

# 授業データ解析による授業改善策発見を目指して

## — 努力・成果・評価の関連性からのアプローチ —

### Toward Discovering Improvement Ideas of Teaching from Lecture Data Analysis

– An Approach to Correlation Analysis for Effort, Achievement, and Evaluation Data –

南 俊朗, 大浦 洋子

Toshiro Minami and Yoko Ohura

#### 【要 約】

多くの大学において教師達は、学ぶ意欲が十分ではない学生を相手に如何に教育効果を上げるか日々努力を続けている。学力の低い学生達を観察すると、彼らは学力以前に学ぶことへの動機づけが十分でなかったり、学ぶとはどういうことかを意識していなかったりと、学びの基礎となるべき、知識に対する好奇心やノートを作成したり、復習や予習をしっかりとやるなどの心の姿勢に問題がある場合が多い。本稿では、学生の学力やその基礎となる学びへの意欲などを授業データに基づいてモデル化することを試みる。授業データとしては平常点の基礎となる出席や宿題提出状況などのほか試験の採点データ、そして学期末に実施された学生自身と授業への評価アンケートの結果を用いる。これらのデータを基に、学生の努力・成果・評価の相互関連を分析する。個々の学生に対する主観的な印象情報に加えて、このようなデータに基づく客観的な知見を併用することによって、より精密な学生モデルを得ることができ、ひいてはより良い授業改善に繋がることが期待できる。

**キーワード:** 授業データ解析, 教育データマイニング, 授業アンケート, 学習者モデル

#### [Abstract]

University professors are struggling with improving their lectures to be more effective against the students who have not enough eagerness to learn in the everyday lectures. As we observe the students who have low performance in learning, we find that their problems are not in their learning ability but in their lack of motivation to learn, lack of curiosity to knowledge, lack of skills for learning, lack of recognizing the importance of note taking, etc. In this paper we present our trial of modeling student's attitudes such as willingness to learn and other ones based on lecture data. The data we use in this paper are attendance and homework scores for the index of student's effort, and examination scores for the index of their achievement, together with their evaluation data on themselves and on the lectures. We analyze their correlations from these data. We believe the resulting knowledge extracted from such analysis gives useful knowledge for improving the lectures.

**Keywords:** Lecture Data Analysis, Educational Data Mining (EDM), Lecture Evaluation Enquiries, Learner Model

## 1 はじめに

### 1.1 研究の背景

大学生の学力低下が指摘されて久しい[15]。その背景には大学の進学率の増加や少子化の結果少ない大学受験生を多くの大学が取り合う形になっていることや、いわゆる「ゆとり教育」の結果、高校までの教育において学力養成が以前と比べて軽視されてきたきらいがあることなど多くの要因が存在するものと考えられる。そのため大学において、高等教育を受けるのに十分な学力のないまま大学生になってしまった学生が珍しくない状況となっている。

その対策として大学側は、リメディアル教育による学生の基礎学力の底上げを図ったり、初年次教育を強化したり、教員の意識を変え、教育力を増強させるためのFD (Faculty Development) 活動を導入するなど様々な工夫を行ってきた。そのような対策にもかかわらず、状況は決して好転していない。むしろ悪化の一途をたどっているさえある。

その最大の原因は教員の教育力や、学力そのものにあるのではなく、学生自身の意識や意欲などにあるのかも知れない。すなわち、学生の学ぶ意欲や新しい知識に対する好奇心や研究心、探索心などの更に根源的な心の在りようをも含めて原因を探求することが必要であると考えられる。

実際、学力の低い学生達を観察すると、彼らは学力以前に学ぶことへの動機づけが十分でなかったり、学ぶとはどういうことかを意識していなかったりと、学びの基礎となるべき知識に対する好奇心、ノートを作成したり復習や予習をしっかりとやるなどの心の姿勢に問題がある場合が多い。

したがって学生の学力低下という問題への解決策を追求していくためには、教育を与える側である教員に対するFDという観点からの対策だけでは不十分であり、教育を受ける側である学生に対する学生SD (Student Development) という観点からの対策も、それに劣らず重要である。本稿ではこのような問題意識に基づき、授業データの解析を通して学生の学びへの姿勢を分析する。

### 1.2 授業データ解析からの知識発見

様々なデータから役立つ知識を発見する手法はデータマイニング(Data Mining)という呼称で広く研究されてきた[1]。近年はWebシステムの発展および普及により、アクセスログなどのデータが大量に得られるようになったことを背景に、ビッグデータ(Big Data)からのデータマイニングが注目を集めている[20]。

教育への応用を意図したデータマイニング、あるいは、データ解析の技術に関する研究はKDD (Knowledge Discovery and Data Mining) [1]やEDM (Educational Data Mining) [4][21][22]などの名称で研究が行われている。本稿の目的である授業データの解析による有用な知見の獲得を目指す研究もこれらの分野の一部と見なすことができる。

論文[22]はe-Learningシステムから得られるデータを用いて学生を分類する問題に対するデータマイニングアルゴリズムを比較した。その目的は学生の成績を事前に予測することにある([24]も同様)。学生の学習成果と見なせる最終成績を予測することで、問題のありそうな学生を早期に発見し、対策を講じることができる。

本稿の研究目的も概ね同様であるが、我々の主要な興味は、得られたデータから直接的に成績を予測することではない。成績以前の問題として、学生のやる気や勤勉さなどの心の姿勢や性格、あるいは心理的な傾向に関するモデル化を行い、それに基づき、いわば間接的に成績を予測し、問題のあることが予測できる学生に対して根源的にケアすることを目指す。

我々の研究はコンピュータ支援による協調学習(Computer Supported Collaborative Learning, CSCL)の研究領域とも関連する。論文[23]では、学生間の交流(Interaction)を構造化することを目的に協調して学ぶための場を扱っている。本テーマに関する我々のアプローチは、まず学生達の学習行動に関する特徴を見出し、その後、得られた知見をベースに学生達の興味や習得知識のレベルのバランスなどを考慮した上で最適な協同学習グループの構築を提案するというものである。

### 1.3 学生のモデル化による学習支援へのアプローチ

我々のデータ解析へのアプローチは授業改善のための知見を得るという大きな目的に関しては、従来のアプローチと基本的に共通である。しかし前節でも述べたようにいくつかの相違点がある。

#### (1) データ解析の手法

本稿におけるデータ解析は、従来研究で多く行われている授業データに対する直接的なデータマイニングではなく、学びに対する学生の心の姿勢をモデル化することに大きな焦点を当てている。

#### (2) 対象データのサイズ

従来のデータ解析はビッグデータと呼ばれる大量なデータからの知識獲得を目指す場合がほとんどである。それに対して本稿の対象データである授業データは、それらと比べると極めて少量であり、そのようなスモールデータであっても有用な知識獲得ができる手法の開発を目指している。

#### (3) 対象データの獲得手段

前項とも関連するが、大量のデータを獲得するためには、何らかの自動的手段を導入する必要がある。実際、従来研究の多くは e-Learning システムのログデータを用いるなどにより大量のデータを得ている。一方本稿が対象とするデータは特別なコンピュータシステムなどを仮定しない一般の授業データなどのスモールデータである。したがって、手作業により必要なデータを作成することも想定内である。

