Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2014, 6(5):1416-1419

Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

Philosophy of Science and Technology Education Based on Linear Regression Theory

Baoyu Ma

Xinxiang University, Henan, China

ABSTRACT

A valid and efficient method to teach Philosophy of science and technology courses in linear regression analysis is presented in this article. The method includes theory and practice parts, where interactive learning methodologies are created. It adopts case-study teaching, since this teaching method effectively integrates theoretical teaching and practical teaching. The lectures should be not an exhaustive review of regression methodology, but they should focus on how the regression models derived. Moreover, the teacher should pay more attention to the theoretical aspects of models rather than to their implementation using software. Students work in teams of three or four on a problem presented by teachers and choose relevant software to carry out their own projects. Feedback from students indicates that this method of teaching improves students' class attendance and greatly increases their interest in learning.

Keywords: *linear regression*； *Philosophy of science and technology*； *theoretical models*

INTRODUCTION

To study Philosophy of science and technologycurriculum theory, we need to first clarify the meaning of the theory of ideological and political education course which also needs to interpret one of the first core word "ideological and political education" and "curriculum theory". We know that the Philosophy of science and technologycourseseducation is a comprehensive and complex concept, its connotation can smaller of bigger. Broad ideological and political education is almost equivalent to the thought education, political education, moral education, discipline education and shape people's ideological consciousness, such as political ideas, moral character and spirit of citizenship which is related to the comprehensive education category. And narrow sense of Philosophy of science and technology education is a proper noun, and led by the proletariat political party, in order to cultivate qualified builders and reliable successors of revolution and construction the education goal of closely linked together the ideology of class consciousness and the political thought education. The concept of curriculum theory is often seen as a education subject manages research area and branch, in China for nearly a century of research history. So far, education science field discussion of curriculum and curriculum theory connotation is still on. But according to one of the most common explanation, course can be interpreted as a subject, teaching material and teaching related expression, can realize the goal of school education ,curriculum theory is of course the proprietary theory, is the school discipline and curriculum system and content structure.

Reality requires one to deal with a range of variables. Students can encounter many variables for a problem, including relationships between the variables [1].Commonly, these problems can be divided into two kinds according to their relationship. One is the problem of determining, which relationship can be expressed using the functions available. Another is not completely determined by using functions, but with random variables. This kind of relationship is called co-relationship. Regression analysis is a method to study the relationship between variables. So one builds regression models to help understand and explain the relationships. Regression models can also be used to predict actual outcomes.

The earliest form of regression was the method of least squares, which was published by Legendre in 1805 and by Gauss in 1809. Legendre and Gauss both applied the method to problems of determining. The term regression was coined by Francis Galton in the 19th Century to describe a biological phenomenon. However, Francis Galton applied the method to the problem of random variability [2].The phenomenon was that the heights of descendants of tall ancestors tend to regress down towards a normal average, and this is, known as regression toward the mean.

Today, regression methods continue to be an area of active research, and new methods have been developed for robust regression, regression involving correlated responses, such as time series and growth curves, regression in which the predictor or response variables are curves, graphs, images or other complex data objects, etc. There are many examples. Feilong and Yubo made an error analysis for the linear programming support vector regression problem in learning theory. Jose M. proposed a machine learning algorithm for regression problems. Qing, H. et al presented parallel extreme learning machine for regression based on MapReduce[3] .

Chen Y. proposed assessing mathematics learning achievement using hybrid rough set classifiers and multiple regression analysis. González-Recio and Forni proposed Genome-wide prediction of discrete traits using Bayesian regressions and machine learning. Tsuruokaetused logistic regression to study learning string similarity measures. Therefore, almost all institutions in which relevant majors have been set offer a course in regression. For example, MIT offers open courseware. In addition, a large number of open on-line teaching resources can also be found, such as applied regression analysis at Columbia Business School for the MBA major.

In 2010, a statistics major was established in the Mathematics and Computer School at Xinxiang University, Xinxiang, China [4].In this year, the School first offered the course for students whose major came from computer science or mathematics. It is compulsory for mathematics students, while it is an elective for computer science students.

