

医学情報処理演習 (A・B クラス)

中澤 港

2004年10月18日

1 キーワード

データ解析, データの水準, 帰無仮説, 検定, 推定, 信頼区間

2 授業の目標および期待される学習効果

(目標) 統計ソフトウェアを使った実際のデータ解析を通して, 医学統計学の講義で学んだ統計手法を, よりしっかりと身に付ける

(学習効果) 実験データや調査データを適切に分析できる能力を身につけることが期待される

3 授業の概要

実験計画や調査デザインに適したデータ構造の設計を行い, 表計算ソフトで入力したデータを統計処理ソフトで適切な手法を用いて解析し, その結果を解釈する, という一連の流れを, さまざまな種類のデータについて繰り返し実践する。講義と演習を情報処理演習室で行う。

4 授業内容のレベル

教養教育でコンピュータ操作の基礎は身に付いているはずであり, 前期に「医学統計学」で統計学の基礎は理解しているはずなので, 実際の統計処理を実践することを主眼におく。

5 履修資格

医学部医学科2年生(「医学統計学」履修を前提とする)

6 関連授業科目

「医学統計学」(2年前期)

7 テキスト / 参考書

データ及び課題を掲載した講義資料を配布する。統計ソフトウェアとしてはフリーソフトの R を使用するので、以下のテキストが参考になるであろう。

- 中澤 港 (2003) 『R による統計解析の基礎』(ピアソン・エデュケーション) 1,800 円
- 間瀬 茂・神保 雅一・鎌倉 稔成・金藤 浩司 (2004) 『工学のための数学3 工学のための データサイエンス入門 フリーな統計環境 R を用いたデータ解析 』(数理工学社) 2,300 円
- 岡田昌史 (編) (2004) 『The R Book - データ解析環境 R の活用事例集 - 』(九天社)
- 渡辺利夫 (2004) 『フレッシュマンから大学院生までのデータ解析・R 言語』(ナカニシヤ 出版)

また、以下のウェブサイトが参考になるであろう。

<http://aoki2.si.gunma-u.ac.jp/R/>

<http://www.okada.jp.org/RWiki/>

8 授業の形式

臨床研究棟演習室のコンピュータを使って、サンプルデータを分析してみる。

9 評価

毎回、課題の回答を提出すること(可能な場合は専用の web サイトを作り、そこから submit する形にするが、基本的にはその場で印刷して提出してもらう)とし、その結果によって評価する。期末試験は行わない。

10 メッセージ

統計は、実際にデータを扱って初めてわかってくるものです。慣れるまでは大変かもしれませんが、やっているうちにわかってくると思います。

11 オフィスアワー

月曜～金曜・10:00～12:00, 電子メール nminato@med.gunma-u.ac.jp

12 授業の展開

- 第1回 データ入力(10月4日 5-8時限)
- 第2回 基本的な図示(10月18日 5-8時限)
- 第3回 記述統計(10月25日 5-8時限) 基本的な要約統計量を求める
- 第4回 標本統計量と母数推定(11月1日 5-8時限) 不偏推定量, 信頼区間, 自由度の概念などを含む
- 第5回 データの分布(11月8日 5-8時限) 正規確率プロット, 検定の概念, Shapiro-Wilk の検定を含む
- 第6回 平均値の差の検定(11月15日 5-8時限) ストリップチャートを含む。独立2標本の場合, 対応のある2標本の場合, Welch の方法
- 第7回 一元配置分散分析と多重比較(11月22日 5-8時限) 層別箱ヒゲ図を含む
- 第8回 相関と回帰(11月29日 5-8時限) 散布図や信頼区間つき回帰直線の図示を含む
- 第9回 計数データの解析と比率(12月6日 5-8時限) 度数分布図を含む
- 第10回 クロス集計(12月13日 5-8時限) モザイクプロットを含む。カイ二乗検定, フィッシャーの正確な検定, オッズ比, リスク比, 率比, カッパ係数
- 第11回 量的データのノンパラメトリックな分析法(1月17日 5-8時限) Wilcoxon の順位和検定, 符号付順位和検定, 順位検定, Kruskal-Wallis 検定, Spearman の順位相関係数を含む
- 第12回 一般化線型モデル(1月24日 5-8時限) 重回帰分析とロジスティック回帰分析を中心に。
- 第13回 生存時間解析(1月31日 5-8時限): カプラン・マイヤ推定量とコックス回帰を中心に。