教育データ解析入門 I & II

中 山　晃

愛媛大学　教育・学生支援機構
英語教育センター

清 水 栄 子

愛媛大学　教育・学生支援機構
教育企画室

SPODフォーラム
2016.08.26

10:00~12:00
13:00~15:00
共通講義棟Ｂ
404講義室

EHIME UNIVERSITY

すばらし～シラバス

- 午前講義の前半（約50分： 10時50分まで）

  代表値と平均値、標準偏差の関係についての基礎的な理解のためのWarm-Upとミニ講義、作図のグループワークを行います（**事例1と事例2**）。

- 午前講義の中盤（約50分： 11時50分まで）

  相関関係についての理解のためのミニ講義と作図のグループワークを行い、集合データの平均値と関連するデータとの相関関係から一定の命題を推定する際に潜む危険性（例：生態学的誤謬、選択バイアス）について学びます（**事例3と事例4及びクイズとミニ数学**）。

  お昼休み

- 午後の講義（事例5を約80分、事例6を約40分： 15時00分まで）

  因果関係について理解につながる事例検討を、架空のシミュレーションデータを用いて、グループワーク形式で行います（**事例5と事例6**）。

Warm Up: 平均点とは？

表　仮想データ（ランダム）

学生番号	学部	得点A	得点B
1	1	63	42
2	1	44	47
3	1	24	59
4	1	67	49
5	1	91	51
6	1	38	55
7	1	83	41
8	1	95	54
9	1	25	41
10	1	57	48
11	1	20	57
.	.	.	.
.	.	.	.
.	.	.	.
300	3	35	54

ある集団の「平均点」って、その集団内の何%を占めていると思いますか？

n=1899　平均142点
のデータ内で
142点の学生：　87名
（約4%）

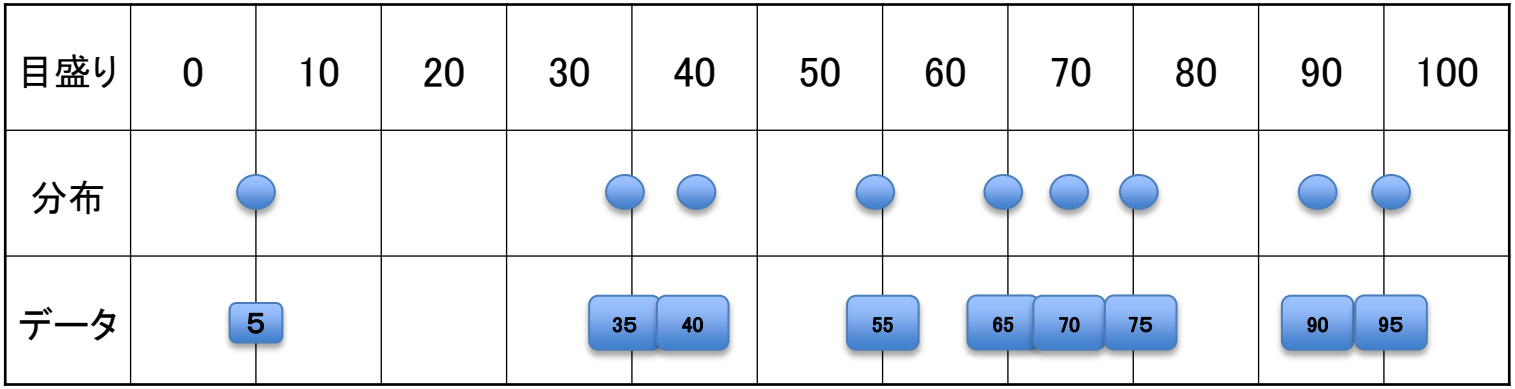
学部	平均点	度数	点数範囲
1	57.07	0	12–100
2	61.58	2	35–89
3	72.42	4	45–100

分布の代表値

- 様々なデータ(変数)の分布について
- その分布の特徴を記述する指標
- すなわち分布全体を1つの値で示すもの

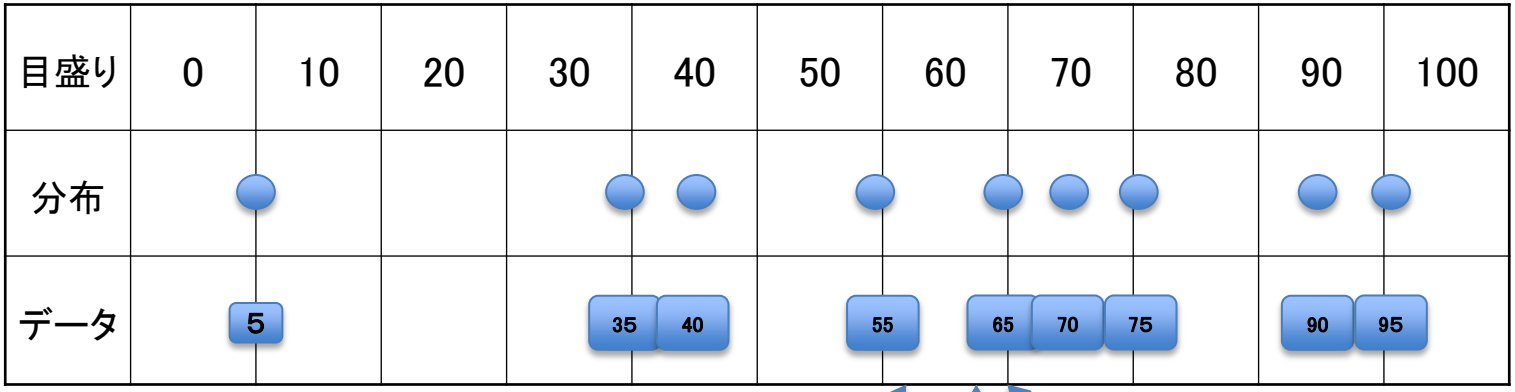
例： 中央値(Median)，平均(Mean)，最頻値(Mode)

中央値と平均



「代表値」はどの辺りにあるべきだと考えますか？

中央値と平均



58.00
平均

65.00
中央値

61.43
調整平均

平均点の外れ値に対する脆弱性に対処するために分布の両端から一定数の値を除いた上で求める平均

分布に含まれる各値にもっとも近い値

極端な例を見てみましょう

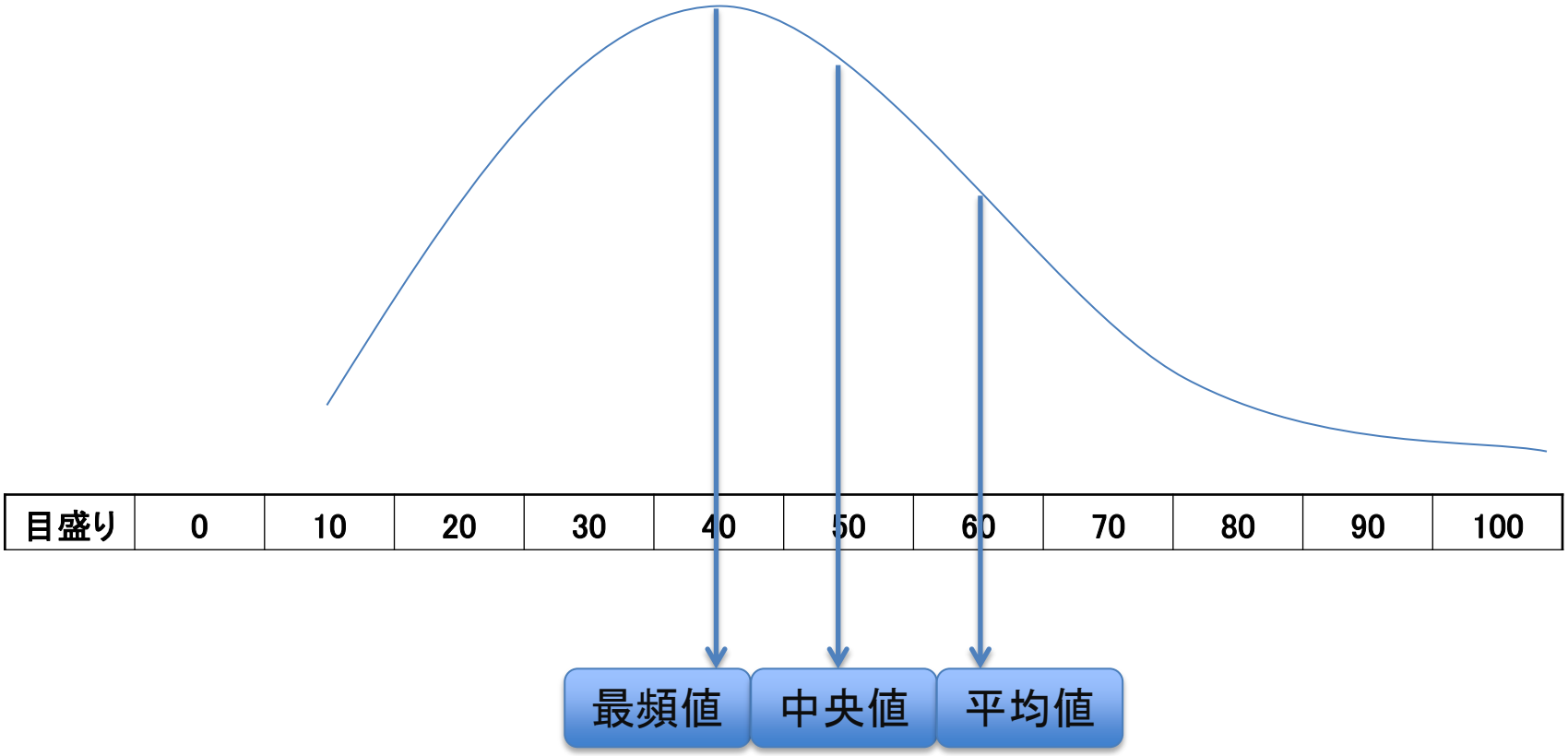
学生ID	英語	国語
1	0	0
2	0	0
3	0	50
4	100	100
5	100	100
中央値	計算してみましょう	
平均		

- どんな値になるでしょうか？
- 分布を代表していると言えますか？

Warm-upのまとめ

- データの分布を簡潔に把握したい（希望・欲求）
- ある点数1つでその分布を把握できるの？（疑問）
- 外れ値（outlier）込みで分布の代表値を求めるなら
  [redacted]を使うことが望ましい
- 外れ値（outlier）の影響を排除したいならば、
  [redacted]を使うことが望ましい

