

1. はじめに

見取りは、授業中に子どもの学びを観察し、その変化を的確に把握する伝統的な方法であり、特に道徳科においては、子どもの授業中の反応や意識の変化を重視する評価が求められる。デューイが提唱した教育理念にもあるように、授業は「学ぶ側（児童・生徒）が伝えられたことに対して反応するかどうかにかかっている²⁾」とし、学ぶ側の反応があるかどうかは授業成立の要件であることを述べている。すなわち、学びは、学ぶ側の経験や意識変化を通じて成立するものであり、道徳科では、授業・教材を通じて子どもがどのように感じ、考えたかといった反応が重要な評価の一指標となる。

子どもの見方・授業観への教師自身の
リフレクション

↑

新たな道徳科授業評価方法

教師の見取り支援

教師の見取り

教師間の対話
補完・融合

統合的な見取り
(子どもの学習状況)

教師間の対話
補完・融合

(新たな情報)
テキスト
マイニング
[探索的概観]

(新たな情報)
ベイズ統計
[検証的確認]

補完・融合

(共起ネットワーク・対応分析) (ベイズ統計【ベイズ推定】)

テキストマイニングツール
(KH Coder)を用いて、授業感想文の可視化から児童の学習状況を探索し、主な特徴を概観する。また、抽出者と関係の深い児童の表現から特徴的な反応を読み取る。

KH Coderによって質的コーディング(畜みたい道徳の基本的な資質として自己理解、他者理解、判断力、心情)を行い、児童と授業における反応の有無の関連についての分割表を作成し、Excel(VBA)によって授業における児童全体および個々の児童の反応の可能性をベイズ推定する。

11

多忙な教師でも取り入れやすい評価手法として、見取り、テキストマイニング⁴、ならびに最近、脳科学、心理学、AI および医療等の分野でも活用されているベイズ統計（ベイズ推定）^{5,6}を組み合わせた新しいアプローチを提案する。ベイズ推定では、児童の授業中の反応の可能性を確率（正確には相対的な確からしさ）によって把握し、教師にとって見えにくい学びの一面を明らかにする。具体的には、「見取り」を含めて「探索的概観」、「検証的確認」の3段階を重視しながら、「見取り」を「探索的概観」と「検証的確認」によって支援することで、道徳科にふさわしい多面的な評価を実現し、教師自身の授業観や教育観の再構築を促すものである。結果的に、この統合的な支援システムは、教師へのリフレクションとして活用することで授業の質の向上に役立てる。このことは、教師が児童の学びをより深く理解し、単に授業改善に資するだけでなく、教師の成長を包括的に支援するための新たな取り組みである。

特に、本稿では、「見取り」を支援するための統合的なシステムにおける「検証的確認」段階において、授業における児童の「反応の可能性」をベイズ推定およびベイズ推定後の残差分析が活用できることを既報¹の結果との比較も含めて明らかにする。さらに、「検証的確認」段階のベイズ推定と授業における児童の反応の可能性に関する評価が特別なプログラミングやソフトウェアの利用を必要とせず、Excel（VBA）のみを用いて簡素化できることを示す。

2. 基本データと分析方法

2. 1 基本データ

ここでは、基本データとして、2020年度、A小学校の2、4、5年（どの学年も欠席者を含めると最大30名）において、道徳科における13授業（13教材[2年の授業：教材①たったさんびきだけのいけ②ぐみの木と小鳥③お月さまとコロ④ないた赤おに⑤森のともだち、4年の授業：教材⑥絵葉書と切手⑦雨のバス停留所で⑧クワガタと少年⑨すれちがい⑩お母さん泣かないで、5年の授業：教材⑪ロレンゾのともだち⑫ぼくは伴走者⑬どうするハインツ]⁷）における児童の授業感想文（制約のない自由記述）を収集した。13授業中、本稿で主に考察する4年の授業における教材の概要を図2に示す。図2中のC-11、B-9などは、平成27年7月文部科学省 小学校学習指導要領解説特別の教科 道徳編に記載されている内容を授業における教師のねらいとして示している。

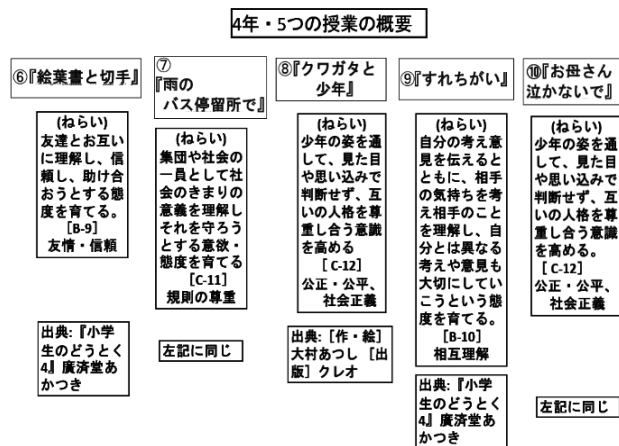


図2 4年道徳科の授業における教材の概要

ちがい⑩お母さん泣かないで、5年の授業：教材⑪ロレンゾのともだち⑫ぼくは伴走者⑬どうするハインツ]⁷）における児童の授業感想文（制約のない自由記述）を収集した。13授業中、本稿で主に考察する4年の授業における教材の概要を図2に示す。図2中のC-11、B-9などは、平成27年7月文部科学省 小学校学習指導要領解説特別の教科 道徳編に記載されている内容を授業における教師のねらいとして示している。

