

<<最終更新日：2024年02月27日>>

English

基本情報

時間割コード／Course Code	290729
開講区分(開講学期)／Semester	集中
曜日・時間／Day and Period	他
開講科目名／Course Name (Japanese)	データ科学特論 II
開講科目名(英)／Course Name	Advanced Course in Data Sciences II
定員／Capacity	999
教室／Room	
ナンバリング／Course Numbering Code	29CHEN5M204,29SSAI5M204,29MASC5M204,29MSS5M204
授業形態／Type of Class	講義科目
単位数／Credits	2.0
年次／Student Year	1,2年
担当教員／Instructor	鈴木 讓,内田 雅之
メディア授業科目／Course of Media Class	該当（学部学生がメディア授業科目を卒業要件に算入できるのは60単位が上限です。）

※メディア授業科目について

授業回数の半数以上を、多様なメディアを高度に利用して教室等以外の場所で行う授業を「メディア授業科目」としています。

学部学生が「メディア授業科目」を卒業要件に算入できるのは60単位が上限です。

なお、非該当の場合であっても、メディアを利用した授業を実施する場合があります。

[授業担当教員一覧](#)

詳細情報

授業サブタイトル／Course Subtitle	WAIC/WBICの数理 (渡辺澄夫ベイズ理論)
開講言語／Language of the Course	日本語
学習方法／Learning Methods	2024年9月に開講するオンライン講義である。Stanというプログラミング言語の操作を学習してから、統計学および代数幾何の知識を得て、WAIC/WBICの本質を理解する
授業の目的と概要／Course Objectives	渡辺澄夫氏は、正則ではない統計モデルに関する情報量基準WAICを提案した。それが正則でない場合でもAICと同等の性能を有することを示すために、独自のベイズ理論を構築した。正則でない場合、最尤推定が収束しないので、AICやBICといった既存の情報量基準が適用できない。本講義では、渡辺氏を始め、この分野で活躍している有数の研究者の方に解説をお願いしている。

9/8(月) 3-5限 13:30-18:20 鈴木讓
9/9(月) 3-5限 13:30-18:20 鈴木讓
9/10(火) 1限8:50-10:20 Andrew Gelman
(Columbia University)
9/10(火)2限10:30-12:00 伊庭幸人
9/10(火)3限13:30-15:00 二宮嘉行
9/12(木)3限13:30-15:00 徳田悟
9/12(木)4限15:10-16:40 矢野恵佑
9/12(木)5限16:50-18:20 Mathias Drton
(Technische Universität München)
9/13(金)3限13:30-15:00 渡辺澄夫
9/13(金) 4限15:10-16:40 *車谷優樹
9/13(金) 5限16:50-18:20 青柳美輝

* 博士課程学生。鈴木讓教授のTFとして話題を提供する

学習目標／Learning Goals

- 1 ベイズ統計学、情報量基準に関する高度な議論を通して、データサイエンスの様々な側面で自信をもって対応できるようになる(難しいと思わなくなる)。

履修条件・受講条件／Requirements, Prerequisites

出欠席及び受講に関するルール／Attendance and Student Conduct Policy

授業計画／Class Plan

題目:渡辺澄夫ベイズ理論の全容

「渡辺澄夫ベイズ理論 with R/Python」第0章、第1章

第1回 担当教員：鈴木 讓

授業時間外学習：講師の話をよく復習して、質問を提出する(課題提出になる)。

題目:Stanによる実装

「渡辺澄夫ベイズ理論 with R/Python」第2章

第2回 担当教員：鈴木 讓

授業時間外学習：講師の話をよく復習して、質問を提出する(課題提出になる)。

題目:正則の場合の一般論

「渡辺澄夫ベイズ理論 with R/Python」第4章

第3回 担当教員：鈴木 讓

授業時間外学習：講師の話をよく復習して、質問を提出する(課題提出になる)。

題目:代数幾何とゼータ関数

「渡辺澄夫ベイズ理論 with R/Python」第6章

第4回 担当教員：鈴木 讓

授業時間外学習：講師の話をよく復習して、質問を提出する(課題提出になる)。

題目:WAICの一般理論

「渡辺澄夫ベイズ理論 with R/Python」第7章

第5回 担当教員：鈴木 讓

授業時間外学習：講師の話をよく復習して、質問を提出する(課題提出になる)。

題目:学習係数を求める

「渡辺澄夫ベイズ理論 with R/Python」第8章

第6回 担当教員：鈴木 讓

授業時間外学習：講師の話をよく復習して、質問を提出する(課題提出になる)。

第7回 題目:"effective Number of Parameters in a Statistical Model"

