

教学 IR 導入ガイド

(初版)



2020 年 3 月

北海道大学 高等教育推進機構 高等教育研究部

はじめに

2009 年、北海道大学は大学間連携事業として教学 IR を本格的に開始しました。具体的には、一年生と三年生を対象にしたアンケート調査により学習状況やコンピテンシーの修得状況の把握を試みたわけです。この調査自体にももちろん意義があるのですが、より重要なのは連携 4 大学全体の平均値を計算し、公開したことです。1 大学単独の集計では、得られた数値にどれほどの意味があるのかわかりませんが、多くの大学の平均値がわかれば、自大学の値と比較することで得られた結果を評価できるようになります。4 大学では全国平均にはなりません、2012 年に発足した大学 IR コンソーシアムに参加すれば、多数の大学の平均値を利用することができます。2019 年 10 月でコンソーシアムへの参加校は 62 校に達し、充分日本の平均値として扱うことができるようになりました。

ところで、実際に教学 IR のデータを扱って感じたのは、その専門性の高さです。統計学やデータ分析の方法については莫大な数のテキストがありますが、教学 IR 特有の処理や考え方について解説した本はありません。ましてや分析結果をいかにして改革に結びつけるかは難しい課題です。これらの点についてまとめた本はありません。日本では、教学 IR に関する活動が始まったばかりです。我々は、これから経験を積み、発展させていくことになります。まだ、学会組織もありません。IR 担当者は、所属大学で孤軍奮闘しています。

このような状況の中で、少しでも同業の方々のお役に立てればと、本書の編纂を企画しました。本ガイドでは、教学 IR の歴史から、アセスメントポリシーとの関係を解説した後、アンケートの作成方法、データの暗号化やクリーニング、前処理の方法を見ていきます。さらには、多変量解析の適用方法について解説し、最後に大学改革への寄与の方法も紹介します。

ここに集約された私たちの経験は、過去 10 年間に北海道大学で行われた IR 活動の成果です。本書が今後の日本の教学 IR の発展に役立てば幸いです。

(執筆者一同)

目 次

はじめに	II
1. IR の歴史と意義	1
2. アセスメント・ポリシーと IR の関連	6
3. データ分析 1 アンケートの作成	9
4. データ分析 2 暗号化	11
データクリーニング	14
多変量解析の前処理	22
5. データ分析 3 アンケートの分析手法—重回帰分析	24
クラスタ分析	27
因子分析	30
6. 年度報告と改革への寄与	35
7. まとめ	36

IR (Institutional Research) は、広義では大学の教育、研究、管理運営すべての分析を意味する。最近では組織の企画、政策策定、意思決定を支援するような情報を提供することと定義されるようになった (Saupe 1990)。IR は米国の大学で古くから導入されていたものであり、その歴史を遡ると以下ようになる。

1. 米国の大学関連データ収集と組織化

米国における大学のデータの収集、分析、頒布に関する法律は 1867 年に制定され、1868 年から収集が始まっている (Lenchner 2019)。これは 1960 年代まで継続した。1968 年には全米を対象にした Higher Education General Information Survey (HEGIS) が開始された。1987 年には Integrated Postsecondary Education Data System (IPEDS) が HEGIS に代わってデータの収集を行っている。このような政府の動きに合わせて、大学側も 1924 年ミネソタ大学に調査部門が設置されたことを嚆矢に 1980 年代には多くの大学に大学内のデータを収集分析する調査部門が設置された (Kinzie 2005)。調査部門の職員も徐々に増加し、1965 年には学会組織である Association of Institutional Research (AIR) が結成された。Saupe の定義が出た 1990 年代以降は大学の意志決定に必須の部門となった。

このような発展の背景には、米国社会からの要請がある (Kinzie 2005)。1980 年以前は財政面から適切な使用と分配が求められていたが、1980 年代には成果に対する accountability (説明責任) が要請されるようになった。すなわち、高度な熟練した労働力を、いかに、より多く輩出しているかを証明する必要がでてきた。その後、アセスメントにも重点がおかれることになった。アセスメントとは、情報を定義し選びだし、そして設計し、収集し、分析し、解釈し、それを使用することによって学生の学習と発達を助けることである。ここでの情報とは資金とか学生、教職員の質である。さらに次の段階では、学生がどれだけ学んだかということに対するアウトカムを評価されるようになった。

2. 最近の状況

21 世紀に入って米国の大学に求められているのは、多面的な assessment プログラムの開発である。そのための手法は、「学生調査」と「同窓生調査」がメインである。公的な機関 (大学) は学習における明示的な目標を確立し、そしてその達成度合いを体系的に evidence として収集することが要求される。しかし、過去の研究から「直接的な尺度を開発するには実質的に相当なコストがかかる」ことが明らかになった。すなわち、学生を対象にした直接的な調査は難しいということである。特に、コンピテンシーの直接測定は簡単にはできない。そこで、入学してくる学生たちの特徴を掴み、学生の満足度や大学での経験を評価する。

さらには、卒業次の成果や同窓生の満足度を分析するわけである。

3. 自律的な調査

このような調査を各大学が独立して行うことは効率的でないし、小さな大学には不可能である。そこで、大学側が中心となって実施する調査組織がいくつか設立された。そこでは、共通アンケートが用いられ、全国の平均値を公開することで、自大学との比較ができるようになった。日本のマスコミが記事とする自習時間などの米国の大学生の学習状況は、この仕組みで得られたものである。

UCLA の Higher Education Research Institute (HERI) が 1966 年から始めたフレッシュマン調査 Cooperative Institutional Research Program (CIRP) は、毎年全米の 1900 の大学から 1500 万人の学生が参加しており、年会費は $(\$825 + \$3.26 \times \text{調査数})$ である。1999 年にインディアナ大学が始めた NSSE (the National Survey of Student Engagement) には、毎年全米の 600 の大学から 32 万人が参加しており、年会費 $\$2000 \sim \8000 (学生数 500 ~ 1200 以上の場合) である。米国にはこれ以外にもいろいろな大学グループの調査組織が存在する。

4. 日本の歴史

日本では 2004, 2005, 2007, 2009 年に同志社大学の山田らが College Senior Survey (CSS) や CIRP の日本版である Japan College Student Survey (JCSS) をいくつかの大学で試験的に実施したものが嚆矢である。山田らは 2008 年には Japanese Freshman Survey (JFS) と Japanese Junior College Student Survey (JJCSS) を実施している。これら 3 つをまとめたものが Japanese Cooperative Institutional Research (JCIRP) である。

2009 年には 4 大学 (北大, 同志社大, 大阪府立大, 甲南大) による連携事業が文科省の支援を受け開始され、毎年 4 大学で JCIRP が実施されるようになった。このプロジェクトは 2012 年からは 8 大学 (4 大学 + 琉球大, お茶の水女子大, 関西学院大, 玉川大) による連携事業となり 5 年間継続されるとともに、一般社団法人大学 IR コンソーシアム (以下コンソと呼ぶ) が組織され希望する大学は本調査に参加できる仕組みができあがった。コンソの加盟大学は原稿執筆時点 (2019 年 10 月) には日本全国の国公立大学 62 大学が加盟し、およそ 10 万人が参加する組織となっている。年会費は一律 30 万円である。

一方、少し遅れて文科省では 2014 年 12 月 22 日の中央教育審議会答申「新しい時代にふさわしい高大接続の実現に向けた高等学校教育, 大学教育, 大学入学者選抜の一体的改革について」の中で、

IR の歴史と意義

(3) 大学教育の質的転換の断行

(略) 大学全体としての共通の評価方針(アセスメント・ポリシー)を確立した上で、学生の学修履歴の記録や自己評価のためのシステムの開発、アセスメント・テストや学修行動調査等の具体的な学修成果の把握・評価方法の開発・実践、これらに基づく厳格な成績評価や卒業認定等を進めることが重要である。

と述べている。2017 年 3 月 31 日 大学改革支援・学位授与機構質保証システムの現状と将来像に関する研究会「教育の内部質保証に関するガイドライン」では、

「教育の内部質保証」とは、大学の教育研究活動の質や学生の学修成果の水準を自ら継続的に保証することを指す。

と定義した。そして、2018 年から実施される認証評価に関する省令の改正(学校教育法第百十条第二項に規定する基準を適用するに際して必要な細目を定める省令[平成十六年文部科学省省令第七号])には、

大学評価基準に定める項目のうち、内部質保証に関することについて、重点的に認証評価を行うものとなることが定められており、次の大学機関別認証評価の重要な評価基準となる。

と明記された。さらに、2018 年度からの私学助成配分ルール変更において、以下のような要件が加わった。

大学等内に IR 機能を整備したうえで、さらなる機能強化を図るための取組を行っていますか。

- 1 IR の企画や実施方法等に関する専門的な高等教育プログラムを履修した者を担当教職員に配置している。：2 点
- 2 1 には該当しないが、IR 担当教職員に IR の企画や実施方法等に関する研修を定期的に受講させている。：1 点
- 3 上記のいずれにも該当しない。：0 点

アセスメント・ポリシーの設定を前提とした、内部質保証がすべての大学に要求されるようになったのである。ここにおいて、日本でも IR の組織化が強く要請されることになった。

