

教学IRにおける データ分析入門

2025年8月29日(金) 10:00~12:00

愛媛大学真鍋亮

講師紹介：真鍋 亮

- 香川県高松市出身
- 愛媛大学教育・学生支援機構教育企画室
- 高等教育論、教育経済学
- 日本スポーツ協会公認コーチ3
- manabe.ryo.gw@ehime-u.ac.jp



■ 最近書いたり発表したもの

- 真鍋亮(2024)「高校での経験と大学教育の関連に関する実証的研究」
- 真鍋亮(2024)「学生時代の経験と社会の関連に関する実証的研究」
- 真鍋亮・島一則(2024)「同一大卒者男子における学業成績と期待生涯賃金の関係性—学習効率と雇用効率に着目して—」
- 真鍋亮(2024)「大学成績の規定要因に関する実証的研究—学校歴と成績の効果に関する比較検証—」
- 真鍋亮(2024)「日本における社会科学系教育の効用—社会科学系学部卒業生の所得に関する分野別比較—」



プログラムの目標と構成

■ プログラムの目標

- ・IRの基本概念を説明することができる
- ・データの基本的な分析方法を知る
- ・分析結果を活用につなげるポイントを知る

■ プログラムの構成

- ・IRの概要
- ・データの整理と可視化
- ・データの分析
- ・活用につながる報告
- ・まとめと振り返り

IRの概要

IR (Institutional Research) の定義

- 機関の計画立案、政策形成、意思決定を支援するための情報を提供する目的で、高等教育機関の内部で行われる調査研究 (Saupe 1990)
- 教育と学習（あるいは学修）を改善するためのデータ収集・分析・報告の実践、あるいは研究が教学IR (松田・渡辺2017)

定義の特徴

■ 中央教育審議会の定義

説明に「能動的な学修」が使われている
ラーニングなのに「教授・学習法」？

さまざまな形態を含む「総称」

広い能力の育成を目指している

■ 定義自体の難しさ

学習の能動性が外部から直接に判断できない
研究者によってさまざまな定義

学習の形態を重視（活動型学習）

学習の姿勢や態度を重視（主体的学習）

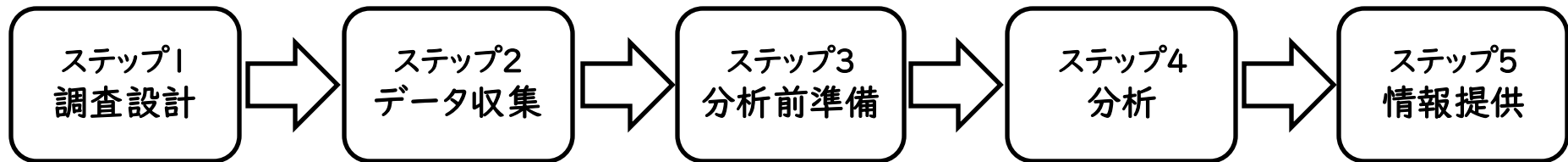
ワーク

■ 自己紹介をしてください

- ・ご所属
- ・最近うれしかったこと（仕事、私生活どちらでもOK）
- ・IR業務との関わり（なぜ興味をもったか、でもOK）

5m 00s

IRの流れ



調査設計

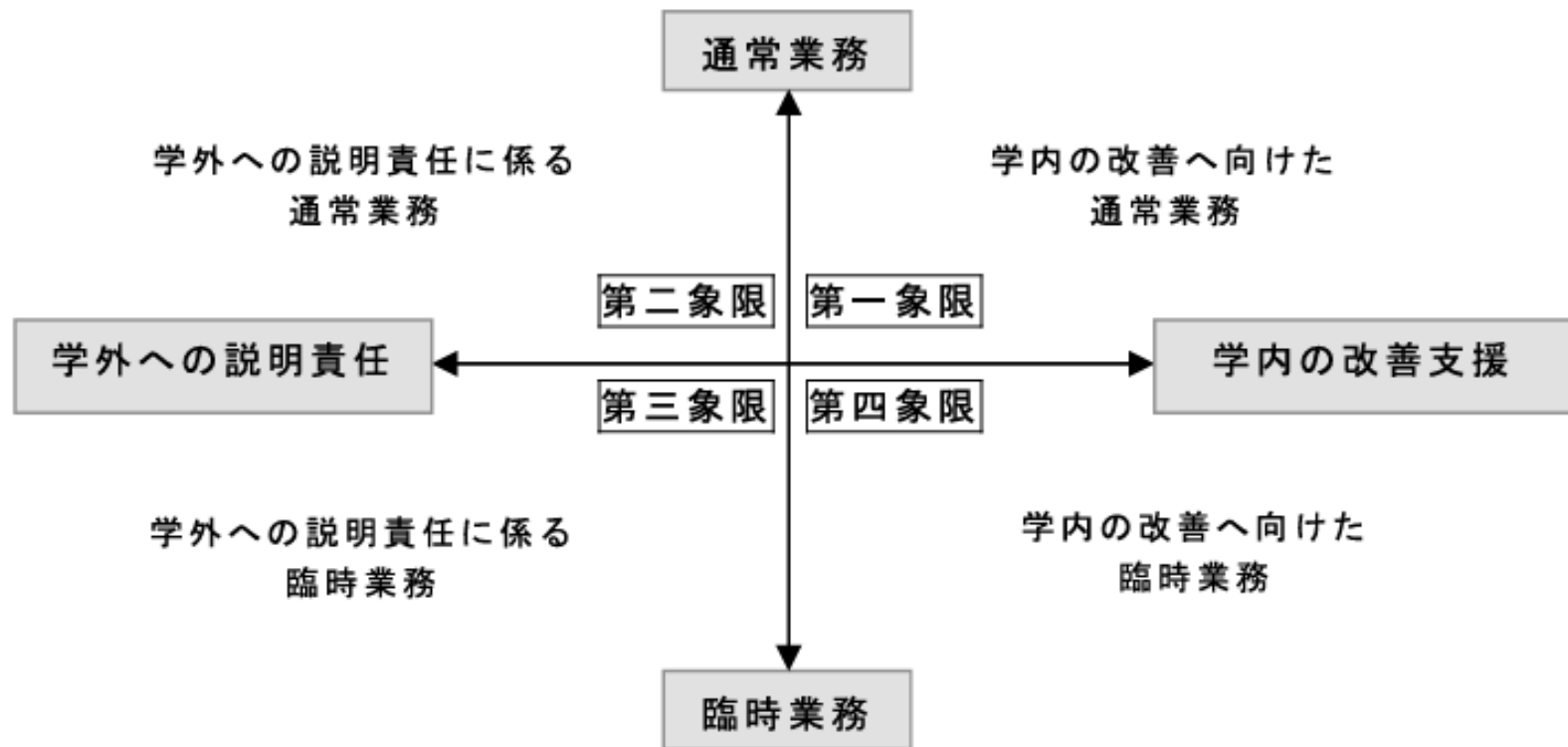
- 目的を明確にする

通常業務か臨時業務か、学外への説明責任か学内の改善支援か（本田ほか2014）

- 仮説検証か、問題の構造や原因を探る探索的な調査か

- 「何のための調査か」を意識して設計する
計画した調査が目的を達成できるか

IR業務の四象限



データ収集

- 学内の既存データ

まずは学内データベースにあるデータを把握し、活用することが重要

- 学内で新たに収集するデータ

対象者の負担を考慮し、無駄のない調査設計を行う

- 学外データ

卒業生、他大学、高等学校、政府統計、企業など、学外の組織が所有するデータも重要

分析前準備

- データクレンジングと整備

収集したデータから不要な情報や重複を削除し、形式を統一する、明らかに不正確なデータ(200歳の学生情報など)を確認・修正・削除

- 欠損値の処理

基本はリストワイズ(削除)、もっともらしい値を計算して割りあてる場合も

- データ統合(複数のDBからデータを収集する場合)

- データの名寄せ

同一人物とみられるデータを統合する

分析

■ 基本的な分析方法

単純集計、時系列の推移、全国平均との比較など
データを意味ある情報に変換する

■ やや高度な分析手法の活用

複数の要因間の関係性を知る（多変量解析）

- ・相関関係：どれくらい似た傾向があるか
- ・因果関係：どれくらい影響を与えているか

■ 定性的データ

データは数値だけではなく、インタビューや文書など質的なデータもある

数値で表される量的データの分析結果の解釈に使うことも

情報提供

- 依頼者への報告

報告の形式（口頭または文書）は状況に応じて選択
データ提供者への完了報告も

- 聞き手に応じた内容の調整

聞き手の知識レベルや報告に何を求めているかを考慮する
意思決定に影響を与える場合は、具体的な結果や影響を明示する

- 簡潔で要点を押さえた報告

執行部など多忙な聞き手のスケジュールを考慮
重要なポイントを簡潔かつ明瞭に伝える工夫が必要

データの整理と可視化

教学マネジメント指針

- 大学の教育活動に伴う基本的な情報であって全ての大学において収集可能と考えられるデータ

各授業科目における到達目標の達成状況、学位の取得状況、学生の成長実感、満足度、進路の決定状況等の卒業後の状況（進学率や就職率等）、修業年限期間内に卒業する学生の割合、留年率、中途退学率、学修時間

