

An Educational Software for the Design of Experiments

Jung Jin Lee and Gunseog Kang

Department of Statistics, Soong Sil University

1-1 Sangdo-dong, Dongjak-ku
Seoul, 156-743, KOREA

Summary: This paper introduces a software called the Information Statistics Provider(ISP), which has many modules on teaching statistics as well as data processing. The ISP is a GUI oriented, interactive, and user-friendly system with multiple sheets. This paper focuses on modules of the Design of Experiments, which show the design sheets of an experiment and enable a post-analysis.

1. Introduction

There are several good statistical softwares such as SAS and SPSS, which have many modules on almost all statistical theories for data processing. However, since these softwares are not designed for teaching statistics, they are not so helpful for a novice to learn statistics. This was a motivation why we decided to develop our own statistical software, which can assist teaching as well as data processing. Our project started in 1990, and we completed an educational software CATS (Computer-Aided Teaching for Statistics) in 1993(Lee et al., 1993). More general-purpose statistical software, the Information Statistics Provider (ISP) running in DOS, was developed in 1995. This package includes the statistical analyses that are frequently used by public (Lee et al., 1994). The ISP was upgraded for running under the Windows system and we announced ISP/BASE in 1999, and ISP/SPC in 2000. The CATS component of the ISP consists of many modules designed for teaching statistical theories that are not easy in a classroom lecture. In this paper we focus on modules for teaching the Design of Experiments(DOE) in the ISP/SPC system.

Since target users of the ISP system are students who are not well trained on statistics, an interactive system design with data sheets was adopted. Spreadsheet like Excel can be a powerful tool in handling data such as input and transformation. We developed our own spreadsheet for the ISP system, which enables fast data handling and further allows using multiple sheets. Due to the interactive system design with multiple sheet capability, we can add many nice features for teaching modules in our software. For example, we can show the designed sheet in experimental design, export other intermediate/final results to the sheet, and perform the post-analysis of the DOE easily.

2. Modules for the Design of Experiments

The modules for the DOE in the ISP are one-way layout, two-way layout, three-way layout, four-way layout, 2^{n-p} factorial design, Latin square design, Graeco-Latin square design, orthogonal array design, Taguchi method, response surface analysis, mixture analysis. The basic structures of the system are similar to other packages, but the ISP has its own features such as the designed sheet of an experiment and post-analysis. For example, Figure 1 (left) shows a dialogue box to select a desired Latin square design for

each level of factor. If one chooses a level of factor, type of Latin square design, and the method of experimental order, either standard or random, then the ISP explains relation between the selected design and the fractional design of three factors experiment as Figure 1 (right) and the corresponding designed sheet as Figure 2 (left). The first column of the designed sheet represents the randomized order of experiments, the second and third columns represent levels of two blocking variables, the fourth column is the factor level and the fifth column is for response data input. Figure 2 (right) shows an example of data input and Figure 3 (left) is the analysis of variance table by using the data in Figure2 (right). The ANOVA table shows that the effect of block 1 needs pooling with the error term, and the pooling process can be done easily with the dialogue box for post-analysis as Figure 3 (right). By using the dialogue box, users can ask sample means of each level or a predicted value at given levels of factors as well as pooling process,. Figure 4 shows the pooled ANOVA table (left) and the confidence interval for each level of factor (right).

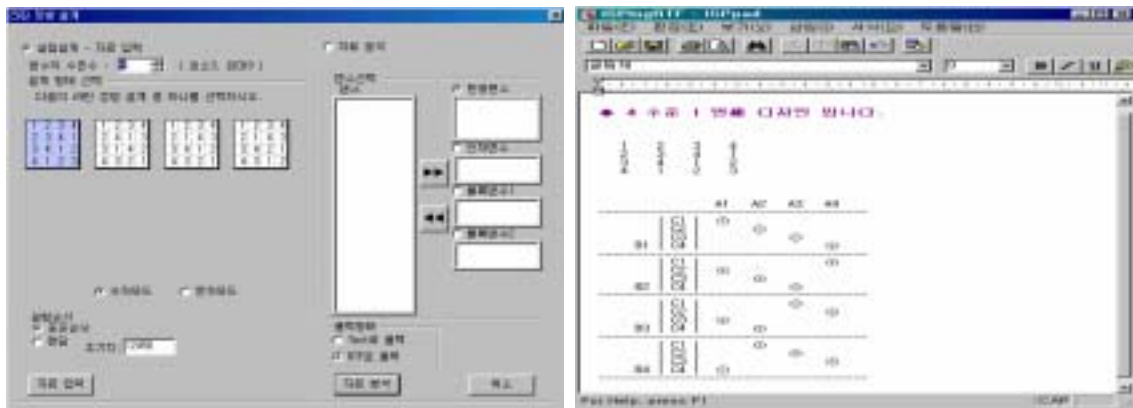


Figure 1. Dialogue box for Latin square design (left) and its explanation as a fractional design of three factors experiment (right)

