

私の「研究」履歴書

研究者になった理由？そりゃ「食うため」ですよ！
自分のオリジナルの技で身を立って食っていく。
ほぼ職人ですね（笑）

■Interview 001

滋賀大学 データサイエンス学部教授
理化学研究所 革新知能統合研究センター
汎用基盤技術研究グループ 因果推論チーム
チームリーダー

清水 昌平 研究分野：因果探索



Introduction

ビッグデータが世の中にあふれ、データ分析を行う人材が求められている現在、同時に有用な知識を得るためのデータ解析法への関心と期待が高まっています。そのような中、清水昌平教授は、理化学研究所AIPセンター汎用基盤技術グループ因果推論チームリーダーとして、因果推論に関する研究を行っています。それでは、この研究がどのように社会の中で役立つのでしょうか。清水教授に聞いてみました。

■手に職をつける

Q. まず、清水先生が研究者を目指したきっかけについて教えてください。

小学校5年生の時「食うためには手に職を付ける必要がある！」と、漠然とそう考えていました（笑）。自分だけのアイデアを作って、みんながそれを使えるようになってくれたら、食えばぐれないぞ！ということですね。でも美術的なセンスがあるわけでもないし…じゃあ、研究者になろう、と。これは大学生になっても変わりませんでしたね。自分は大阪大学の人間科学部というところにいたんだけど、ここは哲学、教育学、心理学、社会学など、いろんな分野があってやることを選ぶんです。自分にとっては心理学や社会学は難しく、数学で統計モデルを作るほうが簡単のように思えた（笑）。数学的なことは、その枠組みの中で真偽を証明したりできますから。特に心理学では構造方程式モデリング（SEM）という手法がよく使われていて、因果にアプローチしていたんだけど、その頃のSEMは前提としている仮定がかなり強かったんです。因果の向きは事前にわかっていなくてはならない、とか、因果分析に必要な変数がすべて観測されている、などをユーザーが保証しなければならない、といったことです。「いやいや、それができるなら苦労しないんだよ。もう少しデータの力でどうにかならんのか」と強く思って、「ここを極めれば手に職がつく！」と、そう思ったんですね。そんなわけで、今の因果関係をデータから探索する研究に至っています。

■相関？ or 因果？

Q. 先生の研究テーマの「因果探索」ですが、どのようなことをしているのでしょうか。

データ分析では、相関関係と因果関係の区別が重要だとよく言われますよね。簡単に説明すると相関は2つの変数が共に変

化する関係のことで、因果は2つの変数の一方を変化させるともう一方が変化するという原因と結果の関係になっているものです。この因果関係をデータから推定するときは、潜在共通原因とよばれる未知の原因の影響を取り除く必要があります。これまでは、潜在共通原因を事前に特定してデータを集める必要がありました。しかし、すべての潜在共通原因を特定するのは困難で、とりこぼしがよく起きます。とりこぼしがあると、因果関係を適切に推定することができないんですね。なので、とりこぼしがあっても、妥当性を失わないような機械学習技術を作っています。

Q. どのような分野で使われていますか？

因果関係が明確に出来ると嬉しいことはたくさんあるので、本当に多様な領域で使われています。疫学、経済学、心理学、社会学、もちろんビジネスの分野でも使われていますね。スーダンで発生した紛争と商品価格の因果関係の探索といった研究でも使用されました。ただ、正直なところ、まだ方法論を開発したところで、応用研究は今後の課題です。皆さん、是非使って下さい（笑）！そのためにフリープログラムの公開といった情報公開や、機械学習のセミナーなど、手法を広める機会は今後も提供していきます。そうそう、去年発売した『統計的因果探索』（講談社）も是非手にとってもらえると嬉しいです。

■連携が大切

Q. これからデータサイエンティストになる方々に一言メッセージをお願いします。

もし統計分析をして、分析モデルに不満を感じる事があれば、是非モデルを開発している我々に不満を伝えて欲しいですね。統計モデルには万能なものはなく、様々な前提や制約のもとに成り立っていて、研究者は日々よりよいモデルを開発しています。そういう意味では、統計分析は一人で行うのではなく、得意な力を出し合って、連携しながら進めていくということが必要なのではないでしょうか。一緒に頑張りましょう！伊達さん！笑（聞き手 DS学部講師 伊達平和）

