

# マルコフ決定過程におけるベイズ管理モデルについて (Bayesian statistical-design model in Markov Decision Processes)

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

## 1 はじめに

先行研究 ([23],[24]) では、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た。本報告では、それらの中で示されている故障確率  $\delta$  のベイズ更新と無情報事前分布での事後分布について、具体例とともに示す。

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, 13, 18]), 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. [18]).

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. [22]). 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, 11, 17, 19, 20]).

本報告では, これまでの我々の先行研究と同様に, V. Makis[10] による逐次決定モデルについて, システムの故障時間分布のパラメータが未知の場合を考察している。[10] では, 具体的には, 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, 12]) 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。[23] では, パラメータ  $\theta$  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  $\theta$  を点推定し, いわゆる推定と制御の原理 (Principle of estimation and control)(cf. [7, 15]) を適用して適応管理図の作成方法を提案した。また, [24] 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  $\theta$  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ついて提案した。本報告では, 事前-事後分析について具体例を挙げて考察する。

## 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。システムの状態空間を  $\tilde{S} = \{0, 1\}$  とおく。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  $\theta$  の指数分布とする。  $\theta$  の真値は未知であると仮定し,  $\theta \in \Theta$  とする。ただし,  $\Theta = \{\theta_1, \theta_2, \dots, \theta_r\}$  は互いに異なる正数とする。  $\theta$  の事前分布  $\rho$  に従う確率変数を  $\hat{\theta}$  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。従って、行動空間は  $A = \{0(\text{to continue}), 1(\text{to stop and search})\}$  である。精査(行動“1”を選択)したとき、システムが正常(状態“0”)であるか不正常(状態“1”)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

$$(1) \quad Y_i = \begin{bmatrix} y_1^i \\ y_2^i \\ \vdots \\ y_n^i \end{bmatrix}, \quad y_j^i = (y_{j1}^i, y_{j2}^i, \dots, y_{jq}^i) \quad j = 1, 2, \dots, n.$$

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、 $y_1^i, y_2^i, \dots, y_n^i$  は互いに独立で各  $y_j^i$  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) $\Sigma$  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  $q$  次元正規分布を表す。ここで、 $\mu_1$  の  $\mu_0$  からの M-距離  $d_1$  について、次を仮定する:

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

以下のようなコスト構造のもとでマルコフ決定モデルとして管理図を考える。**コスト構造:** 次のように、各費用を考える。(1) システムの運用を停止して故障の有無を精査する費用  $A > 0$ , (2) 状態 1(不正常)を状態 0(正常)に取り替える費用  $R \geq 0$ , (3) 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  $M > 0$ , (4) 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、 $X_t = 1$  である確率  $\delta$  を新しい状態とした状態空間  $S = [0, 1](\{\delta | 0 \leq \delta \leq 1\})$  の完全観測のベイズモデルに同値に変換される (cf. [7, 21])。

**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。

(1)  $S = [0, 1]$ : 状態空間, (2)  $A = \{0, 1\}$ : 行動空間, (3)  $\Theta \subset (-\infty, \infty)$ : パラメータ空間, (4)  $c(p, a)$ :  $p \in S, a \in A$  のときのコスト。

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。

$m$  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[10])。

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

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.

$\theta$  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  $\Pi$  で表す.

$\mathcal{P}(\theta)$  を  $\theta$  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  $\pi$  と停止時刻の系列 (停止政策と呼ぶ)  $\tau$  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.

$\tilde{p}_0 = 0, \theta = \theta$  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  $\delta_m$  とする. このとき, 次が成り立つ.

**Lemma 2.1** ([24])

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

ただし,

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

$\bar{\theta} = \theta \in \mathcal{P}(\Theta)$  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:

$$(10) \quad \varphi(\pi|\theta, p_0) = \limsup_{k \rightarrow \infty} \frac{1}{E(\tau_k)} E \left[ \sum_{m=0}^{\tau_k} c(\bar{p}_m, \bar{a}_m) | \theta, p_0 \right],$$

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:

$$(11) \quad v(\pi|\theta, p_0) = \sum_{m=0}^{\infty} \beta^m E_{\pi} [c_{\alpha}(\bar{p}_m, \bar{a}_m) | \theta, p_0].$$

ただし,  $\beta = e^{-\alpha h}$  は割引き率を表し,  $E_{\pi}[\cdot | \theta, p]$  は,  $\theta, p$  および  $\pi$  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  $\theta$ -平均最適,  $\theta$ -割引き最適という. 次の定理が成り立つ.

