観測に費用がともなうマルコフ連鎖の最適停止

竹 山 理
Osamu, TAKEYAMA

状態空間が S の時間的一様なマルコフ連鎖を $(\xi_n)_{n\geq 0}$ とする。$(\xi_n)_{n\geq 0}$ は確率空間 $(\Omega, \mathcal{F}, P)$ 上で実現されているものとし，その遷移確率を $p(x, dy)$，遷移作用素を

$$Tf(x) = E(f(\xi_n) \mid \xi_n = x) = \int_S f(y)p(x, dy)$$

とおく。いま，S 上の非負値関数 $g(x), c(x)$ によって，時点 k における利得が $g(\xi_k)$ で与えられ，追加的に ξ_{k+1} を観測するには $c(\xi_k)$ の費用がともなうものとする。時点 n における純利得の現在価値を G_n とおく。割引係数を定数 α $(0 < \alpha \leq 1)$ とすれば，G_n は

$$G_n = \alpha^n g(\xi_n) - \sum_{k=1}^{n} \alpha^{k-1} c(\xi_{k-1})$$

で与えられる。$(G_n)_{n\geq 0}$ に関する最適停止問題を考察しよう。すなわち，$(\xi_n)_{n\geq 0}$ が生成する増大情報系を $(\mathcal{F}_n)_{n\geq 0}$，有限な $(\mathcal{F}_n)$ 停止時刻の集合を $\mathcal{T}$ とし，観測を $\tau \in \mathcal{T}$ で停止したときの期待値 $E(G_\tau)$ を最大にする問題を考える。

もし $\sup_{\tau\in\mathcal{T}} E(G_\tau) = E(G_{\tau^*})$ をみたす $\tau^* \in \mathcal{T}$ が存在するならば，τ^* を最適停止時刻という。最適停止時刻を一般の場合に特徴付ける方法は Shnell[9] によって確立され，Shiryaev[8] に整理されている。しかし，古典的な問題[3][6][7] のように明示的に最適停止時刻が求まるのは，Chow,Robbins and Siegmund[2] が単調とよんだ場合が大部分であろう。つまり，$\mathcal{F}_n$ に関する条件付期待値を $E(\cdot \mid \mathcal{F}_n)$ と表し

$$A_n = \{\omega \in \Omega \mid E(G_{n+1} \mid \mathcal{F}_n) \leq G_n\}$$

とおくとき，

$$(*) \qquad A_0 \subseteq A_1 \subseteq A_2 \subseteq \cdots \subseteq \bigcup_{n=0}^{\infty} A_n = \Omega$$

をみたす場合である。筆者は[14] において，この単調な場合に最適停止時刻が直接に求まる簡明な十分条件を与えた。ただし，割引係数 α が $\alpha = 1$ と仮定した。この論文では $0 < \alpha < 1$ の場合にも区別なく取り扱えることを論じる。

1 マルチンゲール部分と Dynkin の公式

非負整数全体を $\overline{N}$ とする。$\overline{N} \times S$ 上の関数 $f(n, x)$ が与えられ，$X_n = f(n, \xi_n)$ $(n = 0, 1, 2, \ldots)$ は可積分確率変数列になるとする。$(X_n)_{n\geq 0}$ のマルチンゲール部分（[10][14] の Doob 分解の項を参

照）を求めておこう。ここで S 上の関数 $f(x)$ への作用素 L を

$$Lf(x) = (T-I)f(x) = \int_S f(y)p(x,dy) - f(x)$$

と定め，変数列 $(x(n))_{n\geq 0}$ の差分を $\Delta x(n) = x(n) - x(n-1)$ と定義する。このとき

$$(L+\Delta)f(n,x) = \{Tf(n,x) - f(n,x)\} + \{f(n,x) - f(n-1,x)\} = Tf(n,x) - f(n-1,x)$$

となる。

補題 1　次式で M_n を定めるとき，$(M_n)_{n\geq 0}$ は $(\mathcal{F}_n)$ に関してマルチンゲールである。

$$M_n = f(n,\xi_n) - \sum_{k=1}^{n}(L+\Delta)f(k,\xi_{k-1})$$

証明 $(M_n)_{n\geq 0}$ の差分は

$$\begin{aligned}\Delta M_n &= f(n,\xi_n) - f(n-1,\xi_{n-1}) - (L+\Delta)f(n,\xi_{n-1})\\ &= f(n,\xi_n) - f(n-1,\xi_{n-1}) - \{Tf(n,\xi_{n-1}) - f(n-1,\xi_{n-1})\}\\ &= f(n,\xi_n) - Tf(n,\xi_{n-1})\end{aligned}$$

