

問題提起 淵田吉男（高等教育総合開発研究センター教授）

今日は、問題提起ということでお話をさせていただきます。まず、内容といたしましては、成績評価の在り方をどうして考えていかなければならないかということについて、これはただいま野沢副学長の方から詳しくご説明があったとおりで、私の方から付け加える点はございませんが、中期目標と関連させてお話をさせていただきます。それから現状の全学教育の成績評価を分析したものを、かなりのデータがありますが、かいつまんでお話をさせていただきたいと思います。そして、そこから見えてくる、現在の成績評価に関する問題点を挙げ、時間が許せば一つの成績評価法、先ほど野沢先生から国際的な観点からも成績評価基準が必要だというお話がありましたが、いわゆるGPAといわれるものをご紹介していきたいと思います。

まず、九州大学の中期目標・中期計画というのが9月にも提出されようとしておりますが、その中に「大学の教育研究等の質の向上に関する目標」という項目がございます。この中の「教育に関する目標」に「成績評価に関する基本方針」というものが設けられております。そこでは、「授業の達成目標およびそれに基づく明確な成績評価基準を定める」という目標が掲げられておまして、それに対して「成績評価に関する具体的方策」として、「成績評価基準に基づく成績評価の実施状況を定期的に点検・評価して厳格な成績評価を行う」ということが記されております。この中期目標・中期計画は、いずれ第三者評価機関によって、その達成度に基づいて評価を受けることとなります。九州大学全体の中期目標・中期計画のほかに、各学部・学府の教育に関する中期目標・計画にも具体的に成績評価に関する記述が明確に掲げられているというのも事実でございます。

全学教育における成績評価の難しさ申しますものは、先ほど野沢先生もおっしゃいましたが、同一科目群・同一科目というものが複数の教官で担当、開講されているということに尽きます。担当教官間での成績評価のばらつきが当然予測される訳ですが、それにつきまして、以下に現状のデータをお示し、説明していきたいと思います。

まず、成績区分についてですが、九州大学では優・良・可・不可という四段階の区分を採用しております。ここでは優の部分さらに二つに分けて100点から90点をAaという評価、89点から80点をAbという評価で分類することにし、Aa・Ab・B・C・Dの5段階の区分といたしました。以下、コア教養科目から情報処理科目までのデータ分析についてご紹介いたします。

これが、コア教養科目「人間と文化」の成績分布のグラフです。グラフ中、青で示したものが90点から100点の秀の評価Aa、藤色のものが80点から89点の優の評価Ab、緑が良の評価B、黄色が可の評価C、赤が不可の評価Dを表しています。この「人間と文化」に関しましては、このようにAの比率の高いものあるいはBの評価の高いものというばらつきが見られます。実際の個々の点数分布を見ますと、割合正常な分布が見られます。

「文芸と人間」では、かなりばらつきのあるのが見て取れます。4クラス中の一つでは18%のものがAaの評価を受けていますが、他の2つクラスではAaといものがまったくありません。また、不合格Dは2クラスでは5%以下に対し、残りの2クラスでは25%と33%というばらつきがあります。試験をおこなわなかったあるクラスでは、68%の優があったのですが、実際の評価は10点ごとで行われています。一方、これは試験のあったクラスの点数分布ですが、試験があっても割と荒く採点が行われています。特に、60点での合格の比率が最も高いというところに注目していただきたいと思います。

これは、「歴史の認識」ですが、一見してわかりますことは、Cいわゆる可の評価が非常に多いこ

とです。全クラスともかなり成績評価が厳しいということがわかります。またこの数値は、優の評価の合計を示していますが、2.4%から 23%とクラス間で大きなばらつきがあります。これは、試験を行ったものの点数分布ですが、やはり 60 点での合格が 45%にも上っています。C評価の比率が高いというのはここに原因があると考えられます。試験を行ったクラスは他にもありますが、だいたい同じような点数分布を示しておりました。一方、これが試験を行わなかったクラスの点数分布ですが、これも 5 点刻みの評価になっています。

