

統計-1 (1)

高久雅生

masao@slis.tsukuba.ac.jp

2013年11月13日(水) 1・2時限

本日のお品書き

- 授業内容の解説
 - 概要： 何をやるか？
 - 予定： いつやるか？
 - 目標： 何を目指すか？
 - 評価
- 統計とは？
- データを図や表にまとめる

自己紹介

- 高久 雅生(たかくまさお)
- 職歴:
 - 2004-2008: 国立情報学研究所
 - 2008-2013: 物質・材料研究機構
 - 2013-: 筑波大学図書館情報メディア研究系
- 専門・関心:
 - 情報探索行動
 - 情報検索、電子図書館
 - 学術コミュニケーション
- 連絡先:
 - 研究室: 7D208
 - メール: masao@slis.tsukuba.ac.jp
 - Twitter: @tmasao

(授業概要をシラバスから)

数理統計学の基礎を講義する。

内容は、統計の意味、確率、二項分布、正規分布、統計的仮説検定、推定の考え方などである。

つまり？（要約）

- 数理統計学の基礎
- 以下の応用のための考え方を学ぶ：
 - 統計の意味
 - 確率
 - 二項分布
 - 正規分布
 - 仮説検定
 - 推定

(授業の概要をもう一度...)

数理統計学の基礎を講義する。

内容は、統計の意味、確率、二項分布、正規分布、統計的仮説検定、推定の考え方などである。

教育目標

- 統計的考え方のマスター
 - 統計量についての基本概念、正規分布
- 統計的仮説検定の考え方
- 少なくとも「統計でだまされない」、できれば「統計でだませる」ようになってほしい。

教育目標 (1)

- 統計的考え方のマスター: 統計量についての基本概念、正規分布
- 統計量:
 - どのようなものか？
 - どのような種類があるか？
 - 記述や表現の仕方？
- 正規分布
 - どのようなものか？
 - どのように応用しうるか？
- 上記をマスターすることを目指します。

教育目標 (2)

- 統計的仮説検定の考え方
- 「統計的仮説検定」
 - どのような概念か？
 - 「仮説」とその「検定」の意味が答えられる
 - 検定した結果が統計的にどのような意味を持つか答えられる
- 上記をマスターすることを目指します

教育目標 (3)

- 少なくとも「統計でだまされない」、できれば「統計でだませる」ようになってほしい。
- 「統計でだます?」「統計でウソをつく」
 - どのような事象か?
 - 具体例は?
- 「統計でだます」
 - たまたま間違えて扱うのではなく、分かっている扱いえるようになってほしい。

(脱線) データサイエンティスト

SPOTLIGHT ON BIG DATA

Spotlight

ARTWORK Tamar Cohen, Andrew J Buboltz
2011, silk screen on a page from a high school
yearbook, 8.5" x 12"

Davenport, Thomas H., and D. J. Patil:
“Data Scientist: The Sexiest Job of the
21st Century”. Harvard Business Review,
Vol.90, No.10, p.70-76, 2012

Data Scientist: *The Sexiest Job of the 21st Century*

**Meet the people who
can coax treasure out of
messy, unstructured data.**

by Thomas H. Davenport
and D.J. Patil

W

hen Jonathan Goldman ar-
rived for work in June 2006
at LinkedIn, the business
networking site, the place still
felt like a start-up. The com-
pany had just under 8 million
accounts, and the number was

(脱線2)ビッグデータ

- Twitter: 2.3億ユーザ [\[twitter.com\]](https://twitter.com)
 - 5億ツイート/日
- Facebook: 11億ユーザ [\[fb.com\]](https://fb.com)
 - 100+フレンド/ユーザ
- Google検索: 30兆ページ [\[searchengineland.com\]](https://searchengineland.com)
 - 1,000億検索/月

授業予定

- 11/13
 - データを図や表にまとめる
- 11/20
 - 平均・分散・標準偏差・モード・メディアン
- 11/27
 - 相関、回帰
- 12/4
 - 確率の考え方(1)
- 12/11
 - 確率の考え方(2)
- 12/18
 - 正規分布の使い方
- (冬休み:レポート出題あり)
- 1/8
 - 統計的仮説検定の考え方(1)
- 1/15
 - 統計的仮説検定の考え方(2)
- 1/22
 - 独立性の検定、母平均の検定、等平均仮説の検定
- 1/29
 - 推定
- 2/5? (試験:試験日は後ほどアナウンスします)

