

授業科目名	自然科学データ解析	
時間割番号	GSC505	
担当教員	中村 和幸	
開講学期・曜日・時限	前期・金・II	単位数 2

<対象学生>

修士1年生

<授業の目的>

情報分野のみならず、農学や理学など様々な研究分野において、データを効率的かつ効果的に解析し、得られた結果を適正に解釈できる素養が求められている。この講義では、自然現象の特徴を理解するための様々なデータ解析や可視化などの方法について網羅的に学ぶ。データ解析の基礎となるデータ処理やそのためのプログラミング、確率統計やデータから推測する枠組みについて学ぶとともに、多様なデータへの対応や近年広く使われている機械学習についても学ぶ。

<本授業科目による獲得・涵養が期待されるコンピテンシー（能力・資質）>

医工農学総合教育部向け

記号	コンピテンシー(能力・資質)	説明
M-A	教養 専門に根差した科学的思考力 様々な分野における対象や現象をデータとして捉えその解析方法と結果の解釈に必要な基礎知識を有する	様々な分野の解析方法、結果の解釈に必要な基礎知識を有する
M-B	汎用能力 専門に根差した科学的思考力 自己の専門領域の深い知識と幅広い教養を、現代社会が直面する課題の解決に応用できる	専門領域の知識・教養を課題解決に応用できる
M-C	汎用能力 専門に根差した科学的思考力 現代社会が直面する課題に関するデータについて、様々な研究分野におけるデータ解析手法に精通し活用できる	課題解決に研究分野のデータ解析手法に精通し活用できる

<到達目標>

目標No	説明	修士
1	様々な解析方法の基礎を説明することができる	M-A
2	現象を数値として捉えることができる	M-B
3	データ解析にあまり馴染みのない学生は、データを扱い現象を数値として捉えることができる。データ解析に比較的馴染みのある学生は、解析によって得られた結果の位置づけと意義を考えることができる	M-C

<成績評価の方法>

目標No	割合	説明
1	20%	小テスト/レポート 各回の課題に関するレポートにより理解度を評価する
2	20%	最終レポート 最終レポートにより理解度を評価する
3	60%	受講態度 授業に取り組む姿勢（積極性など、講義中でのコメントや回答など）を評価する

<授業の方法>

本科目の各回の講義は、Zoomを利用したオンライン形式で提供される。各自接続し、講義を受けるようにすること。必要に応じて授業の終わりに課題を出して、提出締切日までに提出してもらう。PCによる演習を事後学習において実施する。

<受講に際して・学生へのメッセージ>

この講義はメディア授業科目（リアルタイム配信型）として開講される。出席はZoomの入室記録で確認する（講義参加前にZoomの表示を名前と所属に変更）。そのため、受講にはZoomを利用できる機器およびネットワーク環境が必要となる。授業中のデータ解析には、無料の統計ソフトRもしくはPythonを用いる（どちらか選択）。RとPythonの環境構築についてはWeek 1で説明する。必要に応じて授業の終わりに課題を掲示し、締切日までに提出してもらう。受講に際し、事前に予習を指示した内容について予習する。特に、各自の研究分野と各回のトピックについて、事前に調べておくこと。また、授業内容を復習して理解を深める。興味関心や自身の研究と関連する解析手法について調べ、研鑽を深める。

<テキスト>

<参考書>

塚本邦尊, 山田典一, 大澤文孝著, 東京大学のデータサイエンティスト育成講座 : Pythonで手を動かして学ぶデータ分析, マイナビ出版 (ISBN:9784839965259) 2019年出版
 金明哲著, Rによるデータサイエンス : データ解析の基礎から最新手法まで, 森北出版 (ISBN:9784627096011) 2007年出版
 村井潤一郎著, はじめてのR : ごく初歩の操作から統計解析の導入まで, 北大路書房 (ISBN:9784762828201) 2013年出版
 森巧尚著, Python1年生 : 体験してわかる!会話でまなべる!プログラミングのしくみ, 翔泳社 (ISBN:9784798153193) 2017年出版
 1年生
 Bill Lubanovic著 ; 鈴木駿監訳 ; 長尾高弘訳, 入門Python3 第2版, オライリー・ジャパン, オーム社(発売) (ISBN:9784873119328)

