AFOSR-TR-77-1004

ADA043463

FINAL ~~SCIENTIFIC~~ REPORT 1 June 72 - 31 May 77

on

PROBABILITY AND STATISTICS AND APPLICATIONS

AF ~~Grant~~ AFOSR-~~72~~-2350-72

Period: June 1, 1972-May 31, 1977

N00014-67-A-0226-00014

Prepared by

Leon Jay Gleser, Principal Investigator

David S. Moore

Marcel F. Neuts

Approved for public release;
distribution unlimited.

2304 A5

Department of Statistics

and Purdue Research Foundation

Purdue University

July, 1977

34 p.

291 730

FINAL SCIENTIFIC REPORT

on

"Probability and Statistics and Applications"

Grant AFOSR-72-2350

June 1, 1972-May 31, 1977

I. HISTORY AND PERSONNEL SUPPORTED

This grant was originally titled "Statistical Problems in Systems, Maintenance, Human Engineering, and Communications". The current title was adopted in June of 1976. Professor D. S. Moore was principal investigator from June 1, 1972 to May 31, 1974, Professor Marcel Neuts was principal investigator from June 1, 1974 to August 31, 1976, and Professor L. J. Gleser was principal investigator from September 1, 1976 to the termination of the grant on May 31, 1977. Faculty members supported at various times on the grant were:

Dr. D. S. Moore (June 1, 1972 - May 31, 1977),

Dr. Marcel Neuts (October 1, 1973 - August 31, 1976),

Dr. L. J. Gleser (August 1, 1973 - May 31, 1977),

Dr. Prem Puri (May 13, 1974 - July 5, 1974)*

Dr. E. Klimko (summer, 1973),

Dr. H. Robbins (consultant: May 9, 10, 1974).

Research assistants whose dissertations or later technical reports and papers were supported by the grant were:

M. Spruill (summer, 1973), advisor: D. S. Moore,

C. C. Carson (1973-1975), advisor: M. Neuts,

A. K. Bhargava (1973-1975), advisor: L. J. Gleser,

M. Wolfson (1973-1975), advisor: M. Neuts,

*Also partially supported by ONR grant N00014-67-A-0226-00014.

S. Fuhrmann (1974-1976), advisor: M. Neuts,

John D. Healy (1974-1976), advisor: L. J. Gleser,

George Casella (1975-1977), advisor: L. J. Gleser,

V. Ramaswami (1976-1977), advisor: M. Neuts,

D. Mihalko (1977), advisor: D. S. Moore,

R. Berger (1977)*, advisor: S. S. Gupta.

Many areas of statistical and probabilistical research were investigated during the period covered by this final report. It seems most convenient to organize the reporting of progress in these areas by separately (and in alphabetical order) summarizing the work of each of the senior faculty members, and their respective students, who were supported under the grant. At the end of this final report a chronological bibliography of publications, technical reports, and dissertations supported by the grant is given.

II. SUMMARY OF RESEARCH AND OTHER ACTIVITIES

A. Research of Dr. Leon Jay Gleser

1. Multivariate Regression With Errors in Variables

The model of regression with errors in both independent and dependent variables is widely applicable. This model is used in biochemical assay, signal detection and tracking, measurement of human performance, studies of crystalline structure, geology and seismology, and instrumental calibration and comparison. The errors-in-variables model is a special case of the linear function equation models studied by econometricians, and of the fixed-factor factor analytic models of psychometrics.

---

*Also partially supported by ONR grant N00014-75-C-0455.

The multivariate errors-in-variables regression model assumes that independent random vectors (x_i', y_i'), $i = 1,2,\ldots,n$, are observed, where x_i is $p \times 1$, y_i is $r \times 1$,

$$x_i = \xi_i + e_i \qquad , \qquad y_i = B\xi_i + f_i,$$

and the error vectors (e_i', f_i') are independent $(p+r)$-variate normal random vectors with mean vector 0 and covariance matrix φ.

Two Ph.D. dissertations written under Dr. Gleser's direction, and with support of the grant, concern statistical inference for the errors-in-variables model. A. K. Bhargava finds maximum likelihood estimators (MLE) for the parameters of the errors-in-variables model under various assumptions about the parameters. When $\varphi = \sigma^2 I_{p+r}$, he also studies weak consistency properties of the MLE of B and σ^2, and for $r = p = 1$ obtains the exact distribution of the MLE of B, while finding an asymptotic representation for the distribution of the MLE of σ^2 when $r = 1 \leq p$. Some of Bhargava's results are published in Bhargava (1977).

The dissertation of John D. Healy concerns tests and MLE for the general linear functional equation model earlier studied by T. W. Anderson in 1951. Healy's results extend those of Anderson in that the matrix of (regression) parameters is allowed to increase with the sample size. In consequence, large sample theory for the MLE's and likelihood ratio tests is both different and more complicated than in the fixed-dimension case. Healy also considers generalizations of errors-in-variables models which include intercepts and take account of concomitant information that can be obtained from design variables.

Healy dissertation appears as a technical report [Healy (Mimeo Series 471)], and portions of his work are being submitted for publication.

