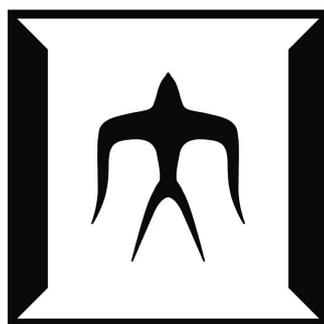


論文 / 著書情報
Article / Book Information

題目(和文)	学生が中退する可能性の数値化と遷移に関する研究
Title(English)	
著者(和文)	白鳥成彦
Author(English)	Naruhiko Shiratori
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第12496号, 授与年月日:2023年6月30日, 学位の種別:課程博士, 審査員:室田 真男,調 麻佐志,猪原 健弘,江原 慶,永原 健太郎
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第12496号, Conferred date:2023/6/30, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

学生が中退する可能性
の数值化と遷移に関する研究

東京工業大学大学院
環境・社会理工学院 社会・人間科学系



白鳥 成彦

指導教員
室田 真男 教授

目次

目次.....	I
図目次.....	IV
表目次.....	IV
第1章 研究の背景：高等教育における中退.....	1
1.1. 高等教育における中退.....	1
1.1.1. 中退者数と中退率.....	1
1.1.2. 中退率と偏差値の関係性.....	2
1.2. 中退学生の特徴.....	3
1.2.1. 大学生が中退する理由.....	3
1.2.2. 学生における中退行動の特徴把握.....	4
1.2.3. 中退学生へのインタビューからの特徴把握.....	4
1.3. 本論文の構成.....	6
第2章 先行研究.....	8
2.1. 中退の理論研究.....	8
2.1.1. 1970年代までの中退研究：心理学的アプローチ.....	8
2.1.2. 1970年代以降の中退理論研究.....	9
2.2. 中退の要因に関する研究.....	12
2.2.1. 中退に関連するマクロ変数.....	12
2.2.2. 中退に関連するミクロ変数.....	13
2.3. 中退の防止施策に関する研究.....	15
2.4. LA・EDM・IR.....	18
2.4.1. LAとEDM.....	18
2.4.2. IR.....	19
2.5. 中退を予測する研究.....	20
2.5.1. 中退を予測する研究の分類.....	20
2.5.2. 代表的な中退予測研究の種類.....	21
2.6. 第2章のまとめ.....	23
第3章 リサーチクエスチョン：先行研究の限界と本研究の貢献.....	24
3.1. これまでの研究の限界.....	24
3.2. リサーチクエスチョン.....	25
第4章 入学から中退までの学修状態の把握.....	27
4.1. 利用するデータと変数.....	27

4.1.1.	利用するデータ	27
4.1.2.	学期と週の表現.....	28
4.1.3.	入学時まで取得できる変数.....	28
4.1.4.	入学後に取得できる変数 1：学期ごとの変数.....	29
4.1.5.	入学後に取得できる変数 2：学期中の変数	29
4.1.6.	目的変数：中退したかどうか.....	29
4.2.	学修状態の数値化：中退確率の定義	30
4.2.1.	中退確率の算出方法	30
4.2.2.	ロジスティック回帰モデルの精度	32
4.3.	中退確率を用いた状態の特性	33
4.3.1.	有意な変数の抽出.....	33
4.3.2.	中退確率を用いた状態.....	37
4.3.3.	中退確率を用いた学生の分類.....	39
4.4.	中退確率を用いた学修状態の遷移	39
4.4.1.	中退確率を用いた遷移の数値化.....	39
4.4.2.	中退確率を用いた遷移の分類.....	41
4.5.	中退確率の遷移を用いたクラスタの作成.....	43
4.6.	中退パターンの導出.....	47
4.6.1.	中退パターンの導出方法.....	47
4.6.2.	導出された中退パターン	47
4.6.3.	5 種類の中退パターンのまとめ.....	49
4.7.	第 4 章のまとめ	50
第 5 章	予測 GPA を用いた学修状態の類型化.....	51
5.1.	1 年目春学期の重要性	51
5.2.	本章の目的.....	52
5.3.	利用するデータと変数	53
5.3.1.	利用したデータ.....	53
5.3.2.	利用した変数	53
5.4.	予測 GPA を用いた学生の類型化	56
5.4.1.	類型化の手法	56
5.4.2.	1 年目春学期 GPA の予測.....	56
5.4.3.	遷移ベクトルの作成	57
5.4.4.	遷移ベクトルを用いた学修状態の類型化	57
5.4.5.	グループ差の抽出.....	57
5.5.	数値実験と結果	58

5.5.1. 実験データ	58
5.5.2. 予測精度.....	58
5.5.3. 類型化の結果	59
5.5.4. グループ差を用いたハイリスク学生群の特徴抽出.....	61
5.6. 第5章のまとめ	65
第6章 まとめと今後の展開.....	66
6.1. 研究のまとめ	66
6.2. 中退防止施策への利用	67
6.2.1. 中退防止施策を誰が行うべきなのか.....	67
6.2.2. パターンごとの中退防止施策.....	68
6.2.3. 中退をしなかった学生のパターン	69
6.2.4. 中退パターンを用いた中退防止施策.....	71
6.3. 本研究の限界.....	72
6.3.1. 有用性における限界	72
6.3.2. 一般性における限界	73
6.4. 今後の展開.....	73
6.4.1. 有用性における展開	73
6.4.2. 一般性における展開	74
謝辞.....	76
本研究に関わる研究発表.....	77
査読付き論文.....	77
査読付き国際会議.....	77
国内口頭発表.....	77
参考文献.....	79

図目次

図 1.1 本論文の構成図.....	7
図 2.1 中退理論研究の流れ.....	11
図 2.2 Tinto の変数群と大学で取得しているデータ.....	13
図 4.1 学期ごとの中退確率ヒストグラム.....	38
図 4.2 入学時の中退確率と1年目春学期終了時の中退確率の散布図.....	41
図 5.1 標準在学年内で卒業した人としない人の1年春 GPA の分布.....	52
図 5.2 週毎の予測精度.....	58
図 5.3 類型化された2018年度入学生の平均予測 GPA.....	59
図 5.4 類型化された2019年度入学生の平均予測 GPA.....	60
図 6.1 中退をしなかった学生の類型化.....	70

表目次

表 2.1 中退に関連する変数の整理.....	15
表 2.2 中退防止施策のまとめ.....	16
表 4.1 学期ごとの中退者数.....	27
表 4.2 入学時まで取得できる変数.....	28
表 4.3 入学後に取得できる変数1.....	29
表 4.4 入学後に取得できる変数2.....	29
表 4.5 目的変数.....	30
表 4.6 精度の分類.....	32
表 4.7 ロジスティック回帰モデルの精度.....	33
表 4.8 有意な変数1 (入学時から3年目春学期まで).....	35
表 4.9 有意な変数2 (3年目秋学期から4年目春学期まで).....	36
表 4.10 中退確率のサンプル.....	37
表 4.11 学期ごとのグループ人数.....	39
表 4.12 学期ごとのグループの遷移.....	42
表 4.13 中退確率の遷移を用いたクラスター.....	46
表 4.14 中退パターン A.....	47
表 4.15 中退パターン B.....	48
表 4.16 中退パターン C.....	48
表 4.17 中退パターン D.....	49
表 4.18 中退パターン E.....	49

表 4.19 5 種類の中退パターン	50
表 5.1 第 5 章で利用したデータ	53
表 5.2 第 5 章で利用した変数	54
表 5.3 授業科目名とカテゴリ	55
表 5.4 良化グループと悪化グループの割合	61
表 5.5 標準修業年月内で卒業ができなかった学生とできた学生の割合	62
表 5.6 2018 年度入学生における下位 2 グループに関連する変数	63
表 5.7 2019 年度入学生における下位 2 グループに関連する変数	64
表 6.1 中退パターンごとの施策	69

第1章 研究の背景：高等教育における中退

1.1. 高等教育における中退

1.1.1. 中退者数と中退率

大学における中途退学率（中退率）は各大学の状況によって大きく異なる。中退率の高さは大学の学生募集や教育の質に大きい負の影響を与えるだけでなく、学生にとっても時間や資金を浪費してしまうことになる。

文部科学省の調査「経済的に困難な学生への支援状況・学生の修学状況等について（調査）」（令和4年3月30日付け文部科学省高等教育局事務連絡）において、令和3年度末（2021年度末）時点において中途退学者の状況がまとめられている（文部科学省2022）。この調査は令和3年度末時点において全国の国公立大学、及び高等専門学校に対して行われた調査であり、回答率は96.0%となっている。この調査によると日本全国の大学、及び高等専門学校における中退者数は令和3年度（2021年4月から2022年3月）において57,875人であり、学生数全体に占める中退者数の割合は1.95%と報告されている。また、令和2年度（2020年4月から2021年3月）においては57,913人の中退者数、1.95%の割合、令和元年度（2019年4月から2020年3月）においては74,129人の中退者数、2.5%の割合と報告されている。

同じ調査の中で国公立専門学校の中退者数も報告されており、回答率が62.4%となっている。令和3年度において24,056人が中退を1年間でしており、生徒数に占める中退者数の割合としては5.81%になっている。令和2年度は中退者数が24,326人、割合が5.53%であり、令和元年度は中退者数が29,038人、割合が6.69%と示されている。大学、専門学校を含めると日本全体としては1年間の間に80,000人から100,000人程度の学生が中退をしている現状となっている。

また、読売新聞によって行われていた大学の実力においても中退率は大きな注目点となっており、大学の評価においても中退率は大きな指標となってきた（読売新聞教育ネットワーク事務局2018）。大学の実力2019の調査では設置校別の中退率を表しており、設置校別の入学者数と中退者数を比べ、標準修業年限（4年間、または6年間）内で中退をする学生の割合を表している。この定義の数値においては国立大学の中退率の平均値が2.9%、公立大学の平均値が3.9%、私立大学の平均値が8.0%と報告されている。

大学、専門学校ごとの中退率はさらに高いものとなっている。1年間の中退者数、中退率を報告している文部科学省の調査では中退者の数と中退率は日本全体の数である。読売新聞の調査において中退率は設置校別ごとに2%から8.0%となっているが、読売新聞における調査の中退率の数値も全体としての数値である。読売新聞の調査では、それぞれの大学、学部によって中退率が示されているが、各大学によって中退率は0%台の大学から30%台の大学までさまざまである。

2014 年度に都内の大学に対して著者らが行った調査では、当該年度の中退率が 30% 超と高い値を示している。このように中退率は日本全体では 1 年間に 2% 弱となるが、それぞれの大学によって大きく異なっている。

アメリカでは中退率は日本よりも高い傾向にある。National Center for Education Statistics (NCES) の調査によると、アメリカにある 4 年制の大学において 6 年間の卒業率は全体で 64%、男性が 60%、女性が 67% であることが報告されている。この調査は 2014 年秋学期に入学した学生が 2020 年までに学位を取得できたのかを調査しており、公立では 63%、私立（非営利）では 68%、私立（営利）では 29% であると述べている (National Center for Education Statistics 2022)。2014 年入学者の卒業率割合 (64%) と比較して、1996 年入学者の割合は 55.4%、2000 年入学者の割合は 57.6%、2004 年入学者の割合は 58.4%、2008 年入学者の割合は 59.6%、2012 年入学者の割合は 62.4% であり、卒業率は徐々に高くなってきている。4 年間の卒業率にすると 2014 年入学者で 46.6%、5 年間の卒業率にすると 60.5% と報告されている。このようにアメリカでは日本よりも中退率は高くなっている。OECD (Organisation for Economic Cooperation and Development) は国際比較ができるように指標を発表しているが、その一つが高等教育卒業率である。OECD の 2019 年データによると 30 歳未満の学生の学士課程または同等レベルにおける 30 歳未満の卒業率は OECD 各国平均で 34%、EU 加盟 22 カ国平均で 32% となっており、日本は 44% となっております。

中退率は大学における学修成果・教育成果を図る指標としても大切なものである。令和 2 年 (2020 年) 1 月に中央教育審議会大学分科会によって出された教学マネジメント指針の中において、大学の教育活動に伴う基本的な情報の一つとして、修業年限内に卒業する学生の割合、留年率、中途退学率があげられている (文部科学省中央教育審議会大学分科会 2020)。日本における中退率はアメリカよりも低く、全体としては 2% 程度となっており低く見えがちだが、中途退学率はすべての大学において収集可能なものであり、学修成果・教育成果の把握の一つとして、学位プログラムごとに把握、可視化、情報公表をすることが求められている。

以上のように日本においても、世界においても中退率はそれぞれの大学によって経営的に、そして学術的に重要な指標であり、大学全体として公表が求められる指標であることが分かる。

1.1.2. 中退率と偏差値の関係性

日本の大学全体における中退の現状を述べてきたが、次に日本において中退と関連する指標の一つである偏差値との関係性について述べる。

清水は社会科学系学部の偏差値と退学率を分析した結果、退学率や就職率は偏差値によってかなりの部分が説明できるとしている (清水 2013, 2021a)。私立大学においては偏差値が低い大学 (偏差値が 44 以下の大学: n=132 校) の退学率の平均は 16% を超えている。偏差値が高い大学ほど退学率は低く、偏差値が低い大学ほど退学率は高いことを指摘している。また、回帰分析の結果の中で清水は偏差値が 1 あがると、退学率は 0.38% 下がることを指摘している。

高等教育への進学率が50%を超えてユニバーサル段階になっている日本の高等教育・大学においては、大学を選ばなければ大学にはどこかの大学には進学できる状態となっている。偏差値のみを基準にして大学を選ぶのではないが、偏差値が低い大学は入学を志望する者にとってだいぶ入学しやすい状況となっている。先程の清水の事例では退学率の四分位点で大学を分割した場合、退学率が低い群（第1四分位群）と退学率が高い群（第4四分位群）、残りの中間群に分けて考察している(清水 2021b)。退学率の高い群と低い群、中間群には様々な差が出てくる。退学率が低い群の偏差値の平均値は59.2であり、退学率が高い群の偏差値の平均は43、中間群の偏差値の平均は50.1となっている。退学率が高いほど偏差値が低く、退学率が低いほど偏差値が高くなる傾向が見られる。退学率が低い群では学生数が多いと中退率が減る結果であるのに対し、退学率が高い群では学生数が多いと中退率が増える結果になっている。また、留学生比率は退学率が高い群で有意に正となっている変数だが、留学生比率が高いと中退率が高い傾向になっている。またST比率（教員学生比率）はどの群においても有意にはなっていない。このように中退率の違いは充足率、学生数、充足率、留学生比率等で関係があり、中退率が高い大学と低い大学では様々な点で異なってくるのがわかる。

1.2. 中退学生の特徴

1.2.1. 大学生が中退する理由

次に大学生がなぜ中退するのかの理由を議論する。学生が大学を中退することは大学にとっても学生にとってもなるべく避けたいことであるが、なぜ学生は中退していくのだろうか。

先述した文部科学省の調査によれば、令和3年度57,875人における中退する理由として転学等が16.3%、学生生活不適応・修学意識低下は15.7%、就職・起業等が14.3%、経済的困窮が13.5%、学力不振が7.7%、心神耗弱・疾患が5.0%（その他は19.5%）と続いている。また、2016年に東京大学による調査では大学設置者別、昼夜別による中退する理由を載せている(東京大学 2016)。私立（昼間）大学に関しては、転学が21.2%と多く、次いで学業不振が18.4%、就職が15.1%となっている。私立大学以外では一身上の都合やその他といった理由が判然としない場合が約20%から40%と多いこと、就職を理由とするケースも同様に13%から23%と多いことも指摘している。

一方で、前述した東京大学の調査において中退理由は複合的なものになることを同様に指摘している。報告書の第5章では、ハローワーク調査からの大学中退者の特徴・背景を喜始照宣（きしあきのり）がまとめている。彼は労働政策研究・研修機構（JILPT）が2015年に出した調査をまとめ（労働政策研究・研修機構 2015）、ハローワークに来所した大学からの中退者（n=758）における中退理由を分析している。この調査の中では中退しようと思った理由を複数選択で、最も重要な中退理由を一つ選択させる形で聞いており、大学生活の累計や居住形態、奨学金の有無等も同様に聞いている。最も重要な中退理由は学業不振・無関心が男女計で42.9%（男性50.1%、

女性 26.9%) と最も多く、次いで進路変更が男女計 15.1%、経済的理由が男女計 14.4%となっている。また複数回答ができる中退理由にすると学業不振は男女計で 72.3%と高くなっており、経済的理由や人間関係・大学生活不適応を最も重要な中退理由として中退した場合も他の理由が関係していることを指摘している。学業不振・無関心を中退理由として他の理由と関連してあげた割合は、人間関係・大学生活不適応の理由を上げた対象者 (n=276) の中で 80.1%、進路変更を理由としてあげた者の中 (n=250) の中では 76.0%、病気・ケガ・休養をあげた者 (n=124) の中では 57.3%、経済的理由をあげた者 (n=223) の中では 61.4%、妊娠・出産・その他をあげた者 (n=122) の中では 43.4%となっている。つまり、学業不振・無関心以外の理由をあげた者の中で 4 割から 8 割程度の方は学業不振を同時に選択していることがあげられる。このように複数選択を可にして中退理由を分析してみると、中退理由である、学業不振・無関心、人間関係・大学生活不適応、病気・ケガ・休養、経済的理由等は複合的に関係していることが分かる。

1.2.2. 学生における中退行動の特徴把握

本節では大学を中退する学生の行動にはどのような特徴があるのかをインタビューから分析する。評価グリッド法、ラダリング法を用いて実際に都内の大学を中退をした学生 2 人にインタビューを行い、実際に大学を中退するまでの行動を明らかにしていく。

評価グリッド法とは讃井らによって提唱されたインタビュー手法であり、建築学分野や心理学分野等多くの分野で利用されている(讃井と乾 1986; 宇治川, 丸山と讃井 1999; 讃井と乾 1987)。辻村は評価グリッド法の分野やその手法について解説しているが、レポートリー・グリッド法をもとに構築された評価グリッド法は研究以外にも商品開発やマーケティング分野においてニーズを発見・把握するために利用されることも多いと指摘している(辻村 2017)。評価グリッド法の中では、ラダリングとよぶ方法を用いて評価項目に対する上位概念と下位概念を抽出する。人の認知構造における価値観といった心理的、抽象的な上位概念から、より具体的な下位概念を結びつけて、人の心理、行動を階層化、構造化して把握していく。上位概念から下位概念を引き出す場合をラダーダウン、下位概念から上位概念を引き出す場合をラダーアップと呼ぶ。本研究においては中退を選択したという具体的な概念、判断はなぜ行われたのかというラダーアップからはじめ、中退に関係する価値観や対象者が持つ行動価値を把握していく。

本研究では、以上の手法を用いて実際に大学を中退した学生 2 名に対してインタビューを行った。一人目は大学を 3 年次に中退し、その後ゲーム会社に就職した当時 21 歳の男性、二人目は大学を 1 年次に中退し、その後フリーターを行っている当時 20 歳の学生である。期間は 2015 年 11 月～12 月、時間は約 60 分間、面接形式で行った。

1.2.3. 中退学生へのインタビューからの特徴把握

中退した学生へのインタビューから得られた中退に関連する要素としては大きく大学関係、大学内の人間関係、大学外の人間関係の 3 種類があった。大学関係としては、大学学費、成績・単

位、大学でわからないことが多かったことといったものがあがり、大学内の人間関係としては友人関係、先輩関係、教員関係があがり、大学外の人間関係としては恋愛関係、アルバイト関係といったものがあがった。

大学学費に関しては「大学でやりたいことがなく、大学に通っている学費が無駄とを感じる」、「もともと大学に通う気持ちはなかったこと」、「大学に通う気持ちはないため、お金（学費）の無駄さを感じた」、といったものがあげられている。成績・単位に関しては「1年時の春学期は大学に来られていたけど、サボりぐせがついてしまったこと」、「バイトのし過ぎで眠くて行かなくてもいいとおもってしまった」、「アルバイトのほうが大学より上だと感じた」といったものがあげられている。大学の学費や成績といったものだけではなく、大学を退学した学生その人だけではなく多くの関係者（複数アクター）に関連する要素が中退を決意したという部分が大きかった。人間関係に関しては「退学すると在学するでは、お金のかかり方もまた違く、そこでやっぱり母親にも相談しなきゃいけなかったの」、「友達と一緒に大学をさぼってしまって」、「ゼミの教員と（会話）をするのは結構充実していた」といったものがあげられている。当該学生以外の人間との関係が中退・在学に影響を与えていることがわかった。以上のように、大学に関係するもの、大学内外の人間関係がインタビューから抽出された。

次に、大学に入学してから、中退を決意し、実際に中退をするまでの時間経過についての把握を行った。大学に入学するときから意識的に中退を決意していたわけではなく、大学に通う気持ちはなくなり、欠席回数が多くなり、一回の欠席があとを引き、そのなかでアルバイトが楽しくなっていき、結果として中退といった形である。学生はある時点でのイベント等で成長・変化し、そのモチベーションも変化する。1時点における問題があとになって関係してくること、つまり中退は時系列と共に行われるプロセスの結果ということが把握できた。

最後の不確実性とは、学生が大学に関して、そして自分に関してわからないことが多く、中退をする結論にいたったのか不明な部分もあるということである。大学でわからないこととは、大学に通う意味がわからないこと、自分の特徴や、やりたい仕事かわからないこと、卒業ができない・わからないということ等があがった。

以上のように、中退には複数のアクター、関係者が存在していること、中退は複数の状態が遷移するプロセスの結論であること、中退のまわりには不確実なことが多いこと、3つの特徴があげられた。

1. 多次元性：個人だけではなく大学、大学内外の複数のアクター
2. 状態の遷移
3. 不確実性：学生ごとの異質性、不明な部分

1.3. 本論文の構成

本論文は6章で構成されており、図 1.1 に全体の構成として章ごとの関係をまとめた。第1章では高等教育における中退の現状と中退率と偏差値の関係性を述べ、本研究における背景と目的を述べた。第2章では本研究に関係する先行研究として、中退の理論研究、中退の要因に関する研究、中退の防止施策に関する研究、LA・EDM・IR、中退を予測する研究の5つに焦点を当てて述べ、本研究の位置づけを明らかにしていく。第3章では先行研究の位置づけから、本研究の焦点としてリサーチクエスチョン 1. 学生ごとの中退確率を設定することで、個々の学生がどの程度中退しそうな状態なのかを明らかにすること、2. 中退確率を時系列ベクトルで整理をすることで確率の遷移を把握可能にすること、3. 時系列ベクトルをクラスタリングすることで中退までの動的特性を明らかにすること、を明確にする。第4章では入学から中退までの学修状態を把握するために、中退確率を用いて学生個人がどの程度中退をしそうなかを数値化するとともに、中退確率を複数用いてベクトルを作成することで状態の遷移を把握し、類型化を行うことで学生の中退までの学修状態を把握する。第5章では学期内の学修状態の遷移を可視化するために、週毎に算出する予測 GPA を用いて学生の状態を把握し、予測 GPA の複数個つなげてベクトルを作成し、学期内の学修状態の把握と類型化を行う。第6章では本研究の限界と今後の展開を述べる。

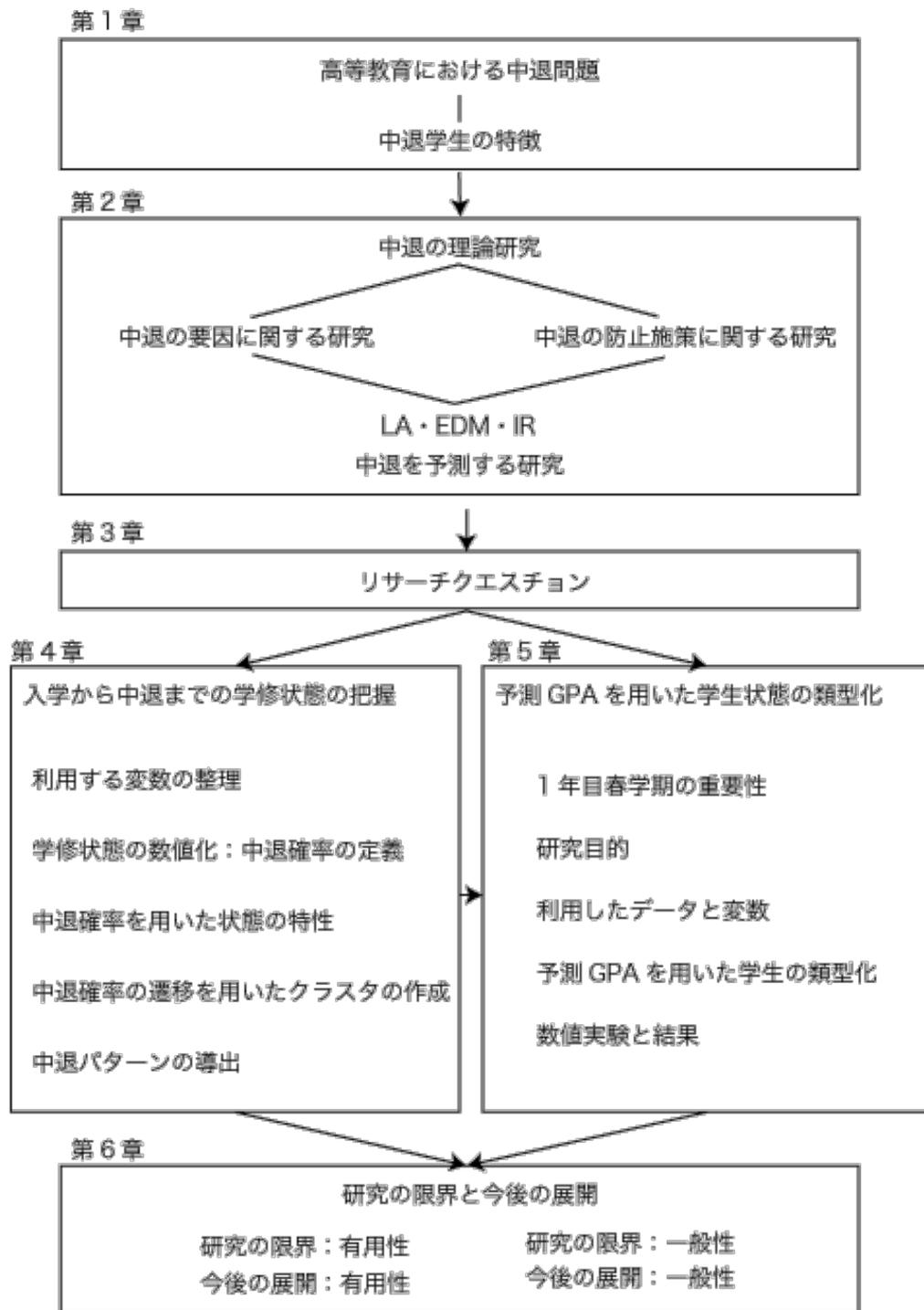


図 1.1 本論文の構成図

第2章 先行研究

本章では中退に関係する研究として、中退理論の研究、中退要因に関する研究、中退の防止施策に関する研究、LA・EDM・IRにおける実践と研究、中退予測に関する研究を述べる。中退の理論研究はアメリカを中心にして発展してきた研究であり、当該学生がなぜ中退をするのかという部分を明らかにするための研究であり、1970年代までの心理学的アプローチと1970年代以降の中退における要因を個人に限定しないアプローチの2つにまとまる。次の中退要因に関する研究は中退理論を基礎に置きながら、どのような要因が中退に関連するのかをまとめた研究である。中退の防止施策に関する研究では、中退理論や中退要因に関連しながら、実際にどのような施策によって学生の中退を防止することができるのかに焦点を当てた研究である。LA・EDM・IRにおける実践と研究は、学内外にある教学データを用いて教学施策を分析し、実践に活かしていく実践・研究である。最後の中退予測に関する研究は、本研究が焦点をあてる研究であり教学データを用いながら、学生がどの程度中退しそうなのか、いつ中退をするのかを分類・予測する研究である。本章では以上の分野の研究をまとめながら、本研究の位置づけを明らかにしていく。

2.1. 中退の理論研究

前章では中退をした学生にインタビューを行いその特徴を把握したが、欧米を中心としてなぜ中退がおこるのかを理論として研究、実践が積み重ねられてきた。1970年代から Tinto, Astin たちによって行われてきた中退がなぜおこるのかといった説明モデルの研究をレビューしながら、関連領域をまとめていくことで、中退における理論研究の流れをみていく。

2.1.1. 1970年代までの中退研究：心理学的アプローチ

中退に関する研究において中退がなぜ行われるのかといった説明モデルの構築が行われたのは1960年代後半以降である。Bergerによれば中退研究は在籍・歩留まり (Retention) として研究されており、1960年代までは歩留まりを意識することはなかったが、1960年代以降から中退防止、そして中退に結びつく理論の確立へとつながっていったと述べている (Berger, Ramirez と Lyons 2012)。1960年代当初、中退に関連する研究は心理学的アプローチを用いていた。心理学的アプローチでは学生の性格、属性、欠点等、個人の要素と中退、歩留まり、学生損耗 (student attrition) の関係を分析しており、心理学を根拠においたものであった。学生個々の要素が中退・歩留まりに影響を与えることになってきたモデルが1970年代までの中退理論研究である。

中退に関する研究は大きく1970年までと1970年代以降に分けられる。Bergerによれば、歩留まり研究の歴史は9つに分割でき、1. 歩留まりの歴史以前 (Retention Pre-History) 1600年代から1800年代、2. 歩留まりの進化 (Evolving Toward Retention) 1800年代中盤から1900年、3.

