

Creditability Evaluation of the Undergraduate's Physical Score

Hua HE

Physical Education Department, Hunan Business College, Changsha, China, 410205

hehua_312@hotmail.com

Keywords: Creditability evaluation, Evaluation indexes, Fuzzy integration, Physical score.

Abstract. Aiming at creditability performance evaluation for physical score of undergraduate students, a general idea is developed. The evaluation indexes and their measurement method are investigated. Then, with the application of fuzzy integration, an opening performance evaluation model is implemented. This model can offer quantitative conclusion in objective ways. At the same time, the operating condition is taken into account. In the end, performance evaluation experiment is performed with field test data.

Introduction

In the field of science and technology, it is an important task to calculate the creditability for the obtained data. The results may demonstrate the objectivity and reliability of these data on revealing the corresponding physical quantity [1]. This is also true in physical education [2].

When the Department of Education is making physical education policies for the universities, the most important basis is the physical score of the students. It shows the students' physical performance and the education achievements [3-4]. However, the physical score sometimes can not reveal the physical built as expected as possible [4]. The reasons are: (1) the test measures may work worse in practical environment [5]; (2) the affection from the testers and the testee, either subjective or objective [6]; (3) the limitation of the technical level and the operating condition in practical [7]. Therefore, the result can not offer a good indication of the students' performance in all scenarios. The worst thing is that the testers do not know the creditability of these results [8-10]. As a result, performance evaluation for the physical score's creditability in undergraduate is of great importance [11, 12].

This work concerns about the problem "evaluation the creditability for undergraduate students' physical score". With the application of probability and statistics, the evaluation indexes can be deduced; consequently, the evaluation model is constructed with the use of fuzzy mathematics. Evaluation test can be organized with the use of these models.

In organizing this paper, the problem and its background are analyzed firstly. The key ideas of this work are explained. These are the main contents in the "Introduction" Section. In the "Problem Formulation" Section, the general idea is discussed. Then, a series of evaluation indexes and their measurement method are developed in the "The performance Evaluation indexes" Section. Consequently, with the application of fuzzy integration, an opening performance evaluation model is formed in the "Performance Evaluation Model Based on Fuzzy Integration" Section. This model can offer quantitative object conclusion. At the same time, the operating condition is taken into account. To confirm the practicability of this work, experiments are performed in the "Experiments" Section. In the end, a brief summary is provided.

Problem Formulation

The creditability evaluation of the undergraduate's physical score is a work that reveals the difference between the physical score and the practical physical performance of a single/group of college student(s) [13, 14]. It (a) analyses the test data and the test condition, (b) calculates the evaluation indexes, (c) works with the evaluation model, (d) tries to offer some information about the precision, robustness and independence of the physical score.

There are three types of work should be finished before the concrete study: analyzing the test data and the corresponding test condition, consulting experts for suggestions; investigating performance evaluation indexes.

The problems and their relation are shown in Fig. 1.

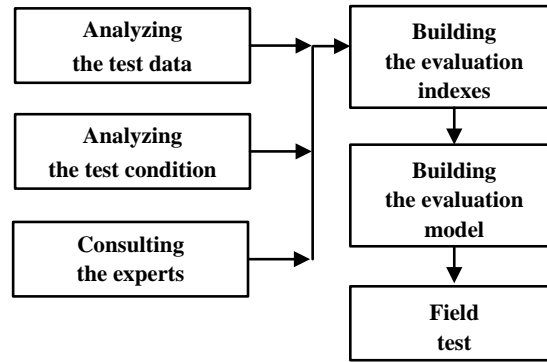


Figure 1. The problems and their relation in creditability evaluation of the undergraduate's physical score

The Performance Evaluation Indexes

The Framework of The Performance Evaluation Indexes

There are some quantities that must be considered: a group of them shows the precision of the physical score; the robustness of the physical score should be interpreted by another group of indexes; and, the independency of the physical score from the test condition should be described by another more group of indexes. The framework of the performance evaluation indexes is shown in Table 1.