**Theorem 2.1 (V. Makis[10])**  $A + R < \frac{M}{\theta}$  ならば, *control-limit* 型の  $\theta$ -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  $f_{\theta}: S \rightarrow A$ ,

$$(12) \quad f_{\theta}(p) = \begin{cases} 0 & \text{if } p < p_{\theta}^*, \\ 1 & \text{if } p \geq p_{\theta}^*. \end{cases}$$

による管理政策が  $\theta$ -平均最適となる.

**Theorem 2.2 (佐々木, 堀口, 蔵野 [23])** *control-limit* 型の  $\theta$ -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:

$$(13) \quad g_{\theta}(p) = \begin{cases} 0 & \text{if } p < \bar{p}_{\theta}, \\ 1 & \text{if } p \geq \bar{p}_{\theta}. \end{cases}$$

### 3 ベイズ事前-事後解析の具体例

ここでは, ベイズ事前-事後解析を基にした管理法について, 故障確率  $\delta$  のベイズ更新と無情報事前分布での事後分布の考察をする.

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  $\theta$  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)  $\sigma$  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 $\sigma$ 期までの全情報は $(Z_1, Z_2, \dots, Z_\sigma, X_{\sigma h})$ である。 $\tilde{\theta} = \theta$ 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、

$$(14) \quad f(z_1, z_2, \dots, z_\sigma, x|\theta) = f_0(z_1, z_2, \dots, z_\sigma|\theta)I_{\{0\}}(x) + f_1(z_1, z_2, \dots, z_\sigma|\theta)I_{\{1\}}(x)$$

ただし、

$$(15) \quad f_0(z_1, z_2, \dots, z_\sigma|\theta) = \int_{\sigma h}^{\infty} \theta e^{-\theta t} dt h_0(z_1) \cdots h_0(z_\sigma) = e^{-\sigma\theta h} h_0(z_1) \cdots h_0(z_\sigma),$$

$$(16) \quad \begin{aligned} f_1(z_1, z_2, \dots, z_\sigma|\theta) &= \sum_{l=1}^{\sigma} \int_{(l-1)h}^{lh} \theta e^{-\theta t} dt h_0(z_1) \cdots h_0(z_l) h_1(z_{l+1}) \cdots h_1(z_\sigma) \\ &= \sum_{l=1}^{\sigma} \left( e^{-(l-1)\theta h} - e^{-l\theta h} \right) h_0(z_1) \cdots h_0(z_l) h_1(z_{l+1}) \cdots h_1(z_\sigma). \end{aligned}$$

である。すなわち、システムを停止して検査を実行する時点までの尤度 $f$ を、システムのそれぞれの状態に応じた尤度 $(f_0(\cdot|\theta), f_1(\cdot|\theta))$ の混合分布で表す。

時点 $kh$ でのパラメータの推定 $\theta_k$ とその真の値への収束性、および適応管理モデルに関して、いわゆる「推定と制御の原理」(Principle of estimation and control)によるMDPsでの最適化手法については、先行研究の[24]を参照されたい。

ここでは、前述の尤度関数 $f(\cdot|\theta)$ とそのベイズ更新、また、故障状態の事後確率について、具体的に見ていく。

V.Makis([10])での一つの例題の設定と同じ正規モデルを扱う。観測データ $\mathbf{y}$ は $q=2$ 次元で標本数 $n=1$ とし、2次元正規分布の分散共分散行列は $\Sigma = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$ とする。 $\mu_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mu_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ とする。このとき、 $d_1^2 = 2$ である。

$z = 2(\mathbf{y} - \mu_0)\Sigma^{-1}(\mu_0 - \mu_1)$ が観測データの十分統計量であることから、 $z_k$ を各時点 $kh$ での $z$ の値とする。 $z$ の分布については、システムが正常ならば $N_1(0, 2\sqrt{2})$ 、不正常ならば $N_1(-4, 2\sqrt{2})$ の(1次元)正規分布に従う。

