

多変量ベイズ管理図の適応手法 (II) (Adaptive methods for multivariate Bayesian Control Chart (II))

神奈川大学・理学部 堀口 正之 (Masayuki HORIGUCHI)
Faculty of Science, Kanagawa University

1 はじめに

本報告は、ベイズ事前-事後分析を基とした品質管理法について考察し、その内容は (佐々木, 堀口, 蔵野 [22] の続編 (sequel) である。

ベイズ推定を用いた適応型の品質管理については、多くの研究があり (cf. [2, 12, 17]), 品質管理の現場でその有効性が報告されている。ベイズ推定を基本とした品質管理では、蓄積された情報を基にして管理限界, サンプルサイズおよびサンプリング間隔を変更して事象や状況の変化に適応していく (cf. [17])。

品質管理図を設計する場合, 統計的手法に重きをおくか, あるいは経費的な側面に重きをおくかは重要であるが, それぞれ一長一短があることが知られている (cf. [21])。そこで, 統計的および経済的な両側面を考慮した管理図の作成が考えられるが, これに答えるためには, 問題を逐次決定過程としてとらえ, その解析結果を管理図に反映させる必要がある。この種の研究も多く行われている (cf. [1, 6, 10, 16, 18, 19])。最近, V. Makis[9] は, システムの状態が既知のパラメータ $\theta (\theta > 0)$ をもつ指数分布に従って正常状態 (state in control) から不正常な状態 (state out of control) に移行する多変量管理モデルをマルコフ決定過程 (Markov decision process, MDP)(cf. [3, 11]) として定式化し, 長時間の平均期待コスト基準のもとでの最適な管理政策を求め, これにより, 多変量管理図の作成方法を提案した。[22] では, パラメータ θ が未知の場合の “Makis Model” を取り扱い, 平均コスト基準および割引きされた総期待コスト最小化問題を定式化して適応最適政策を議論した。その際, §2 では, 最尤推定量を用いて未知のパラメータ θ を点推定し, いわゆる推定と制御の原理 (Principle of estimation and control)(cf. [7, 14]) を適用して適応管理図の作成方法を提案した。

ここでは, ベイズ事前-事後分析の方法を用いて, 未知パラメータ θ の事前分布から各期で得られる情報を基に事後分布を計算し, 事後分布の極限定理を適用して有用な適応政策を求めて管理図の作成に役立たせる。第 2 節では, Makis[9] の管理モデルを解説して, [22] で得られた結果を概観する。第 3 節では, ベイズ事前-事後分析による適応モデルの最適化を考察する。

2 ベイズ管理モデル

本論で取り扱う品質管理モデルを述べ, 同値なベイズモデルによって定式化する。

システムの正常な状態を “0”, 不正常な状態を “1” で表す。状態 0 から状態 1 に移行 (故障原因の発生) する時間分布は, パラメータ θ の指数分布とする。 θ の真値は未知で $\theta \in \Theta$ とする。ただし, $\theta_1, \theta_2, \dots, \theta_r$ は互いに異なる正数とする。 θ の事前分布に従う確率

変数を $\tilde{\theta}$ で表す. 時刻 $t (t \geq 0)$ のシステムの状態を X_t で表す. 与えられた定数 $h > 0$ に対して, h の時間間隔で状態に対する情報 (大きさ n の q 次元データ) を取得して, システムの運用を継続 “to continue” するか (この行動を “0” で表す), システムの運用を停止して故障の有無を精査 “to stop and search” するか (この行動を “1” で表す) を選択する. 精査したとき, システムが正常であるか不正常であるかが正確に分かり, もし不正常ならば正常な状態に瞬間的に取り替え, 正常な状態からプロセスは再スタートする. 意思決定者 (decision maker) が, 状態に関する情報を得て, 0 か 1 の行動を選択する決定時点 (decision epoch) は, $ih (i = 1, 2, \dots)$ である.

時点 $ih (i = 1, 2, 3, \dots)$ で取得する多次元正規情報:

大きさ n の q 次元の標本

$$(1) \quad Y_i = \begin{bmatrix} \mathbf{y}_1^i \\ \mathbf{y}_2^i \\ \vdots \\ \mathbf{y}_n^i \end{bmatrix}, \quad \mathbf{y}_j^i = (y_{j1}^i, y_{j2}^i, \dots, y_{jq}^i) \quad j = 1, 2, \dots, n.$$

によって, システムの状態を推測する情報を取得する.

