

## 『時間的一様でない盗賊問題における最適停止問題』

関西学院大学大学院 理工学研究科

数理科学専攻 千代延研究室 生野貴之

『最適停止問題』とはあるゲームを続けて行っているとき, どの時点でゲームをやめるとプレイヤーにとって一番『もうけ』が多いかを求める問題である. 数式(確率変数や期待値)を用いて表すと可積分な確率変数列  $(X_n)_{1 \leq n \leq N}$  (つまり  $\forall n \in \{1, 2, 3, \dots, N-1, N\}$  に対して  $E[|X_n|] < \infty$ ) が与えられたとする. このとき  $\mathcal{G}$  を停止時刻の全体集合とすると  $E[X_{\sigma_*}] = \max_{\sigma \in \mathcal{G}} E[X_\sigma]$  を満たす  $\sigma_* \in \mathcal{G}$  を求める問題を最適停止問題という. また, この  $\sigma_*$  を最適停止時刻 (**optional stopping time**) という.

最適停止時刻を定めるには *Snell Envelope* という概念が必要になる. *Snell Envelope* の定義を以下に述べる.

### 定義 (*Snell Envelope*)

ある可積分な確率変数列  $\{X_n\} (n = 1, 2, \dots, N)$  に対して  $\{Z_n\} (n = 1, 2, \dots, N)$  が  $\{X_n\}$  のスネル包 (*Snell Envelope*) であるとは

(a)  $\{Z_n\}$  が  $\mathcal{F}_n$ -super-martingale

(b)  $X_n \leq Z_n (1 \leq n \leq N)$

(c)  $\{Z_n\}$  は (a), (b) を満たす中で最小である. すなわち  $\{Z'_n\}$  も (a), (b) を満たせば

$Z_n \leq Z'_n (1 \leq n \leq N)$  が成り立つときにいう

本論文では *Snell Envelope* を用いて以下のゲームについての最適停止時刻について調べた

### ゲームの概要

- 仮定

1, ある家に価値が  $P_1$  の宝が入っている袋が  $N_1$  個, 価値が  $P_2$  の宝が入っている袋が  $N_2$  個あるとする (ただし  $P_1 < P_2$  であり  $N_1 \geq N_2, N_1 + N_2 = N$  であるとする)

2, その家の前には番犬が1匹いるとする

- ルール

1, Aさんは最大  $N$  回, その家から宝を盗むことが出来る. ただし1回毎に盗める宝の個数は1個とする.

2, 宝が袋に入っているので盗むときには, その宝の価値が  $P_1$  なのか  $P_2$  なのかは判別できない.

3, 番犬は盗賊が家に侵入する時は吠えないが盗賊が宝を盗んで家から出る時に吠える事ができる. しかし, いつ吠えるか分からない

#### 4, 吠えられたら今まで盗んだ宝は没収される (逮捕)

以上のルールの下で, 宝の価値の合計の期待値を大きくするには, どこで盗みを止めたら良いのか?

今回の発表では上記の問題に対して確率変数の設定や方針について具体的に定め, そこから *Snell Envelope* を定義し, それらを用いて最適停止時刻がどのように定まるのかについて述べる.

以下に  $N, P_1, P_2, N_1, N_2$  や犬の吠える確率を具体的な数値に当てはめたときの最適停止時刻を紹介する

$N = 7, P_1 = 100, P_2 = 300, N_1 = 5, N_2 = 2$ , 犬の吠える確率  $= \frac{3}{7}$  と定義したとき最適停止時刻に関して以下が成り立つ

##### 最適停止時刻

- 1 個 (1 回) でも『300 円の価値の宝』を盗めば, そこで stop
- スタートから 3 回連続『100 円の価値の宝』を盗んでも, そこで stop

本論文の後半では大きく分けて 2 つの事柄について調べた. 1 つは  $n = [Nt], \lim_{N \rightarrow \infty} \frac{N_1}{N} = \rho_1, \lim_{N \rightarrow \infty} \frac{N_2}{N} = \rho_2$  と定義したとき試行回数  $N$  を十分大きくしていくと前述で求めた最適停止時刻はどのような値に近づくかについて調べた.

$n = [Nt]$  と定義し各  $t$  毎に見ていくとこれまでに盗んだ宝の価値は右連続かつ左極限をもつ関数 (一般に *càdlàg* 関数と呼ばれる) であることが分かる. この関数は一般の距離空間では定義が難しいので新しい距離空間を設定する. それが *Skorohod  $J_1$  topology* と一般に呼ばれている. これを用いて定義の  $X_n$  に相当する確率変数や, その確率変数の *Snell Envelope* が  $N$  を十分大きくしていくときの収束先や, どのように収束するか (例, 概収束や確率収束等) を述べ, それらの事実から最適停止時刻が  $N$  を十分大きくしていくと, どんな確率変数に法則収束に収束するのか, また, その値についての考察を述べる. 結果を以下に紹介する.

##### 結果

$N$  を十分大きくしていくと最適停止時刻  $\sigma_*^N$  は

$$\sigma_*^N \sim \left\lfloor \frac{N}{\beta} \right\rfloor$$

を満たす. ただし, 犬の吠える確率  $= \frac{\beta}{N}$  とする ( $\beta > 1$ )

2 つ目は  $M_t^N \equiv \sqrt{N} \left( \frac{X_{[Nt]} - E[X_{[Nt]}]}{N} \right)$  と定義した (注,  $X_n$ :  $n$  回盗んだときの宝の価値の合計値とする) とき  $N$  を十分大きくしていくと  $M_t^N$  の分布はどのような分布に弱収束するのかについて調べた. この問題については途中段階であり, その経過, そして今後の課題について述べる. また期待値や共分散の結果から  $M_t^N$  の分布はどのような分布に弱収束するのかについての予想を述べる.