MANAGEMENT SCIENCE
Vol. 23, No. 8, April, 1977
Printed in U.S.A.

A MARKOV QUALITY CONTROL PROCESS SUBJECT TO PARTIAL OBSERVATION*

CHELSEA C. WHITE†

This paper studies the problem of optimally controlling a discrete-time production process with countable state space which is subject to one of three control settings at each time interval: produce, inspect while producing, or repair (revise) the process. The cost of the item produced and the inspection and repair costs are assumed dependent on the state of the production process. It is assumed that the inspector-decisionmaker receives imperfect on-line observations of the production process at both times of production and inspection. Bounds on optimal cost are obtained. For the two-state case, several results associated with observation quality are determined which are sufficient for particularly simple characterizations of an optimal policy. Generalizations of several results due to Ross are also presented.

1. Introduction

This paper studies the following generalized formulation of modified versions of the quality control problems presented in [3]–[5] and [8]. A production process is assumed to be in any one of a countable number of states, producing items at times $t = 0, 1, \ldots$. At each time t, the inspector is allowed to choose one of three controls: allow the process to produce, inspect the process while it produces, or repair the process. It is assumed that the cost of the item produced and the costs incurred during inspection and repair are dependent on the state of the production process. Furthermore, it is assumed that the inspector receives imperfect on-line observations of the production process at both times of production and inspection. This assumption generalizes the information pattern assumptions made in [3]–[5] and [8]. The repair decision is assumed to place the production process in a specific state during the next time interval with probability one; thus, the information pattern associated with the repair decision is inconsequential. The problem considered is to determine a nonanticipative control policy, given a priori data, all on-line observations made to date, and all former control decisions made, i.e. the information pattern is classical [10], which minimizes the expected discounted cost accrued over the infinite horizon.

Well-known results useful for the developments which follow are presented in §3 for the quality control problem formulated in §2. §4 presents bounds on optimal cost, which are the optimal costs associated with specific cases of observation quality. The generalized two-state problem is analyzed in §5. Theorem 5.4 states sufficient conditions related to observation quality for more general bounds on optimal cost than those presented in §4. Theorem 5.8 presents conditions related to observation quality which insure that it is optimal never to inspect. §6 considers the special case of the problem discussed in §5, where inspection provides the inspector with the exact state of the production process. §6 is essentially a generalization of similar results in [5], but here the inspector is allowed on-line data at time of production which is dependent on the production process. Conclusions are discussed in the final section.

2. Problem Formulation

Let $(s(t), t = 0, 1, \ldots,)$ be a discrete-time, time-invariant, controlled Markov process model of a production process having countable state space $0, 1, \ldots$. Assume probabilities of the form $p_i^0 = P(s(0) = i)$, $i \in \{0, 1, \ldots\}$, are given, where $p^0 = (p_i^0) \in X = \{x : x_i \geqslant 0, \sum_{i=0}^{\infty} x_i = 1\}$.

* Accepted by Ronald A. Howard; received February 21, 1975.
† University of Virginia.

Let $(u(t), t = 0, 1, \ldots)$ be called the control process, a discrete-time stochastic process having finite state space $U = \{0, 1, 2\}$. The value of $u(t)$ denotes the decision made by the inspector at time t; when $u(t) = 0$, 1, or 2, the inspector at time t allows the production process to produce, inspects the production process as it produces, or revises the production process, respectively. The data on which the inspector bases a decision (the information pattern) will be stated following several additional definitions.

Let $(z(t), t = 1, 2, \ldots)$ be called the observation process, a discrete-time, time-invariant, controlled stochastic process with countable state space $0, 1, \ldots$. The realization of the random variable $z(t)$ will represent the on-line data which becomes available to the inspector at time t.

Define $s_t = (s(0), \ldots, s(t))$, $u_t = (u(0), \ldots, u(t))$, and $z_t = (z(1), \ldots, z(t))$. Assume the production and observation processes are such that

$$P(s(t+1) \mid s_t, u_t, z_t, p^0) = P(s(t+1) \mid s(t), u(t)),$$

$$P(z(t+1) \mid s_{t+1}, u_t, z_t, p^0) = P(z(t+1) \mid s(t+1), u(t)).$$

Let

$$p_{ij}(v) = P(s(t+1) = j \mid s(t) = i, u(t) = v),$$

$$q_{jk}(v) = P(z(t+1) = k \mid s(t+1) = j, u(t) = v) \quad \text{for all } i, j, k \in \{0, 1, \ldots\} \text{ and } v \in U.$$

Let $g(i, v)$ represent the cost accrued at time t, given $s(t) = i$ and $u(t) = v$, which is assumed uniformly bounded in i and v. Costs are assumed to be discounted over the infinite control horizon $\{0, 1, \ldots\}$; hence, $\sum_{t=0}^{\infty} \beta^t g(s(t), u(t))$ is the total cost accrued by processes $(s(t), t = 0, 1, \ldots)$ and $(u(t), t = 0, 1 \ldots)$, where $\beta \in [0, 1)$ represents the discount factor.

