

授業科目名 (英文表記)	データサイエンスのための統計学 (Statistics for data science)		
単位数	2	授業形態	講義・演習
担当教員	大井 達雄		
開講	南紀熊野サテライト	区分	大学院
実施日・時間	第1回	5月23日(土)	10:30~16:00
	第2回	5月30日(土)	10:30~16:00
	第3回	6月27日(土)	10:30~16:00
	第4回	7月4日(土)	10:30~16:00
	第5回	7月11日(土)	10:30~16:00

**【授業のねらい・概要】**

ビジネスに限らず、さまざまな分野で統計的な思考力によって様々な課題を解決していく能力、すなわちデータサイエンス力の高い人材がもてられている。例えば、行政においても政策効果の測定に重要な証拠となる情報や統計などのデータを活用したエビデンス・ベースト・ポリシー・メイキング(EBPM)が推進されている。このような流れはますます重視される傾向にある。本講義ではデータサイエンスの知識やスキルを習得するために、事例を交えながら、ビジネスや行政の現場で使用される実践的なデータ分析や統計分析の手法を習得することを目的としている。

**【授業計画】**

**第1回 オリエンテーション、データサイエンスとは何か。**

データサイエンスが必要とされる背景、データサイエンスに求められるスキルや知識

〔予習〕 データサイエンスに関する情報収集を行うこと

〔復習〕 データサイエンスが進むことによって、私たちの社会がどのようになるか将来を予測すること

**第2回 e-Statを活用したデータ分析事例(1)**

Excelの機能説明(四則計算、関数、グラフの作成、データベースなど)、国勢調査を使用した分析

〔予習〕 国勢調査について事前に調べておくこと

〔復習〕 授業で習った分析方法を別のデータを使用して確認すること

**第3回 e-Statを活用したデータ分析事例(2)**

家計調査を使用した分析

〔予習〕 家計調査について事前に調べておくこと

〔復習〕 授業で習った分析方法を別のデータを使用して確認すること

**第4回 e-Statを活用したデータ分析事例(3)**

経済センサスを使用した分析

〔予習〕 経済センサスについて事前に調べておくこと

〔復習〕 授業で習った分析方法を別のデータを使用して確認すること

**第5回 統計GISによる分析(1)**

地図で見る統計(jSTAT MAP)の機能説明、操作方法の解説

〔予習〕 GISについて事前に調べておくこと

〔復習〕 地図で見る統計(jSTAT MAP)の操作方法について確認すること

**第6回 統計GISによる分析(2)**

地図で見る統計(jSTAT MAP)を使用した活用事例の説明、演習の実施

〔予習〕 地図で見る統計(jSTAT MAP)で分析したいデータを探してこること

〔復習〕 授業で習った分析方法を別のデータを使用して確認すること

**第7回 データ分析の基礎(1)**

データの可視化、度数分布表・ヒストグラムの作成方法の解説

〔予習〕 Excelの機能(四則計算やグラフの作成など)について再度復習すること

〔復習〕 授業で習った分析方法を別のデータを使用して確認すること

**第8回 データ分析の基礎(2)**

代表値と散布度の活用、計算方法や結果の解釈の解説

〔予習〕 Excelの機能(四則計算やグラフの作成など)について再度復習すること

〔復習〕 授業で習った分析方法を別のデータを使用して確認すること

### 第9回 データ分析の基礎(3)

2つのデータの関係性の把握(クロス集計表、および散布図の作成)

[予習] Excelの機能(ピボットテーブル)について予習すること

[復習] 授業で習った分析方法を別のデータを使用して確認すること

### 第10回 データ分析の応用(1)

相関係数・単回帰分析の計算方法、ならびに結果の解釈の説明、活用事例の紹介

[予習] 散布図の内容について再度学習すること

[復習] 授業で習った分析方法を別のデータを使用して確認すること

### 第11回 データ分析の応用(2)

時系列データの内容や分析手法の解説、予測事例の紹介

[予習] 時系列データの概念について事前に調べておくこと

[復習] 授業で習った分析方法を別のデータを使用して確認すること

### 第12回 確率分布

確率、ベイズの定理、離散型・連続型変数、二項分布、正規分布(標準正規分布)の解説

[予習] 確率の概念について事前に調べておくこと

[復習] 授業で習った分析方法を別のデータを使用して確認すること

### 第13回 統計的推定

標本抽出、標本分布、大数の法則、中心極限定理、点推定、区間推定、信頼度の解説、計算方法の説明

[予習] 正規分布について再度復習すること

[復習] 授業で習った分析方法を別のデータを使用して確認すること

### 第14回 統計的仮説検定

帰無仮説、対立仮説、有意水準、片側・両側検定、第1種・第2種の過誤、計算方法の説明

[予習] 区間推定について再度復習すること

[復習] 授業で習った分析方法を別のデータを使用して確認すること

### 第15回 まとめ、最終課題の作成と提出

[予習] これまでの授業の内容を整理すること

[復習] 最終課題を完成させ、提出すること。

#### 【到達目標】

データサイエンスの最近の動向を知り、同時にデータサイエンス社会に必要な能力をパソコンによる演習を通じて理解できることを目標としています。具体的には、データや統計の性質について説明することができること、データを統計的に記述することができること、ならびに推定および検定の考え方について説明することができることなどがあげられます。

#### 【成績評価の方法】

各回の講義で行った課題の提出(50点)および最終課題の提出(50点)により評価します。

#### 【教科書】

特に指定しません。主としてファイル(PDFやExcel)を配布します。

#### 【参考書・参考文献】

- ・大屋幸輔「コア・テキスト統計学 第2版」新世社
- ・生田目崇「マーケティングのための統計分析」オーム社
- ・山中馨, 天谷永, 望月雅光「EXCELで考える統計学」創成社

この他、授業中に適宜紹介する。

#### 【履修上の注意・メッセージ】

授業では、その時間のトピックの説明に続けて、PCを使用して例題や練習問題を解きます。受講生の理解度を見ながら、授業の進度やトピックの一部を調整することがあります(そのつど説明します)。また、PCの操作の習熟もかねて復習も必ず行うようにしてください。特に授業中に理解できなかった箇所については、何度も復習してください。授業内容や個人差によって変わりますが、予習に1時間、復習に1時間程度の学習を計画してください。

#### 【履修をする上で必要な事項】

表計算ソフトMicrosoft Excelの基本的な操作ができることが必要です。また講義で使用しますので、Microsoft Excelがインストールされたパソコンを準備してください。

#### 【授業時間外学修】

授業時間外学習(予習・復習)については授業計画を参照のこと。