

D-STEP 2019 年度講義シラバス

講義名	統計・機械学習モデリング	
担当教員	汪 金芳	
単位数	1 単位	
期間・曜日・時限	後・火 3 限	
科目区分	専門科目	
学習到達目標	線形モデル・ロジスティック回帰モデルを中心とした一般化線形モデルの理論、また意志決定における決定木の基本的考え方を理解し、統計ソフト R や Python を用いて実際のデータ解析を行うことができ、得られた結果を解釈できることを目標とする。	
講義概要	本講義の前半では、線形モデルと一般化線形モデルについての概説を行った後、線形モデル、ロジスティック回帰モデル、対数線形モデル、ベイズ流一般化線形モデルに焦点を当て、一般化線形モデルの理論と R を用いた実際の適用について解説する。後半では、サポートベクトルマシンや、スパース学習、決定木とアンサンブル学習について解説を行う。また Python を用いてこれらの機械学習の具体的適用についても解説を行う。	
成績評価方法	問題演習（6 割程度）とレポート（4 割程度）により成績の評価を行う。	
前提科目		
学習上・履修申請上の留意点及び準備学習		
教科書・参考図書等		
備考		
授業計画	1	線形モデルと一般化線形モデル
	2	ロジスティック回帰モデル
	3	対数線形モデル
	4	ベイズ流一般化線形モデル
	5	スパース学習：L1 正則化とスパース性
	6	スパース学習：スパースロジスティック回帰分析
	7	サポートベクトルマシン
	8	決定木とアンサンブル学習：決定木
	9	決定木とアンサンブル学習：ブースティング
	10	
	11	
	12	
	13	
	14	
	15	

D-STEP 2019 年度講義シラバス

講義名	データマニジング	
担当教員	清見 礼	
単位数	2 単位	
期間・曜日・時限	後・水 3 限	
科目区分	専門科目	
学習到達目標	様々なフォーマットのデータが与えられたとき、それらを統一的に扱うための手法を修得する。	
講義概要	データサイエンスの研究を行うにあたり、データファイルから所望のデータを抽出する、データを扱いやすい形に整形する、データのフォーマットを統一する、などは基本であり、また重要である。本講義では実践的に PC を操作しながらこれらについて学ぶ。	
成績評価方法	毎回の課題 100%	
前提科目	なし	
学習上・履修申請上の留意点 及び準備学習	毎回課題を出すので、それをしっかりこなすこと。もし時間内にできなければ次回までに仕上げること。	
教科書・参考図書等	教科書「Perl データマニジング」(ピアソン・エデュケーション)	
備考		
授業計画	1	イントロダクション
	2	データ、データマニジング
	3	一般的なマニジングの習慣
	4	便利なテクニック
	5	パターンマッチング
	6	ここまでの復習
	7	構造化されていないデータ
	8	レコード指向データ
	9	固定長データとバイナリデータ
	10	ここまでの復習
	11	複雑なデータフォーマット
	12	HTML
	13	XML
	14	独自パーサ
	15	まとめ

D-STEP 2019 年度講義シラバス

講義名	最適化と計算機科学																														
担当教員	藤田 慎也																														
単位数	2 単位																														
期間・曜日・時限	後・水 4 限																														
科目区分	専門科目																														
学習到達目標	<ul style="list-style-type: none"> ・アルゴリズムの基本的な考え方を理解する。 ・組合せ最適化に関する問題解決のための数理的手法を習得する。 																														
講義概要	「コンピュータで効率良く問題を解決するにはどうすればよいか？」という視点から、最適化理論と計算機科学に関する考え方を学ぶ。コンピュータを利用して問題を見通し良く、かつ効率良く解くための最適化問題の定式化と問題解決のための手法について講義する。																														
成績評価方法	試験またはレポートにより評価する。																														
前提科目																															
学習上・履修申請上の留意点及び準備学習																															
教科書・参考図書等	教科書は指定しない。 参考書：IT Text 離散数学（オーム社） その他、学生の理解度に応じて適切な参考書を授業中に紹介する。																														
備考																															
授業計画	<table border="1"> <tr><td>1</td><td>ガイダンス、最適化理論と計算機科学について</td></tr> <tr><td>2</td><td>最適化理論、計算機科学における数学的基礎の確認（1）</td></tr> <tr><td>3</td><td>最適化理論、計算機科学における数学的基礎の確認（2）</td></tr> <tr><td>4</td><td>計算の理論</td></tr> <tr><td>5</td><td>グラフ理論</td></tr> <tr><td>6</td><td>グラフアルゴリズム（1）</td></tr> <tr><td>7</td><td>グラフアルゴリズム（2）</td></tr> <tr><td>8</td><td>オートマトン（1）</td></tr> <tr><td>9</td><td>オートマトン（2）</td></tr> <tr><td>10</td><td>問題の難しさ、難問対策（1）</td></tr> <tr><td>11</td><td>問題の難しさ、難問対策（2）</td></tr> <tr><td>12</td><td>組合せ最適化に関するいろいろなアルゴリズム（1）</td></tr> <tr><td>13</td><td>組合せ最適化に関するいろいろなアルゴリズム（2）</td></tr> <tr><td>14</td><td>これまでのまとめ（1）</td></tr> <tr><td>15</td><td>これまでのまとめ（2）（試験を含む）</td></tr> </table>	1	ガイダンス、最適化理論と計算機科学について	2	最適化理論、計算機科学における数学的基礎の確認（1）	3	最適化理論、計算機科学における数学的基礎の確認（2）	4	計算の理論	5	グラフ理論	6	グラフアルゴリズム（1）	7	グラフアルゴリズム（2）	8	オートマトン（1）	9	オートマトン（2）	10	問題の難しさ、難問対策（1）	11	問題の難しさ、難問対策（2）	12	組合せ最適化に関するいろいろなアルゴリズム（1）	13	組合せ最適化に関するいろいろなアルゴリズム（2）	14	これまでのまとめ（1）	15	これまでのまとめ（2）（試験を含む）
1	ガイダンス、最適化理論と計算機科学について																														
2	最適化理論、計算機科学における数学的基礎の確認（1）																														
3	最適化理論、計算機科学における数学的基礎の確認（2）																														
4	計算の理論																														
5	グラフ理論																														
6	グラフアルゴリズム（1）																														
7	グラフアルゴリズム（2）																														
8	オートマトン（1）																														
9	オートマトン（2）																														
10	問題の難しさ、難問対策（1）																														
11	問題の難しさ、難問対策（2）																														
12	組合せ最適化に関するいろいろなアルゴリズム（1）																														
13	組合せ最適化に関するいろいろなアルゴリズム（2）																														
14	これまでのまとめ（1）																														
15	これまでのまとめ（2）（試験を含む）																														

