http://ansinet.com/itj

ITJ

ISSN 1812-5638

# INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (15): 3325-3329, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.3325.3329


# Quantitative Study and Analysis of Physical Education Teaching Based on Origin

XingZhi Li
Institute of Physical Education, Hainan Normal University, HaiKou 571158,China

**Abstract:** By means of Origin, this study extracts randomly four curriculums of students and conducts descriptive analysis, correlation analysis and regression analysis. The purpose is to better understand the quantitative data analysis in physical education. Besides, cognize and perfect the teaching program from such perspective. Finally, according to the main result, we propose several pieces of advice. We make some complement and extension for physical education area.

**Key words:** Origin, multi-variable linear return analysis, physical education teaching, correlation analysis

## INTRODUCTION

Nowadays, the mode of physical education teaching is based on competition, which is bad for students' personal advantages. It considers rate of reaching standards and scores as main factors of sports abilities. The examination is also based on competition. It has bad effects on students' activations. Now, many students can't find their love sports items on physical education curriculums. Besides, there are huge gaps between real curriculums and their anticipations. Many students hold that the curriculums are just for fun (Long and Xia, 2013). The reason for that is not only due to students. Most of curriculums are in old modes, which can reach up the times. The evaluation methods are also based on competition. The contents are mainly the competitive level and ability of student. It is difficult to promote the total effects of exercises and cultivations in self-exercising ability (Zhang, 2013).

The investigation shows, in the number of college students who exercise for more than 3 times, 35% are female and 60% are male. And the time for exercising is short. Besides, 26% female and 20% male never exercise. 41% students spend less than 30 min on exercising once. Only 40% students can exercise for more than an hour. Therefore, the times for students' active exercises are few. It is relevant to curriculum arrangements (Wang *et al*., 2013).

Nowadays, most papers discuss the existing problems qualitatively and explore the teaching reforms (Li *et al*., 2008). Origin is standard data analysis software, with great power and convenient operations. And it still not used in physical education teaching. Therefore, this paper, by use of Origin, correlation and regressive analysis, studies quantitatively problems in curriculum settings and final examinations.

## ORIGIN SOFTWARE

Origin is developed by American enterprise OriginLab. It has strong functions and convenient operation methods. It is accepted by international science press as standard data analysis and mapping software. It features as briefness, great power and flexibility. The two main functions are data analysis and mapping. The data analysis includes: Giving statistic averages, standard deviation, standard error, sum and groups of data N for data selected; data sort, regulation, calculation, statistics, spectrum transforming (Liu, 2012); linear, polynomial and multiple fitting; fast FFT, correlation analysis, FFT filtering, peak searching and fitting; about 200 internal and self-definition function models for curve fitting and the fitting programs can be controlled; statistic, mathematic and calculus. While analyzing data, you can just choose the data and select corresponding orders (Lin, 2010). The mapping is based on template. There are tens of 2-dimensional and 3-dimensional mapping templates. Choosing the data and clicking relative tool buttons can complete the drawing (Tan, 2013). To 2-dimensional graphics, you can set page, axis, label, symbol and color and many styles of lines. There are more than 100 internal symbols. Regulate data label. Select styles of axis, scale and axis display. Select different signals. There are more than 50 XY axes in each page, which can be exported as varied graphic files or copied to clipboard as objects. The self-definite mathematic functions, graphic styles and plotting templates can connect varied office software, database software and graphic processing software, etc; it can conduct matrix operation such as transpose, inverse, etc conveniently and export 3-dimensional graphics through matrix window; it can be programmed by C language and Lab Talk.

## STUDY METHOD

To collect the data more comprehensively, this study, based on real situations, extracts study objects from college students in different majors. There are 30 students, of whom 20 people are male and the others are female. Besides, collect their performances of four classes: <College English> <basketball>, <comprehensive English> <volleyball>, <practical workplace English> <Ping-pong> and <advanced English> <badminton>.

This paper studies the data in Origin 7.0.

This study uses quantitative research method. It is generally used to get total statistical results. It mainly includes statistical analysis, description and building metering model for analyzing and predicating.

**Data analysis:** Firstly, this study input data (Table 1) to Origin. The analysis methods mainly include:

- **Descriptive analysis:** Arrange, collect and describe the data, such as cost volume, range, etc
- **Correlation analysis:** Including discrimination analysis for test questions and reliability degree analysis
- **Regressive analysis:** Mainly using multi-variable linear return analysis

## MAIN RESULTS

**Descriptive analysis:** Firstly, solve cost volume, range and average, etc., secondly, input the performances of four classes to Origin. Conduct column statistics and get Table 2.

The Table 2 shows, the standard deviation of <volleyball> is 4.54, which is minimum. And its average is 84, which is highest. So, its test questions are simple.

Finally, conduct normal distribution analysis in Origin, shown as Table 3.