Table 1. The framework of the performance evaluation indexes

Type of evaluation indexes	How to measure them
Precision	Measured by the quantities derive from the mean hypothesis test and coincidence hypothesis test for the physical score
Robustness	Measured by the quantities derive from the distribution hypothesis test and variation hypothesis test for the physical score
Independence	Measured by the quantities derive from the independence hypothesis test for the physical score

The evaluation indexes on precision of the physical score

The precision evaluation indexes of the physical score shows the accordance between the test results and the built-in performance. The construction course of the precision evaluation indexes of the physical score is illustrated by the deduction of the Creditability Precision of Vital Capacity Index to Weight (CVCIW).

Step 1, collecting samples. For the college students being tested, we collect the samples of Vital Capacity Index to Weight (VCIW) in a given test condition. Suppose that we have collected n samples, where n is no less than the confidence coefficient requirements (all the following sample size fulfills this requirement thereafter). These samples are recorded as $VCIW_i^{Cum}$, $i = 1 \sim n$. Here, the superscript "Cum" means "cumulative." Then, we collect the samples of VCIW in the same test condition. Suppose that we have collected m samples. These samples are recorded as $VCIW_i^{Test}$, $i = 1 \sim m$. Here, the superscript "Test" means "testing."

Step 2, hypothesis tests. Firstly, we check whether the two series, $VCIW_i^{Test}$ ($i = 1 \sim m$) and $VCIW_j^{Cum}$ ($j = 1 \sim n$) are from the same population by the rank sum test. If the rank sum test fails, then $CVCIW=0$. The literature on rank sum test is extensive [15, 16]. So, the detailed materials are not presented here.

Secondly, if the rank sum test passed, it means they are from the same population. Let

$$\mu^{\text{Cum}} = \frac{1}{n} \sum_{i=1}^n VCIW_i^{\text{Cum}} \quad (1)$$

and

$$\mu^{\text{Test}} = \frac{1}{m} \sum_{i=1}^m VCIW_i^{\text{Test}} \quad (2)$$

denote the means of these two series. We carry out the two-sided test

$$H_0: \mu^{\text{Cum}} = \mu^{\text{Test}}; H_1: \mu^{\text{Cum}} \neq \mu^{\text{Test}}. \quad (3)$$

Here, the significance level α is determined by the tester. CVCIW is determined by the following function

$$CVCIW = \begin{cases} \frac{\min(\mu^{\text{Cum}}, \mu^{\text{Test}})}{\max(\mu^{\text{Cum}}, \mu^{\text{Test}})}, & \text{the two-sided test passed;} \\ \frac{\min(\mu^{\text{Cum}}, \mu^{\text{Test}})}{\max(\mu^{\text{Cum}}, \mu^{\text{Test}})} \alpha, & \text{the two-sided test failed;} \\ 0 & \text{the rank sum test failed.} \end{cases} \quad (4)$$

The evaluation indexes on robustness of the physical score

The physical score's reliability in revealing the students' physical ability is measured by the robustness evaluation indexes of the physical score. They are examined by the distribution characteristic and the variation characteristic. According to the knowledge of statistic, the physical score samples subject to normal distribution when the sample size is no less than necessary. So, whether these results are normal distribution shows the robustness of the physical score. The calculation of Distribution Index of CVCIW (DICVCIW) is used to illustrate the way of deducing robustness evaluation indexes. $I_{\text{CVCIW}}^{\text{Dis}}$ is the robustness evaluation index of CVCIW. Suppose that the threshold of the distribution hypothesis test is λ , the statistic of test in the testing course is w . Then,

$$I_{\text{CVCIW}}^{\text{Dis}} = \begin{cases} 0, & \text{the distribution hypothesis test failed;} \\ (1-\alpha) \frac{w}{\lambda}, & \text{the distribution hypothesis test passed.} \end{cases} \quad (5)$$

The variation of CVCIW can be employed to measure the confidence interval of the physical score. It is named Variance Index of CVCIW (VICVCIW), as is denoted by $I_{\text{CVCIW}}^{\text{Vari}}$. It can be calculated directly from the samples.

