

# 順序尺度データの分析方法に関する考察

櫻井 良樹<sup>\*1</sup>, 根本 淳子<sup>\*2</sup>

<sup>\*1</sup> 熊本大学, <sup>\*2</sup> 愛媛大学

## Study on the Methods to Analyze Ordinal Scale Data

Yoshiki Sakurai<sup>\*1</sup>, Junko Nemoto<sup>\*2</sup>

<sup>\*1</sup> Kumamoto University, Ehime University

教育研修の評価データ収集で用いられるリッカートスケールによるアンケートデータに関し、ピアソンの独立性検定を用いた分析方法を検討した。実データサンプルを題材としてこれらを適用した結果、いくつか新たな知見が得られることが示唆された。

キーワード: Learning Analytics, 順序尺度, リッカートスケール, ピアソンの独立性検定

### 1. はじめに

高等教育と企業内研修のどちらでも、教育研修評価のためにアンケートを収集することが多い。これは、アンケートが手軽に短時間で実施でき、実施者（提供者）と学習者（受給者）の双方にとってメリットがあるからである。通常、アンケートの各項目は、1) 選択肢を提示するクローズド・クエスチョン、2) コメント欄を設ける自由記述オープン・クエスチョンの2タイプを使い分ける。近年、Learning Analyticsが注目されるに従い、2)のような非構造化データに対して統計処理を適用する手法もいくつか考案されている。しかし現時点では、まだ1)を対象としたデータ分析が主流を占めている。アンケートの選択肢の設定方式としては、リッカートスケールを採用する例が非常に多く、特に5段階での選択肢提示が一般的である。ここで、リッカートスケールは順序尺度であるため、定量的なデータ分析には十分な注意が必要である。本稿では、筆者等が取り組んでいるアンケートデータ分析を題材として、5段階リッカートスケールによるアンケートデータ分析手法自体に関して検討した。

### 2. アンケートデータ

#### 2.1 対象科目

今回分析するのは、某大学院大学の某大学大学院の入学予定者向けの準備科目である。年1回、毎年同

時期に実施される。履修期間は約3か月で、受講者は毎年10名～20名程度である。

#### 2.2 アンケート情報

分析対象とするアンケートは、上述した科目に関する学習意欲や満足度を問う19の質問からなり、いずれも5段階リッカートスケール(1:全くxxではない, 2:あまりxxではない, 3:どちらともいえない, 4:ややxxである, 5:とてもxxである)を選択肢として提示している。質問の1つは逆転項目であるため、結果を反転させた。なお、アンケートの後半には、別の視点からの質問に関する自由記述欄を設けているが、今回はこれを対象外とする。

アンケートは、科目開始直後と終了後の2回、同じ内容で実施している。

#### 2.3 属性項目

分析対象とするアンケートデータの属性は以下のとおりである。

- (1) 回答者
- (2) 質問番号
- (3) 実施タイミング

このうち、「(1) 回答者」には「(1) 年度」という関連属性がある。具体的には、アンケートを収集した当該カリキュラムの実施年度である2011～2015年度のいずれかに属している、

## 2.4 データ数

今回の分析では、19の「(2) 質問番号」と、事前（以下 Pre と記す）および事後（以下 Post と記す）の2回の「(3) 実施タイミング」全てに関して回答データが収集できた合計66名の「(1) 回答者」を対象とする。なお、「(1) 年度」別の対象「(1) 回答者」数は以下のとおりである。（）内には各年度の総受講生数を示す。

2011年度	11名	(18名)
2012年度	13名	(18名)
2013年度	11名	(17名)
2014年度	12名	(13名)
2015年度	19名	(22名)

## 3. 分析結果

### 3.1 年度別回答データ (Pre/Post 別)

「(3) 実施タイミング」をパラメータと考え、Pre/Post 別分析を前提とする。各「(1) 年度」に属する全回答者の全「(2) 質問番号」に対する、5選択肢別の頻度値を集計した。表1,2と図1,2にその結果を示す。

表1 年度別回答データ(Pre)

	1	2	3	4	5	合計
2011	2	6	22	94	85	209
2012	2	11	26	102	106	247
2013	0	7	11	95	96	209
2014	0	11	19	77	121	228
2015	2	6	25	147	181	361
合計	6	41	103	515	589	1,254

表2 年度別回答データ(Post)

	1	2	3	4	5	合計
2011	1	3	11	89	105	209
2012	3	8	21	103	112	247
2013	0	0	3	98	108	209
2014	17	9	13	77	112	228
2015	0	3	12	116	230	361
合計	21	23	60	483	667	1,254

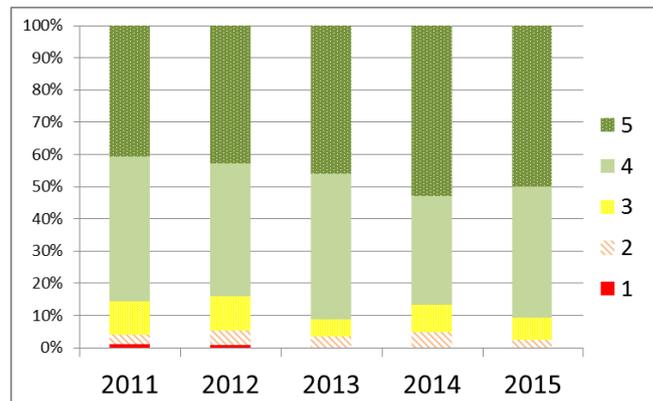


図1 年度別回答データ(Pre)