代表値の相対的な位置（イメージ）の例



平均と標準偏差の理解に向けて

事例の分析

（個別・グループワーク）

事例1　管理職との何気ないSNSでの会話から･･･

うちの学生のTOEICの平均点は、だいたいどのくらいなんだい？

既読
14:32

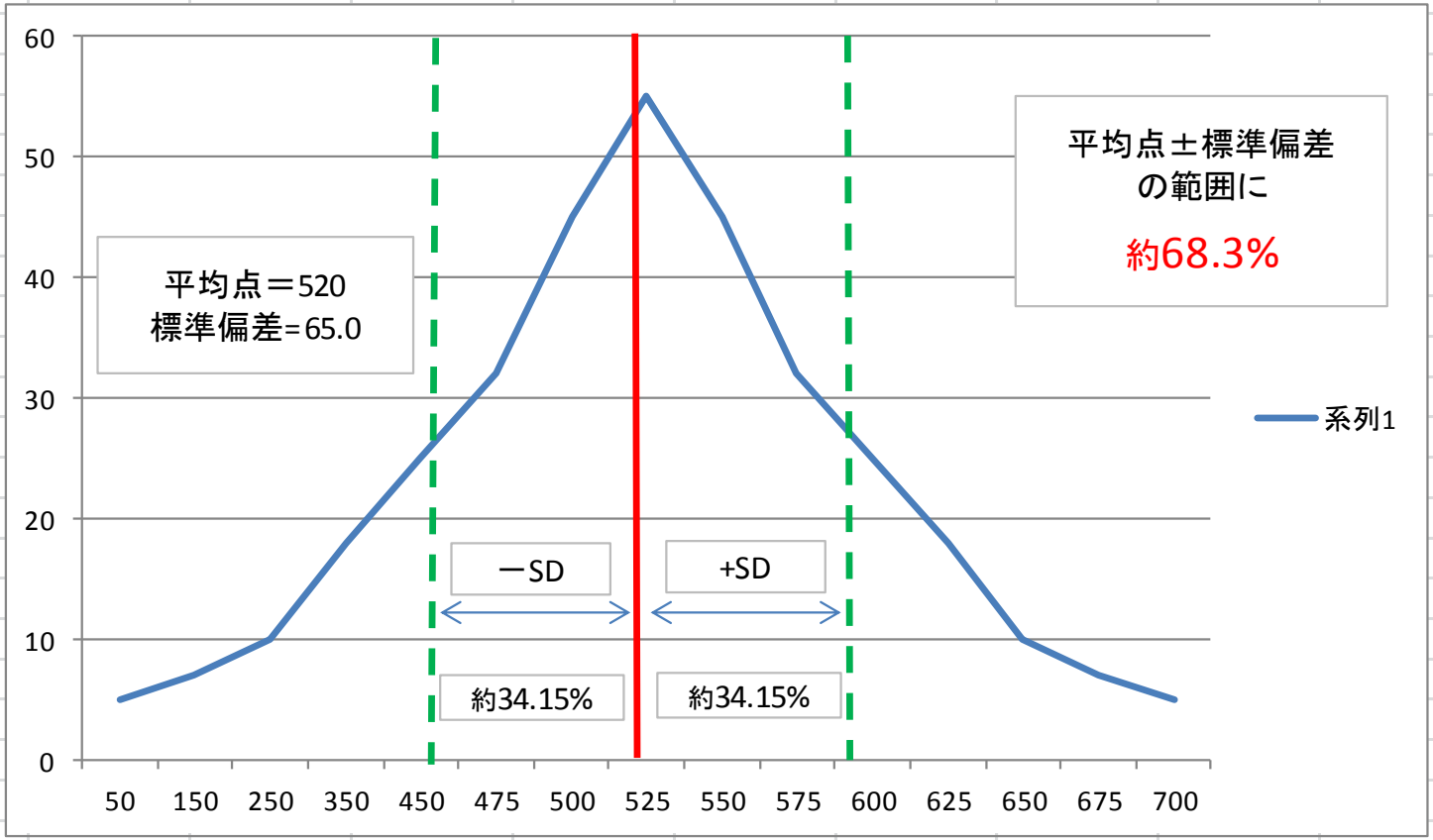
今回の結果だと、440点くらいです。ちなみに近県の同規模のB大やC大さんと比べても、10ポイント近くは上回っています。<(｀^´)>

素晴らしい！次回の昇給は、期待してくれたまえ～！

大学名	平均点 （*M：　Mean*）	標準偏差 （*SD：　Standard Deviation*）
ウチの大学	442.74	15.23
B大学	438.02	190.78
C大学	431.73	197.93