となる。マルコフ性から

$$E(f(n,\xi_n) \mid \mathcal{F}_{n-1}) = E(f(n,\xi_n) \mid \xi_{n-1}) = Tf(n,\xi_{n-1})$$

を得る。よって $E(\Delta M_n \mid \mathcal{F}_{n-1}) = 0$ となる。したがって $(M_n)_{n\geq 0}$ は $(\mathcal{F}_n)$ に関してマルチンゲールである。

次にマルコフ過程の Dyinkin の公式 [4] に相当する等式を示しておく。

補題 2　停止時刻 $\tau \in T$ が $E(\tau) < \infty$ をみたし，さらに定数 C が存在して $\{\tau \geq n\}$ 上確率 1 で

$$\int_S \int_S |f(n,x) - f(n,y)| p(\xi_{n-1},dx) p(\xi_{n-1},dy) \leq C \qquad (n = 1,2,\ldots)$$

であるならば，

$$E(f(\tau,\xi_\tau)) = E(f(0,\xi_0)) + E\left(\sum_{k=1}^{\tau}(L+\Delta)f(k,\xi_{k-1})\right)$$

が成立する。

証明 [15] で示したように，$(\mathcal{F}_n)$ に適合した可積分確率変数列 $(Y_n)_{n\geq 0}$ に対し，ある定数 C が存在して $\{\tau \geq n\}$ 上確率 1 で

$$E(|\Delta Y_n| \mid \mathcal{F}_{n-1}) \leq C \qquad (n = 1,2,\ldots)$$

であるならば，

$$E(Y_\tau) = E(Y_0) + E\left(\sum_{k=1}^{\tau} E(Y_k - Y_{k-1} \mid \mathcal{F}_{k-1})\right)$$

が成立する。特に $(Y_n)_{n\geq 0}$ が $(\mathcal{F}_n)$ に関してマルチンゲールならば $E(Y_\tau) = E(Y_0)$ となる。

この結果を

$$M_n = f(n,\xi_n) - \sum_{k=1}^{n}(L+\Delta)f(k,\xi_{k-1})$$

に適用しよう。$(M_n)_{n\geq 0}$ の差分は

$$\Delta M_n = f(n,\xi_n) - Tf(n,\xi_{n-1}) = \int_S \{f(n,\xi_n) - f(n,y)\}p(\xi_{n-1},dy)$$

であるから，

$$\begin{aligned} E(|\Delta M_n| \mid \mathcal{F}_{n-1}) &= E\left(\left|\int_S \{f(n,\xi_n) - f(n,y)\}p(\xi_{n-1},dy)\right| \;\middle|\; \mathcal{F}_{n-1}\right) \\ &\leq E\left(\int_S |f(n,\xi_n) - f(n,y)|p(\xi_{n-1},dy) \;\middle|\; \mathcal{F}_{n-1}\right) \\ &= \int_S\int_S |f(n,x) - f(n,y)|p(\xi_{n-1},dx)p(\xi_{n-1},dy) \end{aligned}$$

となる。 仮定から，$\{\tau \geq n\}$ 上確率 1 で

$$E(|\Delta M_n| \mid \mathcal{F}_{n-1}) \leq C \qquad (n = 1,2,\ldots)$$

が示された。 よって $E(M_\tau) = E(M_0)$ が成立する。 すなわち

$$E\left(f(\tau,\xi_\tau) - \sum_{k=1}^{\tau}(L+\Delta)f(k,\xi_{k-1})\right) = E(f(0,\xi_n))$$

が導かれた。

2 単調停止問題の解法

まず，$\overline{N} \times S$ 上の関数を $f(n,x) = \alpha^n g(x)$ にとり，$L+\Delta$ を作用すると

$$(L+\Delta)f(n,x) = Tf(n,x) - f(n-1,x) = \alpha^n Tg(x) - \alpha^{n-1}g(x) = \alpha^{n-1}(\alpha T - I)g(x)$$

となる。 ここで $L_\alpha = \alpha T - I$ と定める。 すると補題 1 より， 次式の $(M_n)_{n\geq 0}$ は $(\mathcal{F}_n)$ に関してマルチンゲールである。

$$M_n = \alpha^n g(x) - \sum_{k=1}^{n} \alpha^{k-1} L_\alpha g(\xi_{k-1})$$

すなわち $\alpha^n g(\xi_n)$ の Doob 分解が

$$\alpha^n g(\xi_n) = M_n + \sum_{k=1}^{n} \alpha^{k-1} L_\alpha g(\xi_{k-1})$$

で与えられることが分かる。 これを純利得 G_n の定義式に代入すると

$$G_n = \alpha^n g(\xi_n) - \sum_{k=1}^{n} \alpha^{k-1} c(\xi_{k-1}) = M_n + \sum_{k=1}^{n} \alpha^{k-1}\{L_\alpha g(\xi_{k-1}) - c(\xi_{k-1})\}$$