まず、これらの授業感想文について、テキストマイニングツール KH Coder3(2022 年度当時)を用いて、図 1 の「探索的概観」段階では、抽出語リスト、共起ネットワーク分析および対応分析を行った⁸。

これらの分析結果と考察については、既報¹において詳説している。つぎに、道徳科の授業において、児童に育みたい道徳性に係わる資質・能力をコード(コード 1: 自己理解、コード 2: 他者理解、コード 3: 判断力、コード 4: 心情)として 4 つのコンセプトを設定した。また、クロス集計表(分割表)を作成するために、基本データである道徳科授業における児童の授業感想文を対象に、著者と複数の教師とのブレインストーミングを通して抽出語リスト、共起ネットワークによって語と語の関係や抽出語の原文脈に戻ること(KWIC コンコーダンス⁹)によってコーディングルールを作成した(事例として表 1)。

一般にテキストマイニングにおけるコーディングルールは、まず、対象とする文章(ここでは全児童の授業感想文)における各コードの構成語を選定することである¹⁰。なお、この構成語は、単に単語として捉えるのではなく、文脈の中で、そのコードを代表するものであるかどうかを十分に吟味したうえで選定されている。このようにして選定された構成語(表 1、2 年授業②ぐみ木と小鳥*自己理解(コード 1))を例にすると、児童の授業感想文において『自分 or ふしぎ or すき or お礼 or だいじょうぶ or 心のこる』のいずれか一つ以上が含まれている「文」の数を数えることをコーディングと呼んでおり、各児童の授業感想文は、「文」単位でコーディングルールを適用し、コードごとにコードに当てはまる(あるいは当てはまらない)「文」の数を数える。この「文」の数に応じて、学習者(本稿では児童)は、授業感想文を通して授業に反応していると想定できる。

要約すると、コーディングルールに基づくコーディングとは、コードごとに各児童の授業感想文において、コードを構成する語を含まない(「0」: コードに当てはまらない)「文」

表 1 コーディングルール例(2 年授業: 教材②ぐみの木と小鳥における構成語、語と語は or で接続)

*自己理解(コード1)					
自分	ふしぎ	すき	お礼	だいじょうぶ	心のこる
*他者理解(コード2)					
友だち	ともだち	カッコイイ	人	わかる	知人
分かる	やくそく	ともだち	思い	聞く	話
*判断力(コード3)					
考える	ゆうき	出せる	行く	行きたい	病気
びょうき	びょう気	あらし	嵐	元気	助ける
*心情(コード4)					
心	やさしい	しんぱい	かわいそう	ふりしぼって	いっしょうけんめい

表 2 一般的な分割表の例(29 行 × 2 値列)(4 年授業⑦雨のバス停留所でコード 1: セル[灰色地]の数値は観測度数)

児童 (順不同)	「0」 (コード1に 当てはまらない)	「1」 (コード1に 当てはまる)	ケース数 (文の数の和)
1	1	2	3
2	5	0	5
3	1	2	3
4	1	2	3
5	2	3	5
6	2	0	2
7	2	1	3
8	0	1	1
9	4	1	5
10	0	1	1
11	3	1	4
12	2	2	4
13	2	1	3
14	1	3	4
15	0	3	3
16	0	1	1
17	1	1	2
18	1	2	3
19	2	1	3
20	1	1	2
21	3	2	5
22	2	2	4
23	1	3	4
24	2	2	4
25	1	1	2
26	1	0	1
27	1	2	3
28	2	2	4
29	4	1	5
合計	48	44	92 (文の総数)

の数とコードを構成する語を含む(「1」:コードに当てはまる)「文」の数を数えて、ケース数「文」の数の和」を求めることである。その結果、個々の児童とコードについて、4つの分割表(表2は、4年授業⑦のコード1の場合で、セル「灰色地」の部分は、観測度数[「文」の計数値])を作成する。これらの分割表では、コードごとに、そのコードを構成する語が含まれている「文」(つまりコードに当てはまる「文」)の数、あるいは、含まれていない「文」(すなわちコードに当てはまらない「文」)の数を数えており、重複して数えることはなく、それぞれ独立した「文」である。すなわち、「文」の総数は、統計学における試行回数に相当し、コードに当てはまる「文」の数は、独立の試行(ベルヌーイ試行)の結果(成功回数に相当)であり、当てはまらない「文」の数は、失敗回数に相当することから、統計モデルとして二項分布(二項モデル)を設定できる¹¹。なお、作成した4つの一般的な分割表は、KH Coder3特有の分割表の画面表示(各コードの「1」の列のみをまとめて表示¹²)とは異なっている。

