

# 통계 전문가 시스템에 관한 문헌 연구

염봉진\*    최경미\*\*    김준하\*    김영준\*\*\*

통계 전문가 시스템(Statistical Expert System, SES) 또는 Knowledge-Based Statistical Consultation System은 “통계적 방법에 의한 연구 과정(계획, 자료수집, 분석, 결과해석 등)에서 발생하는 문제를 효과적, 효율적으로 해결하기 위해 통계 전문가 또는 통계 상담자의 경험과 지식을 내장하여 이를 처리할 수 있도록 설계된 전산 시스템”으로 정의할 수 있을 것이다. SES를 구축함으로써 기대되는 효과는

1. 초보 사용자의 통계기법 오용을 방지하고,
2. 기존 통계 software가 지니지 못한 탐색 및 설명기능, 결과해석기능 등을 제공하며,
3. 사용자(초보자 및 전문가)에게 교육 및 통계전략에 대한 실험 환경을 제공하고,
4. 전문가로 하여금 일상적이고 기계적인 업무로부터 벗어나 좀 더 기본적인 문제 해결에 전념할 수 있도록 보조적인 역할을 수행하는 것 등이다.

SES에 대한 연구는 Nelder(1977)에 의해 지능을 갖춘 통계 software의 필요성이 거론되고, Chambers(1980)에 이르러 통계 분야에 전문가 시스템 개념의 도입 가능성이 제시된 이후, 미국과 유럽을 중심으로 활발히 진행되어 왔다. 이는 Gale(1986a)과 Haux(1986)에 소개된 논문들과 계산 통계분야의 정기 간행물, Proceedings(Statistical Software Newsletter, COMPSTAT, ASA Proceedings, Computer Science and Statistics)에 수록된 다양한 논문들로 입증되고 있다.

본 논문의 목적은 여러 곳에 산재해 있는 SES에 관한 문헌들을 모아 그 성격에 따라 분류함으로써 이 분야의 체계적인 연구를 위한 참고자료를 마련하는 데 있다. 본 논문에서는 관련 문헌들을 아래와 같이 분류하였다.

## I. 이론 및 방법론

1. 개관(Overview)
2. 지식습득(Knowledge Acquisition)
3. 통계전략과 지식표현(Statistical Strategy and Knowledge Representation)

## II. 시스템의 구분

1. 통계적 연구의 계획을 위한 시스템
2. 통계기법의 선정과 적용을 위한 시스템

---

\* 한국과학기술원 산업공학과

\*\* KTA 사업지원단

\*\*\* 한국생산성본부 기술지도부

3. 분석결과의 해석을 위한 시스템
4. 기존 package에 접목된 시스템

### Ⅲ. 기타

1. 통계 전문가 시스템에 관련된 도서, 학술잡지 및 Proceedings.
2. 인공지능/전문가 시스템에 관련된 도서 및 학술 잡지

처음, “이론 및 방법론”에는 SES의 개념, 통계전략, 통계지식의 습득과 표현 등에 관한 일반적인 논문 등을 포함시켰다. 이 중, 지식의 습득과 표현은 인공지능 연구자의 도움을 필요로 하는 부분이지만, 통계전략은 통계전문가에 의해 수행되어야 할 연구 분야이다. Hand(1986b)는, 통계 전략이란 “통계적 방법을 사용하는 데 수반되는 선택, 실행(action), 그리고 의사결정에 관한 공식적인 기술(記述)”이라 하였고, Oldford와 Peters(1986a)는 “실질적인 통계문제를 분석하는 데 있어 경험 있는 통계전문가가 사용하는 논리의 근거”라고 정의하였다. 예를 들어, 주어진 연구목표와 경비, 시간 등의 제약조건 하에서 적절한 실험계획을 선정하는 문제나, 회귀분석의 변수선택에서 여러 기준들을 어떻게 이용하여 하나의 회귀모형을 설정할 것인가 하는 문제등을 해결하기 위해서는, 역행렬을 구하는 방법이나 분산분석표의 구성 이상의 포괄적 전략이 요구된다. 유용한 SES의 구축을 위해서는 통계전략의 수립이 반드시 선행되어야 한다는 점을 고려할 때, 앞으로 이 분야의 연구 성과에 따라 SES의 성패가 좌우된다고 보아도 지나치지 않을 것이다.

