The Annals of Statistics
1979, Vol. 7, No. 6, 1329–1332

A KIEFER-WOLFOWITZ THEOREM IN A STOCHASTIC PROCESS SETTING[1]

BY M. C. SPRUILL AND W. J. STUDDEN

Georgia Institute of Technology and Purdue University

In the regression design problem with observations which are second order processes the estimation of the mean function involves function space valued random variables. The best unbiased linear estimator of the mean function is found and an exact analogue of the Kiefer-Wolfowitz theorem in design theory is proved.

Introduction. In the theory of optimal design of experiments with scalar observations Kiefer and Wolfowitz (1960) proved that the design which maximizes the determinant of the information matrix also minimizes the maximum variance of the best linear unbiased estimator of the regression function (see Karlin and Studden (1966) or Fedorov (1972)). They also gave a convenient characterization of the optimum design. Fedorov extended these results to the case of finite-dimensional vector observations (see Fedorov (1972) Theorem 5.2.1).

In this paper the case of observations which are second order processes is investigated. This necessitates dealing with function space-valued estimators. The best linear unbiased estimator is found and a theorem which is exactly analogous to the Kiefer-Wolfowitz theorem is proved. The reader is referred to Mehra (1974) and Viort (1972) for information on optimum designs for estimation of finite-dimensional quantities related to some different stochastic process formulations. Spruill and Studden (1978) discuss some other minimax designs associated with the model of the present paper.

In order to present a more precise statement of our results let $\mathbf{f} = (f_0, f_1, \cdots, f_k)'$ be a vector of mappings from a set X onto a subset of functions on the set T. That is, for each $x \in X$, $f_j(x, \cdot)$ is a real valued function on the set T with value $f_j(x, t)$ at $t \in T$. The points $x \in X$ are possible levels of feasible experiments. For each level some experiment can be performed whose outcome is a stochastic process $\{Y(x, t) : t \in T\}$. It is assumed that the process has mean function

$$\Sigma_{j=0}^{k} \theta_j f_j(x, t), \qquad t \in T,$$

and known proper covariance kernel

$$K(s, t) = \mathrm{Cov}[Y(x, s), Y(x, t)], \qquad x \in X, s, t \in T.$$

Received July 1977; revised July 1978.

[1]Research supported by NSF Grants MCS 76-11049 A01 and MP575-08294.

AMS 1970 *subject classifications.* Primary, 62J05, 62K05; secondary, 62M99.

Key words and phrases. Stochastic process, kernel Hilbert space, function space-valued estimators, *D*-optimum designs, minimax designs.

The constants $\theta_0, \cdots, \theta_k$ are unknowns. The mean function $\Sigma_{j=0}^k \theta_j f_j(x, \cdot)$ is to be estimated on the basis of N uncorrelated observations $\{Y(x_i, t) : t \in T, i = 1, \cdots, N\}$.

Let K be as above and $H(K)$ be the associated reproducing kernel Hilbert space of functions on T with inner product $(\cdot, \cdot)_K$ (see Parzen (1959). The assumptions are as follows.

(A1) The set X can be given a topology so that one point sets are measurable. Arbitrary Borel probability measures will be admitted as possible designs. We denote the class of designs by Ξ.

(A2) The functions $\{f_j\}_{j=0}^k, f_j : X \to H(K)$ are continuous on the compact set X.

It follows from (A2) that the matrix $M(\xi)$ with i, jth entry

$$[M(\xi)]_{ij} = \int (f_i(x), f_j(x))_K \, d\xi(x)$$

exists and is well defined for each $\xi \in \Xi$. If ξ concentrates all mass at x write $M(x)$.

