1996年度日本オペレーションズ・リサーチ学会
春季研究発表会

終端利益の存在する逐次配分問題

02201810　筑波大学　佐藤雅宏　SATO Masahiro

1. はじめに

確率的な価値を持って逐次に出現するターゲットに対し手持ちの弾を何発配分するかという問題, 即ち逐次配分問題についてはこれまで Masatran and Thomas [2], Sakaguchi [4] や Namekata et al. [3] など多くの論文が発表されてきた. また, Derman et al. [1] のように逐次配分問題を経済的投資問題としてモデル化している論文もある.

しかしながらこれまでの研究では, 最終期限に弾等の資源が余ってもこれらを無価値のものとしていた. また Sato [3,4] を除き, 最終期限より前に資源を使い果たしてしまった場合, 決定者は残り計画期間を無為に過ごさなければならないという仮定がなされていた. ところが, 残り資源を別用途に転用あるいは売却できる, または残り計画期間に別の作戦・仕事等を行うことが可能と考えれば, これらを無価値とすることは現実的でない.

本研究では, 残り資源及び計画期間に対しそれらの量に応じた価値を与えてモデル化, 考察を行った. このモデルでは最適政策はある特異な形になることがある. しかしながら, その特異性は我々の直観に反するものでなく, 人間行動のより複雑な側面を表しているのである.

2. モデルの説明及び最適方程式

本研究では離散資源量の有限離散時間逐次配分問題を扱う. 今, 決定者（ハンター）に計画期間 t 期間と弾 i 発が与えられている。尚, 最終時点を 0 として時点を後向きに, 即ち時点が残り計画期間をも表すようにとる.

彼は毎期の初めにターゲットを一つ発見, その価値 $w \in [0,1]$ を直ちに判断, 撃つかどうかを決定する. ターゲットの価値は各期独立であり, 分布関数が $F(w)$ である同じ分布からのランダムサンプルである. 撃つと決めた場合, 確率 $q \in (0,1]$ で命中する弾を 1 発撃つ. 弾が当たればそれを得, その期に w 円で売却する. 外れた場合, 確率 $r \in [0,1]$ でターゲットを見失う. ターゲットを見失っていない場合, 決定者は更にもう 1 発撃つか否かを決める (shoot-look-shoot 戦略). 彼はターゲットを得るか, 見失うか或は弾を使い果たすまで何発でも撃つことができる. 但しターゲットの部分的ダメージは無視, 同一ターゲットに何発撃っても時間の推移はないと仮定する.

その期の狩りの終了後, 彼は次の期以降も狩猟活動を続けるか否かという第二の決定を行う. 今後の狩りをやめる場合, 残りの弾数と計画期間に応じた終端利益 $R_t(i)$ を受け取る. 終端利益は t と i に関し増加, かつ i に関し凹である. 更に狩りを続ける場合は, 次の期に進み新たなターゲットを発見する. このように途中でやめると決定するか, 又は $t=0$ に達するまでハンターは狩猟活動を続ける. 彼は, ターゲットの売却益と終端利益からなる総期待利益を最大化するよう行動する. 一期あたりの割引率を $\beta \in (0,1]$ とする.

時点 t において i 発弾を所持, かつ価値 w のターゲットに出会っているとき, 以降最適に行動した場合の総期待利益を $u_t(i,w)$, その w に関する期待値を $v_t(i)$, 時点 t において i 発弾を所持, かつ現在のターゲットにこれ以上弾を消費しない場合の最大総期待利益を $z_t(i)$ と定義すると, このモデルは以下の最適方程式で表される.

$$
\begin{aligned}
u_t(i,w) &= \max\{z_t(i),\ pu_t(i-1,w)+qw \\
&\qquad +(1-p)z_t(i-1)\},\ t \geq 0, i \geq 1, \\
u_t(0,w) &= v_t(0) = z_t(0) = R_t(0),\ t \geq 0, \\
v_t(i) &= \textstyle\int u_t(i,\xi)dF(\xi),\ t \geq 0, i \geq 0, \\
z_t(i) &= \max\{R_t(i), \beta v_{t-1}(i)\},\ t \geq 1, i \geq 0, \\
z_0(i) &= R_0(i),\ i \geq 0.
\end{aligned}
$$

但し, $p=(1-q)(1-r)$, $R_0(0)=0$ とする.

