Statistical Science
1999, Vol. 14, No. 2, 197–200

Gustav Elfving's Contribution to the Emergence of the Optimal Experimental Design Theory

J. Fellman

Abstract. Gustav Elfving contributed to the genesis of optimal experimental design theory with several papers mainly in the 1950s. These papers are presented and briefly analyzed. The connections between Elfving's results and the results of his successors are elucidated to stress the relevance of Elfving's impact on the development of optimal design theory.

Key words and phrases: A-optimal designs, As-optimal designs, C-optimal designs, optimality criteria, relevant observations, singularity.

1. INTRODUCTION

In general, Gustav Elfving's research activity in mathematics and statistics was characterized by a broad spectrum of interests. His contributions to the development of optimal experimental design theory which were published mainly in the 1950s are discussed in this paper. During that decade I graduated and started as a postgraduate student of Elfving and he guided me into the study of optimal design theory. Therefore, I had the opportunity to closely follow this part of his scientific activity.

In the following I shall give some comments on Elfving's results in optimal design theory and my personal evaluation of their significance for the general development of this part of statistics. Elfving's contributions to other parts of mathematics and statistics have been discussed by Mäkeläinen (1990) and, in this isssue, by Nordström (1999). To everybody familiar with the present status of optimal experimental design theory, Elfving's mathematical style may seem to be somewhat out of date. A heavy arsenal of mathematical technicalities that can be found in almost every paper today is usually not found in his work. This is a consequence of the fact that he was mainly interested in new and fundamental ideas and tried to investigate these with a minimum of mathematical sophistication. Thoughtful reasoning and a search for deep understanding of the problems are central characteristics of his papers. Therefore, a careful reading of the reasoning in his papers shows that the origin of many problems that are still of general interest emanates from, or at least can be found in, his articles.

Some of Elfving's papers are published in journals not easily found by the international community of statisticians. Therefore, the purpose of this paper is to revive his papers and to call young scientists's attention to them. In the following, I neither discuss the papers in chronological order nor intend to give an exhaustive presentation of them.

For readers interested in more general presentations of the whole field of optimal experimental design there are the monographs by Fedorov (1972), Silvey (1980), Pazman (1986), Atkinson and Donev (1992) and Pukelsheim (1993). In this connection, we make special note that Pukelsheim (1993) gives short biographical sketches of Charles Loewner, Gustav Elfving and Jack Kiefer. Recently, Fellman (1997) has discussed the history of statistical science in Finland and the forerunners, including Elfving.

2. ELFVING'S CONTRIBUTIONS

In the history of optimal experimental design theory, Elfving's classical paper of 1952 is probably one of the most central contributions. After more than four decades it is still frequently cited. In this paper he studied C-optimal and A-optimal experimental designs long before these concepts were

J. Fellman is Professor emeritus and former head of the Department of Statistics and Computer Science, Swedish School of Economics and Business Administration, POB 479, FIN-00101 Helsinki, Finland (e-mail: fellman@shh.fi).

coined. For the former the optimality criterion is the variance of the estimate of a linear parametric function $c'\alpha$. For the latter the criterion is the sum of the variances of the estimates of all the parameters. Although he considered the problems mainly in two dimensions and assumed differentiability of the optimality criterion and nonsingularity of the information matrix, he obtained elegant results which proved to be essential for the development of optimal design theory and gave strong stimulus to succeeding research. Within his framework he gave the first optimality criterion according to which one can decide if an observation is relevant or not. Almost all later developed optimality criteria go back to, or have their flavor from, Elfving's initial criterion.

His main results are the following. Consider a set of potential observations

$$y_i = x_i'\alpha + \varepsilon_i, \quad i = 1, \ldots, N,$$

and assume homoscedastic and uncorrelated errors ε_i. The problem is to estimate a linear parametric function $\theta = a'\alpha$ with minimum variance.

To obtain an optimal estimate of θ we select among the possible observations a subset of observations and realize the chosen observations in a specific proportion of the number of observations to be taken. For a given optimality criterion we have to divide the observations in relevant and irrelevant ones. An observation is irrelevant if its inclusion in the subset implies that optimality cannot be obtained; otherwise the observation is relevant. Two questions are raised. Which of the observations are relevant and should be included in the optimal set and what is the optimal proportion of the sample size in which these observations should be realized?