D-STEP 2019 年度講義シラバス

講義名	都市環境データ解析特論	
担当教員	大西 暁生	
単位数	1 単位	
期間・曜日・時限	集中・前期	
科目区分	専門科目	
学習到達目標	<ul style="list-style-type: none"> ・都市環境問題の実態について理解する。 ・人間活動と環境問題の関係について理解する。 ・環境に関する施策や制度について理解する。 ・都市計画について理解する。 ・都市環境解析を事例を通じて知る。 	
講義概要	<p>都市環境問題に関する知識と専門性を深め、その問題解決に向けた様々な対策について学ぶ。具体的には、都市の成り立ちから、現代の都市が抱える環境問題について理解するとともに、将来の持続可能な都市を目指した環境管理、環境デザイン、まちづくりなどを都市計画制度と交えながら学ぶ。さらに、こうした都市環境問題を把握・評価し、対策を講じていくための統計や地理情報等といったデータの入手と解析の方法について理解する。</p>	
成績評価方法	課題 100%	
前提科目		
学習上・履修申請上の留意点 及び準備学習		
教科書・参考図書等		
備考		
授業計画	1	都市環境問題とは
	2	これまでの都市環境問題
	3	都市と環境
	4	環境管理と制御手法
	5	環境影響評価制度
	6	都市計画
	7	都市環境解析の分析事例
	8	まとめ
	9	
	10	
	11	
	12	
	13	
	14	
	15	

D-STEP 2019 年度講義シラバス

講義名	非構造化データ特論	
担当教員	山崎 真見	
単位数	1 単位	
期間・曜日・時限	後・木 4 限	
科目区分	専門科目	
学習到達目標	テキスト、センサ信号、画像などの非構造化データから有用な情報を抽出するための基本的な特徴量エンジニアリング手法を使えるようになることを目標とする。	
講義概要	リレーショナルデータベースに格納されているデータを構造化データと呼び、各列の意味づけが事前に与えられた表形式に変換可能なデータである。この範疇に収まらないデータを全て非構造化データと呼び、テキスト、音声、画像、センサ信号などが該当する。社会インフラとしての Web が整備されるとともにモバイル情報端末の普及により、膨大な非構造化データが利用可能となり、これらデータの活用がデータサイエンスの実践において不可欠となっている。本講義では、非構造化データの管理技術および各種データの特徴と情報抽出のための実践的な技法を理解する。	
成績評価方法	レポート 100%	
前提科目		
学習上・履修申請上の留意点及び準備学習		
教科書・参考図書等		
備考		
授業計画	1	非構造化データの管理
	2	信号処理：特徴量抽出
	3	信号処理：特徴の学習
	4	自然言語処理：形態素解析、構文解析、意味解析、文脈解析
	5	自然言語処理：CBoW 表現の構築
	6	画像データ処理：基本的な前処理技術
	7	画像データ処理：認識と生成
	8	まとめ
	9	
	10	
	11	
	12	
	13	
	14	
	15	

D-STEP 2019 年度講義シラバス

講義名	データ可視化特論	
担当教員	阿部 貴行・大西 暁生	
単位数	1 単位	
期間・曜日・時限	集中	
科目区分	専門科目	
学習到達目標		
講義概要	データの可視化は、大量の多次元データから意味あるシグナルを検出する際や解析結果の説得力を高める際に有用であり、データサイエンスにおいて不可欠である。本講義では、データを可視化する際の文法、色彩や形状の選択や層別化の方法などを学び、空間データの作図と分析、また回帰樹やクラスター分析などの結果の可視化が必須の統計手法についても解説する。地理情報に基づくソフトウェア、プログラミングに基づく R、GUI に基づく双方向の統計解析・可視化ソフトウェア等を用いた演習も行う。	
成績評価方法	課題 100%	
前提科目		
学習上・履修申請上の留意点及び準備学習		
教科書・参考図書等		
備考		
授業計画	1	統計ソフトとグラフ化入門
	2	グラフの種類と特性
	3	グラフ化の文法
	4	層別化とデータの可視化
	5	統計手法とグラフ化（クラスター分析，分類・回帰樹）
	6	統計手法とグラフ化（ネットワークダイアグラム）
	7	地理情報システムの概要と地図化
	8	地理情報システムによる分析
	9	
	10	
	11	
	12	
	13	
	14	
	15	