事例1での留意点

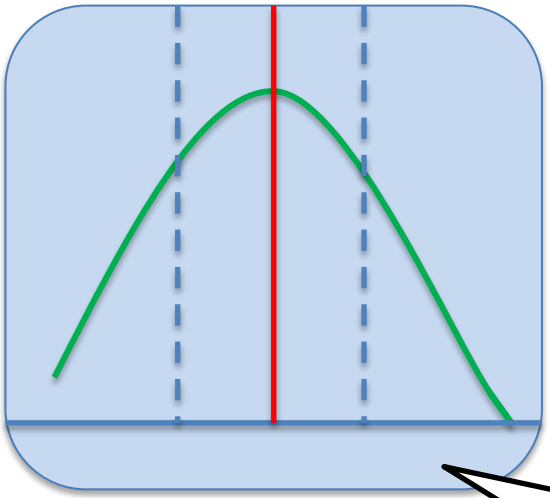
「平均点」だけではわからないデータの様相

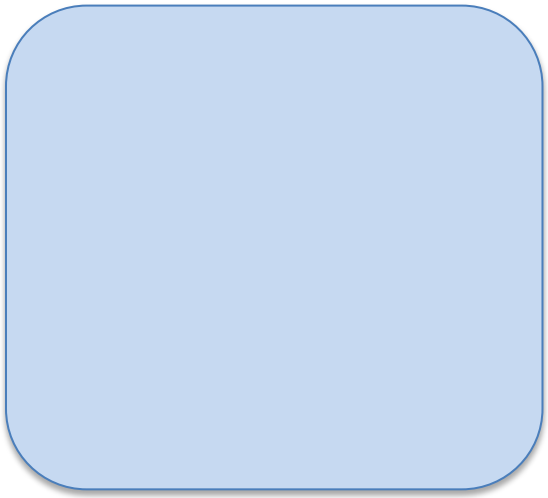


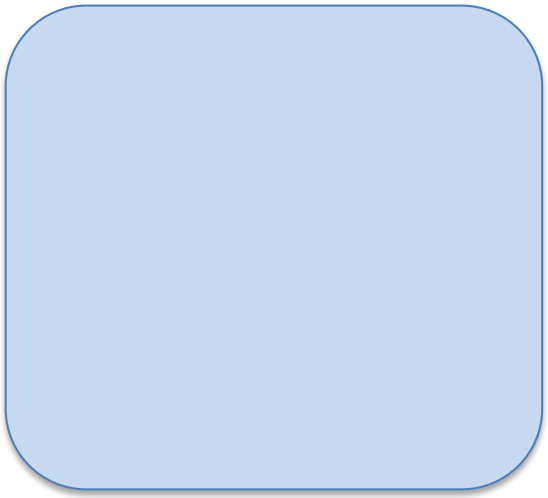
事例1を使ったワーク

ワーク1

事例１の３つの大学の結果をグラフにしてみましょう。



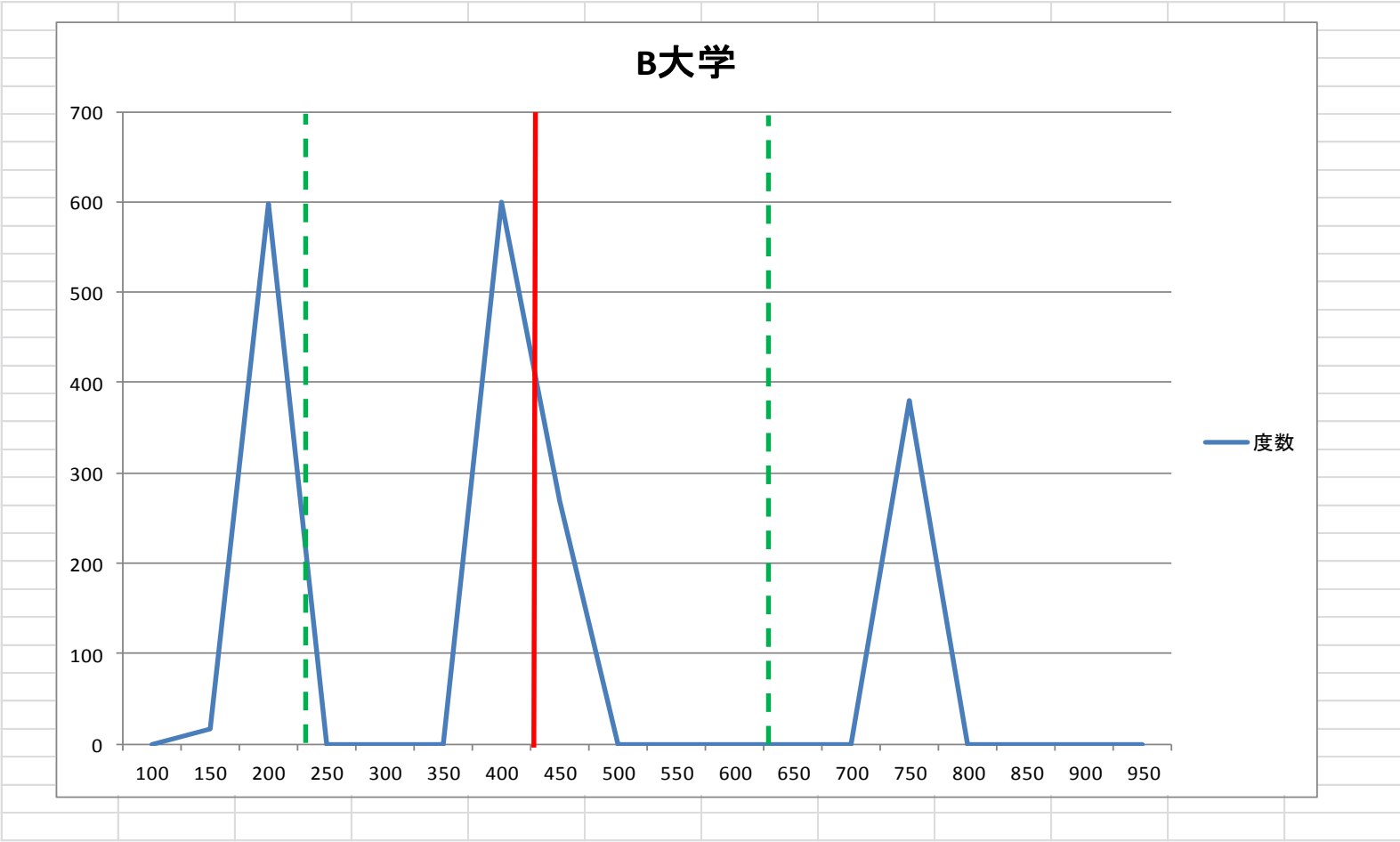




このような感じで3つ作成してみましょう！

結果の共有

実は、データを見て、グラフを書かないとわからないことがあります。



意地悪をしてすみません・・・。

事例1のまとめ

（　　）だけで、全体を把握したような気になってはいけない。

（　　）と（　　）の関係で、ある程度のデータの様相がわかる。

手間はかかりますが、学習成果は、様々な**指標**で多角的に確認しましょう！

余談ですが・・・

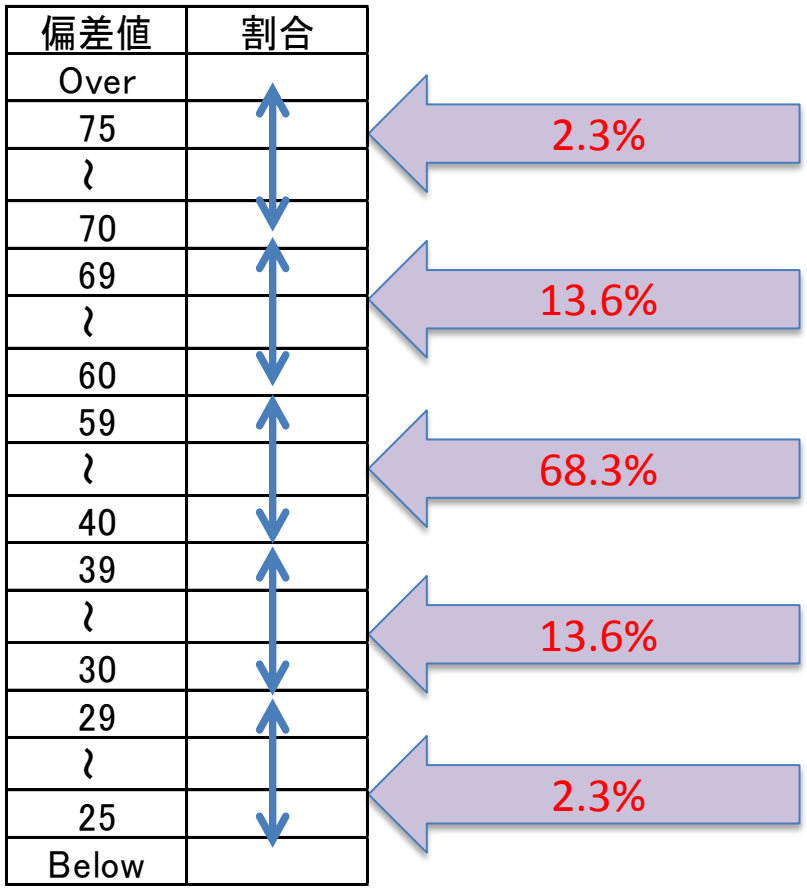
$$= \frac{10 \times (\text{得点} - \text{平均})}{\text{標準偏差}} + 50$$

ちなみに偏差値とは、

①中心の値を50と定め

②75～25の区間にデータの99%が含まれる

指標のことです。学習成果を相対的に、例えば特定集団の中での位置の変化（理解度の停滞や変容）を確認する際は、とても便利な指標といえます。



【　大学生の成績　】

基本的に絶対評価（ある基準を満たしているか、いないかで、秀・優・良・可がきまる）でなされる。

個人で作成した問題で、中間・期末の結果から相対的な学生の学力を把握する（個人内の比較や個人間の比較）場合には、一定の情報を提供してくれる指標と言える。

その他、授業科目間での割合に目を向けると、成績評価が甘い、あるいは厳しいといった分析することもできる。

成績評価基準の平準化・厳格化（GPAへの影響）・明確化を考えると、相対評価の意義が見えてくる。

事例2　とある会議での報告事項･･･

ウチの英語教育はうまくいっているのかね？

既読
14:32

検定試験の結果を見てください！大学全体の平均点は年度末で約1.5ポイント増しです。<(｀^´)>

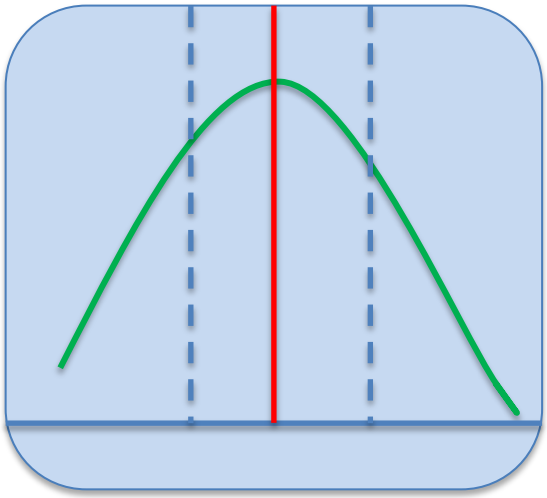
素晴らしい！次回の昇給は、もっと期待してくれたまえ～～！

大学平均と学部	入学直後の平均（SD）	年度末の平均点（SD）
大学全体（*N*=1980）	141.89（15.80）	143.37（18.10）
A学部	141.63（15.95）	139.71（14.17）
B学部	136.57（14.19）	138.88（11.73）

