数理解析研究所講究録
1504 巻 2006 年 177-183

ベルヌーイ過程に付随する最適停止問題

王 琦 (Qi Wang)　　玉置光司 (Mitsushi Tamaki)
愛知大学経営学部
Department of Business Administration, Aichi University

1 はじめに

$Bruss$[1] は次のようなベルヌーイ過程に付随する最適停止問題を考えた。n 個の独立事象 $E_i, 1 \leq i \leq n$ があり、意思決定者は $E_1, E_2, ..., E_n$ を逐次観察する。E_i は i 回目の試行に付随する事象で、2 つの結果「成功」か「失敗」のどちらかを表す。簡単のためにインディケータ I_i を導入し、E_i が成功の時は、$I_i = 1$、E_i が失敗の時は、$I_i = 0$ と定義する。従って、意思決定者は $I_1, I_2, ..., I_n$ を逐次観察することになる。

$Bruss$[1] は最後の成功で停止することを勝ちと見なし、勝つ確率を最大にする最適停止問題を研究した。$Bruss$[2] は $Bruss$[1] を一般化して、最後から m 番目の成功で停止することを勝ちと見なし、勝つ確率を最大にする問題を考えた。$Bruss$[2] で $m = 1$ とおいたものが $Bruss$[1] に他ならない。

今、$p_i = P\Big(I_i = 1\Big), 1 \leq i \leq n$ とおくと $Odds-theorem$($Bruss$[1] の結果) は次のように述べられる。ただし、$q_i = 1 - p_i$、$r_i = p_i/q_i$ と定義する。

Odds-theorem

(a) 最適停止ルール：次式で正整数 s

$$s = \sup\Big\{1 \leq k \leq n : \sum_{j=k}^{n} r_j \geq 1\Big\}$$

ただし、$\sup\{\phi\} = 0$ を定義する。最適停止ルールは、最初の $s-1$ ケの対象をパスし、s 以降最初の成功でストップすることである。

(b) 勝つ確率：最適停止ルールの下で、勝つ確率は次式で与えられる。

$$\Big(\prod_{i=s}^{n} q_i\Big)\Big(\sum_{i=s}^{n} r_i\Big)$$

$Odds-theorem$ の最適停止ルールのように、ある時刻まで対象をパスし、それ以降最初の成功で停止するルールのことを閾値ルールと呼ぶ。より正確には閾値 s の閾値ルールと呼ぶ。$p_i = 1/i,\ 1 \leq i \leq n$ の場合、$Odds-theorem$ は秘書問題 (*secretary problem*) の解を与える。この場合、成功とは相対順位 1 の応募者の出現を意味し、最後の成功で停止することは最良の応募者 (絶対順位 1) の採用を意味する。

$Bruss$[2] は m が一般の正整数の場合も最適停止ルールが閾値ルールとなることを示した。本論では、$Bruss$[1] の別の一般化を試みる。最後の m 個の成功のどれかで停止すれば勝ちと見なす問

題である（$Bruss$[2] では、最後からちょうど m 番目の成功時点で停止したら勝ちである）。2 節では、この問題を取り扱う。この場合も閾値ルールが最適ルールとなることが示される。

秘書問題に拒否確率を最初に導入したのは $Smith$[4] である。そこでは応募者は採用の申し込みを一定の確率 β で受け入れる (確率 $1-\beta$ で拒否する) と仮定した。3 節では 2 節の問題を拒否が許される場合に拡張する。$Tamaki$[5] も拒否確率を考慮した別の秘書問題を取り扱った。

秘書問題においては、$n \to \infty$ とした場合の挙動に興味がある。拒否が無い場合、有る場合についてそれぞれ 2 節、3 節で、漸近的結果が得られているが、これらは NHPP(*non−homogeneous Poisson process*) を用いると、簡単に求めることができる。これについて、4 節で議論する。

2 モデルと定式化

2.1 最適停止ルール

N_k を

$$N_k = I_k + I_{k+1} + \cdots + I_n, \qquad k = 1, 2, ..., n$$

と定義すると N_k は時刻 k 以降に出現する成功の数を表す。我々の問題は、最後の m 個の成功のどれかで停止する確率を最大化する問題であるから、次式を満足する最適停止ルール τ^* を求める問題となる。

$$P\Big\{N_{\tau^*} \le m\Big\} = \sup_{\tau} P\{N_\tau \le m\}$$

$n \le m$ の場合は、明らかに最適停止ルールは最初の成功で停止することであり、勝つ確率は $1 - q_1 q_2 \ldots q_n$ で与えられる。従って、以降は $n > m$ と仮定して議論を進める。DP(*dynamic programming*) を用いて問題を定式化するために次の量を定義する。

