

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: 近持卓 太田遼真

中澤幸大 国方慶吾

米澤祥吾 塩山大空

1 要旨

項目反応理論を複数の項目について適用し能力値を推定することで、より複雑な状況での能力推定を演習として行う。

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 $\mathbf{X}_j = (x_{1,j}, x_{2,j}, \dots, x_{n,j})$ 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。

$$\hat{\theta} = \arg \max_{\theta} P(\theta | \mathbf{X}_j) \quad (1)$$

$$= \arg \max_{\theta} P(\theta) P(\mathbf{X}_j | \theta) \quad (2)$$

ここで $P(\mathbf{X}_j | \theta)$ 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、 $P(\mathbf{X}_j | \theta)$ は以下のようにかける。

$$P(\mathbf{X}_j | \theta) = P(x_{1,j} | \theta) \times P(x_{2,j} | \theta) \times \dots \times P(x_{n,j} | \theta) \quad (3)$$

$$= \prod_{i=1}^n P(x_{i,j} | \theta) \quad (4)$$

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みなすことができる。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、これを局所独立仮定という。すなわち、それぞれの項目が他の項目の答えやヒントになっていない状況である。

また、 $P(x_{i,j} | \theta)$ 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。

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

従って、考えるべき式は以下のようなになる。

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

ただし、これを計算機により計算することは、値域が $[0, 1]$ を取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 $f(x)$ 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。

$$\hat{\theta} = \arg \max_{\theta} P(\theta | \mathbf{X}_j) \tag{7}$$

$$= \arg \max_{\theta} \ln (P(\theta | \mathbf{X}_j)) \tag{8}$$

そのため、実装上は以下を計算すれば良い。

$$\hat{\theta} = \arg \max_{\theta} \left\{ \ln (P(\theta)) + \sum_{i=1}^n \left(x_{i,j} \ln(P_i(\theta)) + (1 - x_{i,j}) \ln(1 - P_i(\theta)) \right) \right\} \tag{9}$$