授業評価

重要

- 提出課題と試験により、成績評価を行います。
- なお、提出課題については、いずれの種類であっても、必ず提出すること。
 - やむをえず欠席した授業回があったとしても、授業資料を確認のうえ、必ず提出すること(たとえ遅れたとしても)。
 - 提出期限に遅れた課題分は提出点のみとなります。
 - 未提出課題がある場合、不合格(D判定)とします。
- 試験は秋C試験週間に実施予定。

教科書

小寺平治. ゼロから学ぶ統計解析. 講談社, 2002, 222p. (ISBN: 4-06-154656-2)

- 授業内の講義内容、予習・復習は、基本的に上記の教科書を用います。
- 必ず入手しておくこと。



授業外の学習

- 教科書に書いてある式・論理を注意深く読んで理解すること
 - 少なくとも教科書の演習問題は自分で解いてみること
- 与えられた課題に対して、課題を満足する問題を自分で設定し、解くこと
- 指定された文献を読むこと

関連する授業科目

- 専門基礎科目 – 【統計とその応用】
 - 量的調査法(2年次)
 - 多変量解析(2年次)
- 専門科目(知識科学主専攻) – 【知識発見】
 - 知識構造化法(3,4年次)
 - データマイニング(3,4年次)
- 卒業研究

(統計-2)...

出席票

- 毎回、必ず出席票の提出を求めます。
 - 気になる点や質問、コメントがあれば、できるだけ記入するようにしてください。
 - 質問については、できるだけ次の授業回にて答えるようにします。
- ※課題についての質問等、間に合わない場合は、直接聞くようにしてください。

授業資料

- 教科書をベースとして学習しますので、教示用スライドの配布は行いません。
- 教示用スライド以外に資料がある場合は随時、配布します。
- ただし、資料すべては筑波大学Moodleサイトにアップロードしておく予定ですので、授業外での学習に際してはそちらも参照してください。
 - <http://moodle.tsukuba.ac.jp>
- 欠席した場合でも、前回以前の資料を授業担当者が持参することはありませんので、各自で授業サイトを閲覧参照しておくこと。

連絡先(コンタクト)

- この授業科目について疑問・質問・意見などあるときは、オフィスアワーに研究室に直接来てください。
 - オフィスアワー: 木曜日5時限
 - 7D 208研究室
- または、メールにてアポイントメントを取ること
 - masao@slis.tsukuba.ac.jp

データを表や図にまとめる



教示用スライドへ

第1回課題

- 以下の数列は2013年度某大学の某授業科目における受講生60名の成績得点の結果である(不合格者除く)
 - 度数分布表・ヒストグラム・累積相対度数折れ線を書け。なお、論理が追えるよう適宜、途中経過を示すこと。すべての途中計算を示す必要は無い。

77	78	80	83	78	80	85	78	76	78	62	80
70	80	65	80	76	81	78	73	78	77	78	68
80	78	92	72	78	90	83	78	78	77	77	80
77	76	78	70	73	77	78	80	78	90	68	78
78	72	77	95	76	78	78	77	73	72	75	76

第1回課題（提出方法）

- 課題はA4用紙（1枚）にて提出すること。
 - なお、2ページにわたる場合は裏面に記載のこと。
- 科目名（統計-1）、課題名（第1回課題）、提出年月日、学籍番号、所属、氏名を提出用紙の一番上に必ず記入すること。
- 提出×切：2013年11月20日（水）
 - 次回の授業時間中に提出を求めます。
 - やむをえず欠席等で当日に提出できない場合は、7D 208研究室前に提出場所を用意するので、そちらに提出すること。

本日のまとめ

- 授業科目の解説
 - 日程
 - 評価
- データのまとめ方
 - データ収集
 - ヒストグラム
 - 度数分布
- 次回予定：統計量