In making a decision at time t, the inspector is assumed to know the a priori probability p^0, the past and present sample path of the observation process z_t, the past sample path of the control process u_{t-1}, and the set $\Xi = (U, \{p_{ij}(\cdot)\}, \{q_{jk}(\cdot)\}, \{g(\cdot, \cdot)\})$, designating the structure of the optimization problem. Note that although the inspector does not know the future sample paths of the observation and control processes, the inspector is assumed to know their future conditional probability distributions. Such a closed-loop policy "causally anticipates" the usefulness of data, allows for the application of the principle of optimality [1], and implies the following definition.

Definition. Call $d_t = (p^0, u_{t-1}, z_t)$ the data sequence at time t. Then a control process is said to be (deterministic) admissible if for each $t = 0, 1, \ldots$, there exists a function $\Psi(\Xi, t, \cdot)$ such that $u(t) = \Psi(\Xi, t, d_t) \in U$. For brevity, the dependence of Ψ on Ξ will be deleted.

The optimization problem is to determine an admissible control policy, given optimization problem structure Ξ, which minimizes with respect to the set of all admissible control policies the criterion $E_{p^0}\{\sum_{t=0}^{\infty} \beta^t g(s(t), u(t))\}$, where E_{p^0} denotes expectation conditioned on a priori data p^0.

The transition and observation probabilities are more precisely defined as follows. Let p_{ij} be the natural unrevised transition probabilities of the production process, i.e. $p_{ij}(v) = p_{ij}$, $v \neq 2$. Revision implies that the production process is put in the 0 state with probability one for the next time interval; thus, $p_{ij}(2) = \delta_{0j}$, where δ is the Kronecker delta. The probabilities $q_{jk}(0)$ represent the natural uninspected observation probabilities, indicating in some sense the quality of data received from the observation process when the system is producing. The probabilities $q_{jk}(1)$ represent the quality of data received from the observation process when the system is producing and being inspected. Since it is known with probability one in which state the production process will be during the upcoming time interval when the decision is to revise, the probabilities $q_{jk}(2)$ are of no consequence and can be chosen arbitrarily.

3. Preliminary Results

It is well known [7] that it is sufficient for the inspector to know only $\tilde{y}(t, d_t) = (\tilde{y}_i(t, d_t))$, where $\tilde{y}_i(t, d_t) = P(s(t) = i \mid d_t)$, in order to determine an optimal policy decision at time t. That is, the probability density vector of the present state of the production process, conditioned on all past and present data, sufficiently summarizes the past and present data for optimal policy decisions. Thus, if Ψ is an optimal policy, there exists a function $\sigma(t, \cdot): X \to U$ for each $t = 0, 1, \ldots,$ such that $\sigma(t, \tilde{y}(t, d_t)) = \Psi(t, d_t)$. Also, the finiteness of the set U implies the existence of a stationary optimal policy $\phi : X \to U$ [7]. Necessary conditions and sufficient conditions for a stationary policy to be optimal can be derived from the dynamic programming equation (DPE)

$$\Gamma(y) = \min_{v \in U} \left[\sum_{i=0}^{\infty} y_i g(i, v) + \beta \sum_{k=0}^{\infty} V(k, y, v) \Gamma(T(k, y, v)) \right] \tag{3.1}$$

for $y \in X$, where $\Gamma(p^0)$ is the unique minimum of the cost criterion, minimized with respect to the set of all admissible control policies, and where

$$N_j(k, y, v) = q_{jk}(v) \sum_i p_{ij}(v) y_i, \quad N(k, y, v) = (N_j(k, y, v)),$$

$$V(k, y, v) = \sum_j N_j(k, y, v), \quad T(k, y, v) = N(k, y, v)/V(k, y, v)$$

for $V(k, y, v) \neq 0$. Note N_j and V are scalars, whereas N and T are vectors. (When $V(k, y, v) = 0$, define $T(k, y, v)$ in X arbitrarily.) More specifically, (3.1) becomes

$$\Gamma(y) = \min \begin{bmatrix} \sum_i y_i C_i + \beta \sum_k V(k, y, 0) \Gamma(T(k, y, 0)), \\ \sum_i y_i I_i + \beta \sum_k V(k, y, 1) \Gamma(T(k, y, 1)), \ \sum_i y_i R_i + \beta \Gamma(e_0) \end{bmatrix}, \tag{3.2}$$

where $e_0 = (1, 0, \ldots, 0, \ldots)$, $g(i, 0) = C_i$, $g(i, 1) = I_i$, and $g(i, 2) = R_i$ (in keeping with notation in [5]), and $N_j(k, y, v) = q_{jk}(v) \sum_i p_{ij} y_i$, $v = 1, 2$. Hence, C_i, I_i, and R_i are the costs accrued at each time interval when the production process is in state i if the decision is to produce, inspect, and revise, respectively.

