

学期末授業評価の分析

山口県立大学情報化推進室

畔津 忠博、吉永 敦征、市村 孝雄

Analysis of End-of-Semester Class Evaluations by Students

Tadahiro AZETSU, Nobuyuki YOSHINAGA, Takao ICHIMURA

Office for Information and Technology of Yamaguchi Prefectural University

要 旨

本稿では、山口県立大学で行われた2008年度における前期と後期の学期末授業評価の分析結果について報告し、まず、質問項目ごとの評価値の統計的性質を平均情報量、尖度、歪度などにより明らかにし、次に授業の満足度に関連する質問項目とそれ以外の質問項目との関係を相関係数と相互情報量により調べる。さらに、授業の満足度を高めるために必要である幾つかの重要な質問項目を回帰分析、数量化I類、主成分分析を用いて示す。

キーワード：授業評価、授業満足度、回帰分析、数量化I類、主成分分析

Abstract

This paper describes the results of the analysis of class evaluations by students, which were conducted at the end of the first and second semester for the academic year 2008 in Yamaguchi Prefectural University. We first clarify the statistical characteristics of the evaluation values of each question item by information entropy, skewness and kurtosis, and then investigate the relations between the question item concerning the class satisfaction and the other question items by correlation coefficient and mutual information. Furthermore, some important question items necessary to increase the class satisfaction are indicated by using multiple regression analysis, quantification theory type I and principle component analysis.

Key words: Class evaluation, Class satisfaction, Multiple regression analysis, Quantification theory type I, Principle component analysis.

1 はじめに

山口県立大学では、2004年度よりWebを利用した授業評価が開始され、現在ほぼすべての授業に対して実施されている^[1]。入力にはパソコンや携帯電話が用いられ、簡単な操作で入力できるようになっている。

授業評価には毎回授業評価と学期末授業評価がある。毎回授業評価は、毎回授業終了後、当日ならびに翌日中に「講義のねらいが明確に示され、おおむねシラバスにそって授業がすすめられましたか？」や「教員に教える熱意が感じられましたか？」など6個の質問項目に対して匿名で3段階評価により行われる。また、評価入力後にコミュニケーションボードが利用できるようになっており、教員と学生が双

方向で意見交換できるようになっている。これらの結果は、次回の授業を改善するために利用される。

また、学期末授業評価は、最終試験終了後の定められた期間内に「教員の毎回の授業準備は十分でしたか？」や「学生の関心や理解度を考慮しながら授業を進めていましたか？」や「総合的に判断して、この授業に満足できましたか？」など、学期を通じた評価を問う質問項目に対して匿名で5段階評価により行われる。加えて、授業の良かった点と悪かった点についての感想も記述してもらい、これらの結果は、実施年度の授業の総括と次年度の授業改善のデータとして利用される。

本稿では、2008年度に行われた学期末授業評価前期2回について、特に講義科目の結果に絞って報

告する。本稿の構成は、以下の通りである。2章では、学期末授業評価における質問項目の内容やデータの入力数、評価値の平均値など基本的な事柄を示す。3章では、各質問項目の5段階評価値の統計的性質について、幾つかの統計量、情報量を示しながらその傾向を述べる。4章では、質問13（満足度）と他の質問項目との関連性を、回帰分析、数量化I類、主成分分析を用いて調べ、どの質問項目が満足度に強く影響を及ぼしているのかを示す。5章で結論を述べる。