仮定:

$X_{ih} = 0$ (または 1) のとき, $\mathbf{y}_1^i, \mathbf{y}_2^i, \dots, \mathbf{y}_n^i$ は互いに独立で各 \mathbf{y}_j^i は同一の分布 $N_q(\mu_0, \Sigma)$ (または $N_q(\mu_1, \Sigma)$) に従う.

ただし, $N_q(\mu_0, \Sigma), N_q(\mu_1, \Sigma)$ は分散共分散行列 (正値) Σ であり, それぞれ平均ベクトル $\mu_0 = (\mu_{01}, \mu_{02}, \dots, \mu_{0q}), \mu_1 = (\mu_{11}, \mu_{12}, \dots, \mu_{1q})$ をもつ q 次元正規分布で表す. ここで, μ_1 の μ_0 からの M-距離 d_1 について, 次を仮定する:

$$(2) \quad d_1 := [(\mu_1 - \mu_0) \Sigma^{-1} (\mu_1 - \mu_0)]^{\frac{1}{2}} > 0.$$

コスト構造: 次のように, 各費用を考える.

- システムの運用を停止して故障の有無を精査する費用 $A > 0$
- 状態 1 (不正常) を状態 0 (正常) に取り替える費用 $R \geq 0$
- 不正常の状態のまま運用したときの単位時間当たりのコスト $M > 0$
- 大きさ n のサンプルをとる費用 $b + nc$ ($b, c \geq 0$)

この決定過程は, 部分観測可能なマルコフ決定過程としてみることができ. 従って, $X_t = 1$ (時刻 t で状態が不正常) である確率を新しい状態として状態空間 $S = [0, 1]$ に拡張して, ベイズの定理により事前状態分布を事後状態分布に変換することにより状態の推移を記述する完全観測のベイズモデルに同値に変換される (cf. [7, 20]).

同値なベイズモデル:

行動の決定時点は $ih (i = 1, 2, \dots)$ で次の要素からなる MDP モデルを考える.

$$\begin{cases} S = [0, 1]: \text{状態空間} \\ A = \{0, 1\}: \text{行動空間} \\ \Theta \subset (-\infty, \infty): \text{パラメータ空間} \\ c(p, a): p \in S, a \in A \text{のときのコスト} \end{cases}$$

標本空間を $\bar{\Omega} = \Theta \times \Omega, \Omega = S \times (A \times S)^\infty$ と表し、プロセスを表す確率変数を $\tilde{\theta}, \tilde{p}_0, \tilde{a}_0, \tilde{p}_1, \tilde{a}_1, \dots$ とする。すなわち、 $\bar{\Omega} \ni \omega = (\theta, p_0, a_0, p_1, a_1, p_2, \dots)$ のとき、 $\tilde{\theta}(\omega) = \theta, \tilde{p}_0(\omega) = p_0, \tilde{a}_0(\omega) = a_0, \tilde{p}_1(\omega) = p_1, \dots$ である。ただし、 $p_0 = 0$ として一般性を失わない。

mh 時点で状態 $\tilde{p}_m = p$ のとき、行動 $\tilde{a}_m = 0$ (または 1) を選択し $(m+1)h$ 時点で $Y_{m+1} = y^{m+1}$ を観測した場合には、 $(m+1)h$ 時点の状態は

$$(3) \quad \tilde{p}_{m+1} = T(p, y^{m+1}, 0) \text{ (または } T(p, y^{m+1}, 1))$$

に推移する、ただし、ベイズの定理により事前-事後ベイズ作用素 T は次のように定まる (Lemma 1[9])。

$$(4) \quad \begin{cases} T(p, z, 0) = (1 - (1-p)e^{-\theta h} h_1(z)) / h(z|p), \\ T(p, z, 1) = T(0, z, 0), \\ \text{ただし,} \\ z = 2 \sum_{j=1}^n (\mathbf{y}_j - \mu_0) \Sigma^{-1} (\mu_0 - \mu_1)^T, \\ \mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_n \end{bmatrix}, \mathbf{y}_j = (y_{j1}, y_{j2}, \dots, y_{jq}), \\ h(z|p) = (1 - (1-p)e^{-\theta h}) h_1(z) + (1-p)e^{-\theta h} h_0(z), \\ h_0(z) = N_1(0, 4nd_1^2), h_1(z) = N_1(-2nd_1^2, 4nd_1^2). \end{cases}$$

