レポート提出票

科目名:	情報工学実験2
実験テーマ:	実験テーマ5 教育システム設計
実施日:	2020年 11月 16日
学籍番号:	4619055
氏名:	辰川力駆
共同実験者:	書かないと！

1 要旨

2 目的

統計モデルを用いた分析は、例えば商品の推薦や迷惑メールの削除機能など、身近な機能を支える基本的な技術となっている。本実験では、このような統計モデルを用いた分析に欠かせない、パラメタの推定の方法について、基本的な技術を習得することを目的とする。

3 理論

3.1 項目反応理論:複数問の反応からの能力値推定

複数の問題に対する正誤反応を得た場合の能力値の推定について考える。今、ある問題系列について正誤反応 $\mathbf{X}_j = (x_{1,j}, x_{2,j}, ..., x_{n,j})$ が与えられたとする。この時の能力値 θ の推定値を考える際には以下のような数式を考えれば良い。

$$\begin{aligned} \hat{\theta} &= \arg\max_{\theta} P(\theta \mid \mathbf{X}_j) &(1)\\ &= \arg\max_{\theta} P(\theta)P(\mathbf{X}_j \mid \theta) &(2) \end{aligned}$$

ここで $P(\mathbf{X}_j \mid \theta)$ はある能力値 θ の受験者が、それぞれの項目に $\mathbf{X}_j)$ のように反応する確率である。そのため、それぞれの項目への正誤が θ のみに影響され定まるとすれば、$P(\mathbf{X}_j \mid \theta)$ は以下のようにかける。

$$\begin{aligned} P(\mathbf{X}_j \mid \theta) &= P(x_{1,j} \mid \theta) \times P(x_{2,j} \mid \theta) \times \cdots \times P(x_{n,j} \mid \theta) &(3)\\ &= \prod_{i=1}^{n} P(x_{i.j} \mid \theta) &(4) \end{aligned}$$

つまり、複数のサイコロを同時に投げたときと同様に同時確率と見なすことができる。それぞれの項目への能力値 θ だけを媒介に独立に正誤反応していると考える。これを局所独立過程という。すなわち、それぞれの項目が他の項目の小 r たえやヒントになっていない状況である。

また、$P(x_{i,j} \mid \theta)$ は正答の場合と誤答の場合両方を以下のように表すことができる。

$$P(x_{i,j} \mid \theta) = P_i(\theta)^{x_{i,j}} \times (1 - P_i(\theta))^{1-x_{i,j}} \tag{5}$$

従って考えるべき式は以下のようになる。

$$\begin{aligned} \hat{\theta} &= \arg\max_{\theta} P(\theta \mid \mathbf{X}_j) &(6)\\ &= \arg\max_{\theta} P(\theta) \prod_{i=1}^{n} P(x_{i,j} \mid \theta) &(7)\\ &= \arg\max_{\theta} P(\theta) \prod_{i=1}^{n} (P_i(\theta)^{x_{i,j}} (1 - P_i(\theta))^{1-x_{i,j}}) &(8) \end{aligned}$$

ただし、これを計算機により計算機により計算することは、値域がをとる関数を複数回かけることになり、計算誤差が生じやすい。そのため、実装上は、このような関数の対数関数を考える。log は単調増加関数であり、ある関数 $f(x) = x$ が $x = x_{\max}$ で最大を取るとき、$\log(f(x))$ も $x = x_{\max}$ で最大を取る。そのため、以下が成り立つ。

$$\hat{\theta} = \arg\max_{\theta} P(\theta \mid \mathbf{X}_j) \tag{9}$$

$$= \arg\max_{\theta} \ln\left(P(\theta \mid \mathbf{X}_j)\right) \tag{10}$$

そのため，実装上は以下を計算すれば良い．

$$\hat{\theta} = \arg\max_{\theta} \left\{ \ln\left(P(\theta)\right) + \sum_{i=1}^{n} \left(x_{i,j} \ln(P_i(\theta)) + (1 - x_{i,j}) \ln(1 - P_i(\theta)) \right) \right\} \tag{11}$$