Degrees-of-freedom adjustment for estimated parameters is a general idea in small-sample hypothesis testing, uncertainty estimation, and assessment of prediction accuracy. The effective number of parameters gets interesting in the presence of nonlinearity, constraints, boundary conditions, hierarchical models, informative priors, discrete parameters, and other complicating factors. Many open questions remain, including: (a) defining the effective number of parameters, (b) measuring how the effective number of parameters can depend on data and vary across parameter space, and (c) understanding how the effective number of parameters changes as sample size increases. We discuss using examples from demographics,

imaging, pharmacology, political science, and other application areas.

担当教員：Andrew Gelman
(Department of Statistics,
Columbia University)

授業時間外学習：講師の話をよく復習して、質問を提出する(課題提出になる)。

題目:計算統計の立場からベイズ統計と頻度論を結ぶ — Bayesian IJKとその周辺

ベイズ的な手法を利用する場合、観測値の背後に頻度論的な母集団を想定して、解析結果の安定性やモデルの妥当性を論じることが行われるが、そのために必要な情報が事後共分散（一般に事後キュムラント）に含まれていることが明らかになってきた。WAICのバイアス補正項の形はそのひとつの現れと考えられるが、本講義では、この流れにそって、事後期待値の感度解析、

第8回

WAICやPCICなどの情報量規準、ブートストラップ法に対する近似、事後期待値の頻度論的共分散を表現する公式を解説する。また、講演者の最近の研究の紹介として、各観測の対数尤度の事後共分散からなる行列が上記との関連で重要な意味を持ち、これを再生核としたカーネルPCAがベイズ推定量の漸近論の理解に有用なことを示す。

担当教員：伊庭幸人 (統計数理研究所)

授業時間外学習：講師の話をよく復習して、質問を提出する(課題提出になる)。

第9回 題目:WAIC に基づく予測情報量規準の新たな展開

近年 WAIC がベイズ解析におけるモデル選択の標準となっていることを踏まえ、その改良の余地を探る。一つの方向として、事前分布の存在を強調するような漸近論を用いることを考える。また、それに基づき、事前分布のクラスが与えられたときに最適な事前分布を与えることを考える。

担当教員：二宮嘉行 (統計数理研究所)

授業時間外学習：講師の話をよく復習して、質問を提出する(課題提出になる)。

題目:回帰問題におけるスケール関係

**第10
回**

回帰問題を題材に、逆温度、ノイズ分散、サンプルサイズをスケール因子とする有限サイズスケール関係式を紹介する。これにより、自己平均性が成り立つが一致性は成り立たない極限において現れる、ベイズ的モデル選択の典型性を示す。渡辺澄夫ベイズ理論との関係についても触れる。

担当教員：徳田悟 (九州大学)

授業時間外学習：講師の話をよく復習して、質問を提出する(課題提出になる)。

題目:さまざまな予測状況におけるWAIC

**第11
回**

近年、尤度以外の損失を用いる擬ベイズやベイジアンニューラルネットワークといった深層学習との融合など、ベイズ学習を取り巻く環境は多様化している。本講義では、そのような多様な設定でのWAICの適用に向けて、WAICの擬ベイズ・重み付きベイズ推論への拡張、そして高次元(特に過剰パラメタモデルと呼ばれる深層学習を解明するためのモデル)での振る舞いを講義する。

担当教員：矢野 恵佑 (統計数理研究所)