#### (4) 例外ケースの取り扱い

データマイニングや統計手法を含め従来のデータ解析においては、通常、例外的なデータを除外した全体の特徴を掴むことが大きな目的である。それらから外れたデータはしばしばノイズデータとして扱われる。一方学生に関するデータにおいては、全体から外れたデータであろうとも、それはそれで人格を持った 1 人の学生に関するデータである限り、それらは尊重されるべきである。

そのような外れたデータをも考慮に入れ、それを如何にして教育に活かすことができるかを考えることが重要である。

#### (5) 解析ツールの開発

データ解析に関する従来の研究の多くは既存のツールを様々なデータに適用したものである。たとえばデータマイニングの研究において統計的な手法と同時に相関規則 (Association Rule) を発見する研究は多い[2]。

本稿における授業データ解析の研究は、データやアプローチに独自性があり、従来の解析手法をそのまま適用する訳にはいかない。したがって、我々は設定した解析課題に適合するような解析手法を考案し、それを改善しながらデータ解析を進めて行くスタイルでの研究を行う。これも本稿の研究において重要性の高いテーマである。

我々と同様にスモールデータを対象とした授業データ解析の研究も少数ながら存在する。合田等は毎回の授業で学生に自己評価コメントを提出させ、それを手動でデータ化し、コメント内容から成績を推定する研究結果を報告している[5][6][7]。

本稿で用いる授業データは平常点の基礎となる出席や宿題提出状況などや試験の採点データ、そして学期末に実施された学生自身と授業への評価アンケートの結果である。我々の場合も学生の自己評価アンケート結果を利用するものの、本稿では学生が与えた評価値を直接利用し、また、評価アンケート中のテキスト情報の利用にあたっては評価値と並んで解析に主観的判断の入り込みにくいと考えられる単語の出現データなどを用いた解析を行っている。したがって、テキストから評価データを解析者の判断を加えて抽出するなどの作業は発生しない。

その他、我々の評価アンケートは期末に 1 度実施するだけであり、毎回実施してはいない。この点だけではなく、データ解析の目的が学生の成績を推定するのではなく、学生モデルの構築にある点も合田等の研究と相違している。

本稿ではこのようにして得られたこれらのデータに基づき学生の努力・成果・評価の相互関連を分析する。個々の学生に対する主観的な印象情報

に加えて、データに基づく客観的な知見を併用することによって、より精密な学生モデルを得ることができ、ひいてはより良い授業改善に繋がることが期待できる。

本稿で用いられるデータ解析へのアプローチと同様な手法により我々は図書館データの解析も行ってきた。論文[12][13][14][16][17]において、我々は図書館の貸出記録データを対象に、図書や利用者に関する専門度などの新しい概念や指標を提案し、解析を行った。論文[11]では図書館内での座席利用データを用いた解析を行った。

これらの経験を通じて、このようなアプローチは、場合によっては不完全で十分な情報を含んでいないかもしれないスモールデータからも有用な結果がもたらされる可能性が高いという確信が得られた。

このような目標や背景を念頭に本稿は以下、次のように構成される。まず第2節では努力と成果の関連性を分析する。その結果を踏まえて、第3節では新たに自己・授業評価アンケートにおける評価データを加え、努力や成果と評価の相互関係を追究する。最後に第4節でまとめと今後への展望を行う。

## 2 努力・成果の関連性分析

本節では、授業に関する努力の指標として出席点と宿題点を、そして成果の指標として試験点を用いて学生の努力と成果の関連性を調べる(本節で述べられる研究に関しては[19]も参照のこと)。

まず2.1節で分析に用いられる授業データの概要を説明した後、2.2節で実際に分析を進める。その分析結果を踏まえた上で3節において評価データを含めた解析および考察を行う。

### 2.1 努力・成果データの概要

本稿で解析に用いるデータはある女子短期大学2年生向けの「情報検索演習」科目における、期末試験の結果や出席状況、そして宿題に関する評価点である。この科目は司書資格の必修となっているため35名の受講者のほとんどは一般の選択

科目と比べ、かなり真剣に受講しているものと考えられる。

本科目の最大の目的は、学生達が司書資格の取得者として情報検索に関する十分な知識やスキルを身につけることである。そのためには、検索の目的を理解し、どういう情報が求められているのかを考えて、適切な検索手段や検索サイトなどを選択し、また、適切な検索キーワードを発想することのできる能力を習得することが肝要である。

本科目の授業回数は15回である。各授業の最初には出席確認を兼ねて簡単なクイズ問題を解いてもらう。これは頭の準備運動と呼ばれ、授業への集中力を高める目的で実施されている。出席点は出席回数をベースに、遅刻状況や頭の準備運動への取り組みなどを評価した結果を反映させるための修正が若干加えられている。

毎回の授業後に出题される宿題の主な目的は、それぞれの授業でどういうことを学んだのかを確認したり、次回に学ぶべき内容の予備調査を行い、次回の学習への導入を容易にすることなどである。宿題点は、出题回数や宿題の質を考慮して評価される。宿題の質の評価に関しては、内容の正確さよりも、問題へ取り組む姿勢、特に、与えられた課題を単にこなすだけではなく、どの程度深く考え、また、検索を工夫しているかなどを反映させた評価点となるよう努めている。

期末試験の問題は3つの問題からなる。問題1は、検索エンジンのサービスを提供している代表的なWebサイトを列挙し、それらの特徴などを要約して説明する問題である。この問題は、適切なキーワードを発想する力と、得られた情報を簡潔に要約する力の両方を評価するための問題であり、普段の宿題に真剣に取り組んできた学生には比較的容易な問題であると考えられる。

問題2は、電子書籍やオンラインの資料を提供しているサイトを見つける問題である。この問題も適切なキーワードを発想する力をはじめ、サイトの特徴を短時間で読み取ることを要求している。

問題3は、インターネットにおける情報犯罪に関して事例を探索し、また、その防止策などに関して議論を行う問題である。本問題の場合は、一般的な検索能力だけではなく、与えられたテーマ

に関して事実に基づいて自分の意見をまとめ、それを表現する能力も問われている。

問題の性格による違いはあるものの、基本的には90分程度の試験時間内で、適切な情報検索を実施し、その結果をまとめる能力への評価が定期試験の評価となるよう考えられている。検索を実施し、その結果を要約するという作業を時間的制約の下で行わないといけない点が普段の宿題とは大きく異なる状況であり、宿題解答の質が高い学生が必ずしも試験解答の質も高いとは限らない要因となっている。

## 2.2 努力点と成果点の関連分析

平常点評価の元になる出席や宿題を行うためには、学生にとって何らかの努力が必要とされる。本稿では、これらを努力の程度を示す2つの指標として用いる。同じく努力を表すとは言え、出席のための努力と宿題のための努力を比較すると、後者の方がより大きな意図的の努力が必要である点で相違する。一方試験点は努力の結果としての成果の程度を表す成果点とみなすことができる。

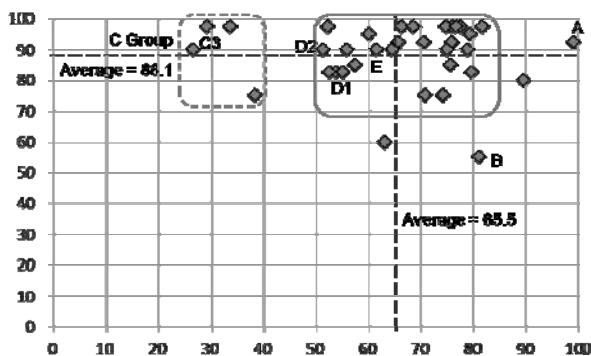


図1. 試験点(x軸)と出席点(y軸)の相関関係。  
相関係数 = -0.1

図1に試験点と出席点の相関関係を示す。図からも容易に読み取れるように、これらの相関関係は強くない。それどころか、相関係数は-0.1と負値になっている。その原因の1つとして考えられるのは図中C Groupとラベル付けられた学生、特に出席点の高い3名の学生の存在がある。彼女達はほとんどの授業に出席しており、一見まじめ