The Table 3 shows, all performances of four curriculums are close to normal distribution.

**Correlation analysis:** Correlation analysis is to study the correlation between random variants, mainly by correlation coefficients. To variants x and y, their sampling values are $x_i$ and $y_i$ and their correlation coefficient is:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}\sqrt{(y_i - \bar{y})^2}}} \in [-1, 1]$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i,\ \bar{y}) = \frac{1}{n}\sum_{i=1}^{n} y_i$$

The abs of $r_{xy}$ closer to 1 indicates their correlation is closer; the abs of $r_{xy}$ closer to 0 indicates their correlation is not closer.

The analysis includes: discrimination analysis and reliability analysis of test questions, Discrimination analysis of test questions.

Discrimination degree is one of indexes for checking test questions, mainly indicated by flair. The higher degree can better divide the levels of students, so the bigger value of test questions is. But the higher discrimination degree doesn't mean more difficult. Neither too hard nor too easy test questions can separate the groups in different levels. Therefore, the questions with middle level of difficulty have best discrimination degree.

The general methods include two ways:

$$D = PH - PL$$

$$D = \frac{XH - XL}{N\ (H - L)}$$

Due to same way, this study only analyzes the discrimination degree of a question in "badminton", shown as Table 4.

Table 1: Performances of 30 students

| No. | Basketball | Volleyball | Pingpong | Badminton | No. | Basketball | Volleyball | Pingpong | Badminton |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 69 | 83 | 69 | 71 | 16 | 71 | 75 | 62 | 90 |
| 2 | 66 | 84 | 61 | 82 | 17 | 63 | 87 | 77 | 88 |
| 3 | 76 | 85 | 70 | 81 | 18 | 70 | 80 | 72 | 85 |
| 4 | 72 | 90 | 80 | 90 | 19 | 60 | 79 | 77 | 77 |
| 5 | 73 | 84 | 74 | 68 | 20 | 69 | 82 | 66 | 85 |
| 6 | 79 | 85 | 68 | 65 | 21 | 72 | 87 | 84 | 87 |
| 7 | 65 | 93 | 81 | 88 | 22 | 76 | 86 | 77 | 76 |
| 8 | 74 | 84 | 77 | 85 | 23 | 75 | 78 | 69 | 87 |
| 9 | 74 | 85 | 70 | 81 | 24 | 71 | 87 | 73 | 70 |
| 10 | 70 | 84 | 70 | 75 | 25 | 68 | 78 | 72 | 90 |
| 11 | 75 | 80 | 75 | 88 | 26 | 75 | 87 | 70 | 83 |
| 12 | 68 | 88 | 67 | 76 | 27 | 60 | 89 | 76 | 88 |
| 13 | 71 | 89 | 77 | 92 | 28 | 69 | 82 | 71 | 76 |
| 14 | 68 | 85 | 80 | 77 | 29 | 77 | 80 | 71 | 81 |
| 15 | 61 | 74 | 85 | 89 | 30 | 64 | 90 | 52 | 71 |

$$\frac{dS(\beta)}{d\beta} = 0$$

The result from Origin shows the discrimination degree of <badminton> is too high, reaching 0.75392.

**Reliability analysis of test questions:** Reliability refers to reliability of test results, which is consistency degree or validity degree. Therefore, the test questions with high reliability can be considered as criterion for students' performances. Reliability degree include test-retest reliability degree, alternate-form reliability degree, split-half reliability degree, Kuder-Richardson reliability degree, Cronbach' alpha and rater reliability degree. This study uses internal consistence to measure reliability, which is split-half reliability. The equation is:

$$r_{xy} = \frac{n.r_{x_1x_2}}{1+(n-1)r_{x_1x_2}}$$

where, $r_{x1,x2}$ Indicates the scores of two half questions and correlation coefficient, n is times of previous questions to later one n = 2.

Firstly, classify the questions into even and odd ones. Secondly, calculate scores of two parts. Finally solve reliability coefficient through correlation coefficients. Then such correlation coefficient is internal consistence coefficient. Take the reliability degree of <basketball> as example, shown as Table 5.

The reliability degree of <basketball> in Origin is not much, only 0.34587. Same as such, the reliability degree of other curriculums can be got, shown as Table 6.

The Table above shows that all the reliability degrees are not high. The lowest one is that of volleyball, only 0.31002; the highest one is that of badminton, reaching 0.57013 but not optimistic.

