最適停止問題とその周辺

――逐次選択過程――

生田誠三

1. はじめに

最適停止問題にはいろいろなタイプがあり，その定義は必ずしも明確ではないが，ここではこの問題を，基本的には“有限時間内に下さなければならない決断をのばしつづけるという行為をいつどのようなときに停止すべきか”ということ，換言すれば，“下さなければならない決断をいつまでのばすか”ということに関連した問題としてとらえることにする．

このような決断は，われわれ自身の日常生活や人生，野球やギャンブルなどのゲームの世界，株や穀物相場などの投機の世界，企業経営，政治・経済社会…等，随所に見られるきわめて人間的な現象のひとつであると言える．しかもこのような決断は往々にしてある種の異常な緊迫感とともにわれわれの目前に迫ってくるものである．場合によっては，それは個人や組織の存亡にもかかわる重大な問題となることさえある．

それゆえにこの種の決断行為は広く人間行動や社会行動を理解していくうえで欠くことのできない重要な要因のひとつであるとも言えよう．

たとえば卒業時までに順次やってくる就職口に対しいつどこで最終的な決断（就職先を決めるという決断）を下すかということは，多くの人々が人生のひとコマとして，少なからず不安と緊張のうちに経験することである．失業保険の有効なうちにできるだけ有利な再就職の口を決めるという問題は失業者にとっては死活の問題であろうし，適齢期のうちに時折やってくるお見合話の中からいかにしてすてきな結婚相手を見つけるかという問題は，友達がつぎつぎに結婚してしまう女性にとってはゆゆしき一大事であり，停年までにできるだけ安くて良い家を探し，購入するという問題は，老後の生活設計を真剣に考えている人々にとっては深刻な問題である．スポーツの世界では，黒星つづきの力士が休場届を出すのをいつまでのばすかという問題，株の世界では上昇気味の株をいつ売却するかという問題，企業では，R＆Dプロジェクトの商品化の可否の決定をいつまでのばすかという問題……等々がある．

このようなきわめて人間的な問題に対しては2通りの研究方向が考えられる．ひとつは，その決断の当事者の心理ドラマを追求し，そこからこのような緊迫した状況下における人間行動についての知見や洞察を得ようという心理学的・行動学的な研究領域であり，もうひとつは，この問題のもっている論理的構造を数学的・確率論的立場から明らかにしていこうという研究領域である．この2つの領域は決して対立するものではない．上記したような現実世界の生々しい最適停止問題に対し真の実証的・学問的解決を得るためには，それぞれの領域から得られる知識，知見，洞察を互いに接触させ，融合し，総合し，より次元の高い有効な理論体系を順次構成していくことが必要であろう．

本稿の目的は，後者の立場から，ある特有の構造をもつ一群の最適停止問題を系統的に取扱うことのできる確率的決定過程のある一般モデルを提唱するところにある．ここではつぎのような研究の方法論をとることにする．まずはじめに釣堀の問題(これを仮に問題Aとしよう)と称する典型的な最適停止問題を例にとり，この問題のもっている特徴的な性格について述べる．つぎに，この問題Aの自然な拡張によって得られる問題 A′, A″を定義し，これらの問題が共通してもっている3つの要因を明らかにする．そして最後にこれらの要因をより一般的なかたちで包含した一般モデルを構築する．この段階までの研究は［8］においてある程度達成されているが，本稿ではそこでのいくつかの不備な点を正す一方，逆にこの一般モデルを，問題 A, A′, A″ およびこれらと似た構造をもつ他のタイプの問題 B, B′,……を系統的に取扱うことができるようある方向に向けて特殊化することにする．これらの問題の中には，A, A′, A″ とは表面上かなり異なったものもあり，最適停止問題というイメージからかけ離れたものもあるが，3節でも述べるように，これらが最適停止問題の典型的な一例である釣堀の問題Aがもっているのと同じような行動学的・経済学的性格（それは他の決定問題には見られない特徴的な性格である）をその中にもっていることが示されるであろう．それゆえに，問題A, A′, A″, B, B′,……は，この一般モデルによって説明可能となる他の多くの問題とともに，あるひとつの明確な研究領域を構成することのできる問題群となり得るであろう．