となる。

次に純利得 $(G_n)_{n\geq 0}$ が単調性の条件 （＊） をみたすように， マルコフ連鎖 $(\xi_n)_{n\geq 0}$ に制限を課すことにする。

$$E(G_{n+1} \mid \mathcal{F}_n) = E(M_{n+1} \mid \mathcal{F}_n) + \sum_{k=1}^{n+1} \alpha^{k-1}\{L_\alpha g(\xi_{k-1}) - c(\xi_{k-1})\}$$

$$= \quad M_n + \sum_{k=1}^{n} \alpha^{k-1}\{L_\alpha g(\xi_{k-1}) - c(\xi_{k-1})\} + \alpha^n\{L_\alpha g(\xi_n) - c(\xi_n)\}$$

であるから,

$$E(G_{n+1} \mid \mathcal{F}_n) = G_n + \alpha^n\{L_\alpha g(\xi_n) - c(\xi_n)\}$$

を得る。これより

$$E(G_{n+1} \mid \mathcal{F}_n) \le G_n \iff L_\alpha g(\xi_n) \le c(\xi_n)$$

となる。ここで

$$D = \{x \in S \mid L_\alpha g(x) \le c(x)\}$$

とおくと,

$$A_n = \{\omega \in \Omega \mid E(G_{n+1} \mid \mathcal{F}_n) \le G_n\} = \{\omega \in \Omega \mid L_\alpha g(\xi_n) \le c(\xi_n)\} = \{\omega \in \Omega \mid \xi_n \in D\}$$

である。いま D への到達時刻を τ^* とおく。つまり

$$\tau^*(\omega) = \inf\{n \mid \xi_n(\omega) \in D\} = \inf\{n \mid \omega \in A_n\}$$

と定める。そこでマルコフ連鎖 $(\xi_n)_{n\ge0}$ の遷移確率が, 次の条件をみたすものとしよう。

$$\text{(H1a)} \quad \begin{cases} x \in D^c \text{ に対して} \quad p^{(k)}(x, D) > 0 \text{ となる } k \text{ がある} \\ x \in D \text{ に対して} \quad p(x, D) = 1 \end{cases}$$

ただし

$$p^{(1)}(x, dy) = p(x, dy), \quad p^{(k+1)}(x, dy) = \int_S p(x, d\xi)p^{(k)}(\xi, dy) \qquad (k = 1, 2, \ldots)$$

である。このとき $(G_n)_{n\ge0}$ が確率 1 で単調性の条件 （*） をみたすことは明らかである。

さらに, 単調性と τ^* の定義から

$$\begin{cases} n < \tau^* \text{ において} \quad L_\alpha g(\xi_n) > c(\xi_n) \\ n \ge \tau^* \text{ において} \quad L_\alpha g(\xi_n) \le c(\xi_n) \end{cases}$$

をみたしている。したがって, 任意の停止時刻 $\tau \in \mathcal{T}$ に対して

$$\sum_{k=1}^{\tau} \alpha^{k-1}\{L_\alpha g(\xi_{k-1}) - c(\xi_{k-1})\} \le \sum_{k=1}^{\tau^*} \alpha^{k-1}\{L_\alpha g(\xi_{k-1}) - c(\xi_{k-1})\}$$

である。ところで $(M_n)_{n\ge0}$ はマルチンゲールであったから, 任意抽出定理より, $\tau \in \mathcal{T}$ に対して $E(M_{\tau\wedge n}) = E(M_0)$ が成り立つ。すなわち, 任意の $\tau \in \mathcal{T}, n \in \overline{N}$ に対して

$$E(\alpha^{\tau\wedge n} g(\xi_{\tau\wedge n})) = E(g(\xi_0)) + E\left(\sum_{k=1}^{\tau\wedge n} \alpha^{k-1} L_\alpha g(\xi_{k-1})\right)$$

である。したがって, 純利得の現在価値 $(G_n)_{n\ge0}$ について

$$\begin{aligned} E(G_{\tau\wedge n}) &= E\left(\alpha^{\tau\wedge n} g(\xi_{\tau\wedge n}) - \sum_{k=1}^{\tau\wedge n} \alpha^{k-1} c(\xi_{k-1})\right) \\ &= E(g(\xi_0)) + E\left(\sum_{k=1}^{\tau\wedge n} \alpha^{k-1}\{L_\alpha g(\xi_{k-1}) - c(\xi_{k-1})\}\right) \end{aligned}$$