The evaluation indexes on independence of the physical score

When the test condition is varying in a given interval, the creditability of the physical score may change accordingly. This characteristic of the physical score is revealed by the independence evaluation indexes. By such kind of evaluation indexes, we have the detailed knowledge of the creditability responding to the changing test condition.

Similarly, the independence of CVCIW is presented as an example. The threshold of the independence hypothesis test is denoted as η . The statistic of test is ρ . Then, the independence evaluation indexes of the physical score

$$I_{CVCIW}^{Ind} = \begin{cases} 0, & \text{the independence hypothesis test failed;} \\ (1-\alpha)\frac{\rho}{\eta}, & \text{the independence hypothesis test passed.} \end{cases} \quad (6)$$

The index system for the creditability of the physical score

Taking the aforementioned materials as an example, the index system for the creditability of the physical score is built.

(a) Evaluation indexes related to precision. creditability of ewight for height score (CESH), creditability of Harvard step score (CHSS), creditability of 50 m run score, creditability precision of vital capacity index to weight (CVCIW), creditability of standing long jump score (CSLJS), creditability of sit-and-reach (CSR), creditability of score of sit-up (CSSU), creditability of grip capacity index to weight (CGCIW).

(b) Evaluation indexes related to robustness. Robust Index of CESH, Robust Index of CHSS, Robust Index of CVCIW, Robust Index of CSLJS, Robust Index of CSR, Robust Index of CSSU, Robust Index of CVCIW, Robust Index of CGCIW.

(c) Evaluation indexes related to independence of test condition. Independence Index of CESH, Independence Index of CHSS, Independence Index of CVCIW, Independence Index of CSLJS, Independence Index of CSR, Independence Index of CSSU, Independence Index of CVCIW, Independence Index of CGCIW.

People can choose some of these indexes in their evaluation work according to the scenario. They can also add more index(es) to this system.

Performance Evaluation Model Based on Fuzzy Integration

Based on the above-mentioned work, evaluation model can be investigated. Here, the performance evaluation model is built based on fuzzy integration [17]. In order to maintain a concise presentation, only three factors are chosen in this model. People can add more factors according to their work in practice.

(a) The factors set. The factors set is $\{CVCIW, I_{CVCIW}^{Dis}, I_{CVCIW}^{Ind}\}$, that is, $U = \{u_i, i = 1 \sim 3\}$.

(b) The weight set. We give a vector on U , $A = \{a_i, i = 1 \sim 3\}$, as the weight set for the factors. They meet $\sum_{i=1}^3 a_i = 1$.

(c) The judgment vector by a single student. Suppose that the excellent level of each factor is no less than μ_{1m} , μ_{2m} and μ_{3m} , respectively.

Further more, suppose that there is only a student to be tested, and the record of her/him are μ_1^{test} , μ_2^{test} and μ_3^{test} , respectively. The judgment vector here is named $R_1 = (r_i^1), i=1\sim 3$. Then,

$$R_1 = \left(\frac{\mu_1^{test}}{\mu_{1m}}, \frac{\mu_2^{test}}{\mu_{2m}}, \frac{\mu_3^{test}}{\mu_{3m}} \right) \quad (7)$$

Specially, if the record in any dimension is equal to or more than the excellent level, it is set to 1.

(d) Performance evaluation result from the judgment vector by a single student. The performance evaluation result from the judgment vector by a single student is named B_1 . According to the fuzzy integration rules [17],

$$B_1 = \int_{\sim U} R_1 \circ \Gamma(\cdot) = R_1 \circ A = \bigvee_{i=1}^3 (r_i^1 \wedge a_i) \quad (8)$$

$\Gamma(\cdot)$ is the probability measurement function on U . $\Gamma(\cdot)$ equals to the weight vector here. R_1 is a measure-able function on U . R_1 is set as the judgment vector here.

