最適停止問題の諸相

坂 口 　 実

最適停止問題は最も狭く定義すれば，2-決定(停止と継続)をもつ統計的逐次決定過程ということになろうが実社会面への応用は広く数学的にも研究に値する諸相をそなえている．本稿では逐次決定過程としての数学的な理論面には触れずに動的計画の見地からの種々の問題内容とその解とを概観する．1，2章では問題のおこりと構造の要点を3～6章ではそれらの展開と拡張につき述べる．

1. 2つの例題

1.1 結婚(あるいは秘書)の問題

n人の女性がいてまったく無作為的に1人ずつ貴君の前に現われるとする．i番目の女性に直面して，貴君は彼女と結婚するか，またはもっとよい女性が将来現われることを期待して，彼女を流してつぎの$(i+1)$番目の女性に対面するか，貴君はどちらかに決めねばならぬ．各女性のよさに順位がつけられて，最良が順位1とする．最良の女性と結婚できる確率を最大にするにはどういう政策をとればよいか？　もちろんリコールはない(1度流した女性にさかのぼって求婚はできない)とする．またn人目(すなわち最後)の女性まで結婚せずにきたならば，貴君は彼女と結婚せねばならない．

i番目の女性のみかけの順位(今までに対面したi人の中での順位のこと)をY_iとすると，

(1.1) $\Pr(Y_i=k)=1/i,\ k=1,\cdots,i$

で$\{Y_i\}_{i=1}^{n}$は独立列である．みかけの順位が1の女性を候補者(candidate)という．候補者でなければ最良であり得ないから，候補者に対面したとき彼女で停止するか流すかの選択が問題である．i番目の女性が候補者であるとき，以後最適にふるまって最良の女性と結婚できる(これを“勝つ”という)確率を$f(i)$とすると，

(1.2) $f(i)=\max[i/n,\ i\sum_{j=i+1}^{n} f(j)/(j-1)j]$
$(1\leqq i\leqq n-1;\ f(n)=1)$

が成立する．右辺カッコ内の第1(2)項は，彼女で停止した(を流した)場合の勝つ確率であり，$i/(j-1)j$はi番目の女性の後に，j番目で初めてまた候補者が出現する条件つき確率である．

(1.2)において〔命題〕$\Gamma\equiv\{i\mid$右辺カッコ内の第1項$\geqq$第2項$\}$とおくと$i\in\Gamma\Rightarrow(i+1)\in\Gamma$が成立することを利用すると，最適政策は：$1\leqq i<s^*$ではすべての女性を見送る．$s^*$番目以後に出現する最初の候補者で停止せよ．ただし$s^*$は$\sum_{j=s^*}^{n-1} j^{-1}<1$を満足する最小正整数である．各$n$に対する$s^*$の値，およびそのときの勝つ確率が，Gilbert-Mosteller[1]に出ている．nが大きいときに漸近的に$s^*\cong ne^{-1}\fallingdotseq 0.368n$, $\Pr(\mathrm{win})\cong e^{-1}$である．現実の人生では年令20才から40才までの20年間が結婚年令で，毎月1人ずつ新しい女性が現われるとすると240人は十分に多勢である．そうすると$20+0.368\times20\fallingdotseq27$才までは結婚せずにようすを見て，それ以後に出現する最初の候補者にproposeするのが最適である．

1.2 逐次抜取りの問題

$cdf\,F(x)$をもつ統計的母集団から1個ずつ独立にn回の逐次抜取りが許されていて，随意のときに抜取りをやめることができる．もしも$(n-1)$回まで停止しないできたならば，最後の$X_n=x_n$は必ず受取らねばならないとする．$X_j=x_j$, $j=1, \cdots, i$, を観察したとする．これを状態$(n, i|x_i)$で表わす．ここでもしも停止すれば利得はx_iであるが，継続すればつぎのX_{i+1}を観察することになる．どちらを選ぶかは(1)x_iがどれほど大きいか，(2)あと$(n-i)$回の抜取りが残っている，ことに依存する．貴君の期待利得を最大にする停止政策を求めよ．最適停止政策により得られる期待利得をv_nとすると，題意により明らかに，

(1.3) $v_n=E\{X\vee v_{n-1}\}$ $(n=2, 3, \cdots;\ v_1=EX)$

が成立する．関数$T_F(z)\equiv E[(X-z)^+]\equiv\int_z^\infty(x-z)dF(x)$を定義すると都合がよくて，これは連続・非負・とつ・非増加関数で，いつも直線$EX-z$の上側にあり$z\to-\infty$のときそれに漸近し，また$T_F(z)\to 0 (z\to\infty)$である．そうすると(1.3)より，

(1.4) $v_n=v_{n-1}+T_F(v_{n-1})$.

