

授業シラバス



詳細情報

2019-7030000459-01

2019-02-12 10:42:28

科目名「クラス名」(サブタイトル)	期別	単位数	開講年次
情報系のための確率・統計	前期	2	3
担当教員	鶴田 直之、高橋 伸弥		
授業形態	実務経験	科目水準	試験実施

概要

「センサーネットワーク」や「SNS」、「IoT」と、インターネット上のデータ量を増やす仕組みが次々に登場している。同時に、これらの大量データを有効活用する試みが活発化し、「データマイニング」や「ビッグデータ」、「オープンデータ」と言うキーワードを生み出している。更には、Deep Learningを火付け役として機械学習や人工知能が再び脚光を浴びる時代になった。これらの一連の流れの基礎にあるものは確率・統計である。このような背景から、確率・統計処理の分析ができる人材の育成が急務になっており、その重要性は小学校から高校までの教育指導要領で統計が強化されたことからもうかがい知ることができる。

本講義では、将来、コンピュータを使って確率・統計処理をする必要が出てきたときに、入門的な内容であれば自学自習ができる程度に、基礎的な数学の諸概念を習得しておくことを目標とする。したがって、高度な証明や導出および確率分布には立ち入らない一方、多変量解析やクラスタリング、識別をはじめとする情報科学の応用技術を具体例としながら数学的な基礎知識を習得する。

到達目標 [▶▶詳細はこちら](#)

近年の確率・計学を応用した情報処理の事例を踏まえ、「データマイニング」、「ビッグデータ」、「人工知能」の用語を用いて確率・統計学の重要性を説明できる（授業計画の1回）。（知識・理解）

基礎的な統計量（平均と分散、標準偏差、中央値）を計算でき、基本的な確率分布（2項分布、ポアソン分布、正規分布、多変量正規分布）の形を知っており、母数パラメータの算出ができる。また、検定の基礎を理解しており、Rのライブラリを使って実例を解くことができる（授業計画の2、3回）。（知識・理解）

多変量の問題に対して、変量間の共分散と相関係数を計算できる。また、これらの応用である主成分分析と潜在的意味解析を理解しており、Rのライブラリを使って実例を解くことができる（授業計画の4～6回）。（知識・理解）

多変量のデータに対して、基本的なクラスタリングの手法（階層的クラスタリング、k-means法、Spectral Clustering）を知っており、Rのライブラリを使って実例を解くことができる（授業計画の7、8回）。（知識・理解）

授業計画		
回	授業計画	学習目標
1	イントロダクション	応用事例の紹介の講義を受け、データ収集、分析、結果の活用という一連の流れを理解する。また、確率・統計学の重要性を理解する（鶴田）
2	基礎的な統計量と確率分布	平均と分散、標準偏差、中央値の復習をし、2項分布、ポアソン分布、正規分布、多変量正規分布を理解し、最尤推定法を用いて母数（パラメータ）の算出を演習する（鶴田）
3	仮説検定	仮説検定の目的と種類について理解し、Excelを使ったt検定の演習を行い、活用方法を理解する（鶴田）
4	多変量データ（その1）	多変量データに対して、独立性、条件付確率、共分散と相関係数の意味を理解して、算出を演習する（鶴田）
5	多変量データ（その2）	固有値と固有ベクトル、主成分分析について理解し、主成分の算出を演習し、結果の読み方、最小二乗法との関係について考える（鶴田）
6	潜在的意味解析	特異値分解と潜在的意味解析について理解し、Rのライブラリを使った簡単なテキストマイニングを演習する（鶴田）
7	クラスタリング（その1）	階層的クラスタリングの手法を理解し、Rのライブラリを使って実例を解くこと演習をする（鶴田）
8	クラス	k-means法およびSpectral