D-STEP 2019 年度講義シラバス

講義名	プログラミング特論	
担当教員	阿部 貴行・立川 仁典	
単位数	1 単位	
期間・曜日・時限	後・火 2 限	
科目区分	専門科目	
学習到達目標	R などのプログラミング言語の基本的な文法を理解し、プログラミングにより基本的な統計手法を使ったデータ解析ができるようになる。更に並列プログラミングの技術を習得し、様々な乱数を用いたシミュレーションを実行できるようになる。	
講義概要	<p>データサイエンス手法を用いてデータ解析をするために必要なプログラミングに関する講義を行う。様々な統計解析手法を具現化するためのツールとして、R 等を用いたプログラム演習も実施する。また近年のスーパーコンピュータは、複数台、高速ネットワークで接続することで、高い処理速度を実現している。そのようなシステム性能を最大限に引き出し活用するために、MPI のような並列プログラミング技術を習得する。プログラム高速化の概念や技術、並列プログラム開発技術なども習得する。</p> <p>(阿部 貴行 / 4 回) 解析結果の信頼性や再現性の保証を目的とし、R 等のプログラミングによる統計解析やシミュレーションを講義・演習する。</p> <p>(立川 仁典 / 4 回) 並列プログラミング技術の習得を目的とし、FORTRAN 等のプログラミングによるシミュレーションを講義・演習する。</p>	
成績評価方法	課題 100%	
前提科目		
学習上・履修申請上の留意点及び準備学習		
教科書・参考図書等		
備考		
授業計画	1	プログラミングに関する入門講義・演習
	2	プログラミングの基本的な文法・演習
	3	R 等を用いたデータの要約・演習
	4	R 等を用いた統計解析入門 1
	5	R 等を用いた統計解析入門 2
	6	R 等を用いたシミュレーション実験
	7	FORTRAN 等を用いた並列プログラミング演習
	8	FORTRAN 等を用いたシミュレーション実験
	9	
	10	
	11	
	12	
	13	
	14	
	15	

D-STEP 2019 年度講義シラバス

講義名	実験とシミュレーション特論	
担当教員	立川 仁典	
単位数	1 単位	
期間・曜日・時限	後・火 1 限	
科目区分	専門科目	
学習到達目標		
講義概要	<p>実験だけでは理解することのできない、様々なシミュレーション技術を学ぶ。具体的には、シミュレーション技術の背景にある理論、すなわち分子力学法、分子動力学法、分子軌道法の基本概念を学び、比較的小さな分子（医薬化合物や有機材料）や生体高分子（タンパク質、RNA、DNA など）の熱力学的性質や電子的性質について学びます。化学的、分子生物学的、生物物理学的現象を理解することです。またデータサイエンス的な手法も駆使することにより、マテリアルズインフォマティクスやバイオインフォマティクスに関しても学ぶ。</p> <p style="text-align: center;">自然科学に基づいたシミュレーション技術である、分子力学法、分子動力学法、分子軌道法の基本概念を学び、様々な実験結果との比較を行う。</p>	
成績評価方法	課題 100%	
前提科目		
学習上・履修申請上の留意点及び準備学習		
教科書・参考図書等		
備考		
授業計画	1	自然科学法則の定式化
	2	分子力学法
	3	分子動力学法 1
	4	分子動力学法 2
	5	分子軌道法 1
	6	分子軌道法 2
	7	マテリアルズインフォマティクスと実験
	8	バイオインフォマティクスと実験
	9	
	10	
	11	
	12	
	13	
	14	
	15	

D-STEP 2019 年度講義シラバス

講義名	多変量データ解析	
担当教員	小泉 和之	
単位数	2 単位	
期間・曜日・時限	前・火 2 限	
科目区分	専門科目	
学習到達目標	多変量データの取り扱い方と様々な解析法について理論と応用の両面から理解することを目標とする。	
講義概要	多変量解析とは、多くの変数を持つデータが持つ特徴をまとめ、各変数間の相互関係を明らかにする統計的手法の総称である。モデリングなどの分野で使われる回帰分析については単回帰分析とその評価方法から始め、重回帰分析においては情報量基準を用いた変数選択も導入し、分析の評価法について理解する。さらには判別分析、主成分分析、多次元尺度構成法等幅広い手法についても学修する。本講義では、多変量解析の主な手法の理論を学び、多変量データに対して適切な解析ができるようにする。	
成績評価方法	演習・レポートにより評価する。	
前提科目		
学習上・履修申請上の留意点 及び準備学習		
教科書・参考図書等	参考書：永田靖・棟近雅彦「多変量解析法入門」（サイエンス社）	
備考		
授業計画	1	多変量データ解析とは
	2	単回帰分析と最小二乗法
	3	重回帰分析
	4	モデル選択と情報量基準
	5	交互作用モデル
	6	線形判別分析
	7	様々な判別分析
	8	主成分分析
	9	主成分回帰分析
	10	数量化理論
	11	多次元尺度構成法
	12	クラスター分析
	13	ファジィクラスターリング
	14	主成分クラスター分析
	15	まとめ

D-STEP 2019 年度講義シラバス

講義名	標本調査法	
担当教員	土屋 隆裕	
単位数	1 単位	
期間・曜日・時限	前・金 1 限	
科目区分	専門科目	
学習到達目標	各標本抽出法の特徴を理解し、適切な標本計画を立案することができる。	
講義概要	公的統計調査や社会調査・マーケティング調査といった統計調査の企画・設計にあたっては、調査票の作成方法や調査の実施方法、標本抽出の方法や母集団推定の方法など、文理をまたいだ様々な技術について精通している必要がある。この授業では、適切な調査の企画・設計ができるようになるため、それらの技術のうち、特に標本抽出の方法や母集団推定の方法について、考え方や理論の概要を、具体的な数値例を用いながら学ぶ。	
成績評価方法	課題 100%	
前提科目		
学習上・履修申請上の留意点 及び準備学習		
教科書・参考図書等		
備考		
授業計画	1	母集団と標本
	2	単純無作為抽出法
	3	不偏推定量と抽出ウエイト
	4	不偏推定量の分散
	5	Basu の象の教訓
	6	層化抽出法
	7	集落抽出法
	8	二段抽出法
	9	
	10	
	11	
	12	
	13	
	14	
	15	