2 学期末授業評価

2.1 質問項目の内容

学期末授業評価は、表1に示す5段階評価により匿名で行われる。

評価値は、本来、質的データであり順序尺度として考えるべきであるが、授業評価の分析で量的デー

表1. 5段階評価

5 = とても（強く）そう思う
4 = ややそう思う
3 = 何とも言えない（分からない）
2 = あまりそう思わない
1 = ほとんど（全く）そう思わない

表2. 学期末授業評価の質問項目（講義科目）

質問1	この授業に興味や期待をもって受講しましたか？
質問2	シラバスは、授業の全体像をつかむのに有効でしたか？
質問3	教員に教える熱意や学生への誠意が感じられましたか？
質問4	教員の、毎回の授業準備は十分でしたか？
質問5	授業と関連した本や資料が適宜紹介されましたか？
質問6	授業では、学生の意欲化や理解を深める上での工夫がみられましたか？
質問7	全体的に教員の話し方（声の大きさ、速さ等）は聞き取りやすかったですか？
質問8	学生の関心や理解度を考慮しながら授業を進めていましたか？
質問9	教室の雰囲気は学習を進める上で適切でしたか？
質問10	全体を通して、授業の内容が理解できましたか？
質問11	学生の毎回授業評価は、この授業改善に役立ったと思いますか？
質問12	試験の内容やレポートのテーマ、分量等は適切でしたか？
質問13	総合的に判断して、この授業に満足できましたか？
質問14	授業を終えた今、自分でさらに学習を深めたいと思いますか？

タとして取り扱っている例も多くあるので^{[2]-[7]}、本稿でも、量的データとして取り扱い、一部質的データとしても考える^[8]。

質問項目の内容と数は、講義、演習、実習、学外実習など授業形態によって多少異なる。表2に講義科目に関する14個の質問項目の内容について示す。

2.2 データ入力数

授業評価の入力データ数を、授業形態ごとに表3に示す。講義科目に関しては、多くのデータ数が得られているが、その他の形態では、あまり多くない。そのため、以降は、講義科目の結果に絞って報告する。

表3. 授業形態ごとのデータ入力数

授業形態	前期	後期
講義	6825	3437
演習	871	604
実習	285	140
学外実習	142	486
実技	424	93
合計	8547	4760

2.3 評価値の平均

図1は、講義科目における各質問項目の評価値の平均を示す。この図より、前期と後期の結果は、全体の傾向についてかなり類似していることが分かる。そこで、次章では、前期と後期で各質問項目の評価値の統計的性質を調べ、質問内容により固有の傾向があるかどうかを確認する。

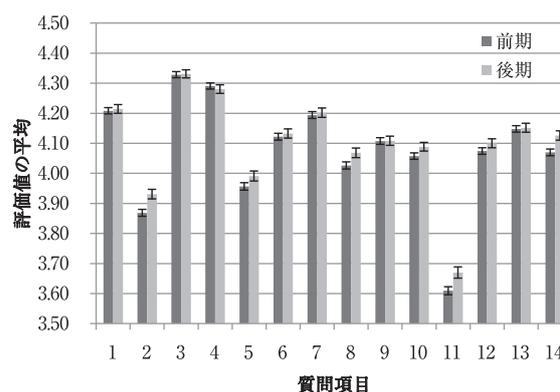


図1. 各質問項目の評価値の平均（講義科目）。誤差バーは標準誤差を示す。

3 各質問項目の評価値の統計的性質

3.1 平均情報量

データ数を N （前期6825、後期3437）とし、 i 番目（ $i = 1, \dots, 14$ ）の質問項目の評価値 k （ $k = 1, \dots, 5$ ）

のデータ数を $N_i(k)$ 、確率を $p_i(k)$ で表すと

$$p_i(k) = N_i(k) / N$$

である。ここで、 $N = \sum_{k=1}^5 N_i(k)$ である。 i 番目の質問項目において評価値をとり得る値とする確率変数を X_i で表すと、 X_i の平均情報量は

$$H(X_i) = -\sum_{k=1}^5 p_i(k) \log_2(p_i(k))$$

で定義される。平均情報量は、情報の無秩序さを示す尺度である。

表4に各質問項目の平均情報量を示す。評価値の数が5個なので最大 $\log_2 5 = 2.322$ (bit) の値を取り得る。

表4. 各質問項目の平均情報量 (講義科目)