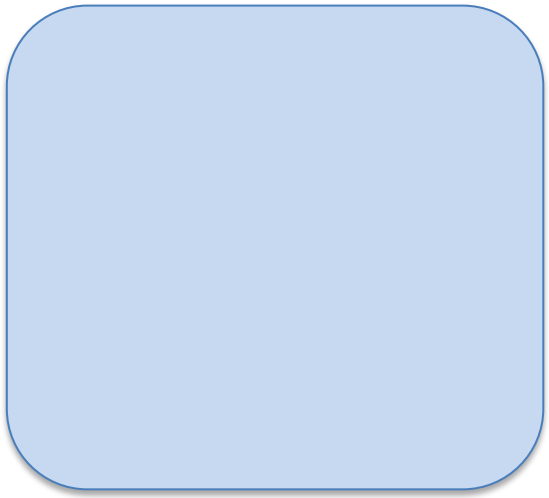
事例2を使ったワーク

ワーク2

事例2の「大学全体の結果」をグラフにしてみましょう。



入学直後



年度末

結果の共有

学生の成績がどのように変化したのか、全体としての可視化を試みる

今回のワークだけでは、「上位層が伸びて、下位層が下がった」のかどうかは、わからない

平均点が上がって、標準偏差が**小さく**なった場合に、学力が向上したと言えるのか　→

平均点が下がって、標準偏差が**大きく**なった場合に、学力が向上したと言えるのか　→

相関関係の理解に向けて

生態学的誤謬

選択バイアス

事例3 管理職との何気ないSNSでの会話から・・・

ワタクシさんは、TOEICで900点を超えているそうだね。ということは、英語がペラペラだし、報告書も英語でサラサラかけちゃったりするんだね。

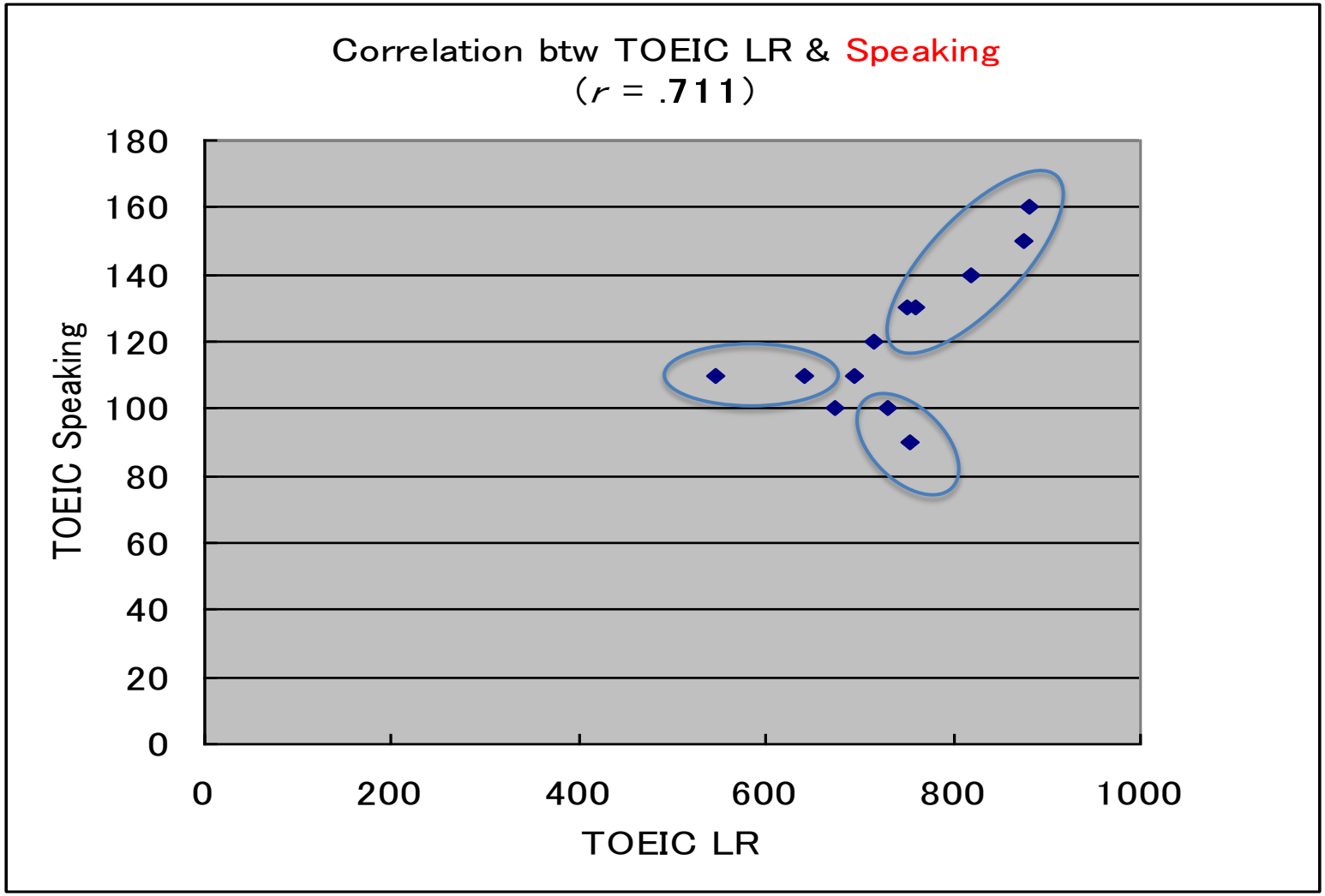
既読
14:32

いえいえ、実は残念なタイプなんです。<(`^´)>

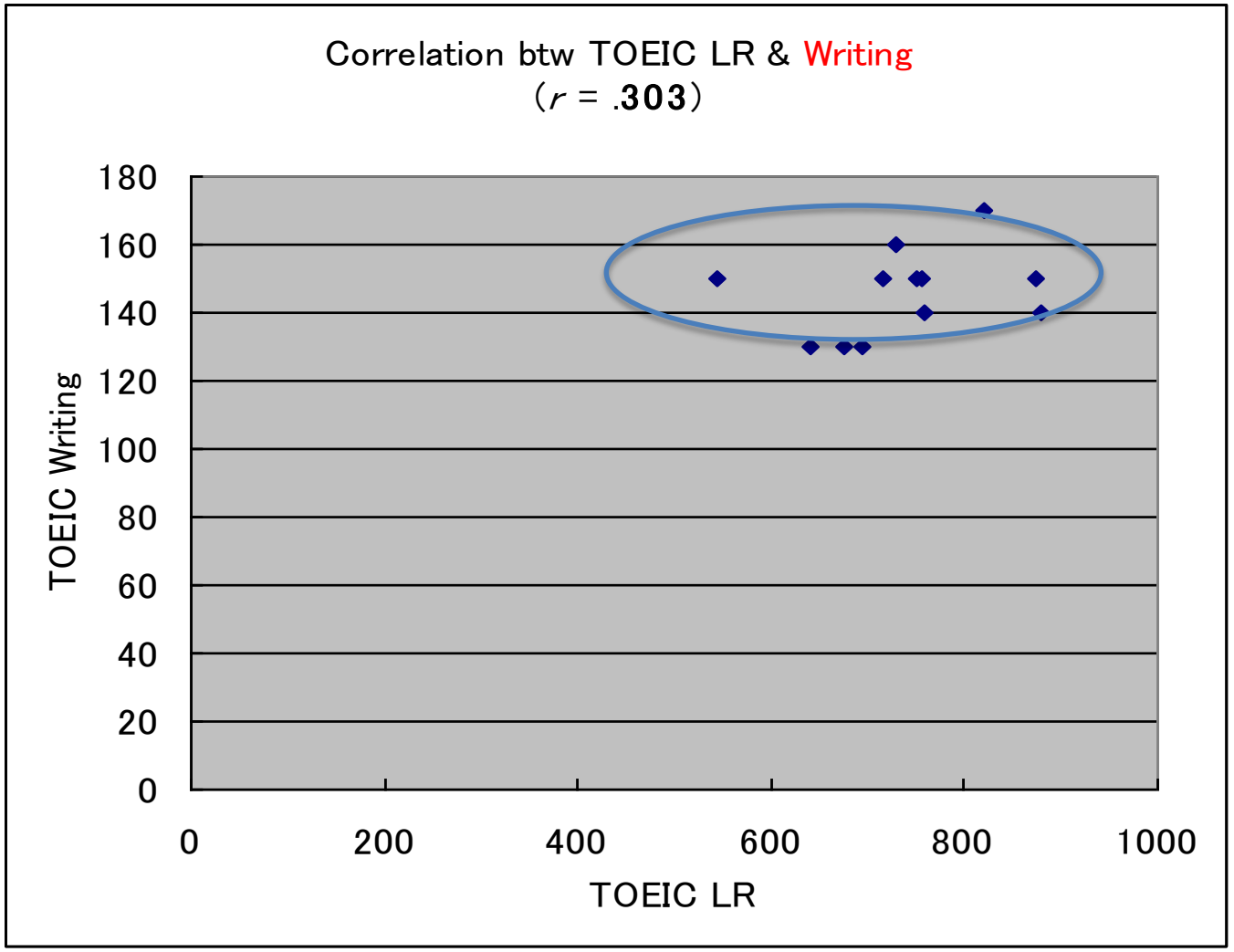
どういうことだね！？

氏名	TOEIC	Speaking	Writing
ワタクシ	900	90	140
Bさん	730	100	160
Cさん	600	110	150

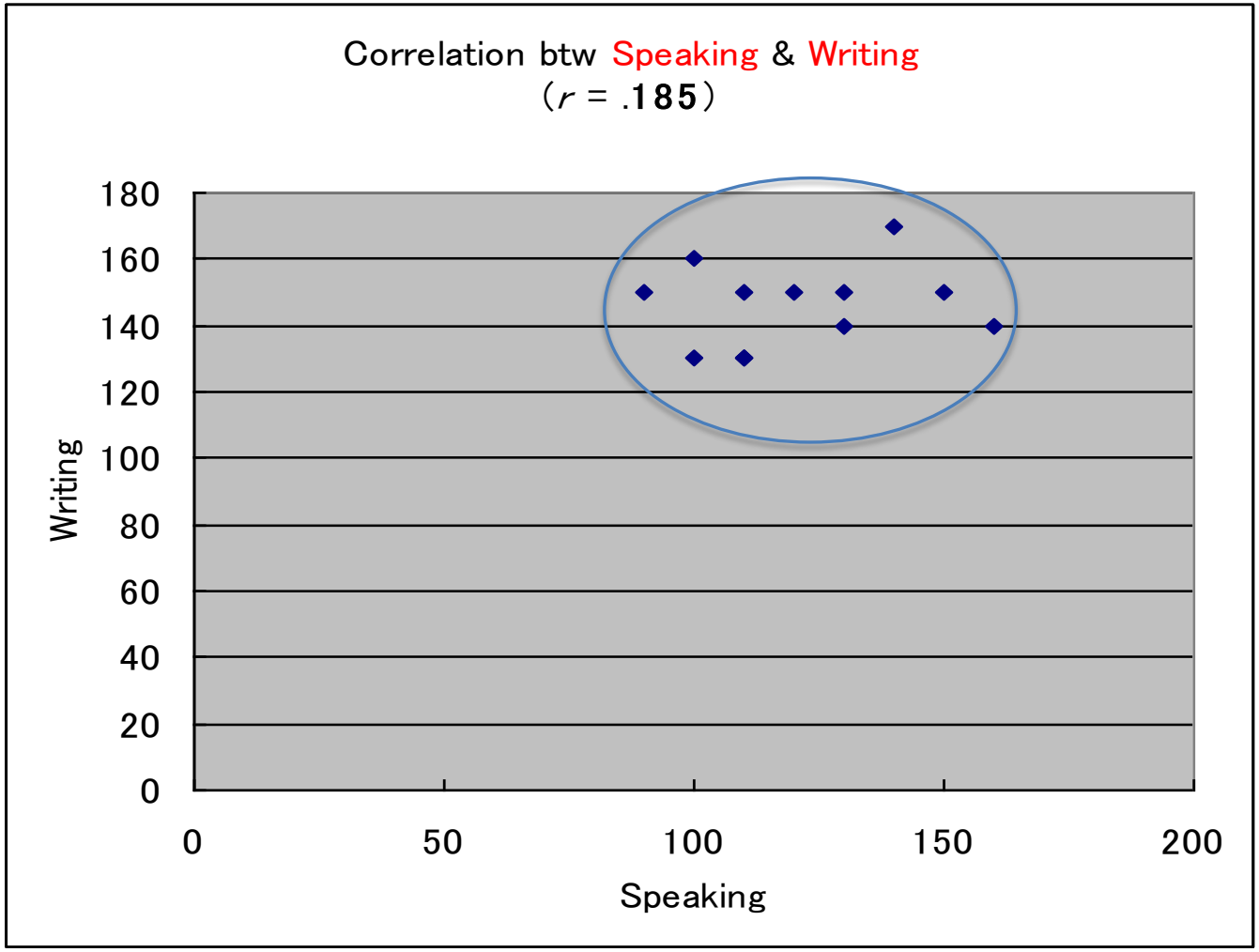
英語力を規定することの難しさ



Writing Skill は LR では規定できない



Speaking と Writing　は　同じ産出系能力でも相関しない



事例3での留意点

- 「L・R」と「S・W」って相関するものなのでしょうか？

  独立したスキル（認知能力や作業能力を必要とする）に相関を見出すことは、困難な場合が多い

  スキルの得点を合成（足し算）した場合、その結果の解釈には注意が必要

- 「Good Test Takers」ってどんな人？

  静穏な環境でしか発揮できないスキルを磨いていないか

  測定したい対象が、知識なのか、スキルなのか明確か

  （剣道やテニスの例）

事例4　とある会議での報告事項(パート2)・・・