The image shows two spreadsheets. The left spreadsheet is a designed sheet with columns V1 (Std.Driver), V2 (Block1), V3 (Block2), V4 (Factor), and V5 (Response). The right spreadsheet shows sample data input with columns V1 (Std.Driver), V2 (Location), V3 (Driver), V4 (Tire), and V5 (Wear).

Run	V1 (Std.Driver)	V2 (Block1)	V3 (Block2)	V4 (Factor)	V5 (Response)
1	2	1	2	2	
2	1	1	1	1	
3	9	3	1	3	
4	7	2	3	4	
5	15	4	2	2	
6	5	2	1	2	
7	10	3	2	4	
8	8	2	4	1	
9	3	1	3	3	
10	6	2	2	3	
11	13	4	1	4	
12	4	1	4	4	
13	14	4	2	1	
14	12	3	4	2	
15	11	3	3	1	
16	16	4	4	3	

Run	V1 (Std.Driver)	V2 (Location)	V3 (Driver)	V4 (Tire)	V5 (Wear)
1	2	1	2	2	17
2	1	1	1	1	14
3	9	3	1	3	18
4	7	2	3	4	9
5	15	4	2	2	14
6	5	2	1	2	12
7	10	3	2	4	11
8	8	2	4	1	13
9	3	1	3	3	12
10	6	2	2	3	11
11	13	4	1	4	13
12	4	1	4	4	8
13	14	4	2	1	13
14	12	3	4	2	14
15	11	3	3	1	13
16	16	4	4	3	9

Figure 2. A designed sheet of Latin square design (left) and sample data input (right)

Latin Square Design

◆ Response : Y5(Wear)
 Factor : Y4(Tire)
 Block : Y2(Location)
 Y3(Driver)

◆ ANOVA & Statistics

<< ANOVA TABLE >>

Effect	SS	df	MS	F-value	p-value
Tire	30.6975	3	10.2292	11.4186	0.0008
Location	6.9375	3	2.3125	2.5823	0.1769
Driver	30.6975	3	10.2292	14.3953	0.0038
Error	5.3750	6	0.8958		
Total	80.9375	15			

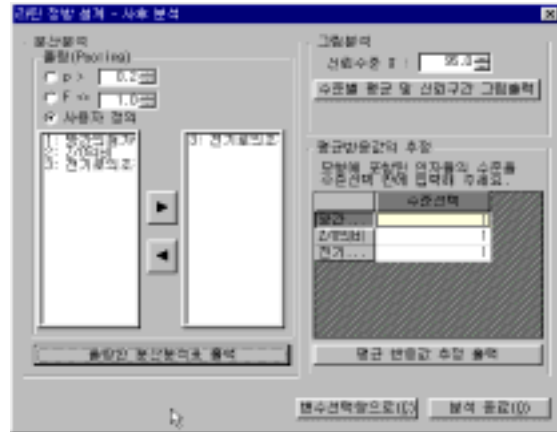


Figure 3. ANOVA table (left) and dialog box for post-analysis (right)