	前期 (bit)	後期 (bit)
質問1	1.663	1.652
質問2	1.877	1.831
質問3	1.565	1.531
質問4	1.606	1.602
質問5	1.902	1.859
質問6	1.767	1.731
質問7	1.728	1.698
質問8	1.849	1.804
質問9	1.768	1.749
質問10	1.736	1.684
質問11	2.044	2.055
質問12	1.763	1.726
質問13	1.716	1.699
質問14	1.792	1.743

確率分布に偏りがあれば平均情報量は小さくなるので、最も値の小さい質問3は分布に偏りがあり、逆に最も値の大きい質問11は偏りが比較的小さいことが予想される。図2、図3に質問3と質問11のヒストグラムを示す。この図より、質問3は右側に偏りがあるが、質問11は質問3と比較すると偏りが少ないことが分かる。

3.2 歪度、尖度

図2、図3より、質問項目によって評価値の確率分布にかなり違いがあることが分かる。確率分布の形状を調べるための基本的な統計量に、歪度と尖度がある。ともに、標準的な確率分布である正規分布からのずれ具合を示す尺度である。歪度は正規分布のように左右対称であれば0になり、右に偏り左に裾が長い場合は負になり、逆の場合は正となる。また、尖度は確率分布が正規分布に比べ尖った形状になると正になり、平坦になる場合は負になる。

X_i の平均値を μ_i 、標準偏差を σ_i とする。このと

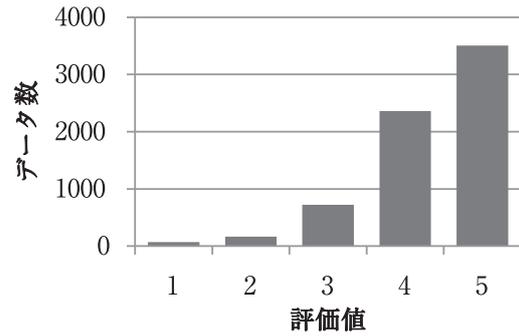


図2. 質問3の評価値のヒストグラム (前期 講義科目)

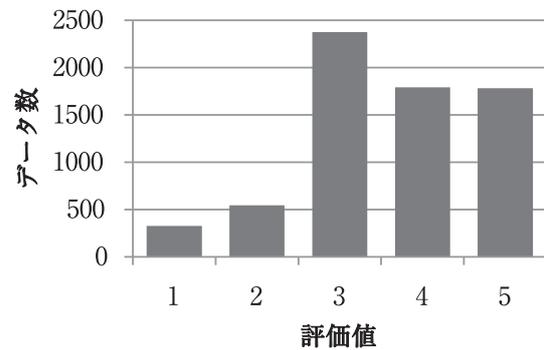


図3. 質問11の評価値のヒストグラム (前期 講義科目)

表5. 各質問項目の歪度と尖度 (講義科目)

	前期		後期	
	歪度	尖度	歪度	尖度
質問1	-1.132	1.280	-1.072	1.095
質問2	-0.592	0.034	-0.605	0.063
質問3	-1.348	1.921	-1.196	1.610
質問4	-1.225	1.438	-1.136	1.190
質問5	-0.785	0.108	-0.782	0.183
質問6	-1.023	0.778	-0.954	0.724
質問7	-1.191	1.085	-1.192	1.273
質問8	-0.879	0.331	-0.841	0.243
質問9	-0.936	0.551	-0.908	0.569
質問10	-0.932	0.895	-0.964	1.191
質問11	-0.412	-0.402	-0.471	-0.444
質問12	-0.870	0.590	-0.901	0.750
質問13	-1.021	0.945	-1.036	1.115
質問14	-0.951	0.697	-1.013	0.897

き歪度 s_i と尖度 k_i は

$$s_i = E \left[\left(\frac{X_i - \mu_i}{\sigma_i} \right)^3 \right], k_i = E \left[\left(\frac{X_i - \mu_i}{\sigma_i} \right)^4 \right] - 3$$

