

講 義 名	データ解析 (Data Analysis)		
開 講 時 期	第 5 学期	単 位 数	2-0-0
担 当 教 員	下平 英寿 西 8 号館 (W) 7 階 707 号室 (内線 3219)		
<p>【講義のねらい】</p> <p>「R を用いたデータ解析入門」. 統計処理ソフトウェアである R を利用して実践的なデータ解析ができるようになること (R に含まれる関数を呼び出してデータ解析を実行する), 背後にある数学, 統計学, アルゴリズムを理解すること (自分自身で関数を記述し, それを用いてデータ解析を行う) を目標とする.</p> <p>【講義計画】</p> <ol style="list-style-type: none"> 1. イントロダクション (社会人口統計データ, バイオインフォマティクス) 2. 期待値, 大数の法則 (ポートフォリオ, ヒストグラム) 3. モンテカルロ法 (MCMC 法, ギブスサンブラ) 4. ベイズの定理 (画像復元) 5. 積率母関数, 中心極限定理 6. 確率モデル (正規混合分布) 7. 判別問題, 分類, 識別 (スパムメール判別) 8. パラメタ推定 (最尤推定) 9. EM アルゴリズム (教師無し学習) 10. 最尤推定量の性質 (クラメール・ラオの不等式, フィッシャー情報行列) 11. 検定と信頼区間 12. 線形回帰分析 (ボストン住宅価格) 13. ロジスティック回帰分析 (スパムメール判別のつづき, ニュートン法) 14. 主成分分析 (ボストン住宅価格のつづき) <p>【成績評価】</p> <p>レポート提出.</p> <p>【テキスト等】</p> <p>講義資料 (PDF 形式) を講義ウェブサイトから各自ダウンロードする.</p> <p>【履修の条件】</p> <p>理論的な理解を深めるためには確率と統計第一・第二を履修していることが望ましいが, 履修していなくても可能. R にシンタックスが類似の言語 (Java 等) の経験があればよい. R プログラミングの詳細は講義中で説明せず, アルゴリズムの説明をとおして多少説明する程度である. 詳細は講義ウェブサイトを参照</p> <p>http://www.is.titech.ac.jp/~shimo/class</p>			