な学生に見えるものの、その成果の状況を見ると試験点が30点前後の最低値の一群を構成している。実際これらの学生を除いて相関係数を求めると0.1と正值になる。

学生達の多くは試験点が50台から80前後であり出席点が70台以上の実線領域に存在している。それらの学生達の多くは出席点が極めて高い値に固まっている一方、試験点は30点以上ばらつく。全体的に見ても、3名中1名程度の学生が極めて真面目に出席しているがC群の学生も含めて、それに見合う十分な成績を収めていないのが現実である。

これは出席すること自体が目的化し、学ぶべきことを十分学んでいない学生が多いことを示唆している。恐らくそのような学生達は、授業に参加すること自体に満足してしまい、学んだ気になっているのであろう。この結果は、教師に対して、学生達が真に学ぶよう一層の努力が必要であることを示していると解釈できる。

学生Aは試験点が最大であり、出席点も満点ではないものの90点以上を確保している。すなわちこの学生は極めて真面目な学生であり、学ぶために十分な努力をしており、またそれ相応の成果をだしている学生であると言える。

学生Bは平常点に関しては最低値の55点に過ぎないが試験点は80点を超えている(上位から4番目)。この学生は、潜在的には十分高い能力を持っているにも関わらず、何らかの理由により欠席が多く、出席点が低くなっているものと推測できる。

こういう学生の存在は、例外的という訳ではなく多くのクラスに存在しているものである。これらの学生達が普段の学習に対して、より一層努力を傾け、本来持っている能力を十分発揮し、より高い試験点、すなわち成果点を上げられるよう指導することが教師に求められる。

次に、x軸を試験点の平均値65.5で、またy軸を出席点の平均値88.1でそれぞれ上下や左右の2つの領域に分割することにより、全体を4つの象限に分けて分析する。第I象限は試験点と出席点の双方とも平均値以上の領域であり、第III象限は双方とも平均値以下の領域である。これら

の領域に存在する学生達は、概ね努力に見合った成果を上げていることになる。

一方第Ⅱ象限は努力の割に成果が低い学生群であり、第Ⅳ象限は努力に比して高い成果を収めている学生群である。特に第Ⅱ象限に位置する学生が第Ⅰ、Ⅲ象限に位置する学生と比較して数多く存在することが、試験点と出席点の相関係数が負になる原因であると見ることもできる。Cグループの学生達はその典型的な例である。

第Ⅱ象限に位置する学生達に関して2つの解釈が可能である。1つの解釈は、彼女達は学び方に問題があり、効果の上がるような学習ができていないというものである。授業中の演習や宿題の課題などを、それ相応の時間をかけてこなしてはいるものの、それが実践的な知識やスキルとして身につけていない。彼女達は与えられた演習課題を単にこなしているだけであり、それから何かを学ぼうという姿勢が欠如しているのかも知れない。これらの学生群の存在は、教師に対して、宿題や演習内容の工夫や、そこから何を学ぶべきであるかを学生達に徹底するための一層の努力が求められていると解釈できる。

可能性のあるもう1つの解釈は、学生達が授業に出席しているのは何かを学ぶためではないというものである。彼女達が欲しいのは実は卒業のための単位や司書の資格だけであり、学ぶこと自体を全く目的とはしていないということである。そのため宿題などの課題に対しても、単に解くだけであり、より深く調べ、真の意味でより多く学ぶことは元々意図していないのかも知れない。こういう学生達に対しては、授業を通して身につけた知識やスキルが卒業後、社会人、あるいは職業人として仕事をしていく上で、いかに役立つか納得できる説明に一層の工夫と努力が必要であろう。

今のところ、授業データなどから学生達の真の意図を推測する手段はない。今後研究すべき課題である。

図2に試験点と宿題点の相関を示す。図1に示した試験点と出席点との相関と同様にこれらの相関は小さい。しいて言えば試験点と宿題点の相関係数(-0.0)は試験点と出席点の相関係数(-0.1)よりわずかに大きい。その差は小さいとはいえ、

宿題点の方が出席点より学習成果の指標である試験点への影響が大きいということになる。これは恐らく、単に出席するだけよりも宿題をする方がより大きな努力を必要とする行為であり、それだけ努力したことによる学習効果が高く、結果として成績への影響がより大きく表れたものと考えられる。

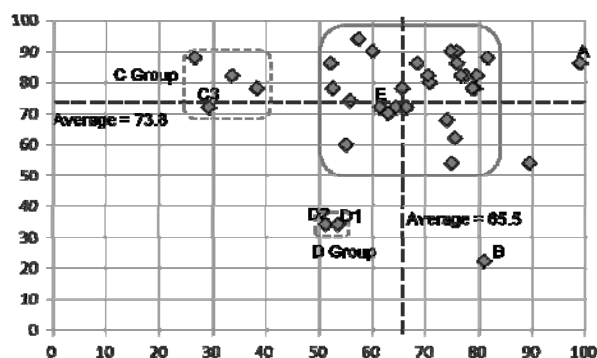


図2. 試験点(x軸)と宿題点(y軸)の相関関係。  
相関係数 = -0.0

更に詳細に調べると、図2ではC群の学生達の値のばらつきが大きくなっている。また図1で中心群に存在した学生の2名が図2では下方に移動し新たにD群を形成している。学生Bも同様である。これらの学生達は出席よりも一層の努力を必要とする宿題に熱心に取り組まなかったものであろう。

学生Aは図2においても高い値を維持している。このことから判断しても彼女は真に優良な学生であり、出席だけではなく宿題もきちんとこなし、恐らくは授業中の演習にも真面目に取り組んだと考えられる学生である。

次に図1と同様に、平均値によるそれぞれの軸を2つの領域に分けて分析する。図1、図2のいずれの場合も、概ね40%は第Ⅰ象限に、そして残りは3つの象限それぞれに20%ずつ学生が存在する。

更に詳しく調べると、努力に対して成果の低い第Ⅱ象限の学生数は、図1(出席)の方が図2(宿題)よりも僅かに多い。この結果も、より努力が必要とされる宿題の実施および提出に対して学生達が不熱心であることを示している。

### 3 自己評価や授業評価に関する分析

本節では、前節の結果を踏まえ、分析を更に深めるために学生の授業・自己評価と努力・成果点との相関関係を調べる。第3.1節で評価データの概要を説明し、第3.2節で実際の比較を行い、その結果を議論する。

#### 3.1 評価データの概要

本稿で用いられる学生の評価アンケートは授業の終盤、期末試験の直前に、当期の授業全体を振り返って実施されたものである。本アンケートは大学においてFD活動の一環として全学いっせいに実施される授業(評価)アンケートとは異なり、記名アンケートであり、無記名のアンケートと比較して学生達はある程度の責任感を持って、質問に回答する。もちろん、自分に関する良い評価にも悪い評価にもある程度の脚色が加えられることは予想できるものの、著者等の観察によると、マイナス評価の程度を弱めることはあってもマイナスの評価をプラスに評価する脚色はほとんどない。