Gleser (Technical Report, Mimeo Series 453) studies the errors-in-variables model with $\varphi = \sigma^2 I_{p+r}$, $r \leq p$, and obtains strong consistency and large sample normality results for the MLE of B and σ^2. Large sample confidence regions for B and σ^2 are also obtained. A paper, Gleser (1976), concerns an improved way of simulating a noncentral Wishart distribution with noncentrality matrix of less than full rank; this new simulation method will be useful in obtaining finite-sample properties of estimators, confidence regions and tests in the errors-in-variables and general linear functional equation models. Unpublished work by Dr. Gleser (see abstract in November, 1973, Institute of Mathematical Statistics Bulletin, #141-61) concerns a theory of intersection-union tests which encompasses most known general methods of constructing goodness-of-fit tests, including those proposed to test the fit of errors-in-variables and linear functional relation models.

Finally, a survey paper by Dr. Gleser presented as an invited talk at the 1977 Central Regional Meeting of the Institute of Mathematical Statistics reviews some of the more important recent results obtained by Drs. Gleser, Bhargava, Healy, and others on the errors-in-variables and general linear functional relationship models. Negotiations are currently in progress to publish an extended version of this paper as an invited paper for the Journal of the American Statistical Association.

2. Minimax Estimation of Means and Regression Parameters

In many physical and biological experiments it can be assumed that the coefficient of variation of the underlying distribution is known.

When the underlying distribution is normai, Gleser and Healy (1976) show that previously suggested estimators of the mean, including the MLE, are inadmissible under squared error loss, being dominated in risk by a certain admissible, minimax, equivariant estimator. A class of Bayes estimators for the mean is also studied.

It has long been an open problem to construct explicit estimators of a mean vector μ of a p-variate normal distribution with unknown covariance matrix Σ which dominate the MLE $\bar{x}$ of μ in risk under the generalized squared error loss:

$$L(\delta(\bar{x},W),\mu,\Sigma) = (\mathrm{tr}Q\Sigma)^{-1}(\delta(\bar{x},W)-\mu)'Q(\delta(\bar{x},W)-\mu).$$

Here, $\delta(\bar{x},W)$ is an estimator of μ based on the sample mean vector $\bar{x}$ and sample cross-product matrix W obtained from (n+1) i.i.d. observations, and Q is a known positive definite matrix. Assuming that the minimum characteristic root of $Q\Sigma$ is bounded below by a known positive constant, Glèser (Technical Report, Mimeo Series 460) finds a wide ciass of estimators $\delta(\bar{x},W)$ which dominate $\bar{x}$ in risk. Making no assumptions about Σ whatsoever, Berger, Bock, Brown, Casella, and Gleser (1977) show that the class of estimators

$$\delta(\bar{x},W) = (I - \frac{\alpha(W)\ c\ Q^{-1}W^{-1}}{(\bar{x}'W^{-1}\bar{x})(n-p-1)})\bar{x},$$

where $\alpha(W)$ = minimum characteristic root of QW, $0 \leq c \leq c_{n,p}$, and $c_{n,p}$ are certain constants tabled through Monte Carlo simulations, dominate $\bar{x}$ in risk. The proof in this paper is parly analytic, and partly based on simulation. In unpublished results, Dr. Gleser has obtained an analytic proof of this result for the case when n is large.

In his Ph.D. dissertation, George Casella has investigated the numerical stability and minimax properties of adaptive generalized ridge regression estimators in the classical linear model:

$$y = X\beta + e,$$

where X: nxp is known, β: px1 is unknown, and e has an n-variate normal distribution with mean vector 0 and covariance matrix $\sigma^2 I_n$. Adaptive ridge regression estimators for the slope vector have the form $\hat{\beta}_K = (X'X + K(\hat{\beta}))^{-1}X'y$, where $K(\hat{\beta})$ is a matrix-valued function of $\hat{\beta}$ having the form

$$K(\hat{\beta}) = P\, D_{k(\hat{\beta})} P',$$

where P is the orthogonal matrix of eigenvectors of X'X, and $D_{k(\hat{\beta})} = \text{diag}(k_1(\hat{\beta}), k_2(\hat{\beta}), \ldots, k_p(\hat{\beta}))$. When σ^2 is known, Casella obtains necessary conditions for minimaxity of the estimators $\hat{\beta}_K$ under general quadratic loss functions:

$$L(\hat{\beta}, \beta, \sigma^2) = \frac{(\hat{\beta} - \beta)'Q(\hat{\beta} - \beta)}{\sigma^2},$$

where Q is a known positive definite matrix. In the case where

$$k_i(\hat{\beta}) = \frac{a_i r(\hat{\beta}'X'X\hat{\beta}/\sigma^2)}{\hat{\beta}'X'X\hat{\beta}/\sigma^2}, \quad i = 1,2,\ldots,p,$$

were a_i, i = 1,2,...,p, are known nonnegative constants, and r(t) satisfies certain regularity conditions, Casella finds necessary and sufficient conditions for adaptive generalized ridge regression estimators to be minimax. Casella also extends his proofs of sufficiency to the case where σ^2 is estimated by the usual estimator based on least squares theory. These last results appear in Casella (Technical Report, Mimeo Series 497).

In his thesis, Casella also considers two definitions of numerical stability of an estimator of β, one based on the condition number of a "design matrix" for the estimator, and the other based on the sensitivity of the estimator to changes $y \to y + \epsilon u$, where ϵ is small and u is a given $n \times 1$ vector. He presents some evidence to indicate that in even moderately well-conditioned designs, minimaxity and numerical stability for adaptive generalized ridge regression estimators may be incompatible requirements. These results are now being prepared as a technical report.