発展初期 (Early Development) 1900 年から 1950 年, 4. 拡大への対処 (Dealing with Expansion) 1950 年代, 5. 中退防止 (Preventing Dropouts) 1960 年代, 6. 理論の確立 (Building Theory) 1970 年代, 7. 全体のマネージメント (Managing Enrollments) 1980 年代, 8. 地平線を広げる (Broadening Horizons) 1990 年代, 9. 現在と未来のトレンド: 初期 21 世紀, である。最初の 4 つの年代はほとんどが歩留まり研究の前史であり, 後の 5 つの年代で在籍研究, 歩留まり研究は共通の問題として認められ, 実践や理論, 知識ベースが発展してきた。以上のように大きく 1970 年代を境に大学生がなぜ中退・歩留まりを選択するのか, どのように防止するべきなのかといった研究・実践が変わってきている。

2.1.2. 1970 年代以降の中退理論研究

1960 年代までの中退に対する心理学的なアプローチにかわり, 学生と大学という組織との関係性に注目した理論が Spady 達の研究からはじまった (Spady 1970)。彼は中退における最初の理論的モデルとして紹介され, 学生と大学の関係が学生の歩留まりに関係するとした。後述する Tinto と同じく Durkheim が提唱した社会 (の統合) における自殺の理論と学生消耗を関係づけた最初のモデル, 理論であり, この考えは後の Tinto のモデルにもつながっている。学生と大学の環境の間のインタラクションによって中退プロセスを説明し, 学生と大学 (学部, 教職員, ピア等) の間のインタラクションは大学に対する学生の統合 (integration) のレベルを決定し, 在学に影響を与えるという仮説を検証している。彼によれば学生の中退・在学は 2 つのシステムに影響を受けているとし, 1 つ目は成績と知的な発展であり, 2 つ目は社会における規範的同一 (normative congruence) と友人のサポートであるとしている。

Tinto は Spady の大学組織と学生間のインタラクションを基礎において大学出学モデル (Institutional Departure Model): 学生統合モデル (Student integration model) を提案した (Tinto 1975)。彼は Van Gennep の社会人類学的な移行の概念 (部族社会における通過儀式: separation, transition, incorporation) を学生の大学社会への統合プロセスに適応している。大学に入る段階, つまり初年次学生は家族や高校等, 大学とは違う価値観を持っているからコミュニティから分離 (separation) される。大学に入るときは大学以前に所有していた価値観を持っているコミュニティからの分離が必要であることを指摘している。新しい価値観を獲得する前に, 彼らの昔のコミュニティからの脱社会化 (disassociating) が起こり, その段階を移行 (transition) と呼んでいる。その後, 統合のプロセスが始まるとしている。

Tinto によると, 大学はアカデミックとソーシャルの 2 つのシステムで構築されているとしており, 学生はこの 2 つのシステムに統合される。大学への統合は成績と知的な発展によって測ることができ, 一方で社会への統合は学生のピアや教職員とのインタラクションによって測ることができる。学生の入学前属性 (家庭の背景や能力, 要素等) は大学での最初のゴールや達成を形作り, そのゴールによって大学・社会への統合の経験は弱くもなり, 強くもなるとしている。そして変化したゴールや達成によって在学・退学に影響を与える。さらに家族や仕事の達成等の外

部の達成レベルは次のゴール・達成に影響を与えている。

Aljohani によれば Spady, Tinto を含めて中退理論研究は構築されてきていると主張する (Aljohani 2016)。彼は中退理論研究の中で大きく下記の 6 つのモデルが重要だと指摘している。6 つのモデルとは、1. Tinto の組織出学モデル (Institutional Departure model) (Tinto 1975, 1993), 2. Bean の学生損耗モデル (Student Attrition) (Bean 1980, 1982), 3. 学生-教職員のインフォーマルなコンタクトモデル (Student-Faculty Informal Contact Model) (Pascarella と Terenzini 1980), 4. Astin の学生巻き込みモデル (Student Involvement Model) (Astin 1999), 5. 伝統的ではない学生の損耗モデル (Non-traditional student attrition model) (Bean と Metzner 1985), 6. 学生歩留まり統合モデル (Student retention integrated model) (Cabrera, Nora と Castaneda 1993) である。図 2.1 は中退におけるモデル研究の流れの概略をあらわしたものである。

2 番目の Bean 1980, 1982 による学生消耗モデル (Student Attrition Model) は (Tinto のモデルと一貫性があるとしているけれども) Tinto や Spady が用いた学生損耗において Durkheim の自殺理論ではなく、Price が 1977 に提唱した仕事における turnover (離職) の理論を退学のモデルに用いた (Bean 1980, 1982)。Bean によれば学生の満足度は組織的な決定要因 (organisational determinant) によるものが大きいとしている。彼は離職における pay 変数に変わって、学生 GPA, 発達, 組織の質, 実際の価値を教育・中退に関連する指標として用いた。目的変数としての中退, 説明変数としての満足と大学との約束 (institutional commitment), 組織的決定要因, 背景変数を Bean のモデルは含んでいる。学生が中退する/在学するに関係するものとして、意図, 背景, 組織, 環境, 姿勢や結果に変数を分類している。

3 番目の Pascarella による学生-教員インフォーマル接触モデル (Student-Faculty Informal Contact Model) である (Pascarella と Terenzini 1980) は大学における社会的、そして大学的なシステムへの統合において教員とのインタラクションは重要であると述べている。Pascarella たちのモデルは Spady と Tinto と同様に大学や大学教員とのインタラクション、関係性を重視したものである。その中で特に学生と教員のクラス外でのインタラクションとその教育効果が重要だと述べている。

4 番目は伝統的ではない学生による損耗モデル (Non-Traditional Student Attrition Model) であり、Bean たちによって提唱された (Bean と Metzner 1985)。伝統的ではない学生を大学外から通学してくる学生とし、伝統的ではない学生に対しては別の学生損耗モデルを適応する必要があると述べた。通学してくる学生は大学内の社会より、外部の環境に影響を及ぼされるという仮説に基づいている。Bean たちが提唱したこのモデルはアカデミックの成績、パフォーマンス、出学の意味、背景、そして環境からの変数によって学生損耗モデルが成立している。

最後は Cabrera たちが提唱した学生歩留まり統合モデル (Student Retention Integrated Model) である (Cabrera, Nora と Castaneda 1993)。学生歩留まり統合モデルは Spady, Tinto における学生統合モデルと Bean たちが提唱したモデルを統合したものになっている。Cabrera モデルにより中退理論研究は結論付けられると Aljohani は述べている。

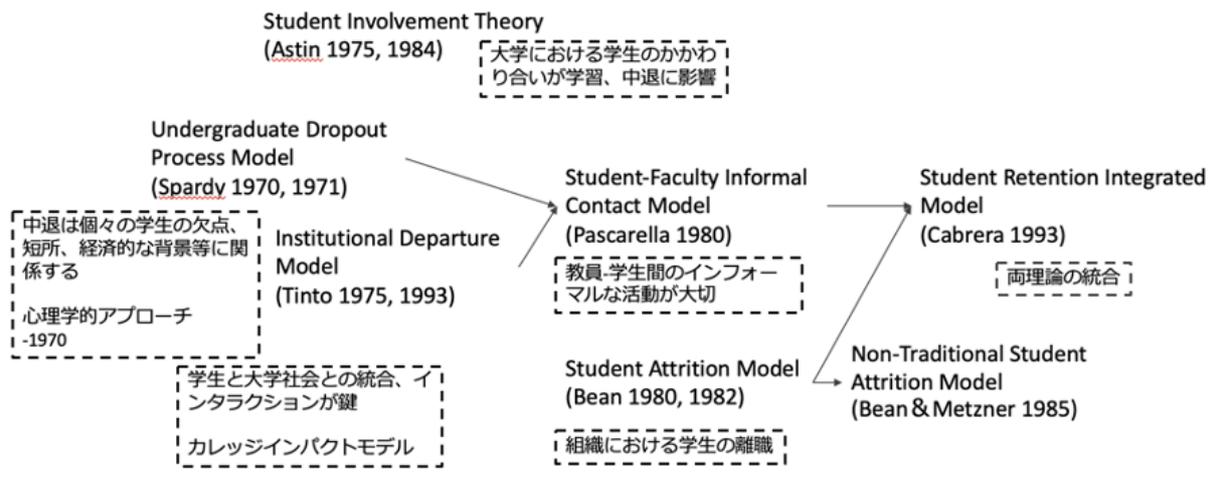


図 2.1 中退理論研究の流れ

また近年では、Lisa Bäumle らは中途退学をする場合、また専攻を変える場合には5つの段階があるとしている(Bäumle, Grunschel と Dresel 2021)。非適合の認識、辞めたい/変わりたいという考え、熟考、情報探索、最終決定、の5つの段階であり、そのフェーズを、中途退学と専攻変更で分離することが可能かどうかを構造方程式モデリング (SEM) を利用して考察している。中退と専攻変更は分けて考えることが妥当であり、5 フェーズの順序は重なっているが妥当であることを指摘している。

日本においては、丸山によって大学退学についてまとめられたものが先駆的な研究となっている(丸山 1984)。丸山は教育社会学という研究背景において、欧米の中退モデルが日本にあてはまるのかを学校基本調査データを用いて明らかにした。欧米で展開された大学退学の説明モデルを教育達成モデル(個人的属性が退学に影響を及ぼす)、カレッジ・インパクト・モデル(大学と学生のインタラクションが影響を及ぼす)、チャータリング・モデル(カレッジインパクトモデルの大学組織が個々の学生に影響を及ぼす時に、学生の変化に及ぼす影響力は大学が外部社会から与えられる意味に左右される)の3分類に分け、退学者数、退学率を目的変数、学生/教員比、平均講義規模、本務教員比、学部規模を説明変数として用い、大学の学部単位で回帰分析を用いて説明を行っている。退学率を目的変数とした場合には、学生/教員比、学部規模のそれぞれが大きくなると、退学率は低下する傾向にあることを指摘している。教員と学生の相互作用を意味する学生/教員比は退学者数、退学率と関係があり、その低下をもたらしているため、この部分に関してはカレッジ・インパクト・モデルが適切である。一方で規模に関しては大学、学部の規模が大きくなると退学者数、退学率が低下する。この意味ではチャータリング・モデル(キャリアモデル)が正しい。日本における大学の退学者数、退学率については学部の大規模化は決してマイナスではなく、必ずしも Small is beautiful とは言えない現状があると指摘している。

このように、中退がなぜ行われるのかといった理論、モデルの研究は欧米を中心になされており、中退に至るまでの心理学的な側面、大学との関係性等社会的な側面についての研究がなされ

てきている。

2.2. 中退の要因に関する研究

本節では前節で述べた中退理論における要因・変数に焦点をあてる研究をまとめる。本節では大学の偏差値や景気等、大学全体に関係するものを中退に関連するマクロ変数、学生個人に関係する変数を中退に関連するミクロ変数として議論する。マクロ変数とミクロ変数をまとめながら中退防止、中退予測に対して考慮すべき変数を整理する。

2.2.1. 中退に関連するマクロ変数

前節で述べた中退理論から導かれることは、中退に関連する要因は学生が属する環境や人間関係、大学等に関連するものと学生個人に関連するもの2つに分類できそうである。白鳥は前者の変数群を中退に関連するマクロ変数、後者の学生個人に関連するものを中退に関連するミクロ変数と呼ぶ(白鳥 2016)。中退におけるマクロ変数には経済的な要因、大学の規模の大きさ、大学の合格率、偏差値、教員学生比などがあげられる。

中退におけるマクロ変数を導く研究としては Manski らの研究があげられる。Manski らは大学からの中退は経済活動であることを指摘し、大学を中退することは大学教育の便益と費用を勘案しながら行われることだと述べている(Manski 1989)。大学在学における便益は正規の職につきやすいことや高卒者より大卒者の方が平均的な収入が高いこと等があげられ、費用としては大学の学費等をあげることができる。

また、丸山によれば大学の規模が小さいほうがより中退率は高く、退学率に関連する変数として、学生/教員比、学部規模をあげている(丸山 1984)。清水は大学レベル・学部レベルにおいて、清水は読売新聞の大学実力を用いて調査を行い、偏差値・一般入試比率、定員充足率と中退率の関係性を分析している(清水 2013)。清水は偏差値、一般入試比率、定員充足率が高いことと中退率が高いことには関係性があると述べている。また、2021年に清水は理工学系と社会科学系の学部を比較し、中退率と偏差値の関係を分析している(清水 2021b, 2021c)。退学率、就職率、決定率(就職率+進学率)に関しては偏差値の影響が大きいことをあげている。ST(教員・学生)比率に関しては、社会科学系では有意ではなく、理工学系では有意であったことを述べるが、理工学系においてもST比率は実質的な影響が少ないことを指摘している。姉川は大学における学習環境と中退率の関係を読売新聞社からだされていた大学の実力や朝日新聞社からの大学ランキング等のデータを利用して分析している(姉川 2014)。学習環境としては大学教員と学生の数の比率、奨学金の給付等をスコアにした生活支援スコア、PBLやゼミ研究室等の設置等をスコアにした学習支援スコア、大学入学時点での成績として偏差値、学生一人あたりの図書貸出数を用いた学習意欲等をデータとして利用している。その中で、大学入学時点での学力が高いほど退学率は低い、学生の学習意欲が高い大学ほど退学率は低い、といった2つの仮説が成り立ち、図書貸出数と学

生教員比率が退学率の抑制に影響をもたらすと述べる。同じ偏差値帯の大学であったとしても学習意欲を高める取り組みをしている大学とそうではない大学では退学率に違いが出ることを指摘している。

立石らは標準修業年限内では卒業できない学生を非標準型として、中退と留年を統合的に扱い、中退が多い/少ない学部レベルの特徴を分析している(立石と小方 2016)。読売新聞の大学の実力におけるデータを用いて、中退率が高い/低い、留年率が高い/低いという4分類に対して、中退、留年共に高い群と中退、留年共に低い群と比較して初年次退学率、ST比、偏差値は有意に差があることを指摘し、教育・学習支援の取組の効果は限定的であることを結論づけている。

このように経済的な指標や大学の規模、偏差値、ST比等は中退率に関係しており、本研究では大学全体に関係する中退におけるマクロ変数群としていく。マクロ変数群は大学全体、学部全体、複数の大学が関連する国全体といった規模感の変数であり、これらの数値によって大学、学部ごとの中退率が計算することができる

2.2.2. 中退に関連するミクロ変数

前節では大学中退におけるマクロ変数について述べたが、個々の学生行動や属性が中退に関連することも多い。中退に関連し、個々の学生毎に違う変数群を本研究では中退に関連するミクロ変数と呼ぶ。

白鳥は前節で述べた Tinto の中退理論をもとに中退に関連するミクロ変数群を整理した(白鳥 2021)。中退に関連する変数を高大接続に関連する変数と大学に入学してからに関連する変数という時間軸と教学の変数, 社会的変数, 経済的変数という変数の種類の2つに分けてまとめている。図 2.2 は中退に関連するミクロ変数を Tinto のモデルと共にまとめたものである。横軸を時系列、



図 2.2 Tinto の変数群と大学で取得しているデータ

縦軸を変数の種類として、大学で取得できるマイクロ変数を配置してある。入試の段階で取得できる変数としては家族関係、スキルと能力、入学前の学習経験がある。家族関係とは家族や学生が所属するコミュニティの背景的な要因であり、社会のステータスや保護者の学歴、コミュニティのサイズ、世帯年収等で計測することができるとしている(Tinto 1993)。スキルと能力とは知的、社会的スキルであり、学習習慣や欠席日数、資格等で判断でき、入学前の学習経験とは高校の評定平均や高校偏差値、課題等で計測できるとしている。

個人に紐づく変数である家族関係、家族背景の変数である保護者、家計支持者の学歴が大学入学後の学生生活に関連する研究は First Generation (大学第 1 世代) の研究としてアメリカを中心に行われてきた。日本においても武内らが親の学歴等の社会階層は大学の学力や友人関係等に影響を及ぼすことを指摘している(武内 2008)。武内によれば、親の社会階層、学歴、教育期待は、当該学生の属性、出身高校等に関連し、それ以降の大学での経験、大学後の職業やライフスタイル等に関連することを学生の社会化のモデルとして提示している。また、学生の入学以前の特性は大学の学力、知識だけではなく、学習習慣、読書、交際・恋愛等にいたるまで大学にはいつても持続することが指摘されており、大学という新しい社会に入るときの準備をする必要性を指摘している。他にも河野はアメリカと日本における第 1 世代研究について NCES (National Center for Education Statistics) からの実態をレビューしており、4 年制大学における第 1 世代の学生は GPA が低く、リメディアル教育の受講率は高かったとしており、大学に入るまでの準備により他学生との差が縮小することを述べている(河野 2003)。

家族関係以外のマイクロ変数の代表的なものは性別、民族、出席、修得単位数、成績等をあげることができる。田尻はビジネス系大学における学修履歴と中退の関連の分析を行い、中退学生に関する変数群を導出している(田尻 慎太郎と白鳥 2013)。彼は女子学生の方が 4 年間で中退をする確率が低く、4 年目以降はその確率が逆転することや、入学時学力テスト、成績、修得単位数なども中退に関連する変数群としてあげている。また、高橋らは琉球大学における退学、休学、除籍、留年等と初年次前期の GPA との関連性を分析し、標準修業年限内に卒業した学生や留学した学生と比較して、退学、休学、除籍、留年となった学生のいずれにおいても初年次前期の GPA が有意に低いことをあげている(高橋、藤本と西本 2019)。他にも、大河内らは 1 学年春学期の成績が以後の学期の成績、未進級、退学に大きい影響を与えることを述べている(大河内と山中 2016)。

表 2.1 は 2.2.1.であげた中退に関連するマクロ変数と 2.2.2.であげた中退に関連するマイクロ変数をまとめたものである。白鳥は中退に関連するマクロ変数とマイクロ変数を用いて、どのように IRer (Institutional Researcher) としてどのように中退防止に関連する作業を行うべきなのかを大学固有、高大接続アプローチ、教学アプローチとしてまとめて報告している(白鳥 2021)。大学固有では中退に関連するマクロ変数を説明変数として用いて大学の中退率を算出することを表している。高大接続アプローチは中退に関連するマイクロ変数の中で特に入学前に判明するものを用いて、学生が中退するかどうかを判別する分析である。教学アプローチでは中退に関連するマイクロ変数の中で入学後に判明するものを用いて、学生が中退するかどうかを判別した分析である。

表 2.1 中退に関連する変数の整理

	大学固有	高大接続アプローチ	教学アプローチ
説明変数	偏差値、ST比率、留学生比率、学生数、一般入試比率、施設の有無、大学住所、等	評定平均値、高校欠席日数、性別、志望順位、家計、等	大学成績（GPA）、単位数、資格等
目的変数	退学率	退学する/しない	退学する/しない
変数の説明	大学に紐づく変数	個人に紐づく変数（入学前に決定する要素）	個人に紐づく変数（入学後に決定する要素）
分析手法	回帰分析：線形回帰等	回帰分析：ロジスティック回帰等	回帰分析：ロジスティック回帰等
期間	長期	中期・短期	中期・短期
特徴	コントロールが非常に難しい	コントロールが難しい	コントロールが容易

本節では中退に関連するマクロ変数として大学固有の偏差値、ST比率、留学生比率等をあげた。マクロ変数は個々の学生の中退に関連するよりも、大学全体、学部全体の中退率に影響を与えていくものだった。次に中退に関連するミクロ変数として、大学前に判別する変数・要因と入学後に判別する変数・要因をあげながら、どのような変数・要因が個々の学生の中退、成績に影響を与えるのかをまとめた。ミクロ変数はマクロ変数とは違って大学に通う学生個々に紐づく変数であり、大学入学前の属性は大学入学後のゴール等に関連することがあることも Tinto のモデルを中心として述べた。

2.3. 中退の防止施策に関する研究

これまでに述べた中退理論・モデルや中退の要因に関する研究をもとに、Alan Seidman や Vincent Tinto たちはこれまでの中退理論研究からどのような施策が中退を防止するのかという理論から行動施策への転換を促している。中退を防止するために必要な式として Center for the Study of College Student Retention の Executive Director である Seidman は下記の中退防止の式を提唱している (Seidman 2012)。

中退防止 (Retention) = 早期発見 (Early Identification) + (迅速 (Early) + 集中的 (Intensive) + 継続的 (Continuous)) な介入

この式では、中退防止を行うためには、早期発見を行い、迅速的で、集中的で、継続的な学生への介入が必要なことを述べている。以上の式を実装する方法としては、個々の学生評価を行うことで早期発見を行い、支援が必要な学生に対しては診断を行い、それに応じてプログラムを調整する事が必要である。プログラムを行うとともに、その学生を観察しながら評価を行い、さら

にプログラムを変更していく，という介入を続けていくことをあげている。

また，Tinto は大学が中退防止のために取りうる行動として，期待 (Expectations)，支援 (Support)，フィードバック (Feedback)，関与 (Involvement) の4つをあげている(Tinto 2012)。中退を防止するためにはオリエンテーション等で学生に期待をし，観察，評価，早期発見をしながら学生にフィードバックを行い，カリキュラム，教授法，課外活動等で関与をし，教学・社会的・費用のサポートをすること，つまり期待，支援，フィードバック，関与の4つのアクションを組み合わせることが中退を防止するために大学が行うべき活動としている。

Seidman と Tinto は共にリスクが高い学生を早期発見し，関与・介入を行い，教学・メンタル・費用・ソーシャルに様々なサポートを行い，それを続けることが大学の活動であるとし，中退防止施策のモデル化を試みている。以上のようにアメリカにおいては中退防止施策をこれまでの研究で積み重ねた理論の上でどのような施策を行うべきかという実践的な中退防止施策のモデル化が行われてきている。

日本では面談を中心とした個々の大学での対応に関する研究が多いが，日本における中退防止施策をまとめている論文としては藤原や岩崎があげられる(藤原，富永と押味 2013; 岩崎ほか 2016; 岩崎 2015)。藤原らは37大学中25大学の中退防止施策をまとめたものとして，教員による対策，学生相談機関による対策，その他の対策と3種類にまとめ，今後の対策として，就学支援，全学的支援，健康支援，生活支援，経済支援，就職支援などをまとめている。現状として行われている対策としては学生指導等が多いことがあげられている。教職員による対策で多いものが面談等の個別対応である。

表 2.2 は藤原らと岩崎の論文内にある中退防止施策(藤原，富永と押味 2013; 岩崎 2015)に対して施策を行う側と施策を受ける側の2軸で分類したものである(白鳥 2020)。施策を受ける側の軸は，中退リスクが高い個人への対応，学生全体に対する対応の2つに分けた。また，施策を行