授業時間外学習：講師の話をよく復習して、質問を提出する(課題提出になる)。

**第12
回**

題目:A Bayesian Information Criterion for Singular Models

The classical Bayesian information criterion (BIC) of Schwarz forms a trade-off between the fit and the complexity of a parametric statistical model by means of the dimension of model. The specific trade-off is derived through asymptotic analysis of Bayesian marginal likelihood. The BIC is commonly employed to address model selection problems across

various applications and is featured in numerous software packages. However, some of these applications involve models whose Fisher information matrices may fail to be invertible along other competing submodels. In such singular models, the refined concept of learning coefficients captures the asymptotic behavior of the marginal likelihood. We present a practical generalization of Schwarz's criterion that leverages theoretical knowledge of learning coefficients in the definition of an information criterion. The method is presented in the context of examples such as determining the number of components in mixture models, the number of factors in latent factor models or the rank in reduced rank regression.

担当教員：Mathias Drton (Technical University of Munich)

授業時間外学習：講師の話をよく復習して、質問を提出する(課題提出になる)。

題目:深層学習と特異学習理論

多層神経回路網は、今日の人工知能の実現において中心的な役割を果たしており、社会や産業への影響が大きくなるにつれて、その適切な設計法の確立が望まれるようになってきている (AIアライメントの問題)。しかしながら、深層学習では、入力と出力が極めて多くの複雑な階層的な構造で結ばれているため、従来の統計的手法だけでは十分な解析を行うことができなかった。特異学習理論は、非線形かつ非正則な学習モデルの数学的な性質を解明するために構成され、現在、AIアライメントの課題の中でどのような役割を果たしているかについて研究が進められている。この講義では、そうした研究の現況と将来の可能性を紹介する。

担当教員：渡辺澄夫 (東京工業大学)

授業時間外学習：講師の話をよく復習して、質問を提出する(課題提出になる)。

第13 回

題目: Blow-upを用いた学習係数の計算例

第14回

ベイズ推測における予測分布の精度を測る評価指標として汎化損失や自由エネルギー等が挙げられるが、これらの漸近挙動を調べる上で、学習係数（実対数閾値）という統計モデルごとに定まる有理数により決定されることが明らかとなっている。本講義では、この学習係数の求め方の1つとして代数幾何学のblow-upという手法を紹介し、混合分布モデルといった統計モデルの学習係数の導出例を紹介する。

担当教員：車谷優樹

授業時間外学習：講師の話をよく復習して、質問を提出する(課題提出になる)。

題目: 特異点解消と学習係数

第15回

学習係数は、特異学習理論において重要な役割を果たす。この講義では、学習係数を理論的に求めるための特異点解消について主に説明する。特に、射影空間やブローアップ、トーリック多様体の概念などについて紹介する。また、近年、得られた、多層線形ニューラルネットワークの学習係数などについて、講義を行う。

担当教員：青柳美輝 (日本大学)

授業時間外学習：講師の話をよく復習して、質問を提出する(課題提出になる)。

鈴木讓「渡辺澄夫ベイズ理論 with R/Stan」(共立出版, 2023年9月)

鈴木讓「渡辺澄夫ベイズ理論 with Python/Stan」(共立出版, 発行予定)

Joe Suzuki, WAIC/WBIC with R Stan, Springer (2023)

Joe Suzuki, WAIC/WBIC with Python Stan, Springer (2023)

教科書・指定教材/Textbooks

参考図書・参考教材/Reference

成績評価/Grading Policy

※学習目標の番号にカーソルをあてると、その学習目標の全文が表示されます。

評価方法/ 学習へ

Evaluation の参加

Methods 度

学習目標1 ○

評価割合 (%) / 100%

Allocation of Marks

成績評価に関する補足情報 / Additional Information on Grading

各担当教員の講義に対する質問をする。その内容によってその回のスコアとし、その合計点で決める。

大学院等高度副プログラム「データ科学」のコア科目

特記事項 / Special Note

障がい等により本科目の受講に際し特別な配慮を要する場合は、基礎工学研究科の大学院係に事前に相談するとともに、初回授業等、早期に授業担当教員に申し出てください。

オフィスアワー / Office Hours

集中講義のため、オフィスアワーはなし

実務経験のある教員による授業科目 / Course Conducted by Instructors with Practical Experience

授業担当教員

教員氏名 / Instructor Name

ふりがな / Name (hiragana)

データがありません

学生への注意書き