Irrespective of the given linear parametric function θ, the convex hull Π of the observation vectors, that is, the convex polyhedron spanned by the set of the coefficient vectors $\pm x_i$ ($i = 1, \ldots, N$), plays a central role. Elfving (1952) proves that only observations whose vectors lie on the border of the polyhedron Π can be relevant. Starting from the parametric function $\theta = a'\alpha$, he defines the scalar k_c such that the vector $a = k_c a_c$. The optimum is obtained if a_c lies on the border of Π. After changing some observations y_i to $-y_i$ and after appropriate renumbering of the observations, the vector a_c can be written

$$a_c = \sum_{i=1}^{k} p_i\, x_i,$$

where $p_i > 0$ ($i = 1, \ldots, k$) and $\sum p_i = 1$ and $x_1, \ldots, x_k$ lie on the border of Π. With this notation, the optimal set for estimation of $\theta = a'\alpha$ is $(y_1, \ldots, y_k)$ and the optimal allocation is $(p_1, \ldots, p_k)$. Using these geometric arguments, Elfving shows that the maximal number of distinct relevant observations k equals the dimension of the observation vectors. He also notes that the number of distinct observations in the optimal experiment may be smaller than k. This is the case when the vector a_c lies on an edge of the polyhedron. These results are geometrically explained in Fig. 1, which is a reprint of Elfving's classical figure. The condition that an observation is relevant and the proportion in which this observation should be included in the optimal experiment are given by Elfving entirely in geometric terms. He also points out how to generalize the two-dimensional results to models with three parameters. For models with more than three parameters, he states that the geometric rule must be replaced by an algebraic procedure. Such procedures were finally given by his successors.

Successors of Elfving have generalized his geometrical findings. In my thesis, I proved that the polyhedron Π discriminates between relevant and irrelevant observations under very general optimality criteria (Fellman 1974, Theorem 2.1.2). Let the optimality criterion be $q(\mathbf{M})$, where $\mathbf{M}$ is the information matrix of the experiment, let $\min q(\mathbf{M})$ denote optimality and let $<_L$ be the Loewner ordering of nonnegative definite matrices. Then the only restriction on the criterion $q(\cdot)$ is that $\mathbf{M}_1 <_L \mathbf{M}_2$ implies $q(\mathbf{M}_1) \geq q(\mathbf{M}_2)$. My theorem holds also when singularity is allowed in the information matrix $\mathbf{M}$. Recently the polyhedron Π and the geometric results concerning optimality criteria of Elfving have

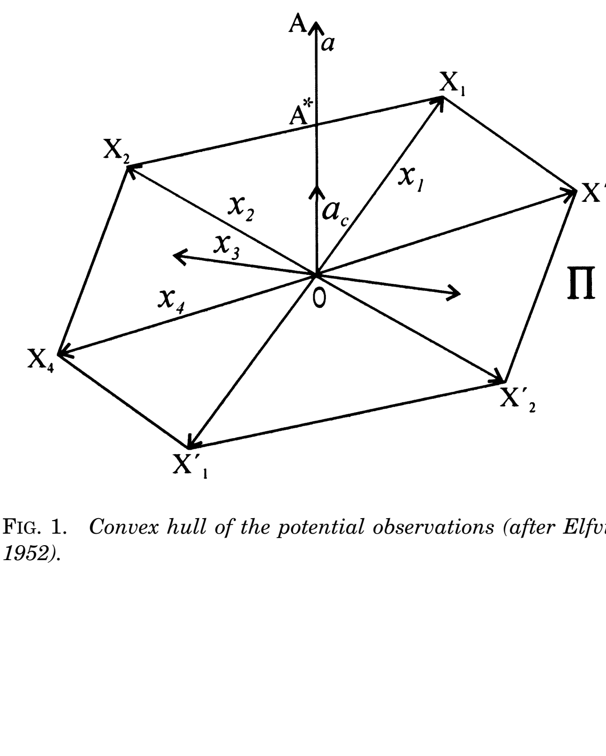

FIG. 1. *Convex hull of the potential observations (after Elfving, 1952).*

been discussed and generalized by Dette (1993a, b), Dette, Heiligers and Studden (1995) and Dette and Studden (1993).

In the paper "Convex sets in statistics" (1953) presented at the Twelfth Scandinavian Congress of Mathematicians and in "A unified approach to the allocation problem in sampling theory" (1954a), Elfving uses similar convexity argument in a more general framework. He introduces a $(1/2)k(k+1)$-dimensional space of symmetric $k \times k$ matrices. In this space the nonnegative definite matrices span a convex cone. Furthermore, the convex hull of the potential observations forms a convex polyhedron in this space. The optimal solution is given by the contact point between this polyhedron and the cone.

In Elfving's last-mentioned papers he also gives the maximal number of different observations contained in the optimal A- and A_s-designs. For the latter design the optimality criterion is the sum of the variances of s (out of k) parameter estimates. These maximal numbers go back to an earlier result given by Chernoff (1953). In my thesis (Fellman, 1974, Chapter 4), I was able to prove that the convexity argument holds under more general conditions. Furthermore, I gave the maximal numbers of different observations in the singular case.