表 2.2 中退防止施策のまとめ

		施策を行う側	
		個人 (教職員)	組織
施策を受ける対象	個々の学生	<ul style="list-style-type: none"> ・ 教学相談 (教学) ・ コミュニケーション支援 (社会・生活) ・ 健康支援、メンタルサポート (社会・生活) ・ 学費相談 (経済) 	<ul style="list-style-type: none"> ・ アラートシステムの構築 (教学、社会・生活、経済) ・ 教学相談のシステム化・ルール化 (教学) ・ ピアカウンセリング (社会・生活)
	学生全体	<ul style="list-style-type: none"> ・ 担任制度 (教学、社会・生活) ・ オリエンテーション、キャリア教育 (教学) ・ 出欠の徹底等、授業における工夫 (教学) ・ 学生と教職員、友人とのコミュニケーション・関係構築支援 (社会・生活) ・ 奨学金制度の連絡 (経済) 	<ul style="list-style-type: none"> ・ 学生データの共通化 (教学、社会・生活、経済) ・ 学生の居場所づくり (社会・生活) ・ 奨学金制度、学費減免制度の改善 (経済) ・ カリキュラムの改善 (教学) ・ 出欠管理システムの改善 (教学) ・ 学部規模の調整 ・ 平均講義規模の調整 ・ 兼務/本務教員比の見直し

う対象の軸は個人（教員・職員等）、組織の2つに分けている。また、それぞれの施策には Tinto が提唱した中退防止施策の3つのタイプの支援（教学、社会・生活、経済）をあてはめている（Tinto 2012）。ただし、全学的な支援に関してはタイプを割り当ててはいない。

左上の個人（教職員）×個々の学生への施策は教学相談、コミュニケーション支援、健康支援・メンタルサポート、学費相談のどれにしても面談という形になる。個人が抱えている課題はそれぞれで、教学的な課題、人間関係的な課題、経済的な課題と複合的に重なっているため、教職員がどの支援が必要なのかを判断して複合的に問題解決をしていく必要がある。左下の個人の教職員×学生全体の施策は、個々の教員が複数の学生・学生全体に行う施策である。授業の改善やオリエンテーション、奨学金情報の連絡等をあげることができる。右上の組織×個々の学生への施策、右下の組織×学生全体としては、個々の学生、学生全体に対して個々の教職員が行っていたことを組織的に行う施策である。3タイプの支援共に組織化した施策になる。右下の組織×学生全体の施策の中には3タイプの支援とは別に、大学全体の改善施策として学部規模の調整等をいれてある。

左上の個々の教職員が学生個別に対する施策の一つである学生相談に関する研究は日本においても藤川や川崎、中島等、多くの事例、積み重ねがある（藤川 2018; 川崎ほか 2014; 中島 2012, 2013）。個別対応は大学不適應に対するカウンセリング等の心理学的なアプローチにおいて多く用いられているが、窪内は日本の大学中退における心理学的アプローチを参照しつつ、どのように学生相談を行うべきかの検討を行っている（窪内 2009）。教員と学生相談担当者が連携し、大学への適應をよりスムーズに行うために学生に働きかけることを提唱している。このように個別対応としての中退防止施策は日本においても学生相談を中心として多くあげることができる。

一方で、教学支援や就学支援、カウンセラーによる相談等の個別対応には、教員の負荷が大きすぎるという問題があげられている。藤原は休退学における対策は個人レベルでの対応から組織レベルの対応に移行する必要性を述べ、教員、学生相談機関、事務部門等による個々の支援をつなげ、体系的に全学的な支援が必要なることを主張している（藤原、富永と押味 2013）。中退防止施策は個々の施策から、表 2.2 でいう右側にある組織としての施策が求められてきている。

個別の対応だけではなく組織的な対応が必要となってくる背景には、東京大学の調査でもわかるように、中退要因の複合的な理由をあげることができる（東京大学 2016）。中退の理由が経済的な要因を学生が選択したとしても、その背後には人間関係の不適應、大学との不適應、教学の準備不足等、複数の要因が存在していることが多い。中退しそうな学生の課題が複合的であるために、中退防止の対処的には個々の面談等の個別対応になることが多い。例えば、中退要因となるデータとしては個人に紐づくUPI、身体的問題、教学的問題、友人関係、経済的な問題と様々であり、またそれらが複合的であるために一律的な対応というよりは個々の面談等による対応に頼る必要が出てくることはやむを得ない。このように複雑な中退要因を解決するためには、面談を中心とした個々の学生ごとの対応と複数の組織が連携した中退防止施策を行う必要性がでてくる。

しかしながら、面談を中心とした個別対応は人の経験によるものが多くなることと、教職員に過度の負担がかかることになる。以上の問題を解決するためには、施策を組織化していくことが

中退施策の方向性の一つである。組織的な施策を行うことで、個々の教職員の負担を減らし、中退防止施策を継続的に行うことができるようになる。一方で、組織的な施策を行うためには、複合的なデータ統合を行う必要があること、データを統合したあとに、共通の施策の策定や組織としての評価の難しさ、短期に結果が出にくい等の課題が続く。

2.4. LA・EDM・IR

本節では中退防止に関連する LA (Learning Analytics), EDM (Educational Data Mining), IR (Insitutional Research) の分野をまとめ、研究の位置づけを確認する。前節までに中退を考える際に、中退の理論に関する研究、中退に関連する要因・変数に関する研究、中退防止施策に関する研究とこれまでに行われてきた研究テーマを中心として述べてきたが、大学に存在する様々なデータを分析し、教学に活かしている分野が LA と EDM であり、その大学等の組織にあるデータを用いた大学の意思決定を支援する分野が IR である。本章ではこれらの分野を概説しながら、この分野におけるデータをもとにした中退防止施策とのつながりまで述べる。

2.4.1. LA と EDM

教育・学習におけるデータ活用について盛んに研究が行われている分野がラーニングアナリティクス (LA : Learning Analytics) とエデュケーションナルデータマイニング (EDM : Educational Data Mining) である。近藤によれば LA, EDM のどちらの分野も 2010 年前後から発展し続けていることを述べている(近藤 2020)。Siemens らは LA, EDM の類似点と相違点を確認しながら、2つの分野のコラボレーションを唱えている(Siemens と Baker 2012)。LAK (Learning Analytics and Knowledge) は LA と同様の言葉として用いられることが多いが、LAK, LA, とともに国際的な動きは 2010 年にカナダのバンフで行われた会議からはじまり、2011 年の SoLAR (Society for Learning Analytics) の設立につながった。SoLAR はラーニングアナリティクスを、学習とその環境の理解と最適化のために、学習者と背景についてのデータを測定、収集、分析、そして、その結果を報告することと定義している。EDM は 2005 年にアメリカのピッツバーグで開催されたワークショップをはじめに、2008 年のカナダのケベックで行われた第 1 回の EDM カンファレンス、2011 年の IDEMS (International Educational Data Mining Society) の発足につながっている。IDEMS はエデュケーションナルデータマイニングとは、教育現場から入手できるユニークなデータを探索するための方法を開発し、学習者とその環境をよりよく理解する新しい分野であるとしている。

LA と EDM の両分野ともに教育に関するデータを利用して教育の実践と研究を発展させており、多くの研究領域が一致している。一方で、Siemens たちは 2つの分野の違いを明らかにしている。1つめの違いは何を重要視するかである。LA は人間の判断に活用することが鍵であるとし、自動化した発見はそのための道具だとしているが、EDM は自動化した発見が鍵であり、人間の判

断への活用はそのための道具だとしている。2つ目の違いは適応分野である。EDM モデルは知的チュータリングシステム等の自動適応 (adaptation) の基礎として利用されるが、LAK モデルは教員や学習者への情報提供や助力 (empowerment) として利用される。3つ目の違いは還元論的アプローチ (reductionistic approach) と全体論的アプローチ (holistic approach) の違いである。EDM は事象を還元し、個別の要素を分析する還元論的アプローチをとることが多いが、LAK は複雑性をもった全体としてシステムを理解しようとする全体論的アプローチをとることが多い。このように LA, EDM 共に共通点としては教育のデータを利用して学習者・環境・教育者等を支援、理解しようとしていることがあるが、方向性が何を重視するのか、適応範囲、アプローチと違っている。

2.4.2. IR

インスティテューショナル・リサーチ Institutional Research (以下 IR) は Saupe の定義によれば高等教育機関が大学の計画の決定、政策決定、意思決定を支援するために情報を提供する調査、分析の活動であるとしている (Saupe 1990)。IR における国際的な組織として Association for Institutional Research (AIR) があり、1966 年に非営利の 501(c) 団体として作られた。AIR は IR における調査や年次大会等を行っている。日本においても 2012 年に IR コンソーシアムが任意団体として設立され、2018 年には非営利型の一般社団法人となっている。IR のコンソーシアムでは学生調査、指標の作成、IR 人材育成のワークショップ等が行われてきており、2022 年 9 月 6 日の段階で国公私立 65 大学が参加している。他には日本インスティテューショナル・リサーチ協会 (JAIR) は IR に関する研究・実践活動の質向上に寄与することを目的とした団体であり、大学情報・機関調査研究会 (MJIR) や International Conference on Data Science and Institutional Research (DSIR) を毎年開催している。

IR においても大学内外にあるデータを活用し、分析していくという方向性は LA, EDM と同様に共通である。日本 IR 協会が出しているガイドブックによれば IR に共通する活動として、リサーチクエスションから、必要なデータの収集、データの分析、分析結果の可視化、報告とされている (日本 IR 協会ほか 2022)。また、同じように中井らは IR の標準的なプロセスとして調査設計、データ収集、分析前準備、分析、情報提供という形で構成されるとしている (中井、鳥居と藤井 2013)。大学内外におけるデータを活用して、高等教育機関における意思決定を支援するという IR の方向性はあるにせよ、LA・EDM と共にデータを収集・分析し、価値を生み出していく部分は類似している。

IR において教育改善に関する機能に焦点をあてたものが教学 IR である。教学データを利用し分析・提案を行う教学 IR と前述している LA・EDM の分野は密接に関連している。松田らは教学 IR において扱う対象は教育機関全体や学部・学科の意思決定支援であることが多いのに対して、LA・EDM では各授業やコースの改善を対象にしていることが多いと指摘している (松田と渡辺 2018)。LA における分析のレベルは、従来のマイクロレベル (授業・学習者レベル) から、メゾ

レベル (機関レベル), マクロレベル (地域・国家・国際レベル) という階層に分かれていること, そしてそのレベル間が連携していることを述べている (Shum 2012).

前節までに述べてきた中退の要因研究と中退防止施策の研究は大学内外におけるデータをミクロレベルからメゾレベル, マクロレベルまで取得して行うようになってきている. 本節で述べた LA, EDM, IR の中のテーマの一つが中退, 中退防止, 中退予測である.

2.5. 中退を予測する研究

2.5.1. 中退を予測する研究の分類

2.3.節で述べたとおり中退しそうな学生や, 出席できなそうな学生, つまりハイリスク学生を早期に発見して, 継続的な介入施策につなげていくことが効率的である. そのために行うのが中退予測, ハイリスク学生の予測である. 中退予測の研究はラーニングアナリティクスや Educational Data Mining の分野が出現し, 大学内のデータが蓄積, 活用されていくなかで行われてきた. 本章では LA・EDM の分野で IR データを活用した予測がどのように行われてきたのかを述べる.

Arto Hellas らは学生のパフォーマンスを予測する (Predicting Academic Performance) レビュー論文の中で, 中退予測をパフォーマンス予測の一つの研究として位置づけている (Hellas ほか 2018). Hellas らは学生のパフォーマンス予測研究をレビューしていく方法として, 1. 予測したい変数 (パフォーマンスの種類・定義) はなにか, 2. 入力する変数はなにか, 3. 手法はなにか, の3つに注目して行っている.

1. 予測したい値 (目的変数) としては Scopus, IEEE, ACM から抽出したパフォーマンス予測に関連する 357 の論文を分析する中において, 88 論文 (24.4%) はコースの得点・成績を用いていることを述べ, コースの得点・成績が一番多いテーマとなっている. 中退予測に関連するものとしては, プログラムからの卒業/歩留まりは 48 論文 (13.4%), コースの歩留まり/中退は 20 論文 (5.5%) であった. 予測研究の中で中退予測に関するものは 20%弱であることがわかる.

次の入力する変数 (説明変数) として 29 のカテゴリに分けて分析しているが, コースのパフォーマンス・成績が 141 論文 (13.1%), 事前コースの成績が 139 論文 (12.9%), 対象コースへの参加 (Engagement) が 113 論文 (10.5%) の割合が高いことを指摘している. 他にもデモグラフィック属性, 性別, 自己効力感などの心理学的要素, 作業時間等のログ等があげられている. 指摘としては少数の変数群を利用しており, 以前のように成績, 人口統計データ, 心理的データ等の複数データ群をまとめて分析した事例は 2010 年代前半とは違い, あまり見られなくなっていることを指摘し, 少数の要因に限った研究がなされていることを述べている.

最後に, 利用されている手法としては統計的手法: 線形モデルが 110 論文 (31.3%) と多く, 分類: 確率的グラフィカルモデルが 80 論文 (12.88%), 分類: 決定木が 74 論文 (11.9%) と次に多くなっている. 他にも統計的手法としては相関関係を利用したもの, 潜在変数モデルを利用し

たもの、分類としてはニューラルネットワーク、SVM、ランダムフォレスト等を利用したもの、クラスタリングやマイニング等が手法としてあげられている。

彼らが把握している予測研究の課題としては、方法論的課題、データの課題、ユーザビリティ的課題の3つを指摘している。1つめの方法論的課題では、データ分析の方法が統計分析から機械学習、深層学習へといった新しい機械学習、人工知能で利用されている手法に移り変わってきていることであり、2つめのデータの課題とは、データ収集の方法、データの量と粒度が変化していることを指摘している。データ収集の方法は紙で取得していたアナログ的な方法からセンサーでの取得やオンラインでの調査のような形に変化してきていることである。また、量としては単一学期、単一コースから、複数学期、複数コースのデータになり、単一のデータだけではなく、複数コンテキストのデータで一般的な妥当性を確認するべきだと主張している。また、データの粒度ではコースの成績等の粒度が粗いデータからLMSなどの測定値等の粒度が細かいデータへの移行を指摘している。最後のユーザビリティ的課題、アプリケーション的課題では、予測で作成したモデルをより一般化するということが、データを使って何ができるという研究へと移行していることを指摘している。

さらに、モデルのリスクや倫理的な配慮についても指摘している。モデルのリスクでは、予測研究がうみだした理想的な学生は偏ったデータを利用している場合が多く、その理想の学生モデル自体を利用することによって、そのバイアス、偏りを助長することになっている形になることを指摘している。倫理的な配慮とは性別データの利用に関して、許諾が取れていない2次的なデータ利用、参加者の匿名化と言った問題があることを指摘している。この問題を解決するためには査読コミュニティにおいて匿名性等を審査し、倫理的な問題を校了するために適切な委員会の審査を受けることを養成している。

2.5.2. 代表的な中退予測研究の種類

以上の予測モデルの分類を用いながら、予測研究について議論をすすめる。近藤・畠中(2016)は当該年度より前のデータ(性別、学部、入試区分、出席率等)を用いることで3年次当初における中退予測が可能だと述べた(近藤と畠中 2016)。Hellasらの分類によれば、目的変数としては3年次までに大学生が中退するかどうか、説明変数として1年次春学期終了時までの性別、学部、出席率等、当該年度等のデータを利用している。利用したモデルとしてはロジスティック回帰モデル、サポートベクトルマシン、決定木モデル、RBF(Radial Basis Function)ネットワーク等の手法を用い、モデル間の予測率の差異を検討している。結果としては1年次春学期終了時までのデータを用いることで、3年次までに中退する学生の50から60%を予測できると述べた。

また、欧米では中退(Dropout)の代わりにRetention Rate(在学率)もよく用いられるが、BINGHAM and SOLVERSON(2016)は親の学歴や民族によって1年次から2年次の在学率が異なることをロジスティック回帰モデルによって説明した(BinghamとSolverson 2016)。また、Guarinら(2015)は入学時のデータと入学後の成績を変数として用い、ナイーブベイズと決定木

を用いて成績不振が原因で学籍を喪失する学生を分類した(Lopez Guarin, Guzman と Gonzalez 2015). 彼らは高校の種類や入試の形式, 年齢, 性別, 民族等の入学時のデータのみを利用し学籍喪失を予測した形と入学時のデータに加え大学入学後のデータを加えた形と 2 種類の実験を行った. リスクのある学生を誤ってリスクのない学生だと予測するコストはリスクのない学生をリスクがある学生だと予測するコストより高いとして, 分類器にコストを設定した meta cost アルゴリズムを用い, 精度に関してもアンバランスなデータを考慮して, True Positive Rate (TP/(TP+FN)) と True Negative Rate (TN/(TN+FP)) の平均値を利用して分析を行った. 結果として, 入学時データのみを利用した場合は決定木を用いた場合で 51 から 52%, ナイーブベイズを用いた場合で 54 から 57% となり, 学業データを利用した場合は決定木とナイーブベイズを用いた両方ともに 70% から 80% 程度の精度となった. 学業データを利用したものはナイーブベイズのほうが決定木より精度が高かったが, 決定木のほうが結果として一貫したものとなっている.

また, 入江・丸岡 (2017) は生存時間解析といった時系列データを用いて中退を予測し, その特徴を導出した(入江と丸岡 2017). 生存時間解析ではロジスティック回帰モデルを時系列データに適用させることができ, 中退分析に用いることで中退する学生の特徴や中退する学生とそれ以外の学生の差異を導出できる. 田尻・白鳥は中退する学生には成績や性別等の特徴があることを指摘し, 入江・丸岡はUPI (University Personality Inventory) 調査の Key 項目と居住形態が中退に対して有意であることを指摘した. 一方で従来の生存時間解析を用いた手法ではイベントまでの時間を解析し, 中退をした人としなかった人の差を統計的に示していることが多い.

このように中退をする／しないという学生を予測し, 分類精度を追求する研究, 分類ごとの差異を抽出する研究はこれまでに多くなされているが, 学期ごとに学修状態の差異を踏まえながら, 中退までに個々の学生がどのような変遷を経て中退に結びつくのかといった学生の動的な状態が分析されることは無かった.

中退予測研究の中で動的な部分にアプローチした研究の代表的なものが, Bonifro たちによるものである(Del Bonifro ほか 2020). 予測する数値としては, 中退するかどうかの 2 値であり, 説明変数としては 1.入学直後に利用できる変数 (性別, 年齢, 高校成績, 高校 ID, 大学 ID), 2.学修の追加条件 (必須課題なし, 必須課題あり & パス, 必須課題あり & not パス), 3. 入学後の初年次単位数, といった 3 つの変数群を利用し, 1, 2, 3 の組み合わせ (1, 1+2, 1+2+3) を用いて実験を行っている. 利用したモデルとしては線形判別分析, Support Vector Machine を用いた判別分析, Random Forest Regressor を用いた判別分析の 3 つを比較している. 用いた精度としては正解率 (精度, accuracy, (TP+TN)/(TP+FP+TN+FN)), 感度 (再現率, Sensitivity, Recall, TP/(TP+FN) : 実際に正であるもののうち, 正であると予測した割合), 特異度 (Specificity, TN/(FP + TN), 実際に負であるもののうち, 負であると予測した割合) を用いている. 結果としては, 入学直後のデータだけを利用した形では予測のパフォーマンスは悪かったが, 中退を防止するための道具としてはデータが完全に揃う前であっても施策に使えることができることを示し

た。中退するリスクを早期の段階で見積もる場合：学修をはじめの前か初年次の学期中に行うことの差異を分析している。今後の課題としては予測ツールを大学にある他のシステム（Learning Management System や学習環境）と連動することで、特定のグループにある学生に施策を効果的にすることができることを指摘している。

2.6. 第2章のまとめ

本章では本研究に関連する中退に関係する研究として、中退理論の研究、中退要因に関する研究、中退の防止施策に関する研究、LA・EDM・IR という分野、中退予測に関する研究を述べた。

1960年代から中退に関する研究は欧米を中心になされ、当初は学生個人の能力に中退の要因を結論づける研究、理論が多かったが、1970年代以降 Tinto たちの研究を発端として、学生個人の能力だけではなく、学生と大学の間、学生と教員、学生と社会との関係性、環境が中退の決定に関連することが述べられてきた。次に、中退理論研究の中で中退に関連する変数の差異に注目し中退に関連するマクロ変数とミクロ変数について述べた。マクロ変数は大学・学部、広くは国全体で決定される変数であり、偏差値やST費、景気等をあげることができる。マクロ変数で分析できるものとしては大学・学部の中退率である。ミクロ変数とは個々の学生に関連する変数であり、性別や家族背景、高校までの成績等の入学前に決定する変数と大学での単位数や成績、プレースメントテスト等の入学後に決定する変数に分割して述べた。次に、中退理論や中退に関連する変数等を用いて中退を防止する施策の研究を述べた。中退防止研究では Seidman が述べる中退防止の式を紹介しながら、日本においては中退要因が複合的なために個人面談が中心となっている現状を述べた。さらに中退に関連する研究・実践等が変化していく背景として21世紀になってから勃興してきた LA・EDM・IR の分野について解説をした。LA・EDM・IR のどこにおいても教育データを収集・分析をし、LA・EDM では教育改善に、IR では様々な意思決定に用いられていることを述べた。これらの分野で行われている一つの研究分野が本研究において焦点を当てている中退予測研究である。中退予測研究では Hellas が用いた分類を用いながら、何を予測するのかという目的変数、何を利用して予測するのかといった説明変数、どのような手法を用いるのかといったモデルの3つに焦点を当てることで予測研究をレビューしてきた。

第3章 リサーチクエスチョン：先行研究の限界と本研究の貢献

第3章では前章までにあげた先行研究における中退予測研究、中退予測を用いた中退防止施策への応用という観点からこれまでの研究の限界を述べるとともに、本研究の貢献の焦点としてリサーチクエスチョンを提示する。

3.1. これまでの研究の限界

中退防止施策を機能させるためには、前章で述べたとおりハイリスク学生を早期発見し、迅速にかつ、継続的、集中的な介入を行っていく必要がある。早期発見のためには教学データ等を利用し中退しそうな学生を予測する必要がある。つまり、中退予測と中退防止施策との連携部分に中退予測研究、中退防止施策の研究の実質化にとって重要な部分がある。

中退予測なしに中退防止施策を行った場合は、どの学生から施策を行っていくべきなのか、どのような施策を行っていくべきなのかは教職員の経験や勘に依存するところが大きくなる。中退防止施策で主な施策は個人面談になるが、ハイリスク学生が明確になっていないと、どの学生から、いつ面談を行うべきなのかを判断することは難しい。さらに中退防止施策を行っていく場合は組織的な対応が必要であることを第2章では述べた。面談を中心とした中退防止施策のみでは、個人の教職員の負荷がどうしても大きくなってしまう。どの程度の人材資源を、どこに、どの程度行うべきなのかは大学の経営、教学マネジメントとしての判断と意思決定の支援がIRとして必要になる。しかし、中退予測を通して現状の組織における中退リスクを判断していかないと、IRとしてどの程度の人材的、経済的資源を費やして教学施策をしていくのかを計算し、リスク評価をすることができない。

以上のことから、中退を防止する施策を行うためには中退予測研究が必要になるが、第1章で述べた中退学生の特徴と中退に関連するマクロ的要因とミクロ的要因から、中退予測する研究の限界・課題があげられる。中退予測研究の課題としては中退を表現するために必要な多次元性、個人異質性に対するもの、学生の状態遷移に対するもの、という2つの中退を表現することに関するものと、中退予測研究が中退防止における施策へ適応が難しいもの、という中退防止に関連するもの、これら3つに対してこれまでの中退予測研究が焦点を当てているとは言い難い。

最初の多次元性、個人の異質性については、これまでの中退予測研究では目的変数を統計的、機械学習的な手法を用いて、説明変数の統計的な差異や予測を行う研究はあるが、全体的な予測精度よりも、個々の学生がどの程度中退しそうなのかに焦点をあてる研究は少ない。統計的な差異を求めて、それに対する施策を提示する方法もあるが、個人ごとの面談が多いなかでは、個々

の学生がどのような状態なのかを判断することが必要である。このように、これまでの中退予測研究は全体的にみて変数の差異を示す研究が多かったのに対して、中退防止につながるためにはそれ以外に個々の学生への判断が必要になってくる。

次の学生の状態遷移に関するものとしては、これまでの中退予測研究がある一時点を対象にし、その段階で中退予測の精度を考える研究が多いことに対して、中退防止の観点からは学生がどの程度状態が変化し、どうなっていくのかといった研究が少なかった。状態の遷移を意識した研究としては先述した Bonifro 2020 等があり、いつ中退しそうなのかを予測する研究としては Aulck 2017 等がある(Aulck ほか 2016)。学生がどのような状態変化を経て、現段階を迎え、今後どのような状態になっていくのかといった部分、状態遷移の部分に焦点を当てている研究は少ない。

個々の学生に焦点を当てて、中退防止を行うことはこれまでも述べたが、中退しそうだとする学生がどの程度いるのかに次に焦点をあてていくことが IR としては求められる。IR はデータを用いて大学における様々な意思決定を行うことが大切になるが、どのような施策を、どの程度、どの学生に対して施策を行うべきなのかに対する客観的なデータを中退予測研究では十分に提供していない。

3.2. リサーチクエスチョン

本研究では中退を予測するモデルを利用し、個々の学生に適応させ、学生ごとの状態を数値で算出する。算出した状態を用いて学生の状態遷移を表し、予測研究と中退防止をつなげるためにパターンを作成することで、学生の中退までの一連のプロセスを表現することができ、中退防止施策にもつなげることができる。下記の3点が本研究のリサーチクエスチョンになる。

1. 学生ごとに中退しそうな状態を数値で算出することで、個々の学生がどの程度中退しそうな状態なのかを明らかにする（個人の異質性）
2. 状態を表す数値を時系列ベクトルで整理をすることで、学生の状態遷移を把握可能にする（遷移の把握）
3. 時系列ベクトルをクラスタリングすることで中退までの動的特性を明らかにする（類型化）

上記の3点について、大学全体での学生状態の把握に焦点をあてたものを第4章、1年次春学期における学修状態の把握に焦点をあてたものを第5章で述べていく。第4章では、どの程度中退をしそうなのかを中退確率として定義し、学期ごとに学生の学修状態を数値化する。学期ごとの学修状態を表す数値を、複数学期つなげ入学時から中退するまでの学修状態の遷移を把握する。その後、学生の状態遷移の数値を類型化し、どの程度の学生数がどのように中退していくのかを把握していく。第5章では、1年次春学期に焦点を当て、週毎の学生の状態を把握するために1年次春学期のGPAを予測した予測GPAを用いる。更に、週毎の数値を学期の最初から学期終了時