これを対数尤度関数と呼ぶ．

4 課題

4.1 課題 2-1

課題 1-3 で解いた項目について、項目反応関数の概形を描く。

表 1: 解いた問題の特性パラメタ

問題	a パラメタ	b パラメタ
3	2.87168	0.69892
9	0.42082	0.22627
21	1.03497	0.31148
24	0.71798	1.22817
30	1.20718	0.65633
35	0.72641	0.14052
40	0.31796	2.23511
51	0.51611	0.94029

4.2 課題 2-2

課題 2-1 で描いたグラフを参考に、それらの項目がどのような項目であったのか考察する。

図 1 のグラフの中央付近の増加量について、Item3 が一番正答率の変化量が大きい。これは、a パラメタが他と比べて大きいからである。つまり、Item3 などの a パラメタが大きい問題は、ある一定の能力値を超えるとほぼ 100 %の正答率となるが、逆にその能力値を越えていないとほぼ 0 %の正答率となることが分かる。

b パラメタについては、Item40 を見ると分かるように正答率が高くなるまでに必要な能力値が高いことがわかる。つまり、b パラメタが低い順に正答率が 50 %になるための能力値が低い。

4.3 課題 2-3

$\theta_j = b_i$ の場合、正答確率は 50 % であることを証明する。

式 (1) に $\theta_j = b_i$ を代入して、

$$\begin{aligned}
P(\theta_j = b_i) &= \frac{1}{1 + \exp\{-Da_i(\theta_j - b_i)\}} \\
&= \frac{1}{1 + \exp\{-Da_i(b_i - b_i)\}} \\
&= \frac{1}{1 + \exp(0)} \\
&= \frac{1}{2} \\
&= 0.5
\end{aligned}$$

よって正答確率は 50 % となる。

4.4 課題 2c-1

$\theta_j = b_i$ での項目反応関数の傾きは a_i に比例することを証明する。

項目反応関数の傾きを知るために、式 (1) を微分すると、

$$\frac{d}{d\theta_j}P(\theta_j) = \frac{Da_i \exp\{-Da_i(\theta_j - b_i)\}}{[1 + \exp\{-Da_i(\theta_j - b_i)\}]^2}$$

この式に $\theta_j = b_i$ を代入して、

$$\begin{aligned}
\frac{d}{d\theta_j}P(\theta_j = b_i) &= \frac{Da_i \exp(0)}{\{1 + \exp(0)\}^2} \\
&= \frac{1}{4}Da_i
\end{aligned}$$

D は固定しているので、a_i に比例している。

4.5 課題 2–4

課題 1-3 で解いた項目から 1 題選びその正誤から描かれる事後分布のグラフを示す。

課題 1-3 で解いた項目から、Item21 を選んだ。Item21 の正答確率と誤答確率は図 2 のようになった。もちろんだが、正答確率 + 誤答確率 = 1 なので 確率 = 0.5 の直線で線対称なグラフになっている。

また、正答確率事後分布と誤答確率事後分布は図 3 のようになった。正規分布は上に凸であるが、それをかけ合わせているので、上に凸なグラフとなっている。

4.6 課題 2–5

課題 2-4 で描いた事後分布から能力値を推定する。またその時の標準誤差を求める。

まず、能力値を推定する。Item21 を正解したので、式 (4) より尤度の最大値を考えて

$$\hat{\theta} = 0.6$$

である。

次に、標準誤差を求める。式 (10) より、フィッシャー情報量は

$$\begin{aligned} I_i(\theta) &= D^2 a_i^2 P_i(\theta)(1 - P_i(\theta)) \\ &= 1.7^2 \times 1.035^2 \times 0.624 \times (1 - 0.624) \\ &\approx 0.726 \end{aligned}$$

であるから、式 (11) より標準誤差は

$$\begin{aligned} \mathrm{se}(\hat{\theta}) &= I_i(\hat{\theta})^{-\frac{1}{2}} \\ &= \frac{1}{\sqrt{I_i(\hat{\theta})}} \\ &\approx 1.174 \end{aligned}$$