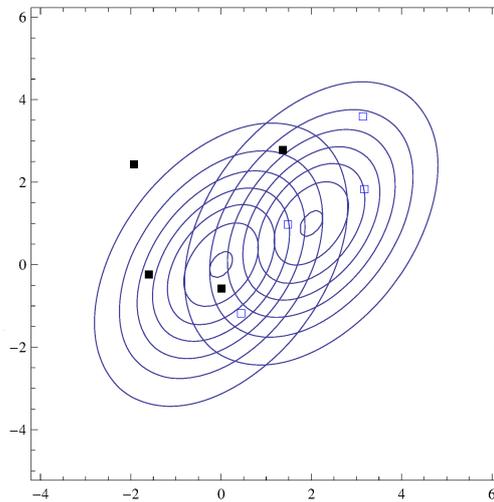
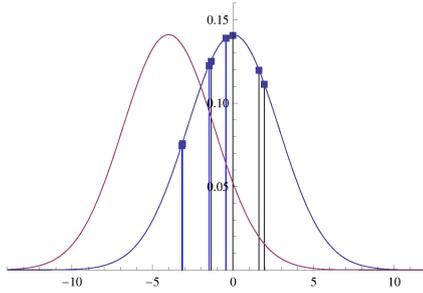
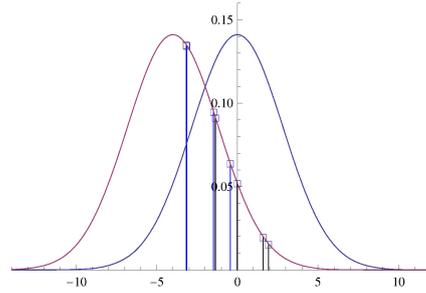


Figure 1: 2次元正規分布の等高線図と生成された乱数の散布図

Figure 2:  $h_0$  の密度関数上の尤度の値Figure 3:  $h_1$  の密度関数上の尤度の値

実際に、2次元正規分布に従う正規乱数を使って、尤度および事後確率の計算を試みる (Figure 1, 2, 3). システムの検査が可能な時間間隔は  $h = 2$  とし、次のデータは時刻  $ih = 2, 4, 6, 8, 10, 12, 14, 16$  のときの観測データを表す. これらのデータは、前半の4つを  $N_2(\mu_0, \Sigma)$ 、後半4つを  $N_2(\mu_1, \Sigma)$  に従うものから得たものである.

$$\mathbf{y}^1 = \begin{pmatrix} 0.43 \\ -0.06 \end{pmatrix}, \mathbf{y}^2 = \begin{pmatrix} -0.80 \\ 2.15 \end{pmatrix}, \mathbf{y}^3 = \begin{pmatrix} -0.27 \\ -1.07 \end{pmatrix}, \mathbf{y}^4 = \begin{pmatrix} 0.26 \\ -2.27 \end{pmatrix},$$

$$\mathbf{y}^5 = \begin{pmatrix} 2.82 \\ 1.09 \end{pmatrix}, \mathbf{y}^6 = \begin{pmatrix} 5.25 \\ 1.76 \end{pmatrix}, \mathbf{y}^7 = \begin{pmatrix} 0.64 \\ -1.21 \end{pmatrix}, \mathbf{y}^8 = \begin{pmatrix} 1.89 \\ -0.62 \end{pmatrix}.$$

これらの値から、 $z_i, h_0(z_i), h_1(z_i)$  の値を求めてまとめたものが Table 1 である. また、 $\theta_1 = \frac{1}{6}, \theta_2 = \frac{1}{8}, \theta_3 = \frac{1}{10}, \theta_4 = \frac{1}{12}, \theta_5 = \frac{1}{14}$  のそれぞれにおける  $f_0(\cdot|\theta), f_1(\cdot|\theta)$  の値の表が Table 2, 3 である.

事後分布  $\delta_k$  については、により、例えば、無情報事前分布として  $P(\theta = \theta_i) = \frac{1}{5} = 0.2$  とした時の  $i = 5$  時点での事後分布  $\delta_5 = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$  は  $(0.23, 0.22, 0.20, 0.18, 0.17)$  として得られる. また、各パラメータが真であるときの時点  $kh$  での故障確率  $\delta_k$  は Lemma 2.1 の式 (8) により、例えば、 $\theta = \theta_1$  であるとき、 $i = 5$  から  $i = 8$  までそれぞれ、0.0024, 0.0287, 0.2803, 0.8842 と変化していることが求まる.

