

科目名：データサイエンス活用の基礎

単位：2

（以下、予定）

■担当者

神永正博、木下勉

■テーマ

統計言語 R の実習を通して一変量統計学・多変量統計学の基礎を学ぶ

■講義内容

Society 5.0 の実現に不可欠なデータサイエンスの基礎を理解し、実際のデータを用いながら活用方法を身につける。データサイエンスの基礎として、一変量統計学から多変量統計学を学ぶ。統計言語 R による実習も取り入れながら議論する。

■達成目標

- （1）R を使って、エクセルファイルなどからのデータの取り込み、データの加工、データの可視化、データの観察、基本統計量の計算ができる
- （2）社会データ等で頻出する分割表の分析ができる
- （3）与えられたデータの単回帰および重回帰分析、一般化線形モデル（ロジスティック回帰等）のあてはめを行い、結果を適切に解釈できる

■授業計画

- 第1回 講義ガイダンス、R, R studio の紹介、R 入門、Society 5.0、社会とデータサイエンス
- 第2回 多変量データの記述
- 第3回 様々な確率分布
- 第4回 正規分布と指数分布と大数の法則と中心極限定理
- 第5回 点推定（不偏推定、最尤推定）
- 第6回 正規分布、特に確率点と区間推定、統計的仮説検定(t 検定)
- 第7回 分割表の検定1（カイ二乗検定、尤度比検定、フィッシャーの正確検定）
- 第8回 分割表の検定2（2 標本の比率の検定、残差分析）
- 第9回 単回帰分析（最小二乗法、決定係数）
- 第10回 多項式回帰分析と赤池情報量基準（AIC）
- 第11回 線形代数学における逆行列、余因子行列およびクラメル公式と重回帰分析1
- 第12回 重回帰分析2（多重共線性、重回帰分析のあてはまりの良さ）
- 第13回 二項分布と二項選択モデル（ロジスティック回帰、プロビット回帰）
- 第14回 ポアソン分布と計数データへの一般化線形モデルの適用（ポアソン回帰、負の二項分布による回帰）
- 第15回 総復習と確認テスト

■成績評価方法・基準

中間レポート 30%、確認テスト 70%で評価する。

■テキスト

神永正博・木下勉「R で学ぶ確率統計学（実データ分析編）」内田老鶴圃 2023