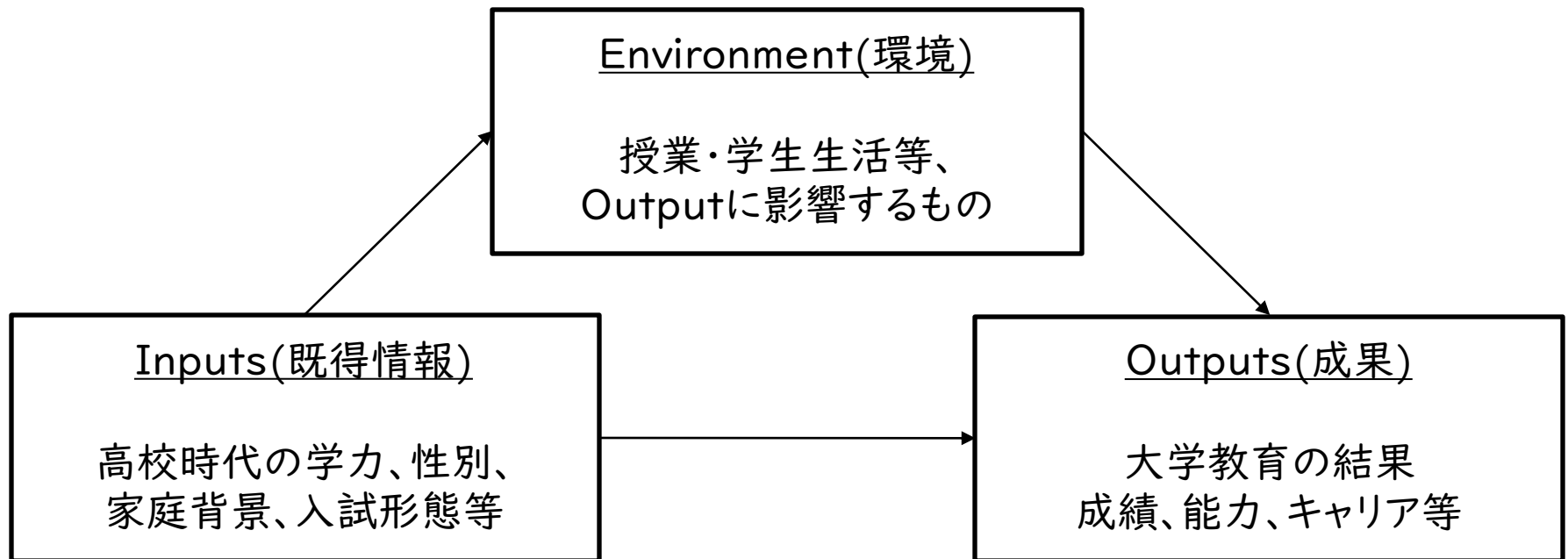
- 教学マネジメントを確立する上で各大学の判断の下で収集することが想定されるデータ

「卒業認定・学位授与の方針」に定められた特定の資質・能力の修得状況を直接的に評価することができる授業科目における到達目標の達成状況、卒業論文・卒業研究の水準、アセスメントテストの結果、語学力検定等の学外試験のスコア、資格取得や受賞、表彰歴等の状況、卒業生に対する評価、卒業生からの評価

教学IRにおけるデータ分析の留意点

- 指標が多く複雑であるため、一度に全てを網羅した分析は難しい
- 大事な部分にだけ注目して、いったん他は無視してみる（モデル化）
- IEOモデルなど、既存のモデルを活用する

学習成果を捉えるIEOモデル



Astin(1993)

今回取り上げる分析事例

- 同一大学卒者男子における学業成績と期待生涯賃金の関係性ー学習効率と雇用効率に着目してー（真鍋、島2024）

キーワードっぽいもの：

- 同一大学卒者（学力などの条件を同じにする）
- 男子（女子よりキャリアが単純）
- 学業成績（高校の評定平均値と大学のGPA）
- 期待生涯賃金（卒業時点の就職先から推計）
- **学習効率と雇用効率**（のちほど説明します）

知の総和答申

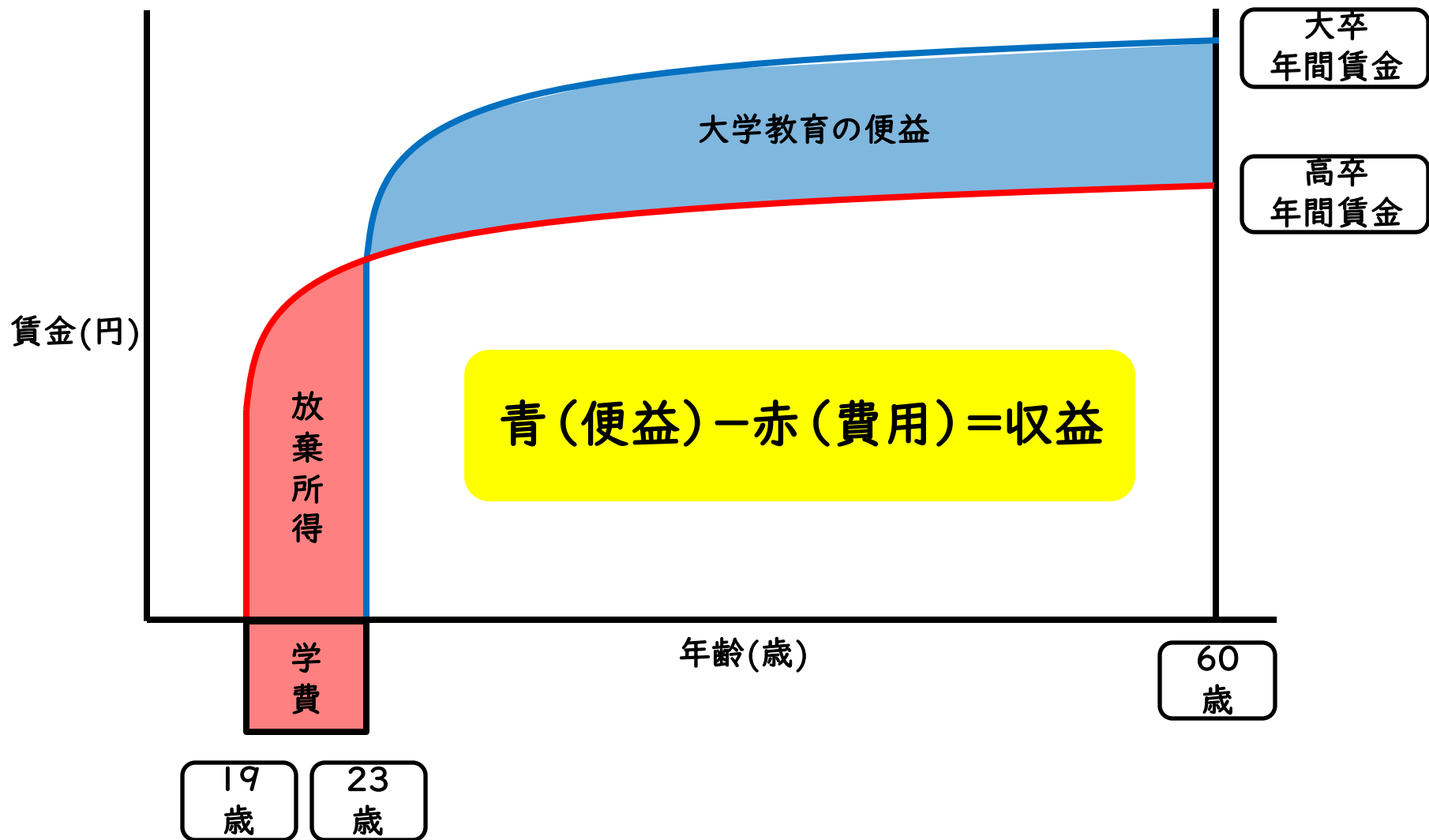
■ 高等教育改革を支える支援方策の在り方

- ・高等教育の価値
→ 高等教育への投資は未来への先行投資
- ・高等教育への信頼
→ 教育研究の成果や効果の情報公表
- ・必要コストの算出
→ 教育コストの明確化
- ・高等教育投資の在り方
→ 財源の多角化

研究の前提：知の総和答申に対する疑問

- 高等教育に投資してもらうための理由は？
 - ・何を根拠に高等教育への投資は未来への先行投資というべきか？
 - ・投資価値（高等教育の成果や効果）をどうやって示すのか？必要コスト（投資額）は算出するけど、コストベースの訴求では、投資の魅力や意義は伝わりにくい
- リターンベースの訴求も必要では？

赤（かけた費用）より青（得られる便益）の面積が大きい＝大学教育の経済的効果！



学習効率

効率がいい



効率が悪くい



成果



労力
時間
能力



雇用効率

効率がいい



効率が悪くい



賃金



成績
労力
費用



この研究の背景と目的

- 近年、政策文書にも出ているように大学教育の経済的価値が社会的に注目されつつある
- すでに大学教育の経済的価値を学習効率と雇用効率で説明しようとしている研究もある
- 大卒といっても学校によって収益（雇用効率）はバラバラだが、このバラつきは学習効率を絡めると説明できそうだ
- だから、学習効率と雇用効率の関係（**学習効率がいい人は雇用効率もいいこと**）を明らかにします