**Regressive analysis:** There is close relation between correlation analysis and regressive analysis. They have common study objects and supplement each other. Correlation analysis is basic and premise for regressive analysis. Regressive analysis is deeply exploring for correlation analysis. This study mainly uses multi-variable linear return analysis. So-called multi-variable linear return analysis is to analyze the linear relation between a variable and several independent variables, the general form is:

$$Y = \beta_0+\beta_{1x1}+\beta_{2x2}+\cdots+\beta_{mxm}+\varepsilon$$

Table 2: Descriptive data analysis

| Curriculum | Cost volume | Average | Range | Standard deviation | Lowest score | Highest score |
|---|---|---|---|---|---|---|
| Basketball | 30 | 70.03 | 19 | 5.09 | 79 | 60 |
| Volleyball | 30 | 84.00 | 19 | 4.54 | 93 | 74 |
| Pingpong | 30 | 72.43 | 33 | 6.96 | 85 | 52 |
| Badminton | 30 | 81.40 | 27 | 7.50 | 92 | 65 |

Table 3: normal distribution analysis

| Curriculum | N | W | p-value | Decision |
|---|---|---|---|---|
| Basketball | 30 | 0.95867 | 0.32715 | Normal at 0.05 level |
| Volleyball | 30 | 0.97482 | 0.71430 | Normal at 0.05 level |
| Pingpong | 30 | 0.95960 | 0.34452 | Normal at 0.05 level |
| Badminton | 30 | 0.92802 | 0.05031 | Normal at 0.05 level |

Table 4: discrimination degree

| No. | Total performance | Score | No. | Total performance | Score |
|---|---|---|---|---|---|
| 1 | 71 | 18 | 16 | 90 | 11 |
| 2 | 82 | 14 | 17 | 88 | 15 |
| 3 | 81 | 15 | 18 | 85 | 16 |
| 4 | 90 | 16 | 19 | 77 | 12 |
| 5 | 68 | 13 | 20 | 85 | 13 |
| 6 | 65 | 14 | 21 | 87 | 16 |
| 7 | 88 | 14 | 22 | 76 | 14 |
| 8 | 85 | 13 | 23 | 87 | 13 |
| 9 | 81 | 16 | 24 | 70 | 16 |
| 10 | 75 | 14 | 25 | 90 | 17 |
| 11 | 88 | 14 | 26 | 83 | 14 |
| 12 | 76 | 15 | 27 | 88 | 17 |
| 13 | 92 | 14 | 28 | 76 | 16 |
| 14 | 77 | 13 | 29 | 81 | 14 |
| 15 | 89 | 17 | 30 | 71 | 15 |

Table 5: reliability degree of basketball

| No. | Basketball | Scores for odd questions | Scores for even questions | No. | Basketball | Scores for odd questions | Scores for even questions |
|---|---|---|---|---|---|---|---|
| 1 | 69 | 39 | 30 | 16 | 71 | 45 | 26 |
| 2 | 66 | 36 | 30 | 17 | 63 | 33 | 30 |
| 3 | 76 | 40 | 36 | 18 | 70 | 40 | 30 |
| 4 | 72 | 39 | 35 | 19 | 60 | 35 | 25 |
| 5 | 73 | 37 | 36 | 20 | 69 | 39 | 30 |
| 6 | 79 | 49 | 30 | 21 | 72 | 40 | 32 |
| 7 | 65 | 29 | 36 | 22 | 76 | 38 | 38 |
| 8 | 74 | 40 | 34 | 23 | 75 | 38 | 37 |
| 9 | 74 | 44 | 30 | 24 | 71 | 42 | 29 |
| 10 | 70 | 35 | 35 | 25 | 68 | 38 | 30 |
| 11 | 75 | 35 | 40 | 26 | 75 | 42 | 33 |
| 12 | 68 | 38 | 30 | 27 | 60 | 36 | 24 |
| 13 | 71 | 40 | 31 | 28 | 69 | 35 | 34 |
| 14 | 68 | 28 | 40 | 29 | 77 | 45 | 32 |
| 15 | 61 | 35 | 36 | 30 | 64 | 29 | 35 |

Table 6: Evaluation data

| Curriculum | Volleyball | Pingpong | Badminton |
|---|---|---|---|
| Reliability degree | 0.31002 | 0.47532 | 0.57013 |

where, $\beta_0$ Is intercept, $\beta_1, \beta_2,\ldots,\beta_m$ is regressive coefficient, $\epsilon$ is residual.

Conduct n independent observations on y and $x_1, x_2,\ldots, x_m$. Get n groups of sampling data $y_i$, $x_{i1},\ldots, x_{im1}$ $(i = 1, 2,\ldots, n)$, then:

$$\begin{cases} y_1 = \beta_0 + \beta_1 x_{11} + \cdots + \beta_{m-1} x_{1m-1} + \varepsilon_1 \\ y_2 = \beta_0 + \beta_1 x_{21} + \cdots + \beta_{m-1} x_{2m-1} + \varepsilon_2 \\ \vdots \\ y_n = \beta_0 + \beta_1 x_{n1} + \cdots + \beta_{m-1} x_{nm-1} + \varepsilon_n \end{cases} \quad (1)$$

of which $\epsilon_1, \epsilon_2,\ldots, \epsilon_n$ are mutually independent and conforming to N $(0, \sigma^2)$.

