

現代日本論演習

3年生対象：2008年度6セメスタ
<火4>

授業の概要

学習目標

さまざまな統計分析手法を理解し、使いこなせるようになる

授業内容

分析の目的とデータの特徴によって、さまざまな統計分析手法が開発されています。この授業の前半では、推測統計学の基本的な概念について解説し、統計的推定・検定とサンプルサイズの決定の方法について学びます。後半では、相関分析、対応のある分析、多変量解析をとりあげて、それらの特徴と使い方を習得していきます。統計解析パッケージを使ってデータ分析の実習をおこないます。

履修要件

5セメスタ開講の現代日本論演習「統計分析の基礎」を履修済みか、それと同等の知識を習得済みの者を対象とする。

テキスト

吉田寿夫(1998)『本当にわかりやすいすごく大切なことが書いてあるごく初歩の統計の本』北大路書房。

成績評価の方法

各回の授業中の課題(50%)、中間試験(20%)、期末レポート(30%)を合計して評価する。

授業の予定

目次

1. 推測統計(10月)
2. 相関係数(11月)
3. 中間試験
4. 変数をキーにした分析(12月)
5. 多変量解析(1月)
6. 期末レポート

()内の日付は、学期前のおおよその計画をあらわしていますが、実際の授業の進行状況によって前後にずれることがあります。

1. 推測統計

- 確率密度と理論分布
- 誤差の対策
- 標本誤差の推定
- 平均値の点推定・区間推定
- 平均値の差の区間推定と t 検定
- 連関係数の区間推定と χ^2 検定
- サンプル・サイズと検定力

2. 相関係数

- 尺度水準についての復習
- 相関図
- Kendall の順位相関係数
- Spearman の順位相関係数
- Pearson の積率相関係数
- 相関係数行列
- 欠損値の処理 (pairwise/listwise)

3. 中間試験

4. 変数をキーにした分析

- 個体間変動と変数間変動
- 対応のある分析
- 2項検定
- ハッセ図の利用

5. 多変量解析

未定 (多元配置の分散分析?)

6. 期末レポート

第1回「確率密度と理論分布」(2008.10.7)

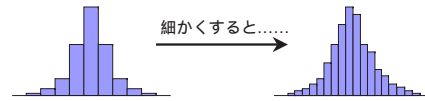
1. ヒストグラムと確率密度
2. 確率の理論分布
3. 2項分布
4. 正規分布

1

【ヒストグラム】

histogram

連続量を階級分けして度数分布を示したもの

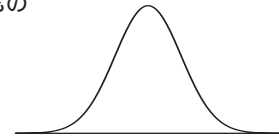


2

【確率密度のグラフ】

Probability density

連続量に対応して、連続的に変化する確率を表したもの



3

【確率の理論分布】

特定の仮定から理論的に導出された確率の分布

例：硬貨を投げるとき

表が出る

裏が出る

4

【2項分布】

binomial distribution

硬貨を n 回投げる。
表が出る回数を x とする。

$n = 4$ のとき、 x はどのような値を
どのような確率でとるか?

5

【計算方法】

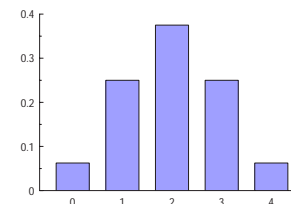
表=1, 裏=0 であらわすと

- 0 0 0 0 ($x=0$)
- 0 0 0 1 ($x=1$)
- 0 0 1 0 ($x=1$)
- 0 0 1 1 ($x=2$)
-
- 1 1 1 1 ($x=4$)

の16通り。それぞれ等しい確率 (1/16) で起こると考える。

6

$n = 4$, 確率=0.5 の2項分布



7

【期待値】

Expected value

値 (x) に確率 (p) を掛けたものの総和:

$$E = (x \times p)$$

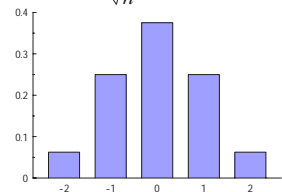
「平均値」と呼ばれることもある

$n = 4$ の2項分布の期待値は?

8

【標準化】

$Z = \frac{(x - E)}{\sqrt{n}}$ に変換すると



9

【標準正規分布】

Standard normal distribution

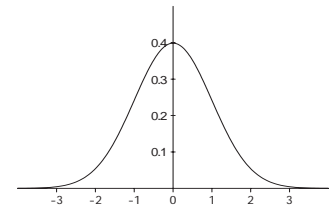
n が大きければ、 Z は
標準正規分布の確率密度関数

$$\frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right]$$

で近似できる

10

標準正規分布の確率密度のグラフ:



11

標準正規分布に定数による加減乗除を加えたものを総称して「正規分布」(normal distribution) という

0.5 以外の確率による2項分布でも、
適当な標準化を行って n を増加させると
正規分布に近づく

12

【正規分布の応用上の意義】

偶然による現象の生起確率や、
その組み合わせで決まる物事は、
正規分布 (またはそのファミリー)
で近似できることが多い

無作為抽出 = 等確率の繰り返し

13

【宿題】

硬貨を実際に投げて、表が出る枚数を数える。

- ・ 4 枚 × 10 回
- ・ 20 枚 × 40 回

それぞれ平均値を求めたうえ、棒グラフも書く

14

【文献】

宮川公男 (1999) 『基本統計学 [第3版]』有斐閣。

15

現代日本論演習 課題

氏名：

学籍番号：

次の2つの場合について、硬貨を実際に投げて、表が出る回数を数える。

- ・ 4枚×10回
- ・ 20枚×40回

下の表を完成させた上で、それぞれ棒グラフを作成する。また、平均値も求める

4枚×10回の場合

回数 _____ 表の枚数

1回目

2回目

3回目

4回目

5回目

6回目

7回目

8回目

9回目

10回目 _____

【平均値】

20枚×40回の場合

回数 _____ 表の枚数

1回目

2回目

3回目

4回目

5回目

6回目

7回目

8回目

9回目

10回目

11回目

12回目

13回目

14回目

15回目

16回目

17回目

18回目

19回目

20回目

21回目

22回目

23回目

24回目

25回目

26回目

27回目

28回目

29回目

30回目

31回目

32回目

33回目

34回目

35回目

36回目

37回目

38回目

39回目

40回目 _____

【平均値】