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 9日

学籍番号: 4619055

氏名: 辰川力駆

共同実験者:

1　要旨

項目反応理論の紹介と、それを用いた場合の受験者の能力値の推定について学習し、課題を通して演習を行う。

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　項目反応理論基礎

項目反応理論では、ある受験者 j がある項目 i に正答する $x_{i,j} = 1$ 確率を以下の式でモデル化する。

$$P(x_{i,j} = 1|\theta_j, a_i, b_i) = P_i(x_{i,j} = 1|\theta_j) = P_i(\theta_j) = \frac{1}{1 + \exp\{-Da_i(\theta_j - b_i)\}} \tag{1}$$

ここで θ_j は受験者の能力値パラメタ、a_i は項目の識別力パラメタ、b_i は項目の困難度パラメタである。また、$D = 1.7$ としたとき、このモデルは以下の特徴がある。

1. $\theta_j = b_i$ の場合、正答確率は 50% となる。
2. $\theta_j = b_i$ での正答率の傾きは a_i に比例する。
3. $a_i = 1, b_i = 0$ のとき、この関数は累積標準正規分布の良い近似となる。

a は識別力パラメタと呼ばれる。a が 0 に近い項目は、能力値によらず一定の正答率となるような、能力に関係ない項目となる。a が 0 に近い値の時は、能力値に関わらず正答率が高い傾向がある。

b は困難度パラメタと呼ばれる。項目反応関数の正答確率が 50 % となる点の θ と同じ値となる。b が大きな項目では正答に必要な能力値が大きくなる。b が大きいほど、小さい能力値での正答率が低い傾向がある。

3.2　最尤推定を用いた受験者能力の推定

項目反応理論を用いた受験者の能力値の推定も最尤推定と同様に計算する。例えば、今ある問題に正答した $(x_{i,j} = 1)$ とする。この時の受験者の能力値は以下の数式を考えれば良い。

$$\hat{\theta} = \arg\max_{\theta} P(\theta|x_{i,j} = 1) \tag{2}$$

ただし、$P(\theta_j|x_{i,j}=1)$ はIRTのモデルに従って直接は与えられていないため、ベイズの定理より、$P_i(x_{i,j}=1|\theta_j)$ を用いて考える。

$$P(\theta_j|x_{i,j}=1) = \frac{P(\theta_j)}{P(x_{i,j}=1)} P_i(x_{i,j}=1|\theta_j) \tag{3}$$

ここで、$P(x_{i,j}=1)$ はその問題に正答できる確率であるが、変数 θ_j と独立な変数であるため、$\arg\max\limits_{\theta}$ を考える上で定数とみなすことができる。従って、以下の式を考えても結果は変わらない。

$$\begin{aligned}\hat{\theta} &= \arg\max_{\theta} P(\theta|x_{i,j}=1) &(4)\\ &= \arg\max_{\theta} P(\theta)P_i(x_{i,j}=1|\theta_j) &(5)\end{aligned}$$

$P(\theta)$ は能力値が一般にどのような分布をしているかを表すので、標準正規分布していると仮定できる。

$$P(\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\theta^2}{2}\right) \tag{6}$$

また、誤答した場合も同様に以下のように考えることができる。

$$\begin{aligned}\hat{\theta} &= \arg\max_{\theta} P(\theta|x_{i,j}=0) &(7)\\ &= \arg\max_{\theta} P(\theta)P_i(x_{i,j}=0|\theta_j) &(8)\\ &= \arg\max_{\theta} P(\theta)(1-P_i(x_{i,j}=1|\theta_j)) &(9)\end{aligned}$$

このような関数を考えることで、ある問題に正答、あるいは誤答した場合の受験者の能力値を推定することが可能である。また、$P(\theta)$ は問題に正答したという事実を受け取る前の確率であり、事前確率と呼ばれることがある。加えて、$P(\theta|x_{i,j})$ は問題に正答あるいは誤答したという事実を受け取ったあとの確率なので、事後確率と呼ばれることがある。

また、項目反応理論では、項目情報量関数と呼ばれる指標が非常に重要となる。項目情報量関数とは、項目反応関数について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を算出したものであり、2パラメタロジスティックモデルでは以下のような式となる。

$$I_i(\theta) = D^2 a_i^2 P_i(\theta)(1-P_i(\theta)) \tag{10}$$

フィッシャー情報量は一般に $\hat{\theta}$ の標準誤差 $\mathrm{se}(\hat{\theta})$ と以下の関係を持つ

$$\mathrm{se}(\hat{\theta}) = I_i(\hat{\theta})^{-\frac{1}{2}} = \frac{1}{\sqrt{I_i(\hat{\theta})}} \tag{11}$$

