

COMPARISON OF COMPUTER PROGRAMS WHICH USED IN STATISTICS TEACHING

HYEYOUNG OH

ABSTRACT. We used SAS program in statistics class as a plan to increase effect of class. Because lease cost of SAS program is high and low-level students felt difficulty of the syntax of SAS program, I planned to introduce other instruction-software. We observe whether it's O.K. to use Excel instead of SAS program and Excel is helpful in having positive attitudes. We try to compare and analyze the merits of two programs. The purpose of this study is to apply Excel to statistics class, cut down the cost of instruction-software, and make students understand statistics more conveniently and easily.

1. Introduction

In twenty-first century called knowledge industry society, the development of Information Technology has much contributed to the field of mathematics-education. Classical class environment to use blackboard and chalks was changed into class environment to make good use of computer by excellent function and performance of computer [7]. Recently in the process of teaching-learning, various technologies such as graphic calculator, spread sheet, or GSP were introduced to mathematics class. So the attempts and researches to try to graft them onto discrete mathematics, the concept of function, graph, and statistics have been achieved continuously [3].

We used SAS program in statistics class as a plan to increase effect of class by making students concentrate on class. The results of computer lab to use SAS program were that all the students felt the interests, they

Received November 4, 2011. Revised December 3, 2011. Accepted December 10, 2011.

2000 Mathematics Subject Classification: 97C80, 97U70.

Key words and phrases: SAS program, excel, command, Ribbon menu, automatic-filling.

This work was supported by the University of Incheon Research Grant in 2011.

were more motivated in computer lab than in classical class, and their understanding, participation, and satisfaction were high [10].

But lease cost of SAS program is high and so it is necessary to cut down on expenses. Because it is hard for students to buy SAS program and unable to be downloaded, they cannot do homework at home. Moreover, low-level students felt difficulty of the syntax of SAS. I planned to introduce instruction-software to do programming easily as well as to cut down on expenses. Students can buy Excel, follow practice, and use it easily. So I became to replace SAS program by Excel.

We observe whether it's O.K. to use Excel instead of SAS program. We observe whether Excel is helpful in having positive attitudes when the students use it in statistics class, too. We try to compare and analyze the merits of two programs. The purpose of this study is to apply Excel to statistics class, cut down the cost of instruction-software, and make students understand statistics more conveniently and easily.

2. Theoretical background

The use of computer in the process of teaching-learning can lessen the difficulty of mathematical concept teaching. Visualization among many functions of computer can relieve the difficulty of mathematics in respects that not only teacher can visualize and teach abstract mathematical concepts but also students control and experience the function of computer directly. In the step of forming the concept or previous step of algebraic formal inference, intuitive activities such as graphic, animation, simulation can highlight the dynamic and genetic aspects of mathematics. It can develop step by step from simple calculation or procedural-oriented education into education which can increase logical thinking [3].

Present secondary and higher curriculum was revised into the direction to develop thinking and problem solution power, emphasize the practicality of mathematics and use calculator or computer by educational instrument, execute mathematics education adapting to changes of the times, and use various teaching-learning method and estimation method [5].

Freudenthal [2] emphasized the use of computer in the aspects of mathematics education methodology and established the question "How to use calculator and computer to induce and develop mathematical

understanding?" as one of important issues of mathematics education. And he asserted to use computer as a strong instrument to motivate and increase mathematical understanding [1]. This assertion could check in statistics class, too. Experimental subjects admitted the class actively through direct and manipulative experience and showed positive attitudes to class.

3. Excel

Excel is spread sheet program to have Windows environment which developed in Microsoft Company in US. It is spread sheet program for personal computer and Macintosh computer [8].

Spread sheet is table calculation program which automates calculation and office work in various table form. It is software that calculator and calculation sheet are integrated. So the table and operation are written and graph is drawn. It connects many spread sheets, integrates, and draws various figures and graphs in Windows environment.

The first spread sheet was VisiCalk which programmed to implement in personal computer Apple in 1978. VisiCalk was program to calculate simply by using only some cells. But database and graphic function as well as improvement of operation efficiency were added to spread sheet of today. Moreover, it became to provide various functions and be equipped with up to communication function.