やっぱり相関するのかね？英語力は？

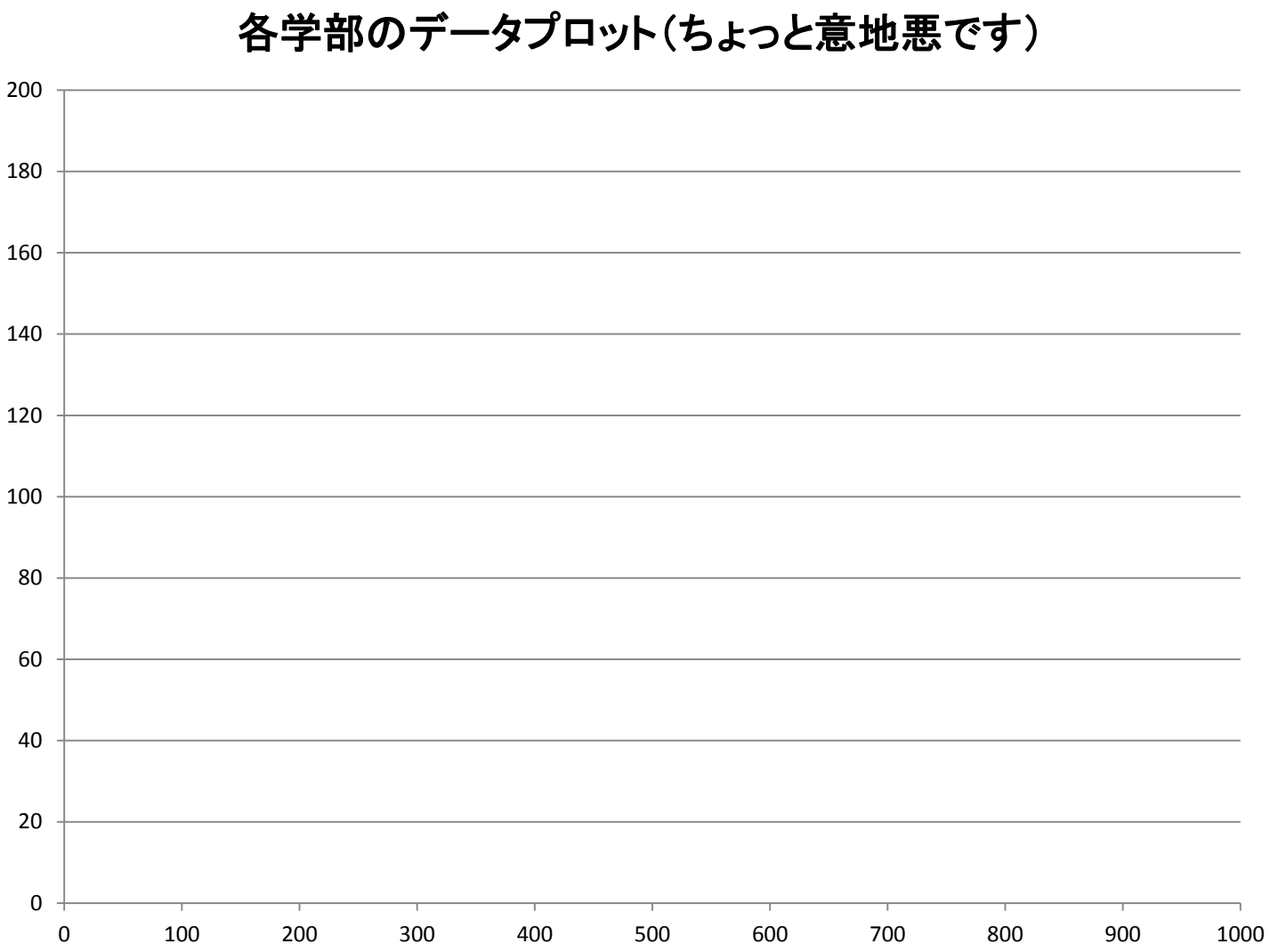
既読
14:32

あれ～、先生、聞いてなかったんですか。だからデータを見ただけではわからない構造があって、その～、え～っと・・・ ٩(•◡•)٩

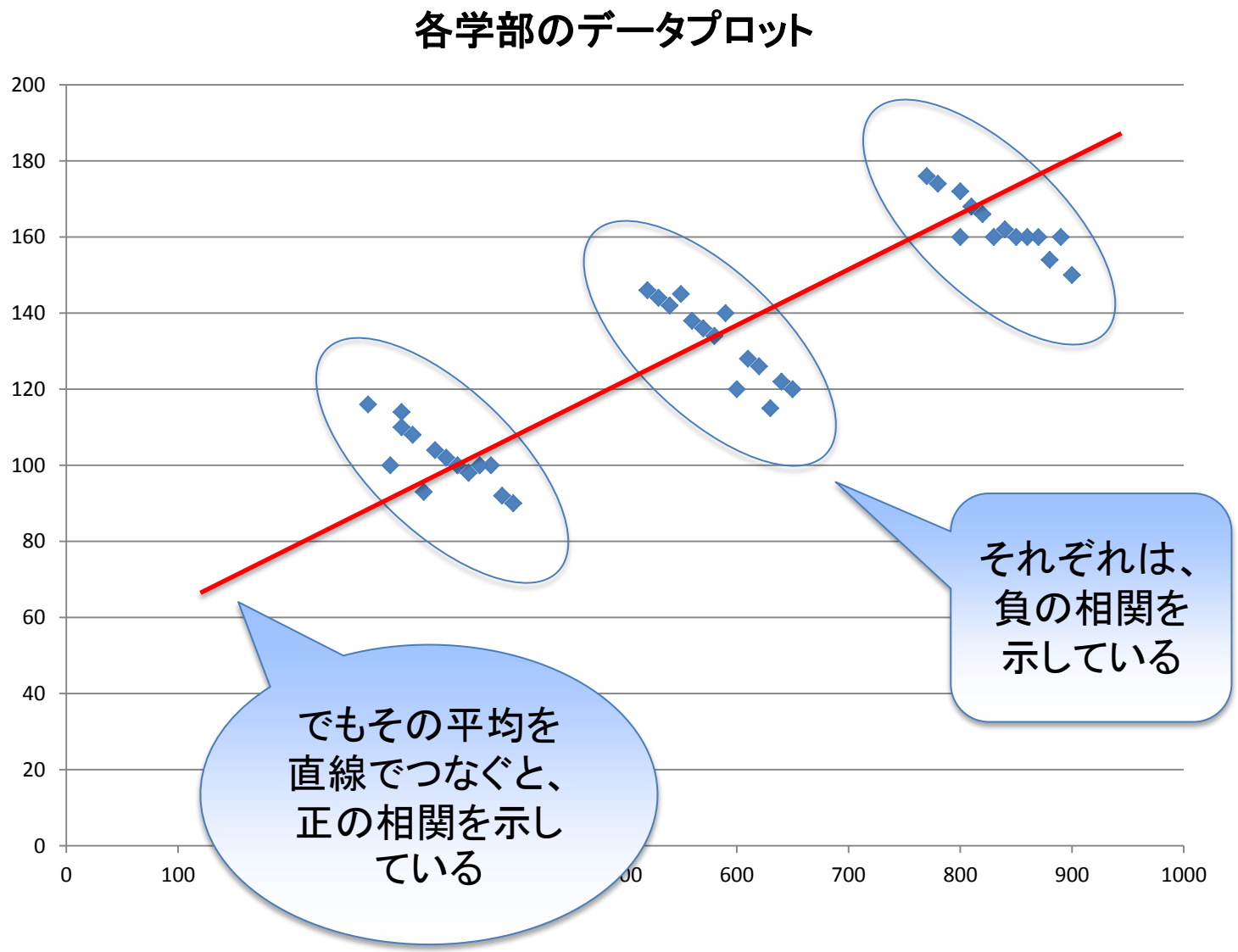
なん～だ、はっきりしたまえ。 (￣－￣)

各学部	LR平均	Speaking平均
A学部	836	163
B学部	585	133
C学部	336	102

事例4を使ったワーク


各学部のデータプロット（ちょっと意地悪です）
200
180
160
140
120
100
80
60
40
20
0
0
100
200
300
400
500
600
700
800
900
1000

結果の共有


各学部のデータプロット
200
180
160
140
120
100
80
60
40
20
0
0
100
600
700
800
900
1000
それぞれは、負の相関を示している
でもその平均を直線でつなぐと、正の相関を示している

結果の共有

中澤（2014, p. 75）を引用し，一部文言を変えて説明します。

（大学）全体の平均点は、学生個人のデータを集積して算出されたものと言えます。一般的に、高い平均点の大学や学部は、その学生集団の学力が高く、優れていると思われます。

しかしながらこのような推論は，学力の高い人は就職や進学に有利であるという私たちの経験則に過ぎず，実際は個人レベルの相関が集合レベルでも得られる保障はありません。

偏差値が高い×ソーシャルスキルが低い （この逆もしかり）
理数スキルが高い×英会話スキルが低い （この逆もしかり）

結果の共有

先ほどの事例では、個人レベルの相関が右下がりの負の相関を見せているのにも関わらず、集合レベルの相関は、正の相関を示しています。

このように、レベルの異なる平均点を基にして、相関についての推論をすると、誤った結論を導く可能性があります。

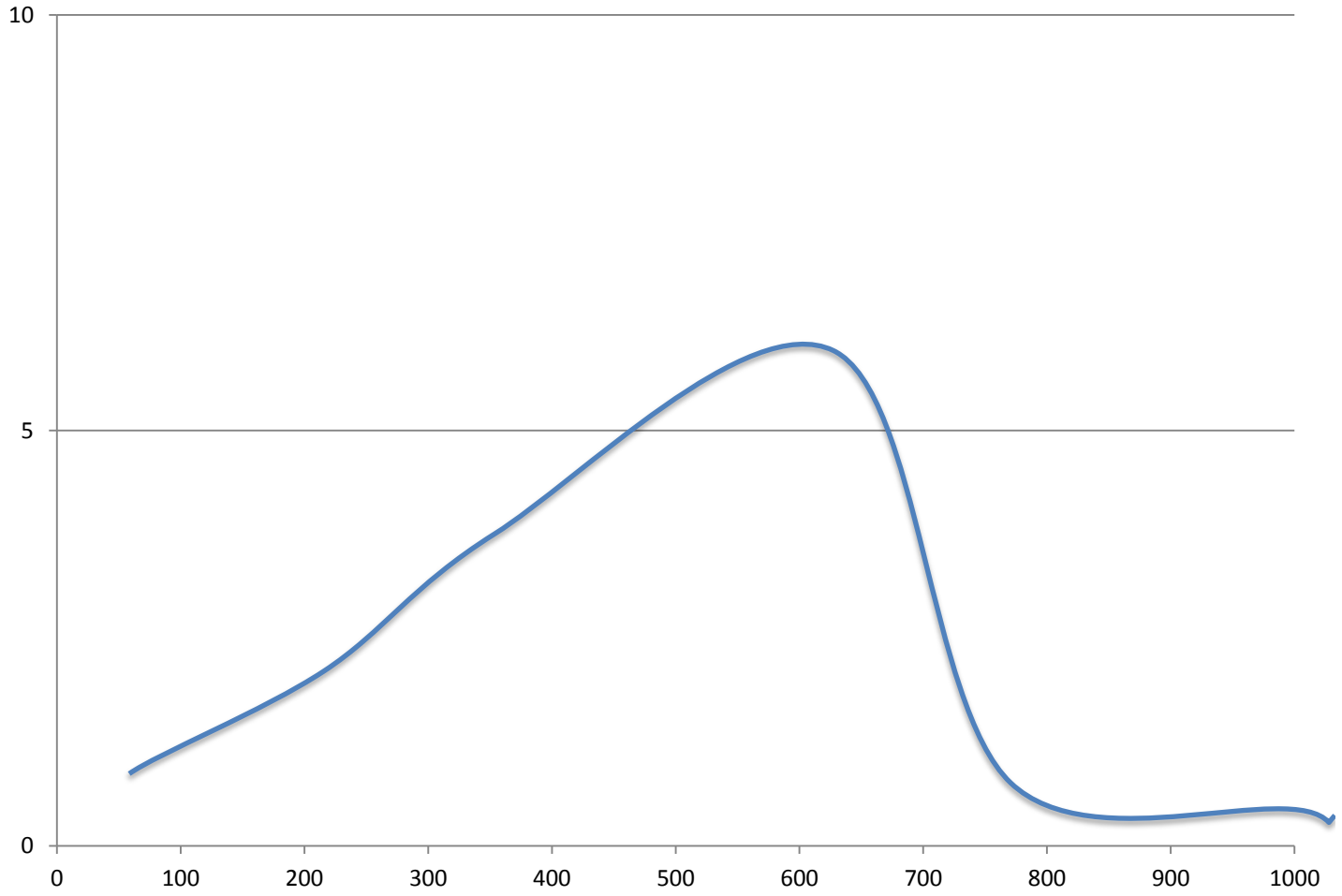
このことを　　と言います。

（Robinson, 2009; Lieberson, 1985; 中澤, 2014)

