



やりたいことは 何でもできる それがStataの 一番の魅力です

関西学院大学 総合政策学部
西立野 修平 准教授

長年Stataをご利用頂いている関西学院大学 総合政策学部 准教授の西立野先生に、Stataの魅力やStataを使ったデータ解析の授業を学部を導入される経緯などについてお伺いしました。

—Stataを使い始めたきっかけを教えてください。

博士号を取得するために、オーストラリア国立大学に留学した時のことです。Stataで統計理論を学ぶ計量経済学の授業があり、その授業がStataを使い始めるきっかけです。留学先ではほとんどの人がStataを使っていて、博士課程の学生はStataを使えるのが当たり前のような環境でした。

—どのようにしてStataを使いこなせるようになりましたか？

留学前に使っていたのは、Excelくらいでした。そのため、コマンド操作に慣れていないですし、授業が英語ですので、最初は訳も分からず使っている感じでした。最初はハードルが高かったのですが、ある水準に達すると、マニュアルやhelp機能を活用したり、自分で新しいコマンドを検索しそれを試したりして、いろいろ使えるようになりました。大学で周りの人に聞くことができる環境だったことも大きかったです。

計量経済学では理論を数式で学びますが、在学中に理論を実用的に使えるようになるかという、時間的に難しいと思います。しかし、Stataを使うと自動的に計算してくれるわけですね。その分理論的な部分を深く追求できないという側面もありますが、実用的に使えるようになるということでは、Stataはすごく有用でした。

大事なことは、どういった分析をしたいのかを自分で考え、まず、Stataでその計算をしてみる。そして、Stataで計算された結果がどういったプロセスで出てきたかについて学び理解を深めていくことだと思います。

adoファイルを使うことで、圧倒的に 分析の幅が広がります

—Stataでよく使う機能や便利だと思う機能はありますか？

よく使うコマンドは、mergeやappendです。どんな統計分析においてもそうですが、まずデータセットを構築し、それを基に分析していきます。色々なデータソースからデータを取り出しデータセットを作るのに、mergeコマンドを使います。最近は取り扱うデータ数が数十億というようなビックデータ分析をすることも増えているので、マージがないと研究を進められてないくらいです。

便利ということ言えば、「Stataではできない分析はない」ということを最近常々感じます。あることを分析したいと考えた時に、Stataにない機能でもadoファイルをインストールすれば新たな

コマンドを使うことができるようになります。adoファイルを使うようになると、圧倒的に分析の幅が広がりますね。

データ分析は、ますますStataに 依存しています

例えば、二地点間の距離を測り地理的な変数を作るため、GIS(Geographic Information System)という地理情報システムを使うことがあります。しかし、私がインストールしたプログラムには距離を測るサービスが含まれていませんでした。このオプションにはさらに費用が必要になることが分かり、困ったなと思ったのですが、もしかしたらStataでできるかもと思い調べてみました。そしたら、そういう機能がStataのadoファイルにあったんです。本当になんでもできるんだなと実感し、最近ますますStataに依存して行っています。やりたいことが何でもできるというのが、私にとっては一番の魅力ですね。

Check! adoファイルとは

adoファイルはStataユーザが作成したプログラムをWeb上で公開している追加機能です。多くのユーザがadoファイルを作成し、そのデータベースが更新されています。Stataからadoファイルをダウンロードし利用することができます。

—Stataで分析する上でお勧めの資料や書籍はありますか？

Stataを使い始めたころ、Stata Press(開発元 Stata LCC)から出版されているパネルデータ分析や生存分析の本などを何冊か読みました。しかし、Stataに付属しているマニュアルに必要なことは載っているので、わざわざ本を買わなくてもマニュアルだけでも十分だと今は思っています。

—Stataで改善点や要望点はありますか？

計算スピードも速いし、シンプルな操作で使いやすいですし、adoファイルを見れば新しいコマンドが作られています。便利になっていく一方という感じがしています。

挙げるとすれば、経済学分野でも機械学習などAIを使った分析がどんどん入っていますので、Stataでもできるようになればいいですね。

※Stataでは人工知能の一部としてStata16から機械学習を用いたモデル選択(Lassoベースの機械学習)が搭載され、2021年4月リリースされたStata17で機能強化されています。

Stataは、ビッグデータの分析に威力を発揮します

— どのような研究をされているのでしょうか？ また、研究の中でStataをどのように利用されていますか？

応用ミクロ計量経済学のデータ分析手法に基づいて、「交通」、「環境」、「貿易」分野における公共政策を定量的に評価することを専門にしています。

「交通」、「環境」分野においては、現在、首都圏に導入されたディーゼル車走行規制が、大気汚染や人々の健康に与えた効果について研究しています。分析の結果、ディーゼル車規制が、大気汚染の濃度を約7%低下させたこと、そして、その結果、規制対象地域の乳幼児の死亡率が低下したことが明らかになりました。大気汚染への政策効果を分析する際には、汚染濃度の観測所別時間値データを分析したので、観測数が膨大になりました。そこで、計算スピードを高めるために、aregというコマンドを使用しました。

「貿易」分野においては、貿易のための援助(Aid-for-Trade)が発展途上国の貿易に与えた効果について研究しています。推計手法に操作変数法を用いており、xtivregというコマンドを使っています。分析や検定の結果が、コンパクトに表示されるので、効率的に分析を進めることができ助かっています。

