Information Theory
Statistical Decision Functions
Random Processes

Volume B

Transactions of the
Seventh Prague Conference
and of the European Meeting
of Statisticians
1974

ACADEMIA PRAHA

CZECHOSLOVAK ACADEMY OF SCIENCES

CZECHOSLOVAK ACADEMY OF SCIENCES

Printed in Czechoslovakia

CONTENTS

Invited Papers

Communications

PREFACE

The Prague Conferences on Information Theory, Statistical Decision Functions, and Random Processes have been organized every three years since 1956. During the eighteen years of their existence the Prague Conference developed from a platform for presenting results obtained by a small group of researchers into a probabilistic congress, this being documented by the increasing number of participants as well as of presented papers.

The importance of the Seventh Prague Conference has been emphasized by the fact that this Conference was held jointly with the eight European Meeting of Statisticians. This joint meeting was held from August 18 to 23, 1974 at the Technical University of Prague. The Conference was organized by the Institute of Information Theory and Automation of the Czechoslovak Academy of Sciences and was sponsored by the Czechoslovak Academy of Sciences, by the Committee for the European Region of the Institute of Mathematical Statistics, and by the International Association for Statistics in Physical Sciences.

Mor than 300 specialists from 25 countries participated in the Conference. In 57 sessions 164 papers (including 17 invited papers) were read, 128 of which are published in the present two volumes of the Transactions of the Conference. Volume A includes papers related mainly to probability theory and stochastic processes, whereas the papers of Volume B concern mainly statistics and information theory.

It gives us a pleasure to express our gratitude to all who have contributed to the success of the Conference, especially to the sponsoring institutions and their representatives, and to those who read papers and delivered their manuscripts for publication in the Transactions.

Sincere thanks are due to Academician Jaroslav Kožešník, the scientific editor of these volumes, to the editorial board the members of which read all published manuscripts, and also to all employees of the Institute of Information Theory and Automation who participated in the organization of the Conference.

The Organizing Committee is very sorry that the premature death of two of its members, who participated in the preparation of this volume, prevented them from seeing the fruits of their labour. Ing. Libor Kubát died on 28th December, 1975, at the age of forty-six and Dr. Zdeněk Koutský died on 3rd June, 1976, at the age of fifty-two. Ing Kubát, as the executive editor, and Dr. Koutský, as one of the main organizers, contributed invaluably to the success of this Conference, as well as of all our previous conferences. The function of the executive editor of this volume was assumed by Dr. Milan Ullrich.

ORGANIZING COMMITTEE

EXPERIMENTAL DESIGN IN DECISION MODELS

Hans Bandemer

Freiberg

1. INTRODUCTION

The possibility of experimental design in a decision model complicates the structure of the problem. When applying the theory of experimental design, as developed by Wald [12], Blackwell [4], Le Cam [9], Kiefer [8] and others, to practical cases and in evaluating optimal decision functions some problems arise not yet treated sufficiently, as it seems. The present note contains some examples for such situations, refers results for special cases and submits proposals for generalization.

2. DECISION MODEL

Let Z be a set of states of nature and B_Z a σ-algebra of subsets from Z, forming a measurable space $[Z, B_Z]$. Further let A be a set of actions a and $[A, B_A]$ a corresponding measurable space. On $Z \times A$ let be defined a loss function $L(z, a)$, the values of which are in the set G_L of a measurable space $[G_L, B_L]$, L is assumed to be $(B_Z \times B_A, B_L)$-measurable.

If it is possible, before the choice of an $a \in A$, to realise a random element Y, defined on $[G_Y, B_Y]$ and taking values in C of a given $[C, B_C]$, the distribution of which depends on the true state $z \in Z$ of nature, then the decision may pass according to a (B_C, B_A)-measurable decision function $d(Y)$. For L is $(B_Z \times B_A, B_L)$-measurable, $L(z, d(Y))$ becomes a random element. The distribution of Y, on the condition that z is the true state of nature, may be known. Then you can evaluate, if it exists, the expectation of the random loss

$$R(z, d(.)) = E_{Y|z}L(z, d(Y))\,, \tag{2.1}$$

the risk function. With a given set D of decision functions $d \in D$ the (statistical) decision problem

$$\{A, D, R\} \tag{2.2}$$

is defined.

3. EXPERIMENTAL DESIGN

If there is the possibility to choose the random element Y from a given set, then the problem of experimental design arises, which is a decision problem, too, for being a problem of choice.

Let $\{Y(v), v \in V\}$ be a set of observable random elements, all defined on the same measurable space $[G_Y, B_Y]$. The distributions of $Y(v)$ depend on the true state $z \in Z$ of nature, but the given set V, the experimental domain, does not.

DEFINITION 1. Every arrangement

$$V_n = (v_1, \ldots, v_n)\,; \tag{3.1}$$

$$v_i \in V, \quad i = 1, \ldots, n\,; \quad n = 1, 2, \ldots,$$

of elements from V is called exact design of size n.

It is not assumed that the v_i in V_n are all distinct from each other. The set of all V_n with fixed n is denoted by V^n, furthermore let be $V^\infty = \bigcup_{n=1}^{\infty} V^n$ and B_{V^∞} an appropriate σ-algebra of subsets from V^∞, forming the measurable space $[V^\infty, B_{V^\infty}]$.

DEFINITION 2. The arrangement

$$\mathbf{Y}(V_n) = (Y(v_1), \ldots, Y(v_n)) \tag{3.2}$$

belonging to V_n, is called observation according to V_n.