5. IR の組織化とは

以上のように IR についての歴史が浅い日本では、ここ数年で大学の業務に新たな機能

が要求された感がある。しかし、米国の歴史を考えれば、本来必須の機能であり、今後とも IR 機能は強化する必要がある。米国の研究大学での IR 部門は、修士以上学位のある職員を 5～10 名擁し学内に積極的なサービスを行う組織である（宮本他 2019）。すなわち、IR を担う職員は分析能力のみならず高等教育のあらゆる側面に詳しく、長い経験が必要な「専門職」であることが期待されている。教職員の兼職などではなく、正職で期限付きではない専門の教職員が望ましい。また、そのような教職員の養成制度や学会の設立まで達成しないと、日本の IR 活動が米国に追いつくことはないであろう。

6. データ収集の合理化

教学 IR の対象となり得るデータは、主に学生個人に付随する。入学前の成績や学習環境・学習態度、入試の成績、入学後の学習環境・学習態度、在学中の成績、就職先、就職後の学習環境・学習態度、入学時・在学時・卒業時・卒業後のアンケート、などである。これらのデータは学生個人にひも付けされていることが望ましいが、日本の現状では、そうっていない。今後、すべてのデータが集約される方向に向かうことが期待される。同時に、これらのデータは大学で一括して管理、更新するためのコンピュータネットワークの整備が必要である。米国では、データウェアハウスを中心とした整備がされてきている。セキュリティ管理とともにデータ管理の統合化が必要である。

参考文献

- CIRP ホームページ(<https://heri.ucla.edu/cirp-freshman-survey/> (2019.10.10 アクセス))
- 大学 IR コンソーシアムホームページ (<http://www.irnw.jp/index.html> (2019.10.10. アクセス))
- Kinzie J. L. (2005) 「米国の学生調査と IR の役割」 同志社大学学習支援・教育開発センター第 2 回講演会(<http://kir013749.kir.jp/content/pdf/0307.pdf> (2019.10.10 アクセス))
- Lenchner, E. (2019) 「Public Use of IR data in the United States, and Applications to the Expansion in the Japanese Higher Educations System」 高等教育ジャーナル —高等教育と生涯学習— 26, 47-58.
- 宮本淳, 山田邦雅, 細川敏幸(2019) 「米国高等教育機関における教学 IR の訪問調査」 高等教育ジャーナル—高等教育と生涯学習— 26, 77-82.
- NSSE ホームページ (<http://nsse.indiana.edu/> (2019.10.10. アクセス))
- Saupe, J. L. (1990) The Functions of Institutional Research 2nd ed. (Florida State)

IR の歴史と意義

University(Tallahassee), FL :Association for Institutional Research.

山田礼子（2006）「転換期の高等教育における学生の教育評価の開発に関する国際研究」[平成 16—18 年度科学研究費補助金研究基盤研究(B)研究成果報告書]

山田礼子（2012）「学士課程教育の質保証へむけて—学生調査と初年次教育からみえてきたもの」, 東信堂（東京）



2

アセスメント・ポリシーと IR の関連

それでは、教育の質保証には、どのような調査が有効であろうか。次の表は、それを整理したものである。手法の中でも青字にしたものは、従来の大学では行われなかった調査で、これからのアセスメントで必要とされるものである。もちろん、最終目的は調査そのものではなく、主にディプロマ・ポリシー（DP）を達成しているか否かの調査の結果を受けて改善をはかり、学生の学習環境が常に向上するように努力する仕組みを構築することにある（細川ら 2018）。

表 1. アセスメントで利用可能な調査

	手 法	目 的	改 善
教育成果	成績評価, 学習状況調査, 卒業生調査, 企業調査	ディプロマポリシーと比較して充分か	カリキュラム改革
教育方法	授業評価アンケート	適切か	授業改善
学習環境	授業評価アンケート, 学習状況調査	良好か	環境改善

この調査対象は、在学生のみならず、卒業生や卒業生を採用した組織にも及ぶ。また、これらの調査結果は、調査項目が一致すれば、学内の学部間、学内の学年間、学外の大学、外国の大学などで比較することが可能となる。いわゆるベンチマークが利用できるのである。

IR の観点から教育の質保証を行うためには、まずアセスメントの実施体制をアセスメントポリシーに記述して責任者と実行組織を明確化する必要がある。ポリシーにはチェックリストを添付し、実施の方法と時期を明示する。例えば、表 2 のようになる。

表 2. チェックリストの例

A学部 B学科 アセスメント・チェックリスト（例）

A学部B学科の学位授与水準（DP）

DP1... について実践できる。									
DP2... について倫理的な判断ができる。									
DP3... に関する資質と能力を修得している。									
番号	名称	実施時期	周期	対象	内容、質問項目等	手法	評価者	結果の活用方法	実施者
1	授業アンケート	7月、1月	毎年	2～4(6)年生	・授業の満足度 ・授業の理解度 ・学習状況	質問紙	学生	結果を学内に公表し、教員に教育方法改善の指針を与える。	A学部
2	成績評価	4月、9月	毎年	2～4(6)年生	・成績の点検	履修指導時に点検する	B学科学生指導関係教員	成績をもとに履修指導を行う。	A学部B学科
3	達成度点検	4月、9月	毎年	2～4(6)年生	・成績等によるDPの達成度確認	履修指導時に達成度表により達成度を点検する	B学科学生指導関係教員	達成度を表すリーダーチャートをもとに履修指導を行う。	A学部B学科
4	1、3年生学習調査※	12月	毎年	1、3年生	・学習状況 ・コンピテンシーの習熟度	質問紙	学生	教職員に開示して今後の教育制度改革参考とする。	総合IR室
5	卒業生調査	2～3月	毎年	4(6)年生	・カリキュラム、卒業研究の満足度等	質問紙	学生	専門科目のカリキュラム改善を行う。	A学部
6	卒業生調査 (OB/OG調査)※	9月	毎年	卒業5,10,15年目の卒業生	・大学教育の有用性と現在の状況	質問紙	卒業生	教職員に開示してカリキュラム改革の参考とする。	総合IR室
7	企業調査 (就職先調査)	随時	〇年に1回	卒業生の就職先	・企業等からのアドバイス	質問紙	企業等の職員	評価に基づき、カリキュラム改革を検討する。	総合IR室
8	外部アドバイザー会議	随時	〇年に1回	教育全体	・社会、企業からのアドバイス	討論	外部アドバイザー	アドバイザーの提言に基づき、カリキュラム改革を検討する。	A学部
9	DPによるカリキュラム点検	随時	〇年に1回	カリキュラム	・DPを実現できるカリキュラムになっているかの点検	ナンバリング、カリキュラムマップなどによるDPとカリキュラムの相互参照	B学科学生指導関係教員	DPとカリキュラムのバランスが悪い場合はいずれかを調節する。	A学部B学科
10～〇	・	・	・	・	・	・	・	・	・
	総合評価	4月	毎年	前年度の教学アセスメント全体	・前年度に実施した教学アセスメントの結果をまとめる	報告書	A学部〇〇委員会委員	年間の教学アセスメントとその成果を総括する。	A学部

※大学IRコンソーシアムの共通アンケートを利用

アセスメント・ポリシーと IR の関連

ここでは、アセスメントに利用できる資料が列記され、調査や会議は時期が明示されている。これは教員側のチェックであるが、学生側のチェックも可能である。例えば表3のように DP 達成への寄与率を各科目で設定し、その修了度合いで達成度を測るものである。いったん e-ポートフォリオなどに設定して、コンピュータ上で行えば、それほど実行は難しい。

表 3. 学生個人の DP 達成度点検表の例
達成度点検表 3年次

科目名	DP1	DP2	DP3	成績	単位数
一般教育演習	2	2	3	A+	2
力学I	3	0	1	A	2
物理学実験	3	1	2	B	2
電磁気学1	3	0	2	A-	2
・	・	・	・	・	・
・	・	・	・	・	・
・	・	・	・	・	・
量子力学	3	0	2		2
合計	92	80	85	GPA	95
目標	100	100	100	3.4	120

教育組織の評価には、まず表4のようなチェックリストを用意する。ここでは、横軸にある学科の DP が、縦軸に学生調査のアンケート項目が列挙されている。○印が、各 DP を評価するために使えるような設問項目を示している。該当する設問を調査して、他大学や他学部との比較で自分の学部の相対的な位置がわかるので、それをベンチマークとして評価する。学部独自、あるいは大学独自のアンケート調査では比較の対象がなく、経年変化を追うことぐらいしかできない。また、比較対象のない数値は評価のしようがない。しかし、コンソの調査に参加すれば、容易にベンチマークが得られる。設問項目は CIRP に準拠しているので、国際比較も可能である。国際比較の例については Lee(2017)を参照。CIRP の設問項目は毎年少しずつ改編されており、現在の CIRP の設問項目については細川ら (2019) を参照してほしい。