したがって$\{v_n\}$は増大列になる．一様・正規・指数など各種分布に対するv_nの数値は直ちに計算できて，坂口[2]に出ている．最適政策は：状態$(n, i|x_i)$においては$x_i>(<)v_{n-i}$ならば停止(継続)せよ，となる．

数値v_nは2通りの役割を担っている：(1) n回抜取りゲームにおける最適政策による期待額 (2) 状態$(n+i, i|x_i)$にいるとき，停止すべきかどうかの閾値である．

2. マルコフ決定過程としての最適停止問題

2.1 離散時刻の場合

前章の2例題を含めて，最適停止問題は離散，あるいは連続時刻のマルコフ決定過程であって，本質的に2個の決定：acception と rejection，あるいは停止と継続，をもつものである．状態i $(i=0, 1, 2, \cdots)$において停止すれば利得$R(i)$が生じ，継続すればコスト$C(i)$がかかるかわりに確率p_{ij}でつぎに状態jに移行する．$[p_{ij}]$は与えられた推移確率行列である．(利益)＝(利得)－(総コスト)を最大にするような停止政策を求めること．状態iから出発して得られる最大期待利益を$V(i)$とおくと，

(2.1) $V(i)=\max\{R(i), -C(i)+\sum_{j=0}^{\infty}p_{ij}V(j)\}$, $i=0, 1, 2, \cdots$

が成立する．たとえば結婚の問題では，停止しそこなったという状態∞と$V(\infty)=0$を追加すると，

$$R(i)=i/n,\ R(\infty)=0,\ C(i)\equiv 0,$$

$$p_{ij}=\begin{cases} i/(j-1)j, & i+1\leqq j\leqq n \\ i/n, & j=\infty \end{cases},\quad p_{\infty\infty}=1$$

のもとで最適方程式(2.1)が成立している．

最適政策や$V(i)$の存在自体が自明でないときはつぎのように考える．状態iから出発して高々n段進行したのち停止するときの期待利益の最大値を$V_n(i)$とすると，(2.1)の左(右)辺のVをV_n (V_{n-1})に変えた式が成立する．$\sup_i R(i)<\infty$, $\inf_i C(i)>0$ならば$V_n(i), n\to\infty$, が収束する(このとき過程は stable であるという)ので，この極限を$V(i)$とおくと(2.1)が成立する．

2.2 OLA停止政策

最適方程式が(2.1)で記述されるマルコフ決定過程が stable であるとする．

(2.2) $B\equiv\{i|R(i)\geqq -C(i)+\sum_{j=0}^{\infty}p_{ij}R(j)\}$

とおく．Bは，今すぐ停止するほうが，あと1段だけ継続してから停止するよりも有利であるような状態iの全体を表わす．そうすると，もしもBが“closed”，すなわち，

(2.3) $p_{ij}=0,\ \forall i\in B, \forall j\notin B$

ならば，$i\in B$のとき，かつそのときに限り停止するのが最適である(Ross の本[3]の§6.5)．この政策を one-stage look ahead (略してOLA)政策という．

OLA政策は導出が容易であって，多くの問題に対して直裁にその最適政策を与える．たとえば結婚の問題では$B=\{i|i/n\geqq\sum_{j=i+1}^{n}\{i/(j-1)j\}\cdot j/n\}=\{i|1\geqq\sum_{j=i}^{n-1}j^{-1}\}=\{i|s^*\leqq i\leqq n\}$であり，$B$は明ら

かに “closed” すなわち (2.3) を満足するから，$i \in B$になり次第停止するのが最適である．もう1つ例題を示そう．**1.2**の逐次抜取りの問題で，$n=\infty$として抜取りコストが毎回$c>0$かかるとする．$p=\langle p_0, p_1, \cdots, p_k\rangle$を所与の確率ベクトルとして，$\Pr(X=i)=p_i,\ i=0,\cdots,k,$とする．今度はリコールが許される（すでに流した値を，さかのぼって accept してもよい）とする．十分多数回やれば最大値kが出てくるであろうが，コストがかさんで損になる．最適停止政策を求めよ．今までの観察値の最大がiであったという状態をiで表わすと，推移確率と最適方程式は，

$$p_{ij}=\begin{cases}0, & j<i\\ p_0+p_1+\cdots+p_i, & j=i\,;\\ p_j, & j>i\end{cases}$$

$$V(i)=\max\{i,\ -c+\sum_{j=0}^{k} p_{ij}V(j)\}$$

になる．$B\equiv\{i|i\geqq -c+\sum_{j=0}^{k} jp_{ij}\}=\{i|c\geqq\sum_{j=i+1}^{k}(j-i)p_{ij}\}$になるが，$T_F(i)=\sum_{j=i+1}^{k}(j-i)p_{ij}$は$i$の減小関数なので$B$は “closed” である．ゆえに$B$は最適停止域である．

2.3 連続時刻の場合とＩＬＡ政策

決定過程が離散時刻でなくて連続時刻の場合もごく普通にあることである．前節の例題をそのまま連続時刻にして説明しよう．家を売る問題：offer あるいは機会(買手が現われること)が到着率λの Poisson過程でやってくるとする．各 offer には “大きさ” (買手が提示する価格)が伴い，それは*cdf* $F(x)$をもち，逐次に到来する offers の大きさ$X(t_i), i=1,2,\cdots$(t_iは到着時刻)は*iid* r.v. 列をつくる．すでに流した offer をも recall してよいとする．時刻tにおいてそれまでに到来した最大の offer をY_tとする．このとき過程を停止すれば利益は，

