

授業評価データに潜在する意外な知見の S-PLUS による掘り起こし

筑波技術短期大学 小池将貴

1. はじめに

筑波技術短期大学の聴覚部では、学生による授業評価を毎年実施している。ここは聴覚障害者を対象とする三年制の国立短期大学であり、平成 2 年 4 月に第 1 期生を受け入れて発足し、本年度は第 12 期生が勉学中である。

学生は受けた授業に対する評価を、複数の質問に対する 5 段階評定を通じて行う。このような 5 段階評定データは、平均値に変換してから統計解析にもちこむのが通常である。この研究では、(1) 5 段階評定データを単純に平均値に変換するのでは問題があることを指摘し、(2) より良い方法で変換することによって安定した統計解析結果が得られることを示す。

2. 5 段階評定による授業評価データ

ここで解析に用いるデータは、平成 11 年度と平成 12 年度の授業評価結果である。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。

1 つの授業は、担当教官名と担当教科名とにより識別される。従って、1 人の教官は担当する複数の教科毎に別々に評価を受ける。また、1 つの授業を受ける学生 (1 つの授業の評価者) の人数は、最大でも全学科共通教科における約 50 名、最小が学科別選択教科の 4~5 名であり、通常は学科別必修教科の約 10 名という場合が多い。

授業を受けた学生は、次に示す表 1 の質問の各々に答えることによってその授業を評価する。質問は 5 段階評定の方式を採り、(良い、やや良い、普通、やや悪い、悪い) という 5 つの水準のいずれかに応えさせるものである。

表 1: 授業評価のための質問一覧

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・教材などがあらかじめ準備されていたか。
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 <u>受けとめる</u> ことができたか) 。
通話	教官はあなたの話を <u>読み取った</u> か (話は <u>通じた</u> か) 。
総合	総合的に見て、良い授業といえるか。

3. 主成分分析への入力データ

3.1 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 (表 1 の 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.

3.2 恣意的な平均値入力

主成分分析法を適用するには,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年度については106件の授業に),それぞれ11項の評価項目(目的・理解・興味・疎通・機器・準備・熱意・態度・受話・通話・総合)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段階評定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段階評定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段階評定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,5段階評定データの代表値として単なる思いつきで平均値を用いてよいのだろうか.

4. 学生の評価スタンス

4.1 5段階評定データの5段階そのままの解析

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調べてみることにした.

4.2 学生反応率データ

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段階評定データがある.つまり,第3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が同じ比率でも受講生が増えると,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すべきである.なお,良いを第5水準,やや良いを第4水準,・・・,悪いを第1水準としている.

4.3 主成分分析

平成11年度については,76件の授業が有る.各授業には,11項の評価項目毎に5水準から成る学生反応率データを対応させた.したがって1件の授業は合計55個(11項 \times 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てみた.

得られた主成分(寄与率が最大25%を持つ第1主成分)の列スコアを図1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要素を11項の評価項目別に5水準に分けて示した.また,水準毎の平均値をAverageとして追加表示した(図1).

主成分分析ではどのようにしてこの列スコアを求めるかということ,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 \times 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個の未知数を定めるのである.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.こうして得られた列スコアを眺めることによって,データの背後に潜在する何らかの知見を見出せばよいということで,主成分分析を利用した.

4.4 学生の評価スタンスの顕在化

図1は,平成11年度の5段階評定データに主成分分析を施して得られた第1主成分の値を,評価項目のコマに分けて示したものである.各コマの横軸は左から右に第1水準(悪い)から第5水準(良い)を表し,縦軸は第1主成分で得られた値である.図1によると,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』と解釈することができる.この評価スタンスは,評価項目11コマ毎に一貫して現れている(図1).全平均(Average)では,それが集約されて示されている(図1).