The concavity of the optimal expected cost function Γ [5] is a particularly important property that will be used throughout the paper. All results to follow for concave functions will eventually be applied to Γ. Other properties related to the concavity of Γ are also of interest. Let $\Gamma^v(y) = \sum_i y_i g(i, v) + \beta \sum_k V(k, y, v) \Gamma(T(k, y, v))$, and let the set of all $y \in X$ such that $\Gamma^v(y) = \Gamma(y)$ be called the production, inspection, and revision region when $v = 0$, 1, and 2, respectively. It follows directly from results in [5] that the revision region is convex. Also, if $q_{ij}(0) = \delta_{jk}$ ($q_{jk}(1) = \delta_{jk}$), the production (inspection) region is convex. The convexity of the vth region, $v = 0, 1$, when $q_{jk}(v) = \delta_{jk}$, results from the fact that Γ^v becomes linear. Note that for the case where $C_i < I_i < R_i$ for all i, if $q_{jk}(0) = q_{jk}(1)$ for all j and k, then the inspection region is null. Note also that these results hold for all $\beta \in [0, 1)$, and hence results in [2] for the expected average cost per unit time criterion are easily extended to the present problem formulation.

4. Bounds on Optimal Cost

It is now shown that for $v = 0$ and/or $v = 1$, if $q_{ij}(v)$ is independent of j ($q_{jk}(v) = \delta_{jk}$), then the associated optimal cost represents an upper (lower) bound on optimal cost. Let $N_j'(k, y, v) = q_{jk}'(v)\sum_i p_{ij}y_i$ and $N_j''(k, y, v) = q_{jk}''(v)\sum_i p_{ij}y_i$, and define the pairs (V', T') and (V'', T'') accordingly. The proof of the following lemma follows from the proofs of [9, Lemmas 1 and 2] and is deleted.

LEMMA 4.1. *Let $D : X \to R$ be a concave function. Assume $q_{jk}'(v) = q_k'(v)$ (that is, $q_{jk}'(v)$ is independent of j) and $q_{jk}''(v) = \delta_{jk}$ for some $v \in \{0, 1\}$. Then, for any $y \in X$, $\sum_k V''(k, y, v)D(T''(k, y, v)) \leq \sum_k V(k, y, v)D(T(k, y, v)) \leq \sum_k V'(k, y, v)D(T'(k, y, v))$.*

It is noted that strict equality will always exist when $v = 3$ for any two arrays $\{q_{jk}'(2)\}$ and $\{q_{jk}''(2)\}$.

The following theorem gives upper and lower bounds on Γ.

THEOREM 4.2. *Assume $\Xi'' = \Xi = \Xi'$ except that for $v = 0$ and/or $v = 1$, $q_{jk}'(v) = a_k(v)$ and $q_{jk}''(v) = \delta_{jk}$. Let Γ', Γ, and Γ'' be the optimal cost functions associated with the system structures Ξ', Ξ, and Ξ'', respectively. Then, $\Gamma'' \leq \Gamma \leq \Gamma'$.*

PROOF. Let $\Gamma_0(y) = 0$ for all $y \in X$, and define

$$\Gamma_n(y) = \min_{v \in U}\left[\sum_i y_i g(i, v) + \beta \sum_i V(k, y, v)\Gamma_{n-1}(T(k, y, v))\right], \qquad n = 1, 2, \ldots .$$

Define Γ_n'' accordingly. Trivially, $\Gamma_0'' \leq \Gamma_0$; assume $\Gamma_{n-1}'' \leq \Gamma_{n-1}$. It then follows from Lemma 4.1 and (3.2) that $\Gamma_n'' \leq \Gamma_n$ for all n. Thus, $\Gamma'' \leq \Gamma$. The inequality $\Gamma \leq \Gamma'$ follows similarly Q.E.D.

5. The Two-State Partially Observed Case: $q_1 \in [0.5, 1]$

This section, and the remainder of the paper, considers the specific case where the general production problem can be in either one of two states. After several preliminary remarks, bounds on optimal cost, more general than those presented in §4, are presented in Theorem 5.4. Theorem 5.10 presents necessary and sufficient conditions for it always to be optimal to produce; also given is a sufficient condition for it to be optimal to revise for x near 1. The section ends with a lemma and corollary which will be useful in the development of results presented in §6.

Let the production process have finite state space $\{0, 1\}$, and assume that the system deteriorates naturally to and is absorbed by state 1 when no revision is allowed. That is, for given $\pi \in X = [0, 1]$, $\pi \neq 1$, let

$$\{p_{ij}(v)\} = \begin{bmatrix} 1-\pi & \pi \\ 0 & 1 \end{bmatrix}, \qquad v \neq 2,$$

$$\{p_{ij}(2)\} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}.$$

Let the observation process also have state space $\{0, 1\}$, where for $v = 0, 1$, $q_v \in [0.5, 1]$ and

$$\{q_{jk}(v)\} = \begin{bmatrix} q_v & 1-q_v \\ 1-q_v & q_v \end{bmatrix}.$$

Let $C_0 = 0$, $C_1 = C$, $R_i = R$, and $I_i = I$, for $i = 0, 1$, where $C < I < R$. Then, the DPE for the partially observed two-state case becomes

$$\Gamma(x) = \min\begin{bmatrix} xC + \beta \sum_{k=0}^{1} V(k, x, 0)\Gamma(T(k, x, 0)), \\ I + \beta \sum_{k=0}^{1} V(k, x, 1)\Gamma(T(k, x, 1)), R + \beta\Gamma(0) \end{bmatrix}, \qquad (5.1)$$