3. Other Research and Other Activities

The paper by Gleser and Kunte (1976), which was revised under the present grant, is a major contribution to the asymptotic theory of sequential interval estimation.

During the period of this grant, Dr. Gleser served as Associate Editor of the *Journal of the American Statistical Association* (1971-1974) and *Psychometrika* (1972-present). He also served (1975-present) as an officer of the Section on Statistical Education of the American Statistical Association, and on the program committee for the 1977 Central Regional Meeting of the Institute of Mathematical Statistics.

B. Research of Dr. Eugene M. Klimko

During the summer of 1973 when he was partially supported by the grant, Dr. Klimko worked in the partially overlapping areas of integer programming, search theory, and numerical analysis of stochastic models.

Klimko (1973) modifies the technique of enumerating indices in Fāa di Bruno's formula to obtain a new, more efficient algorithm for solving knapsack integer programming problems.

Klimko and Yackel (1975) obtain optimality results for a certain restricted class of search strategies in the area of search theory of Wiener processes.

Methods for analyzing the Clearance Problem with two servers in Queueing theory were investigated by Dr. Klimko, and computer programs were written which show that two server clearance problems can be handled with up to 15 customers in each queue. Dr. Klimko also worked with Dr. Neuts on a manuscript for a text on numerical probability, and in connection with this work developed enumeration techniques for evaluation of the joint distribution of points in a bridge hand (Technical Report, Mimeo Series 337), to demonstrate techniques for efficient data storage and enumeration methods in complicated numerical probability problems.

C. Research of Dr. D. S. Moore

1. Chi-Squared Tests of Fit

Chi-squared tests are the oldest and most common tests of fit. They are generally less powerful than special purpose tests or tests based on the empirical distribution function, but have several major advantages over these competing tests. Chi-square tests can be used for data which are discrete, multivariate, grouped, or even censored. Other tests of fit require drastic modification in these circumstances. Moreover, other general tests of fit have non-tabled distributions when data are tested for fit to a parametric model, so that unknown parameters in the model must be estimated from the data. Properly constructed chi-square tests have chi-square limiting null distributions in great generality, enabling the use of standard tables in a wide variety of cases. Modern chi-square tests gain their flexibility from two innovations. The first is the use of data-dependent cells, introduced by A. R. Roy and G. S. Watson in the 1950's. The second is the use of quadratic forms other than the Pearson sum of squares, due to D. S. Robson and his students in the 1970's.

Research under this grant has produced a general theory of chi-square statistics and lead to increased possibilities for applying them. Moore and Spruill (1975) and Moore (1977) are the major papers here. The first gives the large sample theory for the entire class of chi-square statistics with a fixed number of cells. These statistics are nonnegative definite quadratic forms in the standardized cell frequencies, allowing data-dependent cells, multivariate data, and quite general methods of estimating unknown parameters. Asymptotic distributions under both the null hypothesis and contiguous alternatives are obtained. The limiting null distribution of such a statistic is always that of a linear combination of independent chi-square

In Moore (1977), the question of how to find a statistic having a chi-square (and therefore tabled) limiting null distribution is addressed. The form of the statistic depends on the method used to estimate unknown parameters in the model. When minimum chi-square estimation is used, the classical Pearson-Fisher statistic is appropriate. Rao and Robson in 1974 found the proper statistic when maximum likelihood estimation is used, though without adequate proof. Moore (1977) obtains a general recipe for the unique quadratic form having as its limiting null distribution the chi-square law with maximum possible degrees of freedom. The Pearson-Fisher and Rao-Robson statistics are special cases of this.

The thesis of M. C. Spruill (Spruill, 1976a and 1976b)uses Bahadur slope to study both choice of statistic and choice of cells in chi-square tests. Moore (1978) is a modern exposition of chi square theory, to appear in a volume of the Mathematical Association of America's well-known Studies series. Moore (1973) is an unrelated note on another goodness of fit problem.

Work in progress includes application of the general theory described here to tests of fit based on randomly censored data and to tests of fit for multivariate normality, both problems of considerable practical interest. The first of these problems is being pursued by Mr. Daniel Mihalko in his thesis work. This work follows the pattern of Moore and Spruill (1975), but applies to situations in which censored (and hence dependent) observations are used. General results on several chi-square statistics have been obtained, and usable tests of fit for censored data to the negative exponential, normal, Weibull and uniform families of distributions have been derived. This work is being completed with AFOSR support, and will soon be avaible as a technical report.

2. Density Estimation

The use of the nearest neighbor technique in density function estimation goes back to Fix and Hodges' classic work on discrimination. They and subsequent workers estimated a density function $f(z)$ by "empiric measure divided by Lebesque measure" for the sphere centered at z with radius the distance from z to the $k(n)$th nearest of the observations $X_1 \ldots, X_n$. This estimator has the feature that estimation of $f(z)$ is entirely controlled by the $k(n)$ observations nearest to z. This feature has appealed to practitioners, and the estimator has been much used in discrimination, classification and other applications. Work under this grant, done by Dr. Moore in collaboration with Dr. James Yackel, has greatly extended the nearest neighbor technique and given an extensive large sample theory.