Excel is convenient for drawing up formula and calculating in graphic environment of Windows. So it is used widely in tax calculation, report, grade management, and statistics. It is spread sheet program to calculate the content of cell which is the minimum unit adding the concept of row and column to sheet by function automatically.

Excel 2007 was changed from traditional menu form into Ribon menu form. Ribon mean is user interface which is designed to find necessary order easily in order to get desired result [6]. When we use Excel with traditional menu, we should search each menu one by one to find order. Because Ribon menu is changed automatically into menu in accordance with working affairs, is showed in the top of screen, and indicates the menu which used in cell work, we can approach Excel more easily. When we understand simple terminologies to organize the screen, the method to edit worksheet, and function library in Excel, it is possible to apply Excel.

4. A method of study

4.1. Subjects of study. Subjects of this study are students in the department of computer science and those in graduate school of education, who are majoring in mathematics education. There are students who are teachers at high and middle school among the students in graduate school of education. Especially in the case of a middle school teacher, it was helpful for this student to learn statistics with computer lab because she forgot the general parts much except fundamental parts of statistics.

When I taught statistics with SAS program, students in the department of computer science took lessons in theoretical statistics 2 hours and in lab 1 hour a week. I explained the introduction of SAS program and how to use it in lab. Before they execute SAS program, quite much time was necessary in learning the fundamental parts of SAS. Since SAS is language, they should learn its own syntax. To analyze data by using SAS program, they needed data set and so they should learn the generation and application of data set [10].

Statistics class of graduate school of education is 2 hours per week. To teach in lab, it is necessary to use easy software in programming because of short class time. Fortunately much time is not necessary in learning fundamental parts of Excel for statistics class.

4.2. The process of study.

We observe the process of SAS and Excel and compare them chapter by chapter. The chapters of statistics which apply Excel are as follows.

Experiment 1. Frequency distribution

First, we should know fundamental syntax to do programming when we construct frequency distribution by using SAS program. We can get frequency, percentage, cumulative frequency, and cumulative percentage at a time by using `FREQ PROCEDURE`.

But Excel needs no programming and needs to input requirements step by step by using function magician. Frequency, percentage, cumulative frequency, and cumulative percentage are not showed at a time.

Experiment 2. Measure of dispersion

To find average, standard deviation, maximum, and minimum, SAS program uses `MEANS PROCEDURE`. Average, standard deviation, maximum, and minimum are showed in SAS program at a time but Excel uses dialogue box in order to find measure of center. If we calculate

Chapter	Contents
Frequency Distribution	Frequency Distribution, Histogram, Frequency Polygon, Chart, Pie Graph, Bar Graph, Dot Diagram
Measure of Center Measure of Dispersion	Average, Standard Deviation, Maximum, Minimum
Discrete Probability Distribution	Binomial Distribution, Hyper Geometric Distribution, Poisson Distribution
Continuous Distribution	Normal Distribution
Estimation	Confidence Interval on Small and Large Samples
Testing Hypothesis	Testing Hypothesis on Small and Large Samples
Correlation and Analysis of Regression	Scatter diagram, Correlation Coefficient, A straight-line Regression Model

TABLE 1. Chapters of statistics which applies Excel

statistic by hands, calculation would be quite complicated. When we use SAS program and Excel, we can get results easily.

Experiment 3. Discrete probability distribution

To construct binomial distribution in SAS program, we use PROBBNML PROCEDURE. Since SAS program uses cumulative probability, we use difference of the probability to get the probability corresponding to each variable in SAS program. While cumulative probability is given in SAS program, probability corresponding to each variable is given in Excel.

Since there is no difficulty in calculating formula in Excel, we find the probability of binomial distribution by formula. First, we construct binomial distribution table in Excel and then click BINOMDIST function. When function parameter is appeared, we check "true" or "false" in cumulative item. We input "true" in the case of function of cumulative probability and "false" in that of probability density function. Then we can get the desired probability. Once we get one cell, we can get rest of cells easily by automatic-filling function.

New data set should be generated to graph in SAS program and other SAS package is necessary for graph. But it is convenient to graph in Excel by using chart magician.