このように、平均値を取り出して、その多寡を、例えば教師の指導力や学生の学びの成果としてみなすことは、単なる憶測に基づく、乱暴な論調と言っても過言ではありません。なお、（学生が選抜されている）高い平均点をたたき出す学校が「教育効果が高い学校（大学）」とか「良い授業をする学校（大学）」とみなすことを「　　」と言います（中澤, 2014）。

ちょっと息抜き　クイズ！

とあるお茶目な実験から得られた曲線です。



そもそも相関係数って

$$r_{xy} = \frac{S_{xy}}{S_x \times S_y}$$

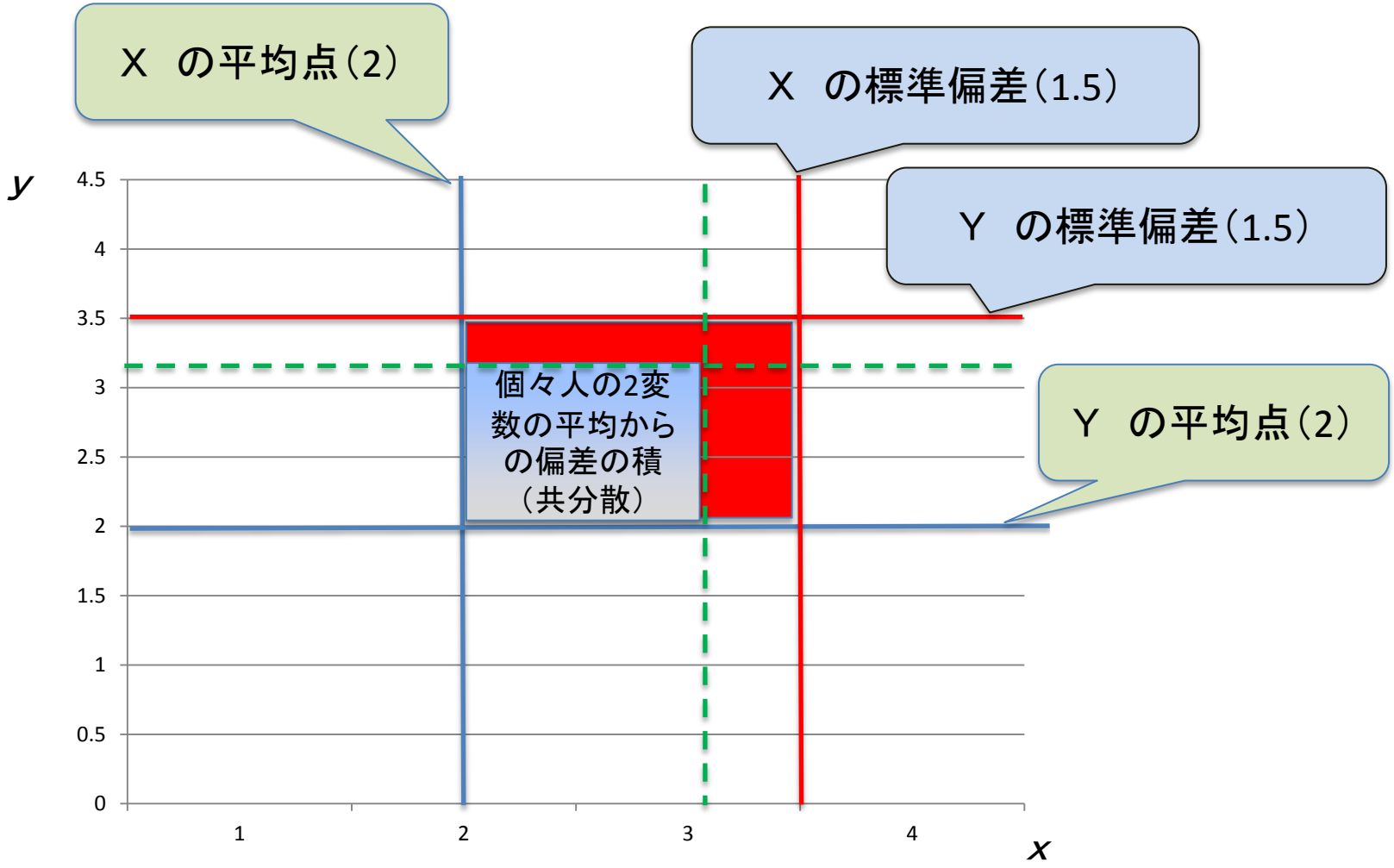
共分散：個々人の2つ変数に対する平均からの偏差の積

「ｉ」の1番からn番までを足す

$$r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

標準偏差

グラフィカルなイメージ図（あくまでイメージです。）



相関？

（ワタクシの例）

体重が重いと、身長が高い

体重が軽いと、体内脂肪が少ない

（満足度の例）

入学前にオープンキャンパスに参加した学生は、卒業時の満足度が高い

因果関係の理解に向けて

事例検討

（個別・グループワーク）

事例5

入試形態で学力差？

最近は、入試方法もいろいろとバラエティに富んでるよね～。で、実際のところ、学力的にはどうなのかね？

既読
14:32

ソレハそれは、バラエティに富んでます。<(`^´)>

こらこら、ちゃんと説明したまえ！

入試方式	とある科目の平均点	最高 / 最低
A方式	880	962 / 763
B方式	730	765 / 480
C方式	600	903 / 230

入試形態で学力差がみられたら

入試方式	とある科目の平均点	最高 / 最低
A方式	880	962 / 763
B方式	730	765 / 480
C方式	600	903 / 230

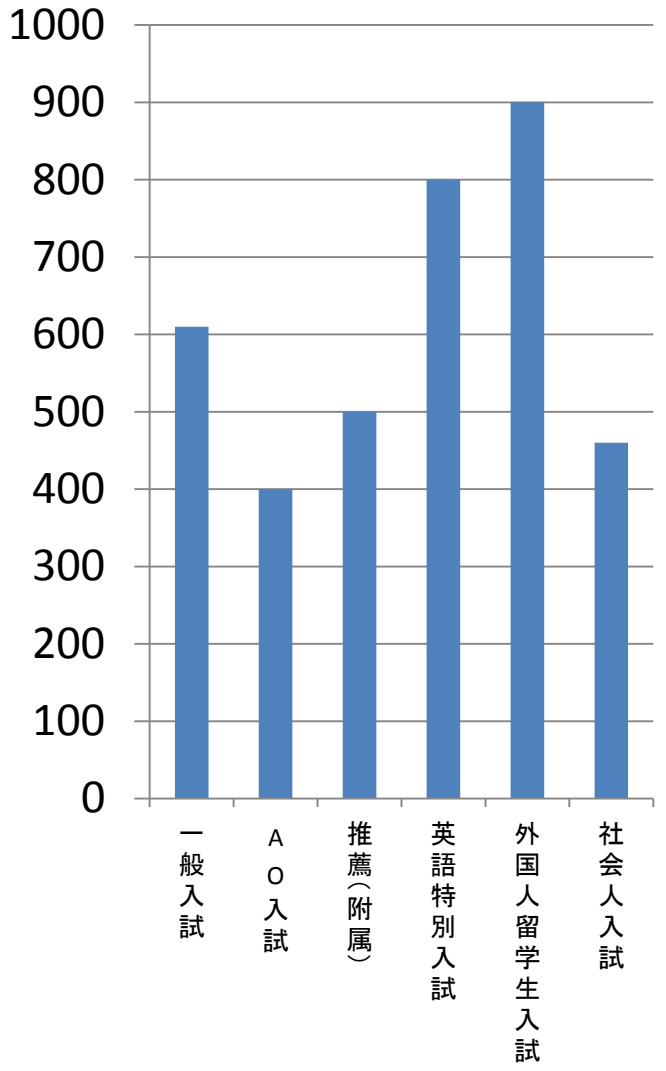
- このような表が会議に提出されても、どのように解釈すべきか、ちょっと困りますね。
- では、この表形式を改善したとして、この結果からどのような教育的な企画・開発（R&D）を提案すべきでしょうか？