Moore and Yackel (1977a) defines a general class of nearest neighbor estimators, using a kernel function to allow the $k(n)$ observations nearest z to be given unequal weights. A heuristic analog between these estimators and the well-known Parzen-Rosenblatt bandwidth estimators is made rigorous to obtain the result that (roughly speaking) any consistency result (pointwise or uniform, in probability or with probability 1) true for the Parzen-Rosenblatt class is also true for the nearest neighbor class of estimators. This allows a large literature to be immediat[illegible] [illegible]plied to nearest neighbor estimators, and even yields some new result[illegible] [illegible]he Fix-Hodges estimator, which is the uniform kernel case of the new [illegible]ss.

In Moore and Yackel (1977b) asymptotic normality and mean consistency are proved for general nearest neighbor density function estimators. In addition, a pointwise almost sure consistency theory for the uniform kernel case is obtained under conditions on $k(n)$ which are shown to be best possible.

3. Other Activities

During the period of this grant, Dr. Moore served as Associate Editor of the _Journal of the American Statistical Association_. He edited, with Professor S. S. Gupta, the volume _Statistical Decision Theory and Related Topics II_ (Academic Press, 1977), the proceedings of a symposium held at Purdue in May, 1976 with the partial support of AFOSR.

Dr. Moore has given numerous lectures on the chi-square results obtained under this contract, including invited lectures at both the 1976 and 1977 annual meetings of the American Statistical Association. He also spoke on this work as one of the lecturers in the American Mathematical Society's "Short Course in Statistics" held in conjunction with its 1977 national meeting in St. Louis.

Dr. Moore was a lecturer in the Visiting Lecturers in Statistics program (1973-1977), and was Program Chairman for the regional meeting of the Institute of Mathematical Statistics held in Madison, Wisconsin in May, 1977.

D. Research of Dr. Marcel F. Neuts

Thanks to the funds and facilities provided by the Grant AFOSR-72-2350, Dr. Neuts has been able to maintain an active research program in Computational Probability and has assumed a leading developmental role in this area. This research is now continuing at the University of Delaware with the support of the Grant AFOSR-77-3236.

In addition, secondary research efforts in the study of limit laws for random variables defined on Markov chains and in the control of epidemics were carried out jointly with doctoral students at Purdue University.

1. Computational Probability

A very large body of literature in probability deals with the study of stochastic models, such as queues, inventories and epidemics. For a number of years prior to 1970, it was pointed out by practitioners that many of the published results were not in a readily computable form and that it was often difficult to extract from them the qualitative information needed in real-life implementations.

Recognizing the validity of this criticism, Dr. Neuts began a program of investigation into the algorithmic aspects of stochastic models. From initial work on certain simple discrete queues, the research has moved on in recent years to the development of stable algorithms for queueing models of substantial complexity.

The three most notable developments obtained in this work are:

a. The introduction of the class of probability distributions of phase type.

b. The solution without transforms of a large variety of queues with semi-Markovian features.

c. The identification of a class of Markov chains, which in the positive recurrent case have a matrix-geometric invariant vector.

These results underlie the numerical solution by real arithmetic of a great variety of queueing problems.

We shall now review each of these in some detail.

The PH-distribuions

A probability distribution on $[0,\infty)$, which can be considered as the distribution of the time till abosrption in a finite Markov chain with m transient states and one absorbing state is said to be of phase type. With the initial probability vector $(\underline{\alpha},\alpha_{m+1})$, the distribution of the time until absorption in the Markov chain

$$\begin{vmatrix} T & T^{\circ} \\ \underline{0} & 0 \end{vmatrix}$$

is given by

$$F(x) = 1-\underline{\alpha}\ \exp(Tx)\underline{e}, \quad \text{for } x \geq 0.$$

The class of PH-distributions contains all finite mixtures of generalized Erlang dis ributions, has a large number of useful closure properties, and is closely related to systems of linear differential equations. Their numerical computation may be easily performed.

The PH-distributions and the related PH-renewal processes enable one to remove from many stochastic models the unrealistic negative-exponential and Poisson assumptions which have often been imposed in the interest of analytic tractability.

By combining results to be reviewed in the next two sections with those obtained for the PH-distributions, Dr. Neuts has been able to evolve algorithmic solutions to such queueing models as PH/G/1, GI/PH/1, SM/PH/1 and others. Since their publication, the papers on PH-distributions have received wide attention. They have been implemented in algorithmic work on stochastic models by several researchers in the US, Canada, Australia, Japan and India.

Jointly with a colleague at the University of Delaware, Dr. Neuts is currently engaged in work on fitting PH-distributions to data by the use of interactive computer graphics and nonlinear optimization. Related work is also in progress, both in West and East Germany.

A Class of Markov Chains and the Markov Renewal Branching Process

There are approximately twelve substantially different complex queueing models of practical interest, which have an embedded Markov chain of the general form

$$P = \begin{vmatrix} B_0 & B_1 & B_2 & B_3 & B_4 & \cdots \\ C_0 & A_1 & A_2 & A_3 & A_4 & \cdots \\ 0 & A_0 & A_1 & A_2 & A_3 & \cdots \\ 0 & 0 & A_0 & A_1 & A_2 & \cdots \\ 0 & 0 & 0 & A_0 & A_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{vmatrix}$$

The entries are finite nonnegative matrices of dimensions mxm for A_ν, $\nu \geq 0$, mxn for C_0, nxm for B_ν, $\nu \geq 1$, and nxn for B_0. The specific definition of the entries of these submatrices depends on the application at hand and can be quite complicated.