授業・学生評価アンケートは全部で12の質問項目からなる。そのうち5問は授業に関する質問であり、6問は学生自身に関する質問、そして残り1問は一般的なコメントを求めている。

本アンケートの質問項目は次のものである：

##### 【授業の評価】

- (1) あなたはこの授業でどういうことを学びましたか？それはあなたの役に立ちましたか？
- (2) 授業全般について、良かったところはどこでしょうか？
- (3) 逆に授業についての問題点は何でしょうか？それらはどう改善したらよいでしょうか？
- (4) 上記の評価を総合して、あなたはこの授業に100点満点の何点つけますか？
- (5) その他、授業に関して気づいた点や感想などを書いてください。

##### 【自分自身の評価】

- (6) この授業に対するあなたの学習態度について良かったところは何でしょうか？どういう努力をしてみました？

- (7) この授業に対するあなたの学習態度の反省点は何でしょうか？今後の自分の学習態度はどのように改めたらよいでしょうか？

- (8) あなたは授業の狙いの理解に努め熱心に取り組みましたか？自分の受講態度を振り返って評価してください。

とても良い、良い、普通、余り良くない、悪い

- (9) あなたは授業中、担当講師に質問しましたか？質問に対して講師は適切に答えたでしょうか？

何度も質問した、1度だけ質問した、質問したかったができなかった、質問なし。

- (10) 授業後に疑問点を解消するための調査をしたことがありますか？

よく調べた、たまには調べた、疑問点を調べたことはない、疑問を感じない。

- (11) 上記の評価を総合して、あなたはこの授業に対する自分の取り組みについて100点満点の何点つけますか？

##### 【その他の質問】

- (12) その他、これまでの項目以外のことについての意見や気づいたことなどを書いてください。

#### 3.2 自己評価と努力・成果点の関連分析

学生によって回答した項目やその数にはばらつきがある。35名の受講生中20名の学生が授業評価点の質問(4)と自己評価点の質問(11)の両方に回答した。これらの回答を用いて自己評価と授業評価の関連性に関する考察を進める。

質問に回答した20名の学生の評価結果を用いて図3に自己評価点と出席点の相関関係を示す。また、図4に自己評価点と宿題点の相関関係を示す。相関係数は図3においては0.4、図4においては0.5であり自己評価との相関は宿題点の方が大きい。

本稿では自己評価点を回答した学生が本授業における自らの努力や成長を評価した点数と捉え、学生が本授業を受けての満足度と解釈することにする。それは、学生の身になって考えると、設問に回答する中で、自分が成し遂げた成果やそのた

めに払った努力などを客観的に評価するよりも、その授業全体を振り返って、自らが十分な努力をしたかどうか、もっと努力すべきではなかったか、もっと努力できたのではなかったか、といった自問の結果として、むしろ主観的に自己評価をする傾向があると考えられるからである。

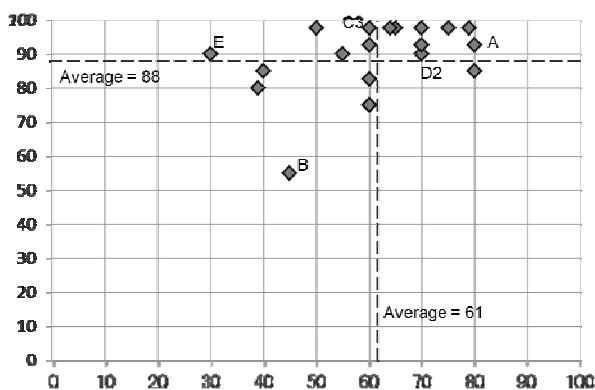


図3. 自己評価点(x軸)と出席点(y軸)の相関関係。相関係数 = 0.4

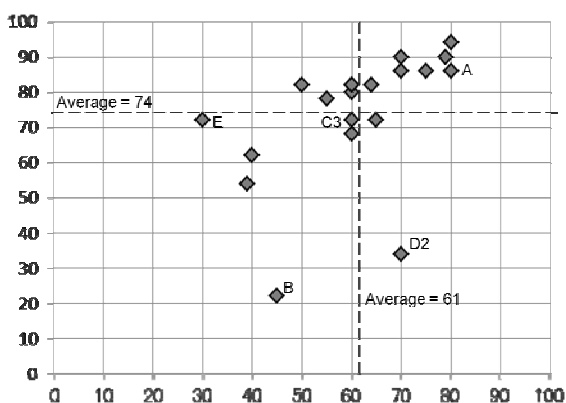


図4. 自己評価点(x軸)と宿題点(y軸)との相関関係。相関係数 = 0.5

このように考えると、学生達も宿題の方が出席と比べてより大きな努力が必要であることを認識しており、十分に宿題に取り組み、それらを提出した自分に満足を感じているため、出席点に対してよりも宿題点の方がより大きな相関があるものと推測できる。

次に個々の学生の自己評価に注目する。学生A, B, C3はそれぞれの状況に見合ったところに位置している。Aは出席・宿題・試験のいずれをとっ

ても高いレベルにあり満足度も高い。C群の学生の中で唯一評価値を回答したC3は高い出席ではなく、幾分低い宿題提出に見合う自己評価をしている。

Bは出席・宿題のいずれも低い値であり、それに合わせて満足度も低く評価している。しかし子細に見ると学生全体の傾向と比較して高めに自己評価している。これは自分の努力が不足していることを自覚しつつも、それでもある程度は努力したと考え、高めに評価したい心理的働きによるものと考えられる。

一方学生D2とEはこれらの学生とは大幅に異なる。D2はD群に属している学生である。すなわち、平均的學生程度に真面目に出席はしているが、宿題提出状況は極めて悪い。それにもかかわらず、自己評価は平均以上である。この学生は、出席のみで満足しているか、あるいは、宿題提出状況が悪いことを十分意識しながら、それを直視するのを避けるために、結果として現実に対する自己評価とは異なる評価を回答しているものとも考えられる。このように自分自身に甘い傾向のある学生であるためか、学生D2は、図2から分かるように成果の指標である試験点は平均以下である。このような学生には、現実を直視し、真に学ぶ努力を促す対応が必要である。

EはD2と逆である。Eの出席や試験の評価値、すなわち努力の度合いは平均程度であるにもかかわらず満足度は最低値である。この学生は自分に対する要求レベルが高いものと考えられる。このような学生もAと並んで良い学生と考えられる。講師には学生のやる気を刺激して、より高いレベルの学習に導くよう指導することが望まれる。

図5に自己評価(満足度)と試験点(成果)の相関関係を示す。興味深いことに相関係数は負値(-0.1)である。すなわち、成果が大きいほど満足度が低下する傾向があることになる。この事実はどう解釈できるであろうか?

