

初年次理系実験の成績評価分析

—入試制度と取組姿勢の観点から—[†]

田島貴裕^{*1}北海道大学高等教育推進機構^{*1}

成績評価が学習への取組姿勢や入試制度とどのような関係にあるのかを検証することを目的に、大学初年次教育における理系実験を事例として、これまでに蓄積された約1万名の学習履歴データの分析を行った。その結果、毎年、履修者の一定数は不可になっており、入学時の入試制度によって、その割合は異なることが明らかとなった。また、実験に対する取組姿勢および成績評価においても、入試制度によりその割合に違いがみられた。取組姿勢と成績評価の関連性については、実験に対する取組姿勢が良いほど、概ね成績評価は高くなる傾向がみられた。さらに、過去の学習履歴データを分析することによって、最初の数回の実験レポート点と、最終成績評価には関連性があることが示され、履修の初期段階で最終成績を予測できる可能性が示唆された。

キーワード：教育評価、入試制度、初年次教育、高等教育

1. はじめに

教育方法の改善や厳格な成績評価を実施するため、「学生の学修履歴等の記録と自己管理のためのシステムを開発することは、学習成果を重視した評価の条件整備として重要」である（中央教育審議会 2008）。そのため、ポートフォリオや履修者による授業評価、アンケート調査など、様々な学習履歴や学習成果を把握するためのデータ収集が試みられてきた。近年では、教育分野における ICT の利活用が進み、このような学習履歴や学習成果に関するデータ収集は容易になっており、各教育現場で盛んにおこなわれていると考えられる。そして、教育現場で収集される膨大なデータ、いわゆる「教育ビッグデータ」に関する研究も盛んにおこなわれており、現在では単にデータを収集する段

階から、いかに意味のある情報を生み出すかに研究の軸がシフトしてきている（寺澤 2015）。教育ビッグデータは、その目的によって大学単位、学部単位、あるいは教員単位により、さまざまな項目のデータが収集されているが、「意味のある情報」を取り出す分析方法や活用方法も多岐にわたると考えられる。

学習支援に資する情報を取り出す分析として、例えば竹生・辻（2016）では、e ラーニングの学習履歴データを用いて、学習行動の類型化と成績の関連性を検証している。履修者の属性の違いや学習行動と成績の分析を行うことで、学習支援の方針が明確になり、成績の改善につながることを示唆している。また、杉山ほか（2016）は、e ラーニング受講者への適切な学習支援を行うため、学習履歴データから最終学習状態を推定し、ドロップアウトの可能性が高い学習者の早期発見を試みている。このような学習履歴データの活用は、講義終了後に実施する履修者アンケートよりも日々の学習者の行動を把握しやすく、履修期間中の早い段階で成績分析や成績評価の推定が可能となり、履修者に対する学習支援の見直しや検討が早期に可能になると考えられる。

そこで、本報告においても履修期間の初期段階で学習支援の検討や改善を行うための基礎資料として、こ

2017年1月30日受理

[†] Takahiro TAJIMA^{*1} : Analysis of the Academic Performance in “Laboratory Exercise in Natural Sciences” : from the Standpoint of a University Admission System and a Learning Attitude

^{*1} Institute for the Advancement of Higher Education, Hokkaido University, Kita 17, Nishi 8, Kita-ku, Sapporo, Hokkaido, 060-0817 Japan

れまでに蓄積された約1万名の実験科目の学習履歴データを解析し、実験に対する取組姿勢と成績評価の関連性や成績評価予測の可能性について検証を行う。ここで取り上げる対象事例は実験科目であり、学習履歴データはeラーニング等で収集されたデータではなく、実験室に設置してある出欠確認システムや実験スタッフによる集計作業により蓄積されたものである。大学初年次教育のなかで実施される理系必修科目であり、年間履修者は約2,000名と多く、教員・TAも例年200名ほどが携わっている。山田ほか(2010)は、履修者が500名を超えるような大規模eラーニング講義において、担当講師が少ない場合でも、学習支援の質を保ちつつ、効率的で効果的な学習支援を実施する必要性を課題として挙げている。本稿で対象とする事例のような多数の教員が携わる大規模な実験科目においても、効果的な学習指導や支援は検討課題である。

一般に、教員が毎週同じ履修者を担当する場合、履修者の毎回の取組姿勢と履修状況を確認し、さらに教員の経験や心証から履修者の最終の成績評価を推測することで、それに応じた学習支援が可能である。しかし、対象事例では教員、履修者ともに非常に多く、同時に複数の実験テーマを多数の教員により実施しているため、各教員は毎回異なる履修者を受け持っている。つまり、各教員は履修者の毎回の実験態度や取組姿勢、履修状況を継続的に観察することが難しく、「経験則」で履修者の最終の成績評価を推測することが難しい。それゆえ、学習履歴データの解析を行い、学習支援の検討が可能な履修者情報を抽出し、教員間で履修者情報を共有する必要性が生じている。