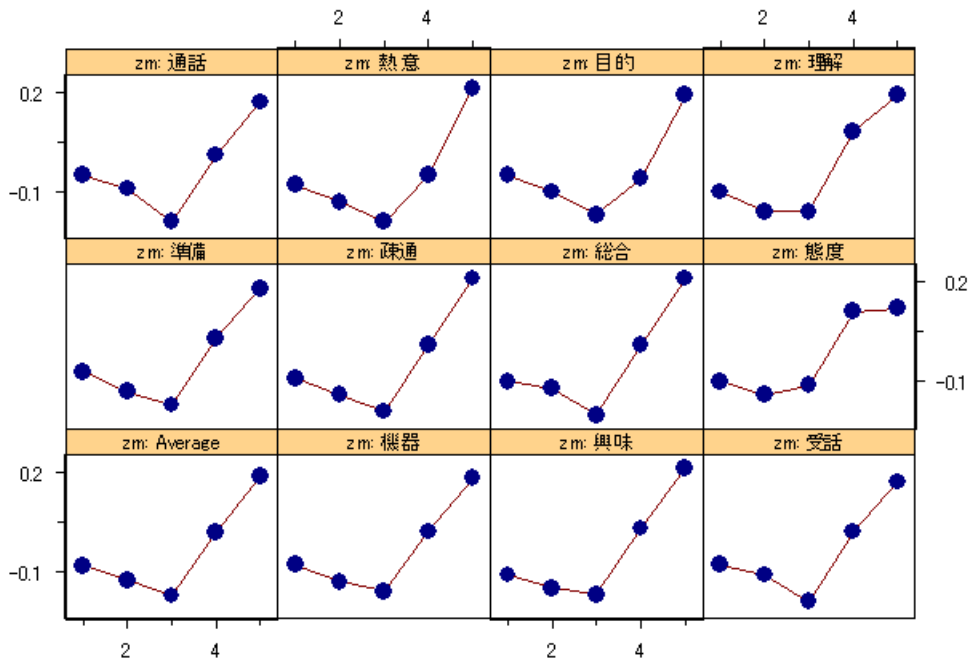


図 1: 学生の評価スタンス(平成 11 年度,授業 76 件)

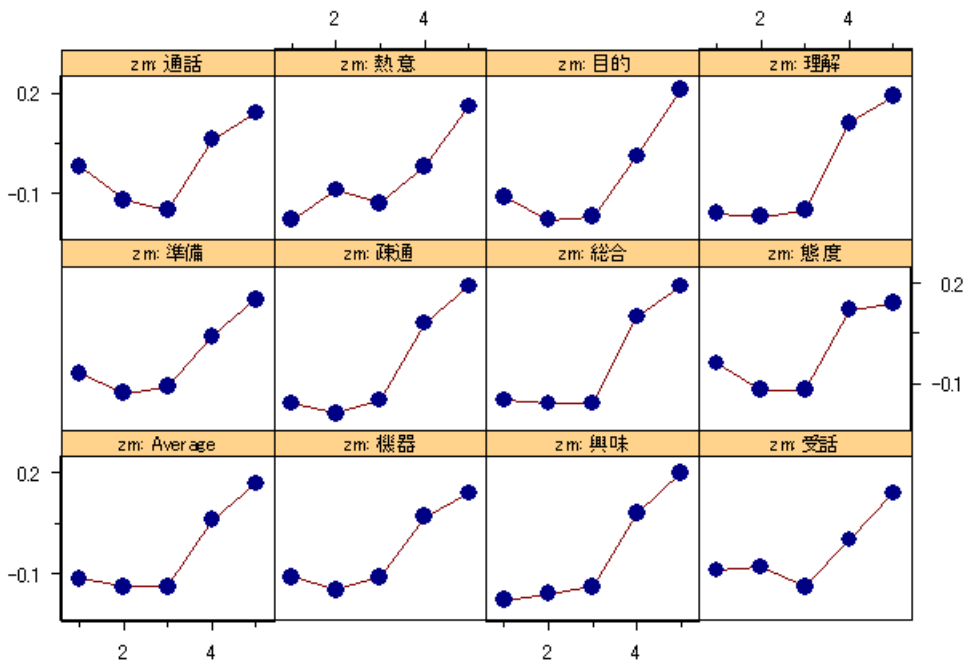


図 2: 学生の評価スタンス(平成 12 年度,授業 106 件)

平成 12 年度についても同様の計算をした(図 2)。図 2(平成 12 年度)においても、『学生の評価スタンスは図 1(平成 11 年度)と同じく、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』のである。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。

5. 主成分分析への入力変換の新しい方法

第 4 章 4 節の主成分分析の結果からわかったことは、5 段階のうちの第 3 水準・第 2 水準・第 1 水準はまとめてマイナス評価、第 4 水準がゼロ評価、第 5 水準がプラス評価ということである。そこで、『最高水準の第 5 水準に反応するときのみ、学生は本音で良いと考えている』と解釈し、入力データ