Set:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{m-1} \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}, x = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1m-1} \\ 1 & x_{21} & x_{22} & \cdots & x_{2m-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nm-1} \end{bmatrix}$$

then Eq. 1 can be demonstrated in matrix as:

$$Y = x\beta+\epsilon$$

$$\epsilon \sim N(0, \sigma^2 I_n)$$

The least squares estimation for model $\beta$ is to choose $\beta$ to make error $\hat{\beta}$ to be the point estimation for $\beta$.

$$S(\beta) = \epsilon^T\epsilon = (y-c\beta)^T(y-c\beta) \quad (2)$$

To solve $\beta$, calculate:

$$\frac{dS(\beta)}{d\beta} = 0$$

due to Eq. 2. Then solve $\hat{\beta}$, which is $\hat{\beta} = (x^T x)^{-1}(x^T y)$. To residual vector $\epsilon$:

$$\hat{\varepsilon} = y - \hat{y} = y - x\hat{\beta}$$

Then the residual sum of squares is:

$$E(\hat{\varepsilon}^T\hat{\varepsilon}) = \sigma^2(n-p)$$

$$\sigma^2 = \frac{1}{n-p}(\hat{\varepsilon}^T\hat{\varepsilon})$$

After building the mathematical model for multiple linear returns, we still need mathematical statistics to test fitting degree, obviousness of linear relations, etc. The general ways are R test and F test.

R test:

$$R = \sqrt{1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}}$$

$\bar{y}_i$ is the average of $y_i$. R is called complex correlation coefficient, mainly used for testing fitting goodness of regressive model. The bigger R means the linear relation is more obvious. Its value is $0<|R|\le 1$.

F test:

$$F = \frac{U/m}{Q/(n-m-1)} \sim F(m, n-m-1)$$

Of which:

$$Q = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2,\ U = \sum_{i=1}^{n}(\hat{y}_i - \bar{y}_i)^2$$

m is the number of independent variables, n is the number of data. F conforms to distribution F (m, n-m-1), set the obvious standard as $\alpha$. $F>F_\alpha(1, n\text{-}m\text{-}1)$. Means the regressive model is obvious. Only the obvious one can be used to predict.

Get the estimation of model factors according to sampling data and the relation between variable and independent variables can be solved, which is the multi-variable linear return equation:

$$Y = b_0+b_{1x1}+b_{2x2}+\cdots+b_{mxm}$$

Next, consider <badminton> as dependent variable and the others as independent variables. Find their correlation. By use of Origin, the multi-variable linear return equation can be solved:

$$Y = 40.=30975-0.10001x_1+0.379x_2+0.0963x_3$$

Due to the equation, the relation between badminton and the others can be seen: if we know performances of any three curriculums, we can estimate the performance of another.

## CONCLUSION

This study, using Origin, extracts randomly four curriculums of students, conducts descriptive analysis, correlation analysis and regression analysis and get following results: the performances of all the curriculum are approach nominal distribution and the indexes like value volume, range, etc are in the reasonable area; the discrimination degrees of some questions are too high. Take <badminton > as an example, some are bigger than 0.7; the reliability degrees of all curriculums are not high, of which the one of badminton is highest but only 0.57013; the correlation among curriculums are small.

The significance of this paper is to provide teaching advice. All in all, this paper proposes several advices: to curriculum setting, it should consider the correlation between curriculums in high and low grades. Hold the discrimination degree of questions, which should be regulated above 0.3.However high discrimination degree doesn't mean complexity. Hold the reliability degree of questions.

## REFERENCES


Li, L.P., S.S. Zeng and Z.H. Liu, 2008. Re-thinking on current reforms of physical education teaching for college and universities. J. Shenyang Instit. Phys. Educ., 4: 89-92.

Lin, B., 2010. Analysis of multi-variable linear return and its application. Basis Frontier Res., 9: 60-61.

Liu, J.X., 2012. Application of origin in physical experiment teaching in colleges and universities. Value Eng., 19: 224-225.

Long, W. and X.C. Xia, 2013. Application of Origin 7.5 in teaching and education. J. Zhangjiakou Vocational Tech. Coll., 2: 63-66.

Tan, Z., 2013. Fuzzy data envelopment analysis and neural network evaluation mathematical applications model which based on martial arts competition. Int. J. Applied Math. Stat., 44: 37-44.

Wang, H. and S. Yang, 2013. An analysis of hurdle performance prediction based on mechanical analysis and gray prediction model. Int. J. Applied Math. Stat., 39: 243-250.

Zhang, B., 2013. Dynamics mathematical model and prediction of long jump athletes in Olympics. Int. J. Applied Math. Stat., 44: 422-430.