注: z は十分統計量で状態の推移は z の値のみに依存する。 $X_{ih} = 0$ (または 1) のとき、 z は平均 0 (または $-2nd_1^2$)、分散 $4nd_1^2$ の 1 次元正規分布に従う。

θ が真のときのコストは次で与えられる:

$$(5) \quad \begin{cases} c(p, 0) = ME(\int_0^h I_{\{X_s=1\}} ds) + b + nc = M \left[h - \frac{1-p}{\theta} (1 - e^{-\theta h}) \right] + b + nc, \\ c(p, 1) = c_1(p) + c(0, 0). \end{cases}$$

ただし、 $c_1(p) = A + Rp$.

レート (rate) $\alpha > 0$ で連続的に割引きされる (割引きコスト基準) 場合:

$$(6) \quad \begin{cases} c_\alpha(p, 0) = ME(\int_0^h I_{\{X_s=1\}} e^{-\alpha s} ds) + e^{-\alpha h} (b + nc) = MB_\theta + \beta(b + nc), \\ c_\alpha(p, 1) = c_1(p) + c_\alpha(0, 0). \end{cases}$$

ただし,

$$(7) \quad \begin{cases} B_\theta = \frac{1}{\alpha} \left\{ \frac{\theta}{\theta+\alpha} (1 - e^{-(\theta+\alpha)h}) - \beta(1 - e^{-\theta h}) \right\}, \\ \beta = e^{-\alpha h}. \end{cases}$$

政策 (policy) は $\pi = (\pi_0, \pi_1, \dots)$ で表し, $\pi_m(H_m) \in A = \{0, 1\} (m \geq 0)$, $H_m = (\tilde{p}_0, \tilde{a}_0, \tilde{p}_1, \dots, \tilde{p}_m)$ とする. 政策の全体を Π で表す.

$\mathcal{P}(\Theta)$ を Θ 上の確率分布の全体とする. 任意の政策 $\pi = (\pi_0, \pi_1, \dots) \in \Pi$ に対して, 停止時刻 (stopping time) の系列 $\tau = (\tau_0, \tau_1, \tau_2, \dots)$ が次によって定まる.

$$\tau_0 = 0, \tau_k = \min\{k | \pi_k(H_k) = 1, k > \tau_{k-1}\}.$$

明らかに, 政策 π と停止時刻の系列 (停止政策と呼ぶ) τ は, 1 対 1 に対応する. 従って, 以降では必要に応じて, 政策と対応する停止政策を同一視して取り扱う.

Z_m を mh 時点での観測データを表す確率変数とする. ただし, $X_{mh} = 0$ のとき Z_m は確率密度関数 $h_0(z) = N_1(\theta, 4nd_1^2)(z)$ に従い, $X_{mh} = 1$ のとき $h_1(z) = N_1(-2nd_1^2, 4nd_1^2)(z)$ に従う.

$\tilde{p}_0 = 0, \tilde{\theta} = \theta$ において,

$$a_0 = a_1 = \dots = a_m = 0, Z_1 = z_1, Z_2 = z_2, \dots, Z_m = z_m$$

のときの故障事後確率を δ_m とする. このとき, 次が成り立つことを証明しよう.

Lemma 2.1

$$(8) \quad \delta_m = \frac{h_1(z_1, z_2, \dots, z_m)}{h(z_1, z_2, \dots, z_m)} \text{ かつ } \tilde{p}_m = \delta_m \quad (m \geq 1)$$

ただし,

$$(9) \quad \begin{cases} h_1(z_1, \dots, z_m) = \sum_{l=1}^m (e^{-(l-1)\theta h} - e^{-l\theta h}) h_0(z_1) \cdots h_0(z_{l-1}) h_1(z_l) \cdots h_1(z_m) \\ h(z_1, \dots, z_m) = h_1(z_1, \dots, z_m) + e^{-m\theta h} h_0(z_1) \cdots h_0(z_m) \end{cases}$$

証明: Z_1, Z_2, \dots, Z_m の確率密度関数を $h(z_1, z_2, \dots, z_m)$ とすれば, 区間 $((l-1)h, lh]$ 内に故障が起こる確率は $\int_{(l-1)h}^{lh} \theta e^{-\theta t} dt = e^{-(l-1)\theta h} - e^{-l\theta h}$ であるから, 式 (9) が成り立つ. ベ

イズの定理により δ_m が定まる. Lemma 2.1 の後半については, 帰納法で示す.