The classical methodology, based on probability generating functions and Rouché's theorem, involves a number of irrelevant technical difficulties and does not in general lead to stable, accurate algorithms.

In several papers, Dr. Neuts has first studied the first passage time from the set of states $\{(i + 1, j), 1 \leq j \leq m\}$ to the set of states $\{(i,j), 1 \leq j \leq m\}$. For technical reasons, this first passage problem is also called the Markov Renewal Branching Process. The key mathematical point in this discussion is the study of the solution to the nonlinear matrix equation, in the set of (sub)stochastic matrices:

$$G = \sum_{\nu=0}^{\infty} A_\nu G^\nu .$$

In the positive recurrent case, the unique solution to that equation is stochastic. A large number of computationally useful formulas for moments and other important quantities involve the matrix G. Once the matrix G has been computed, it is possible to express the steady-state probability vector of P, as well as many queue distributions such as for waiting times and queue lengths, in terms of G.

Several of the papers, published with the support of the grant, deal with the theoretical aspects of the nonlinear equation for G. In order to bring the results also as close as possible to real applications, Dr. Neuts has also written several papers dealing with specific models, such as a study of the effect of change-over times on the behavior of a service system, which handles several types of customers. This problem is relevant in computer systems and also in versatile numerically controlled machine tools.

Further research on applications of this methodology is currently still going on. We are examining the very versatile queueing model PH/SM/1, for which we plan to write a detailed computer program. We have also come across new applications, such as a type of radio-channel subject to fade-out and versions of the two-server-in-series model, with a finite intermediate buffer.

<u>Markov Chains with a Matrix-Geometric Invariant Vector</u>

Under very general conditions, Dr. Neuts has shown that Markov chains of the form

$$P = \begin{vmatrix} B_0 & A_0 & 0 & 0 & \cdots \\ B_1 & A_1 & A_0 & 0 & \cdots \\ B_2 & A_2 & A_1 & A_0 & \cdots \\ B_3 & A_3 & A_2 & A_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{vmatrix}$$

have, in the positive recurrent case, an invariant probability vector $\underline{x} = (\underline{x}_0, \underline{x}_1, \ldots)$ of the form

$$\underline{x}_k = \underline{x}_0 \; R^k, \qquad \text{for } k \geq 1,$$

where R is an irreducible, nonnegative matrix of spectral radius less than one.

This theorem, which generalizes a classical result of Khinchin, is very useful in queueing theory. It leads immediately to the numerical solution of queues of GI/PH/1 and SM/PH/1 types. Waiting time distributions, even those under non-classical queue disciplines, can be computed by well-behaved algorithms.

The paper discussing this result has already been accepted for Advances in Applied Probability. Its applications in the control of queues are now under investigation. Most of the harder problems in control of queues are analytically utterly intractable. Where fast algorithms are available, it becomes feasible however to combine a heuristic search procedure with computations to obtain very good approximations to the optimal procedure within a given class. We believe that in the study of these problems, our algorithmic research will bear its finest fruits.

This concludes our review of the main effort in computational probability. Under partial support of the grant, three students completed their Ph.D. dissertations under the supervision of Dr. Neuts and a fourth student is currently receiving support.

Dr. David B. Wolfson wrote a thesis on limit laws for sums of random variables defined on a Markov chain. Several papers of a theoretical nature have been published by Dr. Wolfson, who is currently an Assistant Professor at McGill University.

Dr. C. C. Carson of Sandia Corporation and Dr. S. Fuhrmann of Rutgers University have completed theses, respectively, on the computational solution of the PH/PH/1 queue and on the control of an epidemic, involving a multistage disease.

2. Other Items of Interest

In order to further research interest in computational probability, Dr. Neuts has prepared an extensive bibliography on algorithmic methods in probability, which was published in the Journal for Computational and Applied Mathematics.

He has also edited a special issue of Management Science on Algorithmic Methods in Probability. This volume of seventeen papers will appear in September 1977 as TIMS-North Holland Studies in the Management Sciences, No. 7.

Dr. Neuts was appointed on September 1, 1976 to the Unidel Professorship in the Department of Statistics and Computer Science, University of Delaware.

With the financial aid of the Grant, Dr. Neuts has been able to attend a number of important scientific meetings, most notably the Fourth, Fifth, and Sixth Conferences on Stochastic Processes in Downsview, Ontario, College Park, Maryland, and Tel-Aviv, Israel, respectively. He presented invited papers at all of these conferences.

During the four year period covered by this report, Dr. Neuts has presented more than sixty invited lectures on computational probability at Universities and research institutions in the United States and abroad.

He was recently invited to give a paper at the Conference on Optimization in Statistics to be held in Bombay, India and also to lecture for five weeks at the Indian Institute of Management in Calcutta during December 1977 - January 1978.

Dr. Neuts has also served as Department Editor for Applied Stochastic Models of Management Science since 1974 and has been active in the organization of several conferences and regional meetings.

