

Discussion on the Teaching of Applied Multivariate Statistical Analysis Course in Local Undergraduate College

Yunhua Ye^{1, a *}

¹School of Mathematics, Jiaying University

^amathyhye@163.com

Keywords: Teaching; Education reform; Applied multivariate statistical analysis; Local undergraduate college; Statistical software

Abstract. Applied multivariate statistical analysis is a branch of mathematical statistics and a professional core curriculum for the students majoring in the applied statistics. The method of multivariate statistical analysis is playing an increasingly wide range of applications in the realm of natural sciences and social sciences. It is of great significance to study how to carry out the teaching of the course of multivariate statistical analysis. Basing on the teaching practice of the author, this paper explores the teaching methods of multivariate statistical analysis and puts forward some useful teaching methods in local undergraduate college.

Introduction

Applied multivariate statistical analysis is a branch of mathematical statistics and a professional core curriculum for the students majoring in applied statistics. Nowadays, with the help of various statistical software, the method of multivariate statistical analysis is playing an increasingly wide range of applications in the realm of natural sciences and social sciences. The course of applied multivariate statistical analysis uses mathematical statistics as a basis and uses mathematics methods as the means to cultivate the students' data analysis ability. The contents of applied multivariate statistical analysis course mainly contains the following five parts [1,2,3]: basic theory of multivariate statistics analysis, the simplification of data structure, classification and discrimination, the interrelationship between variables and statistical inference of multivariate data.

In the era of large data, statistics has become one of the most popular majors nowadays. At present, many local undergraduate colleges and universities have set up statistical majors or applied statistics majors or even big data engineering majors. Compared with the traditional undergraduate colleges and universities, the theoretical basis of the students in local undergraduate college is relatively poor. However, their hands-on ability is relatively strong. Therefore more talent training aim of the local undergraduate colleges is targeting for the cultivation of applied talents. Hence, in the era background of big data, the local undergraduate colleges combine with the actual situation to cultivate the practical ability of the their students. Most of the local undergraduate colleges and universities encounter common problems and difficulties in the teaching of the applied multivariate statistical analysis course. For example, it can be seen that the theoretical basis of the students is relatively poor; learning the cumbersome theoretical derivation problems too much which leads to the results of heavy theory and light practice phenomenon; statistical software practice is not reasonable and not enough, etc. Facing many problems as above, the reform of the applied multiple statistics analysis course and its supporting practice curriculum system is imperative [4]. In the present paper, we give some thinking and discussions on the teaching methods and the teaching reform of the curriculum basing on our teaching practices on the teaching of the applied multivariate statistical analysis course. Basing on the teaching practice of the course, we explore some useful experiences and useful teaching methods. To illustrate these clearly, we summarize them into the following aspects.

The Present Situation in the Teaching of Applied Multivariate Statistical Analysis Course

Nowadays, various methods of multivariate statistical analysis have been widely used in all realms. There are some obvious differences in the teaching models of the courses in China and abroad. China's traditional teaching model in applied multivariate statistical analysis course is "mathematical knowledge + examples explaining + problem solving". This traditional teaching model helps the students mastering the basic knowledge, improving their computing power and raising their logical reasoning ability and application ability to a certain extent. However, it is not difficult to see that there are some problems in teaching and assessment of the course. For example, the teachers only focus too much on theoretical teaching and hence the students are lack of initiative and creativity, lack of practice and so on. As a result, the students feel that they can not use the statistical methods freely. The students' learning time is often spend in coping with the examination and therefore it is difficult to stimulate the students' interest in learning. Compared with this traditional teaching models, foreign teaching model puts emphasis on vividness and universality. The Multivariate statistical analysis courses discuss the data in high dimensional space involving complex calculations such as calculus and matrices. The teaching of statistical method also need to put emphasis on the usage of the statistical software since the manual calculation is almost impossible in the present situation of big data era. Therefore, the introduction of software experiment teaching in local undergraduate colleges and universities is also imperative, so that the students can not only master the theoretical knowledge of statistics, but also can master the rational use of statistical software for practical operation. It is not only the need for personnel training, but also the inevitable trend for the teaching reform of the applied multivariate statistical analysis course.

Show the Students Great Charm of the Course and Stimulate the Learning Interest of the Students in the Teaching

It is useful to stimulate the learning interest of the students by showing them the broad usage of the applied multivariate statistical methods in the real situation. We take the corresponding analysis method in multivariate statistical analysis as an example. The method which proposed by French Benzecri in 1970s is a low-dimensional graphical representation used to study the degree of intimacy between variables and variables. Using the corresponding analysis method in the market research, the market researchers can get useful market information, for example, the consumer brand positioning of the product, which helps them to adjust the marketing strategy in time. This research technique can also be used to test the effectiveness of advertising or marketing activities. If the students know how to use the multivariate statistical methods in practice, they will have more motivations and take more initiatives in the learning of the course.