明らかに, $\tilde{p}_1 = \delta_1$ が成り立つ. $m-1$ のとき, 式 (8) が成り立つと仮定する. 式 (3) より

$$(10) \quad \tilde{p}_m = \frac{\{\tilde{p}_{m-1} h_1(z_m) + (1 - \tilde{p}_{m-1})(1 - e^{-\theta h}) h_1(z_m)\}}{h(z_m | \tilde{p}_{m-1})}.$$

ただし,

$$(11) \quad h(z_m|\tilde{p}_{m-1}) = \tilde{p}_{m-1}h_1(z_m) + (1 - \tilde{p}_{m-1})(1 - e^{-\theta h})h_1(z_m) + (1 - \tilde{p}_{m-1})e^{-\theta h}h_0(z_m).$$

式(11)に $\tilde{p}_{m-1} = \delta_{m-1}$ を代入すると,

$$(12) \quad \begin{aligned} & h(z_1, \dots, z_{m-1})h(z_m|\tilde{p}_{m-1}) \\ &= \sum_{l=1}^{m-1} (e^{l-1}\theta h - e^{-l\theta h})h_0(z_1) \cdots h_0(z_{l-1})h_1(z_l) \cdots h_1(z_{m-1}) \\ & \quad + (e^{-(m-1)\theta h} - e^{-m\theta h})h_0(z_1) \cdots h_0(z_{m-1})h_1(z_m) \\ & \quad + e^{-m\theta h}h_0(z_1) \cdots h_0(z_m) \\ &= h(z_1, z_2, \dots, z_m). \end{aligned}$$

同様に, 式(10)から

$$h(z_1, \dots, z_{m-1})h(z_m|\tilde{p}_{m-1})\tilde{p}_m = h_1(z_1, z_2, \dots, z_m)$$

を得る. 故に,

$$\tilde{p}_m = \frac{h_1(z_1, \dots, z_m)}{h(z_1, \dots, z_m)}$$

が成り立つことが示された. ■

$\tilde{\theta} = \theta \in \mathcal{P}(\Theta)$ と初期状態分布 $p_0 = p \in S$ が与えられたときの政策 $\pi \in \Pi$ の平均期待コスト $\varphi(\pi|\theta, p_0)$ を次で定める:

$$(13) \quad \varphi(\pi|\theta, p_0) = \limsup_{k \rightarrow \infty} \frac{1}{E(\tau_k)} E \left[\sum_{m=0}^{\tau_k} c(\tilde{p}_m, \tilde{a}_m) | \theta, p_0 \right],$$

ただし, $\pi = (\tau_0, \tau_1, \tau_2, \dots)$.

さらに, 割引きされた総期待コスト $v(\pi|\theta, p_0)$ は次で定める:

$$(14) \quad v(\pi|\theta, p_0) = \sum_{m=0}^{\infty} \beta^m E_{\pi} [c_{\alpha}(\tilde{p}_m, \tilde{a}_m) | \theta, p_0].$$

ただし, $\beta = e^{-\alpha h}$ は割引き率を表し, $E_{\pi} [\cdot | \theta, p]$ は, θ, p および π が与えられたときの $\bar{\Omega}$ 上に定まる確率測度 $P_{\pi}(\cdot | \theta, p)$ に関する期待値である.

$\varphi(\pi|\theta, p), v(\pi|\theta, p)$ を最小にする政策 $\pi \in \Pi$ をそれぞれ θ -平均最適, θ -割引き最適という. 次の定理が成り立つ.

Theorem 2.1 (V. Makis[9]) $A + R < \frac{M}{\theta}$ ならば, *control-limit*型の θ -平均最適な政策が存在する. すなわち, $p_\theta^* \in (0, 1)$ が存在して, 決定関数 $f_\theta : S \rightarrow A$,

$$(15) \quad f_\theta(p) = \begin{cases} 0 & \text{if } p < p_\theta^*, \\ 1 & \text{if } p \geq p_\theta^*. \end{cases}$$

による管理政策が θ -平均最適となる.