LINEAR REGRESSION MODEL

The theory-oriented lectures cover single linear regression and multiple linear regressions. Students learn what regression is, how to create its mode, how to estimate the parameters of the model (Estimation Using Least Squares), understanding the assumptions of establishing the conditions for the model, what the regression coefficients are, how to compare the models, and predicting and controlling using regression model. Teachers start theirs lectures with a discussion of simple regression, then, move on to multiple linear regression. This is quite reasonable from a pedagogical point of view, since simple regression has the great advantage of being easy to understand graphically.

Students should place a lot of emphasis on the simple linear regression analysis and understanding its mathematical expressions and be open to more sophisticated concepts. It is difficult for students to study multiple linear regression analysis. However, it is a primary tool in the analysis of real data. Thereby, single linear regression is taught in six sessions, while multiple linear regression requires four sessions. A session is 90 minutes' duration.

Single linear model describes a linear relationship between two variables. One is called the target, response or dependent variable, and is usually represented. Another is called the predicting or independent variables, and is usually represented by x . Given (x1 , x2 , □, xN) , the simple linear r

$$C_{\min} = \sum_{i=1}^{N} \left\{ k_i + C_i X_i + \frac{1}{2}[Y_i(j+1)] \cdot h_i \right\}.$$

$$X_i + Y_i(j) - Y(j+1) - \sum_{i \in R} (1+\delta) q_i = 0,$$

where the data, xi , yi , represent a random sample from a larger population, which consist of n set of observations, the coefficients are unknown parameters, and i are random error or disturbance terms.

FEEDBACK EXERCISES

In the spring semester of the 2012/2013 term, a total of 60 junior students came from two majors, , joined the course at the Xinxiang University, Xinxiang, China. These students already had joined some previous courses, such as mathematical analysis, advanced algebra, probability and statistics, etc [5]. They also demonstrated a certain level of operating computer software capability. A survey questionnaire with six feedback questions was administered to all of these students (shown in Table 1).

Tab.1: Feedback questions

Questions		Majors		Mean
		Computerscience (students30)	Mathematics (students30)	
1	I becameinterestedindataanalysisthroughthisclass.	27 (90%)	28 (93%)	92%
2	I understandtheconceptoflinearregression.	24 (80%)	26 (87%)	83%
3	I performedthelinearregressionusingMATLABprogram.	27 (90%)	26 (87%)	88%
4	I performedthelinearregressionusingMSExcel.	29 (97%)	29 (97%)	97%
5	I thinkthepracticeis usefultounderstandtheconceptof linearregression.	30 (100%)	30 (100%)	100%
6	I canapplymyknowledgeof mathematicsbyperforming theexperiment.	24 (80%)	25 (83%)	82%

There are too many formulations about regression analysis. It is often difficult to understand these concepts of regression analysis, said some students. As can be seen from student feedback, only 80% of computer science students can understand regression, while 87% of mathematics students can do so. However, almost students finished their experiments, and all of them think the practice is very useful to understand those concepts. This information shows that students' interest will be enthused if theory is combined with practice, and as long as theory explanation is not ignored.

Its objective is to strengthen their statistical skills. One of the authors received and accepted the teaching load. Hence, sharing the teaching experience is the objective of this article.

THEORY TEACHING

The theory-oriented lectures cover single linear regression and multiple linear regressions. Students learn what regression is, how to create its mode, how to estimate the parameters of the model (Estimation Using Least Squares), understanding the assumptions of establishing the conditions for the model, what the regression coefficients are, how to compare the models, and predicting and controlling using regression model. Teachers start theirs lectures with a discussion of simple regression, then, move on to multiple linear regression. This is quite reasonable from a pedagogical point of view, since simple regression has the great advantage of being easy to understand graphically.

Students should place a lot of emphasis on the simple linear regression analysis and understanding its mathematical expressions and be open to more sophisticated concepts. It is difficult for students to study multiple linear regression analysis. However, it is a primary tool in the analysis of real data. Thereby, single linear regression is taught in six sessions, while multiple linear regression requires four sessions. A session is 90 minutes' duration.