D-STEP 2019 年度講義シラバス

講義名	時系列データ解析特論	
担当教員	中谷 朋昭	
単位数	1 単位	
期間・曜日・時限	集中	
科目区分	専門科目	
学習到達目標	<ul style="list-style-type: none"> ・時系列データの定常性と非定常性の違いについて理解する ・基礎的な時系列モデルの特徴について理解する ・時系列モデルによる予測の基礎を理解する ・定常性を確認する方法について理解する 	
講義概要	<p>時系列データは、対象を時間軸に沿って観測したもので、気温や湿度、地震波などの気象・気候データ、外国為替レートや株価などの金融データなど、身近に数多く存在している。本講義では、1変数の場合に限定して、これら時系列データを解析するために必要とされる基礎的な内容を教授する。</p>	
成績評価方法	期末レポート 100%	
前提科目	<p>以下の内容を事前に習得しておくこと、講義の理解に有益である。</p> <ul style="list-style-type: none"> ・数理統計学（確率変数、確率密度、期待値、仮設検定など） ・計量経済学または多変量解析（回帰分析に関すること） 	
学習上・履修申請上の留意点及び準備学習	<p>講義内容にあわせて、3年次までの微分積分、線形代数、確率・統計、数理統計学、プログラミング演習の内容を復習して、理解を確固たるものにしておくことが求められる。</p>	
教科書・参考図書等	沖本竜義「経済・ファイナンスデータの計量時系列分析」、朝倉書店、2010年	
備考		
授業計画	1	時系列分析の基礎
	2	定常性、ホワイトノイズ、自己相関
	3	自己回帰モデル
	4	移動平均モデル
	5	自己回帰移動平均モデル
	6	時系列モデルを用いた予測
	7	時系列の非定常性と単位根
	8	単位根検定
	9	
	10	
	11	
	12	
	13	
	14	
	15	

D-STEP 2019 年度講義シラバス

講義名	計算機統計学特論	
担当教員	中谷 朋昭	
単位数	1 単位	
期間・曜日・時限	集中	
科目区分	専門科目	
学習到達目標	<ul style="list-style-type: none"> ・統計解析の過程を記録（プログラム）として残すことの重要性を理解する ・フリーの統計解析環境 R の基本的特性と利用方法を理解する ・他のソフトウェアと R の連携方法について理解する 	
講義概要	<p>大規模なデータは、一般的な表計算ソフトウェアでは処理に困難を生じることがある。また、統計解析の再現性確保のためにも、途中でどのような処理をしたのかを記録することが求められるようになってきている。本講義では、これまで主として表計算ソフトウェアを利用して統計解析を行ってきた学生を対象に、フリーの統計解析環境 R を利用した統計解析の基礎を教授する。加えて、他のソフトウェアとの連携方法についても解説する。</p>	
成績評価方法	期末レポート 100%	
前提科目	<p>初回に紹介する。 R のさまざまな利用方法は、ネット上で多数紹介されている。必要に応じて、検索エンジンなどを利用して自ら調べることで、R を活用する力を身につけることができる。</p>	
学習上・履修申請上の留意点 及び準備学習	<p>講義内容にあわせて、3 年次までの微分積分、線形代数、確率・統計、数理統計学、プログラミング演習の内容を復習して、理解を確固たるものにしておくことが求められる。</p>	
教科書・参考図書等	<p>初回に紹介する。R のさまざまな利用方法は、ネット上で多数紹介されている。必要に応じて、検索エンジンなどを利用して自ら調べることで、R を活用する力を身につけることができる。</p>	
備考	講義中には、PC を利用できるようにしておくこと	
授業計画	1	R と関連するソフトウェアの入手
	2	計算機としての R
	3	R におけるデータの入出力
	4	描画ソフトとしての R
	5	R におけるプログラミング
	6	アドオンパッケージの利用
	7	外部ソフトウェアとの連携
	8	R を用いた文書作成
	9	
	10	
	11	
	12	
	13	
	14	
	15	