Use Association, Analogy and Other Methods in the Course Teaching

When we are talking about multivariate statistical analysis, we will explain the basic concepts of multivariate statistical analysis in a concise and simple way. We also illustrate the students statistical thinking and data processing methods. Multivariate statistical analysis involves random vectors or random matrices, the objects are somewhat difficult for the students. In the teaching of the chapter for the basic theory, we demonstrate the students that the methods are similar to the case of one-dimensional random variables. Firstly, the population is assumed to satisfy the multivariate normal distribution; Secondly, we discuss the definitions, numerical characteristics and properties of multivariate normal distribution; Finally, we discuss the distribution of statistics constructed by the samples of multivariate normal population. The students can learn these facts by the analogy method with one-dimensional random variables. When we talk about the sampling distribution and hypothesis test of multivariate normal population, we also show the students the method of association and analogy with one-dimensional case. We first review one-dimensional case and then generalize it to the multivariate case. For example, when introducing the Wishart distribution, we first review and take the constructive definition of the χ^2 distribution into account, and then use the analogy method to

generalize the constructive method directly to the multivariate case. Therefore, we can obtain the definition of the Wishart distribution naturally. These analogical teaching methods are beneficial for the students to review what they have learned and to accept new knowledge. By using the association and analogy methods, we do not need to derive the construction process in detail in the classroom time when teaching the basis of the multivariate distributions to the students. It is sufficient for the students to know the conditions and conclusions of the multivariate distribution and the origin of the statistical distribution. The most important thing is to teach the students how to use these multivariate distributions in the practical cases.

Fully Combine with the Case Teaching Method in the Teaching of Applied Multivariate Statistical Analysis Course

The applied multivariate statistical analysis course is a combination of theory and practice [5]. Since the basic principle is of mathematical and statistical flavor, therefore the course is not easy to understand for the students. In the teaching of the course, only the principles, methods and practice are related together with the help of the teachers, the students can then understand and consolidate what they have learned in the class. We use the textbook “multivariate statistical analysis” published by China Renmin University Press. The textbook is obviously different from pure mathematical statistics textbooks. The textbook highlight the applications of the cases and penetration of statistical thinking, practical thought on multivariate statistical analysis. These things are also combined with the statistical software, mainly the SPSS software [3]. In the classroom teaching, we use the case teaching method to stimulate the students' interest to understand and grasp the principles and methods of multivariate statistical analysis, so as to improve their ability to analyze and solve practical problems. In the classroom teaching of applied multivariate statistical analysis course, case teaching method is adopted, we ask the students complete many tasks by using various kinds of the multivariate statistical methods. We list some typical cases in the follow table.

Table 1 Some typical case teaching methods

Cluster analysis	Cluster analysis of the level of economic development of Guangdong Province; Cluster analysis of urban residents' consumption expenditure in various regions of Guangdong Province;
Discriminant analysis	Discriminant analysis of the level of economic development of countries or regions; Discriminant Analysis of the Author of the book “A Dream of Red Mansions”
Principal Component Analysis	Principal component analysis of the main civil and urban residents' consumption expenditure in Guangdong province;
Factor Analysis	Factor analysis in the applications of the stock evaluation;
Canonical Correlation Analysis	Canonical correlation analysis of post and telecommunication industry and the national economy

In short, by adopting the case teaching methods, the teachers can stimulate the students' interest in learning, to attract the students to think actively and freely, carry out in-depth study and integrate theory and methods of the multivariate statistical analysis closely with the practice [6].

Combine with the Experiments by Using Statistical Software While Teaching the Courses

"Applied multivariate statistical analysis" course must be combined with the experiments by using the statistical software. The reason is obvious: on one hand, it can strengthen the students' learning and

mastery on the statistical software; on the other hand, it can help the students better understand the contents of the theory and train the students' practical ability and innovation consciousness. Some contents of the multivariate statistical analysis methods, such as regression analysis, analysis of variance, principal component analysis, factor analysis, cluster analysis, discriminant analysis, correspondence analysis and typical correlation analysis are very suitable to combine with the experimental course by using the SPSS, SAS, R language and other teaching software [3]. We apply various statistical software flexibly in the teaching of applied multivariate statistical analysis. The main measures we take are listed as follows :

(1) Since various companies use different statistical software , therefore, in order to adapt the students to the needs of the work and the needs of the companies, we strengthen the students adaptive training of the statistical software. We use the SPSS, SAS, R language, MATLAB and other software for experimental teaching. The the students are encouraged to choose different statistical software to complete the experiments. Through these training, the students are capable of using different software to solve various social and economic problems by using the multivariate statistical analysis methods.