E. Research of Dr. Prem S. Puri

Dr. Puri wrote three papers under partial support of the grant. Puri (1975) surveys work on carrier-borne epidemics, and then turns to study of a generalization of a carrier-borne epidemic model of Dietz and Downton. In the generalized model, the realistic possibility of an infective becoming a susceptible is permitted. Detailed consideration is given to the special case where the epidemic is initiated by a single carrier with no further immigration of carriers, but with immigration of susceptibles allowed. Among other distributional problems, the distributon problem related to the total number of visits of susceptibles to 'infective' state, and total man-units of time spent in the 'infective' state during the time interval [0,t), etc., are studied. A closely related paper, Puri (Technical Report, Mimeo Series 368), was also supported by the grant.

Puri (1976) is a nonmathematical paper intended for sophisticated but otherwise nonmathematical readers. The paper starts with the emphasis on stochastic models in biology and medicine which, because of the large variability among observations in these areas, are considered more appropriate than their determinstic analogs. Again, owing to the basic evolutionary characteristics of living things such as births and deaths, growth and decay etc., one is led in biology and medicine to many dynamic processes of development in time and space. Consequently, the use of stochastic processes for model-building for the study of various biological phenomena becomes quite natural. The paper attempts to describe nonmathematically some of the special processes, which have emerged as useful models in biology and medicine over a period of time. Among others, the processes considered here are Branching processes, Birth and Death processes, Emigration-Immigration processes, Diffusion processes, Quantal Response processes, Competition processes, etc. In each case a brief sketch, frequently historical in character, is followed by a few examples of live situations where these processes arise in practice. The paper ends with about one hundred references of key papers concerning these processes.

F. Other Research

Partial support was provided to Dr. Roger L. Berger for completion of his dissertation "Minimax, Admissible, and Γ-minimax Decision Rules", written under the direction of Professor S. S. Gupta. An abstract of this dissertation follows:

Multiple decision problems are decision theory problems in which the action space has a finite number of elements. Two different types of multiple decision problems are considered, herein. These two types of problems are subset selection problems and robustness of Bayes rules problems.

Subset selection problems arise because the classical tests of homogeneity are often inadequate in practical situations where the experimenter has to make decisions regarding $k(\geq 2)$ populations, treatments, or processes. This inadequacy may be alleviated by formulating the problems as multiple decision problems aimed at ranking or selection of the k populations. A formulation was proposed in which the population of interest is selected with a fixed minimum probability P^* over the entire parameter space. This formulation is called the subset selection formulation.

Chapter I of this thesis considers minimaxity of subset selection rules when the risk is the expected size of the selected subset or the expected number of non-best populations selected. The minimax value of the selection problem is computed for a wide class of distributions. It is found that two classical rules are minimax in location and scale parameter problems if the populations are independent and the distributions have monotone likelihood ratio. Necessary conditions for minimaxity are obtained and are used to show that other proposed rules are not minimax. The minimaxity of rules in a proposed class of rules is also investigated. Chapter II of this thesis considers minimaxity and admissibility of selection rules when the risk is the maximum probability of selecting any non-best population. It is found that, if the restriction is made to non-randomized, just, and translation invariant rules, a classical rule is minimax and admissible in the location parameter problem. The analagous result holds in the scale parameter problem for scale invariant rules. Other rules in a proposed class are found to behave poorly with respect to this risk. But one of these rules is found to have a certain optimality property if the parameters are in a slippage configuration.

A different type of multiple decision problem is considered in Chapter III. The robustness of Bayes rules when the parameter space is finite is considered. The usual Bayes rules are found to be robust in that they are Γ-minimax when the original distributions are replaced by ε-contaminated versions of themselves. To derive this result, some general results involving the relationship of Bayes rules and Γ-minimax rules are obtained. Bounds are obtained on the amount of contamination which may be present with the Bayes rule remaining Γ-minimax.

III. CHRONOLOGICAL BIBLIOGRAPHY OF PUBLISHED PAPERS AND PAPERS ACCEPTED FOR PUBLICATION

1973

[1] Klimko, E. M. (1973). An Algorithm For Calculating Indices in Făa di Bruno's Formula. BIT 13, 38-49.

[2] Klimko, E. M. and Neuts, Marcel F. (1973). The Single Server Queue in Discrete Time - Numerical Analysis II. Nav. Res. Log. Quart. 20, 305-319.

[3] Moore, D. S. (1973). A note on Srinivasan's goodness of fit test. Biometrika 60, 209-211.

[4] Neuts, Marcel F. (1973). The Single Server Queue in Discrete Time - Numerical Analysis I. Nav. Res. Log. Quart. 20, 297-304.

[5] Neuts, Marcel F. and Klimko, E. M. (1973). The Single Server Queue in Discrete Time - Numerical Analysis III. Nav. Res. Log. Quart. 20, 557-567.

[6] Neuts, Marcel F. and Heimann, D. (1973). The Single Server Queue in Discrete Time - Numerical Analysis IV. Nav. Res. Log. Quart. 20, 753-766.

[7] Neuts, Marcel F. and Wolfson, D. B. (1973). Convexity of the Bounds Induced by Markov's Inequality. Stoch. Proc. and Appl. 1, 145-149.

[8] Neuts, Marcel F. (1973). The Markov Renewal Branching Process. Proceedings of a Conference on Mathematical Methods in the Theory of Queues (Kalamazoo, MI), Springer-Verlag, 1-21.

[9] Neuts, Marcel F. (1973). Some Problems in Computational Probability. Proceedings of the Conference on the Impact of the Computer on Mathematical Research and Teaching, Missoula, MT, 145-151.