つまり、こういうこと

■ 高校、大学、社会のつながり

高校ランク・成績、大学成績（学習効率）、期待生涯賃金（雇用効率）に因果関係があるかどうかをみる

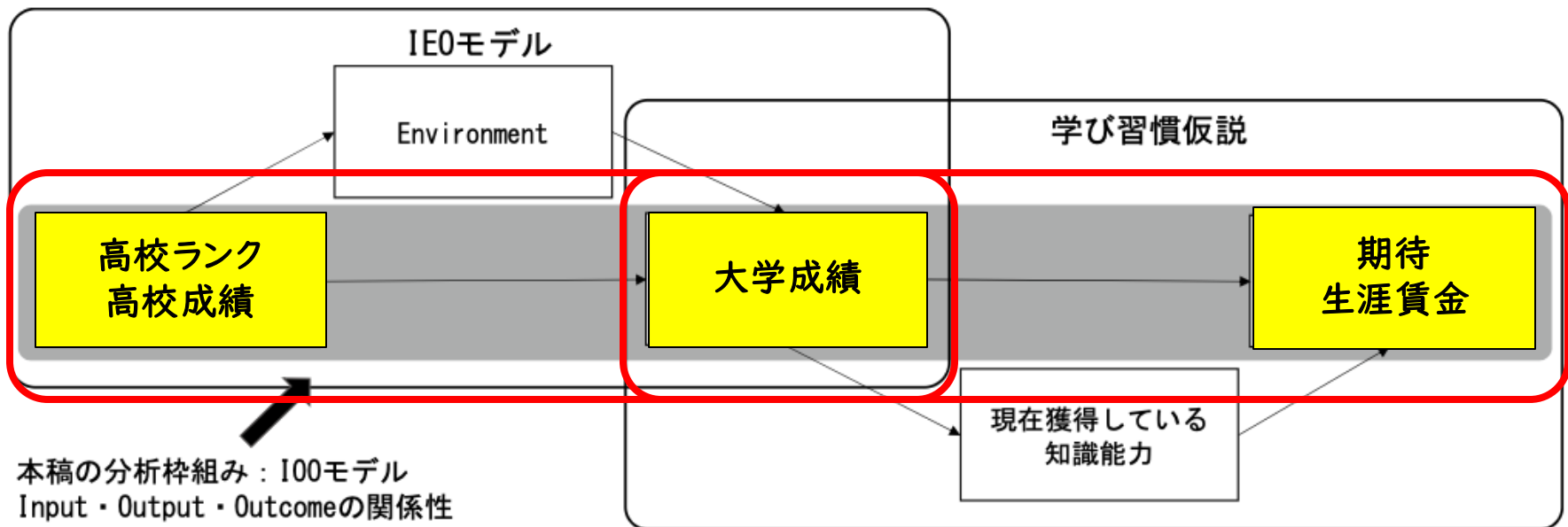


図3 本稿の分析枠組みIOOモデル（筆者作成）

データを整理する

- 教学IRのデータは人（学生や教員など）が持っている
- 表形式に整える（縦に人、横にその人のデータを並べる）

学生番号	学部	氏名	性別	高校名	入試形態	高校評定	GPA	週あたり学習時間	満足度
111001	文	田中 太郎	男	A高校	一般	3.4	3.4	9.3	2
111002	文	鈴木 一郎	男	B高校	一般	3.6	2.4	7.5	3
111003	文	佐藤 花子	女	C高校	総合型	4.2	3.4	10.8	4
211001	経済	高橋 健二	男	D高校	一般	3.4	2.5	7.2	3
211002	経済	伊藤 美咲	女	E高校	学校推薦型	4.3	3.3	10.2	3
211003	経済	山本 優作	男	F高校	総合型	4	3	9.9	2
311001	工	松本 翔平	男	G高校	一般	4.3	4	11.1	2
311002	工	中村 彩香	女	H高校	一般	3.7	2.8	8.4	4
411001	教育	小林 涼子	女	I高校	学校推薦型	3.7	2.4	8.1	3
411002	教育	渡辺 咲良	女	J高校	総合型	4.2	3.1	9.9	2

データの傾向をみる

- 平均値: データ全体の合計をデータの個数で割った値
(データの中心的な傾向を示す)
- 中央値: データを小さい順または大きい順に並べたとき、
中央に位置する値 (外れ値の影響を受けにくい)
- 標準偏差: データが平均値からどの程度散らばっているか
を表す値 (データのばらつきを測る)
- 最大値: データ範囲の上限を示す値
- 最小値: データ範囲の下限を示す値

データを可視化する

- 可視化とは、目にみえない（みえにくい）内容をグラフなどにしてみえる状態、みたい時にみられるようにすること
- 単なる数字の羅列ではイメージしづらい
- グラフは「何をみるのか？」という観点で選ぶ

用途別グラフの種類

- 比較する

 - 1 時点：棒グラフ、複数時点：折れ線グラフ

- 分布をみる

 - 1 項目：ヒストグラム、2 項目：散布図

- 構成をみる

 - 1 時点：円グラフ、複数時点：帯グラフ

- 関係をみる

 - 散布図

ワーク

- 配付資料（論文）のうち、以下のケースでは、どのグラフを使うのが適していますか？

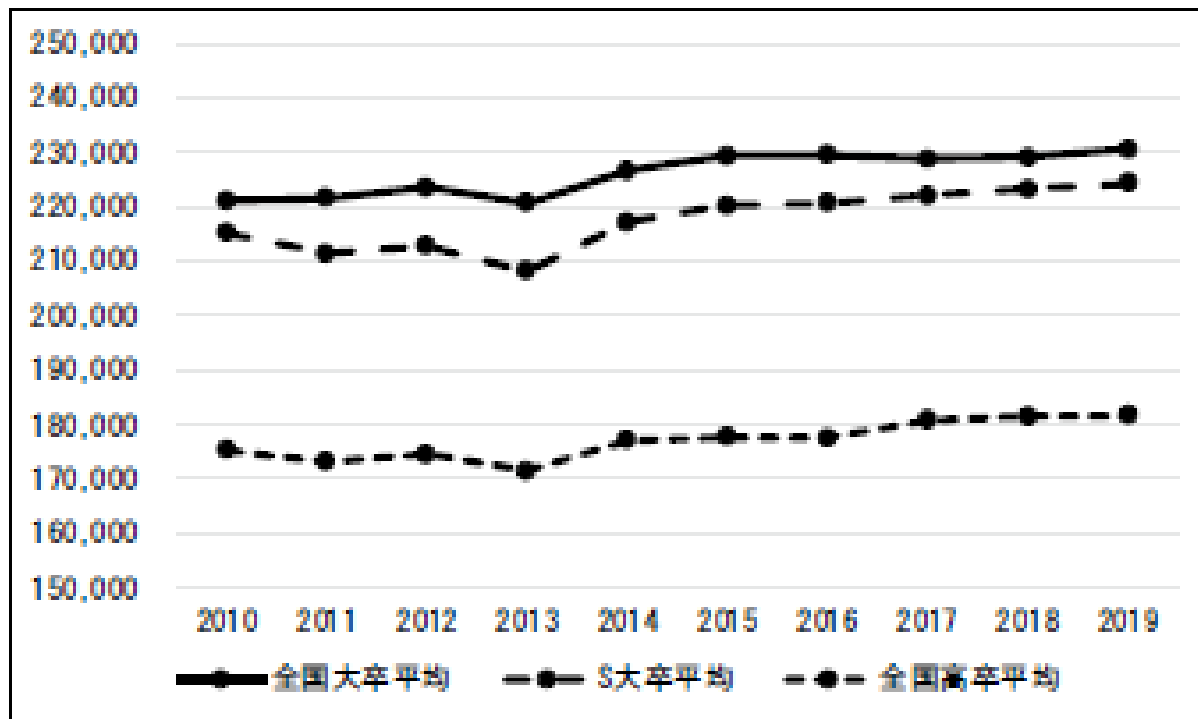
- (1) 10年間の学歴別期待生涯賃金の推移
- (2) 同一大学内での期待生涯賃金の散らばり（偏り）
- (3) 同一大学内で期待生涯賃金「全国平均以上」「高卒以上」「高卒以下」に分類した際の、それぞれの割合
- (4) (3)の10年間の推移

グラフの選び方

折れ線グラフ（比較：複数時点）

■ 10年間の推移（図5）

ほぼ横ばい（高卒と大卒の差もキープ）



(千円)

図5 期待生涯賃金の推移（筆者作成）

ヒストグラム(分布)

■ 期待生涯賃金の散らばり(図6)