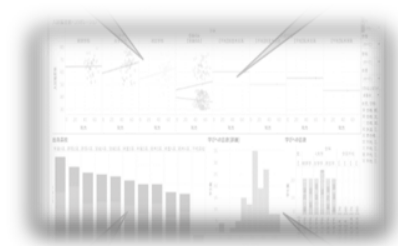


表 4. 組織評価のためのチェックリスト

			文学部		
			【知能・態度】		
			文学部DP1	文学部DP2	文学部DP3
			専門分野に関して十分な学識を身につけるとともに、「人間とは何か」の観点から、人間の歴史、社会、文化及び思想に関して幅広い関心を持ち、知識を修得することができる。	修得した知識を手段として、自己および自己を取り巻く環境が直面する問題を的確に理解し、対処することができる。	日本語ないし外国語を用いて、文献・資料を正確かつ批判的に読み解き、意見を述べることができる。
3年生調査	問8	大学の授業や授業以外の学習に関して、あなたがどのくらい理解しましたか	A 実験、実習、フィールドワークなどを実施し、学生が体系的に学ぶ B 仕事に役立つ知識やスキルを学ぶ C 授業内容と社会や日常生活のかかわりについて、教員が説明する D 授業の一環でボランティア活動をする E 学生自身が文献や資料を読む F 定期的に小テストやレポートが課される G 教員が提出物に添削コメントをつけて返却する H 学生が自分の考えや研究を発表する I 授業中に学生同士が議論をする J 授業で検討するテーマを学生が設定する K 授業の進め方に学生の意見が取り入れられる L 取りたい授業を履修登録できなかった M 出席することが重視される N TAなどの授業補助者から補助を受ける		
	問10	大学の授業や授業以外の学習に関して、あなたがどのくらい理解しましたか	A 授業課題のために図書館の資料を利用した B 授業課題のためにWeb上の情報を利用した C インターネットを使って授業課題を受けたり、提出したりした D 提出期限までに授業課題を完成できなかった E 授業時間外に、他の学生と一緒に勉強したり、授業内容について話したりした F 授業中、教員の考え方や意見に異議を唱えた G 授業を欠席した H 授業に遅刻した I 授業を早退した J 授業中に居眠りをした K 教職員に学習に関する相談をしたり、学内の学習支援室を利用したりした L 単位とは関係のない教員あるいは学生による自主的な勉強会に参加した M 大学の教職員に将来のキャリアの相談をした（卒業後の進路や職業選択など） N 教員に親近感を抱いた		
	問12	入学した時点と比べて、あなたの能力や知識はどのように変化しましたか	A 一般的な教養 B 分析力や問題解決能力 C 専門分野や学科の知識 D 批判的に考える能力 E 異文化の人々に関する知識 F リーダーシップの能力 G 人間関係を構築する能力 H 他の人と協力して物事を遂行する能力 I 異文化の人々と協力する能力 J 地域社会が直面する問題を理解する能力 K 国民が直面する問題を理解する能力 L 文章表現の能力 M 外国語の運用能力 N コミュニケーションの能力		

参考文献

- 細川敏幸, 山田邦雅, 宮本淳 (2018)「アセスメント・ポリシーの考え方: アセスメント・ポリシー研究会報告」高等教育ジャーナル—高等教育と生涯学習— 25, 49-54.
- 細川敏幸, 山田邦雅, 宮本淳 (2019)「学習状況調査アンケートの動向—米国と韓国の現在—」高等教育ジャーナル—高等教育と生涯学習— 26, 59-66.
- Heewon Lee, Atsushi Miyamoto and Jeung Yun Choi (2017) “A Comparative Analysis of the Teaching and Learning Processes of Undergraduate Students in Korea and Japan—Based on the Cases of Seoul National University and Hokkaido University”, Journal of Higher Education, 24, 77-96.

アンケートの作成

1. 調査項目の検討

何かの現状を把握する調査は、調査の目的が明確でなければならない。現在、米国や日本で用いられている高等教育に関するアンケート調査の項目は、学生の学習状況と大学の教育力の測定がその目的である。特に、社会が求めるコンピテンシーの測定は他に手段がないため重要である。

他の多くの調査同様、アンケート項目の検討は常に求められる。それは、社会情勢が大きく変化しているからである。大学を含め社会情勢に影響を受ける組織の調査は、毎年些細ではあるが重要な変更を行わなければならない。

2. 視野を広げる

それでは、視野をどう広げていけば良いのか。社会の動きはニュースが、教育制度改革は学会の情報が重要である。米国では大学教育の組織である Association of American Colleges and Universities (AAC&U) や IR 研究の学会である Association for Institutional Research (AIR) 日本では大学教育学会、高等教育学会、コンソなどが有名である。昨今の日本では、文科省の中教審の教育改革議論や答申も知っておくべきである。多くの場合、数年で実現されるからである。

国内の標準となりつつあるコンソの質問項目も毎年少しずつ変更されている。米国の標準的な IR アンケートである CIRP や NESS も毎年変更を加えている。最近のアンケート項目については細川らの報告 (2019) に詳しい。

3. 項目数の削減

新たにアンケートを考える際には、少し多めの設問を考えて、試行の段階でふり落とすほうが、最初から少数の項目で勝負するより賢明である。しかし、多い項目からどれを落とすかは慎重に考えなければならない。以下の二つの手法が統計学の分野から利用できる。

4. アルファ係数

Cronbach's alpha (クロンバックのアルファ) とは、性格検査の質問項目のように、ある特性に対して複数の質問項目を設け、回答の合計値 (尺度得点) を特性尺度として用いるときに、各質問項目 (変数) が全体として同じ概念や対象を測定したかどうか (内的整合性) を評価する信頼係数。0 から 1 までの値をとり、1 に近いほど信頼性が高いと言える

(BellCurve 統計 Web)。設問の整合性が高ければ、重複した設問の可能性が高い。境界値は 0.8 で、0.8 以上であれば整合性が高いことになる。

5. 因子分析

もう一つの手法は、因子分析である。R コマンドのメニューからは、統計量>次元解析>因子分析 で実行できる。この結果から、典型的な質問項目は近くに分類されるので、そこから重複した設問を選択できる。また因子分析では、各変数間の相関係数が算出される。相関係数の異常に高いものは類似の質問の可能性がある。

参考文献

BellCurve 統計 Web (<https://bellcurve.jp/statistics/glossary/1274.html> (2020.1.30 アクセス))

細川敏幸, 山田邦雅, 宮本淳 (2019)「学習状況調査アンケートの動向—米国と韓国の現在—」高等教育ジャーナル—高等教育と生涯学習— 26, 59-66.

CIRP ホームページ <http://www.heri.ucla.edu/> (2019.10.12 アクセス)

CIRP アンケート http://nsse.indiana.edu/pdf/survey_instruments/2018/NSSE_2018_US.pdf (2019.10.12 アクセス)

大学 IR コンソーシアムホームページ <http://www.irnw.jp/> (2019.10.12 アクセス)

NSSE ホームページ <http://nsse.indiana.edu/> (2019.10.12 アクセス)

NSSE アンケート http://nsse.indiana.edu/html/survey_instruments.cfm (2019.10.12 アクセス)



暗号化

1. 暗号化とは

暗号化とは、データを意味的に関連のない文字列へ一意的に変換することであり、また、その変換されたデータは特定の操作により元に戻すこと（復号）が可能である。

一般に、暗号化結果を利用者だけの固有のものにするために、独自に設定した文字列を暗号鍵として暗号化を行い、また、その暗号鍵でのみ復合できるようになっている。

2. 暗号化を利用するメリット

暗号化はデータの内容を隠すために利用される。特に教学 IR データにおいて、学籍番号は他のデータと組み合わせることにより特定の個人を識別できるデータであるため、個人情報にあたる。個人の特定ができない形でデータを取り扱うためには暗号化が有効である。

特に複数の大学が参加するコンソーシアム型の教学 IR では、各大学が学生データを提供し合う必要があり、さらに、学生個人の追跡を行うタイプの解析が行われる場合には、学生の識別子として学籍番号を暗号化して使用する場合がある。

単に、情報を隠すことが目的であれば、ハッシュ化と呼ばれる方法のような元に戻せない変換を利用することができるが、暗号化のメリットは、下記に示すように復号に依るところも大きい。

以後、暗号化について、1つの文章を暗号化するような場合ではなく、学籍番号のような複数のデータをそれぞれ暗号化する場合を想定して話を進める。

2.1 データ処理の分業化を可能に

データの管理体制として、データのアクセス権を持つ者が必ずしも全ての集計処理を行うとは限らない。その場合、データを暗号化することにより生データへのアクセスの権限のない者にデータ解析を依頼でき、例えば、依頼者は解析によって抽出されたデータを復号することができる。

2.2 データ照会に入り込むリスクを減らす

変換後のデータを直接復号する方法がなく、対応表にアクセスして照会しなければならない場合、その都度、全元データである対応表へアクセスすることになる。各データを直接復号する変換方法があれば、必要なデータのみを扱うだけで済むため、情報漏洩のリスクを抑えることができる。

2.3 誤送信対策

情報漏洩のリスクは、業務として想定されているプロセスにおいて起こるとは限らず、誤ってメールに添付したり、USB メモリを紛失してしまったりするような人的ミスによるリスクが大きい。そのため、普段から暗号化済データを扱う方法も考えられ、その場合は復号手段が必須である。

表 1. AES による暗号化の例

元データ	暗号化されたデータ
001	f nk/Ngxknra6mIhzoDW7uw==
002	/5eRQa1vfWVoUzZosaN8Jg==
003	3U0kyP/uKJN6NA0ZeYYiJw==
ネコ	mroguWI p1NDZt3goDVBS pA==
(半角スペース)	VgFCeIdB2ai1RIL04Lt7oA==