2. 最適停止問題とその周辺

まずはじめに最適停止問題の最も典型的な例としてつぎのような釣堀の問題を取り上げ，その心理学的な意味を明らかにしよう．

釣堀の問題 I 「高い入場料を払って釣堀に入り，まさに最初の一投を振込もうとしている釣人を想定しよう．閉店時までに $N=1$ 匹の鱒を期待重量が最大になるよう釣ることが彼の目的である．ここで鱒の重量はある分布に従っているものとする．ただし釣堀の針の先にはそり返しがなく，針にかかった鱒は簡単な竿さばきで容易に逃がすことができるが，一度釣上げてしまったものの放流は禁止されている」

これは，針にかかった鱒を釣上げないという行為をいつ停止するか，という意味で最適停止問題であると言える．この釣人が合理的な精神の持主であれば彼はおそらくつぎのような作戦を心ひそかに立てるであろう．“入場したばかりの今は時間も十分にあるから，万一相当の大物がかかったならそれを釣上げ早々に引上げるが，それほどの大物でなければ逃してつぎの鱒のかかるのを待とう．しかし大物だけをねらっていると（大物はそれほど沢山いるわけはないから）目標の1匹を釣上げることなく時間が経ち閉店時がどんどん近づいてくるであろう．最悪の場合，手ぶらで帰るなどという不幸なことにならないようまだ時間のあるうちにそれほどの大物でなくてもある程度以上大きければ釣上げてしまおう．それでもなおかつ目標の1匹を釣上げることなく遂に閉店直前まできてしまったら，もう大物・小物などとは言ってはおられないから針にかかったら何でも釣上げてしまおう”

ここで鱒が針にかかるのを就職話がやってくること(R & D プロジェクトのアイデアが出てくること)，その大きさをその就職先の条件の良さ，たとえば給与額（そのアイデアが商品化されたときに期待される総限界利益)，閉店時間を卒業時点（会計年度のおわりの時点）に置きかえてみると，この釣堀の問題が単なる魚とりのお遊びである以上に，意思決定問題における重要ないくつかの側面を含んだ興味ある問題であることが理解できよう．

さて，この釣堀の問題において，閉店までの残余の時間が十分にあるときは大物だけを，それが

少なくなるにつれて小物でも，という釣人の心理は，将来針にかかるであろう鱒の大きさに対する期待が残余の時間が多くあるほど大であるというところからくると言ってよいであろう．

鱒が針にかかるということをチャンスに恵まれる，大きい鱒を大きいチャンス，小さい鱒を小さいチャンス，というようにその表現を日常的な言葉に置きかえてみるとこの釣人の心の動きはまたつぎのようにも言い換えることができる．“時間が十分にあれば将来より大きいチャンスに恵まれる可能性は大であるから現在のチャンスがその可能性以上のものでなければ採択 (accept) しない．逆に時間があまりなければ将来に対しあまり大きいチャンスは期待できないから目前のチャンスがそれほど大きくなくても採択しよう” すなわち“現在の決定に対する態度や姿勢は将来に期待されるチャンスの大きさとの均衡点として定まる”と言うことである．このことは決定問題一般に対しても言いえることであるが，最適停止問題のひとつの際立った特徴は，“将来に期待されるチャンスの大きさが時間の経過とともに，すなわち残余の時間の減少とともに減少する” という点にある．大きいチャンスしか採択しないという態度は，日常的な表現を借ると，強気，高姿勢，安売しない……，逆に小さいチャンスでも採択するという態度は弱気，低姿勢，安売する，……と言うこともできる．このような表現を用いるなら，最適停止問題はまた，“時の経過とともに決定態度が強気から弱気へと徐々に転じていく様をその基本的な性格としてもつ決定過程である”と言うこともできる．

ところでわれわれは，この釣堀の問題Iの自然な拡張としてつぎのような問題を得ることができる．

釣堀の問題 II　「限られた時間内に，M 匹分の餌で $N=1$ 匹を期待重量最大となるように釣ること，ただし鱒がかかるたびに1匹分の餌が失なわれるものとする」

釣堀の問題 III　「限られた時間内に，M 匹分の餌で N 匹 $(N>1)$ を総期待重量最大になるよう釣ること」

捕鯨の問題　「ある一定の期間に，限られた予算，限られた燃料，限られた食料，限られた冷凍スペース，……等で，限られた種類の鯨をそれぞれ限られた頭数内で，総捕獲重量あるいは捕獲された鯨をいろいろな製品にし販売することによって得られる総期待利益が最大になるよう捕獲すること（実際の捕鯨がこのような問題意識にもとづいて行なわれてきたか否かは別にしても，捕鯨制限のきびしい現在，少なくともこのような感覚をもって操業されることが望ましいであろう）」