作成したコードごとの分割表を対象に、図1の「検証的確認」段階におけるベイズ推定は、適切な統計モデル(本稿では二項モデル)において、ベイズの定理¹³を適用し、事前分布(事後分布を推定する前の確率密度分布)から事後分布(推定したい確率密度分布)¹⁴をサンプリング(疑似乱数発生)¹⁵によって母集団を生成する。したがって、ベイズ推定では、頻度統計(χ^2 検定)が適用できない制約条件¹⁶(データ数が少なく、0度数の存在など)があっても分析が可能である。また、ベイズ推定は、適切な事前分布を仮定し事後分布を推定することによって仮説の正しさ(分析者の信念)を確率で表すことができるとされている。ただし、主観的な要素や不確実性の多い統計分析であり¹⁷、ベイズ推定の事後分布の意味を解釈することは難しいともいわれている。

2. 2 分割表のベイズ仮説検定

既報¹では、授業における全体的な児童の反応と個々の児童の反応について、それぞれベイズ仮説検定(ベイズ検定:一般にベイズファクターによる仮説検証)とベイズ推定に分けて分析・考察した。すなわち、分割表のベイズ推定を行う前に、JASP¹⁸(アムステルダム大学心理学部を中心に開発)という統計分析ソフトウェアを用いてg行(最大30名の児童数)×2値列(例えば、表2における「0」と「1」)の分割表におけるベイズ検定を行い、コードごとに、授業における全体的な児童の反応の可能性について分析した。ベイズ検定とベイズ推定のどちらもベイズの定理に基づく分析である。ここでは、ベイズの定理についてのみ概説する¹⁹。

ベイズの定理に基づくベイズ推定は(1)式で表現できる。

$$P(p|D) = \frac{P(D|p)P(p)}{P(D)} \quad \dots (1)$$

本研究では、分割表において二項モデルを仮定できることから、Dを観測度数(データ)とすると、(1)式の左辺P(p|D)は、事後分布における母比率(生起確率:成功確率:反応

確率などと呼称)を表し、 $P(p)$ は事前分布、 $P(D|p)$ は尤度、そして $P(D)$ は周辺尤度である。すなわち、事後分布は、事前分布と尤度の積で求められる^{20,21}。この唯一の定理を用いてベイズ推定およびベイズ検定を行うことができる。

JASPによる分割表のベイズ検定(児童全体とコードとの関係性を評価する場合)においては、適切なサンプリング(JASPでは4形式のサンプリングから1つ)を選択する必要がある。

本研究における分割表[例えば、表2のセルの観測度数、各列の合計(列和)・各行の和(行和)および総度数(児童の授業感想文における文の総数)]のいずれも観測(コーディング)後にしか分からないこと²²から、JASPを用いてベイズ検定を行うには、本質的にポアソンサンプリング(ポアソン分布)のみが選択可能である²³。また、この分割表では、統計モデルとして二項モデルを仮定できるが、ポアソン分布は、概ね二項分布で近似できること²⁴から、JASPにおいてポアソンサンプリングを選択して対数表示のベイズファクター($2\log_e BF_{10}$)を計算し仮説検定を行った²⁵。その結果を踏まえて、ベイズ検定によって、13の道徳科授業における全体的な児童の反応の可能性について概ね評価できることを示した²⁶。

しかしながら、この方法では、道徳科授業における全体的な児童の反応の可能性については、JASPなどの専門的なソフトウェアを用いてベイズ検定を行い、個々の児童の反応の可能性については、別途、ベイズ推定を行う必要があることから、必ずしも効率的ではないといえる。ここでは、ベイズ検定を行うことなく、授業における全体的な児童の反応および個々の児童の反応を一貫して分析できる手法について次項以降で述べる。

2. 3 ベイズ推定における残差分析

一般に、分割表のベイズ推定においても χ^2 検定後の残差分析と同様に残差分析を行うことができる。ベイズ推定の場合、残差の代わりに残差分布(残差の確率分布:ベイズ推定の場合も単に残差と呼称)を求めて、分割表において、残差を評価指標として行・列のカテゴリ変数間(個々の児童と各コードの「0」:反応のない列、「1」:反応がある列)の関係をセルごとに解釈することができる。また、残差から求められる連関係数 V は、分割表におけるカテゴリ(行と列)の全体的な関連性の指標として早くから注目されてきた。その理論的な解説と具体的な計算は、(参考文献および注釈)の^{27,28}に詳しいが、ベイズ推定後の事後分布における残差は、基本的に χ^2 検定後の残差の考え方と同じである²⁹。

二項モデル(本稿における g 行 $\times$ 2値列の分割表)の場合、ベイズ推定後の事後分布(母比率 P_{ij})における残差 e_{ij} は、(2)式に示す通り、「 i 行 j 列における観測度数 O_{ij} とベイズ推定による期待度数 $N_i \cdot P_{ij}$ の差を、期待度数の標準偏差 $\sqrt{N_i \cdot p_{ij} \cdot (1 - p_{ij})}$ で割った値」である。表2の例では、(2)式および(3)式において、 $i=1, 2, 3, \dots, a$, $a=29$ (行)[行:各児童]、 $j=1, b, b=2$ (列)[コードごとの2値列(「0」、「1」)]、 N_i : i 行における行和(「文」の数の和:ケース数)である。

$$e_{ij} = \frac{O_{ij} - N_i \cdot P_{ij}}{\sqrt{N_i \cdot p_{ij} \cdot (1 - p_{ij})}} \quad \dots (2)$$