(e) Performance evaluation result from the judgment vector by a group of students. According to the rules for the judgment vector by a single student, the judgment vector is arrived at for all the students with identical operating condition. Suppose that there are t students, the judgment vector for j -th ($j(1 \leq j \leq t)$) student is $R_j = (r_i^j), i=1 \sim 3$. With regard to the fuzzy integration rules [17], the average judgment vector is attained first, $R = (r_i), i=1 \sim 3$. Here,

$$r_i = \frac{1}{t} \sum_{j=1}^t r_i^j, i=1 \sim 3. \quad (9)$$

The final results is named B ,

$$B = \int_{\sim U} R \circ \Gamma(\cdot) = R \circ A = \bigvee_{i=1}^3 (r_i \wedge a_i) \quad (10)$$

Experiments

Experiments with the proposed evaluation methods are performed to ensure that it can be used in the practice. 100 students in Grade 2008 from a college served as the subjects. They are divided into 6 groups according to gender and the physical score of the college admission test. There are Group M1~M3 for boy student and Group F1~F3 for girl student.

(a) The parameters in the evaluation model. According to the experience, the “excellent” values of the factors are set as in Table 2. Here, the name of the factor is represented by its sequence number to save space.

Table 2. The threshold of “excellent” for the factors

	CVCIW	I_{CVCIW}^{Dis}	I_{CVCIW}^{Ind}
Threshold	≥ 0.85	≥ 0.90	≥ 0.90

(b) The weight set. The weight set is set according to the importance of each factor. It is set as [0.82 0.81 0.75] here.

(c) Evaluation results. The evaluation results are given in Table 3,

Table 3. The evaluation results

	M1	M2	M3	F1	F2	F3
Grade of excellent	0.96	0.92	0.93	0.90	0.94	0.92

To have an in-depth view of the creditability of the physical score of the undergraduate students, we train the testers with different investment. This affect factor is called “investment.” We give the points according to train time and train fees. The sample size here is the sample number for each student on a given test condition. Fig. 2 is the performance evaluation results with respect to investment and sample size; Fig. 3 is the gradient and the contour of the evaluation function.

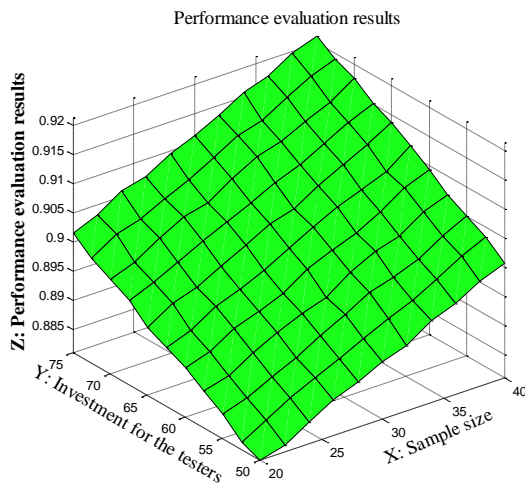


Figure 2. Performance evaluation results with respect to investment and sample size

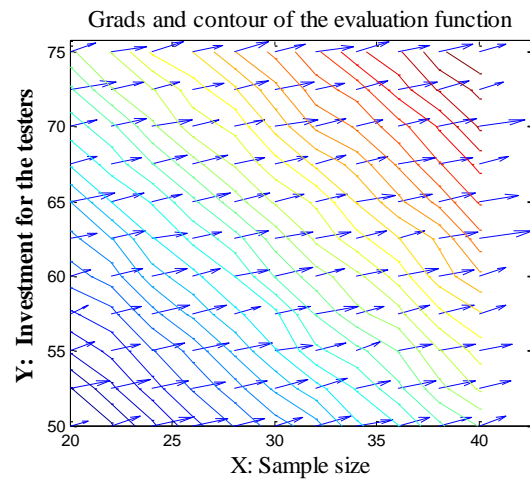


Figure 3 The gradient and the contour of the evaluation function

We can see by these results that, the tester's capability has some affection on the evaluation results. If the sample size is no less than requirement by the significance level, then, the affection is slight.