D-STEP 2019 年度講義シラバス

講義名	数理医学																														
担当教員	中村 和幸, 田栗 正隆, 大家 義登																														
単位数	2 単位																														
期間・曜日・時限	2019 年度休講																														
科目区分	専門科目																														
学習到達目標	医学の現場における統計解析によるアプローチや数理モデルを用いたアプローチの役割と重要性を理解するとともに、いくつかの手法の基礎的な内容について実施できるようになる。																														
講義概要	医学研究においては、統計科学分野との協働研究が広くなされてきたが、さらに数理科学分野や情報科学分野との協同研究も急速に進展している。これは、もちろん生体という機能が生理学だけでなく物理・化学・電気化学的な問題とも本質的に関連しているからであるが、統計化学数理モデル的な手法とコンピューターの発展が協同研究に必要な環境を提供していることによる。また医療や医学研究の現場で使用されている画像情報は年々増加の一途をたどり、もはや人間の視覚だけでは処理しきれない水準まで達している。このような現状で、統計処理・数理モデルを用いたアプローチが必須である。本講義では、統計科学・数理科学と臨床医学との協働研究の現場で行われている最先端の研究を紹介する。																														
成績評価方法	講義内容に関するレポートにより評価する。																														
前提科目																															
学習上・履修申請上の留意点及び準備学習	講義内容にあわせて、3年次までの微分積分、線形代数、確率・統計、数理統計学、プログラミング演習の内容を復習して、理解を確固たるものにしておくことが求められる。																														
教科書・参考図書等	特になし。																														
備考																															
授業計画	<table border="1"> <tr><td>1</td><td>ガイダンス (中村 1)</td></tr> <tr><td>2</td><td>因果効果の定義 (田栗 1)</td></tr> <tr><td>3</td><td>観察研究における交絡調整 (田栗 2)</td></tr> <tr><td>4</td><td>回帰不連続デザイン (田栗 3)</td></tr> <tr><td>5</td><td>直接効果・間接効果の推定① (田栗 4)</td></tr> <tr><td>6</td><td>直接効果・間接効果の推定② (田栗 5)</td></tr> <tr><td>7</td><td>時間依存性交絡の調整 (田栗 6)</td></tr> <tr><td>8</td><td>操作変数法 (田栗 7)</td></tr> <tr><td>9</td><td>臨床試験における共変量調整 (田栗 8)</td></tr> <tr><td>10</td><td>ベイズ型数理モデルとパラメータ推定 (中村 2)</td></tr> <tr><td>11</td><td>パラメータ推定アルゴリズムの数理 (中村 3)</td></tr> <tr><td>12</td><td>医療画像可視化とその原理 (中村 4)</td></tr> <tr><td>13</td><td>創薬におけるデータ解析 (大家)</td></tr> <tr><td>14</td><td>まとめ (中村 5)</td></tr> <tr><td>15</td><td></td></tr> </table>	1	ガイダンス (中村 1)	2	因果効果の定義 (田栗 1)	3	観察研究における交絡調整 (田栗 2)	4	回帰不連続デザイン (田栗 3)	5	直接効果・間接効果の推定① (田栗 4)	6	直接効果・間接効果の推定② (田栗 5)	7	時間依存性交絡の調整 (田栗 6)	8	操作変数法 (田栗 7)	9	臨床試験における共変量調整 (田栗 8)	10	ベイズ型数理モデルとパラメータ推定 (中村 2)	11	パラメータ推定アルゴリズムの数理 (中村 3)	12	医療画像可視化とその原理 (中村 4)	13	創薬におけるデータ解析 (大家)	14	まとめ (中村 5)	15	
1	ガイダンス (中村 1)																														
2	因果効果の定義 (田栗 1)																														
3	観察研究における交絡調整 (田栗 2)																														
4	回帰不連続デザイン (田栗 3)																														
5	直接効果・間接効果の推定① (田栗 4)																														
6	直接効果・間接効果の推定② (田栗 5)																														
7	時間依存性交絡の調整 (田栗 6)																														
8	操作変数法 (田栗 7)																														
9	臨床試験における共変量調整 (田栗 8)																														
10	ベイズ型数理モデルとパラメータ推定 (中村 2)																														
11	パラメータ推定アルゴリズムの数理 (中村 3)																														
12	医療画像可視化とその原理 (中村 4)																														
13	創薬におけるデータ解析 (大家)																														
14	まとめ (中村 5)																														
15																															

D-STEP 2019 年度講義シラバス

講義名	行政課題解決 PBL	
担当教員	田栗 正隆・窪田 和巳	
単位数	2 単位	
期間・曜日・時限	前・金 13:30-18:00	
科目区分	PBL 科目	
学習到達目標	公共領域における課題に対し、データ分析を用いた課題の深堀から解決策の策定・提言までの問題解決の一連のプロセスについて、関係者と連携しながら進められる素地の醸成	
講義概要	本講義では、個票データ（擬似データ含む）分析を用いた問題解決の一覧のプロセスを、主にグループワークを通じて実践的な形で学習する。各グループには、ビジネス経験、データ分析経験が豊富な講師が付き、各回のアウトプットに対してフィードバックを提供する。また、グループワークを通じて問題解決における他者との協働を身に着けるだけでなく、クライアントへの模擬インタビューやプレゼンテーションを実施し、問題解決に必要な能力の素地を総合的に養う。	
成績評価方法	最終報告におけるプレゼンテーション 50%+レポート 30%+授業での貢献度 20%	
前提科目		
学習上・履修申請上の留意点 及び準備学習		
教科書・参考図書等		
備考		
授業計画	1	講義①（データ分析における心構え）
	2	講義②（ロジカルシンキング）
	3	課題説明、グループワーク（アプローチ策定、データ解析）
	4	グループワーク（データ解析、課題深堀）
	5	クライアントインタビューによる追加情報の収集
	6	グループワーク（インタビュー結果を踏まえたデータ再解析）
	7	クライアントへの中間報告及びクライアントとの意見交換
	8	グループワーク（中間報告結果を受けた深堀）
	9	ビジネス専門家によるレビュー及びレビュー結果を踏まえた深堀
	10	講義③（解決策提言に向けたストーリー構築）
	11	最終プレゼンテーションに向けた資料作成、グループワーク（更なる深堀）
	12	最終プレゼンテーションリハーサル
	13	最終プレゼンテーションに向けた資料のアップデート
	14	クライアントへの最終プレゼンテーション及びクライアントからのフィードバック
	15	クライアントからのフィードバックを踏まえたアップデート、授業の振り返り