までつなげ、学期中にどのような状態遷移になるのかを把握する。その後、学期内の学生の状態推移の数値を類型化し、学期中にどの程度の数の学生が、状態がどのように変化するのかを明らかにしていく。

第4章 入学から中退までの学修状態の把握

本章では、前章で述べたリサーチクエスションを入学から中退までの学修状態の把握に利用し、数値実験を行うことで検証をしていく。最初に第4章、第5章で利用するデータと変数を定義し、整理をする。その次に定義した変数を用いてロジスティック回帰モデルを利用して中退確率を算出する方法を述べる。次に、中退確率を用いて学期ごとの学修状態の特徴を述べる。その後、中退確率を複数用いて学期ごとの遷移ベクトルを表現し、遷移ベクトルを用いてクラスタを作成する。最後に作成したクラスタを用いて学修状態の中退パターンを作成する。

4.1. 利用するデータと変数

本節では学生が中退する可能性を数値として定義する際に利用するデータと変数を整理し、定義する。

4.1.1. 利用するデータ

都内にある A 大学の文系学部における 2012 年～2014 年度入学生のデータを対象とし、編入学、再入学、早期卒業、大学院の学生、入学前、入学後のデータが無い学生は除いた。学期毎の通常学生数、次学期以降の中退者数は表 4.1 のとおりである。通常学生とは中退をしなかった学生であり、次学期以降の中退者とは当該学期の次の学期以降に中退をした学生であり、当該学期の中退者はその学期に中退をした学生を表している。

表 4.1 学期ごとの中退者数

	通常学生数	次学期以降の中退者数	当該学期の中退者数
入学時	600	249	
1年目春学期_終了時	600	237	12
1年目秋学期_終了時	600	189	48
2年目春学期_終了時	600	143	46
2年目秋学期_終了時	600	92	51
3年目春学期_終了時	600	71	21
3年目秋学期_終了時	600	51	20
4年目春学期_終了時	600	35	16
4年目秋学期_終了時	600	13	22

4.1.2. 学期と週の表現

$$s = 0, 1, 2, \dots, 8$$

$$w = 0, 1, 2, \dots, 15, 16$$

学期と週を表現するための変数を定義する。 s は大学における学期を表し、最初の学期、つまり1年次春学期が $s = 1$ で表され、1年次秋学期が $s = 2$ で表され、 $s = 8$ は4年生秋学期を表す。また、 $s = 0$ は、1年次春学期開始前、入学直後を表す。 w は学期内の週を表す。 $w = 0$ は0週目、つまり学期前を表し、 $w = 1$ は当該学期の1週目を表し、15週目を $w = 15$ で表す。また、16週目は $w = 16$ で表すが、これは授業（第15回まで）が終わったあとを指す。

4.1.3. 入学時までに取り得る変数

入学時までに取り得る変数 x_i を定義し表4.2にまとめた。変数の選択は第2章で説明した中退に関連するマイクロ変数をもとに選択した。入学時までに取り得る変数 x_i は学生ごとの変数のため、学生 m を個で指定して議論したい場合、個々の学生のデータを表す場合は $x_i[m]$ と書く。

入学時までに取り得る変数は性別、留学生、高校時欠席日数、出身高校課程である。性別は x_1 で定義し、男性は $x_1 = 1$ 、女性は $x_1 = 0$ で表した。性別が不明な場合は `null` を許可した。留学生は x_2 で定義し、留学生の場合は $x_2 = 1$ 、それ以外は $x_2 = 0$ で表した。留学生かどうか不明な場合は `null` を許可した。高校時欠席日数は x_3 で定義し、整数で表した。高校時欠席日数が不明な場合は `null` を許可した。高校時評定平均は x_4 で定義し、小数で表した。高校時評定平均が不明な場合は `null` を許可した。出身高校課程は全日制高校か、通信制高校かで表し、全日制高校かどうかを x_5 を用いて、通信制高校かどうかを x_6 を用いて定義した。全日制高校に通学していた場合は $x_5 = 1$ 、それ以外は $x_5 = 0$ で表した。通信制に通学していた場合は $x_6 = 1$ 、それ以外は $x_6 = 0$ で表した。

表 4.2 入学時までに取り得る変数

x_i : 入学前までに取得できる変数	変数解釈	変数タイプ	NULLの許可
x_1 : 性別	男性=1, 女性=0,	INTEGER	NULLABLE
x_2 : 留学生	留学生=1, それ以外=0	INTEGER	NULLABLE
x_3 : 高校時欠席日数		INTEGER	NULLABLE
x_4 : 高校時評定平均		FLOAT	NULLABLE
x_5 : 全日制高校	全日制高校=1, それ以外=0	INTEGER	NULLABLE
x_6 : 通信制高校	通信制=1, それ以外=0	INTEGER	NULLABLE

4.1.4. 入学後に取得できる変数 1：学期ごとの変数

入学後に取得できる変数 $x_{s,j}$ を定義し、表 4.3 にまとめた。入学後に取得できる変数は s （当該学期）までの総取得単位数，当該学期の取得単位数，当該学期の GPA である。入学後に取得できる変数 $x_{s,j}$ は学生ごとの変数のため，学生 m を個で指定して議論したい場合は $x_{s,j}[m]$ と書く。

表 4.3 入学後に取得できる変数 1

x_1 ：入学後に取得できる変数	変数解釈	変数タイプ	NULLの許可
$x_{s,1}$ ： s （当該学期）までの総取得単位数	単位数：整数	INTEGER	NULLABLE
$x_{s,2}$ ： s （当該学期）までの取得単位数	単位数：整数	INTEGER	NULLABLE
$x_{s,3}$ ： s （当該学期）までのGPA	GPA：少数	FLOAT	NULLABLE

当該学期 s までの総取得単位数は $x_{s,1}$ で定義し，整数で表した。総取得単位数が不明な場合は null を許可した。当該学期 s の取得単位数は $x_{s,2}$ で定義し，整数で表した。当該学期 s の取得単位数が不明な場合は null を許可した。当該学期 s の成績は GPA (Grade Point Average) で表し，GPA は $x_{s,3}$ で定義し，小数で表した。当該学期 s の GPA が不明な場合は null を許可した。

4.1.5. 入学後に取得できる変数 2：学期中の変数

入学後，とくに学期中に取得できる変数 $x_{s,w,k}$ を定義し，表 4.5 にまとめた。学期中に取得できる変数 $x_{s,w,k}$ は当該学期 s ，当該週 w ，当該授業 k の出欠である。 $x_{s,w,k}$ は学生ごとの変数のため，学生 m を個で指定して議論したい場合は $x_{s,w,k}[m]$ と書く。

表 4.4 入学後に取得できる変数 2

$x_{s,w,k}$ ：入学後、学期中に取得できる変数	変数解釈	変数タイプ	NULLの許可
$x_{s,w,k}$ ： w （当該週）の授業 k の出席	出席、欠席、不明、未履修	INTEGER	NULLABLE

当該学期 s における，当該週 w ，授業 k の出欠は $x_{s,w,k}$ で定義し，出席，欠席，不明，未履修の文字列で表した。当該学期 s における，当該週 w ，授業 k の出欠が不明な場合は null を許可した。

4.1.6. 目的変数：中退したかどうか

目的変数として中退したかどうかを表す変数 y を定義し，表 4.5 にまとめた。

中退したかどうかを表す目的変数 y は学生ごとの変数のため，学生 m を個で指定して議論したい場合は $y[m]$ と書く。

表 4.5 目的変数

y：中退したかを笑わず変数	変数解釈	変数タイプ	NULLの許可
y：学生が中退したかどうかを表す	中退した=1、中退していない=0	INTEGER	NULLABLE

中退したかどうかを表す変数はyで定義し、中退した場合は $y = 1$ 、中退していない場合は $y = 0$ で表した。

4.2. 学修状態の数値化：中退確率の定義

本節では前節までに示した学期、週の表現，入学前に取得できる変数，入学後に取得できる変数を利用して学生の学修状態を中退確率として数値化する。中退確率は当該学期 s までのデータを用いて，当該学生 m が，当該学期 s 以降にどの程度中退をしやすいかを中退確率 $p_s[m]$ と表す。

4.2.1. 中退確率の算出方法

中退確率 $p_s[m]$ の算出方法は下記の3つの手順で表される。

1. 現在のデータから学期ごとにロジスティック回帰モデルの係数を決定する
2. 推定した係数を利用して，学期ごとの中退確率モデルを決定する
3. 学期ごとの中退確率モデルを利用し，学生の学期ごとの中退確率を算出する

1. ロジスティック回帰モデルの係数を決定する

1のロジスティック回帰モデルの係数を決定する，では当該学期 s ごとに作成する下記のロジスティック回帰モデルの係数 $\alpha_{s,0}$ ， $\alpha_{s,i}$ ， $\beta_{s,k,j}$ を推定する。

$$\ln\left(\frac{p_s}{1-p_s}\right) = \alpha_{s,0} + \sum_{i=1}^6 \alpha_{s,i} x_i + \sum_{k=1}^s \sum_{j=1}^3 \beta_{s,k,j} x_{k,j}$$

右辺の第1項 $\alpha_{s,0}$ は定数項を表し，第2項における $\alpha_{s,i}$ は入学時までに取得できる説明変数 x_i にかかる係数であり，第3項における $\beta_{s,k,j}$ は入学後に取得できる説明変数 $x_{k,j}$ にかかる係数である。推定方法は最尤推定法を利用する。ロジスティック回帰モデルに関連する尤度 $l(\alpha_{s,0}, \alpha_{s,i}, \beta_{s,k,j})$ は当該学期 s ごとに下記で表される。下記の尤度を最大にする最尤推定を学期ごとに行い，ロジスティック回帰モデルの係数 $\alpha_{s,0}$ ， $\alpha_{s,i}$ ， $\beta_{s,k,j}$ を求める。

$$\prod_{m=1}^M \frac{1}{1 + \exp\left(-y[m]\left(\alpha_{s,0} + \sum_{i=1}^6 \alpha_{s,i} x_i[m] + \sum_{k=1}^s \sum_{j=1}^3 \beta_{s,k,j} x_{k,j}[m]\right)\right)}$$

上記の尤度 $\ell(\alpha_{s,0}, \alpha_{s,i}, \beta_{s,k,j})$ の対数のマイナスをとった、下記の $\ell(\alpha_{s,0}, \alpha_{s,i}, \beta_{s,k,j})$ を最小にする係数を求める。

$$\ell(\alpha_{s,0}, \alpha_{s,i}, \beta_{s,k,j}) = \sum_{m=1}^M \ln(1 + v_m)$$

$$v_m = \exp\left(-y[m]\left(\alpha_0 + \sum_{i=1}^6 \alpha_i x_i[m] + \sum_{k=1}^s \sum_{j=1}^3 \beta_{k,j} x_{k,j}[m]\right)\right)$$

以上の最尤推定を用いて、学期ごとのロジスティック回帰モデルの係数 $\alpha_{s,0}$, $\alpha_{s,i}$, $\beta_{s,k,j}$ を決定した

2. 学期ごとの中退確率モデルの決定

1 で作成した学期ごとのロジスティック回帰モデルの係数 $\alpha_{s,0}$, $\alpha_{s,i}$, $\beta_{s,k,j}$ を用いて、学期ごとに下記のロジスティック回帰モデルを作成する。ただし、 $s=0$ の場合、つまり入学直後の場合は下記の第3項部分は無い。

$$\ln\left(\frac{p_s}{1 - p_s}\right) = \alpha_{s,0} + \sum_{i=1}^6 \alpha_{s,i} x_i + \sum_{k=1}^s \sum_{j=1}^3 \beta_{s,k,j} x_{k,j}$$

3. 学期ごと、学生ごとの中退確率の算出

2 で作成した学期ごとの中退確率モデルを利用して、学生ごと、学期ごとに中退確率を算出する。学生 m ごとに $x_i[m]$, $x_{s,j}[m]$ を学期ごとの中退確率モデルに代入して、学生ごとに学期ごとの中退確率 $p_s[m]$ を算出する。

$$\ln\left(\frac{p_s[m]}{1 - p_s[m]}\right) = \alpha_{s,0} + \sum_{i=1}^6 \alpha_{s,i} x_i[m] + \sum_{k=1}^s \sum_{j=1}^3 \beta_{s,k,j} x_{k,j}[m]$$

ただし、学生ごとに学期ごとの中退確率 $p_s[m]$ は下記の条件（0 以上、1 以下）が付される。

$$0 \leq p_s[m] \leq 1$$

以上の1から3の手順を経ることで、学生ごと、学期ごとの中退確率 $p_s[m]$ が算出される。

4.2.2. ロジスティック回帰モデルの精度

4.2.1 で利用したロジスティック回帰モデルの精度は表 4.6 にまとめた。精度は正解率 (accuracy), 適合率 (precision), 再現率 (recall), F 値 (f1) で求めた。

精度は下記の表で示された TP (True Positive: 真陽性), FP (False Positive: 偽陽性), FN (False Negative: 偽陰性), TN (True Negative: 真陰性) の4つの組み合わせで求められる。TP は実際に中退した学生を中退すると予測できた場合, FP は実際には中退しなかった学生を中退すると予測した場合, FN は実際に中退した学生を中退しないと予測した場合, TN は実際に中退しなかった学生を中退しないと予測した場合を言う。

表 4.6 精度の分類

	Real Positive	Real Negative
Predicted Positive	TP	FP
Predicted Negative	FN	TN

正解率 (accuracy) は下記で表され、予測結果がどの程度真であったのかを表した指標になっている。

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

適合率 (Precision) は下記で表され、正事例だと予測したもののなかで、実際に真陽性、正事例だった値の割合を表す指標になる。

$$Precision = \frac{TP}{TP + FP}$$

再現率 (Recall) は下記で表され、実際に正事例なものなかで、実際に真陽性、正事例だった値の割合を表す指標になる。

$$Recall = \frac{TP}{TP + FN}$$

F 値 (F-Value) は下記で表され Precision と Recall の調和平均で求められる

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

学期ごとのロジスティック回帰モデルの精度を表したものが表 4.7 になる。各値は分布の比率、つまり中退した学生と中退しなかった学生の比率を維持したままデータを訓練用とテスト用に分割した層状 K 分割法 (k=3) を用いて算出した。なお、各値は小数点第 4 位を四捨五入して求めている。

入学時において正解率は 7 割を超えているが、適合率、再現率、F 値は 5 割から 6 割程度になっている。1 年目春学期終了時以降は正解率、適合率、再現率、F 値ともに 7 割を超えてきているため、モデルとしては妥当である。適合率が再現率をどの学期においても超えているが、FN (中退したのに、中退していないと予測した場合) が発生していることを許容していることを表している。

表 4.7 ロジスティック回帰モデルの精度

学期	accuracy	precision	recall	f1
入学時	0.715	0.669	0.528	0.482
1年目春学期_終了時	0.799	0.769	0.702	0.721
1年目秋学期_終了時	0.849	0.810	0.750	0.773
2年目春学期_終了時	0.871	0.824	0.723	0.756
2年目秋学期_終了時	0.912	0.840	0.732	0.770
3年目春学期_終了時	0.912	0.798	0.691	0.722
3年目秋学期_終了時	0.940	0.821	0.707	0.746
4年目春学期_終了時	0.959	0.808	0.761	0.781
4年目秋学期_終了時	0.972	0.603	0.619	0.607

4.3. 中退確率を用いた状態の特性

4.3.1. 有意な変数の抽出

学期ごとに説明変数をデータが 0 以上 1 以下になるように正規化した後に χ 二乗検定を用いた独立性の検定を行った。各変数の VIF はすべて 10 以下であるため、多重共線性は発生していない。

表 4.8, 表 4.9 は統計的に 5% で有意である説明変数, オッズ比, 偏回帰係数, p 値をオッズ比の絶対値を利用し降順で並べ替えを行った形で学期ごとに記した。

入学時の段階では合計欠席日数のみが有意になっており, 偏回帰係数の符号は正になっている。つまり, 高校時での欠席日数が多いことと中退することに関連があることがわかる。

1 年目春学期終了時と 1 年目秋学期終了時の段階において, 偏回帰係数の符号が正になっている変数は合計欠席日数, 負になっている変数は GPA (成績) と単位数である。高校時の欠席日数が多いこと, 1 年目の GPA が低いこと, 単位数が少ないことが中退することと関連していることがわかる。以上のように 1 年目春学期終了時と 1 年目秋学期終了時では, 合計欠席日数, GPA, 単位数と中退することの関係は同じ傾向にある。

一方で, 2 年目春学期終了時以降は合計欠席日数, GPA, 単位数と中退することの関係は大きく変わってくる。

表 4.8 有意な変数 1 (入学時から 3 年目春学期まで)

学期	変数名	オッズ比	偏回帰係数	p値	学期
入学時	合計欠席日数	1.014	0.014	0.015	入学時
1年目 春学期 終了時	合計欠席日数	1.007	0.007	0.022	1年目春
	1年目春修	0.933	-0.069	< 0.001	1年目春
	1年目春_GPA	0.528	-0.639	< 0.001	1年目春
1年目 秋学期 終了時	合計欠席日数	1.003	0.003	0.036	1年目秋
	1年目春修	0.99	-0.01	< 0.001	1年目秋
	1年目春_GPA	0.97	-0.031	< 0.001	1年目秋
	1年目秋_GPA	0.941	-0.061	< 0.001	1年目秋
	1年目秋修	0.925	-0.078	< 0.001	1年目秋
	総単位数	0.916	-0.088	< 0.001	1年目秋
2年目 春学期 終了時	1年目秋_GPA	1.246	0.22	< 0.001	2年目春
	2年目春_GPA	1.075	0.072	< 0.001	2年目春
	1年目春_GPA	1.016	0.016	< 0.001	2年目春
	1年目春修	1.007	0.007	0.001	2年目春
	合計欠席日数	1.003	0.003	0.041	2年目春
	1年目秋修	1.002	0.002	< 0.001	2年目春
	総単位数	0.914	-0.09	< 0.001	2年目春
	2年目春修	0.906	-0.099	< 0.001	2年目春
2年目 秋学期 終了時	2年目春_GPA	1.727	0.547	< 0.001	2年目秋
	1年目春_GPA	1.463	0.38	0.01	2年目秋
	1年目秋修	1.031	0.03	0.002	2年目秋
	1年目秋_GPA	1.028	0.028	0.001	2年目秋
	1年目春修	0.997	-0.003	0.02	2年目秋
	2年目秋修	0.975	-0.026	< 0.001	2年目秋
	総単位数	0.931	-0.071	< 0.001	2年目秋
	2年目春修	0.929	-0.073	< 0.001	2年目秋
	2年目秋_GPA	0.441	-0.818	< 0.001	2年目秋
3年目 春学期 終了時	1年目春_GPA	1.718	0.541	0.026	3年目春
	2年目春_GPA	1.241	0.216	0.001	3年目春
	1年目秋修	1.057	0.056	0.007	3年目春
	2年目秋修	1.013	0.013	< 0.001	3年目春
	1年目春修	0.992	-0.008	0.048	3年目春
	2年目春修	0.945	-0.056	0.001	3年目春
	総単位数	0.94	-0.062	< 0.001	3年目春
	3年目春修	0.935	-0.067	< 0.001	3年目春
	2年目秋_GPA	0.861	-0.15	< 0.001	3年目春
	1年目秋_GPA	0.742	-0.299	0.004	3年目春
	3年目春_GPA	0.643	-0.442	< 0.001	3年目春

表 4.9 有意な変数 2 (3 年目秋学期から 4 年目春学期まで)

3年目 秋学期 終了時	2年目春_GPA	1.628	0.488	0.005	3年目秋
	1年目春_GPA	1.224	0.202	0.044	3年目秋
	3年目春_GPA	1.111	0.105	< 0.001	3年目秋
	3年目春修	1.038	0.037	0.001	3年目秋
	1年目秋修	1.008	0.008	0.018	3年目秋
	2年目秋修	0.978	-0.022	0.001	3年目秋
	総単位数	0.953	-0.048	0.001	3年目秋
	3年目秋修	0.939	-0.062	< 0.001	3年目秋
	2年目春修	0.929	-0.073	0.003	3年目秋
	2年目秋_GPA	0.852	-0.16	0.001	3年目秋
	1年目秋_GPA	0.752	-0.285	0.01	3年目秋
	3年目秋_GPA	0.399	-0.92	< 0.001	3年目秋
	4年目 春学期 終了時	2年目春_GPA	1.506	0.409	0.018
3年目春_GPA		1.153	0.142	0.006	4年目春
1年目秋_GPA		1.135	0.126	0.041	4年目春
3年目春修		1.05	0.048	0.008	4年目春
2年目秋修		1.008	0.008	0.005	4年目春
総単位数		0.963	-0.037	0.009	4年目春
3年目秋修		0.95	-0.051	0.004	4年目春
2年目春修		0.9	-0.105	0.01	4年目春
3年目秋_GPA		0.752	-0.285	0.001	4年目春
2年目秋_GPA		0.636	-0.452	0.005	4年目春
4年目春_GPA		0.203	-1.594	< 0.001	4年目春

2 年目春学期終了時の段階において、偏回帰係数の符号が正になっている変数には合計欠席日数、1 年目と 2 年目の GPA、1 年目の単位数、負になっている変数には 2 年目春学期の単位数と総単位数がある。前学期までとの違いは GPA に関する偏回帰係数の符号が負から正になっていることである。つまり、1 年目秋学期以前は GPA が低いことが中退することと関連があったのに対し、2 年目春学期終了時の段階では GPA が高いことが中退することと関連がある。つまり、1 年目の GPA が良い学生が中退をしまっていることになる。

2 年目秋学期終了時の段階においても GPA と中退の関係性は同じように続く。2 年目春以前の GPA が高いことが中退することに関連する。また、2 年目秋学期終了時以降からは入学時までに取得できる変数である合計欠席日数は有意ではなくなる。

3 年目春学期終了時以降は 1 年目春学期 GPA, 2 年目春学期の GPA の高さは中退することに関連し、2 年目秋学期終了時と同じ傾向が続く。また、当該学期の GPA と単位数は概ね偏回帰係数の符号が負になる。ただし、4 年目春学期終了時は当該学期の単位数との関係はない。

まとめると共通して下記の特徴があげられ、学期ごとに中退する学生の特徴が変化することがわかる。

1. 高校時の合計欠席日数が中退と関係するのは2年目春学期終了時まで
2. 1年目では、1年目のGPAと単位数は偏回帰係数の符号が負になっているが、2年目以降は、1年目春学期のGPA、2年目春学期のGPAは偏回帰係数の符号は正の場合が多い。ただし、4年目春学期終了時の1年目春学期GPAは除く
3. 当該学期のGPAと単位数は概ね偏回帰係数の符号が負になる

4.3.2. 中退確率を用いた状態

4-2で述べた手法を用いて、学生がどの程度中退しそうな状態なのかを学生ごと、学期ごとに算出された中退確率を用いて表した。表4.10は中退確率の一例だが、学生ごと、学期ごとに中退確率が与えられている。また、数値がない場合は実際に中退をしまっていることを表している。

以上の中退確率を学期ごとにヒストグラムで表したものが図4.1になる。階級幅は0.1に設定している。入学時から1年目春学期終了時、1年目秋学期終了時と学期を経ていくごとに、中退確率の分布が0と1の両極化していくことがわかる。

表 4.10 中退確率のサンプル

学生ID	1年目春_中退確率	1年目秋_中退確率	2年目春_中退確率	2年目秋_中退確率	3年目春_中退確率	3年目秋_中退確率	4年目春_中退確率	4年目秋_中退確率
1	0.308	0.899	0.970	0.996	1.000	1.000		
2	0.166	0.072	0.096	0.049	0.004	0.005	0.002	0.008
3	0.138	0.056	0.072	0.014	0.001	0.003	0.001	0.001
4	0.195	0.352	0.573	0.244	0.165	0.041	0.018	0.001
5	0.188	0.032	0.095	0.031	0.008	0.002	0.001	0.000
6	0.247	0.109	0.480	0.602	0.685	0.761	0.230	0.022
7	0.453	0.862	0.924	0.966	0.995	0.997		
8	0.081	0.051	0.169	0.036	0.018	0.011	0.035	0.003
9	0.219	0.077	0.132	0.059	0.002	0.030	0.002	0.000
10	0.046	0.518	0.853	0.968				
11	0.308	0.796	0.461	0.982	0.995	0.994		
12	0.135	0.454	0.857	0.964				
13	0.912	0.927	0.972	0.974				
14	0.088	0.060	0.043	0.024	0.001	0.001	0.000	0.000
15	0.315	0.300	0.147	0.012	0.002	0.003	0.010	0.001
16	0.163	0.257	0.118	0.139	0.314	0.046	0.001	0.000
17	0.248	0.159	0.091	0.109	0.007	0.005	0.003	0.001
18	0.788	0.627	0.273	0.034	0.680	0.552		
19	0.176	0.102	0.126	0.050	0.018	0.017	0.009	0.001
20	0.091	0.048	0.033	0.037	0.019	0.014	0.006	0.001