「異文化の理解」に関しましては、割とクラス間でまとまりがあるように見受けられます。このクラスだけ試験を実施しています。このクラスの点数分布を見ますと、正規分布に近いと申しますか、そういうデータが出ております。60 点での合格も比較的少ないことがわかります。こちらは試験を実施しなかったクラスの点数分布ですが、5 点刻みの評価が行われています。

次に「現代の政治と法」についてですが、かなりのばらつきがあるということが見て取れます。可の多いクラス、良の多いクラス、優の多いクラスと担当者によって成績評価基準が違うということがここから見えてきます。試験があったクラスでは、やはり 60 点評価がかなり多い。C評価に関して、65 点と 60 点しかないわけですが、この差がどこにあるかという問題が出てきます。

「現代の社会と経済」につきましても、大きなばらつきがあるというのがすぐ見て取れます。あるクラスでは 82%のものに対して優と評価していますが、このクラスでは優はわずか 4%で、可中心の評価であるという違いが出てきています。これはあるクラスの点数分布を示していますが、試験があったにもかかわらず、このように極めて荒い点数刻みの評価がなされています。

ここに文系のコア教養科目の成績評価をすべてまとめました。これは、それぞれの科目の平均の分布を示しております。このように同じ科目群であっても、成績の分布は様々です。特に「歴史の認識」の評価が厳しいというのがわかります。このように同一科目群で成績に大きなばらつきが生じてきています。また、可が多いのは、60 点評価が多いということから出てくる問題かと思えます。

次に理系のコア教養科目を見ていきます。「地球と生命」についてですが、ここに示した数値はそれぞれのクラスの優の合計を表しています。8 割程度とするもから 6 割程度とするもの、中でもこのクラスは 3 割と評価が厳しいというのがわかります。同一科目でこのように大きなばらつきが生じてきています。試験を行っても、やはり 5 点刻みの評価があります。従って、全体の点数分布は、このように極めて単純なものとなります。同じく試験を行った他のクラスでは、5 点刻みの評価はされておらず、結果としてこのように極めて自然な点数分布を示しています。担当教官による評価の仕方も様々であるということです。

「物質の世界」では、わりとまとまりがあるような分布を示しています。3 クラスとも試験は行っていません。点数分布を見ますと、自然な点数分布の中に、60 点が突出しているのが大きな特徴です。60 点の駆け込み合格と申しますか、そのようなものが多いという事実です。60 点の駆け込み評価を他の C 評価と同一と見なすのも検討する必要があります。

全ての理系のコア教養科目の成績分布をまとめました。ここに優の割合を示しておりますが、全体の平均値と比較しますと割合まとまりがあるようですが、中身に関しては見てきましたように様々であるということです。

次に、個別教養科目の少人数ゼミナール B について見ることにします。ここには散布図を示しています。これは例えば、A a の散布図を見てみますと、19 クラスの内 A a の比率が 0% のものが 6 クラスあり、8%位のもので 1 クラス、以下 9% が 1 クラスあるといったことを示しています。A b を見ますと一つのクラスでは優をまったく出していない、逆に 100% が優であるという評価が行われ

ているクラスのあることがわかります。例えば、成績評価を相対評価とし、A aを5%なり10%と設定したとしますと、散布図は当然グラフの横軸に対して平行線になるべきものです。このように傾きが大きいものほど、A a成績のばらつきが大きいということを示しています。また、5つの成績区分の散布図を比較することによって、例えばA aの折れ線グラフが他の成績区分のグラフと比べて上の方にあることは、A aの評価の比率が他の評価区分と比べて高いということを示しています。個別に成績評価を見てみますと、ここでは3例挙げていますが、すべてA bで優を出しているクラス、あるいはBの多いクラス、全体の平均に近いような分布を示すようなものとさまざまな例があります。この中で、A aの比率が67.4%というクラスの点数分布を見ますと、100点が全体の45%も占めているという評価も見られます。また、この例では、C評価44人(29.7%)うち28人、つまりC評価の64%が60点で評価されている例があります。また、ここに示す例のようにかなり細かく評価されている例もあります。全体的には、評価はかなりバラエティに富んでいます。

