

授業科目名 Course Name	数理・データサイエンス特論 Advanced Studies in Mathematics and Data Science	教員名 Course Instructor(s)	渡邊 耕二 Koji Watanabe 保田 昌秀 Masahide Yasuda
		Eメールアドレス E-mail	渡邊 kwatanabe@miu.ac.jp 保田 myasuda@miu.ac.jp
授業形態 Class Format	講義 Lecture	オフィスアワー Office Hours	講義終了後 After lectures
科目番号 Course Code	LAI609	担当形態 Mode of Instruction (Solo / Omnibus / Jointly)	オムニバス Omnibus
単位数 No. of Credits	2	配当年次 Allocated Year	1年前期 Spring, 1st Year
		卒業要件 Required or Elective to Graduate	選択 Elective
到達目標 Goals	1. データの種類（分類）、データの特徴を示す指標の把握方法を理解している。 2. データに対してのグラフの選び方が適切である。 3. 2変数データとグラフの書き方が一致している。 4. エクセルでの解析やグラフ表現が習得できている。 5. データの評価方法が適切である。 1. Understand the types (classifications) of data and how to grasp indicators that indicate the characteristics of data. 2. How a graph is chosen for the data is appropriate. 3. The variable data and the graph can be drawn in the same way. 4. Mastery of analysis and graph expression in Excel. 5. The data evaluation method is appropriate.		

授業の概要 Course Overview	情報通信技術や測定技術の発展によるビッグデータの出現など、データが溢れる時代となった。国際文化の多様性を理解する上で、データから有益な情報を抽出する知識と技術は増々重視される。本特論では、データの要約・可視化とそれを可能にする数理に関する講義、身近なソフトであるエクセルを用いたデータ分析の実践、数理・データサイエンスについて学ぶ。 With the advent of big data due to the development of information and communication technology and measurement technology, we have entered an era of overflowing data. Knowledge and skills to extract useful information from data are increasingly emphasized in understanding the diversity of international culture. In this advanced studies course, we will have a lecture on data summarization / visualization and the mathematics that makes it possible, practice data analysis using a familiar software Excel to learn about mathematics and data science.
ディプロマ・ポリシーとの関係 Diploma Policy Objectives	DP1-2 国際コミュニケーションの観点から専門的な研究を行うために必要な学術的思考力を身に付け、国際社会の課題や諸問題を理解し、論理的かつ批判的に分析する能力や問題・課題の解決に向けて提言・実行する能力を有する。 Acquire the academic cogitative skills necessary to conduct specialized research from the perspective of international communication, understand and logically as well as critically analyze issues and topics of the international community, and has the ability to make proposals and take action to solve these issues. DP2-2 現代のグローバル社会のさまざまな課題に対し学際的研究ができる能力を有する。 Acquire the ability to conduct interdisciplinary research on a range of issues that exist in the modern global society.
履修条件・注意事項 Prerequisites /Remarks	学部教育において基本的な統計学が習得していることが望ましい。 It is desirable to have studied basic statistics in undergraduate level education. 大学院において「数理・データサイエンス特論」を受講していることが望ましい。 It is desirable to be taking "Advanced Topics in Mathematical and Data Science" in graduate school.
授業計画 Course Schedule	(渡邊 耕二/ Koji Watanabe) 前半部分 1. データの種類（名義尺度、順序尺度、間隔尺度、比率尺度）とデータの可視化（ヒストグラム、累積相対度数、ローレンツ曲線）について解説する。 Types of data (nominal scale, ordinal scale, interval scale, ratio scale) and data visualization (histogram, cumulative relative frequency, Lorenz curve) will be explained.

2. データの分布の特徴を表す指標（平均値、中央値、最頻値、分散、標準偏差、変動係数、範囲、四分位範囲）と1変数データの要約・グラフ表現（モザイクグラフ、箱ひげ図）について解説する。

Indicators (mean, median, mode, variance, standard deviation, coefficient of variation, range, quadrant range) that represent the characteristics of the data distribution and a summary / graph representation of one-variable data (mosaic graph, boxplot) will be explained.

3. 2変数データの要約（共分散、相関係数）とグラフ表現（クロス集計表、散布図）について解説する。

Summary of variable data (covariance, correlation coefficient) and graph representation (crosstabulation table, scatter plot) will be explained.

4. 2変数データの要約として、相関係数と単回帰分析について解説する。

As a summary of variable data, the correlation coefficient and simple regression analysis will be explained.

5. 事象と確率、確率変数、確率分布（期待値と分散、正規分布、二項分布）について解説する。

Events and probabilities, random variables, and probability distributions (expected value and variance, normal distribution, binomial distribution) are explained.