다음, “시스템의 구분”에서는 특정 문제해결을 위해 개발된 구체적인 SES에 관한 논문들을 열거하였다. 시작 단계에 있는 연구분야에서 흔히 보듯이, 현재 개발되어 있는 SES는 대부분 실험적인 시스템들이며, 보다 나은 시스템의 개발을 위한 시행착오-학습의 과정이 반복되고 있다. 현재 개발된 대부분의 SES는 통계기법의 선정과 적용에 관한 것이며, 계획과 결과 해석을 위한 시스템의 개발은 아직 미비한 상태이다. 이는 응용 분야(application domain)에 관한 지식을 다루어야 하는 데 따르는 어려움 때문이라고 믿어지며, 앞으로 인공지능 분야의 발전(자연어 처리 등)에 따라 활발한 연구가 기대되는 분야이다. 그리고, 기존의 GLIM, BMDP등과 같은 통계 package에 SES를 접목시켜 기법의 선정과 적용등에 관한 편의를 제공하고자 하는 노력도 활발히 시도되고 있다.

“기타”에는 SES와 인공지능/전문가 시스템에 관한 기본적인 도서, 학술 잡지, Proceedings 등을 수록하였다.

이상의 분류는 물론 절대적인 것은 아니며, 또한 미처 수록되지 못한 논문들도 있을 것이다. 이 점에 대해 관심있는 독자들의 많은 의견과 지적을 바란다.

## 참 고 문 헌

### I. 이론 및 방법론

#### 1. 개 관(Overview)

- Altman, N. S.(1985). Expert Systems and Statistical Expertise, Part I : Statistical Expert Systems, Technical Report No. 17, Department of Statistics, Stanford University.
- Chambers, J. M.(1981). "Some thoughts on expert software", in Proceedings of the 13th Symposium on the Interface of Computer Science and Statistics, Springer-Verlag, New York, 36-40.
- Chambers, J. M., Pregibon, D., and Zayas, E. R.(1981). "Expert software for data analysis-An initial experiment", in Proceedings of the 43rd Session of the ISI, Buenos Aires, 279-293.
- Chowdhury, S. I.(1987). Statistical Expert Systems-A Special Application Area for Knowledge -Based Computer Methodology, Thesis, Linkoping University, Sweden.
- Elliman, A. D. and Wittkowski, K. M.(1987) "The impact of expert systems on statistical database management", Statistical Software Newsletter, 13, 14-18.
- Gale, W. A.(ed.)(1986a). Artificial Intelligence and Statistics, Addison -Wesley, Reading.
- Gale, W. A.(1986b). "Overview of artificial intelligence and statistics", in Gale (1986a), 1-16.
- Gale, W. A. "Applications of artificial intelligence in statistics", to appear in Handbook of Statistics, 9, Elsevier, Amsterdam.
- Gale, W. A. and Pregibon, D.(1984a). "Artificial intelligence research in statistics", The AI Magazine, 5, 72-75.
- Gottinger, H. W.(1988). "Statistical expert systems", Expert Systems, 5, 186-196.
- Hahn, G. J.(1985). "More intelligent statistical software and statistical expert systems : Future directions", The American Statistician, 39, 1-16 (with discussions).
- Hand, D. J.(1984). "Statistical expert systems : Design", The Statistician, 33, 351-369.
- Hand, D. J.(1985a). "Statistical expert systems : Necessary attributes", Journal of Applied Statistics, 12, 19-27.