Let $\mathbf{Y}(V_n)$ be a mapping into $[C(V_n), B_{C(V_n)}]$. The dependence of the chosen $a \in A$ on the result $\mathbf{y}(V_n)$ of $\mathbf{Y}(V_n)$ is given by a $(B_{C(V_n)}, B_A)$-measurable function, the decision function $d(\mathbf{Y}(V_n), V_n)$. The set $D(V_n)$ of possible decision functions depends, in general, on the chosen design V_n.

In practical cases the realization of observations according to different designs will be connected with different expenses (time, money, etc.). Therefore it may be taken for expedient to assume the loss function depending on V_n explicitly. The loss function $L_V(z, a, V_n)$ is then defined on $Z \times A \times V^\infty$. Usually it is assumed that the expenses are superposed to the loss function $L(z, a)$ and a cost function $K(V_n)$ is looked for with

$$L_V(z, a, V_n) = L(z, a) + K(V_n) \tag{3.3}$$

(v. e. g. Wald [12]). But it is not always expedient and possible to decompose decision loss and cost in this manner, maybe even impossible to measure them in the same unit.

As in the preceding paragraph the risk function, on the supposition of its existence, is formed by the expectation operator

$$R_V(z, d(\cdot, V_n), V_n) = E_{\mathbf{Y}|z} L_V(z, d(\mathbf{Y}(V_n), V_n), V_n) \,. \tag{3.4}$$

Now the risk depends on both the choosable elements V_n and $d(\cdot, V_n)$.

In order to value the expediency and goodness of the decision $(V_n \in V^\infty, d(\cdot, V_n) \in \in D(V_n))$ and to get rid of the explicit dependence on the unknown state z the risk function (3.4) (like (2.1) in the usual case) must be mapped into R^1 by a suitable functional Q. In the case $G_L = R^1$ Q is chosen usually as the expectation operator with respect to an a-priori-distribution on Z (Bayes' theory) or $Q = \max_z$ (which leads to the minimax theory).

Definition 3. The function

$$R_{VQ}(V_n, d(\cdot, V_n)) = Q R_V(z, d(\cdot, V_n), V_n) \tag{3.5}$$

is called Q-risk.

The aim of a decision theoretical treatment is the minimizing of the risk (3.5) by choosing of an optimal design V_n and an optimal decision function $d(\cdot, V_n)$. Let be V^b a subset of interest from $V^\infty : V^b \subset V^\infty$.

Definition 4. The design V_n^* and the decision function $d^*(\cdot, V_n^*)$ are called (Q, V^b)-optimal, if

$$R_{VQ}(V_n^*, d^*(\cdot, V_n^*)) = \inf_{V_n \in V^b, d(\cdot, V_n) \in D(V_n)} R_{VQ}(V_n, d(\cdot, V_n)) \,. \tag{3.6}$$

The problem (3.6) is a basic one for statistical experimental design (v. also Bandemer [1], Bandemer/Jung [2]).

Remarks. (i) The generalization of (3.1) to design measures (in the sense of Kiefer [8]) suggests itself, but it is omitted here for brevity.

(ii) Since Wald [12] the problem of experimental design is considered in the sequential decision model, too. Some generalized formulation, with regard to estimation in the regression model, may be found e.g. in Hohmann [6].

4. DECOMPOSITION THEOREMS

A difficulty of (3.6) lies in the necessity to evaluate $V_n^* \in V^b$ and $d^*(., V_n^*) \in D(V_n^*)$ simultaneously. In some applications the problem becomes more complicated, when Z has a structure $Z = P \times H$, too, where, e.g., P is a set of probability measures and H is a subspace of R^k (as it is in the regression model, v. paragraph 5).

Therefore it will be interesting to investigate on what conditions a decision problem of the form $U = \{Z_1 \times Z_2, D_1 \times D_2, R\}$ can be decomposed into two partial problems $\{Z_1, D_1, R_1\}$ and $\{Z_2, D_2, R_2\}$. For the case $Q = \max_z$ (i.e. the minimax problem) Näther [10] proved a so-called decomposition theorem, which may be quoted here for example.

THEOREM (Näther). If

(a) for each $z_2 \in Z_2$ and each $d_2 \in D_2$, where $R_1(z_1, d_1) = R((z_1, z_2), (d_1, d_2))$, $U_1 = \{Z_1, D_1, R_1\}$ is a decision problem, for which a minimax strategy $d_1^*(z_2, d_2)$ exists,

(b) in the decision problem $U_2 = \{Z_2, D_2, R_2\}$, where $R_2(z_2, d_2) = \sup_{z_1 \in Z_1} R((z_1, z_2), (d_1^*(z_2, d_2), d_2))$, exists a minimax strategy d_2^*,

(c) in U_2 a maximal strategy z_2^* with respect to d_2^* exists,

then $(d_1^*(z_2^*, d_2^*), d_2^*)$ is minimax in U.

More special decomposition theorems are due to Richter [11].

5. EXPERIMENTAL DESIGN IN THE REGRESSION MODEL

On the probability space $[G, B, P]$ a family of random variables $Y(\boldsymbol{x})$ let be defined, where $\boldsymbol{x} \in V$ and V a bounded and closed subset of R^k. The components x_i of $\boldsymbol{x}$ are the regressors and the expectation function $E\,Y(\boldsymbol{x}) = g(\boldsymbol{x})$, the existence of which is presupposed, is called response surface. A problem of regression theory is the estimation of the unknown response surface by means of a sample, the observation vector $\boldsymbol{Y}(V_n) = (Y(\boldsymbol{x}_1), \ldots, Y(\boldsymbol{x}_n))'$, where usually further informations are given,