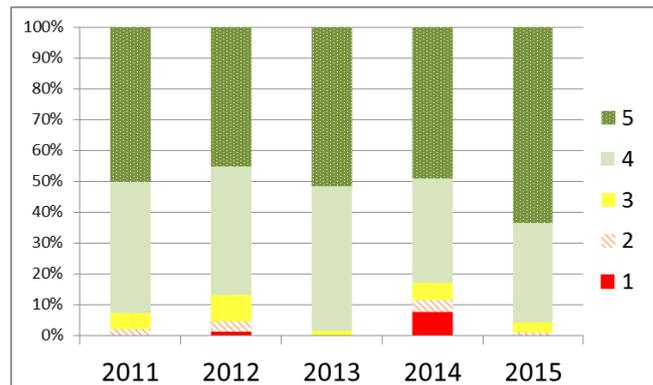


図2 年度別回答データ(Post)

表1および表2に関してピアソンの独立性検定を行ったところ、p値はそれぞれ0.0646と0.0000であり、以下の結果となった。

- (1) Post (表2) では、明らかに回答データと年度とは独立していない（年度ごとに回答データの分布に差がある）
- (2) Pre (表1) では、  
有意水準10%で考えれば独立していない（年度ごとに回答データの分布に差がある）  
有意水準5%で考えれば独立している（年度ごとに回答データの分布に差がない）

### 3.2 特異的な年度の抽出

#### 3.2.1 分析方法の提案

前項で“独立ではない”，すなわち年度間で回答データ分布に差があると判明した場合，さらに分析を進めるためには特異的な分布を示す年度を同定できることが望ましい。そこで今回，この“特異的”な年度を同定する方法として以下に述べる4方式を試行した。

- (1)  $\chi^2$  計算の個別（期待度－観測度数）<sup>2</sup>データ判定  
ピアソンの独立性検定では，クロス集計表にまとめられた個別頻度データから以下の計算式で $\chi^2$ 値を算

出する。(1)

$$\chi^2 = \sum \sum \frac{(\text{観測度数} - \text{期待度数})^2}{\text{期待度数}}$$

ここで、クロス集計表の各要素に対応した右辺の個別データで年度ごとの特徴を見極めることとする。個別データの判定閾値として、以下の2つを検討する。

当該データの自由度(今回は16)における設定有意水準での“ $\chi^2$ 判定値”をt, 個別データ個数をNとしたとき、

- ① 閾値 A = t
- ② 閾値 B = t/N

閾値 A は、個別要素1つで設定有意水準における閾値を上回る(すなわち、帰無仮説を却下する)場合、閾値 B は全個別要素がこれを上回ると設定有意水準において帰無仮説が却下される場合を意味する。

(2) 「単年度 vs 他年度合計」での独立性検定

例えば、2014年度のPreについて分析する場合、表1の2014年度とその他年度合計で2行×5列のクロス集計表に編集し、ピアソンの独立性検定を施す。

(3) 「当該年度以外」での独立性検定

例えば、2014年度のPreについて分析する場合、表1の2014年度以外で4行×5列のクロス集計表に編集し、ピアソンの独立性検定を施す。

(4) 単年度同士での独立性検定

全ての単年度同士の組み合わせ(2行×5列のマトリックス)それぞれに対してピアソンの独立性検定を施す。結果は5行×5列の対称マトリックス(ただし、対角成分は無効)で表現できる。

3.2.2 分析結果1

方式(1)による $\chi^2$ 計算の個別(期待度数-観測度数)<sup>2</sup>データを表3と表4に示す、なお、閾値B以上のデータセルを赤字赤色塗りつぶしで示す。また、閾値A以上のデータセルを橙色塗りつぶしで示す。

表3 個別(期待度数-観測度数)<sup>2</sup>データ(Pre)

	1	2	3	4	5	合計
2011	1.000	0.102	1.361	0.777	1.766	5.005
2012	0.566	1.059	1.608	0.003	0.865	4.101
2013	1.000	0.004	2.215	0.979	0.048	4.246
2014	1.091	1.686	0.004	2.956	1.807	7.543
2015	0.043	2.853	0.730	0.011	0.772	4.408
合計	3.700	5.704	5.918	4.726	5.257	25.305

表4 個別(期待度数-観測度数)<sup>2</sup>データ(Post)

	1	2	3	4	5	合計
2011	1.786	0.181	0.100	0.898	0.342	3.306
2012	0.312	2.657	7.134	0.650	2.858	13.612
2013	3.500	3.833	4.900	3.804	0.090	16.128
2014	45.509	5.551	0.401	1.333	0.709	53.502
2015	6.045	1.980	1.610	3.820	7.514	20.969
合計	57.152	14.204	14.144	10.504	11.514	107.518

