

統計学の社会科学での意義とニーズ、今後

慶應義塾大学経済学部・理化学研究所 API センター 星野 崇宏

本セッションは統計学の若手研究者に対して今後の研究人生の参考になるような話題提供を行えとのことですが、当日ご登壇の他の先生方のお話や議論に刺激を受けて私自身もまだまだ自分の研究業績を積み重ねていきたいと考えております。

【これまでの研究者人生と影響を受けた人物・著作など】

私は大学入学時点では理系でしたが、心理学に関心を持ち学部後期は教育心理学コースに進学しましたが、そこで南風原朝和先生の心理学研究法の講義で「内的妥当性」や「外的妥当性」といった研究デザインの重要性を学び、以後ずっとこれらの話題が自分の中での大きなテーマの一つとしています。丁度修士課程で進学した先の繁榊算男先生が日本行動計量学会の大会のために招聘された Harvard 大学の Donald B. Rubin 教授の研究を勉強し、RCM や EM アルゴリズム、Missing at Random の概念などの研究は彼が在職していた Educational Testing Service（民営化された入試センターのような機関）で必要とされた「もし生物学選択者が物理学の試験を受けた場合」のテスト得点の等化という教育統計学の発想に端を発していること、選択バイアスや因果効果といった社会科学における非常に重要な概念が、統計学的には欠測データ解析という枠組みで統一的に議論できるということを知り因果効果や欠測を研究テーマにしました。但し、これとは別に心理学や行動経済学など意思決定メカニズムの数理モデルも引き続き研究をしており、これがマーケティング分野の研究ともつながっております。

【他分野との関連】

私の論文を読んだ企業の方からお声がけを頂き、博士課程 2 年のころからマーケティング・サイエンス分野の研究も始めましたが、実務での問題意識はある種明確であり、まだまだビッグデータなどと言われる前の時代から大量のデータを利用した分析を行うことができたのは幸運でした。その後因果効果や欠測の研究を通じて医学系の研究にも参加させて頂いていますが、統計学の本質的な部分の一つであるこういった基礎研究は非常に汎用性が高いので、いろいろな分野の方々と共同研究をさせて頂く機会に恵まれました。

【社会科学でのニーズの今後】

社会科学での統計学のニーズについては、メカニズムを踏まえた構造推定や行動経済学的な知見に基づく行動変容のための研究デザインなど様々あり、これについては当日紹介いたします。但し実は社会科学に限らず現在統計学が直面している問題は、いかに本来あるべき研究デザインや解析法を応用分野に示すか（誤った方法の問題点を示すか）にあると考えます。残念ながら国内の統計学やデータ分析の認識が薄い中で AI ブームが起きたことが、ビッグデータ時代の統計学の環境を諸外国と全く違ったものにさせているようです。

私の研究分野の因果効果や欠測なども含めこれまで統計学が長年かけて発展させてきた様々な考え方の枠組みや手法は現実社会で得られるデータ解析の場面でも非常に有用ですが、こういった概念を理解せずに、汎用的なパッケージを誤用した分析がさも最先端の方法であるかのように宣伝されることがあまりにも多いと感じます。残念ながらそのような解析から得られた知見では精度の高い効果測定や予測ができず、なにより説明ができない矛盾した結果を与えることになります。企業や応用研究者にデータ分析はあまり意思決定に役立たないと認識されてしまうと、AI 分野だけではなく統計学に対してもその有用性への疑義が生じます。皆さんには積極的に機械学習分野にも進出していただき、悪貨を駆逐していただくべきなのではないでしょうか？