(2.4) $f(t, Y_t)=Y_t-ct$

であるとする．$c>0$は単位時間当りの待ち費用である．期待利益を最大にする停止政策を求めよ．

この問題では底流する確率過程W_tが Poisson 過程であったが，その他にも Wiener 過程や一般の独立増分の過程である場合にも，類似の問題は多い．一般に時刻tにおいて$W_t=x$のときに，今すぐ停止すれば利得$f(t,x)$が得られるとする．これは$\{(t,x)|t,x>0\}$において連続・有界とする．

(2.5) $T_hf(t,x)\equiv E[f(t+h, W_{t+h})|W_t=x]$,

(2.6) $Af(t,x)\equiv\lim_{h\to 0+} h^{-1}\{T_hf(t,x)-f(t,x)\}$

を定義する．前節でのＯＬＡ政策の類似をとって，$B\equiv\{(t,x)|Af(t,x)\leqq 0\}$を停止域にするのを infinitesimal look ahead (略してＩＬＡ)政策という．［定理］Bが “closed”，すなわち，

(2.7) $\Pr\{$ある有限な$t>0$において

$(t, W_t)\notin S|(0,x)\in S\}=0$

ならば，Bは最適停止域を与える(Prabhu[4])．前記の家を売る問題では$E[f(t+h,Y_{t+h})|Y_t=y]=\lambda h\cdot\{F(y)f(t+h,y)+\int_y^\infty f(t+h,z)dF(z)\}+(1-\lambda h)f(t+h,y)+o(h)$となるから$Af(t,y)=\partial f/\partial t+\lambda\int_y^\infty(f(t,z)-f(t,y))dF(z)$．これに(2.4)を代入すると$Af(t,y)=-c+\lambda T_F(y)$を得る．そこで$B\equiv\{(t,y)|T_F(y)\leqq c/\lambda\}$となるから，$0<c<\lambda E(X)$を仮定しておけば$B=\{(t,y)|y\geqq T_F^{-1}(c/\lambda)\}$となり，これは “closed”，すなわち(2.7)を満足するから，Bは最適停止域である．

3. 結婚の問題への再訪

3.1 断わられる確率のある場合

結婚の問題(1.1節)で女性に propose しても彼女が断わるかも知れなくて，その確率が$0\leqq q=1-p<1$とする．もしも断わられたならばつぎの女性に進む．propose した女性が承諾したとき初めて彼女は available になる．最良の女性がそうなることを win として Pr(win)を最大にせよ．Smith[5]によれば最適政策は**1.1**におけると同じで，ただs^*をより小さいs^0に取替えたものである．漸近的に$s^0\cong np^{1/q}$, $\Pr(\text{win})\cong p^{1/q}$になる．

3.2 順位 i の女性のもつ効用

順位iの女性と結婚することの効用をu_iとし，$1\geqq u_1\geqq u_2\geqq\cdots\geqq u_n\geqq 0$とする．**1.1**では$u_1=1$,

$u_2=\cdots=u_n=0$ であった．r-th girl がみかけの順位 $k(1\leqq k\leqq r)$をもっていることを状態(r,k)で表わす．いま$(r-1)$人を reject してきて，状態(r,k)にきたとする．この girl がみかけの順位 k をもつという条件のもとで，順位 i をもつ条件つき確率は $n^{-1}\left\{\binom{i-1}{k-1}\binom{n-i}{r-k}\Big/\binom{n-1}{r-1}\right\}\Big/r^{-1}$ であるから，この girl を accept することの期待効用は $Q(r,k)\equiv\sum_{i=k}^{n-(r-k)}u_i\binom{i-1}{k-1}\binom{n-i}{r-k}\Big/\binom{n}{r}$ となる．状態(r,k)から出発して以後最適政策を用いて得られる期待効用を $v_n(r,k)$とおくと，

$$(3.1)\quad v_n(r,k)=\max\{Q(r,k),\ (r+1)^{-1}\sum_{k'=1}^{r+1}v_n(r+1,k')\}$$
$$(1\leqq k\leqq r,\ 1\leqq r\leqq n-1,\ v_n(n,k)\equiv u_k)$$

が成立する．みかけの順位 k をもつ girl を k-候補者(記号でC_k)ということにする．［定理］$r_1^*\leqq r_2^*\leqq\cdots\leqq r_n^*$ が存在して最適政策は：$r\in[1,r_1^*)$ ではすべての girl を見送る．$[r_1^*,r_2^*)$に現われる最初のC_1を accept，$[r_2^*,r_3^*)$に現われる最初の C_1 or C_2 を accept，…，$[r_n^*,n)$に現われる最初の $C_1,\cdots,$ or C_n を accept せよ(Mucci[6])．だから r-k 平面で図示すると最適停止域は高々 n 段をもつ階段になる．$\{r_j^*\}$ の値は $\{u_i\}$ から算定され，たとえば **1.1** の問題では $r_2^*=\cdots=r_n^*=1$ である．

