自然科学科目群

データ科学

現象を観察し分析して法則性を導くことにより現象の裏に潜む原理を解き明かす、という過程は学術における基本です。自然科学において実験により取得したデータだけでなく、人文科学・社会科学を含めた広範囲な学術において、フィールド調査により取得したデータ、収集によって得られた資料の分析は学術を進展させるための貴重な資源です。最近では、集積された大量のデータを二次的に活用するデータ駆動型研究も行われています。

日常生活においても、ICT（Information and Communication Technology）の発達によって、データを最大限に生かした新しい価値やサービスが次々と創出されていることは実感されます。例えば、ネットショッピングを行えば、購買データが自動的に蓄積され、客の嗜好を分析して販促方法を決定し、新商品を開発するための基礎資料として活用されることとなります。データの利用は、企業活動だけでなく、法律、金融・保険、健康・医療、災害対策など社会における様々な分野の発展に大きく寄与しています。情報ネットワークとサイバー空間が飛躍的に発展している今日では、データを利用することの重要性は高まる一方です。

データを収集して管理し、必要に応じて流通させ、さらに数理的手法により分析することにより結論を導き、将来の推測を行う学問は近年では「データ科学（Data Science）」とよばれています。データを扱う学問としては古くから統計学が発達してきました。特に数理統計学は数理的手法を用いることにより展開される学術分野です。さらに ICT の発達によって収集可能なデータが大規模化したため、それを扱う計算を高速に行うための情報学の修得も要請されています。すなわち、データ科学は、統計学、数理科学、情報学が融合した学問とみなすことができます。データ科学を学ぶことは、データを数理科学に裏付けられた理論に基づいてコンピュータを用いて分析し、その結果から適切な判断を下す、という一連の手法を学ぶことです。

全学共通教育においてデータ科学を修めることは、各自の専門分野において学術を修得し進展させるための基礎を身に付けるだけでなく、これからの社会において必要とされる知識の基本をも学ぶことにつながります。実際、わが国では、新産業創出や企業存続、社会サービスの向上、国際化進展などのためにはデータ科学の知識を持った人材が近い将来に不足することが懸念されており、データ科学を修めた人材の確保が喫緊の課題とされています。学術を発展させるだけでなくこのような社会からの要請に応えるということも鑑み、全学共通教育においては、データ科学科目として以下の科目を提供し、統計学・数理科学・情報学を横断的に学ぶことができるようにしています。

なお、文部科学省が定める「数理・データサイエンス・AI 教育プログラム認定制度（リテラシーレベル）」において認定を受けたプログラム（https://ds.k.kyoto-u.ac.jp/literacylevel-2/）により、「統計入門」の単位（2 単位）を修得することで、同プログラム修了証の取得が可能です。さらに、「数理・データサイエンス・AI 教育プログラム認定制度（応用基礎レベル）」において認定を受けたプログラム（https://ds.k.kyoto-u.ac.jp/ouyo-kiso/）により、「統計と人工知能」「データ分析基礎」「データ分析演習I」「データ分析演習II」のいずれかの単位（2 単位）を修得することで、同プログラム修了証の取得が可能です。各科目の概要は下記のとおりです。

1．統計入門　（英語対応科目：Introductory Statistics-E2）
データの解析手法についての学問である統計学は様々な分野において重要です。この講義では、データ科学が広範な分野に関わることを鑑み、より多くの学生が統計学の基本的な考え方を理解することを目標にしています。数学的な理論を厳密に理解することよりも、データを扱うエンドユーザーとしてデータの性質に応じた適切な分析方法を選択できるようになることを目指します。そのため、生活の身近な話題についての応用例を数多く紹介し、自然な形で統計的思考法が身に付くようにします。さらに統計解析ソフトを自習形式で取り入れ、実際的なデータ処理を通した感覚的な理解も深めます。

2．統計と人工知能（旧科目名「続・統計入門」）　（英語対応科目：Second Course in Statistics-E2）
文理問わず全ての分野の学生がエンドユーザーとして理解しておくべきリテラシーレベルの統計学をまとめた「統計入門」の発展版で、「統計入門」で扱い切れなかったより高度な話題（分散分析、回帰分析、因果推論、深層学習、画像処理など）を扱います。昨今、様々な分野の問題解決に大活躍している人工知能・機械学習の基本的な考え方や手法を理解し、統計学がどのような形でその基礎を形作っているかを学びます。

3．数理統計　（英語対応科目：Mathematical Statistics -E2）
統計学は様々な数値データの処理手法を提供しますが、ややもするとその手法の意味を理解することなく機械的な計算に陥ってしまいがちです。手法の意味するところを理解して初めて得られた解析結果を適切に利用することが可能となります。数理統計学は確率モデルに基づいた数理的手法による統計学であり、この講義では推定や検定な

どのデータ処理手法を裏付ける数学的な論理の構造を理解することにも力点が置かれます。

４．データ分析基礎　（英語対応科目：Basic Data Analysis-E2）
ビッグデータを分析し知識発見をするスキルは､今や文理や学問分野を問わず求められる時代です｡この講義では、ビッグデータ分析を視野に入れた、統計入門の一歩先の分析法（重回帰分析・主成分分析など）や推測アルゴリズムについて学習します。

５．データ分析演習I、II　（英語対応科目：Data Analysis Practice I, II-E2）
データ分析では、分析の手続きを覚えることはもちろんのこと、分析結果を正しく解釈し、そこから何らかの価値を見出すことが重要です。本講義では、さまざまなデータを統計入門で学習した手法で分析を行いながら、データ分析の基礎技術を実践的に体得することを目標としています。利用されるデータ分析手法や統計解析ソフトは学術分野によって異なるため、履修者が今後専門とする分野や興味に合わせてできるだけ幅広い選択ができるように、異なる専門分野を背景に持つ教員、異なる講義内容を揃えています。

６．数理・データ科学のための数学入門I、II
データ科学は統計学、数理科学、情報学の融合した学問であるため、エンドユーザーとして手法を学ぶだけでも、最低限の数学的知識が必要です。データ科学を学ぶための基礎としての数学の中から、Iでは条件付き確率、統計の基礎、多変量のデータ分析の基礎、データ解析の線形代数の基礎について、IIでは微分積分の内容を総合的に講述します。数学的な理論の完全な体系ではなく、データ科学への応用を重視する内容とします。文系学部学生が高校数学IIIの知識を持たないことに配慮します。