Experiment 4. Estimation

When we estimate 95% confidence interval on mean, we can get confidence interval of required variable by MEANS PROCEDURE. Although we don't know the process of solving confidence interval, results are showed in the output of program. But it is different from the case of Excel.

When we use the method of descriptive statistics in data analysis tool, we assign the same output range as input range to dialogue box. When we assign confidence interval on mean, standard error is appeared in the output. It is the result of $\frac{s}{\sqrt{n}}$. In the case that level of confidence is 95%, $t_{\alpha/2, (n-1)} \frac{s}{\sqrt{n}}$ is found. We can find the confidence interval of unknown population variance by using $\bar{X} \pm t_{\alpha/2, (n-1)} \frac{s}{\sqrt{n}}$.

Experiment 5. Testing hypothesis

When sample is small and we test hypothesis on unknown data, we compare SAS program with Excel. SAS program uses MEANS PROCEDURE and analyzes variables by using "value of variable-value of test statistic".

Excel uses TTEST function by using function magician. The value of test statistic $T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$ should be calculated first and we click the parameter to represent T value, freedom, testing with an one-sided alternative, and testing with a two-sided alternative. Then we get the output. Both SAS program and Excel give calculation of significance probability p-value. We decide and test whether it belongs to the reject region or not by significance probability.

Experiment 6. Correlation analysis and regression analysis

We compare Excel with SAS program in Correlation Analysis and Regression Analysis. To draw a scatter diagram in SAS program, we use PLOT PROCEDURE. In the case of Excel, after we choose chart magician and sort of graph in the dialogue box, and assign the remarks and position to draw, we draw the scatter diagram. When we draw the scatter diagram, to use SAS program is more convenient than to use Excel.

When we use CORR PROCEDURE in SAS program, the output of correlation analysis is appeared. In the case of Excel, we choose

Item	Satisfaction	Interest	Concentration	Understanding
Average score	4.15	4.37	4.07	3.75

TABLE 2. Reaction of students in graduate school of education

Item	Satisfaction	Interest	Concentration	Understanding
Average score	4.17	4.4	3.5	3.4

TABLE 3. Reaction of students in the department of computer science

correlation analysis in statistics data analysis and use the dialogue box of correlation analysis. Then correlation coefficient is appeared. In this case, we assign the input range and output option.

To analyze regression analysis in SAS program, we use REG PROCEDURE. Then we can get regression coefficient and the least squares regression line. In Excel, regression analysis is chosen in the dialogue box of data analysis. The least squares regression line can be solved by y-intercept and regression coefficient.

While many unnecessary results are given in SAS program, only necessary results are presented in Excel. So it is easy to find the result in Excel.

4.3. Analysis of class to use SAS program and Excel. Experimental subjects are students in graduate school of education who entered in 2011 and students in the department of computer science who entered in 2008. Students in graduate school of education used Excel and those in the department of computer science used SAS program. We try to compare computer programs. And we try to analyze satisfaction, interest, concentration, and understanding of students of each group. We observe whether it's O.K. to use Excel instead of SAS program.

4.3.1. Analysis of computer program. Excel and SAS program have the respective merits. Let's observe the merits of SAS program. The merits of SAS program are as follows.

1. Programming is simple.

To analyze large data, SAS program is more convenient than Excel. SAS program has its own command. It takes quite long time to be familiar with the syntax. But once we are familiar with the syntax, we can get the output easily by simple command. On the contrary, since we use dialogue box in Excel and input step by step, it takes long time.

2. It is convenient to draw a scatter diagram.

When we draw the scatter diagram, to use SAS program is more convenient than to use Excel. When we use Excel, we use chart magician and dialogue box step by step. But it is possible to use PLOT PROCEDURE in SAS program.

The merits of Excel are as follows.

1. Students are able to follow Excel easily.

Although students have never used Excel, they are able to follow it easily once they use it after they learn the fundamental parts of it. When they use Excel, they use function magician or data statistics analysis. Since students click them and Excel doesn't require other grammar, they have no problem. All of high- or low-level students can adapt to Excel easily.