そのため、この項目情報量を用いることで、推定された $\hat{\theta}$ に対してどの程度標準誤差があるかを見積もることができる。

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。

課題 1-3 で解いた問題のパラメタについて表 1 にまとめた。そして、それを基に項目反応関数の概形を描くと図 1 のようになった。

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

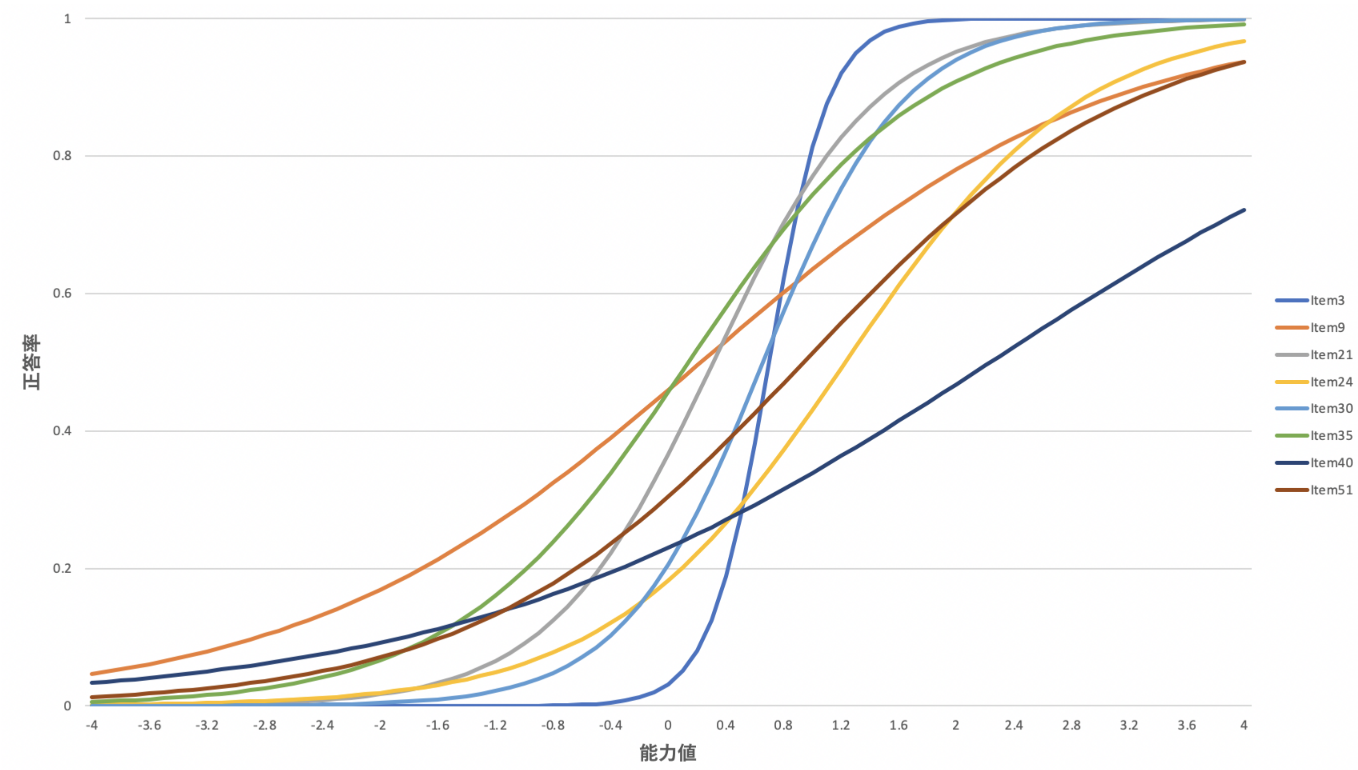

図 1: 項目反応関数

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。

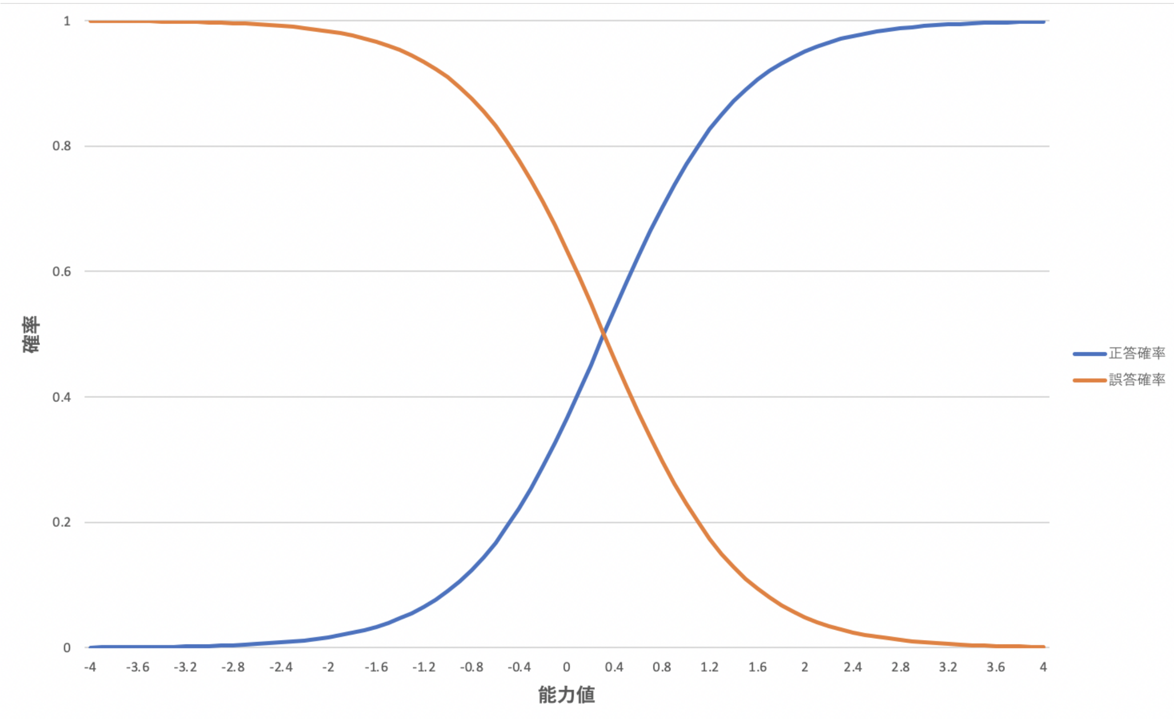

図 2: Item21 の正答確率と誤答確率のグラフ

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。