$v_i^{(m)}$: 最初の i ヶの対象をパスし、それ以降最適に振舞って勝つ確率
$g_i^{(m)}$: 時刻 i で $I_i = 1$ を観測した時、そこで停止して勝つ確率

このとき、次の DP 方程式が成立する。勝つ確率は $v_0^{(m)}$ で与えられる。

$$v_{i-1}^{(m)} = p_i \max\Big\{g_i^{(m)}, v_i^{(m)}\Big\} + q_i v_i^{(m)}, i = 1, 2, ..., n \tag{2.1}$$

ただし、$v_n^{(m)} = 0$。

(2.1) を解くためには $g_i^{(m)}$ を最初に求めなければならない。まず、以下の記号を定義する。

$$Q_k = \prod_{j=k}^{n} q_j$$

$$R_{k,j} = \sum_{k \le i_1 < \cdots < i_j \le n} r_{i_1} \cdots r_{i_j}$$

$R_{k,0} = 1$ と約束する。Bruss[2] より、次の関係が成り立つ。

$$P\Big\{N_k = j\Big\} = Q_k R_{k,j}, \qquad 1 \le k \le n - j + 1 \tag{2.2}$$

この時、$g_i^{(m)}$ は次のように与えられる。

補題 1

$$g_i^{(m)} = \begin{cases} 1 & \text{if } i > n-m \\ Q_{i+1}\sum_{j=0}^{m-1} R_{i+1,j} & \text{if } i \leq n-m \end{cases} \tag{2.3}$$

証明

$i > n-m$ の時は明らかである。$i \leq n-m$ の時、定義より

$$g_i^{(m)} = P\{N_{i+1} \leq m-1\} = \sum_{j=0}^{m-1} P\{N_{i+1} = j\} \tag{2.4}$$

と書ける。従って、(2.2) より

$$\begin{aligned} g_i^{(m)} &= \sum_{j=0}^{m-1} Q_{i+1} R_{i+1,j} \\ &= Q_{i+1} \sum_{j=0}^{m-1} R_{i+1,j} \end{aligned}$$

となる。

補題 2

$g_i^{(m)}$ は i に関して、非減少関数である。

証明

I_j が非負であるから、N_i は明らかに i の非増加関数であり、$N_i \geq N_{i+1}$ となる。従って、$N_i \leq m-1$ であれば、$N_{i+1} \leq m-1$ である、$P\{N_i \leq m-1\} \leq P\{N_{i+1} \leq m-1\}$ が成立する。従って、(2.4) より $g_{i-1}^{(m)} \leq g_i^{(m)}$ が成立する。

(2.1) から

$$v_{i-1}^{(m)} \geq p_i v_i^{(m)} + q_i v_i^{(m)} = v_i^{(m)}$$

なので、$v_i^{(m)}$ は i の非増加関数である。従って、最適停止ルールは閾値ルールになる。

定理 1

(a) 最適停止ルール：正整数 s^* を次式で定義する。

$$s^* = \sup\left\{i : R_{i,m} \geq 1\right\}$$

この時、最適ルールは閾値 s^* の閾値ルールとなる。

(b) 勝つ確率：最適停止ルールの下で、勝つ確率は次式で与えられる。

$$v_0^{(m)} = Q_{s^*} \sum_{j=1}^{m} R_{s^*,j}$$

証明

最適停止ルールが閾値ルールとなることは分っているので、その閾値を s^* とすると

$$s^* = \inf\{i : g_i^{(m)} \geq v_i^{(m)}\} \tag{2.5}$$

で与えられる。さて、$i \geq s^*-1$ の時、$v_i^{(m)}$ は時刻 $i+1$ 以降の成功の出現個数が 1 以上、$\min(m,\ n-i)$ 以下である確率に他ならない。従って、(2.2) より、

$$\begin{aligned} v_i^{(m)} &= P\Big\{1 \leq N_{i+1} \leq \min(m,\ n-i)\Big\} \\ &= \sum_{j=1}^{\min(m,n-i)} P\Big\{N_{i+1} = j\Big\} \\ &= Q_{i+1} \sum_{j=1}^{\min(m,n-i)} R_{i+1,j} \end{aligned} \tag{2.6}$$

となる。従って、(2.5) は (2.3) より

$$\begin{aligned} s^* &= \inf\{i : Q_{i+1} \sum_{j=0}^{m-1} R_{i+1,j} \geq Q_{i+1} \sum_{j=1}^{m} R_{i+1,j}\} \\ &= \inf\{i : 1 \geq R_{i+1,m}\} \\ &= \sup\{i : R_{i,m} \geq 1\} \end{aligned}$$

となり (a) が示される。又、勝つ確率は $s^* \leq n-m$ より、(2.6) から

$$\begin{aligned} v_0^{(m)} &= v_{s^*-1}^{(m)} \\ &= Q_{s^*} \sum_{j=1}^{m} R_{s^*,j} \end{aligned}$$

