

東京大学大学院法学政治学研究科における 統計分析・データサイエンス関連科目群について

2026年2月14日

大学院法学政治学研究科教授 福元健太郎

社会科学研究所准教授 勝又裕斗

東京大学大学院法学政治学研究科における統計分析・データサイエンスに関する授業科目を、2026年度から次のように体系化する。これらの授業を通じて、データを駆使して社会を読み解き、課題を発見・解決する実践力を養う。計量分析を用いた実証研究を行うおうとする大学院生はもとより、計量分析を行う予定がない大学院生にとっても、統計学やデータサイエンスが発展させてきた推論のロジックを理解することは有益である。計量分析に批判的、ないしは、懐疑的な立場からの参加も歓迎する。

授業科目

授業科目は、データサイエンスⅠ～Ⅳの4科目から構成される。ⅠおよびⅡでは、数式展開よりも概念的な理解を重視し、全体の核となる本質的なロジックを習得する。続くⅢおよびⅣでは、この基礎の上に発展的内容を積み上げる。いずれも1 Semester 週1回の2単位である。各科目の概要は、以下の通りである。

データサイエンスⅠ（入門） 数学：☆☆☆☆☆

この授業は、これまで統計分析を学んだことがない人を対象に、回帰分析という方法を中心に学ぶ。回帰分析とは、「ある要因が変わると、別の要因はどのように変わるのか」を調べるための考え方である。分析には、無料で使える R というソフトウェアを用いる。できるだけ多くの人に取り組めるよう、数学（特に確率論）の知識はほとんど使わない。加減乗除ができれば十分である。その一方で、社会の出来事をデータを使って調べるときに、「何が原因で、何が結果なのか」をどのように考えるべきかという点を重視する。また、実際にデータを扱い、分析する力を身につけることにも重点を置く。担当教員の専門が政治学であるため、授業で扱う例やデータは政治に関するものが多い。

但し、ここで学ぶデータ分析の考え方や方法は、政治に限らず、さまざまな分野で応用できる。

データサイエンス II（課題の発見と解決） 数学：★☆☆☆☆

紛争、貧困、移民問題、教育格差、ジェンダー不平等、気候変動。わたしたちの社会は、このような無数の複雑な課題を抱えている。この授業では、こうした現実世界における課題を発見し、解決へと導くためのデータサイエンスについて学習する。この授業の目標は、課題の現状や構造を理解し、解決策の評価を行う能力を獲得することである。そのためには、与えられたデータを分析するだけでなく、どのような問いを立て、どのようなデータを用い、それをどのように分析して、その結果をどのように解釈するか、あなた自身がデザインできるようにならなければならない。その第一歩として、現実の社会的課題に挑んだ研究例を追体験する。データによって、何をどのように明らかにしたのか、そのプロセスを辿ることで、統計学とデータサイエンスの実践について理解を深める。

データサイエンス III（因果推論） 数学：★★★★☆☆

社会科学における多くの問いは因果的である。これは複雑な世界や社会の因果的な理解を通じて、人間行動と制度の理論的探究や社会的課題の要因と解決策の探究が試みられているからである。この困難な課題に対して、既存の知識と新たな情報から統計理論の力によって適切な推論を行おうとするのが因果推論である。この授業では因果的問いについての統計的推論である因果推論について、社会科学への応用を念頭に学習する。具体的には、潜在的結果フレームワークのもとで、無作為化比較試験、傾向スコア、マッチング、重み付け法、回帰分析、機械学習の応用、操作変数、回帰不連続デザイン、差分の差分法、固定効果、および、合成統制法といった様々なリサーチデザインと統計手法を学び、これらの長所と短所を学ぶ。特に、各手法が必要としている仮定とその妥当性の検討方法および結果の解釈について正確に理解し、実証分析に適切に応用できるようになることを目標とする。

データサイエンス IV（統計的モデリング） 数学：★★★★☆☆

社会現象を数字で表したデータは、単純なルールで生まれているとは限らない。そのため、よく知られている線形回帰モデルだけでは、現実を十分に説明できないことがある。そこで重要になるのが、データが現実の中でどのような仕組みによって生み出されたの

かを考え、それに合った統計モデルを用いることである。本講義では、政治学をはじめとする社会科学でよく用いられる統計モデルを紹介し、それらを組み合わせながら、受講生が自ら統計モデルを考え、実際に分析できるようになることを目標とする。具体的には、潜在変数、一般化線形モデル、混合分布、生存分析、多変量分布、時系列分析、因子分析などを扱う。

履修モデル

III および IV は、II の履修を前提とする。一方、I と II に順序の制約はないため、関心に応じて II から先に履修することも可能である。以下に代表的な履修モデルを示す。修士論文や博士論文で計量分析を行う場合には、少なくともモデル 3 以降の履修モデルを強く推奨する。

◆ モデル 1：統計リテラシー習得型 (I のみ)

提示された分析結果を読み解き、その手法を批判的に検討できるようになりたい大学院生向け。

◆ モデル 2：課題解決デザイン型 (II のみ)

社会的課題を発見し、解決へ導くために、統計学やデータサイエンスの実践的な活用を学び、自らリサーチデザインを習得したい大学院生向け。

◆ モデル 3：定性研究メイン型 (I→II (あるいは、II→I))

定性研究をメインとしつつ、計量分析も行いたい大学院生向け。

◆ モデル 4：定量研究メイン型 ((I→) II→III→IV)

計量分析をメインとする研究を行いたい大学院生向け。

2025 年度までの科目との関係

これらの科目の一部は、2025 年度まで我々が開講してきた授業科目を再編したものである。したがって、対応する科目をすでに履修している場合には、あらたに開講される科目を履修することはできない。2025 年度までの科目との対応関係は、おおむね次の通りである。

新科目名	旧科目名
データサイエンスⅠ（入門）	社会科学のデータ分析入門
データサイエンスⅡ（課題の発見と解決）	政治分析方法論Ⅰ
データサイエンスⅢ（因果推論）	政治分析方法論Ⅱ，因果推論
データサイエンスⅣ（統計的モデリング）	政治分析方法論Ⅲ

なお、移行期である都合上、2026年度は政治分析方法論Ⅲを開講し、データサイエンスⅣは2027年度から開講する予定である。また、ここで紹介した授業科目以外にも、統計学やデータサイエンスについて学習することができる授業科目が提供されている。