で定義される。ここで E は期待値を表す。表5に各質問項目の歪度と尖度を示す。すべての質問項目で

歪度が負なので、分布は右に偏っている。すなわち、すべての質問項目で、高い評価が与えられる傾向があることが分かる。

また、尖度は質問2、質問5、質問11が他の質問項目と比べると小さくなっている。これらの質問項目は1、2などの低い評価もある程度与えられる傾向があることを示している。この特徴は、これらの質問項目の平均情報量が高いことから裏付けられる。

4 質問13（満足度）と他の質問項目との関連性

4.1 相関係数と相互情報量

授業評価の主要な目的は、学生の授業に対する満足度を知ることである^[2]。そのため、表2の質問項目の中で、質問13「総合的に判断して、この授業に満足できましたか？」は、特に重要な指標と考えられる。そこで、どの質問項目が質問13を高めるために重要であるのかを調べるために、まず質問13と他の質問項目との相関係数と相互情報量を求めた。

相関係数は2変数間の線形的な関連性を示す指標であり、質問13と*i*番目の質問項目の相関係数は

$$R(X_{13}, X_i) = E \left[\frac{(X_{13} - \mu_{13})}{\sigma_{13}} \frac{(X_i - \mu_i)}{\sigma_i} \right]$$

で定義される。

また、相互情報量は2変数間の統計的独立性を示す尺度である。質問13と*i*番目の質問項目の相互情報量は

$$I(X_{13}, X_i) = H(X_{13}) + H(X_i) - H(X_{13}, X_i)$$

で定義される。ここで、 $H(X_{13}, X_i)$ は X_{13} と X_i との結合エントロピーである。相互情報量は0以上の値をとり、0のとき2変数は完全に統計的独立、すなわち一方の値が分かっても、他方の値は全く予測できない。

表6に、質問13と他の質問項目との相関係数と相互情報量を示す。前期と後期ともに質問13と相関が0.7以上のものは、質問6「授業では、学生の意欲化や理解を深める上での工夫がみられましたか?」、質問8「学生の関心や理解度を考慮しながら授業を進めていましたか?」、質問10「全体を通して、授業の内容が理解できましたか?」、質問14「授業を終えた今、自分でさらに学習を深めてみたいと思いますか?」であった。また、これらは、他の質問項目と比較すると相互情報量も大きな値をもつことが確認できる。

一方、質問2、質問5、質問11は、前期、後期ともに相関係数と相互情報量の値が低いことが分かる。これらは、質問13と関係性が低いことが分かる。

表6. 質問13（満足度）と他の質問項目との相関係数と相互情報量（講義科目）

	前期		後期	
	相関係数	相互情報量 (bit)	相関係数	相互情報量 (bit)
質問1	0.661	0.458	0.652	0.460
質問2	0.538	0.298	0.524	0.310
質問3	0.658	0.417	0.664	0.457
質問4	0.626	0.383	0.625	0.392
質問5	0.479	0.255	0.510	0.291
質問6	0.716	0.529	0.704	0.513
質問7	0.647	0.416	0.648	0.434
質問8	0.705	0.514	0.704	0.520
質問9	0.676	0.475	0.682	0.504
質問10	0.752	0.608	0.742	0.588
質問11	0.508	0.313	0.551	0.346
質問12	0.653	0.449	0.670	0.473
質問14	0.715	0.534	0.716	0.556

4.2 回帰分析

回帰分析は、ある1つの変数（従属変数）を、その他の幾つかの変数（独立変数）で予測するための手法である。ここでは、質問13「総合的に判断して、この授業に満足できましたか?」を従属変数、他の質問項目を独立変数として回帰分析を行った。ただし、質問14「授業を終えた今、自分でさらに学習を深めてみたいと思いますか?」は、授業を受けてみて、その後感じた感想であるので独立変数から除外した。