1975

[10] Klimko, E. M. and Yackel, James (1975). Optimal Search Strategies For Wiener Processes. Stoch. Proc. and Appl. 3, 19-33.

[11] Moore, D. S. and Spruill, M. C. (1975). Unified Large Sample Theory of General Chi-Squared Statistics for Tests of Fit. Ann. Statist. 3, 599-616.

[12] Neuts, M. F. (1975). Computational Uses of the Method of Phases in the Theory of Queues. Computers and Mathematics with Appl. 1, 151-166.

[13] Neuts, M. F. (1975). A Survey of Exact Computational Methods for Queues. Proceedings of the Eight Symposium on the Interface, Los Angeles, CA, 201-205.

[14] Neuts, M. F. (1973). Probability Distributions of Phase Type. Liber Amicorum Professor Emeritus H. Florin, Dept. Math., U. of Louvain, Belgium, 173-206.

[15] Neuts, M. F. and Carson, C. C. (1975). Some Computational Problems Related to Multinomial Trials. Canad. J. Statist. 3, 235-248.

[16] Neuts, M. F. (1975). Computational Problems Related to the Galton-Watson Process. Proceedings of the Actuarial Research Conference, Brown University, in press.

[17] Puri, Prem S. (1975). A Stochastic Process under the Influence of Another Arising in the Theory of Epidemics. Statistical Inference and Related Topics, Vol. 2, Proceedings of the Summer Research Institute on Statistical Inference on Stochastic Processes, Bloomington, IN, Academic Press, 235-255.

1976

[18] Gleser, Leon Jay (1976). A Canonical Representation for the Non-central Wishart Distribution Useful for Simulation. J. Amer. Statist. Assoc. 71, 690-695.

[19] Gleser, Leon Jay and Healy, John D. (1976). Estimating the Mean of a Normal Distribution with Known Coefficient of Variation. J. Amer. Statist. Assoc. 71, 977-981.

[20] Gleser, Leon Jay and Kunte, Sudhakar (1976). On Asymptotically Optimal Sequential Bayes Interval Estimation Procedures. Ann. Statist. 4, 685-711.

[21] Kiefer, J. and Moore, D. S. (1976). Large Sample Comparison of Tests and Empirical Bayes Procedures. On the History of Probability and Statistics [D. B. Owen (Ed.)], Marcel Dekker, 347-365.

[22] Moore, D. S., Whitsitt, S. J. and Landgrebe, D. A. (1976). Variance Comparisons for Unbiased Estimators of Probability of Correct Classification. IEEE Transactions on Information Theory 22, 102-105.

[23] Neuts, Marcel F. (1976). Moment Formulas for the Markov Renewal Branching Process. Adv. Appl. Prob. 8, 690-711.

[24] Puri, Prem S. (1976). Biomedical Applications of Stochastic Processes. On the History of Probability and Statistics [D. B. Owen (Ed.)], Marcel Dekker, 379-399.

[25] Spruill, M. C. (1976a). Cell Selection in the Chernoff-Lehmann Chi-Square Statistic. Ann. Statist. 3, 375-383.

[26] Spruill, M. C. (1976b). A Comparison of Chi-Square Goodness of Fit Tests Based on Approximate Bahadur Slope. Ann. Statist. 3, 409-412

1977

[27] Berger, J., Bock, M. E., Brown, L. D., Casella, G. and Gleser, L. (1977). Minimax Estimation of a Normal Mean Vector for Arbitrary Quadratic Loss and Unknown Covariance Matrix. Ann. Statist. 5, 763-771.

[28] Bhargava, A. K. (1977). Maximum Likelihood Estimation in a Multivariate 'Errors in Variables' Regression Model with Unknown Error Covariance Matrix. Comm. Statist. Th. and Methods A6, 587-601.

[29] Moore, D. S. (1977). Generalized Inverses, Wald's Method, and the Construction of Chi-Squared Tests of Fit. J. Amer. Statist. Assoc., 72, 131-137.

[30] Moore, D. S. and Yackel, J. W. (1977a). Consistency Properties of Nearest Neighbor Density Function Estimators. Ann. Statist. 5, 143-154.

[31] Moore, D. S. and Yackel, J. W. (1977b). Large Sample Properties of Nearest Neighbor Density Function Estimators. Statistical Decision Theory and Related Topics [S. S. Gupta and D. S. Moore (Eds.)], Academic Press, 269-279.

[32] Neuts, Marcel F. (1977). Algorithms for the Waiting Time Distributions under Various Queue Disciplines in the M/G/1 Queue with Service Time Distributions of Phase Type. In Algorithmic Methods in Probability, TIMS North Holland Studies in Management Science.

[33] Neuts, Marcel F. (1977). Some Explicit Formulas for the Steady-State Behavior of the Queue with Semi-Markovian Service Times. Adv. Appl. Prob. 9, 141-157.

[34] Neuts, Marcel F. (1977). An Annotated Bibliography of Computational Probability - Part I. J. Computational and Appl. Math. 3, 53-74.

To Appear

[35] Moore, D. S. (1978). Chi-Square Tests. Studies in Statistics [R. Hogg (Ed.)], Mathematical Association of America.

[36] Neuts, Marcel F. The M/G/1 Queue with Several Types of Customers and Change-over Times. Adv. Appl. Prob.

[37] Neuts, Marcel F. Markov Chains with Applications in Queueing Theory, Which Have a Matrix-Geometric Invariant Probability Vector. Adv. Appl. Prob.

