

5段階評価データの入力変換

筑波技術短期大学電子情報学科情報工学専攻
小池将貴

要旨：聴覚部の授業評価ワーキンググループが毎年実施する学生への授業評価アンケートでは、学生は受講した授業に対してあらかじめ用意された複数の質問の5段階評価に応えることにより授業を評価する。このような5段階評価データを統計手法へ入力するときは、通常は平均値に変換して用いる。この研究では、5段階評価データの入力変換において単純に平均値を用いたのでは問題があることを指摘し、より良い方法を提案する。

キーワード：授業評価 5段階評価データ 入力変換 主成分分析

1. 評価データ

平成11年度は76件、平成12年度は106件の授業について評価データが得られた。ひとつの授業を受ける学生の人数は、最大でも一般教育科目で約50名、最小が専門教育の選択科目で4～5名であり、通常は専門教育の必修科目で約10名という場合が多い。

各授業は、表1の11項（目的・理解・興味・疎通・機器・準備・熱意・態度・受話・通話・総合）の評価項目毎に受講学生の5段階評価を受けることによって評価された。つまり、評価項目の質問は5段階評価の方式を採り、（良い、やや良い、普通、やや悪い、悪い）という5水準のいずれかに応えさせるものである。

表1： 授業評価のための評価項目一覧

目的	授業内容が目的と整合していたか。
理解	理解できたか。
興味	興味を覚えたか。
疎通	あなたと教官との双方向コミュニケーションが成立していたか。
機器	黒板・ビデオ・OHPなどの情報補償機器の使用が適切か。
準備	テキスト・教材・実験機器などが準備されていたか。
熱意	教官の熱意を感じたか。
態度	あなたの受講態度は積極的だったか。
受話	あなたは教官の話を聞き取れたか(話を受けとめることができたか)。
通話	教官はあなたの話を読み取ったか(話は通じたか)。
総合	総合的に見て、良い授業といえるか。

2. 主成分分析への入力データ

2.1 主成分分析の適用

授業評価の本質は授業同士の差別化である。そこで、多変量（11項の評価項目）の線形結合を主成分（授業の代表評価指標）とし、その分散を最大化させる主成分分析法を適用することにした。

2.2 恣意的な平均値入力

そのためには入力データとして、平成11年度の主成分分析については76件の各授業に対して（平成12年度の主成分分析については106件の授業に）、それぞれ11項の評価項目に関するデータが必要になる。つまり、11項の各質問毎に、5段階評価データから主成分分析用の入力データを作成せねばならない。

まず思いつくのは、5段階評価データから平均値

を計算することである。例えば、或る授業を40人の学生が受講し、その授業の或る評価項目について、（良い、やや良い、普通、やや悪い、悪い）の各水準に対する反応の内訳が（7人、10人、6人、9人、8人）という5段階評価データが得られたとしよう。その平均値は、 $(5 \times 7 + 4 \times 10 + 3 \times 6 + 2 \times 9 + 1 \times 8) \div (7 + 10 + 6 + 9 + 8) = 2.98$ というようにして計算される。しかし、単なる思いつきで平均値を用いてよいのだろうか。

3. 学生の評価スタンス

3.1 5段階評価の入力実態

ここで、5つの水準への受講学生の反応実態を調べてみよう。つまり、1つの授業には11項の評価項

目があり、その授業を受講した複数の学生の各々は、水準5（良い）、水準4（やや良い）、水準3（普通）、水準2（やや悪い）、水準1（悪い）という5つの水準のいずれかに応える行為を11回繰り返す。全受講学生の反応行為の全授業延べ回数を水準別に集計し、全水準に対する各水準の反応比率（%）を計算したものを図1（平成11年度）と図2（平成12年度）に示す。