3. 暗号化に関する諸注意

3.1 組み合わせての情報漏洩を意識する

暗号化は、現時点で信頼性が高いといわれているアルゴリズムのものを利用する場合でも、完璧な暗号化は存在しないという認識で扱わなければならない。将来的に弱体性が見つかることや、コンピュータの処理能力が飛躍的に高まってパスワード解析にかかる時間が大幅に短縮されるといったことが起こっても不思議ではない。特に、大学 IR では学籍番号を暗号化する場合があるが、学籍番号と個人情報結びけられた情報を日ごろから公開しないようにすることが重要である。氏名と学籍番号を併記したものを掲示するなどは控えなければならない。

3.2 テキストエディタで表示されないタイプの文字のデータクリーニング

暗号化が適切に完了したかどうかの判断は、一般的に見た目では分からない（表 1）。よって、元データが適切ではないことを、暗号化後に気づくことは難しい。

例えば末尾にスペースが付いていることに気が付かずに暗号化を行った場合、スペースが付いていない場合の暗号化結果とは異なる結果となってしまう。特に、表示されないタイプの文字（スペース、タブ文字など）を取り除くデータクリーニングを心がける必要がある。表 1 の AES による暗号化の例では、半角スペースのみでも他と同様の暗号化結果となっており、暗号化後に見た目で気づくことはできない。

3.3 暗号化するアプリケーションについて

不正ログインは、単に推測されやすいパスワードが使用されていることが原因であることが多い。しかし、一般的なログインの場合を考えると、そもそもユーザ ID がわからなければパスワードを試すことはできず、ユーザ ID を漏らさないことがリスクを抑えることになる。

暗号化する場合には、例えば AES で暗号化する場合でも、使用するアプリケーションによって暗号化結果は異なる。これは、アプリケーションレベルで暗号鍵の扱いが異なるからである。暗号化に使用するアプリケーションは暗号鍵を推測するツールにもなるため、ログインのユーザ ID を漏らさないのと同様に、暗号化に使用するアプリケーションとして独自のものを利用し、アプリケーション自体を漏えいさせないようにすることも重要である。

4. 暗号アルゴリズムの選定

暗号技術評価プロジェクト CRYPTREC にて推奨暗号になっているものを利用するとよい。通常、暗号化に推奨されるアルゴリズムはその時点で様々なところで活用されており、名前が知られていることが多い。逆に、暗号化アルゴリズムは、広く使われているにもかかわらず弱体性が見つかっていないという事実が重要である。2020 年現在では、AES が推奨されるアルゴリズムの 1 つである。

参考文献

CRYPTREC ホームページ, <https://www.cryptrec.go.jp/> (2019.10.12 アクセス)



データクリーニング

はじめに

データクリーニング（データクレンジングとも言う。以下本稿では「クリーニング」と記載する）とは、エラーデータを削除・修正し、一定の書式にデータを整形することを言う。クリーニングは、分析に向けてデータがデジタル化されてから始めることになる。アンケート調査等により得られた各種データのクリーニングについては、確立された一定の方法があるわけではない。一般的にアンケート調査は、設問の設計、調査の実施、結果のデジタル化（表計算ソフトへ回答値が入力された状態）、クリーニング、集計・分析の過程を経る。調査設計時より、クリーニング作業が少なくなるよう、つまりエラー回答や無効回答が少なくなるように考慮することが重要である。ここでは筆者の経験に基づき、クリーニングに関して注意する点などについて述べる。

1. アンケート調査設計時の注意点

1.1 単一選択

「1. ○○, 2. △△, 3. □□, ...」と選択肢が並び、ひとつのみを選択する回答形式である。Web アンケートシステムであれば、ラジオボタンを押すことで回答するので、ひとつのみしか選択することができない。マークシートや手書きの回答用紙であれば、

- ①「2 つ以上選択されたもの」
- ②「(例えば) 選択肢 1 と 2 の間を選んでいるもの」
- ③「回答無し」

がクリーニングの対象となる。

①, ②については、機械的に無効回答として扱うのが基本ではある。しかし、なるべく回答を活かすという方針で、例えば満足度などの段階評価の選択肢が 2 つ選択されていた場合、より高い満足度を回答として採用するというクリーニングも考えられる。いずれにしても、同一の調査内では統一したクリーニング方針を用い、その方針を記録として残すことが重要である。

③については、無回答をなるべく減らす工夫が必要である。すべての回答者が漏れなく回答できるような選択肢を設けなければならないが、多くの場合は、選択肢に「その他」を設けておくことで対応できる。「その他」と自由記述の組み合わせにより、選択肢を再考するうえでの資料となり得る。ただし、自由記述の内容が、既設の選択肢に含まれる内容であ

るのかどうかの検討はクリーニング作業として行う必要がある。特に回答選択肢の数を絞っている場合などでは、「該当なし」の選択肢を設けておくことで、該当する選択肢がない場合と無回答を区別することができる。

1.2 複数選択可

単一選択に比べると、複数選択は回答対象者数と回答数が一致せず、単純集計やクロス集計作業が煩雑であり、後の分析、解釈までを考えるとできれば単一選択としたい。以下、具体的に複数選択可の短所を挙げる。

- ①データ入力作業や集計に手間がかかる。単一選択であれば、問に対してひとつの回答があり、それを入力、集計すれば良いが、複数選択可の場合、回答者によって選択数が異なり、例えば、カンマ区切りで複数選択番号をひとつのセルに入力し集計する場合や、選択肢の数と同数の入力セルを準備し、選択されていれば1を、選択されていなければ0を入力する場合などがある。いずれも入力や、集計作業が単一選択とは異なり、注意が必要であり、手間も多くなる。
- ②複数選択では回答対象者数と回答数（選択肢の全チェック数）は一致しない。集計時には、回答対象者数を集計母数とし、回答者の何割が当該選択肢を選んでいたのかを見る分析が一般的である。回答数を集計母数としないよう注意が必要である。
- ③一般的には、複数選択の場合、それぞれの回答に重み付けができない。例えば、進学理由を複数選択でたずね、2つ以上にチェックがあった場合、どの選択肢を最も重視したのかを測ることはできない。重み付けを行いたい場合、最も重視した理由と次に重視した理由のように2つ以上の単一選択に分ければよい。また、選択肢全てを選択するような回答も見られる。これは、稀なので、全体集計に大きな影響はないと考えられるが、選択数に制限を設ける（上記例では、主な進学理由を2つまで選択する）工夫が有用な場合もある。

特に③については、単一選択の方がより詳細な情報を得られることがわかるが、回答回収率向上のために極力質問数を減らしたい（質問数が少ないように見せたい）場合や質問内容から複数選択可を採用せざるを得ない場合など、「複数選択可」が有用な場合がある。設計時に慎重にどのような選択肢にするのかを考慮しなければならないが、複数選択可採用時にも、単一選択と同様に「その他」と「自由記述」を設け、すべての回答者が漏れなく回答できるように工夫をしておく必要がある。

1.3 条件分岐

「問1で, [1] と答えた方は問2に進み, [2] と回答した方は問5に進んでください」というような条件分岐を行う設問がある。たずねる内容によっては, 便利な方法ではあるが, 経験上以下のような理由で極力避けたい回答形式である。ただし, 高機能の Web アンケートシステムで, 条件に沿って答えるべき設問へ自動的に導いてくれるような場合は, 問題はない。一方で, 集計や結果の解釈は複雑になる。

- ①回答者が指示通りに回答しておらず, 誤回答が多くなり (上記例では, 問1で [2] と回答したにもかかわらず, 問2, 3, 4にも回答している), 後でクリーニングもできず, 指定外回答や無効回答とせざるを得ない。アンケート調査実施者の立場からすれば, きちんと設問を読んで指示通り回答すれば, 矛盾もなく誤るはずがないような条件分岐であっても, 誤回答が多くなる。一方, 回答者の立場からすれば, 設問はあまり読まず, 条件分岐であることを認識せず主に選択肢のみを見て回答するので, 結果的に条件分岐を無視することになる場合が多くなるのは, 必至であると考えられる。
- ②慎重に設計しても, 回答者にとって条件に合わない場合が出てくる。また, 回答者によって分岐条件の解釈が異なる場合もある。これは, 調査実施後に主に回答者からのコメントで気づくことが多い。どのような回答者にとっても唯一の解釈となる条件分岐を作成するのは, 例えば年齢のように回答者のすべてがある集合に漏れなく, 重複なく属することができる分岐以外, 非常に難しいと感じている。
- ③条件に沿って集計を分類する必要があるので, 集計作業に手間がかかり, 解釈も複雑になる。

①にあるように回答者は設問を注意深く読んでいるとは思えず, 調査実施者の意図や工夫は伝わらないことが多々ある。これは, 回答者を非難しているわけではなく, 多忙な回答者が調査に協力する際にどうしても起こり得ることであり, そういうものであると認識して調査を設計する必要があるという注意喚起である。設問は極力単純化することが, 回答者, 調査実施者両者にとって有益である。



1.4 選択肢

選択肢の数や文言について、注意すべき点を列挙する。

①満足度などを段階評価としてたずねる場合、例えば4段階の選択肢にするのか、5段階の選択肢にするのかを吟味する必要がある。奇数の選択肢にすると、真ん中を選択することが多くなる傾向がある。

②段階評価の選択肢番号を平均して数値として比較する場合、肯定的回答を大きくする必要があるが、設問用紙の表記と分析時に扱う数字が異なる場合もあるので、十分注意する必要がある。以下、例を示す。

【設問用紙】		【分析時】
[1] とても満足している	→	4
[2] やや満足している	→	3
[3] あまり満足していない	→	2
[4] まったく満足していない	→	1

