*The Annals of Statistics*
1979, Vol. 7, No. 2, 432–445

# ON DYNAMIC PROGRAMMING AND STATISTICAL DECISION THEORY[1]

BY MANFRED SCHÄL

*University of Bonn*

The main aim of the present work is to establish connections between the theory of dynamic programming and the statistical decision theory. The paper deals with a nonMarkovian dynamic programming decision model that includes Markovian decision models and Markov renewal decision models as special cases. The analysis is based on the total cost criterion where the convergence condition on the expected total cost is such that the discounted and the negative (unbounded) case are included. The striking feature of the present model is the fact that the law of motion is not completely known, which leads to a treatment of the model by the approach of statistical decision theory. The assumptions of the present paper are discussed for a sequential statistical decision problem.

**1. Introduction.** The paper presents a dynamic statistical decision model that generalizes both the decision model of the theory of dynamic programming and the sequential (or nonsequential) model of statistical decision theory. The model is based on a nonstationary nonMarkovian dynamic programming model. The statistical aspect arises from the fact that the model allows for situations where the law of motion $q$ is not completely known. As usual in a statistical decision problem, a leading feature will be the assumption that the law of motion is merely known to be an element of a given class $\{q^{\vartheta}, \vartheta \in \Theta\}$. Thus the total expected cost $R(\vartheta, \pi)$ depends not only on the policy $\pi \in \Delta$ but also on the parameter $\vartheta \in \Theta$ and is, therefore, called the risk function.

The main aim of the paper is to provide sufficient conditions about the action spaces, the class of admissible laws of motion, and the cost functions for the following properties.

1. compactness of the space of policies $\Delta$—with respect to an appropriate topology;
2. a sort of convexity of $\Delta$;
3. lower semicontinuity of $\pi \to R(\vartheta, \pi)$ on $\Delta$ for every $\vartheta \in \Theta$;
4. (equi-) continuity of $\vartheta \to R(\vartheta, \pi)$ on $\Theta$, $\pi \in \Delta$.

These properties are known to supply the foundation for a series of theorems in statistical decision theory, some of which are given in the present paper. More precisely, we prove

(a) the existence of minimax and Bayes policies;

---

Received April 1976; revised August 1977.

[1]Work supported by 'Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 72 at the University of Bonn.'

*AMS* 1970 *subject classifications.* Primary 62C05, 90C40; secondary 62M99, 90C45.

*Key words and phrases.* Compactness and convexity of the space of policies, lower semicontinuity of the risk function, continuity and stability of the optimization procedure, strict determinateness, least favourable a priori distributions, minimax and Bayes policies.

(b) the strict determinateness of the decision problem viewed as a zero sum two-person game;
(c) the convergence of the minimal total expected cost from finite stage play to the minimal total expected cost from infinite stage play;
(d) the existence of a least favourable a priori distribution;
(e) the continuity of $\vartheta \to \inf_{\pi \in \Delta} R(\vartheta, \pi)$.

The compactness of the space of policies was already proved by Schäl (1975b) for the case where the law of motion is completely known. However, the topology used there depends heavily on the law of motion. Hence this result does not immediately apply to the present more general situation. This difficulty is overcome in this paper by a reduction of the dynamic statistical decision model to a dynamic decision model where the law of motion is known and only the cost functions may depend on the parameter $\vartheta \in \Theta$. After this reduction, the analysis of the present paper can be based upon the results of Schäl (1975b). In Schäl (1976) an application to inventory models is given.

**2. The general dynamic statistical decision model.** We use $\mathbb{N}$ to denote the set of the positive integers and we use $\mathbb{R}$ (resp. $\hat{\mathbb{R}}$) to denote the set of the real numbers (resp. augmented by the point $+\infty$). For any set $S$ endowed with some $\sigma$-algebra, $\mathcal{P}(S)$ stands for the set of all probability measures on $S$.

The basic decision model is given by a tuple $(S_n, A_n, K_n, \Theta, \lambda_n, (1_n^\vartheta, \vartheta \in \Theta), (c_n^\vartheta, \vartheta \in \Theta); n \in \mathbb{N})$ having the following meaning:

(i) $S_n$ stands for the *state space* at time $n$ and is assumed to be a standard Borel space, i.e., $S_n$ is a nonempty Borel subset of a Polish space and is endowed with the $\sigma$-algebra of Borel subsets of $S_n$.
(ii) $A_n$ is the *space of actions* available at time $n$ and is assumed to be a standard Borel space. We write $H_n = S_1 \times A_1 \times \cdots \times S_n$ and $H_\infty = S_1 \times A_1 \times S_2 \times A_2 \times \cdots$.
(iii) $K_n$ is a measurable subset of $H_n \times A_n$ and specifies the set of admissible histories. It is assumed that $K_n \subset K_{n-1} \times S_n \times A_n$ and that for every $h \in K_{n-1} \times S_n$ the section of $K_n$ at $h$ is non-empty, where $K_0 \times S_1 = S_1$.

We write $K_\infty = \bigcap_n K_n \times S_n \times A_n \times \cdots$. A *strategy of nature*—usually called *law of motion*—is a sequence $q = (q_n)$, where $q_1 \in \mathcal{P}(S_1)$ is the initial distribution and $q_n : K_{n-1} \to \mathcal{P}(S_n)$, $n > 1$, is a transition probability. Then $q_n(h, \cdot)$ is the conditional distribution of the state of the system at time $n$ given the admissible history $h \in K_{n-1}$. We write $\mathfrak{q}$ for the set of all laws of motion. A *strategy of the statistician* – usually called *policy*—is defined as a sequence $\pi = (\pi_n)$ of transition probabilities $\pi_n : K_{n-1} \times S_n \to \mathcal{P}(A_n)$, such that $\pi_n(h, \cdot)$ assigns probability one to the section of $K_n$ at $h \in K_{n-1} \times S_n$, $n \in \mathbb{N}$. We write $\Delta$ for the set of all policies.

REMARK 2.1. We may also allow for situations where the statistician is the first to take a decision. This situation is included by the present model upon defining $S_1$ as a singleton.

An application of the theorem of Ionescu-Tulcea (cp. Neveu (1965)) yields that any $q \in \mathfrak{q}$ and $\pi \in \Delta$ uniquely define a probability measure $P_\pi^q = q_1 \otimes \pi_1 \otimes q_2 \otimes \pi_2 \otimes \cdots$ on the product space $H_\infty$ and thus a random process $(\zeta_1, \alpha_1, \zeta_2, \alpha_2, \cdots)$ (cp. Hinderer (1970) page 80) where $\zeta_n$ and $\alpha_n$ describe the state of the system and the action at time $n$, respectively. Then $P_\pi^q(K_n \times S_{n+1} \times A_{n+1} \times \cdots) = P_\pi^q(K_\infty) = 1$, $n \in \mathbb{N}$, $q \in \mathfrak{q}$, $\pi \in \Delta$.