同一大学であっても散らばりが見られる

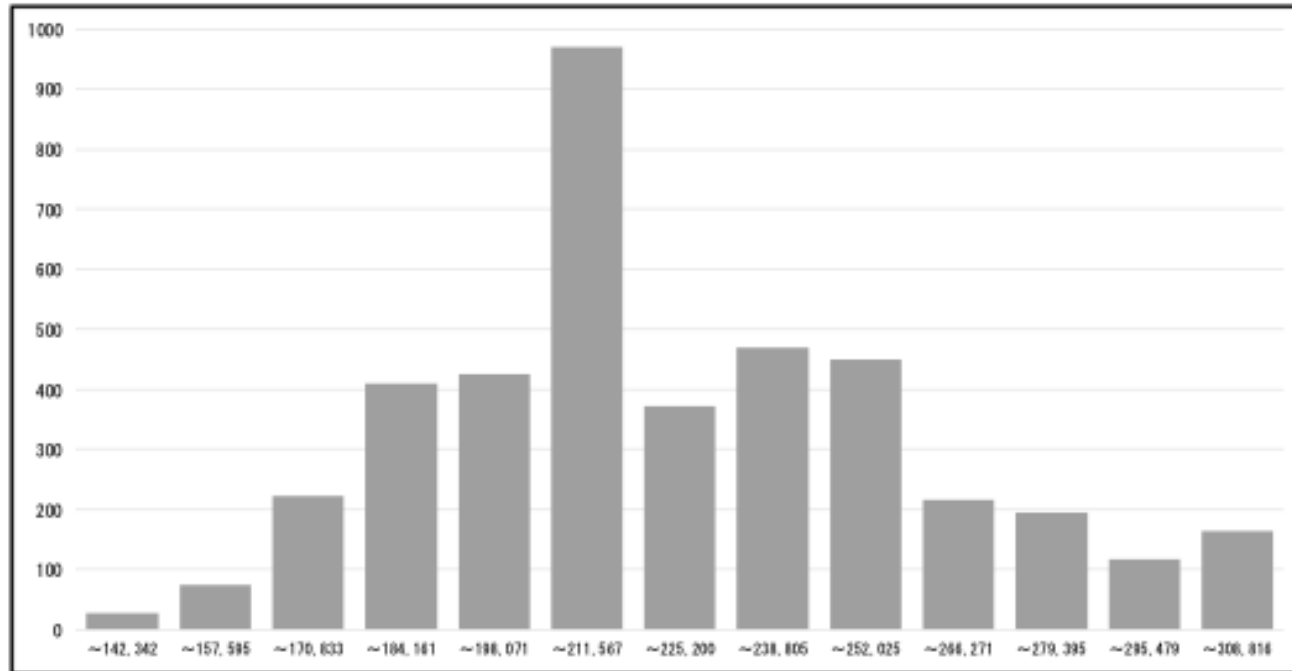


図6 S大学における期待生涯賃金に関するヒストグラム (筆者作成)

円グラフ（構成）

■ 賃金格差（図7）

同一大学でも賃金格差がある

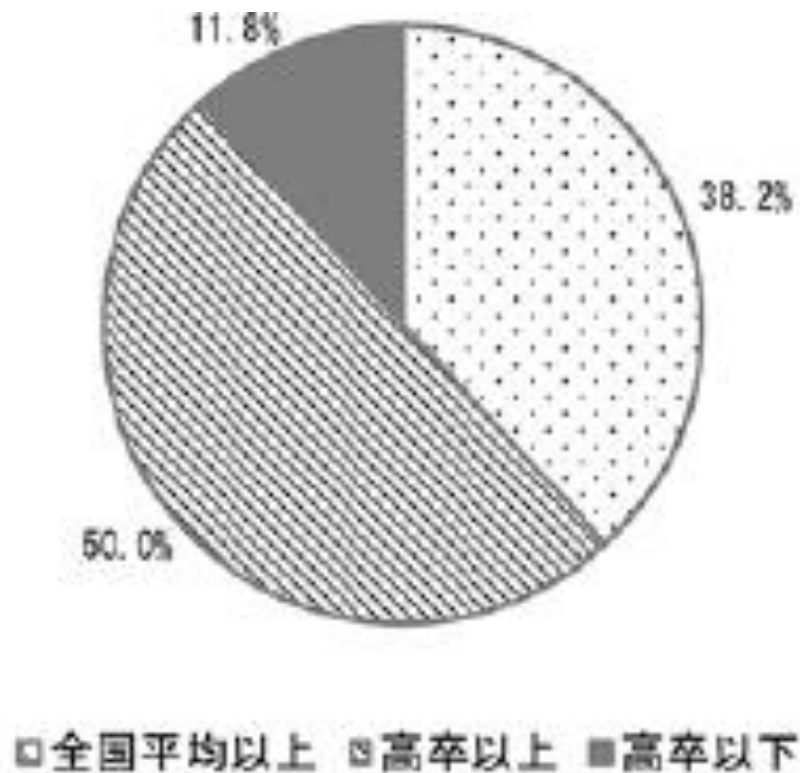


図7 S大学における期待生涯賃金の分散割合（筆者作成）

帯グラフ(構成の推移)

■ 賃金格差の推移(図7)

賃金格差は同程度の割合で推移している

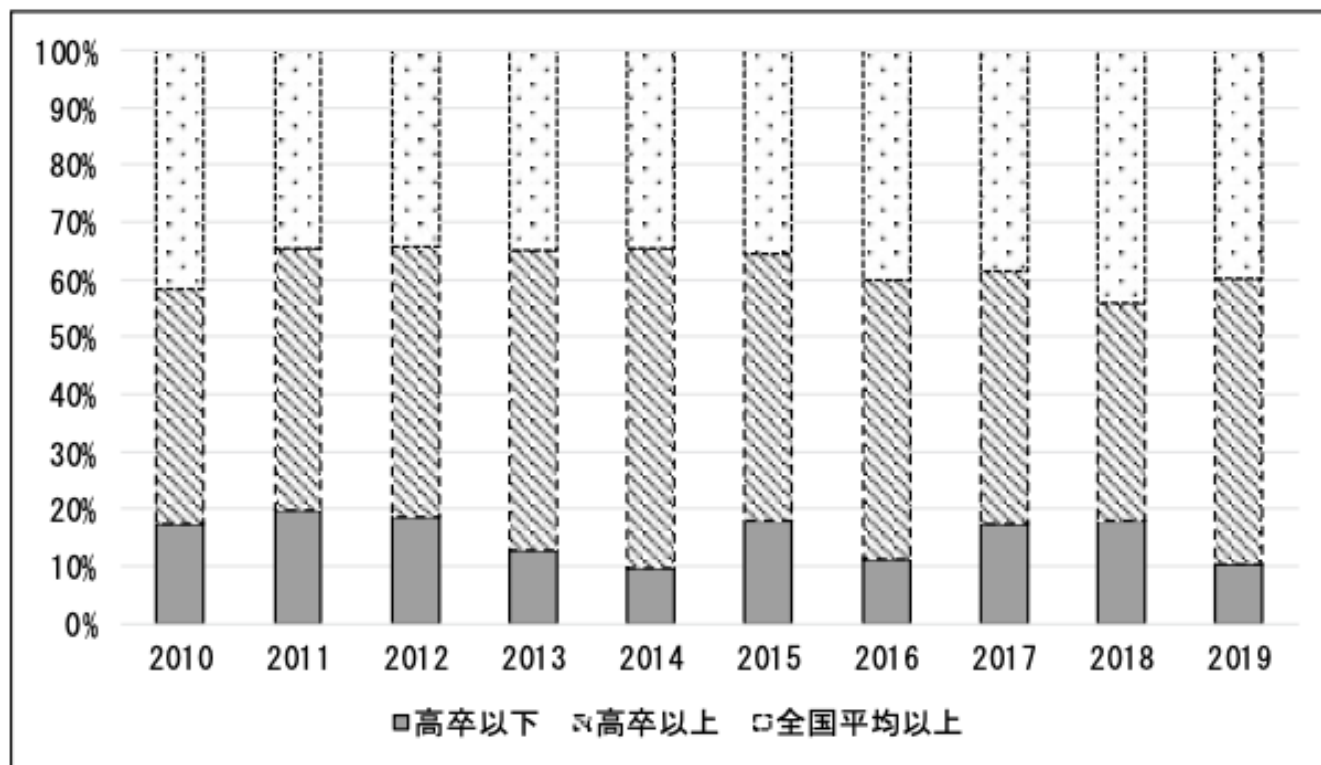
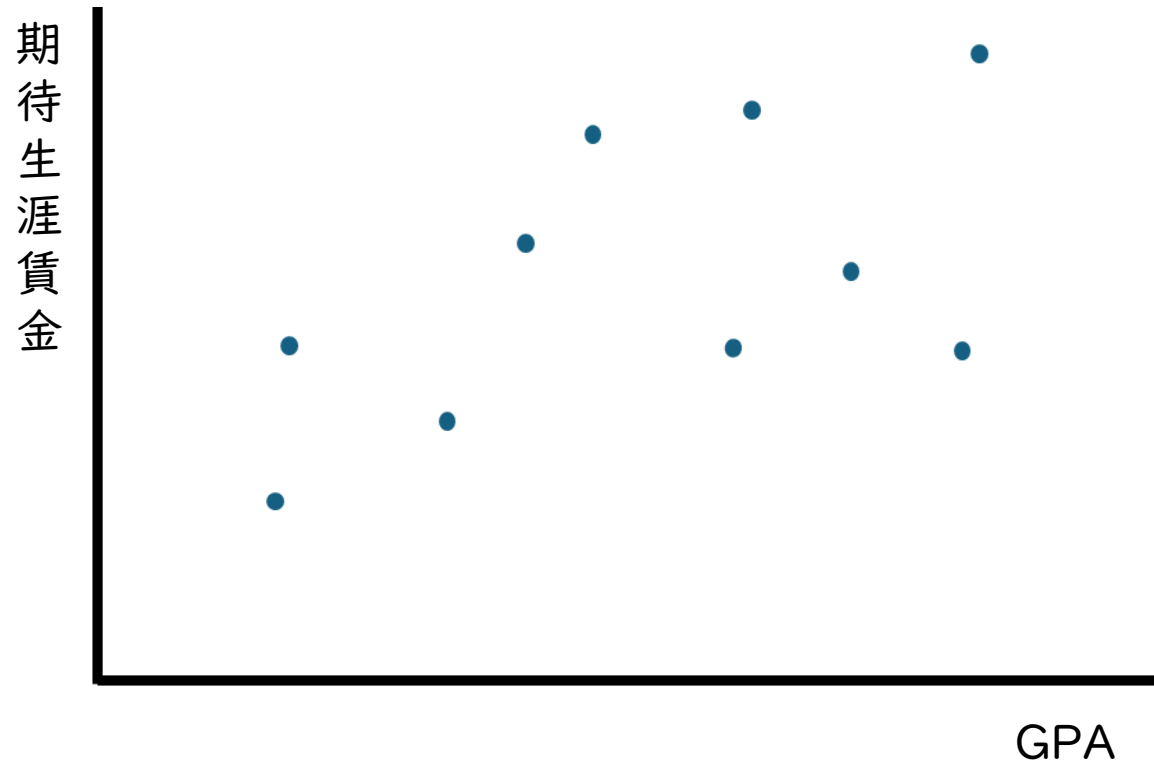