継続 (action '0') を取り続けて停止 (action '1') を行動選択する時点までの尤度に関する最尤推定  $\tilde{\theta}_k$  (cf. [23]) については、次の式から、 $i = 5$  時点での尤度が調べられる.

任意の停止政策  $\tau = (\tau_1, \tau_2, \dots)$  と  $\theta \in \Theta$  に対して、次を定義する:

$$f_k^{\tau k}(0|\theta) := P(X_{s_k h} = 0|\theta, p_0) = e^{-\theta h \tau_k},$$

$$f_k^{\tau k}(1|\theta) := P(X_{s_k h} = 1|\theta, p_0) = 1 - e^{-\theta h \tau_k}.$$

時点  $s_k h$  での  $\theta$  の最尤推定  $\tilde{\theta}_k$  は、 $X_{s_1 h} = x_1, X_{s_2 h} = x_2, \dots, X_{s_k h} = x_k$  のとき、

$$\tilde{\theta}_k(x_1, x_2, \dots, x_k) = \arg \max_{\theta \in \Theta} \prod_{l=1}^k f_l^{\tau_l}(x_l|\theta)$$

で与えられる. ただし、 $\tau_l$  は  $(x_1, x_2, \dots, x_{l-1})$  に依存してもよい. 例題ではこの時、 $\tilde{\theta}_5 = \frac{1}{2} \log \frac{5}{4}$  を得る.

## References

- [1] J. A. Bather. Control charts and minimization of costs. *J. Roy. Statist. Soc. Ser. B*, 25:49–80, 1963.

|         | $z_i$ | $h_0(z_i)$ | $h_1(z_i)$ |
|---------|-------|------------|------------|
| $i = 1$ | 1.60  | 0.12       | 0.02       |
| $i = 2$ | 1.93  | 0.11       | 0.02       |
| $i = 3$ | -0.01 | 0.14       | 0.05       |
| $i = 4$ | -1.37 | 0.13       | 0.09       |
| $i = 5$ | -3.14 | 0.08       | 0.13       |
| $i = 6$ | -0.44 | 0.14       | 0.06       |
| $i = 7$ | -1.48 | 0.12       | 0.09       |
| $i = 8$ | -3.17 | 0.08       | 0.14       |

Table 1:  $z_i, h_0(z_i), h_1(z_i)$  の値

|         | $\theta_1$            | $\theta_2$            | $\theta_3$            | $\theta_4$            | $\theta_5$               |
|---------|-----------------------|-----------------------|-----------------------|-----------------------|--------------------------|
| $i = 1$ | 0.09                  | 0.09                  | 0.10                  | 0.10                  | 0.10                     |
| $i = 2$ | 0.007                 | 0.008                 | 0.009                 | 0.010                 | 0.010                    |
| $i = 3$ | 0.0007                | 0.0009                | 0.0010                | 0.0011                | 0.0012                   |
| $i = 4$ | $6.27 \times 10^{-5}$ | $8.74 \times 10^{-5}$ | $\times 10^{-4}$      | $1.22 \times 10^{-4}$ | $1.34293 \times 10^{-4}$ |
| $i = 5$ | $3.42 \times 10^{-6}$ | $5.19 \times 10^{-6}$ | $6.66 \times 10^{-6}$ | $7.87 \times 10^{-6}$ | $8.86 \times 10^{-6}$    |
| $i = 6$ | $3.41 \times 10^{-7}$ | $5.63 \times 10^{-7}$ | $7.60 \times 10^{-7}$ | $9.28 \times 10^{-7}$ | $1.07 \times 10^{-6}$    |
| $i = 7$ | $3.01 \times 10^{-8}$ | $5.39 \times 10^{-8}$ | $7.64 \times 10^{-8}$ | $9.65 \times 10^{-8}$ | $1.14 \times 10^{-7}$    |
| $i = 8$ | $1.62 \times 10^{-9}$ | $3.15 \times 10^{-9}$ | $4.70 \times 10^{-9}$ | $6.13 \times 10^{-9}$ | $7.42 \times 10^{-9}$    |

