

## 報告要旨

タイトル：Nonparametric Learning Rules from Bandit Experiments: The Eyes Have It!

Yingyao Hu, Yutaka Kayaba and Matthew Shum

発表者： カリフォルニア工科大学 博士候補生（経済学） 萱場 豊

JEL codes: D83, C91, C14

本研究は、不確実性下の学習の実験課題として典型的な2本腕バンディット（スロットマシン）選択問題における人の信念の更新過程（学習）を、ノンパラメトリックに推定している。従来のスロットマシンの選択のデータに加え、目の動きのデータを同時に推定に用いることで、信念の更新過程をより柔軟な仮定の下で推定することが可能となっている。近年、様々な行動経済学のモデルの検定のために、目の動きのデータ（広くは神経経済学的データ）が利用されてきているが、これら新しいタイプのデータをモデルの推定に直接利用しているのは、我々の知る限り本研究が初めてである。

本研究における実験課題をより正確に述べると、2本腕バンディット選択問題の拡張であり、認知神経科学等で学習の研究に標準的に用いられる **probabilistic reversal task** である。この課題では、2つのスロットマシンの内、必ず一方が報酬が出やすいスロットマシン、他方が損失が出やすいスロットマシンとなっており、それが時間を通じて確率的に入れ替わる仕組みとなっている。被験者は、どちらのスロットマシンが良いスロットマシンであるかを試行錯誤を通じて学習、すなわち、どちらのスロットマシンが良いのかについての信念を更新していく。

スロットマシンの選択のデータのみを用いた従来の研究では、信念の更新過程についてベイズ更新等の強い仮定を置かざるを得なかったが、こうした強い仮定から導かれた信念が、実際のスロットマシンの選択を上手く説明できていないことが頻繁に報告されている。本研究では、スロットマシン選択時の目の動きを測定し、それを信念の **noisy measure** として利用することで信念の識別性を強め、さらに産業組織論において開発された最新の計量経済学的手法を利用することにより、信念の更新過程をノンパラメトリックに推定することに成功している。

こうしてノンパラメトリックに推定された信念の更新をベイズ更新と比較したところ、良いスロットマシンと考えていたものを選択することから損失が発生した場合、信念の更新が遅く、元の信念に執着する傾向があることが明らかになった。また、他の代表的な信念の更新モデルである強化学習や“win-stay” heuristic も含めて利得を比較したところ、ノンパラメトリックな信念の更新は、最適な意思決定であるベイズ更新に25%ほど劣り、他の意思決定モデルとほぼ同等の利得を得ることが明らかになった。

（ 以 上 ）