となっている。よって，単調性と τ^* の定義から

$$E(G_{\tau\wedge n}) \le E(g(\xi_0)) + E\left(\sum_{k=1}^{\tau^*} \alpha^{k-1}\{L_\alpha g(\xi_{k-1}) - c(\xi_{k-1})\}\right)$$

が導かれた。

いま，非負値関数 $g(x)$ は S で下半連続であると仮定する。すると $g(\xi_\tau) \le \varliminf_{n\to\infty} g(\xi_{\tau\wedge n})$ である。これと Fatou の補題から

$$E(\alpha^\tau g(\xi_\tau)) \le E\left(\varliminf_{n\to\infty} \alpha^{\tau\wedge n} g(\xi_{\tau\wedge n})\right) \le \varliminf_{n\to\infty} E(\alpha^{\tau\wedge n} g(\xi_{\tau\wedge n}))$$

となる。また，単調収束定理から

$$E\left(\sum_{k=1}^{\tau} \alpha^{k-1} c(\xi_{k-1})\right) = \lim_{n\to\infty} E\left(\sum_{k=1}^{\tau\wedge n} \alpha^{k-1} c(\xi_{k-1})\right)$$

である。したがって

$$E(G_\tau) = E\left(\alpha^\tau g(\xi_\tau) - \sum_{k=1}^{\tau} \alpha^{k-1} c(\xi_{k-1})\right) \le \varliminf_{n\to\infty} E(G_{\tau\wedge n})$$

となる。これと前述の $E(G_{\tau\wedge n})$ に関する不等式から

$$E(G_\tau) \le E(g(\xi_0)) + E\left(\sum_{k=1}^{\tau^*} \alpha^{k-1}\{L_\alpha g(\xi_{k-1}) - c(\xi_{k-1})\}\right)$$

が得られた。

最後に停止時刻 τ^* について，補題 2 が適用できる条件を与えよう。まず，$(\xi_n)_{n\ge 0}$ の遷移確率について

$$\text{(H1b)} \quad \text{ある } N \text{ が存在して } \inf_{x\in D^c} p^{(N)}(x, D) > 0$$

となることを仮定する。このとき[15] でみたように，$E(\tau^*) < \infty$ である。次に $\{\tau^* \ge n\}$ 上確率 1 で $\xi_{n-1} \in D^c$ であるから，$f(n,x) = \alpha^n g(x)$ のとき

$$\begin{aligned}
&\int_S \int_S |f(n,x) - f(n,y)| p(\xi_{n-1}, dx) p(\xi_{n-1}, dy) \\
&= \alpha^n \int_S \int_S |g(x) - g(y)| p(\xi_{n-1}, dx) p(\xi_{n-1}, dy) \\
&= \le \sup_{\xi\in D^c} \int_S \int_S |g(x) - g(y)| p(\xi, dx) p(\xi, dy)
\end{aligned}$$

である。したがって

$$\text{(H2)} \qquad \sup_{\xi\in D^c} \int_S \int_S |g(x) - g(y)| p(\xi, dx) p(\xi, dy) < \infty$$

をみたすならば，補題 2 より

$$E(\alpha^{\tau^*} g(\xi_{\tau^*})) = E(g(\xi_0)) + E\left(\sum_{k=1}^{\tau^*} \alpha^{k-1} L_\alpha g(\xi_{k-1})\right)$$

が成立する。さらに

$$\text{(H3)} \qquad \sup_{x\in D^c} c(x) < \infty$$

を仮定すれば，

$$E\left(\sum_{k=1}^{\tau^*} c(\xi_{k-1})\right) < \infty$$

である。このとき

$$\begin{aligned} E(G_{\tau^*}) &= E\left(\alpha^{\tau^*} g(\xi_{\tau^*}) - \sum_{k=1}^{\tau^*} \alpha^{k-1} c(\xi_{k-1})\right) \\ &= E(g(\xi_0)) + E\left(\sum_{k=1}^{\tau^*} \alpha^{k-1}\{L_\alpha g(\xi_{k-1}) - c(\xi_{k-1})\}\right) \end{aligned}$$

となる。したがって，任意の $\tau \in \mathcal{T}$ に対して

$$E(G_\tau) \le E(G_{\tau^*})$$

が示された。すなわち τ^* が最適停止時刻である。

以上から，次の結果が得られた．

定理 状態空間が S の時間的一様なマルコフ連鎖 $(\xi_n)_{n\ge 0}$ に対して，停止問題

$$\sup_{\tau\in\mathcal{T}} E\left(\alpha^\tau g(\xi_\tau) - \sum_{k=1}^{\tau} \alpha^{k-1} c(\xi_{k-1})\right)$$