Since we were not able to spend in lab much, I gave the students in graduate school of education a demonstration. After that, I assigned the project to them and let it do at home. They said that one-sided demonstration wasn't helpful to form statistical concepts. It was not understandable and helpful until they learned in lab through the project. Some students said that visual effects which computer gives result in helping to form mathematical concept. Some students who learned SAS program appealed difficulty of it since they didn't understand programming even though SAS programming was not complicated. Students who used Excel adapted to it easily.

2. Convenience of automatic-filling function.

While SAS program ought to be imputed necessary parts one by one, Excel need not to input all the cells. If we input one cell, we can get the rest of cells by automatic-filling function. So it is convenient to input and time is reduced.

3. Functions in excel need to be learned.

SAS program was used as supplementary tool. So high-or low-level students were helped to understand statistics class as well as to be motivated to statistics class. But low-level students had difficulty in understanding concept and adapting to problem in both the theoretical class

and the lab. Because low-level students should study in two ways, computer lab showed the burdensome status to them instead of helping the theoretical class [4]. Low-level students couldn't understand grammar of SAS program.

Programming which was used in statistics class was simple but to understand concept about this was not achieved well. They followed the problem which was presented in lab and so printed out the result in the screen. They felt excited for short time but had difficulty in applying to other problem.

On the contrary, regardless of students' levels, they admitted Excel easily, because it needs no programming, clicks menu, inputs in dialogue box, and then got the result.

4. It is easy to understand the process of derivation.

Because SAS PROCEDURE shows the desired output, students can find result to problem even though they don't know the process of derivation to theoretical class. In the case of SAS program, we can find confidence interval in estimation by using MEANS PROCEDURE. In the case of Excel, we become to understand the process of derivation by using dialogue box and inputting step by step.

5. It is easy to find the result.

In the experiment of regression analysis, many terminologies regardless of result were showed in SAS program and few unnecessary terminologies were showed in Excel. So while it is hard to find the result in SAS program, it is easy to find it in Excel.

6. Cutting cost.

We feel cutting down the budget of school because of expensive lease cost of SAS program. Since students were not able to buy the program and to download it, they were not able to perform project at home. It was inconvenient to perform it at empty time just at school. But it is possible to perform it at home because it is easy to get Excel.

SAS program and Excel have common merits except their respective merits. They are as follows.

1. To be familiar with the fundamental concept of statistics, it is not enough to use just mechanical calculation. We used material which was not complicated to analyze statistics data in the theoretical class. That's why students want to get the result easily without tools. We gave many realistic problems to them in lab. They are complex calculation problems which they undergo hardships when they solve by hand.

Students said that they felt unrealistic because they dealt with only artificial problems which did not occur in realistic situation in theoretical class. But they felt realistic because they dealt with complex problems which were related with real life in lab. Students who experienced that the problems given up were solved easily by computer were impressed.

2. Even students who were not excited at statistics in theoretical class concentrated on the class because of visualization which computer appeared. As they checked directly the fact which learned only by theory, they became to know the power of computer. They admitted spontaneously the contents of class and showed the positive attitudes through direct and manipulative experience.

4.3.2. Analysis of students' response. To analyze students' response on computer lab, survey on experimental units through questionnaires were held. Items of questionnaires were satisfaction, interest, concentration, and understanding on lab class. Scores about satisfaction, interest, concentration, and understanding of students in graduate school of education are 4.15, 4.37, 4.07, and 3.75 points (on the basis of 5 points), respectively. Scores about satisfaction, interest, concentration, and understanding of students in the department of computer science are 4.17, 4.4, 3.5, and 3.4 points (on the basis of 5 points), respectively. Satisfaction and interest among items were the highest. This says that computer lab class using SAS or Excel motivated students.

If I gave students in graduate school of education only theoretical class, they would be bored. By adding computer lab, I made students be motivated and concentrate on the class. They said that computer lab with Excel would be more effective in middle and high school than in university. Also, they said that computer lab class should be achieved in middle and high school if the circumstances are permitted.