D-STEP 2019 年度講義シラバス

講義名	ビジネス課題解決 PBL																														
担当教員	小泉 和之・西内 啓																														
単位数	2 単位																														
期間・曜日・時限	後・金 13:30-18:00																														
科目区分	PBL 科目																														
学習到達目標	企業の現実のデータ課題に対して「どうすれば収益性を向上させられるか」という具体的な施策を提案できるようになること																														
講義概要	<p>本講義ではこれまでに学んだ知識を総動員してデータサイエンスを業務に役立てる一連の流れを身につけるべく、実際の企業の課題とデータを用いて OJT(On the Job Training)形式で演習を進める。具体的には大きく分けて次の 3 つのタスクがこなせられるようになることを目指す。</p> <ul style="list-style-type: none"> ・既存データの入手から分析方針の策定、データ加工、モデリング、予測およびその評価を行うこと ・質的（インタビュー）調査の結果から新規に量的調査フォームを開発すること ・分析結果を具体的な企画書に落とし込みわかりやすくプレゼンテーションすること <p>以上を目的として、グループディスカッションを通して「考える力」を、実習を通して「手を動かす力」を、またその過程で必要な知識を「自分で調べられる力」を身につける。</p>																														
成績評価方法	<p>期末レポート 50%</p> <p>最終報告会におけるプレゼンテーション 50%</p>																														
前提科目																															
学習上・履修申請上の留意点 及び準備学習	<p>受講者は一般化線型モデル、機械学習による予測手法、R によるプログラミングについて基礎的な理解を有することが求められる。</p> <p>また演習においては SQL によるデータ加工を行なうため可能な限り下記に挙げる参考図書を用いて予習をしておくことが望ましい。</p>																														
教科書・参考図書等	『SQL 第 2 版 ゼロからはじめるデータベース操作』（翔泳社刊）																														
備考	講義中には、PC を利用できるようにしておくこと																														
授業計画	<table border="1"> <tr><td>1</td><td>与えられたデータに対するリサーチデザイン</td></tr> <tr><td>2</td><td>SQL によるデータ加工の基本</td></tr> <tr><td>3</td><td>演習：SQL によるデータ加工</td></tr> <tr><td>4</td><td>R を用いたデータ分析とモデリング</td></tr> <tr><td>5</td><td>連携企業を交えた中間レビュー①</td></tr> <tr><td>6</td><td>機械学習手法による予測およびクロスバリデーション</td></tr> <tr><td>7</td><td>演習：R を用いた予測モデルの作成</td></tr> <tr><td>8</td><td>質的調査と量的調査</td></tr> <tr><td>9</td><td>演習：インタビュー調査の実施</td></tr> <tr><td>10</td><td>演習：量的調査フォームの開発</td></tr> <tr><td>11</td><td>演習：量的調査結果の分析</td></tr> <tr><td>12</td><td>施策立案とプレゼンテーション</td></tr> <tr><td>13</td><td>連携企業を交えた中間レビュー②</td></tr> <tr><td>14</td><td>演習：分析・予測および提案内容のブラッシュアップ</td></tr> <tr><td>15</td><td>連携企業を交えた最終報告会</td></tr> </table>	1	与えられたデータに対するリサーチデザイン	2	SQL によるデータ加工の基本	3	演習：SQL によるデータ加工	4	R を用いたデータ分析とモデリング	5	連携企業を交えた中間レビュー①	6	機械学習手法による予測およびクロスバリデーション	7	演習：R を用いた予測モデルの作成	8	質的調査と量的調査	9	演習：インタビュー調査の実施	10	演習：量的調査フォームの開発	11	演習：量的調査結果の分析	12	施策立案とプレゼンテーション	13	連携企業を交えた中間レビュー②	14	演習：分析・予測および提案内容のブラッシュアップ	15	連携企業を交えた最終報告会
1	与えられたデータに対するリサーチデザイン																														
2	SQL によるデータ加工の基本																														
3	演習：SQL によるデータ加工																														
4	R を用いたデータ分析とモデリング																														
5	連携企業を交えた中間レビュー①																														
6	機械学習手法による予測およびクロスバリデーション																														
7	演習：R を用いた予測モデルの作成																														
8	質的調査と量的調査																														
9	演習：インタビュー調査の実施																														
10	演習：量的調査フォームの開発																														
11	演習：量的調査結果の分析																														
12	施策立案とプレゼンテーション																														
13	連携企業を交えた中間レビュー②																														
14	演習：分析・予測および提案内容のブラッシュアップ																														
15	連携企業を交えた最終報告会																														

D-STEP 2019 年度講義シラバス

講義名	統計学基礎	
担当教員	小泉 和之	
単位数	1 単位	
期間・曜日・時限	前	
科目区分	e-learning	
学習到達目標	統計学の基礎的な方法論について例を交えながら理解することを目標とする。	
講義概要	本講義では、統計学の考え方を理解することから始める。その上で、データをまとめる記述統計学で最低限必要である代表値、散布度、外れ値などについて理解する。多次元データの取り扱いに際して、2次元データを導入し、相関と回帰について理解する。さらには確率を改めて定義し、広い視野で確率を見る力を身に着ける。母集団と標本（データ）の考え方と確率の世界をつなげるために確率分布とその性質について理解する。最終的には確率分布のパラメータに対する推定、仮説検定の考え方について理解する。	
成績評価方法	課題 100%	
前提科目		
学習上・履修申請上の留意点 及び準備学習		
教科書・参考図書等	参考書：藤川浩、小泉和之「生物系のためのやさしい基礎統計学」（講談社）	
備考	実際の PC での実践は並行して開講されているデータ分析基礎が大いに役立つ。	
授業計画	1	統計学の考え方と代表値・散布度
	2	多次元データの取り扱い
	3	確率・ベイズ確率
	4	確率分布と現象
	5	確率分布の期待値・分散
	6	パラメータ推定
	7	仮説検定の基礎
	8	まとめ
	9	
	10	
	11	
	12	
	13	
	14	
	15	

D-STEP 2019 年度講義シラバス

講義名	応用線形代数	
担当教員	田栗 正隆	
単位数	1 単位	
期間・曜日・時限	前	
科目区分	e-learning	
学習到達目標	データサイエンスに必要な行列、行列を用いた連立方程式の解法、行列式、ベクトル空間、固有値と固有ベクトル等の基礎的事項を理解する。	
講義概要	線形代数は統計学やデータサイエンス分野で要求される数学的な知識の重要な部分をなしている。本講義では、線形代数の応用にあたり主要な内容を計算例を前面に出して解説し、例題を盛り込む形で分かりやすく教授する。抽象的な概念や定理等の証明は必要最小限にとどめ、重要な内容はレポート課題等で確認できるようにする。	
成績評価方法	毎回のレポート課題 100%	
前提科目	特になし	
学習上・履修申請上の留意点 及び準備学習		
教科書・参考図書等	Alan Jeffrey, 「Matrix Operations for Engineers and Scientists: An Essential Guide in Linear Algebra」、Springer、2010 年	
備考		
授業計画	1	行列
	2	連立一次方程式
	3	行列式
	4	ベクトル空間
	5	線形変換と射影
	6	固有値と固有ベクトル
	7	2次形式、データサイエンスへの応用
	8	最終テスト
	9	
	10	
	11	
	12	
	13	
	14	
	15	