更に分析を進める前に、評価に関する分析対象は、自己評価値と授業評価値の両方を回答した学生だけであることを再度指摘しておきたい。そのような学生は全受講生35名中20名に過ぎない。それら学生の試験点の平均値は71であり、それ



は全学生の平均値 65.5 より 5.5 点も高い。一方、出席点や宿題点に関しては大きな違いは見られない。出席点については全体の 73.8 点に対して 74 点、宿題点については全体の 88.1 点に対して 88 点である。すなわち、図 5 に現れる学生達は概して試験点は高く、努力点も相応である。

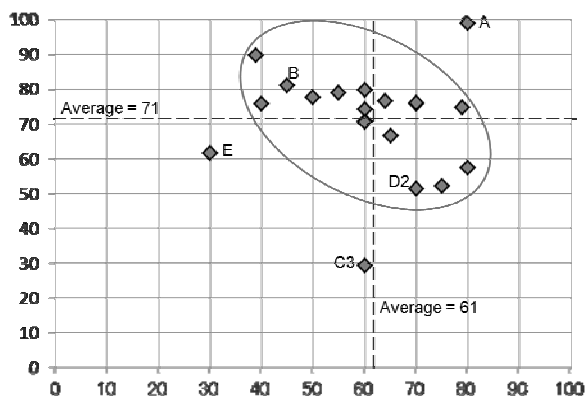


図 5. 自己評価(x 軸)と試験点(y 軸)の相関関係。  
相関係数 =  $-0.1$

次に成績の平均点により全体を 2 つのグループに分けて考察を進める。高成績グループのほとんどの成績点は 70 台に分布する。満足度に関しては 40 程度から 80 程度の範囲内に広く分散している。

高成績グループに関して興味深いのは、成績が高い学生の満足度が低いこと、すなわち成績と満足度の間に負の相関があることである。その意味では学生 A は例外である。この現象は恐らく、彼女達の自分自身に対する要求レベルが高いことを反映しているのであろう。これは自己向上心が高いと言い換えることもできる。

満足度の範囲が広いことを併せて考えると、自分の成果に満足し、自身の努力を満足感を持って振り返る気持ちと、もっと努力することでもっと多くの成果を上げることができたと反省する気持ちの両方が絡み合った結果、満足度ないしは自己評価に関して、結果として前者を強調する判断から後者を強調する判断に分かれたものと見ることも可能である。

一方、低成績グループに属する学生達の多くは平均以上の満足度評価になっている。グループ内

では成績と満足度の関係が弱く、学生個人の判断のばらつきが大きい。

学生 E は例外である。結果的に成績は平均値を少し下回る程度であるにもかかわらず、満足度は最低値を付けている。これは E が低い成績のグループに属しているものの、高成績グループの学生と同様の基準で自己判断を行った結果であると考えることができる。

学生 C3 や D2 は低成績グループの典型例である。たとえば D2 の場合、成績に関しては C3 の次に低いにもかかわらず満足度に関しては平均よりかなり高値である。このような学生は、自分自身を客観化する能力が低いか、事実に基づいて客観的に評価しようという意識が弱いことが、その理由であると推察できる。もしかしたら、自分の状況を十分認識した上で、あえて現実以上に良い評価を与えたいという気持ちが働き、実際の自己認識以上の評価をしているのかも知れない。そうすることで、自身への自信を取り戻したいという補正の気持ちが反映されている可能性がある。

以上の考察は数値化された評価のみに基づいている。学生の評価情報に関して更に詳細な解析を行うためには、テキストによる回答部分を数量化するなど更なる追究が必要である。

### 3.3 自由記述アンケートと試験点の関連

学生が授業で学んだ内容を適切な言葉で記述できるか否かは、授業内容の理解度を反映しており、試験結果にも影響があると考えられる。その可否を確認するために、本節では授業評価アンケートの自由記述項目と試験点の関連性を調べる。

本節では、3.1 節で述べた本アンケートの自由記述項目中、設問「(1) あなたはこの授業でどういうことを学びましたか？それはあなたの役に立ちましたか？」への回答結果を解析対象とする。本設問は授業で学んだ内容を総括し、学生自らの言葉で表現することを求めている設問であり、学習成果が良く表れていることが期待できる質問項目である。

アンケートの自由記述項目への回答のような、文章で表現されたテキストデータを解析対象とす

るためには、何らかの方法により質的データである原テキストを数量化[8]し、解析可能なデータを生成する必要がある。高い質のデータを得るために、人手で内容を評価し、その評価データを分析に用いることは良く行われている。しかし、このような方法において、読み手の主観的解釈に起因するバイアスを完全に排除することができない。そのため得られたデータの信頼性や客観性に統一性がないことになりがちである。

本稿では、客観性の高い手法として、形態素解析と計量分析を組み合わせた計量テキスト分析をフリーのソフトウェアである KH Coder [10]を用い

て行う。KH Coder は各文書の区切りを強制的に与えることができる。今回の解析においては学生毎に回答テキストを1文書として処理した。

一方、文章を形態素解析して得られた「語」は文章を構成する要素に過ぎず、人間の解釈に依存しない客観的なデータである。個々の「語」を抽出するだけでは全体の繋がりを表すことはできないが、計量的分析手法を適用することで量的な語の繋がりを再構成することができる[9]。再構成された情報には、読み手に依存しない客観性と信頼性が期待できる。

抽出語	出現回数	抽出語	出現回数	抽出語	出現回数	抽出語	出現回数	抽出語	出現回数	抽出語	出現回数
検索	88	機能	7	サイト	3	膨大	3	深い	2	すべて	1
授業	37	結果	7	課題	3	たくさん	2	身近	2	ほか	1
情報	37	大切	7	解答	3	デスク	2	数	2	アプローチ	1
思う	34	機会	6	改めて	3	一つ	2	数多い	2	エンジン	1
図書館	33	今回	6	外国	3	引き出す	2	整理	2	カウンター	1
学ぶ	32	タグ	5	基本	3	覚える	2	生活	2	キー	1
知る	30	以前	5	見つける	3	学べる	2	昔	2	サーバ	1
自分	21	疑問	5	工夫	3	楽しみ	2	全く	2	ション	1
方法	21	求める	5	行く	3	簡単	2	増える	2	スムーズ	1
今	17	見る	5	参考	3	技術	2	貸出	2	ツール	1
仕方	16	種類	5	司書	3	教える	2	知れる	2	テーマ	1
調べる	16	多く	5	時間	3	驚く	2	著作	2	テレビ	1
キーワード	13	適切	5	受ける	3	苦戦	2	電子	2	データ	1
様々	11	いろいろ	4	集める	3	駆使	2	導入	2	ネット	1
使う	10	違う	4	宿題	3	見つけ出す	2	特に	2	バリー	1
役に立つ	10	演習	4	出る	3	見極める	2	日本	2	プラス	1
必要	9	活用	4	紹介	3	言葉	2	入れる	2	ボックス	1
利用	9	講義	4	詳しい	3	考え	2	比べる	2	メタ	1
インターネット	8	習う	4	世界	3	今まで	2	物事	2	メニュー	1
パソコン	8	身	4	正しい	3	今後	2	保存	2	ルート	1
考える	8	知識	4	他	3	最新	2	毎回	2	レイアウト	1
行う	8	調べ	4	大きい	3	仕組み	2	歴史	2	レポート	1
得る	8	本当に	4	内容	3	事	2	1つ	1	ワード	1
さまざま	7	良い	4	難しい	3	取り組み	2	HP	1	扱い	1
感じる	7	やり方	3	普段	3	初めて	2	いつか	1	以来	1

図6. 頻出150語のリスト

図6に出現頻度の高い順に並べた抽出語を示す。KH Coder では、動詞などの語尾が変化する語については1種類として処理される。グレーで色分けされている語は、当該科目のキーワードに相当する語であり、色付けされていない語はどの科目や文章にでも出現する一般的な語である。

抽出語は単に頻度だけでなく、文中での用いられ方、他の語との関連性の強さ、どの学生が使っているかなども重要な要素となる。抽出された語に対して対応分析を行うことで、抽出語の主成分を得ることができる。