(iv) The *parameter space* $\Theta$ is a nonempty set endowed with some $\sigma$-algebra. We shall assume that there is given some subset $\{q^\vartheta; \vartheta \in \Theta\}$ of $\mathfrak{q}$ which is known to contain the 'true' law of motion.

For some problems in statistical decision theory, e.g., when dealing with minimax policies, there is no need of a $\sigma$-algebra on $\Theta$. In such situations, $\Theta$ may be thought of as a set endowed with the power set of $\Theta$.

We shall assume that every family $\{q_n^\vartheta(h, \cdot), \vartheta \in \Theta\} \subset \mathfrak{P}(S_n)$, $h \in K_{n-1}$, is dominated by a measure $\lambda_n$ depending on $h$ with likelihood function $1_n^\vartheta(h, \cdot)$.

(v) $(\lambda_n)$ supplies the *dominating measures*. We assume that $(\lambda_n) \in \mathfrak{q}$.

(vi) $(\vartheta, h) \to 1_n^\vartheta(h)$ are the *likelihood functions* and are assumed to be nonnegative measurable functions on $\Theta \times K_{n-1} \times S_n$ such that the law of motion $q^\vartheta$ introduced in (iv) is given through $q_1^\vartheta(ds_1) = 1_1^\vartheta(s_1)\lambda_1(ds_1)$, and $q_n^\vartheta(h, ds_n) = 1_n^\vartheta(h, s_n)\lambda_n(h, ds_n)$, $h \in K_{n-1}$, $n > 1$.

According to (v), $\lambda_n(h)$ is a probability measure for every $h \in K_{n-1}$. If instead $\lambda_n(h)$ is given as a $\sigma$-finite measure for every $h \in K_{n-1}$, then there exists a probability measure $\lambda_n^*(h)$ depending on $h \in K_{n-1}$ such that $\lambda_n(h)$ is dominated by $\lambda_n^*(h)$ with respective finite density $f_n(h, \cdot) = d\lambda_n(h)/d\lambda_n^*(h)$. Then $\lambda_n$ can be replaced by $\lambda_n^*$ and $1_n^\vartheta(h, s_n)$ by $1_n^\vartheta(h, s_n)f_n(h, s_n)$ and we have the situation of (v) and (vi). This consideration neglects the fact that $f_n$ and $\lambda_n^*$ should depend measurably on $h$. However, if $\lambda_n$ satisfies sufficient measurability conditions in $h$, then $\lambda_n^*$ and $f_n$ can be chosen to depend measurably on $h$. Since $q_n^\vartheta$ measurably depends on $\vartheta$ for $n \in \mathbb{N}$, we may infer from the theorem of Ionescu-Tulcea that for any $\pi \in \Delta : (\vartheta, B) \to P_\pi^\vartheta(B)$ is a transition probability from $\Theta$ into $H_\infty$.

(vii) The *cost function* or *loss function* for the $n$th period $(\vartheta, h) \to c_n^\vartheta(h)$ is a measurable function from $\Theta \times K_n \times S_{n+1}$ to $\hat{\mathbb{R}}$ bounded from below.

For $\vartheta \in \Theta$, $\pi \in \Delta$, $m, n \in \mathbb{N}$ let us define

$$R_m^n(\vartheta, \pi) = \Sigma_{t=m}^n \int c_t^\vartheta dP_\pi^\vartheta$$

where $R_m^n = 0$ for $m > n$. More precisely, we should write $\int c_t^\vartheta o (\zeta_1, \alpha_1, \cdots, \zeta_{t+1}) dP_\pi^\vartheta$ instead of $\int c_t^\vartheta dP_\pi^\vartheta$. But we agree to look at any function $f$ defined on $H_n$ as a function defined on $H_\infty$ which depends on the first $2n - 1$ coordinates only. Then $R_m^n$ is a well-defined function from $\Theta \times \Delta$ to $\hat{\mathbb{R}}$ bounded from below and measurable in $\vartheta \in \Theta$. We set

$$z_n = \inf_{\pi \in \Delta, \vartheta \in \Theta, t \geqslant n} R_{n+1}^t(\vartheta, \pi).$$

Then

$$R_1^n \geqslant R_1^m + z_m \quad \text{for} \quad m \leqslant n. \tag{2.1}$$

Since $z_n \leqslant R^n_{n+1} = 0$, $z_n$ is nonpositive. *Throughout this paper we impose*

CONDITION $(A)$. $Z_n \to 0$ as $n \to \infty$.

This condition is satisfied in the *discounted case* (cp. Blackwell (1965)) and in the *negative case* (cp. Strauch (1966)) where the cost functions $c^\vartheta_n$ are nonnegative. Condition $(A)$ implies that the *risk function*

$$R(\vartheta, \pi) = \Sigma^\infty_{t=1} \int c^\vartheta_t dP^\vartheta_\pi = \lim_{n\to\infty} R^n_1(\vartheta, \pi) \tag{2.2}$$

is a well-defined function from $\Theta \times \Delta$ into $\hat{\mathbb{R}}$ bounded from below and measurable in $\vartheta \in \Theta$. The existence of the limit in (2.2) can be justified by Lemma 4.1 in Schäl (1975a). Hence we can extend the domain of $R$ according to

$$R(\mu, \pi) = \int R(\vartheta, \pi)\mu(d\vartheta), \qquad \mu \in \mathcal{P}(\Theta) \ \pi \in \Delta,$$

where we look at $\vartheta$ as a degenerate probability measure giving mass one to the point $\vartheta$. Similarly, we write

$$R^n_1(\mu, \pi) = \int R^n_1(\vartheta, \pi)\mu(d\vartheta), \qquad \mu \in \mathcal{P}(\Theta), \pi \in \Delta.$$

Using (2.1) one may conclude from Lemma 4.3 in Schäl (1975a) that

$$R(\mu, \pi) = \Sigma^\infty_{t=1} \int \left[ \int c^\vartheta_t dP^\vartheta_\pi \right] d\mu = \lim_{n\to\infty} R^n_1(\mu, \pi). \tag{2.3}$$

**3. Optimality criterion.** In the present paper we are concerned with the following concept of optimality (cp. Bierlein (1963), (1967), Bunke (1964), Menges (1966)). Let $\Gamma$ be a subset of $\Delta$ and $\Lambda$ be a subset of $\mathcal{P}(\Theta)$. Then $\pi^* \in \Gamma$ is said to be $\Lambda$-*optimal in* $\Gamma$—we write $\pi^* \in \Gamma^*(\Lambda)$—if

$$\sup_{\mu\in\Lambda} R(\mu, \pi^*) = \inf_{\pi\in\Gamma} \sup_{\mu\in\Lambda} R(\mu, \pi).$$