加えて、本報告では、入試制度の観点から成績評価の分析も試みる。平成20年に中央教育審議会から答申された「学士課程教育の構築に向けて」では、学士課程教育の構築は日本の将来における喫緊の課題と指摘している。同答申は、現行の大学入試制度に関して、「自己決定力の未熟な学生も目立つ中、入学してから時間のゆとりを持って専門分野を選択できる、あるいは柔軟に変更できる仕組みづくり」を指摘しており、入試制度の改善策として、募集単位を「大きく」にすることを提言している(中央教育審議会 2008)。本稿の対象事例も約半数の入学者が「大きく入試」で入学しており、より詳細な履修者像の把握のため、入試制度の違いが取組姿勢や成績評価に対してどのような差異を生み出すのかを検討する。ただし、本稿における入試制度の観点からの分析は、履修者像をより詳

細に把握するためのものであり、入試制度自体の考察や入試制度の優劣を議論するものではない。

したがって、本研究の目的は、学習履歴データを用いて入試制度および取組姿勢の観点から成績評価分析を実施し、履修者像の把握と成績評価予測の検証を行うことで、大規模な実験科目における学習支援へ向けた基礎資料として提示することである。

2. 対象事例

2.1. 分析対象

北海道大学の初年次教育で開講されている「自然科学実験」の履修データ(2011年から2015年の5年間)を分析対象とする。自然科学実験は2006年から開講されている実験科目であり、理系1年次の必修科目である。科目等履修生等をすべて含めた年間履修者数は、約2,000名である。学生は、物理系、化学系、生物系、地球惑星科学系(各6テーマ)から、希望する2カテゴリ(各6テーマ、計12テーマ)を履修する。自然科学実験では、講義により得た基礎知識を確実なものとし、実際に手を動かして自然科学にじかにふれ、考えることにより、理科の素養の実質的向上を目指している(北海道大学自然科学実験編集委員会 2016)。また、物理、化学、生物、地球惑星科学の4カテゴリに共通する科学的な考え、自然に対する知見を養うことのほか、一般的な器具・機械の取り扱い方、測定データの見方、実験ノートの取り方、データの処理方法、レポートの体裁・書き方といった基礎的な実験技術の習得も目的としている。

成績評価は、講義および安全教育を除く12回の実験テーマについて点数化し、相対評価で行われている。全実験テーマの合計点数120点満点中、72点以上で単位取得が可能であり、72点未満は不合格(不可)となる。実験1回分の点数は、10点満点(0点から10点の11段階評価)であり、「実験評価点(5点満点)」と「レポート点(5点満点)」からなる。実験評価点は、全12回の実験テーマにおいて、欠席回数、遅刻回数、レポートの未提出回数、レポート提出の遅滞回数、忘れ物、授業態度など、実験に対する取組姿勢により減点される。レポート点は、各回の実験後に提出する実験レポートの内容について、相対評価で採点される。なお、最終の成績評価は、全12回の実験テーマにおける実験評価点とレポート点を合計し、合計点数の上位から「A 45%(うちA+は5%)、B 50%、C 5%」という割合で評価することとなっている¹⁾。

2.2. 分析項目

入試制度や取組姿勢が成績評価とどのような関係にあるのかを検証するため、以下の項目について比較分析を行う。

分析 (1) 履修者の全体像

- (a) 履修形態 (初修, 再履修)
- (b) 入試制度 (総合入試, 学部別入試)

分析 (2) 実験に対する取組姿勢

- (a) 欠席回数
- (b) 遅刻回数
- (c) レポート未提出回数
- (d) レポート遅滞回数

分析 (3) 取組姿勢と成績評価

分析(1)では、まず履修形態と入試制度の観点から、履修者の全体像を把握する。履修形態は、入学後に割り当てられた授業時間割においてはじめて実験を履修する「初修者」と、それ以外の「再履修者」の2つの履修形態で区分する。入試制度は、入学時から学部が決まっている「学部別入試」と、2年次進級時に学部を決定する「総合入試」により入学した学生の2つの入試制度で区分する。北海道大学では、入学前後における学部学科選択のミスマッチの解消や初年次教育の充実を目的として、2011年より総合入試を導入している(北海道大学 2015)。総合入試は、入学後の1年間は幅広く教養科目や基礎科目を学び、2年次への進級時に、本人の希望と1年次の成績によって学部へ移行するシステムである。また、総合入試には文系と理系の入試枠があり、入学後に文系から理系、理系から文系への移行も可能となっている。

分析(2)では、実験に対する取組姿勢として、(a)欠席回数、(b)遅刻回数、(c)レポート未提出回数、(d)レポート遅滞回数の4つの学習状況を用い、入試制度の観点からそれらの比較分析を行う。入試制度は

分析(1)と同様、学部別入試と総合入試の2区分である。学習状況は、講義および安全教育を除く12回分の実験についてデータを記録している。各学習状況は、初回の講義時にあらかじめ学生に対して実験評価点の対象となることを伝えており、実験レポートを提出するボックス前で全員の学習状況を常に掲示し、周知させている。