ベイズ推定後の事後分布における母比率 P_{ij} から残差 e_{ij} を求めて、どちらの列のカテゴリ（「0」あるいは「1」）に連関（関係性）があるかどうかを後述の理論的あるいは実用的（相対的）な基準を設定して、「1」と連関があれば、授業において個々の児童に反応の可能性が高いと評価する。このように、残差分析結果から、授業における各児童の反応の可能性を読み取ることができる。また、残差分布から行（児童全体）と列のカテゴリとの全体的な連関（すなわち、授業における全体的な児童の反応の可能性）についても、前述のベイズ検定を行うことなく、(2) 式と、つぎの (3) 式を用いて連関係数 V を計算し評価することができる。なお、 N は母比率 P_{ij} の総度数であり、 $N=1$ になるので (3) 式中の N は省略できる。

$$V = \sqrt{(e^2_{11} + \dots + e^2_{1j} + \dots + e^2_{ab}) / N} \quad \dots (3)$$

(2) 式・(3) 式の計算については、WinBUGS²¹（ケンブリッジ大学生物統計学部門で開発されたソフトウェア）によるプログラム²⁸が公開されており、WinBUGS を使用してベイズ推定における事後分布の残差分析から残差や連関係数 V を求めることができる。

次章では、まず、二項モデルを仮定し、WinBUGS とベイズ推定用の Excel (VBA) の結果を比較し、Excel (VBA) のみで反応の可能性を評価できることを示す。つぎに、ベイズ推定の事後分布における残差分析結果と既報¹で提案した「コードごとの事後分布の中央値の平均値³⁰」を用いる評価方法との比較を通して、個々の児童の授業における反応の可能性を事後分布の要約統計量（平均、中央値、95% 信用区間 [事後分布の 2.5% 点から 97.5% 点の区間] など）²¹の中で外れ値の影響が少ない中央値（評価指標）と、その基準の有効性について再検証する。最後に、13 授業において、授業における全体的な児童の反応の可能性に関する評価は、残差から求められる連関係数 V を用いて一貫的・効率的にできることを示す。

3. ベイズ推定における事後分布の中央値の活用

3. 1 Excel (VBA) と WinBUGS による事後分布の中央値

二項モデルにおけるベイズ推定の事後分布の特徴は、通常、要約統計量によって読み取ることができる。この考え方にに基づき事後分布の特徴を読み取るための指標の中で計算が容易で分布の偏りや外れ値の影響が少ない中央値を用いている。事後分布における母比率（反応確率）の中央値が他の母比率の中央値より大きい、あるいは、ある基準（本研究では、コードごとの事後分布の中央値の平均値）より大きい場合、その反応確率は、相対的に確からしさが高いと解釈する。このことは、事後分布の中央値を用いて反応を評価する場合は、確率そのものではなく相対

的な確からしさの高低によって評価できることを表している。既報¹における授業評価(授業における個々の児童の反応の可能性)では、中央値を評価指標とし、コードごとの事後分布における中央値の平均値を評価基準として適用できることを確認した³⁰。

当初、事後分布は、WinBUGSを活用し、その後、生成 AI (OpenAI 社の ChatGPT4o) のアシストにより作成した Excel (VBA) を用いて推定した。いずれの技法も、ベイズの定理 (1) 式のみを適用して、事前分布 (無情報とし一様分布を適用³¹) と尤度の積から事後分布を推定している。WinBUGS では、MCMC 法 (マルコフ連鎖モンテカルロ法という数値積分法) によって近似的にサンプリング (20000 個の疑似乱数の生成) し、事後分布を計算している^{21,28}。なお、WinBUGS では、多くのベイズ推定用の事例プログラムを活用することができる。しかし、実際にベイズ推定を行うには、このソフトウェアの使用方に一定習熟する必要がある。これを軽減するために、二項モデルを仮定したベイズ推定用の Excel (VBA)

によって事後分布の中央値を計算した。その結果、13 授業の 4 コードごとのベイズ推定 (コードごとの事後分布の中央値) において、WinBUGS と本研究で作成した Excel (VBA) の結果を比較すると、最大差は $\pm 0.7\%$ であり、ほとんど平均 100% で一致した (表 3 に例示)。すなわち本ベイズ推定用の Excel (VBA) のプログラムが一定、正確であると結論できる。この Excel (VBA) では、事前分布と事後分布に共役分布特性³²を利用しており、容易にベイズ推定を行うことができる。その結果、児童の授業感想文のテキストマイニングによる「探索的概観」後、4 つの児童に育みたい道徳的な資質に関するコーディングを的確に行うことができるならば、二項モデルを仮定したベイズ推定用の Excel (VBA) を用いて、分割表のベイズ推定は、迅速かつ効率的にできる。

3. 2 事後分布における残差分析結果の活用

分割表のベイズ推定 (事後分布) における残差分析では、残差も分布になるため、分割表の行変数 (個々の児童) と列変数 (「1」の各セル) の非独立 (連関) についての解釈は、残差における (1) 理論的な基準として 95% 信用区間に 0 を含まない場合 (図 3 において箱ひげ図のひげの最下端が基準線 0 より上の場合)、あるいは (2) 実用的 (相対的) な反応の可能性の基準として、(図 3 において箱ひげ図の箱の中央部分が基準線 [コードごとの残差の中央値の平均値] より上に位