これを対数尤度関数と呼ぶ。

4 課題

4.1 課題 3-1

課題 2-4 で描いた事後分布とその対数尤度関数を図示し、それぞれの最大値にどのような関係があるか示す。

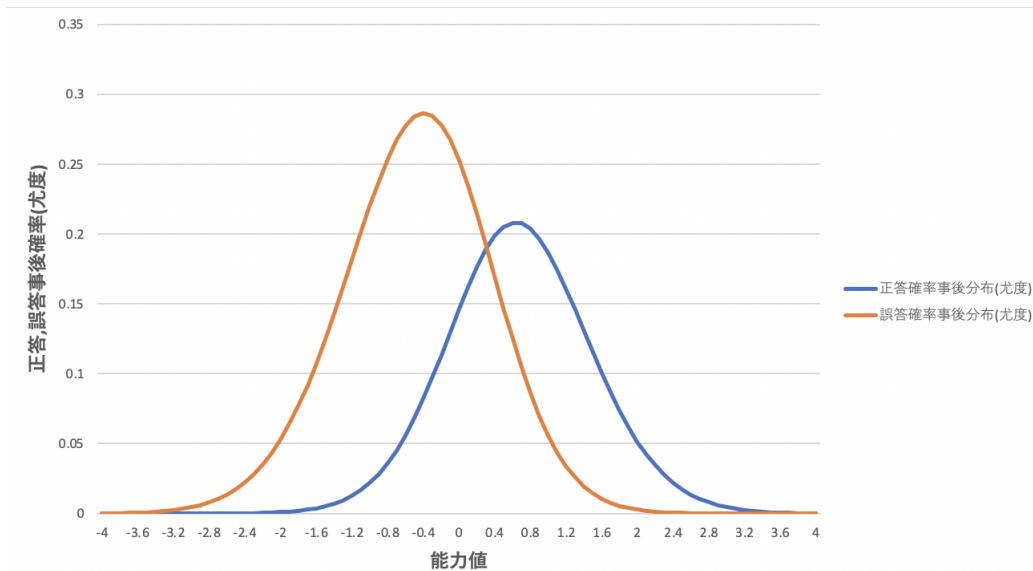


図 1: Item21 の正答確率事後分布と誤答確率事後分布のグラフ

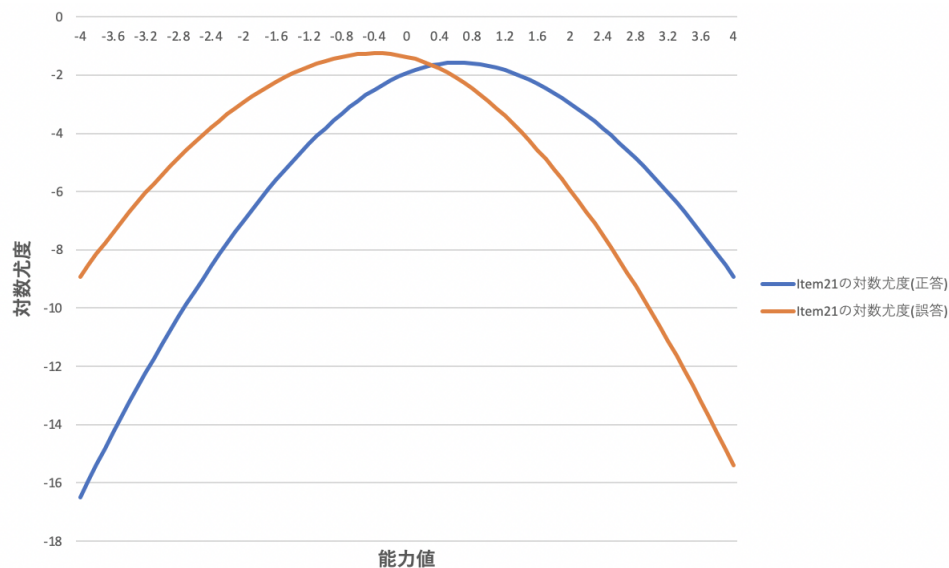


図 2: Item21 の正答と誤答の対数尤度グラフ

Item21 の事後分布のグラフは図1のようになった。また、対数尤度関数のグラフは図2のようになった。尤度の値はすべて0～0.3の間に収まっているので、自然対数を取った対数尤度の値はすべて負の値を取っている。

対数尤度は事後分布の対数をとったものであるが、対数関数は定義域に対して単調増加であるから、ある関数に対し対数をとっても横軸の値に対する縦軸の値の大小関係は変化しない。つまり、どちらのグラフも最大値を取るときの θ は同じで、正答は $\theta = 0.6$ で誤答は $\theta = -0.4$ である。

4.2 課題3-2

課題1-3で解いた全ての項目の反応から自分の能力値 θ の対数尤度関数の概形を示す。

Item3,9,21,24,30,35,40,51を解き全て正答したので、それらの正答確率の対数関数を全て足し合わせると、グラフを求めることができる。結果は図3ようになった。

対数関数なので、和を求めたことにより、上記のItem21だけの対数対数尤度グラフと比べて、最大値を取っている θ の値が高くなっている。これは全て正答したからである。