分析(3)では、入試制度および取組姿勢と成績評価との関連を検証するために、実験レポートの平均レポート点(全12回分のうち、提出されたレポートの点数の平均値)を用いて分析を行う。入試制度は分析(1)と同様、学部別入試と総合入試の2区分、取組姿勢は、分析(2)と同様、(a)から(d)の4つの学習状況を用いる。実験レポートは、一定の履修人数(実験グループ内)で相対的に採点し、点数割合も決められている。また、履修者数の関係上、毎回の実験時に同じ教員が同じ実験グループ(学生)を担当することはほとんどなく、概ね毎回異なる多数の教員が担当している。そのため、実験レポートも毎回異なる多数の教員が採点しており、レポート点は客観性の高い成績指標であると考えられる。

3. 分析と考察

3.1. 分析(1) 履修者の全体像

分析(1)に関して、履修者の全体像を把握するために、履修形態別の履修者数および単位取得状況を表1へ示す。2011年から2015年の5年間の全履修者数は10,204名であり、うち初修者は9,573名(93.8%)、再履修者は631名(6.2%)である。全履修者のうち、単位を取得できなかった履修者(以下、単位を取得できなかった履修者は単に「不可」、単位を取得した履修者は「可」と略)は、5.9%である。履修形態別の単位取得状況は、初修者は3.4%、再履修者は44.8%の履修者が不可となっており、再履修者が再び不可になる割合は極めて多い($\chi^2(1)=1830.4, p<.001$)。なお、各年にお

表1 履修形態別の実験履修者数および「不可」の数

履修形態	全体(2011年~2015年)			全体のうち不可のみ			不可割合 (%)
	人数	平均	(SD)	人数	平均	(SD)	
全履修者	10,204	2,040.8	(10.6)	604	120.8	(16.9)	5.9
初修者	9,573	1,914.6	(8.9)	321	64.2	(16.4)	3.4
再履修者	631	126.2	(6.6)	283	56.6	(7.8)	44.8

表2 「不可」の割合における総合入試と学部別入試の同等性の検定結果

履修形態	入試制度	可		不可			合計		χ^2	df	p	
		人数	(%)	人数	(%)		人数	(%)				
全履修者	総合	5,261	(95.1)	△	273	(4.9)	▼	5,534	(100.0)	21.1	1	**
	学部別	4,339	(92.9)	▼	331	(7.1)	△	4,670	(100.0)			
初修者	総合	5,150	(97.2)	△	148	(2.8)	▼	5,298	(100.0)	11.5	1	*
	学部別	4,102	(96.0)	▼	173	(4.0)	△	4,275	(100.0)			
再履修者	総合	111	(47.0)	▼	125	(53.0)	△	236	(100.0)	10.0	1	*
	学部別	237	(60.0)	△	158	(40.0)	▼	395	(100.0)			

注) △および▼は残差分析の結果を示す。△は有意に多い、▼は有意に少ない ($p < .01$) * $p < .01$, ** $p < .001$

ける再履修者数の割合について違いがあるかを分析した結果、有意差は見られなかった ($\chi^2(4)=1.64, n.s.$)。つまり、毎年、一定数の再履修者が存在していることになる。

つぎに、表1で示した履修形態別の単位取得割合について、より詳細な履修者の全体像を把握するため、入試制度により単位取得割合に違いがみられるかを検証する。つまり、学部別入試と総合入試の2つの入試制度において、「可」と「不可」の比が同じかどうかを確認するため、同等性の検定を行う。分析結果を表2へ示す。全履修者および初修者に占める可の割合は、学部別入試よりも総合入試の方が有意に高い。総合入試の学生は、2年次に学部・学科移行する際、移行点に自然科学実験の成績評価が算入されるが、初修時の成績評価のみが対象であり、再履修の場合は除外される。そのため、初修者の不可になる割合は、学部別入試よりも総合入試の方が低いと考えられる。一方で、再履修者の不可割合は、総合入試よりも学部別入試の方が低い。学部別入試の学生は、学部がすでに決まっていることから成績評価を気にしない履修者が多く、次節に示すように欠席やレポートの未提出等も多く見られる。学部別入試の履修者の中には、「優秀な」学生でも、初修時には安易に不可になる履修者（または不可になっても構わないという履修者）が多く含まれていると考えられる。

履修者の全体像を分析した結果、毎年、一定数の不可が存在することと、履修形態および入試制度により単位取得状況が異なることが示された。以降の分析では、初修者の可の履修者を対象に、取組姿勢と成績評価の検証を行う。

3.2. 分析(2) 実験に対する取組姿勢

自然科学実験では実験に対する取組姿勢の指標として、4つの学習状況に関するデータを収集している。入試制度の違いにより単位取得割合に差があることから、成績評価に加味されるこれらの4つの学習状況にも、入試制度による違いが当然考えられる。そこで総合入試と学部別入試の2つの入試制度において、各学習状況の該当回数構成比に差があるかどうかを確認するため、同等性の検定を行う。分析結果を表3へ示す。