この3つの問題に対しても釣堀の問題Iにおけるのと同じ解釈，すなわち“残余の時間が大のときは大物だけを，小のときは小物でも”という解釈が成立することは容易に理解できよう．ただしこれらの問題においては，リソースの量やそれまでに採択されたチャンスの回数もその決定態度に影響を与えている，という点に注意しよう．

釣堀問題の一般化　「有限期間内に，s 種のリソース $M_1, M_2, \cdots, M_s$ で，それぞれいろいろな価値をもって逐次提示されるチャンスの中から，種類1，2，…，k のチャンスをそれぞれ最大 $N_1, N_2, \cdots, N_k$ 回までを選択して採択し総期待価値を最大にすること．

ただしこれらのリソースは時間の経過とともに，また選択決定のたびごとに減少するものとする」

ところで以上述べてきた問題はそれぞれつぎの3点において共通した構造をもっている．

a.　採択すべきか否かの選択対象がある．ここで選択対象は，釣堀の問題では鱒であり，捕鯨の問題では鯨である．この選択対象を採択する（釣上げる，捕獲する）という行動を A^1，採択しない（逃がす，捕獲しない）という行動を A^0 とし，行動空間を $\boldsymbol{A}=\{A^r : r=0, 1\}$ であらわそう．

b.　行動 $A^r (r=0, 1)$ をとると利得 x^r が得られる．ここで x^r はある cdf F に従う確率変数 θ

の関数である. 捕鯨問題ではθは鯨の重量であり, $x^0(\theta)=0$, $x^1(\theta)=p\theta-c$ (pは鯨をいろいろな製品にして販売する場合の単位重量当りの平均価格, cは銛を1発打つための費用および鯨の解体処理費用などである).

c. 系の状態推移の規則はとられる行動に依存する. たとえば釣堀問題IIIでは, 系の状態は残余の餌の量mとそれまでの釣果nからなるベクトル$i=(m, n)$で与えられる. このとき状態は, 行動A^0(逃がす)をとると$j=(m-1,n)$に, 行動A^1(釣上げる)をとると$j=(m-1,n+1)$に移る. このことを推移確率として表わすと $p_{ij}{}^0=\delta_{i+(-1,0),j}$, $p_{ij}{}^1=\delta_{i+(-1,1),j}$ となる (δはクロネッカーのデルタ).

次節ではこの3点をより一般的なかたちで包含している一般モデルを構築するが, その前にこの3点を基本的に備えている他のタイプの問題を列挙しておこう.

狩人の問題[1] **I** [13] 「ある狩人が獲物を求めN発の弾をもって正午に山に入った. 夕方6時には山を下るものとする. 獲物は30分ごとに現われる. あまり腕のよくないこの狩人は現われた獲物に向け, どれか1発の当るのを期待しつつ盲打に何発かを打ちつづけるものとする. 1発の弾が獲物に当る確率$a(1>a>0)$は獲物とは独立であるとする. 獲物の価値θはあるcdf F(その期待値をEとする)に従う. 目的は6時までにとられる獲物の総期待価値を最大にすることである」

狩人の問題 II 「狩人の問題Iにおいて毎時点(30分ごとにとられている)に$m(\ll N)$頭の獲物が現われるものとする. ここでは, その狩人の腕は百発百中であり, その気になればm発の弾でm頭をとることもできるものとする. 獲物の価値は互いに独立であるとする」

購買問題 [14] 「時々刻々価格変動する一定量の鉄筋を基礎工事のはじまる前日までに期待購入価格最小となるよう買入すること」

売却問題 [12] 「時々刻々価格変動する手持の貴金属を, 半年後にせまった借金の返済にあてるべく期待売却価格が最大になるよう売却すること」

売買問題 I [4] [9] [15] 「一定期間内に, 時々刻々価格変動するある種の貴金属を, "買ってそして売る"という買・売サイクルをN回くりかえすことにより得られる総売買差益の期待値を最大にすること」

売買問題 II 「時々刻々価格変動しているある穀物を, 時々刻々変動する売注文と買注文に対し, どのような売買戦略を立てると, 長期にわたって得られる売買差益の期待値を最大にすることができるか」

