

# レポート提出票

科目名: 情報工学実験3

実験課題名: 課題3 教育システム開発

実施日: 2021年 6月 24日

学籍番号: 4619055

氏名: 辰川力駆

共同実験者:

_____	_____
_____	_____
_____	_____
_____	_____

# 1 要旨

情報工学実験2では、統計モデルを用いた分析におけるパラメタ推定について、項目反応理論を題材に Excel 上で学習・演習を行った。本実験では、前実験で扱わなかった e-Testing システムの実装について、実践を通して学習を行う。今回の実験では、項目反応理論を実装し、シミュレーションを行う。

## 2 目的

問題の特性や問題数が推定精度に与える影響をシミュレーションにて確認する。

## 3 理論

### 3.1 項目反応理論 (IRT)

項目反応理論 (IRT:Item Response Theory) は項目応答理論とも呼ばれ、正答回数のみで能力を測定していた古典テスト理論に対し、受験者の能力値  $\theta$  や問題の品質  $a$ ・難易度  $b$ などを考慮した新しいテスト理論である。確率論を基礎としており、中でも問題の特性や能力値の推定はベイズの定理が応用されている。IRT は今後ますます導入が進んでいくと考えられるので、その中でも基本的な「2パラメタロジスティックモデル」について学習する。

ある能力値  $\theta(\theta \in \mathbb{R})$  を持つ受験者の存在を仮定する。このとき、受験者が問題  $i$  に正答する確率  $P_i(1|\theta)$  は、IRT によると式 (1) のようにモデル化されている。

$$P_i(1|\theta) = P(1|\theta, a_i, b_i) = \frac{1}{1 + \exp(-Da_i(\theta - b_i))} \quad (1)$$

ここで、 $P_i(1|\theta)$  は、受験者の能力が  $\theta$  であった場合に問題に対して正答 ( $x = 1$ ) する条件付き確率を表す。この関数は IRT では項目反応関数と呼称される。

$a(a > 0)$  は識別力パラメタ、 $b(b \in \mathbb{R})$  は困難度パラメタと呼ばれる。 $a_i$  が小さい問題  $i$  の回答  $x$  においては、能力  $\theta$  や難易度  $b_i$  によらず、一定の正答率しか得ることができない。 $a_i$  は非線形 SVM でいう  $\gamma$  に相当し、小さければ項目の機能は落ちるが、大きすぎると項目数が少ない場合に汎化誤差が高まりやすい。そのため、 $a$  は  $0.3 < a < 1.7$  程度であることが望ましいとされている。 $a$  と  $b$  は、 $\theta$  が明らかである場合にのみロジスティック回帰や線形 SVM で求めることができる。 $\theta$  が不明である場合は EM アルゴリズムを用いた周辺最尤推定を使う。 $D$  は項目反応関数  $P_i(1|\theta)$  を正規分布の累積分布関数に近似させるための係数であり、 $D = 1.7$  のとき非常によく近似することが知られている。また式 (1) は正答する確率を表す。ここで、ある問題に正答したとき 1、誤答した場合には 0 となる変数  $x$  があるものとする。すると、反応確率  $P_i(x|\theta)$  は式 (2) のように表せる。

$$P_i(x|\theta) = P_i(1|\theta)^x (1 - P_i(1|\theta))^{1-x} \quad (2)$$

後述するベイズの定理では、これを尤度関数と呼称する。

### 3.2 能力値推定

問題系列についての正誤反応  $\mathbf{X} = (x_1, x_2, \dots, x_N)$  が与えられたとする。このときの能力値  $\theta$  は次式 (3) によって推定することが可能である。

$$\hat{\theta} = \arg \max_{\theta} P(\theta|\mathbf{X}) \quad (3)$$

$P(\theta|\mathbf{X})$  は、ある能力値  $\theta$  の受験者がそれぞれの問題  $i$  に  $\mathbf{X} = (x_1, x_2, \dots, x_N)$  のように反応したという条件の下で、その受験者の能力値が  $\theta$  である確率を示す。 $\max$  は  $P(\theta|\mathbf{X})$  の最頻値であり、 $\arg\max$  は尤も確率が  $P$  が高くなる引数  $\theta$  の集合を表す。