学習状況(a)の「欠席回数」をみると、総合入試、学部別入試ともに「0回」「1回」の合計比率が非常に高い。これは、表3のデータは単位取得者の数値であり、成績評価の基準上、3回以上の欠席は単位取得の合格点(72点)を越えることが困難となっているためである。部活動や自己都合で実験を休んだ場合は「欠席」となり、実験評価点の減点対象となるが、病欠等でやむを得ず欠席した場合は欠席回数には含まれない。つまり、欠席回数が1回以上の履修者は、そのほとんどが病欠等ではなく、部活動や自己都合で欠席していると推測される。欠席回数の割合について、入試制度別で比較した結果、0.1%水準で有意差がみられた。また、残差分析の結果、欠席回数は、「0回」「1回」で有意であった ($p < .01$)。総合入試よりも学部別入試の方が、部活動等で安易に欠席する割合が大きいと考えられる。

(b)の「遅刻回数」では、遅刻をする履修者は総合入試、学部別入試ともに非常に少なく、全体の3~4%程度であった。表3では「2回以上」に該当する履修者は少ないため、「0回」と「1回(以上)」の2区分としている。比較分析の結果、入試制度による違いはみられなかった。

(c) の「未提出回数」では、レポートをすべて提出した履修者（回数が「0回」の履修者）は、学部別入試よりも総合入試の方が7.7ポイント多く、やや開きがあった。未提出回数の割合について、入試制度別で比較した結果、0.1%水準で有意差がみられた。また、残差分析の結果、未提出回数は、すべての回数で有意であった ($p<.01$)。欠席と同様に、総合入試よりも学部別入試の方が、レポートを全て提出しようとする取組姿勢は弱いといえる。

(d) の「遅滞回数」では、レポートを1度も遅れずに提出した履修者（回数が「0回」の履修者）は、学部別入試よりも総合入試の方が8.2ポイント多く、未提出回数の「0回」よりもその差はやや大きかった。入試制度別で比較した結果、0.1%水準で有意差がみられた。また、残差分析の結果、遅滞回数は、すべての回数で有意であった ($p<.01$)。欠席とレポート未提出と同様に、総合入試よりも学部別入試の方が、レポートを全て遅れずに提出しようとする取組姿勢は弱いといえる。

したがって、「遅刻回数」は入試制度による有意な差はなく、「欠席回数」「未提出回数」「遅滞回数」では、学部別入試よりも総合入試の方が該当回数割合は有意に少なく、取組姿勢は良いという結果であった。

また、取組姿勢の良さに着目すると、総合入試、学部別入試ともに、学習状況において「0回」の割合が最も大きいのは遅刻であり、次いで欠席、未提出、遅滞の順であった。そして、この順で総合入試と学部別入試における「0回」の割合の差も大きくなっている。この順序と入試制度間の割合差は、実験の成績評価基

準と関係すると考えられる。実験評価点から減点される点数は欠席が一番大きく、次いでレポートの未提出、レポート提出の遅滞の順である。つまり、減点の程度が軽い学習状況ほど、学部別入試の取組姿勢が悪くなる傾向があるといえる。ただし、遅刻の減点の程度は軽い、遅刻する履修者は非常に少ない。自然科学実験は初年次科目であり、実験前に同じ校舎内で別の講義を履修している学生が多いため、実験室への移動は短時間で可能である。そのため、実験を出席する意思のある履修者で遅刻する学生は非常に少ないと推測される。

以上、履修者の学習履歴データである各学習状況の分析を行うことで、履修者の全体像を具体的に把握することができた。また、実験に対する取組姿勢について検証を行った結果、総合入試と学部別入試という2つの入試制度間において、取組姿勢に明らかな差があることが確認された。

3.3. 分析 (3) 取組姿勢と成績評価

前節では、実験に対する取組姿勢が入試制度により差があることを確認したが、ここではその入試制度により異なる取組姿勢と、成績評価の関連性について検証する。一般に、講義や実験に対する取組姿勢が良いほど、成績評価は良くなることと期待されるだろう。そこで、取組姿勢の度合いと成績評価の関係を確認するため、表3に示した学習状況 (a) から (d) の該当回数別の当該履修者の平均レポート点を算出する。そして、遅刻回数を除く学習状況では、入試制度によって差異がみられたため、入試制度の観点から取組姿勢と成績

表3 各学習状況における入試制度による同等性の検定結果

学習状況	入試制度	0回		1回		2回以上		合計		χ^2	df	p
		人数	(%)	人数	(%)	人数	(%)	人数	(%)			
(a) 欠席回数	総合	4,929	(95.7) △	202	(3.9) ▼	19	(0.4)	5,150	(100.0)	61.1	2	**
	学部別	3,767	(91.8) ▼	311	(7.6) △	24	(0.6)	4,102	(100.0)			
(b) 遅刻回数	総合	4,993	(97.0)	157	(3.0)	-	-	5,150	(100.0)	3.7	1	n.s.
	学部別	3,947	(96.2)	155	(3.8)	-	-	4,102	(100.0)			
(c) 未提出回数	総合	4,564	(88.6) △	439	(8.5) ▼	147	(2.9) ▼	5,150	(100.0)	110.9	2	**
	学部別	3,317	(80.9) ▼	614	(15.0) △	171	(4.2) △	4,102	(100.0)			
(d) 遅滞回数	総合	4,196	(81.5) △	581	(11.3) ▼	373	(7.2) ▼	5,150	(100.0)	92.7	2	**
	学部別	3,005	(73.3) ▼	625	(15.2) △	472	(11.5) △	4,102	(100.0)			

