Environmental Health Perspectives
Vol. 32, pp. 201-202, 1979

Statistical Principle and Methodology in the NISAN System

by Chooichiro Asano*

The NISAN system is a new interactive statistical analysis program package constructed by an organization of Japanese statisticians. The package is widely available for both statistical situations, confirmatory analysis and exploratory analysis, and is planned to obtain statistical wisdom and to choose optimal process of statistical analysis for senior statisticians.

System Modules

The NISAN system has the following diversity functions: (1) highly portable, (2) interactive service, (3) data base, (4) inquiry, (5) documentation and (6) abundant data investigation. The computer language used in the NISAN system is FORTRAN only, and the system consists mainly of four large modules which are interlaced with one another. (*1-8*).

Most STAT modules introduce several adequate processes of statistical analysis, which include current and newly developed statistical methods in exploratory or confirmatory situations. Therefore it may be said at present that the modules attempt to give us the most progressive methodology for the practical use of statistical program packages.

The NISAN system files consist of a data file for analysis and a document file for records. DATA modules control the data file, and DOC modules control the document file. Without considering the file structures, all computational modules can have easy access to any files through DATA and DOC modules. HELP modules have complete inquiry functions for giving answers to the users for dispelling their questions on interpretation and utilization of NISAN system. It is also aimed that even if the user is beginner, he will have easy access to the NISAN system by the guide of HELP modules. The minimum equipment, as a terminal, is a character display or a teletype terminal to access NISAN system.

*Kyushu University, Research Institute of Fundamental Information, Fukuoka, Japan.

Statistical Principle

In view of statistical methodology, the most emphasized functions are as follows: (1) data investigations, (2) graphic representations, (3) generalized methods of optimal scaling, (4) various methods of cluster analysis and multidimensional scaling, and (5) studies on successive processes of statistical inference.

Generally speaking, on applying the statistical methods in practice, it is very important for us to know the features and properties of input data in advance, and we cannot neglect such investigations. Because we believe that the current statistical program packages do not have enough data investigation functions, we studied and strengthened the contents of data investigations, especially the examinations of multivariate normality, the transformation of variables, the check of outliers, and so on. Thus we enable data investigation before statistical analysis.

According to visual features, i.e., inspecting on graphical representation at every stage of analysis, we obtain occasionally important insights, effective suggestions, and interpretation not only on the input data but also on the method of analysis and the results. The NISAN system can give us arbitrarily such graphical representations by users' request.

For categorical input data, the NISAN system includes generalized methods of optimum quantification developed by us, as well as the ordinal nonparametric methods. We are able to choose any of them, depending on the existence of external criterion and on the ordered relations among item-categories.

The methods of cluster analysis and multidimensional scaling give us some of the exploratory methodology in data analysis. These are also widely included with our newly developed system and are applicable with relation to the quantity and properties of input data and to the hierarchical or non-hierarchical structure of clustering.

Concerning the situations of statistical inference, methodologies of testimating, testipredicting and testitesting are included in view of the unspecified mathematical model of data, e.g., pooling methodology is available in cases of ANOVA and MANOVA with the resultant estimations. Facing the statistical data analysis, we can obtain statistical wisdom on the accuracies and precisions for estimates, the powers of test for various alternative hypotheses and so on, by applying a simulation function in NISAN system. Thus, as a whole process of statistical inference, we can consider and examine performances and strategies, and can finally proceed with the optimal procedure of statistical analysis.

Conclusion

The NISAN system was planned and constructed by a statisticians' organization in Japan, since 1976, and the first version will be finished by the end of 1978. The attempts, principle, and methodology of the NISAN system are considerably different from those of the ordinary statistical program package currently being used.

REFERENCES

1. Asano, C., and Jojima, K. An investigation report on statistical program packages in Europe. Report of B-1 Section in the Special Research Project Formation Process of Information System and Organization of Scientific Informations, July (1977).
2. Tarumi, T. Surveys on Statistical Program Packages Used in Japan, Research Report Connected with Grant (A) in Aid for Encouragement of Scientist. March 197.
3. Goto, M., Uesaka, H., and Asano, C. Methodology of data investigation in NISAN system: multivariate normality tests. Res. Rep. No. 90, Res. Inst. Fund. Inform., Sci., Kyushu University, 1978.
4. Goto, M., Inoue, T., and Asano, C. Confirmatory studies of NISAN random numbers. Res. Rep. No. 91, Res. Inst. Fund. Inform., Sci., Kyushu University, 1978.
5. Wakimoto, K., and Taguri, M. Constellation graphical method for representing multidimensional data. Ann. Inst. Statist. Math. 30: 77 (1978).
6. Tanaka, Y., and Asano, C. A generalized method of optimal scaling for partially ordered categories, Res. Rep. No. 76, Res. Inst. Fund. Inform., Sci., Kyushu University, 1977.
7. Matsushita, K., and Ohsumi, N. Evaluation procedure of clustering techniques, Franco-Japon Seminary, Paris, March 13-20, 1978.
8. Hayashi, C. Minimum dimension analysis — MSD. Behaviormetrika 1: 000 (1974).