はじめの$(r-1)$人を見送るという制約つきの政策の中での最適政策による期待効用を $\bar{v}_n(r)\equiv r^{-1}\sum_{k=1}^{r}v_n(r,k)$とおくと，(3.1)より，

$$(3.2)\quad \bar{v}_n(r)=r^{-1}\sum_{k=1}^{r}\max\{Q(r,k),\bar{v}_n(r+1)\}$$
$$(1\leqq r\leqq n-1,\ \bar{v}_n(n)\equiv n^{-1}\sum_{i=1}^{n}u_i)$$

を得る．いま時点 r/n において r-th girl が出現する，と考えて問題を連続時刻に直してしまう．$g(r/n)=\bar{v}_n(r)$とおき $r/n=\alpha$ に保ちながら $n,r\to\infty$ にゆくと，$n^{-1}\sum_i u_i\to 0$ を仮定しておくと，(3.2) は $0\leqq\alpha\leqq 1$ での微分方程式

$$(3.3)\quad g'(\alpha)=-\alpha^{-1}\sum_{k=1}^{\infty}(R_k(\alpha)-g(\alpha))^+,\ g(1)=0$$

に帰する．ここで $Q(n\alpha,k)\to R_k(\alpha)\equiv\sum_{i\geqq k}u_i\binom{i-1}{k-1}\alpha^k(1-\alpha)^{i-k}$ を使っている．方程式 $R_k(\alpha)=g(\alpha)$ は単一根をもち，それを α_k とおくと $r_k^*\cong n\alpha_k$ $(k=1,\cdots,n)$が示される．たとえば$u_1=u_2=1,u_i=0$ $(i\geqq 3)$のときは $\alpha_1\doteqdot 0.35$ は方程式$-\log(3\alpha/2)=1-\alpha$の単一根であり，$\alpha_2=2/3,\alpha_k\equiv 1(k\geqq 3)$，さらにこのときの期待効用は $g(\alpha_1)=2\alpha_1-\alpha_1^2\doteqdot 0.58$ である．

3.3 有限リコールのある場合

結婚の問題で過去m人までさかのぼって求婚できるとする．i-th girl($i+m<n$のとき)が候補者であり，さらにそれに続く$\left\{\begin{matrix}k(1\leqq k\leqq m-1)\\ m\end{matrix}\right\}$人がみな候補者でないことがわかったときに彼女のことを$\left\{\begin{matrix}一時的\\ 真\quad 正\end{matrix}\right\}$候補者といい$\left\{\begin{matrix}TC\\ VC\end{matrix}\right\}$で表わす．$TC$ が VC になったことがわかって初めて，その girl を accept すべきか否かが問題になる．Smith-Deely[7] によると，ある正整数 t^* が存在して最適政策は：はじめの(t^*-1)人はすべて見送る．t^* 番目以後に現われる最初の VC を accept せよ．Yang[8] は同じ問題に対してまた別の接近をしている．

3.4 2回以上 accept する場合

結婚の問題で，r girls を accept してそのどれかが順位1であれば win とする．i-th girl が候補者であるとき，あと最適にふるまって得られる Pr(win)を $f_r(i)$とおくと，

$$(3.4)\quad f_r(i)=\max\left[\sum_{j=i+1}^{n}if_{r-1}(j)/(j-1)j+i/n,\ \sum_{j=i+1}^{n}if_r(j)/(j-1)j\right]$$
$$(1\leqq i\leqq n-1,\ r\geqq 1,\ f_0(i)\equiv 0)$$

が成立する．右辺[…]内の左(右)項は，この girl を accept(reject)したのち最適にふるまって得られる Pr(win)である．たとえば $r=2$ のときの最適政策はつぎのようになる(Gilbert-Mosteller[1])：正整数 $s^0<t^0$ が存在して，はじめの(s^0-1)girls は見送る．s^0-th 以後の最初の候補者で 1st stop，それ以後，かつ t^0-th 以後の最初の候補者で 2nd stop をせよ．漸近的には $t^0\cong ne^{-1}\doteqdot 0.3679n$，$s^0\cong ne^{-3/2}\doteqdot 0.2231n$，$\Pr(\text{win})\cong e^{-1}+e^{-3/2}\doteqdot 0.5910$ である．

欲張っているが，もしも2回 accept して順位

1および2の girls を得たとき win とすればどうであるか？ Nikolaev[9], Sakaguchi[10]により答だけを書くと，$s^*<t^*$ が存在して最適政策は：はじめの(s^*-1) girls は見送る．s^*-th 以後の最初のC_1（定義は**3.2**にある）で 1st stop，それ以後の最初のC_1または t^*-th 以後の最初のC_1 or C_2で 2nd stop をせよ．漸近的には $t^*\cong ne^{-1/2}\fallingdotseq 0.6065n$, $s^*\cong n\alpha$，ただし $\alpha\fallingdotseq 0.2291$ は方程式 $\alpha-\alpha^{-1/2}\log\alpha=(7/2)\ e^{-1/2}-1$ の単一根であり，このとき $\Pr(\text{win})\cong\alpha(2e^{-1/2}-\alpha)\fallingdotseq 0.2253$である．