を考える。ここで $g(x), c(x)$ は S 上の非負値関数であり，さらに $g(x)$ は下半連続とする。また，割引係数 α は $0 < \alpha \le 1$ の定数とする。いま $D = \{x \in S \mid L_\alpha g(x) \le c(x)\}$ と定め，$(\xi_n)_{n\ge 0}$ の D への到達時刻を τ^* とおく。ただし $L_\alpha = \alpha T - I$（T は遷移作用素）である。

このとき次の仮定のもとで，τ^* は最適停止時刻である。

(H1) ある N があって $\inf_{x\in D^C} p^{(N)}(x, D) > 0$, かつ $x \in D$ のとき $p(x, D) = 1$

(H2) $\sup_{\xi\in D^c} \int_S \int_S |g(x) - g(y)| p(\xi, dx) p(\xi, dy) < \infty$

(H3) $\sup_{x\in D^c} c(x) < \infty$

注意 次の不等式は容易に示すことができる。

$$\int_S \int_S |g(x) - g(y)| p(\xi, dx) p(\xi, dy) \le 2 \int_S |g(x)| p(\xi, dx)$$

したがって，$g \ge 0$ のとき，

$$\text{(H2}'\text{)} \qquad \sup_{x\in D^c} \int_S g(y) p(x, dy) < \infty$$

は **(H2)** の十分条件である。

3 適用

最後に[3][6][7]で取り扱われている古典的な問題（職探し問題，盗賊問題，秘書問題）に，定理を適用して最適停止時刻を求めてみよう。

3．1 職探し問題

失業中の労働者が職探しに出かけるとき，一定期間内に1つの就職口が見つかり，就職条件の提示を受けるとする。提示条件のうちの賃金だけを考慮し，労働者はできるだけ高い賃金を支払ってくれる職を望むものとしよう。提示賃金はそれまでに受けた提示額とは統計的に独立な既知の確率分布にしたがうものとする。就職口を見いだすための1回の探索には，一定額の費用がかかるとする。探索の回数には上限を設けず，いったん訪れた就職口は保持されて過去のすべての就職口から選択できるというリコールのある場合に，労働者はどのような方策で職探しにのぞめばよいのだろうか。

労働者が n 回目に訪問した就職口で提示される賃金額 $W_n\ (n=1,2,\ldots)$ は，W と同分布の独立な正値確率変数列で $m=E(W)<\infty$ をみたし，その分布関数 F は既知とする。1回の探索に要する費用を c とするとき，リコールがある場合の純収益は $\max(W_1,W_2,\cdots,W_n)-cn$ で与えられる。ここではさらに，割引がある場合を考えよう。まず

$$\xi_n=\max(W_1,W_2,\cdots,W_n),\quad \xi_0=0$$

とおき，割引係数を定数 $\alpha\ (0<\alpha\le 1)$ とすると，純収益の割引現在価値 G_n は

$$G_n=\alpha^n\xi_n-c\sum_{k=1}^{n}\alpha^{k-1}$$

で与えられる。ただし $0<c<\alpha m$ とする。いま，停止時刻 τ を定めて就職口の探索を止めるものとする。このとき期待純収益 $E(G_\tau)$ を最大にする停止時刻を見いだそう。

はじめに $(\xi_n)_{n\ge 0}$ は $S=[0,\infty)$ 上のマルコフ連鎖であり，
その遷移確率は

$$p(x,dy)=I_{\{y\le x\}}P(W\le x)\delta_x(dy)+I_{\{y>x\}}P(W\in dy)$$

となることに注意する。ただし I_A は集合 A の定義関数，δ_x は $\{x\}$ 上の点確率測度である。一般に $L_\alpha g(x)=(\alpha T-I)g(x)$ は

$$L_\alpha g(x)=\alpha\int_S g(y)p(x,dy)-g(x)=\alpha\int_S\{g(y)-g(x)\}p(x,dy)-(1-\alpha)g(x)$$

であるから，いま $g(x)=x, c(x)=c$ とおき，W の分布関数を $F(x)$ とすると，

$$\begin{aligned}L_\alpha g(x)&=\alpha\int_0^x(y-x)\delta_x(dy)+\alpha\int_x^\infty(y-x)p(W\in dy)-(1-\alpha)x\\&=\alpha\int_x^\infty(y-x)dF(y)-(1-\alpha)x\end{aligned}$$

となる。そこで $D=\{x \mid L_\alpha g(x) \leq c(x)\}$ を決定しよう。上式より

$$L_\alpha g(x) \leq c(x) \iff \int_x^\infty (y-x)dF(y) \leq \frac{(1-\alpha)x+c}{\alpha}$$

