

IAOS Satellite Meeting on Statistics for  
the Information Society  
August 30 and 31, 2001, Tokyo, Japan

## ***On the Joint Research Development of the Educational Web Site for Disseminating Statistics***

**WATANABE, Michiko,**  
*Toyo University*

**SAKURAI, Naoko,**  
*Tokyo University of  
Information Sciences*

**INOUE, Tatsuki**  
*Waseda Univeristy*

**and**

**ASANO, Chooichiro**  
*SOKA University, Japan*

### **Abstract**

This paper introduces an outline of the integrated web site offering such services as statistics education, electronic books, and a search engine for statistics, which we are currently developing. Statistics is a basic and cross-disciplinary study applied to positive analysis in a wide variety of fields. Practical use as a data analysis tool has to be its primary role. Developing statistical analysis software has been the mainstream of disseminating statistics. Computers have become quite popular and the use of data processing software such as spreadsheets is expanding. In this circumstance, demand for basic statistics education is increasing. As the Internet is growing as a large-scale communication media form now and in the future, to increase information about statistical science on the Internet promotes education and the efficient dissemination of statistics.

### **1 Introduction**

Our society has undergone post-industrialization, computerization, advanced computerization, and finally an IT revolution in the 1990's, which brought us a computer-networked society. The challenges society must tackle are to accumulate and to analyze the enormous volume of digitized data. For example, transaction processes between companies (BtoB) and between company and consumer (BtoC) nowadays go through computer networks and are recorded on-line. As many companies own EOS (Electronic Ordering System) and POS (Point Of Sale) databases, there is an increasing expectation for analysis methodology to make an exploratory search for valuable information from the database.

It has become very common for public agencies, foundations, and private companies to publish statistical databases on the web regarding financial, economic, and social events. According to Yamada (2000), the number of sites where private companies regularly update the statistical data has more than doubled from 233 in Sep. 1998 to 512 in May 2000. This trend is expected to continue. Furthermore, the government is considering publishing a microscopic database in addition to the currently published macroscopic one, to meet the strong demand from universities and economic organizations.

In European nations, where microscopic data of economic/social survey results have already been public and its analysis has been fully utilized via the network, the ability to collect and analyze information through computers is regarded as the key factor in income differentials in the society, which is called 'digital divide'.

Thus, in the society where networking has rapidly become popular and computerized data is abundant,

people who have the below-mentioned abilities are widely needed and training opportunities for the statistical analysis ability of administrative officers and business people is regarded as a social need.

Statistics is a basic and cross-disciplinary study applied to positive analysis in a wide variety of fields. Traditional university classes, however, mainly explain the mathematical aspects and offer no practical exercises as a data-analyzing tool. Recent Internet technology makes it possible to develop educational materials for new media, by utilizing multimedia, dynamic charts and hotlinks to data-offering sites. This will change the learning process from passive to active and from theoretical to practical. At universities and other educational and research institutions, it has become easier to use data resources on the Internet in daily classes regardless of discipline, as the network environment has improved. For this reason, various educational resources, which most-utilize the advantages of interactive and multimedia technologies for the Internet are now being developed by individuals and research institutions.

In this paper, we introduce our projects for development of the educational web site for disseminating statistics.

## **2 ITLS(Interactive Text for Learning Statistics)**

An antecedent of this project was a collaborative project to develop/use an educational web site by teaching staff of statistics-related liberal arts at private universities. It aimed at realizing practical analytic education in statistics and econometrics. When it started in 1999, mass education in a large classroom was still popular, although network infrastructure at private universities had been improved to some extent. People did not expect all the educational materials to be on-line. It was a realistic solution to develop materials or give lectures through maximum utilization of the merits of both paper and on-line texts on a PC display. Therefore, our antecedent project first issued a conventional-style paper textbook, and then constructed web sites, which were complementary to the book. We issued a joint work 'Analysis Methodology of Quantitative Economics in the Internet Era - from Basic to New Frontier' from Taga Publishing in April 1999. In this book, we explained the basic tools for analyzing economic data, estimated the finance market, which attracted practical attention in those days and outlined a new frontier style analysis methodology for assessing credit risk.