These ideas proposed by Elfving started a sequence of results concerning design criteria. The numerous papers by Kiefer and Wolfowitz, of which some (Kiefer, 1959, 1961, 1962, 1974; Kiefer and Wolfowitz, 1959, 1960, 1964) are given in the list of references, play a central role in the development of this part of the theory. Essential contributions of these papers are equivalence theorems concerning optimality criteria for different design problems. Further generalizations in this domain of optimal design theory are due to Pukelsheim (cf., e.g., Pukelsheim, 1993).

Generalizations of Elfving's results in order to drop the assumptions of nonsingularity of the information matrices and differentiability of the design criteria appear in a sequence of later papers by other authors—for example, Karlin and Studden (1966), Fellman (1974, 1980, 1985), Silvey (1978), Pukelsheim (1983, 1993), Pukelsheim and Titterington (1983) and Gaffke (1985).

A more general problem was considered by Elfving himself in the paper "Geometric allocation theory" (1954b). In my opinion, Elfving's successors have not paid enough attention to this paper. Here, Elfving combines experiments in a more general way by allocating sets of correlated observations. He also considers singular design matrices and he eliminates this problem by a linear transformation of the parameter vector. After the transformation, there remains a reduced set of estimable parameters. Apparently Elfving was aware of the problems which are caused by singular experiments. However, his proposed transformation is no general solution of the singularity problem in optimal design theory. The optimality criterion can be, and often will be, ill-behaved at the optimal point. Especially, differentiability cannot be assumed. This is a consequence of the fact that an experiment which is optimal for parameter estimation often consists of very few distinct observations. Furthermore, the singularity, that is, the set of non- estimable parametric functionals, may vary over the set of designs. Under such circumstances, a universal transformation for the whole set of designs usually cannot be found. My opinion is supported by the fact that the problems connected with singularity have occupied so many of Elfving's successors. In Fellman (1985) the singularity problem is reviewed in greater detail.

The other papers by Elfving are less general. In Elfving (1956) and in Elfving (1957a), Elfving considers the optimal experimental design when the observations are nonrepeatable. The main result is that the solution of this design problem is a hyperplane condition quite similar to the well-known condition obtained for the standard C-optimal design.

Elfving (1957b) considers the estimation of main parameter contrasts in incomplete block designs. His result is that balancedness is the necessary and sufficient condition for the incomplete block design to be a minimax design.

I conclude this short review of Elfving's contribution to optimal experimental design theory with his own review. In a paper in the Harald Cramér Volume (Elfving, 1959), Elfving gives a short but well-written and comprehensive synthesis of his own contributions and of the state of the art at that time. I can warmly recommend this paper to everyone who wishes to become acquainted with the early history of optimal design theory in general and especially with Elfving's contributions.

3. CONCLUSIONS

Elfving was a central member of the small group which laid the foundations of the theory of optimal design. This group left their marks on the entire theory of today. Elfving was primarily interested in fundamental ideas. He was less interested in attempts to obtain the utmost generality. He marked the path and the generalizations were left to his successors. Therefore, he was able to carry through his studies with a minimum of mathematical sophistication. Often standard calculus and basic matrix algebra were enough. In consequence, the central ideas in his papers are not obscured by overwhelming outworks of mathematical technicalities. This fact and his personal talent for writing lucid texts make his papers elegant and highly readable.

ACKNOWLEDGMENT

This paper is a revised and enlarged version of an invited paper (Fellman, 1991) presented at the Gustav Elfving Meeting (organizer Ingram Olkin) at the International Workshop on Linear Models, Experimental Designs, and Related Matrix Theory, August 6–8, 1990, Tampere, Finland.

REFERENCES

ATKINSON, A. C. and DONEV, A. N. (1992). *Optimum Experimental Designs*. Clarendon, Oxford.

CHERNOFF, H. (1953). Locally optimal designs for estimating parameters. *Ann. Math. Statist.* **24** 585–602.

DETTE, H. (1993a). Elfving's theorem for D-optimality. *Ann. Statist.* **21** 753–766.

DETTE, H. (1993b). A new interpretation of optimality for E-optimal designs in linear regression models. *Metrika* **40** 37–50.

DETTE, H., HEILIGERS, B. and STUDDEN, W. J. (1995). Minimax designs in linear regression models. *Ann. Statist.* **23** 30–40.

DETTE, H. and STUDDEN, W. J. (1993). Geometry of E-optimality. *Ann. Statist.* **21** 416–433.