Theorem 2.2 (佐々木, 堀口, 蔵野 [22]) *control-limit*型の θ -割引き最適な政策が存在する. すなわち, $\bar{p}_\theta \in (0, 1)$ が存在して, 最適な決定関数 $g_\theta : S \rightarrow A$ は次で与えられる:

$$(16) \quad g_\theta(p) = \begin{cases} 0 & \text{if } p < \bar{p}_\theta, \\ 1 & \text{if } p \geq \bar{p}_\theta. \end{cases}$$

3 適応管理モデル

ここでは, ベイズ事前-事後解析を基にした管理法を考察する. 扱うモデルは m 期で, システムの稼働を停止し精査する (to stop and research, $a_m = 1$) と, システムは必要ならば新品と取り換えてシステムの状態 (故障している確率) は 0 (ゼロ) になる. “Stop” と “stop” の期間を one cycle としてみると, この cycle の繰り返し (renewal) となる. 従って, one cycle での未知パラメータ θ の情報をひとまとまりとして処理する事後分布の計算は, 逐次的になされるので便利である.

観測データの系列 Z_0, Z_1, \dots に対する停止時刻 (stopping time) σ で上に有界, すなわち, $P(1 \leq \sigma \leq M) = 1$ なる $M > 0$ が存在するとする.

“Stop” すると, $X_\sigma h$ の値を知ることができるので, 1 期から σ 期までの全情報は $(Z_1, Z_2, \dots, Z_\sigma, X_{\sigma h})$ である. $\hat{\theta} = \theta$ のとき, $Z_1 = z_1, Z_2 = z_2, \dots, Z_\sigma = z_\sigma, X_{\sigma h} = x$ の確率密度関数は,

$$(17) \quad f(z_1, z_2, \dots, z_\sigma, x | \theta) = f_0(z_1, z_2, \dots, z_\sigma | \theta) I_{\{0\}}(x) + f_1(z_1, z_2, \dots, z_\sigma | \theta) I_{\{1\}}(x)$$

ただし,

$$(18) \quad f_0(z_1, z_2, \dots, z_\sigma | \theta) = \int_{\sigma h}^{\infty} \theta e^{-\theta t} dt h_0(z_1) \cdots h_0(z_\sigma) = e^{-\sigma \theta h} h_0(z_1) \cdots h_0(z_\sigma),$$

$$\begin{aligned}
(19) \quad f_1(z_1, z_2, \dots, z_\sigma | \theta) &= \sum_{l=1}^{\sigma} \int_{(l-1)h}^{lh} \theta e^{-\theta t} dt h_0(z_1) \cdots h_0(z_l) h_1(z_{l+1}) \cdots h_1(z_\sigma) \\
&= \sum_{l=1}^{\sigma} (e^{-(l-1)\theta h} - e^{-l\theta h}) h_0(z_1) \cdots h_0(z_l) h_1(z_{l+1}) \cdots h_1(z_\sigma).
\end{aligned}$$

(Lemma 2.1 参照.)

記号の簡略のために、次を定義する:
 政策 $\tau = (\tau_1, \tau_2, \dots)$ に対して,

$$(20) \quad \sigma_l := \tau_l - \tau_{l-1} (l \geq 1),$$

$$(21) \quad D_{\sigma_l} := (Z_{\tau_{l-1}+1}, Z_{\tau_{l-1}+2}, \dots, Z_{\tau_l}, X_{\tau_l h}),$$

$$(22) \quad = (Z_{\tau_{l-1}+1}, Z_{\tau_{l-1}+2}, \dots, Z_{\tau_{l-1}+\sigma_l}, X_{(\tau_{l-1}+\sigma_l)h}),$$

$$(23) \quad D'_{\sigma_l} := (Z_{\tau_{l-1}+1}, \dots, Z_{\tau_l}).$$

ここに, $D_{\sigma_l} = (D'_{\sigma_l}, X_{\tau_l h})$ と表せる.