図8 S大学における期待生涯賃金の分散割合の推移 (筆者作成)

散布図(2つの要素の関係)

GPAが原因で生涯賃金は上がるのか？
上がるとしたらどのくらい？



散布図を作ってみると・・・

- 以下のデータセットで散布図を作るとします

①		②		③		④	
x	y	x	y	x	y	x	y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

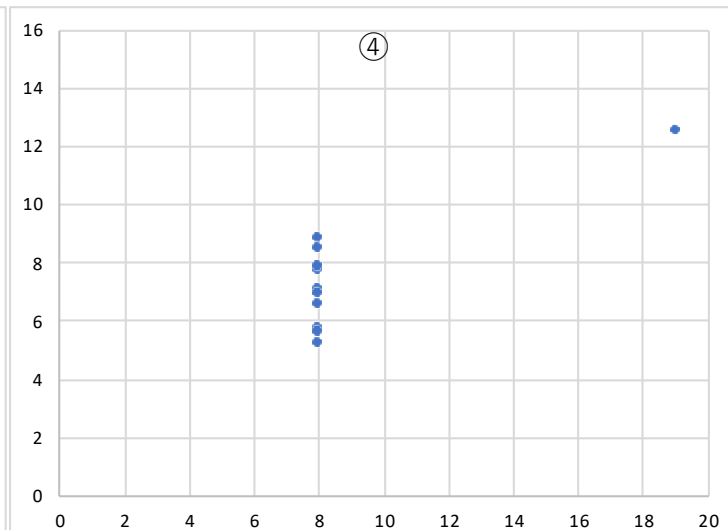
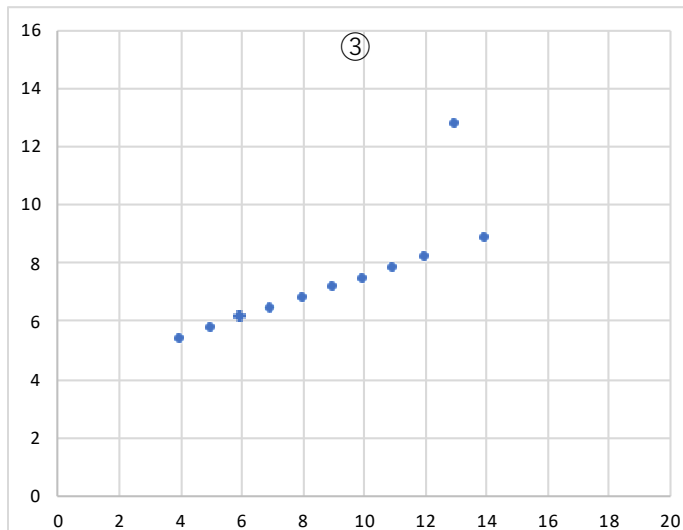
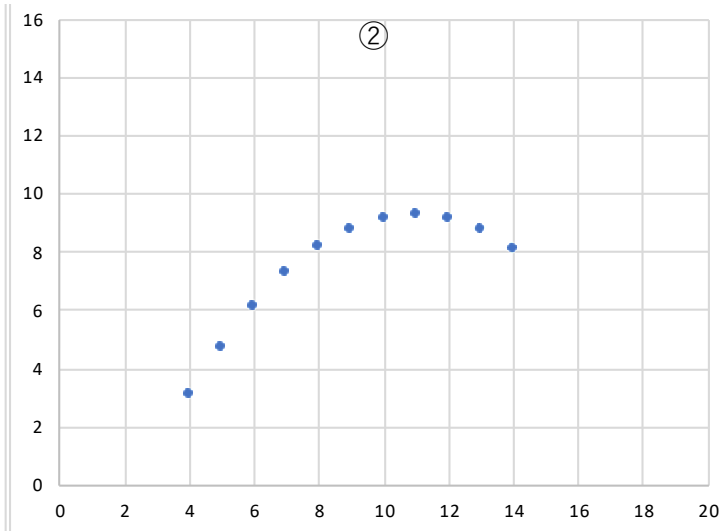
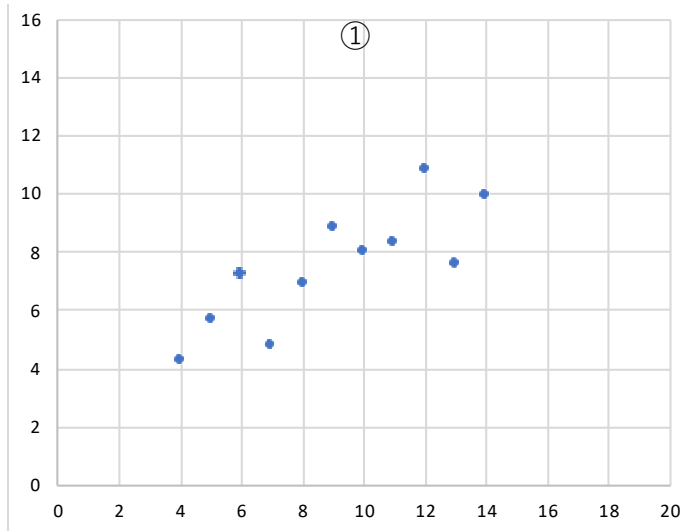
散布図を作ってみると・・・

- この4パターンのデータは、
平均値、標準偏差、相関係数が全て同じです

		平均	標準偏差	相関係数
①	x	9.00	3.16	0.82
	y	7.50	1.94	
②	x	9.00	3.16	0.82
	y	7.50	1.94	
③	x	9.00	3.16	0.82
	y	7.50	1.94	
④	x	9.00	3.16	0.82
	y	7.50	1.94	

アンスコムの四重奏

- 分布が全く異なっている! → 分布をみるのは大事



表作成の基本

- データの推移は上から下へ
- 数値は桁をあわせて右揃え
- 罫線は少なめに

表1 本稿の分析対象者（筆者作成）

卒業年度	偏差値	A学部	B学部	C学部	D学部	計
2010	47.94	199	160	73	95	527
2011	46.48	209	190	68	109	576
2012	48.48	236	175	53	102	566
2013	52.23	220	145	70	85	520
2014	52.84	216	152	61	107	536
2015	53.21	214	171	55	105	545
2016	53.16	208	178	64	104	554
2017	53.17	245	166	76	95	582
2018	53.06	220	154	71	120	565
2019	53.67	205	172	84	104	565
計	51.36	2,172	1,663	675	1,026	5,536

データの分析

データが活きる場面

■ データを活用する意義

- 現状を把握するための根拠
- 先入観や思い込みを修正する
- さまざまな他者との議論の基盤

■ データ分析の主な目的

- 現状把握
 - 今の状況を確認する
- 因果・相関関係の把握
 - 2つ以上の要素の関係を明らかにする
- 予測
 - 与えられたデータから今後について推測する

ワーク

- 以下のケースは偶然か、意味のある違いといえるか、理由とともに考えましょう

1. コインを10回投げて8回表がでた
2. 100人に新薬を投与したら5人に副作用がでた
3. 新しい学習法を取り入れたクラスで、テストの点があがった人が9割いた

偶然か、意味のある違いか

どこで線引きするか？

- コイン投げに何か仕込まれている可能性

どのあたりからイカサマを疑う？

6回 約20.5% 7回 約11.7%

8回 約4.4% 9回 約1.0%

10回 約0.1%

- 新薬の安全性

1人でも副作用でるのは・・・危険じゃない？

- 新しい学習法の効果

テストの点があがった人が9割いたら、効果ありそう？

統計の世界では5%が基準

- 統計学者フィッシャーの提案（1920年代）
- 基準が甘すぎると偶然の誤認が生じやすくなり、基準が厳しすぎると本当の効果を見逃す可能性がでてくる中で、ちょうど現実的で絶妙な線引き
- 検証前に反論（帰無仮説：検証に対する反論）を立ててそれを論破する、という方法をとる
※帰無（なかったことにする）
- P値： $P = \text{probability}$ （確率）、
5%未満で帰無仮説が起こらないと判断する場合が多い

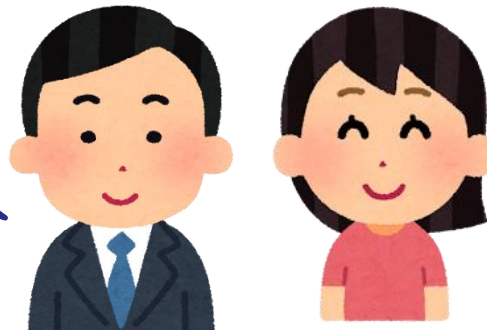
P値のイメージ

対立仮説

GPAと生涯賃金は関係
あるっぽいよ。
勉強がんばってね。

帰無仮説

いやいや、それは偶然で
しょ。



でも、この結果が偶然で
る確率は5%未満だよ。
(論破)

えっ?
そうなの...