投機的在庫問題 「毎日何個かずつ使用される, しかも時々刻々価格変動するある資材を, 1日当りの期待購入費用最小となるように購入すること」

確率的割当問題 [2] 「利益に対する貢献度の異なるn人の作業者に順次やってくるn件の仕事を割当てる. ただし1人に1仕事. 仕事の価値はあるcdf Fに従っている. ここで利益に対する貢献度は$p(1\geqq p\geqq 0)$で与えられ, 価値θの仕事を貢献度pの作業者に割当てると$p\theta$の利益が得られるものとする. 総期待利益が最大となるよう仕事を割当てること」

確率的割当問題 II [1] 「確率的割当問題Iにおいて, 仕事は継続的にやってきて, 仕事の完了したフリーの作業者にはその後にやってくる仕事を割当てることができるものとする」

受注選択問題 [5] [6] 「限られた生産能力の中で, 逐次やってくる注文の中から, 長期にわたって得られる限界利益の総和の期待値を最大にするよう, その採算性の高さから判断し受注すべき注文を選択していくこと (やってくる注文をすべて受注していると受注残をたえず能力いっぱいもつことになり, その後にやってくる採算性の高い注文も納期の都合上受注できないということになり, 結局長期にわたって得られる限界利益の総

和を減少させる，ということが起り得る)」

表面上，これらの問題はどれもが鈎堀の問題とはかなり異なっているが，実はどれもが鈎堀の問題における と同じような行動学的・経済学的性格をもち合わせている．たとえば割当問題Ⅰにおける最適戦略は，先にやってくる仕事は後にやってくる仕事よりも，より貢献度の低い作業者に割当てるような傾向をもつことが証明されている．これは後になるほど大きな価値の仕事（大きいチャンス）が現われる可能性が大であり，その仕事をより高い貢献度の作業者に割当てたほうが得策であろうという直感的な解釈からも理解できよう．また受注選択問題においてもその最適戦略は，受注残の多くあるときほど採算性の高い注文（大きいチャンス）だけを受注するという強気な態度に，逆に受注残の少ないときほど採算性の低い注文（小さいチャンス）でも受注するという弱気な態度になるような傾向をもつことが証明されている．このような強気・弱気という決定機構をもつ一群の問題が，次節で与えられる一般モデルによってどこまで説明され得るかは将来の研究に待つとして，これらの問題がひとつの明確な問題領域を構成することだけは確かである．

3. 逐次選択過程

ここでは，前節で述べた鈎堀タイプの問題の3つの特徴をより一般的なかたちで包含している確率的決定過程のひとつの一般モデルを構築する．

逐次選択モデル　有限・離散時間の確率的決定過程を考える．便宜上，時点 t は過程の終了時点を $t=0$ として逆向にとる．状態空間 $I=\{i\}$ は有限とし，各状態 i に対して有限な行動空間 $A_i=\{A_i^r : r=0, 1, \cdots, k_i\}(k_i\geqq 0)$ が対応している．各時点において m 次元ベクトル $\boldsymbol{\theta}=(\theta_1, \theta_2, \cdots, \theta_m)$ が cdf F に従って提示され，その後に行動 A_i^r（状態 i にいるとしよう）を選択すると利得 $x_i^r(\boldsymbol{\theta})$ が獲得され，次期の状態は確率 p_{ij}^r で j に推移する．目的は，状態 i の時点 t よりスタートしたとき，過程の終了時点までに得られる総期待利得を最大にすることである．この最大値を $v_i(t)$ としよう（もし必要ならば，このモデルは各状態 i に複数個の行動空間 A_{iw}, $w=1, 2, \cdots$ を対応させ，状態 i の各時点にそのうちのひとつが確率 p_i^w で提示される，というように拡張することもできるが，それはあまり本質的な議論とはならないのでここでは扱わないことにする）．なお，スペースの節約のため，今後 $v_i(t), v_i(t-1), x_i^r(\boldsymbol{\theta})$ をそれぞれ v_i, v_i', x_i^r と書き，$z_i^r=\sum_{j\in I} p_{ij}^r v_j'$ と定義する．z_i^r は状態 i の時点 t に行動 A_i^r をとったときつぎの時点 $t-1$ より過程の終了時点 $t=0$ までに得られる総期待利得の最大値である．