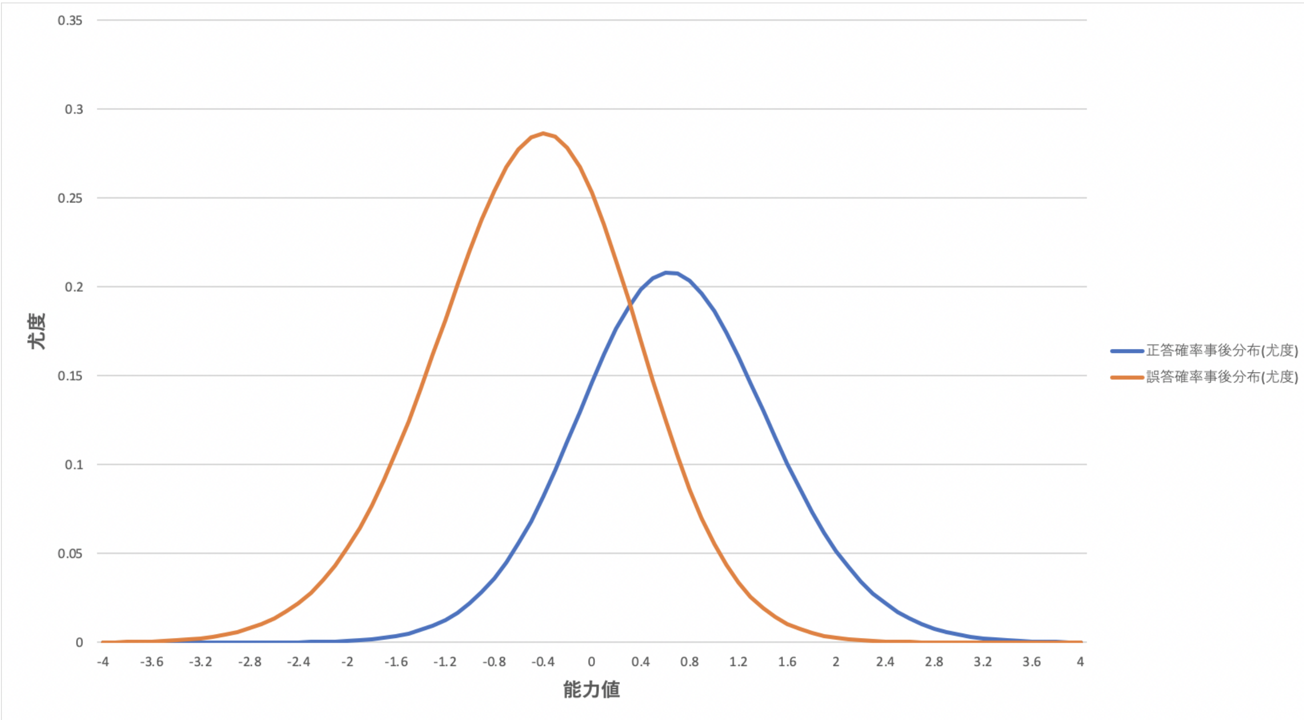

図 3: Item21 の正答確率事後分布と誤答確率事後分布のグラフ

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。

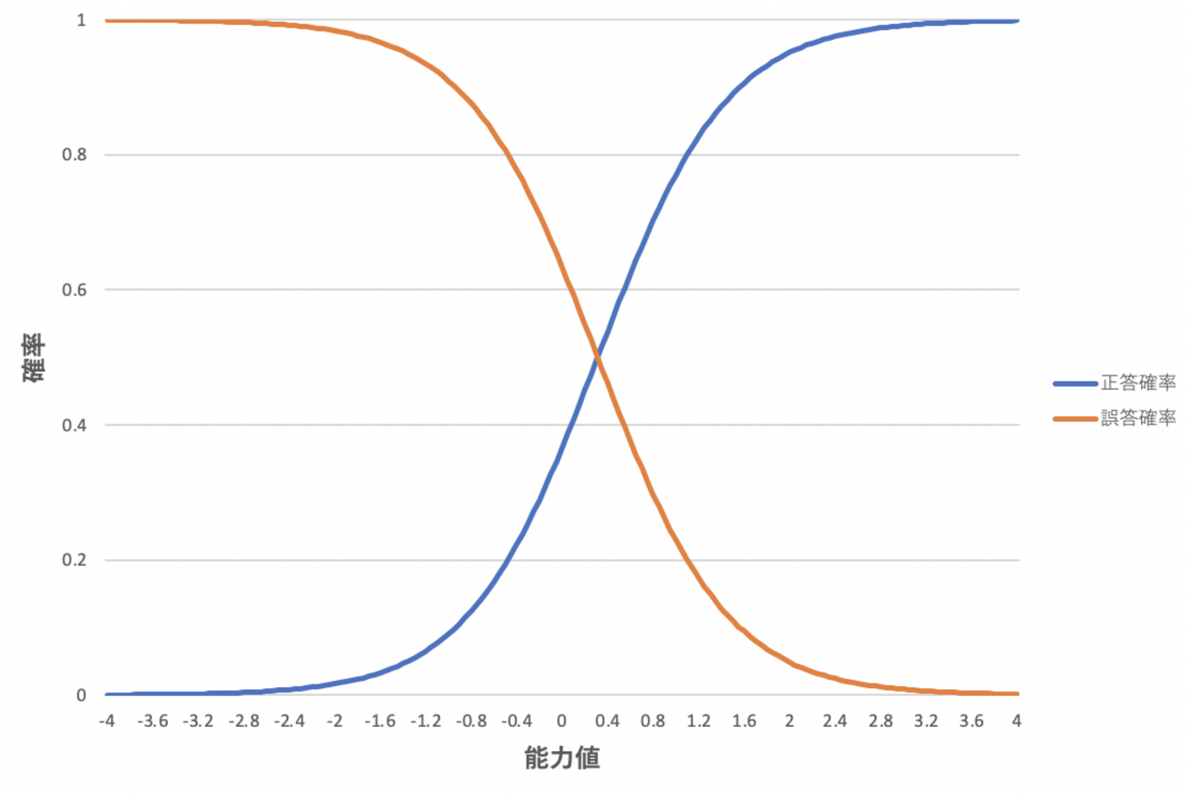

図 4: Item21 の正答確率と誤答確率のグラフ (θ の刻み幅 0.05)