(上記の様に、設問用紙では、選択肢の並び順から肯定的回答の数字が小さくなり、否定的回答の数字が大きくなるが、この4段階の選択肢番号を平均して満足度として見る場合、分析時には、【分析時】に示したように数字を変換し、肯定的回答の数字が大きくなるようにしなければならない。特に記述式回答とWeb回答形式を併用している場合、前者は変換が必要だが、後者は設定により変換不要となることもあるので、注意が必要である。)

③数字とともに「以上、以下、未満」などの文言を使う場合、矛盾のないように注意する。

④複数選択肢の内容は吟味し、なるべく少なくする。実際には多くの回答者が1～2程度の選択肢を選ぶことが多いが、稀にすべてを選択するような回答も見られる。また、選択肢が余りに多いと回答者にとっては回答意欲を削がれ、結果的に無回答やすべてを選択するような回答の誘因となる。



2. データのデジタル化と結合

2.1 データのデジタル化

データがデジタル化される過程は主に次の4つである。

①手書き回答を手入力

回答者が、回答用紙に回答を記入し、それを分析者がエクセルなどの表計算ソフトに手入力することによりデジタル化する。当然のことではあるが、後で回答用紙と入力情報の対応を確認できるよう、それぞれに共通番号を振っておく必要がある。

②マークシート

回答者はマークシートに回答し、デジタル化されたデータはマークシート読み取りの結果として出力される。

③手書き回答をOCRで読み取り

回答者は、フォーマットが整った回答用紙に手書きで回答を記入し、OCR（Optical Character Recognition：文字認識技術）処理によりデジタル化する。

④Webシステム

Web アンケートシステムにより、回答者が回答すると同時にデジタル化される。

①について、例えば選択肢1から5までの質問に、6を記入している場合、都度どのようにクリーニングをするのかを検討するのはなく、入力時はそれらの回答にマークをしておき、デジタル化後にどのようなエラー回答があるのかをまとめ、クリーニングを行うと効率よく作業が進む。

②、③については、基本的には読み取り時の条件設定により、エラー回答を削除、修正することができるが、その設定内容の検討は十分に行う必要がある。また、読み取り不可の回答用紙や、部分的に読み取り不可の回答用紙が存在する場合、手入力により不足分を補う必要がある。

④については、無回答は存在するが、エラー回答は基本的に存在しない。

2.2 データの結合

同じ調査において、記入式と Web システムを併用する場合や例えば学年毎に調査を行い、デジタル化も学年毎に行う場合があると思われる。分析に向けてそれらのデータを結合する必要がある。その際の注意点を挙げる。本稿においては、デジタル化された書式は、列に質問項目、行に各回答が並ぶものを想定しているが、結合するそれぞれのデータ項目、順序が同じであるかの確認は必須である。回答方法により、やむを得ず設問順を入れ替えている場合、一部の方法にだけ追加質問がある場合が想定される。コピーや貼り付けを行い結合する際には、慎重に操作しなければならない。1.4 の②で述べた数値の変換が必要な場合は、必要な変換を終えた後にデータの結合を行うと良い。

3. クリーニング

3.1 エラー回答の確認

ここからは、エクセルにデータが入力された状態を想定して述べる。エクセルのフィルター機能を使い、すべての質問の回答を確認する作業が基本である。図 1 はエクセルのフィルター機能で Q1（設問の選択肢は、1, 2, 3 とする）のデータを表示した状態である。Q1 には、1, 2, 3, 5 の数字の回答と空白セルがあることが分かる。以下、クリーニングが必要な回答を列挙する。

①Q1 は、1, 2, 3 を選択する質問なので「5」は回答者の誤記入、あるいはデータ入力者の誤入力である。

②空白セルは、回答者が無回答であったか、あるいはデータ入力者の入力漏れである。以下、同様に Q2 以降も設問内容に沿ってエラーの確認が可能である。

③設問に 1.3 で示したような分岐がある場合は、次の様に確認を行う。例えば、Q1 で「2」と回答した場合 Q5 へ飛ぶような指示がある時、Q1 のフィルターのチェックマークで選択肢「2」のみを抽出する。その状態で、Q2, Q3, Q4 に回答のあるのは、指示通りの回答していないか、回答誤入力でありクリーニングする必要がある。



	A	B	C	D	E	F	G	H	I	J
1	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
昇順(S)						3	2	2	2	3
降順(O)						2	1	2	3	1
色で並べ替え(I)						3	3	2	2	2
"Q1" からフィルターをクリア(C)						1	1	2	2	2
色フィルター(I)						2	2	2	2	2
数値フィルター(E)						2	1	2	2	2
検索						1	1	1	1	1
☒ (すべて選択)						2	1	2	2	2
☒ 1						1	1	2	2	2
☒ 2						2	2	2	2	2
☒ 3						2	2	2	2	2
☒ 5						2	3	2	3	2
☒ (空白セル)						1	1	1	1	1
OK						2	1	2	2	2
キャンセル						1	1	1	2	2
28	1	1	2	2	2	2	1	1	2	2
29	1	1	1	1	1	1	1	1	1	5

図1. エクセル上でQ1のフィルター画面を開いた状態

3.2 クリーニング

実際のクリーニング作業は、設問毎にエラー回答の確認とクリーニングを繰り返すことになる。以下、3.1の事例で挙げたそれぞれの場合におけるクリーニングの概要を述べる。

- ① フィルターのチェックマークで「5」のみを選び、フィルタリングする。手入力によりデータ化したレコードであれば、元の回答を確認し、適宜修正する。その後も残った「5」は「指定外（エラー）回答」となる。指定外回答をその他の選択肢と重ならないように例えば2桁で「88」と定義し、「5」を「88」で上書きする。
- ② ①と同様に「空白セル」をフィルタリングする。空白セルは無回答であるので、例えば「99」と定義し、空白セルに入力する。

データ分析2 クリーニング

③Q1で「2」と回答した場合Q5へ飛ぶような指示があり、Q2、Q3、Q4に回答のある場合は、手入力によりデータ化したレコードであれば、元の回答を確認、修正し、残った回答は指示通りの回答ではないので「88」を上書きし、無効回答と定義する。この場合のQ2、Q3、Q4の空白セルは、無回答ではなく、回答する必要がない設問であるので、空白のままとしておく。

上記のように指定外回答や無回答を単に「空白」とせず、クリーニングによりきちんと定義しておくことで、「88」と「99」を「無効回答」としてまとめるなどの集計が可能になる。

その他、主に自由記述回答の設問に関する注意点については、以下に箇条書きにする。

- ・調査目的によるが、個人に対する誹謗中傷などの取り扱いは慎重に行う必要がある。自由記述では、調査とは関係のないことが書かれることもあることを想定してクリーニングをする必要がある。共同で行う調査において、取りまとめを行う担当へデータを提出する際や公表する場合などには、特に詳細に確認する必要がある。
- ・出身都道府県名を自由記述でたずねた場合、「東京都」あるいは「東京」と記入・入力する場合が見られる。どちらかに統一する必要がある。
- ・選択肢番号や検定試験のスコアなどの数値を自由記述でたずねた場合、全角で入力していることがあるが、後に計算することなどを考慮して半角に統一する。
- ・使用するソフトウェアによるが、自由記述セルに改行が入力されていると、様々な不具合が生じる。極力クリーニングで削除しておくことを勧める。様々な解説があるので詳細は省くが、エクセルでは「検索する文字列」に[Ctrl] + [J]を入力することにより改行を検索することができる。

4. まとめ

本稿では、アンケート調査データのクリーニングの概要について、設問設計時の注意点から述べてきた。内容に新しい情報はないと感じる読者もいると思われるが、一方で全くクリーニングを行わずに、集計、分析を行っている事例が散見されるのも事実である。クリーニング作業は、100問の設問があれば、100問すべてについてエラーの確認と修正の作業が必要であり、アンケート調査の集計、分析の前処理としては非常に面倒で時間のかかる作業である。自由記述に書かれた内容が、用意した選択肢に当てはまるのかどうかの検討は、時に恣意的な判断になり、際限がなくなることもある。しかし、貴重な回答を1件でも有効回答として取り込むために、クリーニングは必須の作業である。

多変量解析の前処理

多変量解析はアプリケーションがほぼ自動的に実行するが、異常値の検出等は自動的に
は行われない。特に、欠損値がある場合や、数値に数字以外の意味がある場合には、前処理
が必要である。理由の正当な欠損は9とするなど、組織内部での申し合わせで欠損値を入れ
ておく場合もある。

コンピュータに明確に欠損値であることを示すには、数字以外の文字を利用するほうが便
利である。欠損値には、空白や「NA」などが使われる。RではデフォルトでNAが使わ
れる。SPSSでは、欠損値を指示することができる。アプリケーションは欠損値を発見する
と計算から自動的に除外する。ひとつのデータセットで多様な分析、多種類の組み合わせを
計算させる場合には、有効な機能である。

ここでは、Excelを利用して欠損値にNAを入れる手法ならびにRでの処理を紹介する。

1. 欠損値が空白の場合

1. =COUNTBLANK(A485:CB485) 関数で空白をカウントする。括弧内は対象の範囲を
示す。
2. 項目数は77なので、その半分40以上が空白の被験者は削除する。
空白カウントの項目でソートして、40以上の行をすべて選択して削除する。
3. 空白にNAを入れる。
 - a. Ctrl+gの後、セル選択をクリック。空白を選択する。OKをクリック。
 - b. 空白のセルが陰で示されるので、NAとタイプ。Ctrl+Enterで全部の空白がNAに
変換される。