である。ここで $H(x)=\int_x^\infty (y-x)dF(y)$ とおくと，$H(x)$ は減少関数で $H(0)=m>c/\alpha$, $H(\infty)=0$ である。一方で $\{(1-\alpha)x+c\}/\alpha$ は非減少関数だから，方程式

$$H(x)=\frac{(1-\alpha)x+c}{\alpha}$$

は唯一の解 γ をもつ。すると $x\in D \iff x\geq\gamma$ であるから，

$$D=\{x\in S \mid x\geq\gamma\}$$

が導かれた。もとより (H3) は明らかであるから，この D に対して，(H1), (H2′) を確かめよう。$\xi_{n+1}\geq\xi_n$ であるから，$x\in D$ のとき，$p(x,D)=1$ となる。また

$$\inf_{x\in D^C} p(x,D)=P(W\geq\gamma)>0$$

である。よって条件 (H1) はみたされる。次に $x\in D^c$ のとき

$$\int_S g(y)p(x,dy)\leq x+\int_x^\infty ydF(y)\leq\gamma+m$$

である。よって条件 (H2′) はみたされる。

以上から，職探し問題の最適停止時刻 τ^* は，

$$\int_\gamma^\infty (y-\gamma)dF(y)=\frac{(1-\alpha)\gamma+c}{\alpha}$$

をみたす γ によって，$\tau^*=\inf\{n \mid \xi_n\geq\gamma\}$ で与えられることが示された。

3.2 盗賊問題

盗賊は盗みを繰り返して，稼ぎの累積額をできるだけ大きくしたいと考えている。ただし，盗みを犯すときには一定の確率で捕まる可能性があり，捕まればそれまで盗んだ稼ぎをすべて失うものとする。盗賊は盗みから足を洗うタイミングをどのように決めればよいのだろうか。

各回の盗みによる稼ぎ額を $X_n\ (n=1,2,\ldots)$ とし，これは X と同分布で独立な正値確率変数列とする。ここで稼ぎの平均値は $\mu=E(X)<\infty$ で既知とする。また，1回の盗みで捕まる確率は $\beta\ (0<\beta<1)$ とし，捕まるかどうかは過去の履歴に独立と仮定する。いま，独立同分布の確率変数列 $Z_n\ (n=1,2,\ldots)$ で

$$P(Z_n=1)=\beta,\qquad P(Z_n=0)=1-\beta$$

であるものを用意する。$(Z_n)_{n\geq1}$ と $(X_n)_{n\geq1}$ も独立とする。盗賊が n 回目の盗みを遂行後に盗みから足を洗うことにするとき，稼ぎの累積額 G_n は

$$G_n=\left(\prod_{k=1}^n Z_k\right)\sum_{k=1}^n X_k \qquad (n=1,2,\ldots)$$

で与えられる。そこで停止時刻 τ を定めて盗みから足を洗うものとするとき，期待値 $E(G_\tau)$ を最大にする停止時刻を見いだそう。

状態空間を $S=[0,\infty)\times\{0,1\}$ とおき，

$$\xi_n=\left(\sum_{k=1}^{n}X_k,\ \prod_{k=1}^{n}Z_k\right),\qquad \xi_0=(0,1)$$

と定めると，$(\xi_n)_{n\geq 0}$ は S 上のマルコフ連鎖となる。遷移確率は次の 2 つの確率測度の直積測度で与えられる。

$$P\left(\sum_{k=1}^{n+1}X_k\in dy\;\Big|\;\sum_{k=1}^{n}X_k=x\right)=P\left(x+X_{n+1}\in dy\;\Big|\;\sum_{k=1}^{n}X_k=x\right)=P(x+X\in dy),$$

$$P\left(\prod_{k=1}^{n+1}Z_k=t\;\Big|\;\prod_{k=1}^{n}Z_k=s\right)=\begin{cases}1 & (s=0,t=0\text{のとき})\\ 0 & (s=0,t=1\text{のとき})\\ 1-\beta & (s=1,t=0\text{のとき})\\ \beta & (s=1,t=1\text{のとき})\end{cases}$$

したがって遷移作用素は，X の分布関数を F とおくとき，S 上の関数 $f(x,s)$ に対して

$$\begin{aligned}Tf(x,0)&=\int_0^\infty f(x+y,0)\,dF(y)\\ Tf(x,1)&=(1-\beta)\int_0^\infty f(x+y,0)\,dF(y)+\beta\int_0^\infty f(x+y,1)\,dF(y)\end{aligned}$$