3. 最適政策

隣り合う2点 $(i, v_t(i))$ 及び $(i+1, v_t(i+1))$, $i = 0, 1, 2, \ldots$ を, 順に直線で結ぶことにより作られる連続関数を $\bar{v}_t(\bar{\imath})$ とする. ここで $\bar{\imath}$ は非負実数. 同様に $R_t(i)$ を用いて $\bar{R}_t(\bar{\imath})$ を作る. 更に, 関数 $g_t(i, w)$, $\bar{\psi}_t(\bar{\imath})$ を以下のように定義する.

$$g_t(i, w) = pu_t(i-1, w) + qw + (1-p)z_t(i-1) - z_t(i),\ t \geq 0, i \geq 1,$$
$$\bar{\psi}_t(\bar{\imath}) = \beta \bar{v}_{t-1}(\bar{\imath}) - \bar{R}_t(\bar{\imath}),\ t \geq 1, \bar{\imath} \geq 0.$$

これらを用いて, 最適政策は以下のように表される.

(a) もし $g_t(i, w) \geq 0$ なら弾を撃て, さもなくば撃つな.

(b) もし $\bar{\psi}_t(i) \geq 0$ なら狩りを続けよ, さもなくば狩猟活動をやめよ.

方程式 $g_t(i, w) = 0$ を満たすような $w = h_t(i)$ を臨界値と呼ぶ. 臨界値は与えられた t と i に対し一意に定まることがわかっている. 以下 $R_t(i)$ の例として2通り考え, それぞれの例について考察する.

例 1: $R_t(i) = bi$

弾の売却単価を $b \in [0, q)$, 残り計画期間は無価値であるとした場合, 次の結論を得る.

定理 1.

(a) $\beta = 1$ ならば, $h_t(i)$ は i の減少かつ t の増加関数である. 又, 常に最後まで狩りを続けるのが最適となる.

(b) $\beta < 1$ のとき, $0 \leq b \leq \beta q$ ならば $i \leq \rho_1'$ の範囲で $h_t(i)$ は i の減少関数である. 但し, ρ_1' は方程式 $\bar{\psi}_1(\bar{\imath}) = 0$ の0でない解である。なお, $i \leq \rho_1'$ ならば最終時点まで狩りを続ける.

途中で狩りをやめるのが最適な場合には, 臨界値が残り弾数の減少関数とならないケースが存在する.

例 2: $R_t(i) = b_t i + k(t)$

時点 t での弾の売却単価を b_t, $k(t)$ を残り計画期間に対する報酬と考える. 但し, $k(t)$ は t の増加かつ $k(0) = 0$ である. このとき臨界値が残り弾数の減少関数とならないケースがあるばかりか, やめるか否かの最適政策が次のような形になる場合がある. 即ち, 「ある t に対し ρ_1', ρ_1'' $(1 < \rho_1', \rho_1' + 1 < \rho_1'')$ なる正数が存在し, $\rho_1' \leq i \leq \rho_1''$ ならば狩猟活動を続けよ, さもなくば停止せよ」という政策である (Figure 1 参照). これはある程度資源を持っている者には活動を続ける価値があり, ほとんど持たざる者および持ち過ぎている者には活動を続ける意味がないということを示唆している.

参考文献

[1] C. Derman et al.: A Stochastic Sequential Allocation Model, *Operations Research*, Vol.23, (1975), 1120-1130.

[2] D. V. Mastran and C. J. Thomas: Decision Rule for Attacking Targets of Opportunity, *Naval Research Logistics Quarterly*, Vol.20, (1973), 661-672.

[3] T. Namekata et al.: A Sequential Allocation Problem with two kinds of Targets, *J. Oper. Res. Soc. Japan*, Vol.22, (1979), 16-28.

[4] M. Sakaguchi: A Sequential Allocation Problem for Randomly Appearing Targets, *Math. Japonica*, Vol.21, (1976), 89-103.

[5] M. Sato: A Sequential Allocation Problem with Search Cost where the Shoot-Look-Shoot Policy is Employed, *J. Oper. Res. Soc. Japan*, to appear.

[6] M. Sato: A Stochastic Sequential Allocation Problem where the Resources can be Replenished, 日本OR学会1995年度秋季研究発表会アブストラクト集 (1995), 46-47.

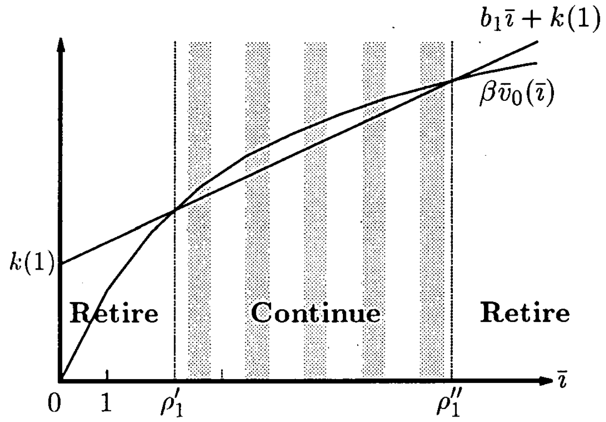

Figure 1 ρ_1' と ρ_1'' の存在の例