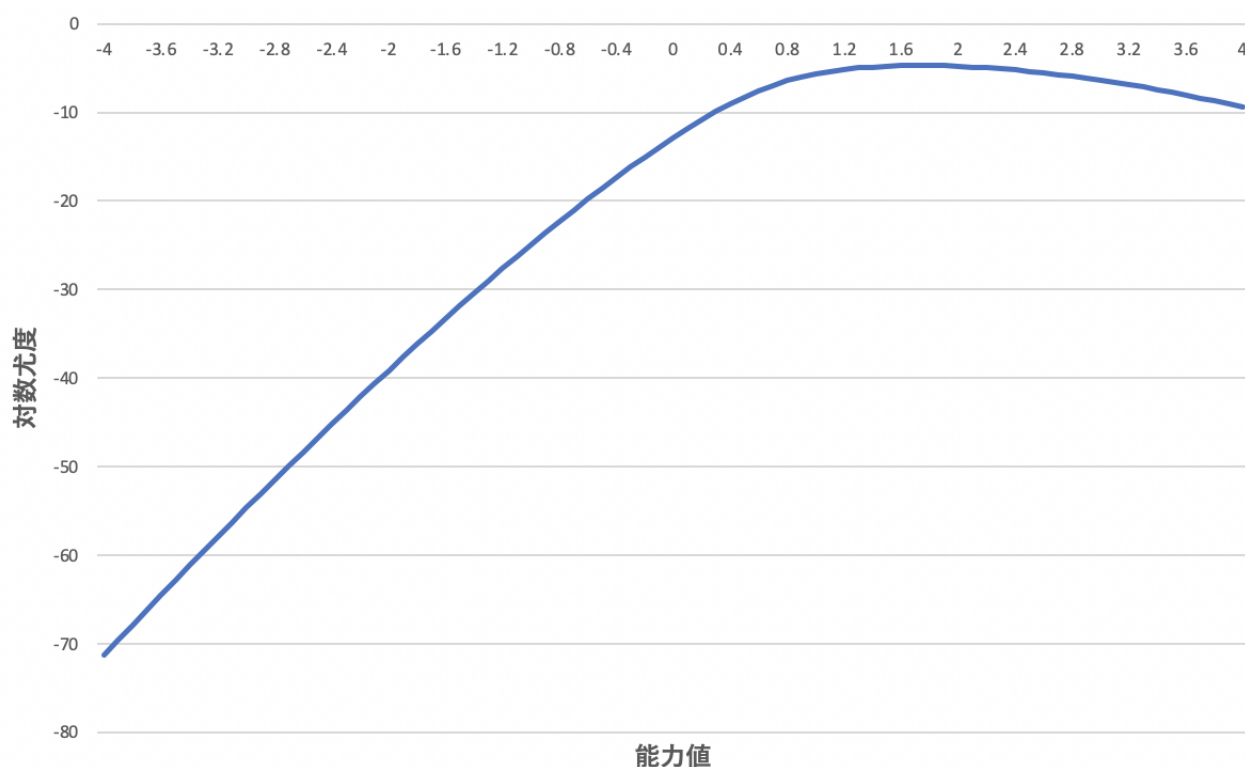


図3: 全ての項目による対数尤度グラフ

4.3 課題 3-3

課題 3-2 で描いた対数尤度関数から能力値を推定する。またその時の情報量 $I(\hat{\theta})$ と標準誤差 $se(\hat{\theta})$ を求める。

図 3 の対数尤度関数で最大値を取っている θ の値は、 $\theta = 1.7$ である。よって自分の能力値は 1.7 であると推定できる。また、情報量 $I(\hat{\theta})$ は各問題に関する情報量 $I_i(\theta)$ の合計である。したがって、8 問の合計を求めたら良く、

$$\begin{aligned} I(\hat{\theta}) &= \sum_{i=1}^8 I_i(\theta) \\ &= \sum_{i=1}^8 D^2 a_i^2 P_i(\theta)(1 - P_i(\theta)) \\ &= \sum_{i=1}^8 1.7^2 a_i^2 P_i(1.7)(1 - P_i(1.7)) \\ &\approx 1.65559 \end{aligned}$$

である。

また、標準誤差は

$$\begin{aligned} se(\hat{\theta}) &= I(\hat{\theta})^{-\frac{1}{2}} \\ &= \frac{1}{\sqrt{1.65559}} \\ &\approx 0.77718 \end{aligned}$$

である。

4.4 課題 3-4

課題 1-3 で解いた項目の a_i が全て「1」だった場合の能力値、情報量、標準誤差を求め、結果について考察する。

a_i が全て「1」だった場合の対数尤度グラフは図 4 のようになる。図 4 の対数尤度関数で最大値を取っている θ の値は、 $\theta = 2.1$ である。

よって、情報量は課題 3-3 と同様にして求めればよいから、

$$\begin{aligned} I(\hat{\theta}) &= \sum_{i=1}^8 I_i(\theta) \\ &= \sum_{i=1}^8 D^2 a_i^2 P_i(\theta)(1 - P_i(\theta)) \\ &= \sum_{i=1}^8 1.7^2 1^2 P_i(2.1)(1 - P_i(2.1)) \\ &\approx 2.22578 \end{aligned}$$