戦略　いま m 次元ベクトル空間 R^m を，$B^0\cup B^1\cup\cdots\cup B^{k_i}=R^m$ かつ任意の $r\neq s$ に対し $B^r\cap B^s=\phi$ となるよう k_i+1 個の部分集合に分割し，集合 $\boldsymbol{B}=\{B^r : r=0, 1, \cdots, k_i\}$ を分割とよび，その全体を $\mathscr{G}_i^m(k_i)=\{\boldsymbol{B}\}$ で表わす．このとき状態 i における戦略は，任意の分割 $\boldsymbol{B}$ に対し $\boldsymbol{\theta}\in B^r$ なら行動 A_i^r をとるというかたちで与えることができる．

基本方程式と最適分割　戦略を上のように与えるとき，v_i は最適性の原理より次式で与えられる．

$$v_i= \max_{\boldsymbol{B}}\sum_{r=0}^{k_i}\int_{B^r}(x_i^r+z_i^r)dF \qquad (1)$$

(1)式の右辺の最大は[11]における定理1より分割 $\boldsymbol{C}_i=\{C_i^r : r=0, 1, \cdots, k_i\}$ で与えられる．ただし，

$$C_i^r=\{\boldsymbol{\theta} : x_i^r-x_i^s\geqq z_i^s-z_i^r \text{ for } 0\leqq s<r,\quad x_i^r-x_i^s>z_i^s-z_i^r \text{ for } r<s\leqq k_i\} \quad r=0, 1, \cdots, k_i \qquad (2)$$

よって(1)式はつぎのように書かれる．

$$v_i=\sum_{r=0}^{k_i}\int_{C_i^r}(x_i^r+z_i^r)dF \qquad (3)$$

今後，分割 $\boldsymbol{C}_i$ を最適分割，(1)と(2)式を基本方程式とよぶことにする．

ところで上記のような最適分割および基本方程式はあまりにも一般的すぎ，このままではこれ以

上の数理的解析はほとんど不可能である．しかしながら幸いなことに，前節で述べてきた問題のほとんどは，その利得関数 x_i^r がつぎのような比較的単純な２つのケースのいずれかで与えることができる．

Case I　$x_i^r=a_i^r\theta$, $r=0, 1, \cdots, k_i(\boldsymbol{\theta}=(\theta)$ すなわち1次元ベクトル)．ここで $a_i^r\uparrow$ in r あるいは $a_i^r\downarrow$ in r とする[2]．θ の cdf を F, その期待値を E とし，$T(g)=\int_g^\infty(\theta-g)\,dF$ と定義する．さらに $0<r\leqslant k_i$ に対して $\Delta z_i^r=z_i^{r-1}-z_i^r$, $\Delta a_i^r=a_i^r-a_i^{r-1}$, $c_i^r=\Delta z_i^r/\Delta a_i^r$, とする．

Case II　$x_i^0=0$, $x_i^r=\theta_1+\theta_2+\cdots+\theta_r$, $r=1, 2, \cdots, k_i\leqslant m$ とする．ただし $\theta_r\downarrow$ とする．$\theta=\theta_r$ の cdf を F^r とし，$T^r(g)=\int_g^\infty(\theta-g)dF^r$ と定義する．さらに，$0<r\leqslant k_i(\leqslant m)$ に対し $c_i^r=z_i^{r-1}-z_i^r$, $c_i^0=-\infty$, $c_i^{k_i+1}=\infty$ とする．

この２つのケースに対する最適分割と基本方程式はそれぞれつぎの定理で与えられる[11]．

定理 1 (Case I)[3]　$a_i^r\uparrow(\downarrow)^*$ in r のとき $c_i^r\uparrow(\downarrow)$ in r なら，最適分割は区間 $C_i^r=(c_i^r, c_i^{r+1}]$ $([c_i^{r+1}, c_i^r))^*$, $r=0, 1, \cdots, k_i$ で与えられ，基本方程式は

$$\left.\begin{aligned} v_i&=a_i^0E+z_i^0+\sum_{r=1}^{k_i}\Delta a_i^r T(c_i^r) \\ (v_i&=a_i^{k_i}E+z_i^{k_i}-\sum_{r=1}^{k_i}\Delta a_i^r T(c_i^r))^* \end{aligned}\right\} \tag{4}$$