*i*番目の質問項目の*j*番目（ $j = 1, \dots, N$ ）の評価値を x_{ij} とする。このとき、重回帰式は

$$x_{13j} = a_0 + \sum_{i=f(1)}^{f(N_r)} a_i x_{ij} + e_j$$

で表現される。ここで、 a_i は偏回帰係数であり、 e_j は誤差、 N_r は採用する独立変数の総数であり、 f は採用する独立変数の番号を指定する関数である。回帰分析では、 j に関する誤差2乗和が最小になるように偏回帰係数を決定する。

表7、表8に前期、後期の回帰分析の結果を示す。 N_r は、ステップワイズ法により得られた結果から、調整済みR2乗の増加率とVIFの値から総合的に判断して、 $N_r = 3$ とした。VIFは他の独立変数との線形的な関連の強さを示す指標である。なお、標準化係数とは、

$$x_{ji}^* = \frac{(x_{ij} - \mu_i)}{\sigma_i}$$

とデータを標準化した後の偏回帰係数である。

この結果により、満足度に影響を及ぼすものは、影響度を示す標準化係数の高いものから順に質問10

表7. 回帰分析の結果（前期 講義科目）

	偏回帰係数	標準化係数	有意確率	VIF
定数	0.296		0.000	
質問10	0.412	0.409	0.000	1.986
質問6	0.304	0.320	0.000	1.901
質問12	0.227	0.230	0.000	1.687
調整済みR2乗			0.686	

VIF: (Variance Inflation Factor)

表8. 回帰分析の結果（後期 講義科目）

	偏回帰係数	標準化係数	有意確率	VIF
定数	0.228		0.000	
質問10	0.393	0.379	0.000	2.067
質問6	0.305	0.311	0.000	1.871
質問12	0.258	0.258	0.000	1.760
調整済みR2乗			0.673	

「全体を通して、授業の内容が理解できましたか？」と質問6「授業では、学生の意欲化や理解を深める上での工夫がみられましたか？」と質問12「試験の内容やレポートのテーマ、分量等は適切でしたか？」の順であることが分かる。

なお、質問13と相関の高かった質問8「学生の関心や理解度を考慮しながら授業を進めていましたか？」は、影響度があまり高くないことが分かった。これは、質問6と相関が高いことが原因の1つと考えられる。質問6と質問8は、質問項目としてまとめることができる可能性がある。

4.3 数量化I類

数量化I類は独立変数が質的データの場合の回帰分析である。ここでは、各質問項目の5段階の評価値に対し、ダミー変数と呼ばれる0か1しかとらない変数を5個用意し、評価値をダミー変数に変換した後、以下の式により回帰分析を行う。

$$x_{13j} = a_0 + \sum_{i=f(1)}^{f(N_j)} \sum_{k=1}^5 a_i^k x_{ij}^k + e_j$$

ここで、 a_i^k はカテゴリウエイトと呼ばれ偏回帰係数に相当する。 i はアイテム、 k はカテゴリを表し、ここでのアイテムは各質問項目の番号であり、カテゴリは評価値を示している。 x_{ij}^k はダミー変数であり0か1の値をとる。また、定数項 a_0 は x_{13j} の平均値である。表9に評価値をダミー変数に変換する方法を示す。

数量化I類では、カテゴリウエイト（偏回帰係数）の値そのものが、回帰式の増分値（減少値）となる。そのため、式の意味が理解しやすい。表10、

表9. 評価値のダミー変数への変換

評価値	x_{ij}^1	x_{ij}^2	x_{ij}^3	x_{ij}^4	x_{ij}^5
1	1	0	0	0	0
2	0	1	0	0	0
3	0	0	1	0	0
4	0	0	0	1	0
5	0	0	0	0	1

表10. 数量化I類の結果（前期 講義科目）

評価値	カテゴリウエイト		
	質問6	質問10	質問12
1	-0.973	-1.480	-0.621
2	-0.649	-0.831	-0.479
3	-0.320	-0.440	-0.248
4	-0.048	0.004	-0.030
5	0.270	0.359	0.224
定数		4.15	
調整済みR2乗			0.687