である。

また、標準誤差は

$$\begin{aligned} se(\hat{\theta}) &= I(\hat{\theta})^{-\frac{1}{2}} \\ &= \frac{1}{\sqrt{2.22578}} \\ &\approx 0.67028 \end{aligned}$$

である。

図4のグラフからわかるように図3とあまり概形に変化はない。これは困難度パラメタ (b_i) を変えていないからである。

情報量に関しては識別力パラメタ (a_i) を2乗するので $a_i = 1$ にすることで増加している。だから、分母が大きくなるので標準誤差は小さくなっている。

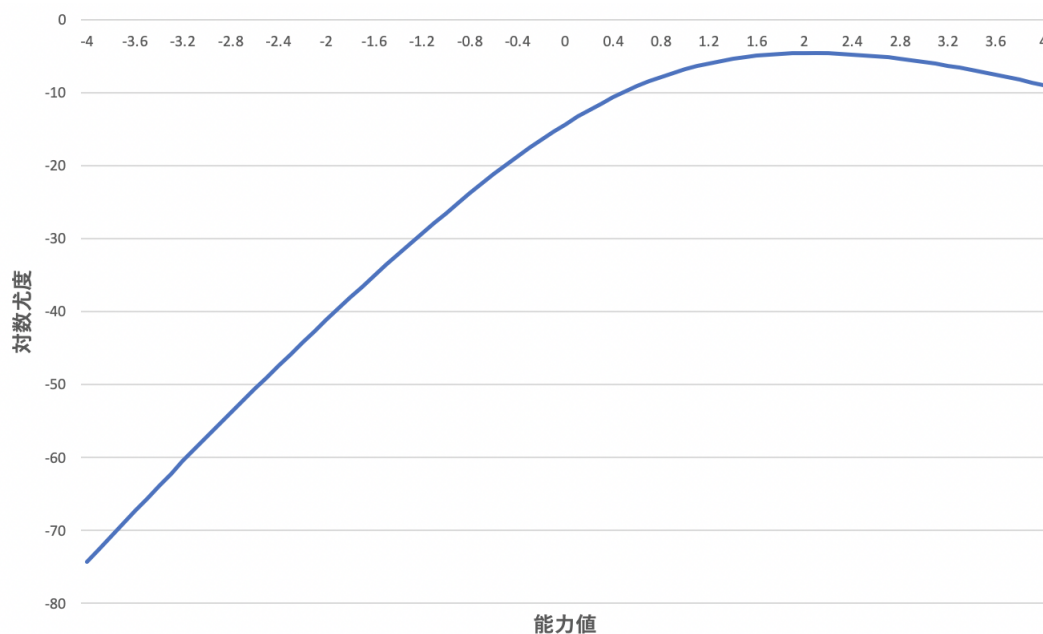


図 4: 全ての項目による対数尤度グラフ ($a_i = 1$ のとき)

4.5 課題 3-5

解いた項目と、能力値 θ の推定結果と、課題 1-4 で行なった偏差値 S の結果を同じグループ内で共有し、比べた結果を考察する。

表 1: G9 メンバーの推定結果

学籍番号 (46190)	15	28	38	55	58	64	94
能力値 θ	0	-0.1	1	1.7	0.4	1	1.3
偏差値 S	45	45	60	70	61.7	70	70
解答した問題	5	1	3	3	8	1	3
	11	5	8	9	20	15	6
	36	20	16	21	28	28	20
	49	25	49	24	35	46	34
			52	30	42	52	45
				35	46		
				40			
				51			

グループ 9 のメンバーの結果を表 1 にまとめた。解いた問題数を比べると、自分が一番多かった。また、能力値も一番高かった。偏差値が 70 のとき 100 点なので自分 (4619055) 以外に 4619064 と 4619094 が満点だが、能力値が違っている。これは問題のパラメタの違いも関係しているが、解いた問題数が多いと能力値が上がる傾向があると考えられる。

5 まとめ

今回の演習では、複数項目の正誤反応が与えられたときの能力値の推定を行った。複雑な状況の下で正確な能力値を推定できた。また、対数関数を用いて、問題全てを解いた合計の推定を行った。