Table 2:  $f_0(\cdot|\theta)$  の値

|         | $\theta_1$             | $\theta_2$             | $\theta_3$             | $\theta_4$             | $\theta_5$             |
|---------|------------------------|------------------------|------------------------|------------------------|------------------------|
| $i = 1$ | $2.82 \times 10^{-10}$ | $2.20 \times 10^{-10}$ | $1.80 \times 10^{-10}$ | $1.52 \times 10^{-10}$ | $1.32 \times 10^{-10}$ |
| $i = 2$ | $1.72 \times 10^{-9}$  | $1.44 \times 10^{-9}$  | $1.23 \times 10^{-9}$  | $1.07 \times 10^{-9}$  | $9.50 \times 10^{-10}$ |
| $i = 3$ | $4.51 \times 10^{-9}$  | $4.01 \times 10^{-9}$  | $3.56 \times 10^{-9}$  | $3.18 \times 10^{-9}$  | $2.87 \times 10^{-9}$  |
| $i = 4$ | $7.25 \times 10^{-9}$  | $6.76 \times 10^{-9}$  | $6.18 \times 10^{-9}$  | $5.63 \times 10^{-9}$  | $5.15 \times 10^{-9}$  |
| $i = 5$ | $8.37 \times 10^{-9}$  | $7.97 \times 10^{-9}$  | $7.39 \times 10^{-9}$  | $6.80 \times 10^{-9}$  | $6.27 \times 10^{-9}$  |
| $i = 6$ | $1.01 \times 10^{-8}$  | $1.00 \times 10^{-8}$  | $9.55 \times 10^{-9}$  | $8.96 \times 10^{-9}$  | $8.38 \times 10^{-9}$  |
| $i = 7$ | $1.17 \times 10^{-8}$  | $1.21 \times 10^{-8}$  | $1.18 \times 10^{-8}$  | $1.13 \times 10^{-8}$  | $1.07 \times 10^{-8}$  |
| $i = 8$ | $1.23 \times 10^{-8}$  | $1.30 \times 10^{-8}$  | $1.29 \times 10^{-8}$  | $1.24 \times 10^{-8}$  | $1.19 \times 10^{-8}$  |

Table 3:  $f_1(\cdot|\theta)$  の値

- [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] M. Horiguchi. Bayesian inference in Markov decision processes. In A.B. Piunovskiy, editor, *Modern Trends in Controlled Stochastic Processes: Theory and Applications, Vol. 2*, pages 177–189. Luniver Press, UK, 2015.
- [9] Masami Kurano. Adaptive policies in Markov decision processes with uncertain transition matrices. *J. Inform. Optim. Sci.*, 4(1):21–40, 1983.
- [10] Viliam Makis. Multivariate Bayesian control chart. *Oper. Res.*, 56(2):487–496, 2008.
- [11] E. L. Porteus and A. Angelus. Opportunities for improved statistical process control. *Management Sci.*, 43:1214–1228, 1997.
- [12] Martin L. Puterman. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons Inc., New York, 1994. A Wiley-Interscience Publication.
- [13] 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.
- [14] Ulrich Rieder. On stopped decision processes with discrete time parameter. *Stochastic Processes Appl.*, 3(4):365–383, 1975.
- [15] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA, 1998.
- [16] George Tagaras. A dynamic programming approach to the economic design of  $X$ -charts. *IIE Trans.*, 26(3):48–56, 1994.
- [17] George Tagaras. Dynamic control charts for finite production runs. *European J. Oper. Res.*, 91:38–55, 1998.
- [18] George Tagaras and Yiannis Nikolaidis. Comparing the effectiveness of various Bayesian  $\bar{X}$  control charts. *Oper. Res.*, 50(5):878–888, 2002.
- [19] Howard M. Taylor. Markovian sequential replacement processes. *Ann. Math. Statist.*, 36:1677–1694, 1965.
- [20] Howard M. Taylor. Statistical control of a Gaussian process. *Technometrics*, 9:29–41, 1967.
- [21] K. M. van Hee. *Bayesian control of Markov chains*, volume 95 of *Mathematical Centre Tracts*. Mathematisch Centrum, Amsterdam, 1978.
- [22] William H. Woodall, Thomas J. Lorenzen, and Lonnie C. Vance. Weaknesses of the economic design of control charts. *Technometrics*, 28(4):408–410, 1986.
- [23] 佐々木稔, 堀口正之, 葦野正美. 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決定の理論と応用」, pages 181–192, 2014.
- [24] 堀口正之. 多変量ベイズ管理図の適応手法 (II). 数理解析研究所講究録 1939 「不確実性の下での数理モデルとその周辺」, pages 152–161, 2015.