となり、(b) が示される。

注

$m = 1$ の時、$R_{i,1} = \sum_{j=i}^{n} r_j$ であり、定理 1 は $odds - theorem$ に一致する。

2.2 漸近的挙動

以後は秘書問題 ($p_i = 1/i$ の場合) において $n \to \infty$ の場合を詳しく調べよう。

補題 3

(a) 最適停止ルール：$n \to \infty$ のとき、

$$\frac{s^*}{n} \to t^* = \exp\left(-(m!)^{\frac{1}{m}}\right)$$

(b) 勝つ確率：$n \to \infty$ のとき、

$$V_0^{(m)} = \exp\left(-(m!)^{\frac{1}{m}}\right) \sum_{k=1}^{m} \frac{[(m!)^{\frac{1}{m}}]^k}{k!}$$

証明

$p_i = 1/i,\ q_i = 1 - 1/i$ より $r_i = 1/(i-1)$ である。従って、

$$\begin{aligned} R_{j,m} &= \sum_{j \le i_1 < \cdots < i_m \le n} r_{i_1} r_{i_2} \ldots r_{i_m} \\ &= \sum_{j \le i_1 < \cdots < i_m \le n} \left(\frac{1}{i_1 - 1}\right)\left(\frac{1}{i_2 - 1}\right) \cdots \left(\frac{1}{i_m - 1}\right) \end{aligned}$$

となり、$x_j = i_j/n,\ t = j/n$ として、$n \to \infty$ とすると $R_{j,m}$ は次の積分 $R_m(t)$ でリーマン近似できる。

$$\begin{aligned} R_m(t) &= \int \cdots \int_{t \le x_1 < \ldots < x_m \le 1} \left(\frac{dx_1}{x_1}\right)\left(\frac{dx_2}{x_2}\right) \cdots \left(\frac{dx_m}{x_m}\right) \\ &= \int_t^1 \frac{dx_1}{x_1} \int_{x_1}^1 \frac{dx_2}{x_2} \cdots \int_{x_{m-1}}^1 \frac{dx_m}{x_m} \\ &= \frac{(-\log t)^m}{m!} \end{aligned}$$

t^* は $1 = R_m(t)$ の根であるから、(a) が得られる。勝つ確率は

$$Q_{s^*} \sum_{k=1}^{m} R_{s*,k} = \left(\frac{s^* - 1}{n}\right) \sum_{k=1}^{m} R_{s^*,k}$$

となり、$n \to \infty$ とすると、これは次式で近似できる。

$$\begin{aligned} & t^* \sum_{k=1}^{m} R_k(t^*) \\ &= t^* \sum_{k=1}^{m} \frac{(-\log t^*)^k}{k!} \\ &= \exp(-(m!)^{\frac{1}{m}}) \sum_{k=1}^{m} \frac{(m!)^{k/m}}{k!} \end{aligned}$$

これより (b) が得られる。

3　拒否を考慮したモデル

成功の出現を観測し、停止を望んだ場合確率 β で停止可能であるが、残りの確率 $1-\beta$ で停止不可能である問題を考える。停止不可能の場合はさらに観測を継続することになる (2 節のモデルは $\beta = 1$ に対応している)。停止可能な成功をアベイラブル (*available*)、停止不可能な成功をアンアベイラブル (*unavailable*) と呼び、最後の m 個のアベイラブルな成功で停止すれば勝ちとなる問題を考察する。2 節と同様に

$v_i^{(m)}$: 最初の i ヶの対象をパスし、それ以降最適に振舞って勝つ確率
$g_i^{(m)}$: 時刻 i 成功がアベイラブルのとき、そこで停止して勝つ確率

と定義すると

$$v_{i-1}^{(m)} = p_i \max\{\beta g_i^{(m)} + (1-\beta) v_i^{(m)},\ v_i^{(m)}\} + q_i v_i^{(m)} \tag{3.1}$$

が成立する。ただし、今の場合 $g_i^{(m)}$ は補題 1 で p_i を $p_i^* \equiv p_i\beta$ で置き換えたものとなる。また、(3.1) の右辺前 1 項は

$$p_i\beta \max\{g_i^{(m)},\ v_i^{(m)}\} + p_i(1-\beta)v_i^{(m)}$$

と書き直すことができるので (3.1) は

$$v_{i-1}^{(m)} = p_i^* \max\{g_i^{(m)},\ v_i^{(m)}\} + (1-p_i^*)v_i^{(m)} \tag{3.2}$$