Obviously for any $\mu \in \mathcal{P}(\Theta)$ we have $\pi^* \in \Delta^*(\{\mu\})$, if and only if $\pi^*$ is a *Bayes policy against the* a priori *distribution* $\mu$, i.e.,

$$R(\mu, \pi^*) = \inf_{\pi\in\Delta} R(\mu, \pi).$$

And $\pi^* \in \Delta^*(\mathcal{P}(\Theta))$ if and only if $\pi^*$ is a *minimax policy*, i.e.,

$$\sup_{\vartheta\in\Theta} R(\vartheta, \pi^*) = \inf_{\pi\in\Delta} \sup_{\vartheta\in\Theta} R(\vartheta, \pi),$$

where use is made of

$$\sup_{\vartheta\in\Theta} R(\vartheta, \pi) = \sup_{\mu\in\mathcal{P}(\Theta)} R(\mu, \pi), \qquad \pi \in \Delta. \tag{3.1}$$

Further, upon defining for any $0 \leqslant \rho_0 \leqslant 1$, $\mu_0 \in \mathcal{P}(\Theta)$, $\Lambda = \{\rho_0\mu_0 + (1 - \rho_0)\nu, \nu \in \mathcal{P}(\Theta)\}$, we obtain the optimality criterion of Hodges and Lehmann (1952) where

$$\sup_{\mu\in\Lambda} R(\mu, \pi) = \rho_0 R(\mu_0, \pi) + (1 - \rho_0)\sup_{\vartheta\in\Theta} R(\vartheta, \pi).$$

If $\Theta = \bigcup \Theta_i$ is a measurable partition and we choose, for some $p_i \geqslant 0$ with $\Sigma p_i = 1$, $\Lambda = \{\mu \in \mathcal{P}(\Theta); \mu(\Theta_i) = p_i\}$, we obtain the optimality criterion of Menges (1966) where

$$\sup_{\mu\in\Lambda} R(\mu, \pi) = \Sigma p_i \cdot \sup_{\vartheta\in\Theta_i} R(\vartheta, \pi).$$

The optimality criterion of this paper is based on the total cost criterion. The interesting paper of Mandl (1974) is also concerned with a decision model where the transition law is not completely known to the statistician. However, Mandl uses the average cost criterion and a sort of uniform optimality.

In order to guarantee the existence of $\Lambda$-optimal policies in $\Delta$ it suffices to find a topology on $\Delta$ such that $\Delta$ is compact and the mappings $\pi \to R(\mu, \pi)$, $\mu \in \Lambda$, are lower semicontinuous (Bunke (1964) Satz 1). (Henceforth we write l.s.c. instead of lower semicontinuous.) In the case where the state spaces are countable discrete spaces, several authors used the topology of pointwise convergence (cp. Derman (1965), Rieder (1973), Wessels (1968), Wald (1950)). Martin ((1967), proof of Theorem 3.2.1) proposed another topology for the discrete case. In the case where the state spaces are Borel subsets of the real line, Wald (1950) attacked the problem by introducing the notion of regular convergence for a sequential statistical decision problem.

We shall prefer a different approach guided by the observation that it is sufficient to find a factorization

$$\Delta \to \phi \to \hat{\mathbb{R}}^{\Theta}$$
$$\pi \to \varphi_\pi \to \tilde{R}(\cdot, \phi_\pi) = R(\cdot, \pi)$$

of the mapping $\pi \to R(\cdot, \pi)$ such that $\phi$ is better structured than $\Delta$ and it is easier to find an appropriate topology on $\phi$ than on $\Delta$. Of course, every topology on $\phi$ induces a topology on $\Delta$ through the mapping $\pi \to \varphi_\pi$.

For a sequential decision problem, LeCam (1955) found a representation of the decision functions by families of linear mappings. He assumed, however, that (in the terminology of the present paper) the loss functions $c_n^{\vartheta}(s_1, a_1, \cdots, s_{n+1})$ are linear combinations of functions $w_n(\vartheta, a_1, \cdots, a_n) \cdot h_n(\vartheta, s_1, \cdots, s_{n+1})$. This assumption—though being fulfilled for the usual sequential procedures—is very restrictive for dynamic programming procedures.

During the revision of the present paper, the author became aware of the paper of Brown (1977) building on the approach of LeCam. Brown gets rid of LeCam's separability assumption on the loss functions and generalizes the model of sequential analysis such that it covers the dynamic programming model. The conclusions of Brown and those of this paper concerning the compactness of the space of policies are similar; however, Brown's hypotheses do not imply, and are not implied by, the assumptions of this paper.

In the present paper, we shall use a factorization of the risk function such that $\phi$ is a subset of $\mathcal{P}(H_\infty)$. For that purpose $\mathcal{P}(H_\infty)$ is endowed with the so-called ws$^\infty$-topology as defined in Schäl (1975b) and $\phi$ is endowed with the relativization of the ws$^\infty$-topology.

**4. Some measure theoretic requisites.** We list some requisite notations and relations. Let $S$ and $A$ be standard Borel spaces. Let $\mathcal{Q}(S, A)$ denote the set of Carathéodory functions, i.e., the set of bounded measurable real-valued functions $u$ on $S \times A$ such that every $S$-section $u(s, \cdot)$ is continuous on $A$. Let $\mathfrak{C}(A)$ denote the set of nonempty compact subsets of $A$. For a mapping $\psi : S \to \mathfrak{C}(A)$, we define, as usual, $\text{graph}(\psi) = \{(s, a) \in S \times A;\ a \in \psi(s)\}$.

The following result can be proved as Proposition 11.6 in Schäl (1975a) where the measurability of the sup-function used there can be justified by results of Brown and Purves (1973) or Himmelberg, Parthasarathy, and Van Vleck (1976). See also Brown (1977) Theorem 3.10, Kertz (1977) Lemma 1.19, Schäl (1977).

(4.1) For any $\psi : S \to \mathfrak{C}(A)$ such that graph($\psi$) is a measurable subset of $S \times A$ and for any $u$ : graph($\psi$) $\to \hat{\mathbb{R}}$, the following statements are equivalent:

(a) $u(s, a)$ depends measurably on $(s, a)$ and l.s.c.ly on $a$ and $u$ is bounded from below;

(b) there exists an increasing sequence $(u_n)$ in $\mathfrak{L}(S, A)$ such that the restriction of $\sup_n u_n$ to graph($\psi$) coincides with $u$.

The *s-topology* on $\mathfrak{P}(S)$ is the coarsest topology rendering the mappings $P \to \int f dP$ continuous for every bounded measurable function $f$ on $S$. Thus, a mapping $\mu : A \to \mathfrak{P}(S)$ is $s$-continuous if and only if for such functions $f : a \to \int f(s)\mu(a, ds)$ is continuous. The *$ws^\infty$-topology* on $\mathfrak{P}(H_\infty)$ is the coarsest topology for which the mappings $P \to \int f dP$ are continuous for every $f : H_n \to \mathbb{R}$ contained in $\mathfrak{L}(S_1 \times \cdots \times S_n, A_1 \times \cdots \times A_{n-1})$ and for every $n \in \mathbb{N}$.

**5. Compactness and continuity conditions (B).** In this section we shall formulate conditions which will turn out to be sufficient for the existence of optimal policies.