Summary

In order to evaluate the creditability of the physical score of the undergraduate students, a novel evaluation frame is developed. The evaluation index system is given with detailed measurement methods. With the application of fuzzy integration, an opening performance evaluation model is formed.

There are some capabilities of these evaluation methods: (a) to calculate the creditability of the students' physical score; (b) to offer suggestions on how to perfect physical education in colleges; (c) to calibrate the insufficient points of the physical situation of a college student.

The evaluation model here is an opening model. People can modify the factor set and the weight set according to their work. This paper only offers an opening of this work. Much more efforts are needed to perfect it.

Acknowledgement

The author wishes to extend her sincere thanks to Dr. Yanpeng Li for his good suggestions on the grammar of the manuscript.

References

- [1] Y.L. Li, The reformation of the physical education in Chinese Colleges, J. of the Beijing Sport Univ. (in Chinese), 25(June, 2002) 798-799.
- [2] S.H. Tan, C. Liu, R.Q. Jiang, The analysis of the problems in reformation of the physical education, J. of the Chengdu Sport Univ. (in Chinese), 21 (April, 2005) 69-70.
- [3] Y.L. Li, Z.Z. He, Y.Q. Jin, Z.W. Gong, Study of the performance evaluation of physical education, J. of Sport Science (in Chinese), 22 (May, 2002) 56-61.
- [4] Q. Chen, The evaluation of the physical performance of the students, Trans. of Sport Sci. (in Chinese), 9 (June, 2002) 12-16.
- [5] H.Y. Xiao. The compare between physical education in China and Japan, J. of the Shanghai Sport Univ. (in Chinese), 23 (Jan., 1999) 72-77.
- [6] L.D. Hansley. The introduction of test standards for American children, J. of Sport Sci. in Zhejiang (in Chinese), 20 (Feb., 1998) 58-60.

- [7] K. Stronks, J.P. Mackenbach. Evaluating the Effect of Policies and Interventions to Address Inequalities in Health: Lessons from a Dutch Program, *European J. of Public Health*, 16 (April, 2005) 346-353.
- [8] K. Backett-Milburn, S. Wilson. Understanding Peer Education: Insights from a Process Evaluation, *Health Education Research*, 23 (Jan., 2008) 85-96.
- [9] D.L. Stanistreet, N.G. Bruce, I. Taylor. Developing Distance Learning in Postgraduate Public Health Education: Promoting Quality in Design and Delivery of Postgraduate Public Health Training through Open Learning, *European J. of Public Health*, 16 (April, 2007) 296-300.
- [10] M. Petticrew, M. Whitehead, S.J. Macintyre. Evidence for Public Health Policy on Inequalities: The Reality According to Policy Makers, *J. of Epidemiol Community Health*, 58 (Sept., 2007) 811–816.
- [11] G.Y. Qi, J.J. Zhang, Z.K. Guo. The foundation and the content of the Standards of Physical Built, *J. of the Beijing Sport Univ.* (in Chinese), 30 (Feb., 2007) 226-228.
- [12] G.Y. Qi. The sampling methods and the data processing measurements of the Standards of Physical Built, *J. of the Beijing Sport Univ.* (in Chinese), 29 (Sept., 2006) 1218-1228.
- [13] S. Silverman. Teaching for student learning in physical education. *J. of Physical Education Recreation and Dance*, 82 (June, 2011) 29-34.
- [14] D.R. Young, A. Steckler, S. Cohen. Process Evaluation Results from a School- and Community-linked Intervention. *Health Education Research*, 23(June, 2008) 976-986.
- [15] E. Suhir, *Applied Probability for Engineers and Scientists*, Mcgraw-Hill, New York, 1997.
- [16] A. Leon-Garcia, *Probability, Statistics, and Random Processes for Electrical Engineering*, third ed., Pearson Education Inc., New Jersey, 2009.
- [17] D. Dubois, H.M. Prade, ed., *Fundamentals of Fuzzy Sets (The Handbooks of Fuzzy Sets Series 7)*, Springer, New York, 2000.