

Title	確率的逐次割り当て問題に関する研究
Author(s)	中井, 達
Citation	大阪大学, 1985, 博士論文
Version Type	
URL	<a href="https://hdl.handle.net/11094/34774">https://hdl.handle.net/11094/34774</a>
rights	
Note	著者からインターネット公開の許諾が得られていないため、論文の要旨のみを公開しています。全文のご利用をご希望の場合は、 <a href="https://www.library.osaka-u.ac.jp/thesis/#closed">〈a href="https://www.library.osaka-u.ac.jp/thesis/#closed"〉</a> 大阪大学の博士論文について <a href="https://www.library.osaka-u.ac.jp/thesis/#closed">〈/a〉</a> をご参照ください。

***Osaka University Knowledge Archive : OUKA***

<https://ir.library.osaka-u.ac.jp/>

Osaka University

氏名・(本籍)	なか 中	い 井	とおる 達
学位の種類	工	学	博 士
学位記番号	第	6687	号
学位授与の日付	昭和60年1月22日		
学位授与の要件	学位規則第5条第2項該当		
学位論文題目	確率的逐次割り当て問題に関する研究		
論文審査委員	(主査)		
	教 授	坂口	実
	(副査)		
	教 授	丘本 正	教授 竹之内 脩 教授 高木 修二
	教 授	西田 俊夫	

### 論 文 内 容 の 要 旨

確率的逐次割り当て問題は、統計的多段決定問題の1つであり、行動決定者の取り得る行動の数が、決定を行なうごとに1つつつ減少する問題である。この問題で、行動決定者は逐次に観測を行ない、その値が得られるごとに、取り得る行動のうちから1つを選択しその値に割り当て、総期待利得を最大にすることを目的とする。一度ある値に割り当てられた行動は、以降の決定において再び用いることは出来ないと考える。

この問題は、Hardyの定理と呼ばれる性質の確率論的一般化と考えることができる。また、この問題は、利得が得られた観測値の和であるような、複数回の停止が可能な、最適停止問題を含み、したがって、この最適停止問題に関する性質は、逐次割り当て問題の最適政策および、その解の構造を求めることにより得られる。ここでは、以下にあげる諸問題について考察し、これらの問題を動的計画法の考えを用いて定式化する。これらの問題における最適政策は、取り得る行動に付随した値が $p_1, \dots, p_n$  ( $p_1 \geq \dots \geq p_n \geq 0$ ) の $n$ 個のとき、それらの値の全ての組み合わせに対して同じ政策となる。

第2章では、各期ごとの観測値の従う分布がマルコフ連鎖に従うと仮定するが、その連鎖の状態は、直接に行動決定者には知らされない場合を解析する。この状態に関する情報は、確率分布により示され、各期ごとの観測値を得るたびに、ベイズの定理に従い学習されるものとする。最適政策、およびその政策のもとでの総期待利得の主な性質は2.2節で得られる。特に学習過程をベイズの定理を用いて記述するために、最適政策は critical number policy となるとは限らない。

行動決定者が逐次に1つつつ観測値を得る問題に対して、第3章では他の型の問題を考える。毎回、行動決定者により観測される、観測値の数が常に決まっているのではなく確率変数である場合である。

ここでは、2つの場合を扱う。またこの問題を複数回停止の可能な最適停止問題として考え、最適選択の問題として示し、最適政策および、その政策のもとでの総期待利得を得た。特に最適政策のもとで、観測値が選択されるには、大きい値を持つだけでなくその期に出現した値の中での順位が高くなければならないことが示される。

しかし観測値の全体の数は、大部分の問題では既知である。この数が予め知らされていない確率変数である場合を考えるのは自然である。第4章で観測値の実際の数について事前知識を持つ場合を考え、各観測値がポアソン過程に従って出現する場合を解析した。

この報告の終わりに、ゲーム論的接近を考える。特に2つの場合を考え、5.1節では拒否権のある2人0和ゲームを扱う。このゲームは最小化プレイヤーによる拒否権を持つ逐次割り当て問題であり、両プレイヤーに対する最適戦略は、このゲームに関連した逐次割り当て問題に対して得られた同じ critical number によって定まる。他の問題は、5.2節で扱われ、“Goofspiel”として知られるゲームの一般化となっている。前のゲームとの相異点は、観測値を得るごとに、各プレイヤーが各々取り得る行動から1つを選び割り当てることである。ここでは特別の場合を扱わず、最適戦略およびゲームの値は、相当複雑となることが示される。

#### 論文の審査結果の要旨

確率的逐次割り当て問題の研究は、Hardy - Littlewood - Polya の補題として、1930年代からよく知られた簡単な割り当て問題を、Derman - Lieberman - Ross (1972) が確率的逐次決定問題へ拡張したことから始まる。本論文は著者が過去5年間、この問題を逐次的、確率的のみならず、さらに連続的、適応的、競合的あるいは部分情報的モデルにもできる限りの種々の拡張を試みて得られた結果をまとめている。

第2章では部分的観察マルコフ (POM) 過程の上の最適割り当て-jobsが level  $i$  と size  $X^i$  とをもっていて、 $i$  は推移行列  $[P_{ij}]$  の単純マルコフ連鎖、 $X^i$  は cdf  $F^i(x)$  をもつ r.v. のとき、level  $i$  が観察できなくて、size  $X^i$  だけが観察できる場合を論じている。第4章では総数未知の jobs が1つずつ逐次に random に到着する。それははじめは頻々と、そのうちに間遠に到着する。time - deadline が与えられて、それまでに最適割り当てをしなければならない。第5章では“Goofspiel”とよばれるカードゲームにおいて、“center card”を jobs に対応させて考える2人0和ゲーム論的割り当て問題を扱っている。著者はこれらの諸問題においてそれぞれ最適解を見出しているが(定理2.2, 4.1, 命題5.2) これらはいずれも既往の結果を拡張する新知見である上に、この分野の今後の発展に新しい視野を提供するものであり、博士論文に値すると認められる。