ELFVING, G. (1952). Optimum allocation in linear regression theory. *Ann. Math. Statist.* **23** 255–262.

ELFVING, G. (1953). Convex sets in statistics. In *Tolfte Skandinaviska Matematikerkongressen*, Lund, August 10–15, 1953, 34–39.

ELFVING, G. (1954a). A unified approach to the allocation problem in sampling theory. Abstract from *Proc. Inter. Math. Congr., Amsterdam*, Sept. 1954.

ELFVING, G. (1954b). Geometric allocation theory. *Skandinavisk Aktuarietidskrift* **37** 170–190.

ELFVING, G. (1956). Selection of nonrepeatable observations for estimation. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **1** 69–75. Univ. California Press, Berkeley.

ELFVING, G. (1957a). A selection problem in experimental design. *Societas Scientiarum Fennica Commentationes Physico-Mathematicae* **20**(2) 3–10.

ELFVING, G. (1957b). Minimax character of balanced experimental designs. In *XIII Congr. Math. Scand.* 69–76.

ELFVING, G. (1959). Design of linear experiments. In *Probability and Statistics: The Harald Cramér Volume* (U. Grenander, ed.) 58–74. Almquist and Wiksell, Stockholm.

FEDOROV, V. V. (1972). *Theory of Optimal Experiments*. Academic Press, New York.

FELLMAN, J. (1974). On the allocation of linear observations. *Societas Scientiarum Fennica Commentationes Physico-Mathematicae* **44** 27–78.

FELLMAN, J. (1980). On the behavior of the optimality criterion in the neighborhood of the optimal point. Working Paper 49, Swedish School of Economics and Business Administration.

FELLMAN, J. (1985). Topics on singular linear models with special reference to experimental design. In *Proceedings of the First Tampere Seminar on Linear Models* (1983) 101–122.

FELLMAN, J. (1991). Gustav Elfving and the emergence of the optimal design theory. Working Paper 218, Swedish School of Economics and Business Administration.

FELLMAN, J. (1997). Glimpses at the history of the statistical science in Finland. *Theory of Stochastic Processes* **3**(19) 177–182. (Proceedings of the Second Scandinavian–Ukrainian Conference in Mathematical Statistics, Umeå, 6–13 June 1997.)

GAFFKE, N. (1985). Directional derivatives of optimality criteria at singular matrices in convex design theory. *Statistics* **16** 373–388.

KARLIN, S. and STUDDEN, W. J. (1966). Optimal experimental designs. *Ann. Math. Statist.* **37** 783–815.

KIEFER, J. (1959). Optimum experimental designs. *J. Roy. Statist. Soc. Ser. B* **21** 272–319.

KIEFER, J. (1961). Optimum designs in regression problems, II. *Ann. Math. Statist.* **32** 298–325.

KIEFER, J. (1962). Two more criteria equivalent to D-optimality of designs. *Ann. Math. Statist.* **33** 792–796.

KIEFER, J. (1974). General equivalence theory for optimal designs (approximate theory). *Ann. Statist.* **2** 849–879.

KIEFER, J. and WOLFOWITZ, J. (1959). Optimum designs in regression problems. *Ann. Math. Statist.* **30** 271–294.

KIEFER, J. and WOLFOWITZ, J. (1960). The equivalence of two extremum problems. *Canad. J. Math.* **12** 363–366.

KIEFER, J. and WOLFOWITZ, J. (1964). Optimum extrapolation and interpolation designs. I. *Ann. Inst. Statist. Math.* **16** 79–108.

MÄKELÄINEN, T. (1990). Gustav Elfving 1908–1984. Address presented at the Gustav Elfving Meeting (organizer Ingram Olkin) at the International Workshop on Linear Models, Experimental Designs, and Related Matrix Theory, August 6–8, 1990, Tampere, Finland.

NORDSTRÖM, K. (1999). The life and work of Gustav Elfving. *Statist. Sci.* **14** 174–196.

PAZMAN, A. (1986). *Foundations of Optimum Experimental Design*. Riedel, Dordrecht.

PUKELSHEIM, F. (1983). On optimality properties of simple block designs in the approximate design theory. *J. Statist. Plann. Inference* **8** 193–208.

PUKELSHEIM, F. (1993). Optimal Design of Experiments. Wiley, New York.

PUKELSHEIM, F. and TITTERINGTON, D. M. (1983). General differential and Lagrangian theory for optimal experimental design. *Ann. Statist.* **11** 1060–1068.

SILVEY, S. D. (1978). Optimal design measures with singular information matrices. *Biometrika* **65** 553–559.

SILVEY, S. D. (1980). *Optimal Design*. Chapman and Hall, London.