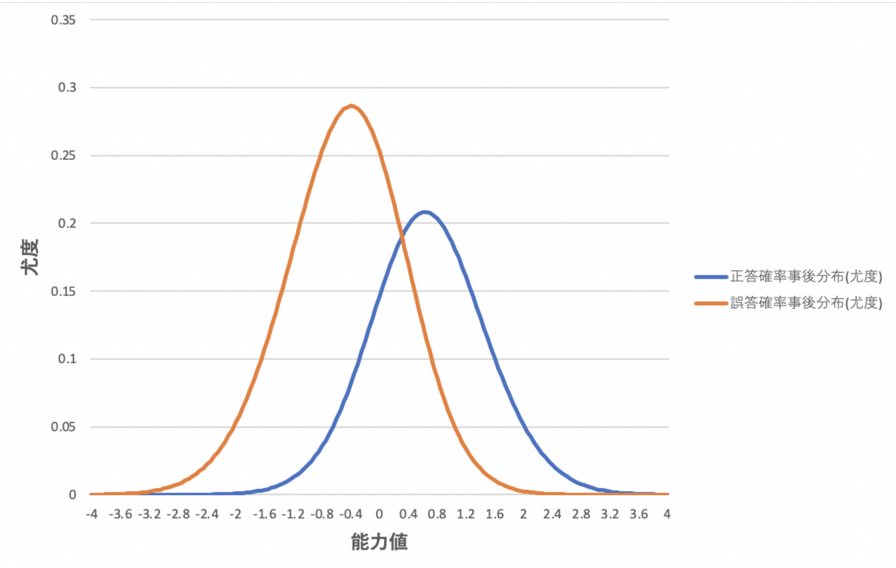

図 5: Item21 の正答確率事後分布と誤答確率事後分布のグラフ (θ の刻み幅 0.05)

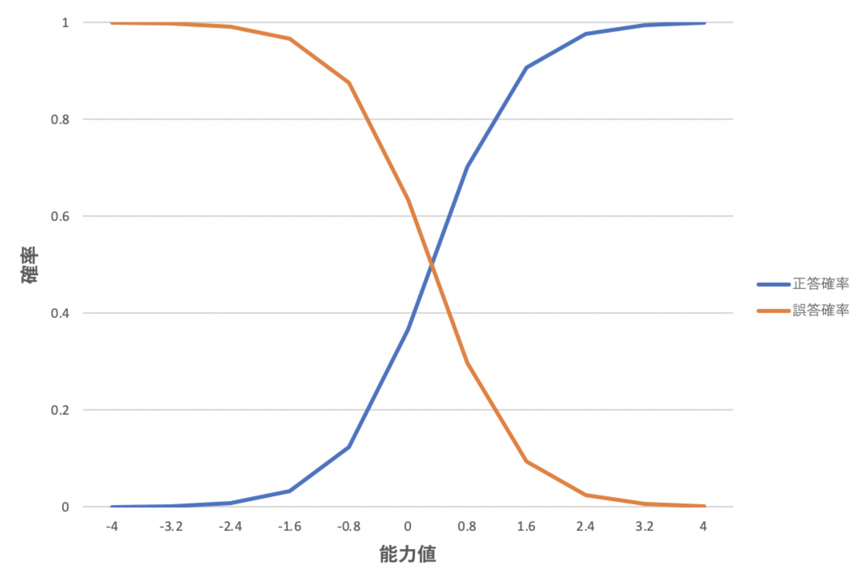

図 6: Item21 の正答確率と誤答確率のグラフ (θ の刻み幅 0.8)