有意差あり

統計的に
意味がある



帰無仮説を棄却

分析事例のP値（表4、表5）

表4 高校ランク・高校成績・大学成績・期待生涯賃金の相関係数（筆者作成）

	高校ランク	高校成績	大学成績	期待生涯賃金
高校ランク	1.	-.34**	.040**	.003
高校成績		1.	.284**	-.004
大学成績			1.	.082**
期待生涯賃金				1.

** 相関係数は 1% 水準で有意(両側)

表5 大学成績の規定要因（筆者作成、値は偏回帰係数）

	大学成績		期待生涯賃金	
	係数	標準誤差	係数	標準誤差
高校ランク	.011 ***	.001	210.391 *	153.998
高校成績	.196 ***	.009	-1105.18	1312.59
大学成績			8249.01 ***	2045.48
(定数)	1.412 ***	.043	201348. ***	6372.54
R ²	.098 ***		.004 ***	
N	4,110		4,110	

注：*** <.001 ** <.01 * <.05

分析の考え方

■ 分析の主な視点

- 全体の傾向（分布）をみる
- 条件の違いやグループに分けて比較する
- 指標間の関係を見る（因果、相関など）
- 時間経過による変化を見る
- 対象を分類する

■ 代表値の計算

- 平均値、中央値、最頻値

■ 表やグラフを作ることによって分析が進む

因果と相関

■ 因果と相関の違い（イメージ）

因果（回帰分析）＝予測

→Aさんがここに来たってことはBさんももうすぐ…

相関（相関分析）＝仲の良さ

→AさんとBさん、よく一緒にいるね

■ 因果と相関の違い（方法）

回帰分析＝**線を引いて予測式をつくる**（一次関数）

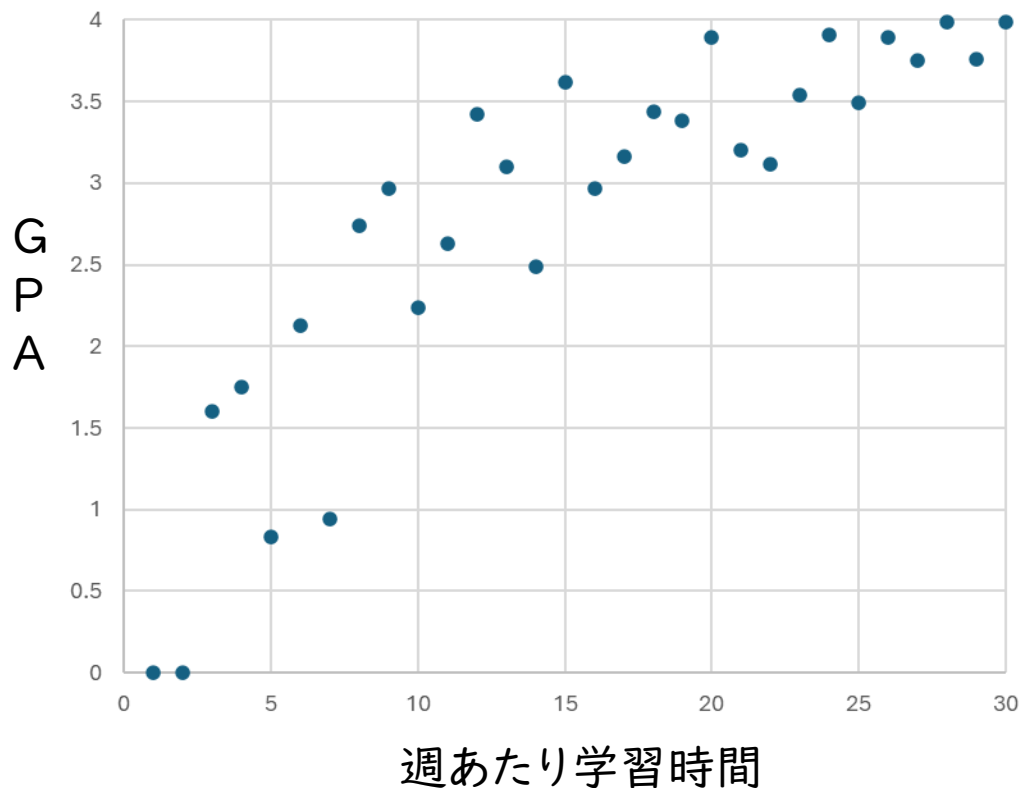
相関分析＝**散らばり具合から関係の強さ**をみる

■ 分析では、相関関係→因果関係の順で、確認される場合が多い

ざっくり相関分析のイメージ

- データとデータに何らかの関係がありそうなこと、関係がある場合にその強さを数値でビシッと表す

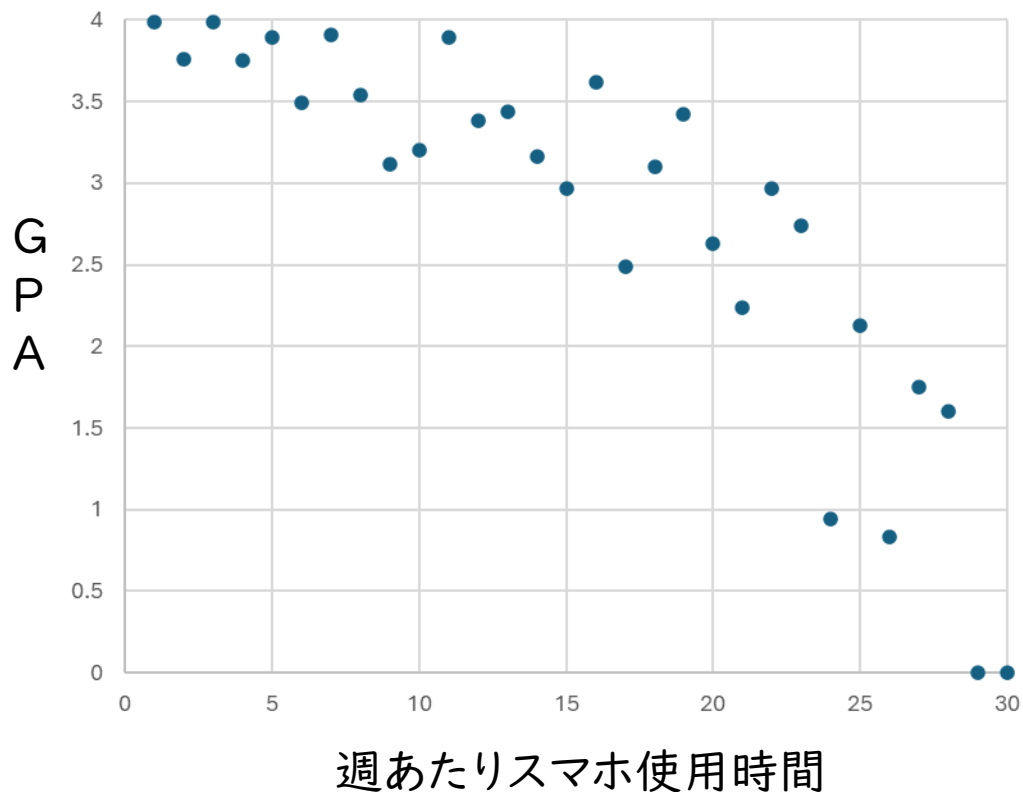
週あたり学習時間とGPA（正の相関関係）



ざっくり相関分析のイメージ

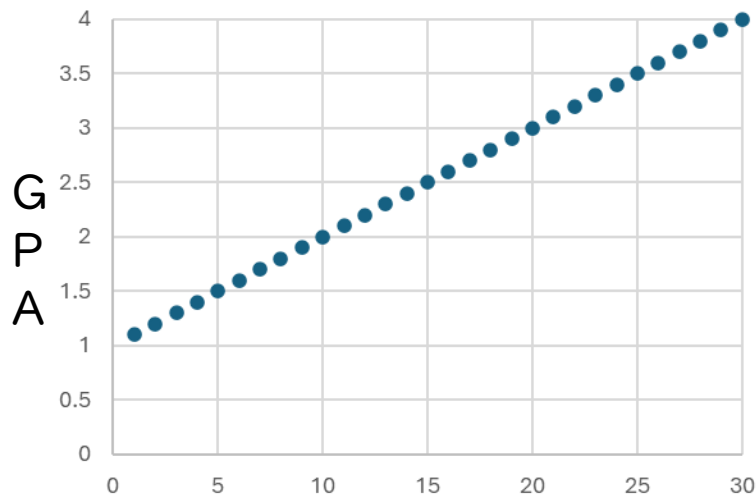
- データとデータに何らかの関係がありそうなこと、関係がある場合にその強さを数値でビシッと表す