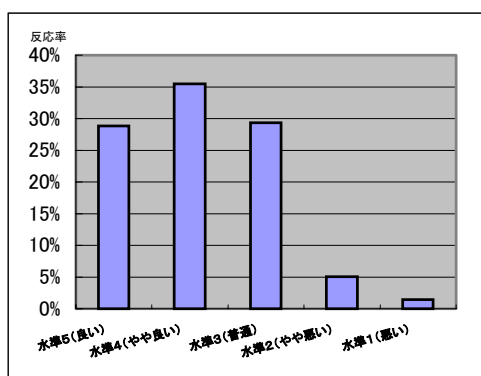


図1：水準別の反応率（平成11年度）

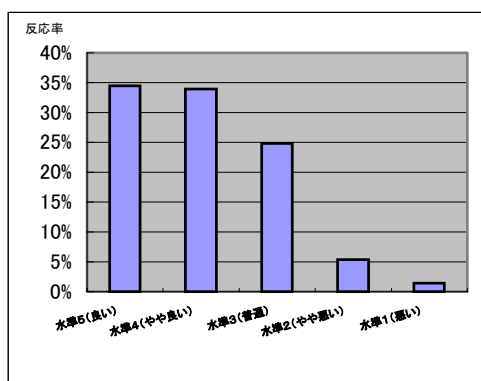


図2：水準別の反応率（平成12年度）

この図1と図2を見ると、水準2（やや悪い）と水準1（悪い）には極めてわずかしか反応していない。全体的には学生は、水準5（良い）、水準4（やや良い）、水準3（普通）という3段階で評価しているのである。

このように5つの水準への学生の反応が2極化しているのに、各水準を等しくウェイト付けする平均値を主成分分析への入力に採用するのは疑問である。せっかく5段階に分離されているデータを思いつきで平均値という1つのデータに集約する前に、5段階に分かれたままの状態が何を語りかけているかを別の方向からも調べてみよう。

3.2 5段階評定データの5段階そのままの解析

ひとつの授業に対して11項の評価項目の質問があり、各質問の回答として受講学生による5段階評

定データが対応している。つまり、第2章2節で例示したデータをここでも引用して例解すると、或る授業を40人の学生が受講し、その授業の或る評価項目について、（良い、やや良い、普通、やや悪い、悪い）の各水準に対する反応の内訳が（7人、10人、6人、9人、8人）であったとする。この5つの人数の組がその授業のその評価項目に対する5段階評定データである。

仮に授業の評定反応の内訳を同じ比率にして受講生を増やすと、内訳の人数も増えてしまう。授業同士を同一の基準で比較するためには、その授業の受講人数で基準化する必要がある。つまり、各水準の人数を受講人数40人で割り算して、（18%、25%、15%、23%、20%）という学生反応率データにして議論しなければならない。

平成11年度については、76件の授業がある。各授業には、11項の評価項目毎に学生反応率データ（評価項目毎に5水準）を対応させた。したがって1件の授業は合計55個（11項×5水準）の数値データから成る。このデータ行列（76行55列）に主成分分析を施した。

得られた主成分（寄与率が最大25%を持つ第1主成分）の列スコアを表2に示す。これは本来55個の要素からなるベクトルであるが、わかりやすいようにベクトルを5個ずつの要素ごとに区切って示した（表2）。

主成分分析ではどのようにしてこの列スコアを求めたかという、構成する55個の要素は最初は未知と考えて、それらと任意の授業の学生反応率データ（11項×5水準=55個の既知データ）との線形結合を作る。これが、その授業の主成分である。すると76件の主成分が得られるので、その分散を最大化させるように列スコアの55個の未知係数を定めたのである。つまり76件の授業の差別化を最大にするものとして列スコアが得られるのである。こうして得られた列スコアを眺めることによって、分析対象に関するヒントを思いつけばよいというのが、主成分分析の趣旨である。

3.3 学生の評価スタンスの顕在化

表2では、同一水準を縦に眺めるとほぼ同じ値が並ぶ。水準別に横に眺めると一定の傾向が見える。

すなわち、学生は5段階評定において『第5水準はプラス評点、第4水準はほぼ零点、第3水準・第2水準・第1水準はすべてマイナス評点という実質的には3段階で評定していた』ことがわかる。この評価スタンスは、評価項目11項毎に一貫して現れて

いる(表2)。

それに続けて平成12年度についても同様の計算をした(表3)。表3(平成12年度)においても、『学生の評価スタンスは表2(平成11年度)と同じ』である。なお、平成12年度データの第1主成分の寄与率は27%であった。

表2：学生の評価スタンス(平成11年度；授業76件)

評価項目	第5水準	第4水準	第3水準	第2水準	第1水準
目的	0.21	0.01	-0.17	-0.18	-0.11
理解	0.19	0.11	-0.15	-0.17	-0.16
興味	0.20	0.08	-0.14	-0.16	-0.18
疎通	0.19	0.08	-0.15	-0.19	-0.16
機器	0.14	0.07	-0.11	-0.15	-0.11
準備	0.15	0.04	-0.11	-0.13	-0.07
熱意	0.16	-0.02	-0.13	-0.09	-0.18
態度	0.14	0.12	-0.12	-0.12	-0.04
受話	0.14	0.00	-0.14	-0.08	-0.09
通話	0.14	0.06	-0.15	-0.12	-0.02
総合	0.19	0.10	-0.16	-0.16	-0.15
平均	0.17	0.06	-0.14	-0.14	-0.12

