_0(i)} \qquad i = 1, \ldots, N$$

And let $\overline{q}$ be the mean of $q(i)$, let $\overline{Y}$ be the mean of $Y_0(i)$. Then after examination error and small errors frequency are

$$C=\frac{S_1}{S_0}=\frac{\frac{1}{N}\sum_{i=1}^{N}[q(i)-q]^2}{\frac{1}{N}\sum_{i=1}^{N}[Y_0(i)-\overline{Y}]^2}$$

$$P=P\{|q(i)-\overline{q}|<0.6745S_0\},$$

So we can examine the gray model of sports result according to table 1. And determined reliability if we use this model to predict the sports result.

Table 1

Precision scale	P	C
Good (first scale)	>0.95	<0.35
Preferably (second scale)	>0.80	<0.50
Qualification (third scale)	>0.70	<0.65
Can (fourth scale)	>0.7(or ≤ 0.70)	≥ 0.65 (or <0.65)
Failed	≤ 0.70	≥ 0.65

Table 2 Athens Olympic Games 'prediction' and error $|q|$

Project (men's)	Actual result	Predict result	error
100 m	9.85	9.85	0.03%
200 m	19.79	19.76	0.17%
400 m	44.00	43.56	1.01%
800 m	01:44.45	01:43.38	1.03%
1500 m	03:34.18	03:35.02	0.39%
10000 m	27:05.10	26:54.25	0.67%
110 m hurdle	12.91	12.92	0.10%
400m hurdle	47.63	47.13	1.06%
3000m handicap	08:05.81	08:02.05	0.77%
long jump	8.59	8.57	0.19%
high jump	2.36	2.42	2.33%
hop step and jump	17.79	18.05	1.44%
Shot	21.16	21.95	3.71%
discus	69.89	69.53	0.52%
hammer	82.91	84.18	1.53%
4×100 relay race	38.07	37.48	1.55%
4×400 relay race	02:55.91	02:55.76	0.09%

Table 3. Some methods comparison

Method / comparison		This text	Linear regression	Square regression	Cube regression	Log regression
$\sum q_t^2$		0.003078	0.004652	0.005518	0.0133	0.002032
$\frac{1}{17}\sum\|q_t\|$		0.98%	1.25%	1.23%	1.97%	0.84%
Error higher than this text's number and percent			13/76.47%	11/64.71%	10/58.82%	7/41.18%
The max error		3.71%	4.18%	5.00%	7.13%	3.26%
Error higher than	2%	11.76%	17.65%	41.18%	41.18%	5.88%
	1%	47.06%	47.06%	58.82%	58.82%	29.41%
	0.50%	64.71%	64.71%	70.59%	70.59%	82.35%

Where

$$q_t = \frac{actualvalue - \mathrm{mod}elvalue}{actualvalue}.$$

Table 4. 2008 Beijing Olympic Games prediction

Project (men's)	$\overline{Y_0}(i+1) = (Y_0(1) + \frac{\overline{L}}{\overline{K}})(e^{\overline{k}i} - e^{\overline{k}(i-1)})$		Precision prediction index		Precision scale	2008 prediction
	K	L	P	C		
100 m	-0.00292258	10.159	0.82	0.003228	second	9.82
200 m	-0.00148226	20.079	1	0.000712	first	19.74
400 m	-0.00165527	44.475	1	0.000064	first	43.64
800 m	-0.00083999	104.66	1	0.000020	first	01:43.66
1500 m	-0.00125126	217.57	1	0.000007	first	03:34.46
10000 m	-0.00587706	1720.8	1	0.000000	first	26:48.43
110 m hurdle	-0.0037958	13.445	1	0.000304	first	12.87
400m hurdle	-0.00171589	48.162	1	0.000070	first	47.22
3000m handicap	-0.00540451	511.56	1	0.000001	first	08:00.73
long jump	0.00049700	8.5347	1	0.004105	first	8.58
high jump	0.0066878	2.2335	0.45	0.010223	fourth	2.41
hop step and jump	0.00445054	17.137	1	0.000384	first	18.04
Shot	0.00320877	20.956	1	0.000636	first	21.74
discus	0.00788544	64.147	1	0.000025	first	70.21
hammer	0.01084233	74.775	1	0.000015	first	84.64
4×100 relay race	-0.00175951	38.385	1	0.000077	first	37.61
4×400 relay race	-0.00196629	179.48	1	0.000008	first	02:55.46

3. 2004 Olympic Games results 'prediction' and compare examination

According to 1964-2000 Olympic Games primitive data, we use GM method to predict 2004 Olympic Games champion's results. Then we compare GM method with other methods (table 3). The conclusion indicates that the GM method is better than other methods.

4. 2008 Beijing Olympic Games prediction and the error test of post-validation

Finally, according to this article's method, we establish sports result development model of partial single item sports, then carried on the error test of post-validation. And has carried on prediction to 2,008 Beijing Olympic Games' partial single items result. Furthermore, if we considered other more information, the predict accuracy will be higher.

5. Conclusion

This article only carried on discussion on gray model prediction of single item sports result. In fact, the gray model method of this article may be used to predict other questions in the sports fields. Such as funds prediction, the demand of sports teacher in college, high school, elementary school in the long run and so on.

6. Reference

[1] Q. Ma, etal. Sports result prediction mathematics method discussion. *Sports Science*. 1987, **7(**2).

[2] Y. Jiang, etal. Sports result index return prediction, *The second conference of national sports science*. 1987,12.

[3] J. Deng. *Fuzzy mathematics*. 1982, **5**(2).

[4] X. Wang, etal. *Gray system analysis and practical computational procedure*. Wuhan: Huazhong university of

science & technology publishing company. 2001, 9.

[5] K. Feng, etal. *Value computational method*. Bejing: Science publishing company. 1978.

[6] Y. Jiang, etal. Gray Models for the Prediction of Results in Track and Field Events, *Sports statistics*. 1987,12.