2. 欠損値が数字などの場合

特殊な数字を入れ替える。

- a. =COUNTIF(B2:CB2, ">5") 関数で5を超えるセルを検出する。括弧内の前半は対象
の範囲を、後半は検出の関係式を示す。
- b. 発見された欠損値はexcelの検索、置換機能でNAに置換する。

3. Rにおける処理

Rの多変量解析では、欠損値をNAとしておけば、自動的に対象から除いて分析してくれ
る。データを読み込んだ後、欠損値のあるデータを削除することもできる。以下のコマンド
で、dataからNAの含まれるデータ行を除外できる。

データ分析2 多変量解析の前処理

```
>data[complete.cases(data)]  
>df <- subset(Df, complete.cases(Df))
```

4. 質問がネガティブな場合

「欠席することがあった」など、質問がネガティブな内容の場合、回答は数字が大きいほど負の意味を持つ。そこで値を反転させる必要がある。選択肢が1～4の整数で構成される場合を考える。まず、表1のような変換のための表を作成し、表内のどこかに置いておく。左の値を右側の値に変換する。この表の左上端のセルをK1, 右下端のセルをL6 とすると、H2にあるデータをI2に変換して表示したい場合、I2に次のような関数を記入する。

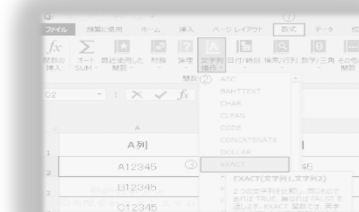
最後の2は戻り値が2列目にあることを示す。

=VLOOKUP(H2,\$K\$2:\$L\$6,2)

表1 修正のための表

元	修正後
1	4
2	3
3	2
4	1
NA	NA

H3 以下に連続してデータがある場合は、これをコピー＆ペーストすればよい。



重回帰分析

1. 教育力評価のために

教学 IR の重要な役割の一つは、機関の教育力評価である。IR の先進国米国では、1993 年 A. W. Astin が、大学教育が学生に与える効果の過程を理論化し（カレッジ・インパクト理論）I-E-O モデルを示した（図 1）。Astin はこのモデルを証明するために全米の学生を対象に、GPA を Outcome、各アンケート項目の結果を Input や Environment として、重回帰分析を行っている（Astin 1993）。日本では、細川ら（2020）が一大学のデータを用いて同様な分析を行っている。

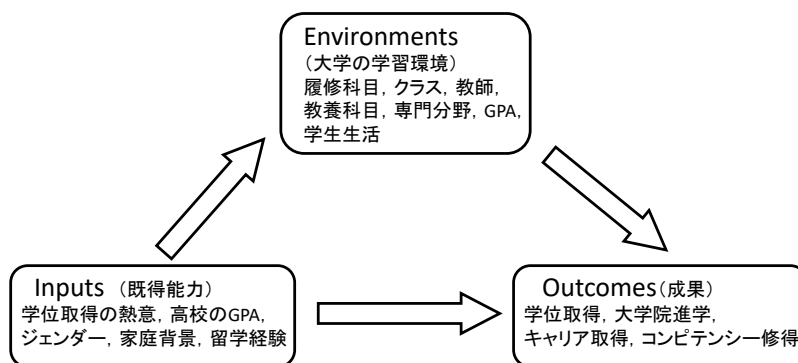


図 1. Astin の I-E-O モデル

この分析により、Outcome に影響する因子は何であるかが明らかになる。例えば、表 1 はある大学における大学 1 年時の GPA と高校での経験との重回帰分析の結果のうち、有意な偏回帰係数を示したもののリストである。

表 1. GPA と高校での経験との重回帰分析（2012 年 1 年生調査）

質問番号	質問内容	偏回帰係数	標準誤差
Q21R	入試形態	-0.0508	0.0113 ***
Q23R	高校の成績	0.0804	0.0134 ***
Q24G	予習復習宿題の経験	0.0961	0.0182 ***
Q24J	困難に挑戦	0.0538	0.0229 *
F値	8.9510 ***		
R-squared	0.1398		
Adj R-squared	0.1242		
n	898		

*:p<0.05, **:p<0.01, ***:p<0.001

データ分析 3 重回帰分析

重回帰分析の実行自体はやさしい。R(v3.5.3)と R コマンドー(v2.5-2)で計算する場合以下のような手順になる。

Excel などではデータを整理して、R に読み込む (R コマンドーのメニュー、以下同様)。
データ>データのインポート>エクセルファイルから を選択して必要な指示をする。R のコマンドとしては以下のように記述される (例)。

```
Dataset <- readXL("E:/F112/HighSchool/data2012_1HSReverse.xlsx",  
+   rownames=FALSE, header=TRUE, na="", sheet="R データ", stringsAsFactors=TRUE)
```

2. 重回帰分析の実行

統計量>モデルへの適合>線形回帰 を選択する。R のコマンドとしては以下のように記述される (例)。

```
> RegModel.1 <-  
+  
lm(GPA~Q20R+Q21R+Q22R+Q23R+Q24A+Q24B+Q24C+Q24D+Q24E+Q24F+Q24G+Q24H+Q24I  
+Q24J+Q24K+Q24L,  
+   data=Dataset)
```

偏回帰係数の表を得るためには、以下のコマンド summary を実行する (例)。

```
summary(RegModel.1)
```

Call:

```
lm(formula = GPA ~ Q20R + Q21R + Q22R + Q23R + Q24A + Q24B +  
    Q24C + Q24D + Q24E + Q24F + Q24G + Q24H + Q24I + Q24J + Q24K +  
    Q24L, data = Dataset)
```

Residuals:

Min	1Q	Median	3Q	Max
-2.10842	-0.30023	0.00893	0.31226	1.50133

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.1274246	0.1369684	15.532	< 2e-16 ***
Q20R	0.0003745	0.0329635	0.011	0.9909
Q21R	-0.0508017	0.0112781	-4.504	0.00000755031 ***

```
Q22R      -0.0195700  0.0352860  -0.555      0.5793
Q23R      0.0803735  0.0134152   5.991 0.00000000303 ***
:          :          : 以下省略。
Q24L      0.01733    0.03457    0.501 0.616734
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 0.4392 on 202 degrees of freedom

Multiple R-squared: 0.1788, Adjusted R-squared: 0.1137

F-statistic: 2.748 on 16 and 202 DF, p-value: 0.0005307

3. その他

最後に全体を対象に VIF (Variance Inflation Factor 多重共線性) を計算する。

```
> vif(RegModel.1)
```

```
Q20R      Q21R      Q22R      Q23R      Q24A      Q24B      Q24C
1.112140  1.048782  1.055758  1.200280 1.673560  2.074378  1.690168
```

この値が 10 を超える場合は多重共線性ありと判定して、その変数を分析対象から除外する。

文系理系に分けるには、データの数値を使って以下のように指示する (例)。ここでは SH という名のパラメータに理系文系の区別を数字で入れてある。