これは言語文化科目Iの「インテンシブ英語演習」の成績の散布図です。クラスは52クラスあります。やはりかなり差があるというのがこれからも見て取れるかと思えます。特に、このようにA a、A b、B、Cのグラフがかなり接近していますが、これはそれぞれの評価の割合が近いということを示しています。しかし、個別に評価を見てみますと、A aのみのもの、A aが非常に多いものがあります。また、B評価が非常に多いもの、C評価が非常に多いもの、更にはD、不可が多いものとかかなりのバラエティが見られます。

「健康スポーツ科学実習」に関しても、かなりばらつきがあるのもこの散布図から見て取れます。個別に見ますと、このようにA a評価が80%以上のもの、A b評価が90%近いもの、あるいはBとC評価が中心のもの、優評価の他には不可しかないクラスもあります。

「線形代数A」の散布図では、グラフは割とフラットに近い形をしていますが、一部ばらつきがあることを示しています。実際に個別に成績分布を見てみますと、この2クラスでは優が75%以上ですが、それと比べ、良の比率の高いものもあれば、可の比率が66%と高いもの、不可の比率の非常に高いものと様々な実態が見られます。

「電磁気学基礎」に関しましては、全体として一見かなりまとまりがあるように見えますが、個別に見ますと一部に大きなばらつきが見られます。A aの評価が非常に高いものからCである可の評価が非常に高いもの、更には極端な例として、不可の評価が極めて高いものもあります。もちろんこれは対象クラス・対象学生の質、レベルにもよるのかもしれませんが、あまりにも大きな差がありすぎるのではないかと思います。

「基礎科学結合論」は、グラフは割とフラットに近くなっており、ばらつきがないように思えます。しかし、これも実際中身を見てみますと、A a評価が非常に高いものから可の評価が50%以上を占めるもの、不可が非常に多いものと実に様々な成績評価の実態が見られます。この科目について、各クラスの点数分布について詳しく見ますと、このようにやはり60点評価がかなり突出している状況が見られます。この科目については、昨年度「学生による授業評価」を実施しました。試験時と授業終了時の理解度について見てみますと、「理解できた」とするものは、授業終了時では9%、試験終了時では33%でした。理解度に関して、このように低い数値が上がっているにもかかわらず、その成績は、100点が58%でそれは全体の優評価の59%を占めています。このような理解度と成績評価のギャップは、何を意味しているのでしょうか。またこれまで見てきましたように、60点評価が突出しています。これは到達度を再調査されるなど、担当された先生がかなりご苦労しておられるとは思いますが、このクラスでは二人を残してすべて60点で合格させるという、このような事実も見

えてくるわけです。

「地球科学概論」を見ますと、一見するだけで、かなり厳しい評価だなということがわかります。可の評価がどのクラスも多い。この場合にもやはり、点数分布を見ますと 60 点評価が 25 名、全体の 25%にも達します。これは、C 評価の中の 41%を占めています。このような評価のあり方についても、もう少し真剣に考えなくてはいけないのではないかと思います。

実験科目について、「基礎生物化学実験」を見てみましょう。対象クラスの問題はあるかと思いますが、この場合にもやはり A a と A b 評価が 79%と 21%とすべて優の評価のものから、A 評価が 10%程度で C 評価が 32%というかなり厳しい評価まで見られます。

「図学」に関しましては、見事にばらついているとしか言いようがない結果です。可の非常に多いもの、優、特に A a 評価が非常に高いもの、A a から C 評価まで均等な分布を示すものなど様々です。

「情報処理基礎演習」につきましても、散布図ではグラフの勾配が大きく、ばらつきがあるということを示しています。これも、中身を見てみますと、優の比率の非常に高いもの、その中でも A a、A b の評価が非常に高いものから、B の評価あるいは可の評価が 60%にも達するようなものまでかなりのばらつきがあります。