注) △および▼は残差分析の結果を示す。△は有意に多い、▼は有意に少ない ($p<.01$)

** $p<.001$

注) (b) 遅刻回数における「1回」は1回以上をすべて含む。

表4 学習状況別、入試制度別、平均レポート点の平均値および標準偏差

入試制度	(a)欠席回数			(b)遅刻回数		(c)未提出回数			(d)遅滞回数			全体
	0回	1回	2回以上	0回	1回以上	0回	1回	2回以上	0回	1回	2回以上	
総合	3.34 (0.62)	2.96 (0.58)	2.81 (0.47)	3.33 (0.62)	2.91 (0.66)	3.37 (0.61)	2.95 (0.61)	2.81 (0.52)	3.40 (0.61)	3.07 (0.59)	2.82 (0.59)	3.28 (0.63)
学部別	2.99 (0.63)	2.84 (0.54)	2.98 (0.54)	3.00 (0.62)	2.63 (0.63)	3.04 (0.63)	2.76 (0.56)	2.73 (0.50)	3.07 (0.62)	2.84 (0.59)	2.61 (0.55)	2.93 (0.63)

注) 表中に示した数値は、上段：平均値、下段：SD.

表5 学習状況別平均レポート点の2要因分散分析の結果

学習状況	主効果		交互作用
	入試制度	回数	入試制度×回数
(a)欠席回数	2.26	7.23 *	4.93 *
(b)遅刻回数	72.95 *	121.73 *	0.76
(c)未提出回数	56.72 *	100.18 *	8.15 *
(d)遅滞回数	188.23 *	255.80 *	3.71

注) 表中に示した数値は、分散分析で得られたF値.

* $p < .01$

評価の検証を行う。

まず、表3に示した学習状況(a)から(d)における該当回数別の平均レポート点の平均値を表4へ示す。平均レポート点の全体の平均値は、総合入試は3.28、学部別入試は2.93であり、総合入試の方が学部別入試よりも有意に高い($t(8787)=25.80, p < .01$)。2つの入試制度間では、平均レポート点に違いがみられる。

つぎに、入試制度の観点から取組姿勢と成績評価の関連性を検証するため、表4に示した学習状況別の平均レポート点に関して、入試制度(2)×回数(3, (b)は2)の2要因分散分析を行った。分散分析の結果を表5へ示す。

学習状況(a)の「欠席回数」では、回数($F(2,9276)=7.23, p < .01$)の主効果が有意、入試制度と回数の交互作用が有意であり($F(2,9276)=4.93, p < .01$)、入試制度で欠席回数の影響が異なることが示された。回数の単純主効果の検定結果では、総合入試において有意であった($F(2,9246)=11.00, p < .01$)。下位検定として回数の多重比較を行った結果、「0回」の平均レポート点は、「1回」および「2回以上」よりも高かった($p < .05$)。つまり、学部別入試では欠席回数にかかわらず同程度の平均レポート点であるが、総合入試では、1度も欠席していない履修者の平均レポート点は高くなるといえる。

(b)の「遅刻回数」では、入試制度($F(1,9251)=72.95, p < .01$)と回数($F(1,9251)=121.73, p < .01$)の主効果が有意であった。遅刻したことがある履修者は全体の3~4%程度と割合は小さいが、遅刻したことがある履修者の平均レポート点は低くなる傾向があるといえる。

(c)の「未提出回数」では、入試制度($F(1,9251)=56.72, p < .01$)と回数($F(2,9251)=100.18, p < .01$)の主効果が有意、入試制度と回数の交互作用が有意であった($F(2,9251)=8.15, p < .01$)。回数の単純主効果の検定結果では、総合入試($F(2,9251)=81.44, p < .01$)、学部別入試($F(2,9251)=26.89, p < .01$)とも有意であった。下位検定として、回数の多重比較を行った結果、総合入試では「0回」の平均レポート点は、「1回」および「2回以上」よりも高かった($p < .05$)。学部別入試では、「0回」の平均レポート点は、「1回」および「2回以上」よりも高かった($p < .05$)。つまり、総合入試と学部別入試のいずれの場合も、未提出のレポートが無い履修者の平均レポート点は、未提出のある履修者よりも高いといえる。

(d)の「遅滞回数」では、入試制度($F(1,9251)=188.23, p < .01$)と回数($F(1,9251)=255.80, p < .01$)の主効果が有意であった。下位検定として回数の多重比較を行った結果、「0回」の平均レポート点は、「1回」および

び「2回以上」よりも高く、「1回」は「2回以上」よりも高かった ($p < .05$)。つまり、レポートの遅滞回数が多くなるほど、点数は低くなる傾向があるといえる。