5. Conclusion

Since there were many students who feel theoretical statistics class difficult, we applied the class using SAS program to statistics. As computer lab to use SAS program in 2008 was supplementary class to demerits of theoretical class, it played an important role in developing interest and understanding class. But low-level students felt difficult even though SAS program was not complex. Besides, since it costs expensive, it is

impossible to use it at home. So I planned to introduce instruction-software to program easily as well as to cut down on expenses.

Experimental subjects of 2011 are students in graduate school of education. They can't use SAS program because of short time. Since Excel is not hard to program, it plays a role in developing the effect of statistics class for short time.

Previous data shows that Excel helps to study statistics, provides motivation of learning, and aids to understand mathematical connection. Four main regions which was emphasized in NCTM(1989) are mathematical connection, mathematics as problem-solving, mathematics as communication, and mathematics as inference. Mathematical connection plays the same role as connection ring in developing schema [9].

Students themselves became to understand statistical concept through project activity because to use Excel is not hard. They didn't realize true meaning to practice in lab class which professor showed the demo one-sidedly. But they became to realize the importance of mathematical concept in doing project. When they study mathematics focusing on not calculating by memorizing simple formula but understanding, they become to realize and apply the usefulness of mathematics, and have the desire to study.

To be the effective class when they use computer program, they should deal with many complex problems, apply to realistic problem, and achieve learning target by making the level of programming easy. Statistics class to use Excel satisfying above conditions showed that students felt easy, excited, concentrated, and satisfactory. Class which used Excel was made best use as an important tool to teach meaningful mathematics by having maximal effects as minimal time. And it showed more merits than class which used SAS program. For this reason, to use Excel stands comparison with to use SAS program.

By making best use of Excel, statistics class in graduate school of education gives positive change to class, becomes to experience manipulatively, and helps to understand mathematical concept. Computer program becomes to test hard mathematical contents in guessing intuitively through experiment and makes development of mathematical fact easy. To teach lessons with Excel should be made the most use actively for cutting cost and effective class.

This study is a study on comparison between inhomogeneous two groups but additional study on comparison between homogeneous two groups is required afterwards.

References

- [1] E.J. Choi, *A Study on the development of education curriculum applying college calculus level education case and numerical operational software*, J. Korea Soc. Math. Ed. Ser. A: The Mathematical Education, **48** (4) (2009), 213–234.
- [2] H. Freudenthal, *Major Problems of mathematical education*, Educational studies in mathematics, **12** (2) (1981), 133–150.
- [3] W.H. Hwang and S.K. Cha, *A Study on the Effectiveness of Dynamic Geometry Software in Solving High School Analytic Geometry Problems*, J. Korea Soc. Math. Ed. Ser. A: The Mathematical Education, **41** (3) (2002), 341–360.
- [4] H.J. Kang, D.H. Kim, S.H. Seo, H.J. Ahn, and K.S. Choi, *On the analysis and improvement of calculus for life science*, J. Korea Soc. Math. Ed. Ser. E: Communications of Mathematical Education, **20** (4) (2006), 503–521.
- [5] H.K. Kim and I.J. Lee, *A Study on development of mathematics teacher's qualification criterion and details of evaluation region, and evaluation of class ability*, KICE, 2008.
- [6] J.H. Lee, *Excel 2007-Basic and Application-*, Jungik Publishing Co., Seoul, 2009.
- [7] H.S. Lee, J.L. Park, and I.C. Jeong, *A Study on the proof of additive law of sine function using technology-A case study focused on mathematics for the gifted-*, J. Korea Soc. Math. Ed. Ser. A: The Mathematical Education, **48** (4) (2009), 387–398.
- [8] *Naver encyclopedia*, <http://100.naver.com/100.nhn?docid=716037>.
- [9] NCTM, *The Curriculum and Evaluation Standards for School Mathematics*, Reston, VA: Author, 1989.
- [10] H. Y. Oh, *A Study on effective statistics teaching using SAS program*, Korean J. Math. **16** (4) (2008), 471–480.

Department of Mathematics Education
University of Incheon
Jemulpo Campus
Nam-ku Dowha 2-dong 235
Incheon 402-808, Korea
E-mail: hyoh@incheon.ac.kr