表3：学生の評価スタンス(平成12年度；授業106件)

評価項目	第5水準	第4水準	第3水準	第2水準	第1水準
目的	0.19	-0.06	-0.17	-0.10	-0.05
理解	0.19	0.08	-0.16	-0.16	-0.10
興味	0.21	0.03	-0.17	-0.15	-0.11
疎通	0.21	0.01	-0.19	-0.14	-0.09
機器	0.18	0.02	-0.16	-0.13	-0.08
準備	0.18	0.03	-0.17	-0.13	-0.07
熱意	0.21	-0.05	-0.19	-0.13	-0.08
態度	0.12	0.11	-0.11	-0.14	-0.10
受話	0.17	0.02	-0.19	-0.11	-0.08
通話	0.17	0.01	-0.19	-0.09	-0.05
総合	0.21	0.01	-0.20	-0.12	-0.10
平均	0.18	0.02	-0.17	-0.13	-0.08

4. 主成分分析への入力変換の新しい方法

まず、第3章1節における授業評価データの5水準別分類によりわかったことは、5段階評定では悪いほうの2つの水準は無視して差支えなく、水準5(良い)、水準4(やや良い)、水準3(普通)に対する反応が主体を成して、実質的には3段階評定になっているということである。

さらに、第3章3節の主成分分析によりわかったことは、5段階のうちの第3水準・第2水準・第1水準はひとまとめにすることができて、ここでも実質

的には3段階評定になっているということである。

ただし、3つの評価段階の較差の均等性の判定が難しいので、『最高水準の第5水準に反応するときは、学生は本音で良いと考えている』という観点から入力データを作ることにした。つまり、授業の評価項目の5段階評定(良い、やや良い、普通、やや悪い、悪い)のうちの良いという水準に反応した学生数をその授業の全受講学生数で除算した値(良いという水準に対する学生反応率)を以って平均値に代わる主成分分析用の入力データとするのである。例えば、第2章2節での計算に例示したデータでは、 $7 \div 40$ (17.5%)となる。

5. 新入力データによる主成分分析

5.1 新入力データによる計算結果

平成11年度の76件の授業について、その11項の評価項目のデータをすべて平均値の代わりに『良いという最高水準への学生の反応率』という新提案に基づくデータにして主成分分析を適用した(表4)。同様に平成12年度についても計算した(表5)。

表4と表5とに示されている数値は、評価項目と主成分との相関係数を表している。その値がマイナスの評価項目の名称は、その後 * 印を付けておいた。

5.2 表4と表5とに見られる多数の一致点

ここで表4(平成11年度)と表5(平成12年度)とでは一致点が多いことに注目されたい。

まず、第1主成分は、『すべての数値が正で1.0に近い。さらにその大きい順に並べた評価項目の表4の先頭の4つの序列(疎通・総合・目的・理解)が表5と興味を除いて一致している。』

第2主成分は、『プラスの評価項目同士が表4と表5とで一致している。』当然、『マイナスの評価項目同士も一致している』ことになる。さらによく見ると、『大きい順に並べた評価項目のプラス側の通話・受話とマイナス側の機器・準備が表4と表5とでその序列まで一致している。』これは主成分の意味を解釈するときに役立つ。

第3主成分も『プラスの評価項目同士が表4と表5とでほとんど一致している。例外は疎通のみであるが、その値はゼロに近くて無視してもよい違いである。』といえる。当然、『マイナスの評価項目同士もほとんど一致している』ことになる。さらに『大きい順に並べた評価項目の序列において、プラスで大きい値の4つの群が表4と表5とで一致している。マイナスで絶対値の大きい3つの群も表4と表

5とで一致している。』これは主成分の意味を解釈するとき役立つ。

平成11年度と平成12年度の5段階評定データは、採択された授業や受講学生の構成が当然異なっている。それにもかかわらず、得られた表4と表5とがきわめてよく似ているのは、それぞれの表の基になる平成11年度（76件の授業）と平成12年度（106件の授業）の授業評価データが共通のデータ構造を有しているためと考えられる。

表4：主成分分析の列スコア（平成11年度；授業76件）

評価項目	第1主成分	評価項目	第2主成分	評価項目	第3主成分
疎通	0.85	通話	0.59	機器	0.46
総合	0.83	受話	0.45	受話	0.38
目的	0.83	疎通	0.25	準備	0.34
理解	0.8	態度	0.23	通話	0.26
興味	0.76	総合*	-0.03	疎通	0.09
受話	0.74	熱意*	-0.05	熱意	0.09
熱意	0.7	目的*	-0.09	総合*	-0.14
通話	0.7	理解*	-0.17	目的*	-0.2
態度	0.64	興味*	-0.23	理解*	-0.3
準備	0.62	機器*	-0.51	態度*	-0.35
機器	0.6	準備*	-0.57	興味*	-0.45
寄与率	55%	寄与率	12%	寄与率	9%