- Hand, D. J.(1986a). "Expert systems in statistics", The Knowledge Engineering Review, 1, 2-10.
- Haux, R. (ed.)(1986). Expert Systems in Statistics, Gustav Fischer, Stuttgart.
- Jockel, K. -H.(1986). "Statistical expert systems and the statistical consultant-considerations about the planning stage of clinical studies", in Haux(1986), 27-43.
- Jones, B.(1980). "The computer as a statistical consultant", BIAS, 7, 168-195.
- Nachtsheim, C. J.(1987). "Tools for computer-aided design of experiments", Journal of Quality Technology, 19, 132-160.
- Nelder, J. A.(1977). "Intelligent programs, the next stage in statistical computing", in Recent Developments in Statistics, eds. J.R. Barra et al., North-Holland, Amsterdam, 79-86.
- Nelder, J. A.(1984). "Present positions and potential developments : Some personal views(on statistical computing)," Journal of the Royal Statistical Society, A147, 151-160(with discussion).
- Nelder, J. A. and Baker, R. J.(1985). "Statistical software : Progress and prospects", in Proceedings of the 16th Symposium on the Interface of Computer Science and Statistics, 33-37.
- Remus, W. E. and Kottemann, J. E.(1986). "Toward intelligent decision support systems : An artificially intelligent statistician", MIS Quarterly, 10, 403-418.
- Streitberg, B.(1988). "On the nonexistence of expert systems : Critical remarks on artificial intelligence in statistics", Statistical Software Newsletter, 14, 55-74(with discussion).
- Thisted, R. A.(1986a). "Computing Environments for data analysis", Statistical Science, 1, 259-274.
- Tukey, J. W.(1986). "An alphabet for statisticians' expert systems", in Gale(1986a), 401-409.

## 2. 지식 습득(Knowledge Acquisition)

- Gale, W. A.(1985a). "Knowledge acquisition in REX and Student", in Proceedings of the 17th Symposium on the Interface of Computer Science and Statistics.
- Gale, W. A.(1986c). "Student phase 1 : A report on work in progress", in Gale (1986a), 239-265.
- Gale, W. A.(1987a). "Knowlege-based knowledge acquisition for a statistical consulting system", International Journal of Man-Machine Studies, 26, 55-64.

- Gale, W. A.(1987b). "Student : A tool for constructing consultation systems in data analysis", in proceedings of the 46th Session of the ISI, Tokyo.
- Gale, W. A. and Pregibon, D. (1984b). "Constructing an expert system for data analysis by working examples", COMPSTAT 84, 227-236.
- Hand, D. J.(1984) "Statistical expert systems : Design", The Statistician, 73, 351-369.

### 3. 통계 전략과 지식 표현(Statistical Strategy and Knowledge Representation)

- Blum, R. L.(1982a). Discovery and Representation of Causal Relationships from a Large Time-Oriented Clinical Database : The RX Project, Springer-Verlag, New York.
- Blum, R. L.(1982b). "Discovery, confirmation, and incorporation of causal relationships from a large time-oriented clinical data base : The RX Project", Computers and Biomedical Research, 15, 164-187.
- Cox, D. R.(1981). "Theory and general principle in statistics", Journal of the Royal Statistical Society, A144, 289-297(with discussion).
- Cox, D. R. and Snell, E. J.(1981). Applied Statistics-Principles and Examples, Chapman and Hall, New York.
- Ellman, T.(1986). "Representing statistical computations : Toward a deeper understanding", in Gale(1986a), 229-238.
- Gale, W. A.(1983). "Using expert systems for developing statistical strategy", invited paper presented at the Annual Meetings of the ASA, Toronto.
- Gale, W. A.(1985b). "Knowledge representation in data analysis," in Proceedings of the 4th International Symposium on Data Analysis and Informatics, Versailles.
- Gale, W. A.(1985c). "Knowledge representation in statistics", in Proceedings of the 17th Symposium on the Interface of Computer Science and Statistics.
- Gale, W. A.(1986d). "REX review", in Gale(1986a), 173-228.
- Gale, W. A. and Lubinsky, D. J.(1986). "A comparison of representations for statistical strategies", in Proceedings of the ASA Statistical Computing Section, 88-96.
- Hajek, P. and Havranek, T.(1982). "GUHA-80-an application of artificial intelligence to data analysis", Computers and Artificial Intelligence, 1, 107-134.
- Hakong, L. and Hickman, F. R.(1985). "Expert system technique : an application in statistics", in Expert Systems 85, ed. M. Merry, Cambridge University Press, Cambridge, 43-63.