表 3 Excel (VBA) と WinBUGS における点推定 (中央値) の比較、灰色地: 各中央値の平均値より大)

お母さん泣かないで (コード I)	Excel VBA (要約統計量)	WinBUGS (要約統計量)	結果の比較
児童 (順不同)	中央値 E	中央値 W	E/W (%)
1	0.2059	0.2063	100.2
2	0.8698	0.8706	100.1
3	0.8407	0.8409	100.0
4	0.7938	0.7937	100.0
5	0.7924	0.7937	100.2
6	0.4207	0.4214	100.2
7	0.3633	0.3641	100.2
8	0.8409	0.8409	100.0
9	0.0735	0.0741	100.9
10	0.7935	0.7937	100.0
11	0.3551	0.3551	100.0
12	0.5010	0.5000	99.8
13	0.4996	0.5000	100.1
14	0.2655	0.2644	99.6
15	0.5000	0.5000	100.0
16	0.1589	0.1591	100.1
17	0.6870	0.6862	99.9
18	0.4988	0.5000	100.2
19	0.6142	0.6143	100.0
20	0.2933	0.2929	99.9
21	0.2064	0.2063	100.0
22	0.6142	0.6143	100.0
23	0.6867	0.6862	99.9
24	0.5018	0.5000	99.6
25	0.4991	0.5000	100.2
26	0.5005	0.5000	99.9
27	0.5003	0.5000	99.9
28	0.3126	0.3138	100.4
平均値	0.5068	0.5069	100.0

置する場合)、これらのいずれかの基準を満たす場合^{27,28}、「検証的確認」段階では、個々の児童と列（「1」：コードに当てはまる）に連関（関

係性：反応の可能性）があるとする。

これらの基準をコードごとの分割表に適用した結果例（4年授業の教材⑩「お母さん泣かないで」：コード1「自己理解」

の場合）を表4に示す。表4における灰色地のセルの○

印は、事後分布の残差および事後分布の中央値による評価に基づき反応の可能性が高い児童である。

また、表5では、4年授業の教材⑥から⑩において、コードごとに反応の可能性の高い児童数を表している。表5欄(1)における数字は、コードごとの事後分布における残差分析

結果による反応の可能性の高い児童数「例えば、授業⑧において、コード1が14名、コード2が12名、コード3が16名、コード4が11名であること」を示している。欄(2)における事後分布の中央値の平均値による評価も同様である。計(1)と計(2)は、異なる評価方法において各授業で各コードにおける反応の可能性の高い児童数の累計、計(1) / 計(2)は、異なる評価方法において各コード反応の可能性の高い児童数の累計比率を表している。

このように、事後分布の中央値の平均値を基準にして反応の可能性がある児童を読み取ると、コードによっては、僅かに基準を超える（あるいは、超えない）、すなわち、どちらともいえない場合の数が増減する。結果として、表5の「欄(1)と欄(2)」では、各授業におけるコードごとに反応の可能性の高い児童数は、多少増減することになるが、コード全体「計(1)と計(2)」としては大差がないといえる³³。

最後に、授業における全体的な児童の反応の可能性については、コードごとに残差から連関係数Vの中央値を求めて評価した。連関係数Vは0～1までの数値で表され、0.1より大で弱い連関（関連性）を示し、1に近いほど連関が大きくなる²⁷。

表6から連関係数Vの中央値の最小値は、4年授業の教材⑥4年「絵はがきと切手」（薄い灰色地）

表4 事後分布の中央値と事後分布の残差による評価の比較（4年教材⑩お母さん泣かないで、コード1の例）

児童 (順不同)	授業における児童の反応の可能性が高い(○)	
	事後分布の残差による評価	事後分布(中央値)による評価
1		
2	○	○
3	○	○
4	○	○
5	○	○
6		
7		
8	○	○
9		
10	○	○
11		
12		
13		
14		
15		
16		
17	○	○
18		
19	○	○
20		
21		
22	○	○
23	○	○
24		
25		
26		
27		
28		

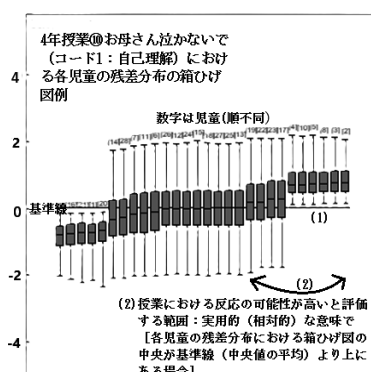


図3 残差分布の箱ひげ図例（4年授業⑩お母さん泣かないで、コード1）各児童を昇順並びで表示

のコード2・コード3の0.445、最大値は、2年授業の教材③2年「お月さまとコロ」(濃い灰色地)のコード1・コード4の0.722である。この13授業では、連関係数Vによって評価すると、児童全体として授業における統計学的な反応の可能性がかなり高いことを示している。この結果は、JASPを用いたベイズファクターによるベイズ検定の結果²⁶と概ね同様の傾向を示して