(2) According to the characteristics of different software, we let the students choose appropriate software to achieve the specific aim by using multivariate statistical analysis methods. For example, there is no distance discrimination method for the SPSS. In order to make up for this deficiency, we ask the students to use the EXCEL or R software to achieve it directly according to the principles of distance discrimination method. For the students who are more familiar with the MATLAB, we also guide the student to use the MATLAB software programming to achieve the same goal. By these kinds of training, our students can use a variety of statistical software correctly to complete different projects.

(3) We use statistical software to strengthen the programming thinking of the students, so that the students can fully understand the principles and methods of multivariate statistical analysis methods. For example, in the study of the principal component analysis, we ask the students use R software programming to achieve this goal. We let the students write their own programming after mastering the principles of the principal component analysis, rather than simply ask them use the command of `princomp ()` function and `pca ()` function. We then ask the student deal with the urban residents' consumer spending of Guangdong Province by using the principal component analysis. Some stutents choose to use the R software. The functions `scale ()` and `var ()` are used to standardize the data and calculate the correlation coefficient matrix; while `eigen ()` is used to calculate the eigenvalues and eigenvectors of the correlation coefficient matrix, etc.. Then the contribution rate of principal component variance and the expression of the principal component can be calculated according to the eigenvalues and eigenvectors of the correlation coefficient matrix. Through these procedures, the students fully understand the principle and procedures of the PCA techniques through completing the principal component analysis of the urban residents' consumption expenditure by using the R software programming.

Combine the Teaching of Applied Multivariate Statistical Analysis Course with Curriculum Papers, Innovation and Entrepreneurship Projects of the College Students Closely

The "Applied multivariate statistical analysis" course contains a wealth of theoretical knowledge, and the course is also highly practical. Therefore there is a high demand that the examination of the course needs to be combined with the actual examination of the practical problem. The purpose of comprehensive assessment can be achieved by letting the students write the course paper and the experiment report. The teachers should give guidance to the students in choosing research topics, conceiving and consulting relevant information and collecting data so as to enable the students to apply the multivariate method to solve the practical problems. Under the guidance of the teachers, the students in local undergraduate college have completed the project of innovation and entrepreneurship by using multivariate analysis of variance, multiple regression analysis, principal component analysis, factor analysis and other multivariate statistical methods. These projects in our college includes college students' sexual concept survey, college students' part-time job survey and college students' travel demand and behavior analysis survey, etc.. Through the completion of these projects, the students have

tremendously improve their comprehensive abilities and hence they can achieved good grades in the course and in the same time, the students' data collection and data analysis capabilities are exercised.

Organize the Students to Participate in Various Competitions and to Complete Various Practical Problems by Using Multivariate Statistical Analysis

In the teaching of the applied multivariate statistical analysis course, we require the students participate in various types of competitions. The students of applied statistics in our college then actively participate in various competitions, such as China Mathematical Contest in Modelling , Mathematical Contest, Teddy Cup data mining challenge, data innovation competition of the Guangdong University students and other events. Through the participation of these competitions, our students not only have achieved good results, but also have improved their ability to collect and analyze data.

Conclusion

Basing on the practical situation of local undergraduate college and the characteristics of applied multivariate statistical analysis course, this paper mainly discusses some teaching practices of the course and its accompanying curriculum system reform. As one part of the teaching, it is in line with our local undergraduate colleges to cultivate the applied talents. Some procedures taken by our college ensure that the students can better understand the statistical ideas, better use the statistical methods. The teachers should grasp how to better stimulate the students' interest in learning. By integrating the theory with the practice and strengthening the use of the statistical software, the students in our college like to find problems, ask questions and solve problems in the learning. Therefore these procedures we have taken can lay a solid foundation for the future work of our students.

Acknowledgements

The work is supported by 2014 Teaching Quality and Teaching Reform Project of Undergraduate University in Guangdong Province (No.415B0202) and Guangdong Natural Science Foundation (No. 2016A030307008).

References

- [1] R.A. Johnson and D.W. Wichern, Applied multivariate statistical analysis, second ed., Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 2012.
- [2] W.K. Härdle and L. Simar, Applied multivariate statistical analysis, third ed., Springer-Verlag, 2012.
- [3] X.Q. He, Multivariate statistical analysis, fourth ed., China Renmin University Press, Beijing, 2015.
- [4] L. Ma and Z.F. Zhang, Reform on the teaching method of the multivariate statistical analysis, Yinshan Academic Journal (Natural Science Edition) (2016) 1-3.
- [5] M. Han, Teaching research and practice of multivariate statistical analysis, Mathematics Learning and Research (2014) 12-13.
- [6] S.L. Ma, Y. Qian and Y.C.Hou, Case teaching mode of the course of Multivariate statistical analysis on SPSS, Journal of Chaohu College (2014) 147-148.