6. データ分析と予測を汎用ソフトであるエクセルとRを使った実習で理解を深める。条件が同じでも、個々の特性で変化する変量は、分布を持っている。テストの成績、実験データ、観測値などがこれにあたる。そこで、エクセルを使って変量を解析する方法について講義する。最初に、得られた変量から度数分布（ヒストグラム）・標準偏差・平均値を計算する方法について解説する。

Deepen understanding of data analysis and prediction using a general-purpose software Excel and R in practical training. Variables that vary with unique characteristics, even under the same conditions, have a distribution. This includes test results, experimental data, and observations. Therefore, a lecture on how to analyze variables using Excel will be given. First, how to calculate the frequency distribution (histogram), standard deviation, and mean value from the obtained variables will be explained.

7. 代表的な分布である正規分布について、エクセルとRを使って平均値・標準偏差を変化させた時の分布のシミュレーションを行う。

For a normal distribution, which is a typical distribution, use Excel and R to simulate the distribution when the mean and standard deviation are changed.

8. 母集団の変量から無作為抽出で得た標本（サンプル）を使った検定方法について解説をする。

Test method using a sample obtained by random sampling from variables of the population will be explained.

(保田昌秀)

9. 汎用的データ解析ソフトエクセルを使ったデータ分析方法を学ぶ。例として、条件にあったデータだけを抽出するIf、Countif等について学び、さらに度数分布を計算するFrequency関数について学ぶ。

Learn how to analyze data using Excel, a general-purpose data analysis software. For example, learn about If and Countif, which extract only data that meet certain conditions, and the Frequency function, which calculates frequency distribution.

10. データ分析と予測を汎用ソフト：エクセルを使った実習で理解を深める。条件が同じでも、個々の特性で変化する変量は、分布を持っている。テストの成績、実験データ、観測値などがこれにあたる。そこで、エクセルを使って変量を解析する方法について講義する。最初に、得られた変量から度数分布（ヒストグラム）・標準偏差・平均値を計算する方法について解説する。

Deepen understanding of data analysis and prediction using a general-purpose software Excel in practical training. Variables that vary with unique characteristics, even under the same conditions, have a distribution. This includes test results, experimental data, and observations. Therefore, a lecture on how to analyze variables using Excel will be given. First, how to calculate the frequency distribution (histogram), standard deviation, and mean value from the obtained variables will be explained.

11. 代表的な分布である正規分布について、エクセルを使って平均値・標準偏差を変化させた時の分布のシミュレーションを行う。

For a normal distribution, which is a typical distribution, use Excel to simulate the distribution when the mean and standard deviation are changed.

12. 母集団の変量から無作為抽出で得た標本（サンプル）を使った検定方法について解説をする。

Test method using a sample obtained by random sampling from variables of the population will be explained.

13. 条件を変化させて得られる変量について、エクセルを使って条件と変量の二つの変量間の相関分析（回帰直線、相関係数）を行う。

For the variables obtained by changing the conditions, Excel is used to perform correlation analysis (regression line, correlation coefficient) between the two variables, the conditions and the variables.

14. 研究として行ってきた「光増感剤を使った酵母菌・大腸菌・がん細胞などの死滅実験（光線力学療法）の結果」を通じて、データ分析の応用例を紹介する。

Through data analysis on "Results of inactivation experiments (photodynamic therapy) of yeast, *E. coli*, cancer cells, etc. using photosensitizers" which was conducted as a project, an

	application example of data analysis is discussed. 1 5. 用意したデータ（銀行の融資額、気象庁の気象データ、小学校児童数の変遷）などを使って、エクセルとワードによる報告書作成について学ぶ。 Learn how to create reports in Excel and Word using prepared data (bank loan amounts, weather data from the Japan Meteorological Agency, changes in the number of elementary school students, etc.).
学生に対する 評価 Assessment Criteria	講義内容に関するレポートにより評価する。 Evaluated by reports (5 times) which are on the content of the lectures.
時間外の学習 について Preparation and Revision outside Class	復習を必ず行う事。また、講義に関連する事前学習と講義後のレポート作成を行うこと。 Revision is essential. In addition, pre-study of lectures and report writing after lectures are required.
テキスト Textbooks	講義前に関係テキストや資料を配布する。 Relevant texts and materials will be distributed before the lecture.
参 考 書 ・ 参考資料等 References	「Rによるやさしい統計学」、山田剛史・杉澤武俊・村井潤一郎、オーム社 「わかりやすい統計学データサイエンス基礎」、松原望・森本栄一著、丸善出版 「科学計測のための波形データ処理」：南 茂夫：CQ出版 「人工知能概論」谷口忠大、講談社