図7は対応分析の結果を2次元の空間布置図で表したものである。出現頻度が3以上の抽出語を示し、アンダーラインは学生各人を表す。ただし学生IDは元々の学生番号ではなく、個人情報保護のためにst0\*\*の形式に変更されている。

表1に図7の学生を互いの距離によって①から⑤までの5つのグループに分け、その試験点についての分散分析結果を示す。学生数は最小で3名、最大16名と、グループ間に3~5倍の開きがある。

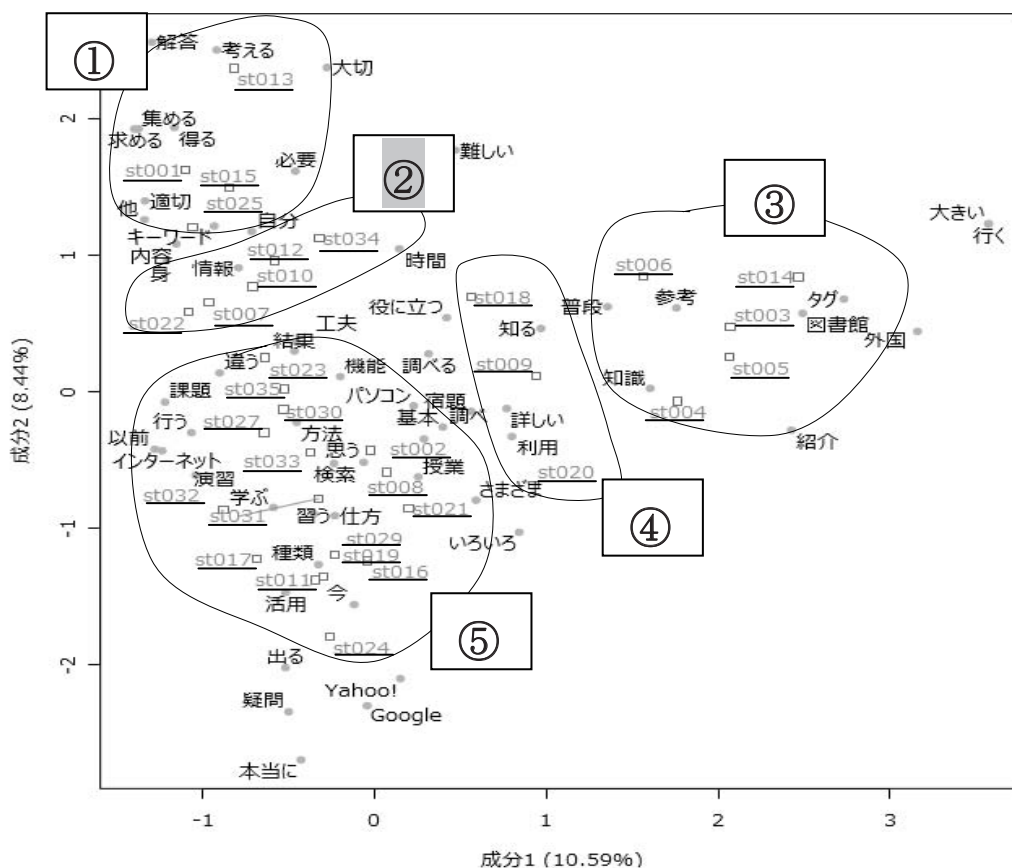


図 7. 学生と関連語の対応分析図

(グループ分けの線と番号ならびにアンダーラインは筆者による。また、アンダーラインは学生を表す。)

表 1. 5 グループの分散分析

グループ	標本数	平均	分散
①	4	65.2	27.3
②	5	70.5	98.8
③	5	83.5	107.2
④	3	69.8	68.7
⑤	16	59.3	335.3

変動要因	変動	自由度	分散
グループ間	2399.34	4	599.84
グループ内	6072.57	28	216.88
合計	8471.91	32	

分散比	P-値	F 境界値
2.76	0.0469	2.71

各グループの平均点は、図 7 において右上に位置するグループ③が 83.5 点と最も高く、

左下のグループ⑤が 59.3 点で最も低い。また、P-値は 0.0469 であり、5 グループの間に試験点の平均に差があることが分かる。

3.2 節で行った数値分類で、いわゆる外れ値にあたる位置にある学生について、①から⑤までのどのグループに属するかを表 2 に示す。標準化数値は各学生の点数を平均が 0 と標準偏差が 1 となるように標準化したものである。負の値はそのクラスにおける平均以下、薄い灰色は下位 30%、濃い灰色は下位 2.5%に位置することを表す。

試験、宿題、出席のいずれかの項目において下位 30%となった学生のうち、その殆どがテキスト分類の⑤グループに属している。一人だけ②グループに属していたが、図 7 に示す通り②グループ自体が⑤グループに非常に近い場所に位置している。よって、数値分類とテキスト分類では大きな差異がないことが分かる。

表2. 学生の数値分類とテキスト分類

数値分類 学生	標準化数値			テキスト 分類
	試験	宿題	出席	
A	2.08	0.79	0.47	③
B	0.97	-3.06	-3.18	⑤
C1	-1.94	0.55	0.96	
C2*	-1.65	0.31	-1.23	
C3	-2.21	-0.05	0.96	
C4	-2.37	0.91	0.23	
D1	-0.70	-2.34	-0.50	②
D2	-0.85	-2.34	0.23	⑤
E	-0.23	-0.05	0.23	①

\*はアンケートなし(5グループに所属しない)

次に、各グループに属する学生がアンケートで使用した特徴的な語と試験の平均点の関連性について調べる。各学生の Jaccard の類似性測度の高い 10 語から、講義内容と関連深い語を抽出する。Jaccard の類似性測度  $J$  は次のように定義される[3]。学生 P の文章中の語の数を  $p$ 、学生 Q の語の数を  $q$ 、両者に共通する語の数を  $r$  とすると、

$$J = \frac{r}{p+q-r} \quad 0 \leq J \leq 1$$

となる。 $J$  の値は、P と Q の学生の文章が完全に一致すれば(すなわち、両者の語が一致すれば) 1 となり、全く異なれば(共通の語がなければ) 0 となる。Jaccard の類似性測度を用いて出現パターンの似通った語と各学生を特徴づける語を関連付けることもできる。

表3に Jaccard の類似度測度を用いた学生毎の特徴語を示す。一般的な情報に関する語を「一般情報●」、講義内容と関連性の深い高頻度の語を「頻出(多)▲」、科目を特徴づける語を「科目■」と3種類に属する語を選別した。学生の番号(st0\*\*)の右側の数値は当該学生の試験点を表す。

表4に各グループの特徴語数と使用比率、そして試験平均を示す。試験点の高いグループ③は、他のグループと比べると特徴語の使用比率

が高い。その他のグループでは、①・②グループは使用比率が低く、④・⑤グループは中間的な値である。

具体的な語を見ると、①グループは特徴語も少なく、個々の学生間の語についても関連性が低い。②グループの学生に共通する語は「課題」、「時間」、「調べる」などであり、課題に関する語が多い。それに比べ、③グループは特徴語のなかでも専門用語の使用が多く、特徴語以外では「外国」、「海外」、「日本」、「全国」など日本と世界を比較するような場所に関する語が多い。また、課題に関する語は現れていない。授業の単位取得のみに囚われるのではなく、広い視野で授業に臨んでいることが窺える。④グループは専門語が多く、頻出語はなかった。⑤グループは一般情報語や頻出語が多く、専門用語は現れていない。共通した語として、「学ぶ」「覚える」「習う」、「役に立つ」、「使う」などの語があり、授業で習ったことを活かすことを考えている学生が多いようである。