CONDITION (B). For every $n \in \mathbb{N}$

(B1) for every $(s_1, \cdots, s_n) \in S_1 \times \cdots \times S_n$, the section $K_{n(s_1, \cdots, s_n)}$ at $(s_1, \cdots, s_n)$ is compact;

(B2) $\lambda_n$ depends $s$-continuously and $1_n^\vartheta$ depends l.s.c.ly on the actions, i.e., $\lambda_{n+1}(s_1, \cdot, \cdots, s_n, \cdot) : K_{n(s_1, \cdots, s_n)} \to \mathfrak{P}(S_{n+1})$ is $s$-continuous; $1_{n+1}^\vartheta(s_1, \cdot, s_2, \cdots, s_{n+1})$ is l.s.c. on $K_{n(s_1, \cdots, s_n)}$ for every $s_i \in S_i$, $1 \leqslant i \leqslant n + 1$;

(B3) $c_n^\vartheta$ depends l.s.c.ly on the actions.

Obviously, Condition (B) is always satisfied if $S_n$ is a countable discrete space and $A_n$ is a finite set for $n \in \mathbb{N}$. Hence, the results of this paper generalize results of Wessels (1968) and Rieder (1973).

REMARK 5.1. The assumption on $1_n^\vartheta$ implies that $1_n^\vartheta$ depends continuously in measure on the action. More precisely, one can prove the following proposition: let $\lambda$ be a measure on any space $S$ endowed with a $\sigma$-algebra $\mathfrak{S}$ and let $1(a, \cdot)$ be probability densities with respect to $\lambda$ for each $a$ where $a$ runs through any metric space. If 1 depends l.s.c.ly on $a$ then each of the following equivalent statements holds as $a \to a^0$:

(a) $\sup_{F \in \mathfrak{S}} |\int_F [1(a, \cdot) - 1(a^0, \cdot)] d\lambda| \to 0$;

(b) $\int |1(a, \cdot) - 1(a^0, \cdot)| \, d\lambda \to 0$;

(c) $\lambda(|1(a, \cdot) - 1(a^0, \cdot)| > \delta) \to 0$ for every $\delta > 0$.

REMARK 5.2. The assumption on $1_n^\vartheta$ remain valid if $1_n^\vartheta$ is multiplied by a function depending only on the states. Such a multiplication may be necessary if one replaces a $\sigma$-finite measure $\lambda_n$ by a probability measure.

REMARK 5.3. One can give assumptions based on the weak topology rather than on the $s$-topology. Then, in Section 7 below, one will refer to Theorem 5.6 instead of Theorem 6.6 in Schäl (1975b) (cp. also Nowak (1975), Kertz and Nachman (1977a)). However, the only advantage gained thereby will concern the assumptions on $\lambda_n$. All other assumptions have to be strengthened. Since in any application of the theory $\lambda_n$ is likely to not depend on the histories, it does not seem worthwhile to give alternative conditions on $\lambda_n$.

**6. Example: the sequential statistical decision model.** We consider a sequential statistical decision problem where a stochastic process $(\zeta_n, n \in \mathbb{N})$ is observed. For the sake of simplicity of expression we shall consider only such experiments in which only one observation is made at each stage. Let $S_n$ denote the state space at time $n$ and suppose that there is given a $\sigma$-finite measure $\mu_n$ on $S_n$ and nonnegative measurable functions $f_n^{\vartheta}(s_1, \cdots, s_n)$ on $\Theta \times S_1 \times \cdots \times S_n$ such that

$$(\vartheta, s_1, \cdots, s_{n-1}, B) \to \int_B f_n^{\vartheta}(s_1, \cdots, s_{n-1}, s_n)\mu_n(ds_n)$$

is a regular conditional probability of $\zeta_n$ given $(\vartheta, \zeta_1, \cdots, \zeta_{n-1})$.

We write $A_n = A_n^t \cup \{e_n\}$ where $A_n^t$ is the set of terminal decisions available after $n$ observations and $e_n$ represents the decision to observe $\zeta_{n+1}$. Further there are given nonnegative measurable functions $b_n$ on $\Theta \times S_1 \times \cdots \times S_n$ and $L_n$ on $\Theta \times S_1 \times \cdots \times S_n \times A_n^t$ specifying the cost of observing $\zeta_{n+1}$ and the loss implied by some terminal decision after $n$ observations, respectively. If the statistician is allowed to take a terminal decision without any observation, we have the situations of Remark 2.1 and choose $S_1$ as a singleton.

The sequential statistical decision problem is a special dynamic statistical decision problem upon setting $S_n$ and $A_n$ as above,

$$\begin{aligned} c_n^{\vartheta}(s_1, a_1, \cdots, s_n, a_n) &= b_n(\vartheta, s_1, \cdots, s_n) && \text{if } a_i = e_i, \quad 1 \leqslant i \leqslant n; \\ &= L_n(\vartheta, s_1, \cdots, s_n, a_n) && \text{if } a_i = e_i, \quad 1 \leqslant i < n,\ a_n \in A_n^t; \\ &= 0 \text{ otherwise}; \end{aligned}$$

$1_n^{\vartheta}(s_1, a_1, \cdots, s_n) = f_n^{\vartheta}(s_1, s_2, \cdots, s_n)\, d\mu_n/d\lambda_n(s_n)$ where $\lambda_n$ is any probability measure that dominates $\mu_n$. In a similar way, every stopping problem or every stopped decision process can be formulated by use of a dynamic decision model (cp. Rieder (1975b)).

Since $c_n^{\vartheta}$ is nonnegative for $n \in \mathbb{N}$, Condition (A) is satisfied. Condition (B) is satisfied if for every $n \in \mathbb{N}$ : $A_n^t$ is a compact metric space and $L_n$ depends l.s.c.ly on the terminal action. Of course, $S_n$ should be a standard-Borel space, $A_n$ should be considered as topological sum of $A_n^t$ and the singleton $\{e_n\}$, and $K_n$ should be chosen as $H_n \times A_n$.

The risk function depends on the policy only through the stopping rule and the terminal decision rule. The model of the present paper allows for stopping times which are not necessarily finite. If it is assumed as usual that the statistician pays an infinite amount if he does not reach a terminal decision in a finite number of steps, he may restrict himself to a.s. finite stopping times.