- Hand, D. J.(1985b). "Choice of statistical techniques", Bulletin of the ISI, Proceedings of the 45th Session, 3, Amsterdam, 21. 1-1 to 21. 1-16.
- Hand, D. J.(1986b). "Patterns in statistical strategy", in Gale(1986a), 355-387.
- Huber, P. J.(1986). "Environments for supporting statistical strategy", in Gale (1986a), 285-294.
- Jocker, K.-H.(1986). "Statistical expert systems and the statistical consultant-considerations about the planning stage of clinical studies", in Haux(1986), 27-43.
- Johnson, P. E. Meyer, R. K., Nachtsheim, C. J., and Zualkernan, I. A.(1987). "Strategies in experimental design", in Proceedings of the ASA Statistical Computing Section, 76-84(with discussion).
- Lubinsky, D. J.(1987). "Tess : A Tree based environment for statistical strategies", in Proceedings of the 19th Symposium on the Interface of Computer Science and Statistics, 565-567.
- Mallow, C. L. and Tukey, J. W.(1982). "An overview of techniques of data analysis, emphasizing its exploratory aspects", in Some Recent Advances in Statistics, eds. J. Tiago del Oliveira and B. Epstein, Academic Press, London, 111-172.
- Oldford, R. W. and Peters, S. C.(1986a). "Implementation and study of statistical strategy", in Gale(1986a), 335-353.
- Oldford, R. W. and Peters, S. C.(1986b). "Object-oriented data representations for statistical data analysis", COMPSTAT 86, 301-308.
- Ostermann, R.(1986). "The use of expert systems in different fields of statistics", in Haux(1986), 99-110.
- Pregibon, D.(1986) "A DIY guide to statistical strategy", in Gale(1986a), 389-399.
- Schach, S.(1986). "Computer support for the design and analysis of survey samples", in Haux(1986), 99-110.
- Thisted, R. A.(1985a). "Knowledge representation for expert data analysis systems", in Proceedings of the 17th Symposium on the Interface of Computer Science and Statistics.
- Thisted, R. A.(1985b). "Representing statistical knowledge and search strategies for expert data analysis systems", in Gale(1986a), 268-284.
- Thisted, R. A.(1986b). "Representing statistical knowledge for expert data analysis systems", in Gale(1986a), 267-285.
- Wittkowski, K. M.(1986a). "Generating and testing statistical hypotheses : Strategies for knowledge engineering", in Haux(1986), 139-154.

Wittkowski, K. M.(1986b). "An expert system for testing statistical hypotheses", in Proceedings of the 18th Symposium on the Interface of Computer Science and Statistics, 438-443.

## II. 시스템의 구분

### 1. 통계적 연구의 계획을 위한 시스템

- Baines, A. and Clithero, D. T.(1986). "Interactive user-friendly package for design and analysis of experiments", COMPSTAT 86, 320-325.
- Blum, R. L.(1982a). Discovery and Representation of Causal Relationships from a Large Time-Oriented Clinical Database : The RX Project, Springer-Verlag, New York.
- Blum, R. L.(1982b). "Discovery, confirmation, and incorporation of causal relationships from a large time-oriented clinical data base : The RX Project", Computers and Biomedical Research, 15, 164-187.
- Haaland, P. D., Yen, D. and Liddle, R. F.(1985). "An expert system for the design of screening experiments", in Proceedings of the ASA Statistical Computing Section, 223-228.
- Haaland, P. D., Yen, D., and Liddle, R. F.(1986). "An expert system for the experimental design", in Proceedings of the ASA Statistical Computing Section, 78-87.
- Olivero, R. A., Seshadri, S., and Deming, S. N.(1987). EXPERIMENTAL DESIGN, Statistical Programs, Houston.
- Rasch, D., Guiard, V., Nurnberg, G., Rudolph, E., and Teuscher, F.(1987). "The expert system CADEMO : Computer-aided design of experiments and modelling", Statistical Software Newsletter, 13, 107-114.