式(17)より, 次を得る.

$$f(D_{\sigma_l} | \theta) = f_0(D_{\sigma_l} | \theta) I_0(X_{\tau_l h}) + f_1(D_{\sigma_l} | \theta) I_1(X_{\tau_l h}).$$

D_{σ_l} の実現値を $d_{\sigma_l} = (z_{\tau_{l-1}+1}, z_{\tau_{l-1}+2}, \dots, z_{\tau_l}, x_{\tau_l h})$ で表す.
 k cycle までのデータ $(d_{\sigma_1}, d_{\sigma_2}, \dots, d_{\sigma_k})$ による θ の尤度は

$$(24) \quad p(\theta | d_{\sigma_1}, d_{\sigma_2}, \dots, d_{\sigma_k}) = \prod_{l=1}^k f(d_{\sigma_l} | \theta).$$

従って, $\tilde{\theta}$ の分布 ρ (θ の事前分布) に対する θ の事後分布 ρ_k は次で与えられる.

$$(25) \quad \rho_k(\theta | d_{\sigma_1}, d_{\sigma_2}, \dots, d_{\sigma_k}) \propto p(\theta | d_{\sigma_1}, d_{\sigma_2}, \dots, d_{\sigma_k}) p(\theta), \quad (k \geq 1)$$

明らかに, 次の漸化式が成り立つ.

$$(26) \quad \begin{cases} \rho_k(\theta | d_{\sigma_1}, d_{\sigma_2}, \dots, d_{\sigma_k}) \propto f(d_{\sigma_k} | \theta) \rho_{k-1}(\theta | d_{\sigma_1}, d_{\sigma_2}, \dots, d_{\sigma_{k-1}}), & (k \geq 2) \\ \rho_1(\theta | d_{\sigma_1}) \propto f(d_{\sigma_1} | \theta) \rho(\theta). \end{cases}$$

有界な停止時刻 σ に対して separation property:

$$(27) \quad P_\tau(f(Z_1, Z_2, \dots, Z_\sigma, X_{\sigma h}|\theta_1) \neq f(Z_1, Z_2, \dots, Z_\sigma, X_{\sigma h}|\theta_2)) > 0, \quad (\theta_1 \neq \theta_2)$$

が成り立つ。ただし、 P_τ は、policy τ が与えられたときの process 全体の確率分布を表す。

$0 < a < b$ に対して、 $\Theta = [a, b]$ として閉区間 $[a, b]$ の確率分布全体を $\mathcal{P}[a, b]$ で表す。このとき、次が成り立つ。

Lemma 3.1 (cf. Th. 2.4 in [20]) $\rho \in \mathcal{P}[a, b]$ とする。 $[a, b]$ 上の任意の有界な可測関数 s に対して、

$$(i) \quad \lim_{k \rightarrow \infty} \int_a^b s(\theta) \rho_k(d\theta) = s(\tilde{\theta}), \quad P_\tau\text{-a.s.},$$

$$(ii) \quad \lim_{k \rightarrow \infty} \bar{\theta}_k = \lim_{k \rightarrow \infty} \int_a^b \theta \rho_k(d\theta) = \tilde{\theta}. \quad P_\tau\text{-a.s.},$$

$\tilde{\theta}$ の分布 ρ (事前分布) は $\rho \in \mathcal{P}[a, b]$ とする。政策 π が平均最適な適応政策であるとは、 ρ -a.s. に次が成り立つ場合をいう：

$$(28) \quad \varphi(\bar{\pi}|\tilde{\theta}, p_0) \leq \varphi(\pi|\tilde{\theta}, p_0), \quad \pi \in \Pi.$$

任意の $k \geq 1$ に対して、 k cycle までの θ の事後分布 ρ_k の平均値 $\bar{\theta}_k$ に対して、 $(k+1)$ cycle では $\bar{\theta}_k$ -平均最適な停止時刻 $\tau^*(\bar{\theta}_k)$ を用いる政策を $\bar{\pi}^*$ とする、ただし、 $\tau^*(\theta)$ は θ に依存して Theorem 2.1 で定まる θ -平均最適停止時刻である。Lemma 3.1 を応用して次が得られる。

Theorem 3.1 $\bar{\pi}^*$ は平均最適な適応政策である。

$k \geq 1$ に対して、次を定める：

$$v_k(\pi|\theta, p_0) = E_\pi \left[\sum_{t=k}^{\infty} \beta^{t-k} c(\tilde{p}_t, \tilde{a}_m) | p_0, \theta \right].$$

$\bar{\pi} = (\bar{\pi}_0, \bar{\pi}_1, \bar{\pi}_2, \dots)$ が asymptotically discount optimal (漸近的割引き最適) であるとは、 ρ -a.s. に次が成り立つときをいう。

$$(29) \quad v_{\bar{\pi}_k}(\bar{\pi}|\tilde{\theta}, p_0) \rightarrow v(\tilde{\theta}|p_0) \quad (k \rightarrow \infty),$$