**7. Results under condition (B).** Throughout this section it is assumed that Condition (B) is satisfied.

*7.1. Compactness of the space of policies and lower semi-continuity of the risk function.* Since $\lambda = (\lambda_n)$ is an element of $\mathfrak{q}$, $P_\pi^\lambda$ is defined. The factorization of $\pi \to R(\cdot, \pi)$ described in Section 3 will be chosen such that $\varphi_\pi = P_\pi^\lambda$ and $\phi = \Pi^\lambda$ where we write

$$\Pi^q = \{P_\pi^q;\ \pi \in \Delta\} \qquad \text{for any} \quad q \in \mathfrak{q}. \tag{7.1}$$

Such a factorization is possible because of the following relations:

$$\int c_n^\vartheta dP_\pi^\vartheta = \int \tilde{c}_n^\vartheta dP_\pi^\lambda, \tag{7.2}$$

where

$$\tilde{c}_n^\vartheta(s_1, \cdots, a_n, s_{n+1}) =$$
$$1_1^\vartheta(s_1) \cdots 1_{n+1}^\vartheta(s_1, \cdots, a_n, s_{n+1}) c_n^\vartheta(s_1, \cdots, a_n, s_{n+1}).$$

$$R(\vartheta, \pi) = \Sigma_n \int \tilde{c}_n^\vartheta dP_\pi^\lambda = \tilde{R}(\vartheta, P_\pi^\lambda). \tag{7.3}$$

$$R(\mu, \pi) = \Sigma_n \int \tilde{c}_n^\mu dP_\pi^\lambda = \tilde{R}(\mu, P_\pi^\lambda) \tag{7.4}$$

where $\tilde{c}_n^\mu(s_1, \ldots, a_n, s_{n+1}) = \int c_n^\vartheta(s_1, \ldots, a_n, s_{n+1})\mu(d\vartheta)$.

THEOREM 7.1. *Whatever $\mu \in \mathcal{P}(\Theta)$, the function $P \to \tilde{R}(\mu, P)$ is* l.s.c. *on $\Pi^\lambda$ in the* ws$^\infty$*-topology.*

PROOF. For a proof of the lower semicontinuity of $P \to \int \tilde{c}_n^\mu dP$ on $\Pi^\lambda$, there is no loss of generality in assuming that $c_n^\vartheta \geq 0$. This may be seen from the fact that $c_n^\vartheta \geq -M$ for some $M \geq 0$ and $\int \tilde{c}_n^\mu dP_\pi^\lambda = \int \tilde{c}_n^\mu[M]\, dP_\pi^\lambda - M$ where $c_n^\mu[M] = \int 1_1^\vartheta \cdots 1_{n+1}^\vartheta(c_n^\vartheta + M)\mu(d\vartheta)$. Now, it is clear that $\tilde{c}_n^\vartheta$ depends l.s.c.ly on the actions. Hence, by Fatou's lemma, $\tilde{c}_n^\mu$ depends l.s.c.ly on the actions. Thus we know from (4.1) that $\tilde{c}_n^\mu$ admits an extension defined on $H_{n+1}$ which can be written as the limit of an increasing sequence of functions in $\mathfrak{L}(S_1 \times \cdots \times S_{n+1}, A_1 \times \cdots \times A_n)$. Therefore, $P \to \int \tilde{c}_n^\mu\, dP$ is l.s.c. As a consequence, $P \to \Sigma_{i=1}^n \int \tilde{c}_i^\mu\, dP$ is l.s.c. Further, by (2.1) and (7.2) we have the inequality $\Sigma_{i=1}^n \int \tilde{c}_i^\mu\, dP \geq \Sigma_{i=1}^m \int \tilde{c}_i^\mu\, dP + z_m$ for $m \leq n$, $P \in \Pi^\lambda$. Now, an appeal to (7.4) and Proposition 10.1.1 in Schäl (1975a) completes the demonstration. $\square$

The following theorem is one of the main results in Schäl (1975b, Theorem 6.6). There it is assumed that $K_n = H_n \times A_n$, i.e., all histories are admissible. However, the result carries over to the present more general situation upon making use of (4.1) (cp. also Schäl 1972), Kertz and Nachman (1977b)).

THEOREM 7.2. *$\Pi^\lambda$ is compact in the* ws$^\infty$*-topology.*

Given the relativization of the ws$^\infty$-topology on $\Pi^\lambda$, we endow $\Delta$ with the inverse image under the mapping $\pi \to P_\pi^\lambda$ of the topology on $\Pi^\lambda$. This is the coarsest topology on $\Delta$ for which the mapping $\pi \to P_\pi^\lambda$ is continuous (cp. Bourbaki (1960) I Section 2.3). Then we can rewrite Theorems 7.1 and 7.2 as

COROLLARY 7.3. (a) *If* $\mu \in \mathcal{P}(\Theta)$, *then* $\pi \to R(\mu, \pi)$ *is a* l.s.c. *function on* $\Delta$.
(*b*) $\Delta$ *is compact.*

7.2. *Existence of optimal policies.*

THEOREM 7.4. *Let* $\Lambda$ *be a nonempty subset of* $\mathcal{P}(\Theta)$.
(a) *There exists a* $\Lambda$*-optimal policy in* $\Delta$.
(b) *If* $\Gamma$ *is a nonempty closed subset of* $\Delta$, *then the set* $*(\Lambda)$ *of all* $\Lambda$*-optimal policies in* $\Gamma$ *is a nonempty closed (and hence compact) subset of* $\Delta$.
(c) *For any* $\alpha : \Theta \to \hat{\mathbb{R}}$, $\Delta_\alpha = \{\pi \in \Delta;\ R(\vartheta, \pi) \leqslant \alpha(\vartheta),\ \vartheta \in \Theta\}$ *is a closed subset of* $\Delta$ *and hence a candidate for the set* $\Gamma$ *in part* (b) *if* $\Delta_\alpha \neq \emptyset$.

PROOF. (a) is a consequence of (b). On the other hand, (b) follows from Corollary 7.3 since the lower semicontinuity of $R(\mu, \cdot\cdot)$, $\mu \in \Lambda$, implies the lower semicontinuity of $\sup_{\mu \in \Lambda} R(\mu, \cdot)$ and the closed subset $\Gamma$ of the compact set $\Delta$ is compact. Further, by virtue of the lower semicontinuity of $R(\vartheta, \cdot)$, $\{\pi;\ R(\vartheta, \pi) \leqslant \alpha(\vartheta)\}$ is closed for any $\vartheta \in \Theta$. Hence $\Delta_\alpha = \bigcap_{\vartheta \in \Theta} \{\pi;\ R(\vartheta, \pi) \leqslant \alpha(\vartheta)\}$ is closed. □

REMARK 7.5. Theorem 7.4 obviously remains true if $R(\mu, \pi)$ is replaced by $R(\mu, \pi) - \rho(\mu)$ for some $\rho : \Lambda \to \mathbb{R}$. Hence, upon setting $\rho(\vartheta) = \inf_{\pi \in \Delta} R(\vartheta, \pi)$, we know that there exists a policy $\pi^*$ that is optimal in $\Delta$ with respect to the *minimax regret criterion,* i.e., with respect to the criterion function $\pi \to \sup_{\vartheta \in \Theta}\{R(\vartheta, \pi) - \inf_{\pi' \in \Delta} R(\vartheta, \pi')\}$.

REMARK 7.6. Theorem 7.4 contains as a special case a theorem on the *existence of optimal tests*. However, this result is known. It is proved in Witting (1966) for an arbitrary state space and generalized by Landers and Rogge (1972) to the undominated case.