多変量のデータに対する統計的な識別法および学習について確率分布を仮定するパラメトリックな方法と確率分布を仮定しないノンパラメトリックな方法（ニューラルネットワークを含む）の原理を理解しており、Rのライブラリを使って実例を解くことができる（授業計画の9～11回）。（知識・理解）

事後確率に基礎を置くベイズ統計によりモデル化や予測、シミュレーションを行う一連の処理の流れを理解しており、Rのライブラリを使って実例を解くことができる（授業計画の12～15回）。（知識・理解）

✦ 授業時間外の学習（予習・復習）

ほぼ毎回、e-learning（項目URLを参照）を用いて予習復習の課題を与えるので、受講に当たっては、しっかり時間（2時間）を割いて取り組むこと。その際、授業の前後に級友と授業に関する話をした時間、興味を持って調べものをしてきた時間、試験対策の勉強をした時間も予習復習の活動と考え、平均すると2時間程度になる分量を積極的に行うこと。

✦ 成績評価の基準

1. 近年の確率・計学を応用した情報処理の事例を踏まえ、「データマイニング」、「ビッグデータ」、「人工知能」の用語を用いて確率・統計学の重要性を説明できる
2. 簡単な例題に対して適切な検定方を使い分けて行うことができ、原理や得られた結果を説明できる
3. 多変量の問題に対して主成分分析と潜在的意味解析を行うことができ、原理や得られた結果を説明できる
4. 多変量のデータに対して階層的クラスタリング、k-means法、Spectral Clusteringを行うことができ、原理や得られた結果を説明できる
5. ベイズ統計によるモデル化や予測、シミュレーションを行うことができ、原理や得られた結果を説明できる

✦ 成績評価の方法

定期試験（60%）と平常点（40%）で評価する。成績評価の基準に挙げた統計処理を行えるかどうか、得られた結果を説明できるかどうかは、授業中の演習とほぼ毎回の予習復習課題にて評価し、これらを平常点とする。原理を説明できるかどうかは定期試験により評価する。

✦ JABEE学習・教育到達目標

(B-1) 数学と物理学を中心とする自然科学についての基礎知識を持ち、実際の問題に応用する能力に対応している。

	タリ グ（そ の2）	Clusteringの手法を理解し、Rのライブラリを使って実例を解くこと演習をする（鶴田）
9	識別 （その 1）	ベイズの決定則と正規分布を仮定したときの事後確率の計算方法について理解し、簡単な2クラスの識別問題を演習する（鶴田）
10	識別 （その 2）	確率分布を仮定しない識別法の体系と、その代表例であるサポートベクタマシンについて理解し、簡単な2クラスの識別問題を演習する（鶴田）
11	識別 （その 3）	確率分布を仮定しない識別法であり、近年、再注目を集めているニューラルネットワークの原理と動向について理解する（高橋）
12	ベイズ 統計 （その 1）	ベイジアンネットワークとその応用例について理解し、Rを使った学習と予測を演習する（鶴田）
13	ベイズ 統計 （その 2）	ベイズ統計とベイズの定理との関係について理解する（高橋）
14	ベイズ 統計 （その 3）	確率分布に従う乱数を生成する方法、およびMCMC法について理解する（鶴田）
15	ベイズ 統計 （その 4）	ベイズ統計によるモデル化や予測、シミュレーションの一連の分析な流れについて理解し、簡単な例題をRを使って演習する（鶴田）

この科目の授業時間数は、試験時間を含めて23.5時間である。各教員が担当する回および回数は、都合により変更することもある。

✦ URL

[Moodleシステムサービス
\(https://moodle.cis.fukuoka-u.ac.jp/\)](https://moodle.cis.fukuoka-u.ac.jp/)

✦ テキスト

授業中に講義資料を配布するのでテキストを購入する必要はない。講義資料はMoodleシステムからダウンロードできる。

✦ 参考書

ほぼ毎回の予習復習課題は学内のMoodleシステムサービスを用いる。



Copyright (C) 2006 Fukuoka University. All Rights Reserved.