したがって、学習状況別に平均レポート点を分類して検証した結果をまとめると、学部別入試よりも総合入試の方が平均レポート点は良く、また、取組姿勢が良いほど平均レポート点は良くなる傾向が明らかとなったといえる。特にレポートの未提出回数や遅滞回数が少ないほど、平均レポート点は高くなる傾向があった。レポートの未提出や遅滞は、実験に対する興味や理解度、あるいは学習そのものに対する意欲が低いために生じると考えられるため、該当回数が多くなるほど平均レポート点が低くなるのは順当な結果である。遅刻に関しても、その該当者は非常に少ないが、遅刻する履修者の平均レポート点は明らかに低い。逆に言えば、全体の該当者が少ない中で遅刻する履修者は、レポート作成に対する姿勢も良くはないと推測される。

入試制度の観点から平均レポート点をみると、前節の実験に対する取組姿勢と同様、明らかな差がみられた。学部がすでに決まっている履修者とこれから移行する履修者では、レポート作成に対する姿勢も異なっていると考えられる。特に総合入試では、2年次進級時の学部・学科移行の際、実験の成績評価が移行点へ算入されるため、すべての実験を休まずに、レポート提出の締め切りを厳守し、良いレポートを書いて提出

しようという意識は高いと思われる。

以上の結果から、取組姿勢と成績評価には関連性があり、入試制度によらず取組姿勢の程度によって成績評価にも差があることが確認された。また、入試制度の観点からみると、入試制度の違いが実験に対する取組姿勢や成績評価へ影響していることが示唆された。

4. 成績評価予測

4.1. レポート点と成績評価の関係性

蓄積された学習履歴データを分析する意義の一つとして、データの傾向を検証し予測することで、今後の学習指導や教育支援に役立てられることがあげられる。特に、履修期間の初期段階で個々の学生の最終成績評価が予測できれば、個々の学生の成績に応じた実験グループ編成の見直しや、実験指導、レポート指導の改善など、実験の履修期間中に可能となりえる。

取組姿勢と成績評価の分析では、レポートの未提出回数や遅滞回数といった取組姿勢が良いほど、レポート点が良くなる傾向が明らかとなった。欠席回数や未提出回数などの取組姿勢は最終成績評価の判定基準点数に組み込まれているため、その回数が多くなるにつれて成績評価が低くなるのは明らかであろう。このため、取組姿勢から成績評価予測は可能であると考えられるが、取組姿勢はある程度、実験回を重ねなければ判明しない指標であり、履修期間の早い段階で成績評価を予測することは困難である。

そこで、実験履修期間のはじめから利用可能なレポート点に着目し、履修の初期段階で最終成績評価の予測が可能か検証を試みる。まず、各回のレポート点と成績評価がどのような関係にあるのかを確認するため、各回のレポート点と、最終成績評価の判定基準となる全実験テーマの合計点数（実験評価点とレポート点の合計点数）の相関分析を行う。自然科学実験の各テーマの内容は独立しているため、履修者のテーマへの興味の度合いなどによって、各回のレポートの「出来栄」に差があり、毎回のレポート点に開きがあることも考えられるためである^{2) 3)}。

表6は各回のレポート点と最終成績評価の分析結果である。相関係数は0.5から0.6を示しており、初回のみやや低いものの、レポート点と成績評価は正の相関がある。入試制度による違いはあまり見られない。履修者のテーマへの興味の度合いや得意、不得意などで各回のレポート点にばらつきがあると思われたが、各回のレポート点は履修者ごとにある程度同じ水準で推

表6 歴各回のレポート点と最終成績評価の相関分析の結果

実験回数	総合入試	学部別入試
1回目	.53 *	.51 *
2回目	.56 *	.56 *
3回目	.55 *	.55 *
4回目	.59 *	.58 *
5回目	.58 *	.60 *
6回目	.60 *	.58 *
7回目	.56 *	.57 *
8回目	.60 *	.59 *
9回目	.57 *	.56 *
10回目	.60 *	.58 *
11回目	.60 *	.56 *
12回目	.57 *	.57 *

注) Pearson, 両側検定 * $p < .01$

表7 記述統計量

変数名	レポート点 1回目	レポート点 2回目	レポート点 3回目	レポート点 4回目	レポート点 5回目	レポート点 6回目	入試制度 ダミー	合計点数
平均	3.20	3.17	3.16	3.18	3.16	3.18	0.56	96.51
標準偏差	1.03	1.06	1.05	1.05	1.06	1.06	0.50	9.15
最大	5.00	5.00	5.00	5.00	5.00	5.00	1.00	120.00
最小	0.00	0.00	0.00	0.00	0.00	0.00	0.00	72.00
観測数	9160.00	9158.00	9166.00	9141.00	9121.00	8999.00	9252.00	9252.00