ただし $c_i^0=-\infty(\infty)^*$, $c_i^{k_i+1}=\infty(-\infty)^*$ とする．

定理 2 (Case II)　$c_i^r\uparrow$ in r なら最適分割は $C_i^r=\{\boldsymbol{\theta}: c_i^r<\theta_r,\ \theta_{r+1}\leqslant c_i^{r+1}\}$, $r=0, 1, \cdots, k_i$ で与えられ，基本方程式は，

$$v_i=z_i^0+\sum_{r=1}^{k_i}T^r(c_i^r) \tag{5}$$

となる．

この２つの定理をある与えられた問題に適用するためには，すべての i に対し $c_i^r\uparrow$ in r あるいは $c_i^r\downarrow$ in r が成立していなければならない．もし可能ならば，それが成立するような必要十分条件を求めるということは今後の興味ある研究課題のひとつとなるであろう．

つぎにこの２つの定理の具体的な問題への適用の仕方について，狩人の問題 I，II を用いて説明しよう．

定理 1 の狩人の問題 I への応用

この問題は逐次選択モデルとしてつぎのように説明される．

- 状態空間 $I=\{i: i=0, 1, \cdots, N\}$ i は手持の弾数
- 行動空間 $\boldsymbol{A}_i=\{A_i^r: r=0, 1, \cdots, k_i\}$ A_i^r は状態 i のときに見つけた獲物に r 発を連射するという行動．明らかに $k_i=i$.
- 利得関数 $x_i^r=(1-(1-a)^r)\theta$, すなわち $a_i^r=1-(1-a)^r$.
- 推移確率 $p_{ij}^r=\delta_{i-r,j}$

時点 t は6時を $t=0$ として30分ごとに逆向にとる．ここで $v_i(0)=(1-(1-a)^i)E$ であることおよび $a_i^r\uparrow$ in r であることに注意．いま $\Delta v_i=v_i-v_{i-1}$, $\Delta^2v_i=\Delta v_i-\Delta v_{i-1}$, $h(i)=a(1-a)^{i-1}$ と定義しよう．このとき $z_i^r=v_{i-r}'$ であるから $c_i^r=\Delta v_{i-r+1}'/h(r)$ となる．ここで $h(r)\downarrow$ であることおよび $\Delta a_i^r=h(r)$ であることに注意．$t=1$ のとき $\Delta v_i'=h(i)E$ であるから明らかに $\Delta v_i'\geqslant 0$ かつ $\downarrow$ である．よって $c_i^r\geqslant 0$, $\uparrow$ in r, かつ $\downarrow$ in i となる．

いま任意の t に対して $\Delta v_i'\geqslant 0$ かつ $\downarrow$ とすると $c_i^r\geqslant 0$, $\uparrow$ in r, かつ $\downarrow$ in i となるから，定理 1 より基本方程式

$$v_i=v_i'+\sum_{r=1}^{i}h(r)T(c_i^r) \tag{6}$$

を得る．これより，

$$\begin{aligned}\Delta v_i&=\Delta v_i'+h(i)T(c_i^i) \\ &\quad+\sum_{r=1}^{i-1}h(r)\{T(c_i^r)-T(c_{i-1}^r)\} \end{aligned} \tag{7}$$

$$\begin{aligned} &=h(1)\tilde{T}(\Delta v_i'/h(1)) \\ &\quad+\sum_{r=1}^{i-1}T_{h(r+1),h(r)}(\Delta v_{i-r}') \end{aligned} \tag{8}$$

$$\begin{aligned}\Delta^2v_i&=h(1)\{\tilde{T}(\Delta v_i'/h(1))-\tilde{T}(\Delta v_{i-1}'/h(1))\} \\ &\quad+\sum_{r=1}^{i-2}\{T_{h(r+1),h(r)}(\Delta v_{i-r}') \\ &\quad-T_{h(r+1),h(r)}(\Delta v_{i-r-1}')\} \\ &\quad+T_{h(i),h(i-1)}(\Delta v_1') \end{aligned} \tag{9}$$