3.5 期待順位を最小にする問題

Chow-Moriguti-Robbins-Samuels[11] は結婚の問題に対して別の接近をしていて，そこでは目標は期待順位を最小にすることである．r-th girl のみかけの順位を Y_r で表わすと，$Y_r=k$ であったときの彼女の順位の条件つき期待値は，(1.1) および**3.2**の $Q(r,k)$ より $\sum_{i=k}^{n-(r-k)} i\binom{i-1}{k-1}\binom{n-i}{r-k}\Big/\binom{n}{r}=\frac{n+1}{r+1}\cdot k$ になる．はじめの r girls はすべて見送るという政策の中での期待順位の最小値をv_rとおくと，最適方程式

$$(3.5)\quad \begin{aligned} v_{r-1}&=E\ \min\{(n+1)Y_r/(r+1), v_r\}\\ &=\frac{n+1}{r(r+1)}\sum_{k=1}^{r}\min\{k, (r+1)v_r/(n+1)\}\\ &\qquad(1\leqq r\leqq n-1, v_{n-1}=EY_n=(n+1)/2)\end{aligned}$$

が成立する．ゆえに $Y_r\leqq(r+1)v_r/(n+1)$になり次第停止するのが最適である．たとえば $n=4$ のとき v_3, v_2, v_1, v_0 の値を算定したのち最適政策は：1st girl は見送る．2nd girl は $Y_2=1$ ならば accept, 3rd girl は $Y_3=1$ or 2 ならば accept せよ．このときの期待順位は $v_0=1.875$である．v_0の値はもちろんnの関数であるが，$n\to\infty$ のとき有限確定値に収束することは驚くべきことと言ってよい．［定理］$\lim_{n\to\infty} v_0=\prod_{k=1}^{\infty}(1+2/k)^{1/(k+1)}\fallingdotseq 3.8695$.

また最近 Rubin-Samuels[12]は memory の観点から，この問題に別の面白い接近をしている．そこでは，いつも今までに出現した候補者だけしか記憶できなくて，対面する各 girl は記憶のgirl と比べてよいか悪いかのどちらかだけが観察される．したがって可能な決定に今までの reject, accept の他に remember が入ってくる．

4. 逐次抜取りの問題への再訪

4.1 逐次割当の問題

$X_1,\cdots,X_n$ がこの順に1個ずつやってくる．n 個の数値$(0<)\leqq p_1\leqq\cdots\leqq p_n(\leqq 1)$がある．到来した各 X_i にどれかの p_j を掛け合わせて $\sum_{i=1}^{n} X_i p_{j_i}$ を最大にせよ．ここで$(j_1,\cdots,j_n)$は$1,\cdots,n$の順列である．もしも全部の r.v. $X_1,\cdots,X_n$ が同時に到来するのならば，それらの観察値を大きさの順に並べて $x_{i_1}\leqq\cdots\leqq x_{i_n}$ として $\sum_{j=1}^{n} x_{i_j}p_j$ を作ればよい（周知の Hardy-Littlewood-Polyá の補題）．だから $\{X_i\}$ の到来が逐次なことが重要である．この問題は**1.2**の逐次抜取りの問題の拡張になっている．実際$p_i=0(1\leqq i\leqq n-k), =1(n-k+1\leqq i\leqq n)$にとると，$k$回acceptできて，これらaccept した観察値の和を最大にするのを目的とする問題になる．さて所求の最大値を$g_n(p_1,\cdots,p_n)$とおくと，

$$(4.1)\quad g_n(p_1,\cdots,p_n)=E[\max\{p_jX_1+g_{n-1}(p_1,\cdots,\overset{j}{\check{*}},\cdots,p_n)\}]\ (n\geqq 2, g_1(p_1)=p_1EX),$$

ただし $\overset{j}{\check{*}}$ は j 番目が欠けていることを示す，が成立する．［定理］$a_{1,n}\leqq\cdots\leqq a_{n-1,n}$ が存在して最適割当政策は：まず $X_1\in[a_{i-1,n},\ a_{i,n}]$ ならば X_1 には $p_i(i=1,\cdots,n; a_{0,n}\equiv-\infty, a_{n,n}\equiv+\infty)$を割当てよ．以降は$(p_1,\cdots\overset{i}{\check{*}},\cdots,p_n)$に対して最適にふるまえ．このとき期待利得は $\sum_{i=1}^{n} p_i a_{i,n+1}$ であり，

$$(4.2)\quad a_{i,n+1}=\int_{a_{i-1,n}}^{a_{i,n}} x\,dF(x)+a_{i-1,n}F(a_{i-1,n})+a_{i,n}(1-F(a_{i,n}))$$

が成立する(Derman-Lieberman-Ross[13])．$a_{i,n}$などは$F(\cdot)$だけに依存してp_jなどに依存しない．$p_1,\cdots,p_n$ が残っているとき p_i が割当てられるような X の期待値が $a_{i,n+1}$ である，と解釈すれば (4.2)は分りやすい式である．数値例：$F(x)=x$, $0\leqq x\leqq 1$, のとき(4.2)は $a_{i,n+1}=a_{i,n}-(1/2)(a_{i,n}^2-a_{i-1,n}^2)$ $(n\geqq 2, a_{1,2}=1/2)$．そこで $a_{i,n}$ は $n=2$

のとき 1/2; $n=3$ のとき 3/8, 5/8; $n=4$ のとき39/128, 1/2, 89/128; …などになる.

