

## 期待効用仮説による学習者の自己評価計量技法(6)

- 多肢選択問題への拡張 -

### Quantitation of Learner's Self-Evaluation by Expected Utility Hypothesis (6)

- Expanded Application to Multiple-Choice Questions -

養老 真一<sup>\*1</sup>, 田中 規久雄<sup>\*1</sup>, 西本 実苗<sup>\*1</sup>, 下倉 雅行<sup>\*2</sup>

Shin-ichi YORO<sup>\*1</sup>, Kikuo TANAKA<sup>\*1</sup>, Minae NISHIMOTO<sup>\*1</sup>, Masayuki SHIMOKURA<sup>\*2</sup>

<sup>\*1</sup>大阪大学

<sup>\*2</sup>大阪電気通信大学

<sup>\*1</sup> Osaka University

<sup>\*2</sup> Osaka Electro-Communication University

あらまし：多肢選択問題を対象に、期待効用仮説に基づき、学習者の確信度を自信度として測定する手法を開発した。これにより、学生の理解度を測り、出題の妥当性について検討することができる。

キーワード：教育評価、確信度、期待効用仮説、自信度、出題の妥当性

#### 1. はじめに

従来、学習者の理解度は客観テストによって測るというアプローチがとられてきた。しかし、学習者自身が「自分がどの程度の確信をもって理解できているか」という、いわば主観的な理解度も教育評価にとって重要な情報であると考えられる。そこで、筆者らはこれを計測する手法の開発に取り組んでいる<sup>(1),(2)</sup>。

これまでの我々の研究では設問が2択であるものに限定されていたが、今回の報告ではこれを多肢選択問題に拡張した手法を提案する。

#### 2. 手法

我々が対象にするのは、設問に対して、複数の選択肢から学生が正しいと考える選択肢をひとつ選択するタイプの問題である。学生はかならずしも確信をもって選択肢をひとつ選ぶのではなく、例えば、ある選択肢は「80%くらい正しいだろう」、また別の選択肢は「20%程度正しいだろう」という感覚をもち、そのうち、もっとも正しい可能性が高いものを選びと考える。我々の手法は学生の「正しいだろう」という確信度を、その間が正解である主観確率と見做し、これを測定するものである。

試験は以下のような方法で行う。

- (1) 学生には、設問に対する解答の全ての選択肢に、「自分がその選択肢が正しいという事にどの程度確信を持っているか」という「確信度」を0～100%の範囲で記入させる。
- (2) 各選択肢に記入する「確信度」の合計は100%になるように、学生に指示する。
- (3) 正解の選択肢に「確信度」 $p$ が記入されている時、得点を $\log(p)$ とする。この事は確信度と正解した時の得点对応表の形で学生に明示しておく。

選択肢の数を $n$ とし、学生が各選択肢にもつ確信度を $q_i (i = 1, \dots, n)$ 、また、実際に学生が選択肢に「確信度」として記入する値を $p_i (i = 1, \dots, n)$ とする。一般には、 $q_i$ と $p_i$ は等しいとは限らない。学生の得点の期待値は $\sum_i q_i \log(p_i)$ となるが、学生がこの期待値

を最大にするように行動すると仮定すると、 $q_i = p_i$ となり、学生は自分の確信度をそのまま記入する事になる。この方法により、我々は学生の各選択肢についての確信度を測定する事ができる。

現実の試験においては、この手法はそのままでは使えない。学生がある選択肢に確信度0をつけると、 $\log(0) = -\infty$ となってしまうためである。そこで、得点の与え方を以下のように改良する。

- (1) 学生の記入する確信度を10%刻みや5%刻みの値に制限する。(これを $\delta$ とする)
- (2) 確信度0の選択肢が正解だった場合の得点を、十分小さな値にとる。(これを $A$ とする)

確信度0についての得点を $-\infty$ から $A$ に変更したため、確信度が0%ではない選択肢に、学生が偽って0%を確信度として記入した方が得になる可能性がある。そのような事が生じないように $A$ を決定する。 $A = \log(\delta) - 2$ とすれば、おおよそ問題はない。得点は $A \sim 0$ の範囲になるが、実際の試験ではこれを0～100になるように換算し、学生に示している。

#### 3. 実験

この手法を用いた実験を、大学生45名に対し、20問の設問をだす形で実施した。各設問における選択肢はすべて4つで、学生の記入する確信度は10%刻みとした。

実験結果を解析するのに、

$$N = \sum_i p_i \log_2(p_i)$$

という量を導入する。ここで $p_i$ は特定の学生が選択肢 $i$ にもつ確信度である。この量 $N$ はエントロピーの逆符号になっている。どれかひとつの選択肢に100%の確信度をもてば最大値0になり、逆に、全ての選択肢に等しい確信度を持っていれば、最小値 $-\log_2(n)$ となる。言い換えると、もっとも自信があれば最大になり、全く自信がなければ最小となる。したがって、これは学生のその設問についての自信の強さを表す量だと考えることができる。以後、 $N$ を「学生の自信度」と呼ぶ。選択肢の数が4の時、 $N$ は-2から0の値をとる。