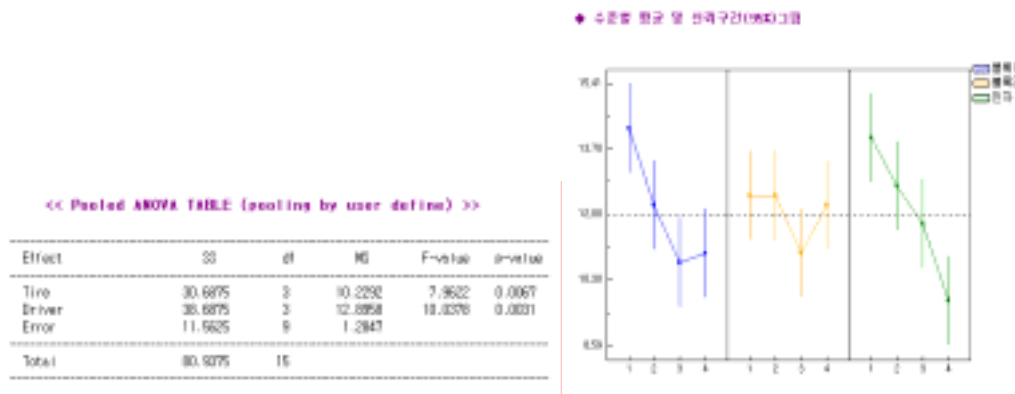


Figure 4. Pooled ANOVA table (left) and mean graphs with confidence interval of each level (right)

This kind of process, design selection → designed sheet for data input → preliminary ANOVA → post-analysis → final analysis, is applied to all modules of the DOE. For example, Figure 5 shows dialog boxes for Taguchi method (left) and response surface analysis (right). Figure 6 shows an example of contour plots from response surface analysis and mixture analysis.



Figure 5. Dialog box for Taguchi method (left) and response surface analysis (right)

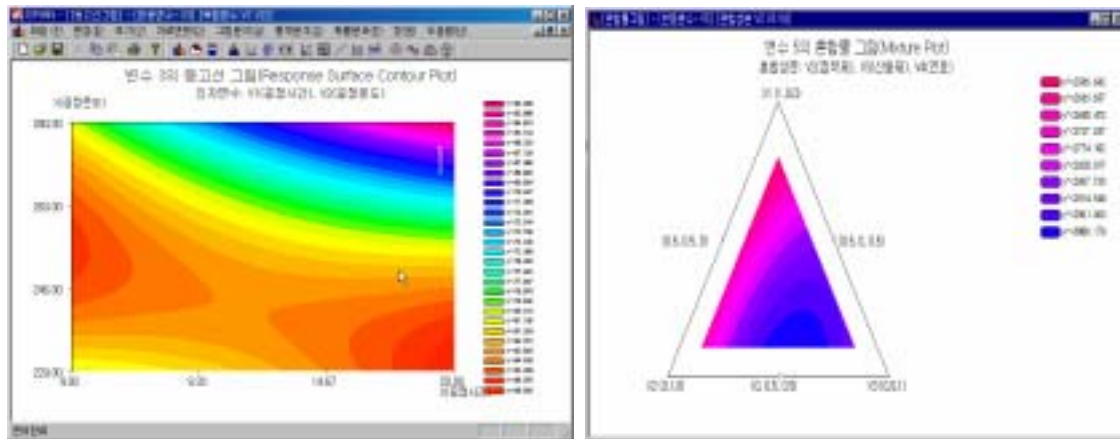


Figure 6. An example of contour plots from response surface analysis (left) and mixture analysis (right)

3. Discussion

This paper introduced the ISP system focused on the DOE, which includes several ideas to teach and to practice the complicated statistical theories. The ISP has a plan to include a hyper-linked textbook, multimedia components, and interactive experiment as traditional statistical software in order to help teaching/learning statistics. We will continue this project and hope to share the ideas and results with those who have common interests.

References:

- Box, G. E. P, Hunter, W. G., and Hunter, J. S. (1978): *Statistics for Experimenters*, Wiley.
- Lee, J., and Kang, G. (1994): Developing a Korean Statistical Package. *The Korean Journal of Applied Statistics*, **7**, 279-288 (in Korean).
- Lee, J., Kang, G., and Lee, Y. (1993): A Study of Computer Aided Teaching for Statistics. *Proceedings of the Asian Conference on Statistical Computing, IASC*, 112-115, Beijing, China.
- Lee, W. and Choi, H. (1998): Dynamic analysis system for graph models operated in Web. *The Korean Communications in Statistics*, **5**, 755-766 (in Korean).
- Oh, B. and Kang, G. (1999): An Introduction of JSP, a Statistical Package Implemented Using JAVA. *Proceedings of the Autumn Conference, Korean Statistical Society*, 132-137 (in Korean).
- Park, S. H. (2000): *Modern Experimental Design*, Min-Young Sa, Seoul.

[Http://www.sas.com](http://www.sas.com)

[Http://www.spss.com](http://www.spss.com)