入試タイプ別平均スコア（英語）

【　設問例　】
「本学入学時の外部英語検定試験の得点を記入し、また、入試形態を選択してください。」

【　回答例　】
1. 一般入試　2. AO入試　3. 推薦（附属）
4. 英語特別入試　5. 外国人留学生入試
6. 社会人入試

【 分析　（IRの部署の人達が考察する部分 】

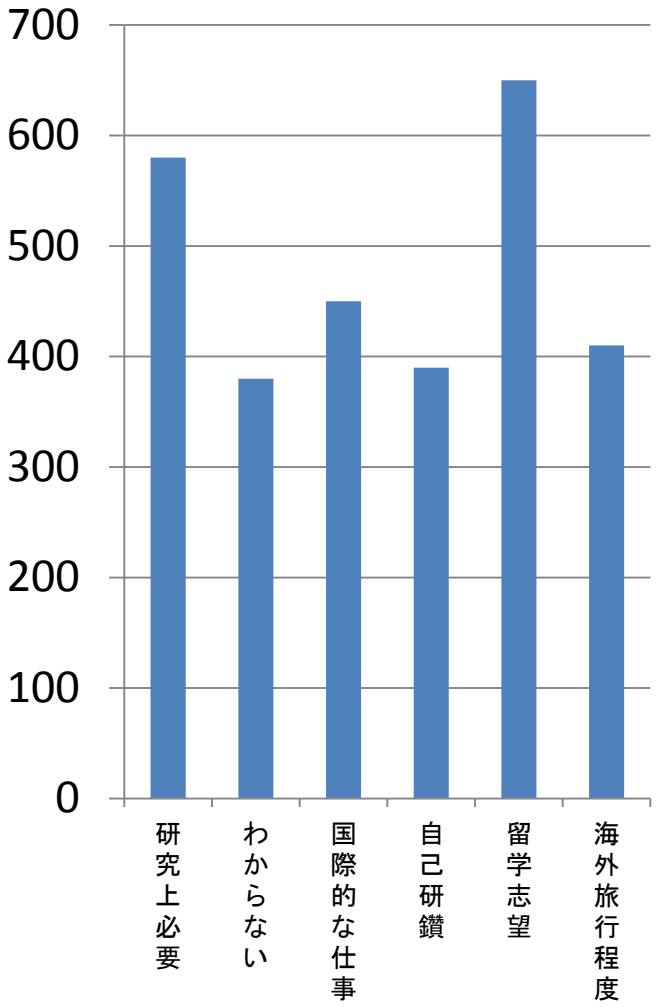


学習目的とスコア相関

【　設問例　】
「将来、英語をどのように活用する予定ですか。」

【　回答例　】
1.　研究上必要　2. わからない
3.　国際的な仕事がしたい
4.　自己研鑽　5. 留学志望　6. 海外旅行程度

【 分析 】

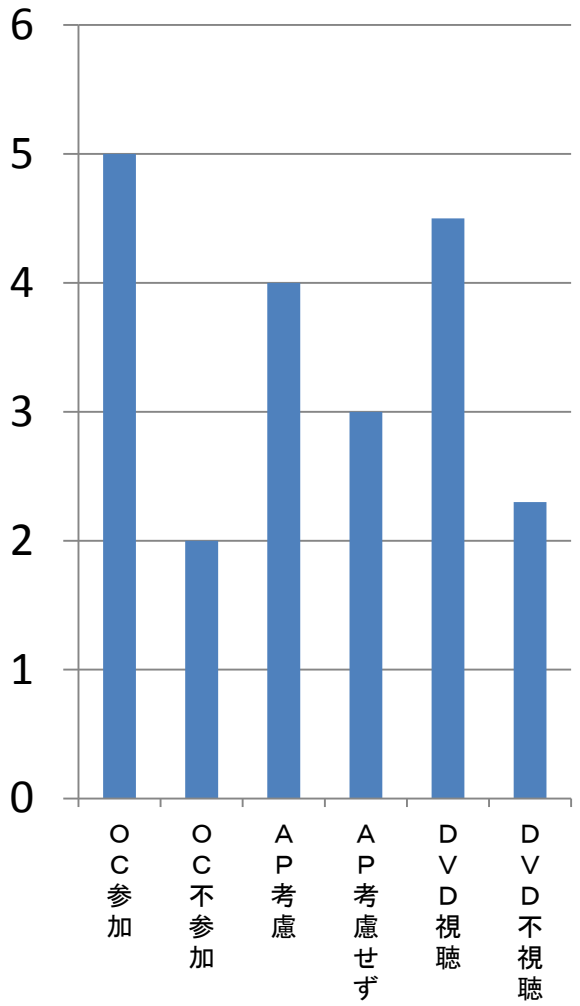


満足度（満足度が別の項目でわかっている・質問している場合）

【　設問例　】
「当てはまるものを選んで、回答してください。」

【　回答例　】
1. オープンキャンパスに参加した　2. していない
3. APを考慮した　4. APを考慮していない
5. DVDを見た　　6. DVDを見ていない

【 分析 】



中間まとめ

相関：　共変動する変数の関係
（例　Listening と Readingの得点など）

因果：　背後にある属性、あるいは時間的に差がある変数間の関係

IRアナリストとしての仕事！

満足度（直前の事例のクロス・データ）

【　設問例　】
「当てはまるものを選んで、回答してください。」

【　回答例　】
1. OC・AP・DVD（有）　2. OC・AP（有）　3. OC（有）
4. AP・DVD（有）　5. DVD（有）　6. すべて無
7. AP（有）　8. OC・DVD（有）

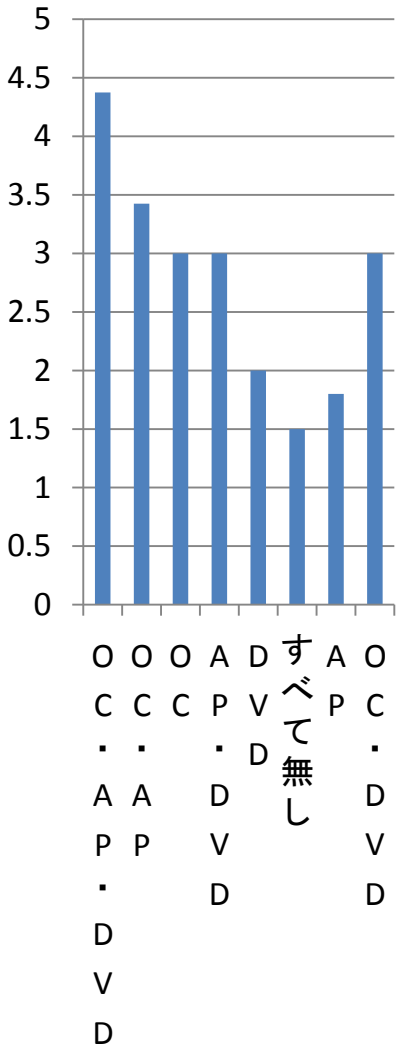
OC: オープンキャンパス、　AP：アドミッションポリシー
DVD：大学の紹介VTR

【 IRアナリストとしての課題】
Q　満足度を向上させるためにはどうするべきか？

↓

コスト（予算配分）とエフォート（人的資源の活用）

しかし、このグラフだけでは読み取りに限界が・・・。

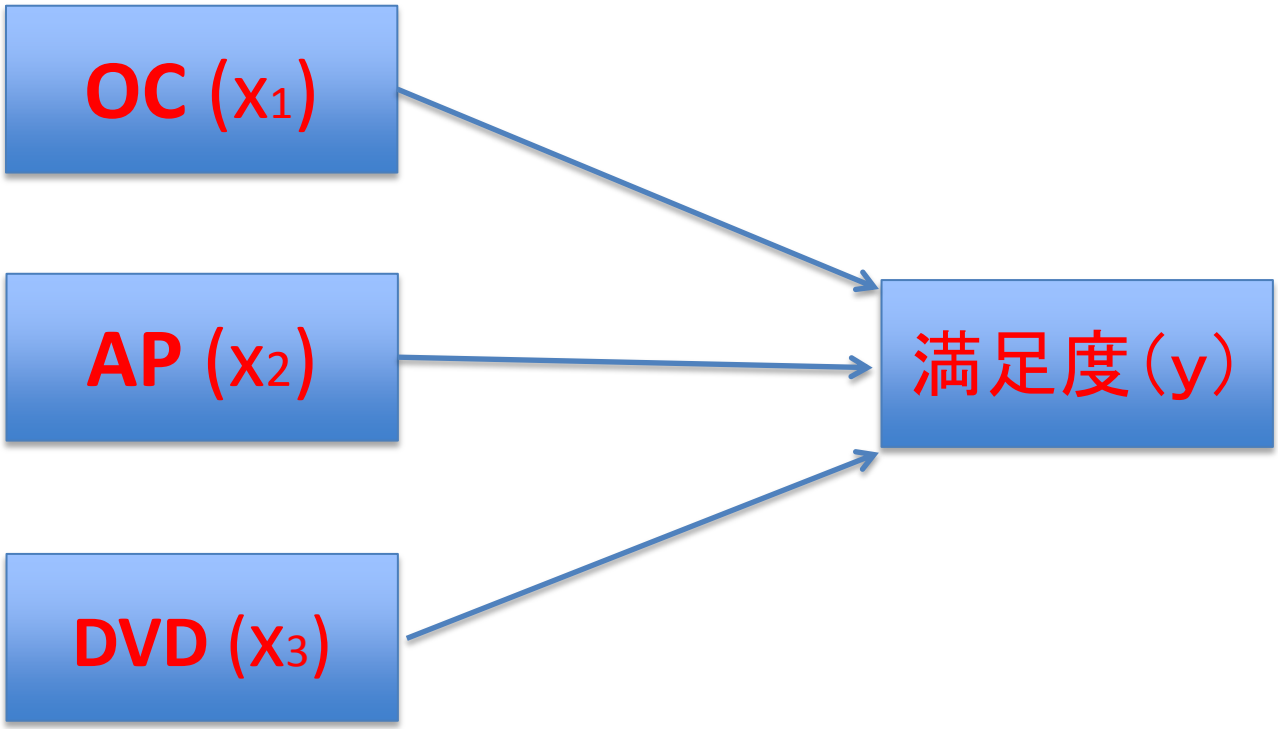


満足度に影響を与える要因（満足度を予測する要因）

問い

- どうやら、満足度をある程度規定する要因としては、「OC（オープンキャンパス）に参加したかどうか」、「AP（アドミッション・ポリシー）を考慮したかどうか」、「大学案内のDVDを視聴したかどうか」が挙げられるようであるが、では、どの要因にコスト（資金）とエフォート（人的資源）をかけることが好ましいだろうか？

グラフィカル・イメージ



相関関係と因果関係の違い

$$y_1 = \beta_1 x_{11} + \beta_2 x_{21} + \beta_3 x_{31} + e_1$$
$$y_2 = \beta_1 x_{12} + \beta_2 x_{22} + \beta_3 x_{32} + e_2$$
$$\vdots$$
$$y_n = \beta_1 x_{1n} + \beta_2 x_{2n} + \beta_3 x_{3n} + e_3$$

（誤差がすくなくなるように）
この連立方程式を
解くことで下記の
回帰式を得る！

$$Y = b_1 x_1 + b_2 x_2 + b_3 x_3 + e_n$$

（満足度）＝（OCの係数) x1 ＋ (APの係数) x2 ＋(DVDの係数） x3 ＋ e1

回帰式のデータ型

学生ID	満足度 (y)	OC (x_1)	AP (x_2)	DVD (x_3)
900xxx1	5	2	2	2
900xxx2	3	2	2	1
900xxx3	2	1	1	2
900xxx4	1	1	1	1

従属変数： 満足度＝5段階（5～1）
独立変数： OC，AP，DVDは、それぞれ2段階（1無、2有）

回帰分析結果

	標準化係数（β）	標準誤差（推定値の標準偏差）	自由度	F	有意確率
OC	.654	.086	2	57.526	.000
AP	.267	.095	2	7.936	.002
DVD	.265	.091	2	8.430	.001

R^2 = .733（分散説明率：自由度修正済み決定係数）

事例5のまとめ

- 記述統計量（グラフ化を含め）だけでは、意思決定には情報が不足する。
- 統計分析を使い、意思決定のための情報を増やすことが必要となる。
- 相関分析とその結果はあくまで2変量の共変動についての情報であるが、回帰分析とその結果は、時間的に差異のある変数間に潜む因果関係についての情報を提供してくれる。