4.2 Random walk の上の最適停止

戦争をやめるには過去の戦闘で引続いて何度も勝っているうちがよい. $\{X_t\}$を各X_tが$\Pr(X=1)=p$, $\Pr(X=0)=q=1-p$, という Bernoulli 分布に従う *iid* r.v. 列とする. $X=1$ が実現することを success という. k 回の試行ののち停止したならば, 利得は (そのときの success run) $-kc$ であるとする. $0<c<1$ は 1 回当りの試行コストである. 期待利得を最大にするような停止政策を求めよ. あと j 回の試行が残されていて, いま success が r 回続いているという状態を(j,r)で表わす. この状態から出発して最適にふるまって得られる期待利得を $V_j(r)$とおくと,

$$(4.3)\quad V_j(r)=\max\{r, -c+pV_{j-1}(r+1)+qV_{j-1}(0)\},\quad V_0(r)\equiv r$$

が成立する. $V_j(0)$を単にV_jと書く. いま問題を少し修正して $X=0$ が起ったならば stop を強制される. もし状態$(j,0)$で stop すればV_jをもらうとする. この修正問題に対しては最適方程式は $W_j(r)=\max\{r, -c+pW_{j-1}(r+1)+qV_{j-1}\}$ となり, OLA停止政策(2.2節)は停止域$B=\{(j,r)\,|\,r\geqq -c+p(r+1)+qV_{j-1}\}=\{(j,r)\,|\,r\geqq(p-c)/q+V_{j-1}\}$ をもつ. V_j は単調増大のゆえに B は "closed" になるから最適停止域である. $n=\infty$ である場合は (4.3) は意味をもたないが, このときも上と類似の解を導くことができる (Starr[14], Ross[15]). 成功ならば 1 つ上の状態に進み失敗ならば最悪の状態に落ちる, という事態は quiz や gamble で多く見られる.

4.3 One-armed-bandit の問題

2 台のパチンコ機械A, Bがある. 当りの確率をそれぞれ p, q として, q の値は既知であるが, p の値はまったく未知であるとする. 全部で n 回パチンコをやるのに, 当りの回数の期待値を最大にするには, どのように逐次に A or Bを選べばよいか? もしも $p>(<)q$ であることがわかっていれば, 機械A(B)ばかりを n 回やれば最適であるに決まっている. しかし p の値が未知であるから機械Aをやりながら p の真値を推定しつつ, 同時にそれを選択のために活用しなければならない. この種の問題は標記(略してOAB)のようによばれ, 推測統計学において広範な研究がある. two-armed-bandit (略してTAB)の問題は, この自然な発展である. 参考文献を 1 つずつだけ挙げると Woodroof[32], Feldman[33].

5. 連続時刻確率過程の上の最適停止

5.1 取替過程

ある機械の作動状態が実数値 x で表わされ, 時刻 t のときの状態 X_t は平均速度 0, 分散速度 1 の Wiener 過程で記述されるとする. $\Pr(X_0=0)=1$, 状態 x のときに単位時間当り運転費用が ax^2とする. 理想状態 0 から外れる程運転費がかさむので適当に取替えをする. この費用は 1 回につき$K>0$とする. (時間)平均的な期待費用を最小にするには取替政策をいかにすべきか? 状態 x のときから出発して最適取替政策により得られる割引期待費用を $V(x)$とおくと, 最適方程式

$$(5.1)\quad V(x)\doteq\min[K+V(0),\ ax^2\Delta t+e^{-\alpha\Delta t}E\{V(x+\Delta x)\}]$$

が成立する. 右辺において取替えをすると瞬時に状態0にもどること, 時間Δt後にかかる費用は$e^{-\alpha\Delta t}$倍に割引されるとしている. Δx は時間 Δt における変動だから正規分布 $N(0,\Delta t)$に従う r.v. である. この式で $\alpha\to 0$ のとき $V(x)-V(0)\to f(x)$, $\alpha V(0)\to\gamma$ として $\alpha\to 0$ にゆくと,

$$(5.2)\, f(x)=\min[K, (ax^2-\gamma)\Delta t+E\{f(x+\Delta x)\}]$$

になる. γ は(時間)平均的な期待費用を表わす. $E\{f(x+\Delta x)\}=f(x)+1/2\Delta tf''(x)+o(\Delta t)$を代入すると, 取替えをしない間: $|x|\leqq\lambda$, では $1/2f''(x)=-ax^2+\gamma$, したがって, $f(0)=f'(0)=0$ のもとで積分して $f(x)=-(1/6)ax^4+\gamma x^2$. これに条件 $f(\lambda)=K$ を代入すると $\gamma=(1/6)a\lambda^2+K\lambda^{-2}$. これを $\lambda>0$ につき最小にすると,