週あたりスマホ使用時間とGPA（負の相関関係）



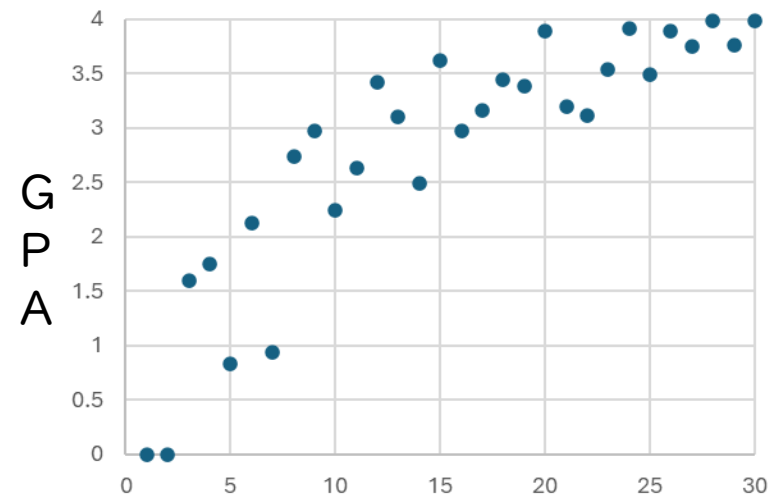
ざっくり相関分析のイメージ

強い正の相関 ($r=1$)



週あたり学習時間

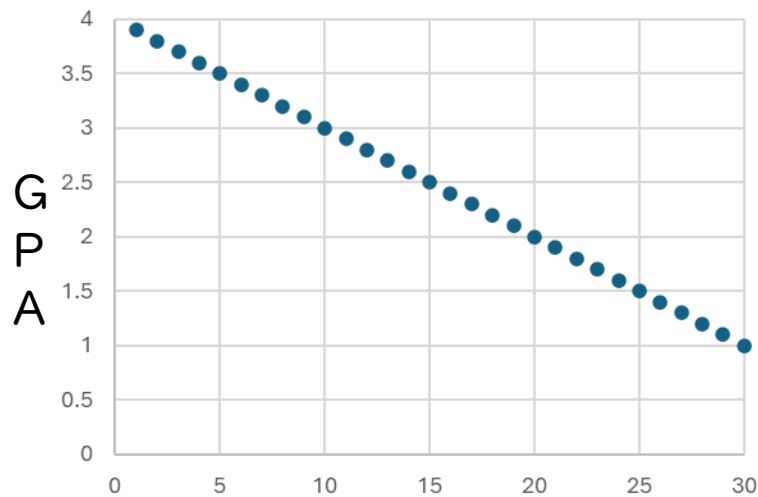
正の相関がある ($r=0.5$)



週あたり学習時間

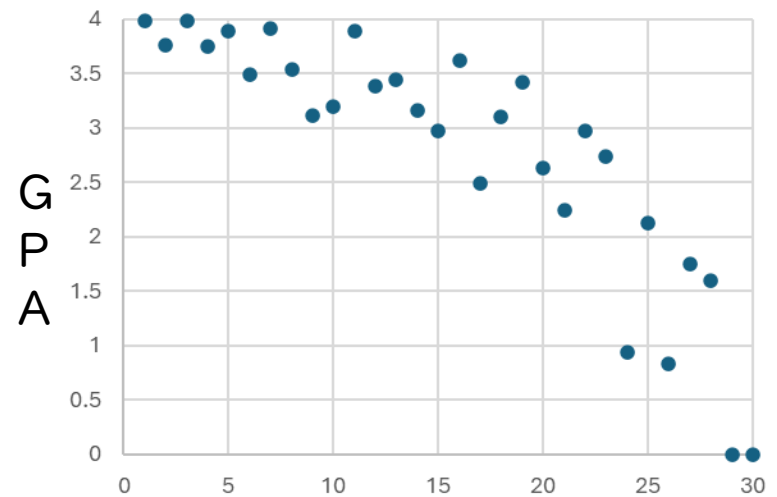
ざっくり相関分析のイメージ

強い負の相関 ($r=-1$)



週あたりスマホ使用時間

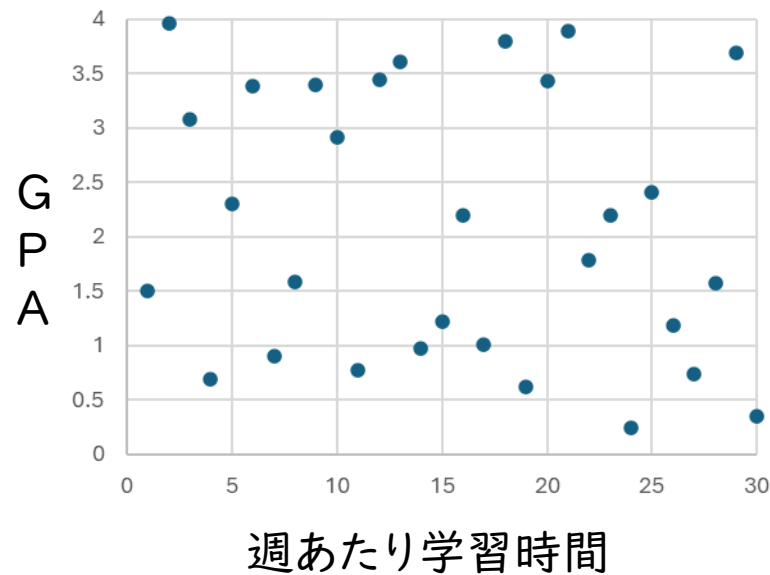
負の相関がある ($r=-0.5$)



週あたりスマホ使用時間

ざっくり相関分析のイメージ

相関がない ($r=0$)



ざっくり相関分析のイメージ

相関係数の値	相関関係の強さ
0.7~1	強い正の相関
0.4~0.7	正の相関
0.2~0.4	弱い正の相関
0	相関がない
-0.2~-0.4	弱い負の相関
-0.4~-0.7	負の相関
-0.7~-1	強い負の相関

相関関係は因果関係ではない

■ 授業出席率と試験の成績

相関：授業に多く出席する学生ほど試験の成績が高い

→ 出席そのものが成績を上げるのではなく、**元々学習に熱心**な学生だから出席率が高いだけかもしれない

■ アルバイト時間と就職内定獲得率

相関：アルバイト時間が長い学生ほど就職内定を獲得する割合が高い

→ アルバイトが直接内定獲得率を高めているのではなく、**積極的に行動する**学生がアルバイトも就職活動も頑張る傾向があるだけかもしれない

ワーク

- 大学成績と期待生涯賃金の増減は連動してそうですが、そのほか期待生涯賃金の増減に関わりそうな第3の要因（**交絡因子**）を考えてください

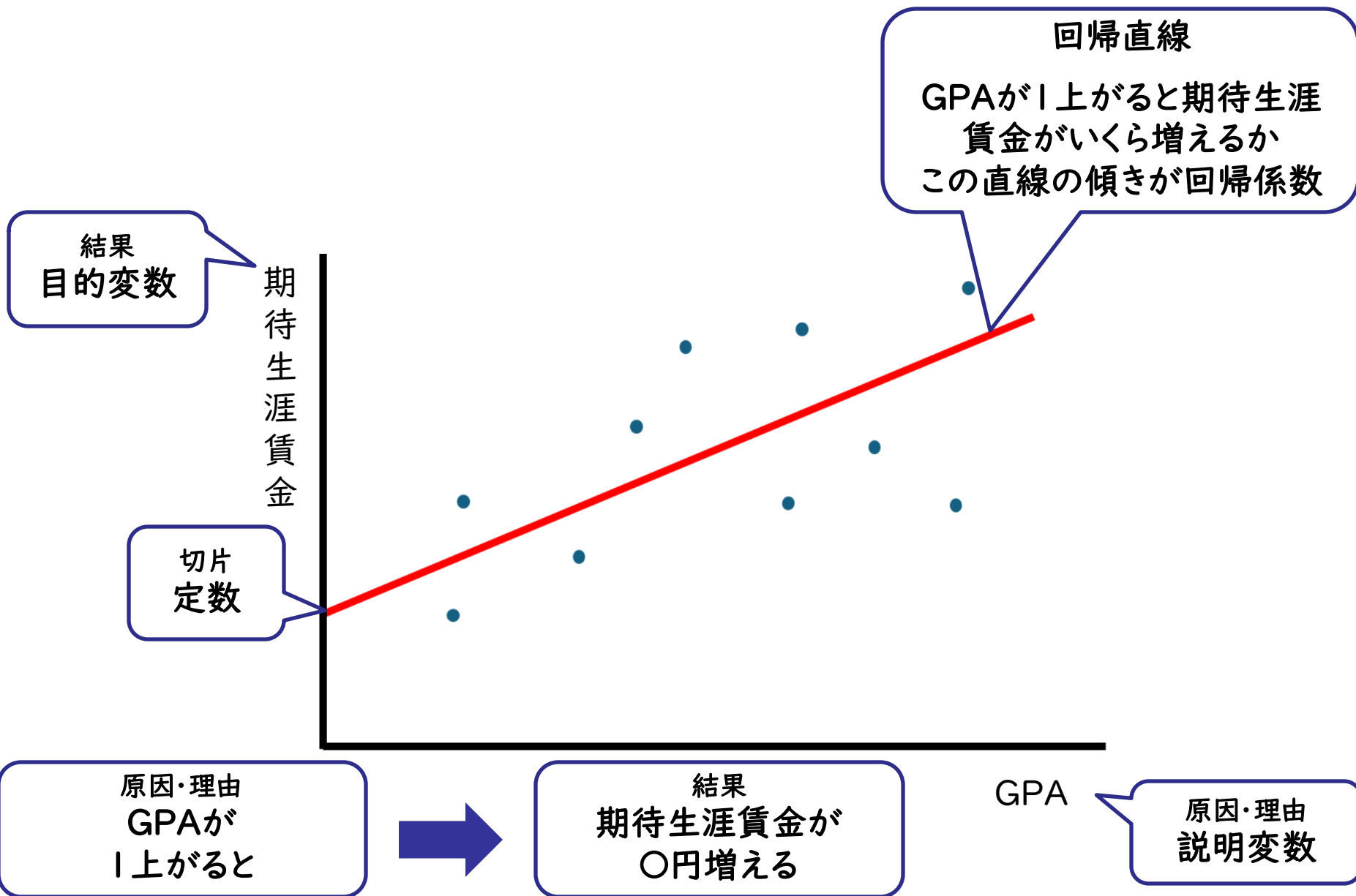
例：コミュニケーション能力、学部などの分野、
就職先地域、景気、才能

期待生涯賃金における第3の要因

執筆中の論文（未発表）

- 社会人2,000人対象とした追試（国公立・文理含む、実際の年収で調査）の結果、**交絡因子**として浮かび上がってきたもの
 - ・未知の課題に直面した（高校）
 - ・課題を解決した（大学）
 - ・学習に熱心に取り組んだ（高校・大学）
 - ・教師との人間関係（高校・大学）
 - ・仕事に熱心に取り組んだ
 - ・仕事での人間関係