By using materials on the Internet through a PC, students were able to actively participate in the learning process. In this environment, students could maintain their will to study, compared with their passive attitude toward conventional style lectures. As a result, it became possible to give participatory and practical education to liberal arts students, who tended not to enter this field.

A problem we had in this project was that we could not give detailed explanations on the web due to the copyright on the published book. So, we set out new ITLS project to develop/publish an interactive textbook for learning statistics. We avoided having common contents with the published book and restructured the web site as an independent textbook. In this process, we standardized the total design, added dynamic charts and improved slides for lectures using Java Applets, and expanded the search function. We offered these materials for self-study by business people and for lecture notes for teaching staff.

ITLS consists of 'Search Engines and Links', 'Statistics and Excel', 'Data Analysis (Description part and Estimate part)', 'Basic Theory of Regression Analysis', 'Analysis of Time-Series and Seasonal Adjustment', 'Lectures on Multivariate Analysis' and 'Statistical Analysis by Statistica'. 'Search Engines and Links' shows a brief explanation of the Internet, how to retrieve economic and social information (statistical data), and how to download data; 'Statistics and Excel' teaches the fundamentals of data analysis procedures on the PC through the operation of spreadsheet software, such as Excel; 'Data Analysis' shows the basics of statistical analysis; and 'Statistical Analysis by Statistica' introduces software for statistical analysis.

In order to promote the understanding of practical data analysis technologies explained in the textbook, the contents of each category offer such functions as keyword searching, supplementary explanation on

keywords, browsing/downloading slides for lectures, links to reference sites, databases for exercises, and Q&A, which cannot be covered by conventional paper textbooks. We try to use short and easy sentences, colors, hyperlinks, images, and dynamic charts, and not to be dependent on texts and numerical formulas, to promote intuitive understanding of the concept of statistics. On the other hand, appropriate explanations using sentences will be made by 'Electronic Books for Statistical Analysis', which will be explained in the next section.

Furthermore, we prepare a page for PC beginners to give some simple explanations with actual screen images of basic operations from keyboard entry to saving files, web browsing/searching, setting up e-mail, using WORD, and creating web sites. This will work collaboratively with basic information education. We are currently offering approx 330 slides from lectures, 120 keyword explanations and 17 dynamic statistical charts in total, using Java Applets.

By making them open to the public on the Internet, we can offer updated accumulation of educational materials without posing a limitation on users, place, and time. At the same time, teaching staff at universities and private companies can share some of the data for their lectures. In the near future, the virtual space where teachers can exchange their opinions freely will be required on the Internet.

This site intends to offer educational materials for teaching staff as well as offering educational services to students. For this purpose, a wide variety of materials open to the users will be required. It is not efficient, however, if the materials are collected by the project members only. So, we are going to publicize the usefulness of this project through publishing papers at domestic/international academic conferences in order to obtain cooperation in sharing educational materials. Furthermore, international cooperation is also important. For example, an international multimedia database including animation and Java Applets which is under preparation by Hagen University, registers and exhibits language-wise multimedia Applets including English, German, and Spanish. We are preparing to publish the English and German versions of the Java Applet created under this project and the Japanese version of the Applet by Mittag (2000). We will register them when they are complete and try to share the resource internationally.

When we use plenty of images and dynamic charts and intend to create a multimedia site with voice data, the response will be slow due to the limitation of line speed. In the future, however, the inter-university network will be improved and we have to prepare the contents in accordance with this trend. For the immediate problems such as access congestion caused by the heavy use during the lecture, we will place the contents in the server offered by other organization and make it work as a mirror site. We are also considering the distribution of CD-ROMs to students for self-study.