where $x \in X$ represents the probability of the production process in state 1, $N(k, y, v) = N(k, x, v)$ with $y = (1 - x, x)$, and V and $T = T_1$ are defined accordingly. Specifically, for $v = 0, 1$,

$$V(0, x, v) = q_v(1 - x)(1 - \pi) + (1 - q_v)(x(1 - \pi) + \pi),$$
$$V(1, x, v) = (1 - q_v)(1 - x)(1 - \pi) + q_v(x(1 - \pi) + \pi),$$
$$T(0, x, v) = (1 - q_v)(x(1 - \pi) + \pi)/V(0, x, v), \quad \text{and}$$
$$T(1, x, v) = q_v(x(1 - \pi) + \pi)/V(1, x, v).$$

It will be useful to note that $V(0, x, v) = -\alpha_v x + \beta_v(0)$ and $V(1, x, v) = \alpha_v x + \beta_v(1)$, where $\alpha_v = (1 - \pi)(2q_v - 1)$, $\beta_v(0) = q_v(1 - \pi) + (1 - q_v)\pi$, and $\beta_v(1) = 1 - \beta_v(0) = (1 - q_v)(1 - \pi) + q_v\pi$.

Several useful properties of $T(k, x, v)$ are presented after the following lemma.

Lemma 5.1. *Let $f::X \to R$ be defined as $f(s) = (as + b)/(cs + d)$. Let c and d be such that $cs + d > 0$ for all $s \in X$. Then, f is monotone nondecreasing on X if $bc \leqslant ad$.*

Proof. Let $s_1, s_2 \in X$ such that $s_2 - s_1 = \epsilon > 0$. Then, $bc \leqslant ad$ implies $bc\epsilon \leqslant ad\epsilon$. Adding the term $(acs_1^2 + (ad + bc)s_1 + bd + acs_1\epsilon)$ to both sides of the inequality and rearranging, it follows that $(cs_2 + d)(as_1 + b) \leqslant (as_2 + b)(cs_1 + d)$, which implies the result. Q.E.D.

Corollary 5.2. *The function $T(k, x, v)$ is such that* (i) *$T(k, x, v)$ is monotone nondecreasing in $x \in X$ for each k and v,* (ii) *$T(0, x, v) \leqslant T(1, x, v)$ for each $x \in X$ and v,* (iii) *$q_1 \leqslant q_0$ implies $T(0, x, 0) \leqslant T(0, x, 1)$ and $T(1, x, 1) \leqslant T(1, x, 0)$, for each $x \in X$, and* (iv) *$x \leqslant T(0, x, v)$ when either $q_v \neq 0.5$ and $x \leqslant (1 - q_v)\pi/[(1 - \pi)(2q_v - 1)]$ or $q_v = 0.5$, for $x \in X$ and $v = 0, 1$.*

Proof. Part (i) follows from Lemma 5.1 directly. Part (ii) follows from Lemma 5.1 by making the substitutions $s_1 = (1 - q_v)$ and $s_2 = q_v$. Similar substitutions and the use of Lemma 5.1 imply (iii). Part (iv) is proved by examining the roots of the function $x - T(0, x, v)$. Q.E.D.

The next lemma will be of use in developments to follow.

Lemma 5.3. *Let $q_1 \leqslant q_0$. Then, for any concave function $D : X \to R$,*

$$\sum_k V(k, x, 0)D(T(k, x, 0)) \leqslant \sum_k V(k, x, 1)D(T(k, x, 1)) \quad \textit{for all } x \in X.$$

Proof. From Corollary 5.2(iii) and by the concavity of D,

$$D(T(0, x, 1)) \geqslant \lambda_0 D(T(0, x, 0)) + (1 - \lambda_0)D(T(1, x, 0)), \quad \text{and}$$
$$D(T(1, x, 1)) \geqslant \lambda_1 D(T(0, x, 0)) + (1 - \lambda_1)D(T(1, x, 0)),$$

where $\lambda_i = (T(1, x, 0) - T(i, x, 1))/(T(1, x, 0) - T(0, x, 0))$, $i = 0, 1$. It is easily shown that for $\xi = (q_0 - q_1)/(2q_0 - 1)$, $\lambda_0 = (1 - \xi)V(0, x, 0)/V(0, x, 1)$, $\lambda_1 = \xi V(0, x, 0)/V(1, x, 1)$, $1 - \lambda_0 = \xi V(1, x, 0)/V(0, x, 1)$, and $1 - \lambda_1 = (1 - \xi) \cdot V(1, x, 0)/V(1, x, 1)$. The result then follows directly. Q.E.D.

The theorem below indicates that the closer the probabilities q_v, $v = 0, 1$, are to 1, the lower the optimal cost function Γ.

Theorem 5.4. *Let $\Xi = \Xi'$ except that $q_v' \leqslant q_v$, for both $v = 0$ and $v = 1$. Associate optimal cost functions Γ and Γ' with optimization problem structures Ξ and Ξ', respectively. Then, $\Gamma \leqslant \Gamma'$.*

Proof. Assume $\Gamma_0(x) = 0$, and define $\Gamma_n^v(x) = (1 - x)g(0, v) + xg(1, x) + \beta\sum_k V(k, x, v)\Gamma_n(T(k, x, v))$, where $\Gamma_n(x) = \min_{v \in U}\Gamma_n^v(x)$. Define Γ_n' and $\Gamma_n^{v'}$ similarly.