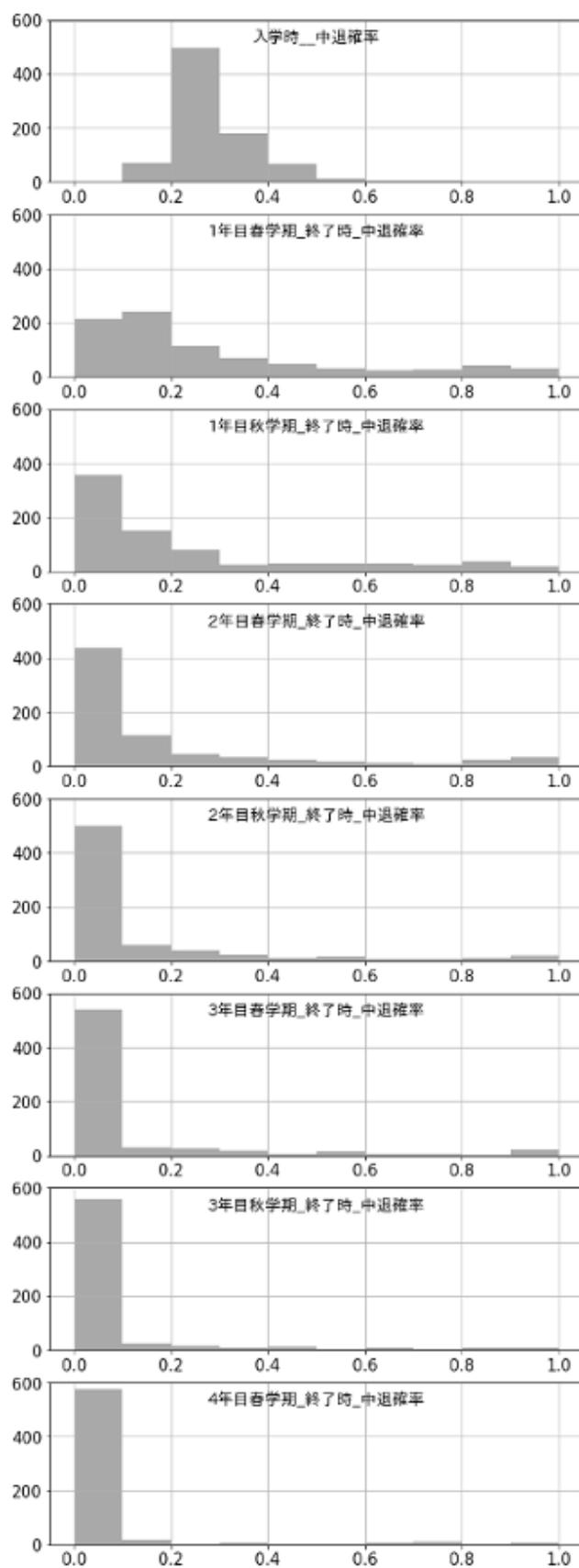


図 4.1 学期ごとの中退確率ヒストグラム

4.3.3. 中退確率を用いた学生の分類

算出した各学期の中退確率を用いて、学期ごとに学生を 5 グループに分類した。グループ番号 I は中退確率が 0 以上 0.25 以下の学生、グループ番号 II は中退確率が 0.25 より大きく 0.5 以下の学生、グループ番号 III は 0.5 より大きく 0.75 以下の学生、グループ番号 IV は 0.75 より大きく 1 以下の学生である。今回のグループはわかりやすさを重視し 0.25 刻みとした。また、グループ番号は付していないが中退した学生のグループも追加し、合計 5 グループである。グループ番号が大きくなればなるほど中退の可能性が高くなることを示している。表 4.11 は学期ごとのグループの人数と中退者数を表している。

表 4.11 学期ごとのグループ人数

学期	グループ番号				
	I	II	III	IV	中退
入学時	319	499	24	7	0
1年目春_終了時	520	166	65	86	12
1年目秋_終了時	551	94	74	70	60
2年目春_終了時	574	78	32	59	106
2年目秋_終了時	582	51	23	36	157
3年目春_終了時	586	35	22	28	178
3年目秋_終了時	591	26	14	20	198
4年目春_終了時	596	11	11	17	214

4.4. 中退確率を用いた学修状態の遷移

4.4.1. 中退確率を用いた遷移の数値化

4-3 で述べた学生ごと、学期ごとの中退確率を用い学修状態の遷移の数値化を行う。中退確率を用いることで、学生は学期 s ごとに 1 つの中退確率を保持することになる。学部の標準修業年限が 4 年であり、入学後 4 年間で卒業する学生 m は $s = 0$ から 8 までの $p_s[m]$ を持つ。各学生の $p_s[m]$ を用いて、学期ごとの学生 m の中退確率を用いた遷移 $\Delta p_s[m]$ を下記のように定める。ただし、当該学期 s においてすでに中退してしまった学生は $p_s[m]$ を持たないため、 $\Delta p_s[m]$ を null としている。

$$\Delta p_s[m] = p_s[m] - p_{s-1}[m]$$

Δp_s の特徴は p_s の定義から下記になる。

・ $\Delta p_s[m] \geq 0$ の場合は状態が悪化，つまり学生 m が当該学期の 1 学期前と比べてより中退に近づいている遷移を表す

・ $\Delta p_s[m] < 0$ の場合は状態が良化，つまり学生 m が当該学期の 1 学期前と比べて中退しそうな状態から離れている遷移を表す

図 4.2 は横軸に入学時の中退確率，縦軸に 1 年次春学期の中退確率を用いた散布図である．赤色の点線は $\Delta p_1 = 0$ であり，中退確率が変わらなかったことを表している．赤色の点線より下の学生は $\Delta p_1 < 0$ であり，学修状態が入学時より良化した学生を，赤色の点線より上の学生は $\Delta p_1 > 0$ であり，学修状態が入学時より悪化した学生を表している．

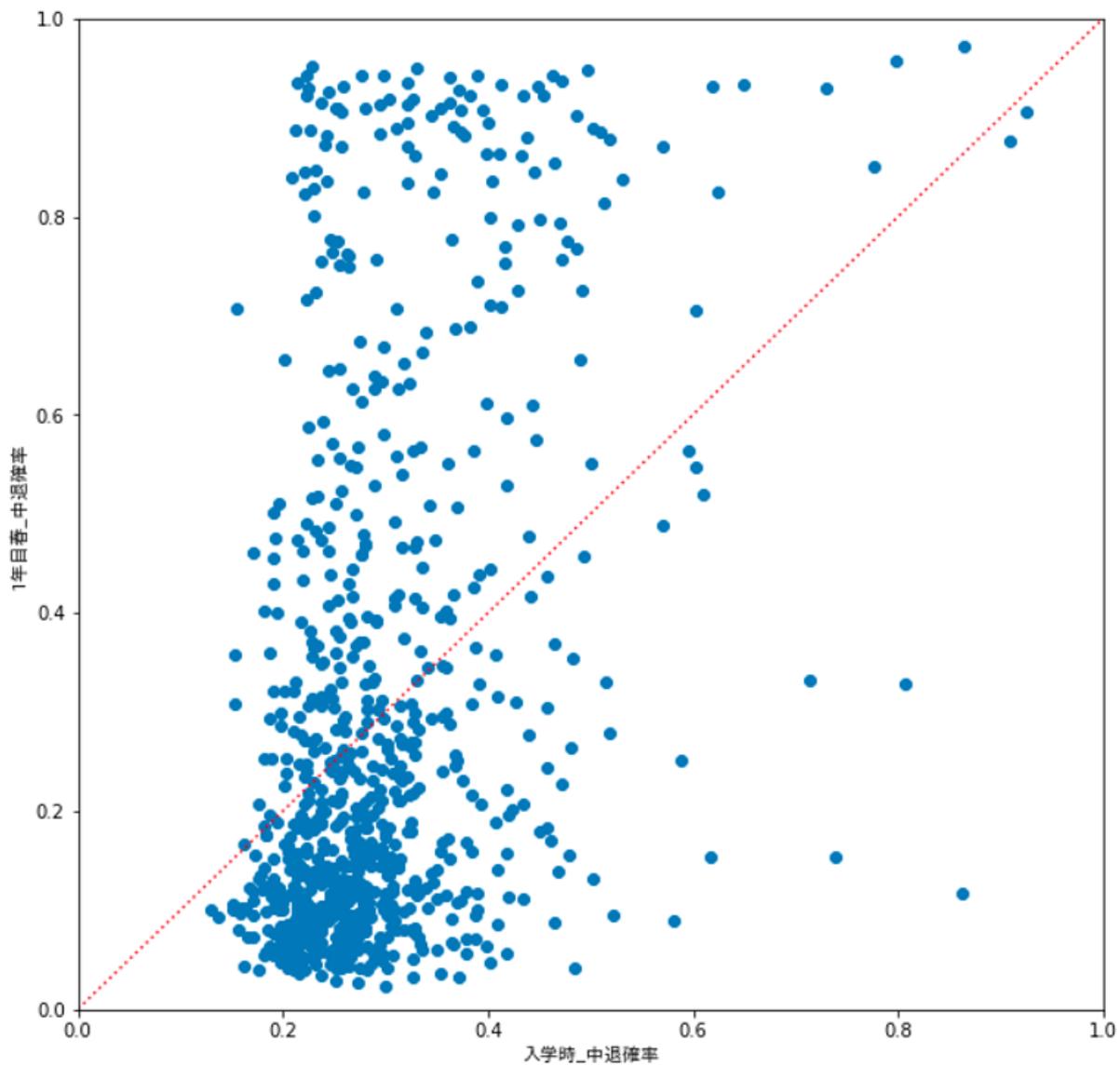


図 4.2 入学時の中退確率と 1 年目春学期終了時の中退確率の散布図

4.4.2. 中退確率を用いた遷移の分類

学期間の中退確率の遷移を用いることで、どの程度の学生が悪化/良化しているのかが分かることを示したが、学期ごとにどの程度の学生が悪化/良化しているのかを把握する。

表 4.12 学期ごとのグループの遷移

学期/状態	グループ番号Ⅰ			グループ番号Ⅱ			グループ番号Ⅲ			グループ番号Ⅳ				
	そのまま	悪化	中退	良化	そのまま	悪化	中退	良化	そのまま	悪化	中退			
入学時->1年目春	72.4%	27.0%	0.6%	56.7%	21.0%	20.6%	1.6%	41.7%	16.7%	33.3%	8.3%	28.6%	71.4%	0.0%
1年目春->1年目秋	90.0%	8.5%	1.5%	45.8%	24.7%	23.5%	6.0%	35.4%	35.4%	18.5%	10.8%	14.0%	59.3%	26.7%
1年目秋->2年目春	93.5%	5.6%	0.9%	48.9%	30.9%	16.0%	4.3%	41.9%	12.2%	24.3%	21.6%	15.7%	54.3%	30.0%
2年目春->2年目秋	94.3%	4.0%	1.7%	46.2%	29.5%	12.8%	11.5%	34.4%	21.9%	18.8%	25.0%	10.2%	49.2%	40.7%
2年目秋->3年目春	96.2%	2.9%	0.9%	43.1%	31.4%	17.6%	7.8%	34.8%	43.5%	4.3%	17.4%	5.6%	72.2%	22.2%
3年目春->3年目秋	97.4%	2.2%	0.3%	45.7%	22.9%	20.0%	11.4%	36.4%	31.8%	9.1%	22.7%	10.7%	57.1%	32.1%
3年目秋->4年目春	98.5%	0.8%	0.7%	50.0%	15.4%	19.2%	15.4%	21.4%	28.6%	21.4%	28.6%	15.0%	65.0%	20.0%
平均	93.2%	5.9%	1.0%	51.8%	23.8%	19.8%	4.5%	37.0%	25.2%	19.7%	18.1%	12.7%	58.2%	29.1%

良化/悪化を把握するために、4-3-3で行ったグループを利用し学期間の遷移を把握する。グループ番号がIからIVになるにつれて状態が悪化したグループであり、グループ番号が次学期に大きくなった場合は悪化、小さくなった場合は良化となる。表 4.12 は前節で表した学期間のグループの遷移を表している。状態がそのままはグループが学期をまたいでグループ番号が変わらず、先学期と同じグループだった学生のグループ内での割合、悪化はグループ番号が大きくなり、先学期より中退に近づいた学生数の割合、良化はグループ番号が小さくなり、先学期より中退から離れた学生数の割合、中退はその学期に当該グループから中退した学生数の割合を示している。

状態がそのままの平均値の割合が悪化、良化、中退に比べて一番高いグループはグループ番号IとIVである。特にグループIのそのままの平均値は93.2%となっており、一度状態が良いグループIに入った場合、そのままグループIで4年生までつながることが多いことが分かる。入学時から1年目春の遷移ではそのままの割合は72.4%と平均93.2%と比較して小さいが、それ以降はそのままの割合が9割を下回ることはない。グループ番号IVでは、そのままの割合が平均で58.2%と良化、中退と比較して一番高いが、中退につながる割合も平均で29.1%になり、良化の割合12.7%と比較して高い。入学時から1年目春の遷移では良化する割合は28.6%と他の学期間に比べて高くなっているが、1年目春以降、良化の割合が20%より高くなることはない。

グループ番号IIとグループ番号IIIは良化する割合が高い。グループ番号IIの場合は良化の平均割合が51.8%と高く、入学時から1年目春の遷移時が56.7%と他の学期に比べて一番高くなっているが、1年目春学期以降も40%より低くなることは無い。グループ番号IIIの場合は良化の平均割合が37.0%とそのまま(25.2%)、悪化(19.7%)、中退(18.1%)と比べて高いことが分かる。グループ番号IIとIIIはグループ番号IとIVと比べて、そのままの割合が低く、良化、悪化していき、結果としてグループ番号I、グループ番号IV、そして中退のグループの数値が高くなっていくことが分かる。

中退に結びつく割合は、グループ番号Iが平均で1.0%、グループ番号IIが4.5%、グループ番号IIIが18.1%、グループ番号IVが29.1%であり、中退に結びつく割合が高いグループはIIIとIVであることがわかる。グループ番号IVから中退に結びつく割合が他のグループと比べて一番高く、次に高いグループがグループ番号IIIであることから、状態が悪化した結果、グループ番号III、もしくはグループ番号IVになってから大学を中退していくことが分かる。

4.5. 中退確率の遷移を用いたクラスタの作成

前節までは学生の学修状態の特徴を学期ごとの断面において遷移を導出したが、実際に中退をするまでにどのような過程を経るのかは学生ごとに大きく異なる。例えば、2年目春学期以降の中退者においても1年目から中退しそうだったのか、それとも2年目以降で中退しそうになったのかといった遷移に注目していない。本節では中退した学生に焦点を当て、入学時から中退するまでの中退確率の遷移に注目することで、中退学生の状態遷移の特徴を抽出する。

中退する学生の遷移の特徴を判断するために、教師無し学習を用いて中退時期ごとにクラスタを作成した。利用する手法は k の数を同時に推定できる X-means 法(Pelleg と Moore 2000)を用いた。X-means は Python のライブラリである pyclustering (バージョン 0.10.1.2) を用い、X-means のパラメータとして最初のクラスタ数には 1、クラスターの最大数は 20、反復の停止条件は 0.025、分割停止基準は BIC (ベイズ情報量基準) を用いた。

クラスタリングの方法は下記である

1. 中退した学生の中退確率を中退した学期まで算出する
2. 中退した学生の中退した学期までの中退確率を遷移ベクトルとして保存する
3. 中退した学生を中退した学期ごとに分類する
4. 2で作成した個々の学生の遷移ベクトルを用い、X-means を行いクラスに分類する

上記の 1 では、4.4.までに説明してきた方法により、中退した学生の入学から中退した学期までの中退確率を算出する。2 では学生の学期ごとの中退確率を遷移ベクトルとして保存する。入学時から中退した学期までの中退確率を用いて遷移ベクトルを作成する。例えば、2 年目春学期に中退した学生の場合は入学時、1 年目春学期終了後、1 年目秋学期終了後の中退確率の 3 変数を用いて遷移ベクトルを作成する。この場合は 2 年目春学期 ($s = 3$) に中退した学生 m の遷移ベクトルは $(p_0[m], p_1[m], p_2[m])$ を保存することになる。3 では中退した学期ごとにクラスタリングを行っていくために、中退した学期 (1 年次春学期から 4 年次秋学期) ごとに学生を分類する。4 では 2 で作成した中退した学期ごとの学生の遷移ベクトルを 3 で作成した学生の分類ごとに用いて、X-means を行い、学生をクラスに分類する。

上記の 1 から 4 を行い中退した学生を学期ごとにクラスタリングを行ったあと、クラスタごとに学期の中退確率平均値を算出する。表 4.13 は中退した時期 (1 年目春学期から 4 年目秋学期まで) ごとに中退学生のクラスタを縦軸に、横軸はクラスタごとの中退確率を用いたグループ番号を記したものである。前節までと同じように中退確率が 0 以上 0.25 以下の場合はグループ番号 I, 0.25 より大きく 0.5 以下の場合はグループ番号 II, 0.5 より大きく 0.75 以下の場合はグループ番号 III, 0.75 より大きく 1 以下の場合はグループ番号 IV をつけた。例えば、2 年目春学期に中退した学生のクラスタは 2 つになり、2 年目春学期に中退した 46 人の学生の内、学期_クラスタ 1 は 31 人、クラスタ 2 は 15 人に分類できた。

時系列に沿って複数回算出した中退確率を用いてクラスタ作成を行うことで、学生がいつから中退しそうになっているのかを判断することができる。例えば、1 年目春学期に中退した学生のクラスタはクラスタ番号 2 から 4 の 3 つに分類される。クラスタ番号 2 のクラスタは入学前、1 年目春学期終了時、共に中退確率は II である。クラスタ番号 3 のクラスタは入学前の中退確率は II であり、1 年目春学期終了時の中退確率は IV になっている。クラスタ番号 4 のクラスタは入学時、1 年目春学期終了後共に中退確率は IV である。このことから、クラスタ番号 2 は中退する学期まで中退する傾向は見られず、クラスタ番号 3 は入学時ではそれほど中退しそうではなかったが、1 年目春学期に中退しそうな状態になって中退をした形であり、クラスタ番号 4 は入学時か

ら中退しそうな傾向であることがわかる。このように、学期ごとに作成されたクラスタから、同じ学期に中退するとしても、中退しそうな最初と中退までの過程が異なることがわかる。

表 4.13 中退確率の遷移を用いたクラスタ

クラス 番号	中退時期	学期_ クラス	人数	入学時_中退 確率	1年目春学期 終了時_中退 確率	1年目秋学期 終了時_中退 確率	2年目春学期 終了時_中退 確率	2年目秋学期 終了時_中退 確率	3年目春学期 終了時_中退 確率	3年目秋学期 終了時_中退 確率	4年目春学期 終了時_中退 確率	パターン
1	1年目春	1	12	II	II							C
2		1	20	II	II							C
3	1年目秋	2	25	II	IV							B
4		3	3	IV	IV							A
5	2年目春	1	31	II	III	IV						B
6		2	15	II	II	II						C
7	2年目秋	1	31	II	III	IV						B
8		2	20	II	II	II						C
9	3年目春	1	13	II	II	II						C
10		2	8	II	III	IV	IV					B
11	3年目秋	1	10	II	III	IV	IV					B
12		2	10	I	I	I	II	II	II			D
13	4年目春	1	16	II	II	II	II	II	III			C
14	4年目秋	1	10	II	III	IV	IV	IV	IV			B
15		2	7	II	II	I	II	III	III			E
16		3	5	II	I	I	I	I	I			D

4.6. 中退パターンの導出

4.6.1. 中退パターンの導出方法

前節では算出した中退確率を用いて学期ごとの中退者のクラスタを作成した。本節では作成されたクラスタを用い、中退確率の推移に着目した5種類の中退パターンを導出する。

中退パターンの導出の際には3つの時間幅（入学前、大学1年目、2年目以降）に注目して行う。濱名は中退を防止するためには初年次教育が重要だとし、高校と大学における移行の難しさがあると述べる(濱名 2006)。大学に適用するための期間として初年次と初年次以降を比べているように、中退防止のためには初年次教育でどの程度大学に適用させることができるのかが一つの鍵になる。以上を踏まえて中退パターンの類型化において、注目すべき時間軸を入学前、1年目、2年目以降とした。

4.6.2. 導出された中退パターン

① 中退パターン A

中退パターン A は表 4.14 で表されるが、入学前から中退確率が高く、中退まで一度も下がらないパターン（クラスタ番号 4）である。入学前から中退までの中退グループがⅢ以上のグループであり、中退人数の合計 214 人中、パターン A に属する人数は 3 人（1.3%）である。

表 4.14 中退パターン A

クラス タ番号	中退時期	学期_ クラス タ	人数	入学時_中退 確率	1年目春学期 終了時_中退 確率	1年目秋学期 終了時_中退 確率	2年目春学期 終了時_中退 確率	2年目秋学期 終了時_中退 確率	3年目春学期 終了時_中退 確率	3年目秋学期 終了時_中退 確率	4年目春学期 終了時_中退 確率	パターン
4	1年目秋	3	3	Ⅳ	Ⅳ							A

② 中退パターン B

中退パターン B は表 4.15 で表され、1年目春学期から中退まで継続して中退確率が高いパターン（クラスタ番号 2, 5, 7, 10, 11, 14）である。

1年目春学期から中退までの中退グループがⅢ以上のグループをパターン B とした。入学時はグループⅡと中退確率は低いが、1年目春学期から中退確率が高く（中退確率が0.5以上）、中退する直前の学期まで中退確率が0.5から下がることなく中退してしまうパターンである。4つのパターンの内、一番人数が多いパターンであり、中退人数の合計 214 人中、パターン B に属する人数は 110 人（46.6%）である。

表 4.15 中退パターン B

クラス スタ番号	中退時期	学期_ クラス	人数	入学時_中退 確率	1年目春学期 終了時_中退 確率	1年目秋学期 終了時_中退 確率	2年目春学期 終了時_中退 確率	2年目秋学期 終了時_中退 確率	3年目春学期 終了時_中退 確率	3年目秋学期 終了時_中退 確率	4年目春学期 終了時_中退 確率	パターン
3	1年目秋	2	25	II	IV							B
5	2年目春	1	31	II	III	IV						B
7	2年目秋	1	31	II	III	IV	IV					B
10	3年目春	2	8	II	III	IV	IV	IV				B
11	3年目秋	1	10	II	III	IV	IV	IV	IV			B
14	4年目秋	1	10	II	III	III	IV	IV	IV	IV	IV	B

③ 中退パターン C

中退パターン C は表 4.16 で表され、中退まで中退確率が高くないパターン（クラス番号 1, 2, 6, 8, 9, 13）である。

このパターンは入学時には中退確率は高くなく、中退までの途中過程ではグループ II に属しており中退確率が低く中退するか、もしくは中退直前だけ確率が高いパターンである。パターン C に属する人数は 101 人（42.8%）である。

表 4.16 中退パターン C

クラス スタ番号	中退時期	学期_ クラス	人数	入学時_中退 確率	1年目春学期 終了時_中退 確率	1年目秋学期 終了時_中退 確率	2年目春学期 終了時_中退 確率	2年目秋学期 終了時_中退 確率	3年目春学期 終了時_中退 確率	3年目秋学期 終了時_中退 確率	4年目春学期 終了時_中退 確率	パターン
1	1年目春	1	12	II								C
2	1年目秋	1	20	II	II							C
6	2年目春	2	15	II	II	II						C
8	2年目秋	2	20	II	II	II	II					C
9	3年目春	1	13	II	II	II	II	II				C
13	4年目春	1	16	II	II	II	II	II	II	III		C

④ 中退パターン D

中退パターン D は表 4.17 で表され、中退まで中退確率が高くないパターンの 2 種類目である。（クラス番号 12, 16）

このパターン D は中退まで中退確率が高くないパターンで中退パターン C に似ているが、途中過程にグループ I に所属したことがあるパターンである。パターン D に属する人数は 15 人（6.4%）である。

表 4.17 中退パターン D

クラス スタ番号	中退時期	学期_ クラス	人数	入学時_中退 確率	1年目春学期 終了時_中退 確率	1年目秋学期 終了時_中退 確率	2年目春学期 終了時_中退 確率	2年目秋学期 終了時_中退 確率	3年目春学期 終了時_中退 確率	3年目秋学期 終了時_中退 確率	4年目春学期 終了時_中退 確率	パターン
12	3年目秋	2	10	I	I	I	I	II	II			D
16	4年目秋	3	5	II	I	I	I	I	I	I	I	D

⑤ 中退パターン E

中退パターン E は表 4.18 で表され、良化・悪化繰り返しパターン（クラス番号 15）である。

このパターン E は途中学期グループ I，II に所属しており、途中過程で良化・悪化をするパターンである。パターン E に属する人数は 7 人（3.0%）である。

表 4.18 中退パターン E

クラス スタ番号	中退時期	学期_ クラス	人数	入学時_中退 確率	1年目春学期 終了時_中退 確率	1年目秋学期 終了時_中退 確率	2年目春学期 終了時_中退 確率	2年目秋学期 終了時_中退 確率	3年目春学期 終了時_中退 確率	3年目秋学期 終了時_中退 確率	4年目春学期 終了時_中退 確率	パターン
15	4年目秋	2	7	II	II	I	I	II	III	III	III	E

4.6.3. 5 種類の中退パターンのまとめ

表 4.19 は 5 つの中退パターンの中退確率の特徴を示したものである。注目すべき時間幅として入学時の中退確率，1 年目，2 年目以降の中退確率を低い（学生グループ I，II），高い（学生グループ III，IV）で分類した・中退パターン B，C，D で 95% 超の割合を示すため，調査した A 大学ではこのような中退パターンが多いことが分かる。

表 4.19 5 種類の中退パターン

パターン	入学時_中退確率	1年目_中退確率	2年目以降_中退確率	人数	割合
A	高い	高い	該当なし	3	1.3%
B	低い	高い	高い (該当なしの場合もある)	110	46.6%
C	低い	低い	低い (該当なしの場合もある)	101	42.8%
D	低い	低い	低い	15	6.4%
E	低い	低い	高い (上下する)	7	3.0%
合計				236	100.0%

4.7. 第4章のまとめ

本章では学生ごとの中退をするまでの学生の状態推移に注目し、ロジスティック回帰モデルを用いて中退確率 $p_s[m]$ を学期ごとに算出し、入学前や1年目春学期終了時等において学生がどの程度中退をしやすいのかを数値化することができた。学生の状態が悪化/良化しているのかどうかを学生 m ごとの中退確率の遷移 $\Delta p_s[m]$ を用いて数値化した。また、中退確率を用いて学期ごとのクラスタを作成し、5つの中退パターンの導出を行い、中退する学生のパターンを考察した。

学生の状態把握、遷移を把握し、中退のパターンを導き出すことができたが、中退を防止する施策を行うためにはできるだけ早い対応が必要になってくる。特に調査した大学においては1年目春学期終了時には、精度が8割に近い形で中退する学生を推定できてしまうため、当該学期中の対応が求められてくる。

第5章 予測 GPA を用いた学修状態の類型化

本章では予測 GPA を用いて学期内における学修状態を週毎に推定する。次に、学期の最初から最後まで予測 GPA を用い、学修状態の遷移ベクトルとして定義する。その後、個々の学生の遷移ベクトルを用いクラスタリングを行うことで学修状態の類型化を行っていく。

5.1. 1 年次春学期の重要性

令和元年度の学校基本調査によれば、4 年制大学において標準修業年限で卒業した学生の比率（以降、卒業率と呼ぶ）は 81.8%であり、平成 8 年度入学生から平成 27 年度入学生までの卒業率は概ね 80%を推移している（文部科学省 2019）。つまり、約 2 割の学生は 4 年間では卒業できないことを表している。卒業率に大きな関係がある 1 つの指標が成績、GPA (Grade Point Average) である。特に、1 年次春学期の GPA と卒業率、中退率との関係はこれまでの研究で多く述べられており、1 年次春学期の GPA が良い学生は、標準修業年限で卒業する割合が高くなる傾向にある。

図 5.1 は筆者らが調査した A 大学における標準修業年限内に卒業した人（右側：2342 人）としなかった人（左側：1591 人）による 1 年次春学期 GPA の分布を箱ひげ図で表したものである。4 年間で卒業した人の 1 年次春学期 GPA の平均は 2.36、標準偏差は 0.81、4 年間で卒業しなかった人の 1 年次春学期 GPA の平均は 1.12、標準偏差は 0.99 であった。2 群の平均値は有意な差（ウェルチの t 検定 $p < .01$ ）があり、1 年次春学期の GPA は 4 年間で卒業するための重要な指標であることがいえる。

このように 1 年次春学期の GPA は大学を退学、休学せず、標準修業年限内において卒業するための重要な指標である一方で、1 年次春学期の GPA が確定してからでは教学施策としての対応は遅く、早期に対応することが求められる。大河内らは前述した論文において、1 学年春学期の成績が確定してからでは補習授業などにおける学修指導を実施しても成績の回復は難しいことを指摘する。入学直後に実施するプレースメントテストを利用し成績不良の学生を早期に絞り込み、対策をたてる施策を提案している。また、白鳥らは 1 年次春学期以降の GPA や単位数を用いて、中退確率を算出・類型化し、教学施策との連携を提案している（白鳥 et al. 2020）。1 年次春学期の成績や単位数が出ることで 8 割近くの中退者が予測できることを述べ、中退リスクが高い学生には 1 年次春学期 GPA 確定よりも前に対応する必要性を述べている。他にも高岡らは従来、学生支援において学生の状態が悪化したあと事後的に対応することが多かったことを指摘し、修学デー

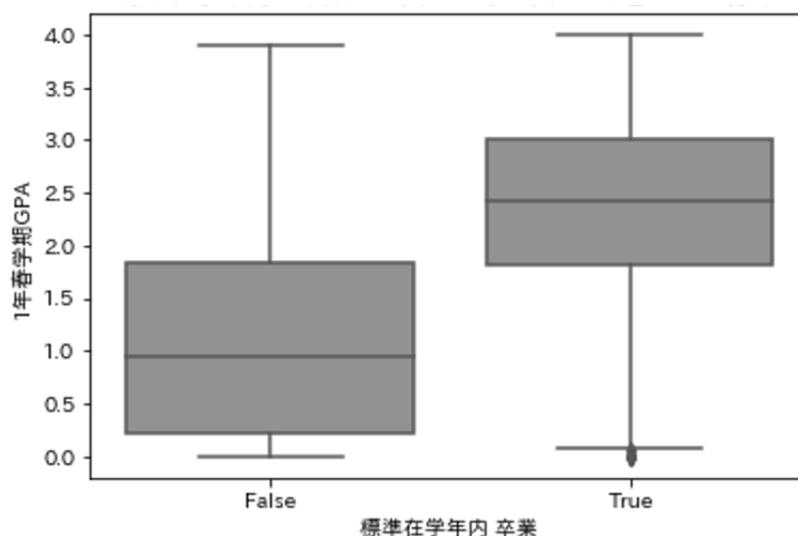


図 5.1 標準在学年内で卒業した人としいない人の 1 年春 GPA の分布

タを用いて学生の状態を事前に予測，判断した形で組織的な施策を行うことを提案している(高岡祥子 et al. 2017).

