

授業評価項目の適応的出題における出題項目と推定精度の関係性分析 The Relationship between the Evaluation Item and the Estimation Accuracy in Adaptive Class Evaluation

中村 修也[†] 石井 隆稔[‡] 赤倉 貴子[‡]
Shuya Nakamura Takatoshi Ishii Takako Akakura

1. はじめに

本研究では、学生による授業評価アンケートの項目を、授業に応じ適応的に出題する方法を提案する。

著者らは、数題の回答結果から出題していない項目の回答結果を確率的に推定可能な階層型アイテムバンクをベイジアンネットワーク[1]により構築した[2]。これにより、少ない出題項目数でも、多くの回答結果を教員に提供できる可能性を示した。本研究では、毎回の授業で階層型アイテムバンクを利用することを想定し、階層型アイテムバンクを発展させる。特に、収集された授業評価結果及び推定結果の推移に着目し、推定精度が良い適応的な出題方法について検討する。先行研究[3]では、「出題項目を選択する方法」として、授業評価の推移から回答分布の変化が最も大きい項目を選択候補とする方法、「推定方法」として、その授業で得られた過去の回答・推定結果を次の推定にも用いる方法を提案し、シミュレーション実験により、推定精度の良い出題方法について検討した。一方で、分析対象とした授業は 1 科目のみであり、また、提案法の推定精度に関する検証は十分ではなかった。

そこで、本稿では分析対象とする科目を増やし、先行研究で提案された「出題項目選択方法」と「推定方法」ごとの推定精度の差を統計的検定により示すことを目的とする。

2. 先行研究

2.1 階層型アイテムバンクの構築

著者ら[2]は授業評価アンケートの項目間の関係性を、ベイジアンネットワークの構造学習によって示した。ベイジアンネットワークとは複数の確率変数の間の定性的な依存関係をグラフ構造によって表し、個々の変数の間の定量的な関係を条件付確率で表した確率モデルである。複雑な因果関係を事前確率と事後確率によって推定し、不確実性を伴う事象の推定に用いることが可能である[1]。著者らは、実際の調査データに基づき、収集した 75 項目の授業評価項目の確率的な因果関係を構造学習により明らかにした。また、確率的な推定により、数題の出題から、出題していない項目の回答結果を推定することが可能であることを示した。

2.2 経時的な変化を利用した適応的出題方法について

先行研究[3]では、工学系学科の統計の授業で全 8 回収集した 8 項目の授業評価アンケートの回答結果を、前項で述べた階層型アイテムバンクを利用し、シミュレーション実験を行った。シミュレーション実験の際には、8 項目中 2 項目を出題した時、残りの 6 項目の推定結果と実際の回答

[†] 東京理科大学大学院工学研究科経営工学専攻

[‡] 東京理科大学工学部情報工学科

結果との分布間の擬似距離（非類似度）を Jensen-Shannon ダイバージェンス[4]（以下 JSD）により評価した。適応的な出題方法として、出題項目を選択方法 2 パターン(I, II)と、推定方法 2 パターン(A, B)それぞれの組み合わせ計 4 パターンの結果を比較した。

- I. 授業の回答・推定結果の分布と、階層型アイテムバンクにある事前分布との JSD が、最も大きい項目を次の出題候補とする。
- II. 前回授業と今回授業の回答・推定結果との JSD が最も大きい項目を出題候補とする。
 - A. これまでに構築した階層型アイテムバンクを推定に用いる。
 - B. 前回授業の回答から階層型アイテムバンクに基づく標本を生成、条件付き確率を生成した標本の分布に書き換えた階層型アイテムバンクを推定に用いる。

