教師の実態把握力を解析する数理モデルに関する考察

－KL 情報量との比較を中心に－

〇上西秀和*・植阪友理**・山口一大**・中川正宣***

*獨協医科大学　基本医学情報教育部門、**東京大学　大学院教育学研究科、

***東京工業大学　大学院社会理工学研究科

１．はじめに

教師が授業を計画する際に、生徒にとって何がどのくらい難しいのかという実態を事前に推し量り（「困難度査定」と呼ばれる）、それらを授業設計に生かすことの重要性が指摘されている（市川、2014）。一方で、「もっと分かっていたと思ったのに、テストをやってみたら、全く分かっていなかった！」など、教師の予測が大きくずれることも少なくない。こうした予測は、教師や課題などによっても異なると考えられる。本プロジェクトでは、表１に示すようなデータ構造を想定し、教師が予測した離散確率分布と生徒の実態を示す離散確率分布のあり方をパラメータ α として表現する、以下のような数理モデルを開発している（植阪・中川、2012）。このモデルでは、課題iにおいて、教師jがカテゴリkについて予測した割合をQ_{ijk}、課題iにおけるカテゴリkに属している児童生徒の割合をP_{ik}として表している（なお、β については、$\sum_k Q_{ijk} = 1$であるという特徴から、α の関数として表現することが可能である。このためパラメータは α のみと捉えている）。

$$Q_{ijk} = \beta_{ij} \, {P_{ik}}^{\alpha_{ij}} \tag{1}$$

表１　データ構造の例

課題1 計算問題（データは人数）

カテゴリ	子どもの実態	教師A	教師B	教師C
～10 点	10	5	15	30
11～20 点	20	10	20	25
21～30 点	30	60	25	10
31～40 点	20	10	20	25
41～50 点	10	5	15	30

課題2 文章問題（データは人数）

カテゴリ	子どもの実態	教師A	教師B	教師C
～10 点	60	80	50	10
11～20 点	40	50	45	20
21～30 点	30	15	30	30
31～40 点	20	10	20	40
41～50 点	10	5	15	60

このモデルでは、教師の予測した確率分布Q_{ijk}を児童生徒の実態を示す確率分布P_{ik}の関数と捉え、α は両者の関係性を記述するパラメータとなっている。このパラメータの値と、確率分布間

の関係性については、植阪ら（印刷中）や仲谷ら（印刷中）を参照されたい。

パラメータ α は、（課題の影響は受けているものの）ある教師の予測の仕方の特徴を記述するパラメータである。それと同時に、別の見方をすると、本数理モデルのパラメータ α は、2 つの確率分布の類似度を評価する指標とも捉えられる。２つの確率分布の比較を行う場合に、情報科学の分野では一般的に、以下の式で示されるような KL（Kullback-Leibler）情報量が用いられる。この指標は、２つの確率分布間（この場合には、P_k、Q_k）の一般的な距離を表現している。

$$KL_P = \sum_k P_k \log \frac{P_k}{Q_k} \tag{2}$$

本稿では、α と KL 情報量との関係性について数理的に考察し、α が従来の KL 情報量では十分に区別できなかった情報を区別しうる指標であることを明らかにする。

２．α と KL 情報量との関係性の解明

以下に述べる証明を通じて、ある KL 情報量に対して、複数の α が対応しており、α を一意に定めることができないことが明らかとなった。すなわち、本数理モデルで提案した α は、従来用いられてきた KL 情報量では区別できなかった側面を表現することができていると考えられる。本節では、ある KL 情報量に対して、複数の α が対応しており、α を一意に定めることができないことについて、数学的に証明する。

事前の式展開　分析に先立ち、いくつかの式変形を行う。なお、以下では、ijの添字を省略して表記する。まず、本モデルにおいてQ_kはカテゴリ kに関して教師が行った予測（割合）であり、$\sum_k Q_k = 1$である。この制約を考慮し、かつ$z = \sum_k e^{-\alpha E_k}$、$E_k = -\log P_k$とおくと、(1)は以下のようにかける（なお、$z$は β の逆数となっている)。

$$Q_k = \frac{e^{-\alpha E_k}}{z} \tag{3}$$

次に、教師が予測した確率分布の平均点について考える。平均をEとすると、$E = \sum_k E_k Q_k$となる。ここで、$\log z$を α で偏微分すると、以下のように$-E$となることがわかる。

$$\frac{\partial}{\partial \alpha} \log z = \frac{1}{z} \frac{\partial}{\partial \alpha} \sum_k e^{-\alpha E_k} = -\sum_k E_k \frac{e^{-\alpha E_k}}{z} = -E \tag{4}$$

さらに、教師が予測した確率分布の分散を考える。V_Eは分散なので、$V_E \geq 0$である。また、$\sum_k Q_k = 1$に注意すれば以下のようにかける。

$$V_E = \sum_k (E_k - E)^2 Q_k$$

$$= \sum_k E_k^2 Q_k - (\sum_k E_k Q_k)^2$$

$$V_E = \frac{\sum_k E_k^2 e^{-\alpha E_k}}{z} - \frac{(\sum_k E_k e^{-\alpha E_k})^2}{z^2} \tag{5}$$