ここまで成績評価の現状分析をご紹介してきましたが、成績評価に関する問題点は、大きなばらつきがあることに尽きます。同一科目群、例えばコア教養科目あるいは個別教養科目を学生はある区分に従って選択していくわけですが、そういった同一科目群の間にばらつきがあるのと同時に、科目群を構成する各科目の中にも大きなばらつきが見られたました。また、同一科目、例えば「インテンシブ英語演習」、「健康スポーツ科学実習」、これらは同じ到達目標をもった内容の科目であるはずですが、その評価にあまりにも大きなばらつきがありました。この成績のばらつきが、対象クラスの学生の質、レベルから生じるという問題があるのかもしれませんが、それにしてもレベル差を考慮してもそれ以上の大きなばらつきがあるのではないかと考えられます。

この問題を解決するためのキーワードとして、「相対評価」、「絶対評価」というものがでてくるのかもしれませんが。また、もう一つの問題点というのは、60 点合格者というのがかなり目立ったという点です。つまり、再調査をして合格にする場合、成績を 60 点とする。あるいは定期試験であまりにも不合格者が多い場合、再調査はせず、腹をくくって 60 点として合格させてしまうという実態が裏にはあるのではないかと気がします。従って、60 点合格者の取り扱いも、どのようにするか検討する必要があると考えられます。成績を 5 段階区分で見てきましたが、秀・優・良・可・不可という成績区分も考える必要があるのではないかと考えます。上智大学などでは全学的に新しく 5 段階評価を取り入れています。また、九州大学では学生が履修を途中で放棄しても、教官が提出する成績報告書に斜線が引かれれば、履修放棄の事実が学生の成績には一切出てきません。そのような履修実態を、学生が放棄したということで、それを withdraw, 「W」無判定という評価をする方法も見られます。これは、履修制限につながる一つの方法になりうると思われます。また、他大学を卒業あるいは中途退学して入学してきた学生に対して、履修教科の単位認定を行っていますが、それに対しても「無判定」、つまり合格という判定だけをしていく方法が採れます。こういった新しい評価法は、GPA にも見られます。

全学教育では、「学習到達度再調査」という制度があります。定期試験の成績が合格基準に到達しない学生に対して、もう一度試験を実施するというものです。これは強制ではなく、担当教官の自由意思に任されています。通常はこの再調査期間として、定期試験終了後の一週間が設けられ、同じ曜日時限に実施していただくようにしています。また、もう一つの再調査の方法が認められています。

それは、全学教育科目担当教官要項に記されていますが、「成績提出最終締切り後の成績訂正」です。これは、十分な学習指導を行ったうえで、不合格になった成績を翌学期初めの一定の期日までに訂正できるものです。正確に申しますと、翌学期が始まって2週間後、つまりその学期の履修手続きが確定する前であれば訂正が許可されます。対象科目には制限がありますが、60点として単位を認めることができます。

「学習到達度再調査」の実態について触れさせていただきます。データとしては不十分なものになっています。それは、この調査が担当教官の自由意思ということで、教務掛に必ずしも実施の報告があるとは限らないからです。試験に使う教室の設定を教務掛に依頼した場合、あるいは再調査の対象者を学生に周知させるために公用掲示板への掲示を教務掛に依頼した場合のみしか把握できないということです。全学教育の全クラス数、13年度には926、14年度には921ありましたが、そのうちの約半分クラスで定期試験期間中、試験が実施されています。しかし、再調査が実施されているのは、教務掛に届けられた分だけで、80クラス、定期試験をしたクラスのたった2割弱しかないという実態があります。ただ、先生によっては、一回の再調査だけではなくて何度もされるという、そういった熱心な先生もおられるというのも事実です。先ほど述べました、「成績提出最終締切り後の成績訂正」に関しては、年間にせいぜい40名程度しかないのが実態です。