IV. TECHNICAL REPORTS, OTHER PAPERS AND DISSERTATIONS SUPPORTED

Only those reports and dissertations which are not published or accepted for publication (see Section III) are listed. The listing of Purdue technical reports is by mimeo series number.

Purdue Technical Reports

312. Neuts, Marcel F. Computer Power--or-The Liberation of Applied Probability. (Expository invited address, U. Western Michigan, Kalamazoo, MI, 1973).

337. Klimko, E. M. Numerical Computation of Bridge Probabilities.

341. Gleser, L. J. and Kunte, S. On Asymptotically Optimal and Asymptotically Pointwise Optimal Stopping Rules. (Parts of this report were published in Gleser and Kunte (1976)).

364. Wolfson, D. B. Limit Theorems for Sums of a Sequence of Random Variables Defined on a Markov Chain. (Ph.D. Dissertation)

368. Puri, P.S. A Simpler Approach to Derivation of Some Results in Epidemic Theory. To appear in the Proceedings of an IASPS Symposium held January 3-10, 1974 in honor of J. Neyman.

396. Gleser, L. J. and Kunte, S. On Sequential Asymptotically Optimal Bayes Confidence Interval Procedures. (Parts of this report were published in Gleser and Kunte (1976).)

415. Carson, C. C. Computational Methods for Single Server Queues with Interarrival and Service Time Distributions of Phase Type. (Ph.D. Dissertation)

423. Fuhrmann, S. Control of an Epidemic Involving a Multi-Stage Disease. (Ph.D. Dissertation).

435. Bhargava, A. K. Estimation of a Linear Transformation, and Associated Distributional Problems. (Currently under revision).

441. Neuts, Marcel F. The Effect of Change-Over Times on the M/G/1 Queue With Several Types of Customers. (Revised version to be published in Adv. Appl. Prob.)

453. Gleser, Leon Jay. Estimation of a Linear Transformation: Large Sample Results. (Submitted to Ann. Statist.)

459. Moore, D. S. Recent Developments in Chi-Square Tests for Goodness of Fit. (Invited lecture at 1976 annual meeting of American Statistical Association, Boston, MA.)

460. Gleser, Leon Jay. Minimax Estimation of a Multivariate Normal Mean with Unknown Covariance Matrix (Currently under revision.)

471. Healy, John D. Estimation and Tests for Unknown Linear Restrictions in Multivariate Linear Models. (Ph.D. Dissertation).

497. Casella, George. Minimax Ridge Estimation. (Submitted to Ann. Statist.)

Other Papers

Gleser, L. J. (1977). Multivariate Regression with Errors in Variables: Linear Functional Relationships. (Invited paper presented at Central Regional Meeting of the Institute of Mathematical Statistics, Madison, Wisconsin, May 17, 1977. Submitted to Journ. Amer. Statist. Assoc.)

Neuts, Marcel F. (1977). Solving Queues Without Rouchè's Theorem. (U. Delaware Technical Report. Submitted to Operations Research.)

Neuts, Marcel F. (1977). Renewal Processes of Phase Type. (U. Delaware Technical Report. Submitted to Nav. Res. Log. Quart.)

Ph.D. Dissertations Not Reproduced In Their Entirety as Technical Reports

Berger, Roger L. (1977). Minimax, Admissible, and Γ-Minimax Multiple Decision Rules.

Bhargava, A. K. (1975). Estimation of a Linear Transformation and Associated Distributional Problems.

Casella, George (1977). Minimax Ridge Regression Estimation.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 77-1004	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) PROBABILITY AND STATISTICS AND APPLICATIONS		5. TYPE OF REPORT & PERIOD COVERED Final
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Leon Jay Gleser, David S. Moore & Marcel F. Neuts		8. CONTRACT OR GRANT NUMBER(s) AFOSR 72-2350
9. PERFORMING ORGANIZATION NAME AND ADDRESS Purdue University Department of Statistics West Lafayette, IN 47907		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332		12. REPORT DATE July 1977
		13. NUMBER OF PAGES 31
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Errors in variables, maximum likelihood, likelihood ratio, large sample theory, linear functional equation; minimax estimation, ridge regression, stability of estimation; sequential interval estimation; integer programming; search theory; chi-squared tests of fit, Bahadur slope, tests on censored data; density estimation, nearest neighbor, Parzen-Rosenblatt (Con't. on back of form)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The report ~~is~~ covers a final scientific report covering the research of Drs. Gleser, Klimko, Moore, Neuts, Puri and various research assistants during the period from June 1, 1972 to May 31, 1977. Progress ~~was~~ made in the areas of inference for multivariate errors-in-variables regression models, minimax estimation for normal means and for slopes in linear regression models, sequential interval estimation, integer programming, search theory, chi-squared goodness of fit tests, density estimation, computational probability, queueing theory, and

20. Abstract

various models for stochastic processes. The major results obtained in each of these areas are summarized.

(19) cont'd:

band width estimation, asymptotic normality and consistency; computational probability, stochastic models, queues, distributions of phase type, semi-Markovian queues, Markov renewal branching processes, matrix-geometric invariant vectors, sums of random variables on a Markov chain; carrier-borne epidemic models, stochastic processes in biology and medicine; ranking, subset selection, robustness of Bayes rules, finite decision problems.