おり、(3)式に示したとおり、ベイズ推定における事後分布の残差から求められるクラメールの連関係数Vは、授業における児童全体として反応の可能性の評価指標に活用できる。したがって、残差分析からクラメールの連関係数Vを一貫的に求めることができ、別途、ベイズ検定を行う必要はなく、効率的な検証につながると考えられる。

4. おわりに

本稿では、常に教師の見取りを支援し、教師へのリフレクションを促進するため、新たな授業評価支援システムの構築を最終的な目標にしている。

そのために教師にとって大きな負担になりかねない統計学的技法を用いた「検証的確認」段階において、ほとんどの教師が使用経験のないWinBUGSやJASPなどの統計分析用ソフトウェアではなく、Excel (VBA) のみでベイズ推定を行い、ベイズ推定の事後分布の中央値だけで、授業における児童の反応の可能性を推測できることを再確認した。

結果として、Excel (VBA) を用いたベイズ推定によって分析の簡素化とシステム構築の「検証的確認」段階を概ね完成することができたと考える。今後、本システムにおいて、授業感想文のテキストマイニングを中心とした「探索的概観」段階を含めて、児童に育みたい道徳的資質に関する「コーディング」部分を生成AIとExcel (VBA) を組み合わせで自動化し、教師の作業負担をさらに軽減して行きたい。

本研究を実施するにあたり、武庫川女子大学大学院 臨床教育学研究科 前教授(現、昭和女子大学名誉教授)押谷由夫先生ならびに大阪経済大学経営学部教授 金井猛徳先生から

表5 4年全授業教材⑥～⑩における残差分析結果と事後分布の中央値による反応の可能性の高い児童数の比較

授業 (4年)	反応の可能性が高い児童数の比較				
	(1)コードごとの事後分布における残差分析結果による評価	計(1)	(2)コードごとの事後分布における中央値の平均値による評価	計(2)	計(1)/計(2)
⑥	14・13・9・11	47	14・13・12・9	48	1.0
⑦	12・16・17・16	61	15・16・10・10	51	1.2
⑧	14・12・16・11	53	12・12・13・14	51	1.0
⑨	15・15・13・13	56	15・15・13・11	54	1.0
⑩	10・17・16・9	52	10・18・11・10	49	1.1

(1)・(2)における数値は、コード1・コード2・コード3・コード4において反応の可能性が高い児童数を表している。計(1)・計(2)は、各コードにおいて反応の可能性が高い児童数の類計値、計(1)/計(2)は、累計値の比率である。

表6 2年・4年・5年全授業①～⑬授業における各コードの連関係数Vの中央値

学年	授業	コード1	コード2	コード3	コード4
2年	①	0.595	0.593	0.595	0.599
	②	0.649	0.647	0.643	0.649
	③	0.722	0.714	0.719	0.722
	④	0.562	0.561	0.561	0.563
	⑤	0.562	0.563	0.563	0.564
4年	⑥	0.446	0.445	0.445	0.448
	⑦	0.536	0.541	0.542	0.542
	⑧	0.454	0.457	0.455	0.458
	⑨	0.477	0.478	0.478	0.478
	⑩	0.528	0.530	0.528	0.529
5年	⑪	0.542	0.539	0.542	0.541
	⑫	0.624	0.622	0.621	0.621
	⑬	0.522	0.526	0.520	0.527

有意義な示唆をいただき感謝いたします。

(参考文献および注釈) (各 Web サイトの URL は、すべて 2024 年 9 月 1 日に参照・確認)