< 授業計画の概要 >

第1回	
タイトル	イントロダクション
事前学習 事後学習	<ul style="list-style-type: none"> ・授業前に、各自のPCに関する環境（OS、インストール済ソフトウェアなど）について事前に調査しておく。 ・講義で示されたプログラミング環境から各自に適した環境を検討し、採用する。オンラインサービスとローカル環境の両方が可能であれば、両方使用できるようにする。
授業内容	<ul style="list-style-type: none"> ・授業の最初に、この授業全体の到達目標、授業の進め方、課題提出の仕方、成績評価の仕方に関するガイダンスを行う。 ・本講義の概要について説明する。 ・本講義で使用するプログラミング言語であるPythonならびに R の実行環境について説明する。
第2回	
タイトル	プログラミングの基礎
事前学習 事後学習	<ul style="list-style-type: none"> ・前回実施したプログラミング実施環境について、再度確認する ・授業中に示したプログラミング課題例を実行し、内容を理解する
授業内容	<ul style="list-style-type: none"> ・プログラミングの考え方の基礎である、変数、代入、関数、メソッドといった概念についての説明を行う ・Python におけるJupyter Notebook, RにおけるR Studio の利用について説明する ・簡単なデータ解析例について実際に実行する
第3回	
タイトル	データ処理と可視化
事前学習 事後学習	<ul style="list-style-type: none"> ・これまでに学んだ、あるいは実施したことのある可視化についてまとめておく ・授業中に示した課題を実施し、まとめる
授業内容	<ul style="list-style-type: none"> ・記述統計の位置づけについて説明する ・プログラムを利用したデータ処理と可視化の方法を解説する。 ・データ処理結果、ならびに可視化結果の解釈について説明する。 ・具体的なデータ例による課題を示す。
第4回	
タイトル	確率統計の基礎
事前学習 事後学習	<ul style="list-style-type: none"> ・これまでに学んだ確率に関する知識を整理しておく。統計学を学んだことがある人は、統計学についても整理しておく ・講義で説明した確率計算の例、統計の例について調べてまとめる。できるだけ、各自の研究分野で調査する。
授業内容	<ul style="list-style-type: none"> ・推測統計の基本となる確率について説明する ・学生の理解状況に応じて、記述統計について、改めて確率の観点も交えながら再度説明する
第5回	
タイトル	推測統計の考え方
事前学習 事後学習	<ul style="list-style-type: none"> ・講義で説明した推測の枠組みについて復習する。
授業内容	<ul style="list-style-type: none"> ・推測統計の枠組み、考え方について説明する。
第6回	
タイトル	推定
事前学習 事後学習	<ul style="list-style-type: none"> ・推測統計の考え方について予習するとともに、各自の研究分野での推定の使用例について調べておく ・講義で示した課題を実施し、提出する。
授業内容	<ul style="list-style-type: none"> ・点推定、区間推定について説明する ・プログラムを通じてどのような結果が出るかについて理解を進める
第7回	
タイトル	様々な検定
事前学習 事後学習	<ul style="list-style-type: none"> ・各自の研究分野における検定の使用例について調べておく ・講義で示したプログラムを実行し、内容を理解する。
授業内容	<ul style="list-style-type: none"> ・検定の考え方、有意水準・p値の意味合い、過誤、多重検定などについて説明する。 ・プログラムを通じて検定を理解できるよう説明する。履修学生の分野や基礎知識にあわせて、多重性、各種検定などの中から内容を選択して解説する。
第8回	
タイトル	線形回帰モデル
事前学習 事後学習	<ul style="list-style-type: none"> ・各自の研究分野における線形回帰の使用例について調べておく ・講義で示したプログラムを実行し、内容を理解する。
授業内容	<ul style="list-style-type: none"> ・線形回帰モデルの考え方、カテゴリデータの取り扱い、変数選択について説明する。 ・プログラムを通じて線形回帰モデルについて説明する。
第9回	
タイトル	一般化線形モデル
事前学習 事後学習	<ul style="list-style-type: none"> ・各自の研究分野におけるカテゴリデータの使用例について調べておく ・講義で示したプログラムを実行し、内容を理解する。

授業内容	<ul style="list-style-type: none"> ・一般化線形モデル，特にロジスティック回帰について説明する． ・プログラムを通じて一般化線形回帰モデルについて説明する．
第10回	
タイトル	ベイズ統計学
事前学習 事後学習	<ul style="list-style-type: none"> ・ベイズ統計学の基盤となっているベイズの定理の考え方と活用の仕方，応用例についての講義中の内容について復習する．
授業内容	<ul style="list-style-type: none"> ・ベイズ統計学の基礎的な考え方について説明する． ・ベイズ統計学の実問題への応用例について説明する．
第11回	
タイトル	時系列・時空間データ解析
事前学習 事後学習	<ul style="list-style-type: none"> ・各自の研究分野における時系列・時空間データとしてどのようなものがあり，何を知りたいかについて調べておく ・講義で示した応用例を通じて，各自の分野での適用について検討する．
授業内容	<ul style="list-style-type: none"> ・時系列・時空間データの特徴と分析法について説明する． ・実問題への応用例について説明する．
第12回	
タイトル	機械学習Ⅰ・分類・識別
事前学習 事後学習	<ul style="list-style-type: none"> ・各自の研究分野において求められる分類問題ならびに識別問題について，調べておく ・講義中に示された識別課題を行う．
授業内容	<ul style="list-style-type: none"> ・分類・識別問題とはどのようなものであるかについて説明する． ・ごく簡単な識別問題を実施し，理解を進める．
第13回	
タイトル	機械学習Ⅱ・生成
事前学習 事後学習	<ul style="list-style-type: none"> ・各自の研究分野において生成モデルが利用される対象について調べておく ・講義で説明する生成モデルの構造について復習する．
授業内容	<ul style="list-style-type: none"> ・生成モデルと分類・識別問題の違いや共通点について説明する． ・オートエンコーダ，GANなどについて説明する．
第14回	
タイトル	まとめと総括
事前学習 事後学習	<ul style="list-style-type: none"> ・前回までの内容について，講義の前に各自で整理をしておく． ・講義の内容を踏まえ，最終レポートを作成する．
授業内容	授業全体を通して身に付いた知識や技能の確認を行うとともに、それらの中でどの知識や技能が各自の大学院やその後の研究に活用できるか、つながっていくかについて議論する．
第15回	
タイトル	
事前学習 事後学習	
授業内容	

< 前年度授業に対する改善要望等への対応 >

より多様な分野におけるデータ解析についての理解につながるよう，応用例に関する説明を増やす。

< 備考 >

本科目は明治大学との単位互換科目です。そのため、講義時間が本学とは異なりますので、ご注意ください。

【講義時間：10：50～12：30（100分授業×14回）】