を得る((9)式は(8)式より得られる)．ここで $\tilde{T}(g)$

$=g+T(g)$, $T_{ab}(g)=aT(g/a)-bT(g/b)$ である．性質 $T(g)\geqq 0$, $T(g)\downarrow$, $\tilde{T}(g)\uparrow$, $0<a\leqq b$ のとき $[0, \infty)$ 上において $T_{ab}(g)\uparrow \text{in } g$ かつ $\leqq 0$, を(7)と(9)に適用することにより $\Delta v_i\geqq 0$ かつ $\Delta^2 v_i\leqq 0$ すなわち $\Delta v_i\downarrow$ であることが容易に証明される．よって $c_i{}^r(t+1)\geqq 0$, $\uparrow \text{in } r$, かつ $\downarrow \text{in } i$ が得られる．かくて帰納法によりすべての t に対して $c_i{}^r\geqq 0$, $\uparrow \text{in } r$, かつ $\downarrow \text{in } i$ となる．よって定理1より，すべての t に対して基本方程式は(6)で与えられ，最適分割は区間 $C_i{}^r=(c_i{}^r, c_i{}^{r+1}]$, $r=0, 1, \cdots, k_i$, で与えられる．すなわち手持の弾数が i のとき，現われた獲物の価値 θ が $c_i{}^r<\theta\leqq c_i{}^{r+1}$ なら r 発を連射せよ，ということになる．ここで $c_i{}^r\downarrow \text{in } i$ は手持の弾数が多いときほど沢山連射せよ，ということを示している．さらに(7)式より $\Delta v_i\geqq \Delta v_i'$ を得る．すなわち $\Delta v_i\uparrow \text{in } t$, よって $c_i{}^r\uparrow \text{in } t$ である．このことは，残余の時間の多いときほど同時に打つ弾数を少なくすることを意味している．これは将来現われるであろうより大きな獲物(大きいチャンス)を得るために弾数を将来に残しておくということを意味している．

定理2の狩人の問題IIへの応用

この問題は逐次選択モデルとしてつぎのように説明される．

- 状態空間I　狩人の問題Iと同じ
- 行動空間 $A_i=\{A_i{}^r : r=0, 1, \cdots, k_i\}$, $A_i{}^r$ はあらわれた m 頭の獲物を価値の大きい順に並べたとき上位から r 頭をとるという行動．明らかに $k_i=\min\{m, i\}$.
- 利得関数 $x_i{}^0=0$, $x_i{}^r=\theta_1+\theta_2+\cdots+\theta_r$, $r=1, 2, \cdots, k_i$. θ_r は r 番目の獲物の価値であり，その cdf F^r は順序統計量のよく知られた公式

$$F^r(\theta)=\sum_{k=0}^{r-1}\binom{m}{k}F(\theta)^{m-k}(1-F(\theta))^k \tag{10}$$

で与えられる．ここで $F(\theta)$ は獲物の価値の cdf である(その期待値を E とする)．ここで $v_i(0)=k_iE$, $c_i{}^r=\Delta v_{i-r+1}'$ であることに注意．前問と同様，差分方程式 $\Delta v_i=\cdots$ と $\Delta^2 v_i=\cdots$ を求め，これに対して性質 $T^r(g)\geqq 0$ かつ $\downarrow \text{in } g$, $\tilde{T}^r(g)\uparrow \text{in } g$, $T^{r-1}(g)-T^r(g)\downarrow \text{in } g$ かつ $\geqq 0$ などを適用することにより(数学的帰納法を用いて)容易に $c_i{}^r\geqq 0$, $\uparrow \text{in } r$, $\downarrow \text{in } i$, かつ $\uparrow \text{in } t$ であることを証明することができる．よって最適戦略は，$c_i{}^r<\theta_r$ かつ $\theta_{r+1}\leqq c_i{}^{r+1}$ なら上位 r 頭をとれ，ということになる．ここで $c_i{}^r\downarrow \text{in } i$ は手持の弾数が多いときほど上位から沢山とり，少ないときほど上位からわずかをとることを，また $c_i{}^r\uparrow \text{in } t$ は残余の時間の多いほど上位からわずかをとり，少ないほど上位から沢山とる，ということを意味している．

4. 今後の研究課題

1．定理1，2に対する2つの適用例からも，この定理をある与えられた問題に適用していくとき，関数 T のいろいろな性質を“巧妙”に組合せて使っていくということが重要なポイントであることが理解できよう．しかしこれは“理論の簡潔さ”という点からすれば多少不自然なことである．これに対する著者の見解は，関数 T の性質をより深く研究するとともに，問題そのものをもっと高い立場からとらえ(たとえば利得関数を $x_i{}^r=a_{i1}{}^r\theta_1+a_{i2}{}^r\theta_2+\cdots+a_{im}{}^r\theta_m+a_{i0}{}^r$ のように一般化するなどして)，そこから系統的な解析手続の可能性を検討し，この“巧妙さ”をルールにまでもっていくようにすべきである，というところにある．

2．Case I は売却問題，購買問題，売買問題などによく使われる．そこで Case I を，θ がマルコフ性をもつ場合[16]，マルチンゲールの場合，傾向変動・周期変動する場合……等についてさらに深く研究するのも今後の興味ある課題となろう．

3．最適停止問題には，本稿では扱わなかったが結婚の問題［3］，や破産回避の問題［7］のように，目的関数が確率で与えられるようなものも多くある．このような問題に対しても，もし可能

ならばその一般化を計るということも今後の研究課題となろう.