1. 小山久子 (2024), 「道徳科授業の見取りを支援する評価方法に関する研究—テキストマイニングとベイズ統計の併用—」, 2023 年度博士 (臨床教育学) 学位論文, pp.1-149 の pp.9-25 授業評価方法の構築と提案, 武庫川女子大学大学院 臨床教育学研究科, 道徳科授業の見取りを支援する評価方法に関する研究—テキストマイニングとベイズ統計の併用—
https://mukogawa.repo.nii.ac.jp/search?page=1&size=20&sort=-createdate&search_type=2&q=75
2. Dewey, J. (1916) . *Democracy and education* (デューイ J. 松野安男 (訳) (1975) . 民主主義と教育 (上) 岩波文庫 岩波書店, p.297.)
3. (参考文献および注釈) の 1, p.7 序章 第 4 節 本研究の特徴と独自性, 図 0-1 を修正・加筆.
4. 樋口耕一, 中村康則, 周景龍 (2022), 『動かして学ぶ! はじめてのテキストマイニング: フリー・ソフトウェアを用いた自由記述の計量テキスト分析 KH Coder オフィシャルブック II (KH Coder OFFICIAL BOOK 2)』, ナカニシヤ出版, pp.1-126. 自由記述を対象に KH Coder というテキストマイニングツールによる計量テキスト分析について, 事例を挙げて, その使い方と解釈が詳述されている.
5. ニュートンムック編集部編 (2021) . 『Newton ライト 2.0 ベイズ統計』, ニュートンプレス, pp.2-79. ベイズ統計の基礎, ベイズの定理, 確率の意味, ベイズ統計による暗号解読, 病気の因果関係, Ai の学習および脳のはたらきの解明などが紹介されている.
6. 涌井良幸・涌井貞美 (2016), 『ファーストブック STEP 身につくベイズ統計学』, 技術評論社, pp.12-13. 多様化の時代に応えるベイズ統計学は, 人間の行動や物の動きに対応し, 曖昧さ, 経験, 個々の性質や特質を把握できる可能性があることについて解説されている.
7. (参考文献および注釈) の 1, 附表資料 1, 資料 1 本研究で用いた教科書・教材の概要, pp.1-4. 2021 年 5 月時点の 13 教材の概要を示している.
8. (参考文献および注釈) の 4, 第 3 章 テキストマイニングは意外と簡単そうだ!, pp.22-25, p.38-39, pp.58-65. 抽出語リストとは, 形態素 (名詞, 形容詞, 副詞などの分類) 解析と呼ばれる技法を用いて, 自由記述文 (テキストデータ) から自動的に語を取り出し語数から特徴的な語 (キーワード) を読み取る. 共起ネットワークとは, 語と語の結びつき (共起性) の強さをネットワークで表示し, 共起の多い組み合わせから文章中の主な話題 (トピック) を読み取り, 対応分析とは, 語と外部変数 (例えば, 児童) を平面上に同時付置することで児童の近くにある語を読み取って, 児童の思いを概観する技法.
9. (参考文献および注釈) の 4, pp.24-25. 共起ネットワークの語から, その語の前後の文脈を読み取ることで, 文章中のその語の意味 (使われた方) について確認できることが解説されている.
10. (参考文献および注釈) の 4, 第 9 章 コンセプト (本稿では児童に育みたい 4 つの道徳的資質: 4 つのコード) を取り出すコーディングルール (各コードを構成する語) の作成, pp.73-84.
11. (参考文献および注釈) の 6, 5.2 二項分布とベイズ統計学, pp.124-127.
12. (参考文献および注釈) の 4, 第 10 章 コンセプトを使って分析を深める, pp.85-89. KH Coder を用いて外部変数とコンセプト (コード) の分割表における分析方法と一般的な分割表との相違が解説されている.
13. (参考文献および注釈) の 6, 2.1 ベイズの定理, pp.42-51. ベイズの定理を解説.

14. (参考文献および注釈) の 6, 4.1 ベイズ統計学の基本公式, pp.88-95. ベイズの定理の定式化が解説されている.
15. (参考文献および注釈) の 6, 付録 G MCMC 法のしくみ, pp.222-229. 事後分布において疑似乱数を用いた数値積分法 (MCMC 法) におけるサンプリングの意味が解説されている.
16. 田中敏・中野博幸 (2013), 『R&STAR データ分析入門 5.9 カイ二乗検定の制約と対策』北大路書房, p.68. 分割表の χ^2 検定における分析上の制約条件が概説されている.
17. (参考文献および注釈) の 6, 序章, ベイズ理論の考え方, pp.10-12. ベイズ統計の基本的な特徴 (信念, 曖昧さ, 柔軟性など) について概説されている.
18. 清水優菜・山本光 (2020), 『研究に役立つ JASP によるデータ分析—頻度論的統計とベイズ統計を用いて— 7.2.4 ベイズファクター』, コロナ社, pp.127-137. JASP を用いて分割表の事例におけるベイズの仮説検定の方法が解説されている.
19. (参考文献および注釈) の 1, 第 1 章 道徳科の新たな授業評価方法の構築と提案, 2. ベイズ統計 (ベイズ検定・ベイズ推定) について, (1) ベイズ統計を用いる理由, pp.31-42.
20. 豊田秀樹 (2022), 『統計学入門 I —生成量による実感に即したデータ分析—』, 朝倉書店, pp.1-16.1 データの要約と理論分布において, 事前分布, 事後分布および尤度など, p.39, において 3.2.7 確信区間について解説されている.
21. Lee, M.D. & Wagenmakers, E.J. (2014). *Bayesian cognitive modeling: a practical course*, Cambridge University Press. (リー M.D.& ワゲンメーカーズ E.J. 井関龍太 (訳)・岡田謙介 (解説) (2017). ベイズ統計で実践モデリング—認知モデルのトレーニング— 第一章 ベイズ分析の基礎, pp.2-13. 第二章 WinBUGS ではじめよう, pp.14-30. 北大路書房,) この文献も (参考および注釈) の 20 同様, ベイズの定理の関する用語とベイズの定理に基づく WinBUGS によるベイズ推定の具体例が詳説されている.
 関連補足: 母比率 (ばひりつ) は, 母集団の試行回数に対する成功回数の比で表され, 成功確率や反応確率とも呼ばれる. しかし, 母比率を直接求めることは難しく, 標本集団 (標本比率) から推定する必要がある. 尤度 (ゆうど) は, 仮定した統計モデルで観測データがどの程度尤もらしいかを示すものであり, 例えば天気予報で「晴れの確率 80%」が実際に晴れた場合, その尤度は高いと評価される. 周辺尤度は, 観測データに基づき複数の統計モデルについて全体的な確率を算出し, モデルの良さを示す指標となる. 観測度数は実際のデータで, 期待度数は独立仮定 (統計分析において, 変数間に関連や影響がないこと) に基づき理論的に計算される値である. 標準化された残差 (観測度数—期待度数) が 0 ならば独立 (非連関) と評価され, 正の残差が大きいほど連関 (関係性) が強いと判断される. ベイズ推定では, 残差も分布であるため 95% 信用区間や中央値などの要約統計量で評価する.
22. 倉田博史 (2019). 『大学 4 年間の統計学が 10 時間でざっと学べる』, 角川文庫, 角川出版, pp.180-181. 分割表の総度数, 列和, 行和が観測する前から分っているか, 観測後にしか分からないかによって, その後の分割表の分析に関わる取り扱い方 (本稿では, JASP によるベイズ検定のサンプリングの選択肢) が異なることを概説している.
23. (参考文献および注釈) の 18, 13.2 カイ 2 乗検定の実行 (2) ベイズ的分析表 13.4, p.129. JASP を用いたベイズ検定におけるサンプリングの選択条件が解説されている.
24. 例えば, スイス連邦工科大学チューリッヒ校の資料,
https://ethz.ch/content/dam/ethz/special-interest/mtec/chair-of-logistics-mgmt-dam/documents/courses/14-fall-quality_management/material/Probability_Distributions.pdf