Stataをマスターするメリットを学生にも伝えたい

— 総合政策学部では、2022年4月からStataを使ったデータ解析の授業が始まるということですが、その経緯を教えてくださいませんか？

本学部では、2021年4月から新カリキュラムがスタートしました。目玉の一つがデータサイエンス教育の充実です。ビッグデータ、人工知能、機械学習などの重要性が高まる中、既存の統計教育にプログラミング(R, Python, Stata)を組み合わせることにより、より高度なデータ分析/データビジュアライゼーション/シミュレーションに基づく政策提言を行える人材を育成することが狙いです。Stataを用いたデータ分析は、特に経済学や医学の分野では、今やグローバルスタンダードになっています。本学部の教員にもStataユーザが多いです。Stataによるデータ解析を通して、一人でも多くの学生がデータ分析の面白さ、Stataの有用性を実感して欲しいと考えています。そして、授業で得たスキルをゼミでの研究に活かして欲しいです。

— Stataを用いてどのような授業を予定されていますか？

2年生以上を対象とした授業になります。Stataの基本的な使い方からグラフ(棒グラフ、折れ線グラフ、散布図など)の作成、単回帰分析、重回帰分析まで扱うことができると考えています。時間に余裕があればパネルデータ分析までやりたいですね。

— 先生のゼミ生が、ISFJ日本政策学生会議*の政策フォーラムで入賞されたとお聞きしました。

2年前から参加しているのですが、昨年、Stataを使って分析した研究発表で3年のゼミ生が入賞することができました。私のゼミには、15人の3年生がいて、3グループに分かれて参加しました。しかし、グループごとにテーマも違いますし、データや分析手法も違うので、指導も難しかったです。実際は指導というよりも、手取り足取りという表現が正しいですね。データをダウンロードし、データセットを作り、分析・推計し、分析結果をまとめて解釈するということを夏休みも含め時間をかけて行いました。

*ISFJ日本政策学生会議

学生への政策立案の支援を目的とした政策提言を発表する会議

— 学生の方は、ある程度Stataを習得できたのでしょうか？

時間的に習得までは難しいですね。一通りStataでやってみたという感じです。Stataに限りませんが、一定水準に達しないと習得するレベルまでいかないと。これもカリキュラム化する理由の一つです。ただ、これをきっかけに、興味を持ちさらに研究してみたいという学生もいて、今ではStataを自分で使って分析しています。

— 総合政策学部は、どのような学部でしょうか？

総合という名前のとおり、文理融合で色々な先生がいます。そのため、それぞれの先生によって定義が異なると思います。私の解釈で言うと、政策を評価することだと思います。国の政策もあるし、民間企業でいうとマーケティングや広告などいろんな政策があります。その政策の効果を計量的に分析することにより政策を評価し、よりよい政策を考えていくというのが弊学の総合政策学部の学びの本流であると思います。

— 本日はありがとうございました。

Check! adoファイルコマンド geodist

2地点間の距離を測るadoファイルコマンドgeodistを紹介します。次のように2地点間の緯度経度を示す変数からそれぞれの距離を新しい変数として作成できます。

・ geodist 緯度1 経度1 緯度2 経度2, gen(距離)

では、実際にアメリカ ミシガン大学にあるMichigan Stadiumに一番近い距離の公園を探してみましょう！

① はじめにadoファイルgeodistコマンドをStataにダウンロードします。

・ ssc install geodist

② 周辺に位置する 4つの公園のID、名前および緯度・経度データを入力します。

・ set obs 4

・ input long parkid str17 parkname double(lat lon)

・ 1 "Gallup Park" 42.27317 -83.694174

・ 2 "Argo Park" 42.291516 -83.744604

・ 3 "Hudson Mills" 42.382194 -83.911197

・ 4 "Nicholas Arboretum" 42.821123 -83.725575

③ Michigan Stadiumの緯度・経度と4つの公園の距離を測り、変数distという名前で保存します。

・ geodist 42.265837 -83.748696 lat lon, gen(dist)

④ 結果を表示させます。

・ list

	parkid	parkname	lat	lon	dist
1.	1	Gallup Park	42.27317	-83.694174	4.5712343
2.	2	Argo Park	42.291516	-83.744604	2.8722922
3.	3	Hudson Mills	42.382194	-83.911197	18.613799
4.	4	Nichols Arboretum	42.821123	-83.725575	2.5536404

*Nichols ArboretumがMichigan Stadiumに一番近いことが分かりました。

西立野先生のゼミ生がISFJ日本政策学生会議政策フォーラムで優秀政策提言賞を受賞しました。

詳しくは、下記URLをご覧ください。

https://www.kwansei.ac.jp/s_policy/news/detail/20201224_isfj_nishitaten



開発元

StataCorp LLC

正規国内代理店

LightStone®
株式会社 ライトストーン

25年以上の経験と実績でお客をサポートします。

〒101-0031 東京都千代田区東神田2-5-12 龍角舎ビル7F

TEL 03-3864-5211 FAX 03-3865-0050

E-mail: sales@lightstone.co.jp <https://www.lightstone.co.jp/stata/>