初回の授業で出題する項目の組み合わせを変えながら、シミュレーション実験を行った結果を図 1 に示す。

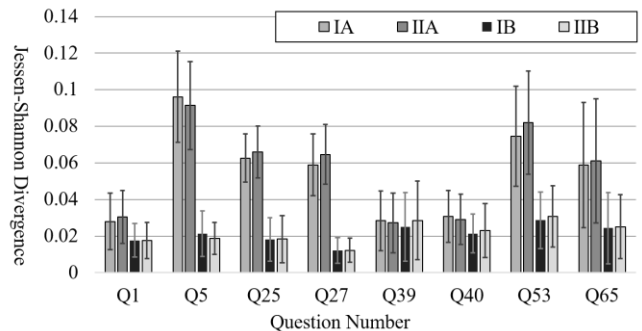


図 1 シミュレーション実験結果[3]

図 1 の縦軸は JSD の平均値とその標準偏差、横軸は項目を表している。各棒グラフは出題方法のパターンを表しており、各組み合わせを便宜上、「IA・IIA・IB・IIB」と記述している。図 1 より、推定方法 B の推定方法の JSD が 0 に近く、推定結果が実際の回答結果によく似ていることが示された。以上の結果から、推定方法 B の、その授業の過去の回答結果も利用した推定方法が、推定精度の良い方法であることが示唆された。

3. シミュレーション実験

3.1 実験概要

前節で示した通り、先行研究では分析対象が 1 科目であり、また、推定精度に関する検討は平均値と標準偏差の概観により示しており、推定精度の検証として十分とは言えなかった。そこで、本稿では、以下の様に適応的な出題方法の検証について見直す。

先行研究と同様の出題項目選択方法, 推定方法を踏まえ, 出題項目選択方法に「Ⅲ. 常にランダムな出題」を加えた。以下, 出題項目の出題方法 3 パターン(I, II, Ⅲ)と, 推定方法 2 パターン(A, B)それぞれの組み合わせ計 6 パターンについてシミュレーション実験を行う。さらに, 分析対象とする科目数を 10 とし, 計 76 回分の授業で収集された 7~9 項目の授業評価アンケート結果を用いて, 各出題方法を適用した。評価方法は先行研究と同様の JSD を用いた。6 パターンの JSD の平均値に対して多重比較における, 平均値の差の検定 (ノンパラメトリック検定 Games-Howell 法, $p<0.01$) を適用し, 推定精度が良い出題方法について考察する。

3.2 シミュレーション実験結果

図 2 に, 統計系授業を対象としたシミュレーション実験結果の例を示す。図 2 は「Q25. あなたは授業に対して意欲的でしたか」についての推定結果を示す。

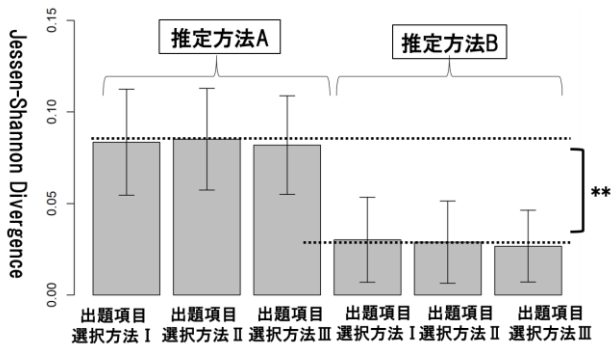


図 2 シミュレーション実験結果の例

図 2 より, 平均値は推定方法 A よりも B の方が低く, 推定精度が良い傾向にある。また, 多重比較の結果, 推定方法 A の出題項目選択方法 (I, II, Ⅲ) と推定方法 B の出題項目選択方法 (I, II, Ⅲ) 各々について, 有意差がある場合がほとんどであった。

また, この結果を踏まえ, 推定方法 B について, 出題項目選択方法 (I, II, Ⅲ) の中で推定精度が良い方法について検討する。そこで, 多重比較の結果, 有意差があり, かつ JSD の平均値が低い出題項目選択方法を集計した。集計方法を図 3 に示す。