をすることにした。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)の良いという第5水準に対する学生の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3章2節での計算に例示したデータでは、新しい入力データ(第5水準学生反応率)は、 $7 \div 40$ (17.5%)となる。

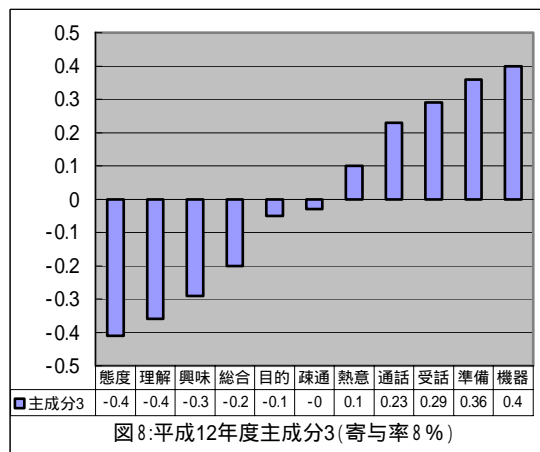
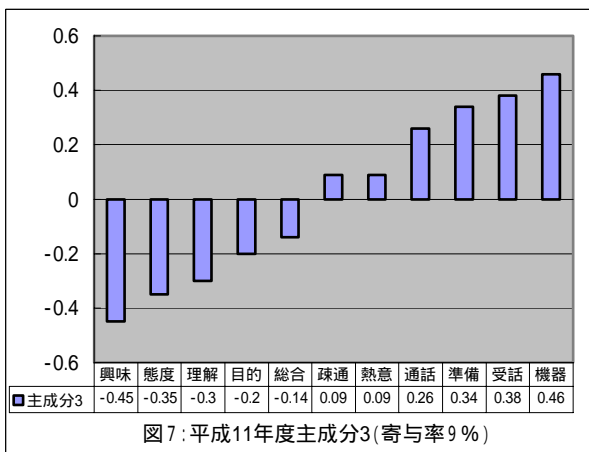
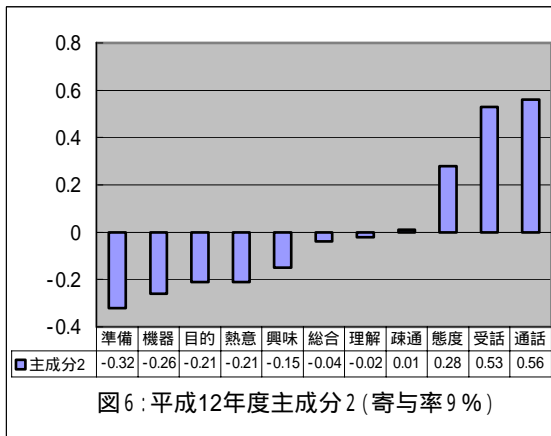
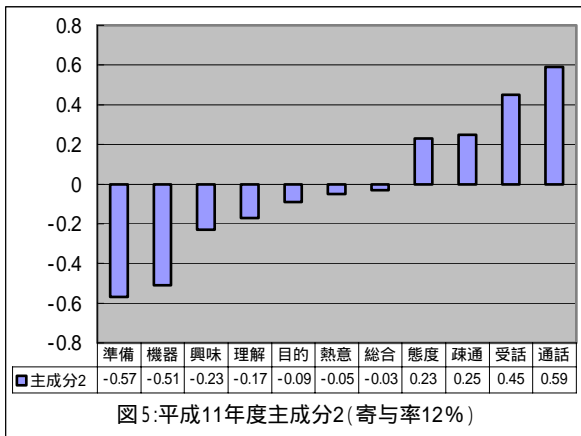
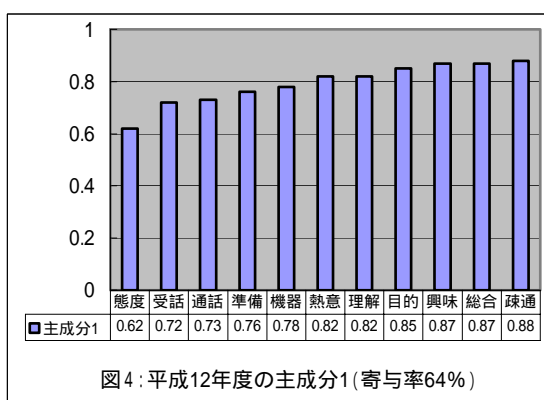
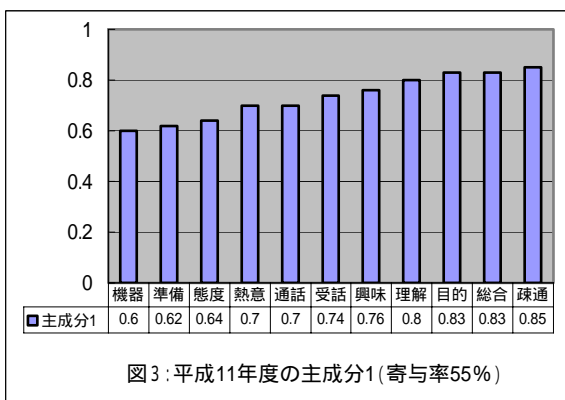
6. 新入力データによる主成分分析