おまけ

ID	OC（参加2・不参加1）	AP（考慮2・考慮せず1）	DVD（視聴2・視聴せず1）	満足度
N1	2	2	2	4
N2	2	2	2	4
N3	2	2	2	5
N4	2	2	2	4
N5	2	2	2	5
N6	2	2	2	4
N7	2	2	2	5
N8	2	2	2	4
N9	2	2	1	4
N10	2	2	1	3
N11	2	2	1	4
N12	2	2	1	3
N13	2	2	1	4
N14	2	2	1	3
N15	2	2	1	3
N16	2	1	1	4
N17	2	1	1	4
N18	2	1	1	2
N19	2	1	1	3
N20	2	1	1	2
N21	1	2	2	3
N22	1	2	2	2
N23	1	2	2	3
N24	1	2	2	2
N25	1	2	2	3
N26	2	1	2	2
N27	2	1	2	4
N28	1	2	1	2
N29	1	2	1	1
N30	1	1	2	2
N31	1	1	1	2
N32	1	1	1	1
N33	1	1	1	2
N34	1	1	1	1
N35	1	1	1	1
N36	1	1	1	1
N37	1	1	1	1
N38	1	1	1	1
N39	1	1	1	2
N40	1	1	1	1
M	1.55	1.55	1.40	2.78
SD	0.50	0.50	0.50	1.27

データを標準化

OC（参加2・不参加1）	AP（考慮2・考慮せず1）	DVD（視聴2・視聴せず1）	満足度
0.8931558	0.8931558	1.2093387	0.9640161
0.8931558	0.8931558	1.2093387	0.9640161
0.8931558	0.8931558	1.2093387	1.7509681
0.8931558	0.8931558	1.2093387	0.9640161
0.8931558	0.8931558	1.2093387	1.7509681
0.8931558	0.8931558	1.2093387	0.9640161
0.8931558	0.8931558	1.2093387	1.7509681
0.8931558	0.8931558	1.2093387	0.9640161
0.8931558	0.8931558	−0.806226	0.9640161
0.8931558	0.8931558	−0.806226	0.1770642
0.8931558	0.8931558	−0.806226	0.9640161
0.8931558	0.8931558	−0.806226	0.1770642
0.8931558	0.8931558	−0.806226	0.9640161
0.8931558	0.8931558	−0.806226	0.1770642
0.8931558	0.8931558	−0.806226	0.1770642
0.8931558	−1.091635	−0.806226	0.9640161
0.8931558	−1.091635	−0.806226	0.9640161
0.8931558	−1.091635	−0.806226	−0.609888
0.8931558	−1.091635	−0.806226	0.1770642
0.8931558	−1.091635	−0.806226	−0.609888
−1.091635	0.8931558	1.2093387	0.1770642
−1.091635	0.8931558	1.2093387	−0.609888
−1.091635	0.8931558	1.2093387	0.1770642
−1.091635	0.8931558	1.2093387	−0.609888
−1.091635	0.8931558	1.2093387	0.1770642
0.8931558	−1.091635	1.2093387	−0.609888
0.8931558	−1.091635	1.2093387	0.9640161
−1.091635	0.8931558	−0.806226	−0.609888
−1.091635	0.8931558	−0.806226	−1.39684
−1.091635	−1.091635	1.2093387	−0.609888
−1.091635	−1.091635	−0.806226	−0.609888
−1.091635	−1.091635	−0.806226	−1.39684
−1.091635	−1.091635	−0.806226	−0.609888
−1.091635	−1.091635	−0.806226	−1.39684
−1.091635	−1.091635	−0.806226	−1.39684
−1.091635	−1.091635	−0.806226	−1.39684
−1.091635	−1.091635	−0.806226	−1.39684
−1.091635	−1.091635	−0.806226	−1.39684
−1.091635	−1.091635	−0.806226	−0.609888
−1.091635	−1.091635	−0.806226	−1.39684

関数 **STANDARDIZE**

（標準化したいセル，M，標準偏差）

= STANDARDIZE(C3,C$43,C$44)

データの標準化とは

平均値 を「0」，標準偏差「1」となるように変換すること

標準化を行うことで、様々なデータを統計学的に解釈しやすくすること

架空データ.xlsx - Microsoft Excel

ファイル ホーム 挿入 ページ レイアウト 数式 データ 校閲 表示 PDF エクセル統計

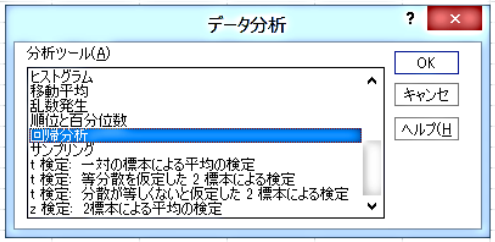
外部データの取り込み 接続 並べ替えとフィルター データ ツール アウトライン 分析

I2 OC

ID	OC（参加2・不参加1）	AP（考慮2・考慮せず1）	DVD（視聴2・視聴せず1）	満足度
N1	2	2	2	4
N2	2	2	2	4
N3	2	2	2	5
N4	2	2	2	4
N5	2	2	2	5
N6	2	2	2	4
N7	2	2	2	5
N8	2	2	2	4
N9	2	2	1	4
N10	2	2	1	3
N11	2	2	1	4
N12	2	2	1	3
N13	2	2	1	4
N14	2	2	1	3
N15	2	2	1	3
N16	2	1	1	4
N17	2	1	1	4
N18	2	1	1	2
N19	2	1	1	3
N20	2	1	1	2
N21	1	2	2	3
N22	1	2	2	2
N23	1	2	2	3
N24	1	2	2	2
N25	1	2	2	3
N26	2	1	2	2
N27	2	1	2	4
N28	1	2	1	2
N29	1	2	1	1
N30	1	1	2	2
N31	1	1	1	2
N32	1	1	1	1
N33	1	1	1	2
N34	1	1	1	1
N35	1	1	1	1
N36	1	1	1	1
N37	1	1	1	1
N38	1	1	1	1
N39	1	1	1	2
N40	1	1	1	1
M	1.55	1.55	1.40	2.78
SD	0.50	0.50	0.50	1.27

OC（参加2・不参加1）	AP（考慮2・考慮せず1）	DVD（視聴2・視聴せず1）	満足度
0.8931558	0.8931558	1.2093387	0.9640161
0.8931558	0.8931558	1.2093387	0.9640161
0.8931558	0.8931558	1.2093387	1.7509681
0.8931558	0.8931558	1.2093387	0.9640161
0.8931558	0.8931558	1.2093387	1.7509681
0.8931558	0.8931558	1.2093387	0.9640161
0.8931558	0.8931558	1.2093387	1.7509681
0.8931558	0.8931558	1.2093387	0.9640161
0.8931558	0.8931558	-0.806226	0.9640161
0.8931558	0.8931558	-0.806226	0.1770642
0.8931558	0.8931558	-0.806226	0.9640161
0.8931558	0.8931558	-0.806226	0.1770642
0.8931558	0.8931558	-0.806226	0.9640161
0.8931558	0.8931558	-0.806226	0.1770642
0.8931558	0.8931558	-0.806226	0.1770642
0.8931558	-1.091635	-0.806226	0.9640161
0.8931558	-1.091635	-0.806226	0.9640161
0.8931558	-1.091635	-0.806226	-0.609888
0.8931558	-1.091635	-0.806226	0.1770642
0.8931558	-1.091635	-0.806226	-0.609888
-1.091635	0.8931558	1.2093387	0.1770642
-1.091635	0.8931558	1.2093387	-0.609888
-1.091635	0.8931558	1.2093387	0.1770642
-1.091635	0.8931558	1.2093387	-0.609888
-1.091635	0.8931558	1.2093387	0.1770642
0.8931558	-1.091635	1.2093387	-0.609888
0.8931558	-1.091635	1.2093387	0.9640161
-1.091635	0.8931558	-0.806226	-0.609888
-1.091635	0.8931558	-0.806226	-1.39684
-1.091635	-1.091635	1.2093387	-0.609888
-1.091635	-1.091635	-0.806226	-0.609888
-1.091635	-1.091635	-0.806226	-1.39684
-1.091635	-1.091635	-0.806226	-0.609888
-1.091635	-1.091635	-0.806226	-1.39684
-1.091635	-1.091635	-0.806226	-1.39684
-1.091635	-1.091635	-0.806226	-1.39684
-1.091635	-1.091635	-0.806226	-1.39684
-1.091635	-1.091635	-0.806226	-0.609888
-1.091635	-1.091635	-0.806226	-1.39684

ここに回帰分析のツールがあります



ランダム 代表値 最頻値 重回帰(簡易データ) 重回帰(簡易非標準化) 重回帰(簡易標準化) 重回帰(フルデータ) 重回帰(フルノーマル) 重回帰(フル購入ソフト) 在籍状況データ クロス集計 在籍状況データ分析 Sheet1

コマンド 平均: -1.40166E-16

結果の読み取り（概略）

概要

回帰統計	
重相関 R	0.875
重決定 R2	0.766
補正 R2	0.747
標準誤差	0.503
観測数	40

独立変数の数の影響を取り除き，見かけ上の当てはまりの良さを差し引いた自由度調整済決定係数によって回帰式を評価する

分散分析表

	自由度	変動	分散	観測された分散比	有意 F
回帰	3	29.8928512	9.96428372	39.38820155	0.0000
残差	36	9.10714884	0.25297636		
合計	39	39			

0.05以下なら回帰式は予測に使えると判断

	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	0.00	0.07952615	2.0438E-15	1	-0.16129	0.161287	-0.16129	0.161287
X 値 1	0.65	0.08423482	7.74181382	0.0000	0.481294	0.822966	0.481294	0.822966
X 値 2	0.24	0.09263159	2.63783911	0.0122	0.056482	0.432213	0.056482	0.432213
X 値 3	0.29	0.08924692	3.20725069	0.0028	0.105236	0.467238	0.105236	0.467238

標準化されているので、係数を比較することが可能

0.05以下なら統計的に有意な変数と判断

最後に

- 本プログラムの到達目標は達成できましたか？
- エクセルでも回帰分析は、そこそこ簡単にできます。
- ただし、事前に、どのような内容でアンケートや調査を行い、どのような結果が得られるのかしっかりとシュミレーションをしておかないと、得たいものが得られず、また調査・分析をやり直さなくてはならず、コストやエフォートがかかってしまいます。
- （このコースの続きとなる）実際の分析は、10月に行われる「IRer養成講座@椙山女学園大学」（実践講習）で行います。来年も本講座があるなら、エクセルを使った実務（実習）を扱いたいです。

引用・参考文献

- 中澤渉. (2014). 教育データを解釈する– 教育社会学における計量分析. 生産と技術, 66(1), 75-77.
- 南風原朝和. (2002). 心理統計学の基礎 統合的理解のために. 東京: 有斐閣アルマ.
- Lieberson, S. (1985). Making it count: The improvement of social research and theory. Univ of California Press.
- Robinson, W. S. (2009). Ecological correlations and the behavior of individuals. International journal of epidemiology, 38(2), 337-341.
- TOEIC®スコアデータ活用ブック. (2005). 財団法人国際ビジネスコミュニケーション協会TOEIC運営委員会