となる。

利得関数を $\xi=(x,s)$ のとき $g(\xi)=xs$ とすれば，稼ぎの累積額は $G_n=g(\xi_n)$ と表せる。そこで作用素 $L=T-I$ に対して，$D=\{(x,s)\in S\mid Lg(x,s)\leq 0\}$ を決定しよう。

$$\begin{aligned}Tg(x,0)&=0,\\ Tg(x,1)&=\beta\int_0^\infty(x+y)\,dF(y)\;=\;\beta(x+\mu)\end{aligned}$$

であるから

$$Lg(x,1)\leq 0\iff\beta(x+\mu)-x\leq 0\iff x\geq\frac{\beta\mu}{1-\beta}$$

となる。したがって

$$D=\left\{(x,s)\in S\;\Big|\;s=0\text{のとき}\,x\geq 0,\ s=1\text{のとき}\,x\geq\frac{\beta\mu}{1-\beta}\right\}$$

である。このとき条件 (H1) は $(\xi_n)_{n\geq 0}$ の定義から，また条件 (H3) は $c=0$ より成立する。条件 (H2′) を確かめよう。

$$\sup_{\xi\in D^c}\int_S g(\eta)p(\xi,d\eta)=\sup_{s=1,x<\frac{\beta\mu}{1-\beta}}\int_0^\infty(x+y)\,dF(y)=\frac{\mu}{1-\beta}<\infty$$

である。以上より，盗賊問題の最適停止時刻 τ^* は

$$\tau^*=\inf\left\{n\;\Big|\;\sum_{k=1}^{n}X_k\geq\frac{\beta\mu}{1-\beta}\right\}$$

で与えられることが示された。

3．3 秘書問題

会社の社長が1人の秘書を求人中である。社長は募集者を面接することで，それまでの募集者を資質の高い順にランク付けできる（同ランクはない）ものとする。応募者はランダムな順序で1人ずつ逐次に会社を訪問し，社長は面接が終了しだい直ちに採用の可否を決定して，採用を決めたときには求人活動を打ち切る。ただし以前にいちど不採用とした応募者を採用すること（リコール）はできない。さらに面接をおこなう応募者は n 人までとし，最後の n 番目まで面接したときには，この最後の応募者を採用するものとする。このとき最大の確率をもって n 人中のベストランクの応募者を採用するために，社長はどのようにすればよいのだろうか。

会社を k 番目に訪問した応募者を，番号 k の応募者とよぶことにする。いま，n 人の応募者を資質の高い順に $1,2,\ldots,n$ とランク付けして，番号 k の応募者のランクを ω_k と表し，n 人の応募者のランクの順列を $\omega=(\omega_1,\omega_2,\cdots,\omega_n)$ とおく。これらの全体を標本空間 Ω にとる。標本総数は $n!$ であるから，応募者がランダムな順序で1人ずつ会社を訪問するとき，$P(\{\omega\})=1/n!$ である。番号 k の応募者を面接するとき，$\{\omega_1,\omega_2,\ldots,\omega_k\}$ においてランク ω_k が相対的に最も良いとき，番号 k の応募者は相対的ベストであるという。n 人中でベストランクになるのは，相対的ベストの応募者に限られ，しかも最後に現れる相対的ベストの応募者である。そこで相対的ベストとなる応募者の番号を増大する順に $\xi_0,\xi_1,\xi_2,\ldots$ とおく。ここで $\xi_0=1$ である。さらに，番号 ξ_i より後に相対的ベストが現れない（番号 ξ_i の応募者のランクが $\omega_{\xi_i}=1$ となる）とき，$\xi_{i+1}=\infty$ と表すことにする。たとえば $n=10$ で $\omega=(6,7,5,8,10,4,9,1,2,3)$ のとき，相対的ベストとなる ω_i は $\omega_1,\omega_3,\omega_6,\omega_8$ であるから，$\xi_0(\omega)=1,\xi_1(\omega)=3,\xi_2(\omega)=6,\xi_3(\omega)=8,\xi_4(\omega)=\infty$ となる。

このように定めた $(\xi_i)_{i\geq 0}$ は状態空間 $S=\{1,2,\ldots,n,\infty\}$ 上のマルコフ連鎖であり，その遷移確率は

$$p(k,\ell)=P(\xi_{i+1}=\ell\mid\xi_i=k)=\begin{cases}\dfrac{k}{\ell(\ell-1)} & (1\leq k<\ell\leq n)\\ \dfrac{k}{n} & (1\leq k\leq n,\ell=\infty)\\ 0 & (\text{その他})\end{cases}$$

となることが分かる（[5] を参照）。したがって，遷移作用素は S 上の関数 $f(k)$ に対して

$$\begin{aligned}Tf(k)&=\sum_{\ell\in S}f(\ell)p(k,\ell)=k\sum_{\ell=k+1}^{n}\frac{f(\ell)}{\ell(\ell-1)}+\frac{k}{n}f(\infty)\qquad(1\leq k\leq n)\\ Tf(\infty)&=0\end{aligned}$$