ここには、成績評価に関する問題点をまとめました。まず、学生から見た問題点です。「科目内容をどれくらい理解すればよい成績に結びつくのかわからない」、「同一科目でも担当教官により授業内容の難易度が異なる」、「同一科目でも担当教官により授業内容が異なる」という訴えが「学生による授業評価」の自由記述欄などであります。これは、結局到達目標がきちんと明確にされていないということに起因するものです。また、「1回の試験だけで評価されるのは納得がいかない」、「出席率、レポート、発言等、授業への積極的な取組も評価してほしい」との訴えもあります。出席しようがしまいが関係なく、最後の試験だけで裁かれてしまう。自分はずっと出席して聞いていたのに、試験の時、たまたま体調をこわし、十分に勉強できなかったので落ちてしまった。まったく出席しなかった学生が、なんだか知らないけれど通っている。特に少人数ゼミナール等だと思いますが、授業への積極的な取組みも評価してほしいという訴えがあります。この二項目は、成績評価の厳格性が問われているものと考えられます。また、「自分の成績がクラスでどの位置にあるのかよくわからない」、「同じ優評価でも科目によって重みが異なる」というものもあります。これはデータ分析で見えてきたとおりですが、同じ科目でもいわゆる楽勝教官とか鬼教官がいて、評価の重みが全然違うという訴えです。この担当教官による評価の重みの違いが成績に反映されることがなく、社会にも通用しないということです。つまり、これは絶対評価とか相対評価に関連する問題です。また、到達度再調査の申し合わせ事項に「再調査による点数は60点」という記載がないことに問題があるのですが、学生から再調査で優あるいは良がもらえるのはおかしいという訴えがあります。再調査とならなかったが成績が可であった学生が、再調査を受け合格したものに可以上の成績がだされることに非常に不満を持っているケースがあります。2年次後期以降の学科振り分け、農学部や工学部が該当するわけですが、その振り分けが、全学教育の成績で決定されています。全学教育に対し農学部などの学生は一生懸命取り組むわけですが、先生によってその評価がかなり違って来る、つまり成績評価の基準がはっきりしていないため不利益を被っているとの不満がでているのも間違いのない事実です。

また、教育効果の面から見た問題点といたしまして、成績評価が学生の主体的な学習への取り組みを促しているかというものがあります。つまり1回の試験で成績が決められていいのかという問題です。学習への主体的な取り組みを促しているのか、つまり、普段から何らかのレポートを課すなりし

てその科目に対する学習意欲を向上させている努力を教官がしているのかという問題が一つはありますし、成績評価がその後の科目履修に対する動機付けに影響を与えているのかという問題もあります。これは厳正な成績評価が行われれば、そして適正に成績評価がなされれば、自分はこの科目に対して優を取ったんだという自信から、もっとこの周辺のことについて勉強してみたいという学習のモチベーションを高める効果、あるいは一生懸命やったのだけど可でしか合格できなかったから、やっぱり自分はここが弱いので、もう少しこの部分を勉強し直してみようとかいう、そういう動機付けにはやはり現状ではなかなかないということがあげられるかと思えます。