Trivially, $\Gamma_0 \leqslant \Gamma'_0$; assume $\Gamma_{n-1} \leqslant \Gamma'_{n-1}$. It follows from (5.1) and Lemma 5.3 that $\Gamma^v_n \leqslant \Gamma^{v'}_n$, $v \in U$; thus, $\Gamma_n \leqslant \Gamma'_n$ which is true for all n by induction. Since $\Gamma_n \to \Gamma$ and $\Gamma'_n \to \Gamma'$, the result is proved. Q.E.D.

LEMMA 5.5. *Let $D : X \to R$ be monotonically nondecreasing in x. Then, $\sum_k V(k, x, v) D(T(k, x, v))$ is monotonically nondecreasing in x, $v = 0,1$.*

PROOF. Assume $x_1, x_2 \in (0, 1]$, $x_1 \leqslant x_2$. From Corollary 5.2(i) and (ii),

$$
\begin{aligned}
D(T(1, x_1, v)) - D(T(0, x_1, v)) &\leqslant D(T(1, x_2, v)) - D(T(0, x_1, v)) \\
&\leqslant x_2(D(T(1, x_2, v)) - D(T(0, x_2, v)))/x_1 + \beta_v(0)(D(T(0, x_2, v)) \\
&\quad - D(T(0, x_1, v)))/(x_1\alpha_v) + \beta_v(1)[D(T(1, x_2, v)) - D(T(1, x_1, v))]/(x_1\alpha_v).
\end{aligned}
$$

The result then follows from straightforward algebraic manipulation. Proof of the $x_1 = 0$ case is direct and is deleted. Q.E.D.

COROLLARY 5.6. *The optimal cost function Γ is monotone nondecreasing in x.*

PROOF. Trivially, Γ_0 is m.n.d.; assume Γ_{n-1} is m.n.d. Clearly, Γ^2_n is m.n.d. It follows from Lemma 5.5 and (5.1) that Γ^v_n is m.n.d., $v = 0, 1$. Thus, Γ_n is m.n.d., which by complete induction is true for all n. Since $\Gamma_n \to \Gamma$, the result holds. Q.E.D.

Proof of the following result is similar to the proof of Lemma 3.2 [5] and is therefore deleted.

LEMMA 5.7. *Let $\phi : X \to U$ be an optimal policy. Then, $\phi(0) = 0$.*

The next result states that $q_1 \leqslant q_0$ implies essentially that the quality of observation at time of production is at least as good as the quality of observation at time of inspection, and hence it is of no benefit ever to inspect.

THEOREM 5.8. *Assume $q_1 \leqslant q_0$. Then, an optimal policy ϕ can be characterized by a number $x_1 \in X$ such that $\phi(x) = 0$ for $x \in [0, x_1)$ and $\phi(x) = 2$ for $x \in [x_1, 1]$.*

PROOF. Lemma 5.3 implies that $\Gamma^0 \leqslant \Gamma^1$. The result then follows from Corollary 5.6, Lemma 5.7, and the concavity of the revise region. Q.E.D.

Proof of Theorem 5.10, presented below, is similar to the proof of Theorem 3.4 [5] and is therefore deleted. (It is noted, however, that in the proof the continuity of Γ can follow directly from its concavity [6, p. 109].) The following lemma states a result necessary for the generalization of the proof of Theorem 3.4 [5] to the case considered in Theorem 5.10.

Define $W(x) = xC + \beta\sum_{k=0}^{1} V(k, x, 0) W(T(k, x, 0))$, the expected discounted cost to be accrued using the policy $\phi(x) = 0$, for all $x \in X$. The following result gives an explicit solution for W.

LEMMA 5.9. *For all $x \in X$, $W(x) = C(x + \beta\pi/(1 - \beta))/(1 - \beta(1 - \pi))$.*

PROOF. Note $W_n \to W$, where $W_0(x) = 0$ and $W_n(x) = xC + \beta\sum_k V(k, x, 0) \cdot W_{n-1}(T(k, x, 0))$. An induction argument shows that W_n is of the form $W_n(x)/C = \alpha'_n x + \beta'_n$, where $\alpha'_0 = \beta'_0 = 0$, $\alpha'_n = 1 + \beta(1 - \pi)\alpha'_{n-1}$, and $\beta'_n = \beta(\alpha'_{n-1}\pi + \beta'_{n-1})$, for $n = 1, 2, \ldots$. Limiting arguments produce the result. Q.E.D.

THEOREM 5.10. (a) *The policy $\phi(x) = 0$ for all $x \in X$ is optimal if and only if $R \geqslant C(1 + \beta\pi)/(1 - \beta(1 - \pi))$.* (b) *If $R < C(1 + \beta\pi)/(1 - \beta(1 - \pi))$, then every optimal policy ϕ is such that $\phi(x) = 1$ for x near 1.*

The next result gives bounds on $|\Gamma_n(x_1) - \Gamma_n(x_2)|$ as a function of n and $|x_1 - x_2|$ for the cases where q_0 and $q_1 \neq 1$ are both sufficiently close to 0.5 and where $q_1 = 1$ and q_0 is sufficiently close to 0.5.