表4から、2015年度Postでは5つの評価値全てにおいて全般的な頻度分布傾向からの偏差が大きいため“特異的”であると考えられる。また、2014年度Postの「回答肢1」が個別データとしては特異値となっていることにも注目すべきである。

3.2.3 分析結果2

方式(2)および方式(3)の分析結果である各独立性検定におけるp値を表5と表6に示す。

表5 年度別Preデータの独立性検定p値

	方式(2)	方式(3)	(参考)総合
2011	0.0664	0.1987	0.0646
2012	0.0442	0.2765	
2013	0.0739	0.2777	
2014	0.1844	0.0558	
2015	0.0912	0.1854	

表6 年度別Postデータの独立性検定p値

	方式(2)	方式(3)	(参考)総合
2011	0.0000	0.41039	0.0000
2012	0.0000	0.00198	
2013	0.0000	0.00067	
2014	0.0000	0.00000	
2015	0.0000	0.00001	

2011年度Postは、5年間総合でも方式(2)でも“独立ではない”と判定されているにもかかわらず、同データを除外した方式(3)では“独立である”と判定された。このことから、2011年度Postは他年度Postとは頻度分布傾向が異なり“特異的”であると考えられる。

3.2.4 分析結果3

方式(4)による単年度同士の独立性検定結果を表7と表8に示す。なお、p値が0.01を下回るデータを赤字赤色塗りつぶしで示す。

表7 単年度同士の独立性検定(Pre)

	2011	2012	2013	2014	2015
2011		0.864	0.170	0.028	0.170
2012	0.864		0.159	0.146	0.087
2013	0.170	0.159		0.073	0.334
2014	0.028	0.146	0.073		0.075
2015	0.170	0.087	0.334	0.075	

表8 単年度同士の独立性検定(Post)

	2011	2012	2013	2014	2015
2011		0.341	0.060	0.001	0.021
2012	0.341		0.000	0.005	0.000
2013	0.060	0.000		0.000	0.002
2014	0.001	0.005	0.000		0.000
2015	0.021	0.000	0.002	0.000	

表8からは、Postでは2011年度以外の各年度が互いに“独立ではない”ことを示唆している。

#### 4. 考察

今回は、2章で示したデータに対して複数の分析方式を適用した。これらの結果を総合的に俯瞰して得られる推定を以下に示す。

表2に示すPostデータ全体(5行×5列)に関する独立性検定ではp値が非常に小さく“独立ではない”と推定された。すなわち、各年度の分布は何かしら異なっていることを示唆した。そこで分析粒度を「(1)年度」へと小さくしたところ、2011年度に関しては方式(1),(3),(4)の結果で他の4年度との違いが存在することが認められた。方式(1)の結果(表4)からは、2011年度が5年度全体の分布傾向に基づく“期待値”に近いことを示している。また、単年度比較である方式(4)の結果(表8)では2012,2013,2015年度とは“独立である”となった。これらから、2011年度は各年度の合計から導出される“平均的な分布”となっている可能性が高いと推定できる。

方式(3)でも2011年度は他のpost年度とは異なった結果を示している。方式(3)の結果では、2011年度データを除外した4年度分のデータ(4行×5列)では“独立である”となった。これとPostデータ全体の独立性検定が“独立ではない”という結果を合わせて考えると、2011年度のデータが加わることによって独

立性が“なし”から“有り”に変化したことを意味すると考えられる。つまり、2012年度から2015年度がそれぞれ“ばらばら”な分布傾向を示しているのに対し、全体の“平均的な分布”である2011年度データが加わることによって分布のばらつきが縮小されたと考える。

Pre全体データの独立性検定結果に関しては熟慮する必要がある。当初の想定では、Pre全体では“独立である”ことを期待していた。すなわち、各年度で受講者は異なるため、年度間でのサンプル集団の均一性は担保されていない。そこで、この全体データの独立性検定によって“独立である”ことをもって、年度間でのサンプル集団の均一性を検証しようと考えた。今回の結果でp値は0.06となった。これは、有意水準を5%(0.05)に設定すれば帰無仮説“年度と評価値は独立である”が却下されず、当初想定どおり“独立である”ことを意味する。しかし、“有意水準5%”については米国統計学学会でもその“絶対視”に対して懸念が表明されている<sup>(2)</sup>。有意水準10%で判定すれば前述の帰無仮説が却下され、“独立ではない”となる。有意水準の設定値に関しては今後の検討課題とする。

#### 5. おわりに

今後、本稿で提案したピアソンの独立性検定に基づく分析方式を他のサンプルデータにも適用し、方式としての堅牢性を高めるとともに、様々な判定閾値や考慮すべき条件(例えば、クロス集計表のマトリックスサイズ、個別データの許容最小値など)の明確化を目指す。

#### 参考文献

- (1) 栗原 伸一: “入門統計学”, オーム社 (2011)
- (2) The ASA's statement on p-values: context, process, and purpose, <http://amstat.tandfonline.com/doi/abs/10.1080/00031305.2016.1154108> (2016年4月11日確認)