$\lambda=(6K/a)^{1/4}$, $\min \gamma=(2/3)^{1/2}(Ka)^{1/2}$

になる．機械運転の制御（Taylor[16]，Bather[17]）の他に，ダムの水量制御(Faddy[18])，企業の株主への配当(Foster[19]）などの問題もこれに属する．

5.2 配分過程

潜水艦が魚雷をもって作戦に従事している．targetが到着率λのPoisson過程で出現する．各targetには大きさXが伴っていて，それはcdf $F(x)$をもち，逐次に到来するtargetの大きさ$X(t_i)$, $i=1,2\cdots$, はiid列をつくる．到来したtargetにj発の魚雷を一斉射撃すれば，それをとる確率は$1-q^j$, $0\leqq q=1-p<1$, とする．作戦時間tとn発の魚雷とが与えられたとき，とったtargetの大きさの和を最大にするような政策を求めよ．所求の最大値を$V_n(t)$とおくと微分方程式

$$(5.3)\quad V_n'(t)=\lambda E\left[\max_{1\leqq j\leqq n}\{(1-q^j)X+V_{n-j}(t)\}-V_n(t)\right]^+,\quad V_n(0)=0$$

が成立し最適政策は：全区間$[0,-\infty)$が$(n+1)$個の部分区間に分割されて，出現したtargetの大きさが左からj番目の部分区間に落ちたならば$(j-1)$発の魚雷を発射せよ，となる（Sakaguchi[20]）．Donis-Pollock[21]，Albright[22]，Mastran-Thomas[23]，Samuel[24]などの論文で扱っている問題は，すべてこの問題の特別な場合か変形である．

5.3 湖で魚をとる

湖に魚がN尾いるのを1尾ずつ捕える．どの魚も釣り上げられるまでの時間Zは指数分布$\Pr\{Z\leqq t\}=1-e^{-\mu t}$をもつ$iid$ r.v.である．n-th capture timeをu_nとおくと$u_{n+1}-u_n$は平均値μ^{-1}の指数分布r.v.の$(N-n)$個の最小値であり，したがって平均値$\{(N-n)\mu\}^{-1}$の指数分布に従う．釣り始めてから時間tで$Y_t=n$尾をとったときに停止すれば利得は$f(t,n)=n-ct$であるとする．$c>0$は単位時間当りの待ち費用である．期待利得を最大にする停止政策を求めよ．(2.3節参照)．

$E[f(t+h, Y_{t+h})\,|\,Y_t=n]=f(t+h,\ n+1)\ (N-n)\mu h+f(t+h,n)\{1-(N-n)\mu h\}+o(h)$であるから(2.6)において$Af(t,n)=\partial f/\partial t+(N-n)\mu\{f(t,n+1)-f(t,n)\}=-c+(N-n)/\mu$となり，$B\equiv\{(t,n)\,|\,Af(t,n)\leqq 0\}=\{(t,n)\,|\,n\geqq N-c/\mu\}$になる．この$B$は明らかに"closed"だから最適停止域を与える．現実問題としてはNが未知であろうから，このときの統計的接近をもStarr[25]はやっている．

6. いろいろな拡張

6.1 多次元確率変数およびゲームへの拡張

1.2節の逐次抜取りの問題において，$F(x)$を2変量cdf $H(x,y)$にしてみるのが拡張の第1歩である．①今度は停止時のxもyもどちらも大きくしたいので，そのように目標を設定する必要がある（Sakaguchi[26],[27]）．②playerが2人いて，player Ⅰ，Ⅱは停止時のそれぞれx,yを最大にしたい．停止の規則として2人のうち両方とも停止したときにprocessが停止するとか，あるいはどちらか一方が停止すれば相手がたとえ継続したくともprocessが停止するとか決めておく．問題は逐次ゲームになる．本稿に続く蔵野氏他の稿があるので詳しくはそちらに譲る．

6.2 陪審員の選定

検察側と弁護側(それぞれplayer Ⅰ，Ⅱとする)とでJ人の陪審員を選定したい．地域住民からランダムに何十人かの有資格者を抽出し，各人に逐次に面接して陪審員として適当であるか否かを査定する．Ⅰ，Ⅱは同時に，しかし独立に，彼をacceptかrejectかを決める．両方ともacceptのときだけ彼は陪審員として"選定"される．Ⅰ，ⅡはそれぞれA，B回reject（忌避）する権利をもっている．X_i $(i=1,2,\cdots)$はi番目の人がもしも陪審員となったら，彼が有罪に投票する確率である．$\{X_i\}_{1\leqq i\leqq J+A+B}$は$iid$で共通の$cdf$を$F(x)$とする．選定された陪審員たちを$\{i'\}$，それらのもつ確率を$\{X_{i'}\}$で表わすと，この逐次ゲームの目標は，Ⅰ(Ⅱ)にとって$E(\Pi X_{i'})\rightarrow\max(\min)$にすることである．あと$j$人の陪審員を選定しなければならなくて，そしてⅠ，Ⅱはそれぞれあとa,b