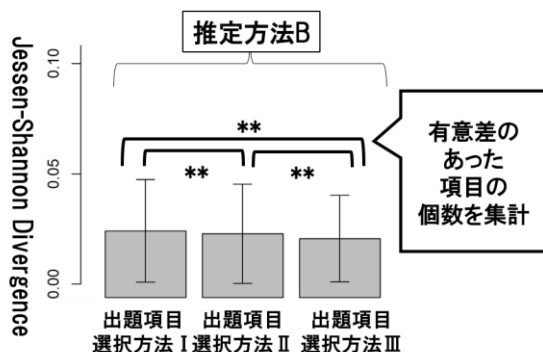


図 3 推定方法 B における出題項目選択方法間の JSD の有意傾向比較図

図 3 に示すように, 出題項目選択方法 (I, II, Ⅲ) 間で有意差がある項目を集計する。その際には, JSD の大小関係に着目し, 分類する。分類した結果を表 1 に示す。

表 1 出題項目選択方法間の有意傾向

①	個数	②	個数	③	個数
有意差なし	43	有意差なし	60	有意差なし	47
I B < II B	6	II B < III B	0	I B < III B	3
I B > II B	15	II B > III B	4	I B > III B	14

表 1 より, どの組み合わせでも有意差がない場合が半分以上を占めている。また, 出題方法Ⅲのランダムな出題方法について, 3 つの出題項目選択方法の中でも, 有意に JSD が小さい (推定精度が良い) 頻度が最も多く, 逆に, 有意に大きくなる頻度は最も少ないことがわかった。

3.3 考察

図 2 に示す通り, 推定方法 B の方が JSD の平均値は小さく, 推定方法 A との有意差も確認できた。この傾向は, 全ての授業評価項目について見られた。この結果から, 推定方法 B の方法による回答結果の推定は, 推定方法 A よりも推定精度が良いことを統計的に示された。

また, 出題項目選択方法について, ランダムな出題 (Ⅲ) の推定精度が良い傾向であった。これは, 出題する項目はできるだけ偏りなく出題することで, 結果的に階層型アイテムバンクにある項目全体を精度良く推定可能であると考えられる。そのため, 先行研究[3]及び本稿で検討した, その授業で過去に得た回答結果と推定結果の分布の差分を利用し出題項目を決定する方法については, 別のアプローチをとる必要があると考えられる。

4. まとめと今後の課題

本稿では, 授業評価アンケートの項目間の関係性をベイジアンネットワークにより構築した階層型アイテムバンクを利用し, 授業評価の経時的な変化を利用した適応的な出題方法について検討した。特に, 推定精度が良い出題方法について, その授業の過去の回答結果も利用し推定する方法の推定精度が良い傾向にあることをノンパラメトリックな多重比較により示すことができた。また, 出題項目を選択する方法については, ランダムな出題をすることで, 階層型アイテムバンクにある項目全体を精度良く推定できる可能性が示唆された。

一方で, ランダムな出題では, 教員の要望や要求を反映した出題ができない可能性がある。そこで, 今後は出題のルールを設計することや, 回答の分布ではなく, 階層型アイテムバンクの構造に着目し, 出題方法を決定する方法について検討したい。また, 階層型アイテムバンクを利用した授業改善支援システムの開発・実験を今後予定している。

謝辞

本研究の一部は, 平成 28~31 年度科学研究費補助金基盤研究 (B) (課題番号 16H03086: 研究代表者赤倉貴子) の助成によるものである。

参考文献

- [1] 植野真臣, “ベイジアンネットワーク,” コロナ社, 東京, 2013.
- [2] 中村修也, 石井隆稔, 藤森進, 赤倉貴子, “適応的授業評価アンケートシステム開発に向けた階層型アイテムバンク,” 電子情報通信学会技術研究報告, vol.115, no.492, pp.159-164, Mar. 2016.
- [3] 中村修也, 石井隆稔, 赤倉貴子, “授業評価の経時変化を利用し授業評価項目を適応的に出題する方法,” 電子情報通信学会技術研究報告, Vol.117, No.65, pp.55-58, May, 2017.
- [4] D.M.Endresand, J.E.Schindelin, “A New Metric for Probability Distributions,” IEEE TRANSACTIONS ON INFORMATION THEORY, vol.49, no.7, pp.1858-1860, 2003.