5.2. 本章の目的

以上のように 1 年次春学期 GPA が決定してからでは教学施策としての対応は遅いため，学期中に学生の状態を把握し，何かしらの対応が求められる．しかしながら，どのような施策を，いつ行うべきなのかは教職員の経験や直感に依存しがちだった．

本章では予測 GPA の推移を用いて 1 年次春学期学修状態の類型化を行う．今回は過去のデータを利用し予測 GPA の推移を類型化することで，過去の学生における学期中の学修状態を推定・把握する．週毎の出欠席データを用いて，1 年次春学期 GPA の予測を週毎に行い，予測 GPA の値を用いて学生ごとに推移ベクトルとしてまとめる．その後，学生ごとの推移ベクトルに対してクラスタリングを用いて類型化を行い，1 年次春学期において，学生がどの程度の数のグループに分割され，どのように状態が推移しているのかを明らかにする．

本手法によって，どの程度の学生数が，どのような学修状態の推移を起こしていたのかを把握することができる．次年度以降に同じ傾向が続くという仮定を置くことにはなるが，過去の学期における学修状態とその推移を把握することで，次学期に，どの程度の学生規模，授業に対して，いつ改善施策を行うことが必要なのかを，教職員はある程度客観的に判断していくことができる．また，大学経営陣や，IR を行う部署 (Institutional Research) としては，過去の年度の施策を振り返り，次年度にどのような教学リソースを，どの程度用意すべきなのかを把握することが可能

になる。

5.3. 利用するデータと変数

5.3.1. 利用したデータ

利用したデータは都内にある A 大学の文系学部にて 2018 年度、2019 年度に入学した学生データである。A 大学は比較的小規模な文系単科大学であり、学生数と 1 年次春学期 GPA の平均値、及び標準偏差を表 5.1 に示す。

表 5.1 第 5 章で利用したデータ

入学年度	2018	2019
サンプル数	281	345
1年次春学期GPA_平均値	2.31	2.35
1年次春学期GPA_標準偏差	1.1	0.95

5.3.2. 利用した変数

利用した変数は 5 種類であり、第 4 章の表 4.2 で示した入学時までに取得できる変数、表 4.4 で示した入学後に取得できる変数 2：学期中の変数を利用する。第 5 章で利用する変数をまとめると表 5.2 で表される。

表 5.2 第 5 章で利用した変数

変数名	値
X_1 : 高校時欠席日数	順序尺度, 4クラス +1クラス (不明他)
X_2 : 評定平均	順序尺度, 4クラス +1クラス (不明他)
X_3 : 全日制高校	名義尺度, 2クラス +1クラス (不明他)
X_4 : 授業の出欠席 (第1週 $X_{4_1,1}$ から第15週 $X_{4_1,15}$)	第1週から第15週まで 名義尺度, 4クラス (出席, 欠席, 不明, 未履修)
Y : 1年次春学期_GPA	連続値, [0,4]

白鳥は中退に関連する変数を大学全体, 国全体に関係する変数であるマクロ変数と個々の学生ごとに変化する変数であるマイクロ変数に分けて整理をしている(白鳥 2016). 今回は大学内において個々の学生状態を把握するために上記のマイクロ変数を用いる. マイクロ変数の中で特に入学時点において取得できる変数として, 高校時欠席日数, 評定平均, 全日制高校を用いた. 以上の変数を用いた理由は, 調査した大学において多くの学生において取得できるものであり, 継続的に取

得できている変数であるために用いた。

x_1 は高校時における欠席日数を学生数で4等分し、順序尺度として割り当てた。2018年の場合は0回以上1回未満、1回以上4回未満、4回以上14回未満、14回以上で分割し、2019年の場合は欠席日数が0回以上1回未満、1回以上4回未満、4回以上12回未満、12回以上で分割した。 x_2 は高校時の評定平均を量で4等分し、順序尺度として割り当てた。 x_3 は全日制高校の出身か、全日制高校の出身でないかの2クラスのいずれかを名義尺度として割り当てた。ただし、 x_i から x_3 のいずれの変数において、データの入力がない場合は不明としてクラスを割り当てた。 x_4 は授業の出欠席を表す変数であり、出席、欠席、不明、未履修の4クラスのうち、いずれかを名義尺度として割り当てた。不明とは授業の履修はしているが、出欠席データの登録がない場合、未履修とは授業の履修登録をしていない場合とした。 x_4 の変数は授業を c 、週番号を w とすると

表 5.3 授業科目名とカテゴリ

年度	科目名	カテゴリ
2018	1.基礎ゼミナール	基礎ゼミ
	2.ICTリテラシ a	ICT
	3.ICTリテラシ b	ICT
	4.英語 1	語学
	5.英語 2	語学
	6.中国語 1	語学
	7.中国語 2	語学
	8.留学生日本語 1	語学
	9.留学生日本語 2	語学
	10.簿記入門	専門基礎
	11.経営学入門	専門基礎
	12.経済学入門	専門基礎
	13.商学入門	専門基礎
2019	1.基礎ゼミナール1	基礎ゼミ
	2.ICT1	ICT
	3.ICT2	ICT
	4.英語1	語学
	5.英語2	語学
	6.中国語1	語学
	7.中国語2	語学
	8.留学生日本語1	語学
	9.留学生日本語2	語学
	10.キャリア設計論基礎1	キャリア

$x_{4,c,w}$ となり、1 番目の授業（2018 年度の場合は基礎ゼミナール）の第 1 週から第 15 週までの出欠席は $x_{4,1,1} \sim x_{4,1,15}$ と表される。 x_4 で利用した授業の科目名とカテゴリ、開講した年度は表 5.3 で表した。

2018 年度入学生の場合は 13 科目、4 カテゴリ、2019 年度入学生の場合は 10 科目、4 カテゴリとした。また、 x_1 から x_3 は入学前カテゴリとして両年度で用いた。開講年度で授業が違う理由は 2018 年度と 2019 年度の間カリキュラム改定があり、1 年次学生が履修する科目が変更になったためである。科目の選定理由は自動登録科目のため、どの 1 年次学生も意識的に履修をやめない限りは、ほぼすべての学生が履修をするためである。また、語学カテゴリの科目は英語、中国語、日本語の 3 つの授業の中からいずれか 1 つを選んで履修する。そのため英語、中国語、日本語のうち履修した科目の 1 つは出席、欠席、不明のうちいずれかが割り当てられ、履修してない語学科目は未履修が割り当てられる。Y は 1 年次春学期の GPA を表し、0 以上 4 以下の連続値として定義した。

5.4. 予測 GPA を用いた学生の類型化

5.4.1. 類型化の手法

前節までに利用したデータと変数を用いて、予測 GPA を算出し、学生の状態を類型化する手法は下記である

- ① 週毎に学生の 1 年春学期の GPA を予測する
- ② 週毎に予測した GPA を用いて遷移ベクトルを作成する
- ③ 遷移ベクトルを用いて、学修状態を類型化する
- ④ 検定を用いて、グループごとの差を抽出する

5.4.2. 1 年目春学期 GPA の予測

前節までに述べた変数を用いて、1 年次春学期 GPA を入学直後（第 0 週目とする）、第 1 週目終了時、第 2 週目終了時と求めていき、第 15 週目終了時まで週毎、学生毎に算出する。第 0 週目の予測の場合には、入学前に利用できる変数 x_1 , x_2 , x_3 を説明変数として用いる。第 1 週目終了時の予測の場合には、 x_1 , x_2 , x_3 に加えて全授業の $x_{4,c,1}$ を説明変数として用い、第 2 週目終了時には x_1 , x_2 , x_3 に加えて、第 1 週目と第 2 週目の出欠席データ $x_{4,c,1}$, $x_{4,c,2}$ を説明変数として用いる。同じように第 15 週目まで学生ごとに GPA の予測を行った。

今回は利用するデータのうちテストデータを 1 つ選び、それ以外を訓練データとして用い予測モデルを作成し、テストデータを用いて予測 GPA を算出する方法 LOOCV (Leave-One-Out Cross Validation: 一個抜き交差検証法) を用いた。LOOCV はモデルの学習に 1 個以外の全てのデータを利用するため、予測精度が高くなるというメリットがある。一方で、デメリットとして全ての

データに対して予測モデルを作成し、予測値を算出するために時間がかかることがあげられる。今回は過去のデータを用いて、当該回時点の GPA を推定し、過去の学生における学修状態を推定・把握していく研究であり、データは 2 年分 ($n=281, n=345$) と多くないため、データをすべて利用したいため、汎化性能を重視する K-分割交差検証法ではなく予測精度を高めるため、LOOCV を用いた。

回帰モデルには線形回帰モデル、サポートベクトルマシンを用いた回帰モデル、ニューラルネットワーク回帰モデル等があるが、本研究ではアンサンブル学習の 1 つであるランダムフォレスト回帰モデルを用いた(Breiman 2001)。このモデルを用いた理由は目的変数に対する説明変数の解釈が可能であること、説明変数と目的変数における非線形な関係を表現できることの 2 点である。これまでに示したように説明変数である出席データと目的変数である 1 年次春学期 GPA は単純な線形関係ではないために本モデルを利用した。

5.4.3. 遷移ベクトルの作成

5-4-2 で求めた週毎、学生毎の予測 GPA 値を用いて学生ごとに遷移ベクトルを作成する。週毎の予測 GPA は i を個人番号、 w は週番号 (0 から 15) とすると $Y_{i,w}$ で表される、学生 i は予測 GPA 値の遷移ベクトルとして $(Y_{i,0}, Y_{i,1}, \dots, Y_{i,15})$ を持つ。

5.4.4. 遷移ベクトルを用いた学修状態の類型化

5-4-3 で作成した学生ごとの予測 GPA の遷移ベクトルを用いてクラスタリングを行い、学生の予測 GPA 遷移の類型化を行う。クラスタリングの手法はクラスタ数を自動で推定する X-means 法を用いた。X-means 法を採用した理由としては学生のグループがどのようになるのかに関する事前知識が存在せず、大学や年度等によってもクラスタ数が変わることが想定されるためである。また、学生ごとの遷移ベクトルは時間と共に変化する時系列データの一部として考えることもできるが、全学生が同じ長さのベクトルを持ち、周期シフトや位相シフト等を考慮する必要がないため k-shape 法や階層 DBSCAN 法等の時系列クラスタリングではない方法を利用した。このように X-means 法を用いて学生の遷移ベクトルを類型化したことで、学生はいずれかのグループに所属することになる。

5.4.5. グループ差の抽出

次に、グループ毎、授業回毎に平均予測 GPA を算出し、グループ毎の予測 GPA の推移の特徴を抽出する。さらに、グループ間で説明変数に差異があるのかをクロス集計を用いて探る。方法としては説明変数ごとに 2 グループ間でクロス集計表を作成し、グループ間の差異があるのかを χ^2 乗検定を用いて行う。3 種以上のグループから 2 種のグループを比較検討していくことになるため、多重性への配慮を Benjamini-Hochberg 法を用いて行う(Benjamini と Hochberg 1995)。Benjamini-Hochberg 法では帰無仮説が棄却される可能性を調整するために調整 P 値 (Q 値と呼

ぶ場合もある) を利用する。本研究では調整 P 値を有意水準 0.05 として利用した。

予測 GPA の算出, 推移ベクトルの作成, 推移ベクトルを用いた類型化, グループ差異の抽出を行うことで, 学期中の予測 GPA の推移を類型化することができ, グループ間の差異をみることでグループごとの特徴を抽出することができる。

5.5. 数値実験と結果

5.5.1. 実験データ

数値実験は下記の 2 パターンで行い, それぞれの年度での類型化の結果を検証する。

1. 2018 年度入学生 (281 人)
2. 2019 年度入学生 (345 人)

5.5.2. 予測精度

それぞれの年度ごとのデータを利用して, 週ごとに回帰モデルを作成する。各週の予測 GPA を算出する回帰モデルの精度を求める際には二乗平均平方根誤差 (RMSE: Root Mean Squared Error) を用いた。

機械学習のツールとしては Python のモジュールである scikit-learn を用いた。Python のバージョンは 3.7.12, scikit-learn のバージョンは 1.0.1, ランダムフォレストのハイパーパラメータはデ

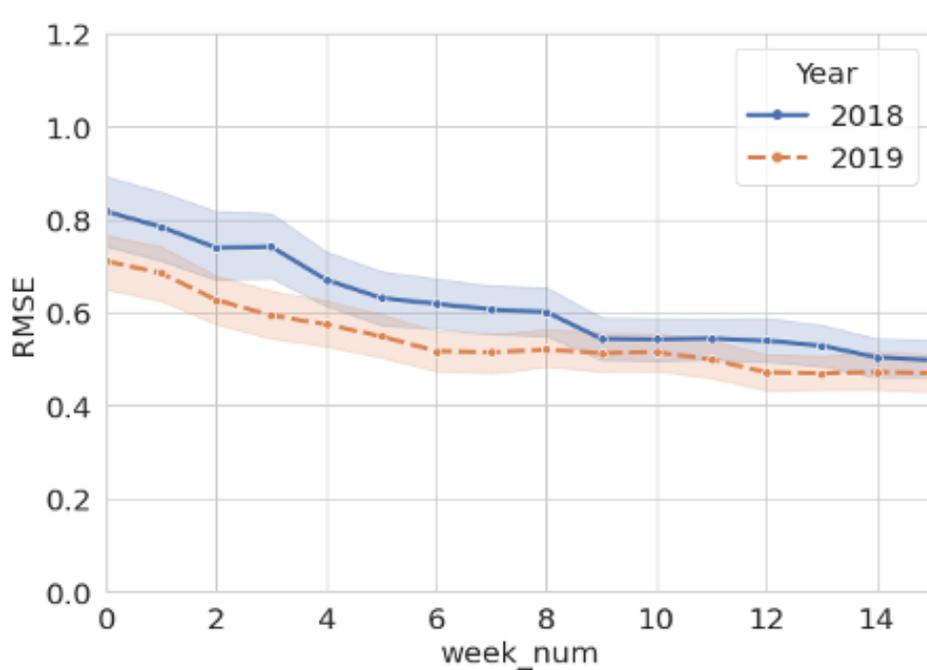


図 5.2 週毎の予測精度

フォルトの設定（決定木の数は 100，最大の深さは None，ノードにおける最小データ数は 1）で行った。

図 5.2 は授業回数ごとの RMSE を表したものである，縦軸は RMSE の平均を表し，横軸は授業回数を表し，入学直後を第 0 回，第 1 週目終了時を第 1 回，第 2 週目終了時を第 2 回とし，第 15 回まで求めた。図を見ると，これまでの研究結果と同じように週毎のデータが増えていくごとに精度は改善されている。2018 年度，2019 年度ともに第 0 回目は RMSE が 0.8 程度から，徐々に下がり，第 8 回目には 0.6 を切り，第 15 回目には約 0.5 まで下がる。どの週においても 2019 年度入学生のほうが 2018 年度入学生よりも精度が若干良い傾向にある。RMSE が 0.8 から 0.5 程度ずれて，信頼区間も 0.1 程度あることから，予測 GPA は 1 程度ずれることが考えられるが，精度が週を追うごとに改善されており，学期全体の推移は把握できるため誤差は許容できると考える。

5.5.3. 類型化の結果

これまでに作成した第 0 回から第 15 回までの予測 GPA 値を用いて，学生ごとに遷移ベクトルを作成し類型化を行った。X-means 法を用いる機械学習のツールとしては Python のライブラリである pyclustering (バージョン 0.10.1.2) を用い，X-means のパラメータとして最初のクラス

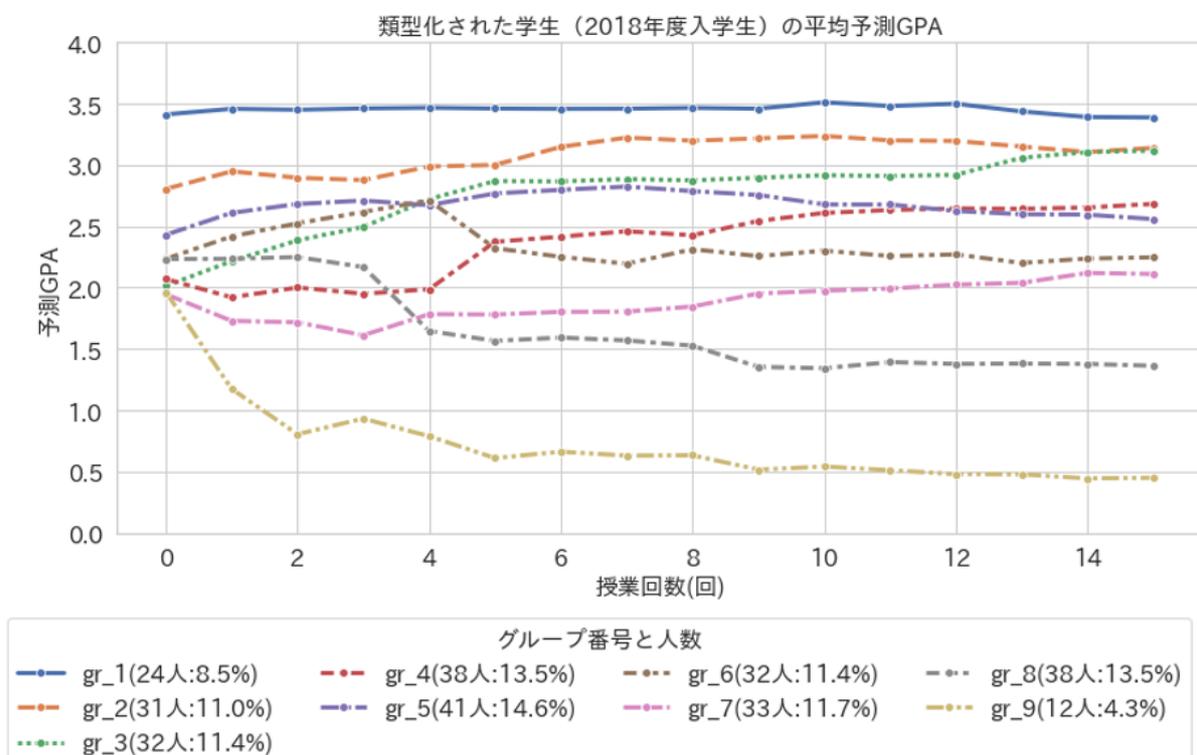


図 5.3 類型化された 2018 年度入学生の平均予測 GPA

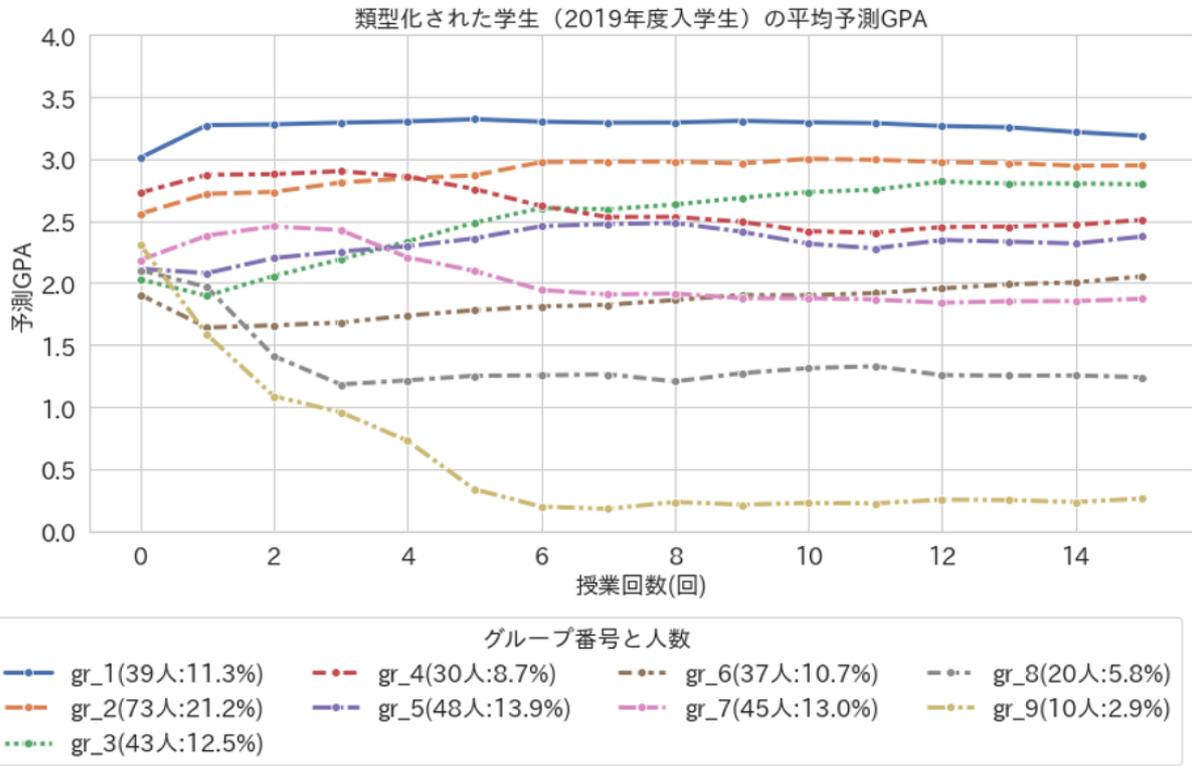


図 5.4 類型化された 2019 年度入学生の平均予測 GPA

数には 1, クラスターの最大数は 20, 反復の停止条件は 0.025, 分割停止基準は BIC (ベイズ情報量基準) を用いた。

2018 年度入学生 281 人を類型化した結果は 9 グループに分かれ, 2019 年度入学生 345 人を類型化した結果は 9 グループに分かれた。図 5.3 は 2018 年度入学生を類型化し, グループごとに各回の予測 GPA 平均値を求め, その推移を表したグラフである。同様に図 5.4 は 2019 年度入学生を類型化し, グループごとに各階の予測 GPA の平均値の推移を表したグラフである。縦軸はグループごとの予測 GPA の平均値, 横軸は各回 (0 回から 15 回) を表している。グループ番号は第 15 回時点での予測 GPA を用いて降順に番号を 1 から付した。

類型化の結果を見ていくことで, 学生の予測 GPA を用いた状態遷移が, どの程度の学生数にあるのかを発見することができる。表 5.4 良化グループと悪化グループの割合に第 0 回時点の予測 GPA と第 15 回時の予測 GPA を比べて, 予測 GPA が上がっていく良化グループと下がっていく悪化グループ, ほぼ変わらないグループ (第 0 回時と第 15 回時の差の絶対値が 0.05 未満) のグループ番号と割合を表した。良化グループ, 悪化グループ共に下の段は予測 GPA が 0.5 以上変化した顕著に良化したグループと顕著に悪化したグループを記した。

良化したグループの割合が高いことは 2018 年度, 2019 年度共に共通して見られる。2018 年度では約 6 割の学生が学期末に向け良化し, 2019 年度では約 7 割の学生が良化した。2018 年度は