回の忌避権をもっている，という状態を$\sigma=(j,a,b)$で表わす．状態δにおけるゲームの値を$V(\sigma)$，また$X=x$を観察したときのゲームの条件つき値を$V(\sigma|x)$とおくと，

$$V(\sigma)=EV(\sigma|X),$$

$$V(\sigma|x)=\mathrm{val}\left\{\begin{array}{c|cc} & \mathrm{rej} & \mathrm{acc} \\ \hline \mathrm{rej} & V(\delta) & V(\alpha) \\ \mathrm{acc} & V(\beta) & xV(\gamma) \end{array}\right\}$$

が成立する．ここで$\alpha=(j,a-1,b)$，$\beta=(j,a,b-1)$，$\gamma=(j-1,a,b)$，$\delta=(j,a-1,b-1)$はσから移り得る4つの状態である．Brams-Davis[29]はこのゲームを解き数値例を与えている．もしも両人の決定が同時ではなくて，先にIが，つぎにIIが決定する規則の場合は逐次双辺ゲームになる．先手のIがacceptしたときだけIIはrejectかacceptかを選ぶ．今度はIは故意に相手の忌避権を使わせるようふるまえるからIがIIより有利である(Roth-Kadane-Degroot[30]，Sakaguchi[28])．

6.3 連続時刻の非協力ゲーム

1.1節の結婚問題(これをA_nで表わす)をゲームに直すことを考える．m人の競争者がいて各人が独立に別個の問題A_nに直面する．最良の女性をacceptしたものの中で最も早くacceptしたものが勝ちとする．誰がいつstopしたかは他人に一切知らされない．各人は自分の勝つ確率を最大にしようとする．この非0和ゲームを$A_n{}^m$で表わす．さて各時点i/n $(i=1,2,\cdots,n)$においてi-th girlが出現すると考えて，$i/n=x$に保ちながら$i,n\to\infty$にゆくと$[0,1]$での連続時刻ゲームになる(3.2節参照)．たとえばA_nにおける最適方程式(1.2)はA_∞では，

$$V(x)=\max\{x,\int_x^1 xV(y)/y^2dy\},\quad V(1)=1$$

という積分方程式になり，解$V(x)=x\cup e^{-1}$をもつ．［定理］問題$A_\infty{}^m$の解はつぎの通り：方程式$-z\log z=1-(1-mz)^{1/m}$の$(0,1)$での単一根を$z_m{}^*$とおく．各playerが時点$z_m{}^*$以後に出現する最初の候補者でstopするのが，ただ1つ平衡戦略である．$z_1{}^*=e^{-1}=0.3679$，$z_2{}^*=0.2953$，$z_5{}^*=0.1659$，$z_{10}{}^*=0.0915$である．Presman-Sonin[31]は，これを一般に到着密度$\lambda(x)$のPoisson流(ただし$0\leqq x\leqq 1$)に対して拡張している．

参考文献（雑誌名は再出以降は略記を用いた．）

1) *J. Amer. Stat. Assn*, **61** ('66) 35-73.
2)「経済分析と動的計画」東洋経済新報社，昭45.
3) *Applied Probability Models with Optmization Applications*, Holden Day, 1970.
4) *Nav. Res. Log. Q.* **21** ('74), 411-418.
5) *J. Appl. Prob.* **12** ('75), 620-624.
6) *Ann. Stat.* **1** ('73), 104-113.
7) *JASA* **70** ('75), 357-361.
8) *JAP*, **11** ('74), 504-512.
9) *Th. Prob. Appl.* **22** ('77), 187-190.
10) *Math. Jap.* **23** ('78), 647-653.
11) *Israel J. Math.* **2** ('64), 81-90.
12) *Ann. Prob.* **5** ('77), 627-635.
13) *Manag. Sci.* **18** ('72), 349-355.
14) *Ann. Math. Stat.* **43** ('72), 1884-1893.
15) *AS*, **3** ('75), 793-795.
16) *Technometrics* **9** ('67), 29-41.
17) *Math. Oper. Res.* **1** ('76), 209-224.
18) *JAP*, **11** ('74), 111-121.
19) *ibid* **12** ('75), 457-465.
20) *MJ*, **21** ('76), 89-103.
21) *NRLQ*, **14** ('67), 513-527.
22) *MS*, ('74), 60-67.
23) *NRLQ*, **20** ('73), 661-672.
24) *JAP*, **7** ('70), 157-164.
25) *ibid* **11** ('74), 294-301.
26) *J. Opns. Res. Soc. Jap.*, **16** ('73), 186-200.
27) *ibid* **21** ('78), 45-58.
28) *ibid*, **21** ('78), 486-508.
29) *Oper. Res.* **26** ('78), 966-991.
30) *ibid*, **25** ('77), 901-919.
31) *TPA*, **20** ('75), 770-781.
32) *Sankhyā* **38** ('76), Ser. *A*, 79-91.
33) *AMS*, **33** ('62), 847-856.