### 2. 통계 기법의 선정과 적용을 위한 시스템

- Altman, N. S.(1985). Expert Systems and Statistical Expertise, Part I : Statistical Expert Systems, Technical Report No. 17, Department of Statistics, Stanford University.
- Andrews, F. M., Klem, L., Davidson, T. N., O'Malley, P. M., and Rodgers, W. L. (1981). "A guide for selecting statistical techniques for analysing social science data", Survey Research Centre, Institute for Social Research, University of Michigan.

- Dambroise, E. and Massotte, P.(1986). "MUSE : An expert system in statistics", COMPSTAT 86, 271-276.
- Froeschl, K. A. and Grossmann, W.(1986). "Knowledge base supported analysis of longitudinal data", COMPSTAT 86, 289-294.
- Gale, W. A.(1986d). "REX review", in Gale(1986a), 173-228.
- Gale, W. A. and Pregibon, D.(1982). "An expert system for regression analysis", in Proceedings of the 14th Symposium on the Interface of Computer Science and Statistics, Springer-Verlag, New York, 110-117.
- Hajek, P. and Ivanek, J.(1982). "Artificial intelligence and data analysis", COMPSTAT 82, 54-60.
- Hietala, P(1986). "How to assist an inexperienced user in the preliminary analysis of time series : First version of the ESTES expert system", COMPSTAT 86, 295-300.
- Jida, J. and Lemaire, J.(1986). "Expert systems and data analysis package management", COMPSTAT 86, 251-258.
- Nugent, W. H.(1986). "Artificial intelligence techniques for retrospective help in data analysis", in Proceedings of the 18th Symposium on the Interface of Computer Science and Statistics, 118-121.
- O'Keefe, R.(1985). Logic and Lattices for a Statistical Advisor, Thesis, University of Edinborough.
- Oldford, R. W. and Peters, S.C.(1984). "Building a statistical knowledge based system with MINI-MYCIN", in Proceedings of the ASA Statistical Computing Section, 85-90.
- Oldford, R. W. and Peters, S. C.(1986b). "Object-oriented data representations for statistical data analysis", COMPSTAT 86, 301-308.
- Oldford, R. W. and Peters, S. C.(1986c). "Statistically sophisticated software and DINDE", in Proceedings of the 18th Symposium on the Interface of Computer Science and Statistics, 160-167.
- Oldford, R. W. and Peters, S. C.(1986d). "Data analysis networks in DINDE", in Proceedings of the ASA Statistical Computing Section, 19-24.
- Portier, K. M. and Lai, P.(1983). "A statistical expert system for analysis determination", in Proceedings of the ASA Statistical Computing Section, 309-311.
- Prat, A., Martii, M., and Catot, J. M.(1985). "Incorporating expertise in time series modeling : The STATXPS system", Statistical Software Newsletter, 11, 55-62.



- Pregibon, D. and Gale, W. A.(1984). "REX : An expert system for regression analysis", COMPSTAT 84, 242-248.
- Streitberg, B. and Naeve, P.(1986). "A modestly intelligent system for identification, estimation, and forecasting of univariate time series : A<sup>4</sup> : ARIMA, artificial intelligence, and APL2, in Haux(1986). 111-137.
- 김영준, 염봉진(1987). "RMEX : 회귀모형 설정을 위한 전문가 시스템", 춘계 인공지능 학술발표회 논문집, 37-53.
- 전용진(1988). "AS/TS : Insight2를 이용한 통계기법 선정자문 시스템", 경영과학, 5, 45-55.