表 5.4 良化グループと悪化グループの割合

	2018		2019	
良化グループ	gr_2,gr_3, gr_4,gr_5, gr_7	62.30%	gr_1,gr_2, gr_3,gr_5, gr_6	69.60%
	(gr_3)		(11.4%)	
	そのまま	gr_1,gr_6	19.90%	
悪化グループ	gr_8,gr_9	17.80%	gr_4,gr_7, gr_8,gr_9	30.40%
	(gr_8,gr_9)	(17.8%)	(gr_8,gr_9)	(8.7%)

そのままの学生グループがいたのに対して、2019年度はそのままの学生グループはなくなり、良化グループと悪化グループに2極化したことがわかる。2018年度と2019年度を比較すると、顕著に悪化したグループの学生数が10%程度減ったことや、良化したグループの学生数が7%程度増えたことから2018年度と2019年度で1年次春学期の学生時の状態が変化したことが確認できる。

予測 GPA の推移の類型化を見ていくことで、どの程度の学生がどこで状態が悪くなっているのかを発見することができる。2018年度、2019年度ともに授業回の中盤（第5回、第6回）までは予測 GPA が安定しないが、中盤以降は安定していくことが分かる。例えば、2018年度第0回時と第15回時の予測 GPA の差の絶対値が一番大きいグループはグループ9 (gr_9) だが、第0回時の予測 GPA は1.96、第15回時の予測 GPA は0.45 となり、第0回時と第15回時との差は1.51 となっている。回が進むにつれてその差は徐々に縮まり、第1回時の差は0.72、第2回時は0.35、第5回時は0.16 となっていく。このように両年度ともに授業中盤までには平均予測 GPA は安定し、ほぼ最終的な平均予測 GPA に近づいたものとなる。

調査した大学における2018年度学事日程を見ると、第4回から第5回の間がゴールデンウィーク (GW) の休暇期間であり、2019年度学事日程を見ると、第3回から第4回の間がGWの休暇期間であった。推移結果を見ると、上下する学生はGW前後までに殆どの結果が決まってくることがわかった。

5.5.4. グループ差を用いたハイリスク学生群の特徴抽出

これまでに行った類型化の結果であるグループを用いてグループ間差を全グループ間で比較することができるが、グループ間の差を用いた1つの例としてハイリスク学生群の特徴抽出について述べる。

表 5.5 標準修業年月内で卒業ができなかった学生とできた学生の割合

	できなかった	できた
GPA1.12以上	24.4%	75.6%
GPA1.12未満	82.0%	18.0%

表 5.5 は調査した A 大学において標準修業年限で卒業ができなかった学生の 1 年次春学期の平均 GPA 1.12 を基準にして、卒業できた学生とできなかった学生の割合を示したものである。1 年次春学期 GPA が 1.12 以上の場合には 75.6% の学生が卒業できるが、1.12 未満の場合には 18.0% の学生しか卒業ができない。この場合 GPA 1.12 を分ける差となる授業は何か、どのような特徴を持っているのかを確認するときに、今回作成したグループを用いることができる。1 年次春学期の GPA が 1.12 未満の GPA 下位群に属するグループと他グループの差を抽出することで、ハイリスク学生群がどの段階で学修状態が悪化し、どの授業で差が出るのかを発見することができる。

前節で述べたとおり 2018 年度入学生、2019 年度入学生の両方ともに全グループで予測 GPA が安定するのが第 6 回（授業中盤）以降のため、第 5 回まで（授業序盤）の変数の差を用いる。

最初にどの段階で学修状態が悪くなっていくのかを確認する。2018 年度入学生の場合は第 15 回時点で平均予測 GPA が 1.12 未満のグループはグループ 9 (gr_9) であり第 15 回時点の平均予測 GPA は 0.45 となっている。またグループ 8 (gr_8) の第 15 回時点の平均予測 GPA は 1.12 を超えた 1.36 である。グループ 8 とグループ 9 間の平均予測 GPA を比較すると、第 0 回時点の差は 0.27 であるのに対し、第 1 回時点では 1.06、第 2 回時点では 1.44 と差が広がり、第 4 回目以降は少し差が縮まるものの約 0.9 程度の差が最後まで維持される。また、2019 年度入学生の場合も第 15 回時点で予測 GPA が 1.12 未満のグループは gr_9 である。2018 年度とは少し違い、gr_8, gr_9 間の平均予測 GPA の差は第 3 回目までは 0.3 程度とあまり開かないが、第 4 回時点では 0.49、第 5 回時点では 0.91 と差が開き、第 6 回目以降は約 1.0 程度の差が維持される。両年度ともに授業中盤以降（第 6 回以降）になると学修状態が悪い形で固定化し、戻ることはないことが分かる。

次に、どの授業において差が出ているのかを確認する。両年度共にグループ 8、グループ 9 における平均予測 GPA の差を出している変数は何かを探ることでハイリスク学生群の特徴を探ることができる。

表 5.6 は 2018 年度入学生のグループ 8, 9 において, グループ間の差が有意な変数を週数, カテゴリ, 調整 P 値と共に並べたものである. 2018 年度入学生においてグループ 8, 9 間で有意な差を示した変数には入学前カテゴリ, 基礎ゼミカテゴリ, ICT カテゴリ, 専門基礎カテゴリの授業が入っており, 語学カテゴリ以外の全カテゴリで有意な差がでている. グループ 8, 9 は授業が

表 5.6 2018 年度入学生における下位 2 グループに関連する変数

比較グループ	週数	カテゴリ	変数名	調整P値
グループ8と9	0	入学前	高校時欠席日数	0.010
	0	入学前	全日制高校	0.049
	1	基礎ゼミ	基礎ゼミナール_Attend_1	0.043
	1	ICT	ICTリテラシ a_Attend_1	0.025
	1	ICT	ICTリテラシ b_Attend_1	0.000
	1	専門基礎	簿記入門_Attend_1	0.017
	1	専門基礎	経営学入門_Attend_1	0.007
	1	専門基礎	経済学入門_Attend_1	0.030
	1	専門基礎	商学入門_Attend_1	0.000
	2	ICT	ICTリテラシ a_Attend_2	0.008
	2	ICT	ICTリテラシ b_Attend_2	0.000
	2	専門基礎	経営学入門_Attend_2	0.002
	2	専門基礎	経済学入門_Attend_2	0.002
	2	専門基礎	商学入門_Attend_2	0.001
	3	ICT	ICTリテラシ a_Attend_3	0.004
	3	ICT	ICTリテラシ b_Attend_3	0.013
	3	専門基礎	経営学入門_Attend_3	0.001
	3	専門基礎	経済学入門_Attend_3	0.000
	3	専門基礎	商学入門_Attend_3	0.000
	4	基礎ゼミ	基礎ゼミナール_Attend_4	0.040
	4	ICT	ICTリテラシ a_Attend_4	0.021
	4	ICT	ICTリテラシ b_Attend_4	0.033
	4	専門基礎	簿記入門_Attend_4	0.001
	4	専門基礎	経営学入門_Attend_4	0.000
	4	専門基礎	経済学入門_Attend_4	0.001
	4	専門基礎	商学入門_Attend_4	0.000
	5	基礎ゼミ	基礎ゼミナール_Attend_5	0.009
	5	ICT	ICTリテラシ a_Attend_5	0.011
	5	専門基礎	経営学入門_Attend_5	0.000
	5	専門基礎	経済学入門_Attend_5	0.009
	5	専門基礎	商学入門_Attend_5	0.000

始まる前(0週)から高校時欠席日数と全日制高校が有意な変数としてあげられている。さらにグループ9の学生は1週目から基礎ゼミ、ICT、専門基礎、それぞれのカテゴリにおいて欠席をしているが、グループ8の学生は出席をしていることがわかる。つまり、平均予測GPAの基準である1.12を超えるかどうかの違いは入学前カテゴリと各回の基礎ゼミ、ICT、専門基礎カテゴリの出欠と関連があることがわかる。

表5.7は2019年度入学生のグループ8、9において、グループ間の差が有意な変数を並べたものである。グループ8、9間で有意な差を示したものはキャリア設計論基礎1の第1週、第4週、基礎ゼミナールの第5週である。2018年度とは違いグループ8、9間の平均予測GPAの差は第4週目以降で開いていくが、その要因となる授業がキャリア設計論基礎1と基礎ゼミナール1の出欠にあることがわかる。

表 5.7 2019 年度入学生における下位 2 グループに関連する変数

比較グループ	週数	カテゴリ	変数名	調整P値
グループ8と9	1	キャリア	キャリア設計論基礎1_Attend_1	0.027
	4	キャリア	キャリア設計論基礎1_Attend_4	0.027
	5	基礎ゼミ	基礎ゼミナール1_Attend_5	0.002

両年度の相違点としては、同じグループ9の学生であっても2018年度入学生は、入学前に決まってくる変数(高校時欠席日数、全日制高校)に有意な差が認められ、その後複数のカテゴリの授業の出欠の差が有意に出たのに対し、2019年度入学生は入学前カテゴリの変数は認められず、授業序盤では出席はある程度していることがわかる。この結果から、2018年度のグループ9の学生群は入学前からある程度ハイリスク学生であり、入学後も授業の欠席が多い学生群であるのに対し、2019年度のグループ9の学生群は序盤の授業はある程度出席していることが分かり、同じハイリスク学生群においても2018年度入学生と2019年度入学生の違いを確認することができる。

調査をした大学では基礎ゼミナールの授業にアドバイザーを配置するアドバイザー制度を用いており、基礎ゼミカテゴリの欠席をモニタリングしていた。しかし、以上のことから、基礎ゼミナールの出欠の把握だけでは標準修業年内で卒業をするための1年次春学期GPA1.12以上になるための状態把握は難しい。基礎ゼミナールだけではなく他カテゴリの授業の出欠を学期の序盤からモニタリングすることが必要であることがわかった。

また、アドバイザーは出席状況が悪い学生に対して、面談等の施策を実施し、状況を改善しようとしている。しかし、図5.3、図5.4をみると施策を行っているが一度悪化した予測GPAは授業中盤以降の改善が難しいことがわかる。このことから、授業中盤までの複数授業からのモニタリングだけでなく、悪化した学修状態にさせない別の施策の必要性が示されている。

5.6. 第5章のまとめ

本章では学生の予測 GPA の遷移を用いて学生を類型化する手法を提示し、数値実験の結果によりその有効性を検証した。学期中、週毎に学生の GPA がどのように推移しているのかを予測し、それを推移ベクトルとして捉え、その予測 GPA の推移結果を用いて学生を類型化する手法を提案した。類型化したグループの差異がどこにあるのかを明らかにすることで、学期途中、どの程度学生が、どの回から良化、悪化するのかが明確になり、介入はどこで、いつするべきなのかをハイリスク学生群の特徴を事例として明らかにした。

第6章 まとめと今後の展開

本章ではモデルの利用方法と限界と今後の展開について述べる。最初に、本研究のまとめを第3章で作成したリサーチクエスチョンに答える形でまとめる。次に、これまでに作成したモデルを用いてどのように中退防止策に適応するのかを考察する。その後、本研究の限界を有用性と一般性という2つの観点から述べる。最後に、本研究における今後の展開を有用性と一般性における展開を述べる。

6.1. 研究のまとめ

第3章で提案したリサーチクエスチョンは下記の3点になり、本研究ではリサーチクエスチョンに対して第4章では学期ごとの区分で、第5章では週ごとの区分で解決を試みた。

1. 学生ごとに中退しそうな状態を数値で算出することで、個々の学生がどの程度中退しそうな状態なのかを明らかにする
2. 状態を表す数値を時系列ベクトルで整理をすることで、学生の状態遷移を把握可能にする
3. 時系列ベクトルをクラスタリングすることで中退までの動的特性を明らかにする

第4章では、学期毎の中退確率を用いて状態把握、類型化を行い、中退までの状態把握を行った。個々の学生がどの程度中退をしそうなのかを中退確率として定義し、学生 m の学期 s ごとの学修状態を中退確率 $p_s[m]$ として算出した。ロジスティック回帰モデルを学期ごとに作成し、モデルごとに当該学期 s までの学生のデータをいれることで学生 m が当該学期 s にどの程度中退しやすいのかを算出した。次に、学生の状態が悪化/良化しているのかどうかを学生 m の当該学期 s における中退確率の遷移を $\Delta p_s[m]$ として算出した。学期ごとの中退確率を複数つなげ入学時から中退するまでの学修状態の遷移を把握した。その後、中退した学期ごとに学生の状態遷移を x-means 法を用いて類型化し、どの程度の学生数がどのように中退していくのかを把握した。類型化の結果を用いて5つの中退パターンの導出を行い、中退する学生のパターンを考察した。

第5章では1年次春学期に焦点をあて、学期内における週毎の予測 GPA を用いて状態把握、類型化を行った。学生 m が当該週 w において、どのような状態なのかを1年次春学期 GPA の予測値 $g_w[m]$ を用いて定義した。予測値 $g_w[m]$ を算出する際には、最初に全体のデータから予測したい学生 m のデータを除去したデータを作成した。次に作成したデータを用いてランダムフォレスト回帰モデルを作成し、最後にモデルに対して当該学生 m のデータを入れる、学生 m の週毎の予測 GPA を算出した。次に、週毎の予測 GPA を15週分まとめる形でベクトルとして捉え、学生 m の遷移ベクトルとして定義した。最後に学生 m ごとの遷移ベクトルをデータとして用い、x-means 法

を利用することで類型化した。類型化したグループの差異がどこにあるのかをあきらかにすることで、学期途中、どの程度学生が、どの回から良化、悪化するのかが明確になり、介入はどこで、いつするべきなのかをハイリスク学生群の特徴を事例としてあきらかにした。

以上のように状態の数値化、学生状態の遷移、類型化のリサーチクエスチョン3点について、中退までの学期ごとに学生状態の数値化、遷移、類型化に焦点をあてたものが第4章であり、1年次春学期における学修状態の数値化、遷移、類型化に焦点をあてたものが第5章である。

6.2. 中退防止施策への利用

第4章、第5章では学生を分類し、中退までのパターンと学期内の状態のパターンを作成した。本節では学生グループ、中退パターンごとに、誰が、いつ、どのような中退防止施策を行うべきなのかを考察する。最初に、第4章で作成した中退確率を用いた学生グループの分類によって、誰が中退防止施策を行うべきかを判断する。次に、中退パターンごとの中退防止施策を説明し、中退をしなかった学生のパターンを確かめる。最後に、今回の中退パターンの抽出までをまとめ、他大学でも利用できる道具として提案する。

6.2.1. 中退防止施策を誰が行うべきなのか

藤原ほかによれば、中退防止施策の当事者は行う側の分類により教員、学生相談機関、その他の3種類に分けられる(藤原、富永と押味 2013)。さらに、独立行政法人日本学生支援機構は学生支援の3階層モデルを提唱し、学生支援には第1層：日常的学生支援、第2層：制度化された学生支援、第3層：専門的學生支援があると述べた(独立行政法人 日本学生支援機構 2007)。報告書の中で当事者は心理臨床を主に行うカウンセラー、大学教育に対しての働きかけを行う教員、さらにその中で制度的にアプローチしてくアドバイザー、学生生活支援としての正課外での個別ニーズに応じた働きかけを行う他職員という主に四者がいることも述べている。下記に施策当事者と施策の特徴を記す。

- ・教員（アドバイザー）：教学的、継続的支援
- ・他教員：授業内における教学的支援
- ・カウンセラー：心理的、専門的支援
- ・他職員：正課外における支援

第4章では中退確率を用いて学生グループを作成したが、以上の中退防止施策当事者の施策特徴と連携させ表 4.12 で示した学期間のグループの推移から誰が中退防止施策を行うべきかを考察する。

グループIはそのままの割合が平均で93.2%と高く、また中退に結びつく割合が1.0%と低い。つまりグループの移動が少なく、中退に直接つながる可能性が低いいためグループIに入った学生には他のグループに比べて中退防止施策の優先順位は低い。

一方でグループⅣに入ってしまった学生に対しては中退防止施策を行う必要が大いにある。グループⅣでは推移状態がそのままになっている学生の割合が平均で 58.2%、中退してしまう学生の割合が平均で 29.1%と高いため、優先順位は高く、かつ短期的な支援より良化させるための継続的、かつ集中的な支援が必要となる。以上の理由から一次的な対応を行う教員より、中退防止を専門的に行うカウンセラーを含む学生相談機関の職員が主に対応する必要性が高くなる。

グループⅡとⅢは推移状態がそのままになっているよりも、良化・悪化の割合が平均的に高い。学生の状態を変化させることが短期的に可能なため、一次対応ができる教員（アドバイザー）が中心となって施策を行うべきグループである。

以上のように、中退防止施策は担当者の違いから教員による対策、学生相談機関による対策、その他と分かれているが、グループⅡ、Ⅲの場合は教員が中心に、グループⅣの場合はカウンセラーを含んだ学生相談機関が中心となっておこなうことが必要である。

6.2.2. パターンごとの中退防止施策

次に、中退パターンごとの中退防止施策をいつ、誰が行うべきなのかを考察する。前節で述べたが学生が一度グループⅣに分類されると、中退する割合が高くなるため、如何にグループⅣにさせないかといった施策から検討する。

パターン A は入学時の段階から中退確率が高いため、対策としては高大接続の段階で対策を行うべきである。中退に関連する入学前変数である高校時の欠席日数が多い学生を確認し、入学前に教育を行うことが求められる。担当者は入試担当職員と初年次教育の担当者が適切である。

パターン B は 1 年目春学期の中退確率が高いことから、対策としては初年次教育、特に 1 年目春学期での対策を行うべきである。継続的に 2 年目、3 年目まで中退確率が下がらずに高いままであるため、担当者は初年次教育の担当教員、学生相談機関の職員が適切である。

パターン C は 2 年目秋学期以降の中退確率を下げることを目的とするが、中退学期直前より以前から兆候が出るので、その学期での対策を継続的に行うべきである。担当者は一次対応ができる教員（ゼミナール担当教員）、学生相談機関の職員が適切である。

パターン D は学期の直前まで傾向が出ないため、単位・成績を中心に早期発見する対策というより、成績には現れない学生の状態を把握する学生相談を主に行うべきである。また、学生グループが中程度（グループⅡ）である期間があるため、継続的にグループ 1 に良化させる対策をしていくことが大切である。担当者は一次対応ができる教員（ゼミナール担当教員）、または学生相談機関の職員が適切である。

パターン E は状態が良化・悪化共にするために処理的な対策を行うべきである。パターン D と同じく継続的に見ていく必要があるため、担当者は一次対応ができる教員（ゼミナール担当教員）と学生相談機関の職員が適切である。

表 6.1 は中退パターンごとの中退防止施策の特徴をまとめたものであり、中退パターンごとの割合、時期、担当者を考慮に入れることで、各大学における中退防止施策を判断する有益な情報

になる。全学的に中退防止施策は行っていくことが求められているが、今回の特徴を参考にする
ことで、教職員の人的資源、予算をどの程度割り当てるべきかの参考資料にすることができる。
A 大学では初年次に中退防止施策における 5 割弱の教職員資源を投入し、2 年目以降に 5 割強を
投入することが一つの方法である。

表 6.1 中退パターンごとの施策

パターン	学生割合	施策時期	施策担当者
A	1.3%	入学時以降	入試課職員、初年次教員
B	46.6%	初年次以降 (1 年目春学期)	初年次教員、学生相談機関
C	42.8%	2 年目以降	ゼミナール担当教員、学生相談機関
D	6.4%	2 年目以降	ゼミナール担当教員、学生相談機関
E	3.0%	2 年目以降	ゼミナール担当教員、学生相談機関

6.2.3. 中退をしなかった学生のパターン

中退パターンの把握とともに、中退確率を用いて中退をしなかった学生の状態を類型化するこ
とも有用である。第 4 章では全学生に中退確率の付与を行い、中退した学生の類型化を行ったが、
ここでは中退確率を算出した後、中退をしなかった学生に焦点をあてて類型化を行う。中退をし
なかった学生の中で、中退をしそうだった学生がどの程度いて、どのように卒業まで学業を続け
たのかを把握する。

図 6.1 は中退確率を用いて卒業した学生の類型化をおこない、クラスごとに平均値を表したも
のである。人数が多い順にクラス 1 からクラス 5 と名付けてある。横軸は学期（1_1 が入学時、
1_2 が 1 年目春学期終了時、4_2 が 4 年目春学期終了時）を表し、縦軸はクラスごとの中退確率
の平均値を表している。

1 年目春から卒業まで中退確率が低いままであるクラスが 1 であり 639 人中 364 人（60.0%）
になる。約 6 割の卒業生は中退確率が 0.1 より高くなることなく卒業していくグループである。
クラス 2 はクラス 1 より 1 年目の中退確率は 1 年次春開始時は 0.1 ほど高いが、クラス 1 と同
じように学年が進行していくにしたがって徐々に下がっていくというクラスであり、185 人
（29.0%）いる。クラス 1 とクラス 2 を足すと 549 人（86.0%）であり、卒業生の内約 8 割の学
生は入学してから卒業するまで順調に推移している学生である。

一方でクラス 1 と 2 の学生のように順調に卒業まで結びつかないクラスが 3 から 5 の学生であ
る。クラス 3 は 44 人（6.9%）であり、1 年目春学期の中退確率が 0.44 と高く、そこから 2 年次
まで中退確率が相対的に高いクラスである。クラス 3 に所属する学生は 1 年目こそ高いが、クラ

ス1、クラス2と同じように学期ごとに中退確率が下がっていったグループである。

一方でクラス4と5は両方共1年目春学期から一時中退確率が高くなり、ある学期から下がっていくグループである。クラス4は27人(4.2%)であり、1年目から中退確率が高くなっていき、2年目春学期にピークを迎え、2年目秋以降に良くなっていったグループである。クラス5は19人(2.9%)であり1年目は中退確率が高く、2年目以降に持ち直していくグループである。

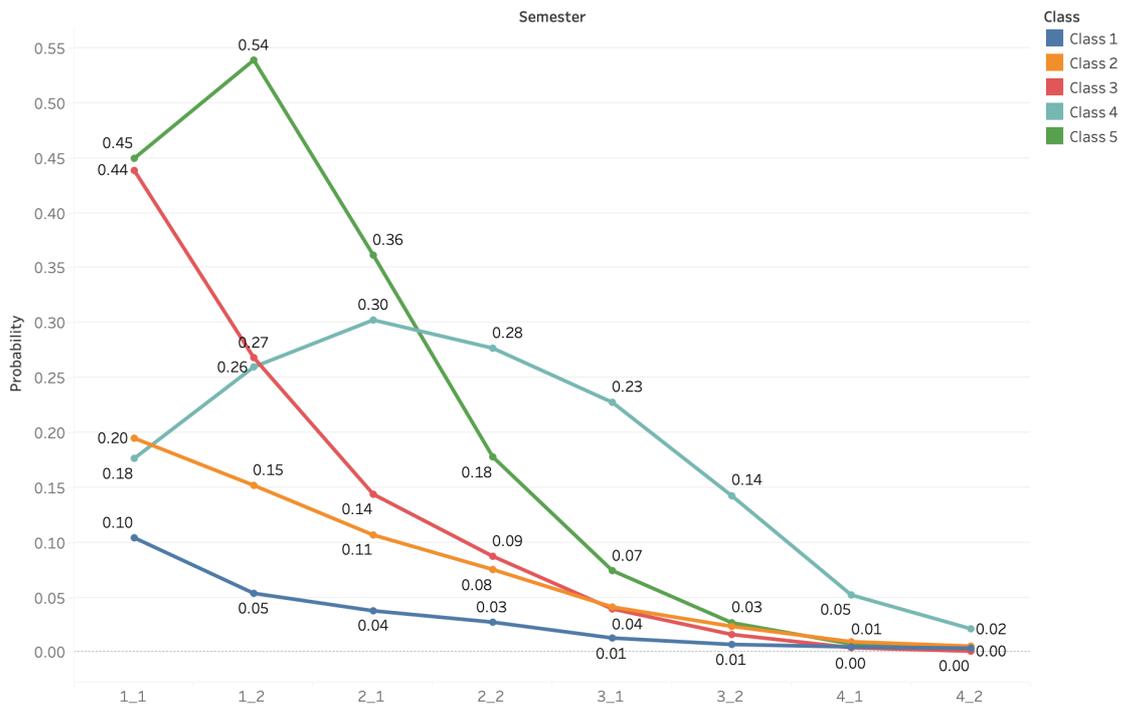


図 6.1 中退をしなかった学生の類型化

中退確率0.3を超えた学生はある程度リスクがあるとする、クラス3(6.9%)、クラス4(4.2%)、クラス5(2.9%)、合計14%の学生はハイリスクな状態であったにもかかわらず、卒業まで結びつけたことが分かる。

最初からハイリスクだったクラスはクラス3とクラス5である。クラス3は学期ごとに徐々に下がってきているため、入学時点ではハイリスクだったが、大学への移行が順調で、初年次教育がうまくいった学生である。クラス3と同じく1年次春学期時点ではハイリスクで、最初の学期にリスクが下がらなかった群がクラス5である。クラス5は最初はクラス3と同じ程度の中退確率であり、特に最初の学期にリスクがあがり、初年次の中で持ち直した学生であった。最初からハイリスクな学生に対しては、初年次の段階で下げていることが分かる。

入学時にはリスクが低かったが在学中に高くなったクラスはクラス4である。クラス4は最初こそクラス2と同じ程度の数値であったが、初年次の学期から2年次までである程度高い学生であ

ったが、卒業に結びついた学生であった。

このように大学に入ってから卒業するまでの学生状態の遷移を類型化することで、卒業した学生でありながら中退に近づいたのか学生がどの程度の数いたのかを明らかにし、どのように卒業に結びついたのかを考察した。

6.2.4. 中退パターンを用いた中退防止施策

中退パターンを利用した中退防止施策を実施するためには各大学において4つのステップが必要になる。

1. 過去データを用いて学期ごとに中退確率を算出する
2. 学生の中退確率に対して学期ごとにクラスタリングを行い、学期ごとにどのようなクラスタがあるのかを確かめる
3. 各クラスタから各大学の中退パターンを抽出する
4. 中退パターンごとに中退防止施策のリソース配分を行う