25. (参考文献および注釈) の 18, p.72. 7.2.4 ベイズファクター, ベイズファクターについて定式化を含めて解説されている.
26. (参考文献および注釈) の 1, 第 3 章「新たな道徳科授業評価方法」の有効性の臨床的確認, p.85 の表 3-4, p.99 の表 3-9 および p.111 の表 3-12 において, 学年ごとのベイズ検定結果(ベイズファクターによる仮説検証)を評価基準とともに提示している.
27. 豊田秀樹(2016),『はじめての統計データ分析—ベイズ的<ポスト p 値時代>の統計学—』,朝倉書店, pp.154-160. 6.5.3 および (参考文献および注釈) の 20, pp.187-199. 13.2.1 カテゴリ間の比較, 両文献は, ほとんど同じ内容であるが, ベイズ推定による分割表(カテゴリ A 行×カテゴリ B 列)の事後分布(母比率)において, A 行と B 列が連関している(非独立:つまり関係性がある)状態の程度の指標として, ピアソン残差に関する一般式を提示し, 残差が正のセルは, 独立の場合よりも高い比率で観察されると説明されている. すなわち残差が正になるセルは, カテゴリ A 行(本稿では各児童)とカテゴリ B 列(本稿では 4 つのコードの「1」の列)に関係性のあることが説明されている.
さらに, 個々のセルではなく, 分割表全体における行と列の連関の程度を表す指標として残差から連関係数 V を求める式を示し, V は 0 ~ 1 の間の値になり, その値が大きいほど連関の程度が高いと解釈できること, およびベイズ推定では, 残差と連関係数 V も分布になることが計算事例に基づき説明されている.
28. D.Lunn, C.Jackson et al. (2012), *The BUGS Book: A Practical Introduction to Bayesian Analysis* (Chapman & Hall/CRC Texts in Statistical Science), pp.140-143. イギリスの 12 箇所の病院において, 小児の心臓手術における失敗例の有無(二項モデルを想定)との連関, すなわち, 12 病院(カテゴリ A 行)と失敗有りの数(カテゴリ B 列)との関係性について, WinBUGS によってベイズ推定後の残差分析を行い, ピアソン残差を求めて残差の箱ひげ図から各病院の手術に関する検証を行ったことが解説されている.
29. 生物科学研究所 井口研究室 <https://biolab.sakura.ne.jp/chi-square-residual-analysis.html> において, 頻度論的統計における残差分析について詳細に解説されている.
30. (参考文献および注釈) の 1, pp.34-42 第 1 章 道徳科の新たな授業評価方法の構築と提案, 第 4 節「新たな道徳科授業評価方法」の構築過程—ベイズ推定における指標と基準—, 分割表において, ベイズ推定の事後分布におけるコードごとの中央値を評価指標とし, 中央値の平均値を評価基準とすることを示している.
31. (参考文献および注釈) の 20, pp.25-26. 2.3.5 無情報的事前分布, 2.3.6 事前分布としての一様分布について解説されている.
32. (参考文献および注釈) の 6, pp.154-156, pp.157-165. 6.1 ベイズ統計学と自然な共役事前分布, ベルヌーイ分布, 二項分布の自然な共役事前分布について解説されている.
33. (参考文献および注釈) の 1, pp.105-106. 第 3 章「新たな道徳科授業評価方法」の有効性の臨床的確認の「検証的確認」段階において, ある児童の反応の可能性が低く, あるいは高く評価されても, 「見取り」・「探索的概観」を加えて統合的に再評価することから, 授業に反応している児童やその数について過少あるいは過大評価を避けることができる. この事例として, 例えば, 本稿の表 4 の児童 9 は, ベイズ推定において反応の可能性が低いと解釈されているが, 「見取り」と「探索的概観」において授業感想文の原文を検索すると, 命に関わる心に残るニュースを取り上げ, 児童 9 などの視点から生命の尊さを感じ取っており, 必ずしも反応の可能性が低いとはいえない. このような事例は, 13 授業のいずれにおいても見られ, 教師の「見取り」, 「探索的概観」そして「検証的確認」の 3 段階を補完・融合してこそ, より正確な評価ができることを示している.