### 3. 분석 결과의 해석을 위한 시스템

- Klosgen, W.(1986). "EXPLORA : An example of knowledge based data analysis", in Haux(1986), 45-60.

### 4. 기존 package에 접목된 시스템

- Berzuini, C., Ross, G., and Larizza, C.(1986). "Developing intelligent software for non-linear model fitting as an expert system", COMPSTAT 86, 259-264.
- Carlsen, F. and Heuch, I. (1986). "Express-An expert system utilizing standard statistical packages", COMPSTAT 86, 265-270.
- Darius, P. L.(1986). "Building expert systems with the help of existing statistical software : An example", COMPSTAT 86, 277-282.
- Kuik, D. J.(1984). "QUESTOR : A conversational preprocessor to BMDP", COMPSTAT 84, 329-334.
- Nelder, J. A. and Wolstenholme, D. (1986). "A front-end for GLIM", in Proceedings of the 18th Symposium on the Interface of Computer Science and Statistics, 113-117.
- Smith, A. M. R., Lee, L. S., and Hand, D. J.(1983). "Interactive user-friendly interfaces to statistical packages", The Computer Journal, 26, 199-204.
- Wolstenholme, D. E. and Neider, J. A.(1986). "A front end for GLIM", in Haux (1986), 155-177.

## III. 기 타

### 1. 통계 전문가 시스템에 관련된 도서, 학술 잡지 및 Proceedings

- Gale, W. A. (ed.) (1986a). Artificial Intelligence and Statistics, papers presented at Workshop on Artificial Intelligence and Statistics, held in April, 1985

in Princeton, New Jersey, Addison-Wesley, Reading.

Haux, R.(ed.)(1986). Expert Systems in Statistics, selected papers from a workshop organized by the working group "Computational Statistics" of the German Region of the International Biometric Society, held in December, 1985 in Aachen, F.R.G., Gustav Fischer, Stuttgart.

*Statistical Software Newsletter*, periodicals published by International Association for Statistical Computing, ISI.

*Proceedings of the Statistical Computing Section*, ASA.

*Computer Science and Statistics, Proceedings of the Symposium on the Interface.*

*COMPSTAT, Proceedings in Computational Statistics.*

## 2. 인공지능/전문가 시스템에 관한 도서 및 학술잡지

Brownston, L., Farrell, R., Kant, E., and Martin, N.(1985). Programming Expert Systems in OPS5 : An Introduction to Rule-Based Programming, Addison-Wesley, Reading.

Buchanan, B. G. and Shortliffe, E. H.(1984). Rule-Based Expert Systems : The MYCIN Experiments of the Stanford Heuristic Programming Project, Addison-Wesley, Reading.

Clocksin, W. F. and Mellish, C. S.(1984). Programming in Prolog, 2nd ed., Springer-Verlag, New York.

Forsyth, R.(1984). Expert Systems : Principles and Case Studies, Chapman and Hall, New York.

Nilsson, N. J.(1980). Principles of Artificial Intelligence, Tioga, Palo Alto.

Rich, E.(1983). Artificial Intelligence, McGraw-Hill, New York.

Waterman, D. A.(1986). A Guide to Expert Systems, Addison-Wesley, Reading.

Winston, P. H. and Horn, B. K. P.(1984). LISP, 2nd ed., Addison-Wesley, Reading.

Winston, P. H.(1984). Artificial Intelligence, 2nd ed., Addison-Wesley, Reading.

*Artificial Intelligence*

*Artificial Intelligence Review : An International Survey and Tutorial Journal*

*Expert Systems : The International Journal of Knowledge Engineering.*

*IEEE Transactions on Systems, Man, and Cybernetics.*

*The AI Magazine*

# A Bibliography of Statistical Expert Systems

Bong Jin Yum, Joon Ha Kim

Department of Industrial Engineering  
Korea Advanced Institute of Science & Technology

Kyung Mee Choi  
KTA Research Center

Young Jun, Kim  
Technical Consultancy Division  
Korea Productivity Center