1. ではこれまでの過去の学生データを用いて、それぞれの大学における中退確率を算出する。それぞれの大学ごとに利用できるデータは違うが、大学ごとに利用できる入学時までに取得できる変数と入学時から当該学期までに取得できる変数を用いて中退確率を算出する。2. では中退確率を用いて、学期ごとにX-means法を用いてクラスタリングを行い、学期ごとの中退者の特徴を把握する。3. では2で作成したクラスタから特徴を把握し、中退学生の類型化を行い、中退パターンを抽出する。4. では学生のグループ、中退パターンに応じて大学における中退防止施策のリソース配分を決定する。以上の4段階を踏まえることで、各個別の大学ごとに中退防止施策を客観的なデータから判断し、実施することが可能になる。

上記で示したように、3. 中退パターンの抽出までは大学ごとに取得できるデータの違いはあるが、本研究と同様のデータを取得できれば、ある程度客観的に類型化を行うことができると考えられる。中退者に関連するデータが足りない場合は、表6のように中退パターンが分かれな場合もある。その際には中退者のデータ量を増やすか、目的変数を中退に関連する変数から各学年の成績不振者等に変更することで量を確保していくことでパターン分割が可能になると考えられる。

しかしながら、4. 中退防止施策のリソース配分では抽出された中退パターンに応じて誰が、どの程度、どの中退防止施策を行うのかを決めていくが、その際にはそれぞれの大学における状況を考慮に入れ、ある程度恣意的に施策を決定していかなくてはならない。例えば、初年次教育を行える教員数や、カウンセラーの人数等を考慮に入れた上で、中退パターンの数やパターンに対応する中退防止施策をおこなうかを決める必要がある。以上のように、中退パターンの導出まではある程度客観的に可能だが、中退防止施策を行う際には5章で述べたような主観的な部分が残

ることが本研究の限界である。

6.3. 本研究の限界

6.3.1. 有用性における限界

本研究では学生の学修状態に注目し、ロジスティック回帰モデル、またはランダムフォレストモデルを用いて数値化を行い、時系列データとしてまとめ、類型化を行った。第4章では中退確率を学期ごとに算出することで中退までに学生がどのようなパターンがあるのかを明らかにする方法を提案した。本手法を用いて、中退が行われるまでの学生の状態を5パターンに分類し、中退までにどのような動きを経て中退をするのかを明らかにすることができた。学期中、学生の状態を把握する手法として、学期中の授業出欠データを利用して、学期ごと、学生ごとのGPAを週毎に予測し、予測GPAの推移を類型化する方法を提案した。1年次春学期において学生がどの程度の数のグループに分割され、どのように状態が推移していくのかを明らかにした。

第4章で述べた中退確率を用いて行う類型化、第5章で述べた予測GPAを用いて行う類型化、のどちらにしても過去のデータを利用した定量化であるという部分が本研究における有用性の限界である。第4章では学生が卒業・中退したあと、第5章では学期終了後と、過去のデータを用いて、過去の学修状態に対して類型化をしているため、大学内にどのようなグループがいて、どのような特徴を持っているのかは把握できる。定量化を行うことで、現状の大学における学生をある程度定量化することはできた。つまり、一般的にどのような中退予防施策を行うことが必要なのかを判断することができた。

しかし、中退を予防するという観点から本研究を見ると、定量化で行う過去の状態の把握のみでは、現状の学生に対してどのような中退防止施策を行うべきかという部分に結びつけるのは難しい。理由は大きく下記の2つである

どの学生から施策を行うべきなのかは分からない

本研究は過去のデータを利用して、過去の状態を定量化しているため、いま目の前にいる学生がどの程度の状態であり、どのような遷移を経るのかを把握することは難しい。過去にどのような学生状態があり、それを類型化してどのようなグループがあり、どのような遷移をとるのかというところまでである。どの学生から施策を行うべきなのかは本研究の範囲では難しい。

どの施策を行うべきなのかは分からない

学生群がどのような状態の遷移になるのかまでは定量化により把握できたが、個々の学生にはどのような施策が必要なのかを判断することは主観的な部分が残る。中退予防施策を行うためには、大学内にどの程度の予算があり、どの程度の人材がある、といったものが大きく関連する。

このように定量化までは本研究までに可能になったが、中退予防を実質化していくためには、教
学リソースや予算、学生の中退理由、当該学生関係者（保護者や友人）との関係改善、等の他要
素を鑑みて施策を実施するのかを判断し、施策を行ったあとにどのような指標でその施策を評価
していくのかを決め、実施していく必要が出てくる。

6.3.2. 一般性における限界

次は本研究における限界の2つめである一般性における限界である。本研究は調査した1大学
のデータを利用したモデル・事例であるため、他の大学でそのまま利用できるのか、大学全体と
して言えることなのかを判断することは難しい。

しかしながら、本研究では大学で一般的に取得しているデータを利用し、定量化を行っている
ため、本研究で提案するモデルを他の大学で作成することは容易である。第4章、第5章共に大
学にあるIR組織なら通常収集しているデータになっている。大学において、教学関係のデータを
収集し、分析をおこなっていく教学IR組織があることは前提となるが、データを収集している組
織がある場合には、今回のモデルで利用するデータの保持は可能なものとなっている。

また、大学全体として、一般的な大学としてのモデルとして提示することは難しい。今回調査
した大学は社会科学系の学部であり、規模としては小規模である。今回作成したモデルを他の大
学で利用する場合には、今回のモデルで利用していない、大学の分野や、大学の学生規模、留学
生比率、教職員学生比率といった大学全体の変数を利用した大学のデータを準備しなくてはなら
ない。他の大学のデータを用いて、今回のモデルを作成することができれば、より一般的に学生
状態を把握すること、大学固有の特徴を抽出することができる。

さらに、本研究で作成した学生モデルではすべての学生に適用することは難しい。中退には学
生の健康状態や、保護者の学歴、経済的状态等、学生ごとの変数が中退という意思決定に関連す
る。今回の研究では、個々の学生データを利用することは行ったが、より精緻に把握するため
には、上記のような学生に関連した健康データや当該学生以外との関係性等のデータを利用する必
要がある。類型化によってグループにはどのような特徴があるのかまでは把握することができた
が、なぜその学生が中退をするのかを加味した上でモデルを作成していくことが必要である。

6.4. 今後の展開

6.4.1. 有用性における展開

今後の有用性における展開としては、今回提案した状態の数値化、遷移の把握、パターン化を
用いて、1. 状態把握から予測研究につなげること、2. どのような学生がどの段階で変化を起
すのかを明らかにし、中退防止の効率的な施策につなげていくこと、の2点を考えている。

状態把握から予測へ

そのためには定量化による把握から、予測を利用した学生状態のリアルタイムな把握をする必要がある。今回のモデル作成で利用したデータ等を利用して、学期ごと、週ごとに学生の状態をリアルタイムで予測していく。本研究では、昨年度までのデータを用いてどの程度の数の学生が、どのような状態になっているのかまでは把握できたが、学期途中において、当該学生がどの程度中退リスクを保持しているのかは、予測モデルを逐次適応させなくては把握できない。過去のデータを用いながら、現在の学生の状態を数値化すること、そして、数値化した後どのような遷移をとるのかを把握することを行い、今回作成された類型化によるグループのどこに所属しているのかを学生ごとに算出していく。

中退防止施策との結合へ

さらに、週毎に説明変数の解釈を行うことで予測 GPA の良化・悪化の原因を分析し、より中退防止という実務で利用できるものを考えていきたい。今回は予測 GPA を算出する際に入学直後に取得できる高校時欠席日数、評定平均、全日制高校と各回の出欠席データのみを用いたが、課題の提出状況やログイン時間等の LMS データを用いることで予測 GPA をより精度が高く算出することができる。他にも、カウンセラーとの面談記録や健康アンケート等を用いることで、学修状態の類型化ではなく、学生の精神状態の推定も可能になると考える。また、今回は学期終了後にどのような状態だったのかを推測する手法になったが、今後は前年度の予測モデルを利用して、逐次、学生がどのような状態にあるのかを推定していくと共に、週毎に説明変数の解釈を行うことで予測 GPA の良化・悪化の原因を分析し、より中退防止という実務で利用できるものを考えていきたい。

6.4.2. 一般性における展開

中退率はそれぞれの大学によって大きく変わり、平均で約 2%程度ということは第 1 章で述べた。約 2%の中退する学生への施策は、それぞれの大学全体としての施策として考えると効率的にはよくない。大学全体の経営的視点を持つと、中退しそうな学生への施策というより、まずは大きな群に対する教学施策から行っていくことは理解できる。

一方で、日本全体では約 57000 人が大学・短大を中退していく。1 大学としては約 2%なので、例外処理的な対応になることは理解できるが、日本全体としての中退学生の数が多い。さらに平均として 2%なので、第 1 章でも述べたとおり中退率が高い大学は多い。偏差値が 45 程度の大学は中退率が 10%を超えることも多い。このように中退率が高い大学は多くあるにも関わらず、日本においては中退率とその施策は公開されづらかった。中退率が高く、それを低くしてきた大学ではどのような中退防止施策を行っているのか、中退率を下げるためにどのような施策をしているのかといったものが共有されづらい。このように中退防止に取り組むべき大学は多いにもかか

ならず、個々の大学のみの方策では難しい。大学間を連携した中退予防ネットワーク、中退防止方策等を展開する必要がある。

また、学生が中退をしたあとに他の大学に移動、もしくは就職して生活を行う等、より良い生活が待っているならば中退をするという選択もありえる。このような中退は移動であり、中退をすることでよりよい生活を選択する判断である。これは良い中退である。

良い中退なのか、それともそうでないのかの判断は付きづらい。良い中退ではなく、消極的な選択としての中退が多い中では中退を選択しないという方向に促し、中退防止方策を行っていくことが多い。中退リスクのある学生に対して、よりよい方策を考えることが、誰一人とりこぼさない教育という考えからは正しい。中退をしたあとに他の大学に移動することが容易、他の生活にスムーズに移行できるようになるためには、個々の大学のみでは難しい。こちらの意味でも大学間を連携した取り組みをしていく必要がある。

また、教学 IR のみの視点ではなく、個々の大学内でより一般性をもった方策を行う教学マネジメント組織、質保証を担当する組織と連動する必要がある。今回のようなハイリスク学生の把握は教学部門、IR 部門のどちらが行っても問題は無いが、それと連動した中退防止方策までを考えると教学を実際に担当する組織や教学マネジメントを担当する組織、カウンセラー等が所属する学生相談部門、高校等とつながっている入試部門、大学全体のマネジメントをする経営部門等の他組織との連携をすることが大学の中で一般的に方策を行うときに必要である。今回のようにデータを用いたモデル作成をもとに、大学内にある複数の組織が連携して中退防止、教学方策を行っていくことで、一般性と実質性をもった教学方策をおこなっていくことができるようになる。

謝辞

本研究を進めるにあたり、指導教官として懇切丁寧なご指導を頂きました、東京工業大学リベラルアーツ研究教育院の室田真男教授に、心から厚く御礼申し上げます。

本研究の遂行にあたり、終始懇切丁寧なご指導、そしてお力添えを頂きました、東京工業大学企画本部の森雅生教授、九州工業大学学習教育センターの大石哲也教授、北陸大学経済経営学部
の田尻慎太郎教授に、厚く御礼申し上げます。

本研究をまとめるにあたりご指導と貴重なご助言をいただきました、東京工業大学リベラルアーツ研究教育院の調麻佐志教授、東京工業大学リベラルアーツ研究教育院の猪原健弘教授、東京工業大学リベラルアーツ研究教育院の江原慶准教授、東京工業大学リベラルアーツ研究教育院の永原健太郎講師に、厚く御礼申し上げます。

最後に、あたたかい励ましをいつも送り続けてくれ、支えてくれた家族に心から感謝します。

本研究に関わる研究発表

査読付き論文

白鳥成彦, 大石哲也, 田尻慎太郎, 森雅生, & 室田真男. (2022). 予測 GPA の推移を用いた 1 年次春学期学修状態の類型化. 教育システム情報学会誌, 39(4), 440–451. 【第 5 章】

白鳥成彦, 大石哲也, 田尻慎太郎, 森雅生, & 室田真男. (2020). 中退確率の遷移を用いた中退学生の類型化. 日本教育工学会論文誌, 44(1), 11–22. 【第 4 章】

査読付き国際会議

Shiratori, N. (2019, January 1). Analysis of the Influence of Learning State before University Admission to College Dropout Using Hierarchical Bayesian Model. NPSE.

Shiratori, N. (2018). Derivation of Student Patterns in a Preliminary Dropout State and Identification of Measures for Reducing Student Dropouts. 2018 7th International Congress on Advanced Applied Informatics (IIAI-AAI) IS - SN - VO - VL -, 497–500.

Shiratori, N. (2017). Modeling Dropout Behavior Patterns Using Bayesian Networks in Small-Scale Private University. 2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI) IS - SN - VO - VL -, 170–173.

国内口頭発表

白鳥成彦. (2021a). 中退防止に関連する変数の整理. 大学情報・機関調査研究集会 論文集, 10, 100–105. 【第 2 章】

白鳥成彦. (2021b, March 24). データを利用した中退防止へ -中長期の組織的取り組みを目指す. 教育学術新聞, 2836, 3.

白鳥成彦. (2020). 日本における中退防止施策の類型化. 大学情報・機関調査研究集会 論文集, 9, 100–105. 【第 2 章】

白鳥成彦. (2019). 中退理論と教学 IR データのマッチングによる中退防止の検証. 大学情報・機関調査研究集会 論文集, 8, 32–37. 【第 2 章】

白鳥成彦. (2019). 入学前データと出席率を用いた初年次学生の中退予備状態推定. In 計測自動制御学会 システム・情報部門学術講演会 2019 (pp. 915–917).

- 白鳥成彦. (2018). 中退防止における 2 つの IR アプローチ：高大接続アプローチと教学アプローチ. 第 7 回 大学情報・機関調査研究会. 第 7 回 大学情報・機関調査研究会.
- 白鳥成彦. (2016). ベイジアンネットワークによる大学生の中退行動のモデル化. 人工知能学会全国大会論文集, *JS AI2016*, 3F35–3F35.
- 白鳥成彦, & 田尻 慎太郎. (2016). 小規模私立大学における初年次教育データを用いた IR 分析：嘉悦大学と横浜商科大学の事例を参考にして (日本教育情報学会第 32 回年会) -- (教職開発研究会). 日本教育情報学会第 32 回年会, 32, 38–41.

参考文献

- Aljohani, Othman. 2016. 「A Comprehensive Review of the Major Studies and Theoretical Models of Student Retention in Higher Education」. *Higher education studies* 6 (2): 1–18.
- Astin, Alexander W. 1999. 「Student involvement: A developmental theory for higher education」. *Journal of college student development* 40 (5): 518–29.
- Aulck, Lovenoor, Nishant Velagapudi, Joshua Blumenstock と Jevin West. 2016. 「Predicting Student Dropout in Higher Education」. *arXiv.org stat.ML* (6 月). <http://arxiv.org/abs/1606.06364v4>.
- Bäulke, Lisa, Carola Grunschel と Markus Dresel. 2021. 「Student dropout at university: a phase-orientated view on quitting studies and changing majors」. *European Journal of Psychology of Education* 37 (3). <https://doi.org/10.1007/s10212-021-00557-x>.
- Bean, John P. 1980. 「Dropouts and turnover: The synthesis and test of a causal model of student attrition」. *Research in higher education* 12 (2): 155–87.
- . 1982. 「Conceptual models of student attrition: How theory can help the institutional researcher」. *New directions for institutional research* 1982 (36): 17–33.
- Bean, John P., と Barbara S. Metzner. 1985. 「A Conceptual Model of Nontraditional Undergraduate Student Attrition」. *Review of educational research* 55 (4): 485–540.
- Benjamini, Yoav, と Yosef Hochberg. 1995. 「Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing」. *Journal of the Royal Statistical Society* 57 (1): 289–300.
- Berger, J. B., G. B. Ramirez と S. Lyons. 2012. 「Past to present」. *College Student Retention*, 編集者: Alan Seidman, 7–34. books.google.com.
- Bingham, Melissa A., と Natalie Walleser Solverson. 2016. 「Using Enrollment Data to Predict Retention Rate」. *Journal of student affairs research and practice* 53 (1): 51–64.
- Breiman, Leo. 2001. 「Random Forests」. *Machine learning* 45 (1): 5–32.
- Cabrera, Alberto F., Amaury Nora と Maria B. Castaneda. 1993. 「College persistence: Structural equations modeling test of an integrated model of student retention」. *The Journal of higher education* 64 (2): 123–39.
- Del Bonifro, Francesca, Maurizio Gabbrielli, Giuseppe Lisanti と Stefano Pio Zingaro. 2020. 「Student Dropout Prediction」. *Artificial Intelligence in Education*, 編集者: Ig Ibert Bittencourt, Mutlu Cukurova, Kasia

- Muldner, Rose Luckin と Eva Millán, 129–40. Springer International Publishing.
- Hellas, Arto, Petri Ihantola, Andrew Petersen, Vangel V. Ajanovski, Mirela Gutica, Timo Hynninen, Antti Knutas, Juho Leinonen, Chris Messom と Soohyun Nam Liao. 2018. 「Predicting academic performance: a systematic literature review」. *Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education*, 175–99. ITiCSE 2018 Companion. New York, NY, USA: Association for Computing Machinery.
- Lopez Guarin, Camilo Ernesto, Elizabeth Leon Guzman と Fabio A. Gonzalez. 2015. 「A Model to Predict Low Academic Performance at a Specific Enrollment Using Data Mining」. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje* 10 (3): 119–25.
- Manski, Charles F. 1989. 「Schooling as experimentation: a reappraisal of the postsecondary dropout phenomenon」. *Economics of education review* 8 (4): 305–12.
- National Center for Education Statistics. 2022. 「Undergraduate Retention and Graduation Rates」. U.S. Department of Education, Institute of Education Sciences.
- Pascarella, Ernest T., と Patrick T. Terenzini. 1980. 「Predicting Freshman Persistence and Voluntary Dropout Decisions from a Theoretical Model」. *The Journal of higher education* 51 (1): 60–75.
- Pelleg, Dan, と Andrew W. Moore. 2000. 「X-means: Extending k-means with efficient estimation of the number of clusters」. *ICML '00: Proceedings of the Seventeenth International Conference on Machine Learning*, 1:727–34. cs.dal.ca.
- Seidman, Alan. 2012. 「Taking Action: A Retention Formula and Model for Student Success」. *College student retention: Formula for student success*, 1 月, 267–84.
- Shum, Simon Buckingham. 2012. 「Learning Analytics」. *Perspectives*. <http://iite.unesco.org/pics/publications/en/files/3214711.pdf>.
- Siemens, George, と Ryan S. J. d. Baker. 2012. 「Learning analytics and educational data mining: towards communication and collaboration」. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 252–54. LAK '12. New York, NY, USA: Association for Computing Machinery.
- Spady, William G. 1970. 「Dropouts from higher education: An interdisciplinary review and synthesis」. *Interchange* 1 (1): 64–85.
- Tinto, Vincent. 1975. 「Dropout from higher education: A theoretical synthesis of recent research」. *Review of educational research* 45 (1): 89–125.

- . 1993. *Leaving college: Rethinking the causes and cures of student attrition*. The University of Chicago Press.
- . 2012. 「Moving from theory to action」. *College student retention: Formula for student success*, 1月, 251-66.
- 宇治川正人, 丸山玄. と 讃井純一郎. 1999. 「電子メールを用いた評価グリッド法の開発: 魅力あるガソリンスタンドの条件に関する調査研究」. 日本建築学会計画系論文集 64 (518): 75-80.
- 河野銀子. 2003. 「大学大衆化時代における'First-Generation'の位相」. 山形大学紀要 教育科学 13 (2): 127-43.
- 丸山文裕. 1984. 「大学退学に対する大学環境要因の影響力の分析」. 教育社会学研究 39 (0): 140-153-en309.
- 岩崎保道. 2015. 「大学における休・退学防止の検討: 学内組織連携型の学生支援策に注目して」. 関西大学高等教育研究 6: 81-86.
- 岩崎保道, 宮嶋恒二, 蔭久孝政, 福島謙吉と谷ノ内識. 2016. 「中途退学の防止についての一考察」. <http://www.kochi-u.ac.jp/daikyo/publication/pdf/pedagogics05.pdf>.
- 近藤伸彦. 2020. 「教育/学習における予測モデルの活用」. 教育システム情報学会誌 37 (2): 93-105.
- 近藤伸彦, と 畠中利治. 2016. 「学士課程における大規模データに基づく学修状態のモデル化」. 教育システム情報学会誌 33 (2): 94-103.
- 窪内節子. 2009. 「大学退学とその防止に繋がるこれからの新入生への学生相談的アプローチのあり方」. 山梨英和大学紀要 8: 9-17.
- 高橋望, 藤本裕介と西本裕輝. 2019. 「琉球大学学士課程における退学・休学・除籍・留年の早期発見に向けた検討: 退学等に至る学生の初年次前期の GPA と入学動機の特徴の可視化の試み」. 琉球大学大学教育センター報= *University Education Center Bulletin*, no. 21: 89-100.
- 讃井純一郎, と 乾正雄. 1986. 「レパトリー・グリッド発展手法による住環境評価構造の抽出: 認知心理学に基づく住環境評価に関する研究 (1)」. 日本建築学会計画系論文報告集. https://www.jstage.jst.go.jp/article/aij/367/0/367_KJ00004066580/_article/-char/ja/.
- . 1987. 「個人差および階層性を考慮した住環境評価構造のモデル化: 認知心理学に基づく住環境評価に関する研究 (2)」. 日本建築学会計画系論文報告集. https://www.jstage.jst.go.jp/article/aij/374/0/374_KJ00004066727/_article/-char/ja/.
- 姉川恭子. 2014. 「大学の学習・生活環境と退学率の要因分析」. 経済論究, no. 149 (7月): 1-16.
- 松田岳士, と 渡辺雄貴. 2018. 「教学 IR, ラーニング・アナリティクス, 教育工学」. 日本教育工学

- 会論文誌 41 (3): 199-208.
- 清水一. 2013. 「大学の偏差値と退学率・就職率に関する予備的分析:社会科学系学部のケース」. 大阪経大論集 64 (1): 57-70.
- . 2021a. 「大学生の就職・進学要因の考察:社会科学系のケース」. 大阪経大論集 71 (6): 267-82.
- . 2021b. 「大学生の退学要因の考察:社会科学系学部のケース」. 大阪経大論集 71 (5): 1-13.
- . 2021c. 「偏差値と退学率・就職率・進学率の関係:理工系学部のケース」. 大阪経大論集 72 (2): 165-80.
- 川崎孝明, 中嶋弘二, 川嶋健太郎と川口恵子. 2014. 「大学における寄り添い型学生支援体制の構築-中途退学防止の観点からの実践的アプローチ-」. 尚絅大学研究紀要 A. 人文・社会科学編 46: 75-89.
- 大河内佳浩, と山中明生. 2016. 「プレースメントテストや高校の履修状況などのデータを用いた初年時成績不振者の早期発見」. 日本教育工学会論文誌 40 (1): 45-55.
- 中井俊樹, 鳥居朋子と藤井都百. 2013. 大学のIR Q&A. 玉川大学出版部.
- 中島絵美. 2012. 「時代や大学の特色に応じた学生相談の取り組みの検討 大学生不適應の予防的アプローチ」. こども教育宝仙大学紀要, no. 3 (1 月): 131-37.
- . 2013. 「学生相談の取り組みの検討— 本学における大学不適應とは—」. こども教育宝仙大学紀要 4: 87-96.
- 辻村壮平. 2017. 「階層的に構造化された評価を引き出すための評価グリッド法」. 日本音響学会誌 73 (12): 783-89.
- 田尻 慎太郎, と白鳥成彦. 2013. 「ビジネス系大学における学修履歴活動データを用いた生存時間分析」. 日本教育情報学会第32回年会.
- 東京大学. 2016. 「経済的理由による学生等の中途退学の状況に関する実態把握・分析等及び学生等に対する経済的支援の在り方に関する調査研究報告書:平成 27 年度文部科学省大学改革推進委託事業」.
- 藤原朝洋, 富永ちはると押味京子. 2013. 「大学における休退学の現状・対策・課題の検討:37 大学の現状と取組」. 九州共立大学研究紀要 4 (1): 11-18.
- 藤川麗. 2018. 「学生相談におけるコラボレーション—実践と研究の発展に向けての課題と展望—」.

- 教育心理学年報 57: 192-208.
- 独立行政法人 日本学生支援機構. 2007. 「大学における学生相談体制の充実方策について」. 「総合的な学生支援」と「専門的な学生相談」の「連携・協働」, 2007年1月1日.
- 読売新聞教育ネットワーク事務局. 2018. *大学の實力 2019*. 中央公論新社.
- 日本 IR 協会, 井芹俊太郎, 近藤伸彦と松田岳士. 2022. *大学 IR 標準ガイドブック: インスティテューショナル・リサーチのノウハウと実践*. 編集者: 塚本浩太. インプレス r&d.
- 入江智也, と丸岡里香. 2017. 「大学入学時における UPI の key 項目への該当および居住形態が退学リスクに及ぼす影響: 生存時間分析を用いた検討」. *学生相談研究* 38 (1): 1-11.
- 白鳥成彦. 2016. 「ベイジアンネットワークによる大学生の中退行動のモデル化」. *人工知能学会全国大会論文集 JSAI2016: 3F35-3F35*.
- . 2020. 「日本における中退防止施策の類型化」. *大学情報・機関調査研究集会 論文集* 9: 100-105.
- . 2021. 「中退防止に関連する変数の整理」. *大学情報・機関調査研究集会 論文集* 10: 100-105.
- 武内清. 2008. 「学生文化の実態と大学教育」. *高等教育研究* 11: 7-23.
- 文部科学省. 2022. 「学生の修学状況（中退者・休学者）に関する調査（令和3年度末時点）」, 2022年.
- 文部科学省中央教育審議会大学分科会. 2020. 「教学マネジメント指針」.
- 立石慎治, と小方直幸. 2016. 「大学生の退学と留年 その発生メカニズムと抑制可能性」. *高等教育研究* 19: 123-43.
- 労働政策研究・研修機構. 2015. 「大学等中退者の就労と意識に関する研究」. 調査シリーズ No.138.
- 濱名篤. 2006. 「初年次教育からみた教養教育・キャリア教育 (シンポジウム2 初年次教育・導入教育のアイデンティティ--キャリア教育と学士課程教育との関係を考える)」. *大学教育学会誌* 28 (1): 46-52.