それからもう一つは、成績評価が優秀者の勉学意欲をさらに向上させる手立てとなっているか。これは、例えば、報奨制度はあるのかということです。「飛び級」に関しては、現在は成績優秀者に対しては行われているところもありますが、成績優秀者であれば交換留学等留学に対して何らかの特典が与えられるとか、あるいは授業料免除、奨学金取得についても加点されるとか、そういった取組がきちとなされているかということです。また、野沢先生もおっしゃいましたけれども、やはり社会に貢献できる人材を養成する義務を持つ大学として、成績評価が卒業生の質の確保に繋がっているのかという点が今大きく問われているところだと思えます。それで、全学教育に関しては、平成13年度に全学教育企画委員会で九州大学の全学教育科目における成績評価の実施指針というものが制定され、14年度以降全学教育科目担当教官要項に教育憲章と共にこれが掲載されております。読んでいきますと、「成績評価は授業科目の目的と目標に照らし、その達成度について行われる。従って、それぞれの授業科目について、その目的と目標を明確にすることが重要である」、「成績評価は、授業科目ごとの目的・目標の実現に資するとともに、円滑且つ効果的に実施されるものでなくてはならない」、「効果的な教育を実現するために、厳格的かつ説明可能な評価を実施する」、「各授業科目の成績評価は、授業担当教官の責任において行う」、「授業担当教官は履修学生に合格基準を明確に示したうえで、それに基づいて成績評価を実施する」とあります。成績評価に関しましては、全学教育のウェブシラバスで書いていただくようにしていますが、例えば成績評価は、レポート30%、出席率20%、試験が50%のように書いていただくようにしていますが、その書き込みがきわめて少ないのが実態です。それから更に「授業担当教官は不合格者が過度に多くならないように、学生の学習意欲の促進や学習指導に努める」、「同一授業科目の成績評価にあたり、授業担当教官によって極端なばらつきが生じないよう成績評価の基準を設ける」と結んでいます。このように非常に重要な実施指針があるのですが、なかなか授業担当者間での話し合いができていないというのが実情です。

まとめますと、一番大きな問題点は、「評価されるものには決定的な影響を及ぼすが、評価するものには何の影響も及ぼさない」ということです。つまり、評価される学生の方は成績には非常に影響されますが、評価する教官は成績を出してしまっただけで、何のお咎めもないということです。成績評価は、自分の信念に基づいているのだということとされてきていたことから、何の問題もあがってこなかったのが今までの状況でした。今後は、やはり成績評価をきちんと説明できるか、説明責任を持てるかどうかということが問われてくることになり、この点が大きな問題だと思えます。

これから、分科会で討論していただくわけですが、座長の先生にはご無理を申し上げて事前にレポートを作成していただきました。午前中はそれを叩き台として自由に討論していただきまして、午後には引き続き討論していただいた後、全体会の報告に向けて討論のまとめを作成していただくことにしたいと思います。全体会は80分ありますが、ご案内では分科会報告5分と記していますが、なるべく討論を活発にさせていただきたいと思っておりますので、報告は4分程度でお願いしたいと思います。討議の後、時間が余れば、全体討議を行っていく予定です。

それと、成績評価のあり方については、全科目共通する部分と科目特有のあり方があるのではないかと思います。今回は科目ごとに分科会を設定し、その中で討論していただくことにいたしました。今後、この結果を踏まえて、成績評価のあり方について、さらに深い討論ができるよう、その方向性が見出せればよいと考えています。

最後に、GPAについて紹介致します。GPAというのは、Grade Point Average というものの略で、アメリカ式の平均成績値、評定平均値、加重平均点と言われるものです。留学時に提出を求められるあるいは奨学金支給の参考指標、進級要件・退学勧告の基礎指標になっています。日本でも次第に、国立大学では東北大学、大分大学等で導入され始めておりますが、これは主に工学部で、JABEEと関連して導入されてきたものです。全学的に合意され、導入された例はまだございません。この表について説明致します。科目として物理科目などの科目とその単位数を記しています。その成績評価はAである。A、B、Cに対して加重点 3.75、2.5、1.25 が設定されています。加重点を何点とするかは大学によって違って来るわけで、そういった意味でまだ統一性がないというのも事実です。日本全体として統一性がないということです。A評価に対する加重点 3.75 に単位数2を掛け、Grade Point は 7.5 となります。このようにして、それぞれの履修科目に Grade Point を付けていきます。これはいわゆる認定科目です。合格ということだけで、この場合には Grade Point を算出しません。Grade Point を合計すると 42.5 となります。それに対して、総単位数は 14 です。Grade Point の合計を総単位数で割った 3.0 が Grade Point Average となります。GPAは、名古屋大学では留学に利用しています。徳島大学では工学部が実施、上智大学、国際基督教大学では全学的に行われておりますが、加重点、GPAの算出方法も様々な状況です。

以上で、問題提起を終わらせていただきます。