ただし、 $\bar{\pi}^*$ は θ -割引き最適政策である。

任意の $k \geq 1$ に対して、 k cycle までの θ の事後分布 ρ_k の平均値 $\bar{\theta}_k$ に対して $(k+1)$ cycle では $\bar{\theta}_k$ -割引き最適な停止時刻 $\tilde{\tau}(\bar{\theta}_k)$ を使う政策を $\tilde{\pi}^*$ とする。ただし、 $\tilde{\tau}(\theta)$ は θ に依存して、Theorem 2.2 で定まる θ -割引き最適停止時刻である。このとき、次が成り立つ。

Theorem 3.2 $\bar{\pi}^*$ は, *asymptotically discount optimal* である.

注意: 本論文で扱った適応モデルについて, その各決定時刻に対する事前-事後分布の関係にはマルコフ性が存在しないために, 同値な完全情報の一般マルコフ決定過程 (cf. [13]) を構成することは出来なかった. 従って, ここでは「推定と制御の原理」を適用する方法以外に有効な管理図作成方法は提案されていない.

References

- [1] J. A. Bather. Control charts and minimization of costs. *J. Roy. Statist. Soc. Ser. B*, 25:49–80, 1963.
- [2] Robert V. Baxley, Jr. An application of variable sampling interval control charts. *Journal of Quality Technology*, 27(4):275–282, 1995.
- [3] Dimitri P. Bertsekas and Steven E. Shreve. *Stochastic optimal control*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1978. The discrete time case.
- [4] Morris H. DeGroot. *Optimal statistical decisions*. McGraw-Hill Book Co., New York, 1970.
- [5] T. S. Ferguson. Optimal stopping and applications (electronic texts). <http://www.math.ucla.edu/~tom/Stopping/Contents.html>.
- [6] M. A. Girshick and Herman Rubin. A Bayes approach to a quality control model. *Ann. Math. Statistics*, 23:114–125, 1952.
- [7] O. Hernández-Lerma. *Adaptive Markov control processes*, volume 79 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1989.
- [8] Masami Kurano. Adaptive policies in Markov decision processes with uncertain transition matrices. *J. Inform. Optim. Sci.*, 4(1):21–40, 1983.
- [9] Viliam Makis. Multivariate Bayesian control chart. *Oper. Res.*, 56(2):487–496, 2008.
- [10] E. L. Porteus and A. Angelus. Opportunities for improved statistical process control. *Management Sci.*, 43:1214–1228, 1997.
- [11] Martin L. Puterman. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons Inc., New York, 1994. A Wiley-Interscience Publication.
- [12] Marion R. Reynolds, Jr., Jesse C. Arnold, Raid W. Amin, and Joel A. Nachlas. \bar{X} charts with variable sampling intervals. *Technometrics*, 30(2):181–192, 1988.
- [13] Ulrich Rieder. On stopped decision processes with discrete time parameter. *Stochastic Processes Appl.*, 3(4):365–383, 1975.

- [14] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA, 1998.
- [15] George Tagaras. A dynamic programming approach to the economic design of X -charts. *IIE Trans.*, 26(3):48–56, 1994.
- [16] George Tagaras. Dynamic control charts for finite production runs. *European J. Oper. Res.*, 91:38–55, 1998.
- [17] George Tagaras and Yiannis Nikolaidis. Comparing the effectiveness of various Bayesian \bar{X} control charts. *Oper. Res.*, 50(5):878–888, 2002.
- [18] Howard M. Taylor. Markovian sequential replacement processes. *Ann. Math. Statist.*, 36:1677–1694, 1965.
- [19] Howard M. Taylor. Statistical control of a Gaussian process. *Technometrics*, 9:29–41, 1967.
- [20] K. M. van Hee. *Bayesian control of Markov chains*, volume 95 of *Mathematical Centre Tracts*. Mathematisch Centrum, Amsterdam, 1978.
- [21] William H. Woodall, Thomas J. Lorenzen, and Lonnie C. Vance. Weaknesses of the economic design of control charts. *Technometrics*, 28(4):408–410, 1986.
- [22] 佐々木稔, 堀口正之, 蔵野正美. 多変量ベイズ管理図の適応手法. 数理解析研究所講究録 1912 「不確実性の下での数理的意思決定の理論と応用」, pages 181–192, 2014.