ただし、IRT において  $P(\theta|\mathbf{X})$  は明らかでない。そこで、式 (4) に示すベイズの定理を用いる。

$$P(\theta|\mathbf{X}) = \frac{P(\mathbf{X}|\theta)P(\theta)}{P(\mathbf{X})} \quad (4)$$

ベイズ統計学においては、 $\theta$  を仮定 (原因)、 $\mathbf{X}$  を結果 (データ) と解釈する。また、 $P(\theta)$  を事前確率、 $P(\mathbf{X}|\theta)$  を尤度関数、 $P(\mathbf{X})$  を周辺尤度、 $P(\theta|\mathbf{X})$  を事後確率と呼称することがある。

ここで、ある項目が他の項目の解答やヒントにならず、かつ問題群が一つの科目 (例えば数字のみ) で構成されたとした場合、尤度関数  $P(\mathbf{X}|\theta)$  は式 (5) のように求められる。

$$P(\mathbf{X}|\theta) = P_1(x_1|\theta)P_2(x_2|\theta)\dots P_N(x_N|\theta) = \prod_{i=1}^N P_i(x_i|\theta) = \prod_{i=1}^N P(x_i|\theta, a_i, b_i) \quad (5)$$

このような「潜在特性値  $\theta$  を固定したとき、異なる項目への反応  $x$  が統計的に互いに独立になる性質」を局所独立性と呼ぶ。この条件を仮定することで、項目間の正誤反応を同時確率とできる。仮定の成立を積極的に評価するのは難しいが、主成分分析を用い、項目に対する反応の無相関性と、受験者の反応の一次元性を検証するケースが多い。

条件付き確率  $P(\theta|\mathbf{X})$  を尤度関数  $P(\mathbf{X}|\theta)$  のみで表現し最大引数  $\theta$  を推定値とする方法論を最尤推定、周辺尤度  $P(\mathbf{X})$  を考慮した最尤推定を最大事後確率推定 (MAP 推定)、事前分布  $P(\theta)$  を考慮した最尤推定をベイズ推定と呼ぶ。ベイズ推定に準拠すれば式 (6) で能力値が推定できる。

$$\hat{\theta} = \arg \max_{\theta} \frac{P(\mathbf{X}|\theta)P(\theta)}{P(\mathbf{X})} = \arg \max_{\theta} \left( \frac{P(\theta)}{P(\mathbf{X})} \prod_{i=1}^N P(x_i|\theta, a_i, b_i) \right) \quad (6)$$

### 3.3 項目情報量

項目情報量とは、項目への反応から能力値を推定する際のフィッシャー情報量である。式 (7) のようになる。

$$I_i(\theta) = I(\theta, a_i, b_i) = D^2 a_i^2 P(1|\theta, a_i, b_i)(1 - P(1|\theta, a_i, b_i)) \quad (7)$$

つまり、問題の識別力と正答確率・誤答確率の積である。さらに、複数の確率変数  $P(1|\theta, a, b)$  がすべて独立という条件の下で、フィッシャー情報量は加法性が成立する。

$$I(\theta) = \sum_i I_i(\theta) \quad (8)$$

フィッシャー情報量は一般に最尤推定値  $\hat{\theta}$  の標準偏差  $se(\hat{\theta})$  と以下の関係を持つ。

$$se(\hat{\theta}) = I(\hat{\theta})^{-\frac{1}{2}} \quad (9)$$

そのため情報量を算出することで、推定能力値  $\theta$  に対する誤差を見積もることが可能である。

4 課題

5 まとめ

## 参考文献

- [1] 特徴量選択のまとめ - Qiita

<https://qiita.com/shimopino/items/5fee7504c7acf044a521>

最終閲覧日 2021/06/24

## A 付録

ソースコード 1: 1\_functions.php

---