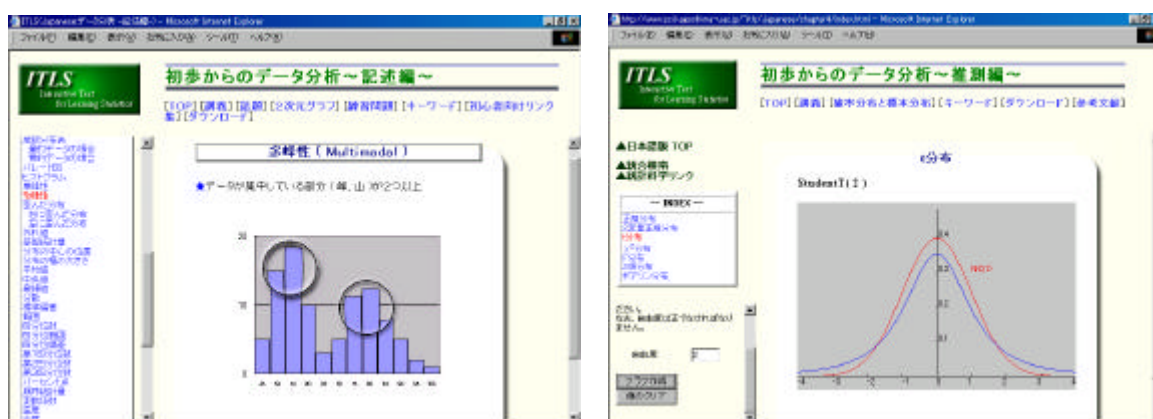


Fig. 1 Examples from ITLS

### **3 EBSA(Electronic Books for Statistical Analysis)**

There are a lot of valuable books which are formally used as technical books for statistical science especially theoretical area or as instruction manual for many fields of application, and are currently out of print or have no plan of reprint. These are valuable for their universality and usability in their statistical science knowledge, regardless of their age, but they are not currently available and usable. To tackle this problem, we created a site named EBSA (Electronic Book for Statistical Analysis) in May 2000 and digitize and publish on the web these out-of-print books on theory and application of statistical science, after receiving official approval from copyright holders (author or the bereaved family, etc.) and publishing companies.

We scanned all pages of the original, converted them to PDF and publish them on the web for browsing and printing. We also collected and integrated all index words in every book and create a glossary database, each of which is referred to in all the applicable pages (group of pages). Therefore, if a user enters a key word of statistical science, the engine searches for the word across all the registered books, or the books the user chose. It often excludes heavy image files and enables users to see and print desired pages only. The system also has text-based index pages so that it can be often retrieved by the prevailing search engines. The opportunities where the original books are read by people will thus be increased.

### **3 Conclusions**

This paper has introduced an outline of the integrated web site offering such services as statistics education, electronic books, and a search engine for statistics, which we are currently developing. Statistics is a basic and cross-disciplinary study applied to positive analysis in a wide variety of fields. Practical use as a data analysis tool has to be its primary role. Developing statistical analysis software has been the mainstream of disseminating statistics. Computers have become quite popular and the use of data processing software such as spreadsheets is expanding. In this circumstance, demand for basic statistics education is increasing. As the Internet is growing as a large-scale communication media form now and in the future, to increase information about statistical science on the Internet promotes education and the efficient dissemination of statistics.

For practical and basic education, a system is required integrating an educational database covering various fields of application, educational software, widely available analytic software, on-line statistics dictionary and analysis sample database. Each site introduced here has not been integrated yet but we believe our effort toward it is very significant.

Web sites introduced here are currently operated on two servers. URLs are:

<http://stat.eco.toyo.ac.jp/~stat/> (Toyo University)

<http://www.sci.kagoshima-u.ac.jp/~stat/> (Kagoshima University)

As the inter-university network system is improved, accumulating and sharing digitized educational materials will be further required in various areas of education. It is hoped that this study would be of some help.

### **References**

Mittag, H-J. (2000). Multimedia and Multimedia Databases for Teaching Statistics , in CD-ROM of Proc. of ICME-9.

Yamada (2000) Report on statistical web sites made by private companies(in Japanese). Proceedings of the 68th annual meeting of Japan Statistical Society 228-229