D-STEP 2019 年度講義シラバス

講義名	データ分析基礎	
担当教員	坂巻 顕太郎	
単位数	1 単位	
期間・曜日・時限	前	
科目区分	e-learning	
学習到達目標	データの特徴や分析目的と分析手法の関係を理解し、各種手法を R でプログラミングでできることを目標とする。	
講義概要	データサイエンスにおいて、データの特徴を理解したうえで、データを分析し、その分析結果を解釈することが重要である。本講義では、データの理解を促したうえで、比較、因果関係の推測、予測などの目的と関連した分析手法を説明する。特に、検定、回帰モデル、機械学習を扱い、それぞれの手法を R により実行する。講義内容の理解度は課題により確認する。	
成績評価方法	課題 100%	
前提科目		
学習上・履修申請上の留意点及び準備学習		
教科書・参考図書等	参考書:今井 耕介, 粕谷 祐子, 原田 勝孝, 久保 浩樹. 社会科学のためのデータ分析入門(上)(下). 岩波書店.	
備考		
授業計画	1	データ分析の実際
	2	R の基礎
	3	様々なグラフ
	4	比較と検定
	5	因果と予測
	6	回帰モデル
	7	機械学習
	8	最終テスト
	9	
	10	
	11	
	12	
	13	
	14	
	15	

D-STEP 2019 年度講義シラバス

講義名	IT セキュリティ・情報倫理	
担当教員		
単位数	1 単位	
期間・曜日・時限		
科目区分	e-learning	
学習到達目標		
講義概要		
成績評価方法		
前提科目		
学習上・履修申請上の留意点 及び準備学習		
教科書・参考図書等		
備考		
授業計画	1	
	2	
	3	
	4	
	5	
	6	
	7	
	8	
	9	
	10	
	11	
	12	
	13	
	14	
	15	

D-STEP 2019 年度講義シラバス

講義名	ノンパラメトリック法	
担当教員	村上 秀俊	
単位数	2 単位	
期間・曜日・時限	後・木 2 限	
科目区分	専門科目	
学習到達目標	<ol style="list-style-type: none"> 1. 統計的検定法の手順を説明できるようにする。 2. 局所最強力検定, 漸近相対効率, 漸近正規性について説明できるようにする。 3. 鞍点近似とその基本的性質について説明できるようにする。 4. カーネル密度推定を理解し, バンド幅を導出できるようにする。 	
講義概要	ノンパラメトリック法や近似法の理論および考え方を学ぶ。ノンパラメトリック検定では局所最強力検定, 漸近相対効率, 漸近正規性について学び, 近似法では鞍点近似を中心に学ぶ。また, カーネル密度推定の理論について学ぶ。	
成績評価方法	レポート課題を総合的に評価する。	
前提科目		
学習上・履修申請上の留意点 及び準備学習	学部での数理統計学の基礎知識は前提として講義を行う。	
教科書・参考図書等	参考書：村上秀俊「統計解析スタンダード ノンパラメトリック法」朝倉書店	
備考		
授業計画	1	ノンパラメトリック検定 (1) 順序統計量およびノンパラメトリック検定の基本的な考え方を理解する。
	2	ノンパラメトリック検定 (2) 基本的な検定統計量および局所最強力検定について理解する。
	3	ノンパラメトリック検定 (3) 漸近相対効率について理解する。
	4	ノンパラメトリック検定 (4) 漸近正規性について理解する。
	5	鞍点近似 (1) 鞍点近似の基本的な考え方を理解する。
	6	鞍点近似 (2) 経験鞍点近似および近似誤差について理解する。
	7	鞍点近似 (3) ノンパラメトリック検定への適応について理解する。
	8	これまでの内容の理解度を確認する。
	9	カーネル密度推定について (1) カーネル密度推定の基本的な考え方を理解する。
	10	カーネル密度推定について (2) カーネル関数の選択, バンド幅の決定法について理解する。
	11	カーネル密度推定について (3) カーネル関数の選択, バンド幅の決定法について理解する。
	12	カーネル密度推定について (4) 高次カーネル密度推定について理解する。
	13	カーネル密度推定について (5) 多変量データに対するカーネル密度推定の基本的な考え方を理解する。
	14	カーネル密度推定について (6) 多変量データに対するカーネル関数の選択, バンド幅の決定法について理解する。
	15	本科目内容の到達度の確認と解説