表11. 数量化I類の結果（後期 講義科目）

評価値	カテゴリウエイト		
	質問6	質問10	質問12
1	-0.995	-1.414	-0.894
2	-0.611	-0.934	-0.560
3	-0.364	-0.412	-0.254
4	-0.034	0.010	-0.048
5	0.264	0.307	0.248
定数		4.15	
調整済みR2乗			0.676

表11に前期、後期の数量化I類の結果を示す。用いたアイテムは、4.2節の回帰分析で得られた「質問6」、「質問10」、「質問12」である。

カテゴリウエイトを比較すると、やはり質問10の影響が強い。すなわち、質問10で高い評価値を与えた人は、質問13の評価値を大きく増加させ、低い評価値を与えた人は大きく減少させるためである。例えば、前期において質問10で5と答えた人は質問13の値を平均値4.15から0.359増加させ、4.509とすることが分かる。

4.4 主成分分析

主成分分析は、データ縮約のための1つの手法である。通常は1人の授業評価の結果を表現するためには質問項目の数と同じ14次元必要であるが、主成分分析を適用した場合、少数の主成分でデータの大半を表現することが可能となる。

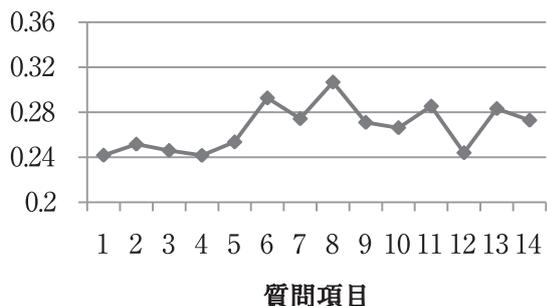


図4. 第1主成分に対応する固有ベクトル (前期講義科目)

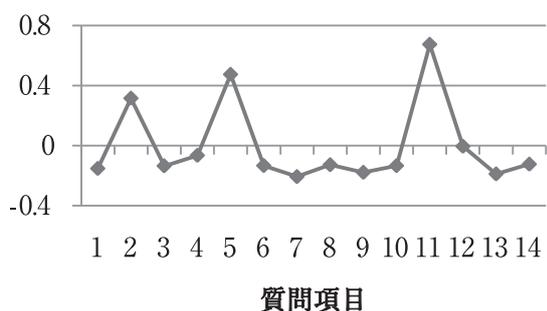


図5. 第2主成分に対応する固有ベクトル (前期講義科目)

表12. 寄与率

	前期	後期
第1主成分	0.597	0.607
第2主成分	0.070	0.065

$X = (X_1, \dots, X_{14})^T$, $\mu = (\mu_1, \dots, \mu_{14})^T$ とする。 X の共分散行列

$$E[(X - \mu)(X - \mu)^T]$$

の固有ベクトルを固有値の大きい順にならべたものを u_i ($i = 1, \dots, 14$) で表す。 $u_1^T X$, $u_2^T X$, $u_3^T X \dots$ はそれぞれ第1主成分、第2主成分、第3主成分と呼ばれる。図4、図5に第1、第2主成分に対応する固有ベクトルを示す。

j 番目の評価値 $x_j = (x_{1j}, \dots, x_{14j})$ は、少数の主成分 $u_i^T X$ を用いて

$$\hat{x}_j = \sum_{i=1}^{N_p} (u_i^T x_j) u_i$$

としても、 \hat{x}_j は x_j の良い近似となる。ここで、 N_p は用いる主成分の数である。表12に第1、第2主成分までの寄与率を示す。第2主成分までで前期、後期ともに67%程度データを表現できることが分かる。

主成分分析を適用した結果、第1主成分に対応する固有ベクトルの成分はすべて正であるので、これは総合的に授業が良かったと評価している指標であると考えられる。第1主成分の因子負荷量 $R(u_1^T X,$