#### 4 まとめと今後の課題

本稿では、ある授業に関して学生達の努力度を表す指標として出席点と宿題点を、学習の成果を表す指標として試験点を、そして、その授業を受講したことに対する学生自身の成果(満足度)評価として期末に実施したアンケートへの回答を取り上げ、それらの関係を分析した。

その結果、出席すること自体に意義を見出し、そこから自分に役立つ知識や技能を学ぶことに重きを置いていない学生が多数存在するらしいことが分かった。そのような学生は成績に関する下位グループに多い一方、上位グループには、自分の現状に満足せず、更なる向上を目指す学生が存在することを示唆する結果が得られた。

また、努力度の高い学生が必ずしも成果度が高いとは限らず、出席や宿題が実質を伴わない表面だけの努力に上滑りしている可能性も示唆された。これは授業の担当者として、授業内容を如何に実質化するかという課題に取り組む必要のあることを示している。

表3. Jaccardの類似度測度による学生の特徴語

① グループ

st001	60	st013	69	st015	71	st025	61
考える	.333	手立て	.500	解答	.500	集める	.333
キーワード ▲	.286	重要	.500	得る	.182	テレビ	.200
求める	.250	考察	.500	データ	.167	実践	.200
得る	.200	進める	.500	整理	.167	ネット	.200
入力	.200	物事	.333	予想	.167	発表	.200
素早い	.200	大切	.125	角度	.167	力	.200
四苦八苦	.200	考える	.125	必ず	.167	質	.200
1つ	.200	自分	.050	抽出	.167	術	.200
付ける	.200	検索 ▲	.016	受講	.167	質問	.200
思いつく	.200			選ぶ	.167	必要	.167

② グループ

st007	75	st010	80	st012	71	st022	74	st034	54
選び方	.333	節約	.250	諦める	.250	スムーズ	.250	大切	.231
基本	.200	効果	.250	用語	.250	増減	.250	機能	.231
課題	.200	短時間	.250	言葉	.200	余裕	.250	簡単	.222
他	.200	繋がる	.250	入れる	.200	生じる	.250	難しい	.200
求める	.143	日常	.250	他	.167	収集	.250	必要	.125
適切	.143	上げる	.250	集める	.167	早い	.250	情報 ▲	.118
得る	.111	身近	.200	調べ	.143	行う	.200	キー ●	.111
行う	.100	時間	.167	違う	.143	数多い	.200	押す	.111
役に立つ	.083	身	.143	適切	.125	課題	.167	PrintScreen	.111
様々	.077	インターネット ●	.111	役に立つ	.077	時間	.167	探す	.111

③ グループ

st005	99	st006	76	st003	77	st004	76	st014	90
外国 ●	.333	図書館 ▲	.148	組む	.200	見せる	.167	行く	.250
図書館 ▲	.222	個性	.143	国	.200	限る	.167	海外	.143
最新	.222	概要	.143	ボックス	.200	興味深い	.167	自動	.143
取り組み	.222	就く	.143	HP	.200	写真	.167	特に	.125
世界	.200	関係	.143	閉館 ■	.200	導入	.143	全く	.125
見る	.167	プラス	.143	蔵書 ■	.200	身近	.143	貸出 ■	.125
IC ■	.167	全国	.143	載る	.200	日本	.143	電子 ●	.125
タグ ■	.167	同時に	.143	ルート	.200	著作 ●	.143	大きい	.125
さまざま	.154	レファレンス ■	.143	取り組む	.200	毎回	.143	普段	.111
感じる	.143	又	.143	県	.200	学べる	.143	図書館 ▲	.107

④ グループ

st009	79	st018	66	st020	64
参考	.286	絞り込み	.333	語	.167
詳しい	.286	増やす	.333	辺り	.167
調べ	.250	生活	.250	初め	.167
各種	.167	増える	.250	語源	.167
目指す	.167	技術	.250	履歴	.167
役立てる	.167	普段	.200	違い	.167
メタ ●	.167	司書 ■	.200	出来る	.167
不要	.167	知識	.167	先生	.167
国内外	.167	IC ■	.143	貸出 ■	.143
何気ない	.167	機会	.143	歴史	.143

特徴語

- 一般情報
- ▲ 頻出(多)
- 科目

表 3. Jaccard の類似度測度による学生の特徴語 (つづき)

## ⑤ グループ

st002	82	st008	77	st011	79	st016	57
レイアウト	.167	画面	.500	ウェブサイト ●	.400	新鮮	.250
項目	.167	開く	.500	疑問	.286	ほか	.250
レポート	.167	メニュー	.500	見る	.250	種類	.143
勉強	.167	ワード	.500	豊か	.200	Yahoo! ●	.143
書き方	.167	保存	.500	気	.200	習う	.143
極める	.167	パソコン ●	.125	いつか	.200	今	.111
サーバ ●	.167	役に立つ	.091	少し	.200	Google ●	.111
始まる	.167	仕方	.071	面白い	.200	さまざま	.111
最近	.167			感覚	.200	使う	.083
苦手	.167			問題	.200	キーワード ▲	.063
st017	66	st019	52	st021	76	st023	56
シオン	.250	昔	.200	活用	.333	選出	.250
苦戦	.250	疑問	.167	扱い	.250	内容	.167
バリー	.250	本当に	.167	上達	.250	宿題	.167
探し出す	.250	今	.130	すべて	.250	適切	.125
結構	.250	思う	.114	多い	.250	役に立つ	.077
的確	.250	学ぶ	.111	使いこなす	.250	キーワード ▲	.063
数	.200	応える	.100	機能	.222	方法	.046
比べる	.200	答え	.100	利用	.200	学ぶ	.030
増える	.200	事前	.100	覚える	.200		
膨大	.167	現在	.100	深い	.200		
st024	75	st027	53	st029	55	st030	27
本当に	.250	仕方	.071	種類	.222	結果	.125
Google ●	.200	情報 ▲	.033	いろいろ	.222	行う	.111
エンジン ●	.167	検索 ▲	.016	ツール	.143	知る	.037
範囲	.167			初めて	.125	情報 ▲	.033
広がる	.167			駆使	.125	学ぶ	.032
役立つ	.167			サイト ●	.111	検索 ▲	.016
教える	.143			基本	.111		
良い	.111			Yahoo! ●	.100		
演習	.111			習う	.100		
Yahoo! ●	.111			今	.095		
st031	34	st032	29	st033	81	st035	51
単語	.167	インターネット ●	.429	方法	.200	引き出す	.333
以来	.167	活用	.333	確認	.200	自宅	.167
アプローチ	.167	st032	.250	半期	.200	作業	.167
意欲	.167	先	.250	行う	.182	多々	.167
刺激	.167	使用	.250	事	.167	慣れる	.167
全然	.167	就職	.250	駆使	.167	聞く	.167
異なる	.167	今後	.200	仕組み ●	.167	触れる	.167
頻繁	.167	受ける	.167	覚える	.167	生かす	.167
デスク	.143	演習	.143	教える	.167	漠然と	.167
たくさん	.143	以前	.143	やり方	.143	小学生	.167