REMARK 7.7. The only property of $\Delta$ used for the proof of Theorem 7.4 is the compactness. When $\Gamma^*(\Lambda)$ contains more than one element, one may start another optimization procedure upon replacing $\Delta$ by $\Gamma^*(\Lambda)$, which is again compact by Theorem 7.4b, and $(c_n^\vartheta)$, $\Lambda$, $\Gamma$ by some other quantities satisfying the same conditions. Then one obtains another nonempty closed set of optimal policies contained in $\Gamma^*(\Lambda)$. More generally, one can start a sequence of optimization procedures determined by some $(c_{ni}^\vartheta)$, $\Lambda_i$, $\Gamma_i$, where $i$ runs through $\mathbb{N}$, and obtain a decreasing sequence of nonempty compact sets of optimal policies which has a nonempty intersection consisting of those policies that are lexicographically optimal for all optimization procedures. Examples can be found in Hodges and Lehmann (1952), Jaquette (1973), Mandl (1971), Quelle (1976).

From the discussion of the optimality criterion in Section 3 it is clear that Theorem 7.4 contains the following result as special case.

COROLLARY 7.8. (a) *Whatever* $\mu \in \mathcal{P}(\Theta)$, *there exists a Bayes policy against the* a priori *distribution* $\mu$.
(b) *There exists a minimax policy.*

Corollary 7.8 was proved for a general statistical decision problem by Wald (1950) Theorems 3.5, 3.7 and Ghosh (1952). Corollary 7.8a is strongly related to a result by Rieder (1975a) Theorem 8.4. Rieder's hypotheses are based on the weak topology (cp. Remark 5.3).

REMARK 7.9. If $\Theta$ is a singleton, then the statistician may restrict attention to nonrandomized policies (cp. Blackwell (1965) Theorem 2, Strauch (1966) Theorem 4.3, Hinderer (1970) Theorem 15.2, Hinderer (1971) Satz 4.1, Schäl (1971) Satz 6.1). This fact remains true if the statistician looks for a Bayes solution (cp. Rieder (1975a) Theorem 8.1), because a Bayesian decision model can be reduced to a model where $\Theta$ is a singleton. This fact can also be seen from relation (7.4) above.

*7.3. Convergence of the minimal cost from finite stage play to the minimal cost from infinite stage play.*

THEOREM 7.10. *Let $\Lambda$ be a nonempty subset of $\mathcal{P}(\Theta)$.*

(a) *Then* $\lim_n \inf_{\pi\in\Delta}\sup_{\mu\in\Lambda} R_1^n(\mu, \pi) = \inf_{\pi\in\Delta}\sup_{\mu\in\Lambda}R(\mu, \pi)$.

(b) *If $(\pi_n^*)$ is a sequence of policies such that $\pi_n^*$ is optimal with respect to the finite horizon n*, i.e., *with respect to the objective function $\pi \to \sup_{\mu\in\Lambda}R_1^n(\mu, \pi)$, then every accumulation point $\pi^*$ of $(\pi_n^*)$ is optimal with respect to the infinite horizon*, i.e., *$\pi^*$ is $\Lambda$-optimal in $\Delta$.*

PROOF. From (2.1) we conclude

$$R_1^n(\mu, \pi) \geqslant R_1^m(\mu, \pi) + z_m, \qquad m \leqslant n, \tag{7.5}$$

$$\sup_{\mu\in\Lambda}R_1^n(\mu, \pi) \geqslant \sup_{\mu\in\Lambda}R_1^m(\mu, \pi) + z_m, \qquad m \leqslant n. \tag{7.6}$$

In view of (2.3) and (7.5), it is readily proved that

$$\lim_{n\to\infty}\sup_{\mu\in\Lambda}R_1^n(\mu, \pi) = \sup_{\mu\in\Lambda}R(\mu, \pi). \tag{7.7}$$

As in the proof of Theorem 7.4, we know that the mapping $\pi \to \sup_{\mu\in\Lambda}R_1^n(\mu, \pi)$ is l.s.c. When these facts are combined with (7.6), we obtain by Proposition 10.1.3 in Schäl (1975a)

$$\lim_n\inf_{\pi\in\Delta}\sup_{\mu\in\Lambda}R_1^n(\mu, \pi) = \inf_{\pi\in\Delta}\lim_n\sup_{\mu\in\Delta}R_1^n(\mu, \pi). \tag{7.8}$$

Combining (7.7) and (7.8) completes the demonstration of (a). Part (b) follows from Proposition 10.1.2 in Schäl (1975a). □

*7.4. Convexity of the space of policies.* In order to obtain further results, one needs some convexity property of $\Delta$.

THEOREM 7.11. *Whatever $q \in \mathfrak{q}$, $\Pi^q$ is a convex subset of $\mathcal{P}(H_\infty)$.*

PROOF. The assertion is a simple consequence of the following characterization of $\Pi^q$ (cp. Strauch (1966) Lemma 7.2, Hinderer (1970) Lemma 13.1, Nowak (1975)). For any $P \in \mathcal{P}(H_\infty) : P \in \Pi^q$ if and only if for $n \in \mathbb{N}$

$$P\circ\zeta_1^{-1} = q_0,\ P\circ(\zeta_1, \cdots, \alpha_n, \zeta_{n+1})^{-1} = P\circ(\zeta_1, \cdots, \alpha_n)^{-1} \otimes q_n. \qquad \square$$

As a consequence of (7.4) and Theorem 7.11, one obtains

COROLLARY 7.12. *Whatever may be* $\pi_1$, $\pi_2 \in \Delta$ *and* $0 < \gamma < 1$, *there exists* $\pi \in \Delta$ *such that* $R(\mu, \pi) = \gamma R(\mu, \pi_1) + (1 - \gamma)R(\mu, \pi_2)$, $\mu \in \mathcal{P}(\Theta)$.

*7.5. Strict determinateness of the decision problem viewed as a zero sum two-person game.*

THEOREM 7.13. $\sup_{\mu \in \mathcal{P}(\Theta)} \min_{\pi \in \Delta} R(\mu, \pi) = \min_{\pi \in \Delta} \sup_{\vartheta \in \Theta} R(\vartheta, \pi)$.

PROOF. This follows from the minimax theorem of Kneser, Fan, Sion (cp. Sion (1958) Theorem 4.2′) where 'inf' can be replaced with 'min' by a lower semicontinuity argument. □

At the end of this section, we remark that Corollaries 7.3 and 7.12 may be used to prove theorems on complete classes of policies. For example, Assumptions (8), (9) and (10) of LeCam (1955) are satisfied and hence his Theorems 3 and 4 apply to the present situation.

**8. Results under condition (C).**

*8.1 Conditions concerning the continuous dependence on the parameter.* Several results in the area of statistical decision theory are based on the compactness of $\Theta$ and the continuity of the mappings $\vartheta \to R(\vartheta, \pi)$, $\pi \in \Delta$. In order to guarantee these properties, we impose the following Condition (C) throughout this section.