Associated with the experiment which takes observations $\{Y(x_i, t) : t \in T, i = 1, \cdots, N\}$ is the stochastic process Z defined on $\{x_1, \cdots, x_N\} \times T = \Gamma$ by $Z(x_i, t) = Y(x_i, t)$. Let B be the covariance kernel of Z, $H(B)$ be the reproducing kernel Hilbert space of functions on Γ generated by B, and $L_2[Z(\gamma) : \gamma \in \Gamma]$ be the usual space of linear random variables defined on the process (see Parzen (1959) and Spruill and Studden (1978)). If we denote the congruence which takes $B(\cdot, \gamma)$ into $Z(\gamma)$ by $\langle \cdot , \rangle_B$, i.e., $Z(\gamma) = \langle Z, B(\cdot, \gamma)\rangle_B$, then we may state our results as follows. The best linear unbiased estimator of the mean $\theta'\mathbf{f}(x)$ of the stochastic process $\{Y(x, t) : t \in T\}$ is

$$\hat{m} = N^{-1}(\langle Z, f_0\rangle_B, \cdots, \langle Z, f_k\rangle_B)M^+(\xi)\mathbf{f}(x), \tag{1}$$

where ξ is the design measure associated with the experiment. Furthermore, indicating the dependence of the estimator defined in (1) by writing $\hat{m}(\xi)$ and defining

$$N^{-1}d(x, \xi) = E\|\hat{m}(\xi) - \theta'\mathbf{f}(x)\|_K^2,$$

one has that $d(x, \xi) = \text{tr}\{M^+(\xi)M(x)\}$ so that the conditions

(i) ξ^* maximizes $\det[M(\xi)]$;

(ii) ξ^* minimizes $\sup_{x \in X} d(x, \xi)$;

(iii) $\sup_X d(x, \xi^*) = k + 1$,

are equivalent.

2. Linear estimators. Linear estimators taking values in an arbitrary Hilbert space may be defined in a manner analogous to Parzen's (1959) for real valued estimators. However, for the sake of simplicity, we shall deal only with the case that they are $H(K)$-valued and $H(K)$ is separable. In this case the space of linear $H(K)$-valued random variables defined on Z consists of elements of the form $\Sigma_{j\geq 1}\langle Z, g_j\rangle_B \phi_j$, where $\{\phi_j\}$ is a complete orthonormal system for $H(K)$, g_j are each in $H(B)$, $\Sigma_{j\geq 1}\langle Z, g_j\rangle_B \phi_j$ is in $H(K)$ with probability one, and $\Sigma_{j\geq 1}E[\langle Z, g_j\rangle_B^2] < \infty$.

In this context an unbiased estimator U of $\theta'\mathbf{f}(x)$ is one which satisfies the following equivalent conditions.

$$\|E_\theta[U] - \theta'\mathbf{f}(x)\|_K = 0$$

$$E_\theta[U(t)] = \Sigma_{j=0}^{k}\theta_j f_j(x, t) \qquad \text{for} \quad t \in T.$$

A "best" linear unbiased estimator U of $\theta'\mathbf{f}(x)$ is one which is unbiased and for any other unbiased estimator V one has

$$E_\theta\|V - \theta'\mathbf{f}(x)\|_K^2 \geqslant E_\theta\|U - \theta'\mathbf{f}(x)\|_K^2.$$

THEOREM. *If an unbiased estimator of $\theta'\mathbf{f}(x)$ exists, then the estimator in* (1) *above is the best unbiased estimator.*

PROOF. Using extensions of Parzen (1959) or Lemma 2.2 of Spruill and Studden (1978) it follows that the estimator in (1) is unbiased and for any finite set $T_n = \{t_1, \cdots, t_n\} \subset T$ $(\Sigma a_j K(\cdot, t_j), \hat{m})$ is the volume of $(\Sigma a_j K(\cdot, t_j), \theta'\mathbf{f}(x))$. Let $U = \hat{m}\text{-}\theta'\mathbf{f}(x)$ and V be the corresponding quantity for any unbiased estimator. For $h \in H(K)$ of the form $h = \Sigma a_j K(\cdot, t_j)$

$$E(h, U)^2 = a' \operatorname{Cov}(U_{T_n})a.$$

Therefore $E(h, U)^2 \leqslant E(h, V)^2$ holds for a dense subset of $h \in H(K)$. From this it can readily be deduced that $E\|U\|_K^2 \leqslant E\|V\|_K^2$.