表5：主成分分析の列スコア（平成12年度；授業106件）

評価項目	第1主成分	評価項目	第2主成分	評価項目	第3主成分
疎通	0.88	通話	0.56	機器	0.4
興味	0.87	受話	0.53	準備	0.36
総合	0.87	態度	0.28	受話	0.29
目的	0.85	疎通	0.01	通話	0.23
理解	0.82	理解*	-0.02	熱意	0.1
熱意	0.82	総合*	-0.04	疎通*	-0.03
機器	0.78	興味*	-0.15	目的*	-0.05
準備	0.76	目的*	-0.21	総合*	-0.2
通話	0.73	熱意*	-0.21	興味*	-0.29
受話	0.72	機器*	-0.26	理解*	-0.36
態度	0.62	準備*	-0.32	態度*	-0.41
寄与率	64%	寄与率	9%	寄与率	8%

ただし、第4章で新しく考案した入力データ変換のおかげで、表4（平成11年度）と表5（平成12年度）のような酷似している表がもたらされたのである。

学生の評価スタンスを無視し思いつきの平均値を入力データとした主成分分析を別途に計算して

みたが、表4と表5のような一致点は見つけられなかった（表6と表7）。

表6：主成分分析の列スコア（平均値入力；平成11年度）

評価項目	第1主成分	評価項目	第2主成分	評価項目	第3主成分
総合	0.91	通話	0.58	機器	0.38
目的	0.89	受話	0.51	準備	0.34
疎通	0.88	疎通	0.27	受話	0.34
興味	0.87	態度	0.04	熱意	0.27
理解	0.83	総合	0.02	通話	0.21
態度	0.74	目的*	-0.01	疎通	0
熱意	0.73	理解*	-0.11	総合*	-0.1
機器	0.72	熱意*	-0.14	目的*	-0.18
受話	0.72	興味*	-0.2	態度*	-0.25
通話	0.69	機器*	-0.42	興味*	-0.35
準備	0.63	準備*	-0.58	理解*	-0.39
寄与率	62%	寄与率	11%	寄与率	8%

表7：主成分分析の列スコア（平均値入力；平成12年度）

評価項目	第1主成分	評価項目	第2主成分	評価項目	第3主成分
総合	0.93	受話	0.51	態度	0.35
理解	0.88	通話	0.47	理解	0.23
興味	0.88	疎通	0.22	興味	0.21
熱意	0.88	機器	0.07	通話	0.18
疎通	0.87	熱意	0.02	総合	0.11
目的	0.84	総合*	-0.08	受話	0.1
機器	0.82	理解*	-0.12	疎通*	-0.01
準備	0.81	準備*	-0.2	目的*	-0.09
受話	0.79	興味*	-0.26	熱意*	-0.2
通話	0.78	目的*	-0.27	準備*	-0.43
態度	0.73	態度*	-0.33	機器*	-0.45
寄与率	70%	寄与率	8%	寄与率	6%

6. 結論

(1) 第3章で明らかにされたように、5段階評定データに対して学生が第5水準（良い）はプラス、第4水準（やや良い）はゼロ、残りの水準はすべてはマイナスという重み付けをしているのに、5水準間の較差を均等とみなす平均値をあえて採用するのはよくない。その代わりに第5水準に対する学生反応率を薦めたい。

(2) それにより、授業科目の違い、受講学生の違い、年度の違いにかかわらず、一貫した分析が可能になる。

Transforming Method of Five-Point Ratings Data for Statistical Method

KOIKE Masayoshi

Department of Information Science, Tsukuba College of Technology

Abstract : The aim of this study is to suggest a transforming method of five-point ratings data into the input of a statistical method. The five-point ratings data of the lectures given in Tsukuba College of Technology for two years were collected. Application of the principal component analysis to these data revealed how the students filled out the questionnaire, i.e. they thought their fifth stage response as a plus evaluation, their fourth stage response as a zero evaluation and their third/second/first stage response as a minus evaluation. So, when five-point ratings data is transformed into the input of a statistical method, you may calculate the rate of the students responding to the fifth stage instead of calculating mean values.

Key Words : instructional evaluation; five-point ratings; transformation of questionnaire data; principal component analysis