LEMMA 5.11. *Let* $\gamma_v = 1 + 2(1-\pi)(2q_v - 1)q_v/(1-q_v)$. *Assume either* (a) *or* (b), *where*:

(a) $q_0 \neq 1$, $q_1 \neq 1$, *and* $\gamma = \max\{\gamma_0, \gamma_1\}$.

(b) $q_0 \neq 1$, $q_1 = 1$, *and* $\gamma = \gamma_0$.

Then, for $\gamma < 1/(\beta(1-\pi))$,

$$|\Gamma_n(x_2) - \Gamma_n(x_1)| \leqslant C|x_2 - x_1|\left(1 - (\beta(1-\pi)\gamma)^n\right)/(1 - \beta(1-\pi)\gamma),$$

for all $x_1, x_2 \in X$ *and all* n.

PROOF. The $n = 0$ case is trivial; assume the result is true for $n-1$. There are three cases of interest. The case where both x_1 and x_2 are such that an optimal policy is to revise for both points is trivial. Consider the $v = 0, 1$ cases, where it is noted that for $D(x, v) = (1-x)g(0, v) + xg(1, v)$, $|D(x_1, v) - D(x_2, v)| \leqslant C|x_1 - x_2|$.

(a) It is easily shown that for $v = 0$ or $v = 1$,

$$\begin{aligned}
|\Gamma_n(x_2) - \Gamma_n(x_1)| \leqslant C|x_2 - x_1| + \beta\{ &\beta_v(0)|\Gamma_{n-1}(T(0, x_2, v)) - \Gamma_{n-1}(T(0, x_1, v))| \\
&+ \beta_v(1)|\Gamma_{n-1}(T(1, x_2, v)) - \Gamma_{n-1}(T(1, x_1, v))| \\
&+ \alpha_v|x_1\Gamma_{n-1}(T(0, x_1, v)) - x_2\Gamma_{n-1}(T(0, x_2, v))| \\
&+ \alpha_v|x_2\Gamma_{n-1}(T(1, x_2, v)) - x_1\Gamma_{n-1}(T(1, x_1, v))|\} \\
\leqslant C|x_2 - x_1| + \beta\{ &(\beta_v(0) - \alpha_v)|\Gamma_{n-1}(T(0, x_2, v)) - \Gamma_{n-1}(T(0, x_1, v))| \\
&+ (\beta_v(1) + \alpha_v)|\Gamma_{n-1}(T(1, x_2, v)) - \Gamma_{n-1}(T(1, x_1, v))| \\
&+ 2\alpha_v|\Gamma_{n-1}(T(0, x_2, v)) - \Gamma_{n-1}(T(0, x_1, v))|\}.
\end{aligned}$$

It is straightforward to show that $|T(0, x_2, v) - T(0, x_1, v)| \leqslant (1-\pi)q_v|x_2 - x_1|/(1-q_v)$ and $|T(1, x_2, v) - T(1, x_1, v)| \leqslant (1-\pi)(1-q_v)|x_2 - x_1|/q_v$. Noting that $\beta_v(0) - \alpha_v = 1 - q_v$ and $\beta_v(1) + \alpha_v = q_v$, it follows that

$$\begin{aligned}
&|\Gamma_n(x_2) - \Gamma_n(x_1)| \leqslant C|x_2 - x_1| \\
&\qquad + C\beta(1-\pi)\{1 + 2\alpha_v q_v/(1-q_v)\}|x_2 - x_1|\left(1 - (\beta(1-\pi)\gamma)^{n-1}\right)/(1 - \beta(1-\pi)\gamma) \\
&\quad \leqslant C|x_2 - x_1| + C\beta(1-\pi)\gamma|x_2 - x_1|\left(1 - (\beta(1-\pi)\gamma)^{n-1}\right)/(1 - \beta(1-\pi)\gamma) \\
&\quad = C|x_2 - x_1|\left(1 - (\beta(1-\pi)\gamma)^n\right)/(1 - \beta(1-\pi)\gamma).
\end{aligned}$$

(b) The $v = 0$ case follows as in the hypothesis (a) case. The $v = 1$ case follows as in the proof of Lemma 3.5 [5]. Q.E.D.

Note that for hypothesis (b) in the above lemma, $q_0 = 0.5$ implies $\gamma = 1$.

COROLLARY 5.12. *Assume either* (a) *or* (b) *in Lemma* 5.11. *Then, for* $\gamma < 1/(\beta(1-\pi))$, $|\Gamma(x_2) - \Gamma(x_1)| \leqslant C|x_2 - x_1|/(1 - \beta(1-\pi)\gamma) \leqslant C|x_2 - x_1|/\pi$, *for all* $x_1, x_2 \in X$.