```
Bunkei <- subset(Dataset, SH==1)
```

```
Rikei <- subset(Dataset, SH==2)
```

得られたサブセットのデータを対象に重回帰分析のコマンドを実行する。

参考文献

Astin, A. W. (1993) "What matters in college: four critical years," Jossey-Bass Publishers (San Francisco)

細川敏幸, 山田邦雅, 宮本淳 (2020) 「日本における《カレッジ・インパクト理論》の有効性の検証—日本の一大学の IR データを使って—」高等教育ジャーナル—高等教育と生涯学習— 27 (印刷中)

クラスタ分析

1. クラスタ分析とは？

クラスタ分析は、似たデータを似た者同士と判定し、似た者同士が集まるクラスタに分類する手法である。「似た」という一般には質的に判断される性質を数学的に扱うことができる個性的な分析法である。

一般的なデータ解析法は、1つの集団の性質を調べたり、複数の集団を比較したりすることが多いが、クラスタ分析では、1つの集団として収集したデータから、その集団には異なるタイプの集団が内在していないかを調べる。

また、一般的な解析法では、データを分類する場合、ある基準に沿って分類するが（例えば、試験の点数が70点以上を「成績高群」）、クラスタ分析では、分類の基準を指定するのではなく、データの分布具合によって、データ自体から分類の基準が判断されるという特徴がある。

2. クラスタ分析の例

例えば、大学一年生が1週間に費やす時間についてのアンケート調査データにおいて、「部活」と「バイト」に費やす時間のデータに注目して、学生はどのようなタイプに分けられるかを調べた場合を考える（山田，2019）。4つにクラスタ化する場合、図1のように、type1：バイトのみのタイプ、type2：どちらもやらないタイプ、type3：どちらもやるタイプ、type4：部活のみのタイプが認められ、2項目のそれぞれの大小が明確なクラスタが形成されることがわかる。

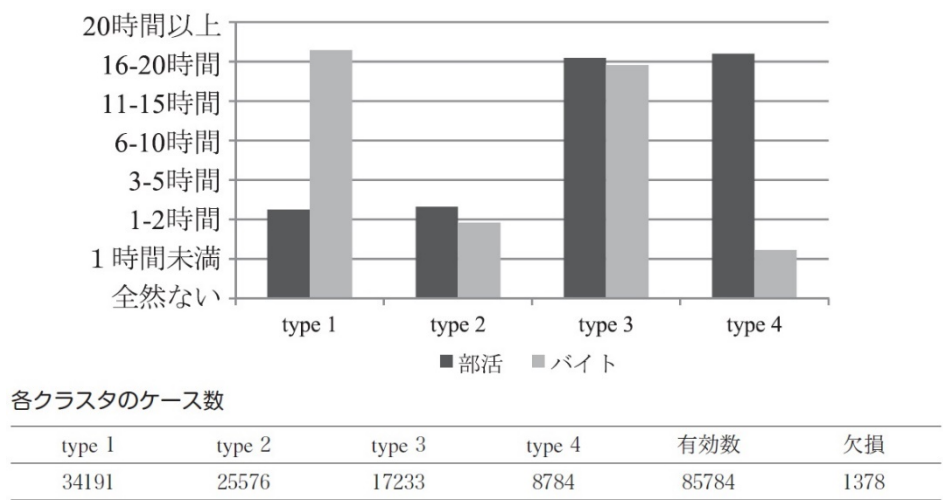


図1. 大学一年生の部活とバイトに費やす時間でクラスタ分析を行った際の各クラスタのケース数と各項目の平均値

3. クラスタ形成の仕組み

クラスタを形成するために用いられる基準にはいくつかの方法が知られているが、クラスタ化を端的に表す方法として、クラスタ分析に使用する変数とする項目を座標軸とした空間を考え、各データが空間の1点として表されるときに、距離的に近いもの同士をクラスタとしてまとめていく方法がある。このようなある種の“距離的な近さ”を数学的にどのように定義するかで、いくつかの手法が存在している。

どの手法を使うかは、実際にデータで試して決めるべきであるが、各手法に絶対的な優劣があるわけではなく、自分が行うデータ解析の展開に合わせて選択すればよい。例えば、クラスタ間の違いを調査するのであれば、互いに距離がより離れているクラスタが形成される方法が差を見出しやすいかもしれない。また、各クラスタに属するデータ数が均等である方が、その後の解析に有用である場合もある。

4. クラスタ分析のメリット

4.1 取組みの効果が表れないときの処方箋として

クラスタ分析によって形成されるクラスタは、ある定義による距離が離れていることによって分類されている。これは、各クラスタに統計的に差があると認められていることを意味するので、その違いが他の項目のふるまいに影響することが考えられる。そのため、クラスタ分析を行った場合には、その続きとして各クラスタを比較する分析を行うことが有効である。全体データとしてデータを扱っているだけではその特徴が埋もれてしまい、発見することができなかった特徴を見つけることができる可能性がある。

例えば、新たな学習サポートを提供している場合に、サポートを受けない層が存在していないか、そしてその層はどのような特性の学生かを調べることができる。

4.2 学生の多様化に対応した分析

大学 IR では、大学、学部、学科別の分析が行われることが多いが、そのような区分においても必ず多様な学生が含まれている。そのような多様な学生の存在を確認し、また、対応するためにクラスタ分析は有効である。

特に、データが大きく2極化している場合、全体データによる分析ではその2極化は平均化される。この場合、平均像というよりも、2極化の中間の実際には余り存在しない層になってしまう。実際に、大学一年生が部活とバイトに費やす時間についての分布(図2)を見てみると、バイトをする学生の分布で「時間ゼロ」は連続的に分布するわけではない(山田, 2019)。このようなデータで平均値を見ることはあまり有益ではなく、一方で、クラスタ分析を行った場合には、図1のように明確なクラスタが認められやすい。

データ分析3 クラスタ分析

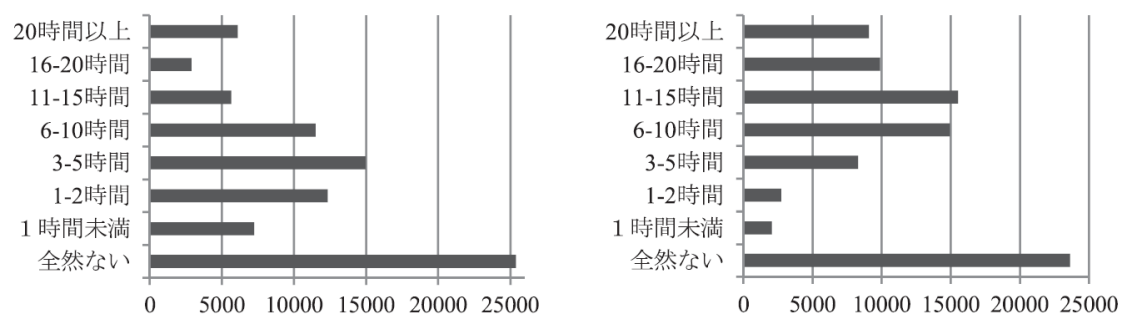


図2. 部活動や同好会に参加する時間/週（左）と大学外でアルバイトや仕事をする時間/週（右）

参考文献

山田邦雅, 宮本淳, 細川敏幸 (2019) 「IR データによる学生層と大学層の形成と評価のシミュレーション」 高等教育ジャーナル—高等教育と生涯学習—, 26, 67-76.



因子分析

はじめに

アンケート調査, 例えば学修状況調査において, 「入学後の能力変化に関する項目」や「授業経験の有無に関する項目」があり, 多くはそれぞれの観点毎に複数の質問項目を設けている。前者においては, 「専門分野や学科の知識」, 「コミュニケーション能力」, 「異文化の人々と協力する能力」, 「外国語の運用能力」などであり, 能力が増えたか減ったかの段階評価を選択肢で示し, 回答を得る。後者は, 「仕事に役立つスキルを学ぶ」, 「学生自身が文献や資料を調べる」, 「学生が自分の考えや研究を発表する」, 「授業中に学生同士が議論をする」などであり, 経験の頻度を選択肢で示し, 回答を得る。これらの観点毎の複数の質問項目間のまとまりを統計学的に見つけ出す方法が因子分析である。因子分析に関する解説書等は, 数多く存在するので, ここでは解釈の流れについての事例を紹介する。

1. 平均値の比較

表1は, 本学で行っている学生に対する学修状況調査(コンソ提供の共通調査)において, 授業経験に関する設問「あなたが受講した大学の授業で, 次のようなことを経験する機会はあるくらいありましたか」に挙げた14の項目である。選択肢は, 「ひんぱんにあった」, 「ときどきあった」, 「あまりなかった」, 「まったくなかった」の4つより回答する形式である。

表1. 授業経験に関する設問項目と選択肢

	ひんぱんにあった	ときどきあった	あまりなかった	まったくなかった
実験, 実習, フィールドワークなどを実施し, 学生が体験的に学ぶ	4	3	2	1
仕事に役立つ知識やスキルを学ぶ	4	3	2	1
授業内容と社会や日常生活のかかわりについて, 教員が説明する	4	3	2	1
授業の一環でボランティア活動をする	4	3	2	1
学生自身が文献や資料を調べる	4	3	2	1
定期的に小テストやレポートが課される	4	3	2	1

データ分析3 因子分析

	ひんばん にあった	ときどき あった	あまりな かった	まったく なかった
教員が提出物に添削やコメントをつけて返却する	4	3	2	1
学生が自分の考えや研究を発表する	4	3	2	1
授業中に学生同士が議論をする	4	3	2	1
授業で検討するテーマを学生が設定する	4	3	2	1
授業の進め方に学生の意見が取り入れられる	4	3	2	1
取りたい授業を履修登録できなかった	4	3	2	1
出席することが重視される	4	3	2	1
TA・SA（上級生や大学院生の授業補助者）から補助を受ける	4	3	2	1

この質問では、経験頻度の高い順に、4, 3, 2, 1 と数値化されているので、これを順序尺度と見て平均値を求め、それぞれ個別の項目の経験頻度の目安とし、比較検討することができる。さらに、項目同士の関係性を確認し、これら授業経験に関する各項目の回答の特徴を明らかにして全体をまとめてより分かりやすく議論する方法が次に紹介する因子分析である。

2. 因子分析

表2は、宮本ら（2016）で行った因子分析を元に、ある年の1年生を対象とした調査結果について3つの因子の因子負荷量を整理したものである。質問項目は表1と同じものである。絶対値で因子負荷量の高い順に4つの数値を太字で示している。以下、この結果をどのように見ていくかの事例を箇条書きで示す。

表2. 授業経験に関する設問項目と因子負荷量

	第1因子	第2因子	第3因子
授業中に学生同士が議論をする	0.737	0.097	0.107
学生が自分の考えや研究を発表する	0.702	0.156	0.250

	第1因子	第2因子	第3因子
授業で検討するテーマを学生が設定する	0.681	0.113	-0.051
授業の進め方に学生の意見が取り入れられる	0.523	0.095	-0.134
仕事に役立つ知識やスキルを学ぶ	0.309	0.643	0.090
学生自身が文献や資料を調べる	0.301	0.225	0.604
取りたい授業を履修登録できなかった	0.293	-0.150	0.114
授業内容と社会や日常生活のかかわりについて、教員が説明する	0.290	0.526	0.228
教員が提出物に添削やコメントをつけて返却する	0.286	0.046	0.083
授業の一環でボランティア活動をする	0.234	0.631	-0.449
TA・SA（上級生や大学院生の授業補助者）から補助を受ける	0.198	0.110	0.135