である。

4.7 課題 2–6

θ の刻み幅を 0.1 以外に 2 パターンを自ら適当に変更した場合の能力値と標準誤差の結果を記述し、結果について考察する。

θ の刻み幅を 0.05 と 0.8 にした場合を考えた。図 4 と図 5 は、θ の刻み幅を 0.05 にした場合で、図 6 と図 7 は、θ の刻み幅を 0.8 にした場合である。

θ の刻み幅 0.05

能力値を推定する。Item21 を正解したので、式 (4) より尤度の最大値を考えて

$$\hat{\theta} = 0.65$$

である。

次に、標準誤差を求める。式 (10) より、フィッシャー情報量は

$$\begin{aligned} I_i(\theta) &= D^2 a_i^2 P_i(\theta)(1 - P_i(\theta)) \\ &= 1.7^2 \times 1.035^2 \times 0.645 \times (1 - 0.645) \\ &\approx 0.709 \end{aligned}$$

であるから、式 (11) より標準誤差は

$$\begin{aligned} \mathrm{se}(\hat{\theta}) &= I_i(\hat{\theta})^{-\frac{1}{2}} \\ &= \frac{1}{\sqrt{I_i(\hat{\theta})}} \\ &\approx 1.188 \end{aligned}$$

である。

θ の刻み幅 0.8

能力値を推定する。Item21 を正解したので、式 (4) より尤度の最大値を考えて

$$\hat{\theta} = 0.8$$

である。

次に、標準誤差を求める。式 (10) より、フィッシャー情報量は

$$\begin{aligned} I_i(\theta) &= D^2 a_i^2 P_i(\theta)(1 - P_i(\theta)) \\ &= 1.7^2 \times 1.035^2 \times 0.703 \times (1 - 0.703) \\ &\approx 0.647 \end{aligned}$$

であるから、式 (11) より標準誤差は

$$\begin{aligned} \mathrm{se}(\hat{\theta}) &= I_i(\hat{\theta})^{-\frac{1}{2}} \\ &= \frac{1}{\sqrt{I_i(\hat{\theta})}} \\ &\approx 1.243 \end{aligned}$$

である。

θ の刻み幅 0.05 および θ の刻み幅 0.8 の能力値と標準誤差は課題 2-5 で求めた値と違っていた。これは、能力値を求める際に、使用できる θ が違うからである。

θ の刻み幅を小さくすれば、さらに使える能力値が多くなるので、正確に推定するならば刻み幅を小さくすればよいと考える。

4.8 課題 2–7

能力値の事前分布が平均-0.3、分散 1 の正規分布に従うと仮定された場合、自分の能力値と情報量について計算し、結果について考察する。

能力値の事前分布が平均-0.3、分散 1 の正規分布に従うと仮定された場合、正答確率と誤答確率のグラフは変わらないので図 2 と同じになる。事前分布が変わると影響されるのは正答確率事後分布と誤答確率事後分布である。そのグラフを図 8 に表した。

能力値を推定する。Item21 を正解したので、式 (4) より尤度の最大値を考えて

$$\hat{\theta} = 0.5$$

である。

次に、標準誤差を求める。式 (10) より、フィッシャー情報量は

$$\begin{aligned} I_i(\theta) &= D^2 a_i^2 P_i(\theta)(1 - P_i(\theta)) \\ &= 1.7^2 \times 1.035^2 \times 0.582 \times (1 - 0.582) \\ &\approx 0.753 \end{aligned}$$

であるから、式 (11) より標準誤差は

$$\begin{aligned} \mathrm{se}(\hat{\theta}) &= I_i(\hat{\theta})^{-\frac{1}{2}} \\ &= \frac{1}{\sqrt{I_i(\hat{\theta})}} \\ &\approx 1.152 \end{aligned}$$

である。

図 3 と図 8 のグラフを比べると分かるが、全体的に少しだけ左の値が大きくなっている。今回は誤答確率事後分布が-0.3 付近で山なりになっているので、その部分が大きくなっている。それらによって、能力値やフィッシャー情報量、標準誤差が変わっている。

5 まとめ

項目反応理論を用いた場合の受験者の能力値の推定について実験を行った。また、課題を通して刻み幅を変更した場合の能力値の推定や、標準正規分布ではなく違った正規分布に従う場合の推定も行った。これらを通して項目反応理論を理解することができた。