6. The Two-State Partially Observed Case: $q_1 = 1$

It is assumed throughout this section that upon inspection, the production process is completely observed, i.e. $q_1 = 1$. The results of this section are essentially direct generalizations of several results in [5]. Due to the fact that $q_0 \in [0.5, 1]$ does not necessarily equal 0.5, many of the hypotheses of the results do, however, differ from the hypotheses of the results in [5].

Theorem 5.8 presented sufficient conditions in terms of q_v for a particularly simple form for an optimal policy. Sufficient conditions for simplified characterizations of optimal policies are now given in Theorems 6.1, 6.2, and 6.4 when $q_0 < q_1 = 1$.

Theorem 6.1 states that in general, there are at most four control regions: an inspection region, a revise region, and two production regions. These four regions are illustrated for the $q_0 = 0.5$ case in [5, p. 591]. Theorem 6.2 gives conditions which imply that the optimal policy is monotonically nondecreasing in x. Theorem 6.4 then gives conditions which guarantee that it is never optimal to inspect. Proof of the following theorem is similar to the proof of Theorem 3.3 [5] and is deleted.

THEOREM 6.1. *An optimal policy ϕ can be characterized by three numbers, $0 \leqslant x_1 \leqslant x_2 \leqslant x_3 \leqslant 1$ such that $\phi(x) = 0$ for $x \in [0, x_1) \cup [x_2, x_3)$, $\phi(x) = 1$ for $x \in [x_1, x_2)$, and $\phi(x) = 2$ for $x \in [x_3, 1]$.*

The above theorem presents a somewhat counterintuitive result in that there may be two produce regions rather than a single produce region containing 0. An example of such a situation, for the case where $q_0 = 0.5$, is presented in [5].

Sufficient conditions for there to be no second region of production are now presented.

THEOREM 6.2. *Assume that either $q_0 \neq 0.5$, and $1 \leqslant (1 - q_0)\pi/((1 - \pi)(2q_0 - 1))$ or $q_0 = 0.5$. Then, the inequality*

$$(1 - \beta)(R + \beta\Gamma(0))/C \leqslant \left[(R - I)/(\beta(1 - \pi)(\Gamma(1) - \Gamma(0))) - \pi/(1 - \pi)\right]$$

implies that an optimal policy can be characterized by two numbers $x_1, x_2 \in X$ where $\phi(x) = 0$ for $x \in [0, x_1)$, $\phi(x) = 1$ for $x \in [x_1, x_2)$, and $\phi(x) = 2$ for $x \in [x_2, 1]$.

PROOF. Let η_1 and η_2 be such that $\eta_1 C + \beta V(0, \eta_1, 0)\Gamma(T(0, \eta_1, 0)) + \beta V(1, \eta_1, 0) \cdot \Gamma(T(1, \eta_1, 0)) = R + \beta\Gamma(0)$ and $I + \beta(1 - \eta_2)(1 - \pi)\Gamma(0) + \beta(\eta_2(1 - \pi) + \pi)\Gamma(1) = R + \beta\Gamma(0)$. If such an η_i does not exist, let it be infinite, $i = 1, 2$. In order that there be a second production region, it is necessary that $\eta_1 > \eta_2$. Then, $\Gamma(\eta_1) = R + \beta\Gamma(0)$. It follows from Corollary 5.2(iv) that $\eta_1 \leqslant T(k, \eta_1, 0)$, $k = 0, 1$, and hence $\Gamma(T(k, \eta_1, 0)) = R + \beta\Gamma(0)$, $k = 0, 1$. The remainder of the proof follows the outline of the proof of Theorem 3.7(a) [5] and is omitted. Q.E.D.

The next result gives a condition which implies an inequality hypothesis in Theorem 6.2 without the determination of Γ. Define $\gamma = 1 + 2\alpha_0 q_0/(1 - q_0)$.

COROLLARY 6.3. *Assume $\gamma < 1/(\beta(1 - \pi))$ and that there exists an admissible control ψ with associated expected discounted total cost Γ' such that*

$$\begin{aligned}&(1 - \beta)(R + \beta\Gamma'(0))/(1 - \beta(1 - \pi)\gamma)\\ &\qquad \leqslant \left[(R - I)/(\beta(1 - \pi)) - \pi C/((1 - \pi)(1 - \beta(1 - \pi)\gamma))\right].\end{aligned}$$

Then, $(1 - \beta)(R + \beta\Gamma(0))/C \leqslant ((R - I)/(\beta(1 - \pi)(\Gamma(1) - \Gamma(0))) - \pi/(1 - \pi))$.

PROOF. The hypothesis on γ allows the use of Corollary 5.12. The result then follows directly, noting for any admissible ψ, $\Gamma \leqslant \Gamma'$. Q.E.D.

The following theorem states sufficient conditions for there to be no inspection region. The proof, utilizing Corollary 5.12, is a direct generalization of the proof of Theorem 3.7(b) [5] and is omitted.