4. (3)式はつぎのようにに書きかえることができる. $v_i=R_i(C_i)+\sum_{j\in I}Q_{ij}(C_i)v_j'$ (11)

ここで,

$$R(C_i)=\sum_{r=0}^{k_i}\int_{C_i{}^r}x_i{}^r dF \quad (12)$$

$$Q_{ij}(C_i)=\sum_{r=0}^{k_i}P_{ij}{}^r\int_{C_i{}^r}dF \quad (13)$$

(11)式はこの過程が,構造をもったマルコフ型決定過程であることを示している.このことは無限計画期間の逐次選択過程(たとえば確率的割当問題IIや受注選択問題)における解の存在と一意性の証明に関し,マルコフ型決定過程の分野で研究されてきた多くのことがそっくり適用できることを意味している.

注1 これは[13]の戦争モデルを平和的な問題に書き直したものである.

注2 本稿では,ある数列 f_n が n に関して非減少なら $f_n\uparrow$ in n と書き,それが強い意味で増加なら $f_n\Uparrow$ in n と書く.混乱のおそれがなければこれを単に $f_n\uparrow$, $f_n\Uparrow$ と書くこともある.非増加,減少の場合にも同様な記号↓,⇓を用いる

注3 文—A(A′)*—B(B′)*—……—C(C′)*— は,2つの文—A—B—…—C—と—A′—B′—…—C′—を表わす.

参考文献

[1] Albright, S. C.: A Markov-decision-chain Approach to a Stochastic Assignment Problem. *Oper. Res.* Vol. 22, No. 1 (1974), 61-64.

[2] Derman, C., Lieberman, G. J. and Ross, S. M.: A Sequential Stochastic Assignment Problem. *Manage. Sci.* Vol. 18, No. 7 (1972), 349-355.

[3] Gilbert, J. P. and Mosteller, F.: Recognizing the Maximum of a Sequence, *J. Amer. Stat. Assoc.* Vol. 16 (1966), 35-73.

[4] Haggstorom, G. H.: Optimal Sequential Procedures when more than one stop is required. *Ann. Math. Stat.*, Vol. 38 (1967), 1618-1626.

[5] 生田誠三:受注選択過程の基礎理論. *JIMA* (日本工業経営学会誌), Vol. 46 (1971), 17-26.

[6] ——:最適受注選択問題の基礎的研究. 学位論文, 慶応義塾大学工学研究科(1975).

[7] ——:生存問題—財産処分によって破産を回避する問題—. 日本OR学会研究発表会アブストラクト集(1976年9月)67-68.

[8] ——:逐次選択過程の理論構成とその応用. オペレーションズ・リサーチ(1977年3月号)164-173.

[9] ——:Structure of the Decision Rules in Optimal Buying-Selling Problem. 日本OR学会研究発表会アブストラクト集(1978年10月), 110-111.

[10] ——:A New Approach to Sequential Stochastic Assignment Problem without using Hardy's Theorem. 日本OR学会研究発表会アブストラクト集(1978年10月), 112-113.

[11] ——:Discrete Time Sequential Selection Process with Linear Reward Functions of Random Variable, 日本OR学会研究発表会アブストラクト集(1979年3月), 165-166.

[12] Karlin, S.: *Studies in Applied Probability and Management Science*. Stanford University Press (1962), 148-158.

[13] Mastran, D. V. and Thomas, C. J.: Decision Rules for Attacking Targets of Opportunity. *Nav. Res. Logist. Q.*, Vol. 20 (1973), 661-672.

[14] Morris, W. T.: Some Analysis of Purchasing Policy, *Manage. Sci.* Vol. 5, No. 4 (1959), 443-452.

[15] Sakaguchi, M.: An Investment Problem: an Optimal Stopping Problem in which two stops are requied, *J. Oper. Res. Soc. Jap.*, Vol. 15, No. 1 (1972), 45-52.

[16] Taylor, H. M.: Evaluating a Call Option and Optimal Timing Strategy in the Stock Market, *Manage. Sci.*, Vol. 14, No. 1 (1967), 111-121.

(いくた・せいぞう 筑波大学社会工学系)