と表すことができる。(3.2) は、この問題が、2 節で p_i を p_i^* に置き換えた問題と等価であることを示している。

4 NHPP(*non − homogeneous Poisson process*)

2.2 節で、秘書問題の漸近挙動を調べたが、本節では NHPP を利用して補題 3 の結果を導く。このアプローチを利用すれば 3 節に対応する秘書問題の漸近挙動も容易に調べることができる。今時間区間 (0, 1) を n ケの等間隔に分割し、k 番目の応募者が時刻 $k/n,\ 1 \le k \le n$ に出現するものとする。*Presman and Sonin*[3] は、このようなセッティングにおいて $n \to \infty$ とすると、時間区間 $(x,\ 1)$ の間に出現する成功の数 (相対順位 1 の応募者の数)$N(x)$ は *intensity rate* $\lambda(t) = 1/t$ の NHPP に従うことを示した。即ち、

$$\int_x^1 \lambda(t)dt = \int_x^1 \frac{1}{t}dt = -\log x$$

であるから、

$$\begin{aligned} P\{N(x) = n\} &= e^{-\int_x^1 \lambda(t)dt}\frac{\{\int_x^1 \lambda(t)dt\}^n}{n!} \\ &= e^{\log x}\frac{(-\log x)^n}{n!} \\ &= x\frac{(-\log x)^n}{n!} \end{aligned} \tag{4.1}$$

となる。閾値ルールが最適であることから、時刻 x 以降の成功の出現個数が 1 以上 m 以下である確率を $f_m(x)$ とし、$f_m(x)$ の最大値を実現する $x = x^*$ を求めればよい。(4.1) より直ちに

$$\begin{aligned} f_m(x) &= P\{1 \le N(x) \le m\} \\ &= \sum_{n=1}^{m} \frac{x(-\log x)^n}{n!} \end{aligned} \tag{4.2}$$

が得られる。これを微分すると

$$f_m'(x) = \frac{(-\log x)^m}{m!} - 1$$

となる。従って x^* は $f_m'(x) = 0$ を解いて

$$x^* = \exp\{-(m!)^{(\frac{1}{m})}\} \tag{4.3}$$

となる。これを (4.2) に代入して

$$f_m(x^*) = \exp\left(-(m!)^{\frac{1}{m}}\right)\sum_{k=1}^{m} \frac{[(m!)^{\frac{1}{m}}]^k}{k!} \tag{4.4}$$

を得る。(4.3)、(4.4) は補題3の結果と一致している。さて、3節に対応する秘書問題は *intensity rate* $\mu(t) = \beta/t$ の NHPP を調べればよい。時間区間 $(x,\ 1)$ の間に出現するアベイラブルな成功の数を $M(x)$ で表すと

$$P\{M(x) = n\} = e^{-\int_x^1 \mu(t)dt} \frac{\{\int_x^1 \mu(t)dt\}^n}{n!} = x^\beta \frac{(-\beta \log x)^n}{n!} \tag{4.5}$$

となる。時刻 x 以降のアベイラブルな成功の出現個数が 1 以上 m 以下となる確率 $g_m(x)$ は (4.5) より

$$g_m(x) = P\{1 \leq M(x) \leq m\} = \sum_{n=1}^{m} \frac{x^\beta (-\beta \log x)^n}{n!} \tag{4.6}$$

で与えられる。従って

$$g'_m(x) = \beta x^{-(1-\beta)} \{ \frac{(-\beta \log x)^m}{m!} - 1 \}$$

となり、$g'_m(x) = 0$ 根 x^{**} は

$$x^{**} = \exp\{-\frac{(m!)^{\frac{1}{m}}}{\beta}\} \tag{4.7}$$

で与えられる。(4.7) を (4.6) に代入すると、最適な閾値 (4.7) の下で、成功する確率は

$$g_m(x^{**}) = \exp(-(m!)^{\frac{1}{m}}) \sum_{k=1}^{m} \frac{[(m!)^{\frac{1}{m}}]^k}{k!} \tag{4.8}$$

となる。(4.3) と (4.7)、(4.4) と (4.8) を比較すると $x^{**} < x^*$ であるが、勝つ確率は β に依存しないことが分り、興味深い結果となっている。

参考文献

[1] Bruss,F.T.(2000a), Sum the Odds to one and stop, *AnnalsofProbability* **28**,1384-1391

[2] Bruss,F.T.(2000b), Selecting a sequence of last successes in independent trials, *J.Appl.Prob.* **37**,389-399

[3] Presman,E.L,and Sonin,I.M.(1972), The best choice problem for a random number of objects, *TheoryProbab.Appl.* **17**,657-668

[4] Smith,M.H.(1975), A secretary problem with uncertain employment, *J.Appl.Prob.***12**,620-624

[5] Tamaki,M.(1991), A secretary problem with uncertain employment and best choice of available candidates, *OperationsResearch* **39(2)**,274-284

[6] Tamaki,M.(2001), A note on the Odds-theorem, 数理最適化の理論とアルゴリズム,**1241**,166-170, 京都大学数理解析研究所