出席することが重視される	-0.004	-0.065	0.150
実験、実習、フィールドワークなどを実施し、学生が体験的に学ぶ	-0.028	0.598	0.309
定期的に小テストやレポートが課される	-0.051	0.159	0.699

それぞれの因子において、因子負荷量（絶対値）上位4位までを太字で示した。

①因子負荷量は、-1 から 1 の値を取り、絶対値で 1 に近いほどその項目が因子に強く影響していることを示す。ここでは、上位4位までを太字で示しているが、例えば因子負荷量 0.4 以上や 0.5 以上を基準として議論をすすめる解説をしている事例も見られる。

②表2では、第1因子の因子負荷量の高い項目が上位になるよう並び替えを行っている。その項目は、「授業中に学生同士が議論をする」、「学生が自分の考えや研究を発表する」、「授業で検討するテーマを学生が設定する」、「授業の進め方に学生の意見が取り入れられる」の4つである。続いて高い因子負荷量は、0.309 の「仕事に役立つ知識やスキルを学ぶ」であり、先に示した4つの項目に比べて小さく 0.4 以下となっている。

データ分析3 因子分析

- ③よって、第1因子と関係が強い項目は、「授業中に学生同士が議論をする」、「学生が自分の考えや研究を発表する」、「授業で検討するテーマを学生が設定する」、「授業の進め方に学生の意見が取り入れられる」の4つであると解釈することができる。さらにこれらの項目の内容から第1因子が持つ意味を考察する。この4つの項目は、アクティブ・ラーニングの要素を含む授業経験と見ることができるので、第1因子を「アクティブ・ラーニング型の授業」と意味付けすることができる。
- ④同様に第2因子を見ていくと、上位4つの項目は、「仕事に役立つ知識やスキルを学ぶ」、「授業内容と社会や日常生活のかかわりについて、教員が説明する」、「授業の一環でボランティア活動をする」、「実験、実習、フィールドワークなどを実施し、学生が体験的に学ぶ」である。これらは、「実社会に関連した授業」と意味付けすることができる。
- ⑤第3因子については、因子負荷量が比較的小さく、第1, 2, 3因子の順に第2因子と重なる項目もあり、関連性の意味付けが難しいと解釈することができる。因子寄与は、第1, 第2, 第3の順に低くなる。
- ⑥以上より、授業経験に関する14項目の質問に対する回答を特徴付ける「アクティブ・ラーニング型の授業」と「実社会に関連した授業」の2つの因子を見つけることができた。これにより近い意味を持つ質問項目をまとめて見ることができ、またこれらの因子に係る項目を合わせて新しい指標として議論することもできる。
- 引き続き、以下は、宮本ら（2016）で行った分析の具体例である。
- ⑦同調査では、3年生を対象にも行われ、またそれぞれ3年分のデータの分析を行った。よって計6組の授業経験に関する回答結果の因子分析を行った。結果は、学年、調査年によらず、6組すべてにおいて、アクティブ・ラーニングの要素を含む授業経験が、第1因子の因子負荷量の高い項目となった。一方、第2因子以降については共通性が見出されず、学年、調査年度によってばらつくため、第1因子「アクティブ・ラーニング型の授業」に注目して分析を進めた。
- ⑧続いて各回答の第1因子得点を求め、正の値を持つ回答を「アクティブ・ラーニング型の授業の経験の度合いが高い群（高群）」、負の値を持つ回答を「同経験の度合いが低い群（低群）」とし、新しい指標にした。

- ⑨新しい指標の高群と低群間において、他の質問項目である能力の獲得状況や成績などの差を見ながら、それらに対する授業経験（アクティブ・ラーニング型の授業経験の度合い）の影響を議論した。

3. まとめ

因子分析により、複数の質問項目間の関係を統計学的に把握することができ、それぞれの項目に共通する内容を議論することができる。多項目にわたるアンケートの質問項目を整理して結果を考察していくうえでは有効な分析手法である。本稿では、既存の質問項目を用いた調査結果の因子分析を行い、共通因子を見出した例を挙げた。逆に因子と質問項目の関係性を予測した仮説モデルを組み込んだ質問項目を設計し、モデルが正しいかどうかの検証にも因子分析を使うことができる。

各因子の意味付けについて、上記例の第2因子では「実社会に関連した授業」としたが、他にも「基礎科目の応用に関する授業」、「職業訓練に関する授業」などが挙げられる。例えば、基礎科目と応用科目の比較にこの因子を使うのであれば「基礎科目の応用に関する授業」とし、大学における就職支援を議論するのであれば「職業訓練に関する授業」として議論を進めるだろう。つまり、この意味付けは分析者が主観的に行っていることになるが、ここはあくまでも解釈であるので、議論の内容に沿って誰もが理解することができる意味付けができれば良い。ただし、この因子は統計学的に見つけ出した潜在的な因子であることを忘れずに議論を進める必要がある。因子の構造が分かりにくい場合に、上記事例で複数の因子において因子負荷量が高くなる項目（「授業の一環でボランティア活動をする」や「実験、実習、フィールドワークなどを実施し、学生が体験的に学ぶ」）を除くなど条件を変えて何度も因子分析を行う必要もある。以上のように試行錯誤が必要になることもあるが、統計学的な根拠を持って調査結果をわかりやすくまとめる有効な方法である。

参考文献

宮本淳，徳井美智代，山田邦雅，細川敏幸（2016）「授業経験の質の差異が学生の学習態度・能力の自己評価に与える影響—2012～2014年学生調査の分析結果より」高等教育ジャーナル—高等教育と生涯学習—，23，79-85.

1. 比較が重要

教育制度改革の担当者に刺激を与えるには、IR 調査で得られた単なる集計データを渡すだけでは不十分である。数字の羅列は、受け取った教員からみれば、あまり意味を成さない。個々の数字の意味するところがわからないからである。改革につながるデータとするには、何ものかとの比較が重要である。比較対象として考えられるものは、世界の大学、米国、日本全体、日本の理系学部、文系学部、医療系学部、自大学全体、自大学の理系学部、文系学部、医療系学部などである。最もふさわしいのは、日本の類似学科の平均値との比較である。

例えば、A 大学工学部建築学科の教育評価を考えてみよう。学内データで比較可能なものは、工学部全体、理系学部全体などである。ところが実際に分析してみると、工学部の学生数が多いため、理系学部全体との比較は差がないことが多い。理系学部には工学部が含まれ、しかも工学部学生の比率が高いためである。この場合は、日本全体の理系学部と比較することで自学の特徴が見えてくる。大学 IR コンソーシアムに加入していれば、日本全体の理系学部と比較できる。

北海道大学の総合 IR 室は、毎年度秋に各学部を訪問し、前年度の 3 年生データ、卒業生データから分析された結果を報告している。そこで、比較の対象にするのは、当該学部の前年度データ、理系、文系あるいは医系データをまとめたものである。場合によっては、上記工学部のように全国平均と比較することもある。

2. 目的を明確に

もうひとつ重要なことは、教育目標（ディプロマポリシー）とアンケートの設問項目との対応である。第 2 章表 4 のような表を作成し、該当するアンケート項目の結果を学内平均などと比較し大きく外れているものについては、カラムの色を変えて指摘している。達成度の低い項目を指摘する一方で、達成度の高い項目も発見できるので、教育組織としての長所も読み取れる。このような分析をすることで、ディプロマポリシーを達成できるような教育システムになっているか否かを推測できる。

以上のように、比較対象や目標を明確にした上での分析を提示することで、各学部に教育改革をうながすことができる。さらに北海道大学では、訪問 2 ヶ月後に各学部にアンケート記入を依頼し、報告した IR データがどう使われ、どのように教育改革に利用されているかを調査している。きめ細かいサービスとその効果測定が次の改革につながるのである。

序章でみたように、日本の教学 IR は各大学で導入が始まったばかりである。しかも、日本の大学は 2004 年の国立大学の法人化以降文科省からの補助金を毎年減額され、予算の減額や教員数の減少を余儀なくされている（岩本 2019）。しかも、これまでお手本にしてきた米国の高等教育機関も、保護者の所得格差の学生への影響、授業料の値上げ、教育ローンの大きな負担（スティグリッツ 2012）、政府補助金の減額（州立大学の補助金は 10% 程度、日本の国立大学は 3～4 割）など、多くの課題をかかえており、参考にならなくなっている。

教学 IR の担当者は、優秀な専任の教員であることが望ましいし、日本でも米国の AIR のような学会の設立が期待される。この困難な時代に、どれくらい IR 組織を充実させるかが、大学の将来を左右するかもしれない。なぜなら、教育改革を支援する強力な支柱の一つが教学 IR だからである。日本の教学 IR が今後とも発展していくよう祈念したい。

参考文献

- 岩本宣明（2019）「科学者が消える：ノーベル賞が取れなくなる日本」東洋経済新報社（東京）
- スティグリッツ、ジョセフ・E、 楡井浩一 峯村利哉訳（2012）「世界の 99% を貧困にする経済」徳間書店（東京）

《本書は科学研究費補助金基盤研究(B)(一般)「IR によるカレッジ・インパクト理論の検証と STEM 教育評価モデルの構築」の成果である。》



教学 IR 導入ガイド（初版）

2020 年 3 月発行

著者：細川敏幸，宮本 淳，山田邦雅

連絡先：〒060-0817 札幌市北区北 17 条西 8 丁目

北海道大学 高等教育推進機構 高等教育研究部

TEL:011-706-7520 URL: <https://high.high.hokudai.ac.jp/>