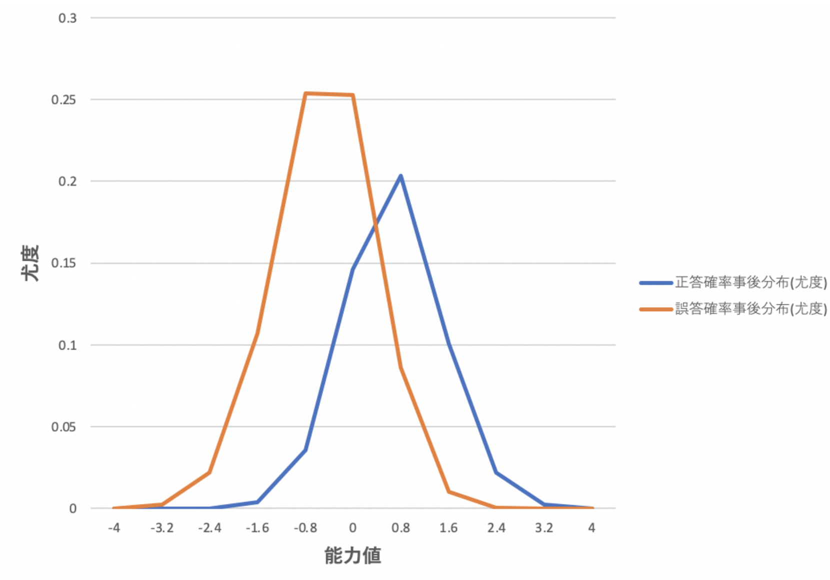

図 7: Item21 の正答確率事後分布と誤答確率事後分布のグラフ (θ の刻み幅 0.8)

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。

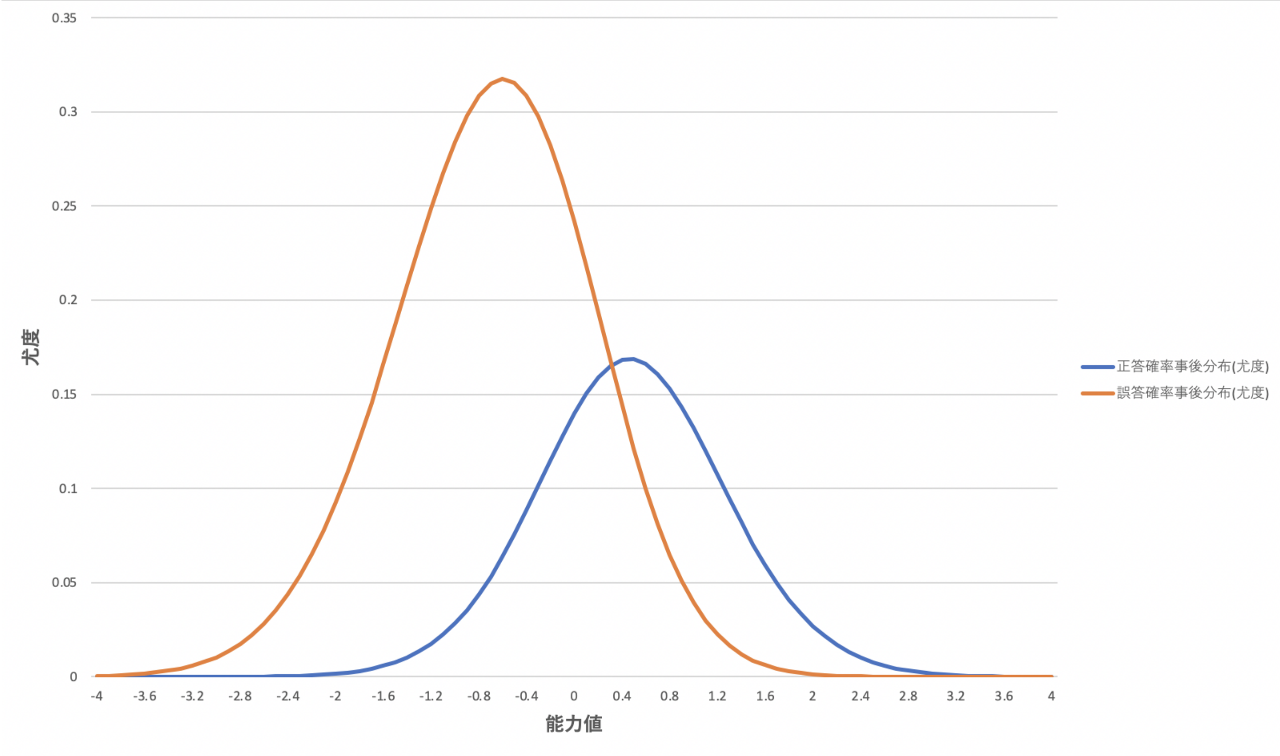

図 8: Item21 の正答確率事後分布と誤答確率事後分布のグラフ (能力値は平均-0.3, 分散 1 の正規分布に従う)

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。