ざっくり回帰分析のイメージ



回帰分析の結果

■ 大学成績の影響（表5）

大学成績（GPA）が1あがると、
生涯賃金がおよそ825万円あがる！

表5 大学成績の規定要因（筆者作成、値は偏回帰係数）

	大学成績		期待生涯賃金	
	係数	標準誤差	係数	標準誤差
高校ランク	.011 ***	.001	210.391 *	153.998
高校成績	.196 ***	.009	-1105.18	1312.59
大学成績			8249.01 ***	2045.48
(定数)	1.412 ***	.043	201348. ***	6372.54
R ²	.098 ***		.004 ***	
N	4,110		4,110	

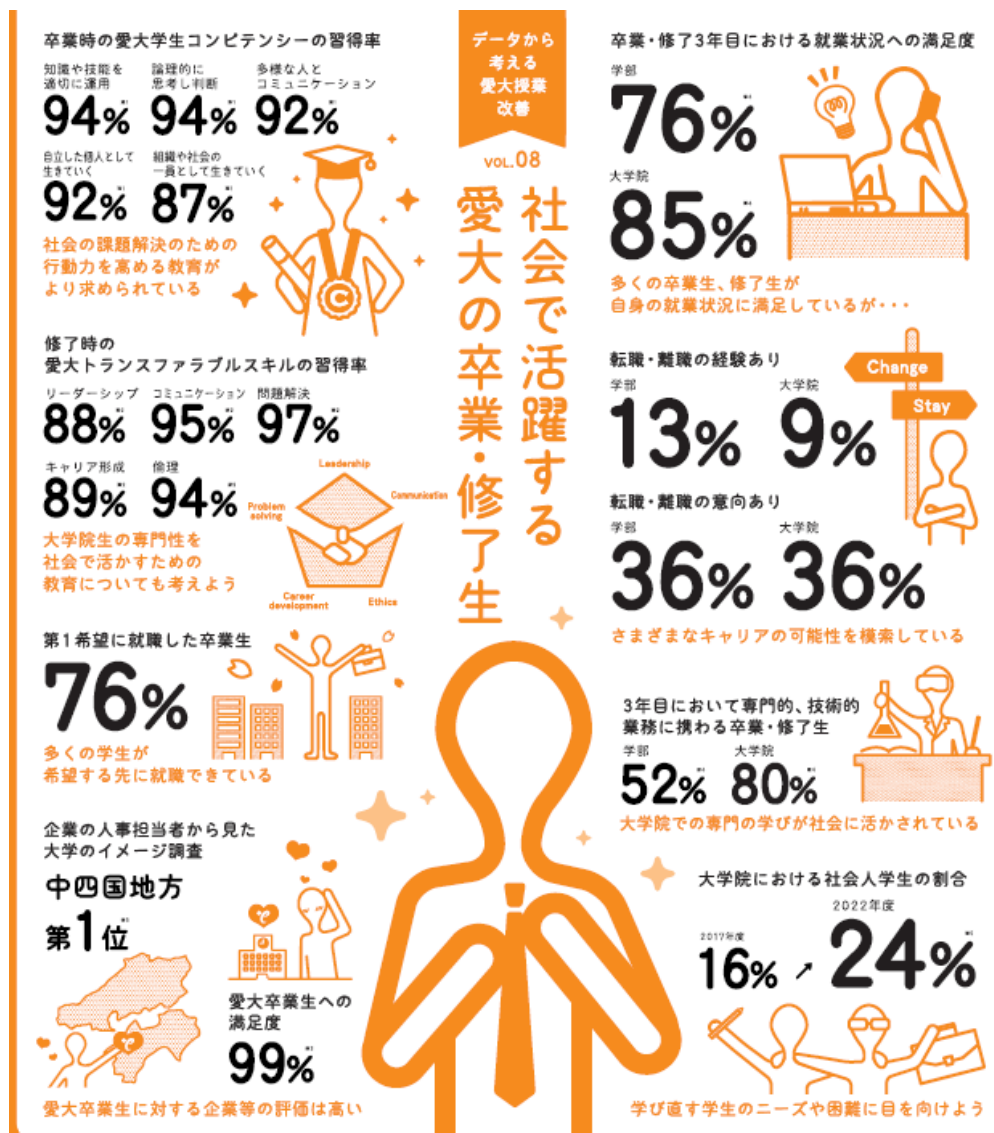
注：*** <.001 ** <.01 * <.05

活用につながる報告

相手にあわせた情報提供

- 相手が何を求めているかを把握する
- 聞き手がどのような情報を持っていて、報告に何を求めているかを理解しておく
- 重要なポイントを簡潔に伝える
- A4一枚ものにまとめて、詳細は資料として末尾につけるなどの工夫
- 適切なグラフを効果的に使う

情報提供（愛媛大学の例）



改善につなげる情報提供のプロセス

- 分析結果を共有し、議論する

議論の場：教授会、学部執行部との打ち合わせ、FD研修会、学生とのミーティングなど

- 問題点を明確にし、改善策を考える
- 改善策に必要な資源の配分を見直す
- 改善策の実行に必要な意思決定を行う

意思決定を支援する視点の例

- 過去や目標と比べて良いか／悪いか
- 学生や市場のニーズが高まっているか／そうでないか
- 他の集団や競合と比べて優れているか／劣っているか
- 設定した仮説が正しいか／そうでないか
- 費用対効果が良いか／悪いか

意思決定支援の例

■ 時系列での比較

志願者数が年々減少し、競争倍率が低下している。新入生調査によると
高校生はSNSを使って大学の情報を得ていた

志願者獲得におけ入試広報活動を強化するために・・・

- ・既存のSNSを拡充しよう・進学相談会などでSNSを周知しよう

■ 目標との関係

大学の目標として地域貢献を掲げているが、地域課題に取り組むプロジェクト型学習に参加した学生は5%未満だった

- ・学生の参加を促す方策を検討しよう・プロジェクトの内容を見直そう

意思決定支援の例

■ 学生のニーズ

学生アンケートによると、データサイエンス科目の開講を望む学生が7割を超えている

- ・新規開講科目を計画しよう・既存科目にデータサイエンスの要素を加えよう

■ 受験市場のニーズ

進学メディアの調査によると、情報系学部を志願する高校生が毎年5%ずつ増加している

- ・新しい学部の設置を検討しよう・もう少し市場の動向を見守ろう

■ 労働市場のニーズ

情報分野の人材需要がここ5年で10ポイント以上上昇している

- ・対応した人材を育成できるカリキュラムを検討しよう・企業にヒアリングしよう

意思決定支援の例

■ 全国平均との比較

国家試験の合格率が全国平均を下回っている。

・学内教員の指導力を向上させよう・外部講師を積極的に活用しよう

■ 競合との比較

同地区の他大学と比較して就職率が低い。

・キャリア支援体制を強化しよう・企業との連携を強めよう

意思決定支援の例

■ 仮説の検証

授業出席率と就職先の満足度には関連があるようだ

・教職員に発信して学生指導に生かそう・学生に発信しよう

■ 費用対効果

スポーツ奨学生の留年率が高い

・奨学生の成績基準を見直そう・スポーツ学生の学習支援を行おう

まとめと振り返り

本プログラムの振り返り

- 今日の内容を振り返って、印象に残ったこと、使えそうと思ったこと、今後実践してみたいことなどがあれば教えてください。小さなことでも構いません

振り返り

参考文献

- 愛媛大学教育企画室(2024)「データから考える愛大授業改善」
九州大学「データで見る九州大学」
中井俊樹・鳥居朋子・藤井都百編(2013)『大学のIR Q&A』玉川大学出版部
本田寛輔・浅野茂・鳶田敏行(2014)「米国のインスティテューショナル・リサーチ
(IR)業務の実態を整理する:説明責任,改善支援,通常業務,臨時業務の観点から」
大学評価・学位研究16:65-81
松田岳士・渡辺雄樹(2017)「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
文部科学省(2020)「教学マネジメント指針」
文部科学省(2024)「大学における教育内容等の改革状況について(令和4年度)」
Astin, A.W.(1993) Assessment for Excellence: The Philosophy and
Practice of Assessment and Education in Higher Education, Phenix,
Arizona: Oryx Press
Saupe, J. L. (1990) The Function of Institutional Research, 2nd Edition,
Association for Institutional Research