THEOREM 6.4. *Assume $\gamma < 1/(\beta(1 - \pi))$. Let*

$$\Lambda = \Gamma(1) - \Gamma(0) - C/(1 - \beta(1 - \pi)\gamma).$$

Then if either

$$((R - I)/(\beta(1 - \pi)(\Gamma(1) - \Gamma(0))) - \pi/(1 - \pi)) \leqslant (I + \beta\pi\Lambda/(C - \beta(1 - \pi)\Lambda))$$

or $I - C \geqslant -\beta\Lambda$, there exists an optimal policy ϕ which can be characterized by a number $x_1 \in [0, 1]$ such that $\phi(x) = 0$ for $x \in [0, x_1)$ and $\phi(x) = 2$ for $x \in [x_1, 1]$.

The following result presents conditions sufficient for the hypotheses of Theorem 6.4 to hold.

COROLLARY 6.5. *Let* $\gamma < 1/(\beta(1-\pi))$. *Assume* $\Lambda' = (R - C - C/(1-\beta(1-\pi)\gamma))$ *and* $R < C(1+\beta\pi)/(1-\beta(1-\pi))$. *Then, if either* $((R-I)/(\beta(1-\pi)(R-C)) - \pi/(1-\pi)) \leqslant (I + \beta\pi\Lambda')/(C - \beta(1-\pi)\Lambda')$ *or* $I - C \geqslant -\beta\Lambda'$, *the hypotheses of Theorem* 6.4 *are satisfied.*

PROOF. It follows from (5.1), Corollary 5.6, and Theorem 5.10 that $\Gamma(0) \leqslant \Gamma(1) \leqslant C + \beta\Gamma(1)$ (hence $\Gamma(0) \leqslant C/(1-\beta)$) and that $\Gamma(1) = R + \beta\Gamma(1)$. Thus, $\Gamma(1) - \Gamma(0) \geqslant R - C$, implying $\Lambda \geqslant \Lambda'$. The result then follows directly. Q.E.D.

7. Conclusions

This paper has considered a quality control model where the inspector, in determining whether to produce, inspect with production, or revise, has been allowed on-line data dependent on the production process at both times of production and inspection. §4 presented bounds on optimal cost which were related to the quality of observation available to the inspector at times of production and inspection. The remainder of the paper considered the two-state state space case: §5 allowed q_0 and q_1 to be members of [0.5, 1], whereas §6 restricted q_1 to equal 1, i.e. inspection always results in a perfect observation. Throughout the paper, however, q_0 was only restricted to the interval [0.5, 1]. This generalization of all previous studies of the production problem allows informative data received by the inspector at time of production to be considered in the decision making process. All of the important results, Theorems 5.4, 5.10, 6.1, 6.2, and 6.4, have given characterizations of the structure of the optimal (stationary) policy and indicate the dependence of the optimal cost and the structure of an optimal policy on observation quality. The extensions of these results, e.g. Theorems 5.4 and 5.8, to more general classes of optimization problems would be of particular interest.

It would also be of interest to determine a simple description of an optimal policy for the case where $q_0 < q_1 \neq 1$, thus generalizing Theorem 6.1. Note, however, that the proof of Theorem 6.1 relies in part on the fact that the minimum of a monotonically nondecreasing concave function and a nondecreasing linear function can have *at most* three different regions characterizing an optimal policy. The minimum of two monotonically nondecreasing concave functions does not necessarily imply such a simple characterization. Hence, a result similar to Theorem 6.1 for the $q_0 < q_1 \neq 1$ case may be difficult to obtain and merits further investigation.[1]

[1] The author was in the Department of Electrical Engineering, Southern Methodist University, Dallas, Texas. This research was partially supported by NSF Grant ENG 75-19692.

References

1. BAR-SHALOM, Y. AND TSE, E., "Dual Effect, Certainty Equivalence, and Separation in Stochastic Control," *IEEE A. C. Trans.*, Vol. AC-19, pp. 494–500.
2. BLACKWELL, D., "Discounted Dynamic Programming," *Ann. of Math. Statist.*, Vol. 36 (1965), pp. 226–235.
3. DERMAN, C., "On Optimal Replacement Rules When Changes in State are Markovian," in *Mathematical Optimization Techniques*, R. Bellman, ed., University of California Press, 1963.
4. GIRSHICK, M. A. AND RUBIN, H., "A Bayes' Approach to a Quality Control Model," *Ann. of Math. Statist.*, Vol. 23 (1952), p. 114.
5. ROSS, S. M., "Quality Control Under Markovian Deterioration," *Management Science*, Vol. 17 (1971), pp. 587–596.
6. ROYDEN, H. L., *Real Analysis*, Second Ed., MacMillan, N. Y., 1968.

7. Sawaragi, Y. and Yoshikawa, T., "Discrete-Time Markovian Decision Processes with Incomplete State Information," *Ann. Math. Statist.*, Vol. 41 (1970), pp. 78–86.
8. Tafeen, S., "A Markov Decision Process with Costly Inspection," Technical Report 107, Stanford U., Dept. of Operations Research and Dept. of Statistics (1968).
9. White, C. C., "Application of Two Inequality Results for Concave Functions to a Stochastic Optimization Problem," *J. of Math. Analysis and Appl.*, Vol. 55 (1976), pp. 347–350.
10. Witsenhausen, H. S., "Separation of Estimation and Control for Discrete Time Systems," *Proc. IEEE*, Vol. 59 (1971), pp. 1557–1566.