また、後ほど、証明の中で利用する関係から、Eの α による 1 階微分を考える。すると、以下の展開からEの 1 階偏微分は、0 または負の値となる。

$$\frac{\partial E}{\partial \alpha} = \frac{\partial}{\partial \alpha}\sum_k E_k Q_k = \frac{\partial}{\partial \alpha}\frac{\sum_k E_k e^{-\alpha E_k}}{z}$$

$$= \left(\frac{\partial}{\partial \alpha}\frac{1}{z}\right)\sum_k E_k e^{-\alpha E_k} + \frac{1}{z}\frac{\partial}{\partial \alpha}\sum_k E_k e^{-\alpha E_k}$$

$$= \left(-\frac{1}{z^2}\frac{\partial z}{\partial \alpha}\right)\sum_k E_k e^{-\alpha E_k} - \frac{1}{z}\sum_k E_k^2 e^{-\alpha E_k}$$

$$= \left\{-\frac{1}{z^2}\left(-\sum_k E_k e^{-\alpha E_k}\right)\right\}\sum_k E_k e^{-\alpha E_k} - \frac{\sum_k E_k^2 e^{-\alpha E_k}}{z}$$

(5)より

$$= \frac{(\sum_k E_k e^{-\alpha E_k})^2}{z^2} - \frac{\sum_k E_k^2 e^{-\alpha E_k}}{z} = -V_E \leq 0 \tag{6}$$

ある KL 情報量に対し、αが一意に定まらないことの証明　以上を用いて、KL 情報量と α を関係づける。より具体的には、KL 情報量を α の関数で表現した後、2 階偏微分を求め、KL 情報量が下に凸の関数になっていることを示す。

まず、P のエントロピーを$S_P = -\sum_k P_k \log P_k$とおき、$KL_P$を$S_P$を用いて表すと、

$$KL_P = \sum_k P_k \log P_k - \sum_k P_k \log Q_k$$

(3)を代入して

$$= -S_P - \sum_k P_k(\log e^{-\alpha E_k} - \log z)$$

$$= -S_P - \alpha \sum_k P_k E_k + \log z \sum_k P_k$$

$$= -S_P + \alpha S_P + \log z$$

$$= (\alpha - 1)S_P + \log z$$

ここから、KL_Pの α による 2 階偏微分を求める。

$$\frac{\partial KL_P}{\partial \alpha} = S_P + \frac{\partial}{\partial \alpha}\log z$$

(4)より

$$= S_P - E$$

（$\alpha = 1$のとき$S_P - E = 0$）

(6)より

$$\frac{\partial^2 KL_P}{\partial \alpha^2} = -\frac{\partial E}{\partial \alpha} = V_E \geq 0$$

2 階偏微分が正であるから、α と KL_P の関係性を示すグラフは図 1 のように下に凸となる。

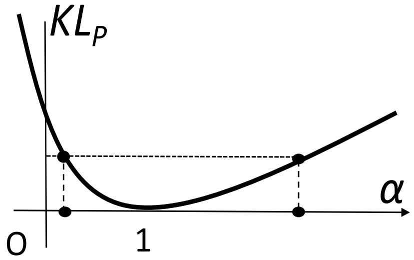

図１　α と KL_P の関係性を示すグラフの概形

このグラフからも、ある KL 情報量に対して、複数の α が対応しており、α を一意に定めることができないことが明らかである。よって、KL 情報量では区別できなかった側面を α は表現していると考えられる。これは、パラメータが示していることとつきあわせてみても納得ができる結果である。KL 情報量は確率分布間の単純な距離を表している。一方で、本モデルのパラメータ α は、本学会において植阪らが論じているように、「高い部分をより高く、低い部分はより低く評定している場合」（$\alpha > 1$）と、「高い部分はより低く、低い部分はより高く評定している場合」（$\alpha < 1$）とを区別している。確率分布間の距離という視点から見ると等距離にあると考えられる 2 つの分布を、α は分布の特徴の違いによって区別しうる指標なのである。

参考文献

市川伸一（2014）『学力と学習支援の心理学』 放送大学教育振興会

仲谷佳恵・山口一大・上西秀和・植阪友理・中川正宣（印刷中）「Web システム”Wits”による教師の実態把握力の解析–算数学力・学習力診断テスト COMPASS を用いた検討–」,『日本テスト学会第 13 回大会発表論文抄録集』

植阪友理・仲谷佳恵・山口一大・上西秀和・中川正宣（印刷中）　「教師の実態把握力を評価する新たな枠組みの提案－新たな数理モデルの開発とパラメータの意味－」,『日本テスト学会第 13 回大会発表論文抄録集』

植阪友理・中川正宣（2012）教師の予測の精度を解析する数理モデルの開発とその適用:見過ごされてきた学力・学習力を検出する実証的方法の提案 認知科学, 19 (2), 236-239.

謝辞

本研究は JSPS 科研 15H02924（基盤研究（B)「教師の『みとり』に関する実証的研究と学校現場への展開」（代表、植阪友理)）の助成を得て行われています。

連絡先

上西秀和：hidekazu@dokkyomed.ac.jp　植阪友理：y_uesaka@p.u-tokyo.ac.jp

山口一大：kazz530@p.u-tokyo.ac.jp　中川正宣：nakagawa@nm.hum.titech.ac.jp