表13 因子負荷量

	前期		後期	
	第1主成分	第2主成分	第1主成分	第2主成分
質問1	0.751	-0.161	0.751	-0.107
質問2	0.712	0.305	0.714	0.298
質問3	0.789	-0.148	0.780	-0.218
質問4	0.763	-0.070	0.758	-0.110
質問5	0.670	0.428	0.690	0.397
質問6	0.845	-0.131	0.838	-0.145
質問7	0.782	-0.201	0.786	-0.241
質問8	0.846	-0.120	0.842	-0.151
質問9	0.793	-0.178	0.808	-0.169
質問10	0.815	-0.140	0.829	-0.090
質問11	0.698	0.564	0.735	0.522
質問12	0.729	-0.005	0.751	0.027
質問13	0.859	-0.195	0.857	-0.175
質問14	0.786	-0.122	0.789	-0.092

x_j)を表13に示す。これより、質問6、質問8、質問10、質問13で高い値が得られている。そのため、これらの質問項目は、学生が授業を評価しているかどうかを判断するために、特に重要であると考えられる。

また、第2主成分に対応する固有ベクトルの成分は、質問2、質問5、質問11が正で、その他は負である。これらの質問項目は、他の質問項目の回答と明らかに違う傾向がある。第2主成分に対応する因子負荷量も表13に示す。これらの質問項目は、授業の満足度に必ずしも強く寄与しない可能性がある。

5 まとめ

本稿では、2008年度に行われた山口県立大学の学期末授業評価について、特に講義科目の結果に絞って報告した。得られた結果は、2008年度においては、以下の通りであった。

(1) 各質問項目の評価値は、2008年度の前期と後期で、ほぼ同様の統計的性質が見られる。これは、前期、後期に関わらず各質問項目に固有の回答傾向があることを示している。また、質問項目によって確率分布にかなり違いがある。

(2) 質問13 (授業の満足度) の評価を向上させるためには、質問10 (学生の理解度)、質問6 (学生の意欲化や理解を深める上での工夫) が特に重要で、加えて質問8 (学生の関心や理解度を考慮)、質問12 (試験の内容やレポートのテーマ、分量) も重要である。ただし、質問6と質問8は、類似した質問であるため、統合できる可能性がある。

(3) 質問2、質問5、質問11は、他の質問項目と

かなり違う評価のされ方をする傾向があり、また質問13の評価にそれほど影響を及ぼさない可能性がある。ただし、これらの項目が重要でないということではない。

今後は、年度ごとの学期末授業評価の比較^[4]や成績データと授業評価との関連の調査^[3]などが考えられる。

参考文献

- [1] 永崎研宣, 「全学毎回授業評価システムの開発と運用」, 平成17年度情報処理教育研究集会講演論文集, pp.109-112, 2005年.
- [2] 星野敦子, 牟田博光, 「学生による授業評価における授業の満足度に影響を及ぼす諸要因」, 日本教育工論文誌, Vol.27 suppl., pp.287-290, 2003年.
- [3] 西田昌彦, 「工学基礎物理・授業評価の分析」, 工学教育研究 No.14, pp.59-79, 2008年.
- [4] 田実潔, 竹原卓真, 「学生による授業評価に基づいた授業改善への探索的研究 (II) - 授業アンケートの分析から -」, 北星学園大学社会福祉学部北星論集, 第46号, pp.65-72, 2009年.
- [5] 長谷川勝久, 原田由香里, 「教育系大学の学生を対象とした授業満足度アンケート項目の開発」, 日本教育工学会論文誌, Vol.30, No.4, pp.447-452, 2007年.
- [6] 森光義昭, 関聡, 「授業評価における改善点の基軸」, 久留米信愛女学院短期大学研究紀要, 第30号, pp.45-57, 2007年.
- [7] 中野良哉, 「学生の授業評価に基づく授業改善の試み - 講義型受動的学習型から能動的学習型への展開 -」, 高知リハビリテーション学院紀要, 第9巻, pp.9-16, 2008年.
- [8] 渡辺晃正, 「学生による授業評価の分析方法 - 数量化Ⅱ類による分析 -」, 桜花学園大学研究紀要, 第4号, pp.105-120 2002年