CONDITION (C). For any $n \in \mathbb{N}$

(C1) $K_{n(s_1, \cdots, s_n)}$ is compact for $s_i \in S_i$, $1 \leq i \leq n$;

(C2) $\lambda_{n+1}(s_1, a_1, \cdots, s_n, a_n; \cdot)$ does not depend on $(a_1, \cdots, a_n)$ and $(\vartheta, a_1, \cdots, a_n) \to 1^{\vartheta}_{n+1}(s_1, a_1, \cdots, a_n, s_{n+1})$ is l.s.c. on $\Theta \times K_{n(s_1, \cdots, s_n)}$ for all $s_i \in S_i$, $1 \leq i \leq n + 1$;

(C3) $(\vartheta, h) \to c^{\vartheta}_n(h)$ is bounded on $\Theta \times K_n \times S_{n+1}$ and $(\vartheta, a_1, \cdots, a_n) \to c^{\vartheta}_n(s_1, a_1, \cdots)$ is continuous on $\Theta \times K_{n(s_1, \cdots, s_n)}$ for all $s_i \in S_i$, $1 \leq i \leq n + 1$;

(C4) $\sup_{\vartheta \in \Theta, \pi \in \Delta} \sum_{i=n}^{\infty} \int |c^{\vartheta}_i| \, dP^{\vartheta}_{\pi} \to 0$ as $n \to \infty$;

(C5) $\Theta$ is a compact metric space (endowed with the $\sigma$-algebra of Borel subsets of $\Theta$).

Obviously Condition (C) implies Condition (B). Condition (C4) implies Condition (A) and is satisfied in the discounted case. Further we infer from (C4) that $R(\vartheta, \pi)$ is a bounded function on $\Theta \times \Delta$.

*8.2. Continuity theorems.*

THEOREM 8.1. $\sup_{\pi \in \Delta} |R(\vartheta, \pi) - R(\vartheta^\circ, \pi)| \to 0$ as $\vartheta \to \vartheta^\circ$, $\vartheta^\circ \in \Theta$.

This continuity theorem is related to results of Boylan (1969), Dubins and Meilijson (1974), Kobayashi, Fujikawa and Kurano (1973), Martin (1967) Theorem 3.3.4, and Whitt (1977). The proof is not given here. For the case $K_n = H_n \times A_n$, i.e., that all histories are admissible, the proof is carried through in Schäl (1976) by use of Remark 5.1 and the observation that the total variation of the difference of the distribution of $(\zeta_1, \alpha_1, \cdots, \zeta_n)$ under $P^{\vartheta}_{\pi}$ and $P^{\vartheta^\circ}_{\pi}$ is not larger than

$$\sum_{m=1}^{n} \int \lambda_1(ds_1) \int \pi_1(s_1; da_1) \int \cdots \int \lambda_m(ds_m) 1^{\vartheta^\circ}_1 \cdots 1^{\vartheta^\circ}_{m-1} |1^{\vartheta}_m - 1^{\vartheta^\circ}_m|.$$

The proof can be extended to the present more general situation by observing that it suffices to show that $\vartheta \to \sup_{\pi \in \Delta} |R(\vartheta, \pi) - R(\vartheta^\circ, \pi)|$ is upper semicontinuous. We endow $\mathcal{P}(\Theta)$ with the weak topology. Referring to Billingsley (1968) 1.2 Problem 8, we obtain the following generalization of Theorem 8.1.

COROLLARY 8.2. $\sup_{\pi \in \Delta} |R(\mu, \pi) - R(\mu^\circ, \pi)| \to 0$ as $\mu \to \mu^\circ$, $\mu^\circ \in \mathcal{P}(\Theta)$.
From Corollary 8.2 we have the following result.

COROLLARY 8.3. *For every subset* $\Gamma$ *of* $\Delta$, *the mapping* $\mu \to \inf_{\pi \in \Gamma} R(\mu, \pi)$ *is continuous on* $\mathcal{P}(\Theta)$.

*8.3. Existence of least favourable* a priori *distributions.* Let $\Gamma$ be a subset of $\Delta$ and $\Lambda$ be a subset of $\mathcal{P}(\Theta)$. Then $\mu^*$ is said to be a *least favourable* a priori *distribution* for $\Gamma$ in $\Lambda$ if

$$\inf_{\pi \in \Gamma} R(\mu^*, \pi) = \sup_{\mu \in \Lambda} \inf_{\pi \in \Gamma} R(\mu, \pi).$$

THEOREM 8.4. *Let* $\Gamma$ *be a subset of* $\Delta$ *and let* $\Lambda$ *be a closed subset of* $\mathcal{P}(\Theta)$. *Then there exists a least favourable* a priori *distribution for* $\Gamma$ *in* $\Lambda$.

PROOF. The compactness of $\Theta$ implies the compactness of $\mathcal{P}(\Theta)$ (cp. Parthasarathy (1967) Theorem II 6.4). Thus $\Lambda$ is a compact subset of $\mathcal{P}(\Theta)$. Further, by Corollary 8.3, $\mu \to R(\mu, \pi)$ is continuous, hence $\mu \to \inf_{\pi \in \Gamma} R(\mu, \pi)$ is upper semicontinuous and attains its supremum on $\Lambda$. □

Theorem 8.4 was proved for a general statistical decision problem by Wald (1950) Theorem 3.14.

The results of Sections 7 and 8 supply the foundation for a series of further results which can be obtained by using only the known decision-theoretical methods. For example, it will be found that the Assumptions 1–3 of Hodges and Lehmann (1952) are satisfied. Furthermore, results corresponding to Theorems 3.9 and 3.20 by Wald (1950) are easily proved for the model of the present paper. Also, the continuity and finiteness of the mapping $\vartheta \to R(\vartheta, \pi)$ implies the admissibility of every Bayes policy against an a priori distribution whose support is $\Theta$ (cp. Blyth (1951), Ferguson (1967) 2.3 Theorem 3, Zacks (1971) Theorem 8.1.2).

**Acknowledgment.** The author is grateful to the referees for several useful remarks.

## REFERENCES

[1] BERGE, E. (1963). *Topological Spaces*. MacMillan, New York.
[2] BIERLEIN, D. (1963). Einheitliche Prinzipien für die Beurteilung statistischer Verfahren. *Blätter der Deutschen Gesellschaft für Versicherungsmathematik VI, 2,* Triltsch, Würzburg.
[3] BIERLEIN, D. (1967). Zur Einbeziehung der Erfahrung in Spieltheoretische Modelle. In *Methods of Operations Research* (ed., R. Henn) Vol. III, 29–54. Anton Hain, Meisenheim.
[4] BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley and Sons, New York.
[5] BLACKWELL, D. (1965). Discounted dynamic programming. *Ann. Math. Statist.* **36** 226–235.
[6] BLYTH, C. R. (1951). On minimax statistical decision procedures and their admissibility. *Ann. Math. Statist.* **22** 22–42.
[7] BOURBAKI, N. (1960). *Elements of Mathematics: General Topology, Part 1.* Herman, Paris.