The theorem has been proved here only for $H(K)$ separable. The theorem has been proved for general $H(K)$ spaces in Spruill and Studden (1977).

3. *D*-optimum designs. In the previous section the designs had rational probabilities at their finite number of support points. The following definitions are made for arbitrary $\xi \in \Xi$.

DEFINITION. The mean $\mathbf{f}'(x)\theta$ is said to be estimable with respect to the design ξ if $\{(y, \mathbf{f}(x))_K : y \in H(K)\} \subseteq \mathfrak{R}[M(\xi)]$.

DEFINITION. Let $d(x, \xi) = \operatorname{tr}\{M^+(\xi)M(x)\}$ if $\mathbf{f}'(x)\theta$ is estimable with respect to ξ and $+\infty$ otherwise.

DEFINITION. The design ξ^* is said to be a minimax design if

$$\inf_{\xi\in\Xi}\sup_{x\in X} d(x, \xi) = \sup_{x\in X} d(x, \xi^*).$$

In this section an additional assumption is made.

(A3) There is a design ξ such that $\int\|\Sigma_{j=0}^{k} a_j f_j(x)\|_K^2\, d\xi(x) = 0$ if and only if $a_0 = a_1 = \cdots = a_k = 0$.

THEOREM 3.1. *Under* (A1)–(A3) *the conditions*

(i) *ξ^* maximizes $|M(\xi)|$;*

(ii) *ξ^* minimizes $\sup_x d(x, \xi)$;*

(iii) $\sup_x d(x, \xi^*) = k + 1$,

are equivalent. The set Γ of all ξ^* satisfying these conditions is convex and closed and $M(\xi^*)$ is the same for all $\xi^* \in \Gamma$.

PROOF. If there is an $x \in X$ such that $\mathbf{f}'(x)\theta$ is not estimable $d(x, \xi) = +\infty$. By (A3) there is a design ξ_0 for which $|M(\xi_0)| \neq 0$. Thus $\mathcal{R}[M(\xi_0)] = R^{k+1}$ and $\mathbf{f}'(x)\theta$ is estimable for all $x \in X$. Therefore attention may be restricted to those $\xi \in \Xi$ for which $|M(\xi)| > 0$. Except that the matrices $M(\xi)$ differ, the remainder of the proof is exactly as it appears in Kiefer (1960), Fedorov (1972), or Karlin and Studden (1966). □

For finding D-optimal designs the iterative process of Fedorov, for example, can be described and shown to converge in exactly the same manner as in Fedorov (1972), Theorem 5.22.

Acknowledgment. The authors would like to thank the referees for their helpful comments.

REFERENCES

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

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

[3] KIEFER, J. (1960). Optimum experimental designs V. *Proc. Fourth Berkeley Symp. Math. Statist. Probability* **1** 381–405.

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

[5] MEHRA, R. K. (1974). Optimal input signals for parameter estimation in dynamic systems-survey and new results. *IEEE Trans. Automatic Control* **AC-19** 753–768.

[6] PARZEN, E. (1961). An approach to time series analysis. *Ann. Math. Statist.* **32** 951–989.

[7] SPRUILL, CARL and STUDDEN, W. J. (1977). A Kiefer-Wolfowitz theorem in a stochastic process setting. Technical report, School of Math., Georgia Inst. Tech.

[8] SPRUILL, CARL and STUDDEN, W. J. (1978). Optimum designs when the observations are second order processes. *J. Multivariate Anal.* **8** 153–172.

[9] VIORT, B. (1972). *D*-optimal designs for dynamic models, part I, Theory; part II, Applications. Technical reports 314 and 316, Dept. Statist., Univ. Wisconsin.

SCHOOL OF MATHEMATICS
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GEORGIA 30332

DEPARTMENT OF STATISTICS
PURDUE UNIVERSITY
LAFAYETTE, INDIANA 47907