6.1 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。同様に平成12年度についても計算した。両年度の第1主成分～第3主成分を図3～図8に示す。その縦軸の値は、評価項目と主成分との相関係数を意味している。

6.2 平成11年度と平成12年度とに見られる多数の一致点

ここで平成11年度と平成12年度との計算結果では一致点が多いことに注目されたい。



<<第1主成分;図3・図4参照>>

- (1) 図3(平成11年度)と図4(平成12年度)共に『すべての数値が正で1.0に近い。』
- (2) さらに『図3(平成11年度)の値が大きい順に右から並ぶ評価項目の先頭の5つ(疎通・総合・目的・理解・興味)が図4(平成12年度)と序列までがほぼ一致している。』

<<第2主成分;図5・図6参照>>

- (1) 『値がプラスの評価項目同士が図5(平成11年度)と図6(平成12年度)とで一致している。』
- (2) 当然, 『マイナスの評価項目同士も図5(平成11年度)と図6(平成12年度)とで一致している』ことになる。
- (3) さらに両極端で見ると, 『大きい順に右から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・準備が図5(平成11年度)と図6(平成12年度)とでその序列までが一致している。』

<<第3主成分;図7・図8参照>>

- (1) 『プラスの評価項目同士が図7(平成11年度)と図8(平成12年度)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。
- (2) 当然, 『マイナスの評価項目同士も図7(平成11年度)と図8(平成12年度)とでほとんど一致している』ことになる。
- (3) さらに『大きい順に右から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(機器・準備・受話・通話)が図7(平成11年度)と図8(平成12年度)とで一致している。
- (4) そして『マイナスで絶対値の大きい3つの群(興味・態度・理解)も図7(平成11年度)と図8(平成12年度)とで一致している。』

平成11年度と平成12年度の5段階評定データは, 教官の定年退職や新規採用等により授業の構成が異なり, 学生の卒業や入学等により受講学生の構成も異なる。それにもかかわらず, 得られた主成分(図3~図8)が双方の年度同士できわめてよく似ているのは, それぞれの主成分の基になる平成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。

ただし, 第5章で新しく考案した入力データ変換のおかげで, 平成11年度と平成12年度の主成分が酷似するという結果がもたらされたのである。なぜならば, 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, 図3~図8のような一致点は見つけられなかった。

7. 結 論

(1) 第4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つの水準間の較差を均等とみなす平均値をあえて統計解析に採用するのは妥当ではない。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) そうすることにより, 教科と学生の双方が入れ替わる年度を通して, 安定した分析結果が得られた。

8. S-P L U Sの活用

(1) エクセルで与えられた入力データをS-P L U Sでの主成分分析計算に結びつけるのに, S-P L U Sのインポート機能が役立った。

(2) やや規模の大きな主成分分析(55変数)はS-P L U Sで問題なく処理できたが, そこから図1・図2のアウトプットに至るまでのさまざまな変数変換は, S-P L U Sコマンドが大いに役立った。統計プログラムを単に集めたパッケージではこうはいかない。

(3) S-P L U Sでの主成分分析結果をエクセルで図3~図8のグラフにするのに, S-P L U Sのエクスポート機能が役立った。