表8 合計点数に対する各レポート点および入試制度の影響に関する分析結果

	1	2	3	4	5	6
定数項	80.29 ** (306.37)	73.76 ** (278.05)	69.96 ** (269.24)	67.12 ** (272.85)	65.43 ** (285.61)	64.28 ** (301.21)
入試制度ダミー	3.74 ** (23.62)	2.88 ** (20.59)	2.41 ** (18.90)	1.99 ** (17.25)	1.64 ** (15.61)	1.40 ** (14.55)
1回目	4.46 ** (58.33)	3.12 ** (43.47)	2.49 ** (37.40)	2.06 ** (33.65)	1.76 ** (31.36)	1.58 ** (30.85)
2回目		3.59 ** (51.43)	2.72 ** (40.93)	2.20 ** (35.92)	1.83 ** (32.49)	1.55 ** (29.92)
3回目			2.82 ** (42.31)	2.12 ** (34.00)	1.76 ** (30.73)	1.51 ** (28.68)
4回目				2.65 ** (43.12)	2.10 ** (36.79)	1.75 ** (33.15)
5回目					2.22 ** (39.41)	1.80 ** (34.27)
6回目						1.93 ** (37.17)
自由度修正済み 決定係数	0.32	0.47	0.56	0.64	0.70	0.75
標準誤差	7.48	6.52	5.89	5.30	4.77	4.29
F値	2170.88	2730.42	2890.60	3166.28	3429.76	3682.38
p値	0.00	0.00	0.00	0.00	0.00	0.00
観測数	9160.00	9073.00	8992.00	8896.00	8785.00	8574.00

注) 表の上段に示した数値は係数推定値. 括弧内は t 値.

** $p < .001$

移していると考えられる。つまり、取組姿勢の良い学生は毎回5点、あるいは取組姿勢の悪い学生は毎回1点など、レポート点数が固定化されている可能性が高いと思われる。

4.2. 最終成績評価の予測分析

つぎに、成績評価予測に向けた検証を行うため、各

回のレポート点と、最終成績評価の判定基準となる全実験テーマの実験評価点とレポート点の合計点数を用いて、最小二乗法により分析を行う。分析に用いた各変数の記述統計量を表7へ示す。表7において、入試制度ダミーは、総合入試による履修者を1、学部別入試による履修者を0とするダミー変数である。合計点

数は、欠席、遅刻、レポートの未提出、レポート提出の遅滞といった実験の取組姿勢に関する減点分を加味した数値である。

表8は、合計点数を被説明変数、入試制度とレポート点を説明変数とした6個の回帰モデルの分析結果である。各モデルにおいて、入試制度ダミーは合計点数に対して正の有意な効果を示している。つまり、総合入試は、学部別入試よりも平均レポート点が高く、取組姿勢も良いことから減点分が小さく、合計点数は高くなっている。モデル1では1回目の実験のレポート点のみ投入しているが、モデル2から6では、実験の2回目から6回目のレポート点を順に投入している。自明ではあるが、各レポート点は合計点数に加算されているため、各回のレポート点をモデルに投入していくと、決定係数は当然大きくなっていく。全12回の実験のうち、6回分のレポート点を投入したモデル、すなわち、前半が終了した時点のモデル6では、決定係数は0.75と大きくなっている。最初の3回目のモデル3においても、決定係数は0.56とモデルの半分の説明力を持っている。したがって、最初の数回のレポート点のみをみても、ある程度、最終成績評価の予測が可能になっているといえるだろう。

5. ま と め

本稿では、大学初年次理系実験の約1万名の学習履歴データを用いて、入試制度と取組姿勢がどのように成績評価と関連しているのかについて検証を行った。履修者の全体像の分析では、毎年、一定数の「不可」が存在し、2年次進級時に学部を決定する総合入試よりも、入学時に学部が決定している学部別入試の方が「不可」の割合は大きいことが明らかとなった。実験に対する取組姿勢では、遅刻回数は入試制度による違いはないものの、欠席回数、レポート未提出回数、レポート遅滞回数といった学習状況においては、学部別入試よりも総合入試の方が該当割合は少なく、取組姿勢は良いことが明らかとなった。成績評価においても、入試制度による差がみられ、学部別入試より総合入試の方が平均レポート点は良いことが明らかとなった。取組姿勢と成績評価では、取組姿勢が良いほどレポート点が良くなる傾向があり、特にレポートの未提出回数や遅滞回数が少ないほど、レポート点が高くなることが示唆された。また、過去に蓄積されたデータ分析を行うことで、ある程度、最終成績評価を予測できる可能性が提示された。

したがって、入試制度や取組姿勢の観点から成績評価分析を行うことで、より詳細な履修者の把握が可能となり、今後の学習支援の改善に資する情報が得られたと考えられる。学習支援の改善を目的として全実験終了後に履修者アンケートは実施されているが、取組姿勢やレポート点など、履修期間中に得られるデータ分析結果も活用可能であれば、実験履修期間中に実験指導の見直しや改善が期待できる。具体的には、入試制度別になぜ実験を履修するのか、履修目的を再確認させる事前講義の充実や、最初の1、2回目の実験レポートの内容が良くなかった履修者に対する指導方法の検討、遅刻やレポートの遅れが多い学生に対する指導頻度・方法の検討が考えられるだろう。また、本稿では入試制度そのものは検討対象としていないが、入試制度の違いにより取組姿勢と成績評価に明確な差が表れており、このような学習履歴データの分析結果は、入試制度の検討場面においても有益な情報の一つとして活用できると期待される。