[8] BOYLAN, E. S. (1969). Stability theorems for solutions to the optimal inventory equation. *J. Appl. Probability* **6** 211–217.

[9] BROWN, L. D. (1977). Closure theorems for sequential-design processes. In *Statistical Decision Theory and Related Topics, II*. Academic Press, New York.

[10] BROWN, L. D. and PURVES, R. (1973). Measurable selections of extrema. *Ann. Statist.* **1** 902–912.

[11] BUNKE, O. (1964). Bedingte Strategien in der Spieltheorie: Existenzsätze und Anwendung auf statistische Entscheidungs-funktionen. *Trans. Third Prague Conf. Information Theory, Statist. Decision Functions, Random Processes* 35–43.

[12] DERMAN, C. (1965). Markovian sequential control processes-denumerable state space. *J. Math. Anal. Appl.* **10** 295–302.

[13] DUBINS, L. E. and MEILIJSON, I. (1974). On stability for optimization problems. *Ann. Probability* **2** 243–255.

[14] FERGUSON, TH. S. (1967). *Mathematical Statistics, A Decision Theoretic Approach*. Academic Press, New York.

[15] GHOSH, M. N. (1952). An extension of Wald's decision theory to unbounded weight functions. *Sankhyā Ser. A* **12** 8–26.

[16] HIMMELBERG, C. J., PARTHASARATHY, T. and VANVLECK, F. S. (1976). Optimal plans for dynamic programming problems. *Math. Operations Research* **1** 390–394.

[17] HINDERER, K. (1970). Foundations of non-stationary dynamic programming with discrete time-parameter. *Lecture Notes Operations Research and Math. Systems.* **33** Springer-Verlag, Berlin-Heidelberg-New York.

[18] HINDERER, K. (1971). Instationäre dynamische Optimierung bei schwachen Voraussetzungen über die Gewinnfunktionen. *Abh. Math. Sem. Univ. Hamburg* **36** 208–223.

[19] HODGES, J. L. JR. and LEHMANN, E. L. (1952). The use of previous experience in reaching statistical decisions. *Ann. Math. Statist.* **23** 396–407.

[20] JAQUETTE, S. C. (1973). Markov decision processes with a new optimality criterion: discrete time. *Ann. Statist.* **1** 496–505.

[21] KERTZ, R. P. (1977). On ws-convergence of probability measures: Portmanteau theorems and the preservation of convergence under mappings. Technical Report, Georgia Inst. Tech., Atlanta.

[22] KERTZ, R. P. and NACHMAN, D. C. (1977a). Optimal plans for discrete-time non-stationary dynamic programming with general total reward function I: the topology of weak convergence case. Technical Report, Georgia Inst. Tech., Atlanta.

[23] KERTZ, R. P. and NACHMAN, D. C. (1977b). Optimal plans for discrete-time non-stationary dynamic programming with general total reward function II: the $ws^{\infty}$-topology case. Technical Report, Georgia Inst. Tech., Atlanta.

[24] KOBAYASHI, Y., FUJIKAWA, H. and KURANO, M. (1973). On efficiency of inferences of transition probabilities to Markovian decision processes. *Bull. Math. Statist.* **15** 113–122.

[25] LANDERS, D. and ROGGE, L. (1972). Existence of most powerful tests for undominated hypotheses. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **24** 339–340.

[26] LECAM, L. (1955). An extension of Wald's theory of statistical decision functions. *Ann. Math. Statist.* **26** 69–81.

[27] MANDL, P. (1971). On the variance in controlled Markov chains. *Kybernetika (Prague)* **7** 1–12.

[28] MANDL, P. (1974). Estimation and control in Markov chains. *Advances in Appl. Probability* **6** 40–60.

[29] MARTIN, J. J. (1967). *Bayesian Decision Problems and Markov Chains.* Wiley and Sons, New York.

[30] MENGES, G. (1968). On the "Bayesification" of the minimax-principle. *Unternehmensforschung* **10** 81–91.

[31] NEVEU, J. (1965). *Mathematical Foundations of the Calculus of Probability.* Holden-Day, San Francisco.

[32] NOWAK, A. (1977). On a general dynamic programming problem. *Colloq. Math.* **37** 131–138.

[33] PARTHASARATHY, K. R. (1967). *Probability Measures on Metric Spaces.* Academic Press, New York.

[34] QUELLE, G. (1976). Dynamic programming of expectation and variance. *J. Math. Anal. Appl.* **55** 239–252.

[35] RIEDER, U. (1973). Stochastische dynamische Optimierung unter Unsicherheit. Lecture presented at the ADOW-meeting 1973 in Rheda/Bielefeld.

[36] RIEDER, U. (1975a). Bayesian dynamic programming. *Advances in Appl. Probability* **7** 330–348.

[37] RIEDER, U. (1975b). On stopped decision processes with discrete time parameter. *Stochastic Processes Appl.* **3** 365–383.

[38] SCHÄL, M. (1971). Ein verallgemeinertes stationäres Entscheidungsmodell der dynamischen Optimierung. In *Methods of Operations Research.* (ed., R. Henn). Vol. X, 145–162. Anton Hain, Meisenheim.

[39] SCHÄL, M. (1972). Dynamische Optimierung unter Stetigkeits und Kompaktheitsbedingungen. Habilitationsschrift, Univ. Hamburg.

[40] SCHÄL, M. (1975a). Conditions for optimality in dynamic programming and for the limit of $n$-stage optimal policies to be optimal. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **32** 179–196.

[41] SCHÄL, M. (1975b). On dynamic programming: compactness of the space of policies. *Stochastic Processes Appl.* **3** 345–364.

[42] SCHÄL, M. (1976). On dynamic programming and statistical decision theory. Preprint No. 84, Univ. Bonn.

[43] SCHÄL, M. (1977). Addendum to Schäl (1975a, b). Technical Report, Univ. Bonn.

[44] SION, M. (1958). On general minimax theorems. *Pacific J. Math.* **VIII** 171–176.

[45] STRAUCH, R. E. (1966). Negative dynamic programming. *Ann. Math. Statist.* **37** 871–890.

[46] WALD, A. (1950). *Statistical Decision Functions*. John Wiley and Sons, New York.

[47] WESSELS, J. (1968). Decision rules in Markovian decision processes with incompletely known transition probabilities. Doctoral dissertation, Technische Hogeschool, Eindhoven.

[48] WHITT, W. (1978). Approximation of dynamic programs, I. *Math. Operations Research* **3** 231–243.

[49] WITTING, H. (1966). *Mathematische Statistik*. B. G. Teubner, Stuttgart.

[50] ZACKS, S. (1971). *The Theory of Statistical Inference.* Wiley and Sons, New York.

UNIVERSITÄT BONN
INSTITUT FÜR ANGEWANDTE MATHEMATIK
WEGELERSTRAßE 6
D-5300 BONN