D-STEP 2019 年度講義シラバス

講義名	欠測データ解析
担当教員	瀬尾 隆
単位数	2 単位
期間・曜日・時限	前・水 1 限
科目区分	専門科目
学習到達目標	<ol style="list-style-type: none"> 1 統計的検定法の手順を説明できるようにする。 2 多変量正規分布とその基本的性質について説明できるようにする。 3 平均ベクトルや共分散行列の最尤推定量を導出できるようにする。 4 尤度比検定を理解し説明できるようにする。 5 データが欠測値を含む場合の平均ベクトルの検定法について理解し，説明できるようにする。
講義概要	統計的推測理論の基礎概念を確認し，統計解析理論や統計的手法の考え方を学ぶ。特に，多変量解析の平均ベクトルに関する検定を中心に，尤度比検定，欠測値データにおける統計解析理論について学ぶことができる。
成績評価方法	到達度評価試験・小テストの結果ならびにレポート課題を総合的に評価する。
前提科目	
学習上・履修申請上の留意点 及び準備学習	学部での数理統計学，および，多変量解析の基礎知識は前提にして講義を行う。
教科書・参考図書等	教科書：資料を配布する 参考書：講義の初めに紹介する。
備考	
授業計画	<ol style="list-style-type: none"> 1 基礎と復習 1 統計的仮設検定，多変量正規分布などについて復習する。 2 基礎と復習 2 多変量正規分布とウィシャート分布標本平均ベクトル，標本分散共分散行列の性質を理解する。 3 欠測データ 1 欠測のメカニズムと単調欠測データとは何かを理解する 4 欠測データ 2 2-step 単調欠測データのもとで平均ベクトルと分散共分散行列の最尤推定量の導出を理解する。 5 平均ベクトルの検定 1 尤度比検定とホテルリングの T^2 検定統計量とその分布を理解する。 6 欠測データ 3 2-step 単調欠測値データのもとでの平均ベクトルの検定を理解する。 7 欠測値データ 4 2-step 単調欠測値データのもとでのホテルリングの T^2 検定統計量の分布を理解する。 8 欠測データ 5 k-step 単調欠測値データのもとで平均ベクトルと分散共分散行列の最尤推定量の導出を理解する。 9 欠測値データ 6 k-step 単調欠測値データのもとでのホテルリングの T^2 検定統計量の分布を理解する 10 これまでの内容の到達度の確認と解説 11 平均ベクトルの検定 2 修正尤度比検定について理解する。 12 欠測値データ 7 単調欠測値データのもとでの修正尤度比検定統計量について理解する。 13 欠測値データ 8 単調欠測値データのもとでの修正検定統計量について理解する。 14 欠測値データ 8 2 標本および多標本問題の場合への拡張について理解する。

D-STEP 2019 年度講義シラバス

	15	本科目内容の到達度の確認と解説

D-STEP 2019 年度講義シラバス

講義名	最適化理論	
担当教員	矢部 博	
単位数	2 単位	
期間・曜日・時限	前期後半・火 4, 5 限	
科目区分	専門科目	
学習到達目標	1. 凸集合・凸関数について説明できるようになる。 2. 双対定理、最適性条件について説明できるようになる。 3. 線形計画問題を解くための主双対内点法について理解できるようになる。	
講義概要	最適化理論のうち、凸関数、線形最適化問題に対する数値解法、双対定理、最適性条件について学習します。	
成績評価方法	出席状況・レポート・試験結果を総合的に評価します。	
前提科目		
学習上・履修申請上の留意点 及び準備学習	数理計画法やオペレーションズ・リサーチに関する基礎知識があることが望ましい。	
教科書・参考図書等	教科書：矢部博 「最適化とその応用」, 数理工学社.	
備考		
授業計画	1	最適化問題 (1)最適化問題の例について学習する。
	2	最適化問題 (2)最適化問題の例について学習する。
	3	凸集合と凸関数 (1)凸集合・凸関数について学習する。
	4	凸集合と凸関数 (2)凸関数であるための必要条件・十分条件について学習する。
	5	凸集合と凸関数 (3)凸関数であるための必要条件・十分条件について学習する。
	6	線形計画問題 (1)単体法について理解できるようになる。
	7	線形計画問題 (2)双対定理について理解できるようになる。
	8	線形計画問題 (3)最適性条件、二者択一定理について理解できるようになる。
	9	線形計画問題 (4)内点法について理解できるようになる。
	10	線形計画問題 (5)中心化 KKT 条件と主双対内点法について理解できるようになる。
	11	線形計画問題 (6)実行可能点列主双対内点法について理解できるようになる。
	12	線形計画問題 (7)実行可能点列主双対内点法について理解できるようになる。
	13	線形計画問題 (8)実行可能点列主双対内点法について理解できるようになる。
	14	制約付き最適化問題 (1)等式制約条件付き最適化問題の最適性条件について理解できるようになる。
	15	制約付き最適化問題 (2)不等式制約条件付き最適化問題の最適性条件について理解できるようになる。

D-STEP 2019 年度講義シラバス

講義名	デザイン思考	
担当教員	佐藤 彰洋	
単位数	1 単位	
期間・曜日・時限	集中・前期	
科目区分	専門科目	
学習到達目標	<ul style="list-style-type: none"> ・デザイン思考のためのワークショップの設計方法 ・デザイン思考をデータ分析に利活用するための枠組み ・政府統計などのオープンデータを用いたワークショップ体験 	
講義概要	<p>デザイン思考は問題解決思考方法であり、人間中心設計思想に基づくユーザーニーズの満足度を重視する問題解決思考方法と位置付けられる。問題発見、問題解決、ニーズとシーズの分析、ユーザー体験のモデル化、プロトタイピング（試作）などを通じて試行錯誤的にたくさんのアイデアから問題解決に対するよい解の探索を行う方法である。本講義では、デザイン思考の発想方法と、デザイン技術、並びに、データ分析にデザイン思考を利活用するための枠組について講義形式で学習するとともに、政府統計などのオープンデータを用いたワークショップを体験することにより、学習した概念的枠組みや技法を実際に参加者とともに活用し体験を通じて学習することにある。</p>	
成績評価方法	<p>ワークショップの設計図 40% ワークショップの成果物 60% または ワークショップのワークフロー観察によるレポート 60%（ワークショップに参加できない場合）</p>	
前提科目	特になし	
学習上・履修申請上の留意点 及び準備学習	パソコンの基本操作方法、パワーポイント、エクセル	
教科書・参考図書等	Aki-Hiro Sato (ed.), "Applications of Data-Centric Science to Social Design: Qualitative and Quantitative Understanding of Collective Human Behavior," Springer, Tokyo, 2019	
備考		
授業計画	1	デザイン思考とは
	2	デザイン思考で利用するデザイン技術
	3	デザイン思考をデータ分析に利活用するための枠組み
	4	デザインワークショップ設計方法
	5	デザインワークショップ実習 1
	6	デザインワークショップ実習 2
	7	デザインワークショップ実習 3
	8	成果発表会
	9	
	10	
	11	
	12	
	13	
	14	
	15	