SINGLE LINEAR REGRESSION MODEL

Single linear model describes a linear relationship between two variables. One is called the target, response or dependent variable, and is usually represented by y. Another is called the predicting or independent variables, and is usually represented by x. Given $(x_1, x_2 \ldots x_N)$, the simple linear regression model is described as:

$$\begin{cases} y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \\ \varepsilon_i \sim N(0, \sigma^2) \; and \; \operatorname{cov}(\varepsilon_i, \varepsilon_j) = 0 \; when \; i \neq j \end{cases} \quad (1)$$

where the data $\{x_i, y_i\}$, represent a random sample from a larger population, which consist of n set of observations, the β coefficients are unknown parameters, and β_i are random error or disturbance terms.

LEAST SQUARE ESTIMATION

A primary goal of a regression analysis is to estimate the relationship between the predictor and the target variables or equivalently, to estimate the unknown parameter β. This requires a data-based rule or criterion that will give a reasonable estimate. The standard approach is least squares regression which is a convex optimisation problem with no constraints. The objective is a sum of squares of terms of the form that are chosen to minimise:

$$\sum_{i=1}^{n} [y_i - (\beta_0 + \beta_1 x_i)]^2 \quad (2)$$

The scatter diagram gives a graphical representation of least squares, which can help students to understand regression graphically. If the fitted regression equation has been obtained, it is a line given by:

$$\hat{E}(y) = \hat{\beta}_0 + \hat{\beta}_1 x \quad (3)$$

Residuals are defined as the difference between the observed value and the fitted value $\hat{y}_i$. Equation (2) minimises the sum of squares of the residuals if the coefficients β take as the fitted coefficient $\hat{\beta}$. By minimising Equation (2), the regression coefficients are obtained by:

$$\begin{cases} \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \\ \hat{\beta}_1 = S_{xy} / S_{xx} \end{cases} \quad (4)$$

where $\bar{x} = \frac{1}{n}\sum x_i, \bar{y} = \frac{1}{n}\sum y_i, S_{xx} = \sum (x_i - \bar{x})^2, S_{XY} = \sum (x_i - \bar{x})(y_i - \bar{y})$

HYPOTHESIS TESTS FOR THE SIMPLE REGRESSION MODEL

Hypothesis testing is used to carry out inferences about the regression parameters β_1 using data from a sample set. The objective is to test the overall significance of the regression. It is necessary for the lectures to place the emphasis on

CONCLUSION

Regression analysis is not only a statistical process for estimating the relationships between variables, but it is also widely used for prediction and forecasting. It is difficult to acquire this knowledge, because it is necessary for undergraduates to have fundamental mathematics and operating computer skills. It includes many techniques for modeling and analyzing several variables.

However, the lectures should be not an exhaustive review of regression methodology, but should focus on how the regression models are derived. In the teaching process, giving hands-on experience to students is necessary. The teacher should pay more attention to the theoretical aspects of models rather than to their implementation using software. In the process of teaching, these will help to improve the teaching effect.

REFERRENCES

[1] Legendre, A.M., Nouvellesméthodes pour la détermination des orbites des comètes. Paris: FirminDidot, Sur la Méthode des moindresquarrés appears as an appendix (**1805**).
[2] Gauss, C.F., TheoriaMotusCorporumCoelestium in SectionibusConicisSolemAmbientum. Ghent University(1809).
[3] Mogull, R.G., Second-Semester Applied Statistics. Kendall/Hunt Publishing Company, 59-59 (**2004**).
[4] González-Recio, O. and Forni, S., Genome-wide prediction of discrete traits using Bayesian regressions and machine learning. *Genetics Selection Evolution*, 43, 7 (**2011**).
[5] Tsuruoka, Y., McNaught, J., Tsujii, J. and Ananiadou, S., Learning string similarity measures for gene/protein name dictionary look-up using logistic regression. *Bioinformatics*, 23, 20, 2768-2774 (**2007**).