表 4. グループ別による特徴語の使用比率

グループ	特徴語数	人数	使用比率	試験平均
①	2	4	0.50	65.2

②	3	5	0.60	70.5
③	12	5	2.40	83.5
④	4	3	1.33	69.8
⑤	18	16	1.13	59.3

学生達の自己評価（満足度評価）を見ると、努力点の中でも宿題点の方が出席点よりも評価に与える影響がより大きい。これはより大きな努力を必要とする宿題の方が、学生の総合評価である自己満足度に与える影響が大きいということであり、十分に納得できる結果である。

興味深いのは、成績と満足度の間には負の相関があることである。1つの解釈は、成績の良い学生達は概して自己向上心が高く、また学習意欲も大きいため、自分の現状レベルに満足せず、もっと高くありたいという欲求が、結果的に低めの満足度評価を与えるというものである。これは十分納得できる説明である。

更に学生が回答したアンケートの文章に対して形態素解析を行い、出現頻度の高い語を抽出し、語の関連度を Jaccard 類似性測度で表現した。言葉や文はその人の考え方を表すものであるため、距離が短い抽出語を使って文章を書いた学生同士は考え方が近いという仮説のもとに事例のデータを 5 つのグループに分け、分散分析を行ったところ試験点の平均に差があることが分かり、数量的に区分できることが判明した。また、各グループ内の学生に対して Jaccard の類似性測度が大きな抽出語の共通性を調べたところ、グループを構成する学生達の考え方が言葉として読み取ることができ、グループ全体を言葉で特徴づけることができた。

今後の研究方向として次のような課題がある。

- (1) 更に新アイデアを考案し、学生の勉学への熱心さ、努力、学びへの態度などを精密化する手法を開発すること、特に、テキスト部分の解析を更に進めること、
- (2) 本稿と異なる授業データを収集し、それらを解析することにより、本稿の研究で示唆された結果を検証したり、比較したりすること、
- (3) 解析手法を一般化、自動化し、授業データ解析システムとして統合化すること。

## 謝辞

本研究の一部は科学研究費補助金（基盤研究(C), 24500318）の助成を受けて実施された。

## 参考文献

- [1] ACM SIGKDD: ACM Special Interest Group on Knowledge Discovery and Data Mining. <http://www.sigkdd.org/>
- [2] R. Agrawal, T. Imielinski, and S. Swami: Mining Association Rules between Sets of Items in Large Databases, Proc. 1993 ACM SIGMOD International Conference on Management of Data, pp.207-216. (1993)
- [3] 土居秀幸, 岡村寛: 生物群集解析のための類似度とその応用, 日本生態学会誌, 61, pp. 3-20. (2011)
- [4] Educational Data Mining Society: <http://www.educationaldatamining.org/>
- [5] K. Goda and T. Mine: PCN: Qualifying Learning Activity for Assessment Based on Time-Series Comments, Proc. 3rd International Conference on Computer Supported Education (CSEDU 2011), 6pp. (2011)
- [6] K. Goda and T. Mine: Analysis of Students' Learning Activities through Quantifying Time-Series Comments, Proc. 15th Annual KES Conference (KES'2011), Part II, Lecture Note in Artificial Intelligence (LNAI 6882), pp.154-164. (2011)
- [7] 合田 和正, 峯 恒憲: PCN法による自己評価コメントの分析からの改善可能な学生の発見, 教育システム情報学会 (JSiSE) 特集論文研究会研究報告, Vol.26, No. 7, pp.51-58. (2012)
- [8] 樋口耕一: テキスト型データの計量的分析 : 2つのアプローチの峻別と統合, 理論と方法, 19(1), pp.101-115. (2004)
- [9] 樋口耕一: 内容分析から計量テキスト分析へ—継承と発展をめざして, 大阪大学大学院人間科学研究科紀要, 32, pp.1-27. (2006)
- [10] 樋口耕一: KH Coder 2.x リファレンス・マニュアル. (2012)  
<http://khc.sourceforge.net/>
- [11] T. Minami and E. Kim: Seat Usage Data Analysis and its Application for Library

- Marketing. Third International Conference on Intelligent Information and Database Systems (ACIIDS 2011), LNAI 6591, pp.238-247. (2011)
- [12] T. Minami: Potentials of Circulation Data Analysis for Library Marketing --- A Case Study in a University Library ---. The 2011 International Conference on Database Theory and Application (DTA 2011), Springer CCIS 258, pp.90-99. (2011)
- [13] T. Minami: Book Profiling from Circulation Records for Library Marketing --- Beginning from Manual Analysis toward Systematization ---. International Conference on Applied and Theoretical Information Systems Research (ATISR 2012) (2012)
- [14] T. Minami: Expertise Level Estimation of Library Books by Patron-Book Heterogeneous Information Network Analysis -- Concept and Applications to Library's Learning Assistant Service --. The 8th International Symposium on Frontiers of Information Systems and Network Applications (FINA 2012), DOI 19.1109/WAINA.2012.184, pp.357-362. (2012)
- [15] 南 俊朗, 大浦 洋子: 学生の成長を助ける学習支援への模索 -授業データ解析による支援方法発見への試み-, 九州情報大学研究論集, 第14巻, pp.39-50. (2012)
- [16] T. Minami and Y. Ohura: Toward Learning Support for Decision Making --- Utilization of Library and Lecture Data ---. The 4th KES International Conference on Intelligent Decision Technologies (KES-IDT' 2012), Springer Smart Innovation, Systems and Technologies 16, pp.137-147. (2012)
- [17] T. Minami and K. Baba: Investigation of Interest Range and Earnestness of Library Patrons from Circulation Records, Proc. International Conference on e-Services and Knowledge Management in IIAI-AAI 2012, IEEE CPS, pp.25-29. (2012)
- [18] T. Minami and Y. Ohura: An Attempt on Effort-Achievement Analysis of Lecture Data for Effective Teaching, Database Theory and Application (DTA 2012), T.-h. Kim et al. (Eds.): EL/DTA/UNESST 2012, CCIS 352, Springer, pp.50-57. (2012)
- [19] T. Minami and Y. Ohura: Towards Development of Lecture Data Analysis Method and its Application to Improvement of Teaching, 2<sup>nd</sup> International Conference on Applied and Theoretical Information Systems Research (2<sup>nd</sup> ATISR 2012). (2012)
- [20] S-H. Myaeng: Towards a memory: human experience mining and semantic social networks, 3rd IIAI International Conference on e-Services and Knowledge Management (IIAI ESKM 2012), Keynote Speech (2012)
- [21] C. Romero and S. Ventura: Educational data mining: A survey from 1995 to 2005, Expert Systems with Applications 33, pp.135-146. (2007)
- [22] C. Romero, S. Ventura, P.G. Espejo, and C. Hervas: Data mining algorithms to classify students, Proc. 1st International Conference on Educational Data Mining (EDM 2008), pp.8-17. (2008)
- [23] L. Talavera and E. Gaudioso: Mining student data to characterize similar behavior groups in unstructured collaboration spaces, Proc. Workshop on Artificial Intelligence in CSCL, 16th European Conference on Artificial Intelligence (ECAI 2004), pp.17-23. (2004)
- [24] N. Thai-Nghe, L. Drumond, T. Horvath, and L. Schmidt-Thieme: Multi-Relational Factorization Models for Predicting Student Performance, in KDD 2011 Workshop: Knowledge Discovery in Educational Data, as part of 17th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. (2011)