

統計学と最適化 (第1週)

講義概要

2021年4月3日

田辺良則 (鶴見大 / NII)

講義予定

- スケジュール
 - 第1回 (4/3 10:30-12:00) ツール・環境の確認・記述統計・データの可視化
 - 第2回 (4/3 13:00-14:30) 母集団と標本, 確率分布
 - 第3回 (4/3 14:45-16:15) 推定・統計的仮説検定
 - 第4回 (4/3 16:30-18:00) 演習
 - 第5回 (4/10 10:30-12:00) 標準線形モデル・
 - 第6回 (4/10 13:00-14:30) 一般線形モデル・リサンプリング
 - 第7回 (4/10 14:45-16:15) 最適化
 - 第8回 (4/10 16:30-18:00) 演習

質問対応

- 講義中，任意のタイミングで質問して良い。
 - マイクをオンにして，講師の話に割り込む。
 - チャットで質問しても構わないが，講師が気づくまでに時間がかかる。
 - ぜひ質問してください。
- 講義後の質問は，LMSの「講義用フォーラム」もしくはメールで行う。
 - 受講生間で共有できるので，できればフォーラムで質問してほしい。
 - 講義用フォーラムの「購読」を勧める。
 - 直接メールで質問しても良い: so@topse.jp

レポートと成績

- 成績は課題レポートの評価による。
 - 60%以上の得点で合格
 - 課題レポートは全2回
 - 提出点 + 評点 (比重は 4:6)
- レポートの提出
 - LMSにアップロード
 - 締切までは、何度でも再アップロード可能
 - 締切後なるべく受け付ける (提出点は下がる) が、保証はできない
- 原則はフィードバック無し
 - 解答が提出された問題には、解答例を配布
 - 明示的な質問が書いてあれば回答する
 - いずれも、時間がかかる。急ぐ質問はフォーラムまたはメールで。

講義概要

- 「データサイエンス」シリーズの諸講義を履修する上で必要となる，統計学及び最適化に関する基本的な知識を学習する
- 数学的な証明の詳細はあまり扱わない
- 実際にデータを操作する
- 想定する典型的な受講生像
 - ソフトウェアに関する業務をしているが，データサイエンスもこれから学んでいきたい。
 - 統計学は学生時代に少し勉強した程度 (あるいは全く勉強していない)
 - 高校では理系の数学を勉強したが，大学では入門程度の数学 (微積分・線形代数) 以上には勉強していない。

演習環境

- 主としてPythonを用いる.
- トップエスイーのEC2インスタンスを利用可
- 自分のPCにローカルに環境を構築することを勧める

参考書

- 特に推薦するというほどのものではないので、参考まで。
 - この講義の内容を理解するにはこれらで十分
- 馬場 真哉「Pythonで学ぶあたらしい統計学の教科書」(翔泳社)
- 山田 剛史, 杉澤 武俊, 村井 潤一郎「Rによるやさしい統計学」(オーム社)
- Peter Bruce, Andrew Bruce, Peter Gedeck「データサイエンスのための統計学入門 第2版」(オライリージャパン) (翻訳はあまり褒められない)
- 東京大学教養学部統計学教室「統計学入門 (基礎統計学I)」(東京大学出版会)
- 東京大学教養学部統計学教室「自然科学の統計学 (基礎統計学III)」(東京大学出版会)
- 川端 一光, 岩間 徳兼, 鈴木 雅之「Rによる多変量解析入門 - データ分析の実践と理論」(オーム社)