で与えられる。

では，相対的ベストである番号 ξ_i の応募者がベストランクである確率はどのように表せるだろうか。$\xi_i=k$ のもとで，番号 k の応募者がベストランクである条件付確率は

$$P(\omega_k=1\mid\xi_i=k)=\frac{k}{n}$$

となる。ここで利得関数を

$$g(k) = \begin{cases} \dfrac{k}{n} & (k = 1, 2, \dots, n) \\ 0 & (k = \infty) \end{cases}$$

と定めると,

$$P(\omega_{\xi_i} = 1) = \sum_{k=1}^{n} P(\omega_k = 1 \mid \xi_i = k) P(\xi_i = k) = \sum_{k \in S} g(k) P(\xi_i = k) = E(g(\xi_i))$$

である。すなわち，ベストランクの応募者を採用する確率を最大にすることは，停止時刻 τ を定めて期待値 $E(g(\xi_\tau))$ を最大にすることと同等である。

そこで作用素 $L = T - I$ について $D = \{k \in S \mid Lg(k) \le 0\}$ を決定しよう。明らかに $\infty \in D$ である。それ以外のとき,

$$Lg(k) = k \sum_{\ell=k+1}^{n} \frac{\ell}{n} \frac{1}{\ell(\ell-1)} + \frac{k}{n} \times 0 - \frac{k}{n} = \frac{k}{n} \left(\sum_{\ell=k+1}^{n} \frac{1}{\ell - 1} - 1 \right)$$

であるから,

$$k \in D \iff \sum_{\ell=k+1}^{n} \frac{1}{\ell-1} \le 1$$

となる。ここで $H(k) = \displaystyle\sum_{\ell=k+1}^{n} \frac{1}{\ell-1}$ は，k に関して減少し，$H(1) > 1, H(n) = 0$ であるから，$H(k) \le 1$ をみたす最小の k が定まる。これを k^* とおくと

$$D = \{k^*, k^* + 1, \dots, n, \infty\}$$

であることが導かれた。このとき，定理の条件 (H1), (H2), (H3) が成り立つことは明らかであるから，最適停止時刻 τ^* は

$$\tau^* = \inf\{k \mid \xi_k \in D\} = \inf\{k \mid \xi_k \ge k^*\}$$

で与えられることが示された。つまり社長は，番号 $k^* - 1$ までの応募者を不採用とし，番号 k^* 以後ではじめて相対的ベストとなる応募者を採用すればよいのである。ただし $\xi_{\tau^*} = \infty$ となるときには，番号 n の応募者を採用することになる。

文　献

[1] Chow,Y.S. and Robbins,H., On optimal stopping rules, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 2, 33-49, 1963.

[2] Chow,Y.S.,Robbins,H. and Siegmund,D., *Great Expectations:The Theory of Optimal Stopping*, Houghton Mifflin, 1971.

[3] Dubins,L.E. and Teicher,H., Optimal stopping when the future is discounted, *Ann. Math. Statist.*, 38, 601-605,1967.

[4] Dynkin,E.B.,*Markov Processes*,Springer-Verlag,1965.

[5] Dynkin,E.B. and Yushkevitch,A.A., *Markov Processes:Theorms and Problems*, Plenum, 1969.

[6] Lippman,S.A. and McCall,J.J., The economics of job search : A survey, *Economic Inquiry* 14, 155-189, 1976.

[7] Samuels,S.M., Secretary problems, *Handbook of Sequential Analysis*, B.K. Ghosh and P.K. Sen, eds., 381-405, Marcel Dekker, 1991.

[8] Shiryaev,A.N., *Optimal Stopping Rules*, Springer, 1978.

[9] Shnell,J.L., Application of martingale system theorems, *Trans. Amer. Math. Soc.* 73, 293-312.

[10] Williams, D., *Probability and Martingales*, Cambridge University Press, 1991.

[11] Zuckerman,D., Job search: The continuous case, *Journal of Applied Probability* 20, 637-648, 1983.

[12] 穴太克則, 最適停止における単調問題とＯＬＡ停止規則の最適性について, 南山経営研究 9, 735-751, 1995.

[13] 穴太克則 , タイミングの数理 , 朝倉書店 , 2000.

[14] 竹山理 , 「情報探索理論」 再訪 , 産業と経済 , 第 13 巻第 3 号 , 1-12, 奈良産業大学経済学会 , 1998.

[15] 竹山理 , 単調停止問題の解法, 数学 53, 422-426, 日本数学会 , 2001.