今後の課題としては、より詳細なデータ収集・分析である。自然科学実験では全部で26テーマが用意されており、1日(1回)の実験では約半数のテーマを同時に実施している。履修者は全12テーマを履修することになるが、履修するテーマ内容、順番はさまざまである。しかし、今回のデータでは、誰が、何回目どのテーマを履修したかまでは収集できておらず、テーマ別の分析には至っていない。より詳細にデータ収集と分析を行うことで、精度の高い成績評価予測が可能となり、教育改善に資するさらなる情報の提供が可能になると考えられる。今後も学習履歴データの蓄積を継続していき、教育改善に寄与しうるデータ分析を行っていきたい。

註

1) 2014年までは「秀15%、優50%、良30%、可または不可5%」であったが、大学全体の成績評価方法の変更に伴い、2015年からは「A+、A、A-、B+、B、B-、C+、C(ここまでが、単位取得)、D、D-、F(以上、不合格または評価せず)」の11段階評価となっている。ただし、自然科学実験では、最終成績評価の判定となる基準点の算出方法(全12回の実験テーマにおける実験評価点とレポート点を合計し、合計点数の上位から相対評価を行う方法)は同じであり、単位取得の条件も120点満点中、72点以上と変更はない。

2) 自然科学実験では、履修者の選択したカテゴリ、

テーマによりレポート点に差がないよう、一定数の実験グループ内で相対的に採点を行っている。レポート点全体の平均値は、実験回数で差はみられていない。

3) レポート点全体の平均値をカテゴリ別にみると、3.14点から3.19点、最終成績評価(全テーマの合計点数)は96.1点から96.8点と若干の幅があった。また、最終成績評価に加味される取組姿勢のうち、未提出回数のみカテゴリ間で若干の開きがあった。ただし、ここではレポート点と最終成績評価の実験回数に着目しているため、全カテゴリ全体の平均値を用いることとする。

付 記

本稿は、田島(2014)に新たなデータを追加し、分析および検証をさらに発展させたものである。本報告の一部は、田島(2013, 2016)および田島・田辺(2013)で発表している。なお、自然科学実験の成績評価分析および分析結果の公開に関しては、北海道大学高等教育推進機構より許諾を得ている。

参 考 文 献

- 北海道大学(2015)総合入試案内【平成27年度版】(PDF).
http://www.hokudai.ac.jp/admission/27sogo_ppt.pdf
(参照日 2017.1.25)
- 北海道大学自然科学実験編集委員会編(2016)自然科学実験. 学術図書出版, 東京.
- 杉山いおり, 渡辺雄貴, 加藤浩, 西原明法(2016)企業内eラーニングにおける社会人の最終学習状態推定. 日本教育工学会論文誌, 40 (Suppl.): 85-88
- 田島貴裕(2013)大学初年次理系実験に対する取組姿勢と成績評価の関連性—入試制度の際の観点から—. 日本教育工学会第29回全国大会講演論文集, 589-590
- 田島貴裕, 田辺大人(2013)大学初年次理系実験の成績評価分析. 日本理科教育学会第63回全国大会大会発表論文集, 429
- 田島貴裕(2014)初年次教育の理系実験に対する取組姿勢と成績評価の関連性—入試制度の観点から—. 日本教育工学会論文誌, 38 (Suppl.): 1-4

田島貴裕(2016)初年次理系実験の成績評価分析—誰が不可になるのか—. 日本通信教育学会第64回研究協議会発表要旨収録, 63-68

竹生久美子, 辻靖彦(2016)eラーニング科目における受講ペースと成績との関連. 日本教育工学会論文誌, 40 (Suppl.): 153-156

寺澤孝文(2015)教育ビッグデータの大きな可能性とアカデミズムに求められるもの—情報工学と社会科学のさらなる連携の重要性—. コンピュータ&エデュケーション, 38 : 28-38

中央教育審議会(2008)学士教育課程の構築に向けて(答申).

http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/1217067.htm (参照日2017.1.25)

山田雅之, 中村信次, 佐藤慎一, 野寺綾(2010)eラーニングにおける学習計画とドロップアウト率の関係. 日本教育工学会論文誌, 34 (Suppl.): 73-76

Summary

We examined the relationship between the academic performance and student's learning attitude or university admissions system, by analyzing approximately 10,000 student's learning log data on "Laboratory Exercise in Natural Sciences" for first year university students.

The data analysis revealed the following: (1) Every year, a certain proportion student gets a failing grade, and the proportion differs among admissions system. (2) The academic performance and learning attitude are also varied by the admissions system. The better the student's learning attitude, the academic performance tends to be higher. (3) Score of several reports from the start is relevant to the performance evaluation. Therefore, it suggests the possibility of predicting the final grade evaluation at the initial stage of the lecture.

EDUCATIONAL EVALUATION, UNIVERSITY ADMISSION SYSTEM, FRESHMAN EDUCATION, HIGHER EDUCATION

(Received January 30, 2017)