さらに、学生の「素点」を次のように定義する。

学生が最大の確信度を与えた選択肢が正解だった場合、「素点」を1、そうでない場合を0とする。また、最大の確信度を与えられた選択肢が複数あった場合、1をその選択肢の数で割ったものを「素点」とする。この量は「通常の採点方法による学生の点数」と考えてよい。

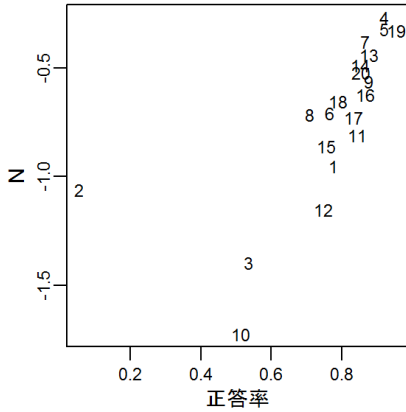


図1 正答率と学生の自信度

図1は、問毎の学生の素点の平均値(これを正答率と呼ぶ)を横軸に、 $N$ の平均値を縦軸にしてプロットしたものである。問2の正答率が5%程度と極めて低い割に、 $N$ は-1.1と比較的大きい。これは学生が「自信をもって間違っている」事を示している。一方、問10は正答率が0.5とそれほどよくはなく、また、 $N$ も-1.7とある程度小さい。これは学生がこの問に対して、自信を持たずに解答し、また、その解答が実際に間違っていることが多い事を示している。

#### 4. 検討

以上は我々の手法により引き出された設問についての情報であるが、以下で実際の問題がどのようなものであったかを確認する。

問2は「次のうち、『個人情報の保護に関する法律』における『個人情報』にあてはまるものを1つ選びなさい」という問題で、選択肢は 1.住所、2.電話番号、3.学籍番号、4.メールアドレスとなっている。正解は2の学籍番号なのであるが<sup>1</sup>、2.の電話番号を正解とした学生が多かった。これは「電話番号は個人情報」という、いわば学生にとっての常識が「間違った選択肢への確信」となっていると考えられる。

問10は「ウイルス対策について、正しい記述を1つ選びなさい」という問題で、選択肢は以下の通りである。

1. ウィルス対策ソフトでは、コンピュータウィルスを検知することはできるが、感染したファイルからコンピュータウィルスを取り除いて修復するには別のソフトウェアが必要である。
2. パターンファイルとは、ウィルスの動きを一時的に抑制するためのプログラムを埋め込んだファ

イルのことである。

3. 「タコイカウイルス」は電子メールの添付ファイルを経由して感染する。感染を防ぐには、添付ファイルの拡張子をよく確認する。
4. Webサイトを閲覧しただけで感染してしまう「ランブラー」は、アプリケーションの脆弱性などを利用して感染する。感染を防ぐには、アプリケーションソフトやウイルス対策ソフトのアップデートを怠らないことである。

考えようによっては、どの選択肢も「間違い」のように思える設問である。半数の学生は正解しているが、それほど自信なく解答したと考えることができる。(正解は4.)

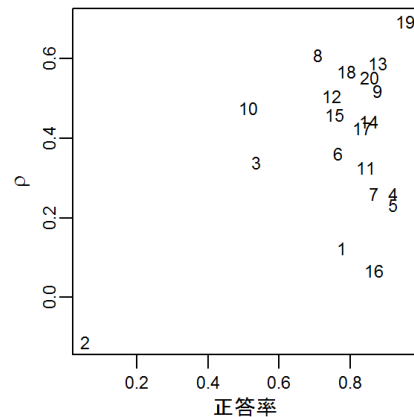


図2 正答率と、素点と自信度の相関係数

図2は、各学生の「自信度」 $N$ と素点の相関係数を $\rho$ として、これを縦軸に、正答率を横軸にとって、プロットしたものである。 $\rho < 0$ は、いわば「自信がない学生ほど、素点が高い」事を意味する。

問2について、この量は-0.1と負になっている。これもやはり、この問について学生が、いわば「間違った自信」をもって解答している事を示している。これもまた我々の方法の有効性を示している。

#### 5. まとめ

我々は、多肢選択問題について、学生の「確信度」を測る手法を提案し、また、これが教師にとって有用な情報を与えうる事を、実験により示した。今後の課題としては、問題のパラエティを広げて実験する事により、この手法の有用性をさらに確認していく事があげられる。

#### 参考文献

- (1) 田中規久雄, 養老真一, 下倉雅行, 西本実苗: 期待効用仮説による学習者の自己評価計量技法(1) - 2択問題を例として -, 教育システム情報学会第36回全国大会講演論文集, pp.388-389 (2011).
- (2) 養老真一, 田中規久雄, 西本実苗, 下倉雅行: 期待効用仮説による学習者の自己評価計量技法(5) - 確信度と正答率から出題の妥当性を量る(続) -, 情報処理学会第75回全国大会講演論文集, vol.4, pp.349-350 (2013).

<sup>1</sup> <https://business.nifty.com/articles/exam/protection/090305/> 「3分スタディ個人情報保護」参照。